White Paper

Mass Distributed Server Consolidation

System z Mainframe Linux-on-z/VM Extreme Virtualization far Outclasses Over-hyped x86/x64 Approaches

About this White Paper

Scale-out distributed servers *(RISC-UNIX and x86/x64 volume types)* vastly proliferated like a fiercely addictive drug; reaching a net live c. 28M distributed servers by end-2007. This drove IT management costs up six-fold to a vast $120B *(2007)*, and IT electrical power/cooling costs to $22B *(2007)*, combined user operating cost more than triple global 2007 new server spending! Such distributed computing proved an **economic catastrophe**, and an **unspeakably wasteful environmental abuse**.

Radical consolidation and elimination of most or all of this "distributed computing disease"/"scourge" is now urgently required at many thousands of large IT users afflicted by these costly, unmanageable, legacy distributed IT infrastructures. Fashionable some years ago, these are now much discredited as more efficient infrastructure options become more widely deployed.

Do not trust the "IT drug dealers" who previously sold your enterprise its distributed computing for unbiased consolidation advice, nor fall too readily for the latest siren x86/x64 "miracle technology cures", several of which are getting high media attention currently. More trustworthy sources of IT consolidation advice/guidance are available!

All IT industry players agree that **virtualization** is the fundamental technology that can now most optimize/improve the efficiency of all IT systems; and of servers in particular. Virtualization is thus the main means to eliminate much of the distributed computing mess. Major benefits flow from better server partitioning and virtualization solutions, but these are far from equal. Absolute, unquestioned virtualization "Gold-Standard", across the whole IT industry, is IBM's System z mainframe, which pioneered virtualization developments over its 40-year evolution. *(See Section 4.)*

IFL-specialty-processor-powered, lower-cost, System z mainframes running Linux-on-z/VM *(IBM's famous, industrial-strength, extreme virtualization hypervisor)*, can easily consolidate the workloads from up to a few hundred *(even a thousand)* stand-alone distributed servers *(x86/x64-UNIX)* onto a single, efficient, compact, secure, and totally-reliable mainframe. Huge cost savings, with rapid ROI paybacks, are readily delivered. Users can thus rapidly escape major parts of their "distributed server disaster" legacy with this immensely powerful, mass-consolidation solution, for qualifying workloads.

This new 2008 Software Strategies White Paper reviews the scale-out, distributed server proliferation catastrophe, and looks at why IT infrastructure optimization is now a business imperative. *(Both outlined in Section 2, with depth back up in Appendices C, D, & E) (Section 2)*. We evaluate the role of virtualization, and the varied approaches that have been implemented in server partitioning/virtualization *(Section 3)*. We recap, from our other System z research, why the 2008 mainframe is now the strongest foundation platform as the central hub of optimized IT infrastructures. This includes its newer, mass-distributed server consolidation role. Its already powerful strengths are also soon to be amplified by arrival of the next-generation, new IBM z6 System z MPU-powered mainframes, due out in Q1 2008 *(Section 4)*.

We also assess *(in considerable depth)* how and why the System z Linux-on-z/VM extreme virtualization & mass consolidation of distributed servers solution is now an superior option over palliative, x86/x64 extensions. *(Such as VMware ESX Server, which we compare in detail with z/VM in Section 5, and Appendixes A & B)*.

Finally, we profile/assess 4 major users who have successfully undertaken mass-distributed server consolidations onto Linux-on-z/VM on System z mainframes, covering *(in all)* 4,700 distributed servers consolidating onto 37 System z mainframes, and reporting extremely impressive benefits *(Section 6)*.

Written by: Ian Bramley

Published: 2nd Edition February 2008

Design & Layout: iok design – e-mail: info@iok-design.co.uk

White Paper

Mass Distributed Server Consolidation – System z Mainframe Linux-on-z/VM Extreme Virtualization far Outclasses Over-hyped x86/x64 Approaches

CONTENTS

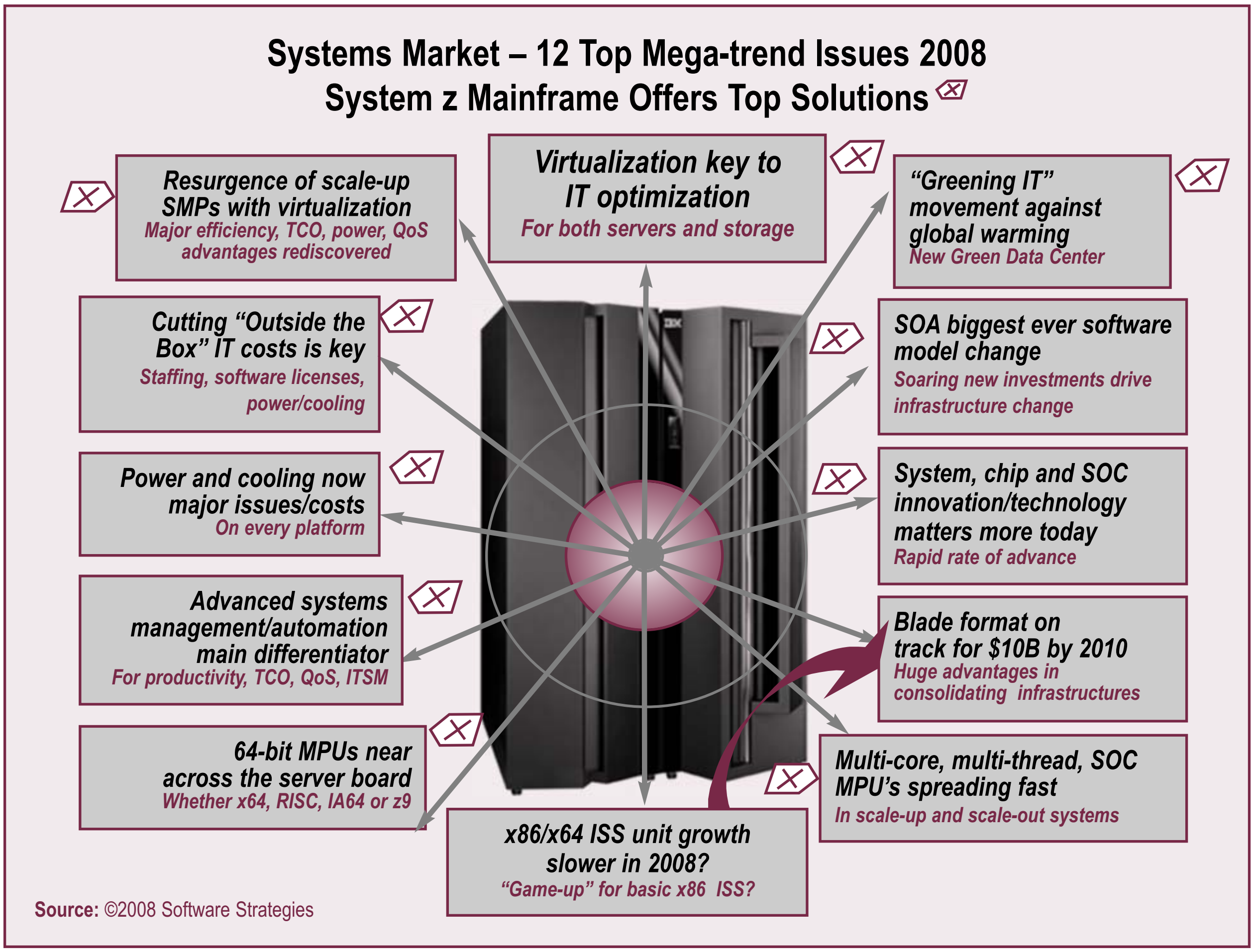
TABLE OF CHARTS

1. Executive Summary

Our White Paper findings/conclusions are highlighted in this Executive Summary, drawn from fuller analysis/discussion in the main sections, links to which are noted below.

1. **Distributed Computing = User Catastrophe:** Distributed servers became an **economic**, **environmental**, and manageability catastrophe for users! Server and software utilization averages 5-7%, one staffer is needed to support every 20-25 distributed servers, and many other severe issues, afflict these sprawls. Their **extreme inefficiency, gross waste of hardware/software capacity, excessive staffing, profligate electrical power/cooling costs, and extravagance with costly data center floorspace**, are now *(belatedly)* clear to all. From green environment/carbon footprint viewpoints, distributed systems were surely an unspeakable abuse. Strong words? Yes, but detailed data analysis in Section 2 and Appendix C amply verifies these truths.

2. **Software Strategies 2008 Systems Market Mega-trends:** With the credit crunch biting, markets falling, and economies now worsening, what are the top dozen early-2008 Systems market mega-trend issues/directions? Figure 1 provides Software Strategies' ranking.



Source: ©2008 Software Strategies

Figure 1: Systems Market –12 Top Mega-trend Issues 2008 – System z Mainframe Offers Top Solutions

IBM's System z mainframe platform, the industry's answer to this distributed server "disease" *(see Appendix C for "epidemic analysis")*, centrally addresses, exploits, or delivers on ten of these twelve 2008 mega-trend issues, and is thus rightly central in this White Paper.

The most powerful "enabling technology" to optimize overly-complex distributed IT infrastructures is virtualization.

3. **Virtualization – Foundation Infrastructure Optimization Technology:** The most powerful "enabling technology" to optimize overly-complex distributed IT infrastructures is virtualization. Thus, virtualization has recently *(and rightly)* grabbed IT industry and media center-stage. Although applicable across all types of IT-based system, this Paper focuses on commercial IT server virtualization.

4. **Server Partitioning & Virtualization:** Sharing host resources over multiple workloads/Operating System *(OS)* images/applications, each isolated in a partition, is the goal. Dynamic partitioning, where partition resources can be changed "on-the-fly", is vital where demands vary rapidly. Two hardware partitioning *(physical and software logical)*, and two hypervisor-based *(Type 1: bare-metal and Type 2: hosted)* software models are seen today. "Type 1 bare-metal hypervisor" solutions *(used for IBM System z's dLPAR & z/VM extreme virtualization hypervisors)* are greatly superior but most difficult to engineer. Of the main ways to implement a hypervisor, "direct hardware support" *(System z)* is much the best, and "translate, trap & emulate" *(VMware ESX)*, needed in x86/x64 virtualization, suffers considerable performance impact.

5. **Virtualization/Server Consolidation Benefits:** Done well, large-scale distributed server consolidation through virtualization delivers five hugely-valuable benefits:

   - **Much-reduced management/support staff costs.**
   - **Much-reduced software costs.**
   - **Sometimes reduced server hardware costs.**
   - **Much-improved flexibility and responsiveness.**
   - **Environment cost savings/green IT benefits.**

   Such benefits *(we found)* are strongest when virtualization champion System z Linux-on-z/VM is used, but are much more limited when restrictive, x86/x64 virtualization software approaches *(VMware ESX Server being the generic)* are used.

6. **IBM System z Mainframe – Virtualization "Gold-Standard":** IBM's System z is today's unquestioned virtualization "Gold-Standard", built via IBM's 40 years of mainframe virtualization pioneering/innovation. Select highlights include:

   - **Dynamic LPAR Partitioning Hypervisor:** Rock-solid, <60 LPARs/system, deep z hardware support, and ultra-secure EAL 5 rated.
   - **z/VM Extreme Virtualization Hypervisor:** Hosts mid-100s to 1,000 significant weight Linux virtual servers/System z.
   - **CPU, Memory, I/O, & Network Capacity:** Flexible, dynamic allocation of all z system resource to LPARs & virtual servers.
   - **HiperSockets™:** High-performance, low-latency, internal System z TCP/IP network links between LPARs.
   - **Intelligent Resource Director** *(IRD)*: SLA-policy-driven, cross-LPAR workload optimizer aids high z utilization and SLA attainment.
   - **Virtual Machine Resource Manager** *(VMRM)*: Manages resources applied to workloads over virtual server groups within z/VM.

IBM's premier "distributed server disaster/disease cure" is its System z mainframe, with the powerful Linux-on-z/VM extreme virtualization/consolidation solution.

7. **System z Linux-on-z/VM Extreme Virtualization:** IBM's premier "distributed server disaster/disease cure" is its System z mainframe, with the powerful Linux-on-z/VM extreme virtualization/consolidation solution. These together, IBM argues, deliver superior distributed server consolidation, for many qualified workload types. Our White Paper evaluates this radical proposition.

8. **Little VMware, Extraordinarily Large Hype:** First-to-market *(2001)* with an x86/x64 server virtualization hypervisor, and later with related Virtual Machine Monitor *(VMM)* management software, was ISV VMware. Resold today by most x86/x64 vendors/channels, VMware became this segment's "generic brand". Although sound and useful, this VMware ESX Server stack suffers limitations/disadvantages *(compared to the System z Linux-on-z/VM approach)*, as well as steep costs. *(Points 3 to 7 are covered in Section 3.)*

9. **Mainframe Transformation + New Workloads = Strong Resurgence:** The IBM mainframe underwent total transformation of its technology, economics, capacity/capability, openness, and in its software stack, from 1993 to date. Modern, new-to-mainframe workloads, including Linux, J2EE™, SOA, ERP, distributed server consolidation, and BI/DW, were resurgence drivers, devouring 70% of post-2001 MIPS sold. Market revival began with 2000's z Series, swelled to resurgence from 2004 *(z990 & z890)*, and was amplified further by the System z9 generation below.

10. **2005-08 System z9 Generation Highly Successful:** The System z9 generation successfully extended the mainframe's now-strong resurgence. First seen as high-end System z9 109s *(September 2005)*, the System z BC and System z9 BC soon followed *(GA July 2006)*, with outstanding virtualization capabilities. Family highlights were:

    - **System z9 Business Class** *(z9 BC)*: For medium business. $100K up, two models at 26-1,786 MIPS capacity.
    - **System z9 Enterprise Class** *(z9 EC)*: For large enterprises. 5 models, up to 54-Way CP, at <17,800 MIPS capacity.
    - **New zIIP** *(System z9 Integrated Information Processor)*: New specialty processor for low-cost DB2, network and XML workloads, on all z9s and z10s.

- **Unique "Hybrid Processor" Architecture:** Includes standard CPs, specialty IFL *(Linux)*, zIIPs *(DB2, networks, XML)*, and zAAPs *(Java and XML)*, mass I/O & channel MPUs, plus crypto MPUs, blended within these "commercial supercomputers", offering huge capacity in a modest footprint.
- **New FICON Express4 4Gbps Channel Speeds:** For improved 2X I/O capacity and performance on z9s.
- **Superb Virtualization/Automated Workload Management**, enabling up to near-100% utilization of capacity.
- **Fullest Capacity on Demand *(CoD)* Options:** For maximum flexibility for customers to rapidly meet varying workload peaks.
- **Leanest, Greenest IT Platform:** Far the greenest solution, smallest data center footprint, lower power and cooling costs per 1,000 users, per virtual server, per workload, etc. *(Versus distributed servers of equivalent capacity.)*

11. **Lowest Cost Platform:** Flatly contradicting old myths, a decade of radical IBM mainframe hardware/software cost reductions, extreme-virtualization-driven efficiency, extensive automation/low staff needs, and excellent "green IT" strengths, now often make System z the lowest cost platform for substantial, mixed commercial workloads today.
12. **Spectacular New IBM z6 MPUs:** IBM recently revealed *(to IT analysts)* its spectacular, new **z6 System z CISC MPU**, that now powers the new System z10 next-generation mainframe, *(Enterprise Class-EC line)* announced February 26th, 2008. These 4.4GHz. + ultra-high-frequency, quad-core z6 MPU chips pack 991M transistors, and use an "extreme-CISC" 894 instruction ISA *(50 new in the z6)*. Together with a huge, new 1.7B transistor, **System z SMP Hub** companion chips, these z6 MPUs have now further extended System z10 capabilities and capacity by wide margins. *(See point 13 below.)*
13. **New IBM z6-MPU-powered System z10 Mainframes Extend Range:** These impressive z6 chips have enabled IBM's new, next-generation System z10 mainframes to offer higher significantly *(+70%)* top-end capacity *(the new System z10 Enterprise Class (EC) just announced)*, further extending their mass consolidation capabilities. IBM will also *(we expect)* offer smaller, denser, and more economical, z6-powered z10 mainframe systems around the z9 BC capacity points *(greater value, yet-to-be-announced, z10 Business Class mid-range system)*.
14. **System z – A "Data-center-in-a-box":** System power and extreme virtualization enable mass consolidation of *(qualified)* distributed server workloads onto this pre-integrated, pre-tested, complete System z "Data-center-in-a-box", that offers:
    - **Hundreds of Processors of Power:** Up to 416 total in a top-end System z9.
    - **Extreme I/O Bandwidth:** And ultra-fast context switching.
    - **Built-in Internal Networking:** High-performance, memory-speed, internal virtual networking, switching, & communications.
    - **"Shared Everything" Resource Model:** Enables favorable Total Cost of Ownership *(TCO)*, increasing returns to scale.
    - **Product of $B Engineering & Software Development:** Sustained, decade-long, massive $Bs IBM R&D effort.
    - **Highest Reliability, Availability, Security:** Of any commercial IT platform.
    - **Lowest Staffing, Software, Space, Power, & Cooling Costs:** Combine to deliver lowest TCO when heavily loaded.

A centerpiece of System z's "Gold-Standard" virtualization is its z/VM V5.3 extreme virtualization hypervisor...

15. **IBM z/VM Extreme Virtualization Hypervisor:** A centerpiece of System z's "Gold-Standard" virtualization is its z/VM V5.3 extreme virtualization hypervisor, latest release of the industry's most sophisticated, scalable, and long-established, bare-metal hypervisor. Now ideal for mass *(hundreds to a thousand)* distributed server workload consolidation onto Linux-on-z/VM, the hypervisor has long also been used by IT groups to host/test multiple OSs, releases, etc., on a single mainframe.
16. **IBM Systems Director Vital Part of Solution:** IBM Systems Director, now a highly-capable, elegant, and comprehensive systems manager for physical and virtual platforms, manages servers, storage, and networking equipment, across all IBM server and storage systems, and major third-party platforms. It complements z/VM on System z with its z/VM Center component.
17. **Freedom From Distributed Computing Scourge With System z:** c. 28M distributed x86/x64-UNIX servers are running today, some 50% *(14M)* on workloads better consolidated onto System z mainframes. Just 14,000 new, IBM z6-powered large mainframes could, we estimate, take over/consolidate all 14M distributed server workloads *(@ c. 1,000 virtual Linux servers/System z)*. This would free users of half of today's distributed computing scourge, as well as saving these users scores of $Bs of costs over the next decade.

18. **System z IFL Specialty Processor for Linux Enabler:** Underpinning soaring Linux workload growth on System z was IBM's Integrated Facility for Linux *(IFL)* specialty processors *(2000 debut)*. These run Linux workloads only, have no software-cost impact, and are currently priced far below standard z9 CP *(@ $95K for z9 BC & $125K for z9 EC respectively)*.

19. **System z All-IFL, Linux-only Mainframes Enabler:** All-IFL Linux-only System z mainframes *(also 2000 debut)* leverage highly favorable IFL pricing, and eliminate regular System z software charging, to provide a highly-competitive Linux hardware/OS platform. Ideal for mass distributed server to Linux-on-z/VM virtual server consolidations, these machines can scale to the same, full, 54-Way, 17,800 MIPS *(System z9 EC)* capacity as a regular mainframe, and can each support up to several hundred mid-weight Linux virtual servers. *(And now far higher with the new z6-MPU powered, System z10 generation of mainframe systems.)*

20. **System z Linux – Success far Above Expectations:** Since 2000, Linux on System z far exceeded expectations, with 2,500+ mainframes already running the OS *(16%+)*. Usually run in a LPAR or under z/VM, Linux has today become an enterprise-grade OS, complementing System z's unique hardware Quality of Service *(QoS)*. It brought many hundreds of new, originally UNIX, Linux, or Windows, software applications, middleware and tools, software products onto System z. Combined with z/VM *(point 15 above)*, IFLs *(point 18 above)*, and All-IFL Linux mainframe hardware *(point 19 above)*, Linux provides the perfect OS foundation for mass distributed server consolidation. *(Points 8 to 20 are covered in Section 4.)*

21. **z/VM Major Features and Capabilities:** In Section 7, and specifically around the following charts:

    - **Figure 15:** Major z/VM V5.3 Hypervisor Features & Capabilities. *(Page 42.)*
    - **Figure 16:** z/VM Extreme Virtualization Leadership – Other Features Summary. *(Page 43.)*
    - **Figure 17:** z/VM V5.3 – Latest Release Advances. *(Page 44.)*
    - **Figure 18:** z/VM Command & Control Infrastructure – Taking Advantage of IBM Software z/VM Portfolio. *(Page 45.)*
    - **Figure 19:** IBM System z z/VM Hypervisor Environment – Advanced, Comprehensive Management Software Suite. *(Page 47.)*

    we summarized and assessed the extensive capabilities under these headings offered by z/VM V5.3. We cannot usefully further abridge those findings here, but comment generally on the tremendous depth and breadth of capabilities the hypervisor offers.

22. **IBM z/VM V5.3 vs. VMware ESX Server V3.1:** In Figure 22 *(and deeply in Appendix B)* we publish a detailed comparison, examining each of these leading virtualization platforms, from the perspective of our distributed server consolidation interest here. One analogy neatly summarizes our comparison results. The Linux-on-z/VM System z solution compares to the VMware ESX Server solutions on x86/x64 like a **Boeing 747 Jumbo jet compares to a fleet of 100 4-passenger Cessna twins** as a means of air passenger transport. Both means, in theory, could be used to transport many people long distance. Where the actual mission is to carry 400 passengers, at the same time, on a 6,000- mile leg over water, economically, rapidly, and safely, the 747 wins easily. It alone has the capacity *(400 seats)*, range *(8,000 miles)*, and far the lowest staffing *(16 vs. at least 100)*. It also pays only one set of take-off/landing fees *(1 vs. 100)*, can fly in most weathers, has an outstanding safety/security record on such flights, and delivers the lowest mission fuel cost, and total cost, per seat mile.

    z/VM on System z *(the 747 in our analogy)* is the heaviest-duty, most industrial-strength, highest-productivity, lowest-unit-cost, longest-perfected, and most trusted extreme virtualization hypervisor in the IT industry, and the only one today capable of mass distributed server consolidation! The Cessna machines *(=VMware ESX Server)* are excellent light planes, designed/optimized for shorter leisure or business flights, usually within one nation, and of a few hundred miles length, for 2-4 passengers only, and in fair weather. *(Points 21 to 22 are covered in Section 5 and Appendix B.)*

23. **4,700 Distributed Servers *(DS)* to 37 System z Mainframes:** 4,700 is the industrial-scale number of distributed servers consolidated onto 37 new System z mainframes, at the four actual mainframe user organizations listed below *(we report/assess on their mass-consolidation successes fully in Section 6.)*:

    - **# 1. IBM Project Big Green – 3,900 DS onto 33 System z, in work**
    - **# 2. Nationwide – 500 DS onto 2 System z to date.**
    - **# 3. Government of Quebec – 200 DS onto 1 System z, so far.**
    - **# 4. Nexxar Group – 80 x86 DS onto 1 System z9 BC, so far.**

24. **Dramatic Benefits for System z Consolidation Users:** Impressively large, and consistent, System z Linux-on-z/VM benefits were delivered, according to these users' distributed server consolidation experiences, and other data. These include:

Impressively large, and consistent, System z Linux-on-z/VM benefits were delivered...

- **Reducing server numbers by 100-200: 1 System z, or more.**
- **Raising CPU/hardware utilization, 10-20: 1 times higher, up to 85-95% on System z from 5% on x86/x64.**
- **Supporting far more virtual servers/CPU core, by 15-30: 1 System z IFL core.**
- **Far fewer CPU cores needed, by 10-20: 1 System z IFL core factor.**
- **10-20: 1-fold fewer CPUs worth of SW licenses to buy, 4-8: 1 lower software cost.**
- **4-8:1 fewer support staff needed, and the same ratios of lower support staff cost.**
- **1 support FTE/100-200 System z Linux virtual servers, vs. 20-25 physical distributed servers/FTE.**
- **Hardware costs actually now often comparable.**
- **Dramatic power & cooling cost savings, 10-20:1-fold less cost, often 90%+ reduction.**
- **Data-centre floorspace, often 25:1-fold less – e.g. slashed from 10,000-sq. ft. *(DS)* to 400-sq. ft. *(System z)*.**
- **Networking equipment/cabling cost savings, 3-4: 1 average reduction.**
- **Far faster server provisioning, 15-1,000: 1 times faster, seconds/minutes not days/weeks.**
- **Enables technically-feasible, affordable Disaster Recovery/Business Continuity *(DR/BC)*, impractical with distributed servers.**
- **30-50% overall TCO savings consistently reported.**
- **Rapid ROIs with paybacks in the15 months to 2.5 years range.**

These are dramatic major benefits and savings. They again highlight just how much wastage distributed systems incur, and how much more cost-effective a System z mainframe mass-consolidation replacement will normally be, for qualified workloads. *(Points 23 and 24 are covered in Section 6.)*

2. The Distributed Computing Disaster – Time To Roll It Back!

Distributed Servers – A "Population Explosion"

...distributed servers – IT's highly addictive, "crack-cocaine-equivalent" drug...

Live global servers jumped more than five-fold from 6M *(1996)* to a stunning 32M *(2007)*, mostly distributed servers. In all, 7.5M new servers were, including 6.9M x86, sold, with c. 4.5M retired, during 2006. This sprawling servers flood had dire effects, see below. Blame the RISC-UNIX and x86/x64 cartel vendors for hyping/pushing these distributed servers – IT's highly addictive, "crack-cocaine-equivalent" drug– to deeply-hooked IT users.

Who Should Read This White Paper?

This White Paper was written for CIOs, CTOs, and senior IT/DP managers in enterprises *(large/medium)* who now need to optimize and consolidate their sprawling, scale-out, distributed computing infrastructures *(including large numbers of x86/x64 servers)* that their enterprises deployed incrementally since the mid-1990s.

IT System Management, Power/Cooling Costs Skyrocket

Soaring distributed server units drove IT management costs skywards *(software, staffing support, and data center)* up 6-fold from <$22B *(1996)* to an unsustainable $120B *(2007)*; and still climbing fast. IT electrical power/cooling costs also shot up from <$10B *(1996)* to $22B *(2007)*, also still growing fast. Combined, these IT system running costs were triple that of global 2007 new server spending!

These dramatic server population and associated IT operating and power costs, are shown in Figure 2 and analyzed below.

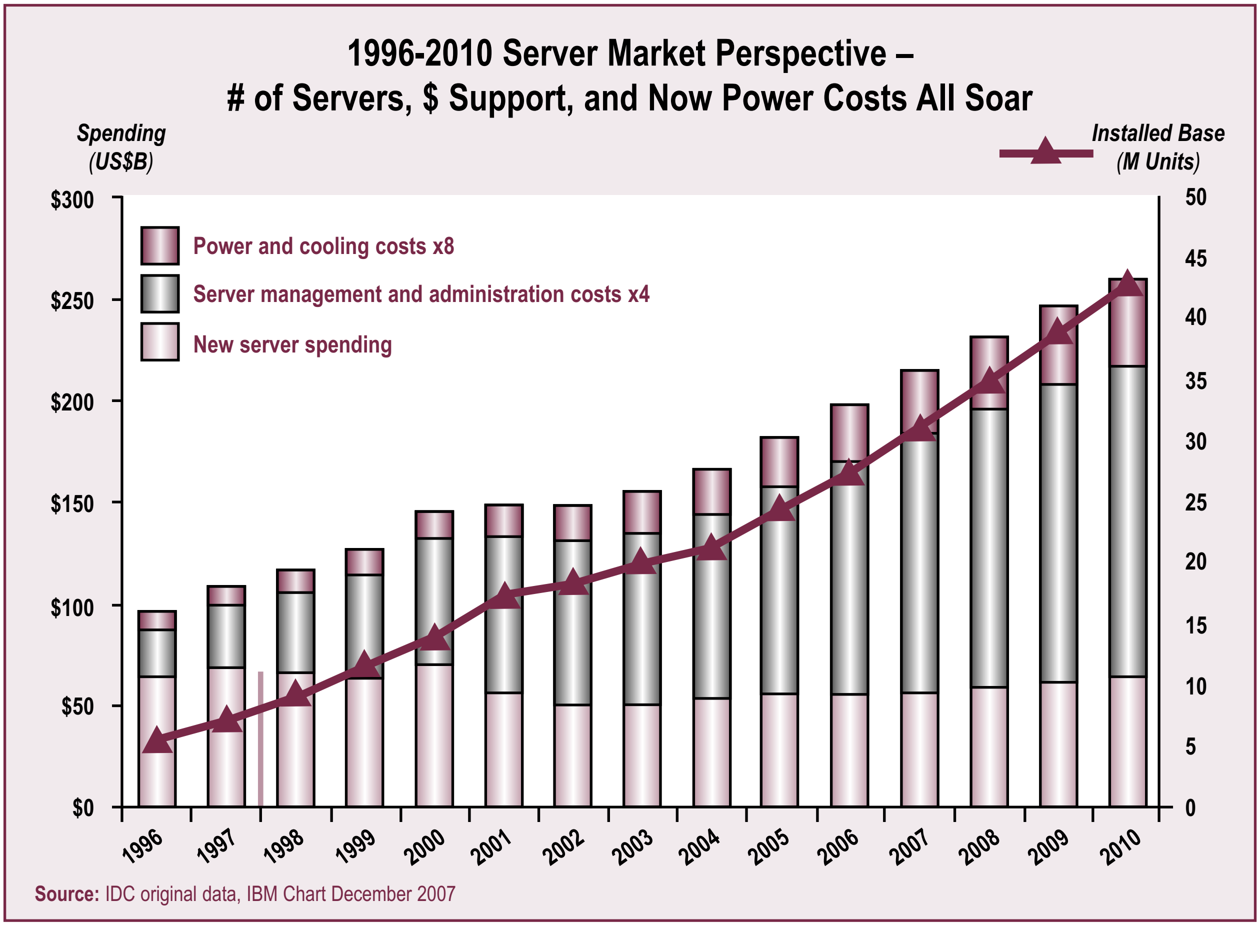


Figure 2: 1996-2010 Server Market Perspective – # of Servers, $ Support, and Now Power Costs, All Soar

This chart shows:

- **Huge Server Population Growth:** Worldwide net servers *(after retirements)* shot up from **6M in 1996** to 35M by 2008, and will hit **45M by 2010**. *(Red line/triangle symbols, RHS vertical axis)*. **7.5M new server units** were shipped into the market *(2006)*, **6.9M** *(92%)* **being x86 type**.
- **4.5M Servers/p.a. Discarded:** Today, 4.5M servers *(we estimate)* fall obsolete and are discarded yearly, overwhelmingly scale-out distributed type. These have shortest service lives at just three years. Such a **"throwaway/disposable servers"** approach has today become intolerably wasteful of our earth's scarce resources, and almost economic madness. But huge replacement business for their vendors!
- **New Server Spending Flat, Capacity Way Up:** For 1996-2008, server spend *(light-pink/lower bars, LHS vertical axis)* was within the $49-$69Bp.a. range. Strong server/storage performance/capacity and price/performance gains of 20%+ p.a. were seen. These near-flat, new server spends thus delivered much-increased new server capacity, and drove hardware costs per transaction/unit of work/user down sharply, each year.
- **System Management & Support Costs Exploded:** Users' System Management, Support, and Administration *(SMSA)* costs *(Grey middle bar, LHS vertical axis)* exploded. These costs *(staff support, software license, and data center space included)* jumped **from $22B in 1996** *(<33% server $B spend)* to **$140B in 2008**, and a projected **$151B in 2010** *(c 2.5X new server $B spend)*. This 7-fold increase from 1996 to 2010 is an unsustainable cost growth that cannot continue.

This 7-fold increase from 1996 to 2010 is an unsustainable cost growth that cannot continue.

- **Power & Cooling Costs Also Rocketing:** Soaring servers numbers *(at denser, higher-frequency)*, sent electrical power/cooling costs skyrocketing. Figure 2 *(dark-pink/top bars and LHS axis)* shows these **soared five-fold** from c $9.6B in 1996 *(<15% of new server $B)* to a projected c. $44.6B by 2010 *(73% of new server $B)*. These jumping power and cooling costs became a major financially burden. But worse, in many areas, it is now impossible to obtain additional, substantial electrical power supplies at all, with utilities unable to deliver. Many data centers outran their electrical power and cooling capacity *(often also out of spare floorspace too)*, and thus now face major $1M-$10M data center expansion investments if they continue on past policies/technologies.

The last decade's wild proliferation of scale-out distributed computing *(majority x86/x64)* drove the above huge increases in IT operating costs users'. Supporting, managing and administering such overly complex IT infrastructures of thousands of scale-out servers, and their associated storage, networks and applications, and of powering and cooling them, became astronomically costly. These high, annually recurring and ongoing, support and power costs combined now far outweigh global customer expenditure on new server systems themselves.

Distributed Computing Debacle – $100Bs Wasted

The distributed server vendors sold c. 50M x86/x64-UNIX servers in 10 years, plus all their software. Yet 90-95% of this hardware/software capacity was completely wasted and never used. Users could not productively benefit from more than a tiny *(5-10%)* part. Yet IT operating costs grew vastly. An astonishingly profitable business for IT vendors, but $100Bs were wasted by "x86/x64-UNIX distributed, server-addicted" IT users!

Yet 90-95% of this hardware/software capacity was completely wasted and never used.

We summarize the primary problems of distributed, scale-out computing with x86/x64 and/or UNIX volume servers in our main Figure 3 chart below, highlighting the impact of each boxed problem area in the text note beneath.

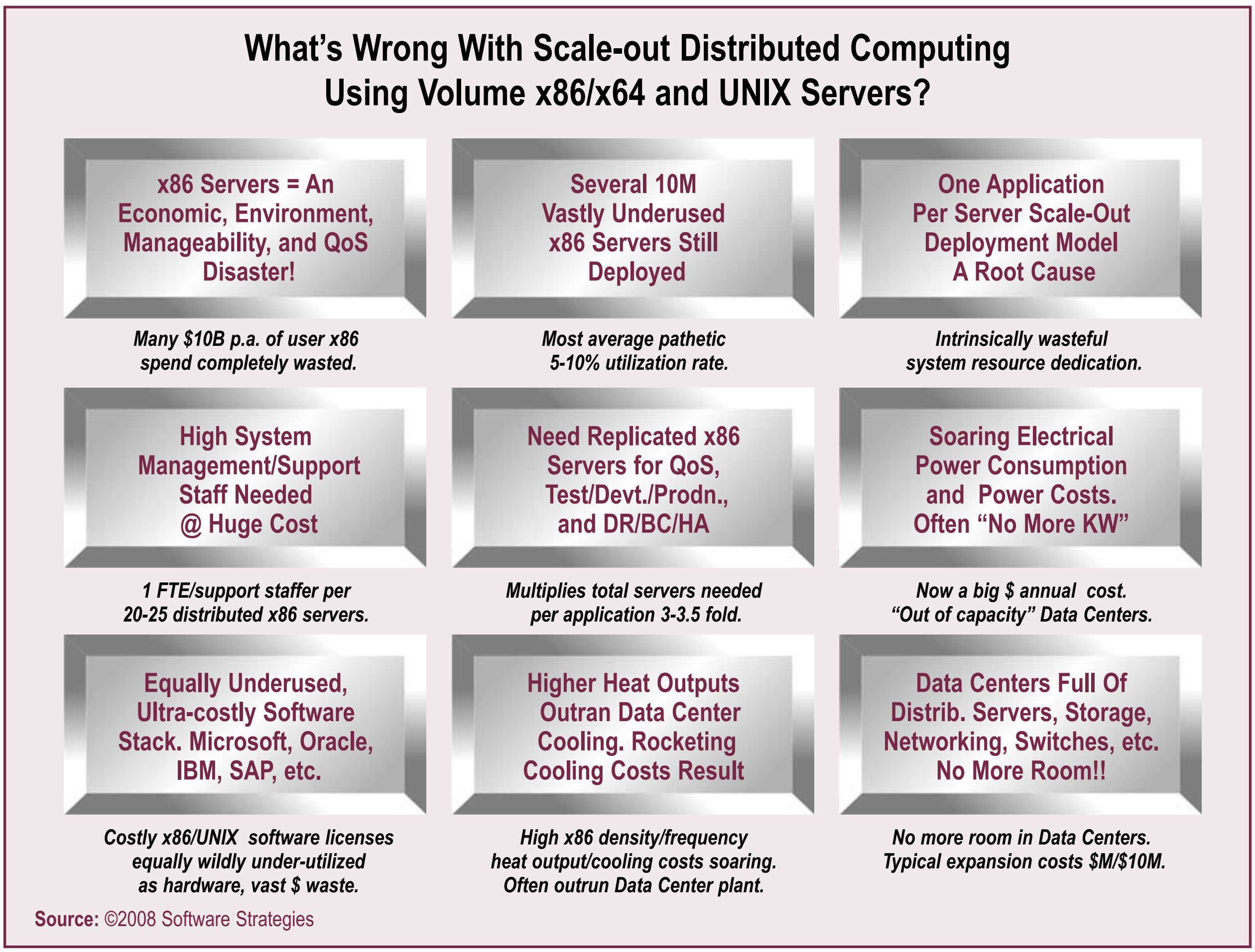


Figure 3: What's Wrong With Scale-out Distributed Computing Using Volume x86 *(and UNIX)* Servers?

...a dramatically different, highly-compelling solution, to these users' distributed servers/x86 nightmares.

This chart summarizes our much more detailed analysis of these distributed computing problems, which is presented fully in Appendix C from page 67.

In this new White Paper, Software Strategies evaluates a dramatically different, highly-compelling solution, to these users' distributed servers/x86 nightmares. What justifies such strong words? Why has this topic become such a central IT issue worldwide?

Radical Distributed Server Consolidation Imperative

With their myriad serious issues, IT user infrastructures afflicted by this "scale-out distributed computing addiction/disease" now imperatively need radical, rapid consolidation and optimization at thousands of enterprises globally.

Radical System z z/VM Workload Consolidation Holds Solution

Whilst the loud voices of "Wintel" major vested interests are currently promoting **"new x86/x64 miracle cures"** to the wide-ranging issues caused by such distributed server sprawl, this White Paper closely examines a rapidly-growing alternative that offers far more compelling benefits. This alternative is **extreme virtualization and mass distributed system workload consolidation** using Linux-on-z/VM *(hypervisor)* on **IBM's legendary System z mainframe** platform

IT Optimization a Continuing Process – Now Accelerating

Distributed system nightmares made IT infrastructure optimization the "number-one job" for most IT groups today. This is the latest, largest step in a decade-long, continuing IT infrastructure improvement process. Four prior consolidation waves predated our current distributed systems focus, and two other, concurrent waves are now often implemented alongside. These, in date order, were/are:

- **Network Standardization & Consolidation.**
- **Data Center Consolidations.**
- **Large Systems/Server Consolidations.**
- **Distributed Servers Physical Consolidation to Server Farms.**
- **Distributed Server Consolidation/Virtualization.**
- **Enterprise Storage Consolidation.**
- **Enterprise Application Instance Consolidation.**

Caution – Some "Consolidation Advisors" Not Trustworthy

x86/x64-UNIX vendors, channels, partners, and their media cheerleaders, are blameworthy for users' overly complex, wasteful, distributed systems-plagued IT estates. These firms sold the distributed "drugs", took shares of $100Bs in users' cash for a decade, but left those addicted users to wrestle with sprawling, wasteful server estates. Today, these self-same suppliers now propose new "x86/x64 miracle cures". Consolidation advice from such deeply culpable suppliers, whose current self-interests are once-again all too clear, cannot and should not be trusted.

Beware Siren "x86/x64-Technology Miracle Panacea" Calls

Now in 2008, the same shameless vendors tout "new, exciting, x86/x64 panacea" solutions. These, they now pledge, will solve the severe user problems that their own, earlier, distributed server sales created for thousands of their customers! The four main new x86/x64 "snake oil flavors" are:

- **x86/x64 Server Virtualization Software.**
- **AMD-V/IDV Virtualization-supportive x64 Servers.**
- **Blade Servers, Mainly x86/x64, and some RISC.**
- **Larger Scale-up x86/x64 Servers with Partitioning.**

These x86/x64 technologies are unquestionably useful, and can be beneficial in some cases. However, we consider more radical "distributed systems disease/disaster" replacement solutions should be compared/chosen, before users throw yet more x86/x64 palliative investments into this bottomless, distributed systems money pit.

Balanced Server Consolidation Advice? – Trust IBM

IBM sells all server types. It also gained wide and deep IT consolidation experience via 1,000s of customer infrastructure engagements in this decade. IBM has just completely reorganized and now offers balanced, customer-best-interest-focused, platform-neutral, IT infrastructure optimization advice, design, and implementation services. Long the most trusted enterprise IT vendor, IBM is a natural choice to advise on best approaches, platforms to use, and virtualization to apply, for most users.

Our Analysis

Scale-out distributed servers are like the **"crack-cocaine of the IT systems business"**, highly addictive and extremely damaging. Once an enterprise got **"hooked on this potent drug"**, it proved extremely hard to wean the "corporate addicts" off. Despite devastating consequences, they just reached for their next x86/x64-UNIX servers" fix for every new workload that arose. **x86/x64-UNIX server addiction proved as desperately bad** indeed for the "health, wealth and welfare" of enterprise IT users *(as detailed above)*, just as crack-cocaine is to its unfortunate human users, their families, and their communities.

Collectively, the "distributed systems" ecosystem vendors sold over 50M distributed x86/x64 & UNIX volume servers in the past decade, plus all their associated storage, OS, middleware, and application software *(the latter often costing far more than server hardware)*. Yet the **vast majority *(90-5%)*** of the capacity of this vast amount of sold hardware/software **was completely wasted**, because it could never be used productively by the customers.

The scale and depth of issues arising from complex today's distributed infrastructures clearly require wide-ranging further IT infrastructure optimization and consolidation at most major enterprise IT users, and on a more modest scale, at many larger SMB firms.

The economic and financial reasons to quickly eliminate the massive wastage of costly hardware and software, support staffing effort, data center space, and power and cooling energy costs, combine to make a hugely-compelling case for consolidation and optimization. IT energy consumption absorbs 2% of global electricity production so that enterprise IT energy reductions, through drastic IT consolidation, are a major and vital part of corporate "green initiatives" to reduce enterprise carbon footprints and help limit global warming in many industries.

3. IT Virtualization – Main Technology for IT Optimization

Virtualization Introduction

One foundation technology, far beyond all others, has today rightly taken IT industry center stage in the important, continuing battle to optimize today's overly complex IT infrastructures. **That technology is virtualization**.

Most of the important IT virtualization concepts, break-though developments, and first practical exploitations, **came on/from the IBM mainframe *(System z today)*** over its 44-year history *(1964-2008)* of success and technology innovation. The result is that today's System z mainframe is the absolute and unquestioned "Gold-Standard" leader most deeply exploiting myriad virtualization technologies, crucial foundations for the amazing capabilities these systems deliver today. *(Overviewed below and in Sections 4 & 5.)*

Virtualization has now become a broad topic, has already been applied to most types of IT systems/subsystems, and is now widely used across the whole industry today.

At its most basic, virtualization provides an **abstraction layer** that enables applications/workloads/users to draw upon **virtual IT resources**, instead of directly accessing **physical system resources**. *(For example, CPU, memory, I/O, networks in server systems. Storage blocks, paths, or volumes, in storage systems, for example.)* These virtual resources become proxies for the underlying real resources, offering architected interfaces and functions that are the same as their physical resource counterparts.

Virtualization cleanly/sharply separates the presentation of IT resources to their users from the actual physical resources themselves. It conversely aggregates pools of resources from multiple, smaller physical sources, enabling their re-allocation to users as one unified virtual resource. So virtual resources may deliver **smaller slices of a larger system's whole resources** to their users ***(partitioning or time-sharing)***. Conversely, such virtual resources may pool or **combine physical resources from several smaller systems** into one larger, logical resource to meet larger capacity needs that one individual source cannot provide ***(clustering)***.

In servers, virtualization gives rise to the capability of **Virtual Machines *(VMs)*** software implementations of a machine *(computer hardware)* that executes allowed programs just like a real hardware machine. The term **"virtual server" *(VS)*** is often used synonymously with VM. A **system VM** provides a complete system platform that supports the execution of a complete OS, akin to how a physical hardware system runs it. Such virtualization therefore multiplexes/shares the underlying physical machine across several, or many, different VMs, each running its own OS copy/type. *(Process VMs are the other main VM class – for example, a Java Virtual Machine (JVM) that just runs the Java process.)*

Server Virtualization – Our White Paper Focus

Whilst virtualization has far-reaching applications in every type of IT system and subsystem, in this White Paper we focus only on server virtualization, because this is the main technology for the IT infrastructure optimization/consolidation of sprawling, scale-out distributed systems that are our central concern.

Storage virtualization, networking systems virtualization, and desktop systems virtualization are also other extremely important, high-benefit areas of application for virtualization technology, but we do not have the space to address these here.

Server Partitioning and Server Virtualization

The key to improved efficiency and fuller usage of server hardware resources is a server system's ability to host and run several or many different workloads, applications, and/or OSs, simultaneously on the one platform, and to efficiently share all its resources between them. Server partitioning and server virtualization, separately and/or in combination, are the related core technologies that enable such crucially, more-efficient use of server hardware resources.

At the bottom extreme on the server sophistication metric spectrum stand traditional, scale-out, distributed, small servers *(typically x86-based)* normally dedicated to a single application/workload, and running one OS copy, in order to avoid any risk of interference with any other application. Such servers traditionally incorporated no hardware, nor standard OS support, for partitioning/server virtualization.

Each System z can support up to 60 dynamically-changeable logical partitions, can also run from 100s to a 1,000 Linux virtual servers of significant weight...

At the other extreme, **the highest pinnacle of partitioning and virtualization** is today's **IBM's System z mainframe**. These scale up to massive 54/64-Way SMP "data-center-in-a-box" system sizes, and up to monster <32-system clusters. Each System z can support up to 60 dynamically-changeable logical partitions, can also run from 100s to a 1,000 Linux virtual servers of significant weight, hosts multiple OS, can easily runs scores of applications, and host tens of thousands of users, all simultaneously, securely and efficiently. All System z9 resources are completely virtualized.

There are four main server partitioning/virtualization approaches in widespread use today, two different forms of server hardware partitioning, and two different, hypervisor-based approaches. A **hypervisor** is a virtualization platform *(thin layer of software and/or firmware)* that allows multiple OSs, and/or multiple virtual servers or VMs, to run on a physical host server computer simultaneously. Hypervisors are also sometimes termed **VMMs *(virtual machine monitors)***.

Server Partitioning/Virtualization Approaches Compared

Factor	Server Hardware Partitioning		Bare-Metal Hypervisor *("Type 1")*	Hosted Hypervisor *("Type 2")*
Basic Operation	Server is subdivided into partitions, each of which can run an OS independently, by some "partition manager" functionality. Two very distinct types.		Hypervisor provides fine-grained timesharing of all server physical resources across multiple virtual servers.	Hypervisor uses host OS services to do timesharing of all server physical resources across virtual servers.
	Physical Hardware Partitioning	**Software Logical Partitioning**		
Description	Server coarsely physically divided, along physical, hardware module boundaries.	Server more flexibility divided into software logical/virtual partitions, of more granular sizes.	Hypervisor software/ firmware runs directly on server hardware *("bare metal")*.	Hypervisor software runs indirectly on the host operating system, not on the server hardware itself.
Well Known Examples	• **HP nPartitions** • **Sun Domains**	• **IBM pSeries LPAR** • **HP p-Partitions**	• **System z LPAR** • **IBM POWER™ Hypervisor** • **VMware ESX Server** *(x86)* • **Xen Hypervisor** *(x86)*	• **VMware Server** *(Was GSX)* • **Microsoft® Virtual Server** • **HP Integrity VM** • **User Mode Linux®**
Main Advantages	Good physical partition isolation. Can support a few larger static workloads/images on one server.	More granular and flexible than physical hardware partitioning. Can be linked to WLM, VMM. Some may offer dynamic change.	Offers highest efficiency and VM availability. Most highly flexible and granular, may offer dynamic LPAR. If co-optimized with hardware support, has almost no performance overhead.	Useful for clients/guests where host OS integration is important.
Main Disadvantages	Limited to physical sub-system boundary configurations. Inflexible, coarse grained, static.	Partitioning software manager processing overhead.	Most difficult to build, best with hardware virtualization support built in.	Reliant upon host OS RAS, and incurs host OS processing overhead.

IBM's z/VM for System z is also a Type 1 bare metal hypervisor, but is unique, completely different type of solution, assessed fully in Section 5

Figure 4: Server Virtualization Approaches Compared

These four main approaches to server partitioning/virtualization are outlined, are exemplified by the leading products using each approach, and compared according to their main advantages and disadvantages, in Figure 4, and discussed briefly below.

- **Physical Partitioning:** Many larger SMP server systems *(>8-Way SMP)* offer at least a physical partitioning capability. This enables a physical server to be divided into a limited number of coarse-grained *(generally in multiples of 4-Way size)* physical partitions *(PPARs)*. These partition divisions lie along physical SMP sub-module boundaries *(cell, CPU board, or module)*. Good physical isolation is achieved on well-designed PPAR implementations, sometimes avoiding any Single Point Of Failure *(SPOF)*. In such systems, failures in one partition cannot affect an OS or application running in another PPAR. However, granularity, flexibility, and adaptability to changing workload needs, are all poor.
- **Software Logical Partitioning:** Partition management software divides the whole server into a number of virtual servers, VMs, or software logical partitions *(LPARs)*, which generally may be more granular and flexible in size than physical partitioning. Software LPAR generally offers less rigorous partition isolation than PPAR. Earlier implementations were mostly static LPAR. The processing overhead and performance impact of the partition software manager can be significant, and is the main disadvantage of this model.
- **"Bare-Metal" Hypervisors:** Often known as Type 1 hypervisors. Now generally agreed to be the best approach to server partitioning and virtualization, but the most difficult to engineer. This method also works best when combined with built-in hardware virtualization support. Then it is the highest-performing, lowest performance-overhead, most flexible, the most dynamically re-configurable, and the most amenable to programmatic control by VMM and WLM software, approach. The longest-established, most-sophisticated, and most-heavily-used in production of this class of hypervisor are the IBM System z mainframe's two bare-metal hypervisors – z/VM and *(d)*LPAR – supported by System z's deeply built-in hardware virtualization support. The two leading x86/x64 virtualizers – VMware's ESX Server and the more recent, open source Xen *(offered by several vendors)* – are also both "bare-metal", Type 1 hypervisors.
- **Hosted Hypervisors:** Often known as Type 2 hypervisors. In this model, virtual servers/VMs are hosted inside a host OS, and use OS services to access the underlying hardware resources. Where tight integration between guest OS/applications and the host OS is needed/beneficial, this approach is useful. Two well-known implementations of this type are VMware's now-free VMware Server *(was GSX Server)*, and Microsoft's free/low-cost Microsoft Virtual Server 2005. Each of these provides an accessible, basic, low-cost entry point for users to first learn about server virtualization, and conduct small Proofs Of Concept *(POCs)*.

This analysis clearly points to the benefits of the hypervisor approach in general, and to the "bare-metal", Type 1 hypervisor in particular, as the best fundamental approach wherever possible.

Static vs. Dynamic Partitioning & Virtualization

The ability to physically partition a server into several smaller partitions that can run multiple workloads, even with the most basic PPAR physical server partitioning, is a useful capability, one that carries no performance overhead, provides excellent partition isolation, but will not usually ensure the highest hardware utilization. PPAR does allow a useful degree of valuable workloads consolidation, and permits multiple OSs be run on a single server, albeit with the partitions relatively coarse-grained, fixed in size to physical system boundaries, and with inflexible, static machine resource allocations to partitions. In this model, any changes to a partition configuration, or its OS, requires that partition to be taken down offline, the configuration changes made, and the partition then restarted, interrupting that partition's service delivery.

Software logical partitioning provides much more granular and flexible sizing of partitions, generally allowing a much wider range of resource allocations to software logical partitions/virtual servers than with PPAR, albeit with a performance overhead for running the partition management software.

The most sophisticated and attractive form of partitioning is **Dynamic Logical Partitioning** ***(dLPAR)***. This is the capability for a dLPAR to be reconfigured dynamically, without shutting it down *(thus enabling service continuity)*, whilst offering the most extensive flexibility in partition/virtual server resource allocation. The pioneering, and most sophisticated dLPAR implementation, is the IBM System z mainframe's LPAR capability, discussed in more detail under System z below.

Over this decade, IBM also added mainframe-inspired dLPAR technology *(called Advanced POWER Virtualization (APV))* into both its POWER-based server families, the System p line *(AIX & Linux OS)* and System i line *(i5/OS, Linux, AIX OS)*. IBM's POWER5-processor-powered versions of these families added micro-partitioning, allowing up to ten LPARs to be configured per processor, so that one IBM multiprocessor POWER5 or POWER5+ onwards *(now POWER6)* server could support a maximum of 254 dLPARS. IBM could and did build hardware virtualization support for APV deep into these chips/systems. APV itself is also a Type 1 "bare-metal" hypervisor, implemented with the "para-virtualization" approach *(see Figure 5 below)*. This was possible here because IBM could and did arrange for all the OSs used on these platforms to be modified to support APV hypervisor calls, making them high-performing and flexible virtualization platforms. *(See the subsection below.)*

Hypervisor Implementation Methods Assessed

Four main methods have been used to date to implement server virtualization hypervisors, and these are explained, exemplified, and compared, in Figure 7 below. Understanding each of these options, and the substantial trade-offs some require, is key to understanding which of actual current virtualization solutions are superior.

Server Virtualization – Hypervisor Implementation Methods Compared				
Hypervisor Implementation Method	**Trap and Emulate**	**Translate, Trap and Emulate**	**Hypervisor Calls *("Para-virtualization")***	**Direct Hardware Virtualization Support**
How Method Works:	The VM runs in user mode. All privileged instructions cause traps.	The VM runs in user mode. Some processor *(e.g. IA-32)* instructions must be replaced with trap operations.	The VM guest OS in the VM calls the hypervisor to access real resources.	The VM runs in normal modes. The hardware does most of the virtualization *(e.g. System z SIE architecture)*. Hypervisor provides control.
Hypervisor Examples:	**IBM CP-67 and VM/370 timesharing OS.**	**VMware ESX Server 3, MS Virtual Server.**	**IBM POWER Hypervisor, Xen Hypervisor.**	**IBM System z dLPAR, IBM System z z/VM.**
Operating System:	Runs unmodified standard OSs.	Runs unmodified, but translated, OSs.	OSs must be modified to issue hypervisor calls *(H calls)*.	Runs unmodified standard OSs.
Main Benefits:	No OS modification and no hardware support needed.	Enables virtualization of "very difficult" ISA CPU systems, e.g. IA-32-x86.	High efficiency, good performance. Can exploit HW virtualization support if any available.	Highest performance, lowest or zero virtualization overhead. Lends itself best to programmatic management, integration.
Disadvantages:	Substantial resource overhead to trap PIs, performance impact.	Substantial resource overhead to translate and emulate, has performance impact.	Need to change operating systems code.	Requires built-in hardware support from system design outset.

Figure 5: Server Virtualization – Hypervisor Implementation Methods Compared

In overall preference order, these methods are:

- **Direct Hardware Virtualization Support:** Provides lowest overhead, highest performance, requires no OS modification and can deliver the most dynamic partitioning, but requires hardware with deeply-embedded, chip/system-level virtualization support, best coupled with a refined, sophisticated Type 1 "bare-metal" hypervisor. IBM System z LPAR and z/VM are the best-known, longest-proven examples.
- **Hypervisor Calls *("Para-virtualization")*:** Requires modification to each OS supported to issue hypervisor calls, can operate on any supported hardware, offers modest overhead and performance impact, and can exploit any available hardware virtualization support. IBM APV and Xen are the best known implementations of this model. It can clearly only be used if the OS concerned is accessible for modification to support the specific hypervisor, and/or to add specific hardware virtualization support exploitation. IBM can naturally do both with its APV on System p and System i, and the Xen vendors co-operate closely with the Linux community to ensure Linux kernel Xen support.
- **Translate, Trap, and Emulate:** This approach must be used with complex, virtualization-unfriendly system ISAs, where some CPU instructions must be trapped and emulated by the hypervisor. OSs run unmodified in this model, but must be translated. The main practical application is to enable any virtualization of the "very difficult" Intel x86/IA-32 and later x64 ISA MPU systems, the only way this can be achieved at all. The disadvantage is the performance overhead the needed translation and emulation processes cause. VMware's ESX Server 3 is the leading current example of a hypervisor employing this method.
- **Trap and Emulate:** This approach, which requires no change to the OS, and no hardware virtualization support, was first used in IBM's pioneering mainframe timesharing OSs that enabled the whole timesharing computer services industry of earlier decades.

Understanding these diverse approaches is crucial in evaluating the potential and/or drawbacks of specific virtualization options.

System z Mainframe – IT Virtualization's "Gold Standard" Platform

IBM's System z mainframe is universally recognized as the unquestioned IT industry "Gold Standard" in virtualization, having pioneered and first developed all main aspects of server system virtualization, starting over 40 years ago. Today, the System z9 mainframes, and their new, February 26th, 2008 announced, System z10 z6 MPU-powered, next-generation successors, are acknowledged to offer the most extensive, scalable, granular, totally secure, highly manageable, high-performance, long refined and optimized virtualization capabilities of any computing platform, by a wide margin. Highlights of the System z virtualization architecture, and major capabilities, are illustrated and commented on in Figure 6, and are broadly described in the paragraphs following.

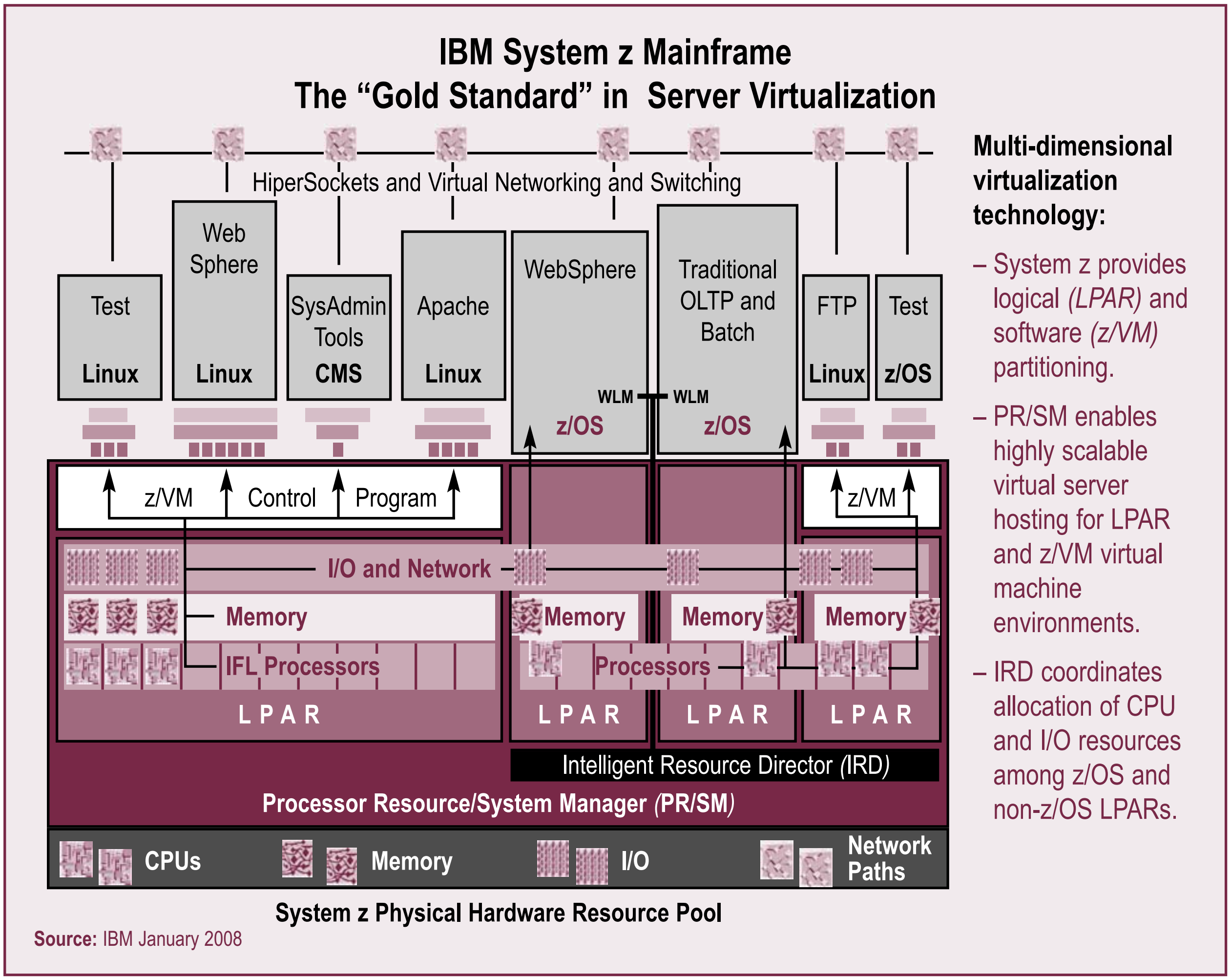


Figure 6: IBM System z Mainframe – The "Gold Standard" in Server Virtualization

The mainframe's highly-developed virtualization capabilities encompass complete virtualization of all mainframe resources, including CPU, to memory, to I/O, and to networking. They also offer sophisticated virtual resource management and monitoring capabilities, all extensively proven on large enterprise workloads at many thousands of mainframe customer sites over many years. The System z mainframe has thus long been the ultimate, "shared everything" enterprise computing platform.

System z mainframes are easily able to support thousands of users, run hundreds of separate applications, across scores of separate computing workloads, with assured high QoS and performance levels for each, with unbreakable security, and at extraordinary levels of reliability, capabilities that no other computing platform can begin to approach.

Deep hardware support for virtualization has long been built into System z at the chip and system level, and its supported OSs are fully virtualization-enabled. The prime partitioning capability is its long-proven, highly flexible dLPAR partitioning. This is implemented with its LPAR "bare-metal" hypervisor *(the PR/SM partition manager)* and fully exploits System z's extensive in-built virtualization hardware support *(SIE)*. This hypervisor enables System z memory, CPU capacity, and I/O interfaces, to be moved between LPARs dynamically in fine-grained increments whilst all the LPAR partitions *(up to 60 LPAR on a top z9 EC system)* continue to run uninterrupted. This mainframe dLPAR partitioning capability is extremely robust, completely secure, and has long run thousands of the world's largest and most demanding transactional applications and databases flawlessly, and so is by far the most "industrial strength" and flexible server partitioning implementation.

System z also supports an additional level of extreme virtualization using its z/VM "bare-metal" hypervisor that uniquely can support up to thousand significant weight Linux OS instances/virtual servers/VMs on a single mainframe system, what we term "extreme virtualization" in this Paper.

Other important, mainframe platform, dLPAR-related features include:

- **System z HiperSockets™:** High-performance, low-latency, "inside-the-box", inter-partition TCP/IP network links. *(See top of Figure 6.)*
- **System z Workload Manager:** Manages the workload within each System z LPAR partition to ensure policy-set partition service SLAs are met, and also collaborates/inter-operates with IRD below.
- **System z Intelligent Resource Director:** Sophisticated, cross-dLPAR workload optimizer that dynamically moves resources across dLPARs to achieve overall, policy-defined SLA performance. *(See Figure 6 positioning.)*
- **Parallel Sysplex:** IBM's System z "full-system-cluster" that enables up to 32-System z mainframes to be loosely coupled, operated, and run as one massive, always-available, mainframe cluster system, supporting up to 1,920 dLPAR partitions.
- **Geographically-Dispersed Parallel Sysplex *(GDPS)*:** IBM service solution *(and software)* that delivers a two- or three-site GDPS distributed System z Sysplex cluster configuration, with extensive failover and recovery capabilities to provide the highest available DR/BC solution to selected/all supported dLPAR workloads.

IBM has also created elegant, common system management and provisioning software ***(IBM Systems Director, IBM Tivoli Provisioning Manager, et al)*** that now supports all IBM Systems server platforms, and which provide extensive support for System z virtualized environments. Such VMM and VS/VM provisioning software is absolutely crucial to achieving the fullest benefits of virtualization, and must be high up on the prospective virtualization customer's "must-have" list.

Over the 2000 decade, IBM also successfully transferred mainframe-inspired virtualization technology over onto its other server platforms *(System p, System i, System x, and BladeCenter)*, some discussed above. We further evaluate the core strengths of the System z mainframe in Section 5, and of its z/VM extreme virtualization hypervisor in Section 6.

x86/x64 Server Virtualization – VMware Generic Brand to Date

The inefficiency and wastage associated with dedicated distributed x86/x64 servers was first recognized at the end of the 1990s. The notion of bringing server virtualization to the fast-growing x86 platform was therefore appealing, for the greater efficiency and utilization it could potentially bring. However, it proved extremely difficult to accomplish this, because of demanding constraints/limitations in the x86 platform's ISA design.

VMware, then a pioneering, small software developer, was first to overcome these severe barriers, and to develop/patent a workable x86 virtualization approach. It brought a family of x86 workstation and server virtualization products to market, starting in 2001. The company grew fast in recent years, reaching revenues of $387M in 2005, and $704M in 2006, up 83% YOY. Today, VMware is the market-leading vendor of x86/x64 platform virtualization software, although its success has now attracted many, newer x86/x64 virtualization competitors. *(See Figure 8 on page 21.)*

With market interest in virtualization burgeoning over the last couple of years, VMware *(an EMC subsidiary)* also attracted quite extraordinary levels of investor/stock market attention/interest, almost at "Google float" levels of hysteria. Parent EMC *(who wisely acquired VMware in 2003 for what now seems a bargain $635M)* raised $1B from a successful, partial Initial Public Offering *(IPO)* of 10%/33M of VMware's shares in August 2007. Floated at $29/share, VMware stock peaked at $125/share and recently *(11th January 2008)* traded at $77.53/share, valuing this relative "software minnow" unit at a staggering $29.7B.

VMware today offers its **"VMware Infrastructure 3"** family of virtualization and related systems management software, exclusively for x86/x64 hardware platforms, running the various Microsoft Windows OSs and Linux versions. Range flagship in our server virtualization context is VMware ESX Server 3, its bare-metal" hypervisor that allows several VMs to be run on a single, physical x86/x64 server, thus allowing limited degrees of x86/x64 server virtualization and consolidation.

The technical architecture of VMware ESX Server, with our explanatory comments on how it works, are shown in Figure 7 below, together with some key constraints of the product *(text boxes)* that are amplified below.

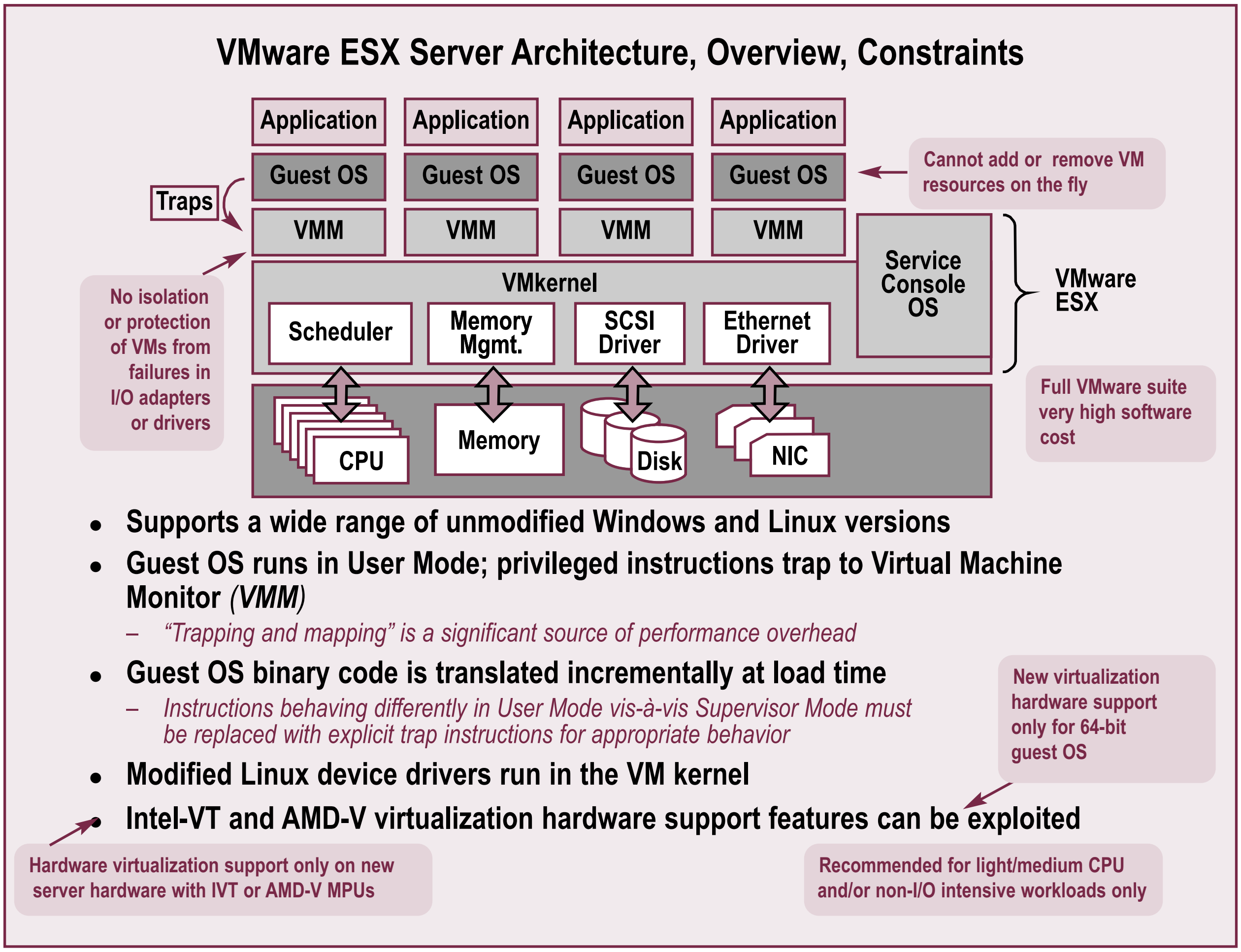


Figure 7: VMware ESX Server Architecture, Operations Overview, and Constraints

ESX Server essentially only offers static server partitioning, because users cannot dynamically add or remove VM resources on-the-fly whilst the VM continues running. Its architecture offers no isolation/protection of VMs from failures in I/O adapters or drivers, which is a serious availability/reliability limitation. Late ESX Server versions can now beneficially exploit the IVT and AMD-V hardware virtualization support now offered in new *(2005/06 on)* Intel and AMD x64 *(only)* MPU-powered servers. However, all older x86/x64 servers obviously provide no such virtualization hardware support, and indeed have arguably all now been rendered obsolete by its arrival on newer x64 MPUs. In addition, the new x64 hardware virtualization supporting facilities only benefit the newest, minority used, full-64-bit guest operating system-supported workloads. The huge majority of 32-bit x86 OSs supporting 32-bit application software workloads thus gain no virtualization benefit from the new hardware.

The full VMware Infrastructure 3 suite stack is today sold as such a high-software-cost offering as to sharply constrain where the VMware stack can economically be deployed. Finally, ESX Server is not generally recommended to host VMs running heavy-duty, CPU-intensive, or especially heavy I/O intensive, application workloads. Users are advised it is better/safer run these directly "on the metal" of an x64 server.

The x86 virtualization/market hype that included VMware mentions made the firm the "generic brand" for x86/x64 virtualization. It is thus today thought of by IT professionals familiar only with the x86/x64 world as the main option for optimizing sprawling, distributed x86/x64 servers. VMware ESX Server 3 is clearly an excellent product, which made the firm a successful market-share-leader in this segment today. As such, it must provide our main market reference point and comparison benchmark for other solutions.

In Sections 4 and 5, we examine the most radical of these alternatives, using the System z mainframe that our research found offers large advantages over the VMware ESX Server virtualization route on x86/x64 hardware.

x86/x64 Virtualization Drawbacks Remain Large

Overall, x86/x64 virtualization remains a relatively young, immature, and a still-evolving technology. Whilst small-scale usage is straightforward, satisfactory, and can be affordable, large-scale enterprise usage is highly problematic, and fails to address many of the serious issues of scale-out distributed computing we enumerated.

Most x86/x64 virtualization solutions currently suffer from numerous limitations and drawbacks when compared with the much more scalable and robust Linux-on-z/VM System z solution, fully assessed in Section 5. Amongst these substantial x86 virtualization issues/limitations of concern are the fact that it today only partly reduces current extreme server sprawl for the reasons above, and that x86 VM integrity and security remains well below enterprise-class, both matters of concern for many applications. That x86 virtual servers cannot be reconfigured without server outages seriously affects the QoS level they can deliver, and their inability to support high-bandwidth I/O virtualization workloads is a severe constraint. In addition, their virtual server workload management, virtual server and total system performance reporting and capacity planning are, as yet, relatively primitive compared to more mature virtualization environments. Provision for VM resource usage measurement accounting, and user chargeback, a vital function on any shared-resource virtualized system, is not widely available nor a standard feature on x84/x86 virtualized systems. The lack of such usage records and analysis also renders forward capacity planning difficult without the needed base data.

With the limited scalability of most virtualized x86/x64 systems, although the need for many discrete servers is partly reduced, it is not eliminated, server sprawls continues but at a lower level. In fact, on the relatively simple x86/x64 platform, adding virtualization actually adds a fairly substantial, extra layer of complexity to the tasks of managing and operating distributed servers, which must be set against the modest gains obtainable.

The recent explosion in the number and diversity of x86/x64 virtualization offerings, reviewed below, has also made choosing such a solution more difficult, and raised the specter of early commoditization of x86/x64 hypervisors, a serious deterrent to further such investment.

Don't Forget Who Pioneered, & has Most Virtualization Experience/Expertise

IBM invented system virtualization, and pioneered, refined, and extended most of the fundamental technologies and approaches of virtualization over almost four decades, mostly on the mainframe platform. This has culminated in the complete virtualization of today's System z9, with its highly sophisticated, long-proven LPAR and z/VM 5.3 bare-metal hypervisors, and the extensive management software that optimizes their usage. IBM also championed, adapted, and extended these mainframe-derived virtualization technologies deeply onto its other platforms *(System p, System i, System x, BladeCenter)* and to System Storage, during the 2000 decade.

It must doubtless be incredibly annoying to IBM, with its unrivalled record of virtualization innovation and leadership, to observe *(relatively small)* VMware *(and its other x86/x64 partners)* cashing in on the current virtualization market "buzz". As usual, the IT giant has been far too polite to say this publicly! Recall also the reality of "coopetition" in today's IT industry. IBM was amongst the earliest major partners/supporters/distributors of VMware software on its System x, and later BladeCenter, systems from 2001/2002 to date.

Fundamental technologies/innovations IBM pioneered have, several times, been first commercialized successfully by other companies, before IBM hit full-speed and maximized its own marketing efforts. Relational Database Management Systems *(RDBMSs)*, exploited faster by Oracle and Ingres, and RISC processors used in UNIX systems first by Sun, HP, SGI/MIPS, etc., are just two well-known examples.

In both cases, IBM later charged back to grab major shares, and/or leadership, in both markets. The firm is now ramping up its efforts to win/lead/dominate in virtualization with all its STG Systems, but now especially with its System z mainframe and z/VM extreme virtualization solutions.

x86/x64 Virtualization Options Mushroom – x86/x64 Hypervisor Commoditization Soon?

Until 2006, VMware could call most of the x86/x64 server virtualization software market its own. However, many new, competing x86/x64 server virtualization products have joined this "feeding frenzy" since 2006. The more important of these *(LHS)*, together with their vendors' associated VMM software *(RHS)*, are shown in Figure 8 on page 21.

x86/x64 Server Virtualization and Management Software
Exploding Hypervisor Market Competition

Main x86/x64 Hypervisors

- VMware ESX Server *(proprietary)*
- Xen *(open source Linux)*
  - ***Distributed by Novell (SLES 10) and Red Hat (RHEL5)***
  - ***Marketed by Virtual Iron Software and XenSource***
- KVM *(open source Linux)*
  - ***Integrated into the Linux 2.6.20 kernel***
- Microsoft Windows *(proprietary)*
  - ***Virtual Server 2005 – Longhorn (2007) – Viridian (2007-2008)***
- Oracle® VM *(open-source, proprietary support)*
  - ***Major, Xen-based new entrant November 2007, certified with many Oracle products.***
- Sun xVM *(open source, proprietary support)*
  - ***Significant new entrant, November 2007.***
- Plus more:
  - ***Virtuozzo, HP Integrity VM, ...***

So many x86 hypervisors – new user challenge!

Virtualization-supportive System Management Toolset

- VMware VirtualCenter
- Virtual Iron Virtualization Manager
- XenSource XenEnterprise
- Microsoft System Center Virtual Machine Manager
- IBM Systems Director
  - ***Including IBM Tivoli Provisioning Manager***
- Oracle Enterprise Manager
- Sun xVM Ops Center

Where the Value is Added!

Figure 8: x86/x64 Server Virtualization/Management Software – Mushrooming Hypervisor Competition

Of these current x86/x64 virtualization hypervisors, we consider Xen of greatest potential significance for the medium term. We base this on its good capabilities/architecture, its open-source, license cost-free pricing, and its distribution via Red Hat, Novell, and now Oracle, with their Linux offerings. Two, specialist, Xen-based virtualization vendors – Virtual Iron Software and XenSource – each also offer Xen-complementary management software for Xen VMs *(Virtual Iron Virtualization Manager & XenSource XenEnterprise)* as well as Xen subscription and support services. Microsoft continues to plug away quietly with its established low-end, Type 2 MS Virtual Server 2005 hypervisor *(acquired with Connectrix)*, and on its next-step Longhorn virtualization developments.

November 2007 saw two further significant new entrants into the market when software-giant Oracle Corporation jumped in *(with its Oracle® VM Xen-based hypervisor)*, and when systems vendor Sun Microsystems launched a high-profile entry *(with its Sun xVM hypervisor and Sun xVM OpsCenter management software)*.

This much wider set of competitors *(and others)* will increasingly give VMware stronger, direct competition as-and-when they gain market traction. They are certain *(according to our research)* to force substantial VMware software price reductions over a period of time, as the x86/x64 hypervisor space *(at least)* becomes a more commoditized, hardware system-embedded, mandatory system component expected as an inclusive part of new servers. Surrounding VM management software is more likely to continue to provide license-fee-worthy, added-value capability, and to thus to provide continuing vendor revenue streams.

This wider choice also raises the questions of which of these solutions offers better long-term technology, and which vendors are trustworthy longer-term virtualization partners. We already saw that IVT & AMD-V chip-level virtualization hardware support essentially rendered all customers' earlier x86/x64 servers obsolete, from a virtualization viewpoint, since their 2005/2006 introduction. When will the next "must-have" x86/x64 chip technology arrive that means customers will again face the need to re-purchase all of their x86/x64 servers?

This should come as little surprise, because such x86/x64 roadmap discontinuities provide convenient, "technology-must-have" drivers, for Intel and its server x86/x64 partners, to force enterprise customers into still-more-rapid replacement cycles for their x86/x64 server bases.

Can users rely upon current, sometimes attractive, virtualization product introductory license fee terms being offered to launch some newer products remaining on offer? Unlikely! Software vendors such as Microsoft and Oracle, for example, are not best known for foregoing the maximum in software license revenues they can extract, albeit with their different strategies.

Our assessment suggests enterprise users should now evaluate more carefully whether large, further investments in the VMware hypervisor, just as its core x86 hypervisor business may become commoditized, are wise or may potentially become the source of future write-offs. This confusing z86/x64 virtualization scene certainly also now calls for careful review of alternative approaches, such as the System z Linux-on-z/VM extreme virtualization solutions assessed in this White Paper.

Our Analysis – Profound, Far-reaching Server Virtualization Roles, Benefits

Virtualization today plays, supports, and enables many important roles/applications in the optimization of server infrastructure IT environments. These roles/applications include server workload consolidations, dynamic virtual server provisioning and hosting, providing a base for better workload management, and offering a mechanism for workload isolation. Virtualization also helps users manage software release migrations, supports development/test/production environments without wasteful multiple sets of servers, and enables multiple types or releases of OS to run on the same physical system. It additionally enables flexible virtual clusters to be configured, and can provide back-up capability at lower costs via virtual back-up servers.

An original motivation for VM, especially on the mainframe, was the wish/need to run multiple OSs *(like and/or unlike types, or of different releases)*, where virtualization allowed users to time-share a single, physical computer between the several OSs concerned. Another common motivation for virtualization was the incomplete resource isolation most contemporary, distributed OSs offered. This rendered these OSs less than fully trustworthy of running multiple applications concurrently, without a real risk of interference occurring between applications/processes. *(So-called QoS isolation, and a main reason behind the widespread, dedicated server deployment pattern so common and so wasteful in distributed computing.)*

Virtualization, through these roles/applications, offers wide-ranging benefits and savings to its users...

Virtualization, through these roles/applications, offers wide-ranging benefits and savings to its users, depending somewhat on the virtualization option selected. These can, however, include: delivering much higher server hardware resource utilization with less wasted capacity, improving the QoS delivered for multiple workloads, and affordably achieving higher availability and better security. Virtualization also reduces the costs of attaining adequate availability, provides much greater flexibility and responsiveness, reduces system management costs, helps support legacy software environments, and thus can protect existing investments.

We also reviewed the main methods of server partitioning and virtualization, highlighting the compelling merits of "bare-metal" hypervisors operating with hardware virtualization support, as the best technology in this space. We also reviewed the main methods with which hypervisors can be implemented, and their pros and cons, to help readers understand why certain virtualization solutions are superior to others on offer.

In summary, and rather more broadly, virtualization/server consolidation therefore offers five major benefits:

- **Much-reduced staff support/management costs:** With many fewer physical servers to manage, and superior management software to hand, common management tasks are easier, and so support staff resources needed are much reduced. Since these people costs became much the largest TCO element in most distributed infrastructures, the TCO savings on support costs from full consolidation using quality virtualization can be huge.
- **Much-reduced software costs:** Again, with many fewer physical servers and processors used more efficiently in a fully virtualized environment, many fewer software licenses, for many fewer CPUs, are needed for OSs, DBMSs, and other key middleware. Since the initial license fees and service/subscription fee costs of software are generally the second-largest TCO component in distributed environments and a higher cost than hardware, virtualization/server consolidation can slash these software costs sharply.
- **Major Environment Cost Savings/Green IT Benefits:** Virtualization also delivers substantial reductions in data center floorspace, IT electrical power demand, and in the cooling energy needed. Both in terms of deferring large new capital investments that would otherwise be needed to extend or add new data centers/capacity and plant equipment, and by providing sharp reductions in operating power and cooling costs, these are today hugely important cost saving items. Perhaps more importantly, they are vital to the enterprise's "green IT" credentials.
- **Much-improved flexibility and responsiveness:** Because virtual resources on the right platform can be changed dynamically, the best virtualized systems can meet changing workload demands in real time, and can reliably deliver SLA mandated performance levels never attainable with distributed systems predecessors.
- **Reduced server hardware costs:** Through much-increased physical resource utilization, on many fewer physical servers than on the previous distributed configuration, hardware cost savings can often also be achieved in distributed server consolidations using virtualization.

These generic virtualization benefits are compelling, but apply to a much greater or a much lesser degree depending upon which approach to virtualization is taken.

The IBM System z mainframe undoubtedly remains the industry "Gold-Standard" for ultimate virtualization amongst all IT platforms. However, its extraordinary capabilities remain much less well/widely known than they should be. Far too many IT community members are still inexperienced/lacking in current knowledge about the latest mainframe technology and system capabilities. In part, this was because of the continuing massive marketing hype focused on distributed x86/x64 computing, from its many wealthy backers. This tends to swamp all other news and technology review coverage. In part, this has also been a generation/age issue; with too many of the younger age groups growing up into IT with only distributed systems experience, and lacking any exposure to mainframes. In recent years, IBM has significantly cranked up its mainframe marketing and promotional efforts with increasing success, and is now deeply established in helping educate a younger IT generation on System z technology via the extensive IBM Academic Initiative for System z universities and customer partnerships.

The IBM System z mainframe undoubtedly remains the industry "Gold-Standard" for ultimate virtualization amongst all IT platforms.

By contrast, little VMware has enjoyed quite extraordinary, and relatively far disproportionate, hype and attention for its first-to-market x86/x64 server virtualization and management platform solutions. These today are resold/offered by many other, larger x86/x64 market vendors/players as a **"x86/x64 panacea solution"** to customers' distributed, scale-out, x86/x64 server sprawl nightmares. Useful and sound though VMware's offerings have been widely found to be, the solutions the vendor offers also have constraints, limitations and disadvantages, as well as steep costs. These drawbacks only become apparent to IT users when VMware solutions are closely compared to deeply different, alternative virtualization solutions.

Extreme virtualization using Linux-on-z/VM on System z mainframes for large-scale, mass consolidation of scale-out distributed systems *(UNIX and x86/x64)* is the most strategically significant, and profoundly different, approach to solving to the same challenges as VMware seeks to address purely within the x86/x64 environment. For this reason, this White Paper is amongst the first *(if not the first)*, to specifically compare the differences between these two main approaches, beginning with a high-level assessment of the previous generation, and the new-generation System z10 EC mainframes *(announced on, and GA from, February 26th, 2008)* in the following Section 4.

4. System z Platform – Strongest Foundation for Optimized IT Infrastructure

Introduction

In Section 3 we introduced the outstanding virtualization capabilities of the IBM System z mainframe around our Figure 6 view on page 17. Experienced mainframe customers/users will be familiar with many of the deep strengths and capabilities of the IBM mainframe platform, and of many of these unique virtualization strengths. For readers newer to the mainframe and seeking to learn more of its capabilities, this Section summarizes select findings from our earlier mainframe research to provide a broad picture of what the platform offers today. We also outline how we expect its now-imminent, next-generation iteration, due out in Q1 2008, will soon further extend this powerhouse commercial system.

Mainframe Transformation Powered Earlier Revival

The 43-year-established IBM mainframe enjoyed real market resurgence, recapturing renewed industry-wide esteem and respect over recent years of this decade. Nine new mainframe generations emerged in the 1995-2006 period, six CMOS G series, and three post-2000 64-bit z/Architecture System z generations, the result of several $B of sustained IBM investment/innovation. These IBM investments transformed the platform's technology, hardware, software, and economics, delivering extraordinary innovation to every level of the mainframe hardware and software stack, and at the fastest rate in the 43-year history.

These IBM investments transformed the platform's technology, hardware, software, and economics...

From their advanced microprocessor technology, balanced/optimized, commercial workload-focused, 64-bit z/Architecture, through sophisticated new storage system complements *(DS8000/DS6000)* were added major OS advances, extended enterprise SOA middleware, and new-generation software tool, portfolios. These, plus System z's wide-margin leadership/advances in virtualization/partitioning, self-optimizing policy-driven workload management, outstanding RAS, and much-superior security/encryption capabilities, were all combined to bring/deliver major advances/innovations to the whole mainframe stack though this decade to date.

Mainframe Revival Swells to Resurgence 2003/2004

Sparked by this sustained, high rate of IBM innovation, mainframe market revival first began with the 2000 zSeries 900 and 2002 z800 systems. It accelerated into resurgence with the z990 in 2003 and the z890 in 2004. IBM's Mainframe Charter *(2003)* laid out IBM's public pledges of continued advances in platform innovation, value and ecosystem development; these have been amply delivered on since, and are continuing apace today. Growth momentum was further increased by the high-end System z9 109 available from September 2005, and again by the System z9 Business Class and System z9 Enterprise Class from July 2006 *(see below)*.

This rapid pace of mainframe developments and innovation, together with major price/performance gains, were continued with the System z9 Business Class and z9 Enterprise Class mainframe generation introduced in 1H 2006. These more granular mainframes saw their maximum capacity doubled once again, as for each of the two prior System z generations, and were/are without doubt the world's most sophisticated business servers. Uniquely, these 64-bit System z mainframes were designed from the beginning to incorporate multiple types of processors *(hybrid computing)*, each optimized to handle specialized tasks, as well as their standard, general-purpose *(CP)* z9 processors. System z mainframes can optionally include "specialty processors", each designed to process eligible Linux, Java, and DB2 database workloads at far lower cost. Other processors dedicated to high-volume I/O and channel I/O, as well as to encrypting/decrypting certain data at hardware speed, are also included. These multiple, hybrid processor resources combine *(System z9 EC up to 416 total processors)* to give the mainframe its immense throughput/capacity for its now-far-wider range of designated business workloads.

Another unique is the mainframe's HiperSockets™ technology. This provides ultra-fast communication among the virtual servers contained in a single mainframe. By contrast, in a scale-out, distributed environment, where many physical servers are connected by networking cables and switches, latencies are far greater, performance/throughput is much lower, and reliability far worse.

System z has also long led the IT industry in its outstanding security capabilities, vitally important to the major enterprise business applications which it is usually deployed to run. The System z mainframe achieved one of the highest levels of security certification on the definitive Common Criteria's Evaluation Assurance Level *(EAL)* – Level 5 – for its logical partitioning *(LPAR)*, one of the mainframe's premier partitioning/virtualization technologies.

For those unfamiliar with these systems, highlights of this highly-successful mainframe generation were:

- **System z9 Business Class *(z9 BC)* entry-midrange mainframes**, ideal for SMB, priced from just $100K, in two models spanning 26-1,786 MIPS capacity.
- **System z9 Enterprise Class *(z9 EC)* high-end mainframes**, for larger enterprises, in five models, with up to 54-way CP, spanning up to a massive 17,800 MIPS capacity.
- **Superb virtualization/automated workload management**, enabling up to 100% utilization of capacity.
- **Full slate of CoD options**, for greatest customer flexibility to meet varying workload needs.
- **Unique "hybrid processor" architecture**, incorporating several types of inclusive, dedicated, and other optional, cost-effective, specialty workload processors, each optimized for major types of processing, into an immensely powerful commercial supercomputer. See examples below.
- **New zIIP data-serving specialty processor engine**, for select DB2 data-serving workloads, sharply improves database workload costs, and supports the enterprise data-serving hub new mainframe role that IBM champions today.
- **Linux specialty processing engine**, which run all Linux processing workloads on these IFLs, for much lower-cost Linux processing on System z. All-IFL *(Integrated Facility for Linux)* System z mainframes can be selected for larger Linux workloads requiring a dedicated system. The IFL is a crucial component in enabling economically attractive, mass-consolidation of distributed workloads on the System z mainframe.
- **Highest levels of system automation**, which slashed the support staff needed, now multi-fold lower than on all distributed platforms for comparable workloads. *(Often 4 to 8 times fewer staff needed on System z.)*
- **New Enterprise-wide roles for mainframe**, extensive additional support delivered apace for these five new roles.
- **Smallest data center footprint, lowest power and cooling costs** per enterprise workload, per 1,000 users, per virtual server, etc. By far the greenest solution, slashing these now huge, and fast-rising, IT costs workload-for--workload.
- **Industry first 4Gbps FICON I/O support**, for higher I/O performance/capacity on both z9 BC & z9 EC systems.

- **Leadership Mainframe Storage:** IBM's winning System Storage portfolio fully refreshed *(August 2006)* with leadership DS8300/DS8100 Turbo high-end, and DS6000 mid-range, disk systems, ideal for use with System z mainframes, and storage virtualization and software leadership gains.
- **Major New System z Middleware Advances:** The extensive suit of optimized System z9 middleware is now extensive. Leading servers, including DB2, IMS, CICS, WAS, WebSphere MQ, NetView, SOA, and Tivoli management middleware software, etc., all saw further advances again through 2006 and 2007, focused on rock-solid J2EE™ support, open standards/Web services enablement, enhancing systems management/control, and on bringing unrivalled SOA support.

These System z mainframes can handle massive commercial/business workloads. For example, a System z mainframe recently achieved the world's largest core banking benchmark result, delivering a record 9,445 business transactions per second *(tps)* in real-time, based on more than 380 million customer accounts with three billion transaction histories. *(See Footnote[1] below)*.

System z Mainframe Now Often Lowest TCO Platform

Further price/performance gains on these mainframe server hardware, storage, OSs and middleware software offerings above, again continued this decade's sharp, downward mainframe TCO trend. Sharp hardware cost reductions averaged 20%, aggregate mainframe software cost continued to fall at the c. 17% p.a. average rate consistently delivered since 1997, and other cost improvements, all combined to give 20% total mainframe price/performance gains per year through this generation once again. Their above capabilities, and such sustained economic improvements, make today's System z mainframe by far the lowest TCO/TCU/Cost-Per-Transaction platform for many medium/large mixed commercial workloads *(both traditional and new)*, including extreme virtual server hosting *(this White Paper's subject)*. This now-well-documented and validated fact absolutely contradicts still-widespread, false perceptions that the mainframe remains "the most expensive" platform, still eagerly spread by its distributed system competitors.

Their above capabilities, and such sustained economic improvements, make today's System z mainframe by far the lowest TCO/TCU/Cost-Per-Transaction platform...

This now-well-documented and validated fact absolutely contradicts still-widespread, false perceptions that the mainframe remains "the most expensive" platform...

The operative measures that must be considered here are the true total costs *(server hardware, software, staffing, power, cooling, data center space, and needed back-up and DR/BC costs)* over a full system service life-span *(can be 5-8 years on System z, versus 3 on an x86/x64 server)*. These total costs, expressed per transaction, per user, per virtual server, and/or per other comparable workload metric, are the correct basis for serious comparison. Fully-loaded System z mainframes have repeatedly proved themselves able to deliver much the lowest TCO results on such rigorous measures, for suitable workloads, compared to their distributed systems competitors.

Like the Boeing 747 long-haul airliner *(that dominated high-density air travel for 3.5 decades)*, although the capital cost of a System z mainframe *(747)* may, at first still seem high *(compared to smaller planes/servers)*, everything is tightly-integrated inclusively "within the box". It is optimized to deliver huge workload throughputs at the lowest total per-unit costs. Just like the 747, the System z mainframe is trusted by *(thousands of)* customers to safely and securely deliver the IT equivalents of "lowest passenger cost per seat mile" and/or "lowest air cargo tonne per mile cost", the two critical economic success metrics that commercial airplane operators must attain for high-profit yields on dense traffic, long-haul routes.

No big surprise then, that no long-haul airline to date has ever even dreamed of, let alone actually operated, a regular, competitive, safe, passenger long route air service using a swarm of scores or hundreds of small, volume-produced, 4 seat airplanes as their route hardware platform. Lack of range, insufficient capacity, huge staffing costs, poor safety with high accident and passenger/crew death rates, inability to operate in adverse weather, and poor overall economics, with almost no economies of scale applying, foredoomed this possible model of commercial air travel before it ever took off.

Almost equally ludicrous/unlikely is to imagine that logically similar scores or hundreds of low-cost, unsophisticated, distributed IT servers could be hooked up into one heap, and that heap can then be expected to deliver a top-quality, always-available, totally-secure, large-scale enterprise IT application workload globally over 7*24*365. First, users must couple multiple tiers of servers and all their needed software.

Footnote[1]: IBM Press Release: "IBM and Financial Network Services Deliver World's Largest Core Banking Benchmark" on http://www-03.ibm.com/press/us/en/pressrelease/21044.wss.

Then they must add the additional DR/BC clustered servers needed to raise availability/protect the platform. They will then need to add a mass of cables, switches, server area, storage area, and Ethernet, network links to connect all these boxes together, to disk and to IP. The power/cooling energy, costly data center floorspace needed, and the army of staff required to operate and support all this must then be costed into TCO per workload metrics. When all done, these costs usually now far exceed those of a System z mainframe that has been specifically built to do just this type of job to perfection "inside one box".

As IBM has systematically driven down the all the TCO cost elements of running System z mainframes year-by-year through this decade *(and before)*, the cost/benefit case for adopting the mainframe has become ever stronger. In addition, the range of workloads that can now best be processed on mainframes has grown sharply because of these hugely improved economics, the wider range of low-cost specialty processing engines offered, and the wider range of mainframe software now offered.

Much detailed TCO mainframe data has also finally been researched, published, and distributed by IBM in 2006 and 2007 covering these sharp improvements in mainframe TCO and overall economics, and interested customers should consult this now-wide body of definitive, favorable TCO evidence

Using many times fewer support staff, several-fold fewer software licenses, much less power, lower cooling, and far less floorspace with far less hardware, these mainframes now offer much lower TCO/TCU/CPT, but far higher reliability, availability, security, scalability, etc., than equivalent power distributed system server farms. These large advantages are today fuelling a fast-growing migration of workloads, from tens of thousands of wasteful distributed servers onto highly virtualized and ultra-efficient System z mainframes.

Every System z An In-house Enterprise IT Service Bureau

Another view of today's System z mainframe is that these systems provide **a complete, in-house enterprise IT service bureau** capability. Each System z can support scores of different applications and workloads, from dozens of internal or partner business entities, and for thousands of enrolled users. These it can serve with complete security, and virtualization- enabled isolation from all others, with every hardware and software resource, power, cooling, and floor space, all most efficiently and completely shared, whilst delivering superior response and performance, highest availability and sophisticated DR/BC where needed. Highly automated, such a mainframe "in-house service bureau" can be operated today with small staffs, delivers lowest IT cost per units of work, and can accurately measure/meter and recharge usage however required. The more groups of users added to the enterprise service bureau mainframe, the lower unit costs become.

System z mainframes are almost invariably used as such shared-everything, multiple applications, and multiple workloads, enterprise hub platforms, supporting many business entities, multiple business units, and multiple locations; all concurrently. In practical terms, each System z mainframe offers a **"data-center-in-a-box"**, with a wide range of "box sizes" offered. Almost everything needed actually runs within one frame/chassis, not strewn all across the data-center.

Therefore, adding a new workload at a mainframe customer site is normally just a question of adding any needed new application software onto existing hardware. When development is complete, and the new application has been tested fully, it is then rolled out into production with both the development and the testing also safely having being performed on the same, single mainframe, using its virtualization capabilities to isolate these environments.

The additional load of a new application *(small-medium)* will often run fine on current mainframe installed capacity unchanged, because a System z can optimize its resource utilization across all application workloads run. If needed, hardware capacity can be instantly increased on-site with installed CoD options, and/or via favorably priced model upgrade paths always offered by IBM. Little or no extra operation/support staff or time is normally required to manage such incremental mainframe loads in production. Therefore, the mainframe approach is "just add software".

The mainframe also provides excellent economies of scale with increasing size of total workload. This drives TCO per unit of processing *(e.g. per transaction, or per user)* sharply down as workload size increases. These economies of scale come from the high hardware utilization achieved *(well up in the 90% range often attained)*, from the discount price curves of IBM mainframe software costs as system sizes increase, and from the only slowly increasing support staff time needed to handle larger workloads on the platform.

The mainframe therefore now offers a dramatically superior economic and environmental model than the scale-out distributed computing model.

The mainframe therefore now offers a dramatically superior economic and environmental model than the scale-out distributed computing model. In that approach, every significant new workload or application requires yet more sets *(usually at least 3 sets – development, test, and production – yet more if any DR/BC is needed)* of dedicated, distributed servers. These need associated storage capacity on SAN or NAS arrays, and all the middleware and application software licenses required on all these servers. They also demand all the complex cabling, network switches and routers needed to interconnect the servers and storage, and to connect them all to the enterprise network, and the substantial support/system administration staff team that is now needed to support, operate, maintain, tune, and troubleshoot all this in production.

Software As A Service *(SAAS)* and System z

Our industry saw climbing adoption of the Software As a Service *(SAAS)* model of application software service delivery, where users of an application solution are serviced/supported from a centrally-run, shared hosting platform, usually operated by, or on behalf of, the SAAS application solution provider. The biggest SAAS success to date has been Salesforce.com, which provides an excellent, modern Web-delivered, extensible CRM application service to its customers, on a classical SAAS per-user, per-month service fee basis. The SAAS model offers many advantages to enterprise users, including "start small-scale at will", lower software support costs *(vendor supports service software image)*, lower capital investments, regular software updates automatically provided, etc., etc., no server hardware needed by user, etc.

The System z mainframe is the perfect host platform for the delivery of SAAS application software services on a large scale, for all the "in-house service bureau" and "data-center-in-a-box" reasons above. It now supports a full gamut of applications, technologies, and modern open standards, to allow the widest range of SAAS applications to run on the platform. Its huge scalability, rock-solid security, outstanding application and data serving performance, unrivalled availability, and automated, low-staffing operations, make System z the ideal host and hub platform to deliver SAAS services. This applies equally if the goal is in-house SAAS delivery across a single enterprise, or is to be commercially provided to many different customer organizations on a SAAS basis.

Precision Usage Analysis, Accounting and Chargeback

To support such above "enterprise IT service bureau", or SAAS-type application service delivery, as well as regular, single customer large-scale internal enterprise-wide, shared usage of a common System z mainframe configuration *(of one or many machines)*, IBM now offers extensive software capabilities. These track and record every form of mainframe resource usage in detail, can analyze this extensively, and can post-process such usage data into usage accounts by user entity, with cost recharge if wished.

Standard, fine-grained, resource usage logging facilities *(SMF/RMF)* at the base OS level are the foundation. Next up, the powerful **IBM Tivoli Decision Support for z/OS *(ITDS for z/OS)*** tool collects all such mainframe usage data *(and other)*, and builds it into a flexible DB2 database on the platform, with extensive, easy-to use Web reporting. Finally, the powerful, modern **IBM Tivoli Usage and Accounting Manager for z/OS** analyses data from that ITDS DB2 database, to provide flexible IT resource usage accounting, by client company, division, department, or even individual users, to account and recharge system usage as required.

The platform can therefore be fully equipped and ready to rock with these sophisticated, flexible, precision usage analysis, and accounting capabilities, that can be deployed quickly to fully support commercial service bureau, SAAS, or internal enterprise shared use accounting, financial chargeback, and cost recovery

Spectacular, New IBM z6 MPU Powers Next-generation System z10 Mainframes, Announced February 2008

In November 2007, just before date of writing *(January 2008)*, IBM presented details of its **spectacular, new z6 System z CISC mainframe MPU chips** to the IT system analyst community, this writer included. These massively-powerful, **ultra-high-frequency, quad-core MPU chips**, together with an immense, new companion **System z SMP Hub chip**, that now power the new, next-generation System z10 mainframes first announced on February 26th, 2008 and to ship from Q1. *(On IBM's accustomed two-yearly, new mainframe generation cycle.)* IBM's pre-announcement of these hugely-powerful, new mainframe z6 MPU and SMP Hub chips, several months ahead of the new mainframe server systems they power, was unusual. Before, IBM usually showed its new mainframe MPU chips at the same time as the new systems they powered, often with limited chip technology disclosure.

By contrast, this time round IBM came out, proudly, publicly, and months ahead of the new System z10 servers, with fairly extensive, and highly-impressive, details of these new mainframe system chips, which are by now ramped-up in fabrication.

The disclosure reflects IBM's high confidence levels/ambitions for the further substantial mainframe growth it now cherishes for all its mainframe-centered businesses *(servers and storage hardware, OS, middleware and tool software, and extensive services)*. These remain highly profitable, and are again expanding. Each complete new mainframe generation typically represents of the order of $1B IBM hardware and software investment, another large indicator of IBM's high confidence in its flagship server platform today.

We assess that IBM took this step firstly because it is proud of its engineering team's accomplishments with these new mainframe chips. Secondly, because this news previewed that IBM would soon offer further leaps in mainframe capacity, performance, and capability, when the next-generation System z10 mainframes built from these chips, that were announced on, and began GA shipping from, February 26th, 2008 *(and whose sales it now wishes to boost)*. Thirdly, IBM needed to definitively kill-off long-running industry rumors and speculations, now plainly shown as wrong, that it would move the System z mainframe onto its POWER6 MPU chip; now definitively not the case.

Because these striking new System z z6 MPU chips now power the new System z10 mainframe generation that clearly continues/accelerates IBM's strong record of mainframe advances *(noted earlier)* by another major step, we discuss some of their highlights here. Figure 9 shows the sharply accelerating mainframe processor frequencies for each mainframe generation, from 1997's G4, up to the new z6 MPU. This new mainframe MPU runs at a far higher frequency, quoted at a stunning 4.4 GHz. level, more than 2.59-times the 1.7GHz. frequency of its z9 EC predecessor that was first seen in 2005. With this large hike in MPU frequency, improved scaling via its new SMP Hub architecture, and other improvements, the capacity/power of the new, System z10 EC top-end mainframes has jumped by 70% over that of their System z9 EC predecessors.

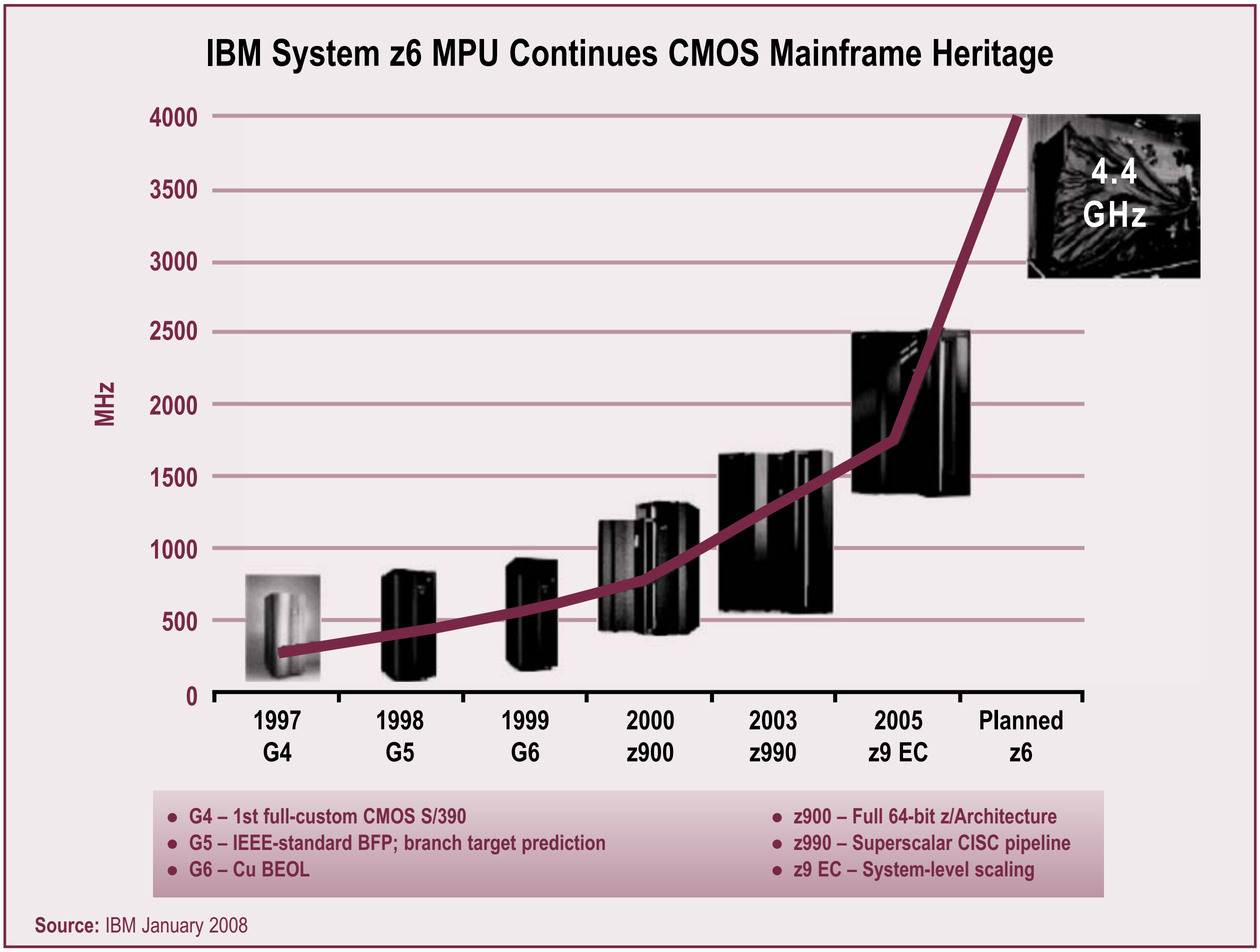


Figure 9: IBM System z6 MPU Continues CMOS Mainframe Heritage

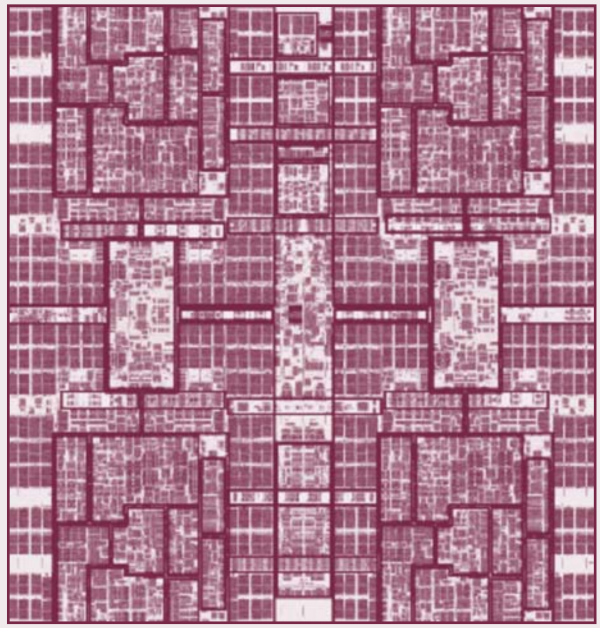
Contrary to widespread, earlier industry speculation, the z6 MPU chip is rather different from the already-released, and also extremely powerful, IBM POWER6 RISC MPU that now powers IBM's System p, System i, and some BladeCenter, servers. The two chips are said to be siblings, not identical twins, sharing considerable DNA, components, and the same IBM SOI 65 n.m. process technology, but have different personalities, characteristics, and target roles. The z6 MPU was naturally heavily optimized for the mainframe's centralized enterprise data-serving hub and extreme virtualization consolidator roles. A pictogram of the z6 MPU chip *(RHS)*, and some of its feature highlights *(LHS)*, are shown in Figure 10 on page 29.

The z6 is IBM's **first quad-core commercial server MPU chip**, with an ultra-high 4.4GHz. core frequency, and with its cores packing 3MB of private L2 cache. Mainframe on-die accelerators for data compression *(heavily used in database work)* and cryptography *(which underpin System z's fantastic encryption performance potential)*, are joined by a sophisticated Decimal Floating Point accelerator *(for ultra-fast and precise financial calculations in hardware)* on the z6 chip similar to that implemented recently on the POWER6.

The z6 MPU packs **almost 1B transistors *(991M)* on the 21.7*20.9 m.m. die**. It also houses highly-sophisticated, integrated SMP communications, including a switch to connect cores to the new **IBM SMB Hub chips**, memory and I/O bus controllers, and the 3rd generation of IBM Elastic Interface bus technology running at up to 3GHz. These bring the z6 the massive bandwidths shown bottom left in Figure 10.

IBM z6 – Next-Generation Mainframe MPU Blockbuster Chip

- **New high-frequency z/Architecture microprocessor core**
  - 4.4GHz. operation in system
- **4 cores per die**
  - Each with 3MB private 2nd-level cache
- **Accelerator engines**
  - Data compression
  - Cryptographic functions
  - Decimal floating point
- **Integrated SMP communications**
  - Switch connects cores to SMP Hub chip
    - Shared cache and SMP fabric
  - Memory bus controller
  - I/O bus controller
  - EI3 technology up to 3GHz bus speeds
- **System interfaces**
  - 2 x 48GB/s SMP Hub
  - 4 x 13GB/s Memory
  - 2 x 17GB/s I/O



- **991M Transistors**
- **138Mb SRAM**
- **6km wire**
- **21.7 x 20.9mm die**
- **1188 signal/8765 total chip I/Os**

Source: IBM January 2008

Figure 10: IBM z6 – Next-Generation Mainframe MPU – Blockbuster Chip

As always, the IBM z6 Instruction Set Architecture *(ISA)* continues the line of upward-compatible mainframe processors that have **preserved full mainframe application compatibility since 1964**, and also **supports all z/Architecture-compliant OSs**. This rich CISC ISA now supports 894 instructions *(668 implemented entirely in hardware)*, 50 of them new in the z6, and supports 24-, 31-, and 64-bit addressing modes.

System z's **industry-leading, chip-level virtualization hardware support** for both high-performance logical partitioning via PR/SM *(dLPAR)*, and for fine-grained virtualization via z/VM – scaling-up to run even 1,000s of virtual server images – is a central z6 feature in our White Paper's context. The z6 MPU also includes the industry's most extensive, extended set of on-chip RAS features, to assure the highest levels of inbuilt error-checking and recovery of any commercial server.

A **massive, new, companion 1.6B transistor, IBM SMP Hub Chip** will connect multiple IBM z6 MPU chips with 48GB/Sec bandwidth per processor, and provide access to shared L3 SRAM cache of 24MB per Hub. Hubs can be paired to increase shared L3 to 48MB per pair. This design provides low-latency cache coherence, robust SMP scaling, and extended SMP scaling when multiple SMB Hub chip/pairs are used, IBM said.

These z6 chip specifications gave early indications of what IBM's next-generation System z10 mainframes are now confirmed as offering customers, after their February 26th, 2008 announcement. These enhanced capabilities include:

- Offer more **compact, more economical, mainframe server implementations** at or near the main existing capacity points of the System z9 range needed to cover the market *(these ran up to 54-Way CP/64-Way total PUs SMP, 17,800 MIPS capacity at the top-end on the z9 EC S54)*. The new System z10 mainframes exploit the z6 MPU's quad core and far higher-frequency specifications, to deliver mainframe models at needed market capacity points over this range, but using fewer MPU chips/MCMs/infrastructure parts than on the z9s, and thus able to be offered at even more competitive prices.

- **Significantly raise high-end mainframe capacity**. The z6 has a 2.59-fold higher-frequency *(4.4GHz., not the 1.7GHz of the z9 MPU)* and is a quad-core *(rather than the z9's dual-core chip)*. It also boasts a sophisticated, high-performing, optimized SMP Hub chip fabric. IBM has chosen to exploit these advances with its new top-model System z10 Enterprise Class, which it rates as offering 70%* more capacity than its z9EC S54 predecessor. *(* LSPR mixed workload, z/OS 1.8.)*

  IBM's new z6-MPU powered, System z10 EC high-end mainframe offers this rated 70%* power hike housed within a similar physical format/footprint 2 bay/frame chassis as its System z9 EC S54 predecessor. The new System z10's industrial design also boasts a tasteful broad green stripe down the front, highlighting its "green credentials". The System z10 EC top-end model's system architecture uses a 4-book, 4-MCM CEC, mounting 16 z6 quad core chips, and providing its maximum 64 z6 MPU cores *(each 2.59 times faster than z9 cores)*, together with the advanced, new z10 SMP scaling fabric. Yet again with the z10, IBM is trying to abolish MIPS ratings *(on the usual ground)*. However, the new z10 EC top end system actually offers about a 30,250 MIPS* capacity up from 17,800 MIPS* on the z9 EC S54, widely varying with workload.

We take the space here to outline the much-extended scaling-up potential of these imminent, next-generation mainframe systems because this will **substantially further extend their extreme virtualization capability** to consolidate large numbers of Linux-on-z/VM virtual server workloads on a single mainframe. The new System z mainframes are, in our view, thus likely to lift today's realistic and practical, 250-400 significant Linux virtual servers per large System z9 consolidation ratio, up to somewhere in **the region of 550-1000 significant Linux virtual servers per large, new, z6-powered mainframe**. Just 1,000 such new mainframes, fully-loaded with Linux workloads migrated over from distributed servers, would thus be able to beneficially replace up to 1,000,000 *(1M)* of these wasteful, distributed, x86/x64 or UNIX servers.

We discuss the huge advantages and savings customers can now gain by doing just this themselves in the remainder of this White Paper.

System z – Ideal Solution for Massive Distributed Server Consolidation

The essence of the argument/case for using System z mainframes to massively consolidate workloads off hundreds or thousands of customer's existing, scale-out distributed systems *(x86/x64 and UNIX)*, and is depicted at a high level in Figure 11A.

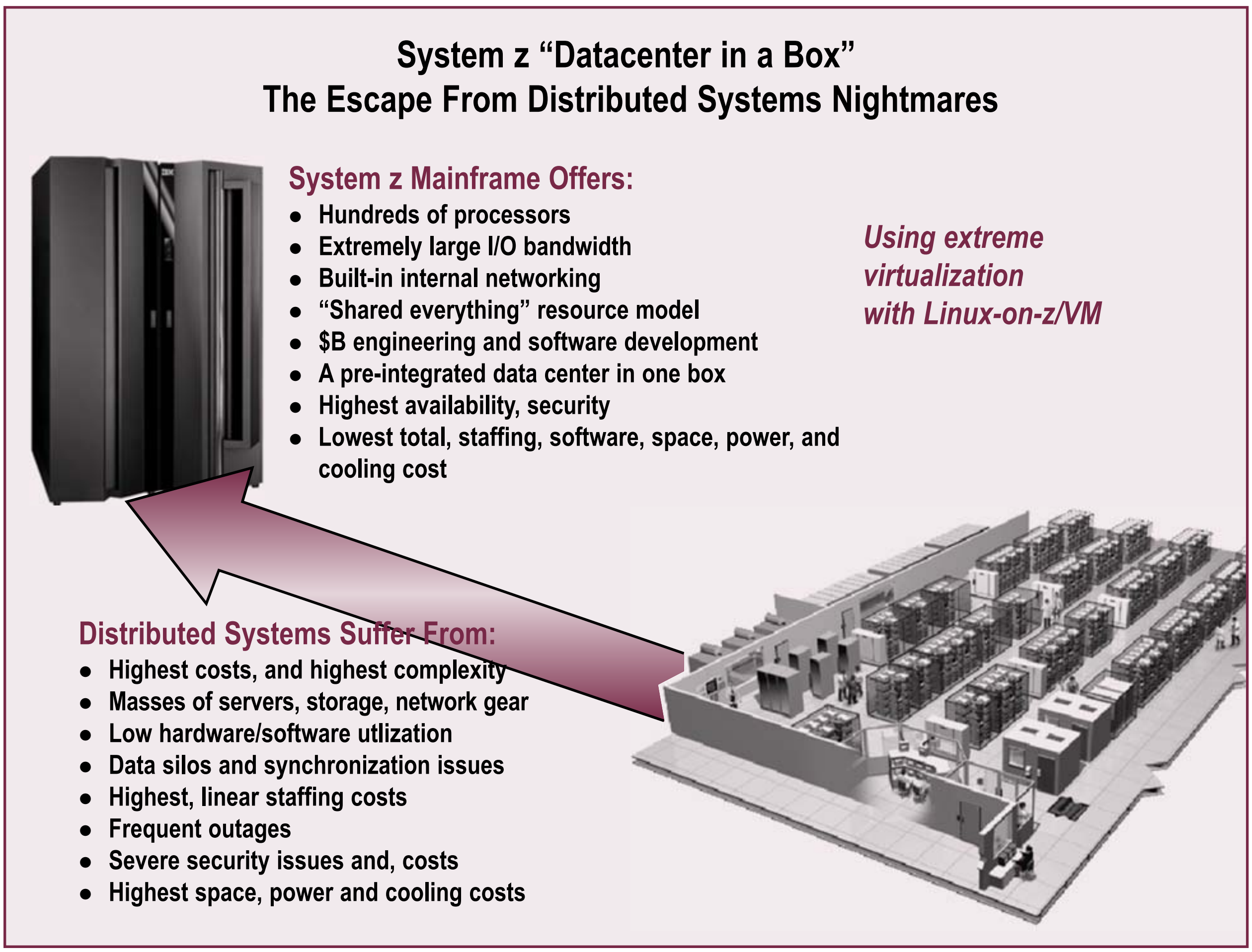


Figure 11A: System z "Datacenter in a Box" – The Escape From Distributed Systems Nightmares

The typical distributed systems data center sprawl *(pictured bottom right)* suffers severely from all the major disadvantages summarized concisely in the bottom left bullets, which we summarized in Section 2 previously, and exhaustively detailed/assessed in Appendix A later in this Paper.

By contrast, the System z mainframe provides literally a complete **"data-center-in-a-box"** that can often handle the same total consolidated set of workloads as previously filled the whole distributed computing data center depicted below, within the single System z mainframe shown. The main features that enable the System z to achieve this extraordinarily massive consolidation, some already discussed above, are summarized in bullet-point form top center in Figure 11, and amplified below:

- **Hundreds of processors worth of power:** Up to 416 total processors can be configured in a top-end System z9, including standard and specialty engine types, and dedicated special-purpose types, in these extraordinarily powerful, hybrid "commercial supercomputers". These already-large processor counts are set to be further increased in the new, 2008 System z mainframe generation.
- **Extreme I/O bandwidth:** System z mainframes were long designed/optimized, from the chip-level upwards and throughout their hardware-software stack, to offer massive I/O bandwidth, ultra-fast context switching, and high degrees of efficient multi-threading. These crucial capabilities together provided System z mainframes with their unrivalled, commercial transaction processing and data-serving performance, and in particular their massive I/O capabilities that are essential to run such large commercial workloads. Mainframes use a sophisticated, long-refined channel I/O architecture *(conceptually similar to recent open systems InfinBand channel I/O model)*. In this, much mainframe I/O workload is offloaded onto dedicated I/O system channel processors to handle much of the I/O heavy lifting off-line the main CP, SA, or specialty processor engines. Up to 1,024 mainframe channels can be supported on a top System z9 EC.
- **Built-in internal networking:** The mainframe offers its ultra-high performance HiperSockets™ links between applications and LPARs, and its ultra-flexible virtual networking and switching, all "inside the box" and operating at memory/bus speeds. This has none of the latency, performance, or reliability issues that are common with the complex physical meshes of cables, switches, and routers, etc., needed to network-connect and inter-connect typical multi-tier distributed server solutions. These capabilities bring superior performance and QoS from such inter-application, inter-process, and inter-partition networking for such mainframe-based solutions.
- **"Shared everything" resource model:** Discussed above, and contributes to the favorable TCO/economics, and the increasing economies of scale, that mainframe platforms offer.
- **$B engineering & software development:** Sustained, decade-long, and massive R&D effort, has been unstintingly continued by IBM, spending many $B total on the advancement of the platform over the past decade. This has not only delivered the extraordinary capabilities we describe here, but has also completely preserved all the long-standing mainframe software, applications, databases, and staff skills of all IBM's 10,000-plus mainframe customers, a proud investment protection claim that no other IT platform can equal.
- **A pre-integrated data center in one box:** Summarized in Figure 11 above, and discussed here.
- **Highest availability and security:** The reliability, availability, and the rock-solid security, delivered by the System z mainframe are legendary, and each exceeds that of any other type of standard commercial system. System z's advanced, chip-level and dedicated processor, hardware encryption support also gives the mainframe platform far the highest levels of encrypted transaction *(e.g. SSL)* performance of any system. In particular, in our distributed systems consolidation onto mainframe context, the System z offers far-superior levels of RAS as standard than any predecessor distributed system can muster by a wide margin.
- **Lowest total staffing, software, space, power, and cooling:** As discussed above.

Superior System z System/Processor Designs Give Wide-Margin Advantages

Mono-MPU-type x86/x64 servers *(which almost all are)* must use their *(single and only type of)* Intel or AMD MPUs for every aspect of processing. This includes not only the basic application processing, but every I/O access to storage, all networking communication and data transmission, all cryptographic processing, any server virtualization support work, any data compression processing required, and decimal floating point math, etc.. The latter six of these common workload requirements can, and often do, consume and absorb large proportions of an x86/x64 MPU's theoretical processor power/capacity. In addition, because these x64 MPUs must therefore be continually switching between all these types of work, to service them in turn, this significantly slows down the MPU's performance on the real, main application processing job. Whilst it may appear to handle all these types of processing tasks concurrently, the whole software context must be swapped in or out of memory at every task switch, absorbing much of the system's bandwidth and the CPU's capacity on such housekeeping.

By contrast, in a System z mainframe, almost all the I/O processing is completely offloaded onto the SA main I/O processors together with the numerous *(up to 336/system)* additional channel I/O processors. This unique System z hybrid processor architecture is shown and quantified in Figure 11B, indicating a top-end z9 could total up to 416 processors in all. Cryptography is performed either on-chip, or with specialized crypto-processors, or a combination of the two, for blinding encryption performance. Most inter-application networking and communication happens "inside the box" at memory speeds, and virtualization and data compression are supported by super-fast, on-die, dedicated silicon hardware support that is magnitudes faster than an Intel or AMD processor doing the same task through its software.

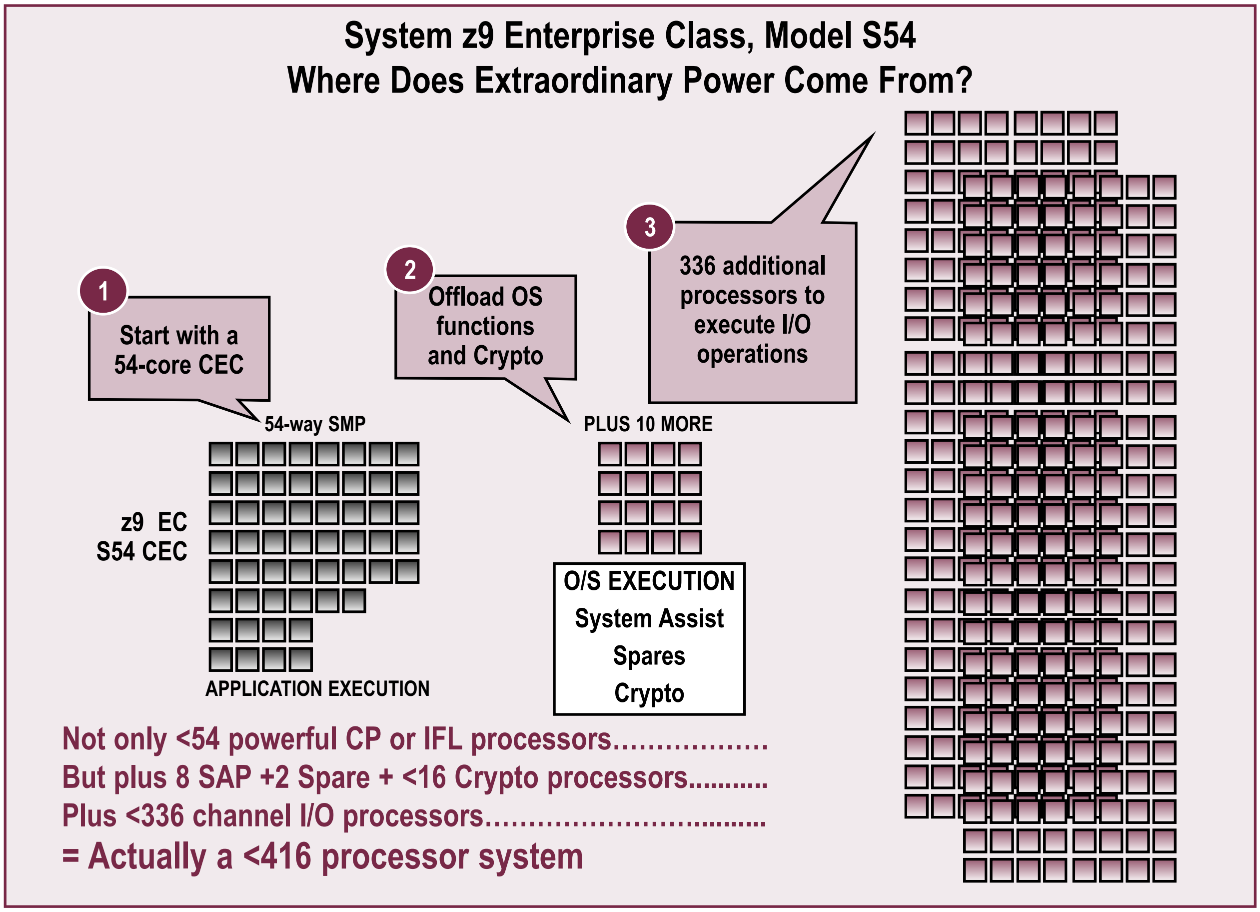


Figure 11B: System z9 Enterprise Class, Model S54 – Where Does This Extraordinary Power Come From?

This is why, in practice, a single System z main CP or IFL processor can do the commercial work of at least 10-20 theoretically quite powerful Intel or AMD MPUs. It is also why one System z IFL can readily run 15-30 significant-weight Linux virtual servers whereas a typical x64 CPU can only support 2-3 VMware virtual servers in practice.

Other profound differences contribute to these wide-margin advantages for System z, and we detail just a couple of these below to reinforce this point:

- **Wasteful Inter-server Communications In Distributed:** In a typical scale-out distributed environment, many different servers are inter-connected, with multiple servers supporting each of the three tiers of logical application processing normally used today. These run the Web serving in the front tier servers, the application serving in the middle-tier, and the database serving in the back-end tier. So intrinsically, every such configuration has a huge amounts of inter-server communication and data-transmission going on across all these "slow wire" links, both between the tiers, and often also to perform the load-balancing within-tier workload distribution to the servers within each tier. As usual, fat TCP/IP software stacks must run on both server ends of each such link, with each server continually chit-chatting with others to maintain communications. Actual TCP/IP planigrams often show an average 65% of all such data transmission is unproductive communications/control data, and only 35% is useful application data. These always-external, cable-connected, and switch or router guided inter-server links are relatively slow, but processing this heavy *(and mostly useless)* inter-server chit-chat absorbs a significant proportion of the entire scale-out configuration's MPU capacity, which is the only place all those fat TCP/IP software stacks needed can be run.

On a System z mainframe, all the application tiers run on LPARs and or virtual servers within one single system, all the inter-application connections use System z virtual networking and virtual switching, inside the box, that all run at memory bus speeds, and which place almost no burden on the main CP or IFL processors, can be dynamically changed, and are far more reliable. The links are far faster, latencies are much lower, and overall application performance far higher and service is more reliable. See also below.

- **Distributed Networking Equipment and Cabling Costs Extremely High:** Often forgotten in comparisons between mainframe and scale-out distributed platforms, the costs of networking equipment/cabling and bandwidth are usually extremely substantial, often either the third or fourth highest cost element, for the distributed alternative. This is because of the extensive numbers of switches, routers, and cabling or fiber links, that are always needed to interconnect up all the tiers and servers, to hook them to whatever storage systems are being used, and to link them up to the enterprise WAN network. Considerable data center floor, and rack, space is also occupied by all this networking gear, which also consumes more power and cooling. By contrast, with a System z solution, most of the networking happen inside the box, and requires no additional equipment, space, power or cooling. Only the links, switches/routers to hook one System z box to the external storage system, and to the WAN are needed.

  One recent, detailed real customer comparison of a large consolidation case that we reviewed put the 5-years networking/connectivity costs for an all-new x64 VMware virtualized distributed server solution at £16M *(12.4% of total solution cost)*, versus £6M for the mainframe *(6.42% of the z total solution cost)*. The distributed solution's networking/connectivity costs here were 2.7* higher than those for the System z solutions, and were a substantial 12.4% and £16M of that solution's £129M 5 year cost. Overall, that example showed the System z solution to be some 35% lower in total costs over that period than the x64/VMware option. The customer's original distributed workload ran on over 2,300 UNIX and x86/x64 servers. The System z solution used five System z mainframes with a total of 255 IFL Linux processors.

System z Mainframe Virtualization Champion

In Section 3, Figure 6 and our nearby assessment, we first introduced the System z mainframe's unrivalled virtualization capabilities. These had been uniquely architected, developed and long-refined to completely virtualize all types of mainframe hardware resource *(CP, memory, I/O, and networking, cryptographic features, cluster coupling processors, etc., etc.)*. The industries' most sophisticated, flexible, dynamic, and industrial-strength mainframe server partitioning is provided by the dLPAR capability delivered by System z's PR/SM hypervisor, enabling up to 60 LPARs per System z mainframe.

System z dLPARs rigorously isolate and secure major OSs, subsystems, applications, and or workloads, on a shared System z mainframe from one another, to the highest EAL5 security rating. Dynamic changes to dLPAR resources can be made when needed, under either operator control, or guided by the System z's amazing Intelligent Resource Director *(IRD)* component, whilst all partitions continue to run. This sophisticated optimizer seeks to balance and optimize the delivery of policy-defined QoS levels to each LPAR running on a System z by reallocating resources dynamically across partitions to achieve the best blend of services delivery and highest overall resource utilization. Many of the world's largest and most demanding commercial IT workloads have run safely and smoothly in System z LPARs for many years now, hence our highest "industrial-strength" assessment above.

System z also supports the z/VM hypervisor that enables hundreds or thousands *(97,943 is the record!)* of Linux virtual servers to be run on a single system z; thus enabling extreme virtualization and consolidation of distributed server workloads.

IBM made coordinated investments in every level of the virtualization technology stack over many years to achieve the current advanced level of virtualization on the System z mainframe. These included engineering virtual server awareness and deep inbuilt hardware support at the chip and system level, extensive support for virtualization in OSs offered for System z *(z/OS, z/VM, zLinux, z/VSE, z/TPF)*, and throughout IBM's System z software portfolio.

- **At the hardware level**, virtualization support is deeply built-in at the MPU hardware instruction level, strong inter-partition communication links are provided, efficient processor and peripheral sharing mechanisms are implemented, the rock-solid System z logical partitioning *(LPAR)* is supported, and all with System z's legendary hardware reliability, scalability, availability, and security.
- **At the hypervisor level**, z/VM offers an efficient, shared-memory-based virtualization model of ultra-scalability, with highly granular resource sharing and simulation, ultra-flexible internal virtual networking, flexible resource controls and usage accounting support. With continuity support for virtual server operations *(virtual server failover)* and an extensive portfolio of fully-featured, mature, server maintenance tools and utilities, z/VM now also provides an unrivalled large-scale virtual server hosting platform for mass consolidation of Linux workloads, as well as for its traditional mainframe site roles.
- **At the application level**, the open, reliable system z OSs are virtualization and virtual-server-aware, to complement and support highly virtualized System z operations. The System z middleware, and the associated IBM System z software tools portfolios, are also engineered to support/exploit System z's virtualization strengths.

Figure 12 summarizes the five broad, unique, functional reasons why IBM can and does, quite legitimately, describe its System z mainframe as today's ultimate virtualization server platform, and what each of those means in business terms.

IBM System z – The Ultimate Virtualization Platform

- **Virtualize everything with up to 100% utilization rates:**
  - CPU, memory, networking, I/O, cryptographic features, system coupling facility, ...
- **Massively scale workload on single System z mainframe**
  - Linux-on-z/VM record is 97,943 virtual machines
  - Each z/VM virtual machine can access < 24,576 devices
- **Security for everything:**
  - World's highest general purpose server security classification
  - System z LPAR technology is EAL 5 certified
- **Non-disruptively add anything:**
  - 54xCPU scalability/mainframe, 32x CPU scalability/z/VM LPAR
  - z/VM designed to support < 8 TB of active virtual memory
  - <32 System z mainframes/Sysplex cluster
- **Optimize and integrate it all with the IBM software portfolio**
  - Comprehensive, leading-edge IBM System z middleware and tools

Consolidate all types of workloads.

Best economics. Start small-grow big in one box.

Secure virtual servers and slash business risk.

Rapidly respond to workload spikes.

Link and connect all application types.

Figure 12: IBM System z – The Ultimate Virtualization Server Platform

The mainframe's ability to **completely virtualize everything**, and to attain hardware utilization rates of up to 100%, enables System z to efficiently, **economically consolidate a wide range of workload types** on a single system.

Massive-scale consolidation on single System z mainframe can be achieved using z/VM's extreme virtualization scaling capabilities, but customers can **also start small at low cost** by just adding small numbers of virtual servers on an existing System z.

Highest security assured by EAL 5-certified LPAR technology and other deeply-embedded security, which enables **virtual servers to be fully and robustly secured** throughout.

All types of additional mainframe **resources can be non-disruptively added** to a running System z, to scale-up capacity over a huge capacity-scaling range, to meet new levels of permanent **demand growth**, or shorter-term **workload spikes**.

Within the last couple of years, IBM has completed a massive effort to deliver and update to the latest SOA technology standards, a quite unrivalled mainframe middleware server software and tools portfolio. We assessed *(in other studies on page 50)* this IBM portfolio as having attained its best-ever strength and breadth in the platform's 43-year history by 2007. We also found that it rates as a well superior stack to the software available for competing platforms. The wide-ranging, world-leadership-class, mainframe software advances crucially enabled the success of many new-to-mainframe growth workloads *(Linux Java/J2EE™, ERP applications, and mass-consolidations, etc.)* that account for much of the renewed growth the platform has enjoyed.

IBM z/VM V5.3 Hypervisor – Extreme Virtualization Enabler for Mass Distributed Server Consolidation

Amongst the central virtualization "crown jewels" on the System z mainframe is IBM's z/VM hypervisor. In its latest release, z/VM Version 5.3, z/VM *(US pronounce "zeeVM")* is the industry's most sophisticated, advanced, scalable, and long-running, bare-metal hypervisor. z/VM is the direct descendant of a 40-year family evolution that began with the famous CP67 mainframe timesharing OS of the earliest System/360 era. The z/VM hypervisor had long been traditionally used on mainframes to run multiple OSs, and/or multiple OS versions, and/or test and development environments, in VMs on separate partitions of the same mainframe that was also running the production workloads. It is, however, also highly capable of running/hosting large numbers of Linux virtual servers on a single z/VM copy. Thousands of these Linux virtual server workloads are now being migrated onto highly consolidated z/VM environments on System z mainframes. These increasing numbers of enterprise IT users can escape from the inefficient, wastefully underused, distributed RISC-UNIX or x86/x64 server platforms that these workloads have been run on to date; gaining huge savings in the process.

Combined with the System z hardware, other OSs, and extensive IBM middleware and tools software for the platform, z/VM is the crucial enabler for such extreme virtualization. The hypervisor enables users to run from 100s to a 1,000 virtual servers *(of significant weight)* under Linux-on-z/VM, to better process workloads mass-migrated from distributed servers far more cost-effectively.

We more deeply assess these large-scale System z z/VM mass-consolidation configurations, comparing why, how, and where they provide superior solutions than the familiar, low-end VMware on x86/x64 virtualization approach later in this Paper.

IBM's Unified Systems Management Suite Fully Supports System z Extreme Virtualization

One of the most important aspects of efficiently operating physical and virtual server environments is the systems management software tools provided by the vendor. IBM Systems now offers a single, comprehensive, ultra-modern, and unified system management tool suite. This highly-capable tool suite, IBM Systems Director, is a first-class, elegant, and comprehensive system manager designed to manage both physical and virtual platforms, to manage servers, storage, and networking equipment, and which now supports all the IBM server and storage systems, as well as most major third-party platforms. An overview of IBM Systems Director is shown in Figure 13.

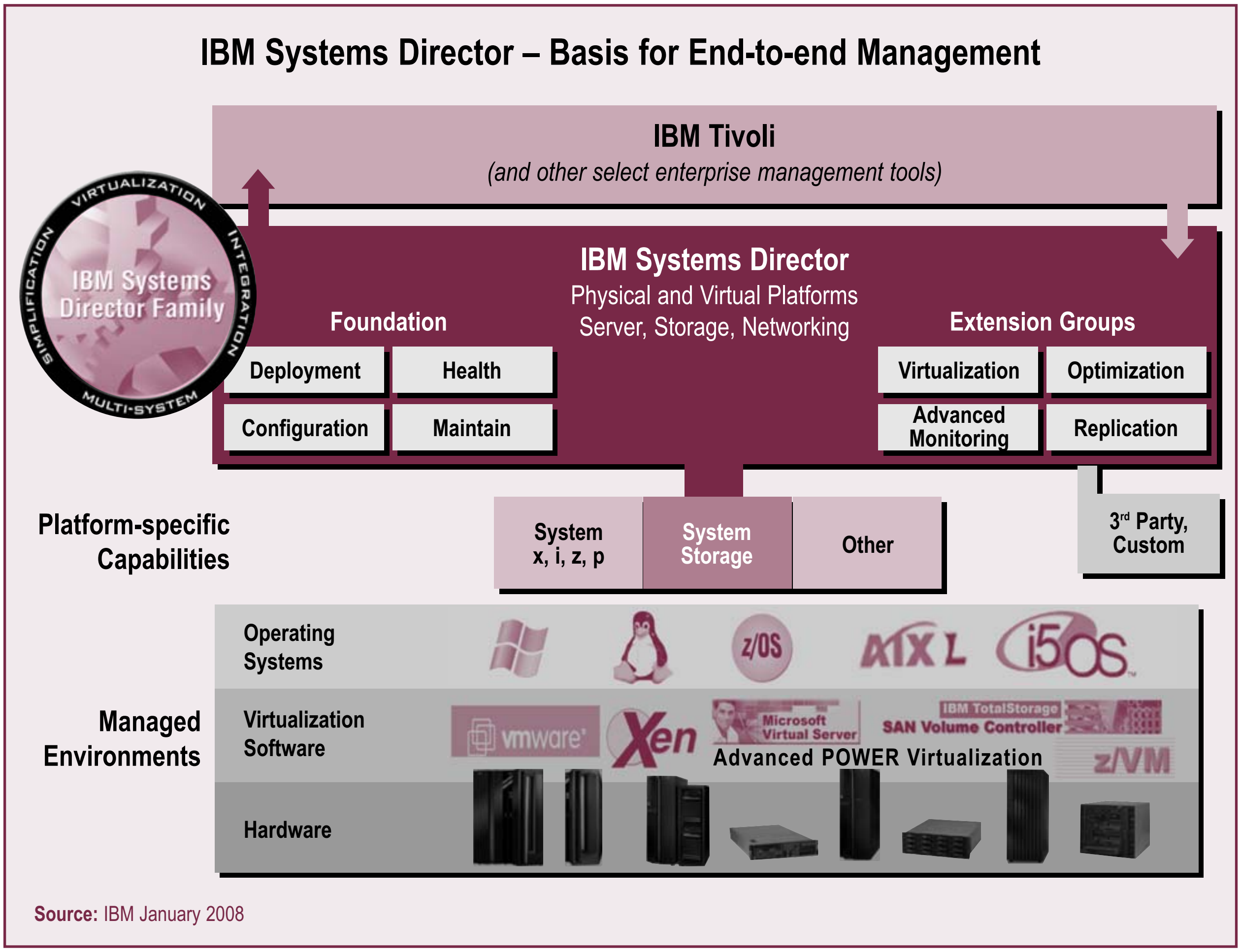


Figure 13: IBM Systems Director – Basis for End-to-end Management

The tool suite incorporates standard, foundational capabilities to manage the system configuration, deployment, maintenance, and ongoing health-checking management processes required in every systems environment. It also offers optional, additional extension groups with more advanced management functionality for virtualization, replication, optimization, and advanced systems monitoring. Specific, in-depth, and full-feature exploitative support for the IBM System z, System p, System i, System x, and BladeCenter server platforms, and for the IBM System Storage systems line, are included, together with basic support for major third-party platforms. All the major OSs offered on the IBM Systems families are also supported.

IBM Systems Director also manages leading virtualization hypervisor platforms, including IBM z/VM *(System z)*, IBM APV *(System p & i)*, VMware ESX Server *(System x & BladeCenter)*, Xen, Microsoft Virtual Server, and the IBM SAN Volume Controller *(IBM System Storage & third party)*.

IBM Systems Director was substantially further developed and extended from the sound basis of IBM's popular and well-regarded earlier IBM Director family, originally the system manager for IBM's System x *(x86/x64)* and BladeCenter platforms. IBM Systems Director can also integrate up into the complete IBM Tivoli enterprise management suite, including the IBM Tivoli Enterprise Portal's *(ITEP's)* complete enterprise management view. Other enterprise management suites *(HP, CA, and BMC)* can also be linked.

Our Analysis

The overview above provided a concise summary of our research-finding highlights on the System z9 mainframe. It also provided as much information as we can currently offer on its imminent, next-generation successor mainframe family, expected in Q1 2008, within the context of how these are likely to extend the System z mainframe's extreme virtualization and consolidation capabilities and capacities.

With some c. 28M distributed servers currently deployed globally, about 28,000 such new, large, z6 MPU-powered System z10 mainframes *(at the consolidation ratios above)* would be needed to consolidate the entire workloads of all c. 28M current standalone distributed servers *(assuming all their current work could technically be migrated)*. If we more cautiously assumed that 50% of total distributed-server workloads are probably well suited to mainframe consolidation, just 14,000 new, z6-MPU-powered generation large mainframes could likely carry this complete load.

This would rid the world, once and for all, of at least half of the existing scourge of distributed computing.

This would rid the world, once and for all, of at least half of the existing scourge of distributed computing. This could save global IT users of the scores of $Bs in ongoing IT infrastructure costs over the decade ahead, and would substantially help human-kind save our planet by slashing IT carbon footprints by large percentages.

For IBM, installing 14,000 additional new mainframes *(on top of its normal run-rate mainframe sales)* would require a big increase in current mainframe production were this strategic move to be achieved within a desirably short time, of say five years.

Success at this level would double the current c.15,000 mainframe systems' global installed base again, and would probably quadruple or quintuple the total installed IBM mainframe MIPS capacity, with the additional, much larger MIPS, new z6-based systems needed.

Enormously beneficial environmental/energy reduction savings would be delivered by mass-migration to the far-more-efficient and environmentally friendly mainframe platform. Top-level national and international political and environmental support for such migrations, can therefore be expected, from boards of directors, governments, and from multi-national organizations.

5. System z Linux-on-z/VM Extreme Virtualization & Consolidation

Introduction

In this Section, we review special System z Linux hardware that allows economically-attractive, mass-distributed server consolidation onto the platform. We then examine Linux strengths, evolution, and success on System z. This grew a now-broad portfolio of Linux mainframe software *(of all types)* that originally ran on other *(usually distributed)* platforms. This portfolio now a supports wide-scale workload migration onto System z software versions, off the original distributed platforms *(UNIX & x86/x64)* where these products ran before. We also take a deep look at IBM's z/VM hypervisor, System z's unique "magic ingredient" for extreme virtualization and mass server consolidation. We finally compare and contrast what z/VM on System z offers today beyond VMware's ESX Server-based virtualization offerings on their native x86/x64 platform.

What do we Mean by Mass Consolidation, Extreme Virtualization?

Newer terms like these are often best be clarified and illustrated by a customer example. A generic reference example illustrating these points is shown in Figure 14. Later, in Section 6 we also review/assess four, real-world, distributed server consolidations to z/VM on System z examples, in considerable detail, providing additional insight and real experience feedback from customers who successfully escaped to System z safety off their perilous and costly distributed server sprawls.

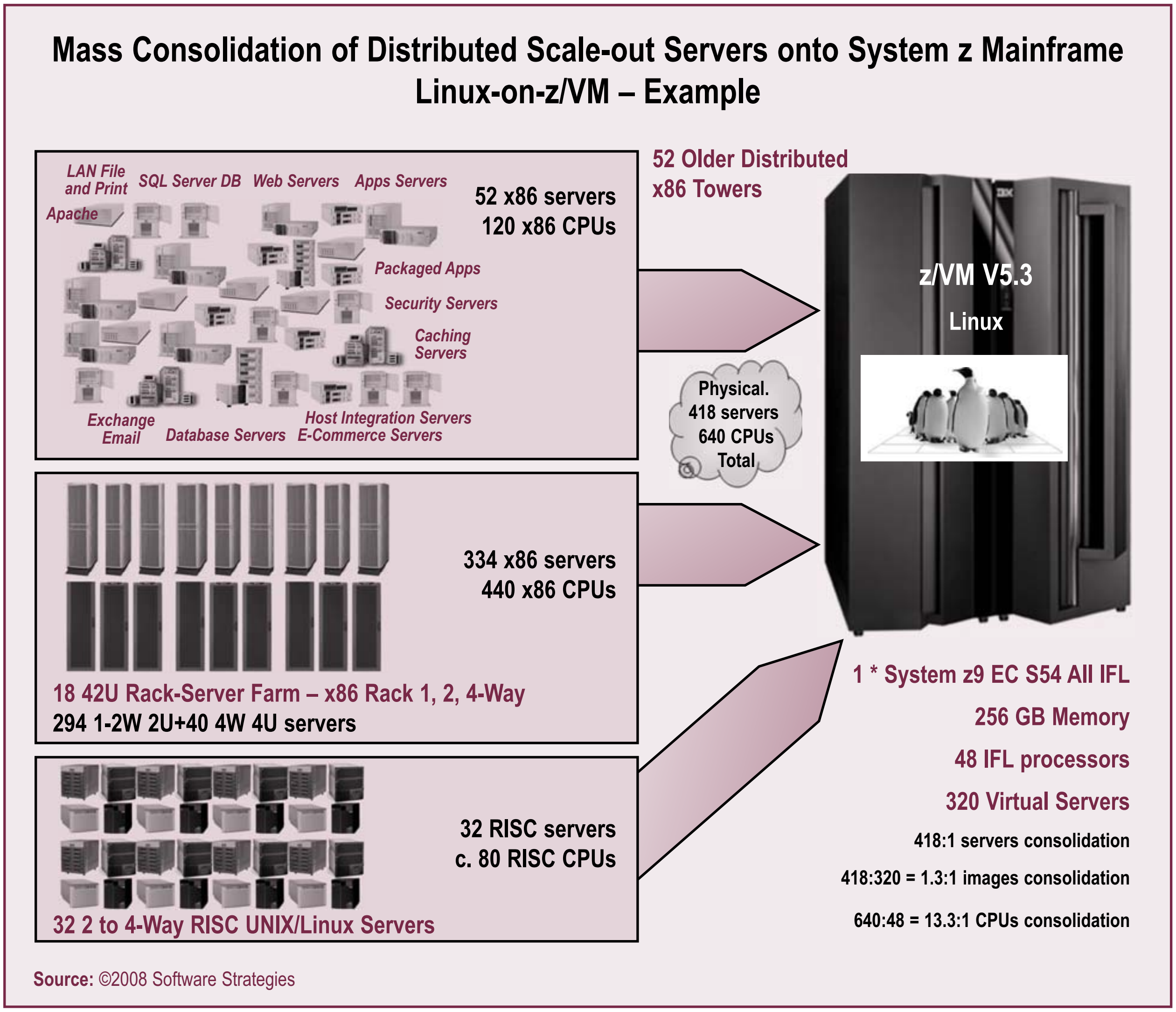


Figure 14: Mass-consolidation of Distributed Servers onto System z Linux-on-z/VM – An Example

This customer "triaged" their existing scale-out distributed server estate into three categories. The first were distributed servers/workloads they decided *(on a number of criteria)* should best be consolidated onto a System z mainframe under Linux-on-z/VM. The second triage category, not shown in Figure 14, were distributed servers they decided should best remain "as-is" on distributed platforms, but with new/refreshed hardware when next due. The third and final triage category were distributed servers/workloads now due/overdue for retirement. These they adjudged not worth either migrating, nor worth retaining, in present form. These were thus to be retired completely with superior, replacement software services assuming their roles on new hardware.

The first triage category, for System z migration and thus of interest to us here, contained the three sub-groups shown left in Figure 14, viz.:

- **52 older distributed servers**, mostly x86 tower types, totaling 120 CPUs, running typical x86 client/server infrastructure workloads, such as collaboration/mail, file and print, and Web-serving, etc., on a mix of Linux and Windows. "Light" Linux virtual servers, running Linux open middleware on z/VM on the System z, took over these workloads. These servers were older machines, last replaced between 2000-2002. These were already well past their service EOL dates, and thus past overdue for replacement anyway.
- **336 newer, rack-dense x86 servers in an 18-data-center-rack-server-farm-format.** A mix of 1-, 2-, & 4-Way servers with a total of 440 x86 CPUs currently. These servers ran larger mid-tier, customer J2EE™ applications, and mid-weight customer database applications, using IBM WAS and DB2 middleware respectively. "Medium" or "Medium-plus"-weight Linux virtual servers, running WAS and DB2 middleware for z Linux, assumed these workloads *(easy migration)*. Mostly deployed in the 2003-2004 period, these servers were also old hardware, then at their normal EOL point.

- **32 RISC-UNIX servers**. 2- to 4-Way servers, running mySAP application and database server roles, underpinning an enterprise-wide mySAP deployment rolled out in early 2005. These server workloads migrated easily to mySAP Linux virtual servers on the same WAS application server and Oracle DB middleware on z Linux, and were also of “Medium” and “Medium-plus” weight. These systems reached their 3-year life span EOL dates 3 months after the above groups, so were moved last to the System z at life-expiry of their old UNIX hardware, which was then sent for recycling.

These three groups of z/VM migration-selected, distributed servers were using 418 physical servers, with 640 assorted x86 and RISC CPUs. They were migrated to one System z9 EC S54, Linux-only model, with 48 IFL Linux specialty processing engines, running the z/VM V5.3 hypervisor. The customer had also determined that software image consolidation was also both possible and desirable. 320 Linux virtual servers replaced the 418 previous physical servers. This gave a physical server consolidation ratio of 418:1, a logical server consolidation ratio of 1.3:1, and a CPU consolidation ratio of 13.3:1. The new mainframe runs the initial total of 320 virtual Linux servers on its 48 IFL processors, at a modest 6.67 virtual servers per IFL processor, because of the fairly substantial workloads amongst the much larger, migrating second and third server groups here.

All the existing physical distributed servers here were past, at, or near their service/depreciation 3-year EOL points, all had zero residual capital value in this customer’s accounts, and carried the near-zero re-marketing value typical of such relatively old, used distributed servers. Had the customer replaced all these distributed servers with an “exactly-as-is-with-latest-equivalents” hardware refresh, they would have needed to repurchase the entire hardware stack again afresh in any event.

By switching to the mainframe solution, this customer now expects to save some $12M over the next five years. Their business case showed 35% of these saving coming from large middleware software license and S&S cost savings *(48 IFLs System z vs. 640 CPUs distributed)*. 40% came from support staff cost reductions, *(down from 21 support FTE distributed to 4 FTE on System z)*, and 12% from reduced connectivity/enterprise networking costs. 5% came from lower power/cooling/data center space, lower security costs, and other smaller savings. In addition, their System z solution is now significantly more reliable and available, delivers much-improved performance and response times to end users, is far more secure, and also now shares full DB/BC provision with the customer’s’ existing System z system at relatively affordable cost. All these advantages have financial value additional to the above base savings.

Since major capital would have needed to be spent anyway to replace the EOL distributed systems, it could easily be argued that most of these $12M mainframe solution savings are pure gain. This hypothetical example was based on average actual ratios and saving percentages. Like each of the four, named-customer, real-world success profiles assessed in Section 6, it makes a convincing business case.

We also present a full financial analysis of a rather larger, actual customer case in Section 6, Figure 21.

Introducing z/VM – Premier Mainframe Virtualization Hypervisor

z/VM is IBM’s premiere, software hypervisor-based virtualization environment for the System z mainframe, offering ultra-efficient management of large numbers of virtual servers *(VS)*, and itself usually running in a secure, mainframe LPAR hardware partition. z/VM fulfils several important roles in helping users extend their mainframe’s business value widely over their enterprise. It lets them integrate applications and data closely on mainframe-hosted Linux virtual servers. These can easily network *(inside the box)* with traditional mainframe applications, and with each other, whilst providing exceptional levels of availability, security, and operational ease. Historically, customer IT groups also long used z/VM to securely host fully isolated mainframe test environments for new OS and middleware releases, to host new mainframe application development and testing, and to support applications/workloads needing separation/isolation for other reasons. Today, z/VM can also now manage/host from 100s significant weight up to a 1,000 virtual Linux servers on a single IBM mainframe used as a large-scale, Linux-only, hosting server. z/VM can also run alongside other System z OSs, such as z/OS, and can also improve site productivity by hosting non-Linux workloads such as z/OS, z/VSE, and z/TPF.

z/VM today employs highly-sophisticated resource utilization algorithms to best share System z hardware resources, and to optimally manage virtual servers in large numbers. Weights and priorities can be assigned to guest virtual servers/machines to determine how allocation of resources will be done. High-priority guests receive resources before low-priority guests. The system’s real processor, memory, I/O and networking resources are completely shared by the guest virtual servers as needed. z/VM can also substantially over-commit resources. It could *(for example)* easily deploy 4 VS needing 8GB of memory each, but actually using much less than 32GB of real memory for all four. The same applies to the allocation of CPU resources. This allows z/VM to drive IBM mainframes up to the ultra-high utilization levels *(up to 95-100%)* they regularly run at both smoothly and safely.

Unlike other hypervisors, z/VM was designed/optimized to *(and does)* easily runs memory-rich and I/O-intensive *(disk and network)* workloads with both extreme data integrity, and with excellent performance. *(These include heavier application server, and database server, workloads.)*

Hardware: The System z IFL Specialty Processor for Linux

One critical success factor underpinning the fast uptake, and the soaring workload growth success of Linux on the System z mainframe, was IBM's 2000 debut of the **Integrated Facility for Linux** ***(IFL)*** specialty processor. These IFL processors ran all Linux workloads. They had the same capacity *(up to 580 MIPS on System z9 EC)*, and access to all the same system hardware capabilities, as standard mainframe CP processors, but were free of regular CP-processor-capacity-related software charges. Priced at far lower-than-standard mainframe CP processors, IFLs have been offered at flat initial purchase prices of $95K *(z9 BC)* and $125K *(z9 EC)* each.

When used with z/VM, one such mainframe IFL processor alone *(supported with its access to often-huge mainframe-shared memory, massive I/O capacity, and huge internal networking capability)* can often smoothly run 15-50 virtual servers, depending on their "workload weights". One IFL, over five years, costs $33K *(z9 BC)* to $43K *(z9 EC)* per annum *(initial capital cost and 4 years maintenance)*. This means the IFL hardware cost per year per virtual server, can be as low as the $790-$2,750 p.a. *(z9 BC)* and $860-$2,900 p.a. *(z9 EC)* ranges, an economical hardware platform, but one delivering outstanding QoS levels far beyond more-costly distributed alternatives.

An existing System z mainframe user can just add a single IFL, plus z/VM and Linux, and start distributed server consolidation relatively modestly, with capacity to consolidate their first 15-50 virtual servers, sharing all their mainframe resources with other existing workloads, a helpful "start-small" route.

Linux-only, All-IFL System z Mainframes

Another important 2000 introduction was the all-IFL System z mainframe. These run purely Linux workloads *(and thus cannot support traditional mainframe workloads at all)*, have no general-purpose CP processors, and cannot run the z/OS flagship OS. With favorable IFL pricing eliminating traditional mainframe capacity-to-software-fees linkages, all-IFL Linux-only System z mainframes are highly competitive in hardware pricing today. They can also scale up to the highest capacity level, System z9 S54 54-IFL mainframe of 17,800 net SMP MIPS total capacity, and capable of supporting low hundreds to a thousand virtual servers *(depending on their "weight")* when running Linux-on-z/VM extreme virtualization. Linux-only, all-IFL-configured, System z mainframes would usually be adopted for mass-distributed server consolidation projects, as was the case in the three larger of the four success profile examples assessed in Section 6.

...providing the crucial capability to run hundreds of Linux virtual servers within one z/VM image under an LPAR.

Linux can run in a System z logical partition, and can communicate with other Linux partitions via an internal TCP/IP socket connection, without using an external network. Linux also takes full advantage of the System z's extensive, multiple-level virtualization capabilities. Hardware-level virtualization includes dynamic logical partitions *(dLPARs)*, allowing a System z mainframe to currently be divided into up to 60 LPARs. Linux can run directly "on the metal" in an LPAR. z/VM provides the second major level of software-hypervisor-based virtualization, providing the crucial capability to run hundreds of Linux virtual servers within one z/VM image under an LPAR. Hardware circuitry assists/accelerates the execution of both the z/VM software and the LPAR with hardware virtualization support, with near-zero overhead.

Linux on System z – Powerful Advantages Combine

Linux enjoyed outstanding market growth/adoption as a server OS over this decade. It now holds third place, with a 13.3% 1H 2007 server OS market share, at $3.4B in factory server revenue, up a healthy 11.0% YOY. For comparison, # 1 OS Windows held a 38.4% share at $9.8B revenue, and # 2 OS UNIX a 32.2% share at $8.2B revenue, for the same period.

Since 2000, running Linux on the System z mainframe succeeded far beyond early expectations. As 2008 begins, over 2,500 mainframe systems already run Linux, over 16% of the platform's installed base. Whilst Linux can be run natively on System z hardware, it is usually run in an LPAR, or under the z/VM extreme virtualization hypervisor *(the option of prime interest in this White Paper)*.

One Linux on System z factor is the ever-stronger capabilities, increasing scale, first-class reliability, and performance of Linux itself, which has continued to mature its enterprise-class OS credentials. This evolution has now positioned Linux as a realistic replacement server OS, well able to safely assume many workloads from both traditional UNIX OS running on their costly, proprietary RISC servers and from larger x86, x64, or Itanium 2 platforms, running Windows applications or Linux workloads.

Linux fitted well onto the System z mainframe platform, bringing many new UNIX-style workloads, and has now been extended to fully exploit unique System z capabilities and QoS. When the 64-bit headroom, large SMP compute capacity, and the massive I/O capabilities, of a System z9 are combined with high-quality, fully-supported, 64-bit enterprise Linux distributions *(RHEL and Novell SLES)*, a open, and flexible, new OS/hardware platform combination results.

If we combine:

- The above **deep Linux OS strengths**.
- With **Linux System z hardware platform affinity**.
- With **low-cost, System z, Linux-only mainframe hardware** *(see above)*.
- With a now-broad **Linux-on-System z software portfolio** *(discussed below)*.
- Using the **extraordinary z/VM hypervisor**.

…we have a strong foundation for mass-distributed server consolidation onto System z. This provides a unique way to simplify IT infrastructures by consolidating hundreds of distributed servers and their networks, fully exploiting excellent System z QoS levels on the workloads, and exploiting the z/VM management tools/features assessed below. Linux-on-z/VM also brings a major improvement in speed-to-market and flexibility, because new servers, networks, and solutions can be provisioned and deployed with amazing speed "inside the box", sometimes in just minutes. Resource allocations can also be changed dynamically, and business responsiveness and flexibility are much improved.

Combined, the z/VM capabilities *(see below)* along with those of Linux on System z itself, allow Linux-on-z/VM to offer advanced, extreme virtualization and virtual server management capabilities that no other system or hypervisor can approach.

Wealth of Linux Software Expands System z Software Portfolio, Enables Consolidation

Linux on System z now offers a wide range of top software – including middleware, tools, and applications – across most segments and categories, from most major ISVs. The success of Linux itself had made the OS a new growth market target for these ISVs. Once their software was out on Linux, a fairly minor further step for an ISV then allowed a System z Linux release to be created/offered. This has made available a crucially wider, still-growing, range of software on System z Linux, including:

- **A wide range of enterprise-quality middleware** servers, including not only all major IBM servers such as WAS, WPS, DB2, WMQ, and ESB, etc., but also some from Oracle, and others.
- A growing **enterprise applications/business packages collection**, from scores of ISVs, large *(including SAP and Oracle)* and small, who moved their products onto System z Linux from other Linux versions.
- Many popular, **open-source Linux IT infrastructure servers and tools** – for example, the Apache Web Server.
- Many **originally UNIX-only software products**, in all categories, previously offered only on the AIX5L, Solaris, or HP-UX proprietary UNIX OSs, but which have now migrated to the close-affinity and UNIX-like Linux cousin, a fairly minor port for their ISV owners. Once onto Linux, System z is a smaller, final step many ISVs have taken, in order to benefit from resurgent mainframe sales as their UNIX platform businesses slow.
- Many leading ISVs of **originally Microsoft Windows server OS** middleware, tools, and applications software did well over the last decade from the huge growth of the x86/x64 distributed hardware market, which dragged many Windows software sales for them. Many of these ISVs recognized the growing size and value of the Linux market, and brought their originally-Windows-software products onto the Linux platform. Microsoft, with its Windows middleware suite was, unsurprisingly, a notable exception here.

In this Paper's mass-consolidation context, these developments now mean a wide range of premier software, long deployed on *(and filling much capacity of)* enterprise customers' scale-out distributed server estates, can now be directly supported by appropriate Linux versions of the same software on System z Linux-on-z/VM. If not the same software, functionally-comparable alternatives *(e.g. SQL Server to DB2 database, MS Exchange to Lotus Domino, IIS to Apache Web server)* that are offered on System z Linux can often be reasonably and easily substituted. It is this wealth of now-available z Linux software that enables more mass-consolidation opportunities *(from distributed servers to System z mainframes)* for larger numbers of customers.

The IBM SWG zLinux product portfolio has also mushroomed from "zero to hero" from 2000 to 2007...

The IBM SWG zLinux product portfolio has also mushroomed from "zero to hero" from 2000 to 2007, and today offers an excellent set of middleware offerings, compatible with its UNIX and Windows versions widely already deployed across the globe. This makes for especially easy migration of applications currently running on this IBM middleware on UNIX and Windows systems onto the same IBM middleware servers running on Linux-on-z/VM.

z/VM Major Features and Capabilities

z/VM is such a rich and extensive virtualized operating environment as to be quite difficult to summarize/assess briefly in a few short pages, as we must do here. In order to accomplish this, we first tabulate/explain a dozen principal z/VM features/capabilities in our densely packed Figure 15 table below.

Major z/VM V5.3 Hypervisor Features and Capabilities

z/VM Feature	Feature/Capability Purpose. Needs, Problems, or Issues Addressed	Feature Detail and Benefits
z/VM CPU Resource Controls	Need flexible, fine-grained mechanisms for allocating, and for controlling the usage of, processing resources to/by all hosted Linux virtual servers in z/VM hosting environments. The broad goals here are to deliver the best VS performance.	• Highly granular sharing of system processing resources. • Allocate system resources per guest image using SHARE command. • Highly flexible, self-managed z/VM Control Program *(CP)* function. • Can reserve CPU capacity for peak usage, and use it only when needed. • Can relinquish processor cycles for other servers when not needed. • "Absolute guests" receive top priority ahead of "Relative guests". • **Virtual Machine Resource Manager** *(VMRM)* may monitor above, and adjust remaining capacity allocated to "Relative guests". • Can set "Hard" and "Soft" resource limits. Latter can be exceeded if un-used capacity is available, the former cannot.
Co-operative Memory Management *(CMM)*	When virtual memory use far exceeds real memory availability, z/VM Control Program *(CP)* paging becomes excessive. Overall z/VM system performance and guest VS throughput suffers. CMM eases this.	• Linux virtual servers/images are signaled to reduce their virtual memory consumption. • Linux memory pages are released. • Demands on real memory, and on the z/VM paging subsystem, are reduced. • Helps improve overall system performance and guest image throughput.
Collaborative Memory Management Assist *(CMMA)*	Addresses common need to help z/VM host larger numbers of virtual servers in the same amount of memory in the Linux-on-z/VM environment.	• Extends co-ordination of memory and paging between Linux and z/VM down to individual page level, using a new System z hardware assist *(CMMA)*. • z/VM knows when a Linux application releases a memory page. • Host Page-Management Assist *(HPMA)*, with CMMA, further cuts z/VM work in resolving page faults. • Helps z/VM host more virtual servers in the same amount of memory. • CMMA is supported on System z9 and in z/VM V5.3. • IBM is working with z Linux distribution collaborators on exploitation support.
Linux Exploitation of z/VM Discontiguous Saved Segments *(DCSS)*	In z/VM Linux virtual server hosting, many or all virtual servers may use one or several of the same software executables. It is excessively wasteful on disk space, and inefficiently slow, for every VS to host its own copy of such shared program executables, and for these to be loaded from disk to memory for every execution.	• Shares a single, real memory location among multiple virtual machines/servers. • Provides high-performance data access to shared programs. Can also reduce real memory utilization. • In Linux exploitation, shared program executables are stored in an execute-in-place file system, then loaded into a DCSS. • DCSS memory locations can reside outside the defined virtual machine configurations. • Access to file system is at memory speeds; executables are invoked directly out of the file system *(no data movement required)*. • Avoids duplication of virtual memory and data stored on disks. • Helps enhance overall system performance and scalability.
Linux Exploitation of z/VM Virtual Disks in Storage *(VDISK)*	Faster disk I/O operating speed is always desirable and a high-speed swap mechanism highly desirable for good Linux performance.	• VDISK support is another z/VM Data-in-Memory technology that simulates a disk device using real memory. • Achieves memory speeds on disk I/O operations. • VDISKs can be shared among virtual machines. • Linux exploitation: VDISK is used as Linux high-speed swap device • Using VDISKs as Linux swap devices instead of real disk volumes: – Reduces demand on the I/O subsystem. – Helps reduce usual performance penalty of swapping operations. • Excellent configuration tool that helps clients to minimize memory footprints required for virtual Linux servers. • Helps improve the efficiency of sharing real resources among virtual machines.

Continued on next page...

...continued from previous page

System z N_Port ID Virtualization *(NPIV)* z/VM support	Without NPIV, hosted Linux images could access all LUNs accessible to the real hardware channels, not wanted/nor secure. With NPIV, each Linux image is separately authorized, via zoning and LUN-masking, with a unique WWPN for each sub-channel or virtual host-bus adapter.	• System z9 FICON Express feature supports FCP N_Port ID Virtualization. • NPIV enables zoning and LUN masking on a virtual machine basis. • Multiple OS images can now concurrently access same or different SAN-attached devices *(LUNs)* via a single, shared FCP channel. • Increases channel utilization, so less hardware may be required. • Helps reduce complexity of physical I/O connectivity. • Supported by z/VM V5.3 and V5.2; z/VM V5.1 supported with PTF.
z/VM Support for Parallel Access Volumes *(PAVs)*	Enabling/providing concurrent I/O to the same disk volume by one or more users/virtual server, user, or job is highly desirable for faster I/O response time and for lower disk device queuing delays.	• PAVs allow multiple concurrent I/Os to same disk volume by one/more virtual servers, users, or jobs. • PAVs offer automatic, co-ordinated Read and Write I/O referential integrity when needed. • z/VM supports PAVs as minidisks for guest OSs exploiting PAV architecture *(e.g. z/OS and Linux for System z)*. • z/VM also provides potential PAVs benefits for I/O to minidisks owned/shared by guests OSs not natively exploiting PAVs *(z/VSE, z/TPF, CMS, or GCS)*. • DASD volumes must be defined to z/VM from supported IBM disks list. • Improves I/O response times by reducing device queuing delays.
IBM System Storage SAN Volume Controller *(SVC)* Software V4.2	z/VM and Linux for System z support for SAN Volume Controller *(SVC)* V4.2. Needed to exploit this IBM powerhouse storage virtualization appliance solution in the mass Linux virtual servers hosted on z/VM environment.	• SVC allows z/VM and Linux access to multiple vendor's SCSI arrays as a single pool of disk capacity. • z/VM FBA emulation lets CMS users access SVC-managed disk space. • New functions in SVC V4.2 include: – Multi-target FlashCopy support *(up to 16 images)*. – Higher number of active FlashCopy relationships at the cluster level. – Designed for improved cluster performance, especially when installed on IBM System Storage SVC 2145-8G4 storage engine. • Support for additional OEM storage devices. • Supported in z/VM V5.3 base product, z/VM V5.2 supported with a PTF.
z/VM Virtual Networking – Point-to-point Comms	Need flexible "within the System z box" point-to-point networking, between z/VM supported Linux servers, and too/from external network*(s)*, without external physical switches, routers, cables, etc. being needed.	• z/VM virtual networking is memory-to-memory, i.e. very fast. • Can define multiple sub-nets and virtual routers in z/VM virtual networks. • IUCV or virtual ICTA used to connect Linux virtual servers to virtual routers. One OSA sub-channel can serve many Linux images. • Virtual routers can access the physical *(real)* network. • Low cost, highly flexible, top-performing, virtual PTP networking capability.
z/VM Virtual Networking – z/VM Guest LANs	Need powerful, flexible virtual networking for z/VM guest virtual LANs too.	• Guest LAN is a "virtual" LAN created by the z/VM Control Program *(CP)*. • OSA Express *(QDIO)* and HiperSockets Guest LANs can be created. • Point-to-point, Multicast, and Broadcast *(QDIO)* connections are supported. • Linux images can connect to one or more Guest LANs. • Can also connect to real network adapters at the same time. • Enables a Linux image to provide external routing and firewall services for other Linux images.
z/VM Virtual Networking – z/VM Virtual Switch	Need virtual switch capability *(L2 and L3)* to build operate z/VM virtual networks without physical routers.	• Eliminates router else needed to connect VSs to physical LAN segments. • May reduce usual overhead with router virtual machines. • Allows VS/VMs to be in the same sub-net with physical LAN segment. • Supports Layer 2 *(MAC)* and Layer 3 *(IP)* switching. • Includes support for IEEE VLAN • Provides centralized network configuration and control. • Easily grant and revoke access to the real network. • Dynamic VLAN topology changes can be made transparent to virtual servers.
z/VM Virtual Switch – Link Aggregation Support	Need scalable, high-capacity, virtual networks with failover support for DR/BC.	• Non-disruptive networking scalability and failover with z/VM V5.3 • Enhances networking bandwidth and Business Continuance *(BC)*. • Uses LACP *(Link Aggregation Control Protocol)*. • Up to 8 OSA ports per VSWITCH. • Requires OSA-Express2 support available with IBM System z9 servers. • Supports recovery of a failed switch where backup OSA configured.

Figure 15: Major z/VM V5.3 Hypervisor Features and Capabilities

These major capabilities are sufficiently described above that we add no further assessments here. An exception is that we do note the sophisticated and advanced nature of many of these features, indicating the strength/depth z/VM has attained as a strategic operating platform and hypervisor today, after its many years of intense development/refinement/extension.

We would particularly stress our assessment that Linux-on-z/VM on System z offers both **outstanding resource sharing** *(of all system resources – CPU, memory, I/O, and networking)* **as well as extreme scalability**. The latter includes the ability to both scale-up vertically *(add more z/VM instances/LPARs)* and scale-out horizontally *(add more virtual servers to each z/VM instance/LPAR)* on the same System z server dynamically as needed.

Other Powerful z/VM Features and Capabilities

Although we already noted some principal features and capabilities, and the associated benefits, of IBM's z/VM hypervisor on System z above, many other notable and important differentiating capabilities mark this rich, extreme virtualization platform. For another, broader-ranging view of z/VM's panoply of capabilities/features, we concisely summarized our "top next 17" of these in our Figure 16 chart, each feature/capability tersely amplified in the note beneath its heading.

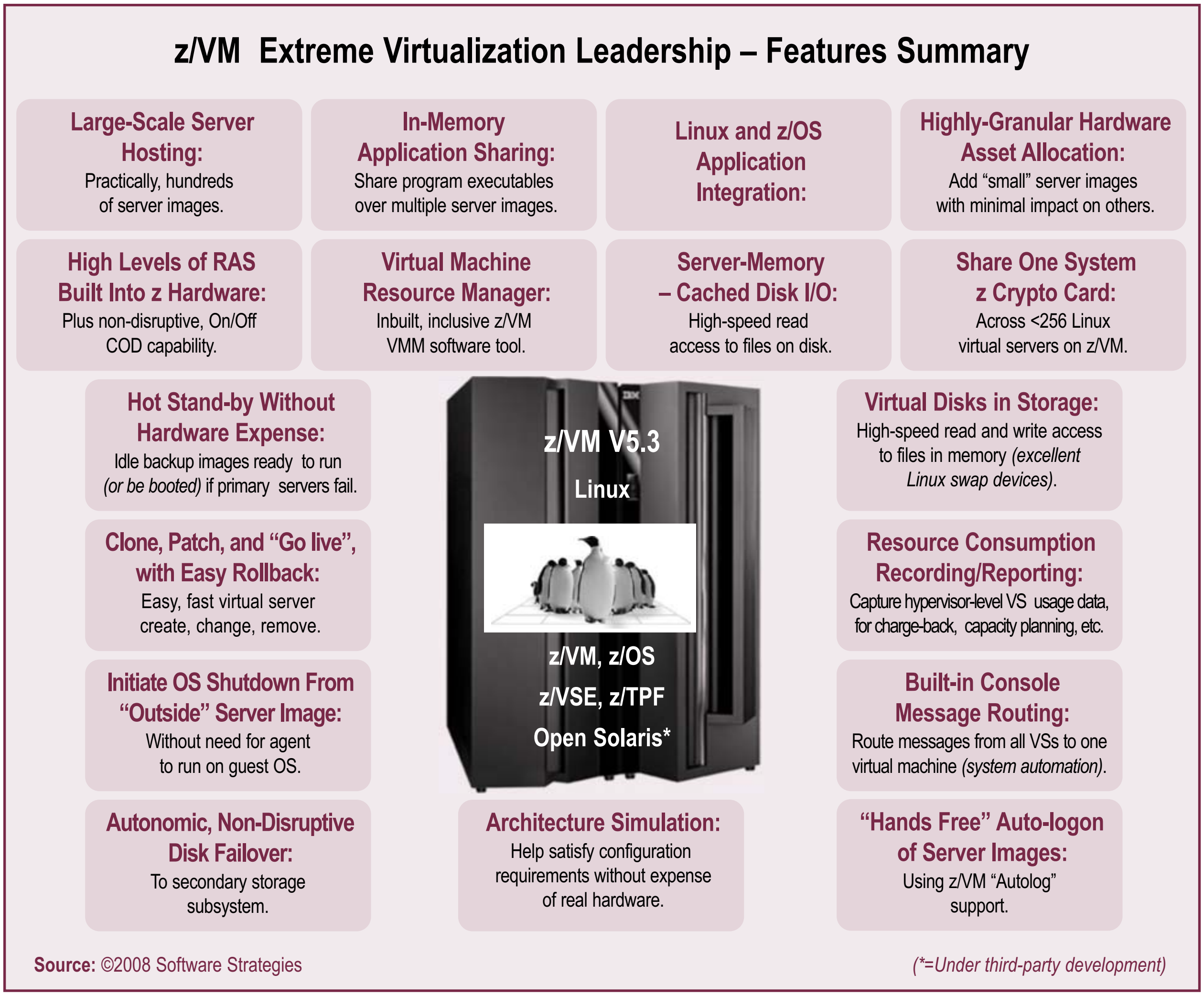


Figure 16: z/VM Extreme Virtualization Leadership – Other Features Summary

Space constraints preclude deeper/fuller discussion/analysis on each of these here, but we consider that our headings and notes speak for themselves.

Latest z/VM V5.3 Release Adds Major Extensions/Improvements

Another useful viewpoint on any major software platform is the breadth of advances added in the last main release. The vibrant, dynamic, advancing status of the z/VM hypervisor is amply illustrated by the numerous, important and substantial advances that were incorporated in the latest major production release, z/VM V5.3. The significant advances in this release, categorized by subject area, are concisely summarized in our Figure 17 chart on page 44:

z/VM V5.3 – Latest Release Advances

Improved Scalability and Constraint Relief:
- Support for more than 128GB real storage.
- <32 real processors in single z/VM image.
- Enhanced memory management for Linux guest.
- Enhanced memory utilization using VMRM between z/VM and Linux guests.
- HyperPAV support for IBM System Storage DS8000.
- Enhanced FlashCopy support.

Network Virtualization:
- Improved virtual network management.
- Enhanced failover support for IPv4 and IPv6 devices.
- Virtual IP Address *(VIPA)* support for IPv6.

Virtualization Technology and Linux Enablement:
- Support for System z specialty processor engines.
- Enhanced VSWITCH and guest LAN usability.
- Modified MIDAW *(Indirect Data Address Words)* for guests.
- Guest ASCII console support.
- Enhanced SCSI support.

Systems Management:
- Enhanced management functions for Linux and other virtual images.
- New function level for DirMaint.
- Enhancements to the Performance Toolkit.
- Enhanced guest configuration.

Security:
- Delivery of LDAP server and client
- Enhanced security via longer passwords.
- Conformance with industry standards.
- SSL server enhancements.
- Tape data protection with encryption support.

Installation, Service, Packaging Changes:
- Additional DVD installation options.
- Enhanced status information.
- RSCS repackaged as optional feature.
- New RACF Security Server for z/VM.
- U.S. Daylight Saving Time effect on z/VM.
- z/Architecture CMS and BOOTP servers.

Figure 17: z/VM V5.3 – Latest Release Advances

This was a rich set of advances, indicating a strategic platform undergoing heavy further development, including substantial efforts to maintain currency, compliance, and new capability exploitation/support for other relevant System z platform advances. An emphasis in this release was also to further extend single-z/VM-image size, and to further strengthen the resource-balanced "scale-up" and "scale-out" support for hosting virtual server workloads. Real memory exploitation was increased up to 256GB real, and up to 8TB virtual, per z/VM instance. A z/VM instance's single-image CPU support was also increased up to 32 processors. OSA-Express2 Link Aggregation enabled enhanced network bandwidth and availability. These, and the many other release advances of Figure 17, comprised a robust latest major release. z/VM is already security-certified to a high EAL 3+ level, and IBM has committed to attaining EAL 4 designation for the hypervisor.

Strong z/VM Management Infrastructure, Software Tool Support

The z/VM environment offers a powerful command/control management infrastructure, the valuable VM management capabilities of IBM Systems Director *(discussed in detail in Section 4, around Figure 13 on page 35)*, and a powerful, rich set of complementary IBM Software z/VM software tools for other functions needed. When running large-scale, virtual server hosting on System z, these z/VM management infrastructure facilities, and z/VM's support software tooling, proved a crucial source of z/VM advantage, enabling the high support staff productivity levels that customers have widely reported.

Typically, with its advanced infrastructure, its powerful automation options, and rich tools portfolio, each support FTE has proved easily able to support 100 virtual servers, and in some cases, up to 200 virtual servers, in a well-organized z/VM environment. These remarkable figures produce huge financial savings in support staff costs when hundreds of distributed servers are consolidated onto Linux-on-z/VM.

With physical distributed servers, each support FTE can typically only support 20-25 physical servers at best, and often as few as 15. So, for example, when 500 distributed servers are moved onto System z at these support productivity averages, 5 support staff *(worst case)* or even 2.5 support staff *(best case)* could support the z/VM environment. This is far lower than the 25 support staff FTEs *(at 20-servers/FTE average)* needed for the previous distributed environment, an five-fold minimum saving in headcount. At typical USA or EMEA $70,000/support FTE base costs, consolidation savings from such staff costs alone would be at least $1.4M per year, or $7.0M over a half-decade. *(20 support staff saved @ $70K p.a. = $1.40M p.a.)*

So what are the productive z/VM infrastructure capabilities that permit such high level of VS support productivity? How did these capabilities evolve? How are they able to scale-up to support many hundreds or thousands of hosted virtual servers with this considerable automation, and with such low support staff levels needed?

The strengths were originally developed for/around the mainframe VM/CMS timesharing OS family *(many names used)* predecessors of today's z/VM hypervisor. VM/CMS family OS were long used to host, deliver, monitor, manage, account for, control and bill for, the running of significant numbers of interactive VMs *(from scores to a hundred or so)*.

These interactive VMs each provided a registered mainframe end user with a private, interactive, online timesharing VM that could securely and interactively process their chosen online workloads. End-user 4GL reporting tools, financial modeling products, and simple end-user database/analysis products *(well-known examples included FOCUS, NOMAD, RAMIS, EPS, SAS, EXPRESS, and AS)* and mainframe application development were major application examples. These and other popular applications were all run in this way under VM/CMS family OSs, on literally thousands of mainframes, supporting low millions of end users, throughout the 1980s and 1990s.

VM/CMS family OS was also the first open source OS, where the early source code freely distributed, and users freely extended the code base. The VM/CMS OS family attracted extraordinary enthusiasm and deep devotion amongst its thousands of user sites, for its amazing capabilities, excellent performance, high efficiency of mainframe hardware usage, and low support costs. No IBM OS family, before or since, attracted such fanatical devotion and enthusiasm as that VM/CMS garnered in its earlier heyday. Linux, in many ways, is today's natural and rightful inheritor of those open source, OS-enthusiasm-sharing aspects of VM/CMS, whilst z/VM directly inherits its core technology strengths in augmented, modern form.

This is the long-proven heritage, and industrial-strength technology base, from which z/VM carried forward the VM/CMS descendant banner throughout this 2000-decade, and into the apparently different role of large-scale virtual server hosting. This newer role is, however, technically very similar to the original outlined above. Today's highly-productive z/VM infrastructure capabilities were evolved, refined, and automated, from those of their forerunners that had so effectively hosted large "interactive end-user" populations. From these origins, this lineage has successfully morphed into today's z/VM hypervisor, ideally equipped for massive Linux-on-z/VM virtual server hosting. The broad operating architecture of today's z/VM environment is depicted in Figure 18.

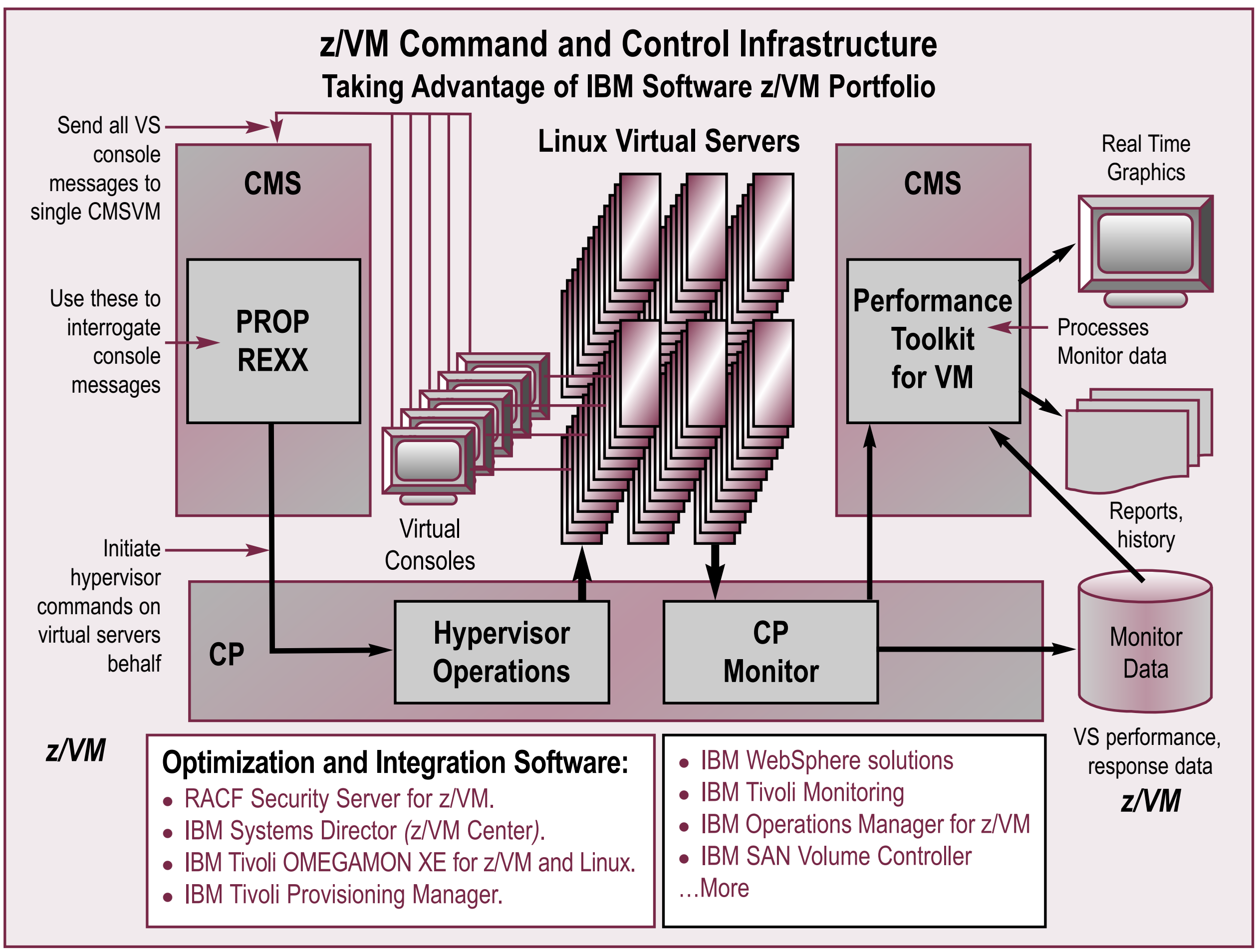


Figure 18: z/VM Command and Control Infrastructure – Taking Advantage of IBM Software z/VM Portfolio

This shows *(center top)* c. one hundred Linux virtual servers running under z/VM *(whole chart)*, each with its own virtual console, under the control of the z/VM core hypervisor shown within the large, lower, CP-labeled box. The CP component of z/VM stands for **"Control Program"**. Also within that box, the **"CP Monitor"** component of z/VM collects performance, resource usage, and response time, data from every running guest Linux virtual server, and writes this data out to the monitor database, shown bottom right. This database is read/analyzed in real time, and/or in batch, by the **Performance Toolkit for VM**, shown in the top right box, *(also labeled "CMS")*, which produces dynamic graphical displays of performance and resource usage, and analysis reports of current or historical usage and performance. This log data also provides the basic data for recharging users by their usage of their virtual servers.

This database is read/analyzed in real time, and/or in batch, by the Performance Toolkit for VM... which produces dynamic graphical displays of performance and resource usage, and analysis reports of current or historical usage and performance.

Top left, we see the many console message streams, one from each Linux virtual server, being compressed into a single CMS VM. This machine runs in the top left-hand box labeled "CMS" *(**originally Conversational Monitoring System**)* under the z/VM CMS component. There, this machine's combined message stream can be automatically operated upon, and efficiently processed by CMS facilities such as VM **PROP** *(**PRogrammable OPerator**)* that allows programmatic message processing, or by **REXX** *(the highly-capable CMS scripting language)*, VM Pipelines, etc. Where needed, these processes can then close the loop, by issuing control and processing requests to the hypervisor to execute on behalf of the Linux virtual server, or servers, affected.

Not shown on the chart, the **"z/VM Center"** task of **IBM Systems Director** also provides fast and easy provisioning of virtual Linux server images on z/VM.

Also not shown on the chart, the **VMRM** has been an important z/VM component since z/VM, V4.3. VMRM dynamically manages the workload performance of one group of *(one or many)* VMs, prioritized against workloads on other VM/VS groups, when there is contention for system resources. VMRM does this with two different interventions, a **resource manager**, and an **I/O priority queuing** capability. VMRM takes-in user-specified workload definitions and goals, compares these with actual achieved live performance from monitor data, and makes resources, and/or I/O priority, adjustments accordingly. The VMRM resource manager also considers workload priority *(Importance Value)*, and whether a workload was selected on the previous sample interval, when choosing workloads for adjustment. Selected workloads are adjusted up or down every other sample interval *(or less)* until they achieve performance within 5% of goal. For obvious reasons, VMRM only "kicks in" in a resource-constrained environment. If the z/VM system is not currently resource-constrained, VMRM cannot further enhance access to resources where these are readily available anyway. VMRM is a foundation component for z/VM's delivery of its excellent, consistent levels of virtual server performance, and of its efficient utilization of CPU and I/O resources to attain user-set service goals.

VMRM is a foundation component for z/VM's delivery of its excellent, consistent levels of virtual server performance, and of its efficient utilization of CPU and I/O resources to attain user-set service goals.

Beyond these invaluable system automation and resource usage monitoring/reporting tools, both described above, z/VM also offers a powerful suite of debug and problem-determination tools/utilities built-in. These include facilities to trace and trap at the instruction level, with no modification of the guest OS needed, and to sniff virtual network traffic among guest systems.

Additional, priced z/VM V5.3 features also now include:

- **DirMaint:** Simplifies task of adding/modifying/deleting users.
- **Performance Toolkit for VM:** Performance recording and reporting *(see Figure 18)*.
- **RACF Security Server for z/VM:** Security services *(including LDAP) (see Figure 19)*.
- **RSCS:** Provides NJE connectivity support for Linux systems.

Beyond these foundation z/VM capabilities, a range of IBM Software management, administration and support products for z/VM are also offered. These are listed in the lower two boxes in Figure 20 *(on page 48)*. With functional descriptions of the roles/services each delivers to the z/VM environment, these are also shown in our more detailed Figure 19 on page 47.

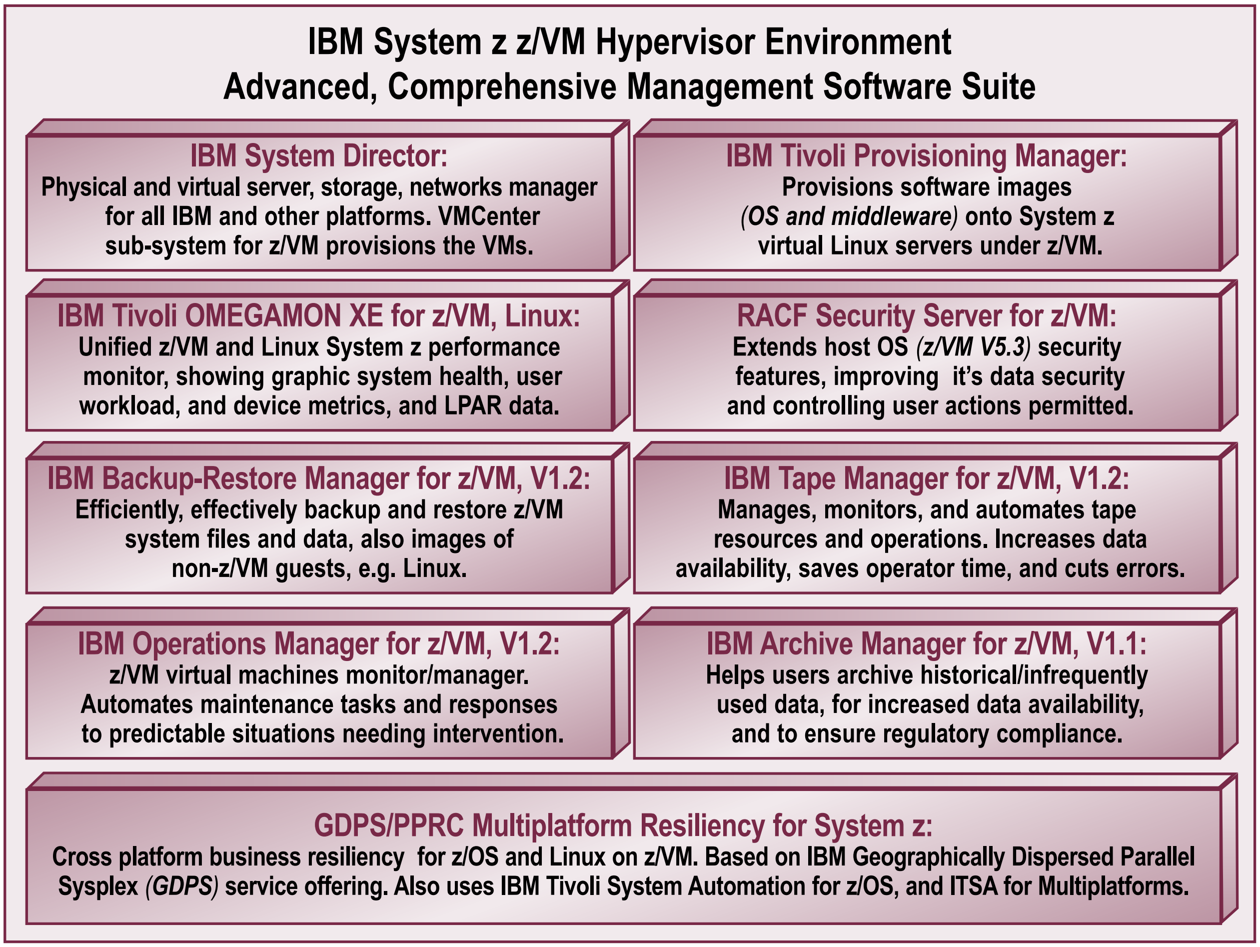


Figure 19: IBM System z z/VM Hypervisor Environment – Advanced, Comprehensive Management Software Suite

We reviewed IBM Systems Director earlier in this Paper, and consider that the Figure 19 descriptions adequately provide brief explanations of the main z/VM support facilities offered. The Performance Toolkit for VM shown above works with OMEGAMON® XE for z/VM and Linux, for enterprise-level performance monitoring. Numerous additional third-party z/VM tool and utility products are also offered by leading USVs, including CA, BMC, and many others.

It is the strength, productivity, power, and automation of all these z/VM capabilities and supporting software, forged by their record of accomplishments in service, that delivers such extraordinary hypervisor productivity for z/VM users.

IBM z/VM V5.3 on System z vs. VMware ESX Server V3.1 on x86/x64

Section 2 overviewed *(and Appendix C & E detail)* the costly, wasteful, distributed server sprawl resulting from a decade of rampant server proliferation, of both RISC-UNIX and x86/x64 servers, the later dominant in recent years. We also reviewed VMware's x86/x64 virtualization family, based around its ESX Server 3 hypervisor. This Paper focuses on the market's major strategic alternative – IBM's System z mainframe running Linux-on-z/VM – a unique extreme virtualization hypervisor that enables mass-consolidation of distributed server workloads onto this more cost-effective, secure, reliable, and manageable enterprise platform. These two **"distributed server consolidation solutions"** come from two such entirely different platform worlds, from two completely different histories, and are championed by two such entirely different platform supporter groups *(System z mainframe champions & x86/x64 enthusiasts respectively)*. For these reasons, and because each camp uses different sets of terms, vocabulary, descriptors and concepts, these two main strategic solutions to the "distributed computing nightmare" are rarely compared back-to-back. This is not surprising, because those knowledgeable enough to cover one candidate usually have little experience of the other, and vice versa.

Software Strategies now smashes this "world's-apart" virtualization knowledge barrier by publishing a comprehensive comparison summarized briefly in Figure 20 *(on page 48)*, backed by an extensive, detailed, comparative analysis, with point-by-point discussion, in Appendix B of this Paper.

Functional Comparison of z/VM 5.3 and VMware ESX Server 3			
Attribute	**z/VM V5.3**	**VMware ESX 3**	**System z Value**
Supported operating systems	Linux, z/OS, z/VSE, z/TPF, z/VM itself	Linux, Windows, NetWare, Solaris 10	z/VM-on-z/VM = adds flexibility
Scalability and Performance			
Hypervisor scalability	Up to 32 CPUs, 256GB of memory, 8TB of "active virtual memory"	Up to 32 CPUs, 64GB of memory	Cost-saving, extremely scalable VS environment
Virtual Machine *(VM)* scalability	Up to 64 CPUs, 1TB of memory, extensive I/O bandwidth	Up to 4 CPUs, 16GB of memory, modest I/O bandwidth	Virtualizes servers on z/VM that cannot run on VMware
CPU sharing	No limit	Up to 8 VMs per CPU	Add servers without HW adds
Architected *(practical)* VM limit	Thousands *(hundreds)* per z/VM copy	128 *(singles)* per copy of VMware	Avoid real server sprawl
CPU capacity on demand	Yes, non-disruptively	No	Fast, easy capacity growth
In-memory support	Minidisk cache; Virtual Disks in Storage; DCSS *(shared program EXEs.)*	Shared virtual memory pages *(find via background operation)*	Enhanced resource utilization
Logical Partition *(LPAR)* support	Yes	No	Secure Linux access to z/OS
Flexible Operations			
Resource over-commitment support *(memory, CPU, network, I/O)*	Extensive	Modest	Absorb workload spikes; add more servers to a "full" system
Reconfiguration of Virtual Machines	Non-disruptive re-config for CPU, I/O, networking; VM reboot for memory	VM reboot required for re-config of CPU, memory, Ethernet, disk	Higher server and application availability; staff productivity
Command and control, monitoring, automation infrastructure	Extensive, robust, time-tested.	Modest	Cost-optimized systems management support
Virtual Machine mobility support	No; single-image scalability of z/VM does not require mobility for management	Yes; essential for workload management across multiple copies of VMware	Can dynamically add/remove resources to meet demand
Integrity and Security			
Fault isolation/hypervisor security	Hardware-assisted isolation*; CAPP/EAL 3+	No I/O virtualization separation; CAPP/EAL 2	Helps avoid security breaches; data security and integrity
Run multiple copies of hypervisor on single server	Yes; share CPU, I/O, and networking resources among z/VM systems	No	Workload isolation; lower-cost failover *(using same hardware)*

z/VM runs in System z LPARs, which have achieved EAL 5 certification; System z HiperSockets provide high-speed, secure connectivity among LPARs.

Figure 20: Functional Comparison of IBM z/VM V5.3 with VMware ESX Server 3

Our findings here were drawn both from our own knowledge of both product platforms, and from research by other leading virtualization experts consulted. Because our comparison analysis and discussions are so fully presented In Appendix B, we do not repeat them. As each product compared is the dominant player in their different virtualization segments, each is undoubtedly a top-class offering, well proven by thousands of customers, so the benchmark levels in any comparison start high. Suffice it to say that the Linux-on-z/VM System z solution to a major distributed server consolidation can only be compared to a VMware ESX Server on x86/x64 solution in the same manner as:

- Comparing **a 747 Jumbo jet, to a fleet of 100 4-passenger Cessna twins**, for a planned mission to carry 400 passengers 6,000 miles over water, in the most economical, rapid, and safe manner.
- Comparing the **latest Cunard passenger liner for 2,000 passengers, to 400, 5-berth, 32' power cruisers**, for a sea voyage to move 2,000 people from Southampton to New York, 3,000 miles, in comfort and safety.

In each example, these are utterly different classes and scales of solution. Our mainframe z/VM equivalents above *(the 747 airplane and the Cunard liner in these analogies)* can each fully deliver the whole mission requirements above for large passenger numbers. They can deliver over the “long distances” required, can do so most economically, most comfortably, and most safely, for their passengers/“guests” *(all operative criteria here)*.

Our VMware ESX Server-equivalent alternatives *(the 4-passenger Cessna twin light plane, and the 5-berth, 32’ power cruiser)* are each desirable, excellent, and useful products in their own domains *(for the far more modest, small-capacity, short-range leisure roles they are capable of/designed for)*. Unfortunately, neither is truly capable of fulfilling the large-scale missions we outlined, through lack of range, the large crew numbers needed, and the serious safety risks, etc., and so are not remotely comparable with the two large transport platforms above.

So how come our IT world has so promoted the absurd, analogous idea that hordes of small, distributed servers can deliver remotely comparable results to those of IT’s equivalent of the 747 Jumbo jet or the Cunard oceanic passenger liner, the System z mainframe? No contest, as in both examples above.

OpenSolaris First Demonstrated on System z under z/VM

This Paper focuses on zLinux because it is today a now well-proven, established, widely used and trusted System z OS. However, what if similar mainframe-mass consolidation solutions could be directly offered to Solaris UNIX server users without needing migration to Linux? Sun Microsystems’ Solaris UNIX OS family remains widely used and well liked in its substantial installed customer base. This is in part because of Sun’s early leadership in UNIX from the mid-1990s, and in part for Solaris’ own mature OS strengths. Solaris versions traditionally ran mostly on Sun’s own costly RISC-UNIX servers *(plus Fujitsu/FSC)* and to a far lesser degree *(and much later)* on x86/x64 platforms.

...first-ever demonstration of the OpenSolaris OS code-base running on an IBM System z mainframe under z/VM.

End-November 2007, at the Gartner Data Center Conference *(in Las Vegas)* witnessed a milestone, first-ever demonstration of the OpenSolaris OS code-base running on an IBM System z mainframe under z/VM. The demonstration was showcased by IBM, Sun, and Sine Nomine Associates, who had collaborated on this effort.

Running under IBM’s z/VM hypervisor *(which already provides good foundations for mass hosting of Linux virtual servers on the mainframe)*, OpenSolaris on z/VM could enable large numbers of OpenSolaris virtual servers to run on one IBM mainframe. This important milestone signals further progress on joint Solaris compatibility between IBM and Sun, since both firms announced *(August 2007)* that IBM was expanding its support for the OS. The porting work is being undertaken by Sine Nomine Associates, an experienced OS research and engineering firm based in Ashburn, Virginia.

No timetable for a “production”, customer-deployable release of OpenSolaris on System z was disclosed. Such a new System z OS option would further extend the System z mainframe’s current mass-distributed servers workloads consolidation options. Beyond today’s onto Linux-on-z/VM consolidations path, an equivalent “onto OpenSolaris-on-z/VM” capability could be added. Sun’s large Solaris installed base, of now often ageing, distributed Sun SPARC RISC-powered servers, remains a tempting target for such mass-consolidation to System z when this port is delivered. Watch this space with interest!

Industrial Strength z/VM vs. Immature x86/x64 Virtualized Management

We can conclude that the unique z/VM platform on System z capabilities, the rich and extensive management and control software portfolio that supports it, and the high support staff productivity these deliver are unequalled. It is clearly apparent today’s z/VM platform is a fully-mature and evolved, enterprise-class, industrial-strength mass-virtual servers operating and management environment proven over many, many years of heavyweight usage.

By contrast, x84/x64 virtualization remains new, and young, borne in 2001 and only becoming widely deployed since 2005/2006, and with much of its associated VMM software still younger, relatively limited, and unsophisticated.

Just to illustrate the wide gulf of understanding between these solutions and their advocates. Some x86/x64 VMM tools and their developers use an elementary server ping test to detect whether a virtual server is running/available. To enterprise IT professionals, it is preposterous that anyone could seriously think that this trivial, near-meaningless, and effectively-useless, test of server/workload availability could serve or be of any value. Enterprise IT professionals know ping tests are a laughable measure of real availability, saying almost nothing useful about the effective QoS of a VM IT service in enterprise computing. x86/x64 virtualization advocates monitoring or using ping tests as a VS availability measure are displaying zero credibility, and cannot be considered serious, in enterprise computing service level management.

Our Analysis

Distributed computing proved a near-unmitigated disaster *(on most metrics)* for enterprise IT users. The sooner these sprawling, staff- and software-cost-intensive, and environmentally profligate legacy infrastructures can be eliminated, the better for all their users. The System z mainframe, with Linux-on-z/VM, offers a uniquely effective mass-consolidation solution, delivering several 100 to one server consolidation ratios *(distributed servers: 1 System z)*, enabled by the powerful capabilities of IBM's z/VM extreme-virtualization hypervisor.

z/VM spreads mainframe business values/QoS *(exceptional levels of availability, security, performance, and ease of operational management)* far more widely over the enterprise, including to the hundreds/thousand consolidated Linux virtual servers that can run on one mainframe, as discussed above. *(With the mainframe used here as a large, Linux-only enterprise virtual-server-hosting platform.)* A single mainframe can also improve IT productivity by also concurrently hosting non-Linux workloads, such as z/OS, z/VSE, and z/TFP, in test or production, and/or at current live or new release levels, all on the same system, with z/VM.

System z's IFL specialty processor for Linux, and all-IFL-Linux-only mainframe models, provide competitively priced mainframe hardware support for all Linux workloads types and sizes. Adding one IFL onto an existing mainframe has a modest cost, but a powerful capacity *(up to 580 MIPS on the z9 EC)* entry point. A large, all-IFL, Linux-only, new mainframe would typically be used for substantial, large-scale migration of several 100 distributed servers to Linux-on-z/VM. *(See Section 6 for four real examples.)*

Linux itself added ever-stronger capabilities, increasing scale, first-class reliability, and good performance, as it has matured over this decade. Combining these core Linux strengths with the unique QoS of the System z mainframe, which Linux on z is increasingly optimized to exploit, has made the duo a powerful combination.

A wide and still-growing range of top software, including middleware, tools, and applications, across most segments and categories, and from most major ISVs, are now available for Linux on System z. The resurgence of the mainframe market, widespread System z Linux adoption by users, and the ease of porting Linux software versions to z Linux, have driven this welcome trend. It also means a now-very-wide-range of distributed, server-resident software can now fairly easily migrate onto System z Linux versions or equivalents, enabling wider-scale consolidation.

z/VM is an IBM flagship mainframe operating environment, and one with a long and truly distinguished history. We summarized select major z/VM features in Figure 15, additional power features in Figure 16, and the latest V5.3 release feature additions/enhancements in Figure 17. This extensive roster, plus our favorable, detailed comparisons with VMware ESX Server *(mentioned below, summarized in Figure 20 and assessed in depth in Appendix B)*, amply indicate the robust strength, depth and capabilities of this powerful hypervisor platform.

z/VM also provides a powerful command/control management infrastructure, *(including IBM Systems Director)* with many inbuilt features *(most visualized in Figure 18)*, several useful, chargeable z/VM options, and a range of external, complementary IBM z/VM software products *(all detailed in Figure 19)*. Built/refined over z/VM's long, distinguished, timesharing history, these integrated tools enable the extremely high support productivity *(100-200 virtual servers/FTE)* that this hypervisor delivers. Virtual server performance/resource use data can be stored, and fully analyzed *(**with Performance Toolkit for VM**)* for SLA compliance, chargeback, or tuning. Powerful scripting/programmable operator facilities enable message-stream handling and virtual server control loops. Debugging and problem determination tools/utilities are also provided. Means to add/delete/change users *(**DirMaint**)*, fast/easy provisioning of z/VM virtual Linux server images from within IBM System Director *(z/VM Center task)*, and the **z/VM VMRM** dynamic workloads performance optimizer, are other major examples.

...z/VM as the leading, large-scale, industrial-strength, heavy-duty, "major-missions" hypervisor.

We presented a back-to-back, functional comparison between z/VM V 5.3 on System z, and VMware ESX Server V3.1 *(x86/x64 only)* in summary form as Figure 20. We also discuss/analyze their differences in considerable depth, and at substantial length *(5 pages)* in Appendix B. Our conclusions can broadly be summarized here by describing z/VM as the leading, large-scale, industrial-strength, heavy-duty, "major-missions" hypervisor. VMware ESX Server V3.1 deserves real credit as the first product to break down the severe barriers to virtualizing x86/x64 servers at all *(2001)*. Its VMware Infrastructure 3 suite today, based around ESX Server, is now a quite broad, capable, but generally light-duty, small-to-medium-scale, and somewhat costly, x86/x64-only virtualization platform. It has also suddenly *(2006-07)* come under heavy fire from a veritable barrage of new "x86/x64 virtualizer" competitors, likely to drive down/commoditize this hypervisor segment soon.

Virtualization on distributed server platforms raises serious new security vulnerabilities that are not yet fully charted or protected against. Gartner has warned users not to assume that existing physical server security techniques will safely protect these new virtual environments.

By contrast, for decades now, z/VM and the mainframe were architected together to deliver highly-secure processing with excellent security and integrity levels. System z servers, and their LPAR partitioning, have achieved the highest EAL 5 certification. z/VM has already achieved EAL 3+ certification, and IBM has stated that it is pursuing EAL 4 certification of z/VM V5.3, offering system solutions that are methodically designed, tested, and reviewed for secure operations.

Finally, work is well under way to port the Sun-donated, OpenSolaris OS to System z under z/VM, with code already demonstrated running live at a public Gartner Group event, in November 2007. When completed/available, this exciting new OS option could allow the large base of Solaris OS, ageing Sun SPARC RISC server, users to directly consolidate these workloads onto this compatible OS on System z, an exciting extension beyond today's strong Linux offerings.

6. Distributed Server Consolidation to System z Linux-on-z/VM – Real Examples

Mass-distributed-servers-to-System-z-consolidation Success Stories Show the Way

IBM has now won numerous consolidations of distributed server *(UNIX & x86)* populations onto System z IFL-specialty-engine-equipped mainframes, using the z/VM advanced virtualization hypervisor, running Linux OS. Here we assess four such significant, distributed-server-to-System z, mass-consolidation named-customer projects The profiles show the breadth/depth of savings/benefits IBM/customers have already achieved with these projects, as well as highlighting important learning from these experiences. The four named customer profiles are:

- **# 1. IBM Project Big Green – Consolidating 3,900 Distributed Servers to 33 System z Mainframes.**
- **# 2. Nationwide – Consolidated c. 500 Distributed Servers to 2 System z Mainframes to Date.**
- **# 3. Government of Quebec – Consolidated 200 Distributed Server Workloads onto one System z9 EC so far.**
- **# 4. Nexxar Group – Sweeps 80 x86 Servers off its Floor onto System z9 BC Mainframe under z/VM.**

The first of these is a massive *(world's largest ever)* distributed to System z mainframe consolidation now well underway inside IBM, the second is a large-scale, major enterprise migration, the third a mid-sized government entity migration, and the last is a fast-growing, smaller, young business consolidating early onto a new mainframe.

Individually and collectively, these substantial, real-world examples Illuminate, and bring to life, the deeply-compelling case for mass-consolidation of appropriate distributed server workloads to System z mainframes using their extreme virtualization capabilities on z/VM. IBM has now scaled up its sales/marketing effort to next win hundreds more, and then thousands more, such migrations over the next several years.

1: IBM 3,900 Distributed Servers Mass-consolidation onto 33 System z Mainframes Biggest Yet

August 2007 saw the unveiling of the largest, most dramatic, project yet to mass-consolidate distributed servers onto System z mainframes, using Linux-z/VM extreme virtualization. In this huge, Project Big Green energy-saving transformation of its global IT estate, IBM is now hard at work consolidating an initial c. **3,900 distributed servers** *(UNIX and x86)* **onto c. 33 System z** mainframes running Linux-on-z/VM, its extreme virtualization hypervisor solution.

IBM stated this new mainframe environment will slash energy consumption by 80%, and bring five-year TCO financial savings in energy, software license and systems support staff costs of over 40%, compared with those of the current, distributed server platforms. Summarized highlights of the consolidation/transformation are shown in the panel below.

IBM Project Big Green – Distributed to System z Consolidation Highlights

- **3,900 distributed servers** *(UNIX and x86/x64)* being consolidated down onto c. **33 System z mainframes**.
- **3,900 servers** *(c 25% of IBM total)*, picked from 8,600 "migration-eligible" of **IBM's c. 16,000 total distributed servers**.
- **23,000 processor cores** on distributed systems to **1,782 System z IFL processors, a 13:1 CPU consolidation ratio**.
- **33 System z EC all-IFL mainframes**, running **Linux-on-z/VM** virtualization, an average **120:1 server consolidation**.
- Application workloads migrating: **WebSphere** Application, Process and Portal Server, **SAP** applications, **DB2®** database applications, and **Lotus Domino** collaboration.
- Distributed servers ran serious application workloads, averaging **5.9 CPU cores/distributed server** capacity each.
- **92% less hardware** needed for System z solution.
- **80% energy savings** over present distributed server set-up anticipated, enough to power a small town.
- **Over 85% data center floor-space reduction** for mainframe solution anticipated.
- **40%+ five-year TCO savings** from energy, software, and system support costs over current systems TCO projected.
- **Commercial TCO model used** to estimate financial savings.

This sweeping "distributed elimination" is enabled by System z's sophisticated virtualization capability, notably the z/VM V5.3 extreme virtualization hypervisor, running the efficient Linux OS. In this case, each System z9 EC *(all-IFL-equipped, large mainframes)* will support about 120 virtual servers *(on average)* under Linux-on-z/VM, each assuming a substantial workload from their distributed server predecessors. The migration will initially use only part of the potential capacity of each mainframe, leaving substantial headroom for future workload growth/expansion on each. It will also provide IBM with far greater flexibility and responsiveness to support future changes in processing needs for these applications, and/or to provide better DR/BC where needed.

This spectacular, mass-distributed servers-to-System z mainframe consolidation initiative is part of Project Big Green, the broad IBM commitment *(announced May 2007)* to sharply reduce data center energy consumption, both for IBM itself, and for its thousands of global IT user customers of all types and sizes. A holistic approach was taken in the consolidation project, including giving careful consideration to application portfolio reduction, asset optimization, and the use of IBM System p™ virtualization for some appropriate work, as well as the System z aspect mentioned previously. The 3,900 distributed servers displaced by this project will be recycled through IBM Global Asset Recovery Services in an environmentally sound manner. The newer servers will be refurbished for resale, older servers will be stripped for spare parts and/or recyclable materials, and any small residues will be properly disposed of.

IBM operates the world's largest, most sophisticated, data center operation, with 8,000,000+-sq. ft. of data center space, to support its 350,000 staff and thousands of IT services customers. Since 1997, the company had already consolidated its strategic worldwide data centers down from 155 to 7 *(a 22:1 data center consolidation ratio)*. The IBM data centers in Poughkeepsie *(New York)*, Southbury *(Connecticut)*, Boulder *(Colorado)*, Portsmouth *(UK)*, Osaka *(Japan)* and Sydney *(Australia)* will participate in this major, distributed-server-to-System z consolidation initiative.

Big Blue promises to report its main project-learning, savings-details, migration efforts, and the challenges and issues encountered/solved over the project life-span, both to guide its own consultants, and also to help guide other IBM customers undertaking similar, large-scale, distributed-servers-to-System z mainframe consolidation moves. IBM's current, extensive consolidation expertise is available through IBM Global Technology Services, and this project will add excellent, additional experience. Specifically, a GTS Implementation Services for Linux Server Consolidation provides clients with just such a Linux mainframe environment for mass-consolidation of distributed servers, to slash costs and raise reliability/performance.

2: Nationwide's Distributed, Linux-on-z/VM System z Consolidation Shows $15M, 3-Year Saving

Nationwide's c. 500 Distributed Servers to 2 System z's Consolidation – Highlights

- **$158B assets, $21B revenue, 36,000 employee** diversified insurance/financial services leader. Totally dependent on IT systems, with **6,000 IT staff** *(1/6 of total)*.
- **Long-time System** z user **for centralized** systems, had also accumulated **5,000+ distributed servers** *(x86 and UNIX.)*
- Own studies showed **distributed servers at 10% av. utilization**: 78% never topped 50% usage at highest peak load.
- Distributed **server provisioning much too slow** for fast-moving business needs – **took from weeks to months**.
- Nationwide's data centers were almost out of floor space, power, and cooling capacity, **faced $10Ms** investment.
- Now* have **477 virtual servers**, now run on **2 System z990** mainframes, **displaced near 500 distributed** servers
- Now* have significantly **better TCO – $15 million savings over 3 years** or more.
- **15 mission-critical applications** *(Nationwide three-tier)* now mainframe deployed, **100,000+ active users/day** usage.
- **80% reduction in data center floorspace** needed. Deferred $10M data center investment otherwise needed
- Now attain **70% average System z CPU utilization**.
- **50% reduction in Web hosting** monthly costs to Nationwide.
- **50% reduction in hardware and OS support staff** efforts/costs.
- **Big middleware cost savings** *(WebSphere, UDB, Oracle)*. Just 25 CPUs System z vs. >1000+ CPUs distributed.
- Significantly faster **provisioning speed** ***(months → days)***.
- **Simple, robust HA and DR for 15 mission-critical applications**, never feasible with distributed.

Columbus Ohio, USA-based Nationwide is one of the world's largest, diversified insurance/financial services firms, with $158B+ assets, $21B+ revenues, and ranked # 104 on the Fortune 500 list *(all 2006)*. Nationwide offers many insurance/financial services, including auto, motorcycle, boat, homeowners, life, and commercial insurance, administrative services, annuities, mortgages, mutual funds, pensions, and savings plans. With 36,000 employees, Nationwide's business depends totally upon its IT systems, where it employs 6,000 staff in IT-related functions.

Nationwide had long used centralized IBM mainframes for its core systems, but had later accumulated a large population of 5,000+ distributed servers *(UNIX and Windows)*. The firm operates a federated IT model, but keeps central architecture governance control, specifying Java/Linux, and Oracle or DB2 database, for all new, custom, application development, for example. Nationwide's own studies showed that its distributed servers averaged 10% utilization, with 78% never exceeding 50% utilization even at their highest workload peaks. Just to provision and bring up a new distributed server for a new application took several weeks of elapsed time, and because of fixed capacity, each had to be heavily over-specified for peak demands. Nationwide was spending valuable resources and time maintaining costly physical servers that took up expensive data center space, consumed much energy, and used only a small fraction *(10%)* of their available computing power. The firm wanted to create a cost-effective, scalable, and 100%-available platform for its WebSphere application infrastructure. Its existing, heterogeneous, distributed systems estate was difficult to handle, and prohibited any useful sharing of hardware and software resources, keeping costs too high.

In a recent analyst briefing *(see Footnote[2])*, Nationwide shared details of its successful mass consolidation, of almost 500 distributed servers onto two additional IBM System z mainframes, using IBM's z/VM extreme virtualization solution, and the SUSE Linux Enterprise Server *(SLES9)* from Novell. This solution has enabled Nationwide to hugely reduce the TCO of this part of its IT environment, with 3-year TCO savings projected at a minimum of $15M. Highlights are summarized in the panel bullets below.

Nationwide chose another path to overcome these severe distributed computing problems.

Nationwide's existing data centers were also rapidly running out of floorspace, power and cooling capacities. If its past, distributed server growth rates had continued, the firm faced investing $10Ms of new capital in data center expansion. Nationwide chose another path to overcome these severe distributed computing problems.

Two IBM System z990 all-IFL mainframe servers, now *(Footnote[2])* running over 477 virtual servers on IBM z/VM V5.2 mainframe virtualization software, each supported by Novell SLES9 Linux, are now in use, replacing nearly 500 distributed servers. One z990 runs all the production workloads, using 4 LPARs, 12 IFL processors, 176GB of memory, and 8 I/O cards. The other z990, located at a second site, runs all development, test, and a full Disaster Recovery facility, for all these mainframe-migrated applications, with data replicated across from the production site. The second site system used 5 LPAR partitions, 13 IFL processors, 224GB of memory, and 12 I/O cards.

...one SA/SE could now easily support 100 virtual servers, a more than 3-fold reduction in its server support staffing required...

Nationwide had previously found that one systems administrator/systems engineer was needed to support every 30 of its physical distributed servers. *(This was considerably better than the industry average of 20!)*. With the mainframe virtual server solution, the firm found one SA/SE could now easily support 100 virtual servers, a more than 3-fold reduction in its server support staffing required; usually one of the two largest sources of the high TCO of distributed systems. *(Software costs are usually the other largest.)*

The rebranded Nationwide.com site runs on this new mainframe platform. To ensure this site could cope with the expected Super Bowl 2006 XL advertising-blitz-induced load spike, an extra IFL CoD processor was rented from IBM for 2 weeks to cover that period, and the system tested ran 22-times the expected peak Superbowl load smoothly. All worked perfectly on the day, with the system handling a large spike easily using this extra CoD capacity.

The striking financial and operational benefits reported are summarized in the panel bullets above. Nationwide also praised its mainframe solution's ability to dynamically change allocation of computer power to specific applications in minutes, and/or to use the sophisticated System z CoD options to increase/reduce compute power for special peaks, as huge advantages. The firm can also now create and provision new Linux-based virtual servers in minutes, so responding far faster as business needs change. Because most enterprise middleware is licensed by the number of CPUs, middleware costs for 25 System z IFLs were also just a fraction of those for the 500 distributed servers *(>1,000 CPUs)* replaced.

Nationwide now plans to place all its future, new Java/Linux applications growth, and migrate all existing major application refreshes, onto its highly-secure and easily-scalable System z...

Nationwide now plans to place all its future, new Java/Linux applications growth, and migrate all existing major application refreshes, onto its highly-secure and easily-scalable System z mainframe virtual hosting platform. The firm took advantage of IBM GTS service *(Capacity Planning/Management)* server and software expertise for some of the best practices used in tuning, capacity management, service management and mainframe resource optimization.

Footnote[2]: Source/Date: Analyst Conference Call – Buzz Woeckener – Manager zLinux/Web Unix, Nationwide, September 6, 2007.

3. Government of Quebec, Canada

Government of Quebec DGTI Distributed to System z Consolidation Highlights

- DGTI – IT service provider to **125 Government of Quebec offices and agencies**.
- Experienced, long-term, **satisfied System z mainframe user, 7 IBM mainframes** before this project.
- Also ran **450 distributed servers** *(UNIX, x86 – HP, Sun, IBM)*.
- **Serious distributed server estate issues**, including:
  - Too many OS, too many versions, much unsupported software, too diverse.
  - Too support intensive, distributed skills hard to find in Quebec.
  - DR/BC not practical/affordable, many applications not covered, etc.
- New strategy: **migrate/consolidate distributed server workloads onto System z** with Linux–on-z/VM.
- **Oracle DB application consolidation** from distributed to a new System z, was 1st major phase
- Added **System z9 EC- 5 IFL processors, ~3,000 MIPS, 8 LPARs**, 40 internal networks. z/VM, SLES Linux, Oracle DB, Velocity and CA software.
- Now* run **200 Oracle DB instances, 150 Linux virtual machines**, on new z9 EC, planning to hit 275 end-year 2, 375 end-year 3.
- Dramatically faster provisioning:
  - New **z/VM Linux virtual server takes 30 minutes**. Distributed was 1 week to 3 months. **336-1,008 times faster**.
  - New **z/VM Oracle DB instance cloned/installed takes 30-45 minutes** on. Was 10-14 hours on Sun. **13-28 times faster**.
- **Highly-successful Oracle DB consolidation**. Now doing WAS and Lotus/Domino distributed-to-System z workloads migration.
- Deep, organization-wide training/briefing on *(System z, z/VM, Linux, virtualization, security, etc.)*, for all staff disciplines/levels, was critical success factor. **200 lecturer/days of education** delivered
- **30% minimum cost savings** to date, 2-year payback on z9 investment 50% faster than expected.
- 2 Linux administrators easily supporting 200 virtual servers. **100 virtual/servers per administrator**.
- Huge software license reductions gave biggest savings. **92% fewer on Oracle DB, 91% fewer on WAS**.
- All above accomplished in **just 4-months**.
- DGTI Oracle to z/VM consolidation project won **Share 2007 Award for Excellence** in technology.

The benefits of mass-distributed server consolidation onto ultra-efficient System z mainframes with Linux-on-z/VM are not limited to the private sector. Most large public-sector/governmental bodies worldwide had also acquired large, equally costly and wasteful, distributed server estates similar to those of their private counterparts. However, political pressures for efficient, low-cost government, and for enhanced, IT-augmented, public services delivery, are becoming louder everywhere. Voter expectations are also increasingly for "low carbon" public agencies, perhaps even more so than for private firms. An excellent example has been set by the DGTI *(La Direction Generale pour le' Traitement de Informatique)*, which provides IT infrastructure/services to 125 offices and agencies *(large and small)* of the Government of Quebec, Canada. The DGTI has long been a committed IBM mainframe user, deploying 5*z Series z890s, 1*z800, 1*System S/390 G5, and is now additionally using a System z9 EC, all-IFL machine, running Linux-on-z/VM *(the subject of this profile)*. In addition, DGTI had also operated/supported some 450+ physical distributed servers running 750+ logical server images. *(Servers from HP, Sun, IBM p Series/System p, mainly UNIX with some x86.)*

The DGTI had experienced considerable problems with its mixed, 450 distributed servers, estate. Numerous OSs were in use, in diverse versions, and ditto their middleware/other software. This diversity needed extensive support staff effort to maintain because of its software complexity. Some older software products still in use were no longer vendor-supported; a substantial vulnerability. Critically, affordable DR/BC coverage was not technically or economically feasible for most of the distributed estate, leaving critical Quebec Government agency/office applications on these platforms with no viable backup/recovery provision.

As its distributed server estate grew, these problems worsened. The DGTI was understaffed in its distributed systems support area, and found these systems increasingly difficult/costly to support. It was also very slow to provision/bring up new distributed, mid-range systems *(complex public procurement process)*, and these systems were completely inflexible.

By contrast, the DGTI's System z mainframe estate had long proved itself a stable, efficient, and mature environment. A good supply of affordable, young, mainframe-skilled staff at early stages in their careers was readily available in Quebec. Operating solid and well-controlled DR processes with its mainframes was a long-proven, straightforward, and affordable option as well.

The DGTI was a large Oracle database user. Many of its distributed servers ran Oracle DB applications. The DGTI began its mass consolidation of distributed server workloads onto its new System z EC mainframe *(acquired especially for this strategy)*, with these Oracle DB applications as the first major workload class to be migrated. Their new IBM z9 EC, all-IFL mainframe, was equipped with 5 IFL Linux specialty processors *(rated at ~3,000 MIPS capacity)*. It configured this system with 8 LPARs to isolate the main software stacks used *(these were Oracle DB (2), WAS (2), a Service zone, a Lab zone, Domino (2), & open source software)*, and set up 40 internal networks within the mainframe. The software used included: Novell SLES *(8, 9, and 10)*, IBM z/VM V5.2 & V5.3, Oracle DB *(9i & 10g)*, WAS, Velocity Software performance tools, and CA automation and scheduling tools.

The DGTI was *(Footnote[3])* running 200 Oracle DB instances on 150 Linux VMs on this z9 EC mainframe, with growth of 100 additional images/per year planned for the next several years. *(300 images in year 2, 400 images in year 3, etc.)*. 25 production server instances were supporting the Government of Quebec's impressive new Portal, for example. Use of a well-tested, "golden Linux+tools+Oracle image" enabled rapid cloning of new virtual server instances. For the first 60 migrations, a rate of one completed migration per working day, 20/25 per month, was achieved.

The DGTI project team overcame many challenges. These included improving technical skills/understanding on the System z z/VM-Linux platform across all teams, ensuring the rigorous isolation, reliability, and security of mainframe applications in a shared resource system, blending in the new virtual networks needed, and providing redundancy-DR/BC affordably where justified. Scores of different Quebec Government offices/agencies were touched by this project, and needed to be both fully-involved/informed and completely satisfied. These challenges were fully met and overcome by the team. The Quebec Government and the DGTI already view this project as a major success. It is continuing to migrate many more Oracle workloads onto its System z9 EC, and is now undertaking similar migrations of WAS and Lotus/Domino workloads off distributed platforms onto the System z mainframe.

The financial case for the migration also proved itself compelling. DGTI benefited from huge cost savings *(CA $1.2M)* in CPU-driven software license costs *(Oracle DB: 4 CPUs on z9 BC vs. 49 distributed – 92% less; WAS 8 CPUs on z9 EC vs. 88 on distributed – 91% less)*. It also benefited greatly from the much lower systems support staff levels needed on the System z solution. Hardware costs were rated as similar, and thus payback on the new z9 EC hardware investment took only 2 years, much faster than the 3 years projected originally. The DGTI reported an over 30% basic cost saving *(software, support staff, and, hardware maintenance)*. This was without even considering the large additional savings on data center floorspace, electric power consumption, cooling costs, and UPS capacity costs, where the mainframe solution additionally produced further large savings.

DGTI has now decided that it will no longer purchase any new distributed systems for Oracle DB or WAS workloads...

Acceptance has been excellent throughout the organization, Oracle database performance on I/O, and on DB loading, is much faster than previously, and affordable DR/BC has been applied for more critical applications. Because of this success, the DGTI has now decided that it will no longer purchase any new distributed systems for Oracle DB or WAS workloads, and will rapidly proceed to migrate all possible, remaining, distributed server workloads onto System z.

Nexxar Group Sweeps 80 x86 Servers off the Floor onto System z9 BC Mainframe under z/VM

Nexxar Group Sweeps 80 x86 Servers off the Floor Onto a System z9 BC Mainframe under z/VM – Highlights

- Needed far more **secure**, **manageable**, **reliable**, **available**, **scalable**, extensible, and flexible IT platform.
- **To support fast growth** via "own label" sales, plus multiple competitor acquisitions, within one, more efficient IT platform.
- To **replace 80 distributed + x86 servers** existing IT infrastructure that was mushrooming with fast growth business.
- Consolidated all **80+ x86 servers into 1 IBM System z9 BC** mainframe that runs **Linux-on-z/VM**, on an **IFL processor**.
- **30% reduction in total IT costs**, operational and administrative, compared to prior distributed model, was achieved.
- Maintenance and support **staffing headcount 75% lower** on Nexxar's mainframe environment.
- **Easier, faster IT integration** of corporate acquisitions on shared z9 BC infrastructure.
- **Far faster times-to-market**, days not weeks, to bring up new customer "white label" services, using rapid virtual server provisioning.
- Nexxar application **software maintenance/support far easier/faster** on one z9 BC system.

Footnote[3]: Source/Date: IBM analyst conference call – D. Kreutzer, VM Resources, DGTI System z Linux-On-z/VM Project Manager, 22nd August 2007.

Nexxar Group, Inc., is a fast-growing financial services firm offering money transfer, money order, bill payment and check-cashing services, across 105 countries through a network of over 40,000 agents and branches. The firm also delivers "white-label" services for other financial organizations to resell, processing these on its own systems. Founded in 2003, Nexxar also acts as an "early-stage consolidator" in its sector, by acquiring, and rapidly integrating, successful, smaller, local competing firms. Nexxar needed an improved IT infrastructure able to provide a high availability, 7*24 service that allowed far-more rapid new server provisioning for its new customers, which offered high customer-specific security, and enabled optimized use of shared resources, easy integration, and high service automation to reduce IT staffing needs. It also needed to be easily able to integrate the IT solutions of its new acquisitions rapidly onto its own IT platform.

The firm's core applications were developed with IBM Rational development tools, used to build its J2EE™-based money transfer/other FS applications suite. These ran on IBM WebSphere Application Server under Linux and were deployed on more than 80 Intel x86 microprocessor-based distributed servers. This infrastructure meant the firm had to add additional dedicated servers, software, networking, cables, and switches, etc., for every new "private-label" customer it won, and/or often to also support each new acquisition that it made.

The firm consolidated all 80 of these x86 distributed server's workloads onto one IBM System z9 BC mainframe running Linux-on-z/VM. It now exploits the IBM mainframe's virtualization capabilities to quickly create secure, custom-tailored computing environments *(virtual servers)* for each new "private label" relationship customer that it wins. The System z9 BC also supports the firm's growth-through-acquisition strategy, by enabling each acquisition's IT environment to be integrated quickly and efficiently onto its shared, unified, mainframe platform.

New virtual servers can now be created on the mainframe in minutes on z/VM by copying/cloning, almost like cutting and pasting. The System z mainframe and z/VM were designed from the ground up to provide an ultra-efficient, shared-resource IT environment, delivering the highest levels of security that financial services firms are mandated *(SEC)* to use.

The IBM System z mainframe with z/VM also streamlines maintenance substantially by reducing by 75% the required head-count to maintain the new operating environment in comparison with the x86 systems it replaced. The System z CPU utilization achieved *(which can near 100%)* is also far superior to that seen on its earlier x86 systems, so Nexxar can now make much more efficient use of its bought server resource. Nexxar also implemented IBM Service Management methodologies to enable easier management of the new mainframe environment.

Test, development and quality assurance systems are now run as virtual servers that can now be delivered far more rapidly and easily to the business. The core customer/transactional databases were seamlessly migrated from distributed DB2 to IBM DB2 for z/OS, and now run in a separate LPAR partition on the mainframe. The built-in cryptographic capabilities of the z9 BC platform, and its proven security features, have enabled Nexxar Group to build completely isolated systems running side-by-side on the same physical hardware. The solution also simplified DR planning at Nexxar Group. The company can restore its entire environment from a single backup on a mainframe in an IBM data center.

In addition to its IBM System z BC mainframe with z/VM & Linux, Nexxar also installed an IBM System Storage DS8000 *(high-end enterprise disk array)* and 3590, and the flagship z/OS OS with IBM's DB2 database. Tivoli OMEGAMON mainframe performance monitoring, IBM Tivoli Composite Application Manager for SOA *(which monitors, manages and controls the Web services layer)*, and IBM Rational mainframe development tools were also added. It also uses the IBM RACF® security software's capabilities. IBM provided services help with this implementation, in the form of its IBM GTS Infrastructure and System Management services.

Nexxar Group thus escaped completely from its earlier increasing sprawl of distributed servers, with attendant swarm of IT workers, and previously skyrocketing software costs.

Nexxar Group was backed by large investor FTVentures – a private equity house focused on software/business services companies and a strategic partner with IBM's Venture Capital Group *(one of more than 120 important VC firms IBM VCG collaborates with)*. A POC, run by IBM at its laboratory in Montpellier, France, first convinced Nexxar Group of the benefit of migrating completely to the new IBM System z9 BC mainframe, advice subsequently validated by the results obtained above.

System z Consolidation – Staggering Financial Savings – An Example

Massive financial savings are delivered by larger distributed scale out servers to System z Linux-on-z/VM consolidations. They also offer major financial savings compared with new x64 server with VMware virtualization replacement alternatives. To illustrate the level of these major savings, we include here a summary financial analysis of an actual customer case recently developed by IBM. The customer, an international process industry company, was not a System z user, but had accumulated a distributed server and workstation population of 7,965 systems globally to run all their applications.

Excluding the 1,792 technical workstations left 5,993 servers, of which 4,425 were located in the major countries of operation, and 1,568 were based in smaller geographies not within the consolidation project's scope. Of the 4,425 major country-located servers that were within scope to be considered for consolidation, 2,339 *(52.6%)* were found suitable for consolidation on Linux on System z, and 2,086 its found could best be consolidated/updated otherwise. *(521 onto new large System p, 455 onto BladeCenters, and 1,110 onto new large x64 with VMware.)* Of the 2,339 servers found suitable for System z consolidation, 689 were UNIX *(IBM, HP, and Sun mix)*, 519 were Linux, 1,118 were Windows, and 13 were VMS.

Figure 21 shows the summarized, comprehensive, five-year total costs of ownership for the three alternative solutions of:

A. Retaining the existing 2,339 distributed UNIX, Linux and Windows server estate "as is".

B. The proposed System z Linux-on-z/VM solution of 5 large, new all-IFL System z mainframes, with 255 IFLs, running 2,339 z/VM virtual server images, based at 2 mega-centers.

C. An alternative, all-new x64 VMware replacement solution of 234 HP DL585 x64 servers, with VMware ESX Server, each running 10 virtual server images.

Figure 21 incorporates both the full capital investment costs *(for all new items)* and all ongoing operating costs, for the existing *(A)* distributed servers and the new proposed *(B-System z, or C-New x64)* configuration alternatives. These each included the full costs of server, connectivity, disk, and enterprise networking hardware and connections, for software licenses, for maintenance on all these, for system support staff, and for energy costs *(electrical power/cooling)* and floorspace over a five-year period.

Alternatives Considered → TCIO Cost Categories↓	A			B			C	
	Existing Distributed Servers	% of A 5 Year TCO	5 Year Saving B over A	New System z z/VM	% of B 5 Year TCO	5 Year Saving B over C	New x64 VMware	% of B 5 Year TCO
Number of servers	2,339			5			234	
System support staff costs	$119.088	54.1%	-$93.176	$25.913	24.5%	-$42.554	$68.466	49.7%
Total software $M	$51.460	23.4%	-$23.053	$28.407	26.8%	-$6.442	$34.849	25.3%
Total server HW $ M	$11.841	5.4%	$28.250	$40.091	37.9%	$31.590	$8.501	6.2%
Enterprise network costs $ M	$11.976	5.4%	-$6.587	$5.389	5.1%	-$6.177	$11.566	8.4%
Power, cooling, floor space $ M	$12.595	5.7%	-$11.651	$0.944	0.9%	-$0.946	$1.890	1.4%
Connectivity total cost $ M	$10.791	4.9%	-$10.452	$0.339	0.3%	-$4.843	$5.182	3.8%
Disk hardware total $ M	$2.368	1.1%	$2.367	$4.735	4.5%	-$2.655	$7.390	5.4%
Total 5 year TCO $ M	$220.119	100.0%	-$114.301	$105.818	100.0%	-$32.026	$137.844	100.0%

Comprehensive 5-year TCO comparisons, all capital and operating costs- hardware, software, data center, networks, staffing.
IBM zRACE analysis- January 2008.

Figure 21: Real Customer – System z Consolidation vs. Existing or New x64 Distributed – Financial Analysis

The 5-year $M costs of the three alternatives A, B & C are shown in the dark shaded columns. The percentage cost breakdowns by cost type for each are shown alongside. We also show the per-cost item $M savings offered by the B-System z solution over both the A-Existing Distributed, and the C-New x64, alternatives.

The B-System z solution offers a dramatic, $114M+*(51.9%)* total 5-year net TCO saving over the A-Existing Distributed Servers retained alternative. $93M of this saving came from far lower staff costs, $23M from lower software costs, $17M from lower connectivity/enterprise network costs, and $11.7M came from lower power/cooling/data center space costs. These were reduced by the $28M higher server hardware, and the $2.4M higher disk hardware, costs incurred to acquire/run the new System z server/storage platforms.

The System z solution also offers a large saving of over $32M *(23.2%)* total 5-year net TCO savings, over the C-New x64 VMware alternative. Here, $42.6M of this saving came from lower staff costs, $6.4M was from lower software costs, $11M came lower connectivity/enterprise network costs, and $2.7M from lower disk hardware costs, less the $31.6M higher server hardware costs to acquire/run the 5 new System z servers.

These dramatic and revealing figures show massive saving sums obtainable by adopting the System z solution versus both the "do nothing" *(continue existing distributed sprawl for 5 years)*, and also versus the best that the x64 world can offer with its VMware alternative. The data clearly confirm the high ROI and NPV available in suitable cases like this. Just over half the considered distributed server population was found most suitable for System z consolidation, with other strong IBM consolidation options deemed better fits for the other just under half.

Our Analysis

Each of the above impressive, mass-distributed server consolidation projects fully exploited the unique System z mainframe's ability to support 100s up to 1,000 significant weight virtual servers. It achieves this through IBM's advanced z/VM virtualization hypervisor, running the open, reliable, and scalable Linux OS, and exploiting low-cost IFL specialty System z processors for Linux workloads at new low levels of mainframe cost. Today's z/VM V5.3 release draws its inspiration, high capacity-scale, and extraordinary depth from its illustrious history of pioneering virtualization developments in the IT industry for almost 40 years. z/VM efficiently deploys *(timeshares)* all available physical mainframe system resources – including CP/IFL processing cycles, memory, networking, and storage I/O– to all of the supported "virtual" servers.

The typical benefits delivered are quite clear from these examples, as well as from numerous other published data on the topic, and are impressively large. We summarize and quantify the range/average levels of all these benefits below.

- **Numbers of servers reduced dramatically:** Server consolidations ratios of 200-300:1 "substantial" distributed server workloads per System z9 EC mainframe can readily be achieved, often more for "light" virtual servers. *(Record tested is over 97,000!!)*
- **Far more efficient CPU/hardware utilization:** 10-20:1-times higher server hardware utilization. System z mainframes running consolidation workloads regularly achieve 85-95% sustained utilization, versus the usual 5-7% on scale-out distributed servers.
- **No. of virtual servers per CPU core:** 15-30:1-fold higher on System z IFL cores *(30-45 VS/core average, range 3-70 VS/core)* than on x64 CPU cores *(2-3 average)*. One mainframe processor *(with z/VM virtualization)* is able to run the work of so many x86 processors through its unique design point. The z9 MPU was extensively optimized to best run the most demanding, multiple, mixed commercial workloads, at the highest rates of CPU utilization, through efficient multi-threading, lightning-fast context switching, and on-chip accelerators *(e.g. cryptographic)*. These System z processors are also much more fully supported by the ample, and high-performance, other mainframe system resources installed.
- **Drastically fewer mainframe CPUs needed:** Typically 10-20:1-times fewer System z processors are needed than on the original displaced distributed servers. *(For example, 1,782 z vs. 23,000 distributed CPUs in case # 1)*. This means far fewer servers and thus much less hardware to buy/run, far less power and cooling, and more particularly, sharply lower software costs *(see below)*.
- **Far Less Software Licenses/Sharply Lower Software Cost:** Today software is often a much larger cost than that of the hardware on distributed solutions using enterprise middleware or applications. Such software is normally licensed on the basis of the number of CPUs installed *(not used!)*. 10-20:1-fold fewer CPUs need to be licensed with System z *(usually at slightly higher per-CPU $ prices)*, translating to huge software license, and S&S cost, savings on the middleware, etc., software, needed. A range of 4-8:1 of total software cost savings have been reported, easily reaching $7-digit savings in larger consolidations.
- **System/Server Support Staffing:** 4-8:1-fold fewer support staff needed/lower support staff cost. Typically, one FTE can support 100-200Linux virtual servers on System z, versus the norm of 20-30 servers/support FTE for most distributed server estates. With average support FTE staff earning c. $70,000 in the USA/Western Europe, the savings here are usually huge in $ value, of similar order to the software savings above.
- **Hardware Costs Now Often Broadly Comparable:** Far less System z hardware is needed, but remains relatively costly per unit. However, System z mainframes have a much longer service life *(5-8 years typical)*, and can also be superbly used/ fully exploited *(at up to 95% utilization)*. A far larger set of distributed hardware is needed, at lower unit costs, but with a much shorter life, and only able to be wastefully used. Another crucial factor here is that, frequently, an existing mainframe can be used and/or extended to support the additional workload of a smaller distributed server consolidation, at low incremental costs. By contrast, dedicated, new, distributed hardware must be bought for every new workload.
- **Power and Cooling Energy Cost Savings:** Often brings a 10-20:1-fold, up to 90%+ reduction, for example, from $600/day for a distributed solution to $32/day for System z, an ever-more significant cost saving and power availability factor today. A single, powerful, mainframe-processing chip executing scores of VS workloads efficiently consumes far less energy than many x86 servers, which use far more power-consuming CPUs and other components. Switching on a 580 MIPS System z IFL processor consumes only about 20 watts.
- **Lower UPS Capacity Costs:** Because of these major reductions in power consumption under a consolidated System z virtual server environment, proportionately large capital and operating cost savings on providing/operating UPS *(Un-interruptible Power Supplies)* plant capacity are delivered.

- **Data Centre Floorspace Savings:** Often 25:1 reductions, from 10,000 sq. ft. for distributed servers, to 400 sq. ft. for one System z full installation. These huge "data center space recovery" gains from System z adoption can often defer or avoid the need for data center extension capital spending of $Ms or $10Ms on the additional data center space and plant costs that would otherwise be needed were distributed servers retained.
- **Far Less Networking Equipment/Cabling Cost:** 3-4:1 average reduction in networking equipment and cabling costs. Internal virtual networking at memory speed inside System z replaces most of the cluttered, slow, unreliable, external networking needed with previous distributed server configurations.
- **Overall TCO Savings:** 30-50% total TCO savings are consistently reported with System z consolidated solutions *(see below)*.
- **Typical Payback Period/ROI:** Experience to date suggests rapid investment cost payback periods on mass-consolidation to System z ranging from 15 months to 2.5 years, which are excellent ROIs. Variability depends on whether existing System z machines can be expanded to accommodate the additional consolidation workload, whether new or 2nd-user mainframes are used, and how much external help is purchased to aid the migration.

Just like that, all prior x86/x64 servers and operating systems are made obsolete!

- **Robust DR/BC becomes feasible and affordable:** With large workloads from scores/hundreds of virtual servers consolidated into one System z mainframe, technically feasible and affordable DR/BC, usually already in place at most mainframe sites, can be easily extended to the new virtual server population.
- **Far faster server provisioning:** 15-1,000:1-times faster *(see examples above)*. On System z, new, cloned, virtual servers can often literally be provisioned in minutes *(or less)* and absolutely within a fraction of a day *(worst case)*. To purchase, have delivered, install, cable, software load, test, and commission a new distributed server usually takes weeks-to-months of elapsed time in most IT shops. *(Unless unused servers are stocked in-house – this is unusual because stocks age/date too fast in x86 space.)*

The mainframe solution offers other, often highly-significant-to-customers, operational, security, performance, and manageability advantages; some of which we cited earlier in this White Paper. The four real-world examples profiled, and the average benefit ratios summarized above, definitively show the strong business case for mass-distributed server consolidation to Linux-on-z/VM on the System z mainframe as highly compelling, and to offer rapid payback on the investment needed.

For our last words in the body of this White Paper, we can do no better than to quote James Stallings – then General Manager, IBM System z mainframe – on announcing the IBM internal consolidation project *(# 1 above)*. He said:

> "The mainframe is the single, most powerful instrument to drive better economics and energy conservation at the data center today. By moving globally onto the mainframe platform, IBM is creating a technology platform that saves energy while positioning our IT assets for flexibility and growth. "

IBM now expects that firstly hundreds, and later thousands, of its enterprise customers will follow this path with their own mass-distributed server consolidations, buying many new System z mainframes to do so. With such a powerful business and technology case highlighted above, we fully expect that these IBM hopes will be fulfilled over the next several years.

Appendix A: Additional Mainframe Virtualization Facilities

Introduction

In Sections 3, 4, and 5 we provided coverage of System z's main virtualization capabilities, to broadly introduce these to readers unfamiliar with today's mainframe, and to allow comparison with the rudimentary virtualization offered today on x86/x64 server platforms *(e.g. VMware)*. Decades of intense development extended the mainframe's refined virtualization, so the System z portfolio now offers numerous other, fine-grained virtualization capabilities. In this Appendix A, we highlight some of these other important virtualization capabilities.

System z Interpretive Execution *(SIE)* – Advanced Virtualization Hardware ISA Support

Providing the foundation System z hardware support for its advanced virtual server hosting is the System z MPU's **Start Interpretive Execution *(SIE)*** instruction. The operand for the SIE instruction is a state descriptor for a System z LPAR or VM. The instruction can equally accommodate fixed-storage guests and page-able virtual server guests, and its interception controls allow hypervisor intervention. The instruction reduces context switch times to switch processing between time-shared virtual servers, which is consequently ultra-fast on System z, a crucial factor for mass virtual server hosting. System z implements two levels of SIE to provide the widest range of benefits, without significant downsides, for a wide range of situations. This required multiple control register sets, and coping carefully with the potential instruction behavior differences at the two levels. To successfully achieve this required considerable IBM architectural and hardware investment. Consequently, SIE now imposes zero performance penalties for running the z/VM hypervisor within a System z LPAR *(for added isolation/security)*, and no shadow page tables are required for DAT-on guest virtual servers.

Additional Mainframe Virtualization Facilities

Figure A1 below itemizes, and briefly describes, other important System z virtualization features and capabilities. These combine with the fundamentals discussed elsewhere in this White Paper, to add broader strength and greater depth to the platform's comprehensive virtualization portfolio:

Additional Mainframe Virtualization Facilities	
System z Virtualization Feature	**Virtualization Feature Description**
Zone Relocation:	• SIE capability that provides multiple zero-origin storage regions *(logical partitions)* on one system. *(See also above SIE sub-section.)* • Enables I/O subsystem to access partition memory directly, without requiring hypervisor intervention.
Translation Lookaside Buffers *(TLBs)*:	• Large allocation of microprocessor space for TLBs directly benefits virtual server scalability. • z9 and z990 provide a TLB arrangement which advantageously uses two buffers. • The second-level TLB feeds address translation information to the first-level TLB when the desired virtual address is not contained in the first-level TLB.
Multiple Image Facility *(MIF)*:	• Enables channel sharing among multiple LPARs. • I/O devices on shared channel paths can be accessed simultaneously by sharing LPARs *(or restricted to a subset of sharing LPARs)*.
Logical Channel Subsystem *(LCSS)* Support:	• Allows a z9 and z990 to be configured with up to 1,024 channels *(512 channels for z890.)* • 256 channels can be configured for each LPAR, with selected channel sharing among LPARs possible.
I/O Priority Queuing:	• Allows high-priority workloads to receive preferential access to I/O subsystem. • Supported by Intelligent Resource Director and virtualized by z/VM.
HiperSockets:	• High-speed, security-rich TCP/IP connectivity among LPARs. • Memory speed communications links.
Adapter Interruption Pass-Through:	• OSA-Express *(Ethernet)* and FCP *(SCSI)* virtual machine I/O can be performed while z/VM guest image is running in SIE mode. • "Thin" interrupt passed to z/VM Control Program when I/O operation belongs to an idle guest system.
QDIO Enhanced Buffer-State Management *(QEBSM)*:	• Two new machine instructions designed to help eliminate overhead of hypervisor interception. • Host Page-Management Assist *(HPMA)*.
Interface to z/VM Paging and Storage Management:	• Designed to allow hardware to assign, lock, and unlock page frames without hypervisor assistance.
Layer 2 *(MAC)* and Layer 3 *(IP)* Network Switching:	• OSA and z/VM support enables virtual IP and MAC network switching without requiring a hosting partition.

Figure A1: Additional Mainframe Virtualization Capabilities

This brief "additional capabilities" summary amply illustrates the extensive richness, depth, the optimized hardware-hypervisor interaction, scalability, and breadth of the long-optimized virtualization capabilities the System z platform offers today, the result of its many years of sustained development.

Appendix B: Anatomy of Industrial-strength, Extreme Virtualization. Comparing IBM's z/VM V5.3 Hypervisor with VMware ESX Server V3.1 on x86/x64

IBM z/VM – Extreme Consolidation/Virtualization Enabler

This White Paper presents the case for IBM System z mainframes *(running up to hundreds of Linux-on-z/VM virtual server workloads migrated/consolidated from wasteful, physical, distributed x86/x64 or UNIX servers)* as today's best way to slash distributed computing costs/wastage *(detailed in Section 2, and in Appendices C, D, and E)*. Our research found the System z, z/VM-based extreme virtualization-based consolidation approach now provides far superior solutions, on all main comparison factors, and by wide advantage margins.

The long-proven, ultra-sophisticated IBM z/VM bare-metal hypervisor *(in its V5.3 latest/current release)* is the core mainframe software enabling these striking System z distributed server consolidations, of which four major customer examples were assessed in Section 6.

Too few IT professionals are yet aware of the compelling advantages of the radical, System z-based solution to the distributed system nightmare. Most remain more familiar with **"many x86/x64-onto-larger-x86/x64"**, **"many-UNIX-onto-larger-x86/x64"**, or **"many-UNIX-onto-larger-UNIX"** server consolidation strategies. This Appendix closely argues why and how System z with z/VM provide a clearly superior solution over VMware x86/x64 server virtualization, the most common alternative.

Competitor VMware ESX Server and VMware Infrastructure 3

Whenever consolidation/virtualization for x86/x64 servers was discussed, one vendor/product line springs to mind. The near-generic name in "x86/x64-onto-x86/x64 server virtualization" was EMC subsidiary VMware and its core product, **VMware ESX Server** *(currently in V3.1 release)*, as we overviewed in Section 3. ESX Server is its flagship x86/x64 bare-metal hypervisor, and the centerpiece of its **"VMware Infrastructure 3"** suite. Pioneer in this market, and a most successful ISV firm, VMware enjoyed a dominant market share in x86/x64 virtualization *(55% last year)*, and related VMM management software. Today it offers its ESX Server-based, **VMware Infrastructure 3** suite *(see Footnote[4] below)* in three editions, "Starter", "Standard" and "Enterprise", each offering differing capabilities, bundled with other VMware target server products, at from $1,000 to $5,750 per 2 CPUs initial license fees. VMware also offers its VMware VirtualCenter management software, used to manage virtual servers hosted on ESX Server.

Because earlier VMware products were widely deployed, many IT professionals had acquired awareness, or some knowledge/experience of, the VMware offerings.

Therefore, in this Appendix B, we highlight major areas where System z's z/VM hypervisor offers large advantages over the comparatively young VMware ESX Server stripling, to help those familiar with VMware to better understand the deeper strengths of the z/VM mainframe solution.

Footnote[4]:

The VMware® Infrastructure 3 family now includes:

- **VMware ESX Server:** Main bare-metal hypervisor to host virtual servers on x86/x64 platforms.
- **VMware VMFS:** Cluster file system for virtual environments.
- **VMware Virtual SMP:** Enables VMs to scale-up to 4 CPUs on ESX Server
- **VMware VirtualCenter Agent:** Virtual server agent to permit management by VMware VirtualCenter.
- **VMware VMotion™:** Moves a live virtual server from one ESX Server instance to another.
- **VMware High Availability** *(HA)***:** High availability solution for virtualized environments.
- **VMware Update Manager:** Patch management/tracking for ESX hosts & VMs.
- **VMware Storage VMotion:** Live migration of VM files across storage arrays.
- **VMware VirtualCenter:** VMware virtual environment management software tool
- **VMware DRS:** Monitors and allocates resources across VMs.
- **VMware Consolidated Backup:** Centralized VM backup.
- **VMware Site Recovery Manager:** Disaster recovery automation solution.

Massively Greater z/VM Scalability

- **Hypervisor Scalability:** With z/VM, users can begin with just a few virtual server *(VS)* images, but can successively grow their VS environment to large numbers of VS on a single copy of z/VM, usually running on their existing mainframe. Over 97,700 VS have been shown to run on a single z/VM copy. Practical deployments of several hundred *(typically 250-450 at upper end)* substantial VS, per z/VM partition on a single mainframe box, have frequently been reported. z/VM is highly efficient in allocating and sharing real system resources on a much more granular basis than VMware ESX. VMware ESX can support a maximum of only 128 powered-on VS on the largest available x86 server, and in practice rarely ever runs even ten VS. This large hypervisor scalability advantage brings the z/VM mainframe approach major TCO and operating advantages. z/VM's advanced, built-in systems management tools can also be leveraged over far large numbers of VS, with far lower support staff efforts, than on an equivalent x86/VMware environment.

- **Far Superior z/VM VM Scalability:** z/VM offers substantially greater VS scalability than VMware ESX Server 3. This is true for the three major dimensions of scalability, the amounts of CPU, memory, and of I/O, which it can support. This is hugely important, because it enables z/VM virtual servers to run a much wider range of sizes of Linux application workloads than VMware ESX, including far heavier-weight workloads.

  - **8-Fold Greater CPU Scalability:** VMware ESX 3 VS cannot be configured with more than 4 CPUs *(with the "VMware Virtual SMP" add-on software)*, whereas z/VM can now support up to 32, much more powerful z mainframe CPUs in a single VS. This severely constrains the maximum image/VS size on VMware to well below one-eighth *(in terms of CPUs, much more in relative power)* of that available to a z/VM VS on a System z mainframe. Consequently, all larger x86 applications/images must be run on a non-virtualized, dedicated x86 server hardware configurations, completely excluding these from any virtualization benefits.

  - **4-Fold Greater Memory Support:** VMware ESX limits each VS to 64GB of memory, whereas up to 256GB can be allocated to a VS on z/VM. This imposes substantial constraints on memory-intensive Linux workloads on VMware ESX Server that are memory unconstrained on z/VM, and which will therefore perform much better on System z.

  - **Massive I/O Capacity on System z & z/VM:** System z and z/VM are optimized to share massive real and virtual I/O capacity with extreme scalability and impeccable data integrity. They fully use the up-to-256 channel I/O paths per LPAR *(and/or z/VM copy)*, and the up-to-336 dedicated I/O channel processors maximum available on a System z mainframe. By contrast, I/O scalability is particularly weak in virtualized x86 systems. VMware ESX users are specifically advised to sharply limit, or to avoid altogether, deploying/hosting I/O-intensive workloads on VSs supported by this hypervisor.

Much Higher z/VM Consolidation Ratio, Per CPU, Per Hypervisor, Per System

- **Greater CPU Sharing, Higher Consolidation Ratios:** VMware ESX sharply limits the number of VS able to share a single x86/x64 server CPU *(generally to well within single figures)*. By contrast, z/VM allows users to allocate a System z CP or IFL's capacity to any number of VSs. Practically, a range of from 3 to 70, and an average of 30-45 significant VSs, can readily be run per System z CP or IFL processor *(each delivering 580 mainframe MIPS on a high-end z9 EC mainframe)*. This means users can host far more virtual servers per CPU with z/VM on System z than with VMware ESX on x86/x64. This is particularly beneficial for consolidating server workloads that have low CPU utilization levels, where far higher consolidation ratios *(physical server processors/VS processors)* can be achieved with z/VM on System z, a major cost-saving factor.

- **Far Higher Architected VS Limit/Hypervisor Instance:** VMware ESX is architecturally limited to supporting a maximum of only 128 VSs per hypervisor instance, but the practical limit is far lower. VMware ESX users typically host single figures of VSs per hypervisor instance, and only for "production-lite" server workloads. By contrast, the architected limit for VMs on z/VM is more than 97,000. Numerous System z customers today are running several hundred Linux VMs on a single copy of z/VM. We expect this already-huge, practical System z VS capacity will more than double with the next-generation, z6-MPU-powered, new System z mainframes, expected soon in Q1 2008, further extending its lead. With VMware, substantial numbers of x86 servers therefore still have to be purchased, commissioned and managed for large environments, because the consolidation ratios *(both per hypervisor instance, and in terms of VSs/CPU)* that can be achieved are far poorer than with z/VM on System z.

z/VM Near-immediate CoD, Far Faster Provisioning, No Disruption

- **Real CPU CoD:** z/VM users can add real, additional CPU capacity to a z/VM LPAR *(or set of LPARs)* dynamically in real time, without suffering any hardware, z/VM, or VS outage. This additional CPU capacity *(or memory)* is immediately applied to the virtual server environment, with no additional configuration steps required. *(The System z also offers the IT industry's most extensive suite of CoD options to add real physical capacity to a system rapidly in a variety of forms)*.

Growing real capacity in an x86 VMware environment means adding more physical x86 servers to an existing infrastructure, and is thus disruptive to the live application service. With System z CoD, users do not have to pre-purchase hardware capacity before it is actually required, but can call it off near-instantly, only when actually needed. In a typical x86 environment, provisioning a new server means the hardware must be bought, delivered, unpacked, configured, connected, powered, cooled, and loaded with software, before it is ready to use, usually taking days or weeks, even if the unit is in stock. If the usual new server procurement cycle times are also added, the whole x86 process can take many weeks, or often months, beginning to end. The availability of new x86 virtual servers is thus vastly delayed, compared to the normal minutes, or sometimes even seconds, actually needed on z/VM to add VSs. These massive x86/x64 delays in provisioning new distributed server capacity can seriously affect business results and performance.

Superior z/VM Performance

- **Extensive z/VM In-memory Support Boosts Performance:** z/VM makes extensive use of advanced, in-memory processing techniques to significantly boost its performance. These go far beyond the limited such exploitations used in VMware ESX. These "accelerators" help enable the large-scale virtual server hosting capabilities of z/VM. The in-memory techniques used in z/VM include:
    - **z/VM Minidisk Cache Support:** z/VM can cache VS disk data in real storage. This is most beneficial when this disk data is primarily read-only, and must be shared among multiple VSs. This enables any or all Linux guest images to enjoy memory-to-memory data transfer speeds on disk read operations, enhancing overall z/VM system scalability, and I/O performance for individual Linux guest images.
    - **z/VM Virtual Disks in Storage:** This feature enables Linux guest images to swap *(page)* to memory instead of to disk. This enhances system scalability and performance by minimizing memory footprint for each Linux guest image, without normal overhead of paging guest memory to physical devices. Requiring Linux to page forces the Linux kernel to be highly-efficient in memory use. This brings performance/scalability gains when running multiple copies of Linux guest VSs on a z/VM hypervisor, especially when the paging volumes are in memory, and not in the I/O subsystem.
    - **Shared Program Executables:** z/VM also supports sharing of program executables among Linux guest VSs using its **Discontiguous Saved Segments *(DCSS)*** feature. This additional, important data-in-memory technique offers major business value where a hypervisor is hosting large numbers of Linux guest images, exactly the z/VM model of large-scale virtual server hosting. Up to hundreds of significant weight Linux guest VSs on z/VM can share one program/set of executables *(or many)*, significantly reducing the amount of real resources *(e.g. memory and disk)* needed to host a large virtual server environment. This reduces hardware resource use costs, and cuts support cost, because IT staff need only manage one single DCSS code base for hundreds of virtual servers.

These advanced in-memory techniques, years of optimization, direct System z hardware assists, plus the much higher power and capacity of the underlying System z hardware, give excellent performance to all z/VM VSs, even when running more substantial workloads that ESX Server cannot practically run.

VMware ESX merely detects pages of memory that are identical across VMs, and can remove these extraneous VM pages from real memory. This improves VMware's real memory utilization, but the process is only done by the VMware ESX Server hypervisor as a background operation when spare CPU capacity is available, so it does not offer instantaneous sharing of memory among guest images. The benefit of this support is lessened when VSs are moved from one copy of VMware to another *(Using VMotion)*.

Additional z/VM-z/OS Co-residence Synergy Benefits

It is important to understand that System z mainframes are mostly run as "central shared service" hosts, normally supporting numerous workloads and applications, including those run under its flagship z/OS. Virtual servers running under z/VM can access and share data with, and communicate directly via high-speed internal networking, with all these other mainframe workloads, all on and within the one system box.

- **System z Logical Partitioning *(LPAR)* Complements z/VM Virtual Hosting:** System z mainframes offer the most sophisticated, dynamic, efficient, and secure partitioning capability of any server, with the renowned System z LPAR *(logical partitioning)* capability, by far the industry's longest-proven, most industrial-strength, and highly-refined server hardware partitioning approach. System z LPAR support allows users to host one or several, large or small, Linux-on-z/VM virtual server hosting environments, side-by-side with z/OS configurations. Such co-residency of Linux-on-z/VM and z/OS offers the following benefits:
    - **High-speed, secure, low-latency connectivity:** Between Linux-on-z/VM and z/OS system applications using HiperSockets™ virtual networking.

- **Major DR/BC operational cost savings:** Existing z/OS data backup *(DR/BC)* facilities can be easily extended to also support Linux disks and data for the co-resident Linux on-z/VM environment, at low cost.
- **Sound Mainframe Processes/Tools Shared:** Long-proven, robust, and highly-disciplined mainframe operational policies and tooling can be extended to include Linux-on-z/VM, offering additional cost savings and data protection.

System z LPAR support also means users can run multiple copies of z/VM on a single mainframe, each hosting large numbers of virtual servers. Most standard x86 servers can only host one copy of VMware.

More Flexible z/VM Operations

- **Resource over-commitment support *(memory, CPU, network, I/O)*:** A distinguishing characteristic of z/VM is its ability to provision virtual server resource levels that significantly exceed the amount of underlying real resources available, and yet deliver excellent QoS and performance to all. This has several major practical benefits:
  - **Highest Resource Use:** Real resource utilization is thus optimized in a Linux-on-z/VM environment, maximizing the business value of customers' System z hardware investments.
  - **Demand Spikes Smoothly Handled:** Unexpected *(or expected)* virtual server workload spikes can be easily met and processed without requiring additional real hardware, whilst maintaining acceptable levels of performance and response times. The QoS and performance systematically delivered by the System z mainframe is legendary, and is far beyond what any x86/x64 platform can deliver today.
  - **New Virtual Server Provisioning Near-Instant:** Generally, users can add lower-utilized virtual servers to an already-configured, highly-utilized Linux-on-z/VM system without requiring any additional real hardware, because of the high efficiency with which real resources are used. Provisioning such new z/VM virtual servers also takes only minutes, rather than weeks or months on x86.

  Whilst VMware ESX supports some modest resource over-commitment, it offers far more limited, and much less sophisticated, resource-use optimization and scheduling than the System z, z/VM platform.
- **Reconfiguration of VSs:** z/VM users can modify VS configurations without requiring a re-boot of the guest Linux OS. This is not allowed/possible with VMware, where the guest system must be stopped and then be re-booted after configuration changes are made. This enhances Linux-on-z/VM virtual server application availability, and improves support staff productivity, by allowing "real-time" configuration changes without having to schedule planned outages.

Superior Tooling and Much Lower Support Staffing With z/VM

- **Top-strength z/VM Command, Control, Monitoring, and Automation infrastructure:** The System z VM environment has been used by thousands of demanding mainframe customers, some for decades. IBM created, matured, and today supports, a broad suite of command & control software tools so that customers can efficiently manage their VM environment. These tools now also help users manage large numbers of virtual Linux servers on z/VM. Their high productivity allows mainframe customers to manage much larger numbers of virtual servers per FTE *(Full Time Equivalent)* with QoS and business value that cannot be realized at all in a virtualized x86 environment *(e.g. highly-granular resource consumption reporting, capacity planning, and automated operations, etc.)*. Many z/VM customers *(including our four Section 6 examples)* have reported that each of their z/VM support staff *(FTEs)* could readily support at least 100, and even up to 200, Linux virtual servers on the platform. Standalone, scale-out, distributed x86/x64 servers typically require one support FTE per 20-25 servers, and even the best VMware virtualized x86/x64 configuration typically improves this only to the 40-50 virtual servers/FTE range. Typically, **z/VM is thus up to four times lower in support staff time/costs per virtual server** than VMware on x86/x64. VMware does offer some useful tools to help manage VMware environments, but these are evidently less comprehensive and productive.
- **z/VM Runs z/VM Guest:** Another important z/VM hypervisor unique is that z/VM can itself run z/VM as a guest. *(Recursion)*. This provides valuable extra flexibility, and further cost savings. It allows customers to cost-effectively test their z/VM systems in guest test z/VM environments. This allows each IT staffer to have a personal copy of z/VM for their own training and development. It also enables additional z/VM hosting services to be securely offered, within or outside the company, without requiring additional LPARs. System z ISVs, remote hosting service provider companies, and DR/BC firms, are all extensive users of this unique z/VM-on-z/VM technology.

  VMware ESX cannot host another copy of VMware at all, so no such useful options are available.

Superior z/VM Integrity and Security

- **Fault Isolation/Hypervisor Security:** Virtualization of all system resources has long been fundamental to the System z architecture. IBM has invested continually in extending mainframe architecture, hardware, firmware and software *(z/VM)* for decades to enable the most highly secure, and the most reliable, VSs. z/VM is an integrated piece of mainframe z/Architecture, and not an "add on".

  Original Intel x86 architecture *(and its initial x64)* successor not only provided no hardware support for virtualization, but also rendered the platform exceptionally difficult to virtualize. VMware was the first company to offer software that broke these constraints to virtualize the x86 platform from 2001. However, because of x86 architectural limits, VMware ESX could only achieve this using an "add-on" approach that added considerable processing overhead, and created security issues. Although now highly fashionable in 2008, virtualization remains a still-recent initiative for the x86 industry.

  The security infrastructure for VMware-on-x86 systems thus remains immature in functional richness, and lacks the years of customer use/verification that has helped shaped the exceptional security support available with z/VM.

  Industry analysts Gartner have warned that companies rushing to deploy Intel-based virtualization technology for distributed server consolidation are facing many security issues, and are exposing themselves to new risks. Gartner also highlighted that *(x86/x64)* virtualization, as any emerging technology, has become a target of new security threats:

  "These security issues are being overlooked, and security tools for virtualization are also lacking. Through 2009, 60% of production virtual machines (x86/x64) will be less secure than their physical counterparts" commented Gartner Vice President Neil MacDonald, in a published statement.

  Intel and AMD recently each added their unique *(and different)* basic virtualization hardware support extensions into their newer x64 server chips. VMware ESX Server can exploit these, which should reduce the current significant VMware processing overhead, but only for users of the newest servers equipped with these new chip-level virtualization support facilities, and only when running the latest, 64-bit guest OSs only. Just like that, all prior x86/x64 servers and operating systems are made obsolete!

- **Run Multiple Copies of Hypervisor on a Single Server:** System z users can run multiple copies of z/VM on the same mainframe while sharing all real CPUs, I/O, and networking resources among the z/VM LPARs. This offers an added level of workload isolation, such as separating development/test servers from production servers. The sharing of real resources among multiple z/VM LPARs means that users can also shut down one z/VM system and use its real system resources in other instances of z/VM on the same mainframe. This supports virtual server failover, and z/VM environment maintenance, without requiring any duplication of the real hardware that is invariably needed with x86/x64 environments.

Needs no VM Mobility

- **VM Mobility Support:** VMware does offer the ability to move a running VSs to another copy of VMware *(VMotion)*. In the x86 world, where demand spikes can easily overwhelm the typical x86 servers running VMware, and where server hardware frequently fails, such a capability is mandatory. It enables a spiking guest VS workload to be moved onto another, larger server's copy of VMware, or permits failing original server hardware to be repaired/replaced, without losing guest service. The running guest virtual server workload is first moved onto another copy of VMware, on another server. Software changes can then be implemented on the original server, or it can be shut down for needed hardware repairs and then restarted. Finally, the original workload can be moved back with VMotion.

  z/VM on the System z mainframe is inherently far more capable of handling workload spikes on any specific guest VM, because it can draw on the much larger resources of the mainframe at most times. Additionally, the sophisticated CoD support on System z allows users to rapidly add real hardware capacity *(CP, memory)* when needed, on a permanent or temporary basis. The underlying hardware reliability *(in terms of MTBF – Mean Time between Failure of hardware)* on a System z is now measured at around in one incident in five decades. In addition, many hardware upgrades, and I/O configuration changes can also be made without requiring any z/VM outage. Additional workload management options are available when running multiple copies of z/VM on the same mainframe, or across multiple mainframes. With such capabilities, there is no need for guest mobility in a z/VM environment.

Favorable z/VM Software Costs

- **Hypervisor and Related Cost:** In the volume x86/x64 server marketplace, relatively low hardware costs have long been the norm/expectation, but software costs are often high for premium-branded middleware, software tools and applications. Nowhere is this truer than with VMware's prime x86 server virtualization software, which carries eye-wateringly-high license and subscription and support fees. Today, ESX Server is offered as the central component of **VMware Infrastructure 3**. This combines select VMware ESX server rights with related VMware server software, into Starter, Standard, and Enterprise bundles. License fees for these bundles now start at $1,000, $3,750, and $5,750 per each 2 processors, with an additional up to 21/25% for annual service and subscription for the top support offering. Thus, when a 4, 8 *(or above)* processor server system *(or multiple equivalent smaller servers)* are licensed for the full-function Enterprise bundle, and the additional site system management software needed *(VMCenter Management Server)*, is added, the VMware license fees rapidly exceed five $ figures. They also head rapidly toward six $ figures for larger configurations, in initial license fees alone, to which the steep service and subscription fees must be added. With typical VMware ESX virtual servers per CPU deployment ratios of 2-4 average, VMware Enterprise users must thus pay an average of $700 in initial license fees per virtual server just for the target system, without adding in the management server software.

  As sector pioneer, last year *(2006)* the 55% x86 virtualization market segment share leader, and a recent stock market darling with investors to please, VMware to date was able to command its high license fees, and thus to post the astonishingly rapid revenue growth it has enjoyed. However, VMware license fees are so high as to constrain quite substantially the range of cases where it may be economic to consolidate multiple x86 workloads onto VMware-virtualized x86 platforms.

  VMware is also now facing far stiffer x86 virtualization market competition, from recently emergent, open-source x86 hypervisors *(Xen and KVM)*, and from major new industry players like Oracle *(Oracle® VM)* and Sun *(xVM)* with license fee-free hypervisor offerings. XenSource, for example, recently stated that its directly VMware-comparable, Xen hypervisor, was currently offered at only 1/10 of the price of its VMware counterpart. Such determined competition seems certain to drive x86 hypervisor prices sharply down, and probably to commoditize the category completely within a year or two.

  z/VM Version 5.3 on System z9 is licensed for the number of System z processor engines it is to run on, at from as little as $6,750/engine *(for 26 engines and above)*, up to $22,500/engine *(for 1-3 engines)* for the whole enterprise *(across all System z machines)*. With each full z9 mainframe processor engine capable of supporting up to 70 virtual servers *(on average)* under z/VM, hypervisor software costs per Linux virtual server can thus be as little as $100 per virtual server, one-seventh of VMware ESX charges stated above.

VMware Constrained to Light-mid-weight Workloads, z/VM Unconstrained

Despite the wide media coverage and large customer base it has attracted, VMware ESX Server is practically constrained in its actual field deployment mainly to running relatively lightweight or mid-scale x86 workloads.

- **VMware – Lightweight Usage the Norm:** Many thousands of customers have tried VMware ESX Server since its first 2001 release, and these now constitute a broad user base. However, the average depth and workload duty of their VMware exploitation remains relatively modest for most users. Whilst customers have been quick to first deploy VMware ESX Server to test x86 server virtualization, and/or to support development or application test deployments, wide-scale production application deployment, or use for substantial workloads, remains uncommon. It should also be borne in mind that VMware ESX Server was designed to host x86/x64 server workloads that are **not CPU-, memory-, or disk-I/O-intensive**. Indeed, users are advised to be most careful about, or to avoid, hosting such workloads on VMware ESX Server. Excluded by this advice, for example, are the common categories of heavy database workloads *(DB2, Oracle, and SQL Server)* and substantial e-mail/collaboration server workloads *(Lotus Domino or MS Exchange)* which users find should not generally be run on VMware. Heavy I/O workloads are particularly to be avoided owing to the generally weak I/O capability of x86/x64 server architectures. The processing overhead incurred from the use of this hypervisor can also be significant, affecting application performance. In addition, widespread concerns remain over security in production VMware environments, another such inhibitor.

  By contrast, in the Linux-on-z/VM virtual server-hosting environment, the robust, high capacity z/VM hypervisor suffers no such limitations on the System z hardware platform with its massive IFL engines, large memory and huge I/O throughput capabilities. These easily allow substantial and heavy-duty Linux virtual servers to perform well under a z/VM environment, but to also concurrently support up to many hundreds of light utilization virtual Linux-on-z/VM servers with outstanding efficiency, and reliability.

It should be noted that VMware recently introduced **VMware ESX Server 3i**, and associated VMware Infrastructure 3i bundles, of similar broad functionality to that described above. However, 3i is delivered in a thin layer, OEM-system-vendor pre-installed, near-firmware-style form *(not as end-user delivered software)*, and comes only with brand-new x86 server hardware *(which has in-built Intel/AMD hardware virtualization support)*.

This comparison between IBM's z/VM V5.3 hypervisor for System z, and VMware's ESX Server 3.1 *(the market-leading x86 hypervisor product family)* clearly highlights many important areas of deep superiority/advantage for z/VM. These make it much the better hypervisor platform for extreme consolidation of Linux-on-z/VM workloads originally running on scores/hundreds of distributed x86/x64 servers, onto the ultra-efficient, far greener, and far lower TCO, System z mainframe alternative.

Appendix C: Scale-out Distributed Servers – Proliferation Catastrophe!

Server Numbers Exploded, Pushed IT Costs Through the Roof

At 2007 close, c. **32M servers were running worldwide**, five times plus more than at end-1996 *(6M)*. About 28M *(87.5%+)* of these were distributed, scale-out, volume servers. The global IT user community pays soaring costs to manage/support *(people costs, software costs, data center space, power, cooling, etc.)* this still-growing, sprawling server population. These costs **exploded 6-fold, from <$22B in 1996 to $120B+ in 2007**. Users' costs for electrically **powering and cooling** these IT systems also rocketed. They jumped from under $10B in 1996, **to c $22B in 2007**, and are **projected to hit $44B by 2010**, if current trends continue unchecked *(see Figure 2 on page 10)*. Combined, such global IT operating costs were near-triple 2007 new server expenditure.

Scale-out, Distributed Servers to Blame!

These exploding distributed server numbers caused the huge IT operating cost increases above *(now absorbing excessive shares of most IT budgets)*. Their **gross inefficiency, massive hardware and software capacity wastage, excessive support staffs needs**, and their wastefully profligate electrical power/cooling consumption/cost use, are now amply documented. This Appendix C analyzes how this nightmare occurred, quantifies the problems, and highlights the urgent need for rapid transition to drastically better IT infrastructures.

Distributed servers finally came to be seen in their true light in recent years. **They have proved to be an economic, environmental, and manageability, catastrophe** for tens of thousands of IT users who had deployed scores, hundreds, and *(in larger enterprises)* often many thousands, of them over the last dozen years.

UNIX Vendors, Then "Wintel Cartel" $Bs, Hyped Distributed Computing

Users were first lured to deploy scale-out, distributed servers by the UNIX server champions *(Sun, HP, IBM, SGI, etc.)*, who began the "client/server" distributed computing push. Next came $10Bs of vendor marketing from the "Wintel Cartel" *(Microsoft – Intel x86 platforms)* with their system vendor *(HP, Dell, FSC, IBM, etc.)* and x86 software *(Oracle, SAP, Microsoft, BEA, IBM, etc.)* accomplices. Both group's distributed systems were claimed to enable the then-popular, new client/server workloads. Both also claimed "lower-cost platform" edge over proprietary competitors. These persistent, but false, claims were comprehensively trashed only much later by myriad 2000-decade analyst TCO studies that later showed their disproportionately high "outside-of-the-box" costs.

These persistent, but false, claims were comprehensively trashed only much later by myriad 2000-decade analyst TCO studies that later showed their disproportionately high "outside-of-the-box" costs.

With average server utilization of 5-10%, one costly FTE/support head supporting just 20-25 distributed servers, and with their costly middleware/applications software equally underused, distributed scale-out *(UNIX & x86 volume server)* computing wreaked the above disastrous impact on overall IT budgets.

Distributed server proliferation **jammed today's data centers/server rooms full**, with no space for new IT equipment. Many now/will soon require $1M-$10Ms-worth of expansion investment. Additional electrical power is now often unobtainable in many locales today, even if it were affordable. CIOs also urgently need to free-up the armies of support staff distributed server estates soak up. They desperately need to re-deploy these precious staff resources onto more productive new business-supportive initiatives, such as major process improvements, and SOA application deployment.

Global warming concerns have rightly forced enterprises to strive much harder for energy efficiency, to attain more carbon-neutral status, and thus to play their part in saving our earth. IT must be major contributor to corporate energy use reductions. IT already uses 2% of global energy, similar consumption to the far more visible airline industry, and a level still rapidly increasing. So getting rid of sprawling, inefficient, distributed IT infrastructure has now also become a corporate strategic imperative for good ecological governance and green compliance, as well as for the major cost-saving, TCO and ROI reasons.

Thousands of IT users today are thus now seeking to escape from their current distributed-x86/x64 & UNIX server nightmares, to:

- **Simplify, consolidate, and optimize their IT infrastructure** *(server and storage)* to eliminate such massive wastage.
- **Use system virtualization and automation on superior platforms**, to improve the flexibility, responsiveness, manageability, QoS, and the DR/BC coverage, of their prime IT infrastructures.
- **Sharply reduce the staffing, and all the other costs** that are wasted supporting current, distributed scale-out IT environments.

Leaving distributed-x86/x64 & UNIX estates unchanged is no longer an option, because their staggering costs and inefficiencies are today rightly seen as intolerable from both the business and ecological viewpoints.

How Today's Scale-out Distributed Computing Disaster Began

From the early/mid-1990s, the scale-out, distributed, client/server computing model emerged, and enjoyed rapid growth. It deployed multiple, high-volume, low-cost, standardized computing servers, **one for each separate workload** *(rather than sharing IT resources for multiple application across larger SMP systems)*. First out were low-end, high-volume, **RISC-UNIX SMP servers**. These basic systems grew out of UNIX workstations at vendors like Sun & HP, initially for technical computing. *(Larger brethren of these small RISC-UNIX servers later scaled-up, and killed off other proprietary mini-computers by the late 1990s.)*

In parallel, a new breed of low-price, high-volume, **Intel x86 processor-powered, PC-technology-based servers** also emerged. First pioneered with Novell's NetWare OS for workgroup computing *(file & print)*, these gained real traction with Microsoft's Windows NT, and were promoted hard by newer PC server vendors, notably Compaq and Dell, who specialized in such x86 servers.

Vendors of both server types promised lower hardware price though standardization and higher sales volumes. They promised rapid technology advances and performance gains by using volume/standard MPUs with faster technology cycles, and touted price/performance advantages over then-competing proprietary servers. UNIX advocates particularly stressed their "open systems", and RISC CPU performance, claims. The x86 champions touted their lower-cost **"Wintel PC market-type economics"**, as well as promoting Microsoft OS and middleware functionality/cost benefits.

What both types of vendors consistently failed to disclose to, or carefully hid from, their prospective customers were the extraordinarily high "outside-of-the-box" costs of deploying and running such distributed scale-out servers on a large scale. These included the mostly wasted server capacity, ditto for all the software licenses, the heavy support-staffing compliments needed, their poor QoS and reliability, and weak security, etc.

Client/Server Boom Saw Scale-out Servers Proliferate Fast

Millions of scale-out distributed servers joined traditional mainframes, proprietary minicomputers, and PC populations that enterprises exclusively used for IT until then, from the mid-1990s onwards. These servers were also well sized/priced for Small-to Medium-sized Businesses *(SMBs)*, who adopted them widely, thus broadening the server market. They supported new applications and workloads, on then-much-vaunted client/server model. These included workgroup computing, e-mail and collaboration, new generations of ERP and CRM applications, and the explosive rise of Web-based computing and e-business. Over the IT boom years up to end-Year 2000, over **10M net additional scale-out distributed servers** were deployed, driving the global server installed base **up from c. 6M *(1996)* to 15M+ *(2000)*, an unprecedented 250% server base increase** *(see Figure 2 on page 10)*. In this breakneck period of business expansion and system investment, few IT users or analysts challenged the wisdom, or the huge real cost/wastage downsides, of this growing server proliferation.

C/S Dedicated Server Approach Multiplied Server Proliferation

This distributed client/server-computing model was based on a **"dedicated server"** approach. Every business application/workload was deployed on its own **dedicated, physical server**, or set of servers. The model also deployed business applications over multi-tiers *(two- or three-tier)* server architecture, with **separate application or database server tiers**, adding further server proliferation over Tiers 2 & 3. Advocates claimed dedicated server deployment gave better performance without resource contention, and at "relatively low cost". However, major enterprise applications deployed on this technology always also needed:

- Not only the **full production system set of servers above**, but also:
  - **Clustered or replicated database and/or application servers** *(plus software, networking)* were essential for most production systems, because the poor reliability/availability of the basic servers delivered much too low a QoS for such applications without. it

- A separate set of **dedicated application development servers** *(plus software, storage and networking)* was also needed to develop the applications, during the design/development or modification/enhancement phases over an application's lifecycle.
- Another separate set of **system testing servers** *(plus software, storage and networking)*, were also required for the integration, load, and stress testing, phases of the application lifecycle.

Enterprise application deployment on distributed servers often thus needed three or more-fold sets of hardware, most grossly under-used for much of the time, needing a large support staff, and occupying extensive scarce and costly data center floorspace. The reality was that the dedicated model **sold far more servers, software and storage for their vendors than were ever really needed**. Their operating systems were unable to provide the secure isolation between applications needed for sharing resources over application. So these quite primitive servers could not offer the more efficient, shared-resource usage modes found in more sophisticated, advanced server platforms.

Exploding Systems Management & Power Costs Resulted

In Section 2 Figure 2 and comments, we revealed the soaring global server population, IT support costs, and IT electrical power/cooling costs, over the 1996 to 2010 timespan. This chart, and our analysis there, graphically illustrated the huge scale, and out-of-control growth of, these IT operating costs, directly attributable to scale-out distributed computing proliferation.

x86/x64 Server Nightmare – Still Getting Worse!

Despite this appalling record of waste/inefficiency, x86/x64 servers continue to sell in extraordinarily large numbers, as Figure C1 below shows. The table shows recent year and quarter IDC Server Tracker market data, with the percentage shares now held by x86/x64 technology servers.

x86 – Dominant Share of Worldwide Server Market – Revenue $B and Server Units						
	2004	2005	2006	2007 Q1	2007 Q2	2007 Q3
Worldwide Server Market – All Server Types						
All Server Revenue $B	$49.0B	$51.3B	$52.3B	$12.4B	$13.1B	$13.1B
YOY Revenue Growth %	+6.2%	+4.4%	+2.0%	+4.9%	+6.3%	0.5%
Server Units M	6.3M	7.0M	7.5M	1.93M	1.94M	2.04M
Units % YOY Growth	+19.3%	+11.6%	+5.9%	+4.6%	+6.1%	6.0%
Worldwide x86 Server Market						
x86 Server Revenue $B	$21.0B	$24.5B	$25.8B	$6.6B	$6.9B	$7.2B
x86 YOY Revenue Growth %	16.7%	+5.22%	+5.2%	+8.7%	+15.5%	+9.1%
x86 Server Units M	5.5M	6.42M	6.9M	1.80M	1.8M	1.9M
x86 Units % YOY Growth	21.2%	+16.7%	+7.4%	+6.5%	+7.8%	+7.8%
x86 Shares of WW Server Market						
x86 % of $B Server Revenue	42.9%	47.8%	49.3%	53.2%	52.7%	55.0%
x86 % of M Server Units	87%	91.7%	92.0%	93.3%	93.3%	93.1%
IDC Server Tracker Data, *= Software Strategies Estimate						

Figure C1: x86 – Major Share of Worldwide Server Market – Revenue and Server Units

The global server market hit $52.3B of revenue, and saw **7.5M server units sold**, for 2006. It also posted healthy c. 5-6% average quarterly growth in revenue and units *(Q1, Q2 & Q3 data)* in 2007. x86 servers already accounted for 53% of server market revenues for 1H 2007, continuing their share climb from 49.3% *(2006)* and 42.9% *(2004)*. Of the total 7.5M server units shipped in 2006, **6.9M *(92%)* were x86 servers**, illustrating their dominance of new server units of sold.

We further review the dynamics that shaped, and created the costly catastrophe of today's scale-out, distributed computing environments in Appendix E.

Shocking Distributed Computing Costs Clarified From 2002

The business recession after Year 2000 *(and the 2001 dot.com bust)*, focused firms on serious IT cost reductions. This forced closer attention to now complex IT infrastructures created by uncontrolled distributed systems deployment. In 2002-2003, analysts first fully quantified, and published accurate total cost studies of, distributed computing environments. These studies quickly uncovered and quantified the deep issues of these platforms.

These main issues were previously summarized in our Section 2, Figure 3 chart on page 11, based on our fuller analysis below:

- **Vastly Underused Servers:** Scale-out distributed servers, most supporting a single application, usually run at 5-10% utilization, or 95/90% wasted capacity, over the full 7*24*365 year. This is a huge waste of valuable server capacity, but also of the costly staff resources that are needed to manage/support these near-empty servers. It also brings high software license costs for little-used software, wastefully consumes large areas of precious data center floor space and large amounts of electrical power/cooling. Bad enough already, the usual replication of distributed servers needed for production, development, test, and DR/BC, multiplies the production server set several-fold, driving down average utilization to these commonplace, feebly low levels.
- **Equally Underused, Wasteful, and Costly Software:** Most distributed servers run database, middleware, and/or application software licenses *(often costing more than the server hardware)* that are equally 90-95% wasted, with such low host server hardware. Utilization. This massively wastes the high initial license fees, and annually recurrent subscription and support charges on such software *(average c. 15-20% of OTL charges per year)*.
- **Systems Management Staffing/Support Costs:** One FTE support staffer is usually needed per 20-25 distributed scale-out servers, to keep such estates running. In Western Europe or the USA, average budgetary costs/FTE support person are at least $70,000 p.a., and often $100,000 p.a. burdened. Annual people support costs per 1,000 distributed servers thus average $3.5M+ in such markets; a stunning amount.
- **Similarly Wasteful Storage Pools:** Most enterprise users face equally pressing problems of poor utilization *(often way under 40%)*, excessive storage diversity, and high support staffing/management costs, with the diverse storage estates that evolved from their distributed servers deployments, and often later extended to multiple SANs. Many also added pools of NAS storage later. In such complex, often multi-vendor, storage environments, many management tools and different support skills, were needed. Enterprise-wide storage management became extremely costly and complex, and gaining access to all enterprise information often remained impossible. With soaring data volumes, this situation rapidly worsened over this decade, and cannot now be permitted to continue.
- **Extremely costly networking/connectivity:** The usual array of distributed servers supporting multiple tier enterprise application and database serving require extensive inter-server, server-storage, and server-network connections, with numerous external switches, routers, cables and/or optical links. This networking and server connection complexity is far higher with distributed platforms that System z *(where most of the networking is internal)*, and the networking/connectivity cost is often 3-4 times higher that with System z.
- **Poor Reliability, QoS:** Scale-out distributed systems, basic in design/technology, were/are extremely weak in the RAS/QoS. Only replication and clustering further increases their resilience, but adds cost and complexity. Their basic designs prevented the reliable, end-to-end, high QoS demanded of enterprise-grade IT services, unless clustering was added at high cost. And yet, the pervasive expansion of enterprise systems out to the Web now demands much higher service-availability levels, more consistent response times, and better application performance, over much more widely-varying demands, which these primitive platforms often cannot meet.
- **Non-existent or Unaffordable DR/BC:** With complex, scale-out distributed IT infrastructures, it was usually impossible, or unaffordable, *(and often both)* to provide robust DC/BC of the type that protects many mainframe-based enterprise systems. *(With automated storage replication, and rapid workload transfer to a hot-standby second site without loss of transactions.)* The multiplicity of distributed servers involved, the often-fragmented storage associated with them, as well as the limited hardware, OS and middleware technology of such systems, precluded DR/BC backup for their applications. And yet today's much tougher regulatory compliance/governance regimes now demands robust DR/BC, and longer-term data retention ILM services, in more industries, and across more major applications.
- **Short-life, High-depreciation, Rapid-obsolescence Hardware:** Scale-out distributed servers are built/engineered for short average lives of about three years, with failures increasing rapidly beyond their design lifetime. Rapid rates of advance in volume MPU, memory, and I/O also render them obsolete in performance, price/performance, etc., rapidly. They therefore incur much higher capital-value depreciation/TCO rates than longer-lived, larger SMP systems *(typically run for up to 5-8 years)*. From the "green IT" perspective, it is environmentally unsound and energy/carbon inefficient for enterprises to be buying, using, and throwing away distributed servers *(despite vendor recycling programs)* at so young an age, when better alternatives are available. With over 4.5M servers currently being de-commissioned per year, this is now a major ecological, as well as a prime economic, consideration.

With over 4.5M servers currently being de-commissioned per year, this is now a major ecological, as well as a prime economic, consideration.

- **Data Center Floor Space Constraints & Costs:** The flood of distributed computing equipment deployed over the last decade has literally filled many data centers. Many users pulled back earlier, user-site-deployed, distributed kit into their data centers, to manage them better in "server farms". Whilst the density of such systems improved *(with more capacity packed into smaller form-factors)*, the number of server, storage and network units housed in enterprise data centers soared. This pressured enterprises with data center space constraints, such as those in high-real-estate-cost locations *(such as major city centers or financial districts)*. From the late 1990s, this drove adoption of rack-dense packaging for server/storage/networking equipment, to increase density and to maximize floorspace usage, and then to the more recent, meteoric rise of the still-denser blade server format since 2002. However, these extreme density solutions then often hit floor loading, and especially power and cooling problems/constraints, discussed below. With investment needed to build a new, top-grade, Category 4 data center measured in the $10M+ range, avoiding these massive new capital spends for as long as possible is imperative.

- **Mushrooming Network Complexity:** User's diverse distributed server, operating, and storage estates, migrations from proprietary to Ethernet-based open networking, and now to the Internet, also caused high complexity in enterprise communications, storage networks, and server area networks. For these, users deployed a profusion of routers, switches, and appliances, copper and Fiber Channel optical wiring, and the multiple management tools/skills needed, to operate, monitor, and manage these over-complex networking fabrics. These further raised high staff costs, floorspace, power consumption, and cooling issues mentioned previously for servers and storage gear. They also brought both lower networking reliability, and performance/latency issues, to the complex, multi-tier applications run over these network fabrics.

- **Troubleshooting Extremely Difficult/Costly:** Now-widespread, complex, multi-tier distributed computing enterprise applications *(especially where run over multiple hardware and OS platforms)* were found far more difficult, time-consuming and costly to troubleshoot, repair, and to optimize performance for, than integrated single-platform environments. With them, much time and energy was wasted tracking down where the source of faults or performance degradation lay, with finger-pointing/blame between the several platform groups involved common.

- **Slow Provisioning/Inflexibility:** In scale-out distributed computing, every new business workload/application *(or sudden demand increase)* requires the provisioning, installation, and commissioning of a new physical server *(or set of servers)*, plus all associated storage, networking gear, and switches, etc. Next, their software stack must be assembled and installed onto the new hardware. In most larger enterprises, this provisioning process, including gaining financial approval, ordering and procuring the new equipment, its installation and commissioning in the data center, and to obtain and install of all the software needed, usually takes many weeks, often several months. This extremely labor-intensive, time-consuming, provisioning process makes it impossible for enterprises to respond rapidly enough to urgent new business needs and requirements, or to the large/rapid upswings in demand that are common with today's Web-based, customer- or partner-enabled enterprise systems.

All those 90-95% underutilized distributed servers are burning vast and increasing (as clock speeds crank up) amounts of electrical power (which all also needs costly cooling)...

- **Soaring Electrical Power Demands/Costs:** Proliferating distributed servers plus related equipment, with increasing processor chip frequencies, and with ever-smaller/denser packaging, escalated electrical power consumption and its costs, to become the massive issues they are today. Electrical power thus became a major operating cost for all IT equipment *(servers, storage, network appliances/routers/switches, etc.)*.

  Today, the IT industry rightly focuses on **processing capability per watt of electrical power**, via "System-On-A-Chip" integration, dual and multi-core designs, etc. All those 90-95% underutilized distributed servers are burning vast and increasing *(as clock speeds crank up)* amounts of electrical power *(which all also needs costly cooling)*, whilst filling up scarce and costly data-center floorspace, **for no benefit or value whatsoever to the business**. This has become an intolerable situation from a "green corporation"/green IT" and financial perspective.

- **System/Data Center Cooling/Thermal Loads:** The electrical power demand growth above, and rapidly increasing density of server/storage/networking gear more tightly packed in data center racks, also bought massive thermal load/cooling issues and costs. Effectively cooling more powerful chips within individual rack system units, the complete racks themselves, and the data center as a whole, all became critical issues. Heat/temperature is the enemy of electronic equipment reliability, and advanced thermal engineering/cooling technologies/skills are essential to achieve stable and reliable cooling, and yet are amongst the scarcest skills and the most difficult issues for vendors and users to master. Customers also need new skills, tools, services, and vendor help, to improve and optimize their data center machine-room areas, under-floor cable areas, and plenums, etc., for optimum cooling.

- **Heterogeneous Operating Environments Add to Costs, Complexity:** New classes of applications now deemed mission-critical to the enterprise were widely deployed, each using different operating environments, over the last decade. Enterprises found themselves running mainframe OS environments such as z/OS and z/VM, different versions of Microsoft Windows, often several flavors of UNIX *(such as AIX5L, Solaris, and/or HP-UX)*, continuing proprietary OSs *(such as i5/OS, OS2200, and MCP, etc.)*, and more recently, surging Linux deployments. Each different hardware platform/OS combination used in the business added complexity and required dedicated staffing/skills for support, increasing costs. Only the growth/spread of the open-standard Linux OS platform is providing cross-hardware platform commonality and major relief to this OS diversity challenge

Our Analysis – Distributed Servers – UNIX & x86/x64 – IT System's "Crack-cocaine"

On a "rational analyst" view today, it may seem hard to understand why such an inefficient, wasteful, and costly computing model enjoyed such success for so long, given its severe downsides analyzed above *(and in Appendices D & E)*. Rational observers would expect such a wasteful and damaging approach to have been "rumbled" by astute IT users far earlier, and/or even to have seen it banned completely on **"bad for business health and welfare"** grounds by now!

On a "rational analyst" view today, it may seem hard to understand why such an inefficient, wasteful, and costly computing model enjoyed such success for so long, given its severe downsides...

But such rational views ignore the **highly addictive properties** of distributed x86/x64 & UNIX platforms, especially to their younger-generation supporters. They also discount the extraordinary power, wealth, and immense market influence, of the "industry cartel" providing the x86/x64 platform's "base-ingredients", as well as the success of their "street distributors" in the IT marketplace over the last decade. The rich, powerful companies promoting these platforms have profited handsomely from user's addictions.

"Colombian cartel" equivalents, making the "base-ingredients" for x86/x64 servers, were Microsoft and Intel, shipping their dominant OS/middleware, and server CPUs respectively. Each made $10Bs in near-monopoly profits from this trade *(plus the adjacent PC space)* in the last decade. Only Linux, the rising open-source OS star, and AMD's winning server processors, limited these Wintel monopolists, providing them with rare but real competition, and some genuine user choices.

"Pushers", hooking enterprise IT users on the "addictive drug", were x86/x64 system vendors Compaq/HP, Dell, IBM, FSC, NEC, and lately Sun, etc. These were abetted by enterprise middleware software vendors "pushing" database, application server, integration, etc., foundations software for x86/x64 platforms, notably Microsoft, IBM, Oracle®, and BEA Systems, etc. Both these groups, plus other ecosystem players like enterprise application ISVs, and SI implementers, all profited vastly from customer "x86/x64 addiction" for more than a decade.

A heroically profitable, "crack-cocaine"-like business for the "cartel and pushers", but hundreds of $B wasted by the enterprise-IT-user community in the last decade!

A heroically profitable, "crack-cocaine"-like business for the "cartel and pushers", but hundreds of $B wasted by the enterprise-IT-user community in the last decade!

Self-evidently, users cannot rely on major vendors *(who are dependent on highly-profitably x86/x64 hardware/software/services sales today)* for unbiased, trustworthy help and advice on changing their situation, because the vested interests in selling "more of the same" are too strong. It is unrealistic to expect such firms to help IT user customers to sharply purge/eliminate grossly wasteful, excessively costly, and environmentally extravagant, distributed computing from their IT infrastructure. The main exception is IBM, who sells all types of server platforms, has more experience of IT optimization/consolidation than any other company, and who has just completely re-organized its global STG system sales organization along a unique "customer's best interests first" structure. IBM will now actively help, advise, and support customers to rapidly simplify and optimize their whole IT infrastructure, using the best mix of technologies available, both IBM and others, offering a raft of services to speed and support this process.

Appendix D: Widespread IT Infrastructure Optimization/Consolidation Now Imperative

Distributed, Scale-out Computing Must be Eliminated

We previously detailed, the bleak yet accurate picture of how over-complex, wasteful and inefficient IT infrastructures, from a decade-plus scale-out distributed server *(x86/x64 and UNIX volume server-based)* proliferation, became so widespread. Today, CIOs/IT managers and leading commentators, now recognize this stark reality. Our collective, dozen-year-long splurge of installing tens of millions of distributed servers proved **an economic, manageability, and environmental/ecological, near-disaster**. "Near-disaster" is a strong term. Is this comment justifiable? We say clearly yes!

A near-disaster in their **appalling hardware utilization** *(average 5-10%)*, **sky-high operational costs** *(particularly staffing, and software license, costs – both often poorly measured)*, their **soaring electrical power demands**, associated **high cooling costs**, and **sprawling data-center floor-space needs**. These costs multiplied explosively in recent years, and many data centers are now full to bursting-point.

These systems are the electricity "Gas Guzzlers" of the IT world!!

Strong public/political global warming concern today adds compelling governance pressure on all enterprises to become "greener". IT accounts for 2%+ of global electricity consumption already, so demands for IT to slash its share of **corporate carbon footprints** have become compelling. These strong social and political governance pressures now absolutely demand big improvements in wasteful, distributed IT infrastructures. These systems are the electricity "Gas Guzzlers" of the IT world!!

Rapid replacement of such economically and ecologically unsatisfactory, distributed scale-out x86/x64 & UNIX IT infrastructures is thus now a top business and IT imperative worldwide. Thousands of enterprises now seek superior, better-optimized future IT infrastructures, and are desperate to eliminate the heavy problems cited in Section 2 and Appendix C above.

The unsustainable IT operations, management, and support cost burdens of distributed computing has also stopped enterprises from implementing newer business solutions now needed to compete better.

The unsustainable IT operations, management, and support cost burdens of distributed computing has also stopped enterprises from implementing newer business solutions now needed to compete better. Too much of their IT budgets today remain wastefully absorbed supporting/managing current inefficient distributed systems infrastructures. Sweeping reductions in unjustifiable distributed system costs are now the main source from which CIOs can release precious resources *(money, staff, space, power, and cooling)* for new IT system developments.

Complex, distributed IT environments are also far harder to integrate, are difficult to make reliable and are thus vulnerable to disruption, and often lack rigorous security. These poor QoS attributes severely constrain users from developing/deploying newer, extra-enterprise collaborative applications their businesses, partners and customers need, if using such platforms.

IT Infrastructure Optimization/Consolidation 1997 to Date, Earlier Momentum

Improving IT infrastructures is not an all-new priority: there have been continuing, accelerating efforts on it for the last decade. What were the earlier waves of infrastructure optimization/consolidation , and what are the next priorities?

Since the late 1990s, we saw waves of **IT infrastructure optimization**, and **server and storage consolidation**, widely implemented. First of these were earlier Data Center and "large-system image consolidations" starting in the mid-later 1990s. Early this decade, users began rationalizing sprawling distributed server and storage estates that had grown so unwieldy. The main waves were:

- **Network Standardization & Consolidation.**
- **Data Center Consolidations.**
- **Large Systems/Server Consolidations.**
- **Distributed Servers to Server Farm Physical Consolidations.**
- **Distributed Server Consolidation/Virtualization Ramps Up.**
- **Enterprise Storage Consolidation.**
- **Enterprise Application Instance Consolidation.**

The approximate adoption timelines of these important, interrelated, types of IT infrastructure optimization and consolidation are shown in Figure D1, with selected, relevant technology landmarks.

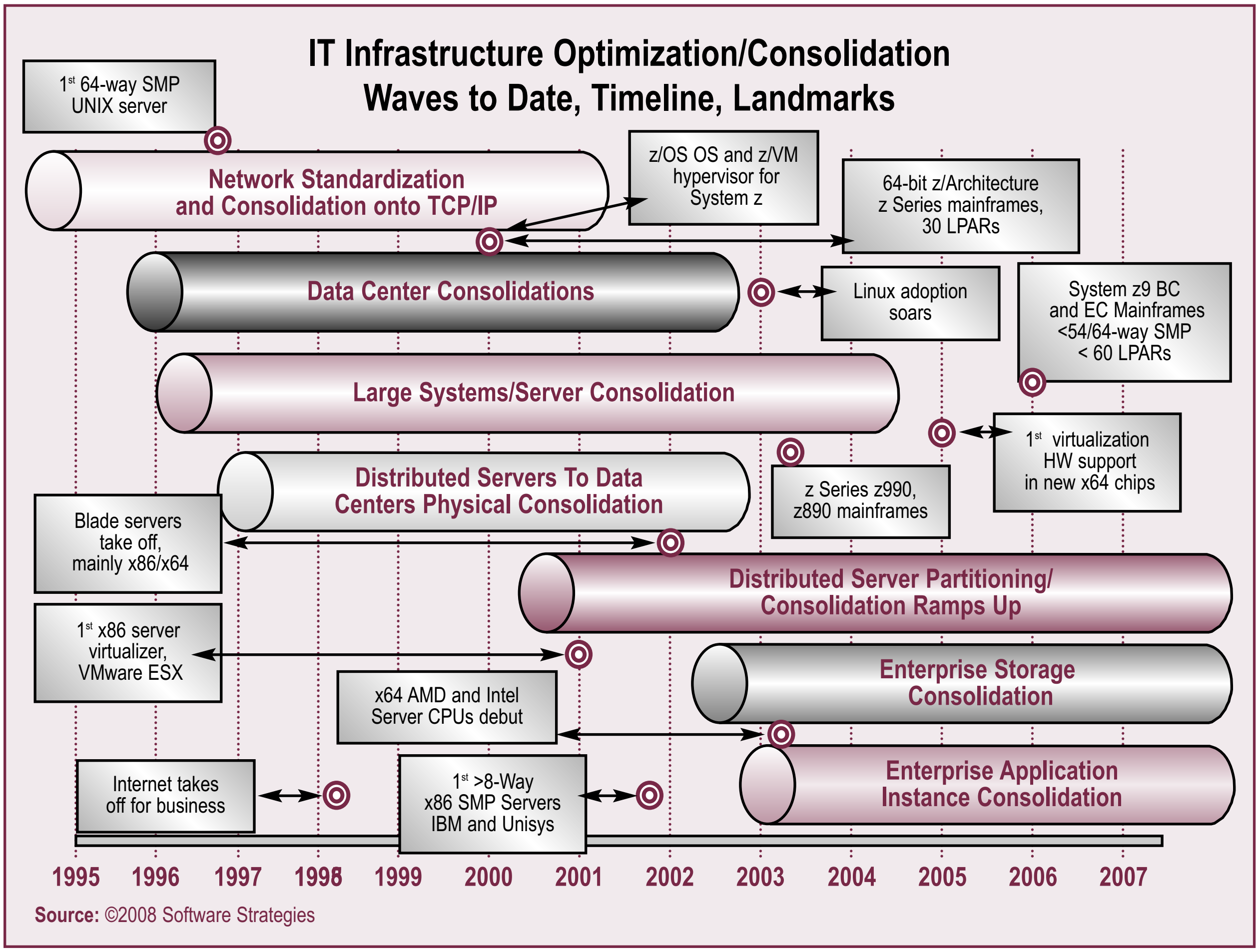


Figure D1: IT Infrastructure Optimization/Consolidation Waves to Date, Timeline, Landmarks

We overview each of these briefly below:

- **IT Network Standardization & Consolidation:** The 1990s saw users move away from proprietary networking architectures *(IBM SNA, DECNet, and IPX, etc.)* used till then. Enterprises consolidated on IP-based networking technology/TCP/IP protocol. These emerged out of Internet developments, and were widely adopted by the UNIX/open systems movement. The explosive growth in business Web use *(late 1990s)* hastened the consolidation of enterprise networks onto TCP/IP. The decade also saw migration from earlier private-data-network-centric approaches to the mixes of private, and public Internet, data networking that are common solutions today.

- **Data Center Consolidations:** These began in the late 1990s. Telecommunications costs first fell low enough, and bandwidth grew enough, by then so one data center could serve a big region, multiple countries, or a continent. The massive scaling-up of high-end systems *(both mainframes and top-end UNIX)* of the past dozen years was the other enabler. *(See next point below.)*

  IBM itself, running the world's largest and most sophisticated data center operations, recently reported that it had now consolidated down from **the 155 separate data centers it operated worldwide in 1999, down to just 7 in late 2007**. In so doing, IBM achieved huge savings it cited at several $B over the period. The 7 IBM "super sites" today provide it a massive 8,000,000 square feet+ of data center space. These support all IBM's internal, and all its customer-outsourcing, processing needs from these strategically-located, major s, including New York, Connecticut, Colorado, the United Kingdom, Japan and Australia.

- **Large-system Consolidations:** Top-end commercial systems *(mainframe, UNIX, and other proprietary)* soared from a typical maximum 10-Way SMP/<450 MIPS capacity in 1997, up to today's top 64-Way/128-Way/17,800 MIPS top-end systems. These, with mainframe-pioneered partitioning/virtualization, now let many large workloads be centralized onto on larger 2007 systems *(that in 1995 would have needed many top-end systems)*. Sophisticated large system clustering *(such as the powerful IBM mainframe Parallel Sysplex, and later enterprise UNIX clustering schemes for AIX, Solaris, and HP-UX)* enabled still-more-powerful and resilient, multiple, large system clusters. These encouraged data center consolidation, and helped lift DR/BC capabilities. Consequently, we saw much large-systems consolidation beginning initially with mainframes from the mid/later 1990s. Five years after, a similar process of UNIX large system consolidation also emerged, and continues today.

- **Distributed Servers to Server Farm Physical Consolidation, Move to Blades Accelerate:** Millions of distributed scale-out servers acquired from the mid-1990s were first often deployed out in the business, near their LOB sponsors/buyers and users. However, after Year 2000, corporate IT was often mandated to take over control of these *(by now sprawling, scattered)* distributed systems. IT often pulled these systems back into the data center, and into more rack-dense formats, where they could be more easily physically maintained, supported, and managed more productively. These "server farms" allowed standard server images and cloning to be used, bringing more productive support from their greater standardization, and from increased automation. This migration into data centers accelerated adoption of the rack-dense volume server format, so rack formats were ubiquitous by early in the 2000-decade. From 2002, blade servers were the next major volume servers development. Sales of blade servers have soared, and they are now projected to hit $10B revenue by 2010/11.

- **Distributed Server Partitioning/Consolidation Ramps Up:** From 2000, as mid-range/larger RISC-UNIX SMP servers first got basic partitioning, consolidations of smaller, distributed UNIX server workloads onto these larger SMPs grew. IBM System p became the clear virtualization leader in UNIX, and also won the UNIX market-share leader crown status from 2005.

  In the x86/x64 world, IBM *(X architecture scale-up x440, and X2, X3, X4 successor generations)* and Unisys *(ES7000 family)* pioneered mid-range to high-end x86/x64 scale-up SMP servers from 2001/2002. These larger *(<16-Way & <32-Way)* x86 systems competed with large, RISC-UNIX servers at lower hardware prices. More sophisticated than volume x86 servers, both vendors applied mainframe/large systems expertise to their higher-end, Intel-CPU-based servers. Initially, these only offered coarse-grained, fixed physical partitioning by 4-Way cell/module boundaries, which first allowed several larger x86 workloads to be supported on a single, large x86 server. They also later adopted VMware ESX Server for software logical partitioning *(see below)*.

  ...the nature of the x86 Instruction Set Architecture (ISA) made it extremely difficult to create any server virtualization solution for the x86 server platform at all.

  Early x86-processor-based volume hardware was originally designed to run only a single OS copy and a single application at a time. The original Intel x86/x64 processor architectures provided no hardware support for virtualization, and the nature of the x86 Instruction Set Architecture *(ISA)* made it extremely difficult to create any server virtualization solution for the x86 server platform at all. It was not until 2001 that a first breakthrough, x86 server virtualization product reached market, with release of VMware's ESX Server. By using a "trap, translate and emulate" workaround, VMware overcome the x86's lack of virtualization hardware support, and its unhelpful x86 ISA. VMware's ESX Server allowed a few **Virtual Machines** *(VMs)* to be run *(at some performance penalty)*, with their OS unchanged, on a suitable x86 server. We discussed this important advance in more detail in Section 3.

  Most of this White Paper focuses on this, the most widespread, important and urgently needed form of IT intra-structure optimization and consolidation.

- **Enterprise Storage Consolidation:** Parallel to, and driven by, complex and sprawling distributed server deployment, most enterprise users acquired equally fragmented, diverse storage environments over the past decade. These began with numerous pools of DAS local to specific distributed server platforms, later often extended to multiple SANs, with many later adding pools of NAS. Heavy data duplication, poor utilization of physical storage capacity *(often way under 40%)*, excessive storage diversity, and high support staffing/management costs resulted. With the soaring data volumes *(c +50% p.a.)* seen this decade, and ever-rising costs of managing such complex storage infrastructures, this situation become critical, and solutions mandatory. So many enterprise storage virtualization/consolidation projects have been implemented since 2003, when the viable storage virtualization solutions became available. Storage market revenue number one, and storage virtualization pioneer/leader, IBM seized a commanding lead with its widely successful IBM SAN Volume Controller *(SVC)* and related storage virtualization software. Over 10,000 IBM SVC engines have already been deployed to date.

- **Enterprise Application Instance Consolidation:** Most early 1990s continent-wide or global, deployments of enterprise applications *(whether ISV packages or custom in-house developed)* for international enterprises used the multiple-instance deployment model. In this, each region, country, or zone, adopting the application was supported with a local application “instance”, running on dedicated server hardware locally-placed. The common fat PC-client/server architecture of most such applications then had poor performance, high complexity, and limited the size of instance that could be run. The applications themselves, on relatively small and unsophisticated distributed platforms, were unable then to scale-up to, or support, multiple customer business entities upon a single software instance/server platform.

  In the current 2000-decade, most enterprise applications moved to Web browser thin-client, fat-server software architectures, which ran well over high bandwidth public or private IP/Internet networks at far lower cost. Rich Web 2.0 and/or Ajax UIs today give such applications the compelling user interfaces now expected, without the dire performance impacts, and costly desktops needed, for the old fat-PC-client model.

  These characteristics, plus much larger, partitioned SMP host servers *(mainframe & UNIX)*, and more scalable enterprise applications packages now able to support multiple business entities in one single software instance, have changed the optimum global deployment model. Now, running a single, global or continental, application instance, on a partitioned large SMP system, is often the best solution. Such centralized sites can support several or many country or local business operations, preserve full isolation and security between them, but consolidate all processing in a single, efficient, central system, which can then affordably be fully protected with a DR/BC secondary site.

Reliable IT Infrastructure Optimization Help/Advice – Who Not to Trust?

IT users wishing to sharply improve, their current, overly complex, distributed systems-plagued infrastructures should today **be especially careful about the vendors/advisors** they take strategic advice, help, and guidance from. Most of the IT industry *(x86/x64 server vendors, x86/x64 software ISV providers, SIs and SPs working on x86/x64 platforms, Microsoft, Intel, AMD, etc.)* are today highly revenue-dependent on their x86/x64 product and services revenues. Similar **“x86/x64 revenue dependency”** applies to many other parties in the wider x86/x64 ecosystem, from whom enterprise IT users might also seek guidance *(channels firms, resellers, consultants, advisors, media, analysts, commentators, promoters)*. Such firms previously advocated and/or sold the enterprise its present assortment of costly, wasteful, distributed x86/x64 systems, causing the serious IT infrastructure sprawl problems they face today.

Most of the IT industry... are today highly revenue-dependent on their x86/x64 product and services revenues.

Today in 2008, such players now offer new **“x86/x64-technology miracle solutions”**, now touted as the panaceas for solving the dire problems their own prior x86/x64 product sales created for customers! The four main such 2008 **“x86/x64-technology miracle solutions”** are:

- **Blade servers** *(in place of rack or tower volume server formats)*, for their higher density, shared and reduced-cost infrastructure and better integration, superior management, and other valuable efficiencies. *(IBM & HP are the clear leaders in blade servers.)*
- **Larger, partitionable scale-up x86/x64 server systems** that allow workload consolidation from several smaller x86/x64 systems onto a physically partitioned, larger x86/x64 SMP, and/or onto multiple VMs running on x86/x64virtualization software *(see below)*. *(IBM and Unisys are leading exponents.)*
- **x86/x64 server virtualization software** technology that enables a few VMs running x86/x64 workloads to be run relatively safely on a single, partitioned/virtualized server.
- **New-generation x64 CPU-powered servers** using new AMD and Intel server chips introduced over 2005/2007. These first included basic hardware virtualization support *(AMD–V (Pacifica) and Intel-VT (Vanderpool))* to improve x64-CPU- based server virtualization. This hardware support made it easier to create virtualizers for x64 platforms, reduced the substantial performance overhead previously seen in all-software x86/x64 virtualizers, and has spawned the many competing x64 virtualizers now in this market today. The advance also rendered all prior x86/x64 servers lacking such hardware virtualization support obsolete, requiring customers to physically replace all their pre-existing x86/x64 systems should they want/need efficient x64 virtualization on their such platforms.

These four x86/x64 technologies are unquestionably important to most platform users. However, they are far from being the only solution to the distributed x86/x64 challenges of Section 2 and Appendix C. We examined, in depth, the leading, arguably superior approach – that of IBM System z mainframe extreme virtualization with Linux-on-z/VM – in Section 5 onwards.

Heavily revenue-dependent x86/x64 platform vendors cannot safely be trusted to provide objective, reliable, and trustworthy guidance on how users should best escape from their existing, inefficient and costly distributed x86/x64 infrastructures. Specifically, those most deeply dependent on this marketplace, and who fall within our **"customers beware"** advisory, are HP, Dell, Fujitsu-Siemens Computers, Unisys, and x86-newbie Sun Microsystems, as well as Microsoft, Intel, VMware/EMC, and AMD. These vendors/suppliers, together with their main ISV and SI x86/x64 cheerleaders *(who include Oracle as a main ISV example, and EDS and Accenture as leading SI examples)* are all deep in x86/x64 ecosystem *(cartel)* dependency.

Heavily revenue-dependent x86/x64 platform vendors cannot safely be trusted to provide objective, reliable, and trustworthy guidance...

To expect unbiased, un-self-interested advice on this from such firms is as unrealistic as asking a local crack-cocaine dealer for help or advice in kicking that addiction. These vendors have overwhelming vested interests in selling more of their "current stock of dope". Each now champions their blend of the above new x86/x64 elements, pressing enterprises to renew/replace "older x86/x64" estate investment by their "newest x86/x64 miracle cure" replacements.

...unrealistic as asking a local crack-cocaine dealer for help or advice in kicking that addiction.

Where customers follow this "pusher's path", the deep x86/x64 problems *(discussed in Section 2 and Appendix C)* remain, and the opportunity for more radical improvement is lost.

We suggest enterprises carefully evaluate, and skeptically review, overtures, proposals, and advice from such biased vendors on this vital topic, and deeply consider more attractive solutions to the distributed x86/x64 challenge.

New IBM-wide Organization, IT Infrastructure Optimization Initiative, to Help Customers

Most enterprise customers today face challenging distributed IT infrastructures urgently needing serious optimization and consolidation to eliminate the issues above. Who can such customers trust for large-scale, objective and professional, deeply-experienced, and technology-neutral, advice/guidance able to deliver IT infrastructure optimization plans/roadmaps tailored specifically to each customer's unique situation? One large company users clearly can so trust for just such help and advice with this vital mission is IBM.

One large company users clearly can so trust for just such help and advice with this vital mission is IBM.

IBM offers a full range of System x servers, extensive IBM software offerings, and wide-ranging IBM services, for the x86/x64 platform. It long held the equal second revenue share position in x86/x64 servers, behind only sector-leader HP. IBM's strongholds here are its advanced, leadership *(mainly)* x86/x64 blade server platform *(BladeCenter)*, and its innovative/unique, modular scale-up, mid-high-end x64 servers *(now on 2007's latest X4 architecture)*. These each play in important x86/x64 server market sub-segments, where IBM has generally been # 1 sub-segment leader in recent years. It also offers fiercely competitive, technology-differentiated, volume x86/x64 servers, together with superb cross-platform system management software *(IBM Systems Director)*. IBM thus offers the most advanced x86/x64 solutions available, plus many years of deep experience and expertise in optimizing, consolidating, and simplifying x86/x64 infrastructures for customers. It thus can offer the very best of x86/x64 technology itself.

However, IBM also offers its highly-successful System z mainframes, popular System i integrated business systems for SMBs, and its UNIX-market-leading System p servers. IBM also remains the dominant market-share leader in High Performance Computing *(HPC)* with its Blue Gene *(top-end)*, System p clusters *(midrange HPC)* and System x/BladeCenter x64/PowerPC clusters *(low-end HPC)*. The firm is also the largest global provider of all storage *(total disk and tape)* and external disk, and is a leading storage software provider, with its extensive IBM System Storage family. Finally, IBM is far the world's largest IT services and advisory firm, with over $40B in services revenue.

The company thus offers the broadest systems, storage, software, and services, portfolio in the industry, with deep experience on all these, and is clearly the best-qualified and experienced organization in the whole IT industry to offer large-scale help on this vital topic.

Hundreds of customers, over recent years, asked IBM to provide them with fuller advice and guidance on how best to optimize and consolidate their IT infrastructures, which platforms to best use, and where. IBM recently announced a new, customer-focused global systems sales organization, and a major, cross-IBM global initiative plus comprehensive services, designed to do just this. IBM made this major change to better offer definitive, authoritative help to enterprise customers worldwide on how to optimize and consolidate their current IT infrastructures. Via this new organization, IBM customers will be guided/helped to determine the best IT infrastructure solutions for their specific needs.

These optimum solutions will exploit the best mix of infrastructure optimization opportunities available from the whole IBM Systems, System Storage, and Software products, portfolios. Each of these are industry-leaders in their infrastructure optimization, virtualization, and consolidation capabilities. IBM's heavyweight help in selecting the best mix for their unique needs, and helping customer to implement the change rapidly, will thus be a powerful advantage.

This important change replaces IBM's previous, product-silo-focused Systems sales approach. Under that prior model, several IBM Systems brands *(System z, System p, System i, System x & BladeCenter, and Storage)* were often each offering their solutions to the same customer at the same time. This left the customer to assess and decide what mix of IBM systems and software *(and other vendors' offerings)* they should best adopt. Large numbers of customers wanted IBM to guide/advise them with an "IBM-recommended" best mix and solution set for their individual requirements.

The new, holistic, cross-IBM system platform approach evaluates/examines the customer's existing IT estate, business goals/directions, needs and issues, and defines an optimum mix of consolidation, virtualization, platforms, and technologies. Services-led implementation engagements then provide IBM's fullest help to actually move the customer's IT estate from its present state to the new, optimized configuration for their unique infrastructure needs/requirements. An IBM brand-neutral advice basis is promised, with extensive IBM implementation help ready in place to speed customer transitions.

The services, methods, special software tools, support, large-scale migration factories, etc., needed to plan, effect and deliver such large-scale, well-planned infrastructure migrations, typically in several phases, are now offered. These are the refined, coordinated and extended new end-products based upon thousands of earlier engagements conducted by IBM under several pre-runner server and storage consolidation programs that ran over the 2000 decade to date.

In an engagement, the joint IBM-customer team will work hand-in-glove to determine the best mix of IBM new platforms, technologies and solutions *(and sometimes other vendor offerings too)* to best fits the customer's overall future IT Infrastructure needs. This will start with careful, in-depth discovery analysis of their whole existing infrastructure, challenges and issues, workloads and applications, skills and resources, and future business needs.

For example, for a customer with 5,000 older distributed x86/x64 servers in their estate, the joint IBM-customer engagement might determine/recommend these should best be consolidated as follows:

- **3,000 2-way server workloads onto 1,500 new virtual Linux servers under z/VM** on 4 new System z9 mainframes *(375 VS/System z mainframe.)* These workloads to including Linux mid-tier application servers and medium-duty database applications running on Linux middleware servers.
- **1,000 4-way Windows-specific, heavy database and application servers** could best be moved onto **125 scale-up IBM X4 architecture System x**, x64-powered SMP servers, each physically partitioned *(8 new images per System x server)*.
- **1,000 1- and 2-Way light Linux server workloads** should migrate onto **20 BladeCenters, with 280 2-Way x64 blade servers**, some virtualized with VMware ESX where needed.
- A complementary **storage virtualization and consolidation project**, with a SAN extension, and improved management software.

IBM can also offer most/all of the middleware software needed, and can also *(IGF)* readily finance all the new hardware, software, and services required as a single deal.

From this major, cross-IBM, long-term effort, we expect to see a fast-growing number of other, large x86/x64 & UNIX servers to System z Linux-on-z/VM consolidation projects being won and deployed, similar to the real cases in Section 6.

Our Analysis

The distributed, scale-out x86/x64 platform *(the cause of many/most of the problems cited)* proved a highly-addictive "drug" to which many enterprise users are severely addicted, to great commercial benefit for their system and software vendors. Kicking this costly and debilitating IT habit, and replacing it with healthier platform solution mixes *(alternatives)* is a vital part of such IT infrastructure optimization/consolidation, in our considered view.

The distributed, scale-out x86/x64 platform... proved a highly-addictive "drug" to which many enterprise users are severely addicted...

Several earlier IT infrastructure optimization and consolidation approaches were already deployed from the mid-late 1990s, and we reviewed each of these above. Whilst these overlap, for the purposes of this White Paper, we focus mostly on **Distributed Server Consolidation/Virtualization** the largest and most widespread current need.

We highlighted the numerous IT vendor exponents of x86/x64 technology who were largely responsible for the distributed nightmare today faced by many enterprise IT users. These vendors collectively now loudly tout a quartet of new "x86/x64 technology panacea solutions" to the problems their earlier mass server sales created at enterprise customers. These culprits now have the effrontery to propose their long-suffering customers should "rip-and-replace" these "old x64/x64" platforms with these vendor's newest x86/x64 technology solutions.

In contrast, IBM offers a broad systems portfolio that today includes its leading-edge System x and BladeCenter x86/x64 platforms, as well as its other leading server families, its leadership System Storage family, and an extensive IBM software and services portfolio.

We suggest enterprise customers be extremely wary, and be toughly critical, of claims for these improved x86/x64 technologies from such self-interested, biased vendors. Users should always actively explore/evaluate alternative major solutions, such as extreme virtualization and consolidation of distributed workloads with Linux-on-z/VM on System z mainframes. *(Discussed in Sections 4 to 6.)*

In contrast, IBM offers a broad systems portfolio that today includes its leading-edge System x and BladeCenter x86/x64 platforms, as well as its other leading server families, its leadership System Storage family, and an extensive IBM software and services portfolio. It has now reorganized into a more customer-facing form, able to offer platform-neutral, "in-customer-best-interests" IT infrastructure optimization advice, design, and implementation services on a global basis and at a large scale. As long the most trusted vendor name in enterprise IT, IBM will be a natural choice to provide this deep, broad view, help and support to enterprises, to select and deploy the best mix of optimized IT infrastructure platforms, systems, virtualization, and software.

Appendix E: Distributed System Market Dynamics

Reviewing the evolution of scale-out, distributed systems in the last decade explains how their unsatisfactory current state arose, and deserved the additional discussion below here.

- **Business Unit/LOB Buying Widespread:** Central IT groups long controlled mainframe operations, but many lost part IT buying control when proprietary mini-computers and PCs arrived in the 1980s. These let business-unit/LOB executives buy application solutions, servers, and/or PCs themselves. Client/server computing *(from the mid-1990s)* saw more business units buy/deploy alone. Many did so without the rigor, discipline, processes, and standards, long enforced in well-run, central IT shops. Many ERP *(SAP, Oracle, and JDE, etc.)*, and CRM/SCM, enterprise applications were bought this way from the 1990s.These ran first on UNIX, but later on x86, platforms. This trend proliferated distributed servers in both large and smaller firms. Scattered LOB system ownership/location prevented accurate IT cost/usage analysis/accounting. After 2000, many enterprises asked central IT to bring LOB distributed systems sprawls back under professional IT control. IT usually moved these systems into their data centers, and rationalized the chaotic patchworks inherited from earlier LOB procurements. This swing back to more centrally-managed, physically-centralized, IT continues to accelerate today.
- **UNIX Early Volume Servers Lead Lost to x86:** From c. 1994, the UNIX camp *(workstation, server, software ecosystem – Sun, HP, IBM, SGI, FSC, and Bull, etc.)* became an "open-systems" force with market muscle and influence. The vendors pushed volume UNIX servers to wide deployment. Later, their servers scaled up, competing and winning against most proprietary SMP contemporaries. Low-end, high-volume RISC-UNIX server sales boomed up to 2001, with their full 64-bit power/performance ahead of x86 servers until c. 2003's x64 debut.

  But Intel's rapid new MPU life cycles and x64 technologies pushed x86 volume server performance up faster. "Wintel cartel volume economics" gave these boxes lower price-points. x86 servers first took, and later widened, their units volume lead. Low-end RISC-UNIX *(1-, 2-, and 4-Way)* server sales dwindled as x86/x64 server now filled many previously UNIX roles. Next came growing installed base migration from now aging, costly RISC-UNIX servers *(Sun Solaris and HP-UX- earlier UNIX leaders)* to new x86/x64 platforms, most running Linux. *(Often tagged as "the UNIX user's liberator!")*
- **2003 On -Low-end RISC-UNIX, IA-64, & x86 Servers Killed BY x64:** The AMD64 64-bit-extended Opteron MPU *(first seen in April 2003)*, then Intel's near-identical EM64T *(now known as Intel 64)* 64-bit extended Xeon server processors changed volume server markets. *(Duo termed "x64" here)*. Their 64-bit addressing, good performance, and full IA-32 software compatibility, finished off much of the volume UNIX server market. One popular move was for users to migrate off more costly, old UNIX hardware, onto these new, fast, low-cost x64-powered volume servers under Linux, a fairly easy migration.
- **Volume Server Form-factors Evolve – Tower to rack to blade servers:** Most early1990s on low-end, volume servers were of floor-standing tower format. Next came a major swing to rack-server formats users could pack higher for greater density in data center server farm racks, or in SMB server rooms. Volume rack servers *(UNIX & x86)* were dominating server unit sales from early in the 2000-decade: they remain a large segment today.

- **Blade Servers Fastest Growing Server Segment/format:** With superior density, better shared infrastructure, better economics, and superior manageability, blade server sales have soared. Blade revenues jumped from $158M *(2002)* to $2.11B *(2005)*, an 83.5% CAGR. Top market researchers forecast this will have tripled by 2009. In Q2 2007, blade servers won 6.7% of total server revenue, posting a striking 36.7% YOY growth. IBM led the blade server market with its extensive, sophisticated BladeCenter family; lately dropping just behind other blade-major HP's *(BladeSystem C Class line)*. HP has been aggressively cannibalizing its huge ProLiant *(tower and rack)* x86/x64 base, turning users over to BladeSystem, in recent months.

- **Multi-core MPUs Move x64 Systems Up-scale, Squeeze Mid-range SMPs:** For years, most volume x86 server sales were traditional, small 1-Way, 2-Way, and fewer 4-Way, -CPU SMP volume servers, powered by single-core Intel x86 *(usually Xeon)* or AMD MPUs. From 2005 on, dual-core AMD and Intel x64 MPUs saw similar 1/2/4-socket physical server designs suddenly morph into 2-Way, 4-Way, and 8-Way-core, 64-bit SMP systems. This lifted their capacity/performance into contention with other low to mid-range SMP servers *(E.g. UNIX, System i, other proprietary)*. 2007 saw top x64 processors moving up to 4-core MPU chips *(Intel Core Xeon 2*2-core units, AMD Opteron "Barcelona" true 4-cores on 1 die MPU)*. 1-, 2-, and 4-socket x64 servers sporting such quad-core MPUs now bring 4-Way, 8-Way, & 16-Way-core SMP systems. These retain similar physical formats to their dual or single-core forerunners, at competitive prices, and power enough to now compete with mid-range SMP servers *(typically 8-Way to 16-/20-/24-Way SMP in recent years)*. Traditional systems are thus now squeezed from below by such fast rising capacity, x64 quad-core-based systems.

- **The Role of Windows:** Majority OS choice on x86/x64 servers was long Microsoft Windows OS *(Windows NT, Windows 2000, Windows 2003)* for x86-only *(32-bit)* and x64 *(64-bit extended)*. Windows has also *(like Linux)* encroached deeply onto largest OS market segment UNIX this 2000 decade. Windows first overtook UNIX server revenues in 2005 *(at $17.7B and 34.5% of 2005 total new server revenue)*. Microsoft garnered several $B profit yearly from its Windows server monopoly franchise. It spent big on aggressive Windows server marketing, in lockstep with fellow Wintel cartel-partner Intel, but faced sharper real competition from Linux in recent years. Over its OS layer, Microsoft's Windows Server middleware stack *(SQL Server, Exchange, MSMQ, and BizTalk, etc.)* was the platform high-volume, low-cost option fighting enterprise middleware *(from vendors IBM, BEA Systems, and Oracle, etc.)* also extensively supporting the Windows platform.

  The huge x86 server market grew a large ecosystem: applications, middleware and tools software ISVs, SIs, BPs, and SPs. These helped Windows succeed, across the SMB segment *(where it is particularly strong)*, and the enterprise. Back a decade, when legacy "fat PC client/server" architectures were fashionable, Windows had the edge of commonality between PC client and server OSs. This decade saw a mass migration off cumbersome, poorly-scaling, and costly "fat-PC client/server" architectures. Most users migrated to a new generation of server-centric, Web-browser-thin-client enterprise applications, where Windows held no special advantage. However, it still held 38.2% of the total server market by revenue in Q2 2007 – and posted a healthy 18.7% growth, still riding the x86/x64 train.

- **Rapid Rise and Strategic Significance of Linux:** Linux saw fast server growth/adoption this decade to date. It now commands third place in the global server OS market. For 1H 2007, Linux held a 13.3% revenue share, with $3.4B in new server factory revenue, up a healthy 11.0% YOY. *(After # 1 OS Windows with a 38.4% share at $9.8B revenue, after # 2 UNIX with a 32.2% share at $8.22B revenue, and well ahead of # 4 z/OS with a 9.5% share at $2.2B revenue.)* Rapid open-source community development advances saw Linux rapidly scale-up, add enterprise capabilities, and to now seriously rival both Microsoft Windows, and proprietary UNIX versions, in many areas of capabilities, applications, and workloads, etc.. The two leading Linux server distributions aimed at serious business and enterprise use *(Novell SuSe Linux (SLES) and Red Hat Linux (RHEL))* enjoyed wide adoption and are respected, modern platforms today.

  Linux's license cost-free, open-source community development model, its excellent record on reliability/availability, first-class performance/exploitation of hardware resources, and the now-widespread availability of Linux skills at reasonable cost, made Linux the "lowest TCO" OS platform – for many different important workloads, and on numerous important hardware platforms. These include:

  - **Applications/Workloads:** High-Performance Computing *(HPC)*, Web serving and front-end infrastructure, application serving, enterprise middleware, database serving, and enterprise applications.

  - **Linux Hardware Platforms:** Runs on all server hardware platforms. Especially successful on all x86/x64 servers, including blade servers, IBM System z mainframes, on RISC servers *(powered by POWER, SPARC, and PA-RISC MPUs)*, IBM System i5 business systems, Itanium platforms, and all leading HPC systems and clusters.

  With such wide Linux popularity and growth, an excellent range of open source and proprietary software tools, middleware of every type, and enterprise applications, is now available on Linux. This portfolio now rivals that on leading UNIX versions *(with which it retains a close affinity)*. Linux thus now often lets software users escape lock-in on high cost, proprietary UNIX systems. Users can also increasingly standardize on Linux across all their server platforms, a major simplification in IT infrastructure and support.

Related Software Strategies Mainframe Research

Earlier, in-depth Software Strategies System z mainframe-focused White Papers/Reports include:

1. **"Managing System z Mainframe SOA Environments – Strong IBM z/OS SOA Software Advances Key."** Software Strategies White Paper, 1st Edition, January 2007, 66 p.p., 6 charts & tables. *(In-depth White Paper assessing and evaluating the System z9 mainframe as the preferred enterprise SOA applications platform from 2007. Particular focus on the "second-generation" SOA adopter issues of securing, managing, and virtualizing the enterprise SOA environments.)*
2. **"Top 15 Reasons Users Should Stay On/Upgrade/Move on to the IBM Mainframe."** Software Strategies White Paper. 1st Edition December 2006, 30 p.p., 6 charts & tables. *(Hard-hitting, outspoken short Paper highlighting the 15 Top Reasons users should stay on the System z mainframe platform, or migrate to it, from 2007.)*
3. **"New System z9 Mainframes Hit Mid-market, Refresh Top End – Powerful New SOA & Data-serving Software Delivers New Enterprise-wide Roles."** Software Strategies White Paper, 2nd Edition July 2006. 60 p.p., 21 charts & tables. *(In depth assessment of System z9 Business Class and Enterprise Class mainframes, HW/SW capabilities, TCO leadership, and 2006 SOA System z middleware software stack advances.)*
4. **"Information as a Service Unfolds – System z9 Mainframe/DB2 Premier Data & Information Server for SOA."** Software Strategies White Paper, February 2005, 74 p.p., 30 charts & tables. *(In-depth assessment of IBM's "Information as a Service (IAAS) strategy, supporting extended information management product portfolio, and the role of the System z9 mainframe as the enterprise SOA data serving and information delivery hub.)*
5. **"SOA Takes Off – New WebSphere SOA Foundation Extends IBM's Lead with New System z9 Mainframes as the Hub of the Enterprise."** Software Strategies White Paper, 2nd Edition, November 2005, 56 p.p., 20 charts and tables. *(In-depth assessment of the case for adopting SOA, evaluation of the IBM WebSphere SOA Foundation software, methods and services, and the role of the System z9/zSeries mainframe as the enterprise SOA hub.)*
6. **"Spectacular System z9 Mainframes Leap Ahead with Doubled Power, Enterprise Hub Roles – Virtualization, Security, Availability, SOA & Value Advances."** Software Strategies Enterprise Server Spotlight Report, September 2005, 72 p.p., 28 charts and tables. *(In-depth assessment of IBM System z9 109 mainframe platform hardware, OSs, middleware, storage and virtualization capability, first new system to emerge under IBM Systems' new "Systems Agenda" strategy for the emergent era of collaborative computing.)*
7. **"New Power-driven, High-end and Modular Enterprise Storage Systems – Game-changing Server Technologies/Advances Supercharge IBM's Storage Market Leadership Bid."** Software Strategies White Paper, 2nd Edition, November 2004, 42 p.p., 21 charts and tables. *(In-depth technology assessment of IBM's new DS6000 and DS8000 enterprise storage systems.)*

About Software Strategies

Software Strategies is a specialist analyst firm focused on e-Infrastructure platform strategies and issues. Specialist expertise on mainframes, servers, OSs and middleware has been a common thread. Several thousand Enterprise IT users have benefited from our authoritative events, presentations, conferences, newsletters, journals, and reports. Since 1997, we have hosted and/or spoken at numerous successful industry events, including our popular Focus Events. We have also worked closely with industry leaders, including: IBM; Microsoft; Intel; ICL; Unisys; CA; BMC; Stratus Computers; NetIQ, and many others.

Author

Ian Bramley
Managing Director – Software Strategies
ian.bramley@software-strategies.net

This new White Paper was researched/written by Ian Bramley, Managing Director of Software Strategies, in January 2008 *(2nd Edition published February 2008)*. The views expressed are those of Software Strategies, and are based on our proprietary research. Ian founded Software Strategies in 1997, and is an experienced enterprise infrastructure analyst, a keynote speaker at many industry events, and has published many reports and papers. He was previously Director of Enterprise Platforms at Butler Group, and Founder/Chairman of the Enterprise NT Management Forum from 1998 to end-2000. Previously, he held a variety of executive positions with several international software vendors over a 25-year industry career.