Server Virtualization Assessment and Business Analysis

Contact: STA   888-477-8246   www.staweb.com

Table of Contents

Introduction

ABC Inc. engaged Systems Technology Associates (STA) to perform a Virtualization Assessment utilizing VMware Capacity Planner software. During this study, STA captured and analyzed performance data on the servers in ABC Inc&s. infrastructure for a minimum of 30 days in order to identify and model server consolidation scenarios utilizing a virtual infrastructure based on VMware vSphere 4.

Scope

The Consolidation Study began in June 2010 when STA installed the VMware Capacity Planner software. ABC Inc. provided a list of servers in their IT infrastructure that was imported into the Capacity Planner database. Afterwards, the servers were validated and the list reduced by eliminating those servers giving invalid data or incomplete data. The result was a list of 18 servers to be included in this assessment.

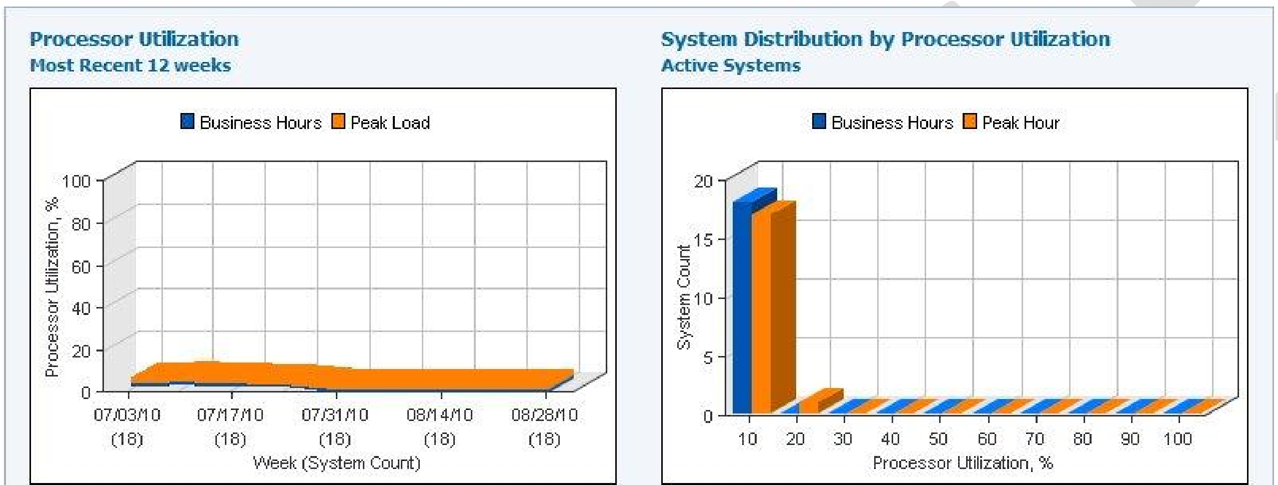
Working against the list of servers, STA used Capacity Planner to perform an inventory of their hardware and software configuration, and to collect performance data for analysis and modeling.

Purpose

After approximately twelve weeks of performance data collection, information from the VMware Capacity Planner Data Warehouse was analyzed to determine the extent to which the sampled servers could be consolidated. The purpose of this document is to report the result of that assessment.

Executive Summary

As a group, the servers in the scope of this assessment showed very low resource utilization rates as indicated in the graphic below. Most IT infrastructures exhibit the same unused capacity, which generally indicates a good fit for server consolidation. Our findings validate this; server consolidation is possible at ABC Inc., yielding its benefits and cost savings.



System Count

CPU(s)	Systems
1	1
2	4
4	10
8	3
Total Systems	18

Processor Summary

CPU MHz	173,283
Avg CPU Utilization	0.42
Peak CPU Utilization	1.18
CPUs	73
Avg MHz/CPU	2,374
Avg MHz/System	9,627

Optimization Opportunities

Consolidation:	18
Anomalies:	6
Alerts:	0
Trend Deviations:	0
Systems with CPU <200 MHz:	0

Conclusions

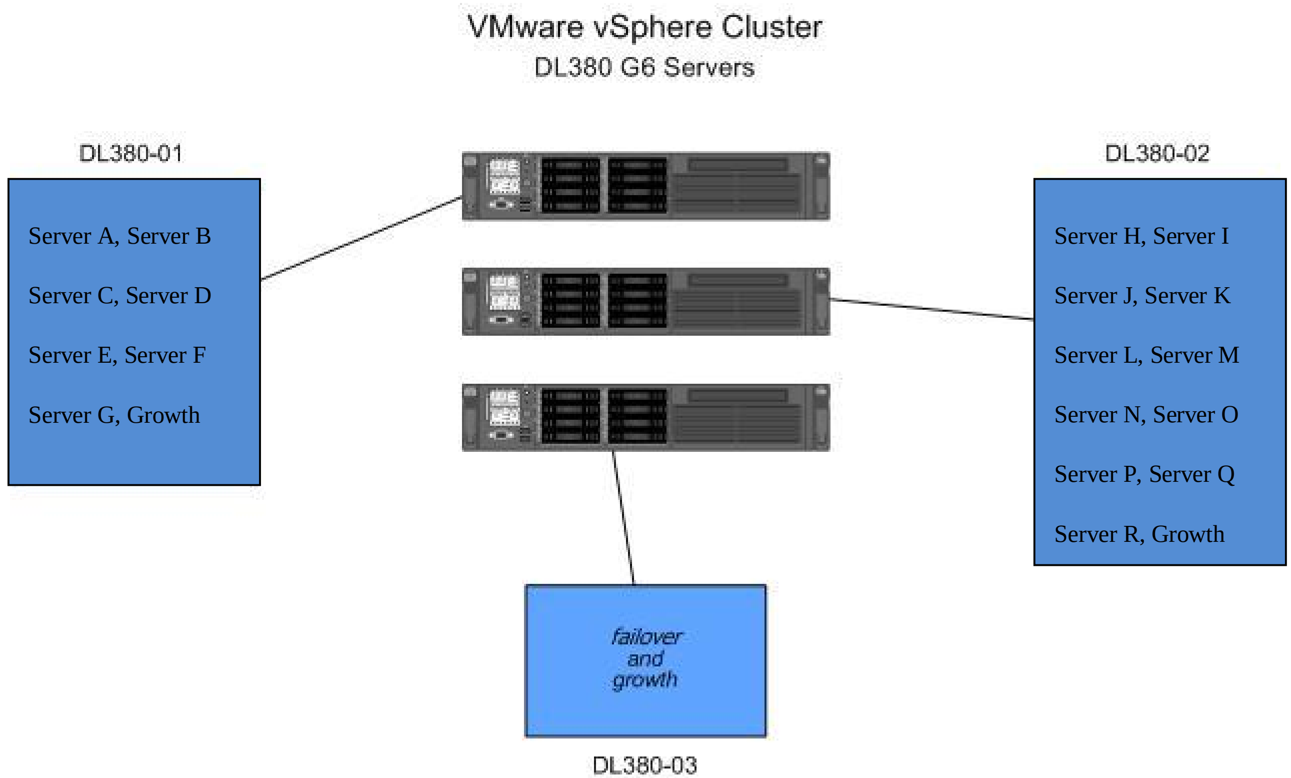
Consolidation makes sense for most servers in an IT infrastructure. Some servers exhibit behavior that suggests they would not be good consolidation candidates, more often than not this behavior is performance related. These servers are called Systems with Exceptions. In our analysis, and subsequent findings, we include Servers with Exceptions in the after virtualization numbers. So when we describe a post-virtualization configuration in this report it includes Virtual Host Servers and Systems with Exceptions.

Consolidation makes sense for the 18 servers that were included in this assessment. ABC Inc. may realize the following benefits from consolidation as described in this document:

- Multiple workloads are run on fewer, larger servers
- Workloads draw on a large pool of memory, processor, disk and network resources
- Unused computing resources are kept to a minimum
- Reduced power and cooling requirements
- Reduced system maintenance

Consolidation at ABC Inc. could happen in several scenarios, depending on a number of factors. In this study, STA offers one consolidation scenario that uses actual workload measurements from ABC Inc. servers to illustrate the extent to which consolidation is possible utilizing VMware vSphere 4. The scenario uses a sampling of 18 servers from the environment.

In this consolidation scenario, workloads of ABC Inc. servers, as measured by the VMware Capacity Planner, are grouped and mapped on the target platform of HP DL380 G6 servers. Mapping occurs within the constraints of the hardware and the virtual host server thresholds. The scenario assumes no boundaries, political or geographic, in mapping the workloads onto the same physical server; the primary consideration is resource utilization. The following graphic represents the effective consolidation, showing server names in workload groups; each server name represents a virtual machine.



This graphic shows the extent to which consolidation is possible. The location of workloads in this context is for sizing purposes only; it is not meant to indicate where an ABC Inc. server will run after it is virtualized. In an actual production vSphere virtual infrastructure the three, host DL380 computers would form a single resource pool in which all virtual servers (workloads) would run. The location of where any given workload runs would be fluid, constantly evaluated and set by the management server to provide the best resource utilization. Workloads that need to run together on the same host, as well as workloads that must never run on the same host, would also be handled by the management server using affinity rules.

In the table below we summarize the impact of the consolidation scenario by showing the change in number of processors, memory and total servers targeted for virtualization (consolidation).

		Before Virtualization	After Virtualization	Saving
Consolidation Scenario: DL380 Servers	No. of Processors	73	24	67%
	Total Memory (GB)	65	96	none
	No. of Servers	18	3	83%
	Power Demand (W)	11,000	2,760	75%
	Thermal (BTU/hr)	36,600	10,290	72%
	Spare Capacity	0	35%	

Recommendations

The current ABC Inc. Windows server infrastructure can be consolidated, achieving significant savings; this is our conclusion. As such, the recommendations made in this document are presented as a means to establishing a virtualized server environment and realizing the benefits it offers:

- **Adopt a Scale-Up Architecture**. This is a pivotal point in changing IT policies by abandoning the one application per server, scale-out architecture in favor of a consolidated, shared resource, scale-up architecture.
- **Enhance SAN Infrastructure**. Enhance the scope the SAN infrastructure to include non-mission critical servers that would be virtualized in a VMware Virtual Infrastructure.
- **Establish a Virtual Infrastructure**. Design and build the infrastructure first. Account for special processing requirements and high availability requirements.
- **Migrate Applications**. Migrate applications into the virtual infrastructure hosted by VMware ESX Server. Benchmark performance after each workload is added to track resource utilization. Enable high availability and distributed resource scheduling.

Next Steps

STA suggests the following as the next steps in implementing a virtual server infrastructure:

- Validate this Consolidation Estimate with your Technology Team
- Adjust the consolidation analysis to include all servers
- Leverage these results to build a business case
- Engage STA to ensure a successful implementation of VMware Virtual Infrastructure.

Current Environment

This section provides some information about the servers being assessed. The information is presented in the following categories:

- System Discovery
- Server Inventory
- Performance Summary
- Performance Anomalies

The information presented in this section was collected directly from the servers using the VMware Capacity Planner tool version 2.7. The information was sourced from the Windows Registry and from Windows Perfmon Counters.

System Discovery

After installing the VMware Capacity Planner, the first step of the Virtualization Assessment is to identify systems within the ABC Inc. network. This is the Discovery Process.

In the Discovery Process, we imported a list of servers imported it into the Capacity Planner. The list was created by ABC Inc. and consists of server names and locations in spreadsheet format. The list contained the 18 servers shown below.

Server A	Server E	Server I	Server M	Server Q
Server B	Server F	Server J	Server N	Server R
Server C	Server G	Server K	Server O	
Server D	Server H	Server L	Server P	

After the server list was imported, the Capacity Planner ran a series of qualifying queries to validate the list. Queries not only test that the server can be reached but also verify that the Capacity Planner can extract inventory and performance data. Servers that cannot be validated are removed from the scope of the assessment. The result of the queries is shown below:

Servers names imported	18
Servers with invalid data	-0
Servers with incomplete data	-0
Servers to be assessed	18

Server Inventory

After identifying the servers that were to be included in the assessment, the next step was to gather more information on the hardware and software configurations of the remaining servers. This step is the Inventory Process.

Although there are hundreds of pieces of information on any given machine, the Capacity Planner focuses on only those elements pertinent to capacity planning. Accurate information such as CPU, NIC, RAM, Disk Drive, and Software, is collected from all authenticated servers. These pieces of information allow us to perform comparative analysis on servers across the ABC Inc. enterprise.

The table below provides a summary of the inventory information collected for each server.

Server Name	OS	CPUs	CPU Speed	Total Mhz	RAM (MB)	Disk Size (GB)	NIC (Mbs)	Power (W)	Cooling (BTU/hr)	Services
Server A	Windows Server 2003, Standard	8	2,261	18,088	4,096	159	2,000	500	1,700.00	101
Server B	Windows Server 2003, Standard	2	2,992	5,984	1,024	2,584	2,000	713.9	2,435.80	102
Server C	Windows Server 2003, Standard x64	4	1,995	7,980	8,192	3,078	2,000	722.2	2,464.10	105
Server D	Windows Server 2003, Standard	2	2,793	5,586	1,024	73	2,000	507.9	1,733.00	121
Server E	Windows Server 2003, Standard	1	2,793	2,793	2,048	73	1,000	442.4	1,509.50	122
Server F	Windows Server 2003, Standard	4	2,493	9,972	4,096	365	2,000	495.1	1,689.30	118
Server G	Windows Server 2003, Standard	8	2,394	19,152	4,096	374	4,000	500	1,700.00	113
Server H	Windows Server 2003, Standard	4	1,995	7,980	4,096	249	2,000	495.1	1,689.30	106
Server I	Windows Server 2003, Standard	2	2,992	5,984	1,024	73	2,000	507.9	1,733.00	104
Server J	Windows Server 2003, Standard	4	2,826	11,304	4,096	146	2,000	495.1	1,689.30	105
Server K	Windows Server 2003, Standard	4	2,660	10,640	4,096	474	2,000	722.2	2,464.10	108
Server L	Windows Server 2003, Standard	2	3,192	6,384	4,096	367	2,000	507.9	1,733.00	101
Server M	Windows Server 2003, Standard	4	2,660	10,640	4,096	1,079	2,000	495.1	1,689.30	117
Server N	Windows Server 2003, Standard	4	1,995	7,980	4,096	3,827	2,000	722.2	2,464.10	119
Server O	Windows Server 2003, Standard	8	1,861	14,888	4,096	73	2,000	495.1	1,689.30	115
Server P	Windows Server 2003, Standard	4	2,660	10,640	4,096	1,263	2,000	722.2	2,464.10	118
Server Q	Windows Server 2003, Standard	4	1,995	7,980	4,096	249	2,000	495.1	1,689.30	139
Server R	Windows Server 2003, Standard	4	2,327	9,308	4,096	511	0	722.2	2,464.10	118

Performance Summary

This subsection of the document provides a look at how much resource the target servers are consuming. The performance counters described below are used in the server performance summaries shown in this section; they represent the core performance counters collected by the Data Collector.

Server Name	The name of the server for which performance statistics have been collected.
CPU Percent Busy	The percentage of time that the processor is executing a non-Idle thread. This counter displays the average percentage of busy time observed during the sample interval. It is calculated by monitoring the time the service was inactive, and then subtracting that value from 100 percent.
Available Memory (MB)	The amount of physical memory available to processes running on the computer. It is calculated by summing space on the Zeroed, Free, and Standby memory lists. Free Memory is ready for use. Zeroed Memory is pages of memory filled with zeros to prevent later processes from seeing data used by a previous process. Standby Memory is memory removed from a process working set (its physical memory) on route to disk, but is still available to be recalled. This counter displays the last observed value only; it is not an average.
Pages/Sec	The number of pages read from or written to disk to resolve hard page faults. (Hard page faults occur when a process requires code or data that is not in its working set or elsewhere in physical memory, and must be retrieved from disk.) The counter was designed as a primary indicator of the kinds of faults that cause system-wide delays. It is the sum of memory as pages input/second and pages output/second without conversion. It includes pages retrieved to satisfy faults in the file system cache (usually requested by applications) non-cached mapped memory files. This counter displays the difference in the last two samples, divided by the duration of the sample interval.
Page File Percent Avg	The percentage of the Page File being used.

Physical Disk Percent Busy	The percentage of elapsed time that the selected disk drive is busy servicing read or write requests.
Physical Disk Avg Queue	The average number of both read and write requests that were queued for the selected disk during the sample interval.
Bytes Tot/Sec	The number of bytes the server has sent to and received from the network per second. The value provides an overall indication of how busy the server is.
CPU Queue	The number of threads in the processor queue. There is a single queue for processor time even on computers with multiple processors. Unlike the disk counters, this counter counts ready threads only, not threads that are running. A sustained processor queue of greater than two threads generally indicates congestion. The counter displays the last observed value only; it is not an average.
Cache Bytes	The number of bytes being cached.

The table below shows the actual values collected during the sampling period. They are presented with the industry averages as well as highlights alerts and anomalies for each performance counter. The first row of the table shows the industry average, which was obtained from the VMware Capacity Planner Information Warehouse and represents the averages of the performance metrics collected.

The values were used in our analysis and are the basis for the recommendations presented in this document. Some of the field values in the performance table are highlighted to draw your attention:

- Yellow highlighted fields indicate anomalies that occur when the performance metric is more than three standard deviations away from the industry average.
- Red highlighted fields indicate alerts that occur when the server exceeds the software vendor threshold.

Server Name	CPU Percent Busy	CPU Queue	Available Memory (MB)	Cache MBytes	Pages/Sec	PageFile Percent Avg	Bytes Tot/Sec	Physical Disk Percent Busy	Physical Disk Avg Queue
Industry Average	**4.91**	**0.68**	**2,697**	**339**	**62.48**	**7.92**	**76,675**	**5.35**	**0.68**
Server A	0.13	0.00	3,585	121	2.23	0.43	11,346	2.67	0.03
Server B	0.16	0.00	252	422	7.15	5.12	16,110	0.32	0.01
Server C	0.11	0.00	6,547	941	3.80	2.11	50,055	0.10	0.00
Server D	0.87	0.00	135	73	8.46	16.62	5,536	0.30	0.00
Server E	2.20	0.00	351	119	39.71	71.58	7,846	31.95	0.32
Server F	0.14	0.00	1,987	257	5.25	1.41	16,729	1.05	0.02
Server G	0.22	0.00	1,918	110	1.10	1.08	6,281	0.13	0.00
Server H	0.02	0.00	3,421	200	0.10	1.89	1,208	0.01	0.00
Server I	0.18	0.00	391	160	0.06	0.53	2,756	0.18	0.00
Server J	0.05	0.00	3,395	200	0.19	2.02	5,531	0.07	0.00
Server K	0.42	0.05	2,793	235	2.78	0.24	41,327	0.57	0.01
Server L	0.05	0.00	3,521	93	0.01	1.21	1,906	0.06	0.00
Server M	0.28	0.00	1,656	168	0.31	1.01	3,870	12.68	0.25
Server N	1.96	0.00	3,141	134	2.14	1.37	11,718	3.14	0.06
Server O	0.43	0.00	1,976	210	0.17	0.39	52,318	0.16	0.00
Server P	0.22	0.00	1,682	215	0.03	1.32	21,711	2.70	0.11
Server Q	2.35	0.05	796	201	43.98	0.74	28,228	19.31	0.19
Server R	0.31	0.00	2,403	177	1.93	3.53	14,382	8.19	0.16

Performance Anomalies

An anomaly indicates performance that is significantly different from the industry performance averages of like-servers provided by the VMware® Capacity Planner Information Warehouse. An anomaly is determined when Capacity Planner detects performance that is three standard deviations from the industry average.

A performance difference indicated by an anomaly can be either good or bad. A good anomaly can provide a comparison to real-world, real-time performance that can be leveraged to configure similar machines, whereas a bad (or undesirable) anomaly can identify a problem and allow it to be addressed proactively before serious performance issues surface.

This subsection of the document presents the anomalies identified by Capacity Planner. The definitions shown below describe the columns used in the Table of Anomalies that follow.

Class (Instance)\Metric	The performance metric that deviates from the industry average.
Industry Average	The average for this metric based on data collected from Capacity Planner client servers and stored in the Information Warehouse.
Threshold	The peak hour average.
Standard Deviations	The numbers of standard deviations away from the industry average for the performance counter in the Object column.
Prime Time Average	The average utilization during prime time hours.
Low Load	The lowest sustained utilization; not the minimum observed values.
Peak Load	The highest sustained utilization; not the maximum observed values.

Performance anomalies for each of the target servers of this assessment are shown in the subsections below:

Server	Class (Instance) \ Metric	Industry Average	Threshold	Standard Deviations	Prime Time Average	Low Load	Peak Load
Server X	Paging File (_Total) \ % Usage	7.92	100	4.49	71.58	67.58	79.98
Server X	Paging File (_Total) \ % Usage Peak	12.57	100	3.87	90.72	90.72	90.72
Server Y	PhysicalDisk (_Total) \ Disk Read Bytes/sec	267437.60	300000000	4.86	11405601.00	6.82	40270639.00
Server Y	PhysicalDisk (1 D:) \ Disk Read Bytes/sec	267437.60	300000000	4.86	11400346.00	4568.87	40262103.00
Server Y	PhysicalDisk (_Total) \ Disk Bytes/sec	382781.70	300000000	4.13	11421069.00	4915.66	40304115.00
Server Y	PhysicalDisk (1 D:) \ Disk Bytes/sec	382781.70	300000000	4.12	11400368.00	0.71	40262159.00

Findings

The following findings were developed while assessing the performance counters extracted from the inventoried servers. The findings are oriented at establishing a virtualized server platform and migrating physical servers into it. As such, this section is presented in the following subsections:

- Gap Analysis
- Consolidation Scenario
- Storage Considerations

Gap Analysis

The section describes the results of the Gap Analysis between the current state (existing technology) versus the future state (future objective) vision. The results are provided in a matrix-format.

Component	Current State	Future State	Gap	Requirement	Action
Application Servers	Commodity servers with varying types and processor count.	Consolidated server environment. Standardized technology. Scalable platform.	Platform/Architecture. Currently one server per application. Varying technology standards.	Reduce application servers to minimum possible. Standardize technology platform.	Build virtual server infrastructure. Address anomalies and migrate physical servers.
Disk Storage	Stable SAN utilized only by mission critical servers. Widespread use of DAS in non-mission critical servers.	Consolidated storage environment utilizing SAN, all servers connected.	Platform/Architecture. Islands of DAS, inefficient use of storage. Varying technology standards.	Improve efficiency and manageability of storage. Standardize storage technology.	Enhance SAN scope to include virtualized servers (critical and non-critical). Eliminate islands of storage.

Consolidation Scenario

Consolidation makes sense for most Windows servers in an IT infrastructure. Some servers exhibit behavior that suggests they would not be good consolidation candidates, more often than not this behavior is performance related. This analysis carefully considered performance behavior in determining which servers could be consolidated. In this consolidation scenario, servers whose performance measurements deviated too far from the industry average have been excluded from consolidation for fear of consuming too much resource. Better to leave these exceptions ñas isò on their current physical servers rather than ñcontaminateò the virtual host server.

There are two controlling elements to a consolidation scenario; the hardware platform for the virtual host servers and the resource utilization thresholds. This consolidation scenario uses an HP DL380 G6 server as the target hardware platform. All inventoried systems have been grouped by workload within the constraints of the target hardware platform configuration. It is important to point out the memory configuration of the target hardware platform is 32GB. While the DL380 can be configured with significantly more RAM and may host more workloads, STA recommends a moderate amount of RAM to limit the number of virtual machines to a manageable number. This is crucial to managing the impact of a host server failure and the resulting recovery; the greater the number of virtual machines the greater the impact.

In building a consolidation scenario, server workloads are added to a virtual host server. Resource utilization thresholds influence how many workloads may be added to a host server. The server workloads (identified by server name within each virtual host) are mapped onto virtual host servers until a threshold is met, at which point the workload is mapped to another virtual host server. The server thresholds for this consolidation scenario were also set moderately. This was done to allow headroom for high availability purposes; there must be sufficient resource available on the surviving ESX hosts to support the virtual machines of a failed host.

Also included in this consolidation scenario are lists of reusable servers and vintage servers. Reusable servers are servers considered new enough to be redeployed elsewhere in the organization; use at a DR facility is not uncommon. Vintage servers are servers considered too old to be redeployed.

Finally, the consolidation summary is presented, showing before and after virtualization numbers and what savings, if any, are expected from the scenario. Systems with exceptions are included in the after virtualization numbers.

Hardware Platform

The target hardware platform for this consolidation scenario is as follows:

Manufacturer	**Hewlett Packard**
Model	**DL380 G6**
Number of CPU cores	**8 (2 x 4)**
CPU Speed	**2800 Mhz**
RAM	**32,768 MB**
NIC	**4 x 100/1000**
HBA	**2 x 4GB**
Hard Drives	**2 x 146GB RAID1**

ESX Server Thresholds

The following thresholds were used in this consolidation scenario. Workloads from the inventoried servers are placed into the virtual environment provided by one ESX server. When a threshold is exceeded, the ESX server is closed, a new ESX server is opened and the server workload is added to the new ESX server.

Max processor usage	**60%**
Processor queue per CPU	**4**
Max memory usage	**80%**
File Sys Cache (MB)	**600**
Page file usage	**70%**
Paging (pages/sec)	**2400**
Disk I/O transfers/sec	**1000**
Disk I/O MB/sec	**100**
Network speed (MB/sec)	**100**

Consolidated Server Layout

The table below described the consolidated server layout for this consolidation scenario. The first column is the target system followed by the source systems included in the group. For each entry, the rest of the table is devoted to the primary resources of processor, memory, disk and network, which are shown in two sections; capacity and estimated utilization.

The basic limiting factor for this consolidation grouping is the memory.

		Capacity						Estimated New Utilization								
		Processors		Memory	Disk	Network		Processor		Memory				Disk		Network
Target System Name	**Source System Name**	**Count**	**Speed (Mhz)**	**Size (MB)**	**Size (GB)**	**Count**	**Speed (MB/ sec)**	**% Used**	**Queue per CPU**	**% Used**	**File Sys Cache (MB)**	**Page File %**	**Paging (pages per sec)**	**I/O (trans per sec)**	**I/O (MB per sec)**	**Speed (MB per sec)**
New Systems																
DL380-01	Server A Server B Server C Server D Server E Server F Server G	8	2,800	32,768	144.00	4	4000.0	3.04	0.00	28.58	786.69	1.21	32.48	31.15	0.48	0.03
DL380-02	Server H Server I Server J Server K Server L Server M Server N Server O Server P Server Q Server R	8	2,800	32,768	144.00	4	4000.0	13.08	0.03	74.13	3,054.08	6.29	849.78	984.58	33.60	0.34
DL380-03	*>> reserved*	8	2,800	32,768	144.00	4	4000.0	0	0	0	0	0	0	0	0	0

Reusable Servers

The following servers contain a single CPU with a speed greater than 2,499 MHz or multiple CPUs with a combined speed greater than 10,000 MHz. They are identified as reusable:

Server A	Server D	Server G	Server I	Server K
Server B	Server E	Server H	Server J	Server L
Server C	Server F			

Vintage Servers

The following servers contain a single CPU with a speed less than 2,500 Mhz or multiple CPUs with a combined speed less than 10,000 Mhz. They are identified as vintage servers and are considered not reusable:

Server M	Server O	Server P	Server Q	Server R
Server N				

Consolidation Summary

The effect of consolidating servers for this scenario is summarized in the table below:

	Before Virtualization	After Virtualization	Saving
No. of Processors	73	24	67%
Total Memory (GB)	65	96	none
No. of Servers	18	3	83%
Power Demand (W)	11,000	2,760	75%
Thermal (BTU/hr)	36,600	10,290	72%
Spare Capacity	0	35%	

Storage Considerations

Direct attached storage (DAS), a side effect of the scale out, one application per server computing model, is the most prevalent storage technology in IT infrastructures today. Unfortunately, it is also the most inefficient and difficult to manage.

Any organization that is considering a virtual server infrastructure should give up its DAS in favor of a storage area network (SAN) before undertaking server consolidation. Only when incorporating a SAN can a virtual infrastructure fully realize the portability, reliability and availability potential of virtual machines.

Definition

SANs are networks of storage devices that are designed to allow data to move more efficiently without adding load to the LAN. SAN technology allows servers to be physically separate from their storage devices at distances up to 100 km, which is far greater than was previously possible and is highly beneficial for business recovery scenarios.

In concept, SANs are similar to LAN architectures, which allow connectivity to many devices through hubs, switches, and even directors for high availability. The main benefit of a SAN is that it virtualizes storage for a number of disparate applications, thereby reducing TCO. In addition, organizations derive significant business benefits from the improved backup and data recovery capabilities that SANs provide.

Consolidation

The consolidation of storage is one of the main functions of a SAN. The "simple" act of gathering business data into a dedicated environment obviates the need to manage disk capacity for individual application servers.

SANs enable organizations to address their overall storage requirements rather than those of departmental systems, but this is only the start of the consolidation process. The true goal of a SAN is to abstract the individual storage demands of different systems, regardless of their architecture, in order to truly realize the storage capacity of the enterprise. Virtualization can be used to maintain storage capacity across all applications with minimal downtime.

Performance

The modular nature of SANs, with storage subsystems and fabric separated and available for incremental development, makes it possible to fine tune storage in a way that is similar to tuning memory within a server. Components of good performance include:

- Appropriately loaded disks
- Appropriate numbers of spindles
- Sufficient throughput
- The ability to segregate demand

Each of these components can be managed not only in significant volumes but also individually, which makes it possible to get the best ñaverageò performance and to provide greater resources where they are required.

Scalability

Scalability within SANs can be achieved in two ways; adding hosts and adding storage.

In the scenario of adding hosts, more hosts may need to be connected to the storage subsystems as the number of enterprise users increases. In data centers, more users are often indicated by more SQL Server systems handling new applications, more Exchange servers, or more storage for file services.

In the adding storage scenario, we assume that as data stores become larger (for databases or messaging systems, for example), storage capacity requirements will increase. More storage can be added to a SAN to meet these requirements, which is a characteristic that makes it more scalable than DAS and some NAS implementations.

Benefits

SANs make efficient use of storage capacity by consolidating widely distributed disk space onto fewer storage arrays. SANs deliver a better return on capital invested in storage and servers with their increased flexibility and reduced administrative costs that result from centralization. SAN capacities vary from manufacturer to manufacturer, from gigabytes to hundreds of terabytes. SANs offer exceptional data reliability and availability.

The high capacity, scalability, and performance of SAN solutions enable service providers to offer value-added storage services such as storage on demand, dynamic storage allocation, online backup services, archiving and recovery, and offline snapshots. SANs are optimized to transfer large blocks of data between servers and storage devices; they are well suited for a wide variety of uses.

Recommendations

The recommendations made in this document are the result of analyzing the performance information gathered on eighteen (18) servers within the current ABC Inc. IT environment. The analysis shows the current ABC Inc. server infrastructure can be consolidated, achieving significant savings. The recommendations are suggested as a means to establishing a consolidated server environment and realizing the benefits it offers.

Adopt Scale-Up Architecture

STA recommends that the ABC Inc. adopt a scale-up architecture for its server infrastructure. This architecture is the basis for a consolidated server environment and represents a turning point in which:

- Multiple workloads are run on fewer, larger servers
- Workloads draw on a large pool of memory, processor, disk and network resources
- Unused computing resources are kept to a minimum

The current architecture model (scale-out) is essentially one application per server. Although this model reduces dependencies by virtue of being autonomous, it does so at a great cost of unused computing resources, increased power and cooling, and increased management.

By adopting a scale-up architecture model the ABC Inc. can more effectively manage its server resources, reduce power and cooling, reduce wasted resource and improve resource utilization.

Enhance SAN Infrastructure

ABC Inc. currently has Storage Area Network (SAN) for its mission critical servers; servers that are not mission critical utilize direct attach storage (DAS). STA recommends that ABC Inc. enhance the scope of its SAN infrastructure to meet storage requirements for non mission critical servers that would be virtualized in a VMware Virtual Infrastructure. Placing all its disk storage in a SAN would enable ABC Inc. to more easily manage storage demands and future growth.

Establish Virtual Infrastructure

STA recommends that ABC Inc. establish a virtual server infrastructure in its data center. The infrastructure must be built to support the workload requirements of the applications currently running in the (scale-out) server farm. Not only does the virtual infrastructure need to satisfy application resource requirements but also scalability and availability requirements as well. As such, the virtual infrastructure must support high availability and dynamic resource allocation.

STA recommends ABC Inc. build their virtual infrastructure using VMware vSphere 4.

Migrate Applications

STA recommends migrating Windows applications to the virtualized operating environment provided by VMware vSphere 4. ABC Inc. should develop a migration plan based upon the application properties, business schedules, age of hardware and other factors. STA acknowledges that ABC Inc. would like to establish a migration schedule based upon the age of the server hardware; servers would be migrated to the virtual infrastructure when they reach retirement. The following servers are candidates to be migrated first:

```
Server A
Server B
Server C
Server D
```

The consolidation scenario presented in this document suggests a grouping of workloads to ESX hosts based on resource consumption. The location of workloads in this context is for sizing purposes only; it is not meant to indicate where an ABC Inc. server will run after it is virtualized.

In a production vSphere virtual infrastructure the host computers would form a single resource pool in which all virtual servers (workloads) would run. The location of where any given workload runs would be fluid, constantly evaluated and set by the management server to provide the best resource utilization. Workloads that need to run together on the same host, as well as workloads that must never run on the same host, would also be handled by the management server using affinity rules. ABC Inc. should evaluate their application requirements may want to consider other factors in their consolidation groupings.

Regardless of the specific consolidation grouping, STA recommends that after a workload is migrated to the virtual infrastructure, a performance measurement (benchmark) be taken on the hosting ESX Server so that the resources are not oversubscribed.

As the workloads are migrated to the virtual infrastructure, STA recommends that the High Availability (HA) feature of VMware be enabled and configured to support those workloads that require an automatic restart after a server failure. This feature will minimize downtime and eliminate the need for dedicated stand-by hardware. This feature can also be used to provide uniform high availability across the entire virtual environment without the cost and complexity of traditional solutions.

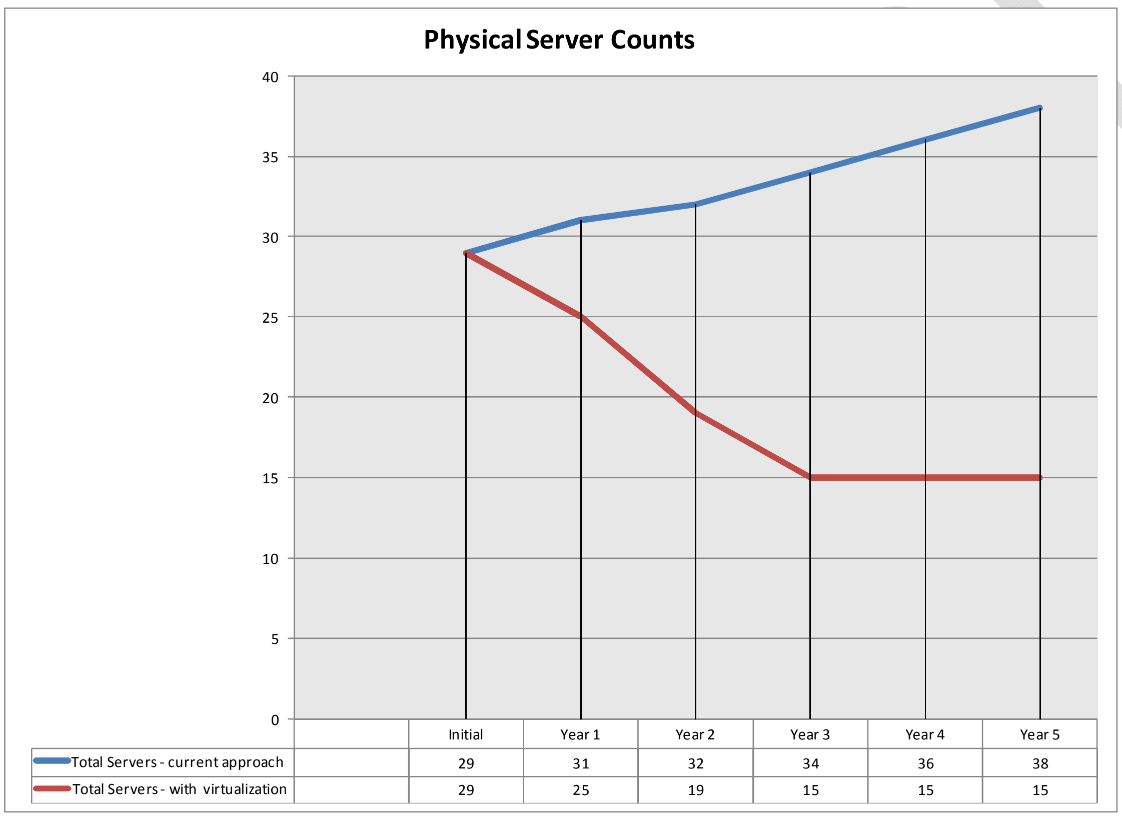
STA also recommends that ABC Inc. enable and configure the Distributed Resource Scheduler (DRS) feature to dynamically allocate and balance computing resources across a collection of hardware resources. VMware DRS continuously monitors utilization of resources and intelligently allocates available resources among the virtual machines based on pre-defined rules that reflect business needs and changing priorities. When a virtual machine experiences an increased load, VMware DRS automatically allocates additional resources by redistributing virtual machines among the physical servers.

By migrating applications to the VMware Infrastructure ABC Inc. can begin to realize the benefits of a consolidated server environment. By enabling the HA and DRS features, the platform provides more flexibility, availability and performance than its pre-consolidation counterparts.

Business Analysis

The technical analysis shows that the current ABC Inc. server infrastructure can be consolidated using virtualization. This consolidation will result in significant savings. In order to understand the magnitude of potential savings STA has developed a Total Cost of Ownership (TCO) comparison between replacing the current physical servers one-for-one vs. implementing server virtualization. Figure 1 shows the estimated number of physical servers required for each scenario with 5% annual growth.

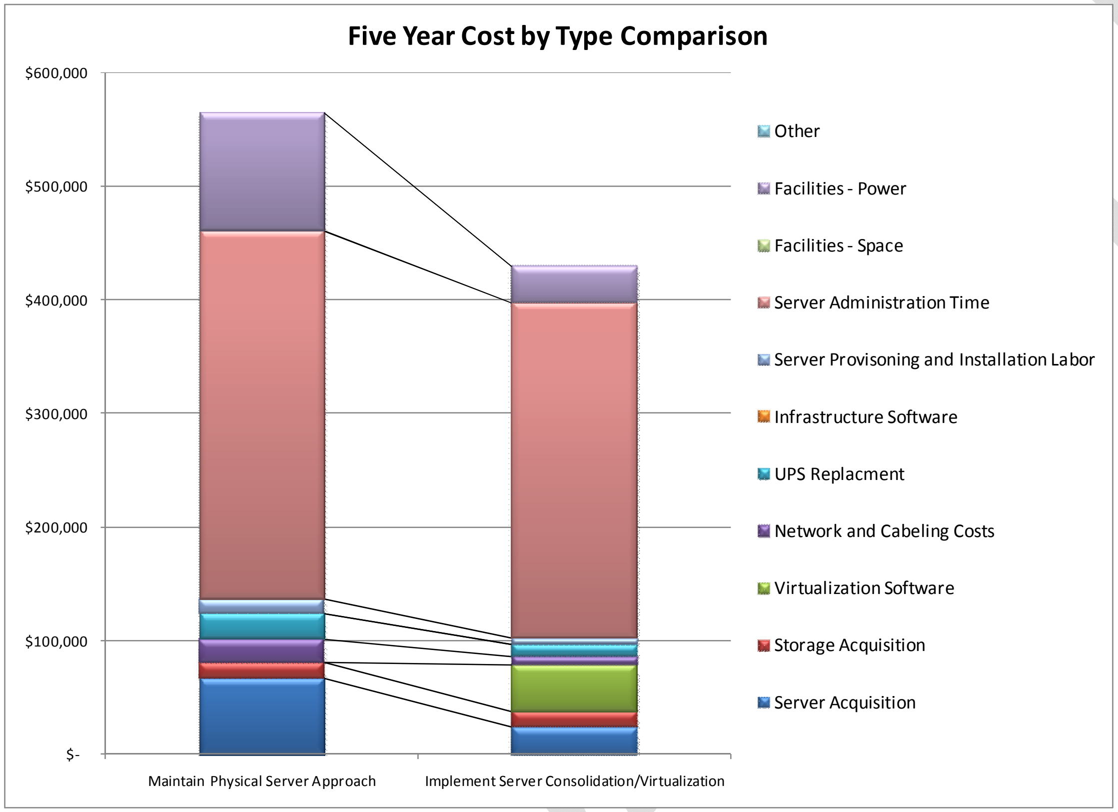
Figure 1



The analysis assumes that the current inventory of servers will be replaced over the next 3 years. This is based on an average useful life of 4 years. With virtualization, 17 of the 29 current servers will be virtualized using 3 physical host servers running VMware. This level of consolidation still leaves plenty of room for growth within the virtual host servers. The analysis assumes a 5% annual growth in the number of servers required to support ABC Inc.ôs business needs. With virtualization the additional servers added can run as virtual machines. This drops the total number of physical servers from 38 to 15 over the next 3 years. This reduction in physical servers will translate into savings on power, future server replacements, future UPS replacement, future network switch replacement, server provisioning costs, and server administration costs.

As virtualization technology improves the remaining dedicated physical servers will likely be virtualized as well. This will provide additional future savings. Figure 2 shows a comparison of total costs over a 5 year period. This includes both capital and operating cash flows.

Figure 2



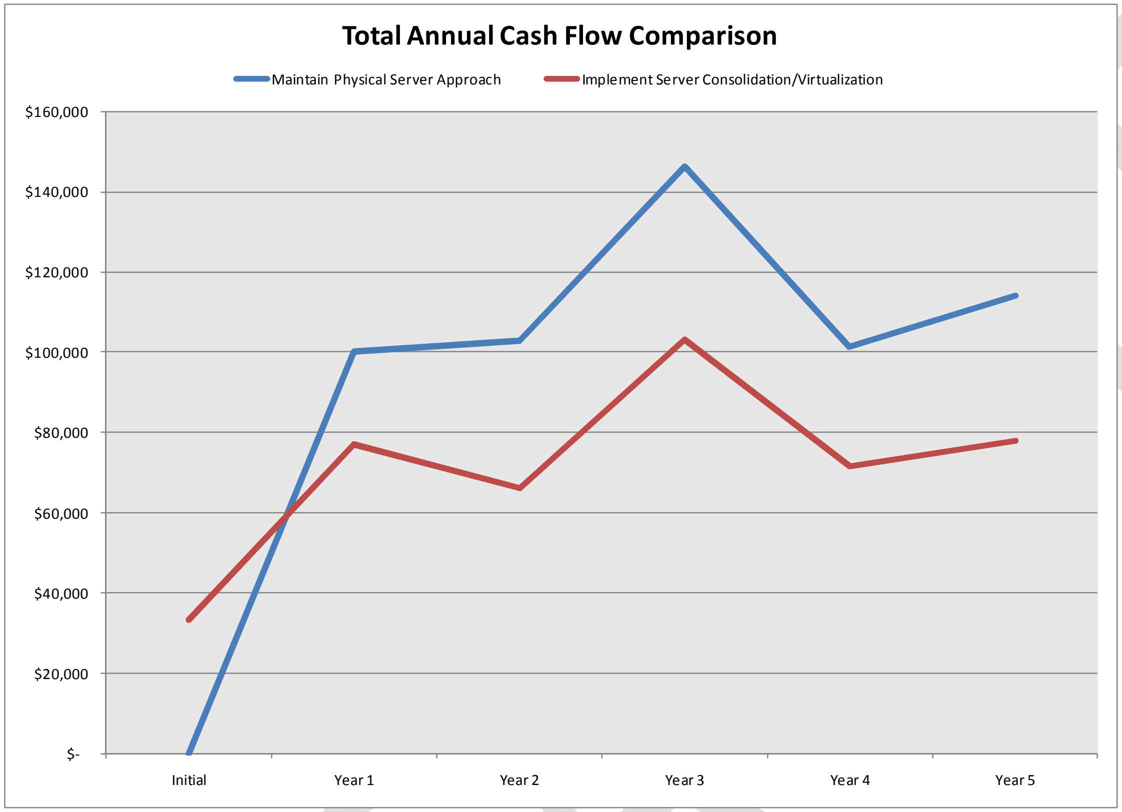
Based on the assumptions in the analysis, implementing virtualization of 17 servers and for future server growth, could result in total net savings over 5 years of 24% and $136,000.

It is likely that the servers not currently identified as virtualization candidates will become candidates within 3 to 5 years. This will increase the financial benefits realized by ABC Inc..

Metric	5 Years
NPV of Savings	$ 93,308
Payback period (months)	15
Total Investment in Virtualization	$ 64,888
Return on Net Investment	109.1%
Discount Rate Used for NPV	10.0%

Figure 3 shows a comparison of total annual cash flows. This includes both capital and operating cash flows.

Figure 3

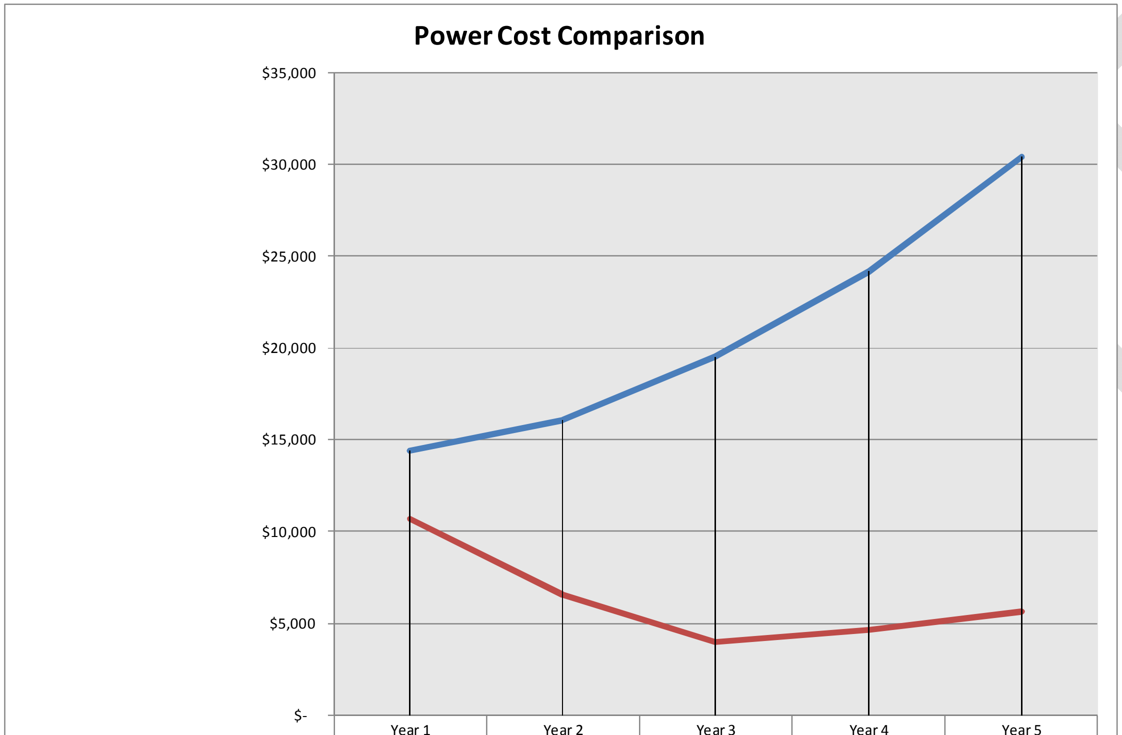


The model assumes that 2 virtual host servers and VMware software is implemented initially or early in year 1. One more virtual host is added in year 3 to complete the consolidation project. The spike in year 3 is due to replacement of servers, UPSs, shared storage, and network switching.

The capacity of the virtual host servers leaves room for high availability failover capabilities. This level of failover does not exist in the current physical only approach. The value of this additional protection from downtime is not quantified in the analysis but it is one for the most important reasons to move to virtualization.

Figure 4 shows a comparison of annual power usage and costs. The virtualization solution not only reduces power usage and costs, it significantly reduces the amount of physical hardware to maintain and dispose of in the future. If more servers are virtualized in the future then the power savings will be greater.

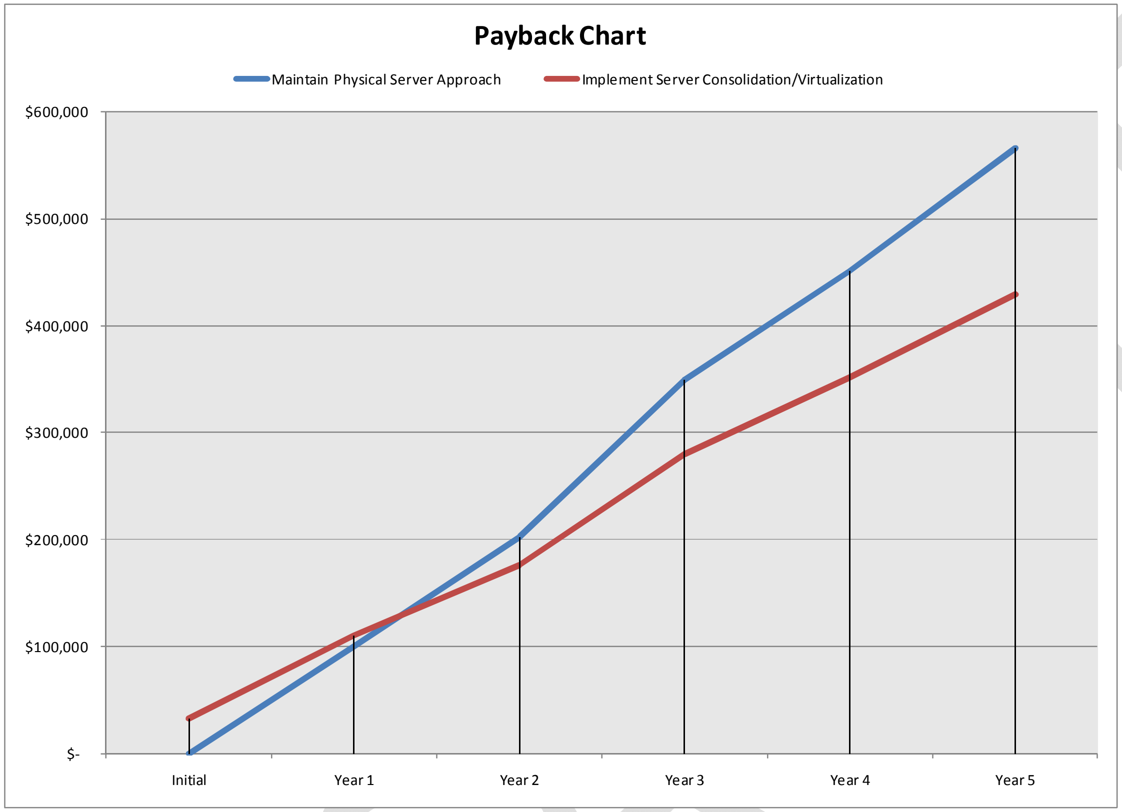
Figure 4



	Year 1	Year 2	Year 3	Year 4	Year 5
Power and Cooling - Current	$14,363	$16,036	$19,533	$24,177	$30,378
Power and Cooling - with Virtualization	$10,667	$6,532	$3,987	$4,616	$5,611

Figure 5 shows a comparison of cumulative costs. Where the lines cross, 1 year, is the point that the investment in the virtualization solution is recovered and begins to provide a return on investment.

Figure 5



Although the analysis shows significant savings and value over 5 years the following key benefits of virtualization were not quantified in the analysis but are the most important reasons to implement virtualization:

- With virtualization ABC Inc. IT will be able to provide high availability clustering capabilities. In order to match this functionality with the 1 to 1 physical server approach, the number of servers would need to be doubled and clustering software would have to be deployed.

- With virtualization ABC Inc. can lay a solid, cost effective disaster recovery foundation. A single server can be deployed in a secondary data center to support multiple, critical applications. With the physical 1 for 1 server approach this capability will require many more servers, more supporting infrastructure, and will be more complex to manage.

- With virtualization ABC Inc. IT will be able to provide a great deal of server and application deployment flexibility and agility. Test and development servers can be deployed and retired within the virtual infrastructure without taking down other servers and applications. This ability is not available with a 1 to 1 physical server infrastructure.

Table 1 shows the numbers on an annual basis. This analysis compares the total overall costs of replacing current physical servers on a 1 for 1 basis with the option of consolidation using server virtualization technology. The timing of cash flows is based on the estimated timing of server replacements and additions.

Table 1

Maintain Physical Server Approach										
	Initial	Year 1	Year 2	Year 3	Year 4	Year 5	3 Year Total	4 Year Total	5 Year Total	Notes
Server Acquisition	$ -	$ 20,500	$ 18,000	$ 18,500	$ 5,000	$ 5,000	$ 57,000	$ 62,000	$ 67,000	Replacement of current servers based on age. Estimate includes prepaid maintenance.
Storage Acquisition	-	-	-	14,059	-	-	14,059	14,059	14,059	Storage growth and drive replacement
Virtualization Software	-	-	-	-	-	-	-	-	-	
Network and Cabeling Costs	-	-	-	20,000	-	-	20,000	20,000	20,000	Replacement of network switches

UPS Replacment	-	7,000	7,000	7,000	1,000	1,000	21,000	22,000	23,000	Replacement of UPS hardware
Infrastructure Software	-	-	-	-	-	-	-	-	-	Clustering software would be required to match Vmware Enterprise functionality
Server Provisoning and Installation Labor	-	3,200	3,296	3,055	1,399	1,801	9,551	10,950	12,751	
Server Administration Time	-	55,104	58,603	64,129	69,935	76,030	177,836	247,771	323,801	
Facilities - Space	-	-	-	-	-	-	-	-	-	Estimated ongoing costs for data center space.
Facilities - Power and Cooling	-	14,363	16,036	19,533	24,177	30,378	49,932	74,108	104,486	Estimated ongoing costs for data center power.
Other	-	-	-	-	-	-	-	-	-	
Totals	$ -	$ 100,167	$ 102,934	$ 146,277	$ 101,510	$ 114,209	$ 349,378	$ 450,888	$ 565,097	

Implement Server Consolidation/Virtualization

	Initial	Year 1	Year 2	Year 3	Year 4	Year 5	3 Year Total	4 Year Total	Total	Notes

Server Acquisition	$ 15,848	$ -	$ -	$ 7,924	$ -	$ -	$ 23,772	$ 23,772	$ 23,772	Investment in virtual host server hardware with prepaid maintenance
Storage Acquisition	-	-	-	14,059	-	-	14,059	14,059	14,059	Storage growth and drive replacement
Virtualization Software	17,544	-	4,014	10,130	4,014	5,414	31,688	35,702	41,116	Investment in Vmware software licenses and support
Network and Cabeling Costs	-	-	-	8,000	-	-	8,000	8,000	8,000	Replacement of network switches
UPS Replacment	-	10,000	-	-	-	-	10,000	10,000	10,000	Replacement of UPS hardware
Infrastructure Software	-	-	-	-	-	-	-	-	-	
Server Provisoning and Installation Labor	-	1,792	1,450	1,494	839	144	4,736	5,575	5,719	
Server Administration Time	-	54,496	54,252	57,645	62,102	66,810	166,393	228,495	295,305	
Facilities - Space	-	-	-	-	-	-	-	-	-	Estimated ongoing costs for data center space.

Facilities - Power and Cooling	-	10,667	6,532	3,987	4,616	5,611	21,186	25,802	31,412	Estimated ongoing costs for data center power.
Other	-	-	-	-	-	-	-	-	-	
Totals	$ 33,392	$ 76,955	$ 66,248	$ 103,240	$ 71,571	$ 77,979	$ 279,835	$ 351,406	$ 429,385	

Difference										
Cost of Category	Initial	Year 1	Year 2	Year 3	Year 4	Year 5	3 Year Total	4 Year Total	5 Year Total	% Savings
Server Acquisition	$ (15,848)	$ 20,500	$ 18,000	$ 10,576	$ 5,000	$ 5,000	$ 33,228	$ 38,228	$ 43,228	64.5%
Storage Acquisition	-	-	-	-	-	-	-	-	-	0.0%
Virtualization Software	(17,544)	-	(4,014)	(10,130)	(4,014)	(5,414)	(31,688)	(35,702)	(41,116)	0.0%
Network and Cabeling Costs	-	-	-	12,000	-	-	12,000	12,000	12,000	60.0%
UPS Replacment	-	(3,000)	7,000	7,000	1,000	1,000	11,000	12,000	13,000	56.5%
Infrastructure Software	-	-	-	-	-	-	-	-	-	0.0%
Server Provisoning and Installation Labor	-	1,408	1,846	1,562	559	1,657	4,815	5,375	7,032	55.1%

Server Administration Time	-	608	4,351	6,484	7,833	9,220	11,443	19,276	28,496	8.8%
Facilities - Space	-	-	-	-	-	-	-	-	-	0.0%
Facilities - Power	-	3,696	9,504	15,546	19,561	24,767	28,746	48,307	73,074	69.9%
Other	-	-	-	-	-	-	-	-	-	0.0%
Totals	$ (33,392)	$ 23,212	$ 36,686	$ 43,037	$ 29,939	$ 36,230	$ 69,543	$ 99,482	$ 135,713	24.0%

Table 2 shows a comparison of the estimated number of physical and virtual hardware components required for each option, including growth. These numbers are used to calculate the timing and amount of cash flows noted in Table 1. Reducing the number of physical servers translates into the need for less supporting data center infrastructure such as UPSs and network switching.

Table 2

Maintain Physical Server Approach						
Item Type	**Initial**	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Total Physical Servers	29	31	32	34	36	38
Physical Servers Added	-	2	1	2	2	2
Total Virtual Servers	-	-	-	-	-	-
Virtual Servers Added	-	-	-	-	-	-
Total UPS Required	19	20	21	22	23	24
UPS Added	-	1	1	1	1	1
Total Network Switches Required	4	4	4	5	5	5
Network Switches Added	-	-	-	1	-	-

Total Telco Boxes	3	3	3	3	3	3
Telco Boxes Added	-	-	-	-	-	-
Total Storage Arrays	1	1	1	1	1	1
Storage Arrays Added	-	-	-	-	-	-
Total Physical Components	56	59	61	65	68	71

Implement Server Consolidation/Virtualization						
Item Type	**Initial**	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Total Physical Servers	29	25	19	15	15	15
Physical Servers Added	-	(4)	(6)	(4)	-	-
Total Virtual Servers	-	8	15	22	24	26
Virtual Servers Added for Growth	-	2	1	2	2	2
Total UPS Required	19	16	12	10	10	10
UPS Added	-	(3)	(4)	(2)	-	-
Total Network Switches Required	4	3	3	2	2	2
Network Switches Added	-	(1)	-	(1)	-	-
Total Telco Boxes	3	3	3	3	3	3
Telco Boxes Added	-	-	-	-	-	-
Total Storage Arrays	1	1	1	1	1	1
Storage Arrays Added	-	-	-	-	-	-
Total Physical Components	56	56	53	53	55	57

Table 3 shows a comparison of the estimated time required to provision physical vs. virtual servers. Once virtual host servers are deployed the task of provisioning a new virtual server requires much less time than the time required to deploy a physical server. The time recovered can be used for other tasks. A virtual infrastructure also allows a great deal more flexibility in deployment and retirement of test and development servers. This provides a much better return on investment from the infrastructure. The benefits of this flexibility were not quantified in this business analysis.

Table 3: Server Provisioning

	Physical	Virtual
Hours to provision 1 server	10	2

Maintain Physical Server Approach	Initial	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Physical Servers Provisioned	0	10	10	9	4	5	38
Hours	0	100	100	90	40	50	380
Virtual Servers Provisioned	0	0	0	0	0	0	0
Hours	0	0	0	0	0	0	0
Total Hours	0	100	100	90	40	50	380
	$ -	$ 3,200	$ 3,296	$ 3,055	$ 1,399	$ 1,801	$ 12,751

Implement Server Consolidation/Virtualization	Initial	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Physical Servers Provisioned	0	4	3	3	2	0	12
Hours	0	40	30	30	20	0	120
Virtual Servers Provisioned	0	8	7	7	2	2	26
Hours	0	16	14	14	4	4	52
Total Hours	0	56	44	44	24	4	172
	$ -	$ 1,792	$ 1,450	$ 1,494	$ 839	$ 144	$ 5,719

IT Staff Work Hours

Weeks per year	50
Days per week	5
Hours per day	10
Hours per FTE per year	2500

Average Annual IT Staff Salary & Benefits	$ 80,000
Cost per hour	$ 32.00
Growth in Labor Costs	3.0%

Table 4 shows a comparison of the estimated time required to manage physical vs. virtual servers. For reasons noted in previous sections of this report virtual servers requires less management time than the time required with physical servers. The analysis numbers in Table 4 are based on increasing the number of servers managed by one IT FTE (full-time equivalent). The time recovered can be used for other tasks. Since not all current servers are being virtualized the overall time savings is not as significant as it could be.

Table 4

	Physical	Virtual
Servers managed per FTE	45.0	45.0
VMware Enterprise Benefits		
DRS	-	9.0
vShield Zones	-	4.5
Update Manager	-	4.5
Data Recovery	-	0.3
Hot Add	-	0.3
	-	18.6
Adjusted servers per FTE	45.0	63.6
% improvement		41.3%

Maintain Physical Server Approach	Initial	Year 1	Year 2	Year 3	Year 4	Year 5
Physical Servers Managed	29	31	32	34	36	38

Server per Admin	45.0	45.0	45.0	45.0	45.0	45.0
Hours per year		1722	1778	1889	2000	2111
Cost per hour	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00
	$ -	$ 55,104	$ 58,603	$ 64,129	$ 69,935	$ 76,030

Maintain Physical Server Approach	Initial	Year 1	Year 2	Year 3	Year 4	Year 5
Virtual Servers Managed	0	0	0	0	0	0
Server per Admin	63.6	63.6	63.6	63.6	63.6	63.6
Hours per year		0	0	0	0	0
Cost per hour	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00
	$ -	$ -	$ -	$ -	$ -	$ -
	$ -	$ 55,104	$ 58,603	$ 64,129	$ 69,935	$ 76,030

Implement Server Consolidation/Virtualization	Initial	Year 1	Year 2	Year 3	Year 4	Year 5
Physical Servers Managed	29	25	19	15	15	15
Server per Admin	45.0	45.0	45.0	45.0	45.0	45.0
Hours per year		1389	1056	833	833	833
Cost per hour	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00
	$ -	$ 44,448	$ 34,806	$ 28,279	$ 29,128	$ 30,002

Implement Server Consolidation/Virtualization	Initial	Year 1	Year 2	Year 3	Year 4	Year 5
Virtual Servers Managed	0	8	15	22	24	26
Server per Admin	63.6	63.6	63.6	63.6	63.6	63.6
Hours per year		314	590	865	943	1022
Cost per hour	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00	$ 32.00

$ -	$ 10,048	$ 19,446	$ 29,366	$ 32,974	$ 36,809
$ -	$ 54,496	$ 54,252	$ 57,645	$ 62,102	$ 66,810

IT Staff Work Hours

Weeks per year	50
Days per week	5
Hours per day	10
Hours per FTE per year	2500

Average Annual IT Staff Salary & Benefits	$ 80,000
Cost per hour	$ 32.00
Growth in Labor Costs	3.0%

Table 5 shows the general assumptions used in the analysis.

Table 5

General Assumptions	
Server Growth Rate	5.0%
Steady-state constant used to convert nameplate power consumption to steady state	1.00
Power Usage Effectiveness (PUE) for the datacenter	1.30
Average price of electricity (cost per kWatt hour)	$ 0.12
Growth in Power Cost	5.0%
Data center operating hours per day	24
Data center operating days per week	7

Data center operating weeks per year	52
Data center operating hours per year (calculated value)	8736
Cost of Capital for NPV Calculation	10.0%
Average Server Admin annual salary & benefit	$ 80,000
Growth in Labor Costs	3.0%
IT Staff Work Hours	
Weeks per year	50
Days per week	5
Hours per day	10
Hours per FTE per year	2500
Average Server Replacement Cost 1 proc. Including 3 yrs of support.	$ 2,500
Average Server Replacement Cost 2 proc. Including 3 yrs of support.	$ 3,000
Average cost of 1500KVA UPS	$ 1,000
Average cost of network switch	$ 4,000