Benchmark Report

Retek*0 Demand Forecasting 4.1*

Creation Date: 01/08/00 2:25 PM
Last Update: 01/13/00
Version: 0.1

Document Information

Approvals

Name	Company	Date	Signature

Change Record

Date	Name	Version	Change reference
08-Jan-00	Dave Vranicar	0.1	Revised with references to Customer deleted
10-Jan-00	Hale Brown	0.2	Proofread

Reviewers

Name	Company	Date	Signature
Dina Davies	IBM	13-Jan-00	

Related Documents

Document	Author
RDF 4.1 Benchmark Report	Retek

Table of Contents

Executive Summary

Inventory replenishment is the process of providing the right product at the right place at the right time. To have all products in stock at all times is not technically difficult, but it is prohibitively expensive. To replenish inventory profitably, a business must weight the costs of stock outages against the costs of holding inventory and of ordering more often.

Advanced forecasting techniques can dramatically improve the accuracy of replenishment decisions. But to preserve profit, businesses must also balance the costs of achieving forecast accuracy against the benefits of doing so. For the vast majority of everyday consumer goods, the savings to be attained on an individual item at a store from one week to the next are counted in cents rather than dollars. (The exceptions are fashion items, innovative products such as consumer electronics, or otherwise highly differentiated products with relatively high price tags and margins.) For most others, the profit potential becomes more obvious only when one multiplies weekly savings across many items, many weeks and many stores. Very large retailers may be able to save tens or even hundreds of millions of dollars through better forecasting and replenishment.

Yet by virtue of the tens of thousands of items they often carry across hundreds or thousands of stores, large retailers must be preoccupied with generating forecasts efficiently. If forecasting systems cannot generate large numbers of forecasts in an acceptable time, retailers must compromise. For example, they may use older data to get around bottlenecks in processing the most recent point-of-sales data. Or they may generate forecasts for fewer products and stores. Both compromises reduce forecast accuracy. By doing so they also reduce potential profit that could be gained through more accurate replenishment.

One large Retek customer believed it could improve profit by tens of millions of dollars a year through increased revenue and reduced inventory cost. In evaluating software to help them achieve such savings, the company's managers identified two primary requirements for a forecasting system:

(1) It must use advanced forecasting methods to generate the highest level of forecast accuracy that the retailer's data would allow.

(2) It must generate forecasts fast enough to allow:

   (a) the system to use the most recent sales history and

   (b) forecasts to be generated for all SKUs in all stores where the retailer's inventory managers determine it is beneficial to do so.

While the retailer's business managers felt confident that Retek Demand Forecasting (RDF) could provide accurate forecasts, its IT managers were concerned that RDF may not be able to generate the volume of forecasts their improved business processes would require. To address their concerns, Retek designed a benchmark test to prove that RDF can process the retailer's data within the time available.

In generating data for the benchmark, Retek compiled sales history from a cross section of retailers. It modified these data to approximate the data volume and density of the client for whom Retek performed the test. Retek then grouped the test data into "domains" or logical subsets that would allow processing in manageable increments. The test included the following processing operations:

- data hierarchies
- loading of sales transactions
- batch forecasting and
- forecast export.

In each case the benchmark tested worst-case scenarios, using the most advanced--and time consuming--forecasting algorithms available within RDF (i.e., AutoES[1]). Such scenarios were not intended to duplicate the production system the retailer would eventually put in place. Rather, they were designed to define an upper limit for the run times the retailer could expect after its final system design had been determined.[2] In the event RDF could not run within an expected 8-hour batch window, additional experimentation would be performed to determine limiting assumptions that would allow the batch process to finish within the allotted time.

Retek performed the benchmark in conjunction with IBM on an RS/6000 - S80 machine, running with 24 processors and 64 GB of RAM. The benchmark tested the volume of 5,000 stores to accommodate current needs as well as future growth of the chain. Retek set the target RDF forecast volume at about 60 million SKU/Store combinations. Assuming a SKU/Store/Week sales density of about 15%, about 20 million non-zero time series would be forecast in the test scenarios.

The parameters and results from the primary worst-case scenario of the benchmark appear in the following table:

Parameters	Scenario 1
SKU/Store Combinations	62,680,800
Active SKU/Stores	20,735,820

[1] It should be remembered that running AutoES entails optimizing over a number of different forecasting models and selecting the best one. Creating a forecast for a given time series actually encompasses the creation of multiple forecasts, plus the calculation of the cumulative prediction intervals used to provide appropriate safety stock to support demand. Each of these algorithms have been coded in an optimal manner on the software side to minimize the number of cycles required from the hardware.

[2]

Domains	60
Data Density	15.68%
Forecast Horizon	13
Batch Process	**Processing Time (mins)**
Loading of Hierarchies	13
Loading of Sales Data	27
Generation of Forecasts	56
Exporting of Forecasts	38
Total	**134**

The total batch processing time was 2 hours and 14 minutes. Generation of forecasts required only 56 minutes of this total, with the remaining time being spent in loading and exporting data. Retek's test team believe the total run time fits well within any batch processing window the retailer's future environment would be likely to impose. The efficiency of processing leaves this retailer with the option of considering many alternative implementation strategies, with little regard for constraints on batch run times.

In addition to the primary scenario, Retek ran additional scenarios to test how different parameter settings would affect run times. The highlights of the findings and their implications appear below:

- Effect of Changes to Forecast Horizon – In addition to the baseline scenario requiring a forecast horizon of 13 weeks, Retek tested performance with a forecast horizon of 20 weeks. Data import and export times remained constant, with forecast generation increasing from 56 to 67 minutes.

  Implication: This result should minimize concerns about the effects on total processing time of extending the forecast horizon or increasing the granularity of forecasts (e.g., exporting forecasts at the daily level).

- Effect of Changes in Data Density – Retek tested four scenarios with data densities that varied from 5% to 10% to 15% and to 20%. In the the first two Retek calculated 167,000 non-zero SKU/Store forecasts; in the last two, 21 million.

  Implication: While run times were significantly different for each of the four scenarios (79, 95, 146, 167 minutes, respectively), run times appear to be more strongly correlated to the number of active SKU/Store forecasts than to data density.

- Effect of Number of Active SKU/Stores and Data Density – Retek tested the effect of dramatically increasing the number of Active SKU/Store combinations over levels tested in prior scenarios. Two new scenarios included about 58 million active SKU/Store combinations each, with data densities of 14% and 29%. Run times were 294 and 336 minutes (5 and 6.5 hours), respectively. (This compares to 2 hours for the baseline scenario.)

  Implications: Run times appear to increase linearly with number of Active SKU/Store combinations, almost independently of data density.

- Effect of Number of SKU/Stores – While the density of each domain (1 million SKU/Stores) remained constant, Retek tested the effect of reducing total data from 60 domains to 40 and then to 20. Run times, as expected, descended in linear increments from 146 minutes to 97 to 56.

Implication: The number of active SKU/Stores appears to be the most important factor affecting batch run times. If the retailer finds that its data density in production is lower or higher than the test estimate of 15%, batch run times should not be significantly affected

- Effect of Number of Domains – The baseline test scenario had 60 domains of 1 million SKU/Store combinations per domain. Retek also tested the effect on performance of larger domain partitions. New scenarios tested 44 domains with 1.4 million SKU/Stores each and 32 domains with1.9 million SKU/Stores each. Run times were nearly identical for all three scenarios (134 min. for 60 domains, 113 min. for 44, and 124 min. for 32).

  Implication: The number of domains should not significantly affect batch performance. Organizations can create domains to increase functional efficiency without signigicant compromise to batch processing performance.

- Effect of Number of Processors – Retek tested two additional scenarious to determine the effects of the number of processors on run time. The scenarios subsequently reduced the number of processors from 24 to 18 and then to 12. Run times were as follows:

Processors	12	18	24
Total Batch Run Time (Import+Forecast+Export)	215 min.	154 min.	134 min.
Total Batch Run Time (Time Series/Processor/Second)	136	126	109
Forecasting Run Time	81 min.	60 min.	56 min.
Forecasting Run Time (Time Series/Processor/Second)	360	324	260

  Implication: Increased processing power clearly increases processing rates. Yet processing speed does not increase at a linear rate with the number of processors used, suggesting that the efficiency of each processor decreases as more are used.

- The Effect of Source Forecasting Level – When RDF peforms source-level forecasting, it uses the most resource-demanding forecast algorithms (i.e., AutoES) only at higher levels in the product/store hierarchies (for example Subclass/Region). RDF generates lower-level SKU/Store forecasts using Simple Exponential Smoothing or Croston's Intermittent Method, both of which run much faster than AutoES. As test scenarios reduced the number of "complex" forecasts from 21 million to 3 million, batch run times for forecasts declined from 56 minutes to 43. Overall batch performance declined by the same seven-minute increment, from 134 to 121 minutes. This occurred because the processes for loading and exporting data were the same in both cases.

Implication: Use of source-level forecasting should not significantly affect forecast batch run times.

- Effect of Execution Approach – Because the number of test domains was larger than the number of processors (by nearly a 3-to-1 margin) the processing of domains could be sequenced in a variety of ways. For example, the domains could be processed either simultaneously (in parallel) or sequentially (in series). Retek tested three approaches:
  1. Two domains from each file system ran in parallel (i.e., a total of 40 domains), with the third domain of every file system launched as soon as one of the first two domains had finished. Retek named this approach {1,2}{3}.
  2. The three domains ran in parallel. Retek called this approach {1,2,3}.
  3. The three domains ran in series. Retek called this approach {1}{2}{3}.

Run times for these three approaches were 134 minutes (baseline), 122 minutes (parallel), and 156 minutes (series).

Implication: A parallel approach appears to offer greatest processing efficiency. But because performance varied by type of batch job performed (data loading vs forecast execution), the retailer should perform additional tests on its production system and production data.

0The rest of this document provides more detail on both the study's methods and its results.

Introduction

This document summarizes the results of the RDF 4.1 performance benchmark that Retek performed for one large U.S. retailer. It details the objectives, the assumptions, the results, and the conclusions drawn from the benchmark.

Objectives

The purpose of the benchmark test was to reduce the retailer's risk in choosing RDF as its primary engine for generating accurate demand forecasts. The specific objectives were as follows:

- To prove that RDF 4.1 could process the retailer's volume of data for demand forecasting within an acceptable time and
- To measure its processing performance so that the retailer and potential hardware vendors could specify computer configurations that can complete processing within the time available.

Hardware Configuration

[This sections needs to be reviewed by IBM to make sure all the information is correct. Most of this information was taken from the IBM web site.]

Retek performed the benchmark on an IBM RS/6000 model S80. It is a 64-bit symmetric multiprocessor system. Details of the configuration appear in the following table:

RS/6000 S80	
Maker	IBM
Model	RS/6000 - S80
Number of Processors	24
Processor Type	450 MHz PowerPC RS 64 III
RAM	64 GB
Hard Disk	1 TB (to be confirmed by IBM)
Operating System	AIX 4.3.3

Appendix A describes the disk layout and the mappings of RDF domains to the hard disks.

Benchmark Scope

This section defines the scope of the benchmark. It identifies the processes evaluated and the characteristics of the data used.

RDF Processes

The benchmark will measure performance for the following RDF processes:

- **Hierload**: batch processing of hierarchy loading into RDF from the Retek Merchandising system (RMS).
- **Dataload**: batch processing of weekly sales loading into RDF from RMS.
- **Gen_frcsts**: batch processing of forecast generation and automatic forecast approval.
- **Export_frcsts**: batch processing of data export from RDF to RMS.

Data Volumes

The benchmark tested data for 5,000 stores. Assuming roughly 12,000 SKUs to be forecast per store, the target RDF benchmark volume was 60M SKU/Store combinations.

Data Density

Data density is the percentage of SKU/Store combinations that register at least one sale a week. It defines the percentage of the total 60M SKU/Store combinations for which forecasts should be run. Typical retail data densities are around 5% or less. Retek assumed data density of 15% in order to test worst-case scenarios, even though this number seemed high from Retek's experience with other retail clients.

Active SKU/Store Combinations

A SKU/Store combination is considered "active" if it has sales history and requires a forecast to be generated. In Retek's experience with other retailers only about one-third of total SKU/Store combinations in the RDF database are active. This is because not all SKUs are sold in all stores, and not all SKU/Store combinations require forecasts to be generated. Retek's systems require forecasts only for SKUs that use the dynamic replenishment method. Consistent with its experience with other retailers, Retek set the target for this benchmark at one-third of total SKU/Store combinations. The test would generate one-third of 60M forecasts, or about 20M.

RDF Configuration & Settings

The Retek benchmark team defined the following standard configuration:

- Four sales measures: Regular, Promotional, Clearance and Total Sales. It loaded only Regular and Total Sales.

- Two location attributes: Store opening and closing dates.
- One final and seven source forecasting levels.

To be sure the benchmark would address the retailer's real-world needs, Retek tested worst-case scenarios for the forecasting parameters that affect the forecast-generation batch process. The following table presents Retek's settings for forecasting parameters. An asterisk indicates each worst-case assumption.

Parameter	Final Level	Source Level
Forecasting Method	Simple/Croston	AutoES*
Source Level	SKU/Store*	N/A
Training Window	All* (110 weeks)	All* (110 weeks)
Forecast Horizon	13 Weeks	N/A
Cumulative Interval	Yes*	N/A
Export Forecasts	All*	N/A

Some of the foregoing parameters changed during this benchmark to measure their effect on performance.

Benchmark Results

The results of the benchmark appear in this section as execution times for the different batch processes.

Retek changed several key parameters from the baseline (that is, the most likely) scenario to measure their effect on execution times:

- Number of active SKU/Store combinations (i.e., number of SKU/Store combinations for which a forecast is generated)
- Number of RDF domains to be run simultaneously on a given server
- Data density
- Number of processors dedicated to RDF batch runs
- Forecast Horizon

By analyzing the effects of selectively changing these parameters, a retailer should be able to estimate accurate batch run-times under different circumstances.

Baseline Scenario

The following tables summarize the characteristics of the most likely or "baseline scenario" and the provide execution times for the Retek batch processes:

Parameters	Scenario 1
SKU/Store Combinations	62,680,800
Active SKU/Stores	20,735,820
Domains	60
Data Density	15.68
Forecast Horizon	13
Batch Process	**Time (mins)**
Hierload	13
Dataload	27
Gen_frcsts	56
Export_frcsts	38
Total	**134**

Effect of Forecast Horizon

Increasing the forecast horizon from 13 to 20 weeks produced the following results:

Parameters	Scenario 1	Scenario 2
SKU/Store Combinations	62,680,800	62,680,800
Active SKU/Stores	20,735,820	20,735,820
Domains	60	60
Data Density	15.68	15.68
Forecast Horizon	**13**	**20**
Batch Process	**Time (mins)**	**Time (mins)**
Hierload	13	12
Dataload	27	29
Gen_frcsts	56	67
Export_frcsts	38	39
Total	**134**	**146**

Effect of Data Density

The benchmark tested data densities of 5%, 10%, 15% and 20%. Results appear in the following table and graph:

Parameters	Scenario 3	Scenario 4	Scenario 2	Scenario 5
SKU/Store Combinations	62,680,800	62,680,800	62,680,800	62,680,800
Active SKU/Stores	167,263	167,263	20,735,820	20,735,820
Domains	60	60	60	60
Data Density	**5.25**	**10.66**	**15.68**	**19.84**
Forecast Horizon	20	20	20	20
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	15	16	12	26
Dataload	10	19	29	36

Gen_frcsts	33	39	67	66
Export_frcsts	21	21	39	38
Total	**79**	**95**	**146**	**167**

Effect of Number of Active SKU/Stores and Data Density

Retek tested extreme cases where almost all SKU/Store Combinations are active. Scenarios 6 and 7 have data densities of 15% and 30%, respectively. The following table presents the results.

Parameters	**Scenario 1**	**Scenario 6**	**Scenario 7**
SKU/Store Combinations	62,680,800	62,680,800	62,680,800
Active SKU/Stores	**20,735,820**	**57,767,880**	**57,812,160**
Domains	60	60	60
Data Density	**15.68**	**14.20**	**29.35**
Forecast Horizon	13	13	13
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	13	19	18
Dataload	27	25	68
Gen_frcsts	56	134	137
Export_frcsts	38	115	113
Total	**134**	**294**	**336**

Effect of Number of Domains

All the preceding scenarios tested 60 domains with about 1 million SKU/Store combinations each. Scenarios 8 and 9 varied the number of domains with the following results:

Parameters	**Scenario 1**	**Scenario 8**	**Scenario 9**
SKU/Store Combinations	62,680,800	60,058,768	60,687,972
SKU/Stores per domain	**1,044,680**	**1,364,972**	**1,896,496**
Active SKU/Stores	20,735,820	19,849,984	19,985,728
Domains	**60**	**44**	**32**
Data Density	15.68	15,68	15,64
Forecast Horizon	13	13	13
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	13	8	9
Dataload	27	26	26
Gen_frcsts	56	45	53
Export_frcsts	38	33	35
Total	**134**	**113**	**124**

Effect of Number of SKU/Stores

With the average number of SKU/Stores per domain held constant at around 1 million, Retek reduced the total number of SKU/Stores to about 40 million and 20 million in Scenarios 10 and 11, respectively:

Parameters	Scenario 2	Scenario 10	Scenario 11
SKU/Store Combinations	**62,680,800**	**41,787,200**	**20,893,600**
Active SKU/Stores	20,735,820	13,823,880	6,911,940
Domains	**60**	**40**	**20**
Data Density	15.68	15,68	15.68
Forecast Horizon	20	20	20
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	12	7	5
Dataload	29	19	12
Gen_frcsts	67	48	25
Export_frcsts	39	24	15
Total	**146**	**97**	**56**

Effect of Number of Processors

In scenarios 12 and 13 Retek reduced the number of processors from 24 to 18 and then to 12:

Parameters	Scenario 1	Scenario 12	Scenario 13
SKU/Store Combinations	62,680,800	62,680,800	62,680,800
Active SKU/Stores	20,735,820	20,735,820	20,735,820
Domains	60	60	60
Data Density	15.68	15.68	15.68
Forecast Horizon	13	13	13
Processors	**24**	**18**	**12**
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	13	13	17
Dataload	27	34	49
Gen_frcsts	56	60	81
Export_frcsts	38	47	68
Total	**134**	**154**	**215**

Effect of Source Forecasting Level

For all of the foregoing scenarios, Retek generated forecasts at SKU/Store level because to impose the greatest processing demands on the RDF system. Scenario 14 tested processing efficiency under less demanding circumstances, with the source level changed to Sub-Class/Store.

Parameters	Scenario 1	Scenario 14
SKU/Store Combinations	62,680,800	62,680,800
Active SKU/Stores	20,735,820	20,735,820
Source Level Forecasts	**20,735,820**	**2,671,740**
Domains	60	60
Data Density	15.68	15.68
Forecast Horizon	13	13
Batch Process	**Time (mins)**	**Time (mins)**
Hierload	13	13
Dataload	27	27

Gen_frcsts	56	43
Export_frcsts	38	38
Total	**134**	**121**

Effect of Execution Approach

In all the foregoing scenarios, each file system contained three domains for a total of 60 domains. In scenario 1 Retek launched parallel processing for two domains from each file system (i.e., a total of 40 domains). As soon as one of the first two domains had finished, Retek launched processing for the third domain of every file system. Retek named this approach {1,2}{3}. Scenarios 15 and 16 tested parallel processing {1,2,3} and serial processing {1}{2}{3}, respectively.

Parameters	**Scenario 1**	**Scenario 15**	**Scenario 16**
SKU/Store Combinations	62,680,800	62,680,800	62,680,800
Active SKU/Stores	20,735,820	20,735,820	20,735,820
Domains	60	60	60
Data Density	15.68	15.68	15.68
Forecast Horizon	13	13	13
Execution Approach	**{1,2}{3}**	**{1,2,3}**	**{1}{2}{3}**
Batch Process	**Time (mins)**	**Time (mins)**	**Time (mins)**
Hierload	13	20	24
Dataload	27	27	35
Gen_frcsts	56	40	53
Export_frcsts	38	35	44
Total	**134**	**122**	**156**

Appendix: Disk Layout and Domain Mapping

Disk layout for the benchmarks was designed to maximize input/output (I/O) throughput to optimize the performance of RDF. It was designed based on the following information:

- Data for 5,000 stores on an Acumate database divided into 64 domains.
- Each domain needs 12G space. 16G is used for calculation. Total database space 1024G.
- Each domain database consists of Sales, Market Parameters, Forecast and Other DB files. They reside in one subdirectory (logically).
- Each database file (e.g. Sales.gem) is a JFS file in a filesystem.
- A total of 192 disks (hdisk 8 .. Hdisk 199) available, this gives a total space of 3264G
- Disk cluster : 8 disks per cluster, this number may change for performance reason

- Number of domains per disk cluster : 3
- Each LV size is 16*3=48G, which contains three domains
- Number of clusters needed: 64/3 = 22 clusters
- 2 clusters not in use, available for other purpose

The following table summarizes the mapping of the domains to the disks. Figure A-1 graphically describes the layout for the S80 machine.

SSA Loop	Dom 01-03	Dom 04-05	Dom 07-09	Dom 10-12	Dom 13-15	Dom 16-18		Dom 19-21	Dom 22-24	Dom 25-27	Dom 28-30	Dom 31-33
1	8	9	10	11	12	13	14	15	16	17	18	19
2	20	21	22	23	24	25	26	27	28	29	30	31
3	32	33	34	35	36	37	38	39	40	41	42	43
4	44	45	46	47	48	49	50	51	52	53	54	55
5	56	57	58	59	60	61	62	63	64	65	66	67
6	68	69	70	71	72	73	74	75	76	77	78	79
7	80	81	82	83	84	85	86	87	88	89	90	91
8	92	93	94	95	96	97	98	99	100	101	102	103

SSA Loop	Dom 34-36	Dom 37-39	Dom 40-42	Dom 43-45	Dom 46-48	Dom 49-51		Dom 52-54	Dom 55-57	Dom 58-60	Dom 61-63	Dom 64
9	104	105	106	107	108	109	110	111	112	113	114	115
10	116	117	118	119	120	121	122	123	124	125	126	127
11	128	129	130	131	132	133	134	135	136	137	138	139
12	140	141	142	143	144	145	146	147	148	149	150	151
13	152	153	154	155	156	157	158	159	160	161	162	163
14	164	165	166	167	168	169	170	171	172	173	174	175
15	176	177	178	179	180	181	182	183	184	185	186	187
16	188	189	190	191	192	193	194	195	196	197	198	199