

IT Audit

Microsoft adCenter Business Intelligence Infrastructure

Powered by EMC Symmetrix VMAX and Microsoft SQL Server 2008

By Brian Garrett, Vice President, ESG Lab

November 2010

This ESG IT Audit was commissioned by EMC
and is distributed under license from ESG.

Contents

Executive Summary	3
Background	4
Introducing Microsoft adCenter	5
adCenter Business Intelligence.....	5
How it Works	6
The Challenges.....	7
The Solution	9
The Benefits	11
The Bigger Truth	13

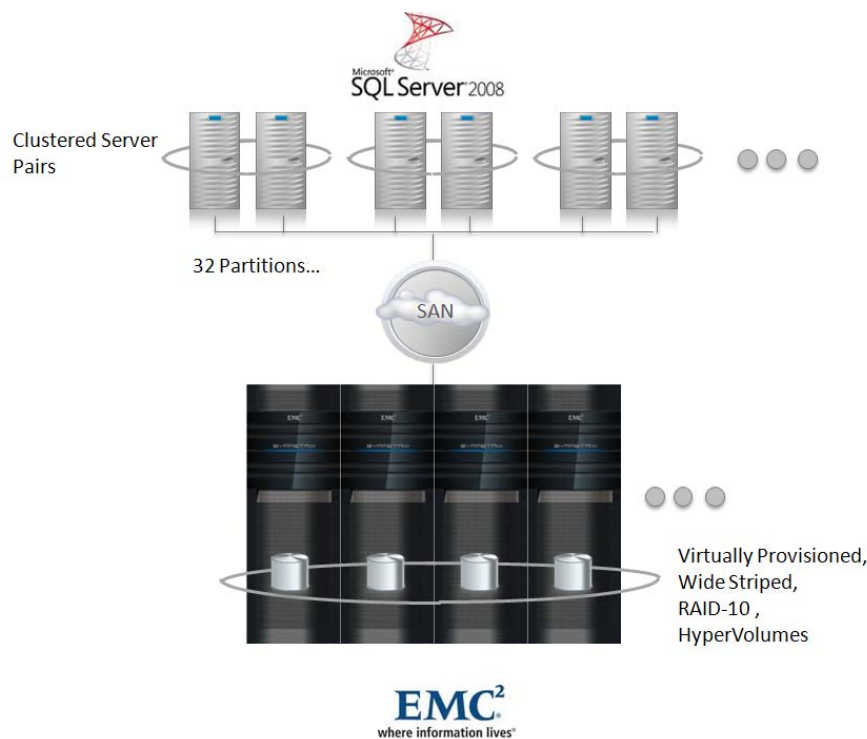
All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. Copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482.0188.

Executive Summary

EMC Symmetrix Storage and Microsoft SQL Server 2008 have been used to create the massively scalable foundation for the Microsoft adCenter Business intelligence (BI) platform. The adCenter platform, which delivered \$1.9 billion in online advertising revenue¹ from Microsoft and partner properties (e.g., Bing, MSN) in 2009, was recently upgraded to accommodate a massive wave of new growth as Yahoo! search and online advertising began moving to the platform in 2010.

This report documents the results of an IT Audit based on an ESG Lab visit with the adCenter BI team in Bellevue, Washington. Interviews with Microsoft and EMC employees were used to assess the extreme scalability of the adCenter BI infrastructure. This report highlights the value of using a building block approach which leverages multiple Microsoft SQL Server 2008 partitions residing on virtually provisioned, wide-striped, EMC Symmetrix VMAX hyper volumes to create a scalable and reliable platform for advanced business intelligence (see Figure 1).

Figure 1. Microsoft SQL Server 2008 and EMC VMAX in the Microsoft adCenter BI Infrastructure



This report also documents the results of a storage consolidation initiative which Microsoft refers to as “Project Snoqualmie” in reference to the fact that the project's energy savings could be used to power all the homes in Snoqualmie, WA. Consolidating with Symmetrix VMAX and leveraging the platform’s virtual provisioning, automated management, and high density drive options, adCenter has implemented a capacity-on-demand model that:

- Increased storage utilization from 34% to 70%
- Reduced storage rack space by 91%
- Reduced storage energy requirements by 92%
- Reduced the total cost of storage ownership over five years by 69%

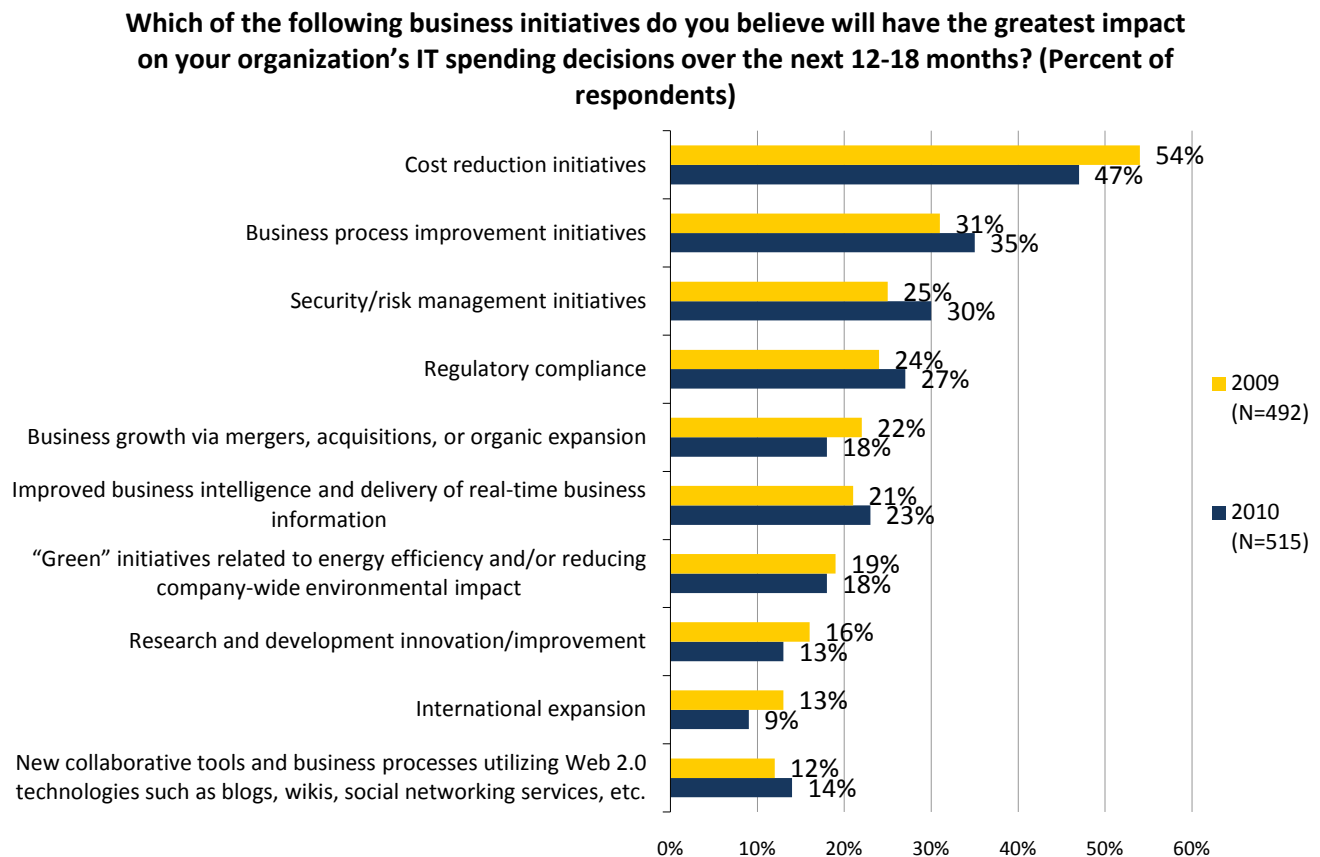
¹ Source: Microsoft 10K, July 2010.

Background

In a business world driven by information, insight is knowledge and knowledge, of course, is power. Business intelligence (BI) applications take data collected from multiple locations and parse it based on strategy and business objectives. Click stream analysis of online advertisements is a particularly interesting example of a business intelligence application: click stream analytics are used to interpret customer behavior in order to deliver more focused marketing and advertising efforts. Like most BI applications, click stream analysis for online advertising is driven by huge amounts of data. Unlike many BI applications, which tend to have a few users requesting long running reports against a large data set, click stream analysis often has many users running reports in parallel.

Providing an infrastructure to support business intelligence becomes more challenging as information volumes grow, users demand faster performance, and costs escalate. Business managers just want answers and don't want to deal with the challenges of data aggregation, how query workloads impact response times, or the difficulty of maintaining and protecting data without interrupting operations. Each challenge becomes evident when we take a look at the results of ESG spending surveys: as Figure 1 shows, while the push for cost reduction has eased a bit in 2010, it remains a top priority for organizations of all sizes. Business process improvements increased in importance, business growth remains in the top five, and "green" initiatives remain in the top ten. Interestingly, "improving business intelligence and delivery of real-time business information" was one of the initiatives that increased in importance between 2009 and 2010.²

Figure 2. Business Initiatives That Will Have The Greatest Impact on Spending Decisions, 2009 v. 2010



Source: Enterprise Strategy Group, 2010

The recent upgrades to the Microsoft adCenter BI infrastructure were driven by similar initiatives: cost reduction, business growth, "green" IT, and improving business intelligence. Business intelligence has become a strategic differentiator in a growing number of industries. BI drives adCenter's mission and revenue stream.

² Source: ESG Research Report, [2010 IT Spending Intentions Survey](#), January 2010.

Introducing Microsoft adCenter

Microsoft adCenter is the platform that enables customers to advertise through Microsoft's Bing search engine, Windows Live, and other properties. Advertisers select keywords and combinations of keywords that will appear during online searches and on content partner websites. Companies bid in an auction for their ad status when particular words are searched; they vie for placement with the same keywords and the auction process determines whose ad is displayed most prominently at any given time. Advertisers pay Microsoft when their ads are impressed on the screen and when viewers click on their ads. adCenter helps customers manage their search ad campaigns; its advanced capabilities enable advertisers to target the right consumers to maximize effectiveness. With Microsoft's recent alliance with Yahoo!, adCenter will become the platform for all Microsoft and Yahoo! search advertising campaigns, representing a huge growth opportunity.

While a significant portion of this infrastructure is dedicated to serving ads for customers, another large portion supports the business intelligence database for constant data monitoring, analysis, and tuning. It is this engine that drives adCenter revenue, tracking all ad placements, impressions, and clicks for analysis and billing.

A project of this size requires an extremely robust and reliable infrastructure as well as enterprise scalability. It is proof of the abilities of the Microsoft SQL Server database and EMC Symmetrix storage that they can handle this massive project, its enormous growth, and its stringent service-level requirements on a moment to moment basis. The ad serving portion delivers five nines availability using clusters in three geographies, each handling one-third of the ad traffic at all times and failing over to each other if needed.

adCenter Business Intelligence

The Business Intelligence portion of Microsoft adCenter performs click stream analysis and provides near real-time feedback to advertisers on how their ad words are being used. Advertisers bid on and pay for keywords and combinations of keywords, and the auction results determine where ads are placed on the screen and how often they appear. The BI system tracks ad impressions, click-throughs, and a multitude of statistics used to generate invoices to adCenter customers and to provide responses to customer database queries.

Data marts within the database are created with information specific to types of users; for example, the advertiser data mart is the largest because of the number of advertisers and enormous reporting demand, but there is also a publisher data mart for properties displaying ads. These include Microsoft search properties (e.g., Bing), Microsoft publishing properties (e.g., MSN, MSNBC, and Windows Live), and non-Microsoft publishers that display content with ads in the margins (such as sites dedicated to news, sports, social networking, etc.).

Some advertisers are individuals, such as mom-and-pop shops with small ad campaigns and only a few keywords; they use the simple GUI to query the database. Others are premium advertisers, such as large online auction sites managing thousands of keywords at a time. The majority of adCenter customers are advertising agencies that develop campaigns and manage online advertising as a whole for their customers. These companies use the GUI and sophisticated API calls to query the database on the effectiveness of their ad campaigns so they can tune and optimize throughout the day. From the database they can find out in near real-time which words are working and which are not paying off.

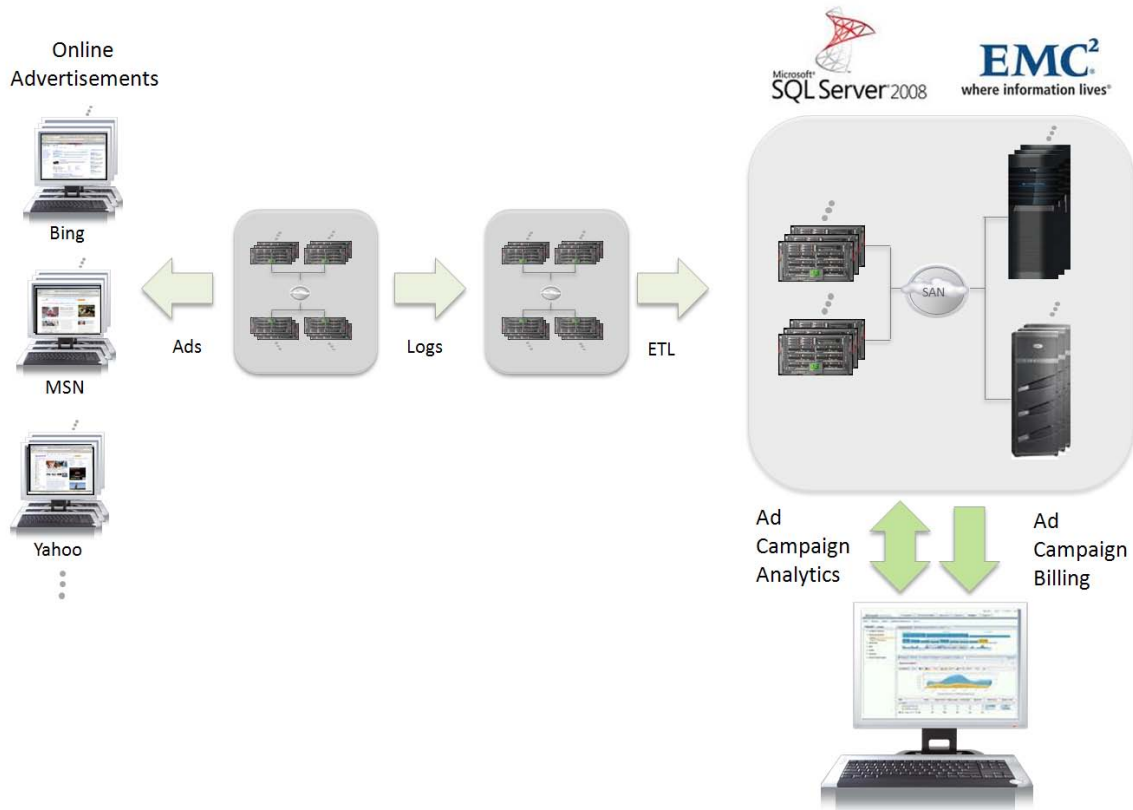
Web logs are created with every ad impression and click-through; these are aggregated hourly into advertiser-specific groups. For example, imagine that you are searching online for a car. Advertisers will bid on key words such as "car," "auto," "automobile," "sedan," etc. When you search these words, if American Car Company paid the most for the keyword "auto," then its ad will be at the top of the page in the right margin; Frank's Used Cars and Town Tire Company may have ads impressed farther down the page due to a lesser bid. If Frank's Used Cars (or its advertising agency) is tracking the campaign and doesn't like where the ad is placed, it may bid more, resulting in the ad moving farther up the page the next time "auto" is searched. Advertisers may have multiple impressions per hour of their ad when these keywords are searched. As bids are made, keywords are added, updated, and deleted, ads are impressed, and ads are clicked on by online searchers, the BI database records all the information from the web logs; data is aggregated for each advertiser hourly, daily, and monthly and is retained for three years.

Advertisers want to be able to target their ads to particular customers; the ability to fine-tune a campaign in real time based on analysis of its effectiveness is what nets advertisers a competitive advantage. The BI division is focused on capturing all the data about ads, impressions, and click-throughs and on getting those reports to customers as quickly as possible—giving adCenter itself a competitive advantage.

How it Works

As shown in Figure 3, advertisements on Bing, MSN, Yahoo!, and a variety of other websites are served by the adCenter infrastructure. Logs of online ad activity are shipped, extracted, transformed, and loaded (ETL) into the adCenter BI infrastructure on an hourly basis. Multiple SQL Server database partitions stored on EMC disk arrays (Symmetrix VMAX and the previous generation Symmetrix DMX) are used to store hourly, weekly, monthly, and yearly analytics. Advertisers use adCenter graphical user and programming interfaces from Microsoft to run ad campaign analytics reports. The adCenter infrastructure is also used to generate revenue as it generates billing information for online ad campaigns.

Figure 3. Microsoft SQL Server 2008 and EMC Symmetrix VMAX in the Microsoft adCenter BI Infrastructure



The Challenges

The challenges of scaling the adCenter infrastructure have created a significant new opportunity for Microsoft. Microsoft's CEO, Steve Ballmer, characterized this well when he said, "scale drives knowledge." As the volume of adCenter search and advertising traffic grows, the relevance of searches and the effectiveness of advertising grow accordingly. To realize the benefits of scale, the extreme size and growth of the Microsoft adCenter business intelligence infrastructure posed a number of significant challenges:

"Scale drives knowledge, which can turn around and drive innovation and relevance."

- Steve Ballmer, CEO, Microsoft

Scalability. The original business intelligence database built in early 2006 was 200 GB and has now grown to 250 TB. Not only does it organically grow at about 30% per year, but when adCenter takes on all of Yahoo!'s online advertising, additional six-fold growth is expected—bringing the database up to more than 1 PB. Managing such a large and growing business intelligence infrastructure is a key challenge for the Microsoft adCenter team.

Operating costs. As the database grows, more and more storage is needed. This leads to higher costs not only to manage the storage systems but also in terms of the additional data center floor space and energy needed for powering and cooling devices; an expensive data center expansion was on Microsoft's horizon. These costs can quickly get out of control in an infrastructure with sizable and sustained growth, stripping away profitability.

Availability. Continuous uptime is critical and the team is constantly trying to improve operations to ensure that ongoing maintenance, upgrades, and data protection activities don't interfere with productivity.

Service Level Agreements. adCenter business intelligence operates under strict SLAs which get tougher with each infrastructure upgrade. It is a challenge to provide faster responses to customers and still keep up with the massive growth of the application

- *The Retrieval Time SLA* is the most critical. When an advertiser queries the database for an update on the number of click-throughs and ad impressions, that response must be returned from the database within four seconds (before upgrading from SQL Server 2005 to SQL Server 2008 and Symmetrix DMX/VMAX, this SLA was six to eight seconds). This fast response time is one of the primary reasons adCenter selected EMC storage. adCenter is committed to meeting this SLA nearly 100% of the time as it is a direct customer interaction and impacts their satisfaction. Users expect information instantly and, as a result, one of adCenter's long-term goals is sub-second response to a database query.
- *The Data Latency SLA* requires that data be no more than three hours old. That is, once an activity takes place—an ad impression is served, a customer clicks through, a bid is revised, a key word is deleted—information about those activities must be included in reports within three hours. This keeps advertisers up to speed on their campaign effectiveness in real time so that they can tune effectively. With click stream analysis, the data is most valuable to all parties as soon as it is rendered. Because advertisers want to tweak their campaigns by depreciating some keywords, upgrading others, and increasing their bid rates throughout the day, real-time information has a high degree of value. adCenter built an infrastructure focused on speed as well as growth so that customers get the fastest performance while the data is most valuable. The most current data is kept on tier-1 and tier-0 storage, and a creative partitioning strategy maximizes performance and availability. As the data ages, it loses value, so longer running queries can be run on lower-cost storage without such high-performance.

"Our number one problem is scale as we try to meet retrieval time and data latency SLAs. With a big step in growth due to the addition of Yahoo! traffic and organic growth on top of that, we've got a really, really large problem to solve. We grew beyond the capabilities of a single SQL Server partition a long time ago."

- Sasha Berger, BI Architect, Microsoft adCenter

BI at OLTP speeds. The performance requirements for this application are unusual. Classic data mining and business intelligence applications are focused on processing historical data so that organizations can make intelligent decisions. Most commonly, these databases have few users and large, complex queries to run; mass quantities of real-time responses are not the norm. Conversely, traditional OLTP applications are built for many users with small, real-time, responsive-time sensitive requests. The challenge for adCenter is to offer business intelligence services at OLTP-like speeds. The system does handle long-running queries (such as providing statistics for particular keywords over the last six months), but 80% of the requests are for real time answers to questions such as, “How did my ad campaign perform in the last 4 hours?” and “Is this keyword having the right impact now?”

Storage automation. Automating routine storage administration tasks is necessary to reduce the cost and complexity of keeping up with storage capacity growth while meeting performance and availability SLAs. Complex storage administration distracts the team from working on developing new system features that will enhance the experience of current customers and entice new ones. The less time developers have to deal with storage tasks, the more time they can spend developing features focused on customer satisfaction that can increase adCenter market share.

Data integrity. Data integrity is a critical concern for the adCenter BI infrastructure team. Corrupted or missing data can result in under- or over-billing a customer. For example, imagine the problems that can occur if the data mart indicated that an advertiser got 25,000 impressions when they actually only had 2,500. This type of error causes operational interruption, reduces credibility, and can drive customers to other advertising vendors.

Green. Reducing the power and cooling requirements of the adCenter BI infrastructure helps Microsoft meet its corporate level “green” initiatives. In a message to Microsoft employees in 2009, Steve Ballmer said, “Microsoft will cut the rate of our carbon emissions and continue to invest in efforts to significantly reduce our use of natural resources. Our goal is to reduce our carbon emissions per unit of revenue by at least 30% compared with 2007 levels by 2012. We’ll continue to invest in programs and search for opportunities to reduce the environmental impact of our operations.”

“Our goal is to reduce our carbon emissions per unit of revenue by at least 30%.”

- Steve Ballmer, CEO, Microsoft

The Solution

To serve its approximately 300,000 advertisers, the business intelligence infrastructure group employs about 150 people. The current advertiser database (280 TB) grows by 19 million rows per day and is expected to grow by 100 TB/year. The database handles about 400,000 report requests daily.

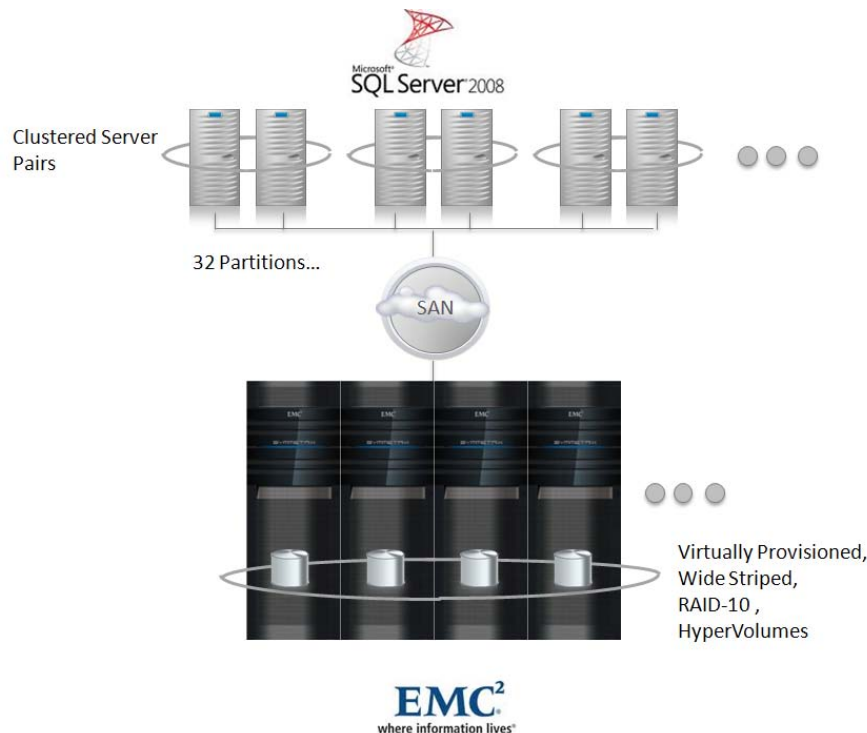
The BI infrastructure operates in two sites, in California and in Chicago, to ensure business continuity. Each site utilizes 64 blade servers in 32 clustered pairs running SQL Server 2008, all attached to enterprise-class EMC Symmetrix storage.

“Buddy system” ensures availability and performance. The total server infrastructure is currently built using 128 server nodes with 32 pairs in each location. As shown in Figure 4, Microsoft Windows Failover Clustering software is used to configure the servers as active-active pairs. To ensure predictable performance and business continuity, the pairs operate in a kind of “buddy system.” During the database’s extract-transfer-load (ETL) action, the process of loading data into the database is the most processor- and storage-intensive. adCenter designed the system so that in each clustered server pair, Server A acts as the primary active node and takes log-shipped data for an hour, while user requests are directed to Server B (which is also available for fault tolerance). Every hour, they switch roles for a short time—B writing shipped logs while A services requests—so that advertisers are never without current data and the load process never interferes with requests. As soon as the load is complete, the roles are once again reversed. By isolating the database load from servicing requests, performance SLAs are ensured and loads operate more quickly, helping to meet the data currency SLA.

Microsoft adCenter Infrastructure

- 200 databases
- 700 servers
- Multiple geographies
- 280 TB BI database
- More than two billion rows
- More than 250,000 reports per day

Figure 4. Microsoft SQL Server 2008 and EMC Symmetrix VMAX in the Microsoft adCenter BI Infrastructure



Partitioning strategy delivers predictable performance at scale. A critical architectural design in the adCenter business intelligence infrastructure combines the partitioning capabilities of Microsoft SQL Server 2008 with Symmetrix hyper-volumes. In the most recent upgrade and reconfiguration, adCenter created 32 partitions (having already increased from eight to 16) to store adCenter BI analytics by day, week, month, and year. Separating the

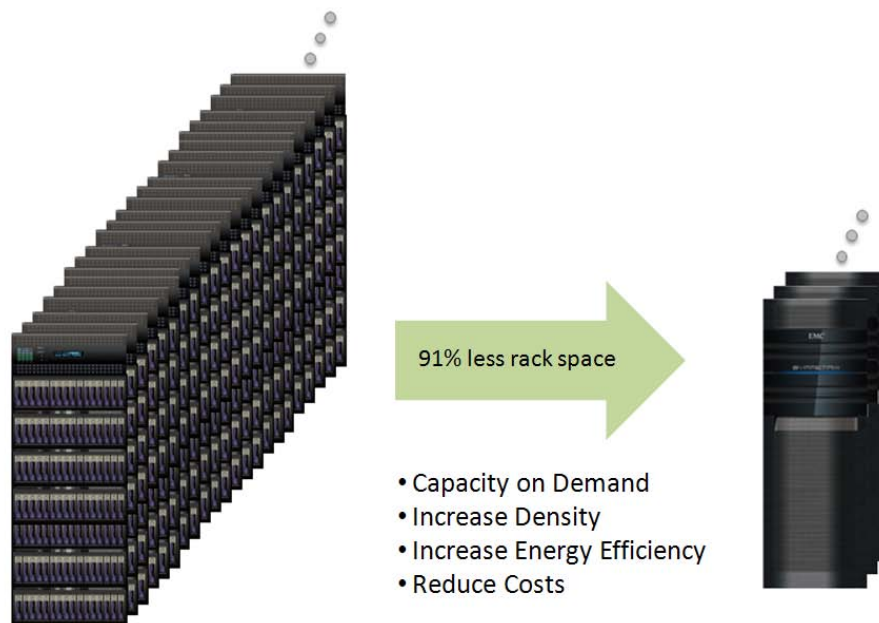
data in this way works to isolate and optimize queries and reports. In this manner, quick reports running against the most recent ad traffic data are separated from long running reports looking for monthly or yearly trends.

Each partition is directed to multiple EMC Symmetrix LUNs built with RAID 10 hyper-volumes that are wide-striped across as many disk drives as possible. This architectural approach provides predictable, scalable performance for pools of disk drives and enables adCenter to do a better job of scaling server processing power. Dividing the workloads helps the adCenter team more predictably meet SLAs.

The adCenter team uses a modular and repeatable storage configuration and layout when adding a new EMC Symmetrix VMAX to the infrastructure. The latest standard uses 400 GB drives split into 55 GB chunks. These are protected with RAID 10 mirrors that are placed in different storage bays and racks to protect against an outage. These 55 GB mirrored chunks are then striped across 70 to 80 drives, which are used to create 2 TB LUNs. Since each drive can conceivably perform 100 IOs per second for read operations, each LUN has the capability to handle 8,000 random reads. While some of the workloads are random and some are sequential, this extreme level of performance that can be achieved with each LUN is used to ensure that servers will run out of processing power before the storage array becomes a bottleneck. By increasing the number of partitions and having those partitions each going to separate LUNs, performance traffic between the most recent data and the longer-running queries is isolated. By striping wide and deep and distributing the workload so that each 2 TB LUN covers multiple drives spread across multiple racks and bays, adCenter has built a highly available system whose theoretical limits are higher than the application itself can reach. The end result is a modular building block approach that simplifies storage deployment which frees the adCenter team to focus on more valuable development efforts.

Storage consolidation reduces space, power, and costs. ESG audited the results of a recent adCenter storage consolidation project. The project migrated adCenter BI data from legacy modular storage arrays to EMC Symmetrix VMAX arrays. As shown in Figure 5, the higher capacity of the newer VMAX arrays combined with the on-demand capacity provided by virtual provisioning, an optimized drive layout scheme, and automated management was used to consolidate from 129 racks of modular midrange storage to only 12 racks of Enterprise-class EMC VMAX storage.

Figure 5. Consolidating the Microsoft adCenter BI Infrastructure with EMC Symmetrix VMAX



As shown in Table 1, one of the biggest benefits of the project was an increase in storage utilization from 34% to 70%. This led to a number of benefits which illustrate the bottom line value of a well-orchestrated storage consolidation initiative. Power requirements were reduced by 92% from 676 kVA to 57 kVA. Microsoft compared

these dramatic power savings to the power used by all of the homes in a suburb near the Microsoft campus called Snoqualmie, Washington. An ESG Lab audit of the costs associated with Project Snoqualmie has confirmed that the consolidation initiative will reduce the cost of the adCenter BI infrastructure by 69% over the next five years.

Table 1. Symmetrix VMAX Consolidation Benefits

	Racks	Utilization	Power (kVA)
Before VMAX consolidation	129	34%	676
After VMAX consolidation	12	70%+	57
Savings	91%	36%+	92%

The Benefits

SQL Server 2008 and EMC Symmetrix DMX and VMAX storage are at the heart of the BI infrastructure. The fact that these products can handle this system given its size and growth is testament alone to their robustness. Upgrading infrastructure components, moving to SQL Server 2008 and Symmetrix VMAX storage, and restructuring the partitioning strategy together have resulted in a 10%-20% performance boost. In addition, the new infrastructure has resulted in fewer problems—the number of issues posted has fallen by 10% - 15%. Also, despite the tremendous growth and division of database workloads into 32 partitions, the BI operations team has not needed to hire additional staff—the same team manages the infrastructure even as it grows.

Scalability. The infrastructure design and Symmetrix features make it possible for the BI database to continue its enormous growth, while still achieving performance and availability requirements. High-density disk drives, storage tiers, and automated management are critical enablers of this growth—without them, the team would spend inordinate amounts of time just trying to keep up. In addition, SQL Server 2008 is highly scalable. While database partitions consist of 15-20 TB of data, CPU usage is currently only about 50%. Peaks, when a flood of load requests comes in, max out at about 95%. This type of scalability shows how able SQL Server 2008 is to handle the huge infrastructure, leaving plenty of headroom for continued expansion.

Availability. A key reason that adCenter deployed Symmetrix arrays is high availability. Continuous uptime is paramount, so the fault tolerant Symmetrix design is a plus. More important, Symmetrix Virtual Provisioning enables expansion without downtime and will enable the addition of Yahoo! advertising as well as annual organic growth without interrupting operations. In addition, Virtual LUN Migration ensures that data can be moved as needed with minimal effort and no downtime.

Performance. The performance capabilities of Symmetrix were another influencing factor. The only way to meet all of adCenter BI's SLAs is to get enough IO traversing those 128 server nodes. Previously, the BI team had a hard time keeping up with user demand because the system was starved for IO. In addition, SQL Server 2008 capabilities have improved performance considerably as its data compression capabilities enable excellent read performance. With reads numbering 100/second, performance is critical. While writes are much less frequent, write latency impacts data availability—when writes take too long, other processes are interrupted. As a result, SQL Server 2008 compression has dramatically reduced the time it takes for hourly BI database updates.

Without the combination of SQL Server 2008 and EMC VMAX, we wouldn't have been able to meet our performance SLAs.

- Guru Kirthigavasan, Project Manager, Microsoft adCenter

Management. With Symmetrix DMX/VMAX, the adCenter BI teams find it easier to deploy new systems and automation makes management simpler. For example, built-in VMAX automation has reduced the time to do LUN masking and mapping by 70%, eliminating the need for multiple spreadsheets and scripts. LUN migration using the wizard saves time and effort. This makes it easier for them to deliver excellent engineering services in an environment that is growing astronomically—and frees up developers to focus on more value-added activities.

Cost. Virtual Provisioning dramatically reduces the cost of storage capacity, which can now be delivered on demand. Allocating storage to hosts only once reduces planned downtime and management costs. Adding capacity to an existing virtual pool of storage can be performed online without interruption.

Recently, the adCenter team performed an evaluation to see if they could achieve their operational objectives and save money with many smaller arrays or direct-attached storage. The evaluation revealed that the best way to reduce overall costs while ensuring SLAs are met is to consolidate on large, enterprise-class disk arrays from EMC.

Utilization. By upgrading to the latest generation Symmetrix VMAX, adCenter was able to consolidate more storage. Virtualization enabled the team to increase utilization from 34% to more than 70% while still meeting their stringent requirements and saving money in the data center.

Continuous improvement. adCenter also has the advantage of quarterly professional services visits by EMC “speed gurus” to evaluate the performance of the business intelligence infrastructure. This way, configurations can be continuously improved based on results from the previous quarter. EMC professional services staff extract database performance metrics and wash them through their

knowledge base and tools to show regions of the array that are under- or over-utilized. This effort has been used to drive a number of significant performance optimization recommendations. A joint effort of EMC professional services, the adCenter BI development team, and the adCenter operations team is used to perform a cost benefit analysis on each recommendation. Some of those efforts have resulted in storage reconfiguration efforts. Others have resulted in database application changes. And others have used a bit of both. The partitioning and LUN placement strategy described earlier in this report is an excellent example of a combined optimization effort.

Power savings. The EMC refresh and consolidation initiative was part of what Microsoft called “Project Snoqualmie,” in reference to the fact that the project's energy savings are enough to power all the homes in Snoqualmie, WA. By using massive consolidation in Symmetrix VMAX storage and leveraging the platform’s Virtual Provisioning, automated management, and high density drive option, adCenter implemented a capacity-on-demand model that addressed low utilization on legacy arrays and delivered substantial cost and energy savings.

This consolidation effort reduced the number of data center racks from 129 to 12, which extended the useful life of the existing data center and avoided the cost of further build out. In addition, it reduced the number of managed systems by 96%. Power consumption dropped from 320x30-amp plugs down to 14x50-amp plugs, saving 619kVA in hardware device consumption. Unused storage capacity and power, previously orphaned behind hosts, are today available in a shared pool.

The combination of power savings, 29:1 storage array consolidation, and extending the useful life of the data center has reduced the total cost of ownership for adCenter BI storage infrastructure by 69% over five years.

Storage consolidation onto EMC VMAX arrays with virtual provisioning increased storage utilization from 34% to 70%. The consolidated system meets performance requirements while saving a lot of money on power and space in the data center.

- Mike Anderson, Consultant, Systems Integration Engineer – Microsoft Databases and Applications, EMC Symmetrix Partner Engineering

The Bigger Truth

Storage vendors are quite adept at showing off their latest extended software capabilities. These are often very valuable, but sometimes we get caught up in exotic new solutions and forget about what a difference basic storage consolidation can make.

The Microsoft adCenter business intelligence infrastructure is a great example of the immensely powerful value proposition of good old-fashioned enterprise storage. Consolidated, enterprise-class, field-proven EMC Symmetrix storage can handle the enormity of this infrastructure and still support the addition of Yahoo!'s advertising and 30% annual organic growth. This massive consolidation is enabled by the combination of high-density, low-power drives and firmware advancements such as virtual provisioning and LUN migration. The system provides adCenter with extreme reliability, all the performance it needs, and management automation that doesn't interrupt operations.

In addition, EMC Symmetrix storage offers these benefits at an attractive total cost of ownership. This lower TCO is not what most would expect—Microsoft adCenter actually went through the exercise of calculating the cost to implement this infrastructure with JBOD and realized that the most cost-effective way to achieve its operational and budget objectives was by consolidating on Symmetrix systems. This is what Symmetrix has always been known for, so it's no surprise to EMC. But this value proposition can sometimes get lost in the quest for the latest and greatest widget. EMC has been a leader in this game for a long time and the adCenter deployment is an example of why that is so. That's not to say that EMC doesn't have advanced capabilities and state-of-the-art software built into the Symmetrix platform, but in this case, it's the field-proven scalability and core storage capabilities of the Symmetrix that's doing the job.

Combined with SQL Server 2008 and its advanced features—including data compression, scalability, and performance—the adCenter BI system is both advanced and enormous. The adCenter team has built a tier 1 data center architecture that delivers predictable scalability, reliability, and performance; automated management; higher resource utilization; and savings in data center space and power. Wrapping SQL Server 2008 and EMC Symmetrix technology in tier 1 process and procedures, the Microsoft adCenter team has created a BI infrastructure with tier 1 uptime and performance. All this, plus the ability to take on Yahoo!'s advertising and continue to support 30% annual growth as the total cost of the storage infrastructure was reduced by 69% over five years.

It's quite a feat. This is clearly an important segment of Microsoft's business—in FY2010, Microsoft's online advertising revenue increased 8% to 1.9 billion.³ It's certainly not something it would be taking chances with.

The infrastructure upgrade and simple, scalable architecture built with SQL Server 2008 and EMC Symmetrix DMX/VMAX systems are expected to support the current and expanded adCenter business intelligence effort expansion for the next 18 months. The adCenter team is currently designing the infrastructure that comes after that. Chances are, that will be something to see as well.

³ Source: Microsoft 10-K, 30 July 2010.



Enterprise Strategy Group | **Getting to the bigger truth.**