

RP
00/01
c.1

SERVERS and STORAGES
CONSOLIDATION AND VIRTUALIZATION

by

Mohammad Badran

Project submitted in partial fulfillment of the requirements for the Degree of
Master of Science in Computer Science

B

Department of Computer Science and Mathematics

LEBANESE AMERICAN UNIVERSITY

July 2008



LEBANESE AMERICAN UNIVERSITY

School of Arts and Sciences - Beirut Campus

Project Approval Form

Student Name: Mohammad Badran I.D. #: 200102211

Project Title : Servers and Storages: Consolidation and Virtualization

Program : Masters of Science

Division/Dept : Computer Science and Mathematics

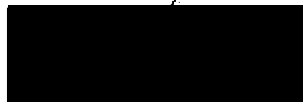
School : Arts and Sciences

Approved by: Ramzi A. Haraty

Project Advisor:



Member : AbdulNasser Kassar



Date July 8, 2008

Plagiarism Policy Compliance Statement

I certify that I have read and understood LAU's Plagiarism Policy. I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.

This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: Mohammad Badran

Signature:

A solid black rectangular box used to redact the signature of the student.

Date: Friday, July 11, 2008

I grant to the LEBANESE AMERICAN UNIVERSITY the right to use this work, irrespective of any copyright, for the University's own purpose without cost to the University or its students and employees. I further agree that the University may reproduce and provide single copies of the work to the public for the cost of reproduction.

To my parents

Acknowledgment

This research would not have been possible without the help and assistance of many persons.

First I would like to express my gratitude to my supervisor Dr. Ramzi Haraty to whom I am particularly indebted for his substantial aid, advice, and assistance from the beginning to the end of my study.

I am also deeply grateful to my faithful colleagues Rose Abrahamian, Wissam Aboushakra and my manager Sonia Haddad.

I would like to express my sincere gratitude to the Lebanese American University whose financial support during my graduate studies made it all possible. Finally, I would like to thank my friends and family for their love and long support.

Abstract

Many companies and academic institutions are seeking a highly available and scalable storage system along with a consolidated cutting-edge high performance and easy to manage servers. These are the corner stones for any modern data center nowadays. This project will try to shed lights on such technologies and present a road map to acquire and successfully achieve and deploy “consolidation and virtualization” using blade servers and fiber channel storage area network solutions.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION.....	1
CHAPTER II: DAR AL-HANDASAH EXISTING DATA CENTER	3
II.1 - COMPANY OVERVIEW.....	3
II.2 - "EXISTING" DATA CENTRE AT DAR AL-HANDASAH AND ITS PROBLEMS.....	4
CHAPTER III: DAR AL-HANDASA NEW DATA CENTER.....	6
III.1 - OVERVIEW	6
III.2 - BLADE SERVERS AND ENCLOSURE SYSTEM: CONSOLIDATION.....	7
III.2.1 - Blade System Overview: Consolidation.....	7
III.2.2 - Benefits of Blade System.....	7
III.2.3 - Requirements of Blade System.....	8
III.2.4 - HP BladeSystem Servers c-Class	11
III.2.5 - HP ProLiant BL460c.....	12
III.2.5.a - HP ProLiant BL460c.....	12
III.2.5.b - HP ProLiant c-Class server blade architecture	13
III.2.5.c - Xeon Quad-core processors	14
III.2.5.d - I/O Technologies in BladeServers HP ProLiant BL460c.....	15
III.2.6 - HP BladeSystem c7000 Enclosure.....	16
III.2.6.a - 2.6.a Overview.....	16
III.2.6.b - Architecture	18
III.2.6.c - Power and Cooling.....	21
III.2.6.d - Managing the c7000 HP BladeEnclosure.....	22
III.2.6.e - LAN and Fiber Connectivity the c7000 HP BladeEnclosure.....	23
III.2.6.f - HP c-Class 7000 BladeSystem AT WORK.....	24
CHAPTER IV: STORAGE AREA NETWORK AND VIRTUALIZATION: HP-EVA SAN	26
IV.1 - SAN STORAGE OVERVIEW.....	26
IV.2 - BENEFITS OF SAN STORAGE	27
IV.3 - DAR AL-HANDASAH HP SAN SOLUTION	28
IV.3.1 - HP Storage Area Network Solution from HP: EVA6100.....	29
IV.3.1.a - HP EVA6100 Overview	29
IV.3.1.b - HP EVA6100 Architecture.....	30
IV.3.1.c - HP EVA6100 SAN Specifications:	30
IV.3.1.d - HP EVA6100 Redundancy.....	31
IV.3.1.e - HP EVA6100 Storage Virtualization: vDisk and vRAID technologies.....	32
IV.3.1.f - The Value of EVA Virtualization	35
IV.3.1.g - HP SAN EVA6100 at WORK	36
CHAPTER V: DAR AL-HANDASAH HP BLADE SERVERS AND SAN SETUP	38
V.1 - OVERVIEW.....	38
V.2 - SYSTEM ARCHITECTURE (WITH BRIEF CONFIGURATION)	38
V.3 - BILL OF QUANTITIES FOR THE SOLUTION	39
CHAPTER VI: CLUSTERED FILE SYSTEM-HP POLYSERVE MATRIX SERVER.....	43
VI.1 - OVERVIEW	43
VI.2 - HP POLYSERVE CLUSTERED FILESYSTEM-ARCHITECTURE	43
VI.3 - HP POLYSERVE CLUSTERED FILESYSTEM-CONFIGURATION AND MANAGEMENT	45
CHAPTER VII: EXTENDING THE ROADMAP.....	46
CHAPTER VIII: PROBLEMS ADDRESSED AND ENHANCEMENTS ACHIEVED	48
CHAPTER IX: CONCLUSION.....	49
REFERENCES.....	50

List of Figures

FIGURE III-1: HP C-CLASS BLADESYSTEM-AT WORK.....	12
FIGURE III-2:HP BLADESYSTEM C7000 ENCLOSURE – FRONT VIEW.....	16
FIGURE III-3: HP BLADESYSTEM C7000 ENCLOSURE – FRONT VIEW.....	17
FIGURE III-4: HP BLADESYSTEM C7000 ENCLOSURE – SIDE VIEW.....	19
FIGURE III-5: HP BLADESYSTEM C7000 ENCLOSURE – REAR VIEW.....	20
FIGURE III-6: FOUR COOLING ZONES IN THE C7000 ENCLOSURE.....	21
FIGURE III-7: EXAMPLE OF POWER EFFICIENCY WITH AND WITHOUT DYNAMIC POWER SAVER.....	22
FIGURE III-8: HOME PAGE FOR ILO2 INTERFACE BLADEENCLOSURE.....	24
FIGURE III-9: ILO2 INTERFACE BLADEENCLOSURE.....	25
FIGURE IV-1: THE HP ENTERPRISE VIRTUAL ARRAY 6100 AND DISK CONTROLLERS.....	29
FIGURE IV-2: TRADITIONAL RAID COMPARED TO VRAID.....	34
FIGURE IV-3: DISTRIBUTED SPARING OR “VIRTUAL HOT SPARE”.....	35
FIGURE V-1: PHYSICAL/LOGICAL LAYOUTS FOR THE HP BLADE SYSTEM AT DAR AL- HANDASAH.....	41
FIGURE V-2 : PHYSICAL/LOGICAL LAYOUTS FOR THE HP BLADE SYSTEM AT DAR AL- HANDASAH.....	42

List of Tables

TABLE III-1: HARDWARE SPECIFICATIONS.....	13
TABLE V-1: BILL OF QUANTITIES.....	39

CHAPTER I: INTRODUCTION

The project will mainly focus on two main topics in the world of information technology infrastructure and its deep complexity which are: servers and storages – consolidation and virtualization. The work will focus on the most recent and advanced technological trends in the servers and storage solutions that the major hardware suppliers are delivering like: HP, DELL, EMC, and IBM. The project will discuss in details the HP blade servers and HP c-Class blade chassis: how consolidation is achieved, advantages of blade servers, hardware component of HP BL460c blade server. It will also focus on the SAN storage solution provided by the world's first leader in storage technology: Hewlett Packard. The project will prove that the Fiber Channel SAN storage solution provided by HP, the EVA-Enterprise Virtual Array, is the best in the SAN market. It will explain how HP achieved virtualization on the storage level: hardware and software.

In addition to the above two major topics the work proposes and digs deep in one of the most reliable and robust file systems: the HP Clustered File System or PolyServe Matrix Server. This clustered file system, unlike most clusters, meets both cluster goals: high availability and load balancing. This project discusses how PolyServe Matrix Server, a new HP product, can provide a clustered file system for many high demanding applications and organizations.

II.1 Scope of the Report

As mentioned above, the project will discuss three main topics:

- Blade servers [consolidation]: HP blade servers c-Class blades,
- SAN storage [virtualization]: HP EVA SAN storage (Enterprise Virtual Array), fiber channel disks, fiber channel switches, and
- Clustered file system: HP PolyServe Matrix Server.

This project presents a roadmap designed for any company or academic institution willing to implement blade servers with its consolidation features and SAN

storage with its virtualization advantages to get advantage of all its benefits from processing power, storage scalability, high availability and much more.

II.2 Organization of the Report

The project report is organized as follows: Chapter 2 discusses the company overview and the existing data center at Dar Al-Handasah with all its deficiencies. Chapter 3 presents and explores the new HP BladeSystem elements: Blade servers and Blade Enclosure. Chapter 4 discusses the fiber channel storage area network and its virtualized storage system. Chapters 5 presents the new blade system setup at Dar Al-Handasah. Chapter 6 discusses a highly reliable clustered file system, which provides high data rates: the HP PolyServe Matrix Server. Chapter 7 and Chapter 8 extend the roadmap for consolidation and virtualization and summarize the problems and enhancements achieved. Chapter 9 concludes the report.

CHAPTER II: DAR AL-HANDASAH EXISTING DATA CENTER

II.1 - Company Overview

Dar Al-Handasah (Shair and Partners) is a leading international consultancy specializing in architecture, engineering, planning, environment and economics. Founded in 1956, the firm is today ranked as one of the world's top international design consultants in the world. With main operations in London, Cairo and Beirut, Dar Al-Handasah is staffed with over 2,500 professionals located in 44 offices throughout the Middle East, Europe and Asia.

Dar Al-Handasah innovative plans, designs and successful implementations are due to a number of factors:

- Half a century of global experience in all type of major large scale developments in engineering and design,
- Talented and dedicated staff of professionals with multi-disciplinary, multi-culture and multi-lingual experience,
- Skills for developing, adapting and transferring latest high-tech planning, engineering and design techniques to indigenous and existing conditions in emerging environments, and
- A commitment to build long-term client-relationships and actively support our clients worldwide.

Dar's early beginnings date back to 1956 to a small engineering practice in Beirut. As a young, innovative firm, it quickly gained recognition with designs in infrastructure, public facilities, housing and industrial plants. Further ground-breaking solutions enhanced the company's reputation internationally and its range of services grew. During the next decades, the firm expanded into a full-service firm, providing master planning, architecture, engineering, and economic services.

In the 1970s, it expanded geographically throughout the Middle East and Africa. In addition to Beirut, two other main design headquarters were opened in Cairo and London.

As economics gained momentum into the 1980s, the Middle East modernized and developed at an incredible pace. Dar Al-Handasah was honored to play a leading role in planning and designing many countries' infrastructure networks, buildings and new communities. Among notable assignments was the principal role of designing and supervising all major road systems, water and sanitation networks for the Kingdom of Saudi Arabia. During this decade of expansion, Dar also diversified geographically and acquired Perkins and Will and T.Y. Lin International in North America and Asia, and Penspen and Pierre-Yves Rochon in Europe.

Today, the multi-disciplinary services and the geographic range allow Dar Al-Handasah to better service clients and the community, as it continues engineering developments throughout Africa, Asia and the Middle East [14].

II.2 - "Existing" Data Centre at Dar Al-Handasah

The existing data center for Dar Al-Handasah was first deployed in year 2000. Briefly, the setup consisted of standalone servers: HP Proliant ML350 + IBM xSeries in addition to a backup library with SDLT tape technology 160 GB. Moreover, the storage technology that Dar Al-Handasah used to store data files, mailboxes and archive files consisted mainly of DAS (Direct Attached Storage). All data files were stored on hard drives directly attached to each server. Thus it hindered the file transfer and access to these files a lot. Storage flexibility was very minimal at Dar Al-Handasah since in order to increase more space one needs to either install new hard drives or a new server with enough hard disk to host our new disk volumes taking into consideration the hassle around taking track of all these disk links.

In summary, the main problems that Dar Al-Handasah is facing are:

- Hardware maintenance for extremely different and heterogeneous servers,
- Hardly scalable storage system,
- Old servers, slow performance hard drives for file access, and
- Unreliable file storage system.

•

CHAPTER III: DAR AL-HANDASA NEW DATA CENTER

III.1 - Overview

In this section the project will provide a roadmap to the new cutting-edge technologies in the world of servers and storage via discussing the implementation of HP blade servers c-Class with the HP-SAN storage solution: the Enterprise Virtual Array in Dar Al-Handasah Data Center. This project will be a comprehensive guide for any company willing to implement and apply consolidation and virtualization.

As mentioned above the project will discuss three main topics:

- Blade servers [consolidation]: HP blade servers c-Class Blades,
- SAN storage [virtualization]: HP EVA SAN storage (Enterprise Virtual Array) ,
Fiber Channel disks, Fiber Channel switches, and
- Clustered File System: HP PolyServe Matrix Server.

~~These technologies were successfully implemented at Dar Al-Handasah data center which permanently solved the resident bottlenecks, deficiencies and problems that the company has been facing for 7 to 8 years: hardly scalable storage system, old servers, slow performance CPU processing, unreliable file storage system, and decentralized storage.~~

III.2 - Blade Servers and Enclosure System: Consolidation

III.2.1 - Blade System Overview: Consolidation

Currently, the cutting-edge, most reliable and cost-efficient server system is the BladeSystem. The blade servers along with its compact enclosure are becoming the defacto standard for setting up any data center for many reasons [3].

As data center technologies have changed the legacy rack-mounted servers' infrastructure has increased the complexity of managing and deploying these servers in any new data center. To avoid and get rid of these deployment and maintenance tasks of servers, most companies started to move to consolidated solutions that decreases the cabling, hardware maintenance, and complex deployment of new servers. The blade servers along with its well-designed enclosure have these features and consist of all the elements of storage, compute and network in one single 10-U enclosure [4].

III.2.2 - Benefits of Blade System

The main benefits of the blade system consist of the following:

- Easy hardware maintenance for extremely different and heterogeneous servers,
- Very scalable storage system,
- High performance hard drives for file access,
- Highly reliable and redundant file storage system,
- Increased cost saving compared to traditional rack-mounted servers: fewer wires, power cables, redundancy from the start,
- Easy and fast change and upgrade for any server: using the HP Virtual Connect modules in the enclosure, one can replace a server in a "plug-and-play": replace and recover from hardware failures easily and quickly, and

- Energy-economical: due to the new design and technologies that HP introduced in its blade enclosure (HP Thermal Logic and FAN Zones), power consumption and cooling is more efficient than rack-mounted servers [6].

III.2.3 - Requirements of Blade System

The core requirements and specifications of the Servers and Storage Products are categorized and summarized by the following points:

A. Servers:

1. Blade Servers:

1.1 General Specifications:

- Support for up to two Quad-Core Intel Xeon processors,
- Support for a minimum of 32GB RAM,
- Support for 4GB RAM memory modules,
- Support for a minimum of two hot pluggable SAS/SATA hard disk drives,
- Support for RAID 0 and RAID 1,
- Support for virtualization,
- Support for a minimum of two x Gigabit Ethernet ports, and
- Support for a minimum of two x multifunction TOE Gigabit Ethernet ports (for iSCSI Connectivity) and/or two x 4/2/1Gbps auto negotiating speeds Fiber Channel ports.

1.2 Required Minimum Configurations:

Server type A:

- Two 2.33 GHz Quad Core Intel Xeon processors,
- 32GB RAM (8x4GB),

- Two 146GB 10K RPM SAS HDDs,
- Two x Gigabit Ethernet ports, and
- Two x multifunction TOE Gigabit Ethernet ports or two x 4/2/1Gbps Fiber Channel ports depending on the corresponding storage networking option.

Server type B:

- Two 2.33 GHz Quad Core Intel Xeon processors,
- 8GB RAM (2x4GB),
- Two 146GB 10K RPM SAS HDDs,
- Two x Gigabit Ethernet ports, and
- Two x multifunction TOE Gigabit Ethernet ports or two x 4/2/1Gbps Fiber Channel ports depending on the corresponding storage networking option.

Server type C:

- One 2.33 GHz Quad Core Intel Xeon processors,
- 8GB RAM (2x4GB),
- Two 146GB 10K RPM SAS HDDs,
- Two x Gigabit Ethernet ports, and
- Two x multifunction TOE Gigabit Ethernet ports or two x 4/2/1Gbps Fiber Channel ports depending on the corresponding storage networking option.

2. Blade Enclosure:

- Redundant, hot swappable power supplies,
- Redundant, hot swappable cooling fans,
- Switched Ethernet connectivity to the LAN,
- iSCSI and FC SAN connectivity options to be included based on the corresponding storage networking option, and

- Connectivity between blade servers and enclosure should provide full redundancy on all active components.

3. Manageability Components:

- Basic systems management, diagnostics, monitoring, and alerting software,
- Image management software and hardware for automated/manual server,
- Deployment and recovery. Server recovery mechanism should be described in details, and
- Physical and/or IP KVM console.

B. SAN Storage:

1. Storage Array:

- Switched architecture between storage controllers and disk shelves,
- Redundant storage array controllers,
- Support for mixed drive types and sizes (FC and FATA/SATA),
- Minimum fully loaded configuration: 100 HDDs (equivalent to 50TB of raw capacity if all disks are FATA/SATA 500GB),
- Support for at least RAID 1 and RAID 5 configurations,
- Support for Microsoft Windows clustering,
- Policy based file migration functionality between FC disks and FATA/SATA disks (tiered storage approach),
- Snapshot and cloning functionalities, and
- Storage management/provisioning software.

2. Storage Networking:

- Fiber channel connectivity to be provided natively on all components of the storage area network (servers, SAN switches, and storage array controllers),

- Two storage array controllers with a minimum of two 4Gbps FC ports/controller,
- A link aggregation mechanism (for bandwidth aggregation) should be supported on all storage area network components, and
- Fiber channel switches (4Gb/s) that support advanced networking features such as link aggregation and VSANs,

3. General Requirements:

- Redundant SAN Switches,
- Redundant Supervisor modules on the SAN Switches (if applicable),
- Redundant power supplies on the SAN Switches,
- Fully redundant connectivity to SAN switches for all servers and storage controllers,
- 30% minimum available SAN switch ports after connecting all servers and storage array,
- Clustering should be supported regardless of the adopted storage networking option, and
- Boot from SAN (optional).

III.2.4 - HP BladeSystem Servers c-Class

In this section, we introduce the cutting edge blade servers and enclosure, the HP c-Class BladeSystem. This section can benefit anyone who needs to know more and deeply about what are blade servers.

The project will start exploration to the blade servers and blade enclosure with a vivid picture about this system at work (see Figure III-1).

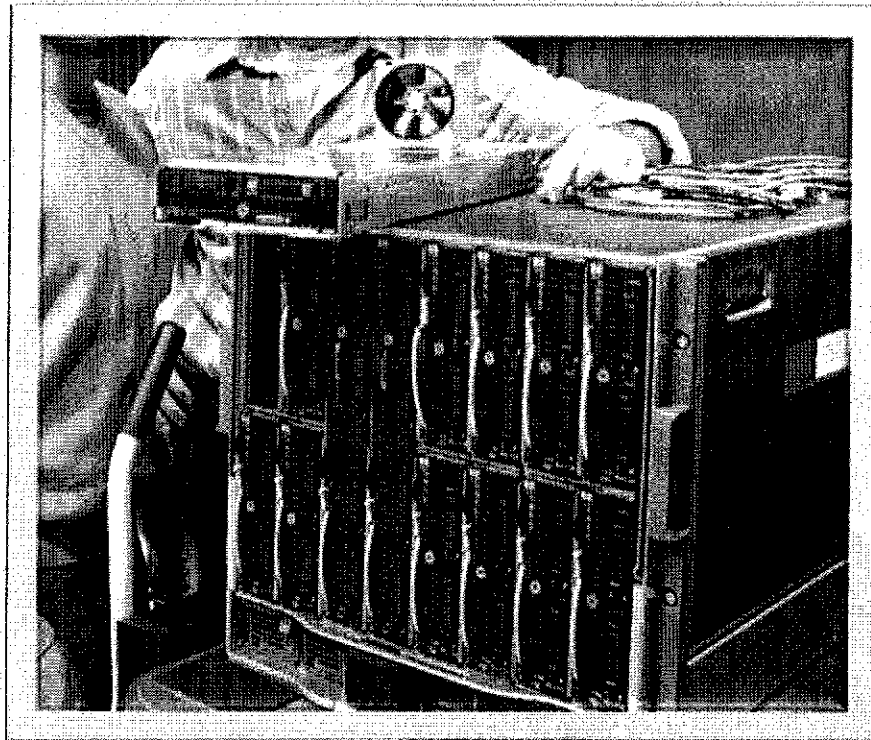


Figure III-1: HP c-Class BladeSystem-at work [3]

The most descriptive sentence that can be said about the HP c-Class Blade system is:

“HP BladeSystem: a smarter way to build an infrastructure. One box, full of possibilities” [8].

Dar Al-Handasah chose the HP ProLiant BL460c server to fill our HP c-Class Blade Enclosure with servers. Each enclosure can contain 16 half-height servers.

III.2.5 - HP ProLiant BL460c

III.2.5.a - HP ProLiant BL460c

This server is an ideal choice for customers looking for flexibility, performance and reliability in an enterprise-class server blade. Refer to Table III-1 to have an overview about its hardware specifications.

Table III-1: Hardware Specifications

Number of processors	1–2
Number of cores	8
Processors supported	Dual- and Quad-core Intel® Xeon® processors: up to 3.0 GHz
Cache	Up to 8 MB L2 (2 x 4 MB)
Maximum memory	32 GB
Network ports	2
Drives supported	Up to 2
Maximum internal storage	292 GB
I/O expansion	2 PCIe Mezzanine Expansion Slots
Form factor	6U or 10U enclosure

III.2.5.b - HP ProLiant c-Class server blade architecture

An HP ProLiant c-Class server blade is a complete server that slides into an HP BladeSystem c-Class enclosure. The most important feature for the blade enclosure is the easy way to add more blade servers to the enclosure. ProLiant c-Class server blades are built in standard form-factors, referred to as *half-height vertically measures* (4U) and *full-height* (8U) vertically measured. Both half-height and full-height server blades fit into any device bay in a BladeSystem c-Class enclosure.

The ProLiant c-Class server blades include enterprise-class technologies [9]:

- Two or four AMD, Intel x86, or Intel Itanium® processors,
- Hot-pluggable internal disk drives,
- Memory modules using advanced memory technologies such as fully buffered DIMMs,
- Multiple slots for I/O cards,

- Embedded multifunction Gigabit Ethernet network interconnects that use TCP/IP offload engine (TOE),
- Thermal logic technologies, and
- Power management.

Half-height server blades (see Figure III.2) support up to two processors, eight DIMM slots, two hot-pluggable drives, two embedded Gigabit Ethernet adapters, and two mezzanine slots for optional cards that provide a variety of different I/O fabric connectivity to the eight interconnect bays.

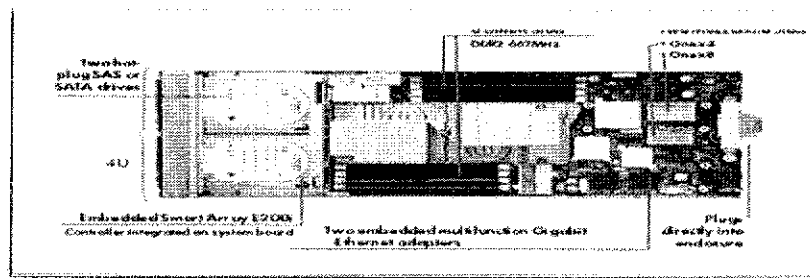


Figure III.2: Internal view of the half-height HP BladeSystem BL460c G5

III.2.5.c - Xeon Quad-core processors

Intel Xeon-core 5300 and 5400 processors have dual socket platforms (Figure III.3). The HP blade servers (half or full height) are equipped with Xeon 5300/5400 series processors giving the server great computing power with small form factor size.

Important specifications that the Xeon-processors have are:

- Two dual cores where each core has a shared L2 cache,
- Operate at a maximum frequency of 3.0-3.16 GHz (1333 MHz FSB) with 4 MB of L2 cache per dual core (8 MB total), and
- Intel 5000P series chipset delivers a significant increase in processing capacity (see Figure III.3).

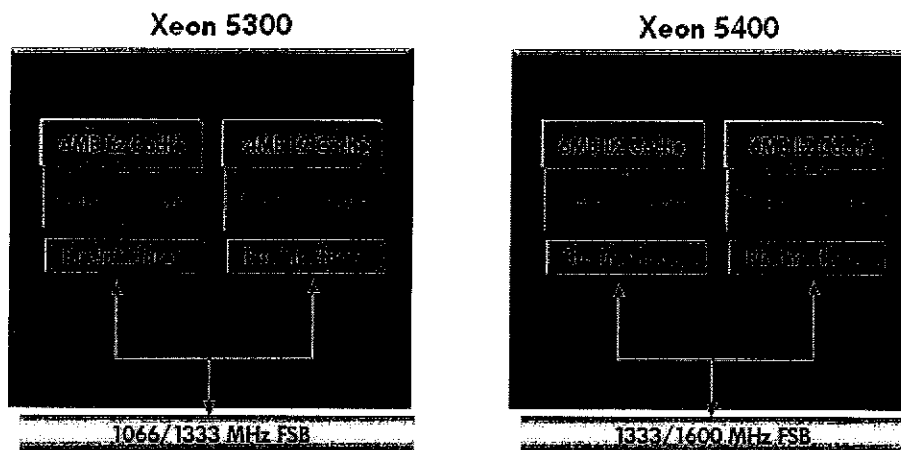


Figure III.3: Major components of Intel Xeon Quad-core 5300 and 5400 Series processors

III.2.5.d - I/O Technologies in BladeServers HP ProLiant BL460c

HP ProLiant c-Class Server Blades support:

- PCI Express (PCIe),
- Serial attached SCSI (SAS), serial ATA (SATA) I/O technologies,
- Multifunction 10 Gigabit Ethernet, Multifunction and standard Gigabit Ethernet
- 4 GB Fibre Channel, and
- 4X DDR (20 Gb) InfiniBand. SAS is a serial communication protocol for direct-attached storage devices such as SAS and SATA Small Form Factor (SFF) disk drives [10].

The project will not dig into deep through the various components of the HP BladeServer BL460c such as:

- Serial Attached SCSI technology,
- RAID controller,
- Mezzanine cards,
- TCP/IP Offload engine,
- Power management, and
- Thermal technologies [9].

This project will focus on exploring the great features and services of the blade enclosure that combines and packages all these blade servers together and provide us a really central point to manage all our servers achieving our goal: consolidation.

III.2.6 - HP BladeSystem c7000 Enclosure

III.2.6.a - Overview

As the report started discussing the HP blade servers with a figure, the project will also repeat the same thing here to better understand the details (see Figure III-2).

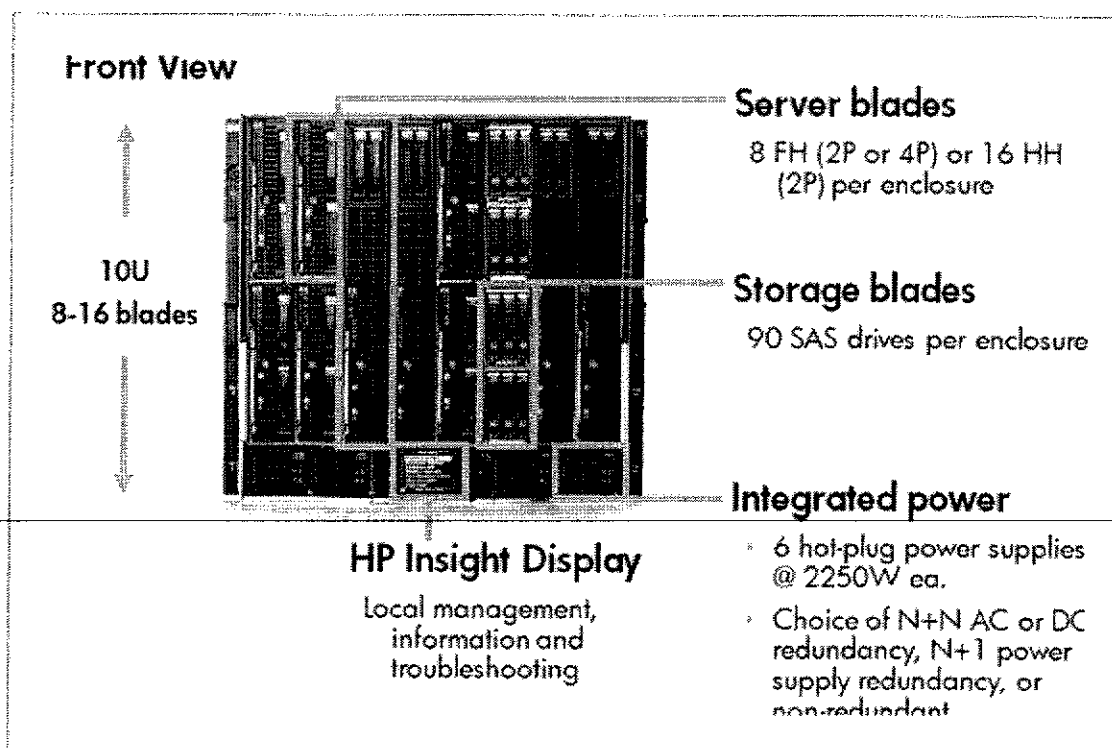


Figure III-2:HP BladeSystem c7000 Enclosure – front view [1]

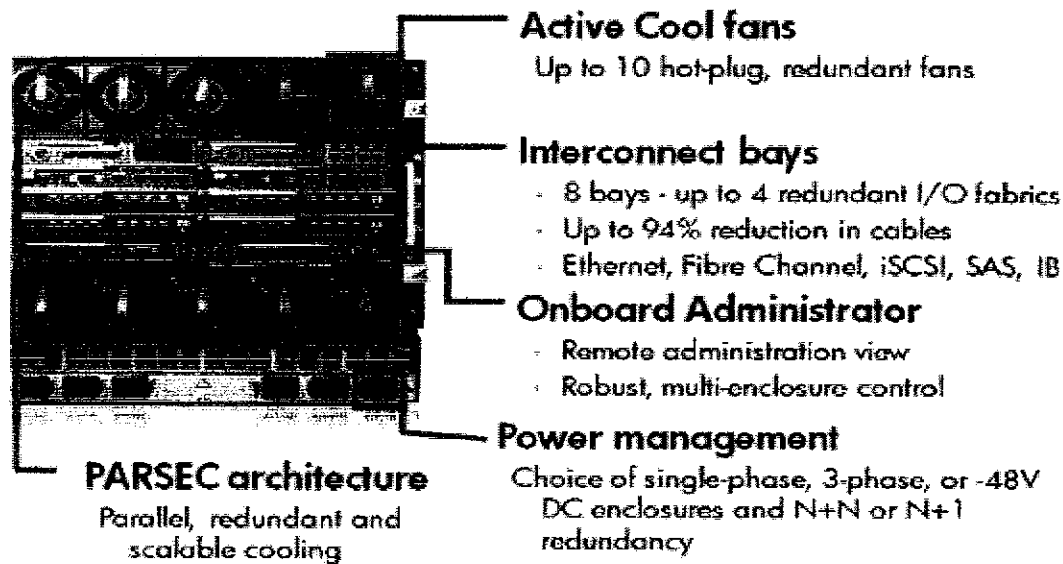


Figure III-3: HP BladeSystem c7000 Enclosure – front view [3]

The blade enclosure in general is a key element in the blade system as a whole, because through it all blade servers operate: power, LAN connectivity, Fiber SAN connectivity, cooling, and management. The blade enclosure hosts all blade servers and provides cooling, power and connectivity to the outside world. Literarily speaking, consolidation all happens in the blade enclosure itself (see Figure III-3).

HP Blade System c7000 Enclosure was the first enclosure to be built from scratch after its predecessor class the p-class blade enclosure. Major enhancements have been inserted to the c-class enclosure to increase redundancy and improve consolidation. It is especially designed to enterprise level and business-critical applications for the corporate.

The HP c7000 blade enclosure provides:

- Power,
- Cooling,
- I/O infrastructure needed to support modular server, and
- Interconnect and storage components.

The key element in the HP c-Class Blade enclosure is seamless manageability that it provides for the data center administrators. This easy management system is really easy and user-friendly because HP created and integrated many technologies and applications such as:

- Integrated Lights-Out 2 (iLO 2) server management processor in every server blade,
- Onboard Administrator that manages the entire enclosure, and
- Insight Control management software provides integrated management of server blades [13].

The HP c7000 Blade system paid much concern for redundancy on all its active components to increase the up time for all servers and to decrease the failover period in case of any hardware failure. Nearly, everything inside the c7000 blade enclosure is redundant. The c7000 blade enclosure is designed to force the signal midplane and the separate power backplane to have no active components.

III.2.6.b - Architecture

The HP c7000 blade enclosure is 10U high. The blade servers that the c7000 blade enclosure hosts are connected to the outside world (LAN and fiber SAN) through the interconnect modules. These interconnect modules can be Ethernet switches or fiber optic switches or any applicable connectivity module (see Figure III-4).

Rear View

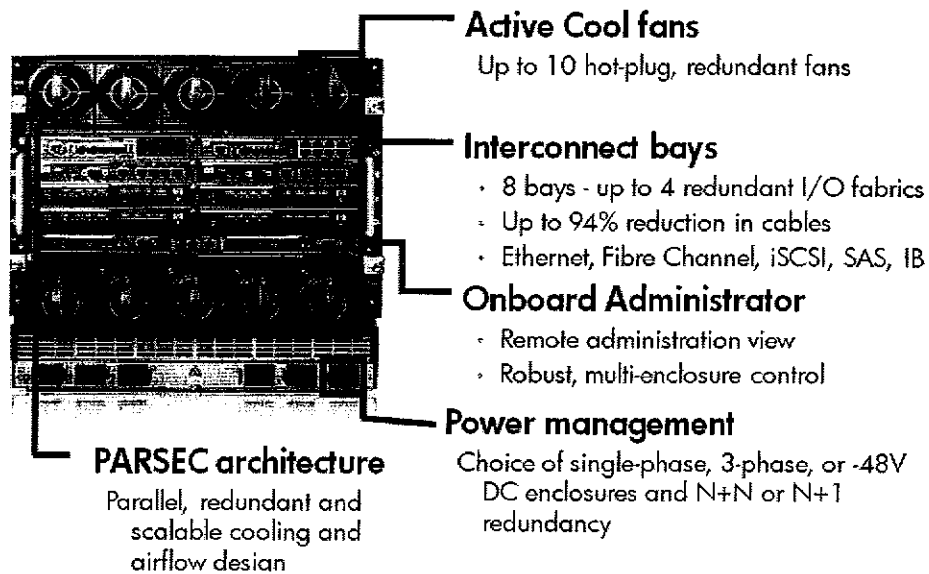


Figure III-5: HP BladeSystem c7000 Enclosure – rear view [10]

As the above image shows, the HP BladeEnclosure has up to 10 cooling fans which keep all the 16 servers cooled along with the interconnect switches that connect the embedded servers through ether channel (see Figure III-5). The enclosure itself is divided into 4 cooling zones where each zone consists of $\frac{1}{4}$ of the hosted servers.

The onboard administrator is a hardware component inside the enclosure which administers the enclosure chassis and all connected servers and components. One can access the onboard administrator of the chassis via a webpage from which we literally have 100% granular access to all servers. Using a technology created by HP the iLO2 (Integrated Lights-Out 2), the systems administrator can remotely access each server's BIOS during setup and boot/turn off the power on any server. Moreover one can mount a CD/DVD drive remotely to the server's hardware. In addition, through the onboard administrator the user can check the chassis's health state:

- Temperature,
- Cooling fans functions,
- Power usage by server,
- Each server status,

- Access to the interconnect Ethernet switch inside the enclosure, and
- Power supplies status and usage.

The key advantage or main feature that the BladeEnclosure provides is consolidation. The consolidation happens at the level of:

- Servers,
- Network cabling,
- SAN cabling,
- Power cabling,
- Hardware maintenance, and
- Server room space in terms of “U” where each HP c-Class BladeEnclosure can hosts up to 16 half-height servers.

III.2.6.c - Power and Cooling

HP uses intuitive cooling strategy which divides the blade enclosure into 4 cooling zones where 10 fans take responsibility of these 4 zones. Each zone contains up to 4 servers. The corresponding fans provide direct cooling to its zone and redundant cooling to adjacent zones (see Figure III-6).

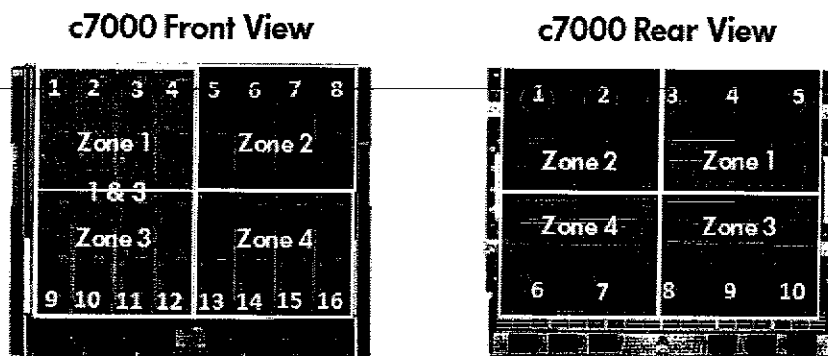


Figure III-6: Four cooling zones in the c7000 enclosure [2]

Regarding the power scheme that HP c-Class blade system uses is a pooled power. Also the power energy is grouped into 1 pool where all of the 6 PSUs (power supply units) provide energy to this pool of power. Dynamic Power Saver tries to operate

less PSU with high use rate to efficiently use the energy provided by the 6 power supplies. For example, 2 PSU running at 80% rate and providing energy to 7 servers is much better and more power saving efficient than running all 6 PSU each at 15 % (see Figure III-7).

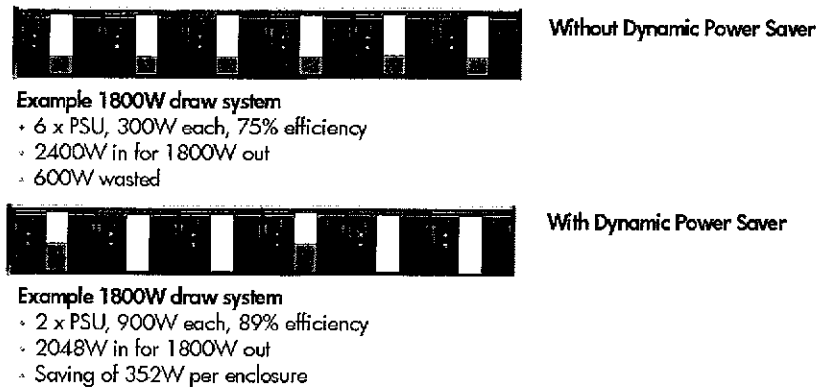


Figure III-7: Example of power efficiency with and without Dynamic Power Saver [1]

III.2.6.d - Managing the c7000 HP BladeEnclosure

The key feature that the HP c7000 BladeEnclosure provides is ease of manageability. Through the OnBoard Administrator and iLO2 (Integrated Light-Out) technology, accessing the enclosure and servers and managing them is very much easy where you only need a web browser. The OnBoard administrator is a hardware component that sits inside the blade enclosure having separate processor and separate LAN connectivity from the servers. The OnBoard administrator will be accessible even if all servers are down. Moreover, HP includes a redundant secondary OnBoard administrator in case the primary one fails. The heart of the c-Class enclosure management is the Onboard Administrator. The Onboard Administrator module in the HP c7000 blade enclosure provides four key services for the entire enclosure: detection, identification, management, and control.

Managing a c-Class enclosure involves multiple functions:

- Detecting component insertion and removal,
- Identifying components including required connectivity,

- Managing power and cooling, and
- Controlling components including remote control and remote consoles.

III.2.6.e - LAN and Fiber Connectivity for the c7000 HP BladeEnclosure

The c7000 BladeEnclosure connects the residing servers to the company's Local Area Connection and to the SAN storage via interconnect modules. These modules are directly plugged into the backplane of each blade server and provide connectivity to the external world.

Regarding the LAN connectivity the customer has 2 options: either to use GbE2c Ethernet switches where the servers are connected through them and use ether UP-Links to the company's corporate switch or to use pass-through modules where each server's NIC is directly connected to the core switch server [7].

The SAN-storage to blade servers connectivity can be done using either Fiber or iSCSI protocol. The defacto standard protocol that many companies are using to connect their blade servers to their storage network is the Fiber Channel protocol. Some companies are using the new emerging and promising iSCSI protocol. According to studies has been made by our technical team at Dar Al-Handasah, the information technology department found that the robust and most rigid protocol to use for SAN-Blades connectivity is the Fiber protocol for many reasons:

- Reliability,
- Maturity,
- Speed: up to 4Gbps compared to current 1Gbps Ethernet,
- Flexibility,
- Robustness, and
- Experience.

III.2.6.f - HP c-Class 7000 BladeSystem at WORK

In this section, the project will demonstrate how the enclosure works and how it can be accessed and managed via the iLO2 technology where nearly every management task of the server can be done remotely.

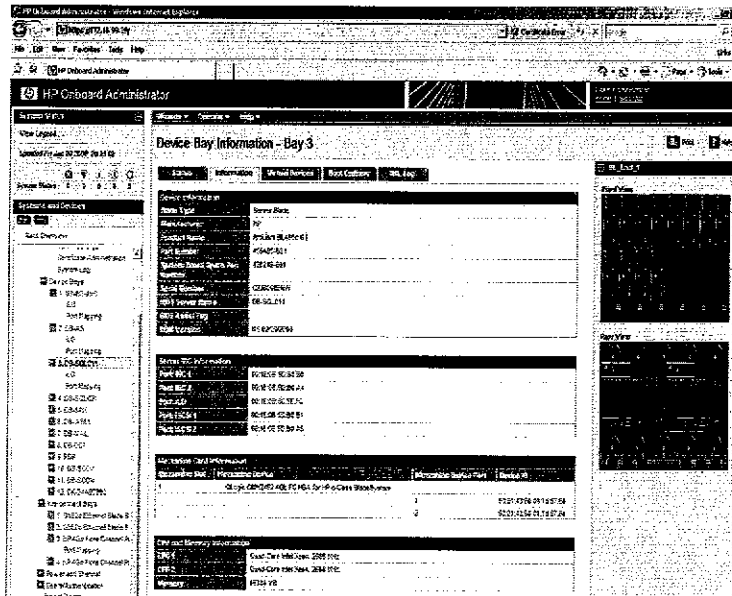


Figure III-8: Home page for iLO2 Interface BladeEnclosure

Notice that the above picture graphically shows the servers residing in our BladeEnclosure and in which port (see Figure III-8). It also shows both a rear and front view for the enclosure which includes: fans, power supplies, interconnect modules, and server.

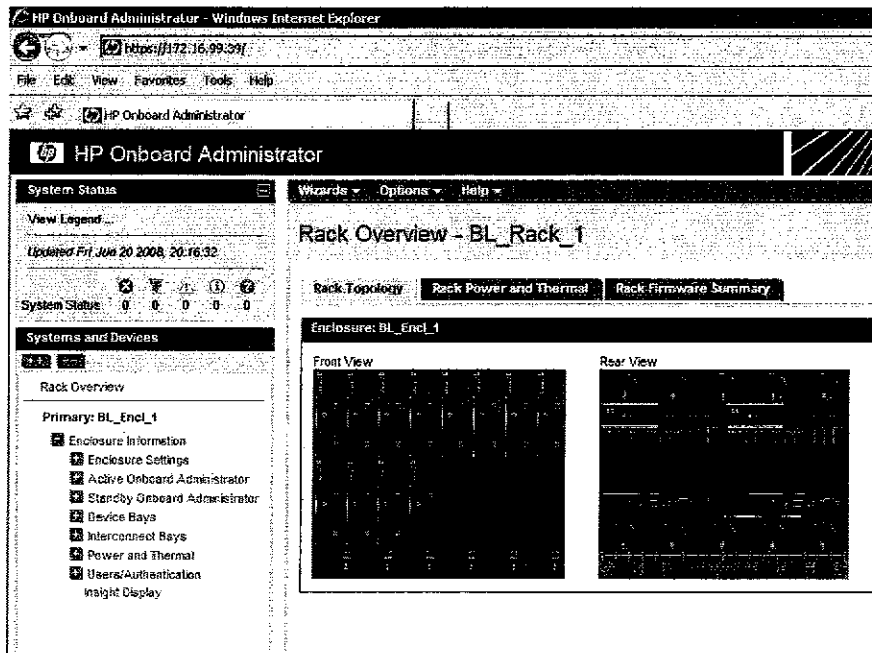


Figure III-9: iLO2 Interface BladeEnclosure

Figure III-9 shows detailed information about the server in bay 3 from which the network administrator can press the power button or mount a virtual CD/DVD ROM.

CHAPTER IV: Storage Area Network and Virtualization: HP-EVA SAN

IV.1 - SAN Storage Overview

“With data doubling in size every year, companies are looking for new ways to manage their growth and control costs. A simple historical glance on the storage technologies shows us that our most lovely and legacy rack-mounted servers where saving their data on a Directly-Attached-Storage (DAS) devices” [3].

The DAS devices are still used till now, but the high demand for storage is pushing companies to find better solutions. Companies and academic labs are becoming literary data intensive and this returns to the fact that cutting edge software applications are demanding more and more storage to store their large data files. For example, HDTV technology or modern CGI animation systems needs to produce huge data files and consumes humongous amounts of storage in tera bytes and even peta bytes.

In addition to that, enterprise-level companies demand not only large volumes of storage but also to be highly available, scalable, reliable and redundant. Due to all of the above reasons, storage inventors and technology specialists started to think how they could meet these expectations and needs. Storage area networking was the solution- to take out storage away from the server hardware itself and consolidate all storage disks in one entity or enclosure which is centralization [11], [6].

The simplest way for understanding a storage area network is to compare it to an already popular type of information technology infrastructure solution—the local area network (LAN), which enables multiple PCs to share key information technology resources such as applications, servers, shared files and printers.

SANs provide similar resource sharing, but are specifically designed for servers to share storage devices such as disk arrays or tape libraries.

The majority of SANs deployed today are built on the fiber-channel (FC) protocol, which offers high levels of performance and availability for demanding users. However, there are currently available IP SAN solutions based on the iSCSI protocol, which enable smaller information technology environments to benefit from shared network storage at a particularly favorable price.

Whether you choose fiber channel or iSCSI SAN, moving to a storage area network and phasing out your DAS storage system will help you to store and manage growing amounts of data more efficiently, while simultaneously cutting operational and management costs.

IV.2 - Benefits of SAN Storage

Benefits of the storage area network include:

- Online scalability,
- High levels of availability,
- Simple and centralized management,

- High utilization of disk capacity, and
- Faster data restoration [3], [9], [6].

IV.3 - Dar Al-Handasah HP SAN Solution

Dar Al-Handasah is adopting the HP storage area network solution: **HP StorageWorks 6100 Enterprise Virtual Array** (see Figure IV-1). The HP EVA6100 offers high performance; high capacity and high availability "virtual" array storage solutions.

The project will dig deep inside this storage master and see how it controls its disks, write data, configure vRAID technology and vDisks, implements the hot spare disk technology. In addition, the work will discuss how this SAN solution implements redundancy at all its active parts: disk controllers, shelves, disks, connectivity with blade servers, and connectivity with FC SAN switches.

Both entities, the SAN storage and the blade servers, need to communicate with each other over a highly available, redundant and high performance links. The SAN storage (disk controllers) and the blade servers are connected to each other through fiber channel (FC) switches where the fiber channel protocol is used as a way of communication over fiber optic media.

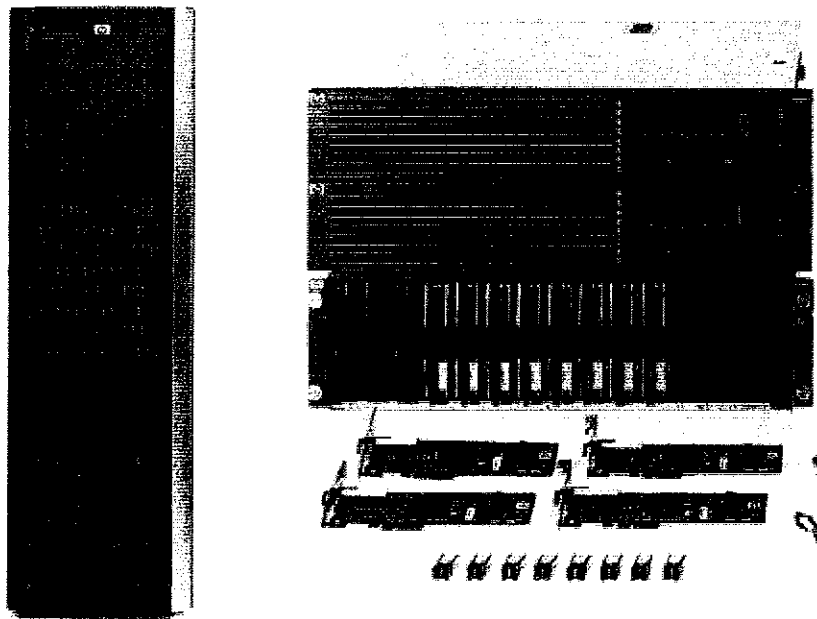


Figure IV-1: The HP Enterprise Virtual Array 6100 and Disk Controllers [2]

IV.3.1 - HP Storage Area Network Solution from HP: EVA6100

IV.3.1.a - HP EVA6100 Overview

As mentioned above, after many studies and meeting with many hardware suppliers in Lebanon and storage experts in Europe, Dar Al-Handasah technical team finally agreed to adopt the HP storage area network solution : HP StorageWorks 6100 Enterprise Virtual Array.

The key features and technical properties that the SAN storage solution form HP, the HP EVA6100, has acquired over many years are:

- Reliable and scalable storage system,
- Intelligent Disk Controllers controlling the hard drives,
- Redundant and highly available storage system,
- Fast data access to disk : 2 disk controllers, each having 2x4Gbps ports,
- Virtualized storage presentation of Disks to servers, and

- Virtualized RAID levels on the created virtual disks.

The HP EVA6100 aligned with HP StorageWorks Command View EVA aggregates and automates storage management tasks to easily administer and monitor storage when capacity changes with just a few clicks.

IV.3.1.b - HP EVA6100 Architecture

The new EVA 4100/6100/8100 arrays support:

- Industry leading automatic LUN growth and shrink with Dynamic Capacity Manager,
- Up to 240TB of connectivity with 450GB High Performance and 1TB FATA drives,
- Configuration and proactively identifies problems to prevent reactive service requests or system downtime,
- Heterogeneous Fiber Channel SAN support for Apple Mac OS X,
- Scalable Enterprise class storage support from 1TB to 240TB for Mac OS X, and
- Local replication support [3], [5].

IV.3.1.c - HP EVA6100 SAN Specifications:

The new EVA 6100 has the following technical specifications:

- RAID levels supported : vRAID0,vRAID1,vRAID5,vRAID10,vRAID50,
- Number of Disk Controllers: 2,
- Cache: 4GB,
- Host Connectivity: Fibre Channel,iSCSI,
- Host ports (per controller): 2,

- Port speed: 4 Gbps (Fibre),
- Drives per 1 shelve: 14 disks,
- Number of shelves : 8,
- Maximum number of Disks per EVA6100-SAN Storage: 112,
- Drive types: High Performance Fibre Channel and Fibre Attached Technology Adapted (FATA),
- Maximum storage capacity: 112 TB,
- I/O Requests per Second (IOPs): 154,000, and
- Maximum Throughput: 1,380 MB/s.

IV.3.1.d - HP EVA6100 Redundancy

Focusing on redundancy and high availability features, HP assures that its SAN-storage EVA family is fully redundant and has no single point of failure from blade servers up to SAN storage system due to multi-pathing in accessing the disk controllers.

The disks themselves are connected to shelves. Each shelve can hold up to 14 disks either SAS or FATA disks. All these disk shelves are connected to each other and to the dual disk controllers through arbitrated loops. For redundancy, HP configures its SAN system with 2 disk controllers for high availability of the system. Both controllers communicate with each other and control the hard disks in the SAN EVA. The disk controllers can be called the maestro of the SAN storage system. Each controller has 2 ports either Fiber or iSCSI. Each of these ports is connected to a separate Fiber Channel switch. And each blade server has 2 Fiber ports through its HBA each of these ports is connected to a separate Fiber channel switch. In this manner the blade server has Fiber multi-paths to access the SAN storage system.

IV.3.1.e - HP EVA6100 Storage Virtualization: vDisk and vRAID technologies

Storage virtualization is a concept implemented through different methods, both alone and in combination. "At the highest level, virtualization is the logical representation of physical resources as pools rather than individual devices. Servers, applications and other agents can draw from these pools what they need without regard for the origin of the resource at hand." [4].

Traditional RAID or redundant arrays of inexpensive disks was first created to add redundancy to data in case of a hardware failure of any disk in the disk array. In addition to that, RAID improved the data rate of file saved on disk arrays by striping data blocks among the disk members of the disk array. For example, if a disk array RAID level 5 is composed of 5 physical disks, the input/output of data blocks across these disks is 5 times faster than 1 disk since data is transferred in parallel mode back and forth to these disks. The HP-EVA implements RAID and extends this technology to virtualized RAID levels called vRAID 0, 1, and 5. The HP EVA stripes the data across all disks spindles in the storage enclosure evenly and equally.

Here HP creates a new technology called vDisk to represent virtual disks. The HP EVA6100 SAN storage apply virtualization on the storage system level disassociating the capacity from its underlying physical disk restrictions to create large pools of storage available to many servers [8].

In addition to the standard features that storage virtualization provides like: high availability, data protection, and fault tolerance, the HP EVA also provides:

- Optimal performance,
- Ease of management,
- Improved capacity utilization,
- Powerful data replication tools, and
- Faster restore times.

As mentioned above the HP EVA tries to virtualize the storage system by separating the physical disk architecture from the logical presentation of the disk to the server host. To create a vDisk on the HP EVA SAN storage, one can open the web GUI called EVA ComamndView and simply specify the disk volume and vRAID level: vRAID0, vRAID1, vRAID5, etc.

The virtual disk is created and presented to the server as a logical disk where the disk controller distribute the vDisk space among all physical hard drives present in the EVA. This technology is created by HP where a server can have a logical disk of size 100 GB where this logical 100 GB disk is physically and in reality scattered over a group of physical Fibre Channel disks like 50 disks. This disk virtualization gives great performance and speed to accessing the data files compared with legacy RAID technology.

Because ordinary RAID on rack-mounted servers is applied on 4 to 6 disks maximum whereas in HP EVA SAN every vDisk is created will be equally distributed over a pool of High-performance disk like 40 to 50 disks. Reading in parallel from 50 disks simultaneously is much faster than reading from 5 disks only. This is the main advantage that storage virtualization has on traditional storage system: Virtualizing the storage system and achieving better performance provides more utilization for disk usage and higher throughput.

This virtual RAID is implemented by the disk controllers that execute on the fly leveling algorithm to evenly distribute the vDisk space to all physical disk spindles across the disk arrays giving a high read/write data transfer rates to this logical vDisk.

Figure IV-2 image illustrates the performance difference between traditional RAID and the new vRAID implemented by the HP-EVA disk controllers across a pool of disks.

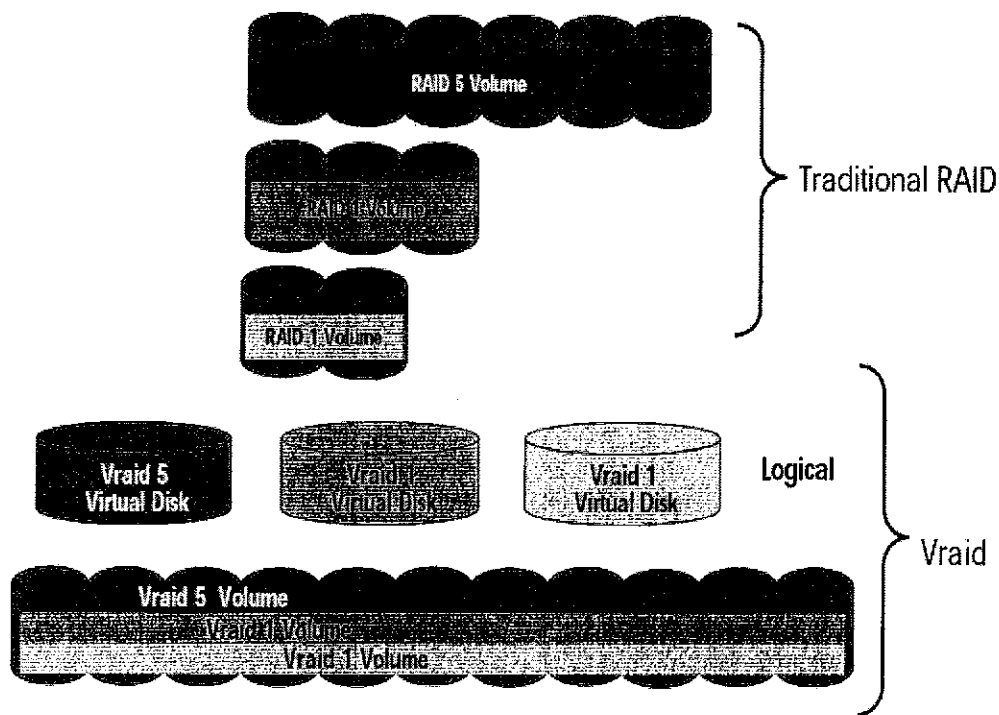


Figure IV-2: Traditional RAID compared to vRAID [4]

In addition to the improved performance that the HP EVA SAN storage provides, it also acquired the below key features:

- Simplified management: Through the use of HP StorageWorks Command View EVA,
- Dynamic Expansion: to expand or shrink vDisks capacity dynamically—without application downtime, and
- Distributed sparing: or virtual sparing of disk capacity provides faster rebuild times in the EVA, which is a benefit because it reduces the possible vulnerability of data to an additional failed drive (see Figure IV-3).

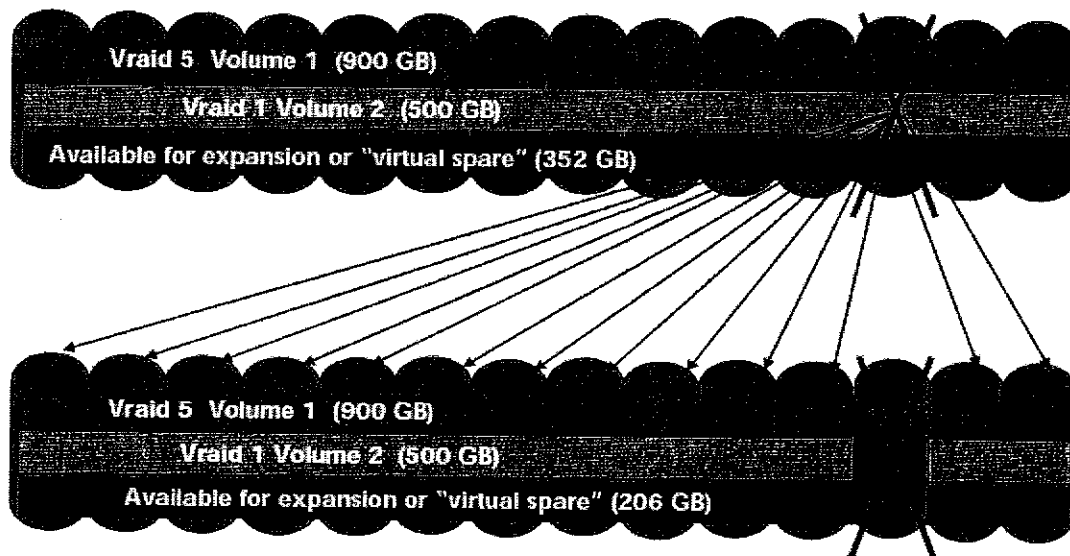


Figure IV-3: Distributed sparing or "virtual hot spare" [4]

The EVA controller provides three types of data replication tools:

- Traditional snapshot,
- Virtually Capacity-Free Snapshot (Vsnaps), and
- Virtually Instantaneous snap clone.

The most important tool is the capacity-free snapshot where only pointers to the disk blocks are backed-up. Whenever a change happens to this block that EVA takes a snapshot for it, the disk controller before writing the change it will copy-out this disk block to external location and then apply the changes.

IV.3.1.f- The Value of EVA Virtualization

Lower management and training costs:

- Easy to use, intuitive interface,
- Easy to create Vraid volumes (LUNs), and
- Unifies storage.

Better application availability:

- Dynamic LUN expansion,
- Enterprise-class availability,
- Dynamic expansion, and
- No storage reconfiguration.

Improved performance:

- Striping across all drives reduces hotspots and bottlenecks, and
- Automatic load leveling.

IV.3.1.g - HP SAN EVA6100 at WORK

In this section the report will demonstrate how the HP-EVA6100 storage system works, how the SAN EVA can create vDisks and present them to server hosts. The core software application that HP uses to control and manage the creation, deletion, and maintenance of vDisk is HP CommanView and HP BusinessCopy. Both are web-based GUI and user-friendly applications. Working with the HP SAN is easy but at the same time very critical where many details should be kept in mind throughout the creation of a vDisk.

The basic steps to create a vDisk on HP-EVA are:

1. Login to the server hosting the CommandView EVA and open the URL.
2. Click on Create a vDisk and enter the parameters:
 - a. Type
 - b. Disk group
 - c. Capacity used
 - d. Redundancy:vRaid0,1,5
 - e. Cache type: write back or through
3. Create a Host or a server to present the disk to it: the host is a blade server having 1 or 2 Fiber channel connectivity to the SAN storage. Each Fibre connection has a

WorldWideName (WWN) which is unique world-wide similar to a MAC address of an Ethernet card. The host is identified by the HP-EVA SAN by its HBA WWN names.

4. Then present the vDisk previously created to the server host and restart the server so that the server can see its new logical disk. In addition to that, the CommandView EVA can control the disk controllers as well as the disk shelves/enclosures and all its physical hard drives as the below figure shows.

CHAPTER V: DAR AL-HANDASAH HP Blade Servers and SAN Setup

V.1 - Overview

After discussing thoroughly the aspects and providing a detailed roadmap for implementing virtualization and consolidation, the report will introduce and explain briefly the DAR implementation of such a system to be a case study for future reference.

V.2 - System Architecture (with brief configuration)

As discussed above, Dar Al-Handasah decided to acquire for its new data center at Beirut the HP BladeSystem solution which consists of:

- HP c7000 Blade Enclosure,
- HP blade servers BL460c to populate the BladeEnclosure,
- HP StorageWorks EVA6100 SAN storage system with two disk controller HSV (Hierarchical storage virtualization),
- FATA and SAS Drives to populate the EVA SAN,
- 2 redundant x Fibre Channel Switches to connect the blade servers to the EVA-SAN,
- HP tape backup Library to perform backup jobs, and
- On each server an HP-Multipath I/O software is installed so that the server host can access the same SAN disk through multi-paths across the SAN switches increasing redundancy.

V.3 - Bill of Quantities for the Solution

Table V-1 below shows bill of quantities for the new data center solution for Dar Al-Handasah in Beirut branch.

Table V-1: Bill of Quantities

No.	Item	Description	Quantity
1	Server Type A *	Fully configured as per specs	4
2	Server Type B *	Fully configured as per specs	6
3	Server Type C *	Fully configured as per specs	14
	BladeEnclosure	Fully configured, including LAN and SAN connectivity components, Power Supplies, Cooling FANs, etc...	TBA
5	SAN Storage Array	Without Hard Drives	1
6	300GB FC Disk	15K rpm	12
7	500GB FATA/SATA Disk	7.2K rpm	32
8	SAN Switches	Fully configured, including power supplies, supervisor modules, etc...	2
9	SAN Cabling	Full SAN Cabling	TBA
10	Backup Solution	With LTO-3 drives	1
11	LTO-3 Tape		10
12	Design Review		2 Days
13	Training	For three persons	16Hr/pers

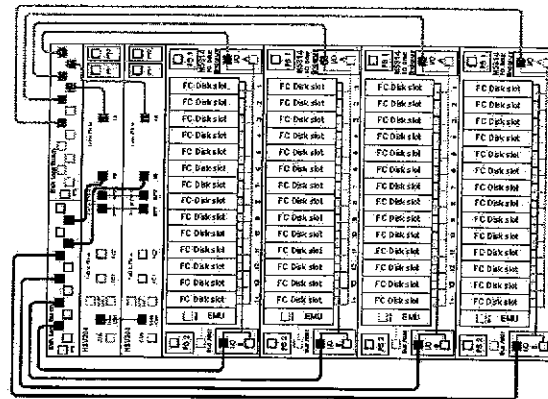
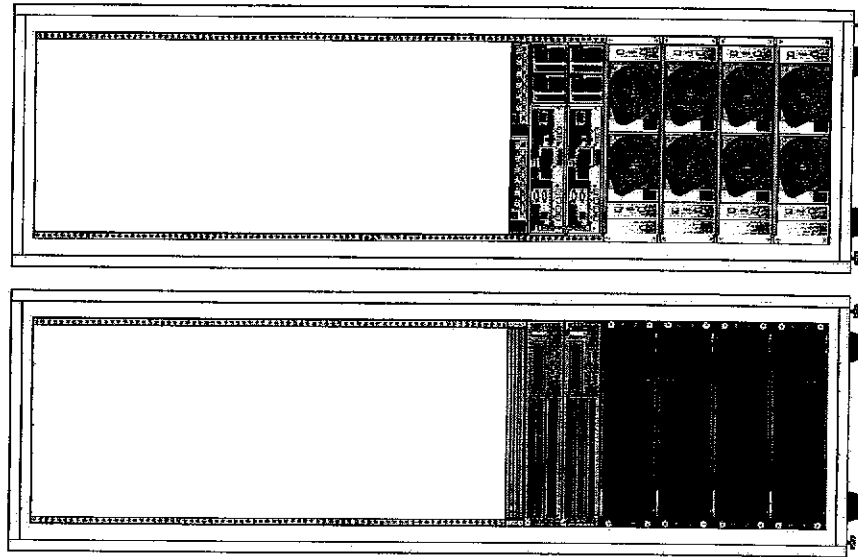
The detailed technical specifications for the HP BladeServers and EVA-Storages solution can be summed as the following:

- 2x HP c7000 BladeEnclosures, each fully loaded with PSU and FANs,
- 11xBlade BL460c servers in first enclosure and 12xBlade BL460c servers in 2nd Enclosure,
- Each blade has 2x1-Gbps Ethernet connectivity AND 1 Fibre QLogic mezzanine card with two HBA-Fiber optic connectivity to the SAN switches,

- Each blade enclosure has 2 xGbE2c Ethernet Blade Switch-16 1 Gbps ports, each server NIC is connected to one of these 2 Ethernet switches,
- All server blades are connected to the corporate LAN by up-links. These Ethernet channel uplinks connect the corporate core switch to the blade enclosures' Ethernet switches thus establishing a trunk between,
- 2xFabric brocade Switches to redundantly connect the servers Fiber HBA ports to the HP-EVA disk controllers,
- HP StorageWorks EVA6100 SAN (see Figure V-1):
 - 2 Disk Controllers HSV
 - 5 Disk Shelves each contains 12 disk drives
 - 12x300GB FC Disk-15K rpm: High performing disk space=7TB
 - 32x500GB FATA Disk-7.2K rpm: cheap, low performing disk space=14TB
 - Each disk controller having 2x4Gbps Fiber ports connectivity to the SAN Fibre switches
- Fiber Channel Zoning: each group of homogeneous servers along with the EVA-SAN fiber channel ports are grouped together in a Fiber Channel Zone. This configuration is done on both Fiber Channel Switches (see Figure V-2),
 - For example: all windows servers FC ports on the FC switches are put together in the same zone with the EVA-SAN fiber ports
 - Clustered servers AND EVA fiber ports are zoned together
 - VMware ESX server AND EVA fiber ports are zoned together
 - Tape backup library AND backup server AND EVA fiber ports are zoned together
 - Any server needs to see other server or EVA SAN storage through the Fiber Channel switch needs to be in the same zone
- Pass-through modules in the BladeEnclosure for fiber channel connectivity between each server up to the FC switches.

EVA6000 2C4D with Loop Switches

EVA6000_2C4D Model's pin	Description
AD5568	4U controller assembly with two HSY200 controllers, four M5314B 3U enclosure, a dual loop switch option, cables, and mounting hw



EVA6000_2C4D Additional components pin	Description
AD542B	The M5314B is a 3U dual-redundant FC Loop 14-bay disk enclosure with mounting hardware, and the necessary copper FC cables for connecting to an HSY200 Controller pair. NOTE: Minimum of four disk drives, any size/type, are required per disk enclosure per factory configured enclosure.
321694-B21	Cable FC Copper SFP .6m
324394-B21	Cable FC Copper SFP 2m NOTE: This cables are used with the M5314 and M5314A, but not necessary for new installations.
AG311A	HP StorageWorks 4Gb Cable Kit
221470-B21	2 Gb/s SFP Fibre Channel Transceiver Kit NOTE: Used on 2Gb switches with SFP Ports.
A7446B	4 Gb/s SFP Fibre Channel Transceiver Kit NOTE: Used on 4Gb switches with SFP Ports.
221691-B21	FC Short Wave 2-Meter Cable, LOSC (1 Gb to 2 Gb)
221691-B22	FC Short Wave 5-Meter Cable, LOSC (1 Gb to 2 Gb)
221691-B26	FC Short Wave 30-Meter Cable, LC/SC (1 Gb to 2 Gb)
221691-B27	FC Short Wave 50-Meter Cable, LC/SC (1 Gb to 2 Gb)
221692-B21	2-meter LC-LC Multi-Mode Fibre Cable
221692-B22	5-meter LC-LC Multi-Mode Fibre Cable
221692-B23	15-meter LC-LC Multi-Mode Fibre Cable
221692-B26	30-meter LC-LC Multi-Mode Fibre Cable
221692-B27	50-meter LC-LC Multi-Mode Fibre Cable

Hard Drives

- 1500 GB SATA 3.5" 7.2K RPM
- 3000 GB SATA 3.5" 7.2K RPM
- 6000 GB SATA 3.5" 7.2K RPM
- 1500 GB SAS 3.5" 10K RPM
- 3000 GB SAS 3.5" 10K RPM
- 6000 GB SAS 3.5" 10K RPM
- 1500 GB SAS 2.5" 10K RPM
- 3000 GB SAS 2.5" 10K RPM
- 6000 GB SAS 2.5" 10K RPM

Figure V-1: Physical/Logical Layouts for the HP Blade System at Dar Al-Handasah [2]

Data Center Physical/Logical Layout Summary

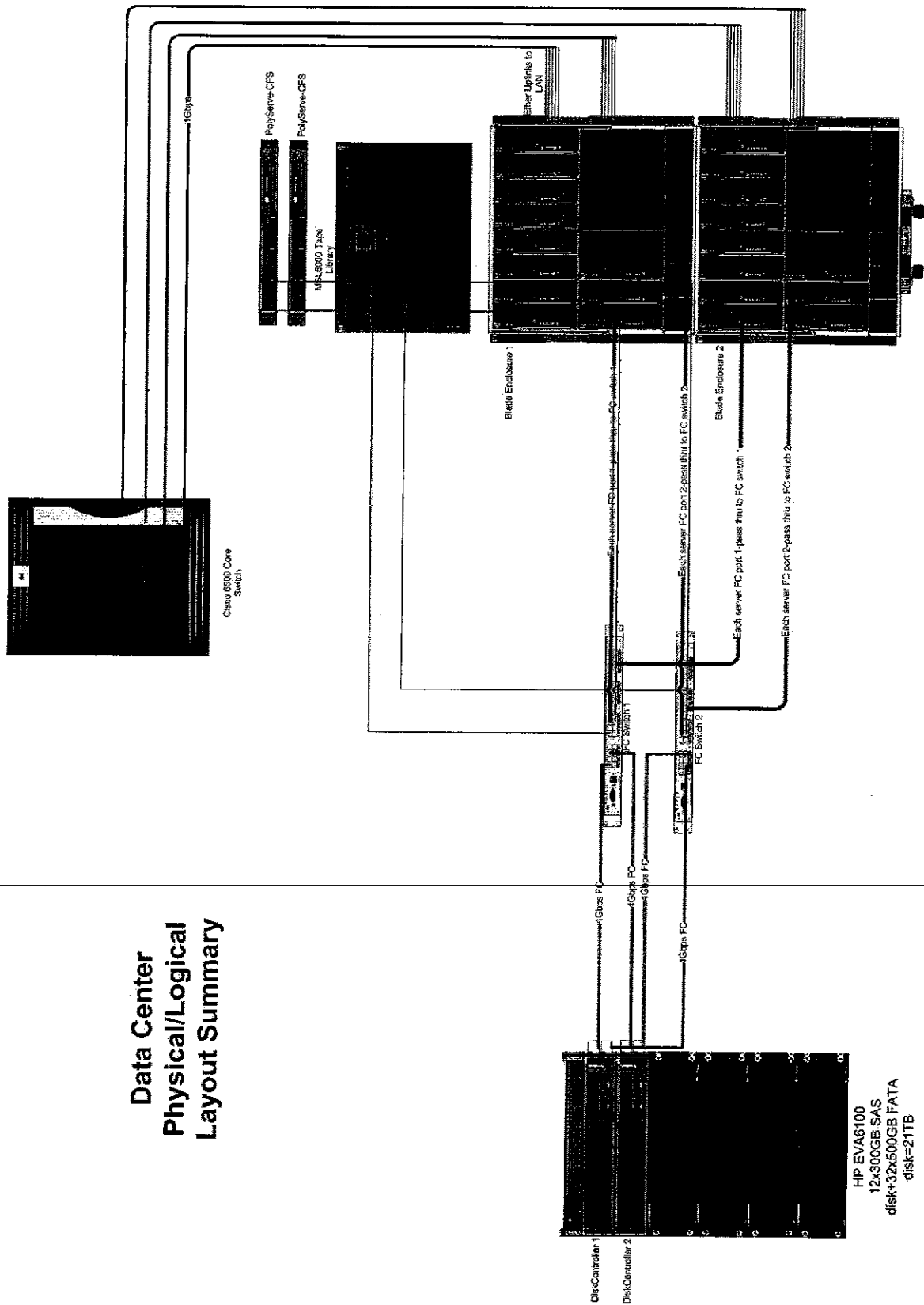


Figure V-2 : Physical/Logical Layouts for the HP Blade System at Dar Al-Handasah

CHAPTER VI: Clustered File System HP PolyServe Matrix Server

VI.1 - Overview

Dar Al-Handasah main concern is always a highly-available high-performance file system. To achieve this high availability, the solution is to adopt a clustered file system where 2+ nodes participate in serving the data files to users, if one fails the other node takes action. To achieve high performance, the nodes should exercise load balancing in a way or other so that requests are equally distributed among all nodes of the cluster, thus achieving not only high availability and redundancy but also load balancing for better performance and higher data throughput.

In this section of the project, the report will introduce a new HP product for file systems for clusters called PolyServe Matrix server. The clustered file system (CFS) is a software application which provides customer to implement and have a highly available and high performance file sharing environment. The main advantage of using this solution is that its nodes work in an Active-Active manner. This means that all file server nodes which are members of the file cluster can access the same shared disk volumes at the same time. This option gives the solution not only high availability but also loads balancing at the same time.

VI.2 - HP PolyServe Clustered File System-Architecture

The software is bundled with 2 HP DL380 nodes already configured with Windows Storage Server R2 from HP. The PolyServe clustered file system should access the EVA-SAN via multi-paths to ensure redundancy and high availability. In addition to

that, the heart-beat between the 2 node clusters is a cross link cable between the servers NICs.

The PolyServe Clustered File System can run on Windows or Linux systems. The application creates virtual host IPs and attach them to the cluster nodes so that clients can access the file system in case one of the node fails the virtual IP failover to the next node.

PolyServe Matrix Server includes the following features:

- **PolyServe Cluster File System (CFS):**

The CFS enables multiple servers to read and write shared data concurrently on shared SAN storage devices. The Matrix Server Cluster File System is designed to be scalable, highly recoverable and highly available.

- **Failover Support for Applications:**

MxS uses virtual hosts and monitors to provide highly available and scalable access to applications running on the matrix.

- **Matrix Wide Administration:**

The Matrix Server Management Console and command line interface enable you to configure and manage the entire matrix either remotely or from any server in the matrix.

The multipath I/O (MPIO) implemented in PolyServe and provides the clustered file system the following features:

- Eliminates single points of failure in the storage fabric,
- HBA, FC Switch, switch port, storage port, cables,
- Can improve performance by load balancing I/Os through the fabric,
- Windows does not natively provide MPIO, third party solutions are required,
- Windows provides a generic DSM (Device Specific Module) infrastructure that allows array vendors to write their own plug-in modules,
- Use a third party MPIO solution to deploy a matrix with maximum availability,

- Third party solutions make loss of path transparent to Matrix Server, and
- Ensure the storage is configured correctly before installing and configuring the cluster.

VI.3 - HP PolyServe Clustered File System- Configuration and Management

To manage the HP PolyServe Clustered file system, one should access or install the management utility and login with granted credentials to the server. The parameters that the clustered file system uses to implement its configuration mainly are:

- Server Fencing,
 - Membership Partitions,
 - Cluster-Wide Device Naming,
 - Matrix Server Disk Usage,
 - Cluster Volume Manager,
 - Cluster File System,
 - Quotas,
 - Multipath I/O,
 - Snapshots, and
-
- Interesting Topics.

Each of the above topics needs detailed knowledge and training so that the storage administrator can realize and configures the clustered file system storage disk. The report will not explain these topics in this project in details but will try to give a short briefing.

CHAPTER VII: Extending the Roadmap

At this point and after exploring the blade servers and virtualized storage, enterprises can build on top of this high-technological and scalable architecture any desired application, configuration or system that benefits the corporate.

This roadmap to virtualization and consolidation about how to implement and deploy BladeSystems with SAN storage can be the basis and corner stone for a new data centre. The report will give a brief overview about the server virtualization environment and its advantages.

As discussed before, virtualization can be applied on different levels: storage, server, network, etc. The report showed and presented a roadmap on how to apply storage virtualization. The report will introduce and briefly discuss Server virtualization by configuring and deploying VMware ESX3.5 Virtual Server.

Server virtualization is to show the same physical server hardware resources (CPU power, memory RAM, disk drives, NICs, etc.) as a shared resource to virtually have multiple virtual machines running on top of this same physical server.

Generally speaking, a physical machine running a specific server role will not use its CPU power all the time with full usage or even its RAM memory size. Here comes server virtualization where it takes advantage of this idle CPU and keeps it always busy serving multiple virtual machines being hosted and running on top of this powerful physical server.

One of the most advanced server virtualization products in the market is VMware ESX3.5 Server. This product runs on top of the hardware layer of any server and its Linux-based kernel. This product which runs as a middle-ware between the physical machine and the virtual machines is responsible for hosting and managing the virtual machines. A key feature of this product is treating the underlying physical hardware

(processor, memory RAM, SAN storage) as a resource pool and distributing these resources among its hosted virtual machines. In other words, if a physical server has 4 Intel Xeon Quad-core 3.0 GHz with 32 GB RAM will serve as a pooled CPU power and memory which will be shared among the hosted 20-25 virtual machines. In other words, even CPU power and memory can be virtualized and shared.

Main benefits of server virtualization are:

- 1 physical server, multiple virtual machines running,
- Cost saving on additional hardware, maximum usage of hardware,
- Fast server deployment,
- High available servers, and
- Portable servers independent from underlying hardware, easy to move.

CHAPTER VIII: Problems Addressed and Enhancements Achieved

The “consolidation and virtualization” roadmap that presented and discussed thoroughly in this project addressed and tried to solve many problems and deficiencies that the Dar Al-Handasah’s data centre was facing. After deploying this new information technology infrastructure, the new data centre at Dar Al-Handasah experienced the below major enhancements where any adoption for this roadmap will take advantage of the following:

- Easy and flexible hardware maintenance for all servers/storage systems,
- Scalable, reliable, highly available storage system,
- Fast data access to all network drives,
- Narrow backup-window,
- High available and load balancing storage Clustered File System,
- Increased CPU power,
- Virtualized Environments, and
- Centralized server management and storage system.

In addition to these enhancements that this roadmap provides, the report briefly mentions who may take benefit from this roadmap:

- Any midsize-large business corporate requesting a high available and reliable server and storage system,
- Any academic institute requesting grid architecture for High Performance Computing application running CPU intensive algorithms like: graph theory or Bioinformatics,
- A video streaming company requesting high data transfer rates with highly scalable and performance storage system, and
- A need for highly available cluster database system for critical business applications such as a banking system.

CHAPTER IX: Conclusion

In conclusion, the work tried in this project to present a detailed road map on how to implement “consolidation and virtualization” in acquiring a BladeServer solution with a scalable storage system, a Fiber Storage Area Network. The solution that discussed thoroughly and that can stand as a good case study for implementing such a solution was the new Data Center at Dar Al-Handasah. This data center is equipped with the cutting edge technology solutions of:

- Blade servers: HP c7000 BladeEnclosure + HP B1460c blade servers,
- SAN Storage: HP EVA6100 SAN,
- Brocade Fiber Channel switches, and
- HP PolyServe Clustered File System.

Any corporate that is in need of a highly available file storage system, high CPU power, scalable storage system, and easy maintenance for servers, high data transfers, and many other features can be achieved by implementing such a road map.

CHAPTER X: References

- [1] Managing the HP BladeSystem c-Class (2008). Retrieved April 21, 2008, from Hewlett Packard website: <http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00814176/c00814176.pdf>.
- [2] Clustering in a SAN Environment (2008). Retrieved June 03, 2008, from Microsoft Corporation website: <http://www.microsoft.com/windowsserversystem/storage/clustering.mspx>.
- [3] HP Blade System c-Class Insight Display (2008). Retrieved April 30, 2008, from Hewlett Packard website: <http://h71028.www7.hp.com/ERC/downloads/4AA0-5762ENA.pdf>.
- [4] Data Storage Technology and Management (2008). “Your online connection to Storage”, Retrieved May 12, 2008, from <http://searchstorage.techtarget.com>.
- [5] Fibre Channel Industry Association, “Fundamentals of Fiber Channel Storage Area Networks”, Eagle Rock, FCIA Association, c2001.
- [6] Rainer Erkens, “Storage Networks Explained: basics and application of fibre channel SAN, NAS, iSCSI and InfiniBand/Uif”, Chichester: John Wiley and Sons, Ltd., c2004.
- [7] Thomas Jepsen, “Distributed Storage networks; architecture, protocols and management”, West Sussex, England: J. Wiley and Sons, c2003.
- [8] Fred Rabert, Senior Sales Network Architect. “Storage Area Network – Evolution”, BLADE Network Technologies, Inc.
-
- [9] Dr. P.T. Ho. “Resiliency Considerations of SAN Storage for High-Availability Central Servers”. The University of Hong Kong, 2nd November, 2007, Guangzhou.
- [10] Bob Violino. “Host bus adapters (HBAs) explained”. 29 Apr 2008, <http://searchstorage.techtarget.com>.
- [11] EMC (2008). Retrieved March 16, 2008, from <http://emc.com/>.
- [12] Dell Corporation (2008). Retrieved May 6, 2008, from <http://dell.com>.
- [13] VMware (2008). Retrieved May 16, 2008, from <http://vmware.com>.
- [14] Dar Al-Handasah (2008). Retrieved June 16, 2008, from <http://www.dargroup.com>.