White Paper

EMC OPTIMIZATION FOR MICROSOFT SQL SERVER 2012 BUSINESS INTELLIGENCE

EMC VNX5300, EMC FAST, EMC Replication Manager, and Microsoft Hyper-V 2012

- Performance optimized by FAST
- Rapid recovery provided by EMC Replication Manager
- Business insight provided by Microsoft Business Intelligence suites

EMC Solutions Group

Abstract

This white paper describes how to design and implement Microsoft SQL Server 2012 Infrastructure for SQL Server based Data Warehousing and Business Intelligence (BI) in a virtualized environment enabled by the EMC® VNX®5300 storage array. It demonstrates the performance improvements of EMC Full Automated Storage Tiering (FAST) technology on EMC VNX5300 and showcases the new BI Semantic Model in SQL Server 2012 integrated with SharePoint 2010. It also shows the accelerated protection for data warehousing through EMC Replication Manager with EMC hardware snapshot technology.

December 2012

EMC² PROVEN SOLUTIONS

EMC²

Part Number H11257

Table of contents

Executive summary **5**
- Business case 5
- Solution overview 5
- Key results 6

Introduction **7**
- Purpose 7
- Scope 7
- Audience 7
- Terminology 8

Technology overview **10**
- Introduction to key components 10
- EMC FAST 10
- EMC Replication Manager 10
- EMC VNX5300 storage array 11
- Microsoft MPIO 11
  - Microsoft PerformancePoint Services 11
- Microsoft SharePoint 2010 12
- Microsoft SQL Server 2012 12
  - Columnar storage 12
  - BI Semantic Model 13
  - xVelocity in-memory analytics engine 13
  - Power View 14
- Microsoft Windows Server 2012 Hyper-V 14

Solution configuration **16**
- Overview 16
- Solution architecture 16
- Test configuration 17
- Hardware resources 19
- Software resources 19
- Multipath design considerations 20
- FAST design for SQL Server 2012 Data Warehouse 20
- Cluster and virtualization design considerations 22
- Reserved LUN pool design considerations 22
- SharePoint 2010 farm configuration for BI 23
- Power View infrastructure configuration 23
- SharePoint PerformancePoint and PowerPivot for SharePoint configuration 24
- PowerPivot for Excel configuration 25

Test and validation **26**
Overview 26
Test scenarios 26
Test objectives 26
Test procedure 27
Test results 27
Effects of columnstore on DSS workloads 27
Enabling EMC FAST on the DSS data pool 29
Hyper-V host server performance 35
BI showcase 35
Power View reports 35
Hierarchies 35
Measures 36
KPI 36
Reports 37
PerformancePoint BI 38
Dashboard 39
PowerPivot for Excel 40
Data protection 41

Conclusion **42**
Summary 42
Findings 42

References **43**
White papers 43
Product documentation 43
Other documentation 43

Appendix **44**
Disabling SCSI filtering 44
Power View report 44
PerformancePoint scorecard and report 46
Scorecard 46
Report 47
PivotTable and PivotChart on Excel 47

Executive summary

Business case

In an increasingly competitive environment, businesses are being driven to optimize processes and improve services. IT organizations must strive for more efficiency and improved quality of services, including:

- Achieving and sustaining high performance IT Infrastructures to support Decision Support System (DSS) workloads.
- Providing accelerated protection and rapid recoverability of critical business data for data warehousing.
- Empowering users to do more and enable them to develop their own Business Intelligence (BI) instances through self-service, reducing the IT department burden, and decreasing Time-to-Solution for the analytics business teams.

Microsoft has released SQL Server 2012 with an array of new and improved capabilities. One new feature is the BI Semantic Model (BISM), which connects to the Power View and PowerPivot self-service BI tools. In addition, the analysis services capability has been expanded, and columnar store in the database engine provides significantly faster performance in scanning data warehouse datasets.

This solution highlights EMC's capabilities for Microsoft SQL Server, especially with its BI features. Microsoft SharePoint is used as the presentation layer while the entire complex environment is simplified with a set of design guidelines supported by EMC® VNX® storage.

Solution overview

This solution illustrates the performance and functionality of a virtualized, medium sized, SQL Server 2012 Data Warehousing and BI environment on the EMC VNX storage platform running SQL Server 2012.

The objectives of this solution are:

- To demonstrate the performance capabilities of SQL Server 2012 Data Warehousing and BI in a virtualized environment enhanced with EMC storage array technologies such as:
  - EMC VNX Virtual Provisioning
  - EMC VNX Full Automated Storage Tiering (FAST) technology
- To showcase EMC applications and development expertise for Microsoft SQL Server and SharePoint in business intelligence applications.
- To demonstrate rapid replication and recovery using EMC Replication Manager and EMC SnapView™ snap technology for SQL Server 2012 Data Warehouse.

The deliverable of this solution is to show the DSS workload performance improvement enabled by EMC FAST auto-tiering, and showcase the BI Semantic Model in SQL Server 2012 BI suites integrated with SharePoint 2010, along with the EMC VNX capabilities for BI storage optimization.

This solution also provides protection for SQL 2012 Data Warehouse by creating snaps with the EMC VNX capabilities:

- EMC Replication Manager
- EMC SnapView snap

The inability to integrate company information from multiple sources impedes current decision support systems. The challenge lies in harvesting the data and turning it into a tool for competitive advantage. This requires data management and analytics expertise. EMC Consultants can help solve these problems. For more information, refer to Business Intelligence and analytics at the EMC website.

Key results

This solution demonstrates the following results:

- EMC VNX5300 is capable of sustaining high volume SQL Server 2012 DSS workloads on SQL Server 2012.
- EMC FAST technology on VNX series array is proven to accelerate performance of SQL Server BI and Data Warehousing environments transparently.
- EMC Replication Manager with EMC hardware snapshot technology significantly accelerates protection for data warehouses.
- This solution also showcases the new BI Semantic Model in SQL Server 2012 Analysis Services and the enterprise BI suites integrated with SharePoint 2010.

Introduction

Purpose

The white paper describes the design, testing, and validation of a SQL Server 2012 BI solution in a virtualized environment using VNX5300 with FAST technology focused on medium-sized business market. Specifically, this solution:

- Validates the ability of EMC VNX5300 to sustain high volume of SQL Server 2012 DSS workloads with both rowstore indexing and columnstore indexing.
- Shows that EMC FAST auto-tiering technology can boost performance for DSS workloads running on SQL Server 2012.
- Demonstrates the Microsoft BI Suite including the enterprise BI suites integrated with SharePoint 2010 and PowerPivot, the personal BI tool with Excel.
- Validates rapid replication and recovery using EMC Replication Manager and EMC SnapView snap technology for SQL Server 2012 Data Warehousing environments, with negligible performance overhead to the application.

Scope

The scope of this white paper is to:

- Showcase the performance and functionality of enterprise-class data warehousing in a virtualized environment running SQL Server 2012 with new columnar store technology.
- Demonstrate storage performance optimization with FAST.
- Show accelerated protection for data warehouse through EMC Replication Manager with EMC hardware snapshot technology.
- Showcase EMC application and development expertise and consulting capabilities in Microsoft BI solutions.

Audience

This white paper is intended for personnel designing and managing enterprise data warehousing and business intelligence environments, and storage architects involved in planning, architecting, or administering an environment with VNX.

Terminology

This paper includes the following terminology.

Table 1. Terminology

Term	Definition
Business Intelligence (BI)	The ability of an organization to collect, maintain, and organize knowledge. BI technologies provide historical, current, and predictive views of business operations.
EMC VNX5300	Designed for the midrange entry space. It is the introductory model for the VNX unified platform. This model provides block only, file services only, or both block and file services.
Decision support system (DSS)	A decision support system (DSS) supports business or organization decision-making activities. The solution uses a TPC-H-like application to mimic DSS workloads.
Device-specific modules(DSM)	Microsoft Multipath I/O (MPIO) is a Microsoft-provided framework that allows storage providers to develop multipath solutions that contain the hardware-specific information needed to optimize connectivity with their storage arrays. These modules are called device-specific modules (DSM).
Fully Automated Storage Tiering for Storage Pools (FAST VP)	A feature of VNX storage arrays that automates the identification of data volumes for the purpose of allocating or reallocating business application data across different performance and capacity tiers within the storage array.
Hierarchy	Groups of columns arranged in levels.
Key Performance Indicator (KPI)	A quantifiable measurement for gauging business success. Business executives frequently consume KPIs that are grouped together in a business scorecard to obtain a quick and accurate historical summary of a business success.
Measure	Represents a column that contains quantifiable data, usually numeric, that can be aggregated. A measure is generally mapped to a column in a fact table.

Term	Definition
TPC-H-like power	Computes the TPC-H-like query processing power at the chosen database size by using the results of the power test. Bigger value for TPC-H-like power means more powerful query execution ability of the system.
TPC-H-like power test	Measures the query execution power of the system when connected with a single user. It runs the analysis in a serial manner–running one of the queries and updating the functions at a time and measuring the elapsed time.
Clustered Shared Volumes (CSV)	A feature that simplifies the configuration and management of Hyper-V virtual machines in failover clusters. Multiple virtual machines that are distributed across multiple cluster nodes can all access their Virtual Hard Disk (VHD) files at the same time.

Technology overview

Introduction to key components

This section provides an overview of the technologies that are used in this solution:

- EMC FAST
- EMC Replication Manager
- EMC VNX5300 storage array
- Microsoft Multipath I/O (MPIO)
- Microsoft PerformancePoint Services
- Microsoft SharePoint 2010
- Microsoft SQL Server 2012
- Microsoft Windows Server 2012 with Hyper-V

EMC FAST

Fully Automated Storage Tiering (FAST) can lower total cost of ownership (TCO) and increase performance by intelligently managing data placement at a sub-LUN level. When FAST is implemented, the storage system measures, analyzes, and implements a dynamic storage-tiering policy much faster and more efficiently than a human analyst could ever achieve.

Storage tiering puts drives of varying performance levels and cost into a storage pool. FAST collects I/O activity statistics at a 1 GB granularity (known as a slice). The relative activity level of each slice is used to determine which slices should be promoted to higher tiers of storage.

FAST is a licensed feature available on EMC VNX series. The VNX series supports a unified approach to automatic tiering for both file and block data.

With FAST, customers can achieve:

- Better performance with lower cost, fewer drives, less power and cooling, and a smaller footprint
- Maximum utilization of Flash drives for high-performance workloads
- Lower cost of storage by placing the less accessed data on SATA drives
- Radically simplified automated management in a tiered environment

EMC Replication Manager

EMC Replication Manager manages EMC point-in-time replication technologies through a centralized-management console. Replication Manager coordinates the entire data replication process, from discovery and configuration to the management of multiple application-consistent, disk-based replicas. Auto-discover your replication environment and enable streamlined management by scheduling, recording, and cataloging replica information, including auto-expiration. With Replication Manager, you can put the right data in the right place at the right time—on-demand or based on schedules and policies that you define. This application-centric product allows you to simplify replica management with application consistency.

EMC VNX5300 storage array

The VNX5300 storage array is a member of the VNX series next-generation storage platform, providing the industry's highest bandwidth. VNX5300 provides high-performing, unified storage with unsurpassed simplicity and efficiency. Organizations achieve new levels of performance, protection, compliance, and ease-of-management.

The VNX5300 storage array delivers a single-box block and file solution, which offers a centralized point of management for distributed environments. This makes it possible to dynamically grow, share, and cost-effectively manage multiprotocol file systems and provide multiprotocol block access.

For VNX specifications for both block and file, refer to EMC VNX Series unified storage system.

Microsoft MPIO

Microsoft Multipath I/O (MPIO) is a Microsoft-provided framework that allows storage providers to develop multipath solutions that contain the hardware-specific information needed to optimize connectivity with their storage arrays. These modules are called device-specific modules (DSMs).

MPIO is protocol-independent and can be used with Fibre Channel, Internet Small Computer System Interface (iSCSI), and Serial Attached SCSI (SAS) interfaces in Windows Server 2008, Windows Server 2008 R2, and Windows Server 2012.

MPIO includes a DSM designed to work with storage arrays that support the asymmetric logical unit access (ALUA) controller model, as well as storage arrays that follow the Active/Active controller model.

The advantages of multipathing are as follows:

- The ability to fail over to the other path and resend any outstanding I/O to address the path failure. For a server that has one or more host bus adapters (HBAs) or network adapters, MPIO also provides:
  - Support for redundant switch fabrics or connections from the switch to the storage array
  - Protection against the failure of one of the adapters within the server directly
- The ability to balance the data load to maximize the throughput
- Higher bandwidth attachment from the host to the storage system

Microsoft PerformancePoint Services

PerformancePoint Services in Microsoft SharePoint Server 2010 is a performance management service that you can use to monitor and analyze your business. By providing flexible, easy-to-use tools for building dashboards, scorecards, and key performance indicators (KPIs), PerformancePoint Services can help individuals across an organization make informed business decisions that align with companywide objectives and strategy. Dashboards, scorecards, KPIs, and reports help drive accountability.

Microsoft SharePoint 2010

Microsoft SharePoint 2010 provides a business collaboration platform for enterprise and commercial organizations and the Internet. SharePoint enables organizations to share content and information through websites, blogs, wikis, and document libraries—some of the many features of a SharePoint environment. Content within these features can be managed collectively from start to finish. Thoughts and ideas can be exchanged and discussed using the same set of capabilities and tools.

The SharePoint 2010 platform is composed of a set of integrated capabilities that can be used as-is or customized to address specific business needs and integrated with other products and solutions. The platform can be deployed both within the organization (through intranets) and outside the firewall (through extranets and the Internet) to enable interaction with employees, customers, and business partners.

Microsoft SQL Server 2012

Microsoft SQL Server 2012 is the latest version of the Microsoft database management and analysis system for e-commerce, line-of-business, and data warehousing solutions. It enhances SQL Server BI capabilities by spanning the spectrum of reporting and analysis requirements as well as new self-service capabilities.

This solution demonstrates these feature enhancements:

- Allowing users to visualize and rapidly explore data with the new ad-hoc query capability in Power View (formerly known as Project Crescent) and new capabilities in PowerPivot.
- Providing users with accessible intelligence across the organization through managed self-service BI, dashboards and scorecards, and the collaboration capabilities of Microsoft SharePoint.

Columnar storage

Columnar storage stores data tables as sections of columns of data rather than as rows of data. This feature improves I/O performance and makes more efficient use of memory on data warehousing queries. Columnar storage is exposed as a new index type, a columnstore index.

Columnstore indexes significantly speed up the processing time of common data warehousing queries. Typical data warehousing workloads involve summarizing large amounts of data. The techniques typically used in data warehousing and decision support system to improve performance are pre-computed summary tables, indexes views, and OLAP cubes. However, these techniques can be inflexible, difficult to maintain, and must be designed specifically for each query. By using columnstore indexes, users can reduce the overheads so that pre-computation is no longer required.

The key characteristics of SQL Server columnstore technology are as follows:

- **Columnar data format**

  Database engine stores one column at a time.

- **Faster query result**

  Columnstore indexes can produce faster results for the following reasons:

- Only the columns needed are read. Therefore, less data is read from disk to memory and later moved from memory to processor cache.
- Columns are heavily compressed. This reduces the number of bytes that must be read and moved.
- It improves buffer pool usage, because most queries do not touch all columns of the table. Therefore, many columns will never be brought into memory.

BI Semantic Model

The BI Semantic Model is a new hybrid data model that offers the powerful analytical capabilities of multidimensional data models as well as the simplicity and familiarity of tabular models. The BI Semantic Model provides a powerful flexible platform for nearly all BI applications, supporting virtually all of the requirements that any growing enterprise might have, including advanced analytical capabilities, professional developer tools, ease of adoption and use, wide spectrum of end user tools, sophisticated business logic, performance, scalability, and time to solution. It also enables a wide spectrum of BI solutions—including reporting, analytics, scorecards, dashboards, and custom applications—that are suitable in a wide range of business contexts, covering business users, BI professionals, and IT.

The BI Semantic Model has three layers:

- The data model layer is a conceptual model that supports either the traditional multidimensional or the new tabular data models. Client tools like Excel, Power View, and SharePoint 2010 Insights consume the data model for their user experience.
- The business logic and queries layer represents the embedded business logic using either Multidimensional Expressions (MDX) or Data Analysis Expressions (DAX).
- The data access layer provides the integration of data from multiple sources such as relational databases, business applications, flat-files, and OData feeds.

For the BI Semantic Model architecture, refer to the Microsoft white paper *New Capabilities Overview - Breakthrough Insight.*

xVelocity in-memory analytics engine

The xVelocity in-memory analytics engine is the next generation of the VertiPaq engine that was introduced in SQL Server 2008 R2, with PowerPivot for Excel 2010 and PowerPivot for SharePoint 2010. Vertipaq is an in-memory columnstore engine that achieves breakthrough performance for analytic queries by employing techniques such as columnar storage, state-of-the-art compression, in-memory caching, and highly parallel data scanning and aggregation algorithms. In SQL Server 2012, the xVelocity in-memory analytics engine has been enhanced to support both self-service BI (PowerPivot) as well as corporate BI (Analysis Services tabular mode) scenarios.

Power View

Power View is a new, highly interactive data exploration and reporting tool that enables users to visually explore their data and answer ad-hoc questions with ease.

Power View is a Web-based Microsoft Silverlight application that can be used directly by users—from business executives to information workers—without having to depend on IT. Using Power View, users can enjoy the interactive views of any data residing in the form of a tabular model, including PowerPivot files published to the PowerPivot Gallery or a Tabular BI Semantic Model on an instance of Analysis Services. With Power View, users can quickly and easily analyze their data, prepare an effective presentation of their discovered insights, and share this information broadly using only a browser.

Power View is available within Microsoft SharePoint Server 2010. It requires SQL Server Reporting Services (SSRS) in SharePoint integrated mode and a tabular instance of SQL Server 2012 Analysis Services (SSAS) or PowerPivot for SharePoint to take advantage of PowerPivot models.

Power View includes the following main benefits:

- Visual design experience. Users can access and analyze information in a completely interactive, web-based authoring environment that has the familiar look and feel of Microsoft Office.
- Filtering and highlighting data. Users can interact with the data to gain insights.
- Presentation-ready. Users can share information quickly and easily, browse and present data at virtually any time without previewing it on any other platforms.

Microsoft Windows Server 2012 Hyper-V

Windows Server 2012 Hyper-V is the latest Microsoft virtual platform. It continues to transform IT infrastructure into the most efficient, shared, and on-demand utility, with built-in availability, scalability, and security services for all applications and simple, proactive automated management.

Windows Server 2012 Hyper-V has the following scalability and performance enhancements, which enable a virtual machine to leverage more resources from the hypervisor:

- Delivered greater scalability with support of up to 64 virtual CPUs per virtual machine
- 1 TB virtual machine RAM
- Non-Uniform Memory Access (NUMA) support in a virtual machine

In Windows Server 2012 Hyper-V, you can control the way a cluster handles virtual machines by assigning a priority.

Priorities can be set as High, Medium, Low, or No Auto Start. When a virtual machine is made highly available, the default priority is set to Medium.

Priorities control the order in which virtual machines are started and ensure that lower-priority virtual machines automatically release resources if they are needed by higher priority virtual machines.

When the No Auto Start priority is applied to a virtual machine, it will not automatically start in the event of a cluster node failure.

Solution configuration

Overview

This solution includes a VNX5300 storage array with three tiers of storage managed and controlled by FAST. The storage array services two SQL Server instances and a small SharePoint farm in a virtualized environment running SQL Server 2012 with DSS workloads.

Solution architecture

This solution design includes the following physical components:

- Two Hyper-V hosts, each hosting one SQL Server virtual machine
- One management server, running Replication Manager server
- EMC VNX5300 SAN storage with three FAST-enabled tiers

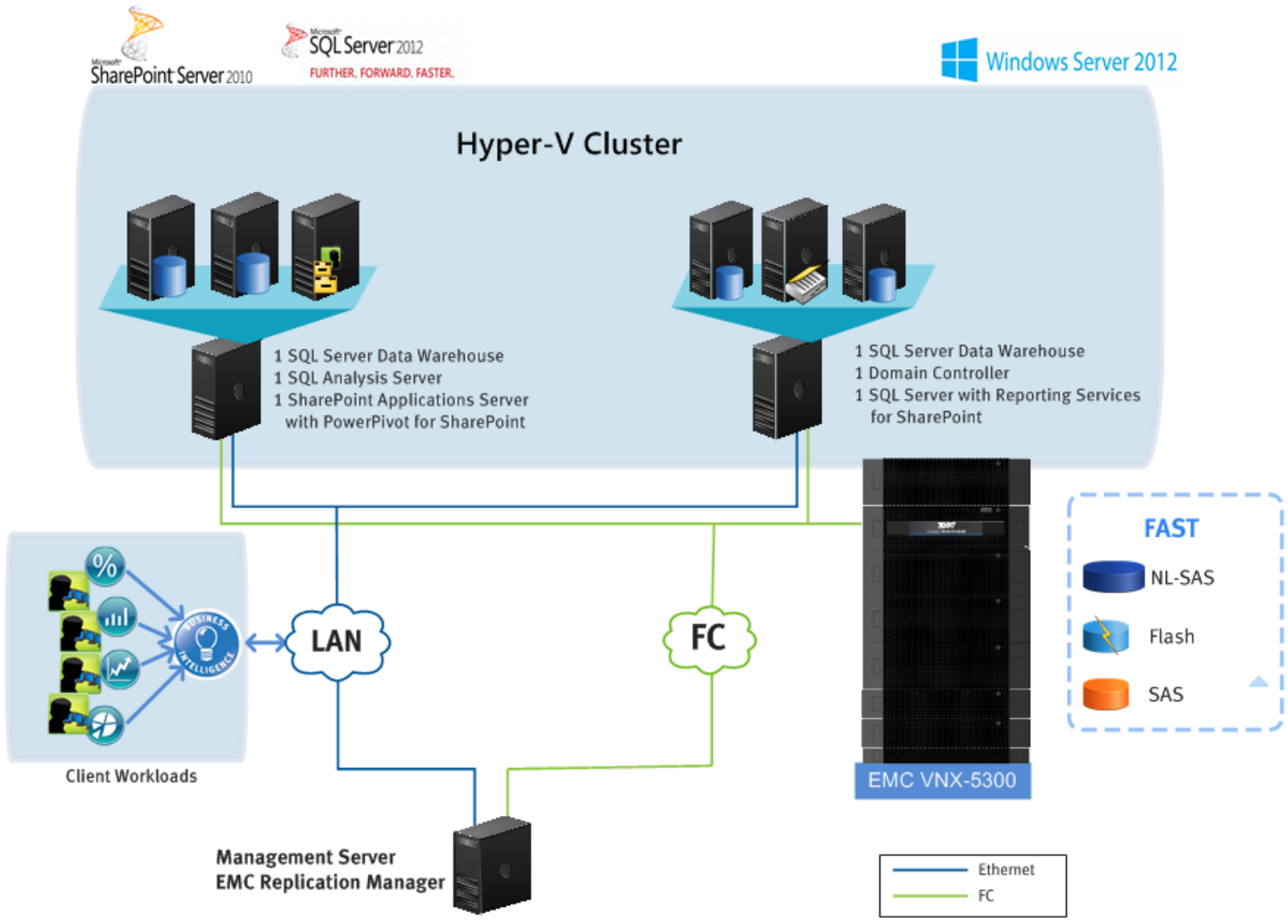


Figure 1. Solution architecture

This solution design includes the following virtual machines on the Hyper-V hosts:

- Two SQL Server virtual machines running DSS workloads
- One SQL Server 2012 Analysis Server in Tabular mode. The server stores the tabular BI Semantic model that is used as the data source for the Power View reports. It is one of the prerequisites for Power View.
- One SQL Server 2012 with Reporting Services for SharePoint Server 2010. We configured the SQL Server Reporting Services in SharePoint integrated mode. It is another prerequisite for Power View.
- One SharePoint Applications Server with PowerPivot for SharePoint enabled. This virtual machine also runs central administration for SharePoint. To enable PowerPivot for SharePoint, we installed one Analysis Server instance in SharePoint mode as required.

Test configuration

Table 2 shows the profile of the SQL Server virtual machines for DSS configuration.

Table 2. SQL Server virtual machines for DSS configuration

Profile	Value
Total SQL Server database capacity	1 TB
Number of SQL Server instances	2
Number of user databases for each virtual machine	1
Number of virtual machines	2
Type of Hyper-V disk	Pass-through
SQL Server virtual machine configuration	18 vCPUs with 48 GB RAM
Database read/write ratio	100/0
Concurrent users	1

Table 3 shows the storage pool design for the data warehouses without FAST.

Table 3. Storage pool design without FAST

Pool name	Drive size/ technology/RPM	RAID protection	Number of drives	Pool capacity
SQL Data warehouses	600 GB SAS 10k	RAID5 (4+1)	25	10 TB

Table 4 shows the SQL Server LUN design for the data warehouses in rowstore.

Table 4. LUN design for the data warehouses in rowstore

Item	Component	LUN capacity (GB)	Quantity	Total capacity (GB)
SQL DSSDW01	tempdb	75	4	300
	tempdb log	50	1	50
	Database	200	6	1,200
	Database log	200	1	200
SQL DSSDW02	tempdb	75	4	300
	tempdb log	50	1	50
	Database	200	6	1,200
	Database log	200	1	200

Table 5 shows the SQL Server LUN design for the data warehouses in columnstore.

Table 5. LUN design for the data warehouses in columnstore

Item	Component	LUN capacity (GB)	Quantity	Total capacity (GB)
SQL DSSDW01	tempdb	40	4	160
	tempdb log	50	1	50
	Database	125	6	750
	Database log	200	1	200
SQL DSSDW02	tempdb	40	4	160
	tempdb log	50	1	50
	Database	125	6	750
	Database log	200	1	200

The following list shows the Windows and SQL Server 2012 configuration of each virtual machine. In this solution, we followed the best practices for optimal performance:

- Use Large Pages for the SQL Server instance by enabling the 834 startup parameters. For more information, refer to *Tuning options for SQL Server when running in high performance workloads* at the Microsoft Support website.
- Use the **Lock page in memory** option for the SQL Server instance. For more information, refer to the MSDN Library topic *Pre-Configuration Database Optimizations.*

- Use the 64K format unit for all LUNs containing SQL Server data and log files. For more information, refer to the MSDN Library topic *SQL Server Best Practices*.

Hardware resources

Table 6 shows the hardware resources needed for this solution.

Table 6. Hardware resources

Equipment	Quantity	Configuration
EMC VNX5300	1	7.2K 2 TB (NL SAS)—Quantity: 39 10K/15K 600 GB (SAS)—Quantity: 64 SATA Flash 200 GB (Flash)—Quantity: 5 Flare Operating Environment 32 FAST enabled (Drive counts reflect what we chose as an efficient storage solution for this customer profile and should not be considered a maximum of the VNX5300 array)
Four-socket server	2	4 × 6 cores, 128 GB RAM (for Hyper-V hosts)
Two-socket server	1	2 × 4 cores, 36 GB RAM (for Replication Manager server)
Fibre Channel switch	1	8 GB SAN switch
GbE network switch	1	Network switch–24 ports

Software resources

Table 7 lists the software resources needed for this solution.

Table 7. Software resources

Software	Quantity	Version	Purpose
EMC VNX5300 code	1	Flare operating environment 32	Support the latest VNX features
Windows Server 2012	6	RTM	Two SQL Server virtual machines One analysis server virtual machine Domain controller Two Hyper-V host servers
Windows Server 2008 R2	1	SP1	Microsoft SharePoint 2010 virtual machine
SQL Server 2012	3	SP1	Two SQL Server instances to run DSS workloads One SQL Server instances to run SharePoint 2010
SQL Server 2012 Analysis Services	1	SP1	Vertipaq Analysis Instance for BI showcasing
SQL Server 2012 Reporting Services	1	SP1	Reporting Services for BI showcasing
SharePoint 2010	1	JUNE 2012 CU	SharePoint for BI showcase

Microsoft SQL Server 2012 PowerPivot for SharePoint	1	SP1	Plug-in to support PowerPivot in SharePoint
Microsoft SQL Server 2012 PowerPivot for Excel	1	SP1	Plug-in to support PowerPivot in Excel
EMC Replication Manager	1	5.4.2	Manages EMC point-in-time replication technologies through a centralized-management console
EMC Solutions Enabler	1	7.4.0.0	Used to discovery of VNX arrary.
EMC Navi Secure command line interface (CLI)	1	7.32	CLI used for management, diagnostics and reporting functions.
EMC Admsnap	1	2.32	Used to create snapshot session

Multipath design considerations

Each hyper-V host server had two HBA ports connected to the SAN switch. Two ports from two engines on VNX5300 were zoned to 8 GB SAN switch and zoned to each port of a hyper-V host. Thus each device on the host had two paths. Through these paths, the workload could be evenly distributed between the two VNX5300 engines.

We configured MPIO to use the Microsoft DSM, which is a generic DSM provided for Windows Server 2012 in a Fibre Channel, iSCSI, or SAS-shared storage configuration.

To configure the MPIO to determine if multiple instances actually represent the same LUN through different paths, you need to discover and configure the devices that the operating system sees through redundant adapters:

1. Open the MPIO control panel.
2. On the **Discover Multi-Paths** tab, choose the device hardware named **DGC VRAID** and click **Add**.
3. Restart the server as required by MPIO.

Note To work with the Microsoft DSM, storage must be compliant with SCSI Primary Commands-3 (SPC-3).

FAST design for SQL Server 2012 Data Warehouse

Outlined below are some general guidelines for sizing a consolidated workload with FAST:

- Flash drives provide the best performance for workloads with larger numbers of random reads. Because of the sequential write nature of the database log LUNs, enable FAST only for data LUNs to effectively utilize Flash drives.
- Perform sizing according to the FAST requirements. The following configurations are recommended:
  - RAID 5 Flash drives for the balanced performance and total cost of ownership (TCO)
  - RAID 6 SATA drives for the best data protection
  - RAID 5 SAS drives for the balanced performance and TCO

- Use separate storage pools for data files and log files.
- Pin the Log LUNs to the SAS disks.

Table 8 shows the FAST tiers used in this solution.

Table 8. FAST tiers for DSS

Tier name	Drive size/ technology/RPM	RAID protection	Number of drives	Tier capacity
Extreme Performance	200 GB SATA Flash	RAID5 (4+1)	5	730 GB
Performance	600 GB SAS 10k	RAID5 (4+1)	25	10 TB
Capacity	2 TB SATA 7.2k	RAID6 (6+2)	8	10 TB

Table 9 shows the LUN design for FAST.

Table 9. LUN design for FAST

Virtual machine	LUN name		Number of LUNs	LUN size	Default tier	FAST tiering policy
SQL DSSDW01	DW data		6	125 GB	SAS tier	Auto-tiering
	DW log		1	200 GB	SAS tier	N/A
	tempdb	Data	4	40 GB	Flash tier	Highest tier available
		Log	1	50 GB	SAS tier	N/A
Total				1.16 TB	N/A	N/A
SQL DSSDW02	DW data		6	125 GB	SAS tier	Auto-tiering
	DW log		1	200 GB	SAS tier	N/A
	tempdb	Data	4	40 GB	Flash tier	Highest tier available
		Log	1	50 GB	SAS tier	N/A
Total				1.16 TB	N/A	N/A

The design incorporates the following recommended practices:

For data warehouse:

- Use multiple data files for better performance. This can avoid Page Free Space (PFS) contention on the SQL Server. By placing each data file on a separate LUN, the data is evenly allocated on disks; therefore the workloads are pushed to the disks more evenly. In this solution, we used six LUNs for each data warehouse.
- Pin the LUNs for log files to the SAS disks (RAID 1+0) from a separated storage pool because of the sequential write nature of the database log.

For tempdb:

- In this solution, we used four data files on four LUNs for each DSS instance tempdb.
- Pre-allocate space, and add a single data file per LUN. Be sure to make all files the same size.
- Pin the log LUNs to the SAS disks (RAID 1+0) from the same pool as the one used for data warehouse logs, and assign each log file to one LUN dedicated to log files.
- Enable autogrow; for common data warehouse large growth size is appropriate. 1 GB is a reasonable value for this setting in this environment.

For more information about optimizing tempdb performance, refer to the MSDN Library topic *Optimizing tempdb Performance*.

Cluster and virtualization design considerations

In this solution, we used the quorum configuration of **Node and Disk Majority** for the hyper-V failover cluster. We configured a pass-through disk as disk witness to sustain failures of half the nodes. To enable virtual machine live migration and failover, we put all the virtual machine operating system virtual disks and configuration files on Clustered Shared Volumes (CSV).

To control the order in which virtual machines are started, here is the priority table (Table 10) we set for the virtual machines.

Table 10. Priority table setting

Virtual machine	Startup priority
Active directory server	High
SQL Server for DSS	Medium
SQL Server for SharePoint	Medium
Analysis server in Vertipaq mode	Low
SharePoint application server	Low

Reserved LUN pool design considerations

In this solution, we provided protection to the data warehouse based on EMC VNX Classic Snaps –SnapView snap technology, but the protection for this solution is optional.

Outlined below are some general guidelines for sizing the reserved LUN pool used by SnapView snap:

- Consider the rate of change on the source LUN when sizing the reserved LUN(s) that will be assigned to the reserved LUN pool.
- Consider the pattern of data writes. When the writes are very small and very random, the 64 KB tracking can cause much more data to be written to the reserved LUNs than what is actually changed on the server. In this situation, larger reserved LUNs should be used.

- Consider the overflow LUN factor. Configure twice as many reserved LUNs as source LUNs.
- Consider using small reserved LUNs to save space and large reserved LUNs to maximize the number of participating source LUNs.

In this solution, we assumed 10 percent sequential data change in the source LUNs and assumed that the overflow LUN factor is two, allowing for a 20 percent data change in source LUNs.

Replication Manager supports Windows 2008 R2 and prior versions. In this solution, we configured a 500 GB data warehouse to store the sales data sample of EMC Flash storage arrays on SQL Server 2012 running on Windows 2008 R2.

Table 11 shows the design of the reserved LUN pool.

Table 11. Design table for reserved LUN pool

Source LUN name	Number	Capacity	Reserved LUN size	Total number of reserved LUNs
Data LUN	6	100 GB	10 GB	12
Log LUN	1	200 GB	10 GB	4
Reserved LUNs	N/A	N/A	N/A	16 (× 10 GB)

To conserve space, we configured small reserved LUN size as 10 GB (100 GB x 20% /2). Therefore, we need 160 GB (10 GB × 16) reserved LUN capacity.

SharePoint 2010 farm configuration for BI

In this solution, we configured a small SharePoint 2010 farm with Service Pack 1 combined with PowerPivot, PowerView, SQL Server Reporting Services, and PowerPivot features for Excel 2010. The small SharePoint farm used one virtualized application server and one dedicated SQL Server 2012 virtual machine.

In this solution, we used Reporting Services in SharePoint mode to provide report generation and delivery based on Microsoft SQL Server and Microsoft SharePoint products. Running Reporting Services in SharePoint mode provides the Power View and data alerting features. For more detailed steps about how to install Reporting services to SharePoint farm, see Microsoft TechNet Library topic *Add an Additional Reporting Services Web Front-end to a Farm.*

Power View infrastructure configuration

We installed and designed the Microsoft Business Intelligence infrastructure integrated with SQL Server 2012 and SharePoint 2010, which focuses on SQL Server 2012 Reporting Services with Power View.

For the detailed steps to build the Power View infrastructure, you can refer to MSDN Library topic *Power View Infrastructure Configuration and Installation: Step-by-Step and Scripts*.

Note To show the defined key performance indicators (KPIs) and hierarchies, SP1 or later versions are required across all SQL components including SQL Server 2012 database engine, SQL Server 2012 Analysis Services, SQL Server 2012 Reporting Services, and SQL Server 2012 PowerPivot for SharePoint.

SharePoint PerformancePoint and PowerPivot for SharePoint configuration

We designed a PowerPivot workbook and saved it to the SharePoint site. In Dashboard Designer, design the dashboard, scorecard, report, and KPI by connecting the PowerPivot workbook as the data source as shown in Figure 2.

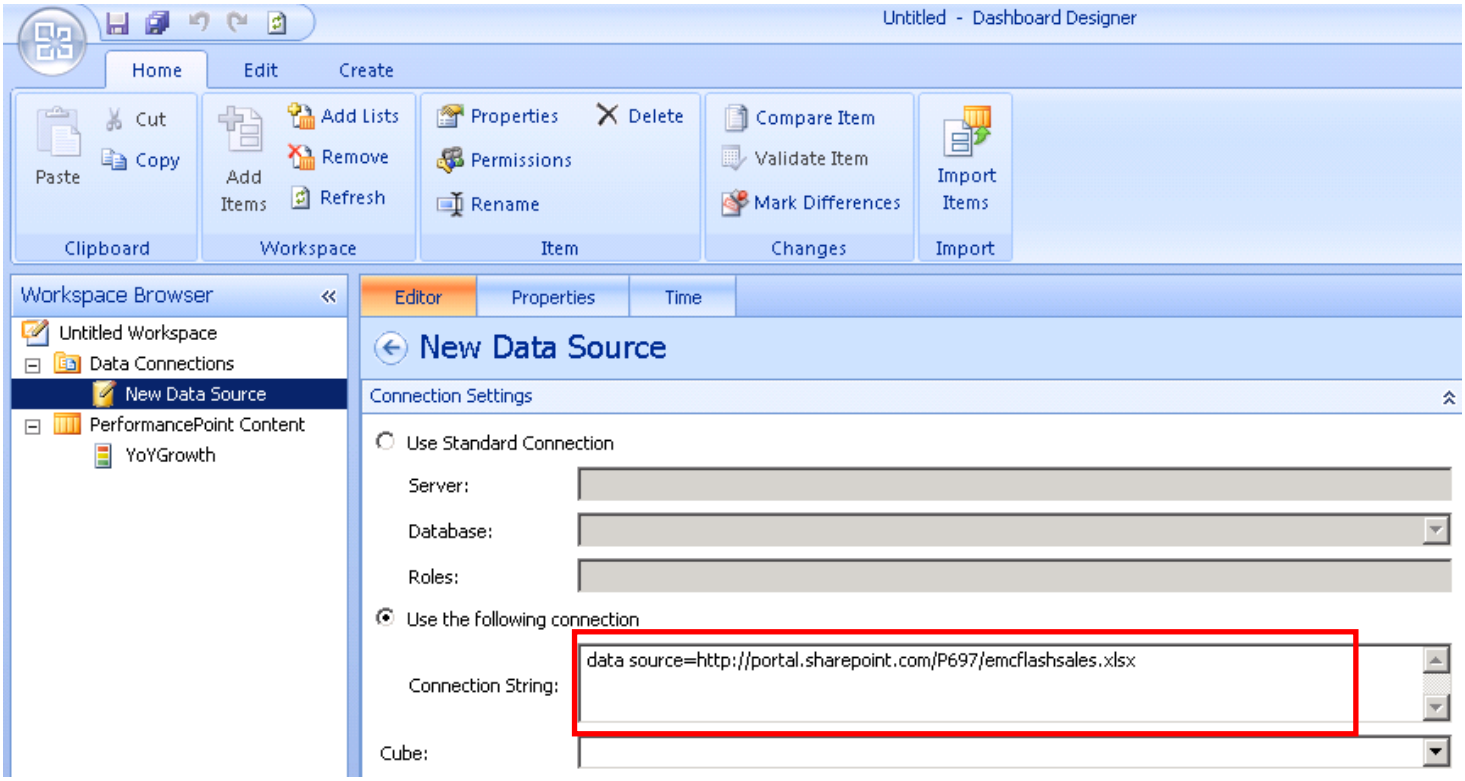


Figure 2. Data source defined in Dashboard Designer

Note To connect to a PowerPivot 2012 workbook as a data source from SharePoint, use SQL Server 2008 R2 ADOMD.NET provider instead of the SQL 2012 version.

For detailed steps on installing PowerPivot for SharePoint, refer to Microsoft Technet Library topic *PowerPivot for SharePoint Installation (SharePoint 2010).*

For detailed steps to build the PerformancePoint environment, refer to Microsoft Technet topic *Configure PerformancePoint Services for a BI test environment.*

PowerPivot for Excel configuration

In this solution, we configured the PowerPivot for Excel to be used for personal BI.

In this case, we pulled approximately 3 GB data from a data source to a PowerPivot workbook. 64-bit versions are required for Excel and PowerPivot.

Note To show the defined KPIs and hierarchies, SP1 or later versions are required for the SQL Server 2012 PowerPivot for Excel.

For detailed steps to build the PerformancePoint environment, refer to Microsoft Technet topic *Configure PerformancePoint Services for a BI test environment*.

Test and validation

Overview

This solution validates the performance and functionality of running DSS workloads on a data warehouse of columnstore data format in a virtualized SQL Server 2012 environment. It also validates the storage performance optimization function with FAST.

Test scenarios

This solution tested the following scenarios:

- **Scenario 1: Baseline performance with rowstore indexing on VNX5300 without enabling FAST**

  Run the baseline performance test on the data warehouse without enabling FAST.

- **Scenario 2: Baseline performance with columnstore indexing without enabling FAST**

  Run the same baseline performance test on the same dataset with columnstore indexing without enabling FAST; compare the performance between the rowstore and the columnstore.

- **Scenario 3: Baseline performance with columnstore indexing with FAST enabled**

  Run the same baseline performance test on the same dataset with columnstore indexing with FAST enabled; compare the performance before and after enabling FAST.

- **Scenario 4: BI showcasing on SQL Server 2012 Data Warehouse**

  Design the BISM tabular model based on the data warehouse storing the sales data sample of EMC Flash storage arrays and showcase the Power View report based on the tabular model, the PerformancePoint dashboard, scorecard, and reports in SharePoint, as well as the model in PivotTable and PivotCharts in Excel.

- **Scenario 5: SQL 2012 Data Warehousing protection by using Replication Manager with SnapView snap**

  Show the protection for data warehouse through EMC Replication Manager with EMC SnapView snap technology.

Test objectives

The objectives of this solution test involve:

- Capturing the benefits of using columnstore indexing for DSS workloads in SQL Server 2012.
- Demonstrating the FAST technology benefits for SQL Server DSS workloads.
- Demonstrating the various BI capabilities on SQL Server 2012 by leveraging the EMC infrastructure.
- Illustrating effective, lightweight data warehousing protection through EMC Replication Manager with EMC hardware snapshot technology.

Test procedure

The EMC solution team conducted a series of tests by running DSS workloads against the target data warehouses:

1. Build two rowstore TPC-H-like data warehouses on SQL Server 2012, and run the baseline performance test without enabling FAST.
2. Load data from each rowstore data warehouse to a new data warehouse, and create columnstore indexes on the tables in the new data warehouses. Run the same baseline performance test, compare the performance between the rowstore and columnstore, and observe how much the performance benefits from using the columnstore.
3. Enable FAST on the storage pool by adding SATA Flash and Near-line SAS (NL SAS) disks, put the tempDB data LUNs on the Extreme Performance tier, and enable auto-tiering for all the data LUNs. Run the same baseline performance test, and compare the performance before and after enabling FAST.
4. Build the Power View infrastructure, configure PowerPivot for SharePoint 2010 and for Excel, design the BISM tabular model based on an artificial sales data sample of EMC Flash storage arrays; design and showcase the model in the Power View report, dashboard, scorecard, and reports in SharePoint, and also show the model and reports in Excel.
5. Install and configure the Replication Manager server and host server running SQL Server 2012 storing the sales data sample of EMC Flash storage arrays, configure the reserved LUN pool for the snapshots, and create and run a SnapView snap job to generate a snapshot for the data warehouse; after the data changes occur, restore the data warehouse from the snapshot back to the original status.

Note The sample sales data of EMC Flash storage arrays used in this white paper is not real and is for testing purpose only.

Test results

Effects of columnstore on DSS workloads

Columnstore indexes significantly speed up the processing time of common data warehousing queries. Typical data warehousing workloads involve summarizing large amounts of data.

This section describes the test results of scenario 1 and 2 listed in the Test scenario section. It compares the performance of the same DSS workloads running on the data warehouse in the rowstore data format and the columnstore data format.

In this test scenario, the data warehouse was transitioned from rowstore to columnstore.

The duration of running the same test was signicantly shortened from 8 hours 28 minutes to 1 hour 22 minutes on SQL DSSDW01, and from 10 hours 51 minutes to 1 hour 30 minutes on SQL DSSDW02. Therefore the TPC-H-like power was 7.8 times of the columnstore on SQL DSSDW01, and 10 times of the columnstore on SQL DSSDW02. The test demonstrated that columnstore is much more suitable for common DSS workloads, achieving a much shorter processing time and providing much more powerful processing capability.

After transitioning to columnstore, the CPU usage increased significantly on both SQL Server instances, from 25.65 percent to 39 percent on DSSDW01 and from 19.46 percent to 37 percent on DSSDW02. The committed memory increased by less than 10 percent on both SQL Server instances.

Compared with rowstore, the data warehouse in columnstore saved approximately 10 percent space and 10 percent time to create columnstore indexes.

For a detailed performance data comparison, refer to Table 13.

In this test scenario, the IOPS of the DSS workloads on the rowstore data warehouse was almost the same as the columnstore, while the average IOPS per disk was 178. At the mean time, the I/O size of DSS workloads on the data warehouse in columnstore was larger than in rowstore. Thus the DSS workloads throughput in columnstore was 33.6 percent higher than in rowstore. Refer to Figure 3 for details.

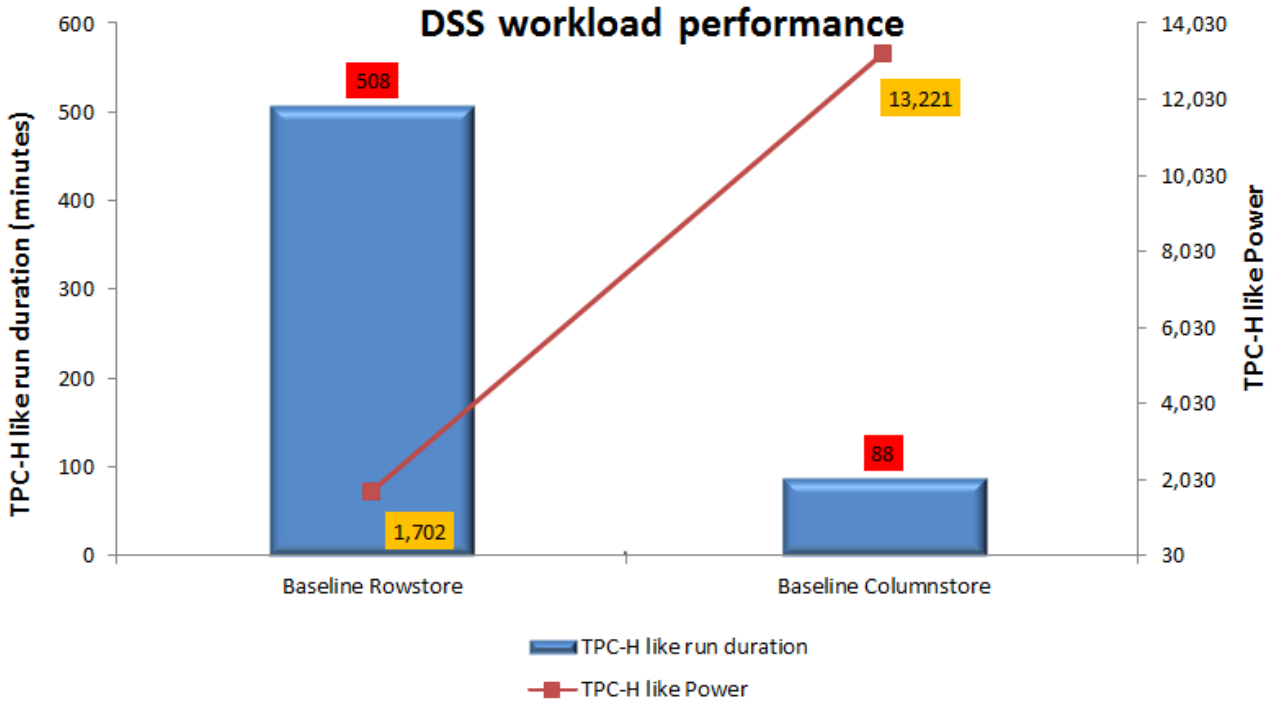


Figure 3. TPC-H-like run duration and TPC-H-like power on SQL DSSDW01

Table 12 shows the space usage and index creation time for data warehouses in rowstore and in columnstore.

Table 12. Space usage and index creation time comparison

Database	Space usage(GB)	Index creation time
Rowstore TPC-H-like data warehouse	990	6 hours 32 minutes
Columnstore TPC-H-like data warehouse	910	5 hours 59 minutes

Enabling EMC FAST on the DSS data pool

EMC FAST enhances performance by leveraging Flash drives. FAST maxmizes the utilization of Flash drives for high-performance workloads and is able to pin LUNs for particular workloads to certain tiers.

This section describes the test results of scenario 3 in Test scenario. It compares the performance of the same DSS workloads running on the data warehouses in the columnstore data format before and after enabling FAST.

Before FAST was enabled, all database LUNs were initially pinned to the RAID 5 SAS tier. After FAST was enabled, the highly active data was automatically moved to the Flash tier, and the inactive data was moved the NL SAS tier.

After enabling FAST, the total storage bandwidth increased from approximately 250 MB/s to 380 MB/s. This presents a 51 percent improvement in storage performance. The test duration running the same DSS workloads decreased from 88 minutes to 54 minutes on DSSDW01, and decreased from 90 minutes to 60 minutes on DSSDW02. The TPC-H-like power increased by approximately 53 percent on DSSDW01 and 23 percent on DSSDW02.

After enabling FAST, the host CPU usage increased from 39 percent to 65 percent on DSSDW01, and increased from 37 percent to 57 percent on DSSDW02. Both were increased by more than 50 percent. This indicates that SQL Server can utilize more CPU because of the increased I/O handling capability enabled by FAST.

Table 13 shows the detailed comparison performance data for scenario 1, 2, and 3 listed in the Test scenario section.

Table 13. Performance data comparison

Performance	Baseline rowstore		Baseline columnstore		Columnstore with FAST enabled	
	DSSDW01	DSSDW02	DSSDW01	DSSDW02	DSSDW01	DSSDW02
Throughput (MB/sec)	105.9	82.4	135.9	115.7	203.4	176
Average read I/O size on data(KB)	8 ~ 120	8 ~ 108	64~150	64~144	64 ~ 115	64 ~ 120

Performance	Baseline rowstore		Baseline columnstore		Columnstore with FAST enabled	
Average read I/O size on tempdb (KB)	8 ~ 64	8 ~ 64	64	64	64	64
Average write I/O size on tempdb(KB)	8 ~ 64	8 ~ 64	64	64	64	64
Host CPU usage (%)	25.65	19.46	39	37	64.18	56.95
Committed Memory in use (%)	52.97	57.98	55.43	61.81	57.19	60.35
TPC-H-like run duration	8h 28m	10h 51m	1h 22m	1h 30m	0h 54m	1h 00m
TPC-H-like power	1,701.9	1,539.7	13,221.2	15,240.1	20,328.7	18,679.6
Storage Processor usage (%)	7		9.9		12.8	
Average total IOPS per disk	177.6		178		SAS:42 SATA Flash: 952	
Average disk utility (%)	SAS: 54		SAS: 56		SAS: 12 SATA Flash: 50	
Total throughput (MB/sec)	188.3		251.6		379.4	

Figure 4 compares the throughput and average hosts CPU usage before and after enabling FAST.

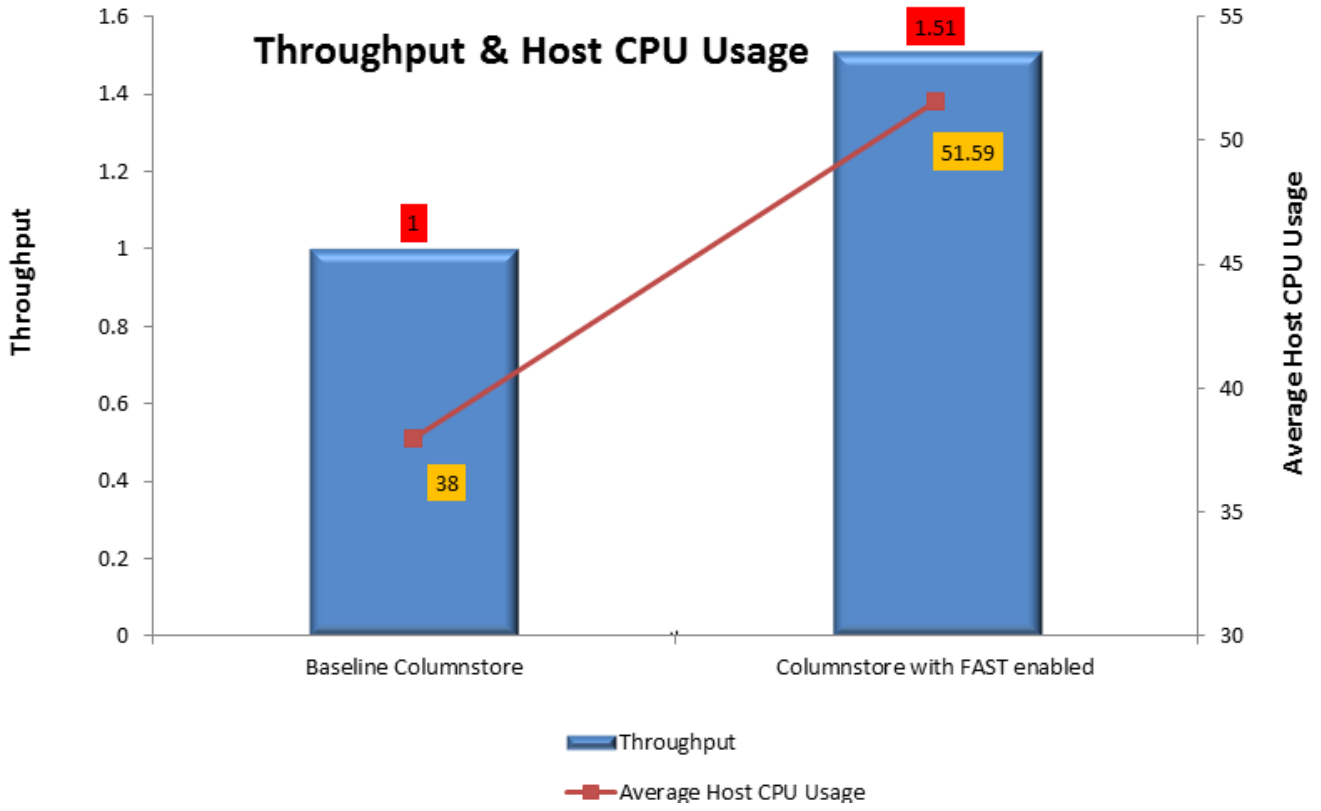


Figure 4. Throughput and host CPU usage comparison

After enabling FAST, the storage reached a balanced state, as shown in Figure 5.

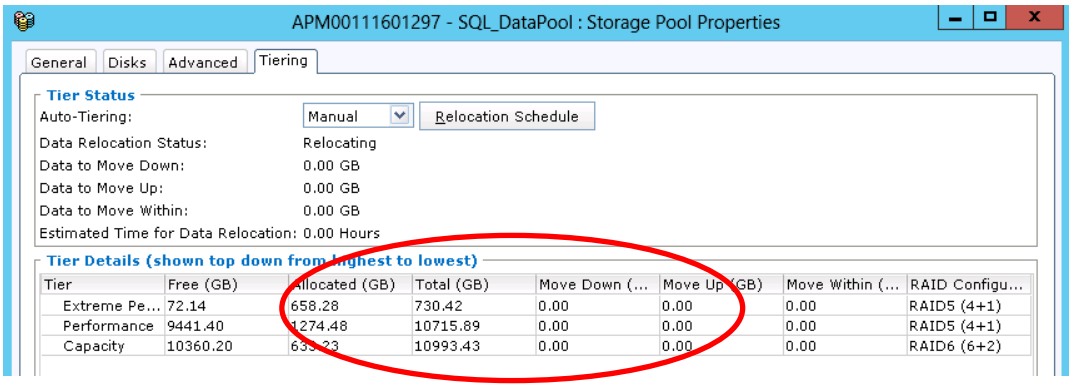


Figure 5. Data pool tier status for data warehouse

Before enabling FAST, the database LUN was bound to the SAS tier. All data was on the SAS tier, as shown in Figure 6.

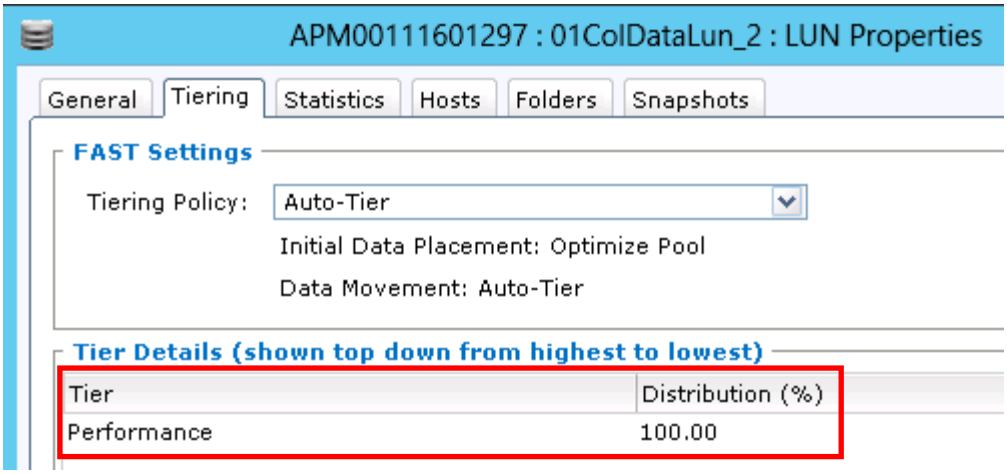


Figure 6. LUN tiering status before enabling FAST

After enabling FAST, the data was on different tiers, as shown in Figure 7.

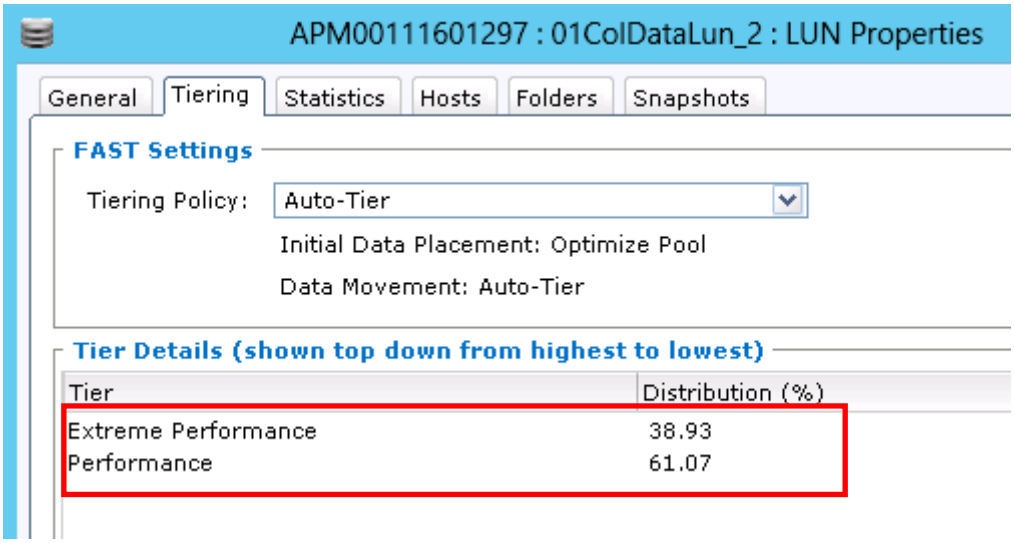


Figure 7. LUN tiering status after enabling FAST

We moved the tempdb data LUNs to the Extreme Performance tier to guarantee the high performance for the random reads/writes, as shown in Figure 8. The FAST design for SQL Server 2012 Data Warehouse section contains detailed storage design information for FAST in this solution.

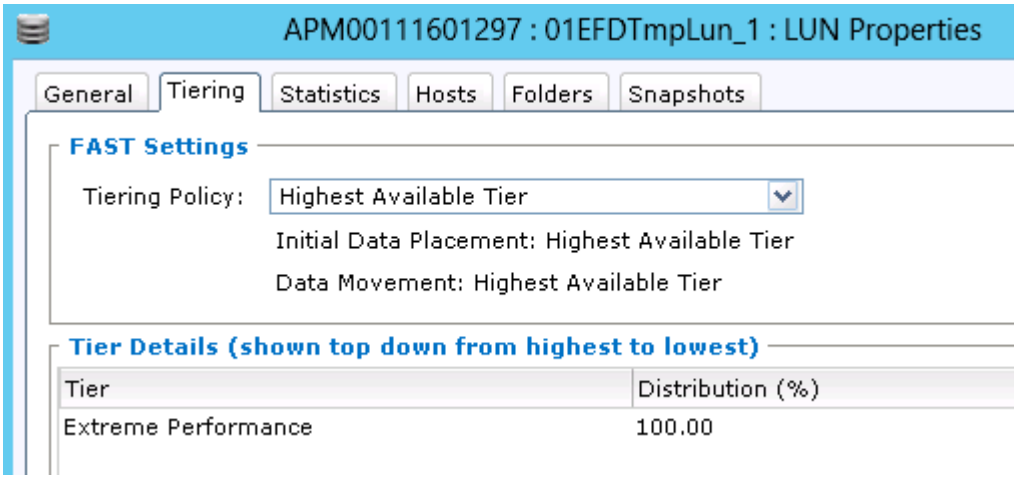


Figure 8. Tempdb data LUN tiering status

After enabling FAST, the disk utilization of the SAS tier dropped tremendously to approximately 10 percent. The average utilization of the Flash tier was approximately 50 percent with spiking to more than 90 percent, which accelerated the performance significantly. Figure 9 shows the utilization of SAS disks and Flash disks after enabling FAST from Unisphere® Performance Analyzer.

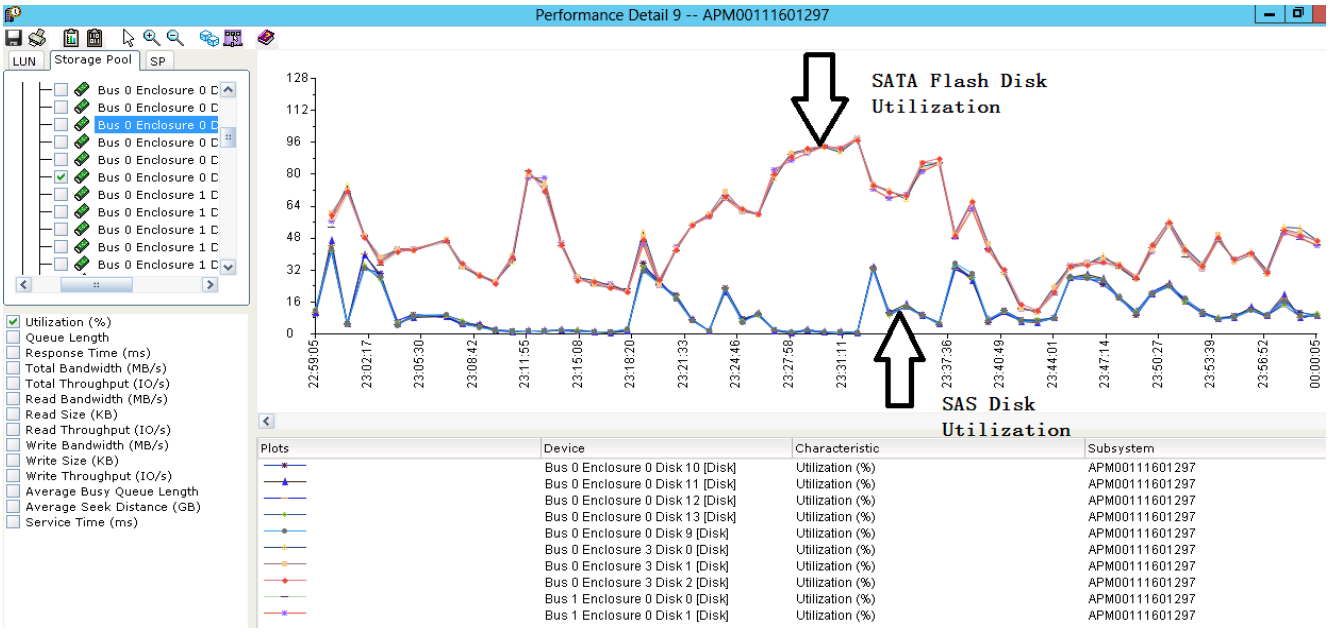


Figure 9. Disk utilization after enabling FAST

After enabling FAST, most of the bandwidth came from the Flash tier. The workloads pushed to the SAS tier dropped notably because the hottest data was moved to the Flash tier automatically.

Figure 10 shows the total bandwidth of SAS disks and Flash disks after enabling FAST from Unisphere Performance Analyzer.

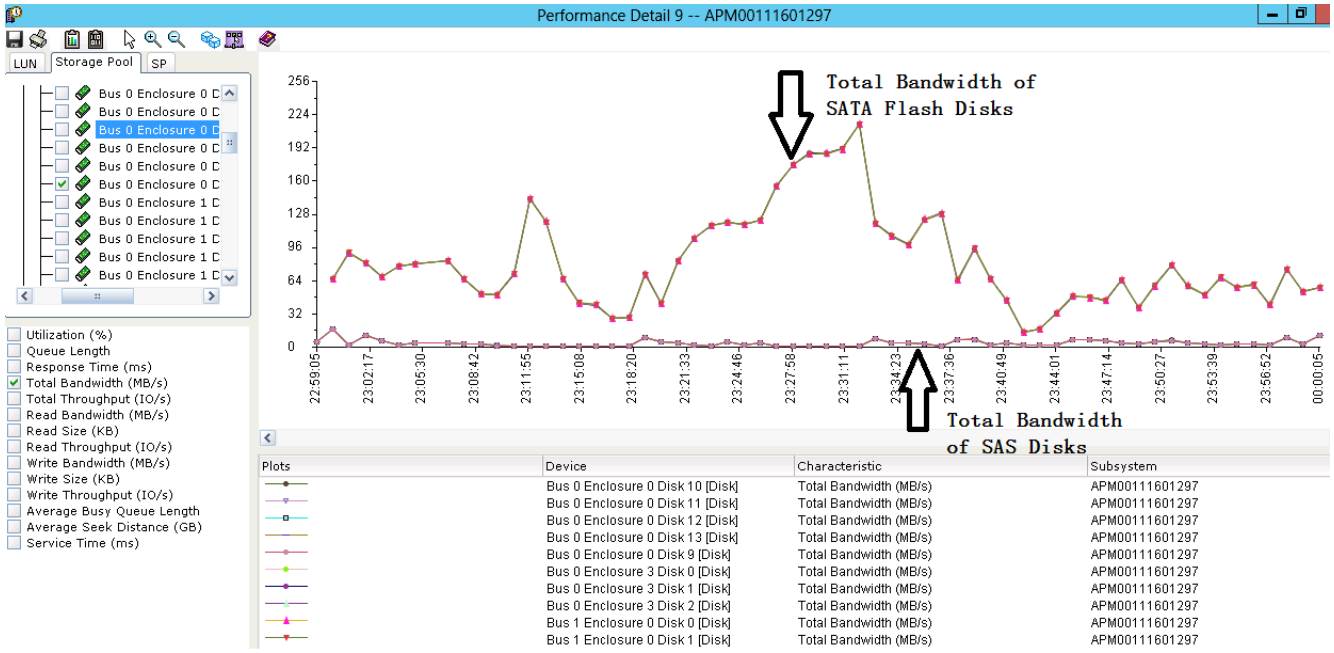


Figure 10. Total bandwidth after enabling FAST

After enabling FAST, most of I/Os went to the Flash tier; each Flash disk handled around 1,000 I/Os with spiking to more than 3,000 I/Os. Figure 11 shows the total throughput (I/Os) of SAS disks and Flash disks after enabling FAST from Unisphere Performance Analyzer.

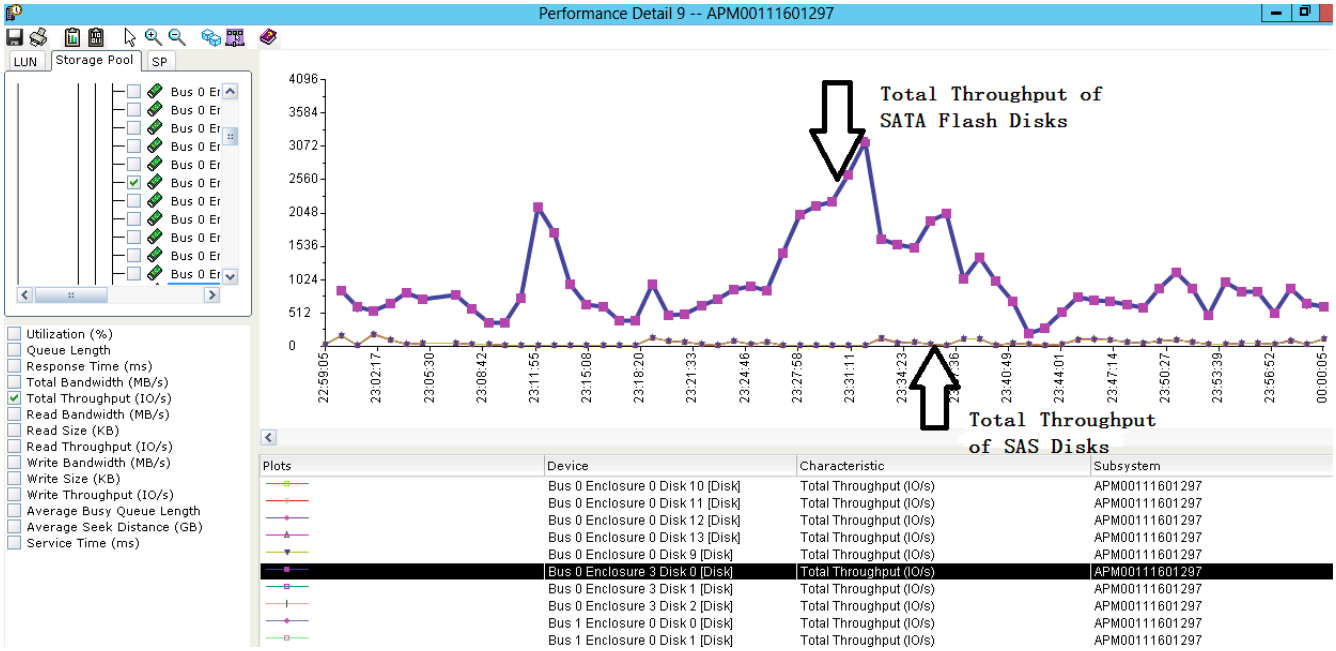


Figure 11. Total throughput after enabling FAST

Hyper-V host server performance

From the Hyper-V perspective, the Hyper-V host servers that hosted the SQL Server 2012 virtual machines consistently operated with excellent performance. In test Scenario 3 listed in the Test scenario section, the Hyper-V host servers pushed higher loads, which achieved higher CPU usage and higher throughput.

Table 14 lists the performance counters that we monitored on the hyper-V host servers including Guest Run Time, Hypervisor Run Time, and Total Run Time, which were used to measure the processor utilization for all virtual machines running on the host servers and the processor utilization for the hypervisor on the entire system. The processor utilization was in a healthy state on all Hyper-V nodes.

Table 14. Hyper-V processor performance in performance test

Performance counter	Target	Node 1	Node 2
Hyper-V hypervisor logical processor\% Guest Run Time	‹65%	48.45%	43.66%
Hyper-V hypervisor logical processor\% Hypervisor Run Time	‹5%	3.11%	3.06%
Hyper-V hypervisor logical processor\% Total Run Time	‹70%	51.56%	46.72%

BI showcase

This section describes the test results of scenario 4 listed in the Test scenario section. In this solution, we used the SharePoint server as the presentation layer for the entire BI environment.

We implemented a BI solution with Power View reports to visually explore data based on a Tabular BI Semantic Model in SQL 2012 Analysis Server that analyzes the sales data sample.

We also designed a complete BI and performance management solution by using PerformancePoint server to report and analyze the sales data sample in a PowerPivot workbook.

Finally, we showcased the defined Tabular model in the PowerPivot workbook through the PivotTables and PivotCharts in Excel 2010.

Note The sample sales data of EMC Flash storage arrays used in this white paper is not real and is for testing purpose only.

Power View reports

In this solution, we configured Power View infrastructure and defined a BI Semantic Model in SQL 2012 Analysis Server in **Tabular** mode including the following items, and generated a Power View report by using the model as the data source.

The model includes the following items:

Hierarchies

We generated a generic time dimension table, and created a hierarchy with year, quarter, month, and day on the time dimension table. Also, we created a Region hierarchy involving Region and Nation levels for slicing.

Measures

In this model, we defined several measures based on the columns within the fact table:

- **Sales total**: Calculates the total sales amount for the current year. Here is the formula:

```
SalesTotal:=CALCULATE(SUM(H_Lineitem[SalesAmount]),
H_Lineitem[l_discount] >= 0)
```

- **Sales total in previous year**: Calculates the total sales amount of the last year. Here is the formula:

```
SalesTotalPrevYr:=CALCULATE([SalesTotal],
DATEADD(DimDate[FullDate], -1, YEAR))
```

- **YOYGrowth**: Calculates the year over year's growth of the total sales amount. Here is the formula:

```
YOYGrowth:=([SalesTotal]-
[SalesTotalPrevYr])/[SalesTotalPrevYr]
```

KPI

In this model, we have defined a Key Performance Indicator (KPI) based on the measure YOYGrowth:

- If the value of this measure is less than zero, the KPI showed the red light.
- If the value is greater than 0.1, the KPI showed the green light.
- Otherwise, the KPI showed the yellow light.

As shown in Figure 12, based on the sales data sample of EMC Flash storage arrays, we defined a BI Semantic Model in Tabular mode to perform data analysis, and viewed and compared the result from various perspectives.

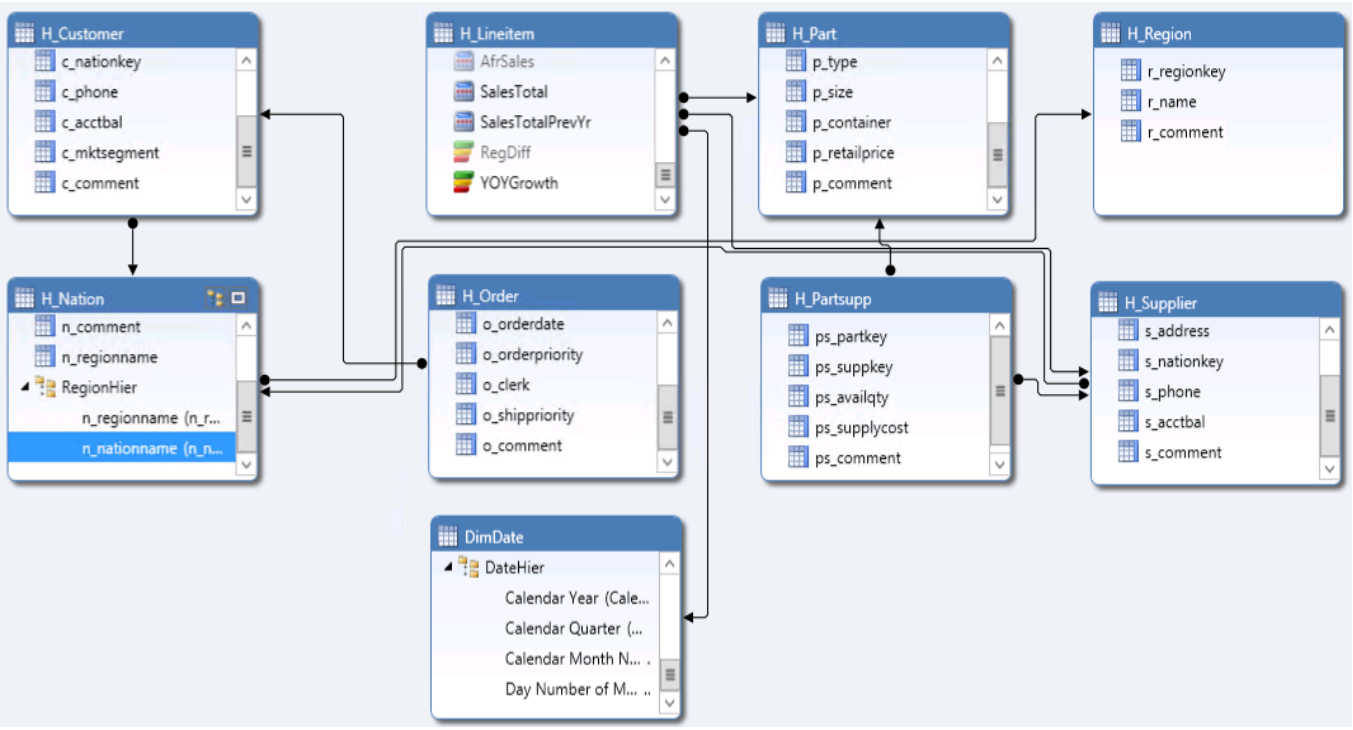


Figure 12. Data model in Tabular mode

Reports

In this solution, we designed three views in a Power View report to show the defined BI Semantic Model.

The following view includes four charts for data from various perspectives, as shown in Figure 13. In the upper-left corner, there is a Card report showing the suppliers, shipmode, and the average discount rate. In the upper-right corner, you can find a tile chart that calculates the available quantity by region. Below that, there is a pie chart showing the supplier costs by nation. And the lower-left corner shows a Microsoft bing map. Similarly, you can drill through the detailed data by using the defined filters on the right-hand side.

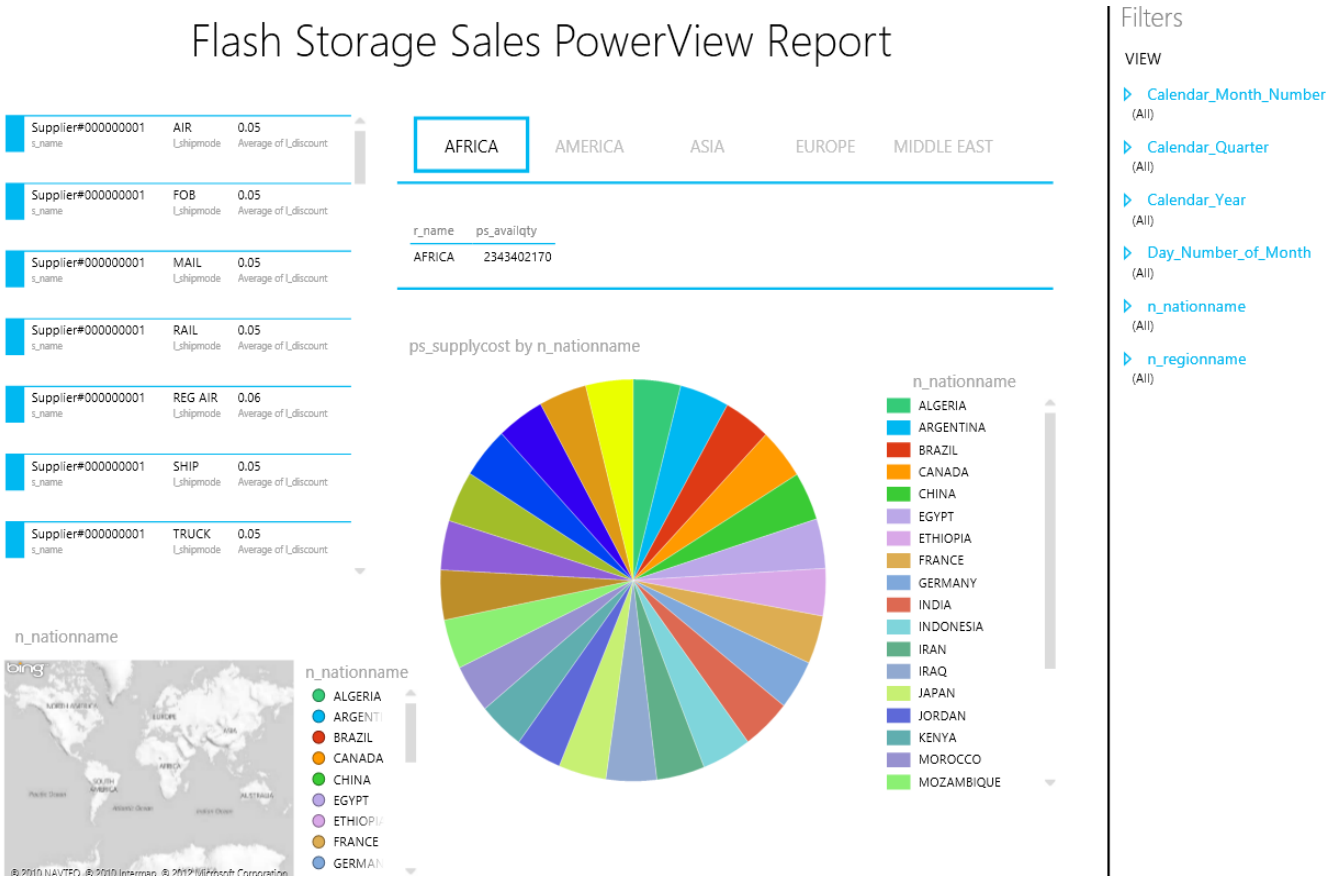


Figure 13. Flash storage sales PowerView report

For the other two views in the Power View report, refer to Appendix.

PerformancePoint BI

In this solution, we created a Business Intelligence Center in SharePoint Server 2010 to store items such as reports, PerformancePoint scorecards, and PerformancePoint dashboards. We defined a tabular BI Semantic Model in a PowerPivot workbook as the data source, and uploaded it to the SharePoint server. We designed the KPI, scorecard, report, and dashboard based on the PowerPivot workbook in the Dashboard Designer, and uploaded it to the SharePoint server.

Note As part of this solution, we used these data examples to illustrate sample results.

We defined same measures, hierarchies, and KPI as in the Power View report.

Based on the sales data sample of EMC Flash storage arrays, we have defined the BI Semantic Model in Tabular mode in the PowerPivot workbook to view and compare the result from various perspectives. Figure 14 shows the data model.

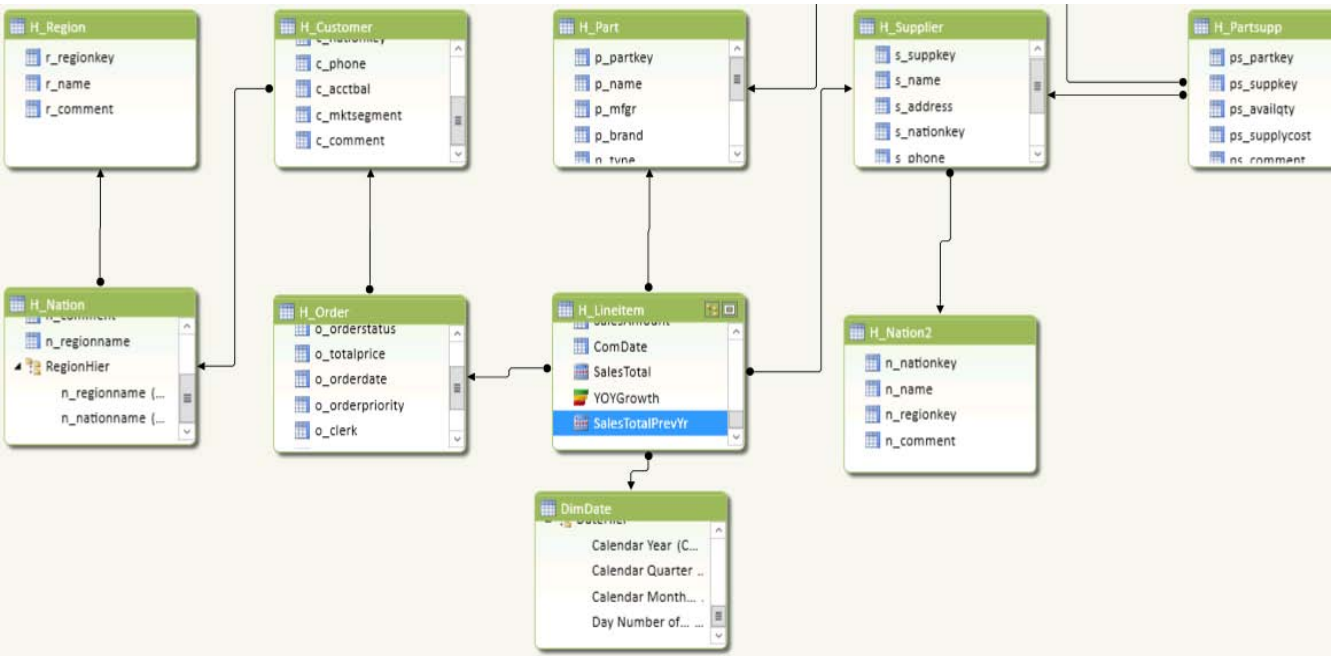
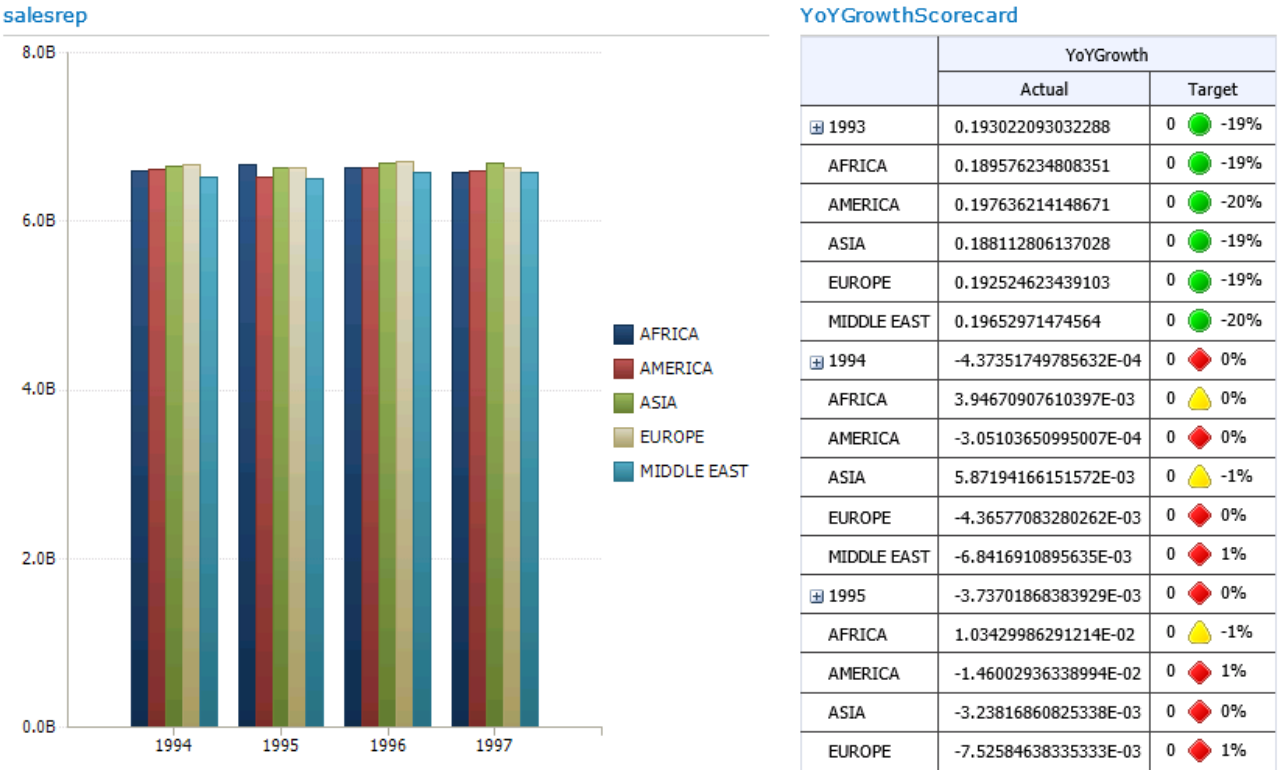


Figure 14. Data model in Tabular mode

Based on the data model, we designed a scorecard, a report, and a dashboard. For the scorecard and report, refer to Appendix.

Dashboard

In the Dashboard Designer, we defined a dashboard to show the sales report together with the year over year growth scorecard, as shown in Figure 15.



	YoYGrowth	
	Actual	Target
1993	0.193022093032288	0 -19%
AFRICA	0.189576234808351	0 -19%
AMERICA	0.197636214148671	0 -20%
ASIA	0.188112806137028	0 -19%
EUROPE	0.192524623439103	0 -19%
MIDDLE EAST	0.19652971474564	0 -20%
1994	-4.37351749785632E-04	0 0%
AFRICA	3.94670907610397E-03	0 0%
AMERICA	-3.05103650995007E-04	0 0%
ASIA	5.87194166151572E-03	0 -1%
EUROPE	-4.36577083280262E-03	0 0%
MIDDLE EAST	-6.8416910895635E-03	0 1%
1995	-3.73701868383929E-03	0 0%
AFRICA	1.03429986291214E-02	0 -1%
AMERICA	-1.46002936338994E-02	0 1%
ASIA	-3.23816860825338E-03	0 0%
EUROPE	-7.52584638335333E-03	0 1%

Figure 15. Sales report with the year over year growth scorecard

PowerPivot for Excel

PowerPivot for Excel is a convenient tool for designing personal BI models and reports. In this solution, we created a tabular BI Semantic Model in a PowerPivot workbook, and designed some PivotTables and PivotCharts that connect the workbook as the data source to view the data.

In this workbook, we defined the same data model as in the previous workbook defined in the PerformancePoint BI environment.

We created some PivotTables and PivotCharts to view and analyze the data. For details about the PivotTables and PivotCharts, refer to Appendix.

Data protection

This section describes test results of scenario 5 listed in the Test scenario section. In this solution, we created a SnapView snap job in the Replication Manager, ran the job to generate a snapshot for the protected data warehouse, and made some data changes on the data warehouse. At last, we restored the snapshot back to the status when the snapshot was created on the data warehouse. Table 15 shows the reserved LUN pool usage.

Table 15. Reserved LUN pool usage

Source LUN name	Source LUN size	Snapshot LUN size	Snapshot LUN usage
Data LUN 1	100 GB	10 GB	21%
Data LUN 2	100 GB	10 GB	21%
Data LUN 3	100 GB	10 GB	21%
Data LUN 4	100 GB	10 GB	21%
Data LUN 5	100 GB	10 GB	21%
Data LUN 6	100 GB	10 GB	21%
Log LUN	200 GB	10 GB x 7	95%

The Replication Manager took only 7 minutes and 27 seconds for a SnapView snap job to create a snapshot. This job included the snapshot preparation, and created a Virtual Device Interface (VDI) session to generate a snapshot. It took only 5 minutes and 39 seconds to restore 800 GB data LUNs with 80 GB data changes as described in Table 16.

Table 16. Snapshot creation and restore duration

DW size protected	Data changed	Snapshot creation time	Snapshot restore time
800 GB	80 GB	7 minutes and 27 seconds	5 minutes and 39 seconds

Conclusion

Summary

EMC VNX5300 is capable of accommodating the storage performance and capacity requirements of the DSS workloads in a virtualized environment running SQL Server 2012. VNX5300 provides protection and rapid data recovery by leveraging EMC hardware snapshot technology.

- Environments with complex business intelligence, data warehousing, and collaboration workloads can be easily managed and automatically optimized in a continuous manner with EMC VNX and EMC FAST technology.
- EMC Replication Manager with EMC hardware snapshot technology accelerated and offloaded protection for data warehousing.
- Columnstore indexes significantly speed up the processing time of common data warehousing queries. Typical data warehousing workloads involve summarizing large amounts of data.
- The BI Semantic Model offers powerful analytical capabilities, simplicity, and familiarity of tabular models. Power View enables users to analyze and visually explore their data quickly and easily. PerformancePoint Services provide a flexible and easy-to-use BI solution.

Findings

The key findings of this proven solution are as follows:

- By enabling FAST, the performance of DSS workloads improved significantly. The total bandwidth increased by more than 50 percent, while the test processing time was shortened by one third.
- EMC Replication Manager with EMC SnapView snap technology provided protection for data warehousing. It enabled rapid data recovery and reduced complexity in setup. It took only 5 minutes and 39 seconds to restore 800 GB data warehouse LUNs with 80 GB data changes.
- The DSS workloads running on the columnstore data warehouse took much less time than on the rowstore data warehouse, and achieved more than 30 percent higher throughput utilizing larger I/O transfer.
- The BI Semantic Model enabled the Tabular model, which was easy to configure and understand. It was widely used by Power View, PowerPivot for SharePoint, and PowerPivot for Excel.
- Power View is a self-service BI tool, which enabled users to customize the reports, configure the infrastructure, and analyze the data from various perspectives easily.
- KPI, PerformancePoint dashboards, and PerformancePoint scorecards in SharePoint Server 2010 were easy to configure, design, and deploy. PerformancePoint enabled users to explore data in a flexible way.

References

White papers

For additional information, see the white papers listed below.

- *EMC CLARiiON Reserved LUN Pool Configuration Considerations*
- *EMC FAST VP for Unified Storage Systems*
- *EMC Mission Critical Infrastructure for Microsoft SQL Server 2012*

Product documentation

For additional information, see the product documents listed below.

- *EMC VNX Series Unified Storage Systems* (http://www.emc.com/collateral/software/specification-sheet/h8514-vnx-series-ss.pdf)

Other documentation

For additional information, see the documents listed below.

- *Add an Additional Reporting Services Web Front-end to a Farm* (http://technet.microsoft.com/en-us/library/hh479775.aspx)
- *Configure PerformancePoint Services for a BI test environment* (http://technet.microsoft.com/en-us/library/hh223268.aspx)
- *Optimizing tempdb Performance* (http://msdn.microsoft.com/en-us/library/ms175527.aspx)
- *PowerPivot for SharePoint Installation (SharePoint 2010)* (http://technet.microsoft.com/en-us/library/ee210654.aspx)
- *Power View Infrastructure Configuration and Installation* (http://msdn.microsoft.com/en-us/library/jj591528.aspx)
- *Pre-Configuration Database Optimizations* (http://msdn.microsoft.com/en-us/library/ee377059(v=bts.10).aspx)
- *SQL Server Best Practices* (http://msdn.microsoft.com/en-us/library/cc966412.aspx)
- *Tuning options for SQL Server when running in high performance workloads* (http://support.microsoft.com/kb/920093)

Appendix

Disabling SCSI filtering

In this solution, we installed EMC Solutions Enabler in a virtual machine hosted by Hyper-V with pass-through storage devices. EMC supports the installation of EMC Solutions Enabler on a child virtual machine with pass-through storage devices only when the parent partitions are running Windows Server 2012, and when the appropriate settings for the virtual machine have been configured.

EMC Solutions Enabler implements the usage of the extended SCSI commands, which are filtered by the parent partition by default. A bypass of the filtering is provided with Windows Server 2012 Hyper-V. To disable the filtering of SCSI commands, the administrator can execute the following PowerShell script on the parent partition of all the Hyper-V hosts in which SQL Server virtual machines reside, and then restart the virtual machines respectively:

```
// PowerShell Script: Set_SCCI_Passthrough.ps1
$Target = $args[0]
$VSManagementService=gwmi MSVM_VirtualSystemManagementService -
namespace "root\virtualization"
foreach($Child in Get-WmiObject -Namespace root\virtualization
Msvm_ComputerSystem -Filter "ElementName='$Target'")
{
$VMData=Get-WmiObject -Namespace root\virtualization -Query
"Associators of {$Child} Where
ResultClass=Msvm_VirtualSystemGlobalSettingData
AssocClass=Msvm_ElementSettingData"
$VMData.AllowFullSCSICommandSet=$true
$VSManagementService.ModifyVirtualSystem($Child,$VMData.PSBase.Get
Text(1))|out-null
}
```

The usage of the script under PowerShell execution environment is as follows:

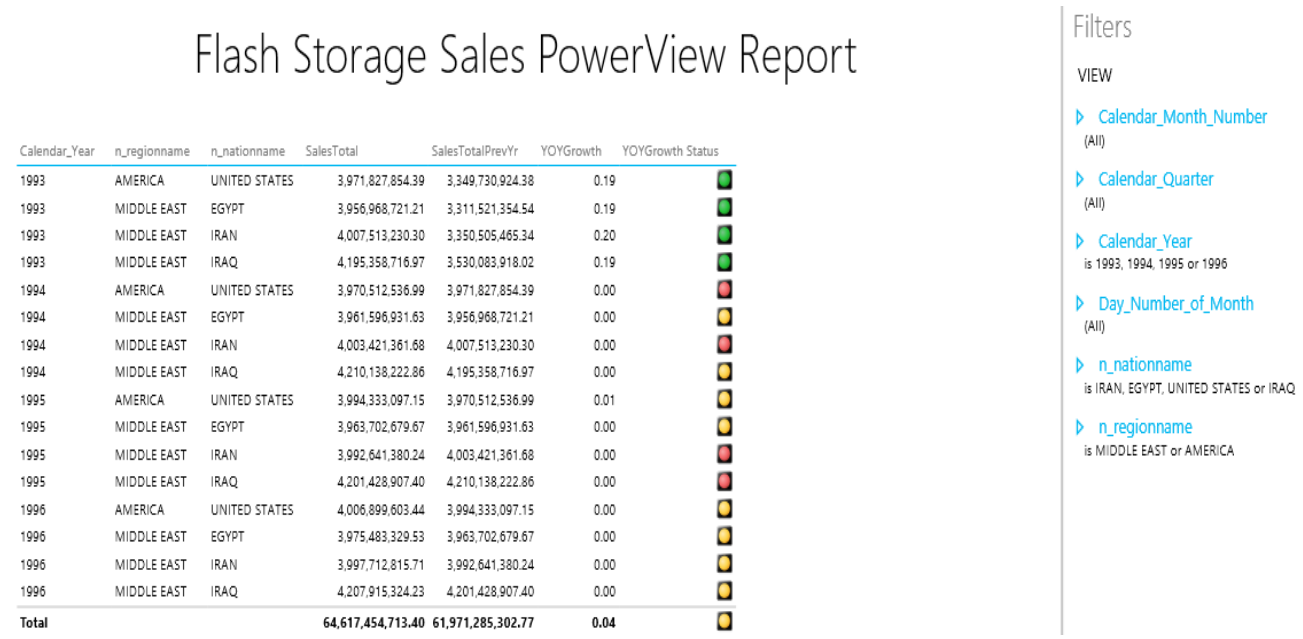
```
.\Set_SCCI_Passthrough.ps1 ManagedVirtualMachineName
```

Use the following script to check the current value of the SCSI filtering. Provide the name of the virtual machine target on which you want to report:

```
// PowerShell Script: Get_SCCI_Passthrough.ps1
$Target=$args[0]
foreach ($Child in Get-WmiObject -Namespace root\virtualization
Msvm_ComputerSystem -Filter "ElementName='$Target'")
{
$VMData=Get-WmiObject -Namespace root\virtualization -Query
"Associators of {$Child}
Where ResultClass=Msvm_VirtualSystemGlobalSettingData
AssocClass=Msvm_ElementSettingData"
Write-host "VirtualMachine:"$VMData.ElementName
Write-Host
"CurrentlyByPassingSCSIFiltering:"$VMData.AllowFullSCSICommandSet
}
```

Power View report

The following view shows the sales amount and year over year growth by date across regions and nations in a table. You can view the detailed sales data by drilling through the date and the nations, as shown in Figure 16.



Calendar_Year	n_regionname	n_nationname	SalesTotal	SalesTotalPrevYr	YOYGrowth	YOYGrowth Status
1993	AMERICA	UNITED STATES	3,971,827,854.39	3,349,730,924.38	0.19	
1993	MIDDLE EAST	EGYPT	3,956,968,721.21	3,311,521,354.54	0.19	
1993	MIDDLE EAST	IRAN	4,007,513,230.30	3,350,505,465.34	0.20	
1993	MIDDLE EAST	IRAQ	4,195,358,716.97	3,530,083,918.02	0.19	
1994	AMERICA	UNITED STATES	3,970,512,536.99	3,971,827,854.39	0.00	
1994	MIDDLE EAST	EGYPT	3,961,596,931.63	3,956,968,721.21	0.00	
1994	MIDDLE EAST	IRAN	4,003,421,361.68	4,007,513,230.30	0.00	
1994	MIDDLE EAST	IRAQ	4,210,138,222.86	4,195,358,716.97	0.00	
1995	AMERICA	UNITED STATES	3,994,333,097.15	3,970,512,536.99	0.01	
1995	MIDDLE EAST	EGYPT	3,963,702,679.67	3,961,596,931.63	0.00	
1995	MIDDLE EAST	IRAN	3,992,641,380.24	4,003,421,361.68	0.00	
1995	MIDDLE EAST	IRAQ	4,201,428,907.40	4,210,138,222.86	0.00	
1996	AMERICA	UNITED STATES	4,006,899,603.44	3,994,333,097.15	0.00	
1996	MIDDLE EAST	EGYPT	3,975,483,329.53	3,963,702,679.67	0.00	
1996	MIDDLE EAST	IRAN	3,997,712,815.71	3,992,641,380.24	0.00	
1996	MIDDLE EAST	IRAQ	4,207,915,324.23	4,201,428,907.40	0.00	
Total			**64,617,454,713.40**	**61,971,285,302.77**	**0.04**	

Figure 16. The first view of Flash storage sales Power View report

In the following view, we designed two charts. As shown in Figure 17, the column chart shows the sales amount by year across regions. The bar chart shows the detailed sales amount by drilling through to each quarter per nation.

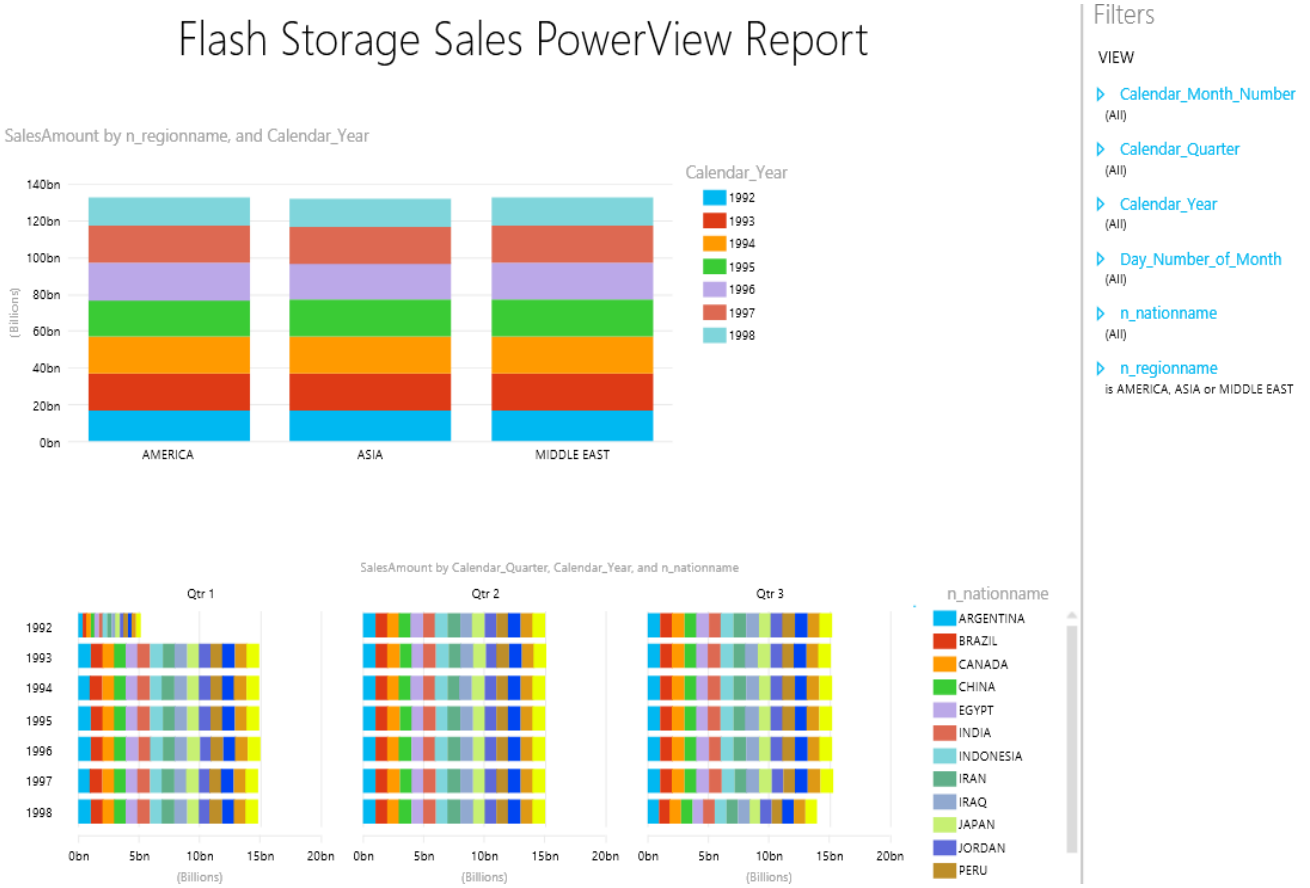


Figure 17. The second view of Flash storage sales PowerView report

PerformancePoint scorecard and report

Scorecard

Based on the data model, we defined a scorecard to show the year over year growth across different regions over the years, as shown in Figure 18. Also you can expand the date to view the data in smaller granularity.

YoYGrowthScorecard

	YoYGrowth	
	Actual	Target
⊞ 1993	0.193022093032288	0 -19%
AFRICA	0.189576234808351	0 -19%
AMERICA	0.197636214148671	0 -20%
ASIA	0.188112806137028	0 -19%
EUROPE	0.192524623439103	0 -19%
MIDDLE EAST	0.19652971474564	0 -20%
⊞ 1994	-4.37351749785632E-04	0 0%
AFRICA	3.94670907610397E-03	0 0%
AMERICA	-3.05103650995007E-04	0 0%
ASIA	5.87194166151572E-03	0 -1%
EUROPE	-4.36577083280262E-03	0 0%
MIDDLE EAST	-6.8416910895635E-03	0 1%
⊞ 1995	-3.73701868383929E-03	0 0%
AFRICA	1.03429986291214E-02	0 -1%

Figure 18. YoYGrowth scorecard

Report

Based on the data model, we defined a sales report over the years across different regions, as shown in Figure 19.

Figure 19. Sales report of different regions

PivotTable and PivotChart on Excel

We used a PivotTable to show the sales amount and the year over year growth, as shown in Figure 20.

1	RegionHier	(Multiple Items)			
2					
3	**Row Labels**	**SalesTotal**	**SalesTotalPrevYr**	**YOYGrowth Value**	**YOYGrowth Status**
4	**AMERICA**	**10436249387**	**9386520484**	**0.111833656**	
5	**ARGENTINA**	**2082617367**	**1870327241**	**0.113504269**	
6	1993	1322399706	1104134088	0.197680355	
7	1994	760217660.9	766193152.8	-0.007798937	
8	**Qtr 1**	**318747941.2**	**327531326.9**	**-0.026816933**	
9	**Qtr 2**	**324117014.3**	**324686730.1**	**-0.001754663**	
10	04	105908959.5	105116914.3	0.007534898	
11	05	110883314.5	113316102.3	-0.021469039	
12	06	107324740.3	106253713.6	0.010079899	
13	**Qtr 3**	**117352705.4**	**113975095.7**	**0.029634629**	
14	**BRAZIL**	**2108324848**	**1892970738**	**0.113765156**	
15	1993	1330741539	1108438449	0.200555197	
16	1994	777583309.1	784532289.3	-0.008857481	
17	**Qtr 1**	**327568655.4**	**339244073.9**	**-0.034415984**	
18	**Qtr 2**	**334716728.9**	**329988256.1**	**0.014329215**	
19	04	111277948.7	108309499.7	0.027407098	
20	05	112456799.9	112908418.9	-0.003999871	
21	06	110981980.3	108770337.5	0.020333143	
22	**Qtr 3**	**115297924.8**	**115299959.2**	**-1.76446E-05**	

Figure 20. PivotTable to show the sales amount and the year over year growth

We also used a pie PivotChart to show the calculated total supplier cost for the selected nations, as shown in Figure 21.

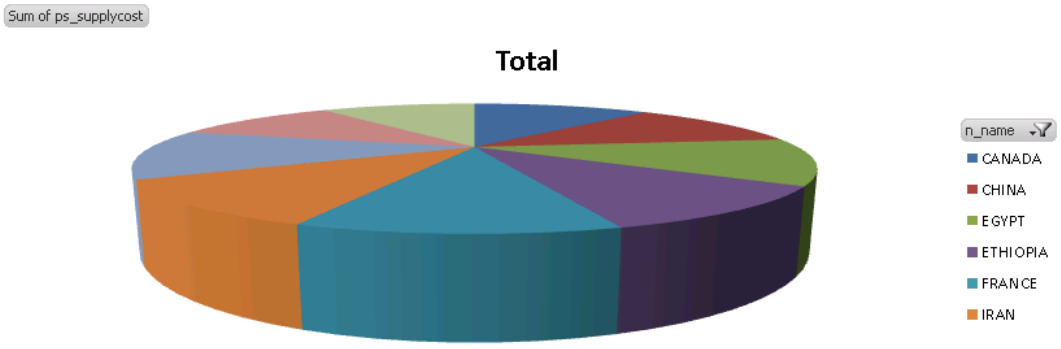


Figure 21. Pie PivotChart to show the calculated total supplier cost