



# Best practices

## Expanding an IBM Smart Analytics System

Garrett Fitzsimons

*IBM Data Warehouse Best Practices Specialist*

James Cho

*Senior Technical Staff Member, IBM Smart Analytics System Core Engineering*

Issued: December 2011

Expanding an IBM Smart Analytics System .....	1
Executive Summary .....	3
Introduction .....	4
Concepts of an IBM Smart Analytics System.....	5
How to expand a system (and when to replace).....	7
What modules can I add? .....	7
Expanding and redesigning for new applications and users .....	8
Recognizing the need to expand.....	10
Documenting Service Level Objectives for comparison.....	11
Gathering data for capacity planning .....	11
Understanding the capacity planning indicators .....	15
Identifying the user module as a bottleneck .....	18
Options for a mixed-generation environment.....	19
Models based on System x and Linux platform .....	19
Models based on Power Systems and AIX platform .....	20
Failover modules and high availability groups.....	20
Planning for an expansion .....	22
Considerations when planning for an expansion .....	22
Major milestones for an expansion project.....	23
Case study: Mixed generation 7600 and 7700.....	27
Case study: Mixed generation D5100v4, 5600 R1 and 5600 R2 .....	29
Conclusion .....	32
Appendix A. Query performance.....	33
Appendix B. Linux command examples .....	34
Further reading.....	35
Contributors.....	36
Notices .....	37
Trademarks .....	38

## Executive Summary

Data warehouse environments continue to experience an explosion in data growth. Growth is also occurring in the number of concurrent users and queries against data warehouses as tactical and operational applications gain a broader audience. As a result, you might find yourself in a situation where you need added capacity to cope with increased enterprise demands. To help you meet these demands, the IBM® Smart Analytics System is a flexible data warehousing solution that supports a building block approach to expansion.

This paper is for anyone involved in planning or implementing an expansion of data warehouses built on DB2® for Linux®, UNIX, and Windows® software. The examples in this paper apply to the IBM® Smart Analytics System based on IBM System x® and IBM Power Systems™ servers. The principles discussed apply generally to systems that are similar to the IBM Smart Analytics System. After reading this document, you will be able to plan for growth in demand for capacity of all the resources in an IBM Smart Analytics System. The intent is not to provide you with a step-by-step upgrade guide, but to cover all the milestone steps involved in expanding and growing an IBM Smart Analytics System. The paper looks at releases up to and including the IBM Smart Analytics System 7700 R1.1 and the IBM Smart Analytics System 5600 R2.

## Introduction

This paper describes best practices in planning for growth in demand for resources and adding capacity to your IBM Smart Analytics System environment. It covers how to introduce capacity planning into your daily operations, what considerations to take, and what the available options are.

The principles discussed apply to IBM Smart Analytics Systems in general, although the primary example is for an IBM Smart Analytics System 7700 R1.1. Examples and references to other models are given where appropriate. In some cases, examples of commands apply to the AIX operating system only; you can find the equivalent commands for the Linux operating system in appendix B.

These best practices are of greater value to you if you have a working knowledge of an IBM Smart Analytics System and of DB2 software, including partitioned database environments. If you have questions about the specific details for your own system, engage your IBM team.

The first section of the paper deals with the concepts behind the IBM Smart Analytics System. The second section of the paper examines system expansion versus system replacement and summarizes the different expansion options. The third section details how to gather and understand performance data for capacity planning purposes in order to recognize when your system needs to be expanded. The fourth section provides a checklist of the expansion options available for both System x and Power Systems based systems and how a high-availability environment can be expanded. The fifth section reviews the key points in planning for an expansion project. Finally, case studies specific to System x® and Power Systems™ servers systems are provided.

This paper builds on and complements a series of best practice papers that examine various aspects of the IBM Smart Analytics System. For more information, see the section "[Further reading](#)".

Although the same principles apply, adding an application or business intelligence module to your IBM Smart Analytics System is outside the scope of this paper.

## Concepts of an IBM Smart Analytics System

The IBM Smart Analytics System is a workload-optimized solution for warehousing that is preintegrated and preconfigured according to best practice guidelines for data warehousing. The system consists of modules that together provide the environment needed to support a data warehouse.

The IBM Smart Analytics System stack undergoes extensive quality testing to help ensure that the system is stable and that all software, firmware, and hardware components function correctly together. Upgrades to this stack are tested and published on a regular basis and support levels are guaranteed on this basis.



**Maintain software and firmware levels at published stack levels and plan upgrades with IBM Lab Services.**

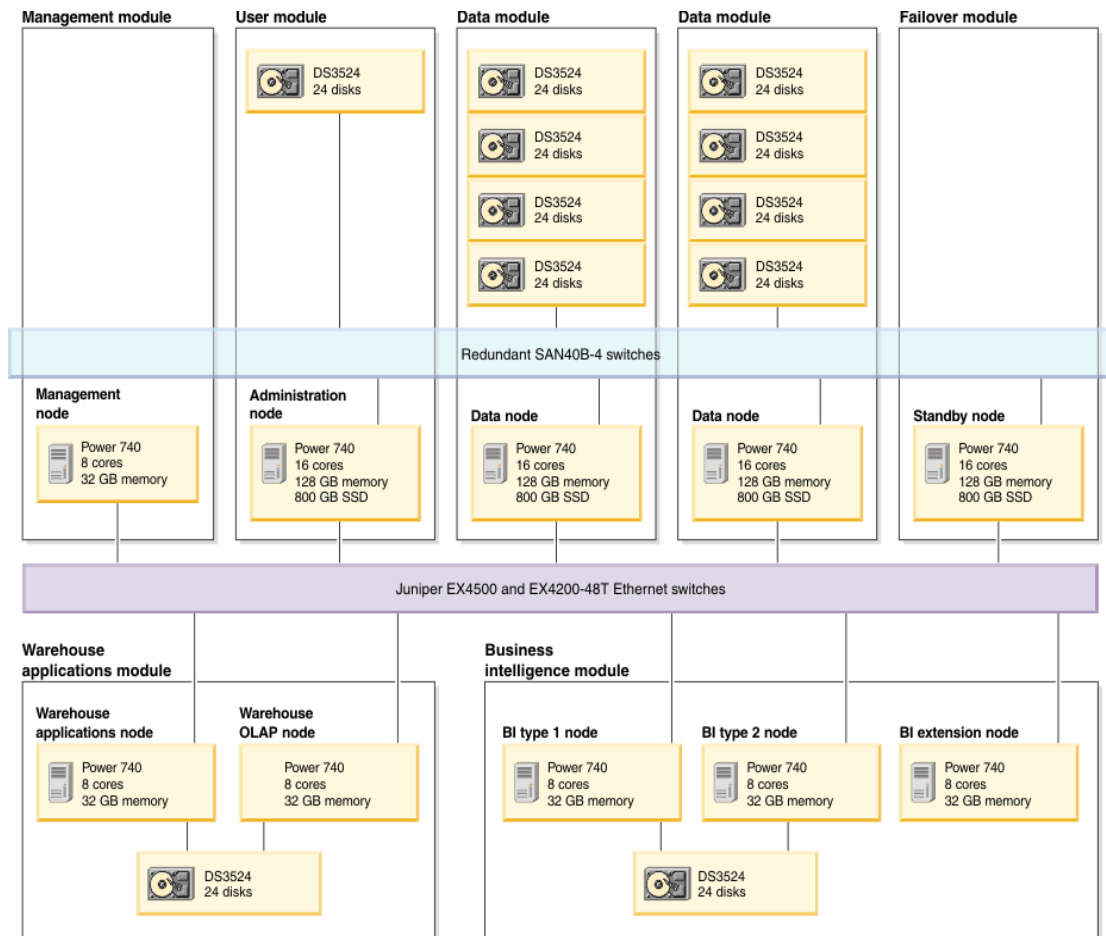


Figure 1. Base components of an IBM Smart Analytics System 7700 R1.1

Upgrades to software and firmware outside of the published stack are handled on a case by case basis. Each system is provided with a standard 3-year support contract, which

covers hardware and licensing. Extended warranties can be acquired on a case by case basis by submission through a special bid to IBM.



**Standard hardware support is three years. Take this timeline into account when planning for expansion to avoid the need to purchase extended support.**

The IBM Smart Analytics System illustrated in Figure 1 shows the management, user, data, and failover modules that make up the system. In this case, the system is an IBM Smart Analytics System 7700 R1.1. These modules can be augmented with application and business intelligence modules. Storage and Ethernet switches provide the network capacity needed. The concept of data modules with the same amount of storage, memory, and processors is consistent for all IBM Smart Analytics System releases.

Each IBM Smart Analytics System has at least one user module that contains the administration node. The administration node performs the function of maintaining the system catalog, coordinating data access queries and managing single database partition group objects.

A typical data warehouse database contains a combination of current and historical data that satisfies various queries. Approximately 30% to 40% of data in a data warehouse is considered active because it tends to be queried more often than historical data.

Deviations from this approximation of active data in turn require more or less system resources in your environment.

## How to expand a system (and when to replace)

An IBM Smart Analytics System can be expanded to accommodate more data and more users by adding a data or user module to the system. If you experience incremental growth in demand for resources, adding data or user modules is the recommended approach.

You can perform rolling upgrades and replace older-generation modules with newer-generation modules over time. Rolling upgrades ensure that the overall solution remains within the standard 3-year support contract. A benefit to rolling upgrades is that advances in technology provide hardware and software with a greater capacity to perform existing queries faster than before. Each IBM Smart Analytics System generation of module takes advantage of new hardware and software releases to accommodate more concurrent queries and larger data warehouses.

Before you decide to expand or replace your system, ensure that your existing system is fully optimized by:

- Optimizing queries.
- Performing prudent database maintenance.
- Implementing DB2 Workload Manager (DB2 WLM).

An IBM Smart Analytics System might also exhibit signs of excessive workload, which can happen when the workload exceeds what the system was originally sized for. Excessive workload can indicate that a system replacement rather than an expansion might be needed. Consider the number of user connections, the number of concurrent queries, the volume of active data, and the storage capacity needed to cater for data growth. If you need to add a significant volume of data, users, applications, or queries, then a replacement system rather than a system expansion might be more appropriate.

### *What modules can I add?*

An existing system can be expanded by adding a data or user module:

- **Data module**  
Adding a data module increases the storage capacity available to the database. Existing data is redistributed across each new data node added and each database partition has proportionately less data. This addition can help increase performance. However, when data growth continues and storage capacity usage returns to levels seen before the expansion, performance can return to the original levels.
- **User module**  
Adding a user module increases the capacity of the system to handle user connections and reduces the network traffic to the existing user module. A user module mirrors the configuration of the administration node in that a single database partition is created and added to the partitioned database. An additional database partition group is created and tables in database partition

group SDPG are replicated from the administration database partition to the user node.



**Add a data module to expand storage capacity and reduce data volume per database partition. Add a user module to increase the capacity of the system to accommodate users.**

It is also possible to add a standby module to introduce high availability into your environment. You can also improve high availability by adding additional standby modules, as the standby configuration is typically smaller than the primary configuration. Mixed-generation high availability (HA) environments, where multiple product generations of user and data modules co-exist, might require an additional standby module to remain highly available.

### *Expanding and redesigning for new applications and users*

A successful data warehouse implementation often leads to an increase in the number of users and applications over time. A transition towards an Enterprise Data Warehouse (EDW) that supports multiple lines of business applications can soon follow.

You can expand an IBM Smart Analytics System to accommodate multiple, related applications within the same database where separate data sets need to interact with each other. However, do not use a single IBM Smart Analytics System to support more than one database or share infrastructure across multiple, disparate data warehouses. Disparate workloads often can become difficult to manage effectively. Instead, implement multiple systems.

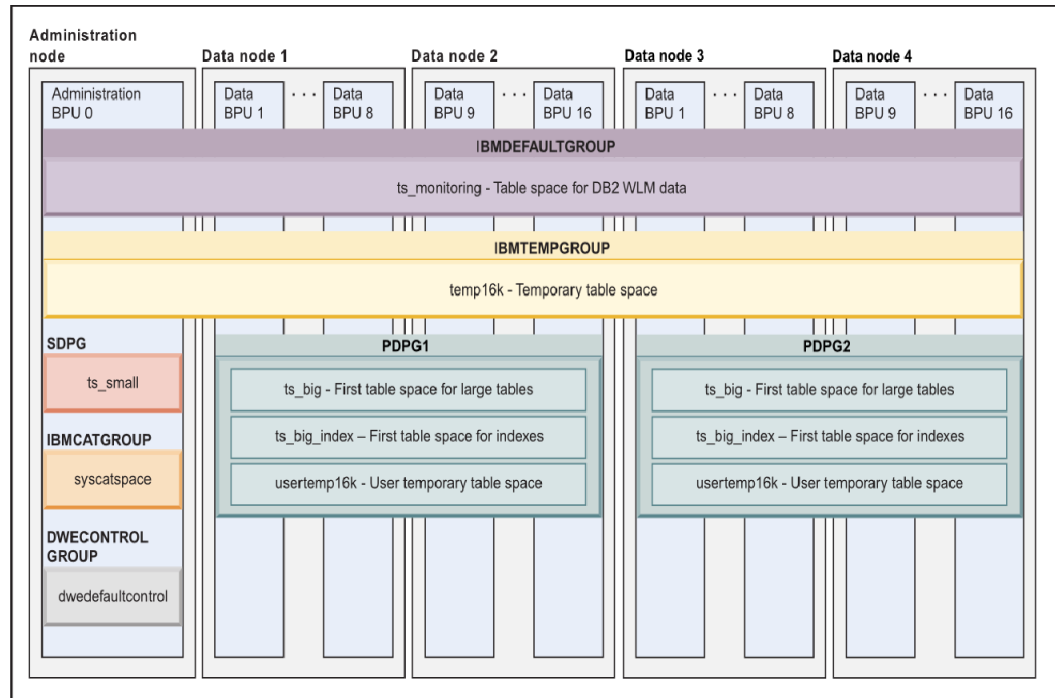


Figure 2. Example of multiple database partition groups in an IBM Smart Analytics System



To separate application data and associated queries to a subset of the database partitions in the partitioned database, use database partition groups. Database partition groups are used to group a number of database partitions together. Table spaces, when created, must be assigned to a database partition group and tables within the table space are partitioned across the database partition group. Figure 2 illustrates how the partitioned database partition groups PDPG1 and PDPG2 separate table spaces and tables into groups of data nodes within one database.

The difference in this type of expansion is that the existing partitioned tables are not expanded across each new data node. Instead, a new database partition group is created across each new data node and new table spaces are created in the new database partition group. In this way, the data on the new and existing data nodes are kept separate from each other. Separate schemas are also recommended to further isolate the data. Note that colocated queries can occur only between tables in the same database partition group.

The advantages of separating application data and associated queries are that you can:

- Build on your existing infrastructure.
- Minimize the cost associated with a new application.
- Soft-isolate the two data sets for performance, maintenance, security, expansion and chargeback purposes.
- Minimize time to production.
- Control the amount of data and data growth per database partition group.

An additional user module can be installed to handle the user connections associated with the new data modules. The existing administration node continues to hold the database catalog.

## Recognizing the need to expand

The timeline between recognizing the need to expand and completing a successful expansion is crucial for the continued operation and availability of your data warehouse. You must monitor, identify, and understand the capacity indicators associated with increased resource usage well in advance of the need to address resource issues.



**Good capacity planning practices can assist in the early detection of trends in resource usage.**

Monitor server, storage, and network statistics on a frequent basis. You can determine resource capacity issues by comparing performance statistics over time against existing and projected performance baselines. Begin to plan for expansion of your IBM Smart Analytics System well before resource capacity is not expected to meet your service level objectives.

The reasons for expansion might be explained through company mergers and acquisitions. Or, you might need to accommodate new data sources, new applications, more users, changes in data retention, compliance or regulatory requirements. In system terms, these requirements can be categorized as:

- **More users**

Adding more users translates into an increase in user connections, concurrent queries, and resource usage on the user node.

- **More queries**

An increase in the number of concurrent queries through new applications or more users increases the demand on resources across each data node.

- **More data**

Predicted or unpredicted data growth that exceeds the existing storage capacity of the system.

Recognizing the need to expand your IBM Smart Analytics System can be determined by:

- Documenting service level objectives for comparison.
- Gathering data for capacity planning.
- Understanding the capacity planning indicators.

## *Documenting Service Level Objectives for comparison*

Service level objectives (SLO) for the query workload include how much time certain categories of queries (simple, medium, and complex) can take for each application. SLOs can also be documented for extract, transfer, and load operations (ETL), backup and other operations.

Create a performance baseline for individual workloads and for the system as a whole. Each baseline must show what resources are used when the system is at nominal and at peak usage. Review the contents of individual workload baseline operations on a regular basis. This review ensures that your baseline continues to reflect the queries and operations that take place in your environment. Periodically check that your workload baseline accurately reflects SLOs in place.



**Create and document a performance baseline for each workload as well as overall system usage and a forecast performance baseline for the next 12 months.**

Determine and document service level objectives and create a current and projected 12 month performance baselines for each SLO. Doing so helps you determine when your system needs to be expanded. Include:

- Average and maximum time for backup operations to complete.
- Average and maximum time for ETL operations to complete.
- Average and maximum response time for a baseline set of queries by application or business unit.

## *Gathering data for capacity planning*

Capacity planning involves identifying the resources needed to satisfy service level objectives for workloads into the future. Consider the following when choosing which data to gather:

- Determine the capacity of resources in your environment at the outset so that you can compare the capacity used against the capacity available. This information does not need to be too granular, because regular alerting and housekeeping tasks can handle day to day operational activities.
- Gather data on a daily basis and collate the output for the current month, and for a rolling 12 months. Doing so enables you to compare your new data against your performance baseline and can help identify trends in resource usage.
- Use software such as IBM InfoSphere Optim™ Performance Manager, IBM DB2® Performance Expert and tools including `nmon`, `vmstat`, `topas`, `sar` and others to gather statistics on resource usage. Use whichever software or tool you are most comfortable with to gather data. The examples in this document are designed to illustrate how statistics can be gathered and understood rather than the exact commands or tools to use.

- Parse and aggregate the data gathered separately for user, data, and standby nodes where applicable to avoid skew when reviewing data. Note where a standby node has failed over and might cause skew in your results. Fail back nodes as soon as possible.
- Incorporate user feedback such as perceived changes in workload performance into the capacity planning process.

## Backup operations and other workloads

Determine the average and maximum backup operation time and collate the data on a monthly basis. The duration of backup and other maintenance operations can increase as the volume of data per data node to be processed increases.

The following example queries the DB\_HISTORY view to determine average and maximum backup times:

```
SELECT AVG(TIMESTAMPDIFF(4,CHAR(TIMESTAMP(END_TIME) -
TIMESTAMP(START_TIME)))) AS Avg_Backup_Time,
MAX(TIMESTAMPDIFF(4,CHAR(TIMESTAMP(END_TIME) -
TIMESTAMP(START_TIME)))) AS Max_Backup_Time
FROM SYSIBMADM.DB_HISTORY
WHERE OPERATION = 'B' AND OPERATIONTYPE = 'F'
```

Collate similar data for ETL operations by using the metadata associated with your ETL application.

## Storage subsystem

The data warehouse table space containers are on the external storage subsystem and have a defined capacity for each model of the IBM Smart Analytics System. A `/db2fs/<instname>/NODEnnnn` or similar file system is created on each user and data node. Refer to your IBM Smart Analytics System User Guide for the exact details of your environment.

Record the storage capacity of each file system together with the minimum, average, and maximum percentage used for each database partition.

Use the following example command to return disk usage statistics in gigabytes for each database partition entry in the `db2nodes.cfg` file, where `bcuaix` is the instance name:

```
rah "df /db2fs/bcuaix/NODE*"
```

## Table space pages used

Record storage subsystem resource usage at the database level in addition to the operating system level. You want to record usage at both levels, because table space containers indicate only the amount of disk space reserved for the database. Space reserved for the database does not necessarily indicate the amount of actual data that resides in each container. Record the total size of all table spaces together with the maximum and average amount of data pages used.

Use the following query to return total space assigned to each table space per database partition and the space used daily.

```
SELECT DBPARTITIONNUM, TBSPACE,
       SUM(BIGINT(TBSP_TOTAL_PAGES*PAGESIZE))
       /1024/1024/1024   GB_TOTAL,
       SUM(BIGINT(TBSP_USED_PAGES*PAGESIZE))
       /1024/1024/1024  GB_USED
FROM SYSIBMADM.TBSP_UTILIZATION U, SYSCAT.TABLESPACES T
WHERE U.TBSP_NAME = T.TBSPACE
GROUP BY DBPARTITIONNUM, TBSPACE
ORDER BY DBPARTITIONNUM, TBSPACE WITH UR
```

## Processor, memory, and I/O statistics

Record processor, memory, and I/O statistics on a daily basis for each user and data node. Use the `vmstat` command as shown to capture processor and memory metrics for 30 consecutive seconds. Reissue the command every 30 minutes throughout the day. This helps ensure that an effective representation of system performance is captured.

```
vmstat -It 1 30
```

Capture input and output statistics for each user and data node. Use the following example command to collect data:

```
iostat -s -Dlt 1 30
```

## User connections

Record the number of user connections for each user node at intervals of 30 minutes. For more detailed analysis, group users according to business unit or other logical grouping. Use the following example to return a count of the number of applications connected:

```
SELECT COUNT(*) FROM SYSIBMADM. MON_CONNECTION_SUMMARY
```

## Concurrent queries

Record the number of concurrent queries in the database every 30 minutes on a daily basis. Use the following query to determine the number of `SELECT` statements that are executing in your environment:

```
SELECT COUNT(*) FROM SYSIBMADM.MON_CURRENT_SQL
```

The `MON_CURRENT_SQL` administrative view returns key metrics for all activities that were submitted on all members of the database and that have not yet completed. These metrics provide a point-in-time view of currently executing static and dynamic SQL statements in the currently connected database.

## Network

Focus on the DB2 Fast Communications Network (FCM) when collecting network data for capacity planning. The FCM is important, because it is dedicated to database

application traffic. Gather statistics for the other networks, such as management, LAN-free, and corporate networks, as required.

Gather statistics every 30 minutes using the following `netstat` command as an example, where `en11` represents the FCM network interface name. The `netstat` command produces cumulative figures since network startup, which you can examine.

```
netstat -s -I en11
```

## *Understanding the capacity planning indicators*

Understand how resources are used and where resource usage is trending with respect to your performance baseline and performance forecast. Understanding resource usage trends in the context of overall capacity can help you plan for expansion in a timely manner.



**Align service level objectives with capacity planning indicators so metrics can be easily compared.**

When reviewing data:

- Distinguish between user and data node groups to avoid skewing results.
- Know your workload schedules and where changes take place.
- Look at data in the context of seasonal or financial calendar cycles.

## **Storage subsystem capacity usage**

Typically, the amount of data per database partition increases during the lifecycle of the data warehouse. Query performance is affected by an increase in the volume of data that each database partition needs to process. Adding a data node can reduce the volume of data each database partition needs to process. Follow these guidelines to recognize when storage capacity might need to be expanded:

- Use the data gathered to predict when storage capacity used within each database partition associated file system reaches 60% and 80%.
- Plan for expansion when storage reaches 60% capacity and the predicted rate of growth shows that 80% capacity will be used within 12 months. This window gives sufficient time to plan and implement a successful expansion project before storage utilization reaches 100%.



**Begin planning for storage expansion when storage capacity reaches 60% and is projected to reach 80% within 12 months.**

When a data module is added to an IBM Smart Analytics System, existing data needs to be redistributed. You redistribute data to each new database partition so that each database partition has an equal amount of data to process. The redistribution process uses the temporary table space `temp space` during processing to sort and rebuild objects. In some situation, a small amount of additional `temp space` might be required on each data node to complete these tasks.

Differences in file system usage across database partitions might indicate data skew or a file system that is also being used for some other purpose.

## Server resources and usage

Review capacity indicators in the context of the node type and in the context of the performance baseline for each group of node types. Capacity indicators for each node group must be reviewed in the context of processor, memory, I/O, network, user connections, and concurrent queries.

### Processor usage

Use the following data output by the `vmstat` command to understand processor usage:

- `%sy` (processor usage by system processes)
- `%us` (processor usage by user processes)
- `%r` (run queue number)

High processor usage in itself is not an indicator of a resource capacity issue. The IBM Smart Analytics Systems module is designed to run at high processor utilization rates even at low concurrency. High concurrency levels can result in higher run queue numbers. Run queue metrics must be understood in the context of your performance baseline and when service level objectives might be compromised.

Where concurrency increases significantly over time, in line with processor usage and run queue numbers, consider implementing workload management with the DB2 workload manager (WLM).

Where total processor usage, represented by `%sy + %us`, is greater than 80% on a consistent basis and the run queue number is more than eight times the number of CPU cores, the system could be constrained by resources available.

Look to establish trends in these indicators over time to help identify where the system is exhibiting stress on a consistent basis. Review the output in the context of your performance baseline, service level objectives and in the context of data, user, and application growth.

### Memory usage

Use the following data output by the `vmstat` command to understand memory usage:

- `%avm` (available virtual memory)
- `%fvm` (free virtual memory)
- `%wa` (CPU waiting)
- `%fi` (paging in)
- `%fo` (paging out)

High memory usage can result in excessive paging which in turns affects processor usage and can result in a degradation of performance. Excessive paging can indicate a larger than expected number of concurrent users.



Compare available virtual memory (%avm) with free virtual memory (%fvm) against baseline figures to understand usage and trend towards capacity:

- If the %fvm value is low or progressively trending lower against your performance baseline, then your system might be experiencing capacity-related performance degradation.
- A decrease in the %avm value can indicate an increase in caching. An increase in caching exists, if the values for %wa, %fi, and %fo are trending higher on a consistent basis when compared to baseline values.

### I/O usage

Use the following output from the `iostat` command to understand I/O usage:

- %wa (io waiting)
- %avg sqsz (average service queue size)
- %sqfull (indicates the number of times the service queue is full per second on AIX only)

Look for a sustained increase in input and output operations against your performance baseline over time. Determine an increase in I/O operations by:

- Comparing %wa trends.
- Comparing service queue values against baseline to determine the trend on I/O performance. When the service queue is full across all nodes, no I/O can occur. This is a rare occurrence. Where %avg sqsz increases over time, it indicates growth in resource usage.

### Network

Network capacity indicators can help determine where a user node is network-bound or where significant database traffic is occurring that is not in line with expectations.

Use the following metrics output from the `netstat` utility. The metrics produced by `netstat` are cumulative and can be reset by the `root` user.

- %ipkts (Input packets for network interface)
- %ierrs (Input errors for network interface)
- %opkts (Output packets for network interface)
- %oerrs (Output errors for network interface)
- %coll (Collision rate)

If %coll is greater than two percent of %opkts and trending higher against your performance baseline, then network traffic is considered very high.

Zero values for packets and non-zero values for errors might indicate hardware failures or no traffic. These values must be assessed as part of normal operational problem determination tasks. Ensure that %ipkts and %opkts have a positive, non-zero value.

If a user node is network-bound, it may result in query performance degradation. The Fast Communication Network (FCM) network is primarily used for internal database traffic. An increase in FCM traffic without an increase in concurrent queries might indicate unnecessary traffic caused by database or query design. Refer to the “Query optimization in a data warehouse” best practices paper to learn more about query optimization, collocated queries and data skew.

### **User connections and concurrent queries**

Each user module coordinates data access queries, collates the results returned from each database partition and passes the result set back to the user connection. Follow these guidelines to recognize when capacity might need to be expanded:

- Observe growth in concurrent user connections and concurrent queries on each user node and predict when capacity will meet or exceed initial sizing estimates.
- Look for sustained increases in the traffic over the FCM network. These increases could indicate that a user node is a bottleneck.

Where more than one user module is implemented, ensure that performance is balanced across both user modules to maximize the resources available.

Implement DB2 WLM to manage workloads to help ensure that each query acquires processor time in relation to the priority of the query.

### ***Identifying the user module as a bottleneck***

Where data nodes are not at capacity and are performing as expected, your system might be constrained by the throughput from the administration node. This bottleneck might be due to many factors, such as a high number of large single partition tables or replicated tables. An additional user module or a replacement of the existing user module with a new generation module might alleviate the symptoms. Newer generation modules provide greater network, processor, and memory resources and can help provide greater throughput in circumstances where the data modules are not fully used.

When replacing an existing user module with a new generation user module, the following requirements must also be considered:

- Standby module for the new generation module
- New HA Group
- New generation storage subsystem
- New generation Ethernet switch
- Option to repurpose existing user module as a data module

## Options for a mixed-generation environment

As new generations of the IBM Smart Analytics System module are introduced, older models are no longer manufactured and might not be available for purchase. The IBM Smart Analytics System design typically delivers next generation modules with greater capacity and improved performance. Each generation of module also provides an expansion path for previous generation modules, protecting your existing investment. A mixed generation environment is one where multiple generations of the user and data module co-exist in the same cluster that supports the data warehouse.

As of the publication of this paper, the most recent and supported models are listed in the following sections. Storage values for each data node are approximate and based on default disk size configurations.

### *Models based on System x and Linux platform*

The modules listed in Table 1 are supported in a mixed generation environment on IBM Smart Analytics Systems based on System x and the Linux platform. Older models on this platform not listed here are supported on a case by case basis. You can contact IBM Support for more information about expanding older systems.

Model	Cores	Storage	Database partition	Key features of data node
D5100 V4 May 2009	4	3 TB	4	<ul style="list-style-type: none"> <li>✓ X3650 server, DS3400 storage</li> <li>✓ Major software / driver / firmware refresh, Minor hardware update.</li> </ul>
5600 R1 Mar 2010	4 8(SSD)	6 TB	4	<ul style="list-style-type: none"> <li>✓ X3650 M2 server, DS3400 storage</li> <li>✓ New SSD option for core warehouse with additional x5570 processor</li> <li>✓ New business intelligence module option</li> <li>✓ Unified network definition</li> <li>✓ Major software / driver / firmware refresh, Minor hardware update</li> </ul>
5600 R2 Nov 2010	6 12 (SSD)	12 TB	8	<ul style="list-style-type: none"> <li>✓ Complete hardware / software / driver / firmware refresh</li> <li>✓ X3650 M3 server, DS3524 storage</li> <li>✓ Additional X5680 6 core processor for SSD</li> <li>✓ Twice the capacity and performance of 5600 R1 module.</li> <li>✓ 10 GbE Network, J48E Switches.</li> <li>✓ SAN zoning</li> <li>✓ DB2 logging design and new file system layout</li> <li>✓ HA configuration updates</li> <li>✓ Active – active warehouse application modules</li> </ul>

*Table 1. IBM Smart Analytics System Models based on System x and Linux platform*

### Models based on Power Systems and AIX platform

The models listed in Table 2 are supported in a mixed generation environment on IBM Smart Analytics Systems based on Power Systems and the AIX platform. Older models on this platform not listed here are supported on a case by case basis. You can contact IBM Support for more information about expanding older systems.

Model	#Cores	Storage	Database Partitions	Key features of data node
7600 R1 Sep 2009	4 (option for 8)	7.2 TB	4	<ul style="list-style-type: none"> <li>✓ Power 550 processor</li> <li>✓ Option to add additional two processors</li> <li>✓ Improved data compression</li> <li>✓ New business intelligence module</li> <li>✓ Hardware refresh</li> </ul>
7600 R1.2 Jun 2010	4 (option for 8)	7.2 TB	4	<ul style="list-style-type: none"> <li>✓ IBM Smart Analytics System Control Console</li> <li>✓ Automated software and firmware maintenance</li> </ul>
7700 R1.1 Nov 2010	16	21.6 TB	8	<ul style="list-style-type: none"> <li>✓ Power 740 processor</li> <li>✓ Introduced solid-state disk technology to house TEMP table space and enhance database performance.</li> <li>✓ Reduced floor space</li> <li>✓ Unified network definition</li> <li>✓ 10x faster network than 7600 R1.2</li> <li>✓ 2x less expensive per TB capacity than 7600 R1.2 module</li> <li>✓ Redundant DB2 logging design</li> </ul>

Table 2. IBM Smart Analytics System Models based on Power Systems and AIX platform

### Failover modules and high availability groups

When the high availability (HA) configuration shown in Figure 3 is implemented, nodes are arranged into HA groups. Each group consists of up to four active nodes for Power Systems and five active nodes for System X environments. Each group also consists of administration or data nodes and one failover module which contains the standby node. Each server must be identical to the other servers in the HA group. For example, a single HA group cannot include both Power 550 and Power 740 servers.

#### HA group configuration

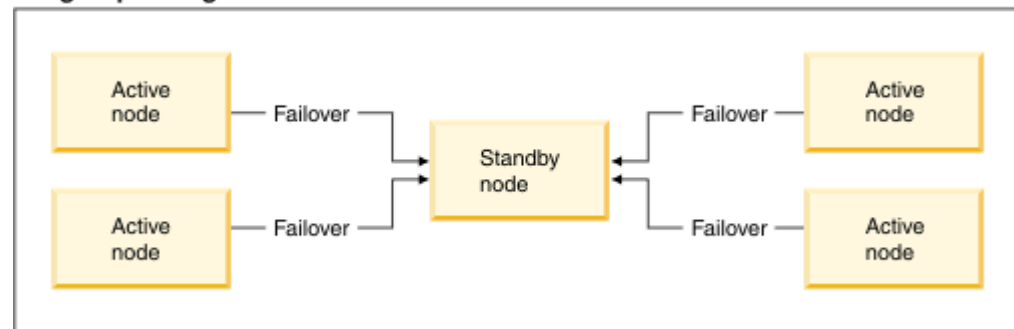


Figure 3. High availability configuration in an IBM Smart Analytics System 7700 R1.1



Each HA group must consist of identical physical servers. Different server types require separate HA groups.

The failover module does not contain dedicated external storage. This module is connected to SAN switches so that it can access the storage subsystem that is allocated to the other modules within the HA group. IBM Tivoli® System Automation for Multiplatform software is used to automate failover.

### High Availability Groups for System x

Use the following table to determine whether a separate HA group might be needed in your expansion project. For example, the 5600 R1 and 5600 R2 modules use different server types and so cannot be placed in the same HA group.

HA Compatible	D5100 V4	5600 R1	5600 R2
D5100 V4	Y	Y	N
5600 R1	Y	Y	N
5600 R2	N	N	Y

Table 3. High availability compatibility matrix for x based IBM Smart Analytics System

### High Availability Groups for Power Systems

Use the following table to determine whether a separate HA group might be needed in your expansion project. For example, the 7600 R1 and 7700 R1.1 modules use different server types and so cannot be placed in the same HA group.

HA Compatible	7600 R1	7600 R1.2	7700 R1.1
7600 R1	Y	Y	N
7600 R1.2	Y	Y	N
7700 R1.1	N	N	Y

Table 4. High availability compatibility matrix for Power Systems



Introducing a new data module does not increase the performance capacity of other data modules. Reducing the volume of data per database partition can help increase overall performance.

## Planning for an expansion

The IBM Smart Analytics System is supported as a preintegrated and preconfigured solution. Although you are free to change the system to suit the changing needs of your business environment, doing so might invalidate your support agreement. To avoid invalidating your support agreement, engage with IBM first. Where you want to remain within the support agreement, ensure that any changes are approved by the IBM Smart Analytics System Change Control Board first. Your IBM contact can assist in this process. You can engage with IBM through the normal channels associated with an IBM Smart Analytics System.

The expansion of the IBM Smart Analytics System can be performed by the IBM Lab Services team. The IBM Lab Services team offers a range of services around the IBM Smart Analytics System, including the data redistribution process.

### *Considerations when planning for an expansion*

Consider the following recommendations when planning to expand your system with same or next generation modules:



**An expanded system might need additional cooling, floor space, network capacity, and backup infrastructure.**

- Mixed generation nodes with different servers cannot be part of the same HA group. In this scenario, you need to implement a new HA group and a new standby module.
- Some features of new generation modules such as solid-state disks (SSD) on the 7700 R1.1, might not be supported on older modules. As a result, you cannot use the new features on the new generation module until all modules support SSD.
- Additional Ethernet switches, SAN switches, and storage rack might be required. Use spare ports where enough are available. Avoid, where possible, implementing both SMC and Juniper switches in an environment.
- Additional cooling, power, network, and floor capacity might be required.
- File systems need to be aligned across all nodes using GPFS or NFS file systems. For example, /db2home, the instance owner home directory needs to be shared across the partitioned database, including each new node.
- Software levels must be consistent across all nodes. A stack upgrade of the existing system might be required in order to be compatible with the new generation module.
- The backup and recovery infrastructure and procedures might need to be expanded to maintain service level objectives.

- Data needs to be redistributed across each new data module and this process will take a period of time to complete.
- Existing applications, for example ETL applications, and administration scripts, need to be modified to include new modules and need to be tested.
- Advances in disk technology mean that newer generations of storage subsystems have greater storage capacity. This capacity can result in some redundant disk space available to the database on the newer modules.

### *Major milestones for an expansion project*

The tasks involved in an expansion project depend on the existing environment and the modules to be added. Many tasks can take place in advance of the database expansion and with no outage needed. Outages are advised for the redistribution process and where some software or firmware upgrades are needed. Any outage window needed depends on the nature of the upgrade and environment. The major milestones are as follows:

- Pre-upgrade planning
- Integration of new hardware
- Stack upgrade of software and firmware levels
- Updating high availability processes
- Expanding the database
- Redistributing data
- Completing the expansion process

### **Pre-upgrade planning**

These tasks can take place without the need for an outage. The reason for these steps is to reduce risk and to avoid an unnecessary outage during the database expansion.

- **Performance baseline**

Standard best practice is to have a performance baseline for a bucket of queries that broadly represent the query workload. Include a copy of the explain plan for each query included. This information can be used for post-redistribution analysis.

- **Health check**

A health check might be advised by IBM to determine whether any symptoms exist that might cause an issue or delay during the expansion project. For example, if there is significant data skew or if software or firmware levels exist that are outside of the published and validated stack. In addition, cabling, port availability, software, and firmware versions needs to be assessed.

- **Verify site requirements**

Expanding a system requires additional flooring, cooling, power, and connectivity. You can work with IBM to understand site requirements.

- **Redistribution pre-check**

The data warehouse database can be assessed before the redistribution phase to determine the characteristics of the database. You can also check whether all database objects are in a nominal state and what the estimated processing time is for the redistribution process.

- **Auxiliary prerequisite data**

Some additional data around component level configuration settings is gathered including SAN storage and operating system kernel levels. The IBM team can assist with these tasks during the expansion process if required.

- **Acquire new licenses**

Where additional licenses are required, they can be obtained before the upgrade project. An example is SAN switch licenses.

## Integration of new hardware

Integrate new rack, server, storage, storage, switch, and other components into the existing infrastructure. IBM Lab Services can provide this service as part of the *New Node Installation Services* offering.

Typically, the new hardware can be positioned without affecting the availability of the existing hardware, dependent on your environment.

The following components, where used, need to be updated to reflect the new hardware:

- Hardware Management Console (HMC) for Power Systems and the AIX platform
- Remote Support Manager for Storage (RSM)
- Systems Director
- IBM Smart Analytics System Control Console

## Stack upgrade of software and firmware levels

Software and firmware levels are tested, published, and supported as a stack on a regular basis. See the section "Further reading" for details. It is recommended that you adhere to a published stack to remain within the support agreement.

New modules will be configured with the most recent published stack. You must upgrade your existing system to this stack before adding the new module. The stack upgrade might involve an upgrade of DB2 software and you must incorporate appropriate testing into your expansion plan.

Where a mixed generation environment is planned, the following applies:

- Stay with the latest firmware recommended for each generation module. For example, you have an IBM Smart Analytics System 7600 and are adding an IBM



Smart Analytics 7700 module. In this case, stay with 7600 recommended firmware for the 7600 module and the recommended firmware for the 7700 module.

- For HMC and RSM modules used in a mixed generation environment, implement the higher level from the 7600 and 7700 validated stack.
- For the AIX platform, DB2 software, GPFS, and TSA implement the higher level from the 7600 and 7700 validated stack.

## Updating high availability processes

Where high availability (HA) is in place, you need to modify the existing failover and failback processes. You need to include new modules, including any new HA groups. You also need to update the HA test cases worksheet documentation to incorporate the new modules and you must complete the individual failover and failback tasks.



**Incorporate each new module and HA group into your failover and failback documentation and test.**

## Expanding the database

When you add a data module or user module, the partitioned database must be expanded to include the new database partitions on each new node. You need to expand the partitioned database before data can be redistributed. You can update the *db2nodes.cfg* file to include the new database partition by issuing the `db2start dbpartitionnum` command for each new database partition. This process does not redistribute any data to the new database partitions and no outage is required.

Include the new database partitions on each new node as soon as the new hardware is installed. This way, you can address any problems that might arise well in advance of the redistribution process. Ensure that a full backup is taken.

## Redistributing data

When adding a data module, existing data must be redistributed to the new database partitions. The data redistribution processing time depends on the volume and type of objects and the number of existing and new data modules. You can reduce the complexity of the operation by planning for an offline redistribution. You can also engage the IBM Lab Services team to perform this task. There are three steps in the redistribution process:

### 1. Pre-redistribution checks and tasks

You can use the `precheck` tool, available through IBM Support, as part of the expansion project to help in validating that the database is in nominal condition before executing the `REDISTRIBUTE` command.

In addition to the database backup ensure that you have a backup of the output from the following commands:

```
db2 get db cfg -- database configuration
db2 get dbm cfg -- database manager configuration
db2look -e -- gather output for DDL
db2look -wlm -- gather output for WLM configuration
db2look -xd -- gather output for authorizations
```

## 2. Redistribution process

Issue the REDISTRIBUTE command in offline mode to ensure that the process completes as quickly as possible. In line with best practice recommendations for all system upgrades, complete a full backup before and after running the REDISTRIBUTE command. By completing a full backup, you have a RESTORE point that refers to the original data nodes in the event of hardware failure.

## 3. Post-redistribution checks and tasks

You can use the `postcheck` tool, available through IBM Support, to help assess the system after the redistribution process is complete. You might need to perform additional tasks to update statistics and compression dictionaries in accordance with baseline query performance and associated query plans.

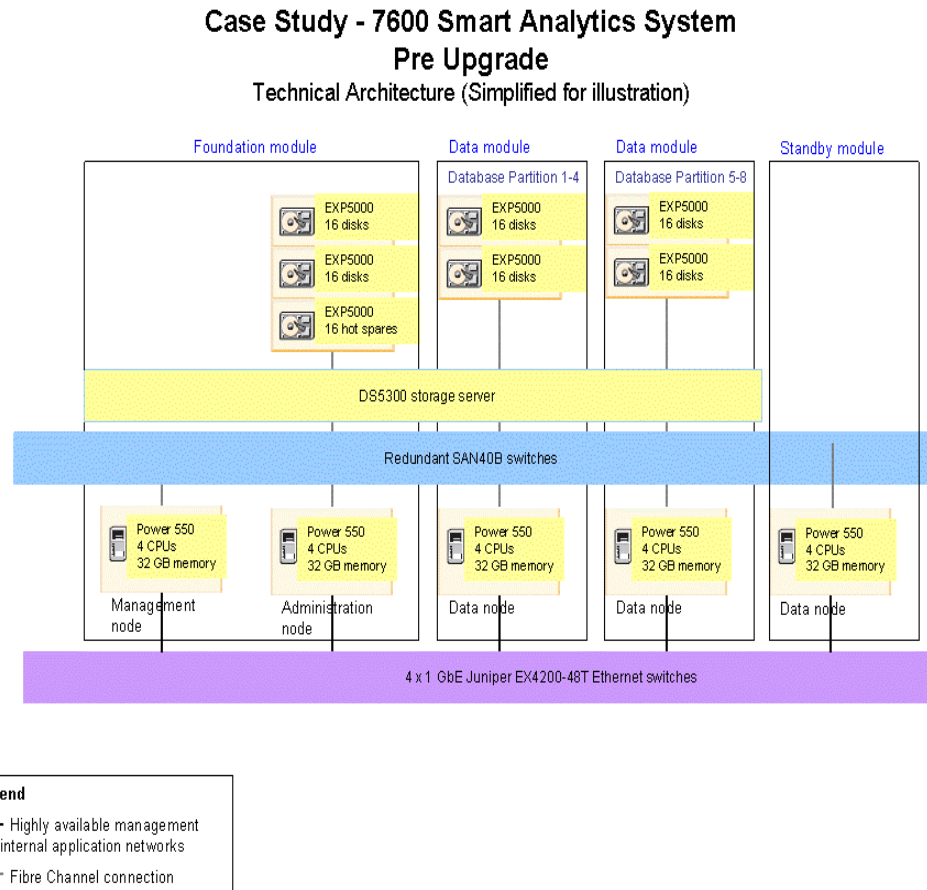
## Completing the expansion process

When you have completed the expansion tasks, the system is expanded and ready for acceptance. Acceptance typically incorporates the following tasks:

- ✓ Performance test and reconciliation against baseline.
- ✓ Installation record and operations documentation updated.
- ✓ Backup operation successful and backup/recovery processes updated.
- ✓ ETL and extract processes test complete.
- ✓ Fail over and fail back testing complete.

## Case study: Mixed generation 7600 and 7700

This case study looks at an existing 7600 configuration, illustrated in Figure 4. The expansion project adds a 7700 R1.1 data module to increase storage capacity.



*Figure 4. Pre-upgrade environment showing a standard IBM Smart Analytics System 7600*

The characteristics of this case study were:

- Storage capacity reached 70% and it was predicted that the remaining capacity would be consumed within 12 months.
- Outages were to be minimized during the expansion project
- The need to replace the entire system was ruled out
- The mixed generation IBM Smart Analytics System environment is supported
- Future capacity growth requirements can be accommodated because further expansion is possible
- Performance levels were maintained.

The chosen expansion, illustrated in Figure 5, involved:

- Adding a 7700 R1.1 data module, storage subsystem, and SAN switches to the existing 7600 R1.2 environment to take advantage of 2xStorage and CPU processing capability.
- Adding a Juniper switch with the capacity to add a module at a future point.
- Adding a storage rack positioned for rolling replacement of older modules with newer modules as needed.

The following upgrade tasks were completed:

- Support agreement approved for mixed-generation 7600/7700 IBM Smart Analytics System.
- All software levels upgraded and matched across expanded cluster in line with the published validated software stack for 7700 R1.1.
- 7600 (FCM) Juniper switches uplinked from 1GbE to 10GbE.
- HMC upgraded to support p740 servers.
- Existing data redistributed across each new database partition.
- Installation record updated to reflect new environment.
- System health check approved and performance acceptance tests completed.

### Case Study – 7600/7700 Smart Analytics System Post Upgrade

Technical Architecture (Simplified for illustration)

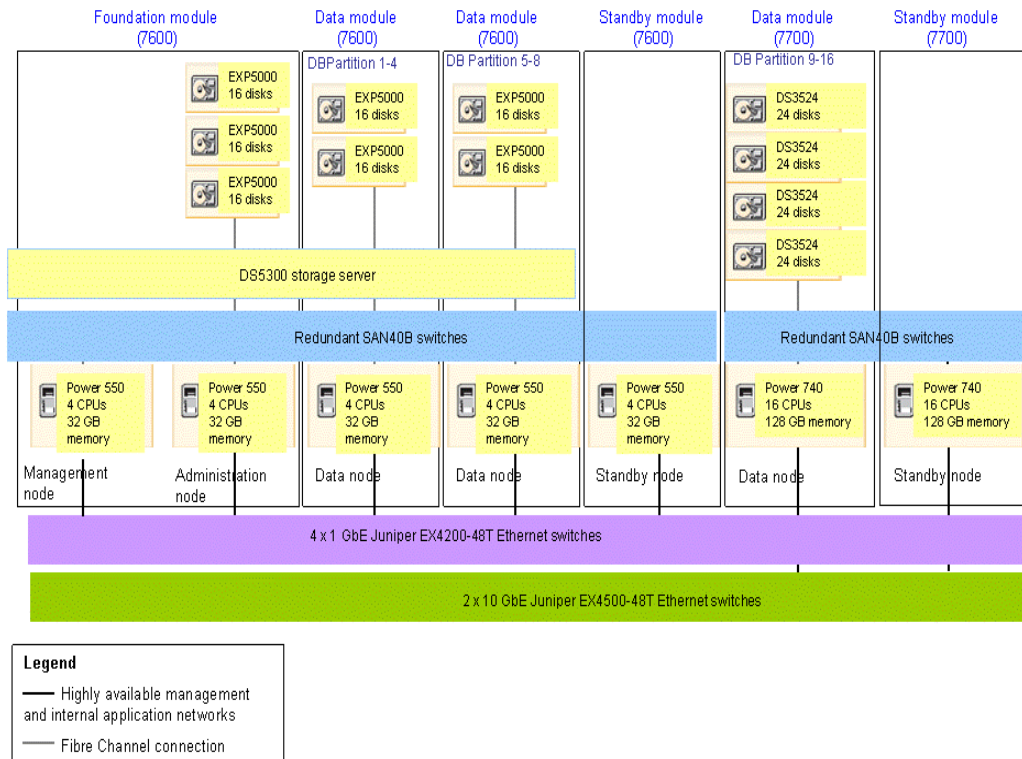


Figure 5. Post-upgrade environment showing supported 7600/7700 IBM Smart Analytics System

## Case study: Mixed generation D5100v4, 5600 R1 and 5600 R2

This case study outlines an expansion that accommodated three generations of System x based modules. The existing system, illustrated in Figure 6, included D5100 v4 and 5600 R1 modules based on the x3650 server. The module to be added was a 5600 R2, which is based on the x3650 M3 server.

The characteristics of this case study were:

- Storage capacity reached 80% and it was predicted that the remaining capacity would be consumed within nine months. In addition, new data sources could not be introduced into the data warehouse until after the expansion.

The goals of the case study were:

- To add data capacity to accommodate predicted data growth.
- Continued support of a mixed generation IBM Smart Analytics System environment.
- To support future capacity growth requirements.

### Case Study – D5100v4 / 5600R1 Smart Analytics System Pre Upgrade Technical Architecture (Simplified for illustration)

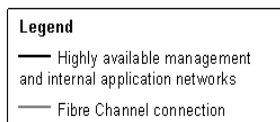
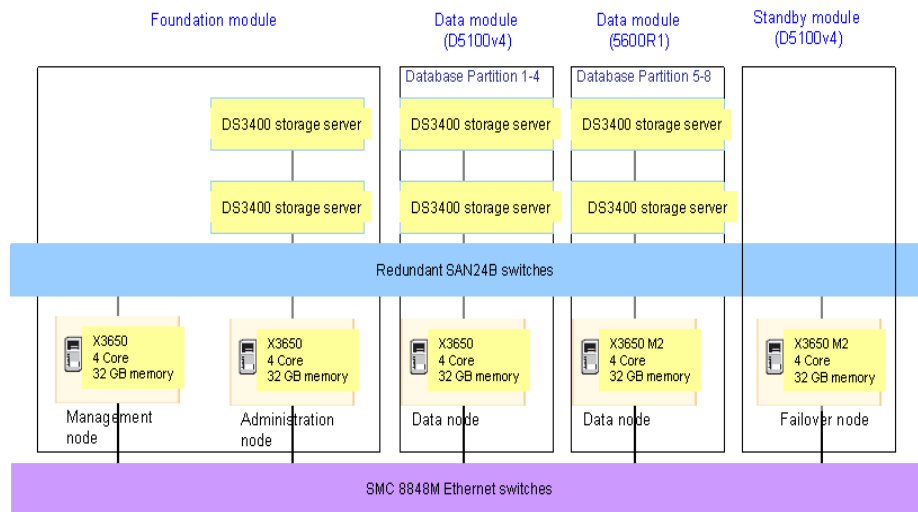


Figure 6. Pre-upgrade environment of D5100v4 and 5600R1 IBM Smart Analytics System

The chosen expansion, illustrated in Figure 7 involved:

- Adding a storage rack positioned for future rolling replacement of older modules with newer modules as needed. RSM and 5600R1 data modules were added previously and took up all available space in the initial storage rack.
- Adding one 5600R2 data module and one 5600R2 standby module to the new storage rack.
- Adding new SAN and Ethernet switches to the new storage rack. The existing SMC8848M switch could have been used but not enough ports were available. An additional SMC8848M switch was added and this avoided having both SMC and Juniper switches in the same environment.

**Case Study – D5100v4/5600R1/5600R2  
Smart Analytics System - Post Upgrade  
Technical Architecture (Simplified for illustration)**

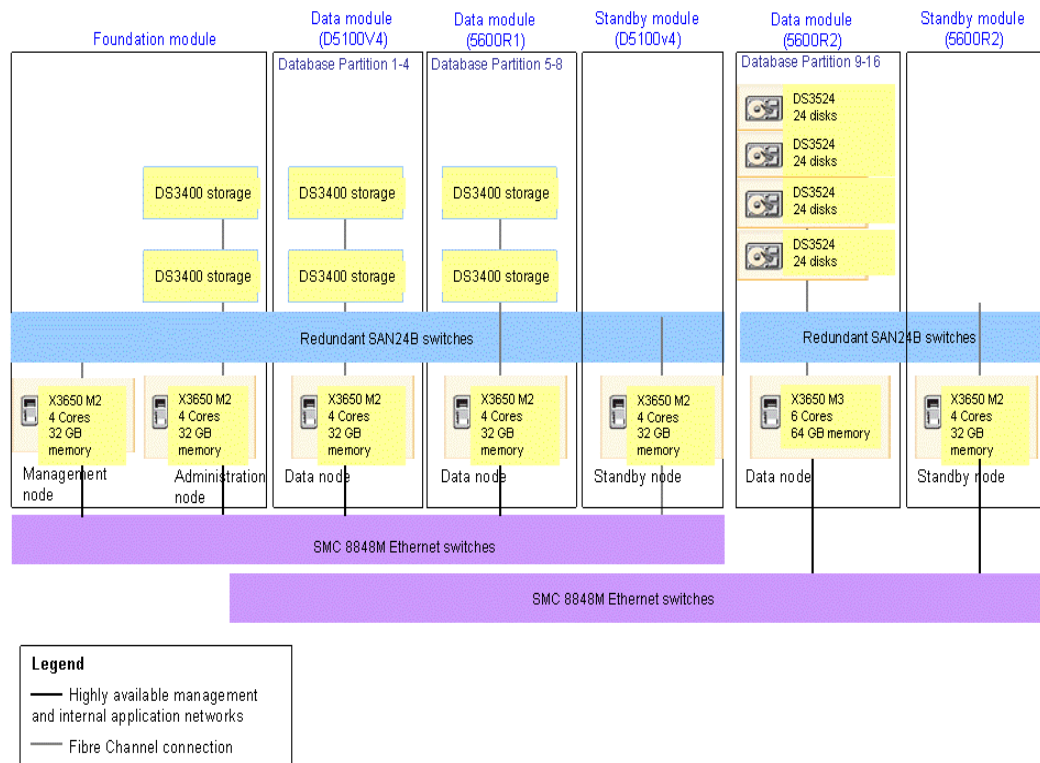


Figure 7. Post-upgrade environment of D5100v4 and 5600R1 IBM Smart Analytics System

The following upgrade tasks were completed as part of the expansion project:

- Support agreement approved for mixed generation D5100v4, 5600R1, 5600R2 IBM Smart Analytics System.
- All software levels upgraded and matched across expanded cluster in line with the published validated software stack.
- Existing data redistributed evenly across each data node, existing and new.

- System health check approved and performance acceptance tests completed.
- High availability test cases updated to accommodate new standby module.
- Installation record updated to reflect new environment.

With respect to outages, the following milestones applied:

- Adding a storage rack – no outage
- Adding new hardware and cabling – no outage
- Adding a HA group and test – no outage
- Adding new data node to partitioned database and test – no outage
- Performing database backup before redistribution – no outage expected
- Redistributing data across each data node – outage expected
- Performing post-redistribution optimizations and backup – outage expected
- Handing over the expanded system to business for performance check and acceptance

## Conclusion

The IBM Smart Analytics System can be expanded to accommodate growth in users and data, providing a scalable and flexible approach to data warehousing infrastructure needs. The following key recommendations apply:

- Good capacity planning practices can assist in early detection of trends in resource usage. You can create and document a performance baseline for each workload and a forecast baseline for the next 12 months. Align your service level objectives with capacity planning indicators so that metrics can be easily compared.
- Begin planning for storage expansion when storage capacity reaches 60% and is projected to reach 80% within 12 months.
- You can add a data module to expand storage capacity and reduce the data volume per database partition. Or, you can add a user module to increase the capacity of the system to accommodate users. When adding modules, remember that multi-generation environments might need multiple high availability groups. Ensure that you incorporate fail over and fail back testing into your expansion plan.

There are other, more general recommendations that apply as well. When planning an expansion, do not forget to include site preparation. Additional cooling, floor space, and network capacity might be needed to accommodate an expanded system.

You also need to take into account the operational implications of your expanded system. Review your backup and recovery infrastructure to ensure that you can maintain service level objectives after the expansion. Also, ETL applications and maintenance scripts need to be reviewed to accommodate the expanded system.



## Appendix A. Query performance

Where growth in demand for resources on your system affects the performance of existing queries, consider the following guidelines:

- Tune queries by reading best practice papers and applying the recommendations suggested.
- Reassess your data model on a regular basis to help ensure that it continues to reflect usage patterns.
- Implement DB2 WLM to assign more resources to those queries that are of higher priority to the business. Avoid allowing low priority queries to gain an equal or larger percentage of processing resources than high priority queries.
- Reduce data volumes through archiving data and pruning of staging and unused tables.
- Review indexing and remove unused indexes.
- Review compression usage and ratios on a regular basis.
- Use database maintenance tools to ensure that statistics and data organization are up to date.
- Move to a later platform to take advantage of technology advances.

## Appendix B. Linux command examples

The syntax of the command examples in the main body of the paper is based on what the AIX operating system support. You can substitute the following examples on the Linux operating system:

Use the `netstat` command as shown to gather statistics every second. Restrict the output to the FCM network by specifying the network via the `-I` parameter:

```
netstat -I enl1 1
```

Use the `iostat` command as shown to gather statistics every second for 30 seconds.

```
iostat -x 1 30
```

Use the `vmstat` command as shown to gather statistics every second for 30 seconds.

```
vmstat -at 1 30
```

## Further reading

- DB2 Best Practices

<http://www.ibm.com/developerworks/db2/bestpractices/>

- Best Practices: Query optimization in a data warehouse

<http://www.ibm.com/developerworks/data/bestpractices/smartanalytics/queryoptimization/>

- Best Practices: DB2 Workload Management

<http://www.ibm.com/developerworks/data/bestpractices/workloadmanagement/>

- Validated stack for IBM Smart Analytics System 5600 R2:

<https://www-304.ibm.com/support/docview.wss?uid=swg21560576>

- Validated stack for IBM Smart Analytics System 7700:

<http://www-01.ibm.com/support/docview.wss?uid=swg21450335>

- Vmstat command usage on AIX

<http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=%2Fcom.ibm.aix.cmds%2Fdoc%2Faixcmds6%2Fvmstat.htm>

- Vmstat command usage on SUSE Linux

<http://www.redbooks.ibm.com/redpapers/pdfs/redp4285.pdf>

## *Contributors*

Alvin Law

*Senior Management Consultant, IBM*

Charles Lai

*IBM Smart Analytics System Core Engineering*

Chub Varga

*Solution Architect, Information Management,  
IBM*

Clare Scally

*SVT Engineer, Information Management, IBM*

John Bell

*Distinguished Engineer & Data Warehouse  
Architect, IBM*

Konwen Kuan

*IBM Smart Analytics System Core Engineering*

Larry Pay

*IT Specialist, WW Lab Services, IBM*

Nikolaj Richers

*Information Architect, IBM Smart Analytics  
System*

Sermasak Sukjirawat

*WW Data Warehouse and Smart Analytics  
System Practices*

Simon Woodcock

*Data Warehouse Consultant, IBM*

## Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

Without limiting the above disclaimers, IBM provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any recommendations or techniques herein is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Anyone attempting to adapt these techniques to their own environment do so at their own risk.

This document and the information contained herein may be used solely in connection with the IBM products discussed in this document.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE: © Copyright IBM Corporation 2011. All Rights Reserved.

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

## **Trademarks**

IBM, the IBM logo, and [ibm.com](http://ibm.com) are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)

Windows is a trademark of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.