ON-DEMAND
OPTIMIZED RESPONSIVE
EFFICIENCY
FAST CLOUD FLEXIBLE
SERVICES

PowerEdge C410x Overview

Name: Alejandro González | Fabián Salamanca

Title, Cloud Solutions and HPC Storage

Last updated: 2/8/2010

SPEEDTCO MODULAR
CUSTOMIZED
STACK HYPERSCALE
POWER

Agenda

- What customers told us

- PowerEdge C410x PCIe expansion chassis

- GPU options

- The PowerEdge C6100 with the C410x and other recommended configurations

- HPC Storage

SPEED TCO MODULAR CUSTOMIZED

DELL

Dell is committed to
Simplifying HPC
by removing cost and complexity

What Customers Told Us

Oil & gas, gaming, scientific research, financial modeling, media & entertainment

- Have to wait too long to see results

- Want to use more GPUs to increase speed

- Demand for compute continues to increase, budget is shrinking and space is the same

- When a GPU goes down, am not notified and have to bring the entire unit down

- Can't add GPUs to my existing servers

- Don't want to buy more servers when can use GPUs for less

- Can't mix PCIe devices inside of a server

- Apps pushing GPU ratios higher
  - Currently 2 GPUs to 1 server (2:1) dominates
  - Shift to 4:1 underway
  - Leading accounts architecting 8:1

- Thermal challenges
  - 190-220 watts per GPU
  - 40mm fans in 1U, or active heat sinks

- Additional requirements
  - Enhanced management capability
  - Improved GPU serviceability
  - Improved power and thermal redundancy
  - Rack density and integration

OPTIMIZED RESPONSIVE EFFICIENCY

DELL

Dell Delivers Through Key Tenets

Open Architecture

Leverage current and future investments in heterogeneous technology, people and processes, without sacrificing integration.

Heterogeneous, Future Proof

Integrated

Linking the heterogeneous components of the data center ecosystem, removing unneeded steps, licensing and consoles.

Integrating the Ecosystem

Automated

Automating entire data center ecosystem, improving consistency, control and dramatically affecting the time and resources needed.

End-to-End Automation

Which Gives Customers

Choice

- Choose your pace
- Choose your technology
- Choose to leverage existing investments in technology, people and process

Improved ROI

Flexibility

- Mix heterogeneous components
- Integrate physical and virtual environments
- Integrate with custom processes, tools and applications

Right-sized Solution, Every time

Agility

- Automating standard and repetitive tasks
- Consolidating tools and consoles
- Managing ecosystem holistically

Respond to Business Faster

PowerEdge C6100

Capacity, performance and flexibility in an eco-friendly package

CLOUD AND CLUSTER OPTIMIZED, SHARED INFRASTRUCTURE SERVER

Great for: Hyperscale-inspired building block for high-performance cluster computing (HPCC), Web 2.0 environments and cloud builders where performance is key

- Up to 4 server nodes in 2U
- 2S Intel 5520 chipset (<100W processors)
- Xeon 5500 series CPUs
- Xeon 5600 series CPUs
- 12 x DDR3 RDIMM
- 24 x 2.5" or 12 x 3.5" HDD
- 2 x GbE Intel 82576 Kawela
- 1 x 16 Gen II (low-profile/half-length)

- 1 x 8 Gen II (custom mezzanine slot)
- IPMI 2.0 management only
- iKVM, DCMI, PXE support
- Hot-swap server nodes & HDD
- Redundant 1100W power supplies
- 470W, 750W, 1400W options*
- Intel Node Manager compliant

∗ These items ship in July 2010

OPTIMIZED RESPONSIVE EFFICIENCY

Why PowerEdge C6100?

High Compute Density

- Up to 4 independent 2-socket server nodes in a 2U chassis
- Higher density than many competitive modular designs

High Storage Density

- Up to 6 x 2.5" or 3 x 3.5" HDDs per node
- Higher per server spindle count than HP and IBM 2U products

Highly Serviceable

- Each independent server node is hot-serviceable
- Hot-serviceable server nodes are not available from HP or IBM

SHARED INFRASTRUCTURE FOR MAXIMUM EFFICIENCY AND DENSITY

- **2x density of traditional 1U servers**
- **20% more dense** than blades
- **30% more storage** than comparable HP servers
- **Shared infrastructure** uses less floor space, power and cooling
- Service individual nodes as needed to **increase uptime by up to 75%**
- **92% efficient** hot-plug redundant power supplies

OPTIMIZED RESPONSIVE EFFICIENCY

DELL

PowerEdge C410x PCIe Expansion Chassis

Maximizing space, weight, energy and cost efficiency, with unprecedented flexibility

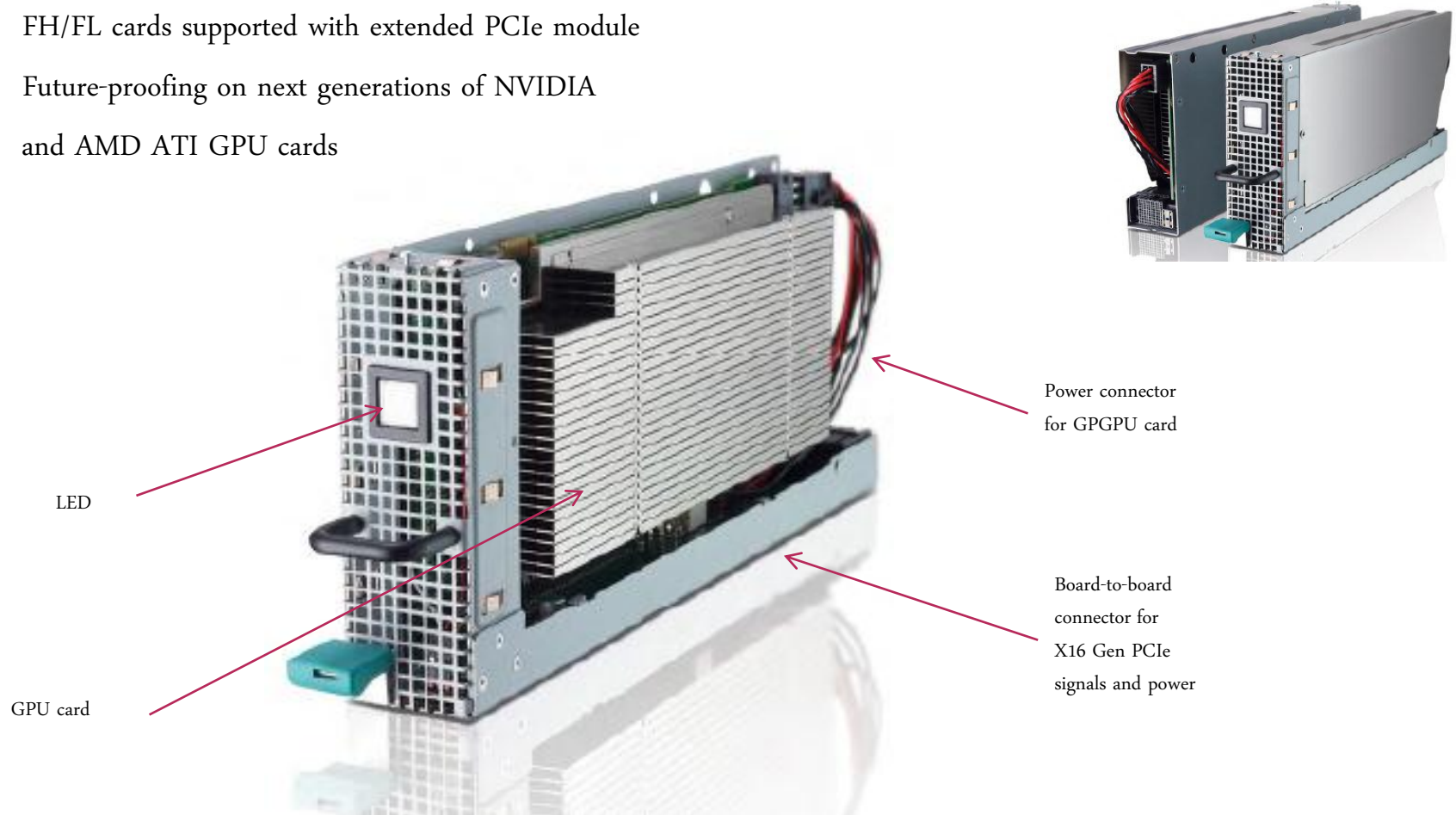**PCIe EXPANSION CHASSIS CONNECTING 1-8 HOSTS TO 1-16 PCIe**

Great for: HPC including universities, oil & gas, biomed research, design, simulation, mapping, visualization, rendering, and gaming

- 3U chassis, 19" wide, 143 pounds
- PCI express modules: 10 front, 6 rear
- PCI form factors: HH/HL and FH/HL
- Up to 225W per module
- PCIe inputs: 8PCIe x16 IPASS ports
- PCI fan out options: x16 to 1 slot, x16 to 2 slot, x16 to 3 slot, x16 to 4 slot
- GPUs supported: NVIDIA M1060, M2050, M2070 (TBD)
- Thermals: high-efficiency 92mm fans; N + 1 fan redundancy

- Management: On-board BMC; IPMI 2.0; dedicated management port
- Power supplies: 4 x 1400W hot-plug, high efficiency PSUs; N+1 power redundancy
- Services vary by region: IT Consulting, Server and Storage Deployment, Rack Integration (US only), Support Services

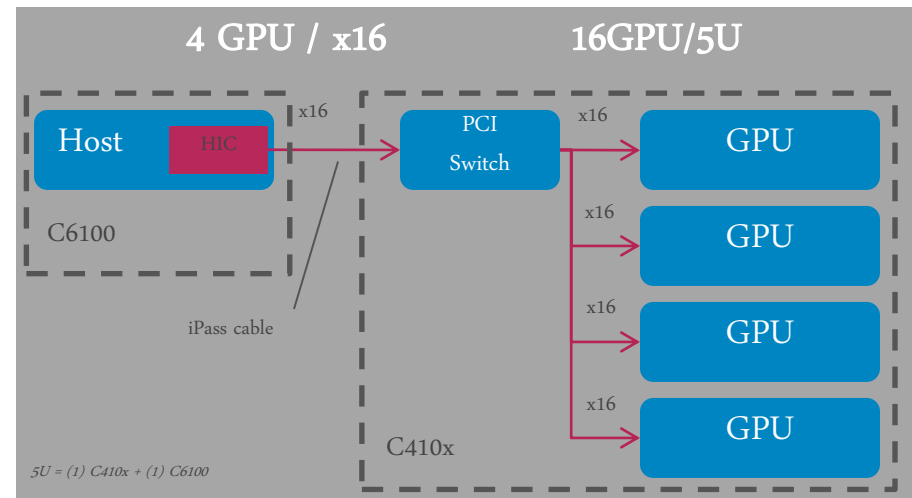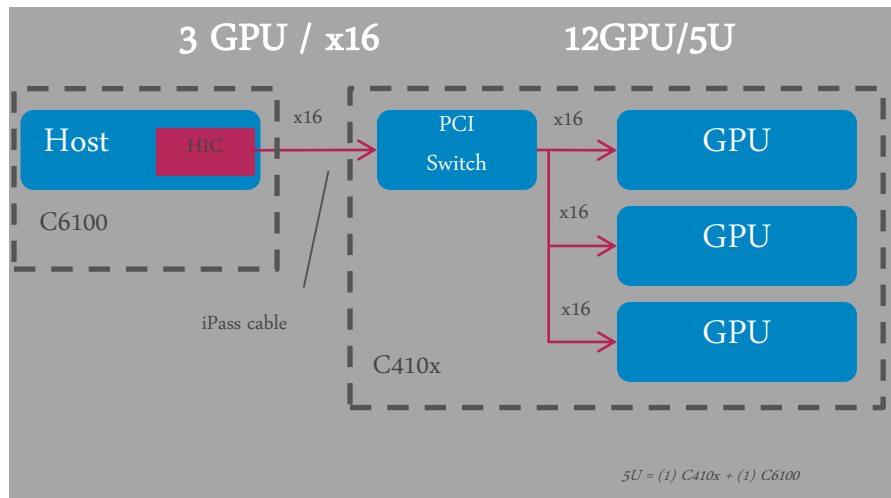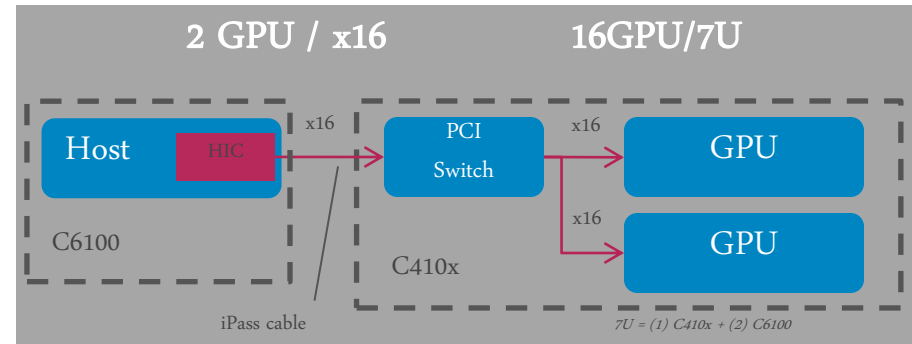**OPTIMIZED** RESPONSIVE EFFICIENCY

DELL

PowerEdge C410x PCIe Module

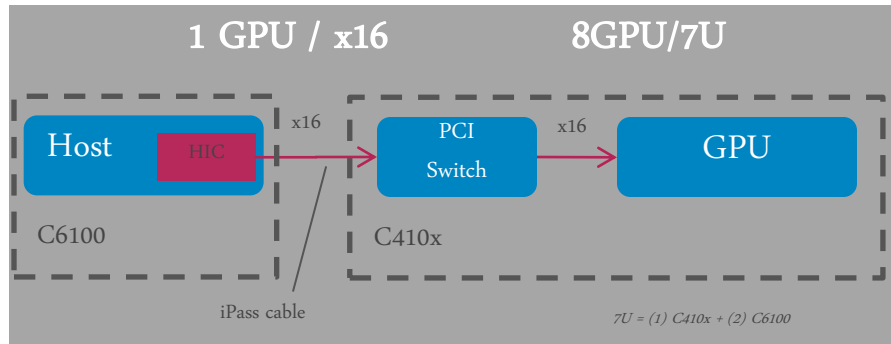- Serviceable PCIe module (taco) capable of supporting any half-height, half-length (HH/HL) or full-height/half-length (FH/HL) cards

- FH/FL cards supported with extended PCIe module

- Future-proofing on next generations of NVIDIA and AMD ATI GPU cards

Power connector for GPGPU card

Board-to-board connector for X16 Gen PCIe signals and power

LED

GPU card

SPEED TCO MODULAR CUSTOMIZED

DELL

Flexibility of the PowerEdge C410x

- Enabling HPC applications to optimize cost / performance equation off single x16

1 GPU / x16   **8GPU/7U**

Host — HIC   x16 → PCI Switch   x16 → GPU

C6100   C410x

iPass cable

7U = (1) C410x + (2) C6100

2 GPU / x16   **16GPU/7U**

Host — HIC   x16 → PCI Switch   x16 → GPU / x16 → GPU

C6100   C410x

iPass cable

7U = (1) C410x + (2) C6100

3 GPU / x16   **12GPU/5U**

Host — HIC   x16 → PCI Switch   x16 → GPU / x16 → GPU / x16 → GPU

C6100   C410x

iPass cable

5U = (1) C410x + (1) C6100

4 GPU / x16   **16GPU/5U**

Host — HIC   x16 → PCI Switch   x16 → GPU / x16 → GPU / x16 → GPU / x16 → GPU

C6100   C410x

iPass cable

5U = (1) C410x + (1) C6100

GPU/U ratios assume PowerEdge C6100 host with 4 servers per 2U chassis

SPEED TCO MODULAR CUSTOMIZED

DELL

Flexibility of the PowerEdge C410x

- Enabling HPC applications to optimize cost/performance equation off dual x16

SPEED TCO MODULAR CUSTOMIZED
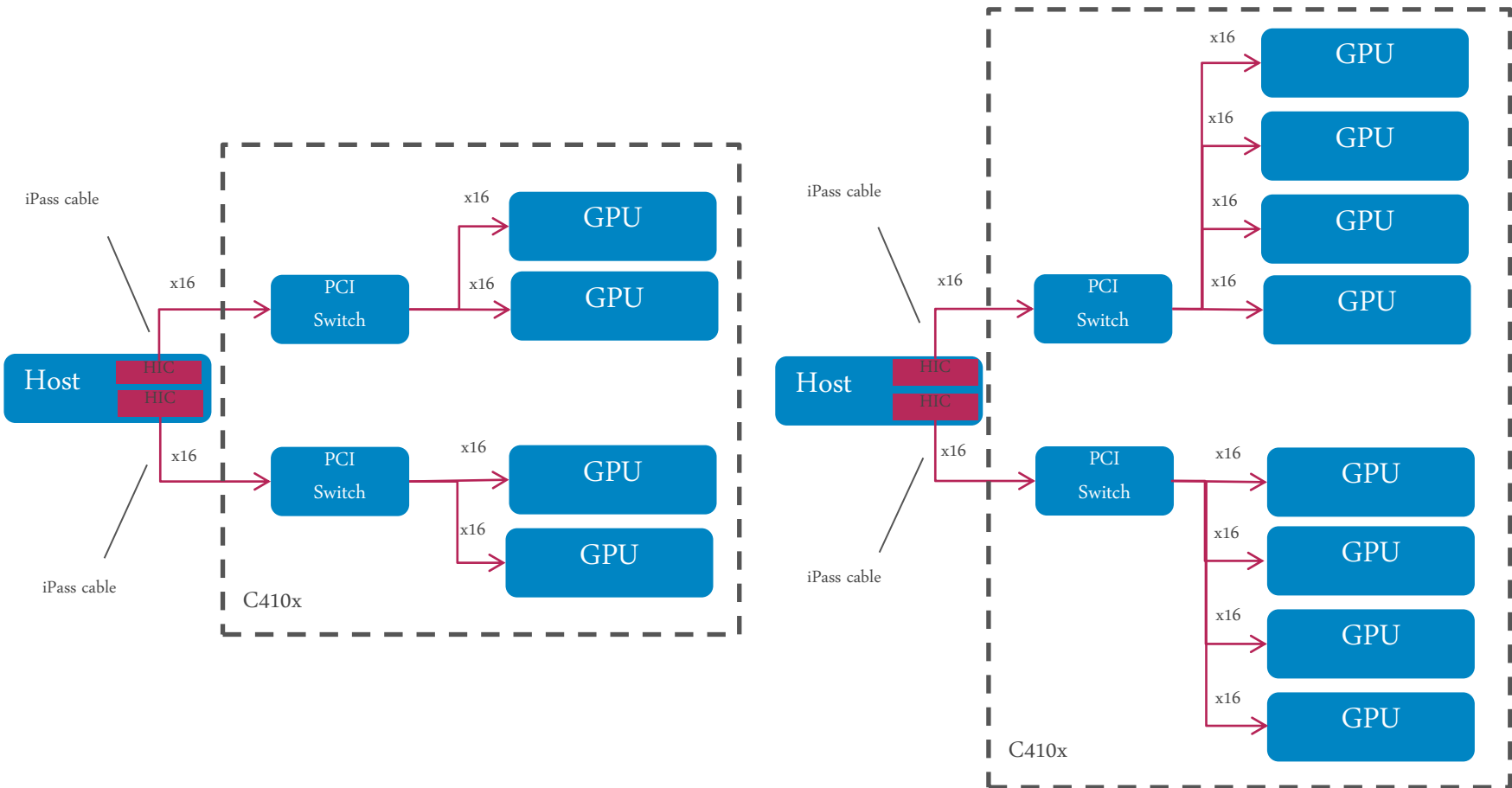
FERMI ready: extra performance

GPU	G80	GT200	Fermi
Transistors	681 million	1.4 billion	3.0 billion
CUDA Cores	128	240	512
Double Precision Floating Point Capability	None	30 FMA ops / clock	256 FMA ops /clock
Single Precision Floating Point Capability	128 MAD ops/clock	240 MAD ops / clock	512 FMA ops /clock
Special Function Units (SFUs) / SM	2	2	4
Warp schedulers (per SM)	1	1	2
Shared Memory (per SM)	16 KB	16 KB	Configurable 48 KB or 16 KB
L1 Cache (per SM)	None	None	Configurable 16 KB or 48 KB
L2 Cache	None	None	768 KB
ECC Memory Support	No	No	Yes
Concurrent Kernels	No	No	Up to 16
Load/Store Address Width	32-bit	32-bit	64-bit

SPEED TCO MODULAR CUSTOMIZED

DELL

Why PowerEdge C410x?

See Results Faster

- 16.5 TFLOPs of computing throughput in single precision performance per server (fully loaded with M2050 cards)
- Configs up to 16 GPUs so you can set the right ratio to maximize application performance

Mix and Match Flexibility

- Enables 1-8 servers to attach to up to 16 PCIe devices
- PCIe devices can include GPUs, IB cards and more

Highly Serviceable

- The only PCIe expansion chassis to provide up to 3x the uptime as competing chassis
- Hot-add serviceable GPUs
- Hot-plug redundant power supplies and fans

PCI e EXPANSION CHASSIS
CONNECTS 1-8 HOSTS TO 1-16PCIe

Save

- 25% more density
- 31% less weight
- 11% less the cost of additional switches, racks and energy over the NVIDIA S1070

OPTIMIZED RESPONSIVE EFFICIENCY

DELL

Recommended Dell Configurations

- PowerEdge C6100/C410x options

  – The PowerEdge C6100 is validated with the C410x at this time

  – Support is for the NVIDIA Tesla M-series (M1060, M2050/**M2070**)

    › AMD ATI Firestream support will be available in the future

  – Many options for configurations

    › Configurations are defined as the [Number of GPUS]:[Nodes]

      - For example "1:1 sandwich"

  – Key differentiators for Dell:

    1. Quad data rate (QDR) Infiniband is available on all configurations

    2. Redundant power supplies on the host node

    3. Redundant power and fans on the GPU chassis

    4. Independent serviceability on the host node and GPUs

SPEED TCO MODULAR CUSTOMIZED

PowerEdge C6100 and PowerEdge C410x Configurations

ON-DEMAND

DELL

PowerEdge C6100 Configurations

"1:1 Sandwich"

C6100
C410x
C6100

Summary

C6100 "1:1 Sandwich"

One Dell C410x (8 GPUs)

Two C6100 (8 nodes)

One x16 slot for each node to 1 GPU

7U total

- 8 GPUs total
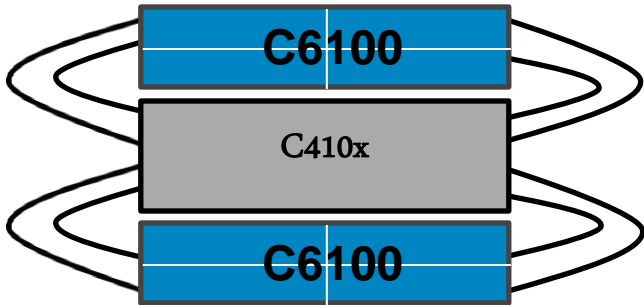
- 8 nodes total (1 GPU per board)

Details

- Two C6100

    - 8 system boards each with:

        - 2S Westmere, 12 DIMM slots, QDR IB, up to 6 drives per host

        - Single port x16 HIC (iPASS)

- Single C410x

    - 8 GPUs (not fully populated)

        - Maxed out number of iPASS connections from C410x

- **PCIe x16 per GPU** (max host-GPU throughput)

- Total space = 7U

- See later table for metrics

SPEED ICO CUSTOMIZED
STACK HYPERSCALE

PowerEdge C6100 Configurations

"2:1 Sandwich"

Summary

C6100 "2:1 Sandwich"

One Dell C410x (16 GPUs)

Two C6100 (8 nodes)

One x16 slot for each node to 2 GPUs

7U total

16 GPUs total
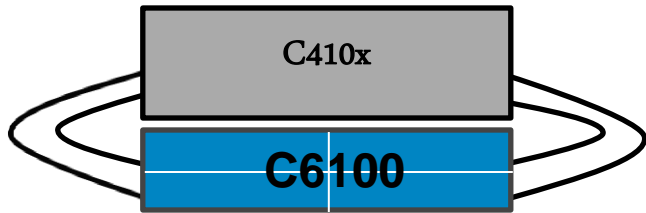
8 nodes total (2 GPUs per board)

Details

- Two C6100
  - 8 system boards
    - 2S Westmere, 12 DIMM slots, QDR IB, up to 6 drives per host
    - Single port x16 HIC (iPASS)
- Single C410x
  - 16 GPUs (fully populated)
- **PCIe x8 per GPU**
- Total space = 7U
- See later table for metrics

Note: This configuration is equivalent to using the C6100 and the NVIDIA S2050 but this configuration is more dense

PowerEdge C6100 Configurations

"4:1 Sandwich"

Summary

C6100 "4:1 Sandwich"

One Dell C410x (16 GPUs)

One C6100 (4 nodes)

One x16 slot for each node to 4 GPUs
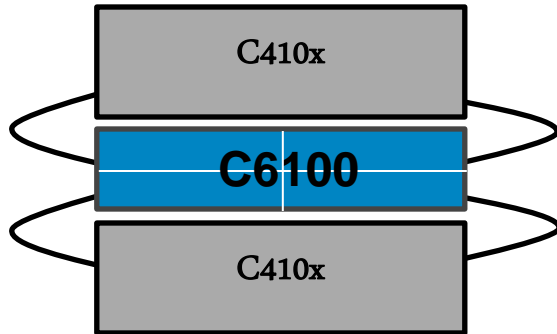
5U total

    16 GPUs total

    4 nodes total (4 GPUs per board)

Details

- One C6100
  - 4 system boards
    - 2S Westmere, 12 DIMM slots, QDR IB, up to 6 drives per host
    - Single port x16 HIC (iPASS)
- Single C410x
  - 16 GPUs (fully populated)
- **PCIe x4 per GPU**
- Total space = 5U
- See later table for metrics

PowerEdge C6100 Configurations

"8:1 Sandwich" (Possible Future Development)

```
┌─────────────────┐
│      C410x       │
├─────────────────┤
│      C6100        │
├─────────────────┤
│      C410x       │
└─────────────────┘
```

Summary

C6100 "8:1 Sandwich"

Two Dell C410x (32 GPUs)

One C6100 (4 nodes)

One x16 slot for each node to 8 GPUs

8U total

  32 GPUs total

  4 nodes total (8 GPUs per board)

Details

- One C6100
  - 4 system boards
    - 2S Westmere, 12 DIMM slots, QDR IB, up to 6 drives per host
    - Single port x16 HIC (iPASS)
- Two C410x
  - 32 GPUs (fully populated)
- **PCIe x2 per GPU**
- Total space = 8U
- See later table for metrics

SPEED TO CUSTOMIZED
STACK HYPERSCALE

PowerEdge C6100 Configuration Summary

- All configurations use C6100+C410x

  - All host nodes have QDR IB and up to 6 drives per host node

  - 2S host nodes and 12 DIMM slots

- Primary metrics:

  - Number of GPUs per U (GPU/U)

  - Host nodes per U

  - Effective number of lanes per GPU

Configuration (GPU:Host)	GPU/U	2S Host Nodes/U	Effective number of PCIe lanes per GPU	Number of GPUs per socket
1:1	1.14	1.14	16	0.5
2:1	2.28	1.14	8	1
4:1	3.20	1.25	4	2
8:1	4.00	0.50	2	4
Supermicro	2.00	1.00	16	1
Appro	4.00	1.00	8	2

SPEED TCO MODULAR CUSTOMIZED

NVIDIA Tesla GPU Computing Products

Data Center Products   Workstation

	Server Module			1U Systems			Workstation Boards		
	Tesla M1060	Tesla M2050	Tesla M2070	Tesla S2070	Tesla 2050	Tesla S1070	Tesla C2070	Tesla C2050	Tesla C1060
GPUs	1 T10 GPU	1 T20 GPU	1 T20 GPU	4 T20 GPUs		4 T10 GPUs	1 T20 GPU		1 T10 GPU
Single Precision	933 GFLOPS	1030 GFLOPS	1030 GFLOPS	4120 GFLOPS		4140 GFLOPS	1030 GFLOPS		933 GFLOPS
Double Precision	78 GFLOPS	515 GFLOPS	515 GFLOPS	2060 GFLOPS		346 GFLOPS	515 GFLOPS		78 GFLOPS
Memory	4 GB (no ECC)	3 GB (2.625 GB w/ ECC)	6 GB (5.25 GB w/ ECC)	4 * M2070	4 * M2050	4 * M060	6 GB (5.25 GB ECC)	3 GB (2.625 GB ECC)	4 GB (no ECC)
# of Cores/ Frequency	240@1.3 GHz	448@1.15 GHz	448@1.15 GHz	4 * (448@1.15 GHz)	4 * (448@1.15 GHz)	4 * (240@1.3 GHz)	448@1.15 GHz	448@1.15 GHz	240@1.3 GHz
Mem BW	102 GB/s	148.4 GB/s	148.4 GB/s	148.4 GB/s		102 GB/s	144 GB/s		102 GB/s
Display	No display IO			No display IO			Single dual-link DVI		No display IO

HPC STORAGE

- Storage solution depends on:

  - Applications, IO patterns, number of users

  - Network

  - Throughput requirements

  - Backups, HSM

- Most Dell HPC solutions are configured with Dell PowerVault MD Series Storage.

- For Specialized solution requirements Dell offers:

  - EqualLogic

  - Compellent

Realities

- HPC storage is about 15-25% the cost of a system but about 90% of the problems

- HPC Storage is about **Solutions** not just hardware

  – Hardware, file system, client, management/monitoring, documentation, support

- There are no one, two, or even three file systems/solutions that satisfy the various requirements

  – Recent IDC study: 25 customers = 13 file systems

- Applications/Processes drive solutions (*just like compute*)

  – Very few understanding about IO characteristics of the apps

  – A very large percentage of data is never touched approximately 2-4 weeks after it is created

SPEED TCO MODULAR CUSTOMIZED

The DELL HPC Storage Solution

- Full Lustre solution, fully configured, tested, tuned, and deployed

  – On-site installation with client deployment and training included

- Simple, linear scalability

  – 2.4 GB/s of throughput per base object solution (redundant), linearly scalable

  – 48TB minimum capacity

  – Up to 336 TB of capacity in standard complete configuration in a single rack
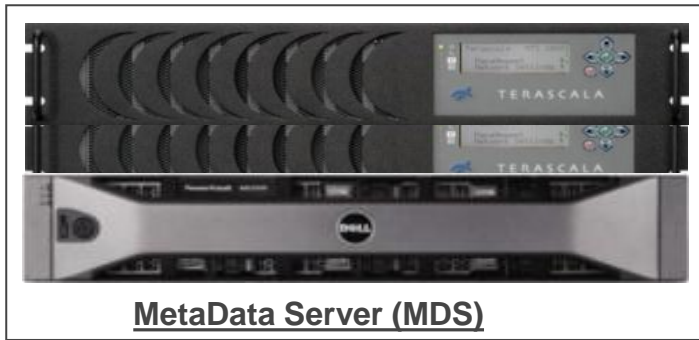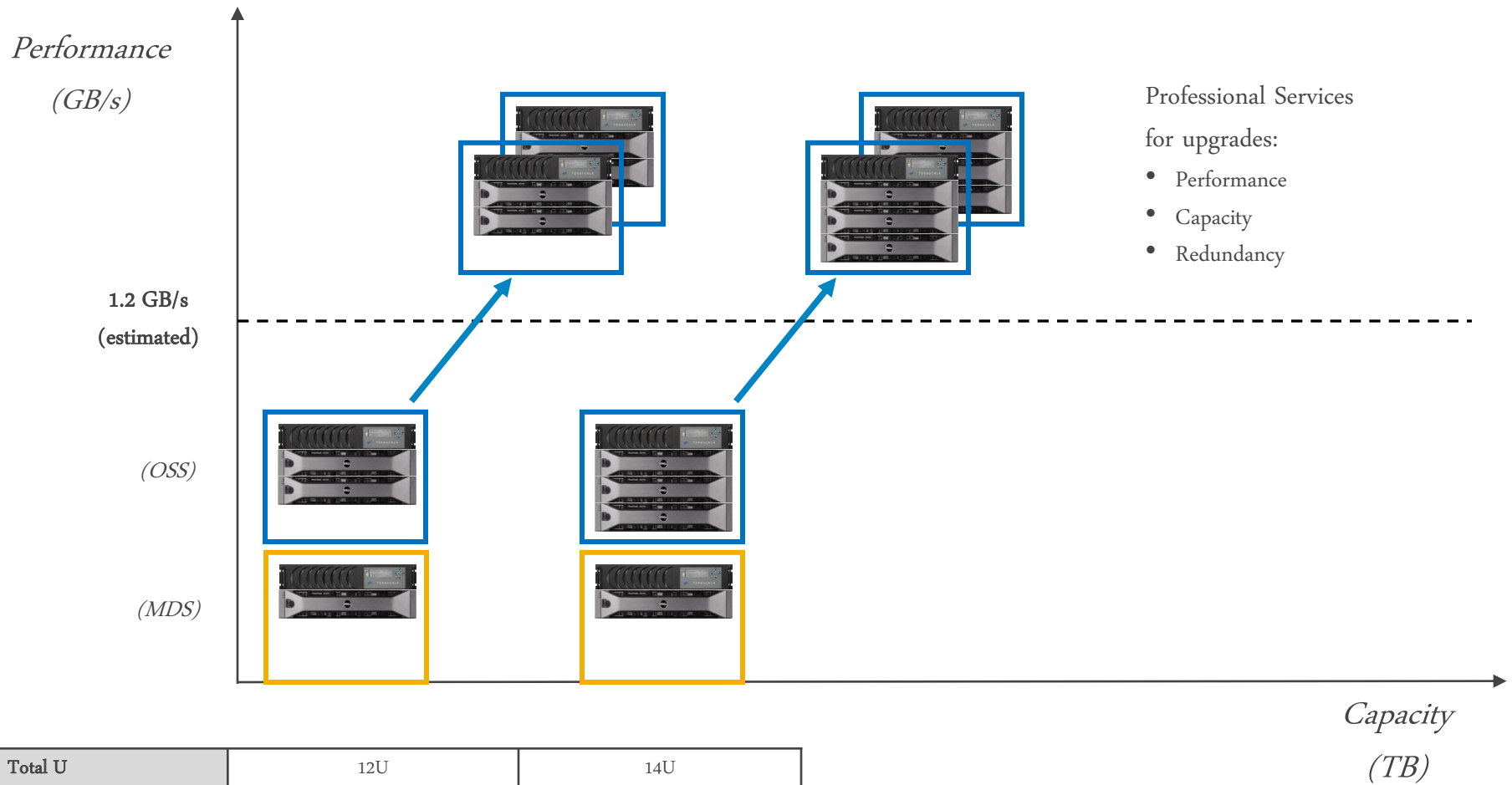
- Full management system

MetaData Server (MDS)

Base OSS

Building Blocks

MetaData Server (MDS)

Base OSS

- 2 redundant active - passive servers with QDR connectivity
- Dell MD3220 12 TB storage array (24 drives)
- Supports multiple Object Storage Servers

- 2.4 GB/sec estimated throughput for active/active OSS node
  - 1.2 GB/s estimated throughput for non-redundant node
- 2 Dell MD3200 arrays providing 48TB of storage
  - Expansion with Dell MD1200 with 24TB per enclosure
- Use the building block to architect as per your needs

Architected to scale in capacity and performance

SPEEDTCO MODULAR CUSTOMIZED

Upgradables solutions

Performance
(GB/s)

1.2 GB/s
(estimated)

(OSS)

(MDS)

Capacity

(TB)

Professional Services for upgrades:
- Performance
- Capacity
- Redundancy

	12U	14U
Total U	12U	14U
Capacity	48TB	72TB
Estimated Performance	~1.2 GB/s	~1.2 GB/s

Q&A
Thanks!!!!

Backup Slides

Competitive Advantages of the PowerEdge C410x

- PowerEdge C410x PCIe expansion chassis is purposefully designed to power, cool and manage PCIe devices

    - Supports GPUs and other PCIe devices

    - Redundancy

        › n+1 PSUs and n+1 fans

    - PCIe modules are individually serviceable

        › Can service a GPU, device, PSU or fan without bringing down the entire chassis.

        › No opening of compute nodes

        › No bumped DIMMs

        › No disturbed dust

        › No cabling / un-cabling issues

    - Onboard BMC to monitor and manage chassis health and configure PCIe switches

SPEED TCO MODULAR CUSTOMIZED

DELL