کتاب

Hitachi Data Store Introduction

به زبان فارسی

نویسنده:

مهندس ابوالفضل هاشمی

فهرست

۱- Hitachi Storage Platform Introduction ۳

۲- Hitachi USP V/VM Architecture ۶

۲-۱- Shared Memory ۹

۲-۲- Cache Memory ۹

۲-۳- Front-End Director (FED) ۹

۲-۳-۱- Fiber Connection (FICON) ۱۰

۲-۳-۲- Enterprise System Connection (ESCON) ۱۰

۲-۳-۳- Fiber Channel (FC) ۱۰

۲-۴- Back-End Director (BED) ۱۱

۲-۵- Data Drive ۱۳

۲-۶- Service Processor ۱۳

۲-۷- Power Supply ۱۳

۲-۸- Battery ۱۴

۲-۹- Control Panel ۱۴

۳- RAID Level ۱۶

۴- Data Drive ۱۷

۵- Device Emulation ۱۷

۶- Physical Installation ۱۸

۷- VSP Introduction ۲۰

۱- Hitachi Storage Platform Introduction

سیستم‌های ذخیره‌سازی Universal Storage Platform (USP) V، USP VM و Virtual Storage Platform (VSP) نام محصولات شرکت Hitachi است که برای مراکز داده طراحی و تولید شده‌اند. بطور مثال سیستم‌های VSP شامل سری‌های F و G هستند که می‌توان به نمونه‌های آن مانند G200، G400، G600، G800، F400، F800 و ... اشاره نمود. شرکت Hitachi در سال ۲۰۰۴ برای اولین بار از محصول USP که این محصول در سطح Enterprise و Midrange طراحی شده بود و در سال ۲۰۰۵ از Network Storage Controller رونمایی نمود. USP یکی از اولین آرایه‌های دیسک برای مجازی‌سازی کردن آرایه‌های دیسک دیگر در Appliance بجای Network بود. در سال ۲۰۰۷ نسل بعدی آنها به نام USP V جایگزین USP و USP VM جایگزین Network Storage Controller گردید. هر دو سری V و VM دارای ویژگی‌هایی همچون تحمل‌پذیری بالا در برابر خطا، عملکرد بالا و Load Balancing هستند که بر اساس طراحی سیلیکون ساخته شده‌اند. این طراحی باعث می‌شود تا پهنای باند مورد نیاز برای پشتیبانی و ادغام زیرساخت سرویس‌های ذخیره‌سازی Block Based و File را برای بستر خود ارائه نمایند. از قابلیت‌های دیگر آنها می‌توان به Thin Provisioning (استفاده بیشتر از منابع سخت‌افزاری موجود)، فعال‌سازی مجازی‌سازی برای تبدیل Hitachi SAN Storage و برندهای دیگر به یک Pool، Copy/Replication بین سیستم‌های ذخیره‌سازی در فواصل بسیار بدون نیاز به نرم‌افزار خاص، ارائه پارتیشن‌های مختلف برای میزبان‌های مختلف برروی یک یک سیستم ذخیره‌سازی، افزایش دسترسی‌پذیری، حفاظت و دسترسی داده و مدیریت منابع ذخیره‌سازی اشاره نمود. سیستم‌های ذخیره‌سازی USP V و VM یک قسمت کامل از معماری راه حل ذخیره‌سازی بر مبنای سرویس شرکت Hitachi است. یکی دیگر از قابلیت‌های این سیستم ارائه Data-At-Rest Encryption است. این قابلیت داده را با استفاده از رمزگذاری سخت‌افزاری بر مبنای AES-256 بدون کاهش توان عملیاتی بر روی بعضی و یا همه درواویوهای Open و Mainframe رمزگذاری می‌نماید. همانطور که گفته شد قابلیت Thin Provisioning در این سیستم‌ها وجود دارد که توسط نرم‌افزار پیشرفته Hitachi Dynamic Provisioning که باعث کاهش هزینه‌ها و اضافه کردن فضای ذخیره‌سازی است، می‌شود. این سیستم‌های ذخیره‌سازی پایه نیازمندی‌های برنامه‌های کاربردی کلاس‌های متفاوت است. بطور مثال یک برنامه کاربردی را می‌توان در کلاس‌های Thin provisioning، I/O load balancing، Data classification و ... قرار داد. نرم‌افزارهای شرکت Hitachi راه حل‌هایی برای ارائه قابلیت‌هایی مانند Data Mining، Database Backup/Restore، Application Testing، مدیریت منابع ذخیره‌سازی و ... ارائه می‌نماید. نمونه‌ای از این نرم‌افزارها Hitachi Storage Navigator، Hitachi Dynamic Provisioning، Hitachi TrueCopy، Hitachi

ShadowImage، Hitachi Universal Replicator، Hitachi Virtual Partition، Hitachi SNMP Agent و ... است.

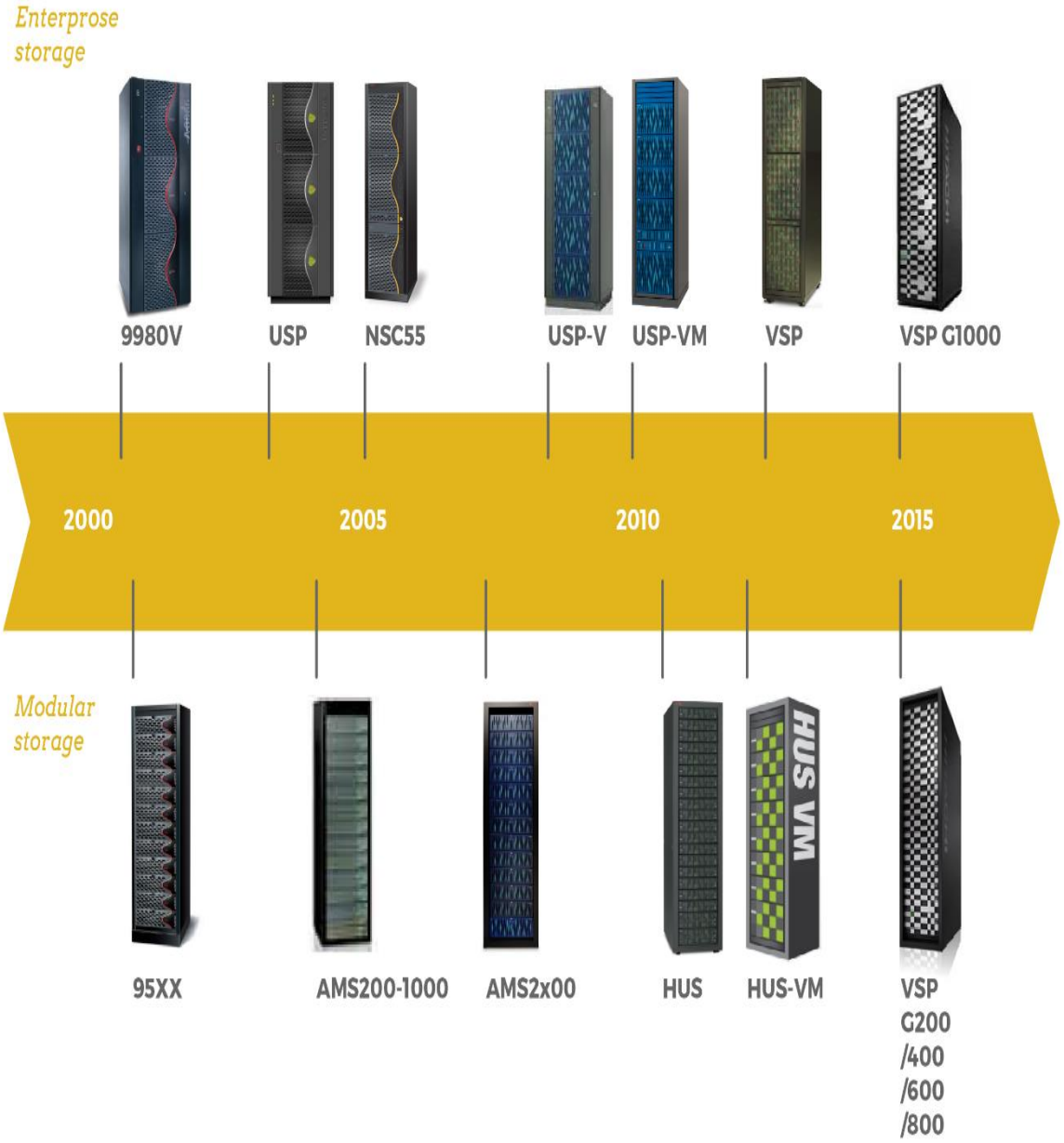
در جدول زیر مقایسه‌ای بر USP V و VM آمده است.

Specification	USP V	USP VM
Cache memory capacity	512 GB	128 GB
Shared memory capacity	32 GB	16 GB
Total storage capacity (internal and external)	248 PB	96.2 PB
Aggregate bandwidth	106 GB/sec	13.3 GB/sec
Fiber-channel ports	224	48
FICON® ports	112	24
ESCON® ports	112	24

در جدول زیر نیز قابلیت‌های USP V آمده است.

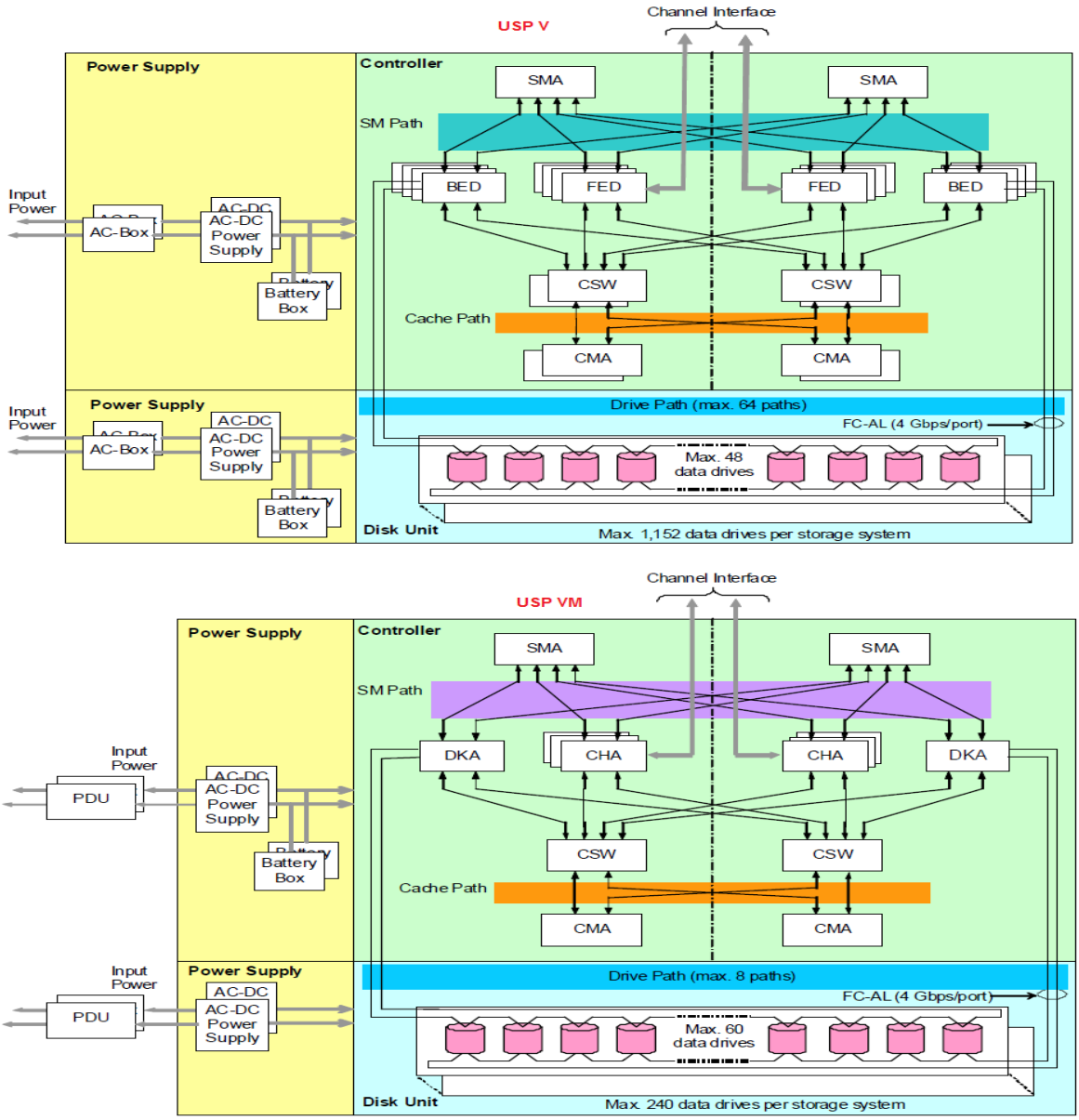
Controller	
Single-rack configuration: controller and up to two drive chassis. Optional second rack: up to two drive chassis. Rack: HDS-supplied rack or customer-supplied rack.	
Universal Star Network Crossbar Switch	
Number of switches	2
Aggregate bandwidth	13.3 GB/sec
Aggregate IOPS	1.2 million
Cache Memory	
Boards	8
Board capacity	4 GB, 8 GB, or 16 GB
Maximum	128 GB
Shared Memory	
Boards	4
Board capacity	4 GB
Maximum	16 GB
Front-End Directors (Connectivity)	
Boards	3
Fiber-channel host ports per feature	8 or 16
Fiber-channel port performance	4 or 8 Gb/sec
Maximum fiber-channel host ports	48
Virtual host ports	1,024 per physical port
Maximum FICON host ports	24
Maximum ESCON host ports	24
Logical Devices (LDEVs)—Maximum Supported	
Open systems	65,536
Mainframe	65,536
Disk Drives and Flash Drives	
Capacity (fiber channel)	73 GB, 146 GB, 300 GB, 400 GB, 450 GB, 600 GB
Capacity (SATA)	750 GB, 1 TB, 2 TB
Capacity (flash drive)	73 GB, 146 GB, 200 GB, 400 GB
Number (minimum–maximum)	0–240
Spare drives per system (min.–max.)	1–16
Internal Raw Capacity	
Minimum (73-GB drives)	0 GB (146 GB)
Maximum (2-TB drives)	472 TB
Maximum Usable Capacity—RAID-5 (7D+1P)	
Open systems (2-TB drives)	386.0 TB
Mainframe (2-TB drives)	366.4 TB

Maximum Usable Capacity—RAID-6 (6D+2P)	
Open systems (2-TB drives)	330.9 TB
Mainframe (2-TB drives)	313.4 TB
Maximum Usable Capacity—RAID-1+0 (2D+2D)	
Open systems (2-TB drives)	232.4 TB
Mainframe (2-TB drives)	220.1 TB
External Storage Support	
Maximum internal and external capacity	96.2 PB
Virtual Storage Machines	8
Back-End Director	1, standard or encrypting
Operating System Support	
Mainframe: IBM OS/390®, MVS/ESA™, MVS/XA™, VM/ESA®, VSE/ESA™, z/OS, z/OS.e, z/VM®, zVSE™, TPF; Fujitsu MSP; Red Hat Linux for IBM S/390® and zSeries®	
Open Systems: Sun Solaris, HP-UX, IBM AIX®, Microsoft® Windows, Novell NetWare, Red Hat and SuSE Linux, VMWare ESX, HP Tru64, HP OpenVMS	



۲- Hitachi USP V/VM Architecture

در تصویر زیر مؤلفه‌های سخت‌افزاری USP V/VM نشان داده شده است.

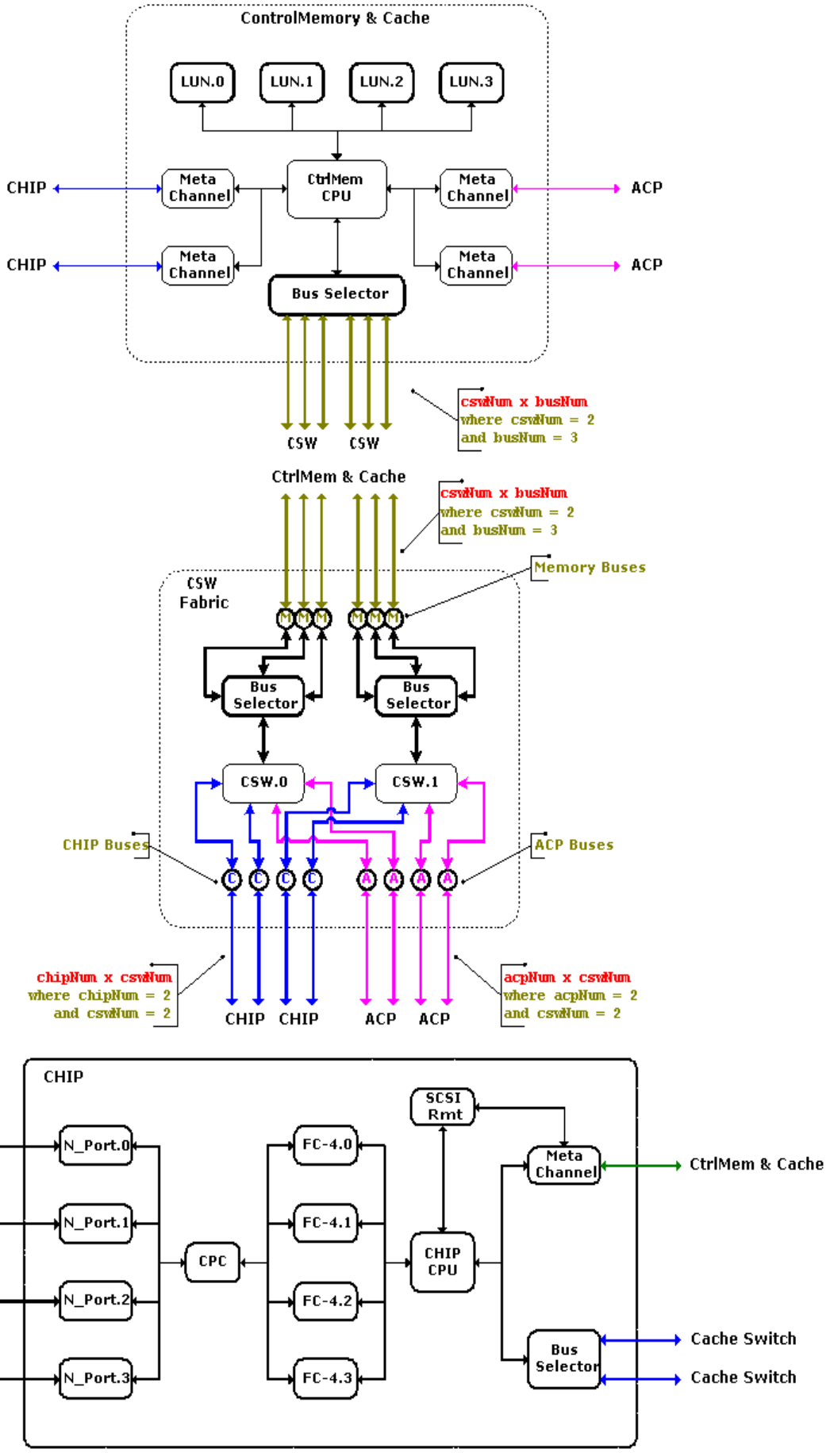


همانطور که قابل مشاهده است تفاوت این دو سیستم در Rack-Mount بودن، نوع قابلیت‌های Controller (BED، FED و ...) و ورودی برق است. در USP V مؤلفه‌های Back-End Director (BED)، Front-End Director (FED)، Shared Memory Adapter (SMA)، Cache Memory Adapter (CMA) و Cache Switch (CSW) به ترتیب در ادامه توضیح داده می‌شوند. همچنین در USP VM نیز مؤلفه‌های Channel Adapter (CHA) و Disk Adapter (DKA) وجود دارد. USP V/VM بر اساس معماری Hi-Star Crossbar Switch یا همان ماتریکس Hi-Star طراحی شده‌اند که از مسیرهای داده (Data) و فرمان (Command) نقطه به نقطه برای افزونگی و بهبود عملکرد استفاده می‌شود. به عبارت

ساده‌تر هر مسیر داده و فرمان مستقل از یکدیگر هستند و با از بین رفتن هر مسیر، مسیر دیگر استفاده می‌شود. مسیرهای اختصاصی بین هدایت کننده‌های جلویی و عقبی (Path Directors) و Cache توسط Cache Switch (CSW) با سرعت بالا هدایت می‌شوند. بنابر این مشکلاتی که در معماری Bus وجود داشته است در اینجا از بین می‌رود. به عبارت دیگر همه داده‌ای که برروی USP V/VM ذخیره می‌شوند، از مسیرهای پر سرعت با افزونگی کامل (Redundancy) به حافظه Cache انتقال می‌یابند.

هر Controller سیستم ذخیره‌سازی USP V/VM از دو Redundant Controller تشکیل شده است. به عبارت دیگر هر Controller نصف شده است که به هر نیمه Storage Cluster گفته می‌شود. هر Storage Cluster شامل همه مؤلفه‌های فیزیکی و منطقی مانند Cache، Front-End/Back-End Directors، Power Supply و ... است. هر Storage Cluster باید به هر Host یا میزبان با استفاده از مسیرهای جایگزین متصل شود تا هنگام بروز خطا در یک Cluster دیگر Cluster بتواند سرویس‌دهی نماید. هدایت کننده‌های جلویی و عقبی بین Clusterها جدا هستند تا پشتیبانی کامل را فراهم کنند. به عبارت کلی افزونگی و پشتیبانی در USP V/VM به گونه‌ای طراحی شده است که Single Point Of Failure یا تحمل شکست در یک نقطه از بین برود.

همانطور که گفته شد مؤلفه‌های کلی سخت‌افزاری USP V/VM شامل Controller، Drive Unit و Power Supply است. هر مؤلفه با توجه به تصویر به مسیرهای Cache، Shared Memory و یا Drive متصل شده‌اند. همانطور که گفته شد با رخ دادن یک خطا مشکلی برای دسترسی کاربران به داده وجود نمی‌آید. در ابتدا به معماری داخلی Controller پرداخته می‌شود که در تصویر بعد قابل ملاحظه است. مؤلفه Cache Switch با دو مؤلفه Control Memory & Cache و CHIP تعامل دارد. CHIP یک Node Module است که مانند Host از یک یا چند N_Port تشکیل شده است. Control Memory & Cache ورودی و خروجی SCSI را به ورودی و خروجی بلوک‌های LUN ارسال می‌کند (بین CHIP و LUN) و همچنین عملیات Cache را برروی بلوک‌های LUN (بین LUN و Cache) انجام می‌دهد. مؤلفه ACP شامل چندین جزء از جمله Array Group است که Array Group Drivers، Meta Channel و Bus Selector را به هم متصل می‌کند. از Meta Channel برای ارسال اطلاعات کنترلی و از Bus Selector برای ارسال اطلاعات داده به Control Memory & Cache استفاده می‌شود. در تصویر زیر می‌توان عملیات هر قسمت را مشاهده نمود.


ControlMemory & Cache
LUN.0
LUN.1
LUN.2
LUN.3
CHIP
CHIP
Meta Channel
CtrlMem CPU
Meta Channel
Meta Channel
Meta Channel
ACP
ACP
Bus Selector
CSW
CSW
cswNum x busNum
where cswNum = 2
and busNum = 3
CtrlMem & Cache
cswNum x busNum
where cswNum = 2
and busNum = 3
Memory Buses
CSW Fabric
M
Bus Selector
Bus Selector
CSW.0
CSW.1
CHIP Buses
ACP Buses
C
A
chipNum x cswNum
where chipNum = 2
and cswNum = 2
acpNum x cswNum
where acpNum = 2
and cswNum = 2
CHIP
CHIP
ACP
ACP
CHIP
SCSI Rmt
F_Port
F_Port
F_Port
F_Port
N_Port.0
N_Port.1
N_Port.2
N_Port.3
CPC
FC-4.0
FC-4.1
FC-4.2
FC-4.3
CHIP CPU
Meta Channel
CtrlMem & Cache
Bus Selector
Cache Switch
Cache Switch

۲-۱- Shared Memory

این حافظه بصورت Non-Volatile Memory (NVM) یا ماندگار است که داده آن در صورت قطع برق از بین نمی‌رود. این حافظه شامل Cache directory و اطلاعات تنظیماتی است که برروی USP V/VM استفاده می‌شود. همچنین Path Group Array یا آرایه‌های گروهی مسیر مانند انتخاب مسیر خودکار نیز برروی Shared Memory قرار دارد. حافظه Shared Memory بصورت دوتایی است که هر کدام از آنها برروی دوتای اول کارت Shared Memory قرار می‌گیرد که اینها برروی Cluster 1,2 قرار دارند. در زمان قطع برق Shared Memory به مدت حداقل ۳۶ ساعت توسط باتری پشتیبان نگهداری می‌شود. UPS V می‌تواند تا 32 GB و USP VM می‌تواند تا 16 GB پیکربندی شود. مقدار این حافظه بستگی به پارماترهای متعددی مانند تمام فضای Cache، تعداد Logical Device، عملیات Replication و ... دارد.

۲-۲- Cache Memory

مقدار این حافظه در USP V تا 512 GB و در USP VM تا 128 GB قابل پیکربندی است. تمام حافظه Cache در USP V/VM حداقل ۳۶ ساعت توسط باتری پشتیبانی نگهداری می‌شود. در سیستم‌های ذخیره‌سازی USP V/VM تمام خواندن و نوشتن داده برروی Cache صورت می‌گیرد. میزان سرعت نوشتن داده برروی Cache بصورت خودکار توسط الگوریتم‌های کنترل Cache مدیریت می‌شود تا مقدار بهینه‌ای صورت گیرد که البته به کارکرد مشخصات مشخصات خواندن و نوشتن ورودی و خروجی (IO) بستگی دارد. Cache به دو نیمه مساوی برروی دو کارت جدا که Cache A و Cache B خوانده می‌شود، تقسیم می‌شود. Cache A برروی Cluster 1 و Cache B برروی Cluster 2 قرار می‌گیرد. نوشتن داده بصورت معمول برروی هر دو Cache A , B توسط یک عملیات Write نوشته می‌شود. بنابراین از داده همیشه دوتا وجود دارد که در صورت معیوب شدن یکی، داده کپی فوراً برروی درایو ذخیره می‌شود. این طراحی Cache جامعیت کامل و همچنین مقاوم‌سازی در برابر قطعی برق را فراهم می‌کند. البته سیستم‌های Mainframe می‌توانند ویژگی‌هایی را نیز اعمال کنند. به عنوان مثال با دستور Cache Fast Write (CFW) می‌توان داده را بدون کپی نوشت.

۲-۳- Front-End Director (FED)

سیستم‌های ذخیره‌سازی USP V/VM از سیستم‌های All-Mainframe، All-Open و Multiplatform پشتیبانی می‌کنند. FED دستورات مربوط به Host را پردازش و همچنین دسترسی Host به Cache را مدیریت می‌کند. در محیط‌های Mainframe مؤلفه FED تبدیل Count Key Data to Fixed Block Architecture (CKD-to-FBA) و FBA -to- CKD را برای داده در Cache اجرا می‌کند. هر قابلیت FED (یک جفت برد و هر برد برروی یک Controller) از یکی از انواع اینترفیس‌های کانال میزبان (Host Channel Interface) ساخته شده است. هر اینترفیس می‌تواند یکی از انواع Fiber Channel (FC)، Fiber Connection (FICON) و Enterprise System Connection

(ESCON) باشد. هر اینترفیس برروی هر برد می‌تواند داده را بطور همزمان و یا مستقل ارسال نماید. قابلیت‌های FC و FICON مؤلفه FED بصورت نسخه‌های موج کوتاه (Multimode) و یا موج بلند (Single-Mode) در دسترس است. زمانی که از موج کوتاه استفاده شود USP V/VM می‌تواند تا فاصله ۵۰۰ متری و زمانی که از موج بلند استفاده شود می‌تواند تا فاصله ۱۰ کیلومتری از Host قرار گیرد.

۲-۳-۱- Fiber Connection (FICON)

مؤلفه FICON می‌تواند داده را با نرخ 4 Gbps ارسال و همچنین از ۸ پورت برای هر Controller پشتیبانی کند. لازم به ذکر است که هر برد برروی هر Controller باید برروی Controller دیگر نیز پیکربندی شود. هر ویژگی FICON برای FICON و High Performance FICON (zHPF) قابل استفاده است. FICON/zHPF یک تکنیک لایه نگاشت (Mapping Layer) بر مبنای FC و تکنولوژی multiplexing است که کمک به بهبود نرخ ارسال داده و کاهش تعداد کانال‌های ارتباطی می‌کند. همچنین از فاصله بیشتر بین USP V/VM و Mainframe پشتیبانی می‌کند. FICON و FICON/zHPF ویژگی‌های یکسانی دارند ولی FICON/zHPF دارای عملکرد بهتری است. نرخ ارسال داده در FICON خیلی به تنظیمات بستگی دارد.

۲-۳-۲- Enterprise System Connection (ESCON)

مؤلفه ExSA از سرعت 17 MB/Sec و ۸ پورت برای هر برد پشتیبانی می‌کند. هر کانال ExSA می‌تواند بصورت مستقیم به Channel Path Identifier (CHPID) یا Serial Channel Director متصل شود.

۲-۳-۳- Fiber Channel (FC)

هر FC می‌تواند تا نرخ 4 GBps و ۸ یا ۱۶ پورت برای هر برد (حتماً هر Controller باید یک برد یکسان داشته باشد) را پشتیبانی کند. هر USP V/VM از نسخه Single-Mode و Multimode برروی همان مؤلفه پشتیبانی می‌کند. از FC برای ارتباط با Mainframe نیز استفاده می‌شود و این در صورتیکه است که مسیرهای کانال FICON میزبان برروی حالت Fiber Channel Protocol (FCP) تعریف شده باشد.

در جدول زیر مشخصات FED برای USP V/VM امده است.

Parameter	**Specifications**
Number of front-end director features	USP V: 1 – 8, 14 when FEDs are installed in BED slots USP VM: 1 – 3
Simultaneous data transfers per FED pair: FICON ExSA (ESCON) Fibre-channel	 8 8 8, 16
Maximum data transfer rate: FICON ExSA (ESCON) Fibre-channel	 400 MB/sec (4 Gbps) 17 MB/sec 400 MB/sec (4 Gbps)
Physical interfaces per FED pair: FICON ExSA (ESCON)	 8 8

Fibre-channel	8, 16
Max. physical FICON interfaces per system	USP V: 112 USP VM: 24
Max. physical ExSA interfaces per system	USP V: 112 USP VM: 24
Max. physical FC interfaces per system	USP V: 224 USP VM: 48
Logical paths per FICON port	2105 emulation: 65,536 (1024 host paths × 64 CUs) 2107 emulation: 261,120 (1024 host paths x 255 CUs)
Logical paths per ExSA (ESCON) port	512 (32 host paths × 16 CUs)
Maximum FICON logical paths per system	2105 emulation: 131,072 2107 emulation: 522,240
Maximum ExSA logical paths per system	8,192
Maximum logical paths per Control Unit (CU)	2,048
Maximum LUs per fibre-channel port	2,048
Maximum LDEVs per storage system	65,280 (256 LDEVs x 255 CUs)

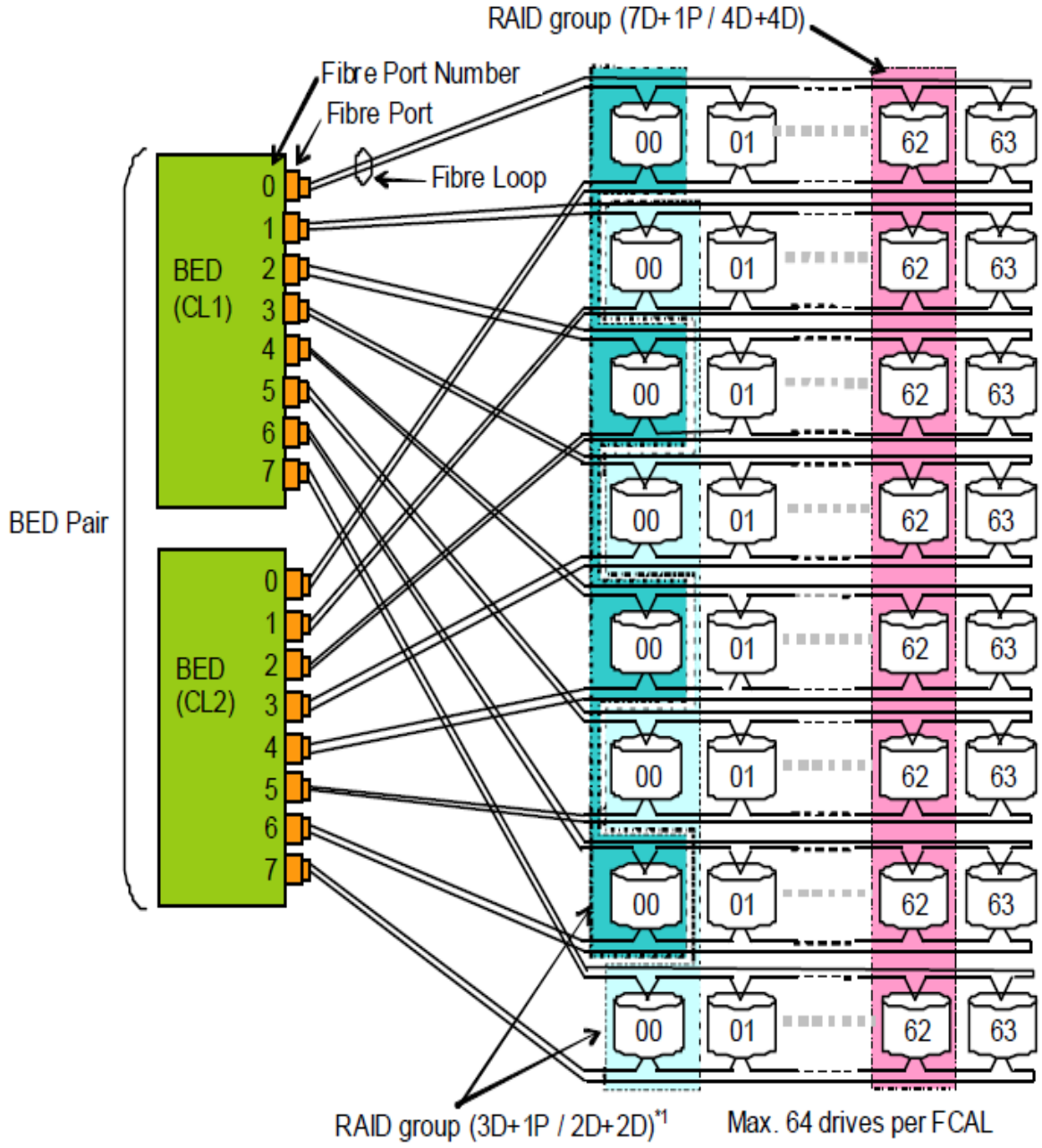
۲-۴- Back-End Director (BED)

مژلفه BED ارسال داده بین Data Drive و Cache را کنترل می‌کند. BED بصورت دوتایی برای افزونگی و عملکرد نصب شده است. سیستم USP V می‌تواند تا ۸ جفت BED را پیکربندی نماید و تا ۶۴ انتقال داده را همزمان از یا به Data Drive انجام دهد. سیستم USP VM تنها ۱ جفت BED را پیکربندی نماید و تا ۸ انتقال داده را همزمان از یا به Data Drive انجام می‌دهد. هر سیستم USP V/VM از دو نوع مؤلفه BED استاندارد (Standard) و رمزگذاری (Encrypting) پشتیبانی می‌کند. Encrypting BED (EBED) داده را برای هر دو سیستم Mainframe و Open رمزگذاری می‌نماید. Data Driveها به دوتایی‌های BED با استفاده از کابل فیبر و توپولوژی Arbitrated-Loop (FC_AL) متصل می‌شوند. هر دوتایی BED دارای ۸ مسیر فیبر Back-End مستقل است که توسط ۸ ریزپردازنده (Microprocessor) کنترل می‌شود. هر دو پورت FC مربوط به درایو به هر دوتایی برد BED از طریق مسیرهای فیزیکی متفاوت متصل شده است. به عبارت دیگر یک سوئیچ حلقه بین درواپوها و پورت‌ها است که عملکرد و افزونگی را بهبود می‌بخشد. هر دوتایی BED شامل ۸ بافر حافظه (Buffer) (برای هر مسیر فیبر یکی) است که داده را از یا به Cache انتقال می‌دهد. همانطور که گفته شد هر دو پورت درایو می‌تواند داده را برروی ۸ پورت انتقال دهد. هر دو مسیر به اشتراک گذاشته شده به هر دوتایی برد BED جدا از هم متصل شده است تا قابلیت افزونگی مسیر را فراهم نماید. همچنین هر دوتایی BED قابلیت این را دارد تا ۸ انتقال داده را همزمان از یا به Data Drive انتقال دهد. در جدول زیر قابلیت‌های BED قابل مشاهده است.

Parameter	**Specifications**
Number of back-end director features	USP V: 1 – 8 USP VM: 1
Types of back-end director features	Standard Encrypting
Back-end paths per BED feature	8
Back-end paths per storage system	USP V: 8 – 64 USP VM: 8
Back-end array interface type	Fibre-channel arbitrated loop (FC-AL)
Back-end interface transfer rate (burst rate)	400 MB/sec (4 Gbps)

Maximum concurrent back-end operations per BED feature	8
Maximum concurrent back-end operations per storage system	USP V: 64 USP VM: 8
Back-end (data) bandwidth	USP V: 68 GB/sec USP VM: 8.5 GB/sec

همه قابلیت‌ها (Function)، مسیرها (Path) و درایوهای داده (Data Drive) توسط یک دوتایی BED کنترل می‌شوند. که به آن دامنه آرایه (Array Domain) گفته می‌شود. هر دامنه آرایه شامل انواع مختلفی از پیکربندی Logical Volume Image (LVI) و یا Logical Unit (LU) است. همه انواع مختلف RAIDها اجازه دارند داخل Array Domain یا تحت دوتایی BED باشند ولی اجازه ندارند داخل Array Group باشند.



*1: A RAID group (3D+1P/2D+2D) consists of fibre port number 0, 2, 4, and 6, or 1, 3, 5 and 7.

Data Drive -۵-۲

سیستم‌های USP V/VM از HDD Drive با سرعت بالا و Flash Drive با سعت فوق‌العاده پشتیبانی می‌کند. در این درایوها از فرمت Fixed-Block Architecture (FBA) استفاده می‌شود. FBA اصطلاحی برای درایوهای HDD است که هر بلوک قابل آدرس‌دهی برروی دیسک اندازه یکسانی دارند. که از 4 Byte برای شماره بلوک استفاده می‌شود. هر Data Drive می‌تواند بدون اختلال در سیستم جانمایی شود. USP V/VM با استفادهاز تکنیک‌های تشخیص خطا و بررسی صحت اطلاعات (Data Scrubbing) خودکار، خطا را تشخیص و اصلاح می‌نماید. همچنین در صورت لازم بودن Dynamic Sparing فراخوانی می‌شود. برای هر Array Group و هر سطحی از RAID، هر درایو زاپاس(Spare) می‌تواند از هر درایو دیگر با همان سرعت، همان نرخ ارسال داده و با همان ظرفیت یا کمتر در هر سیستم ذخیره‌سازی، پشتیبانی بگیرد، حتی اگر در Array Domain متفاوت باشد. به عبارت دیگر حتی اگر در دوتایی BED متفاوتی باشد. در USP V تا ۴۰ درایو Spare و در USP VM تا ۱۶ درایو Spare می‌توان پیکربندی نمود. در تنظیمات استاندارد برای هر نوع دیسک درایو یک درایو به عنوان Spare در نظر گرفته می‌شود. همانطور که مشخص است این درایو Spare به عنوان ظرفیت سیستم ذخیره‌سازی محسوب نمی‌شود. ابزارهای تشخیص خطای دیسک، پیغامی بصورت خودکار به Hitachi Data System Support Center ارسال می‌کند و درایو Spare را وارد سیستم می‌نماید.

Service Processor -۶-۲

سیستم‌های USP V/VM شامل یک کامپیوتر ساخته شده سفارشی است که به آن Service Processor (SVP) گفته می‌شود. SVP با Controller یکی می‌شود و فقط توسط کارکنان مجاز Hitachi Data System (HDS) قابل استفاده است. VSP عامل HDS را برای تنظیمات، تعمیر و نگهداری، سرویس‌دهی و به‌روزرسانی سیستم ذخیره‌سازی فعال می‌کند. همچنین VSP قابلیت‌های Storage Navigator را فراهم و عملکرد داده را برای مؤلفه‌های کلیدی USP V/VM جمع‌آوری می‌کند تا تحلیل و تست‌های تشخیص را فعال نماید. باید توجه داشت که SVP هیچ دسترسی به داده ذخیره شده برروی USP V/VM ندارد.

Power Supply -۷-۲

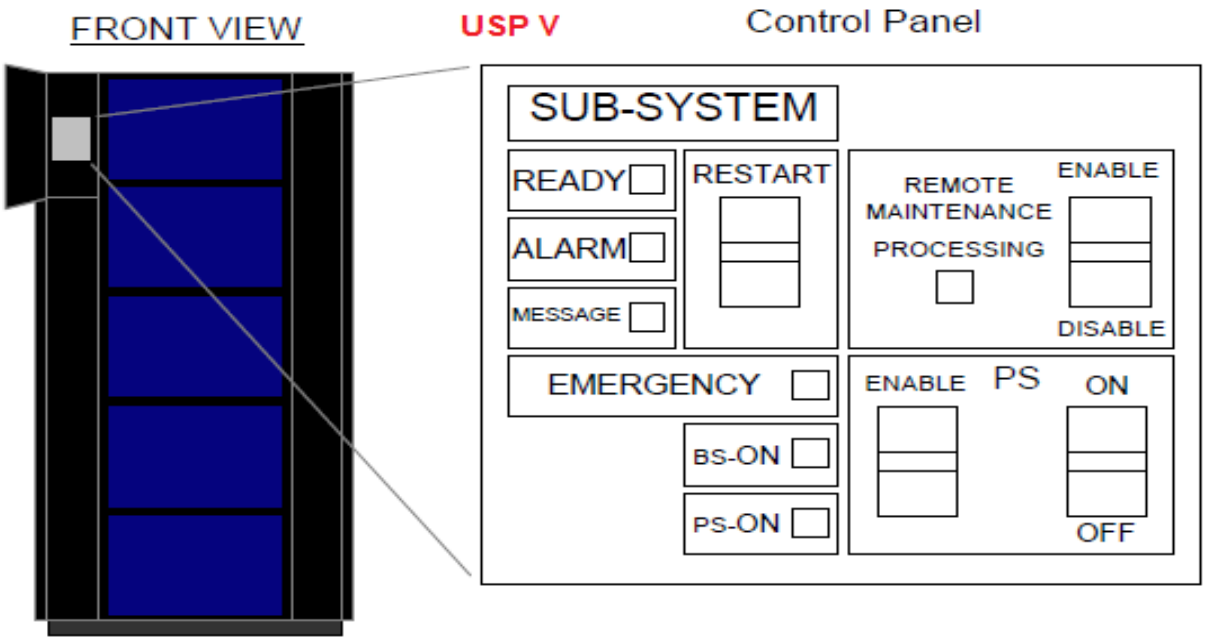
برق هر Storage Cluster توسط مؤلفه برق (Power Supply) دارای افزونگی تأمین می‌شود و هر مؤلفه برق توانایی فراهم کردن برق کل سیستم را در صورت نیاز دارد. به همین دلیل است که USP V/VM می‌تواند با از دست‌دادن چندین مؤلفه برق همچنان پایداری خود را حفظ کند. برای استفاده کامل از این قابلیت پیشنهاد اکید می‌شود که سیستم USP V/VM به دو منبع برق متفاوت متصل شود. برق ورودی USP V/VM از نوع AC است و توسط مؤلفه برق به AC-DC تبدیل می‌شود تا برق 56v/12v را فراهم نماید. و سپس هر مؤلفه سیستم ذخیره‌سازی برای خود یک تبدیل کننده DC-DC دارد تا برق مورد نیاز 56v/12v خود را فراهم کند.

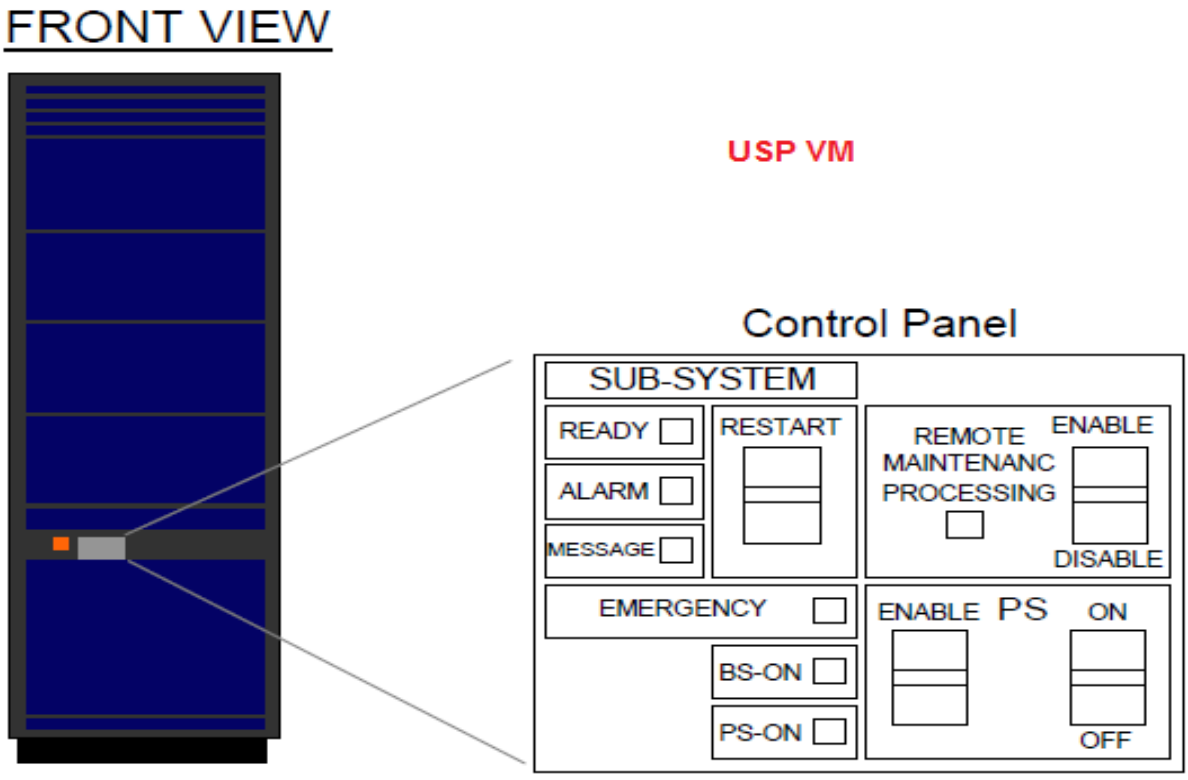
Battery ۲-۸

باتری‌های دستگاه USP V/VM از جنس باتری‌های نیکل-هیدروژن است که برق پشتیبان برای مؤلفه‌های کنترلی و عملیاتی مانند Cache Memory، Shared Memory، FED و BED را فراهم می‌کند. تنظیمات سیستم ذخیره‌سازی و شرایط عملیاتی، تعداد و نوع باتری را مشخص می‌کند.

Control Panel ۲-۹

صفحه‌ای است که در آن وضعیت سیستم نشان داده می‌شود. در تصویر زیر می‌توان مکان Control Panel را برروی USP V/VM مشاهده نمود. برای بازکردن این صفحه باید دکمه PUSH فشار داده شود.

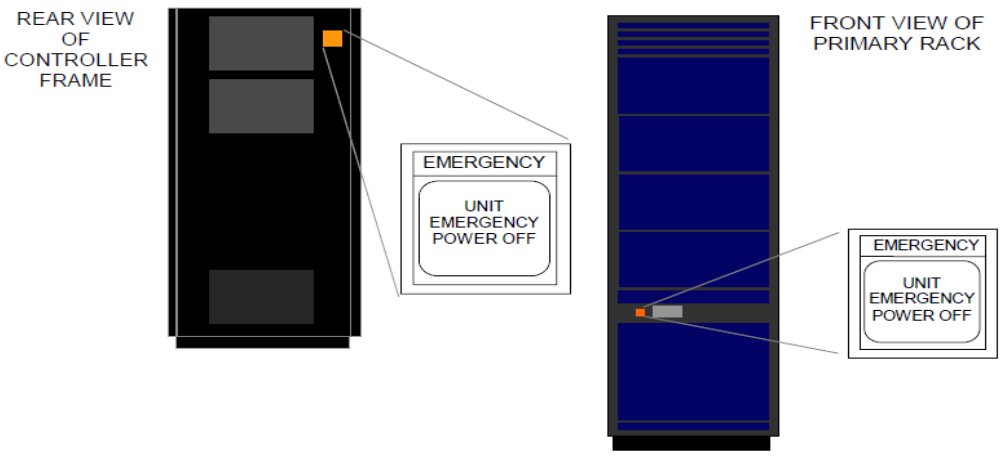




در جدول زیر توضیح داده می‌شود که هر چراغ LED و هر دکمه برروی Control Panel به چه معناست.

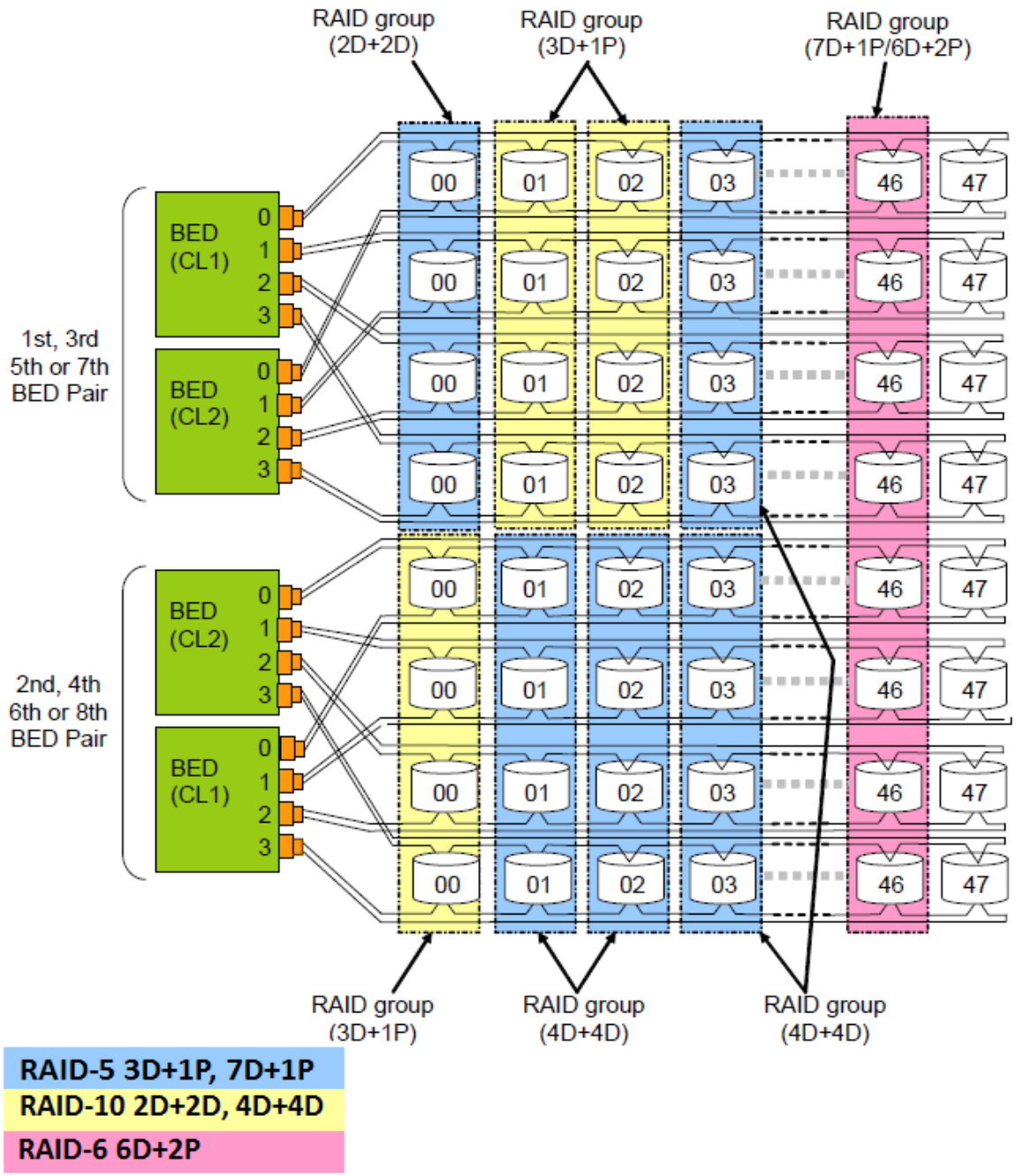
Name	Type	Description
SUBSYSTEM READY	LED (Green)	زمانی که روشن است به معنای این است که ورودی و خروجی برروی Channel Interface در حال کار است.
SUBSYSTEM ALARM	LED (Red)	زمانی که روشن است به معنای این است که ولتاژ DC پایین است، جریان DC بالا است، دما غیر عادی است و یا خطایی رخ داده است.
SUBSYSTEM MESSAGE	LED (Amber)	در صورتی که روشن باشد به معنای این است که Service Information Message (SIM) ایجاد شده است. در صورتی که چشمک زن باشد یعنی در SVP خطا رخ داده است.
SUBSYSTEM RESTART	Switch	با استفاده از این دکمه مسیر درایو حفظ شده، حفظ نمی‌شود و اجازه به اجرای دستور نوشتن نمی‌دهد.
REMOTE MAINTENANCE PROCESSING	LED (Amber)	زمانی که روشن باشد فعالیت تعمیر و نگهداری از راه دور فعال است. اگر تعمیر و نگهداری از راه دور فعال نباشد این چراغ روشن نمی‌شود.
REMOTE MAINTENANCE ENABLE/DISABLE	Switch	این دکمه برای تعمیر و نگهداری از راه دور استفاده می‌شود. وقتی تعمیر و نگهداری از راه دور فعال باشد چراغ گزینه قبل چشمک می‌زند. اگر دکمه به حالت DISABLE برود، تعمیر و نگهداری از راه دور متوقف می‌شود.
BS-ON (Basic Power Supply)	LED (Amber)	این چراغ نشان می‌دهد که ورودی برق در دسترس است.
PS-ON (Power Supply)	LED (Green)	این چراغ نشان می‌دهد که سیستم ذخیره‌سازی روشن است.
PS SW ENABLE	Switch	برای روشن نمودن مؤلفه برق باید دکمه در حالت ENABLE باشد.
PS ON / PS OFF	Switch	این دکمه برای روشن یا خاموش نمودن سیستم ذخیره‌سازی است. این دکمه زمانی معتبر است که دکمه PS REMOTE/LOCAL در حالت LOCAL باشد.
EMERGENCY	LED (Red)	این چراغ وضعیت Emergency Power-Off (EPO) برروی در پشتی دستگاه را نشان می‌دهد. زمانی که OFF باشد یعنی دکمه EPO خاموش است و زمانی که ON باشد یعنی دکمه EPO روشن است.

دکمه Emergency Power-Off برروی Controller قرار دارد و فقط در موارد اضطراری استفاده می‌شود. مکان این دکمه در تصویر زیر قابل مشاهده است.



۳- RAID Level

تکنولوژی RAID قابلیت تحمل‌پذیری بالا در برابر خطا برای درایوهای داده سیستم USP V/VM فراهم می‌کند. الگوریتم‌های مدیریت Cache، سیستم USP V/VM را قادر می‌سازد تا یک RAID Strip (تکنیکی برای تکه‌تکه کردن داده و ذخیره برروی چند درایو است) کامل از داده در Cache پیش از دسترسی فعلی، فعال نماید تا اجازه دسترسی بعدی از Cache در سرعت انتقال کانال میزبان راضی کننده باشد. سیستم‌های USP V/VM از Raidهای ۱۰، ۶ و ۵ و همچنین از Raidهای مخلوط شده (Intermixed RAID) شامل گروه آرایه‌های مخلوط شده (Intermixed Array Group) در دامنه آرایه (Array Domain) پشتیبانی می‌کند. در تصویر زیر Intermixed RAID نشان داده شده است.

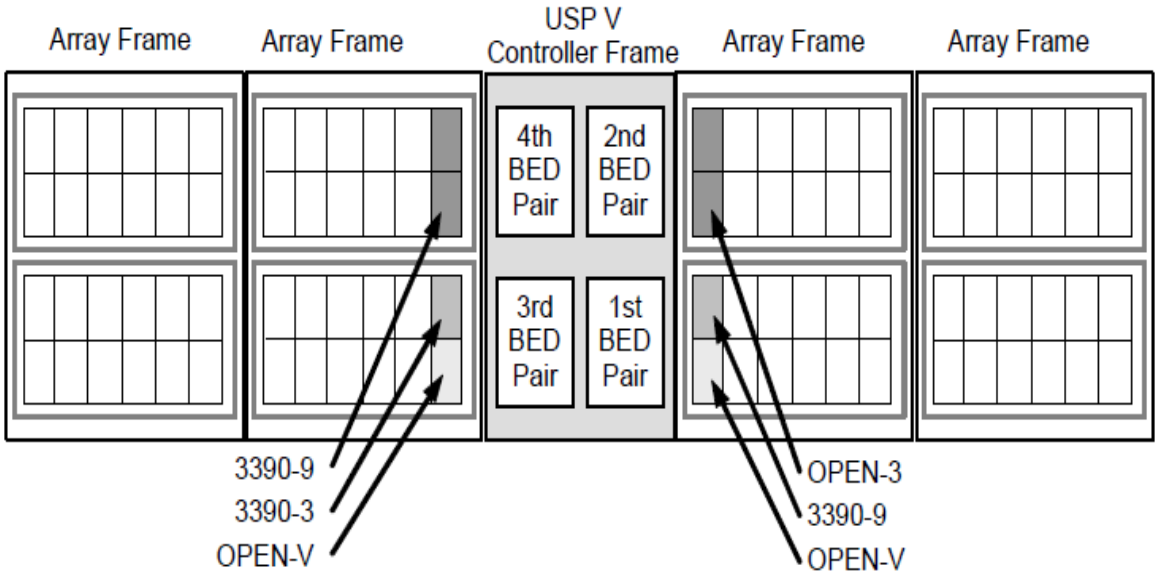


۴- Data Drive

همه درایوهای داده در یک گروه آرایه (Parity Group or Array Group) باید یک ظرفیت داشته و از یک نوع باشند. انواع مختلف درایوها می‌توانند به همان دوتایی BED متصل شوند. همه درایوهای تحت یک دوتایی BED باید در همان نرخ پروتکل Bus یا 200 MB/Sec یا 400 MB/Sec عمل کنند. بطور مثال وقتی یک گروه آرایه درایوی با نرخ ارسال 200 MB/Sec با یک گروه آرایه دیگر با نرخ ارسال 400 MB/Sec آمیخته شود، هر دو با سرعت 200 MB/Sec عمل می‌کنند. درایوهای Flash و Disk می‌توانند با یکدیگر آمیخته شوند و مکان نصب درایو Flash هیچ محدودیتی ندارد.

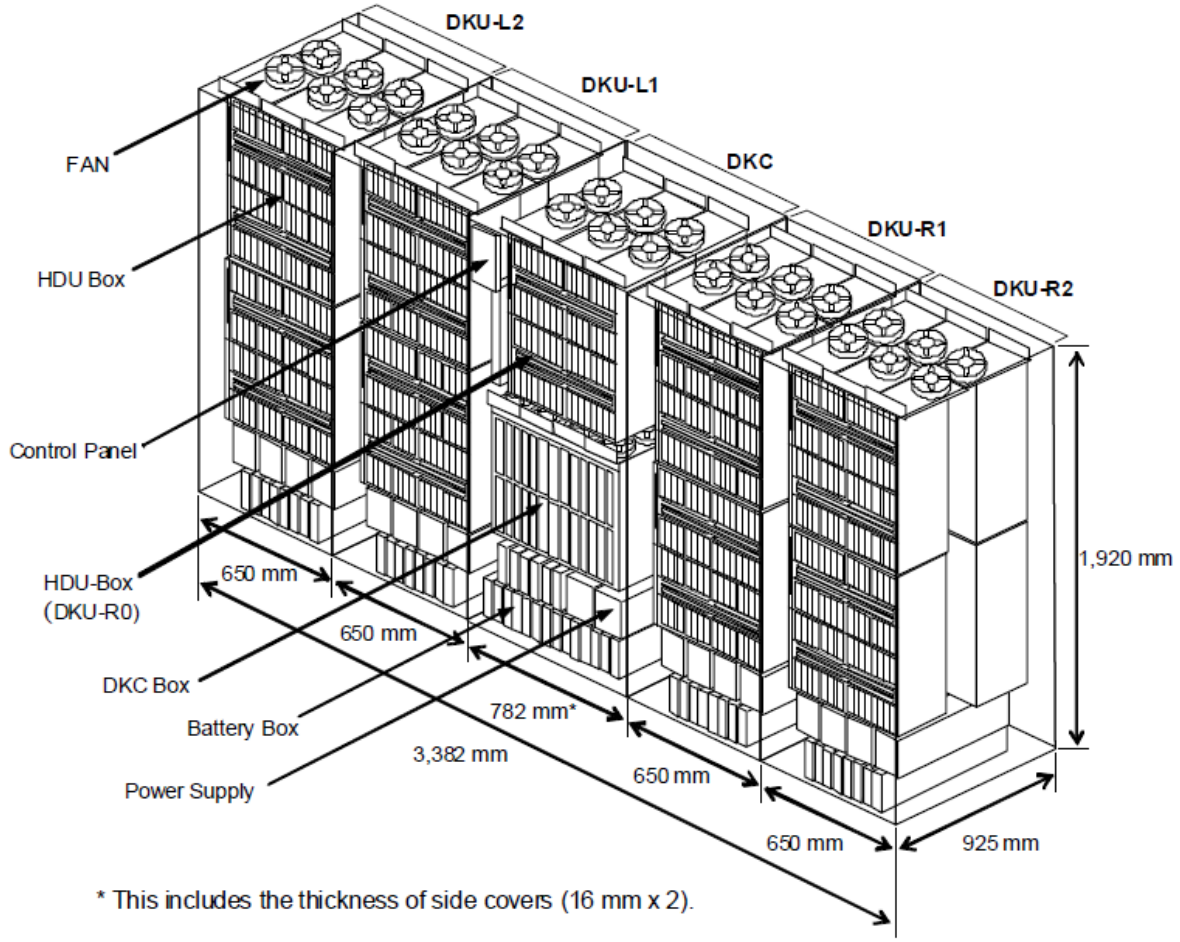
۵- Device Emulation

Device Emulation در سیستم ذخیره‌سازی بدان معنا است که در یک سیستم ذخیره‌ساز تقلید شده (Storage Device Emulation)، دستگاه ذخیره‌ساز دیسکی را به سیستم عامل به عنوان یک منبع ذخیره‌ساز واقعی ارائه می‌کند که ممکن است این دیسک از چندین دیسک تشکیل شده باشد. سیستم‌های USP V/VM از همه Device Emulationهای بهم آمیخته برروی همان دوتایی BED پشتیبانی می‌کند ولی محدودیتی که وجود دارد این است که همه دستگاه‌ها در یک گروه آرایه باید دارای شکل و قالب یکسانی داشته باشند. قابلیت مجازی Logical Volume Image (LVI) و Logical Unit Number (MUN)، انواع فضای منطقی (Logical Volume) متفاوت را فعال می‌نماید. در واقع منبع ذخیره را با یک فرمت خاصی برای میزبان فراهم می‌کند که این منبع ممکن است از چندین دیسک تشکیل شده باشد. زمانی که Virtual LVI/LUN مورد استفاده قرار نگیرد، یک گروه آرایه می‌تواند فقط یک نوع دستگاه را پیکربندی نماید. بطور مثال 3390-3 یا 3390-9 ولی هر دو قابل استفاده نیستند. زمانی که Virtual LVI/LUN مورد استفاده قرار گیرد، می‌توان انواع مختلف 3390 یا انواع مختلف OPEN-x را داشت ولی هر دوی 3390 و OPEN را نمی‌توان داشت.

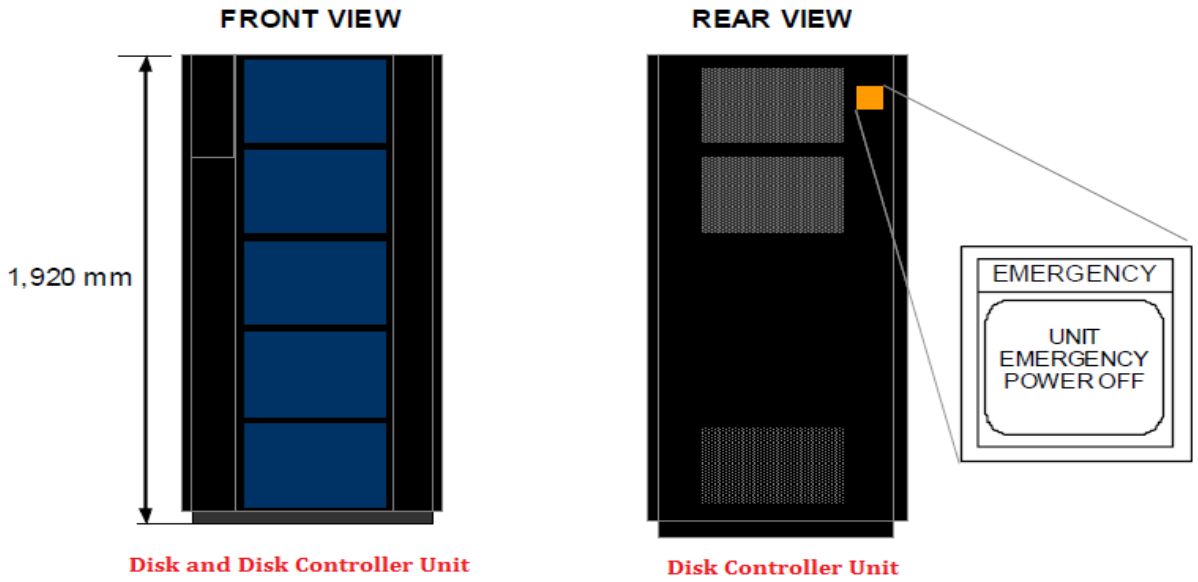


۶- Physical Installation

پیش از نصب تجهیزات USP V/VM باید شرایطی را بررسی نمود. ابتدا باید اتاق سرور یا محیطی که این سیستم در آن نصب می‌شود از نظر استحکام بررسی گردد. وزن هر Disk Controller (DKC) تقریباً حداکثر برابر 757.5 KG و وزن هر Disk Unit (DKU) تقریباً حداکثر برابر 753 KG است. بعد از بررسی وزن باید دما و رطوبت محیط بررسی شود. در وضعیت خاموش تجهیز و نگهداری در انبار دما باید بین 25°C- و 65°C+ و رطوبت باید بین 5-95% باشد. همچنین انحراف دما باید حداکثر 20°C باشد. در آخر نیز باید برق تأمینی سیستم از Uninterruptable Power Supply (UPS) گذشته و دارای افزونگی باشد. در تصویر زیر می‌توان سیستم USP V را با حداکثر پیکربندی مشاهده نمود.

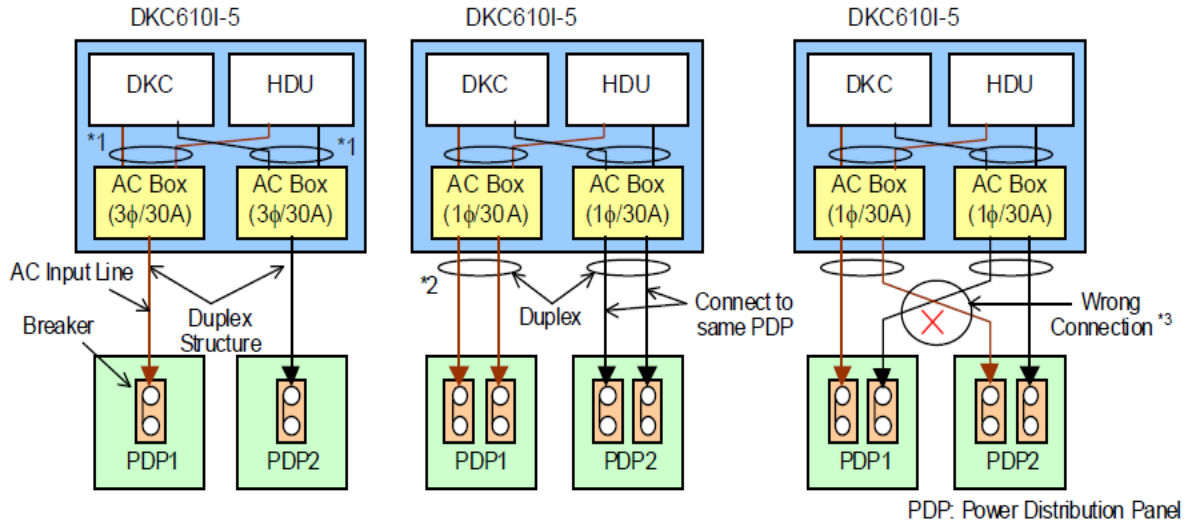


در Hard Disk Unit (HDU) Box، دیسک‌های ذخیره‌سازی قرار می‌گیرد. تنها تفاوت در ابعاد Disk Controller Unit (DKC) و Disk Unit (DKU) در عرض آنها است و در ارتفاع و عمق هیچ تفوتی وجود ندارد. در تصویر بعد این ابعاد کاملاً مشخص و قابل مشاهده است.



نحوه اتصال برق بدون UPS و بصورت مستقیم به سیستم USP V/VM در تصویر زیر نشان داده شده است.

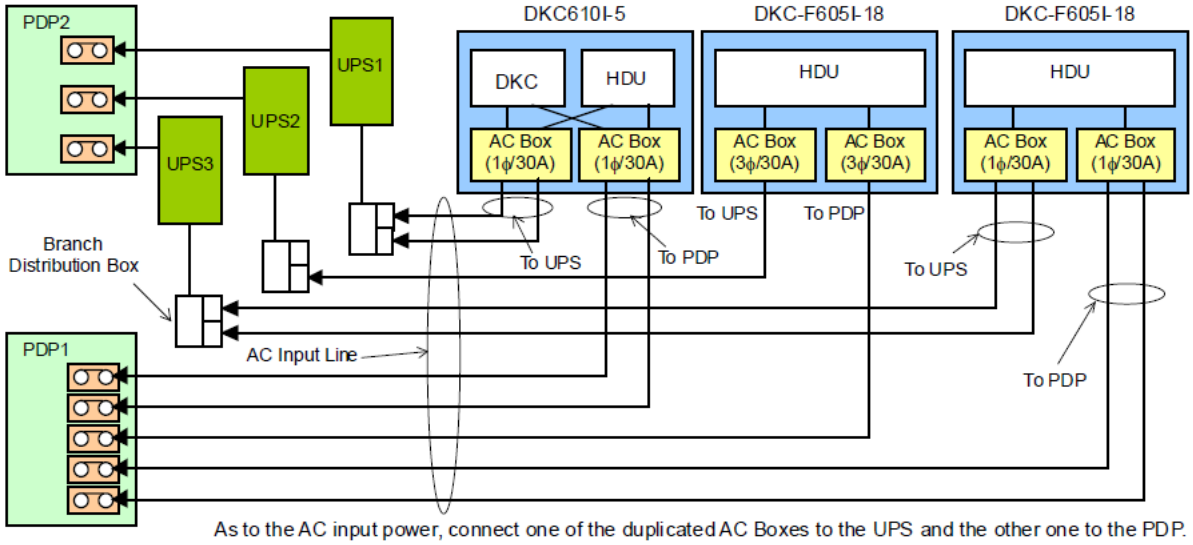
When AC input line is connected to direct Power Facility



باید توجه نمود که توزیع برق در داخل هر واحد انجام می‌شود و در بیرون نیازی به توزیع برق افزونگی نیست.

همچنین در تصویر بعدی می‌توان اتصال برق UPS و مستقیم را مشاهده نمود.

When AC input line is connected to UPS



As to the AC input power, connect one of the duplicated AC Boxes to the UPS and the other one to the PDP.

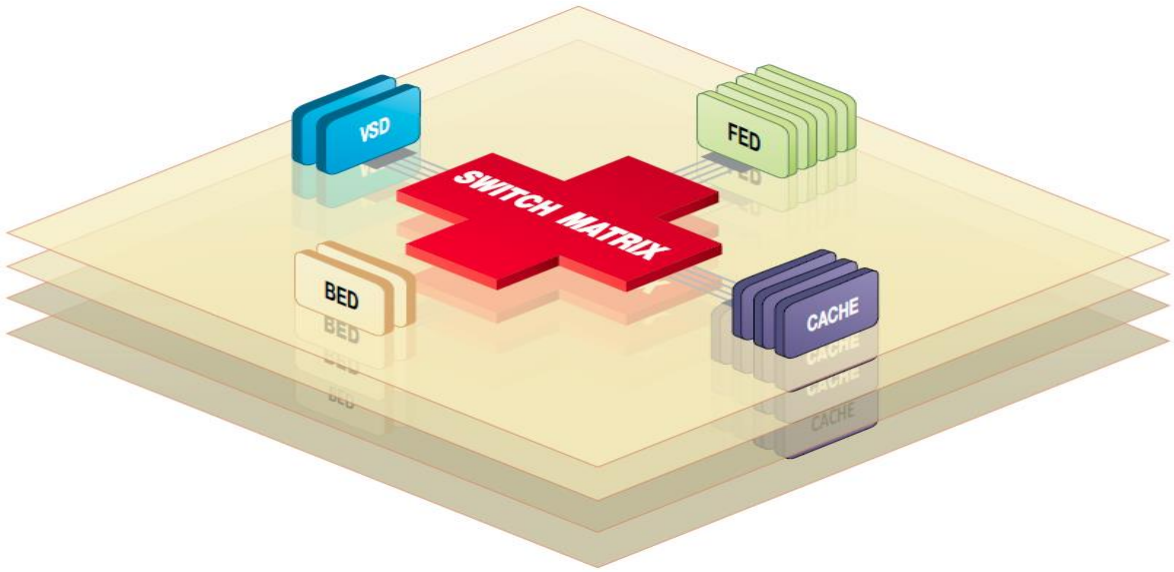
PDP: Power Distribution Panel

در صورتی که سیستم در حال کار، خاموش و یا غیر استفاده باشد باید دما، رطوبت انحراف دما و ارتفاع مورد نظر به شرح جدول زیر باشد.

Parameter	Operating		Non-Operating2		Shipping & Storage	
	Low	High	Low	High	Low	High
Temperature (°C)	16	32	-10	43	-25	60
Relative Humidity (%)	20 – 80		8 – 90		5 – 95	
Temperature Deviation (°C/hour)	10		10		20	
Altitude	-60m to 3,000m				-	

۷- VSP Introduction

معماری سری‌های Virtual Storage Platform (VSP) برخلاف تکنولوژی‌های USP و USP V ساخته شده است. در این معماری طراحی HiStar E-Network استفاده شده است که یک شبکه صلیبی-ماتریسی در پشت آن نهفته است.



ارتباط پشتی دیسک‌ها از طریق لینک 6 Gbit/s SAS به جای 4 Gbit/s FC Loop است. سیستم پردازشی داخلی Intel چند هسته‌ای است. علاوه بر اینها از دیسک‌های LFF 3.5 Inch و SFF 2.5 Inch، دیسک‌های SSD ، SAS و SATA پشتیبانی می‌شود. در طراحی VSP رشد دیسک‌ها از سه طریق صورت می‌گیرد. Scale-Up به معنای اضافه کردن ارتباط، پردازش و ظرفیت به عنوان یک واحد است. Scale-Out به معنای اضافه کردن چند واحد به یک سیستم منطقی است. Scale-Deep به معنای استفاده از مجازی‌سازی برای گسترش VSP و یا برندهای دیگر است.

در VSP قابلیت Dynamic Tiering پشتیبانی می‌شود. این قابلیت بدین صورت عمل می‌کند که برای بهینه‌سازی عملکرد، داده‌ها بصورت خودکار بین Tierهای مختلف ذخیره‌سازی جابه‌جا می‌شوند. Tiering یا لایه‌بندی سیستم‌های ذخیره‌سازی بدین شکل است که در Tier-0 دیسک‌های SSD که سرعت و عملکرد بالایی دارند ، در Tier-1 دیسک‌های HDD که سرعت و عملکرد بالایی دارند (بطور مثال 15K RPM)، در Tier-2 دیسک‌هایی که حجم بالایی دارند قرار می‌گیرند و به همین ترتیب ادامه می‌یابد تا دیسک‌هایی با حجم بالا و سرعت پایین قرار می‌گیرند.

در VSPها برای ایجاد عملکرد حداکثر از یک حافظه Cache عمومی برروی مدیر ذخیره‌سازی‌مجازی قرار می‌گیرد. همچنین بصورت خودکار Wide-Stripping از داده‌ها انجام می‌شود که این قابلیت اجازه می‌دهد تا Pool Balancing و رشد فضای Volumeها یا Shrink صورت گیرد. همچنین VSP از Thin Provisioning (استفاده بیش از منابع سخت‌افزاری)، Storage Reclamation (خالی کردن فضا در صورت حذف فایل)، رمزگذاری داده‌ها و دیگر موارد پشتیبانی می‌کند.

VSP از لحاظ سخت‌افزاری می‌تواند بین ۱ تا ۶ کابینت 19 Inch وسعت داشته باشد. حداکثر دیسک‌ها برای SFF 2.5 Inch برابر ۲۰۴۸ یا نهایتاً 1.2 PB و برای LFF 3.5 Inch برابر ۱۲۸۰ یا 2.5 PB است.

هر VSP Controller شامل Dual In-Line Cache Memory Module (DIMM)، Processor، Battery، Fan، پورت برای ارتباط ماژول‌های iSCSI و FC I/O است. هر Controller دارای یک ارتباط Ethernet برای مدیریت است و در صورت از بین رفتن هر Controller داده‌ها از Controller دیگر عبور پیدا می‌کنند. برای قابلیت اطمینان بالا، سخت‌افزارهای ضروری بصورت افزونگی تعبیه شده‌اند. اضافه کردن و یا جابه‌جا کردن هر یک از این مؤلفه‌ها هنگام به‌روزرسانی Firmware بدون هیچ گونه وقفه در ارسال داده‌ها می‌توان صورت بگیرد. هر دیسک Hot Spare می‌توان بصورت خودکار برای دیسک معیوب تنظیم و جایگذاری شود. سیستم ذخیره‌سازی VSP Block-Level توانایی ارسال داده‌ها را با استفاده از پروتکل FC و iSCSI دارد.

با توجه به انتخاب تنظیمات شاسی دستگاه VSP، می‌توان از دیسک‌های SAS، SSD و درایوهای ماژول Flash استفاده نمود. سیستم ذخیره‌سازی VSP عملکرد بالا را با استفاده از پشتیبانی دو Controller، 1024 Cache

GB، افزایش سرعت پردازش I/O با استفاده از تنظیمات Flash Drive و استفاده از پورت‌های سرعت بالای FC و iSCSI فراهم می‌کند. همچنین می‌توان با استفاده از RAIDهای ۰، ۱، ۵، ۶، ۱۰ قابلیت اطمینان ارسال داده‌های در حال ارسال به Flash Memory در هنگام قطع برق و همچنین ارائه یک سیستم افزونگی کامل را فراهم نمود.

همانطور که گفته شد هر VSP شامل دو Controller است که می‌تواند با میزبان‌ها ارتباط برقرار نماید. همانطور که گفته شد هر Controller شامل ارتباط Ethernet برای مدیریت Controller است که می‌توان با استفاده از Hitachi Device Manager-Storage Navigator آن را مدیریت نمود. همچنین هر Controller شامل چراغ‌های LED جلو و عقب است که وضعیت هر Controller را نشان می‌دهد.

در شکل زیر می‌توان حداقل و حداکثر تنظیمات را مشاهده نمود.

Characteristics	Virtual Storage Platform (minimum configuration)	Virtual Storage Platform (maximum configuration)	
Number of Frames or Racks (basic control frame or rack plus optional array frames or racks and second control frame or rack)	1	6	
Number of Grid Switches in Hi-Star-E Network	2	8	
Raw Cache Bandwidth (GB/sec)	8	128	
Raw Control Memory Bandwidth (GB/sec)	8	32	
Aggregate IOPS (60:40 mix Read:Write – 8Kb)		237,000	
Fibre Channel Port Performance	2/4/8Gb	2/4/8Gb	
Supported Logical Devices (LDEVs)		65,280	
Number of Disks per System (min./max.) (3.5")		1280	
Number of Disks per System (min./max.) (2.5")		2048	
Number of Disks per System (min./max.) (SSD)		256	
Number of Spare Disks per System (min./max.)	1	256	
Max. Internal Raw Capacity (2TB disks)		2.52PB	
Max. Internal and External Capacity (PB)		255PB	
Max. Useable Internal capacity (TB) Open Systems, RAID-5 (7D+1P)		2,192.2 (Open-V)	2,080.8 (IBM® z/OS® 3390M)
Max. Useable Internal Capacity (TB) Open Systems, RAID-6 (6D+2P)		1,879 (Open-V)	1,779.7 (z/OS 3390M)
Max. Useable Internal capacity (TB) Open Systems, RAID-1+0 (2D+2D)		1,256.6 (Open-V)	1,190.2 (z/OS 3390M)
Cache Boards			
Number of Boards	2	16	
Max. Board Capacity (GB)	32 (64)	32 (64)*	
Capacity (GB)	16	512 (1,024)*	
Control Memory			
Number of Boards	2	8	
Max. Board Capacity (GB)	4	4	
Capacity (GB)	8	32	
Front-end Director Boards	2	24	
Fibre Channel Ports	0	192	
Virtual Ports per physical port	1,024	1,024	
IBM FICON® Ports	0	176	
Fibre Channel over Ethernet (FCoE) Ports	0	(88)*	
IBM ESCON® Ports	n/a	n/a	
Max. Hitachi Virtual Storage Machines	1	32	
Back-end Directors (SAS 6Gb)	2	8	

** Cache capacity and FCoE will be added in an upcoming release.*

اجزای VSP در تصویر زیر قابل مشاهده است.

Figure 19. Control Chassis (front view)

The Control Chassis

The control chassis contains the virtual storage directors, FEDs, BEDs, data cache, grid switch boards and up to two disk chassis, as well as power supplies (see Figures 19, 20 and 21).

The disk chassis contains the hard disk drives, power supplies, battery modules, cooling fans and communication interfaces.

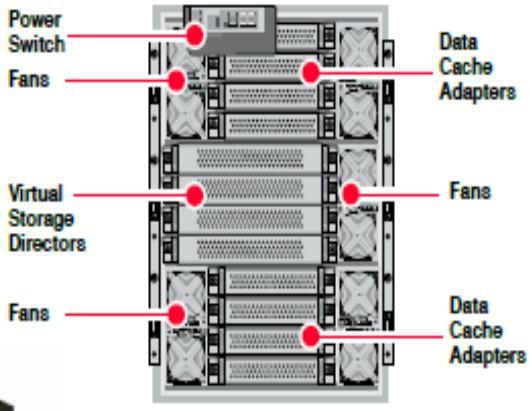


Figure 20: Control Chassis (back view)

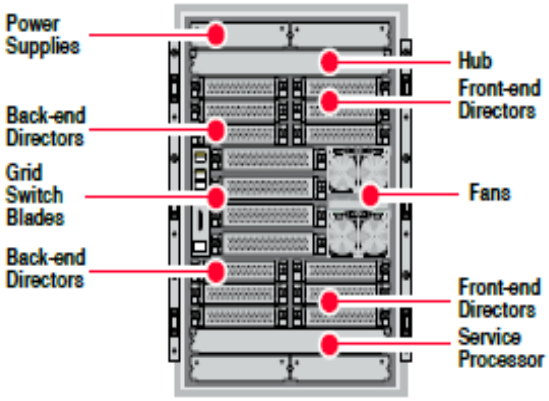


Figure 21. Virtual Storage Platform Racks Showing Control and Disk Chassis