24x7 Availability Study for SAP R/3

Commissioned by:

Lisa Smith
Senior Business Area Manager
SAP Worldwide Deployment
IBM Global Services

Prepared By:

Laura Cooper
IBM/SAP Competency Center
Austin, Texas, USA

PREFACE

ACKNOWLEDGMENTS

Paul Beedham
Henry Bilsland
Paul Chamberlain
Jim Dilley
Terrance "Tar" Druggan
Brian Forney
Greg Greth
Ron Jensen
Mike Mesmer
Hugh Varilly
Tom Wilkinson
Lyn Weyrick
Jim Wright

REFERENCES

IBM SAP 24x7 Operations UK Service Delivery Discussion Document by Henry Bilsland
BC SAP High Availability
Implementing SAP R/3 in an OS/390 Environment Using AIX Application Servers Controlled Availability Version (ITSO Redbook, SG24-4945-00)
Database Administration Experiences: SAP R/3 on DB2 for OS/390 (ITSO Redbook, SG24-2078-00)
IBM DB2 and SAP R/3 Administration Guide (ITSO Redbook, SG24-4871-00)
SAP R/3 in Switchover Environments (SAP AG, 50020596)
SAP R/3 on DB2 for OS/390: Planning Guide SAP R/3 Release 3.1H (SC33-7961-00)

EXECUTIVE SUMMARY

In response to the growing requirements to re-engineer and continuously improve our business process, numerous units are moving to the SAP R/3 application. Several of the business projects have requested extended service hours to meet business requirements. Some of the projects have requested that the systems be available 24 hours a day, 7 days a week, 365 days a year. Unfortunately, this level of continuous availability is not possible with the current versions of SAP R/3. Some of the mandatory systems management tasks require system down time. With well-architected systems management procedures, operational practices, and systems design, the duration and frequency of scheduled down time can be reduced - but not eliminated.

In particular, SAP R/3 release upgrades, the addition and configuration of new functionality (such as new language support), system recovery, and the changing of specific SAP R/3 parameters require system down time. SAP has significantly reduced the down time required for upgrades, however, a minimum of one to two days of down time are still needed for a release upgrade. Adding additional language support requires an outage of several hours. Parameter changes require the R/3 system to be stopped and restarted. While this is only a short outage, it is required to activate any parameter changes. For the AIX platform, down time is also required for database reorganizations.

Backups should be run outside of peak hours to avoid performance problems. Data migration and batch jobs must be carefully scheduled and monitored to prevent performance problems. In addition, custom code must be reviewed and thoroughly tested prior to running in production. A poorly coded program or a batch job scheduled during peak hours can quickly cause severe performance problems.

The OS/390 solution eliminates the necessity for downtime for database reorganizations and reduces the impact of backups. It is also a more stable and reliable platform for the database server. For the OS/390 solution, the SAP R/3 Central Instance and Application servers must still be run on AIX. Only the database server runs on OS/390. If the Central Instance is down, the SAP R/3 system is not available. The critical services running on AIX can be protected from lengthy outages by HACMP. Down time for SAP R/3 upgrades, additional language support, and parameter changes is still required.

It is possible to reduce or eliminate down time for some tasks by setting up and using a hardware and software mirror of the production system, referred to in this document as the "mirror production" system. This can be an expensive option, but it can help eliminate down time for recovery testing and client copies. It can also reduce down time needed for upgrading SAP R/3 releases, testing HACMP, testing batch jobs and testing data migrations.

Some suggestions for enhancements to IBM hardware and software and to the SAP R/3 product to help improve availability are listed at the end of this document.

In summary, while SAP R/3 can be implemented as a highly available system, this does not include a truly continuous operation type environment. Some level of planned outages will be required, but such outages can be minimized through a systematic approach to architecting the production environment and its supporting processes and procedures.

INTRODUCTION

BACKGROUND

In order to decrease cost and cycle time and increase customer satisfaction, IBM has launched eleven Business Process Reengineering initiatives. Five of the reengineering initiatives are using the SAP integrated application suite to provide application support. There are currently 23 SAP R/3 projects within IBM. Nine of these projects are expected to be in production by the end of 1997. Several of these projects are worldwide. Several of the business projects have requested extended service hours to meet business requirements. Some of the projects have requested that the systems be available 24 hours a day, 7 days a week, 365 days a year which is currently unattainable. Maintenance, upgrades, recovery, and other systems management tasks require system down time. This study will address the requirement for continuous availability and describe how to minimize the time when production systems would be unavailable.

This study is intended for IBM internal use for systems included in the SAP Tower Services Offering. It may be expanded in the future to include IBM customers or to identify the areas in which IBM and SAP can work together to improve availability of the R/3 complex. Currently, this study is not intended to be distributed outside the internal IBM SAP Tower Systems.

The recommended SAP Tower Services System solution is an SP using a high node as the database server. For smaller systems, a J50 RS/6000 may be used in place of the SP high node. All production systems are currently (as of 9/97) running on the RS/6000 platform. At least two sites will convert to S/390 on the OS/390 platform in 1998. This study will look at the characteristics and relative merits of the AIX and OS/390 operating systems and the RS/6000 versus S/390 hardware platform in terms of availability.

Some of the IBM SAP R/3 systems are using Oracle as the database, while others use DB2/6000. The sites that are planning to move to the S/390 will be migrating their databases from Oracle or DB2/6000 to DB2/390. This study will also look at the characteristics and relative merits of Oracle, DB2/6000 and DB2/390 databases as they affect availability.

The IBM systems support users world wide. World wide systems add support and availability challenges. In a world wide implementation, there is no such thing as off-shift hours though peaks may still occur. This can complicate planning and scheduling such necessities as backup windows, batch windows, and regular maintenance windows all of which are normally scheduled during off-shift time.

Some issues, such as backups and database reorganizations are present for all the SAP systems within IBM. Other issues such as management backing, funding, availability of telephone lines,

and maintenance windows of service providers only affect some sites. Some of the architecture recommendations and procedural recommendations will require additional expense, such as additional hardware and/or additional support staff. A business cost/benefit analysis will need to be performed for each R/3 internal project to determine whether the benefits justify the cost.

DOCUMENT SCOPE

The goal of this study is to minimize downtime for both scheduled and unscheduled outages. This document identifies the inhibitors to continuous availability in an SAP R/3 environment. It identifies the requirements for reducing the impact of these inhibitors and outlines solution options.

Both production systems and development R/3 systems have requirements for continuous availability. Only production systems are included in this study. Most of the findings will be applicable to the development environment as well as to the production environment. Inhibitors that apply only to development will be noted. However, recommendations that only apply to development systems will be added at a later time if needed. The production systems will be running either on AIX with the RS/6000 platform or on OS/390 with the S/390 platform. Only the database server runs on the S/390 platform. The OS/390 structure for R/3 requires the central instance and application servers run on the RS/6000.

The AS/400 and NT operating systems and the other databases supported by SAP are not included in this study.

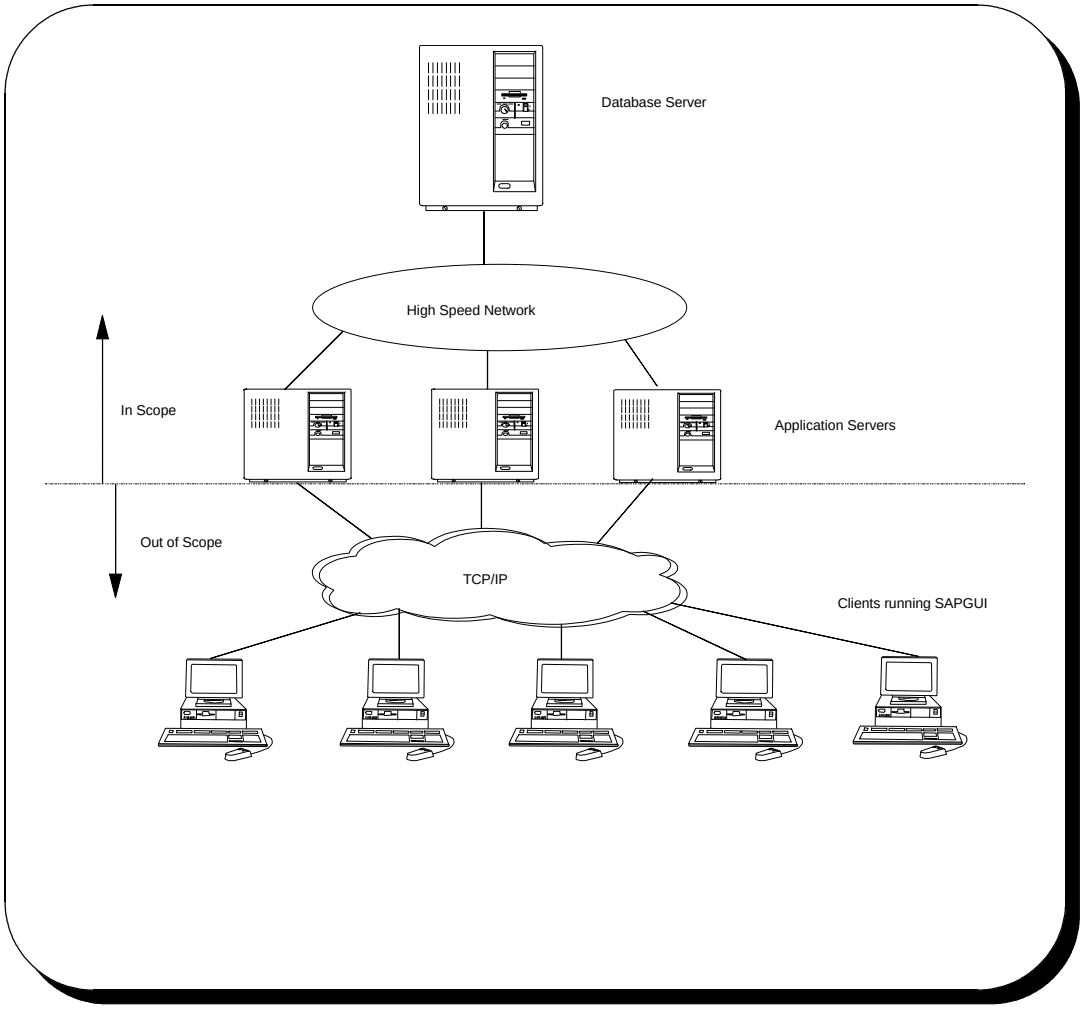
Availability includes the database server, the central instance (if it is separate from the database), associated application server(s), and the network between these systems. It does not include the end user workstations, the SAPGUI, or the network (LAN/WAN) that attaches the client workstations to the application server(s). When appropriate and reliable tools exist to measure and support end-to-end availability, this study will be updated to include the client workstations, the network (LAN/WAN) from the client workstations to the application servers, and SAPGUI. Information on a measurement tool that is currently being implemented is included in this document as an appendix.
Although the network availability from the application servers to the WAN is considered out of scope for this study, the network should be designed, whenever possible, to eliminate single points of failure. An outage in the connection from the application servers to the WAN will be perceived by the end users as a system outage. A network solution that is currently being evaluated in E/ME/A is included as an appendix for reference.

There are five key elements of availability as follows:
- The SAP R/3 Database Server is available and connected to the local network.
- The SAP R/3 Central Instance (if used) is available and connected to the local network.

- At least eighty percent (80% or better) of the SAP R/3 Application Servers are available and connected to the network.
- None of the critical application server groups are down. Critical application server groups are those configured for a specific function or group of users, such as for a specific language.
- The Operating system, Database Management System and SAP R/3 software on the database server, central instance (if used), and available application servers are running and functioning correctly.
- Those network elements required to permit communication among and between these various servers and systems are available and functioning correctly. (This excludes network connections outbound to the end user).

If an application server is down, the end users who normally connect to that application server are able to connect to the R/3 system through another application server. Logon load balancing can, and should, be set up to automatically select an available application server. The end user would automatically be connected to an eligible alternate and would need to login to R/3 again. In this situation, it is possible that there may be a performance impact. If close to half of the application servers configured for a specific customer set are not available during the peak hours, the performance impact on the system could be quite substantial. The actual performance impact depends on the number of users logged on and what those users are doing. Because the IBM projects vary in size and scope, it is difficult to determine how many application servers can be down before performance becomes a problem.

Disaster Recovery is the subject of a separate study.



ASSUMPTIONS

Systems follow design recommendations of IBM SAP Tower Offering
Not applicable to platforms, operating systems, and releases outside the scope.
The reader is assumed to have some SAP R/3 knowledge.

TERMS

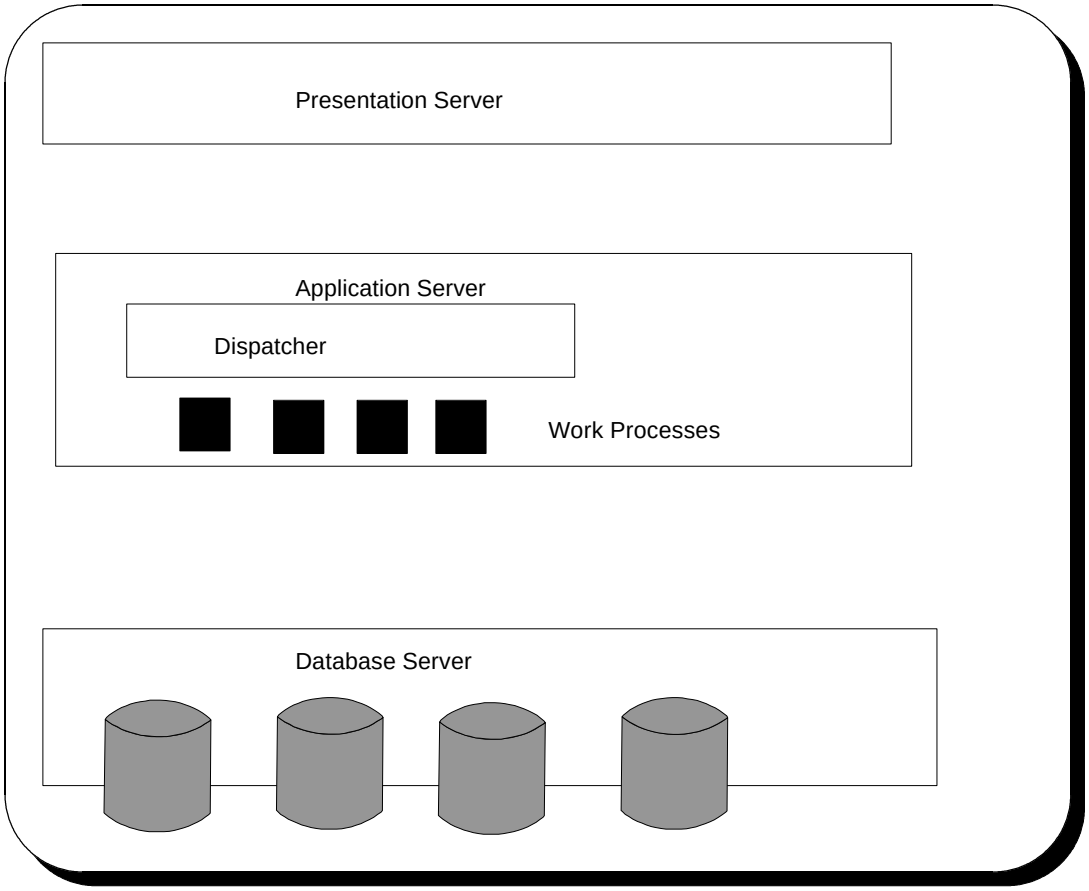
Term	Explanation (as defined by SAP)
R/3 System	The whole system consisting of one or more R/3 instances and the underlying database system seen as entity.

Term	Explanation (as defined by SAP)
Database system	Database software or database management system which underlies R/3.
Central instance	R/3 instance without a database, but including a dispatcher process with at least the two system-critica components: enqueue and message service. In additi to this minimal configuration, the central instance usu contains additional work processes (dialog, update, batch, spool, gateway).
Central system	Fully-fledged R/3 instance including the database. A central system is a standalone R/3 unit.
Application server	R/3 instance, without a database, consisting of a dispatcher process and several kinds of work processes (dialog, update, batch, spool, gateway). An application server does not contain the central, system-critical enqueue and message service components.
R/3 instance	General expression for a single R/3 client server unit that runs on a specified host machine. An instance m be a central system or some sort of application server It is started up together as a unit which is configured common instance profile.

SAP R/3 ARCHITECTURE - MAPPING OF SYSTEM SERVICES

The R/3 system has a three tier architecture which can be divided into the following logical components: database, application, and presentation. The multi-layered architecture of the R/3 system makes it possible to separate the application logic from the database and the presentation logic. The flexibility of this architecture allows for a wide variety of implementations. The R/3 system can be run on a single machine or be distributed across multiple machines to improve performance, increase workload, and improve availability.

The database service can be run on AIX, OS/390, OS/400, and several non-IBM operating systems. For the purpose of this study, only AIX and OS/390 operating systems are considered. If the database services are run on AIX, the database service can only be offered on one host machine. Because the database service can only be offered on a single host machine and because the R/3 system will not run if the database is not up and running, the database service is a single point of failure in the R/3 architecture. HACMP can eliminate lengthy outages caused by hardware problems and some operating system and software problems if the database server is running on AIX. The parallel sysplex provides similar protection against hardware and operating system outages on the database server running on OS/390.



Regardless of whether the database is running on OS/390 or AIX, the application services must be run on AIX. The application services can be run on one or more machines. If the database service is run on AIX, the application services can be run on the same machine as the database. Each application server has a single *Dispatcher* that manages the work load of the instance. The dispatcher assigns end-user requests and units of work to the work processes of the instance for completion. Each work process is assigned a primary role by the dispatcher that will control, to a certain degree, what type of work will be performed by that work process. The number of work processes and the types of work processes that exist for an instance is set in the instance profile. The work processes and the dispatcher appear to AIX as a collection of processes executing in parallel. The work processes consist of the following: dialog (D), update (V), enqueue (E), batch (B), message (M), gateway (G), and spool (S). At least one of each process, except spool, is required in the R/3 system. The R/3 system can run without any spool processes if no printing is required. Only one enqueue and message process can run on an R/3 system. In release 2.2, more than one update service requires careful configuration. SAP recommends only one update service for release 2.2. The enqueue and message services make up what is referred to as the "central instance." Multiple dialog, update, batch, gateway, and spool processes can be run in the R/3 system. However, at most, only one spool process and only one gateway process can be configured per application server.

On the AIX platform the R/3 architecture has three single points of failure: the database, the enqueue service, and the message service. In release 2.2, the update service is an additional single point of failure in the architecture. On the OS/390 platform if the parallel sysplex is used, the enqueue and message services are single points of failure. These services are more susceptible to system failures because of lack of redundancy. SAP's recommendation for addressing these single points of failure is to use "switchover software." For AIX, the switchover software is HACMP.

Service	Description	Layer	Number of services per R/3 System	Number of services per application server
SAPGUI	The user interface module that implements platform-specific input an output functions of the R/ system.	Presentation	>= 1	none
Dialog	Processes the online user input from the SAPGUI and executes the application logic	Application	>= 1	>= 0
Update	Executes changes to the database asynchronously	Application	>= 1 (For release 2.2, only one update service is supported.)	>= 0
Enqueue	Supports R/3 logical lock management of business objects	Application	exactly 1	0 or 1
Batch	Processes programs that require no interaction wit users	Application	>= 1	>= 0
Message	Coordinates requests across R/3 System services within a single R System	Application	exactly 1	0 or 1

Service	Description	Layer	Number of services per R/3 System	Number of services per application server
Gateway	Handles the communication services between R/3 Systems and external systems such as R/3 and others	Application	>=1	1
Spool	Manages spool requests and formatting	Application	>= 0	0 or 1
Database	Database services provided by the underlyin DBMS	Database	exactly 1	0 or 1

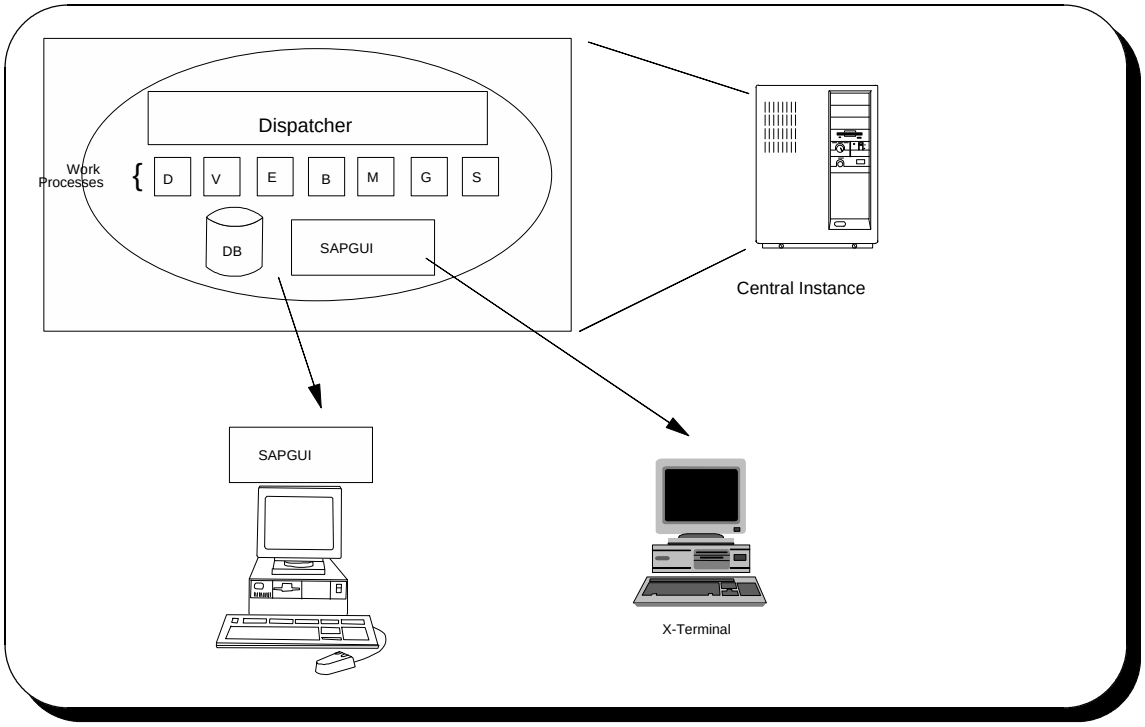
During R/3 installation, four options are supported by the installation tool, R3INST. The options are:

- Database, central instance, and application services on the same host
- Central instance without the database.
- Standalone database
- Dialog instance host

CENTRAL SYSTEM

The simplest configuration is the first option, the central system. With a central system, all R/3 application services and the database are mapped onto one physical machine. A central system is currently not possible if OS/390 is used. For the platforms covered in this study, a central system can only be installed on the AIX platform. The advantages of a central system are:

- Ease of administration - Since there is only one host machine, ease of administration and configuration simplicity are the two main benefits of a central system.
- Availability - A central system depends on the availability of one host machine. SAP recommends use of a standby host and "switchover software" to protect the R/3 system from failure of the central host machine. HACMP provides the "switchover software" to protect from failure of the central host machine.
- Performance - As long as the system has been adequately sized for all R/3 services, a central system can provide a performance advantage. A central system has less communications overhead across the network than configurations which contain application servers or a separate central instance. Typically, small R/3 installations will run on a central system.



CENTRAL SYSTEM WITH ADDITIONAL DIALOG HOSTS

Additional dialog hosts can be added with the fourth installation option: "Dialog instance host." To increase the number of users supported by the R/3 system, additional dialog hosts can be added, as long as the CPU on the database host has enough cycles to support the additional dialog hosts. This configuration is not supported for OS/390. In the OS/390 configuration, the central instance (enqueue and message) must be run on AIX. For the platforms covered in this study, this configuration can only be used if all components are run on the AIX operating system.

The advantages to this configuration are:

- Availability of the central system - The critical components of the R/3 architecture - the database, enqueue service, and message service - are still located on a single machine. This machine can be protected by one standby HACMP system.
- Availability of dialog host machines - At installation the dialog host machines contain only the dispatcher and dialog services. Other services - such as batch, update, spool, and gateway - can be added by modifying the instance profile. These machines are also referred to as application servers. If there are enough application servers to maintain adequate performance after the failure of one application server, it is not necessary to introduce an additional standby host to maintain availability requirements. Logon Load Balancing should be used to manage the application servers.

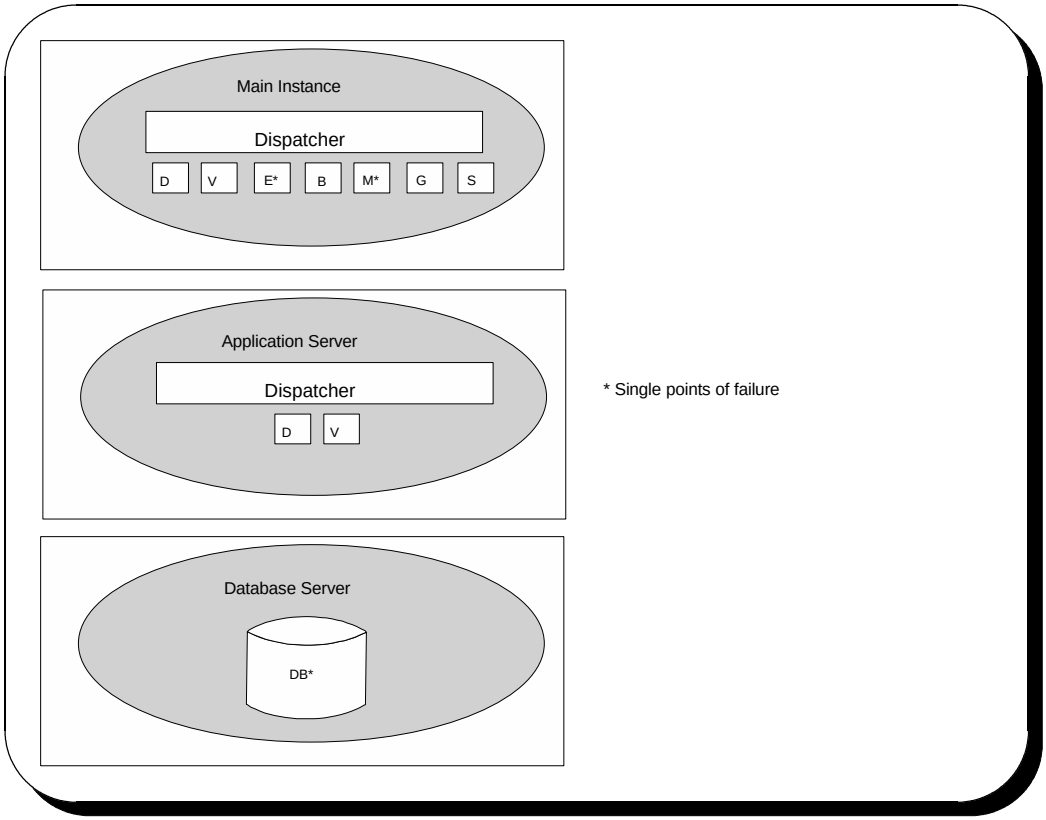
{ Performance - Mapping services onto separate hosts frees the database and central instance from the burden of providing these services. This configuration is generally used for medium to large installations. If the R/3 system is supporting a large number of users, the benefits of off loading application services to other machines will outweigh the overhead needed for network communication. The number of users a central system will support depends on the physical resources, on the number of concurrent users, and on the activity of the users.

STANDALONE DATABASE

The primary reason for introducing a standalone database host is performance. This configuration allows the CPU cycles required by the other services to be off loaded on other hosts and is typically used in medium to large installations. This configuration is used with OS/390 and most large RISC implementations. The database server runs on OS/390 or on AIX. The central instance and application servers run on AIX. To maximize availability, proper mapping of the R/3 system services is essential.

The advantages of this configuration are:

{ Availability of the database server - The R/3 System depends on the availability of the database server running the DBMS. It is important to protect the database server with HACMP or with Sysplex failover support.
{ Availability of the Central Instance - The R/3 system also depends on the availability of the Central Instance. The availability of the Central Instance should be protected with HACMP.
{ Performance - For large installations, the largest machine can be dedicated to providing only database services. A separate machine can be used to provide the critical R/3 services.



GENERAL RECOMMENDATIONS

The reduction of single points of failure will help to increase availability. Consolidate the critical services on the most reliable host machines. Protect these services with HACMP. As stated earlier, the critical services are the database, enqueue, and message.

Distribute non-critical resources. Dialog, update, batch, gateway, and spool services can be offered by multiple host machines. Distributing these services onto multiple hosts can provide redundancy and protect these services from hardware failures of any one host.

Use Logon load balancing. R/3 logon load balancing allows application host machines in an R/3 system environment to be used more effectively. At logon time, users are distributed over the application hosts in a defined group. If an application server has failed, the users will automatically be connected to one of the working application servers in the group.

Use printer groups for printers. If a single printer is down or otherwise unavailable, output will be sent to another printer within the printer group.

SAP R/3 FOR OS/390 STRUCTURE

SAP R/3 for OS/390 implementation supports only the DB2/390 database. The implementation is only available as a three tier structure. Presentation services run on workstations connect to application servers that are running on AIX or Windows NT. (Windows NT is outside the scope of this document.) The supported platforms for the presentation servers are the same as for the AIX solution. Only AIX application servers will be considered for this study. The SAP R/3 database runs on an OS/390 system connect to AIX application servers through an ESCON channel which can be connected through a gateway node.

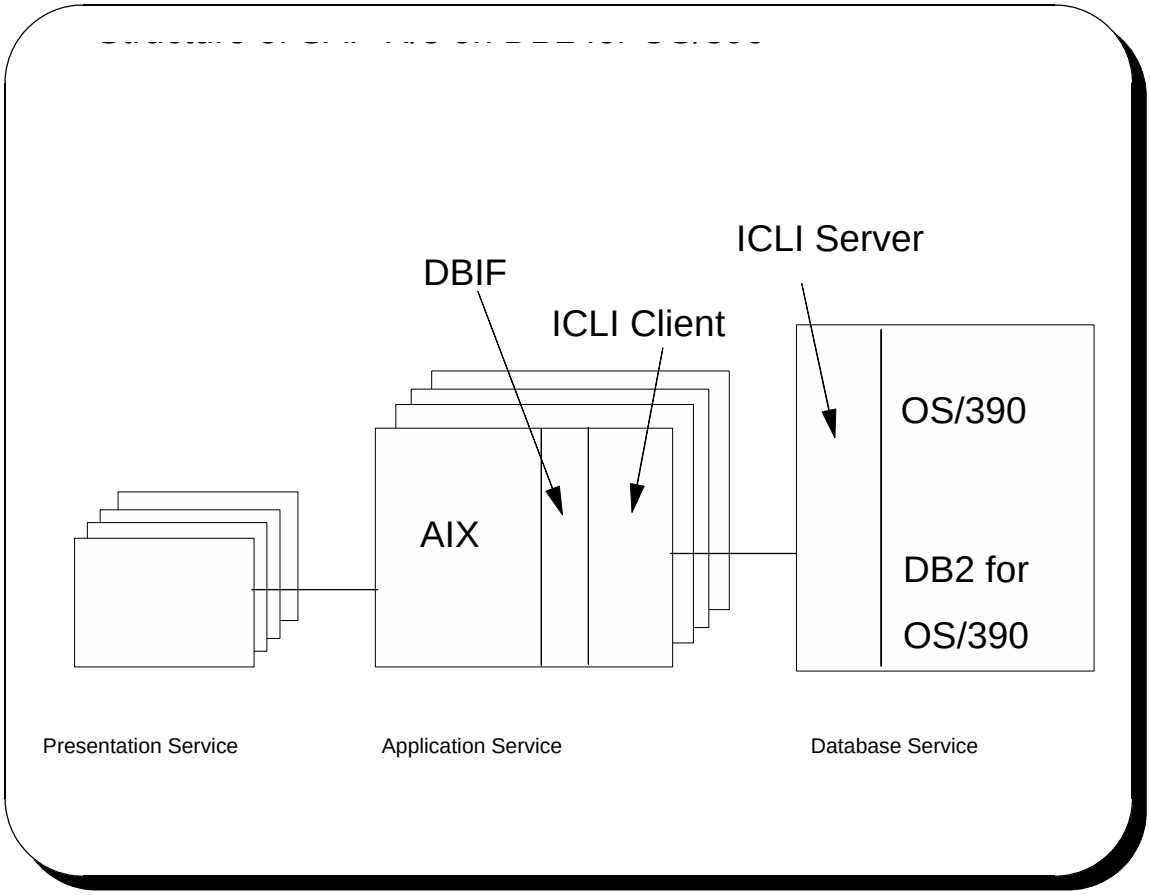


Figure 6. Structure of SAP R/3 on DB2 for OS/390

The Database Interface (DBIF) of SAP R/3 has been modified to support DB2 for OS/390. The DBIF resides on the application server and is responsible for interacting with the database.

The DBIF uses a component called the Integrated Call Level Interface (ICLI) for communication with the database service. The ICLI consists of a client and server component, which allows AIX application servers to access an OS/390 database server remotely across a network. The DBIF uses only a subset of database functions and the ICLI delivers exactly that subset.

There is a high-speed, high bandwidth communications connection between the application server and the OS/390 database server. As a tactical solution, IBM has implemented a connection called Enhanced ESCON, in which an AIX system is the application server connected to an OS/390 database server system through ESCON channels.

As a future strategic solution, IBM will implement a new high-speed communication that supports ESCON and Open Systems Adapter Feature (OSA-2) connections. The new communications support allows multiple ESCON channels or OSA-2 for high availability and throughput.

Figure 6 (Structure of SAP R/3 on DB2 for OS/390) shows the important components for SAP R/3 on DB2 for OS/390:

DBIF
The Database Interface (DBIF) of SAP R/3 has been modified to support DB2 for OS/390. The DBIF resides on the application server and is responsible for interacting with the database.

ICLI
For communication with the database service, the DBIF uses a component called the Integrated Call Level Interface (ICLI). The ICLI consists of a client and server component, which allows AIX application servers to access an OS/390 database server remotely across a network. Multiple ICLI servers can be set up which prevents the ICLI from being a single point of failure.

The server component of the ICLI is an OpenEdition-based program, while the client component is an AIX-shared library.

High-Speed Communications Connection

The high-speed, high-bandwidth communication link that ICLI uses between the application servers and the OS/390 database server is called Enhanced ESCON. IBM has implemented Enhanced ESCON as a tactical solution. It connects the application servers to the OS/390 database server through ESCON channels.

For communication between the System/390 processor and the SP processor, only a subset of the SP nodes running SAP R/3 applications need to be connected to the System/390 through

the ESCON channel. The SAP R/3 applications on the remaining SP nodes access OS/390 by using the ESCON-connected nodes as gateways (which are accessed through the SP High Performance Switch).

The Enhanced ESCON support provides an OS/390 internal communications protocol that is roughly compatible with the standard AF_INET UDP protocol for ESCON-connected RS/6000s.

Enhanced ESCON support consists of the following parts:

- An OS/390 physical file system (PFS) that provides enhanced ESCON communications through a new I/O device driver for the ESCON channel interface. The new PFS is called the AF_UEINT PFS.
- A complementary network interface driver for AIX that supports both the existing CLAW and the new Enhanced ESCON communication interfaces.

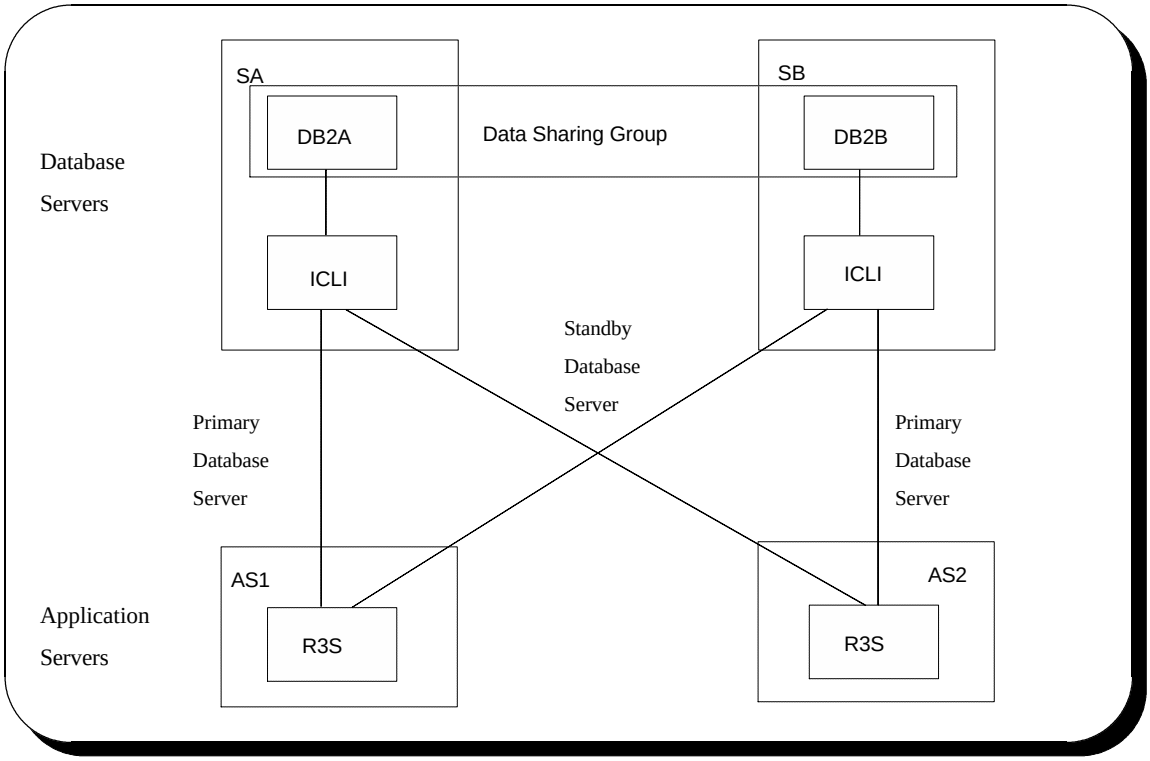
SYSPLEX Failover Support

SYSPLEX Failover Support is the capability of R/3 on DB2 for OS/390 to redirect application servers to a standby database server in case the primary database server becomes inaccessible. When a R/3 work process detects the primary database server is inaccessible, it rolls back the current transaction and automatically performs the redirection. The work process detecting this situation propagates this knowledge to all other work processes on the same R/3 instance. If the standby server becomes inaccessible and symmetric reconnect is allowed, the work processes are redirected back to the primary database server.

Redirection to a standby database server requires the use of DB2 datasharing. All primary and standby database servers must be members of the same data sharing group. A primary and secondary database server can be defined for each application server. You can also define which ICLI server instance on the standby database server an application server connects to after a failure.

The SYSPLEX Failover Support can also be used in a non-data sharing environment to recover the failure of an ICLI server. In this case, two ICLI server instances would be run on the same OS/390 system. The second instance would be used as a standby database server.

The SAP R/3 on DB2 for OS/390: Planning Guide recommends that the primary database server for the application servers should be distributed across all data sharing members. The standby database server should also be distributed across all data sharing members. Each ICLI server instance should be used for both primary and secondary connections. A simple data sharing scenario with two members is illustrated in the following diagram:

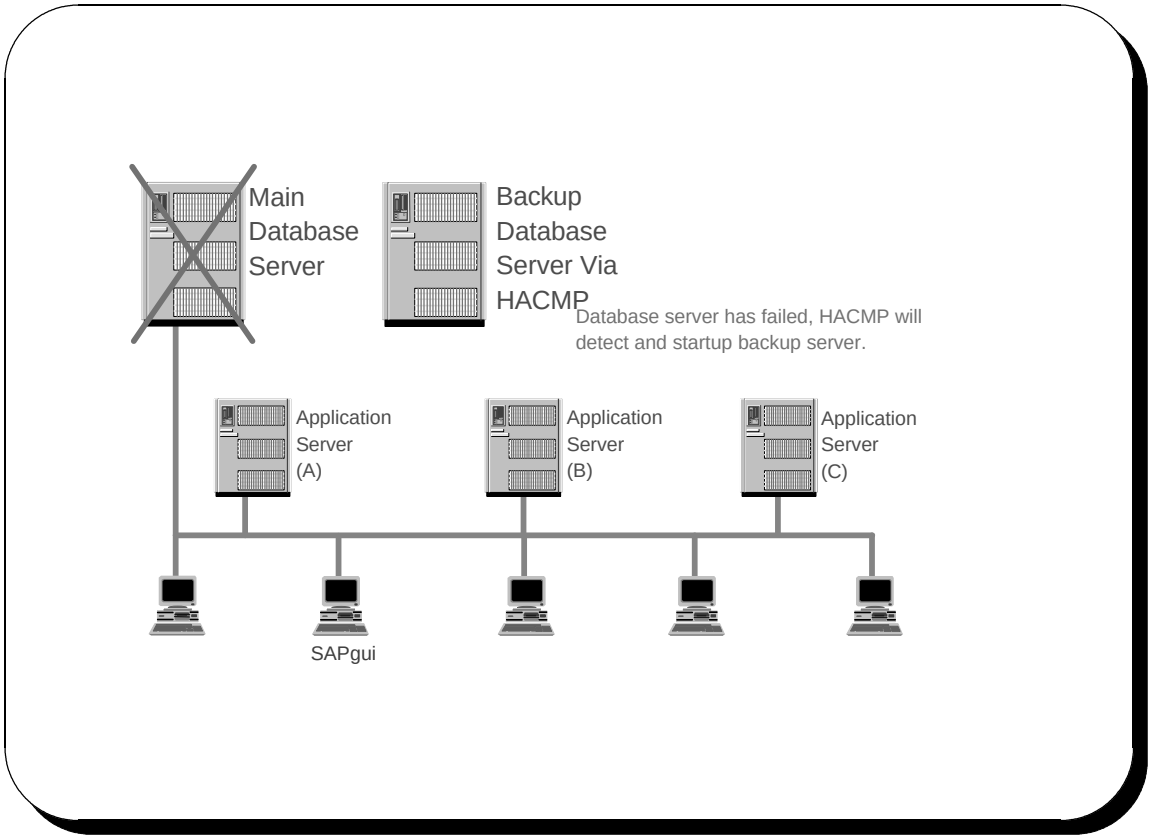


The application servers must be able to reach the primary and secondary database server hosts using both TCP and UDP protocols. The HPDT UDP or Enhanced ESCON communications as well as standard TCP/IP communications must be set up appropriately.

It is important to start a failed DB2 member on the same or other OS/390 member as soon as possible in order to release the retained locks. The Automatic Restart Facility should be used to reduce the time a particular DB2 member is down. When a DB2 member stops abnormally, the surviving OS/390 systems determine if the corresponding OS/390 failed also and restart the DB2 member on the same or different system.

HACMP Failover Support

As discussed in an earlier section, R/3 has some single points of failure. High Availability Cluster Multiprocessing (HACMP) runs on AIX. It monitors machines that run critical resources. If a machine running a critical resource becomes unavailable, the service or services are restarted on a failover or standby machine. The users will notice a short outage while the system restarts on the failover machine. They may have to log back into R/3. The disks attached to the database server or to the central instance are twin-tailed and attached to both the primary and failover systems.



24X7 INHIBITORS - SCHEDULED

Certain activities which must be performed in the normal operation of a production system inhibit the provision of a 24x7 availability of the SAP Tower Server Offering. This section describes those tasks and the effect they have on availability.

As stated in the Service Level Agreement for the IBM SAP Tower Server Offering, a maintenance window is negotiated and established for each of the SAP systems. As projects combine, the competing needs of the users will make planned downtime more difficult to schedule and probably result in less frequent maintenance windows of longer duration. The maintenance window can be used to perform the following tasks:

{ Off-line database backups
{ Client Copy
{ HACMP Testing
{ SAP Database maintenance such as table reorganizations
{ Business Data Loads
{ Disruptive Correction/Transport requests
{ Operating system backups
{ Software maintenance

- Periodic reboot of AIX
- Hardware maintenance, installation and removal
- Network hardware and facilities maintenance, installation and removal
- Raised floor and facilities maintenance

In addition to the obvious impact to continuous availability for scheduled system down time, stopping and restarting SAP R/3 causes a performance impact. When an R/3 instance is shutdown, the buffers are flushed. After the restart, there will be a performance impact until the buffers are reloaded.

An exact hardware and software production mirror is recommended by the US Competency Center for all production systems. This system should be refreshed periodically from a production backup. This system is used to test any changes or upgrades to the system. It will help identify any problems with the maintenance so that these problems can be resolved without affecting the production system. This system is referred to as the "mirror production system" throughout this document.

Scheduled Facilities Outages

Power Outage

Some of the IBM facilities schedule outages for long holiday periods or for maintenance. If the facilities outage is a power outage of very short duration, the system can be run on the UPS. If the power outage is longer than UPS can provide power to the systems, either the system must be shutdown, some other source of power must be provided, or the system must be moved to a remote location.

There are a few options to providing other sources of power. These can be used together or separately. One solution is to run the systems on a generator. When the power is shut off, the systems could run on UPS until the generator can take over and provide power. Local client workstations will lose power if they are attached to the same power source as the database server, central instance, application servers, and active network components unless the workstations are also attached to the UPS. If local workstations can draw power from the generator, the users can log back in to R/3 once power is restored. In a scheduled outage, the end users should be warned if they will lose workstation power in a planned outage. If the workstations are remote, the switch to UPS and then to the generator should be transparent.

Redundant power sources from different transformers can be set up with the local public utility company. If a transformer or the power lines are down due to maintenance or failure, an outage can be avoided by switching to the alternate power source. UPS must be in place to avoid outages when moving to the alternate power source.

High Availability Geographic Clustering (HAGEO) offers real-time data mirroring across LANs and WANs. This offering would allow the production workload to be failed over to an off-site location for the duration of the power outage. This emerging technology currently does not work well with large database products such as R/3, but may be feasible in the future.

Other hardware vendors appear to offer a disk system which allows a remote third mirror. The mirrors are synchronized at the hardware level, so no additional load is placed on the CPU while the disks are synchronized. This would allow the option of running the production system at a remote site for the duration of the outage. Once power was restored to the production site, the production disks could be synchronized again from the remote site. When the synchronization is completed, the production workload could be moved back to the production site.

Another option is to use the BRS service offering to restore the production system off-site and run the production system at the remote location for the duration of the outage. Because of the time involved to take a complete backup, restore the production instance on the remote system, perform a roll-forward recovery, set up the network, this is probably not a feasible solution unless the planned outage is longer than the time estimated to recover from backup, move the users, and then reverse the process when the facility is available again.

Other Facilities Outage

For other facilities outage, the only options may be to move the production workload to a remote location or to shut the production system down for the duration of the outage. This could be done by using the remote third mirror capability provided by other hardware vendors or using the BRS offering to restore the production system off-site.

HARDWARE MAINTENANCE

Hardware maintenance falls into two categories; changes that require the system to be shut down and changes that do not require the system to be shut down. Some hardware maintenance can be done while the system is running. If the same maintenance must be done on the development and consolidation systems, it should be performed on these systems before it is applied to the production system. If there are any problems, they can be identified and resolved without causing an outage on the production system. If this maintenance does not need to be applied to the development and test SAP systems, it should be applied to the production "*system test*." Whenever possible hardware maintenance, whether or not it requires a system shutdown, should occur during the scheduled maintenance window. If there is an unplanned outage caused by the maintenance, it will have minimal impact to the production users.

The SAP Tower System Offering recommends that all production machines on AIX have HACMP set up. For hardware maintenance requiring a system shutdown, the production workload could be moved to the HACMP failover machine. A short outage would need to be scheduled to move the production workload to the failover machine. After the maintenance is complete, the failover machine could either become the primary production machine until the next outage or an outage would need to be scheduled to move the production workload back to the primary machine. While hardware maintenance is in progress on the failover and on the production systems, HACMP failover capability will not be available.

If the database is running on OS/390, the S/390 Parallel Sysplex will help avoid an outage for hardware maintenance. The Parallel Sysplex permits dynamic workload balancing through its datasharing architecture and workload management. Multiple copies of the operating system and SAP R/3 application run at the same time, each having access to all the resources in the Parallel Sysplex. As a result, should any hardware component be taken off-line for maintenance, SAP R/3 will continue to run. The ability to dynamically adjust the workload permits the system to provide a consistent and stable environment.

OPERATING SYSTEM MAINTENANCE AND UPGRADES

Operating system maintenance falls into the same categories as hardware maintenance in that it may or may not require the system to be shutdown. All changes should first be tested on the development, consolidation, and "mirror production system" environment prior to being loaded on the production systems. If possible, all changes should be loaded on the production system during a scheduled maintenance window since there is always a possibility of the system outage caused by applying the maintenance.

Since release 3, HACMP has offered release independence. The operating system does not have to be at an identical levels on the failover system and the primary system. The changes should be loaded on the failover system first. HACMP may not be available while this change is being applied. If HACMP is compatible with the changed and unchanged operating systems, the production workload can be moved over to the failover machine. A short outage would need to be scheduled to move the production workload to the failover machine. The change could then be applied to the primary system. While this change is applied to the primary system, HACMP might not be available. After the change is complete, the production workload can stay on the failover machine until the next outage or it can be brought back to the primary system. If it is brought back to the primary system, a short outage will need to be scheduled.

Maintenance should be applied on the development, consolidation, and "mirror production systems" before it is applied on the production system. No untested changes should ever be permitted on the production system.

For OS/390 with the S/390 Parallel Sysplex, SAP R/3 can continue to be available to users when the operating system on one of the processors is not available. Maintenance can be applied while the system is running. The workload will be dynamically adjusted across the other processors providing continuos availability. DB2 V5 data sharing along with dynamic workload routing allows PTF maintenance to be applied without disrupting service.

SOFTWARE MAINTENANCE AND UPGRADES

Whether or not an outage is required depends on the software package an the maintenance or upgrade being applied. For example, upgrades to ADSM, or to single application servers probably will not require a system outage. An upgrade to HACMP probably will require an outage. All software maintenance and upgrades should first be performed on the mirror production″ machines. If the work can be successfully completed on the "mirror production system" without causing a system outage, it can be scheduled on the production complex. The changes should only take affect during the scheduled maintenance windows if there is any chance that they could cause a system outage or cause an outage to an essential system component.

HACMP upgrades and changes should be tested on the "mirror production" complex. The HACMP failover testing can be repeated and verified in this environment.

NETWORK MAINTENANCE AND UPGRADES

All network maintenance and upgrades should be tested on other systems before the changes are applied to the production system. If possible, maintenance and upgrades can be done while the system is running. If there is a possibility that the changes will require a system outage, the changes should be done during the network maintenance window to minimize the impact on the end users. If it is possible, the network maintenance window should be coordinated with the system maintenance window.

TESTING

HACMP testing

HACMP scripts can be partially debugged and tested on the sandbox system, however, most of the debugging and testing must be done an a system architecture that is the same as the production complex.

A copy of the production system can be built on the "mirror production system" to allow for more extensive debugging and testing. Once the scripts are debugged and tested in this environment, they must be tested on the actual production complex to verify that they will work properly on a production failover. This will require production system down time for the duration of failover testing. If possible, testing should be scheduled during the maintenance window. A complete test, which would verify the system behaved as expected for each possible failure point, could take several days to a week for the initial set up.

Another approach is to temporarily move the production work load to the "mirror production complex. Completely debug and test the HACMP set up. When testing is complete, copy the production database back onto the production complex. With this method, HACMP has already been tested on the production complex. The scheduled outage for the production environment in this scenario is when the production workload is moved onto the "mirror production" complex and again when it is moved back to the production complex.

To minimize the chance of a HACMP failover, planned or unplanned, not working properly, the production workload should be failed over to the backup machine periodically during the scheduled maintenance window. The production workload can be run on the secondary machine until the next scheduled or unscheduled failover or it can be brought back to the production complex in the same scheduled outage. Ideally, HACMP would be set up in a round-robin configuration. The production workload would be failed over weekly during the maintenance window.

Post upgrade verification for all hardware and software changes

Copy the production system to the "mirror production system"." Make all hardware and software changes to this environment first. Once all changes are tested, schedule a production outage if needed and repeat on the production system.

All changes, if they apply, should be made to the development and consolidation systems before they are loaded on the production system. If no *"mirror production"* environment is available, applying changes to development and consolidation before production is even more important.

SAP R/3 UPGRADES

In R/3 releases 2.2 and 3.0, there is no way to completely avoid downtime on the production system for an upgrade. For releases earlier than 3.0, all upgrades must be done entirely offline. Upgrades can be done with archiving turned on or turned off. If archiving is turned off, a complete backup, online or offline, must be done before and after the upgrade. If a complete offline backup is not possible before the upgrade, the upgrade can be performed with transaction logging turned on. Upgrades generate many archive logs. The archive log file must

be monitored so that archive log files can be copied to tape should the archive filesystem fill up before the upgrade is complete. If the archive filesystem fills up, the upgrade will stop. If the database is Oracle, the upgrade will resume once there is space in the archive filesystem. If the database is DB2, the database will crash and must be recovered before the upgrade can continue.

R/3 release 3.0 introduced new upgrade strategies with the "repository switch." With release 3.0 there are three possible upgrade strategies. Each strategy has a different effect on the time a complete backup of the system must be made, recovery options if the system crashes during the upgrade, and overall length of the upgrade, including duration of system down time.

Beginning with release 3.0, the preparation phase is run. This phase (called PREPARE) can be run while R/3 remains online. It identifies some conditions which would cause the upgrade to fail, such as insufficient disk space. These conditions must be resolved before the upgrade can be performed. For example, if there is insufficient disk space, the upgrade should be postponed until enough disk space for the upgrade is added to the system. The script can be run several times until all problems are resolved.

A number of steps are required after the upgrade. These steps can be performed while the system is online.

The three possible upgrade strategies for release 3.0 are:

{ A_off

Database recovery is only possible to the status before the upgrade. A backup (online or offline) must be done before starting the upgrade. Database transaction logging is turned off during the upgrade. A complete offline backup must be done after the upgrade is complete. This strategy has the longest system down time.

{ A_on

Database recovery is possible to the current status during the upgrade. A large amount of database transaction logs are generated. A backup at the end of this upgrade is not required, however, SAP recommends on online backup to avoid having to roll-forward large amounts of transaction logs.

{ A_switch

Database recovery is possible to the status before the beginning of the downtime. An online or offline backup is recommended before the beginning of the downtime. A previous backup can be used along with a roll-forward of the database transaction logs of a recovery is required. An offline backup must be done after the backup is complete.

SAP provides the following table comparing upgrade strategies (BC SAP High Availability). Actual downtime will vary depending on a number of factors such as size of database, number of clients, system activity, disk type, CPU, and so on.

	A_off	A_on	A_switch
Downtime (approx.)	14 hours + time for offline backup	10 hours	9 hours + time for offlin backup
Backup before/during upgrade?	Recommended before - may be online or offline	No	Recommended during - may be online or offline
Backup at end of upgrade?	Obligatory offline	Recommended online	Obligatory offline
Transaction logs generated?	No	Yes	Yes
Database recovery possible?	Yes - to status before upgrade	Yes - to current status during upgrade	Yes - to status before actual downtime commences
Advantages	No transaction logs	Shorter downtime	Shorter downtime
Disadvantages	Long downtime Limited recovery	Danger that transaction logs not secured. This could stop the database. Extra disk space required to store logs	Danger that transaction logs not secured. This could stop the database. Extra disk space required to store logs

The difference between the longest and shortest down time, based on SAP's estimates, is 5 hours. If the extra time can be negotiated with the end users, the A_off method is the easiest method of upgrading the system.

Upgrades should be practiced on a copy of the production system first. Copy the production system to the "mirror production system"" and run the upgrade there first. This process will identify any issues that need to be resolved and help to predict the down time needed to upgrade the production system. After the upgrade on the "mirror production system" is complete, a gap analysis can be performed, functionality can be verified, and users can become familiar with any changes to the system prior to upgrading the production system.

Upgrades should always be performed on the production system last. The development and consolidation systems should be upgraded before the production system is upgraded. There are a number of upgrade strategies. Some strategies are described below:

Copy Development and Consolidation

If the equipment is available, some locations will copy the development and consolidation systems onto other machines and upgrade the copies. This leaves the development and consolidation systems at the same level as production and allows correction and transport to continue. Another advantage is it allows the developers and perhaps the end users to become familiar with the new release before the production system is upgraded. After the upgrade on the production system is complete, the back level development and consolidation systems are no longer needed and the equipment can be used for other purposes.

Preproduction system

A preproduction system can be used as a staging area for emergency fixes and corrections while the development and consolidation systems are at different levels than the production system.

PROJECT MAINTENANCE (CORRECTION TRANSPORTS)

SAP recommends that no direct changes be made on the production or in the consolidation system. All changes should be made in the development system. Each change should be tested on the development system. Once the change is fully tested on the development system, it is transported to the consolidation system and tested again there. After the change has been fully tested on both the development system and the consolidation system, it is transported to the production system. A change control process should be established and followed. All changes to the production system should be approved and signed off that the appropriate level of testing has been successfully completed. No changes should ever be made directly on the production system or transported from the development system directly to production. There may be emergencies or situations (such as the development and integration systems have been upgraded and production has not) which may require the normal testing path to be bypassed, but these cases should be the exception, not the general practice, and should require approval by management.

As part of the change control process, transport windows may be established for the production system. Changes would not be loaded on the production system outside these windows (except in emergencies). The windows should not be on Friday afternoons or prior to holidays.

PROJECT UPGRADES

Adding new country support

When a new country is added to a production system, there may also be requirements for supporting an additional language and new configuration data to be loaded. If the language is not a subset of the code page loaded on the application servers, a new code page must be installed on the application servers that will support the new users. All application servers should run a code page which supports all supported languages. .

In some cases, the new language must also be loaded into the production system. The new language should first be loaded into the development and consolidation systems before it is loaded in the production system. The approximate length of the outage required to load the new language can be determined by loading first on the development and consolidation systems. The production system must be down during the language load. It should be loaded during a scheduled maintenance window.

Analysis should be done to determine whether there is a requirement to load the new language into SAP R/3. It is possible that the new language requirement can be satisfied by providing a code page to support the new language without having to load that language into R/3. For example, the end users have a requirement to print certain invoices in French. The users would need to be able to use the French character set to create the invoices, but may not need to login to R/3 in French. In this example, the French language would not need to be loaded into SAP R/3, but a code page supporting the French character set would need to be installed on AIX.

OPERATING SYSTEM BACKUPS

Operating system backups can be done while the system is online for both AIX and OS/390. For AIX, operating system backups can be done with mksysb or with Sysback/6000. Sysback/6000 may make recovery of the operating system faster because the image can be booted from disk, whereas mksysb must be booted from tape in the standard implementation. SP nodes are booted from disk. For disaster recovery, a Sysback/6000 tape backup must be restored to disk first and then booted from disk. In a disaster recovery scenario, this may slow recovery time. The root volume group and disk space should be kept as small as possible to allow faster backups and recovery. The OS/390 operating system can be backed up online with SMSdss.

R/3 BACKUP

Oracle and DB2/6000 on AIX

The backup window on R/3 systems is usually the most frequent of the scheduled system outages. It is possible to take an online backup, which does not require system down time. Although, online backups on the AIX platform impact performance while the backup is running, stopping and restarting R/3 for an offline backup also causes a performance impact. Shutting down R/3 for an offline backup will cause the R/3 buffers to be flushed. These buffers must be reloaded after the system is restarted. Until the buffers are reloaded, the performance is impacted.

Oracle and DB2/6000 both offer offline and online backups. Only offline backups require system down time. Backups can be done at the database level or at the tablespace level.

The fastest IBM solution for AIX is currently ADSM. ADSM can be configured to support multiple parallel backup sessions to decrease the backup time further. (Although the AS/400 is out of scope for this document, it should be noted that there is no AS/400 client. AS/400 systems cannot be backed up with an RS/6000 ADSM server.) Lab measurements in an SP environment are shown below.

Number of sessions	production throughput	lab measurement
1	12 - 17 GB per hour	18.7 GB per hour
2	24 - 30 GB per hour	33.0 GB per hour
3	35 - 43 GB per hour	47.4 GB per hour
4	41 - 51 GB per hour	56.8 GB per hour

Another vendor appears to have a faster backup system, however, this system seems to take a very long time to load the tape drives. Once the drives are loaded, the backup rate is considerably faster than ADSM. One measurement showed a backup rate of 12 gigabytes in 4 minutes with four tape drives. This measurement does not include time to load the tapes. Including tape load time, this same site backs up a 90 gigabyte database in 2 hours and 5 minutes with 4 tape drives. This falls into the ADSM range.

Another option is to take an offline backup from a third mirror. Other hardware vendors appear to offer a disk system which allows a remote third mirror. The mirrors are synchronized at the hardware level, so there is no performance impact to the R/3 system. A short outage would be required to shutdown the system and break off the third mirror. An offline backup could then be taken from the third mirror. Once the backup is complete, the third mirror can be

synchronized again with no impact to the R/3 system. The advantage to this method is it requires only a very short outage and has no performance impact to the system.

Offline Backups
During an offline backup the whole database or a tablespace will be unavailable for use. Usually offline backups are done at the database level. Offline backups at the tablespace level are risky in R/3 because the application modules require access to data in several tablespaces. If one tablespace is not available, most application modules will not be able to continue.

The major advantage of a full offline backup is it can be restored without a recovery. No redo log files are needed. The major disadvantage is the database cannot be used while the backup is done. The R/3 system must be shutdown for the duration of the backup.

Online Backups
An online backup may be a full or partial backup. Users can continue to use the system while on online backup is done. The log files generated while the backup was running must be available to recover the database.

The major advantage to an online backup is the database is available for normal use while the backup is taken. The disadvantages to an online backup are that performance deteriorates while the backup is running and that the redo log files must be available for a recovery. The database performance is impacted because more redo log information is generated while the backup is running. If used for recovery, the online backup must be supplemented with the redo files archived during the backup. If the log files are not available, the database cannot be recovered.

DB2 V.5 for OS/390

DB2 V.5 database server for OS/390 can backup the database at a much faster rate than AIX backups. In addition, DB2 V.5 has the capability of doing incremental backups. Only changed pages of data are backed up. Incremental backups are not supported for databases running on UNIX.

Currently, there is no ADSM client available for OS/390. ADSM cannot be used to backup and restore a DB2 V.5 database on OS/390. The R/3 application servers on AIX and ICLI server can be backed up with ADSM.

The COPY utility is used for database backups in of DB2 V.5 in OS/390. Copy generates either a full or incremental copy of a tablespace or data set. A COPY done with the FULL YES option results in a complete copy. A COPY done with the FULL NO option results in a copy of only those pages that have changed since the last COPY was done.

The COPY utility is run with either SHRLEVEL CHANGE or SHRLEVEL REFERENCE. SHRLEVEL CHANGE allows read and update access to the tablespace being copied, and bears functional equivalence to an SAP R/3 AIX online image copy. SHRLEVEL REFERENCE prevents update access to the tablespace being copied, and bears functional equivalence to an SAP R/3 AIX offline image copy. One difference in the latter is that an SAP R/3 AIX offline backup is taken after the SAP R/3 system is shut down. In the SAP R/3 OS/390 implementation, DB2 for OS/390 must be running. The application servers may or may not be shut down. SHRLEVEL REFERENCE will have a performance impact on the end users because no updates are permitted for the duration of the backup. SHRLEVEL CHANGE will have a minimal performance impact on the end users, but will run longer than SHRLEVEL REFERENCE.

RESTORE/RECOVERY

The restore and recovery time should be considered when designing a backup strategy. The types of recovery scenarios supported will determine the type of backup required. If the time it takes to restore and recover the system is too long, then the backup strategy may need to be redesigned.

Physical failures are the easiest recovery scenario to evaluate. The problem determination time and the recovery time can be estimated more accurately. Data integrity problems are much more complex. It may take days or even weeks to determine there is a data integrity problem. Once the problem is identified, it may be easier to fix the problem then to recover the database and lose other work completed in the interim. If a point in time recovery is required, the impact of lost transactions plus the impact to other interfacing systems must be evaluated.

Oracle and DB2/6000 on AIX

To recover from an offline backup, only a restore is required to have a consistent database. No database recovery is needed. To recover from an online backup, both a restore and a recovery are required. Recovery is a roll-forward of logs. The roll-forward can take a considerable amount of time. The worst case scenario should be calculated for a recovery. If the worst case scenario is unacceptable, the backup strategy should be changed.

For example, for online backups, the worst case would be if the system failed just before the next online backup. The time to recover would be the sum of the following:

- time to determine there is a problem
- time for decision to be made to recover or fix the problem
- time to locate backup media
- time to recall media from off-site storage
- time to restore system from backup
- time to restore archive redo logs

{ time to apply archive redo logs
{ time to verify the recovery is complete

With one ADSM session, 12 to 16 GB can be restored per hour. If the database is 24 GB, it will take approximately two hours to restore the database. The rate for applying logs is between 1/2 GB and 2 GB per hour for both DB2/6000 and Oracle. If 1 GB of archive redo log files are generated per day and online backups are done once a week. In the worst case, 7 GB of archive files will need to be applied. Assuming the recovery rate is 2 GB per hour, 7 GB of archive redo logs will take 3.5 hours. The recovery window in this case is 5.5 hours plus problem determination and verification time. If an acceptable recovery time is less than this time, the backup solution will have to be modified. One way to modify the backup solution is to do more frequent online backups. This will shorten the recovery time.

Lab measurements for restore using ADSM in an RS/6000 SP environment are:

Number of sessions	production throughput	lab measurement
1	12 - 16 GB per hour	18.0 GB per hour
2	22 - 28 GB per hour	31.6 GB per hour
3	not measured	not measured
4	34 - 43 GB per hour	47.5 GB per hour

DB2 for OS/390

The recovery effort for DB2 on OS/390 is more automated than the recovery solutions for UNIX databases, where more manual effort is necessary. DB2 has two recovery modes, point-in-time recovery and recover to currency. A difference to UNIX databases is the QUIESCE point. QUIESCE is a special user-defined point of consistency. Executing the QUIESCE utility enables recovery not only to a defined time, but also to a specific point (for example, after the last online REORG).

Recover to Currency
Recover to Currency is the activity required to restore a DB2 tablespace or tablespaces to the last complete unit of recovery for the tablespace. Included in this restore are the indexes related to all tables in the specified tablespace(s). For example, this type of recovery would be necessary in the event of a device failure, and the DB2 for OS/390 tablespaces and indexes on the device would have to be recovered to the point at which the device failure occurred.

Levels of Recovery
Database Recovery: DB2 for OS/390 database is backed up and recovered at the tablespace, partition, or data set level. In contrast, a DB2 for AIX system is backed up and recovered at either the database level or the tablespace level. The SAP R/3 database in a DB2

for OS/390 environment is neither backed up nor recovered at the DB2 for OS/390 database level. Rather, image copies are taken at the tablespace, partition, or data set level. Recovery of the DB2 for OS/390 databases that make up an SAP R/3 database is also not done at the database level, as in UNIX implementations, but at either the same level or smaller as that which was used to generate the OS/390 backups copies.

Tablespace Recovery: DB2 for OS/390 can be backed up and recovered at the tablespace level and at the partition level.

Point in Time Recovery
A Point in Time recovery is considerably more complex and time consuming than a recover to currency. Many more objects are involved. Point in time recovery is made to a point of consistency, or quiesce point. A quiesce point can be established in one of three ways:
{ ARCHIVE LOG command
{ The QUIESCE utility
{ The STOP DB2 command

There are two times when system availability will be affected:
When the point of consistency (quiesce point) is established
When the point in time recovery is taking place

ARCHIVING

Increasing quantities of data will eventually impair database performance and make database management more difficult. Data archiving removes data which is no longer required online in the system, but must be retained. SAP R/3 provides a method for archiving. The impact of archiving on system availability differ depending on the release. In releases 2.2 and earlier, archiving requires system down time. From release 3.0, system down time is no longer required.

Release 2.2 on AIX

In release 2.2 of SAP R/3, the archiving procedure requires a full database backup prior to beginning the archive. The backup can be online or offline. During the archive process, no users and no batch jobs can be running on the system. Should the archive process fail before it completes, the database must be restored from the backup taken prior to starting the archive. After the database restore is complete, the archive can be started again from the beginning. Once the archive has completed successfully, SAP recommends backing up the database again and a complete database reorganization. After a database reorganization, another backup should be taken.

If online backups are used, the system is not available during the archive and during the database reorganization. If the archive does not complete successfully, the system will also be unavailable during the database restore/recovery.

In release 2.2, data archived cannot be brought back into the system. This function at this release would more appropriately be named purge rather than archive. The only way to retrieve archived data is a recovery from backups taken prior to the archive.

Release 3.0 on AIX

Beginning with release 3.0 of SAP R/3, there are three major changes in the archiving procedure with respect to availability:

- A database backup is no longer required before beginning the archive.
- If the archive fails before it completes, it can be restarted. It is no longer necessary to restore the database and start the archive from the beginning.
- The archive can be run online. The archive can run parallel to R/3 system operation, however, there will be a performance impact.

In release 3.0, archiving is performed in three steps:

- Create archive file
- Execute delete program
- Store the archive files

After closing the first archive file, a new archive file is created and the archiving procedure continues. The delete program reads the data from the closed archive file and deletes it from the database. This procedure guarantees that only data which has been correctly saved in the archive file is deleted from the database.

Once the archive and delete have completed successfully, SAP recommends backing up the database again and a complete database reorganization. After a database reorganization, another backup should be taken.

If online backups are used, the system is unavailable during the database reorganization. The users can continue working during the archive. If the archive fails, it can be restarted without performing a database restore/recovery.

If the archiving is done on a smaller and more frequent basis, the recovery scenarios should include point in time recovery to before the archive. After the system is recovered, the data that was archived will exist in the archive storage system as well as in the R/3 database. In this case, the R/3 archive must be repeated to remove the out-of-date information again. The archive storage system must be able to handle duplicate data in some manner.

Backups of the archived data and the archive system database should be included in the backup scheme. If there is a requirement to retain archived data, backups of the archived data as well as the archive system database must be considered for disaster recovery and business recovery procedures.

Backup and recovery for archived data and the archive system database not relevant in releases 2.2 and prior because there is no way to retrieve archived data other than a system restore.

Release 3.0 on OS/390

SAP R/3 became available on the S/390 platform with release 3.0. Archiving with release 3.0 can be done online. After the archive and delete are complete, a backup and a database reorganization are recommended by SAP. Both of these tasks can be completed online in the OS/390 implementation. There may be a performance impact to the users during the archive and subsequent backup and reorganization, however the SAP R/3 system will remain available.

CLIENT COPIES

Periodically there may be requirements to copy the production client to the development and consolidation systems. A copy of a client from one system to another system can be done with a database export, with a client export, or with a restore from backup. With the database export and the client export, the system is not available to the users during the export. A restore from backup would not require additional production down time, however, the target system would need the same amount of disk space as the production system.

Client Export from Production

While the client is exported from the production system, the client should be unavailable to users. No changes can be made to the client while the export is running. Once the export is complete, the system can be made available again. The import to the target system does not impact the production system. In some cases, there are problems with the export that are not apparent until the client is imported into the target. In this case, the client would have to be exported again from the production machine. As the production client grows in size, the client will have to be exported and imported in smaller sections. Running a test import on the target client will bring up potential problems which can then be fixed before running the actual import. If there is a problem with just one table on the import to the target system, the entire client will have to be exported again if any changes have been made to the client since it was last exported. The only way to ensure that no changes are made to the client until the entire client is successfully imported into the target is to make the system unavailable to the users. This will result in a lengthy outage as the export and import can take several hours.

Export from a copy of Production

A copy of the production system can be restored from backup onto the *"mirror production"* environment. The client export can be done from the copy of the production system. With this approach to client copy, no outages are required on the production system for the client copy. The system can be kept in a consistent and unchanged state until the export and import are complete without causing an outage on the production system. Because there has been no activity on the copy of the production client, a single table or tablespace can be exported again if necessary. The export can be done over several days because the client is not being changed.

DATA MIGRATIONS AND BATCH JOBS

Data migrations from another system into the SAP R/3 system are usually necessary during a conversion to SAP R/3. Sometimes data migrations continue to be needed as an ongoing practice. These migrations typically do not affect system wide availability, however they will impact performance. It is possible for the performance impact to be so severe that the system appears unavailable to the end users. To minimize the performance impact, whenever possible, batch jobs and data migrations should be scheduled to run during non-peak hours.

Good procedures and coding practices can also minimize the performance impact. The programs that will perform the data migrations should be reviewed, approved, and thoroughly tested in the development and consolidation system prior to being transported to the production system. Stress tests should also be done on the *"mirror production* platform prior to using data migration programs on the production system. If the production system is running on an SMP, it is usually better to run several smaller migrations than to run one large migration job. If only one job is running on the system, the other processors may be idle and waiting for work. Disk I/O monitoring and balancing is also necessary. If all the programs are writing to the same disk, I/O contention will become a performance bottleneck.

Batch jobs will also affect performance rather than availability. If there is a requirement for a medium to heavy batch load or report load, a batch administrator should be assigned to manage scheduling. Batch jobs and reports can place a heavy load on the system and substantially degrade online performance. The batch administrator should, whenever possible, schedule batch jobs to run during non-peak hours. If there is a batch window, the services on one or more application servers can be changed from dialog services to batch services for the duration of the batch window. If the batch jobs cannot complete in the non-peak time, the large batch jobs should not be run concurrently. Instead, large batch jobs should run sequentially to minimize the performance impact to the online users.

On an SMP machine, several small batch jobs will run more efficiently than one large batch job because the small jobs can run in parallel on different processors. When possible, split batch workload into smaller jobs rather than combine them into one large job.

DATABASE MAINTENANCE

Space Management with Oracle and DB2/6000

Proper sizing of the initial system is important. The data files are the actual operating system files used to store data. One file belongs to only one tablespace, but a tablespace may consist of multiple data files. At installation, the data files should be created with enough space to accommodate anticipated growth. The database data dictionary in the R/3 default installation is usually sufficient for a development system. The data dictionary seems to become fragmented when the database grows beyond 15 to 30 gigabytes. If the database is anticipated to grow beyond 15G, the SYSTEM tablespace should be increased in the default installation. Reorganization of the database data dictionary requires a full database export and import. This is a lengthy process and requires the system to be unavailable to the end users. The best way to avoid a data dictionary reorganization is to increase the space at installation.

Monitoring of the space is also important. Additional disk space should be added before it is needed to allow room for growth. Anticipating growth can be challenging, but running out of space can cause system down time.

If proper sizing, monitoring, and capacity planning is neglected, it can eventually lead to system down time. As the production database grows in size, database objects will need space to grow. If that space is not available, SAP R/3 cannot write to the database. When space must be added quickly and has not been anticipated by proper monitoring and planning, it is often not added in an orderly manner. Later on, when more disk space is added, the database files will probably need to be moved to balance I/O or to make administration and management easier.

SAP recommends running several jobs periodically to delete information that is no longer needed. These jobs are referred to as Reorganization jobs. SAP describes the jobs and the recommended frequency they should run in OSS note 16083. These jobs will delete obsolete spool requests, batch jobs logs, ABAP dumps and other information that is no longer needed. When possible, delete old information rather than allocate additional space. These jobs can be scheduled to run automatically and most can be customized to meet the retention periods established for the end users. Running these jobs periodically can help with managing the database and possibly prevent system downtime by keeping the tables which contain this information from filling up.

Database Reorganization

With both Oracle and DB2/6000, SAP R/3 should be down when a reorganization is running. DB2/6000 has a utility which will determine when a reorganization is recommended for a

specific table or index. Oracle requires monitoring and analysis to determine whether a table or index should be reorganized. Situations requiring reorganization are:

{ Extent overflow
{ Tablespace overflow
{ Fragmentation
{ Chained rows
{ Maximum number of files reached
{ File size limitation (in AIX 2 gigabyte file restriction)

Oracle release 7.3 supports online reorganization for some indexes. SAP does not make a recommendation for or against online reorganization. However, SAP will not provide support for any problems encountered during an online reorganization. Because of locking problems and possible data integrity problems, it is best to schedule down time for all reorganizations.

With DB2 V5 for OS/390, there is little or no requirement for reorganization in the SAP R/3 application. If a reorganization is required, the online REORG feature of DB2 Version 5 almost eliminates disruption of service to the SAP R/3 end users. During a share-level change reorganization, there is a brief time during which updates are prohibited. Measurements during lab testing of the online REORG with share-level change showed that updates were prohibited for about three minutes.

Periodic Reboot of AIX

There are different opinions about whether AIX needs to be rebooted periodically. In AIX 3.2.5 and in traditional UNIX systems, there are some problems with "memory leaks" that require periodic reboots of the system. Beginning with AIX 4.1, it is questionable whether periodic reboots of the system are still necessary. At the IBM location in Poughkeepsie, the AIX servers have, in some cases, run for up to six months without a reboot. Other locations choose to reboot servers at least once a month.

24X7 INHIBITORS - UNSCHEDULED

In the new client server computing environment that has evolved in recent years, the systems management practices and procedures and the development practices and procedures have often been overlooked or forgotten. The best way to minimize both the frequency and the duration of unscheduled outages is to establish and maintain systems management procedures and development practices.

Systems management procedures should be established and maintained for change control, problem management, capacity planning, backup, recovery, and disaster recovery. In addition, the R/3 development team should have defined procedures for testing and reviewing new and

changed code before it is put on the production system. A phased development process with check points should be implemented.

Roles and responsibilities should be clearly defined. If roles and responsibilities are not clearly defined, it may take longer to identify who can and should fix a problem than it does to actually fix the problem.

The architecture of SAP R/3 further necessitates establishing and following procedures. All programs - batch jobs, data migration routines, reporting routines - should be reviewed, tested, and have the appropriate signatures before they are used on the production system. Poorly coded programs can quickly cause system-wide performance problems. Running jobs or reports at the wrong time can have a performance impact on the entire system. It is not unheard of for a single user to cause system-wide performance problems because the same report has been resubmitted several times. Performance problems are considered in this availability document because severe performance problems may be perceived by the end user as a system outage. While performance problems may not directly affect availability measurements, there are procedures and policies that can help avoid or minimize these performance issues.

HARDWARE FAILURE

Time to recover from a hardware failure would include the following:

- Time to identify the failing component
- Time to place a service call
- Time for service representative to respond to the call
- Time for the service representative to obtain replacement parts
- Time for the service representative to replace failed components
- Time to restart the system
- Time to restart the application

Whenever possible, "spare parts" should be stored at or near the production location. If there is a hardware failure, replacing the failed component should not be delayed because the replacement parts were not readily available.

The system architecture should be designed to eliminate single points of failure. The single points of failure in the SAP R/3 architecture (database, enqueue, and message services) should be protected with either HACMP or S/390 Parallel Sysplex as appropriate. In addition, the disks should be mirrored. If a disk fails, the system should use the mirror and the system continue to run. If the disks are mirrored, a single disk failure should be transparent to the end users.

Critical system services (database, enqueue and message services) should be protected with HACMP. A hardware failure on the production server should result in a HACMP failover to

the standby system. This may result in a short outage to the users as the disks are moved to the standby system and SAP R/3 is restarted. SAP recommends stopping and restarting the application servers after a failure of the central instance. If the enqueue service on the central instance is stopped and restarted, all open transactions in the entire R/3 system must be aborted to protect data integrity . If the application servers are not stopped and restarted, the consistency of the database may be endangered. Stopping and starting the application servers will cause the buffers to be flushed which will result in performance degradation until the buffers are reloaded.

Failover time will vary depending on the number of disks that need to be moved over to the failover machine and the database restart/recovery time. If the database and central instance are on different physical machines and both are running on the AIX operating system, both will need to be protected with HACMP.

For a database failover, the DB Reconnect feature available with release 3.0E and greater ensures that all work processes of all R/3 instances are automatically reconnected to the database after the database becomes available again. The end users generally keep their sessions and the interactive input they have made up to the occurrence of the failure. In addition, all local memory buffers remain loaded which avoids the performance impact while the buffers are reloaded. (See OSS notes 24806 and 41678 for information on configuring DB Reconnect).

If the database server is running on the S/390 Parallel Sysplex, the database will be protected from single points of failure by the Parallel Sysplex. Should any hardware or software component fail, the S/390 Parallel Sysplex will dynamically adjust the workload to allow SAP R/3 to continue to run uninterrupted. With DB2 for OS/390 data sharing, customers can spread their transaction workload across multiple DB2 systems, all sharing the same data. Workloads can be shifted between processors during unplanned outages. Access to data remains unaffected. The data sharing feature means systems capacity can be added without interrupting service.

For the OS/390 solution, the central instance will be running on AIX. The central instance should be protected from hardware failure by HACMP.

Network Failure

The network should be designed to eliminate single points of failure. A network design that is currently being evaluated in E/ME/A is described in an appendix to this document.

R/3 SOFTWARE FAILURE

R/3 software failures can be divided into two categories: basis and application. Support from the development team will be needed to help identify software failures in the application area. Procedures should be in place for reporting problems to SAP and tracking problems once they are reported.

DATABASE FAILURE

The database should be monitored on a regular basis. Regular monitoring will reduce or eliminate database failures due to storage problems. The system administrator or database administrator should be able to react to space problems before they cause a system outage.

The duration of an outage caused by a database failure is difficult to predict. If the failure can be isolated to a specific tablespace or data file, a partial restore and recovery can be used. A partial recovery will cause a shorter outage than a full database restore and recovery.

If the database failure is caused by data corruption or by a logical error in the database management system, the problem will take much longer to detect and resolve. It is possible for the system to continue to function after the database has been corrupted. The system may continue to run for several hours or days before the problem is discovered. In this situation, a restore/recovery may not be feasible because other changes to the database would be lost. Fixing a data corruption problem may require in-depth knowledge of R/3 and the database.

A technical assessment of the problem will need to be made after a database failure to determine the severity of the problem and the skill level needed to recover.

When a database problem does occur, there should be clear reporting and resolution procedures. The person who is monitoring the system (e.g. the operator) should know who to contact. There should also be procedures for reporting defects to the database vendor.

POWER OUTAGE

UPS (Uninterrupted Power Supply) and CPS (Continuous Power Supply) should be in place for all production systems to prevent system down time because of power outages. Vital business process SAP R/3 production systems should only be placed in sites which have the ability to continue to run during a power outage. Additional precautions described in the Scheduled Outages section of this document can also be used to prevent unscheduled outages due to disruption of power. All active components of the SAP R/3 system and network must be attached to the UPS/CPS to avoid an outage.

Operating system failure

An operating system failure on the production system running AIX should cause an HACMP failover. From a functional perspective, this will operate the same as it would if there were a hardware error on the production server. The users will notice a short outage while the production workload is moved to the failover system.

For OS/390, the parallel sysplex will help avoid outages caused by some operating system failures. Should the operating system fail on a single processor, a failover will occur to an alternate processor. Transactions in progress will be rolled back, but users will not lose their sessions.

For both AIX and OS/390 operating systems, the most recent stable version should be used. The most recent stable version will probably not be the most recent version.

Software failure

To reduce the likelihood of software failures, a software evaluation and selection process should be established. Criteria other than just price should be evaluated. The evaluation process should include operability, interoperability, available training, documentation and stability of the company that develops and maintains the product.

Prior to installing the software on the production system, it should be installed and tested on the "mirror production"" environment.

Support staff should be trained in the software product, so that the appropriate skills are available when problems occur.

Custom code failure

A phased application development, review, and test process should be established. No custom code changes should be allowed on the production system without proper testing. In the event of a custom code failure, the support staff should know who to contact and the escalation procedures. A phased development plan, review, and testing process should reduce the likelihood of custom code failure. Procedures for handling code failures when they do happen will shorten the duration of the outage.

Add on software failure

Add on software or "bolt-ons", ABAP programs purchased from other companies, should undergo the same testing process and procedures as custom code and have similar problem determination and escalation procedures.

Support server failures (bridges, name servers, etc)

There is more than one way to reduce the failures in this area. One approach is to eliminate as many as possible. Architect the solution such that name servers, bridges, etc. are not needed. This may make managing multiple systems more time consuming, but reduces the complexity level and the skill level needed to resolve problems.

Another approach is to use name servers, bridges, etc. to help reduce the overhead of maintaining the systems, but to make sure that the support staff exists, is properly trained, and accountable for these subsystems.

Print server failures

Set up printer groups. If a single printer is lost, the print job is directed to another printer in the printer group. Make use of external spool programs that store spool jobs and make reprinting easier in the event of a printer failure. This also eliminates the necessity of storing print jobs in R/3 which increases database overhead.

Site Disaster

Disaster Recovery and Business Recovery Planning are not discussed in detail in this document. There should be a definition of what is considered a site disaster. Once a site disaster is declared, there should be clear instructions for who to call and who is responsible for the different areas of the system.

Restores/Recovery

Restore and recovery procedures should be established and thoroughly tested. These tests should be repeated periodically on the "mirror production" systems. The best way to identify problems with recovery procedures is frequent testing. Those people who will be called upon to perform a restore should be very familiar with the process before a problem occurs. Documentation should also be reviewed and updated at predetermined intervals. Contact information must be kept up to date. Periodic restores of backups should be done to the production mirror system to track and monitor the existing backup strategy.

Human Error

The chances of human error can be reduced by establishing procedures and clearly defining roles, responsibilities, and accountability. Each of the roles should be evaluated periodically. Those tasks that are repeated most frequently should be automated when possible. The manual steps for any automated tasks should also be documented.

RECOMMENDATIONS

TECHNICAL PROCEDURES TO INCREASE AVAILABILITY

Common practices and procedures should be implemented in each location. Periodic reviews of practices and procedures should be conducted by a technical team.

SAP offers EarlyWatch Services and GoingLive Check. With the EarlyWatch Service, SAP offers the opportunity to have a periodic system analysis through a remote connection. SAP's R/3 specialists examine the system for potential problems and bottlenecks. The results of the EarlyWatch Service is a written report with recommendations to improve performance and availability.

With the GoingLive Check R/3 specialists inspect the configuration of the individual system components and provide recommendations for system optimization. The main goal of this check is to prepare the R/3 system for production.

ORGANIZATIONAL AND OPERATIONAL PRACTICES

Roles and responsibilities should be clearly defined, as well as accountability and escalation procedures. Contacts for different areas should be documented. These practices should be defined for both on-shift and off-shift.

Once roles and responsibilities are established, reviews of most frequent tasks should be conducted. Whenever possible, automate the most frequent or most time consuming tasks. Whenever a task is automated, the manual process should be documented in case the automatic process fails.

As suggested in the Unscheduled Outages section, a problem log should be maintained. Outages should be reviewed periodically. Improving availability is easier if appropriate data has been collected. If information about the most frequent outages as well as the longest outages is collected, it is easier analyze these outages to reduce or eliminate down time in the future. For example, if the outage is caused by a hardware problem, having "spare parts" readily available will reduce down time.

ARCHITECTURE

All production systems with critical business services should be located in sites which have the ability to ride out a power failure. Both UPS (Uninterrupted Power Supply) and CPS (Continuous Power Supply) should be implemented and tested. .

When architecting a solution, the single points of failure for both hardware and software should be listed and evaluated. Whenever possible, eliminate single points of failure. If eliminating single points of failure is not possible, protect these services with the parallel sysplex or HACMP. Because of the increased complexity introduced when HACMP is installed and configured, HACMP should be used only for stable systems that have requirements for greater than 99.9% availability. See the "General Recommendations" section in "SAP R/3 Architecture - Mapping of System Services" in this document for recommendations for R/3 configuration.

The vendor receiving the best high-availability score by the Gartner Group is the Tandom NonStop. IBM MVS and Digital VMS were both in second place. Each of these platforms has a long history of being optimized for continuous operations. AIX received a lower score than MVS for availability. The MVS solution for R/3 is not yet mature enough to give a recommendation based on availability. The MVS solution for R/3 requires that the central instance and the application servers run on AIX. The central instance is a necessary component for R/3 availability.

PRODUCT AND HARDWARE ENHANCEMENTS

AIX Enhancements

- Remove 256M segment size limitation
- Remove 2G file size limitation
- Remove 10 segment limitation

SAP Enhancements

- Reload R/3 buffers after application server/central instance shutdown.
- Allow dynamic parameter changes. Eliminate requirement to cycle system to activate parameter changes.
- Dynamic load balancing for gateways.
- Shorten downtime needed for release upgrades to SAP R/3
- Supported tools for backup, restore, and recovery of remote third mirror
- Allow application services to run on the OS/390 platform, especially enqueue and message services.

IBM Hardware/Software enhancements

- Make HACMP more robust.
- Build features into IBM storage devices to allow mirroring on hardware level so there is no impact to performance when mirrors are synched
- Add remote third mirror capability for IBM disk
- Add online reorg support to DB2/6000
- Add incremental backup support to DB2/6000

APPENDIX A

OUTLOOK OF R/3 4.0

CCMS Enhancements

The next version of Computing Center Management System (CCMS) is scheduled to be available with R/3 Release 4.0 in the second half of 1998. The next version of CCMS delivers new levels of functionality, automation and openness. Highlights include improved job scheduling and output management, enhanced security features and open object-oriented interfaces. Many of the new features will also be available to current customers using R/3 Release 3.1, allowing them to increase system management functionality immediately without upgrading to R/3 Release 4.0.

With R/3 Release 4.0, CCMS offers system administrators a new, object-oriented environment for monitoring and managing single, distributed and even multiple R/3 installations and components from a single tool. Under the new architecture, any software resource inside or outside R/3 can be monitored with the addition of a "data supplier" program and fed into the CCMS monitoring architecture. If problems are detected, CCMS can handle them automatically or notify operators through e-mail or pagers.

The new monitoring architecture is also well-suited to enabling Web-based system monitoring and management for greater customer convenience and ease of use. SAP is supporting the industry in the definition and adoption of the Common Information Model (CIM).

Based on a new object-oriented design, the next version of CCMS will give customers a wide variety of scheduling options, including time-based, event-based and calendar-based scheduling, as well as the ability to schedule background jobs using external scheduling applications. Through support for job streams, customers define a wide range of job execution criteria and conditions.

CCMS also provides enhancements in its management of complex print jobs spanning multiple locations and multiple countries. New capabilities included advanced load balancing and new interfaces to external output management systems that enable customers to use existing job-scheduling tools, lower staff training costs, and provide greater convenience.

New Upgrade Procedure

R/3 Release 4.0 upgrades are based on a sophisticated timetable and action road map. The timetable includes up front preparation time and time set to test upgrades before going live. New substitution methods minimize downtime during the upgrade. Parts of the upgrade can be performed while the production system is up and running - with no impact on stability or security. Automated procedures are available to modify the production system.

Enhanced planning tool, the newly developed Java-based graphical upgrade assistant, comes with database measuring and data volume estimates and enables remote upgrades. Upgrade road map and planning tools support a precise upgrade process tailored to the specific needs of SAP customers. Remote upgrade service provides customers with an inexpensive and resource-saving alternative to performing the upgrade process in-house.

APPENDIX B

END-TO-END SYSTEM AVAILABILITY

IBM Global Services has developed a monitoring tool called AMA (Application Monitoring and Alerting) which was initially implemented to monitor Lotus Notes Servers (a sample AMA screen has been provided as an attachment to this document SAMPLE.PRE). The tool has been modified to target SAP systems consistent with the Notes process. The tool is PC-based and measures the availability and response time of a specific function module, end to end. The tool runs on a dedicated workstation placed on a LAN segment (specified by the customer) with other SAP users at a particular site. It is important that the tool be located on the same LAN segment that the end-users are on to give the best possible view of the environment. The tool wakes up at predefined intervals and executes the SAP function module (RFC_SYSTEM_INFO) and then adds the measurement to a AMA central collection database. If any response exceeds set limits, or problems occur accessing the system (network or database down), the tool has the ability to request assistance via a pager to a contact person established ahead of time.

The tool was developed to work on all internal SAP projects regardless of the release of SAP or database management system. The function module is "noninvasive" (no altering of the production database in any way). The end to end measurements of this function module will provide a clearer view of the time required for the function module to travel from the client workstation, over the network to the SAP system, and return to the client workstation. Minimal data will be transmitted with this function module so that network and SAP production system loads are not negatively affected.

In the future, this study will be expanded to include end-to-end availability.

APPENDIX C

E/ME/A NETWORK SOLUTION

To be provided by E/ME/A

APPENDIX D

AVAILABILITY CHECKLIST

Recommendation	**Yes/No**
Implemented Continuous Power Supply	
Implemented Universal Power Supply	
All disks mirrored or RAID	
Implemented HACMP for Central Instance	
Implemented HACMP or Sysplex for Database Server	
Have production mirror system	
Mapped R/3 services to duplicate dialog, gateway, spool, update, an batch services on multiple machines	
Implemented login groups for application servers	
Implemented regular backup schedule	
Tested recovery procedures	
Tested disaster recovery procedures	
Batch Administrator to monitor and manage batch jobs and data migrations	
Defined systems management procedures	
Defined escalation procedures	
Periodically review systems management procedures and practices	
Defined and documented roles and responsibilities	