Unique Cooling Solutions for Dense HPC Systems

John Lee, VP, ATS

APPRO
Supercomputer Solutions

Industry Trends

- Rack Power Density is going up
  - In 2006, Peloton was ~ 24kW/rack
  - In 2008, TLCC was ~ 27kW/rack
  - In 2010, Edge Cluster is ~ 32kW/rack (GPU Cluster)
- Outside of DOE, we have deployed >34kW/rack densities to commercial datacenters (GPU)
- We have some configurations that is >50kW/rack!
  - No one has bought them yet though - ☺
- Most cluster shipments are still predominantly indirect (air) cooling
- Customers looking for flexible power/cooling solutions
  - Liquid-cooled rear door heat exchanger
  - Hot or cold-aisle containment
  - There is no one right answer to every problem

APPRO Supercomputer Solutions

System Focus

:: Design Objectives

- **Ideal Building Block** for commodity HPC applications
- **Open Standards –** maximum flexibility to support standard 19” rackmount infrastructure eco-system
- **1U Alternative** – All the benefits of standard 1U server without any of its weaknesses
- **Improved power savings** through shared power/cooling infrastructure
- **Improved *Reliability/Availability/ Serviceability*** (RAS) over standard 1U servers
- **Cost-effective design to maintain price-parity with 1U servers** with additional value-add features at no premium

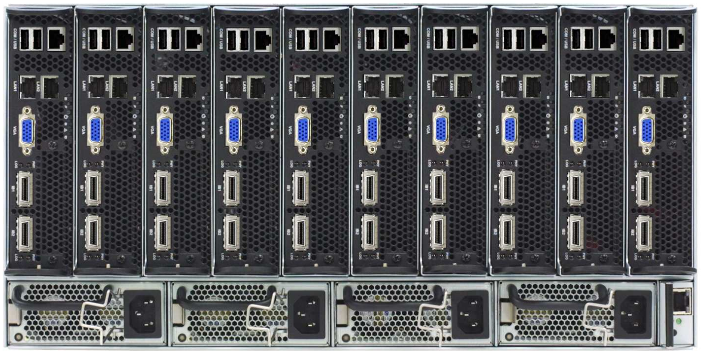
System Focus

:: Feature Summary

GreenBlade

Front View

Rear View



- 5RU Chassis holds up to ten dual socket blade servers or five hybrid GPU blades
- Up to four high-efficient 1625W hot-swappable PS in either 2+2 or N+1 configuration
- Supports three hot-swappable, redundant fan modules
- Shared infrastructure design reduces system power consumption up to 20%
- Integrated chassis management module
  - Monitors & controls individual blades
  - Monitors & controls platform PS
  - Monitors & controls platform fans
  - Supports Powerman

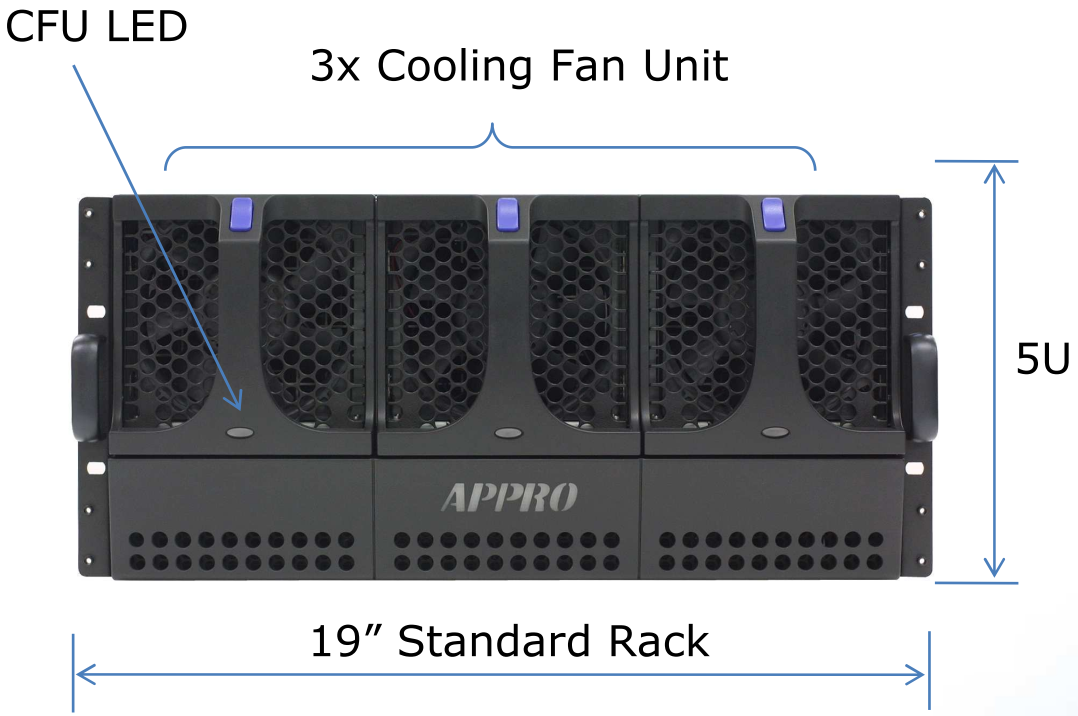
anagement

APPRO
Supercomputer Solutions

System Focus

:: Cooling

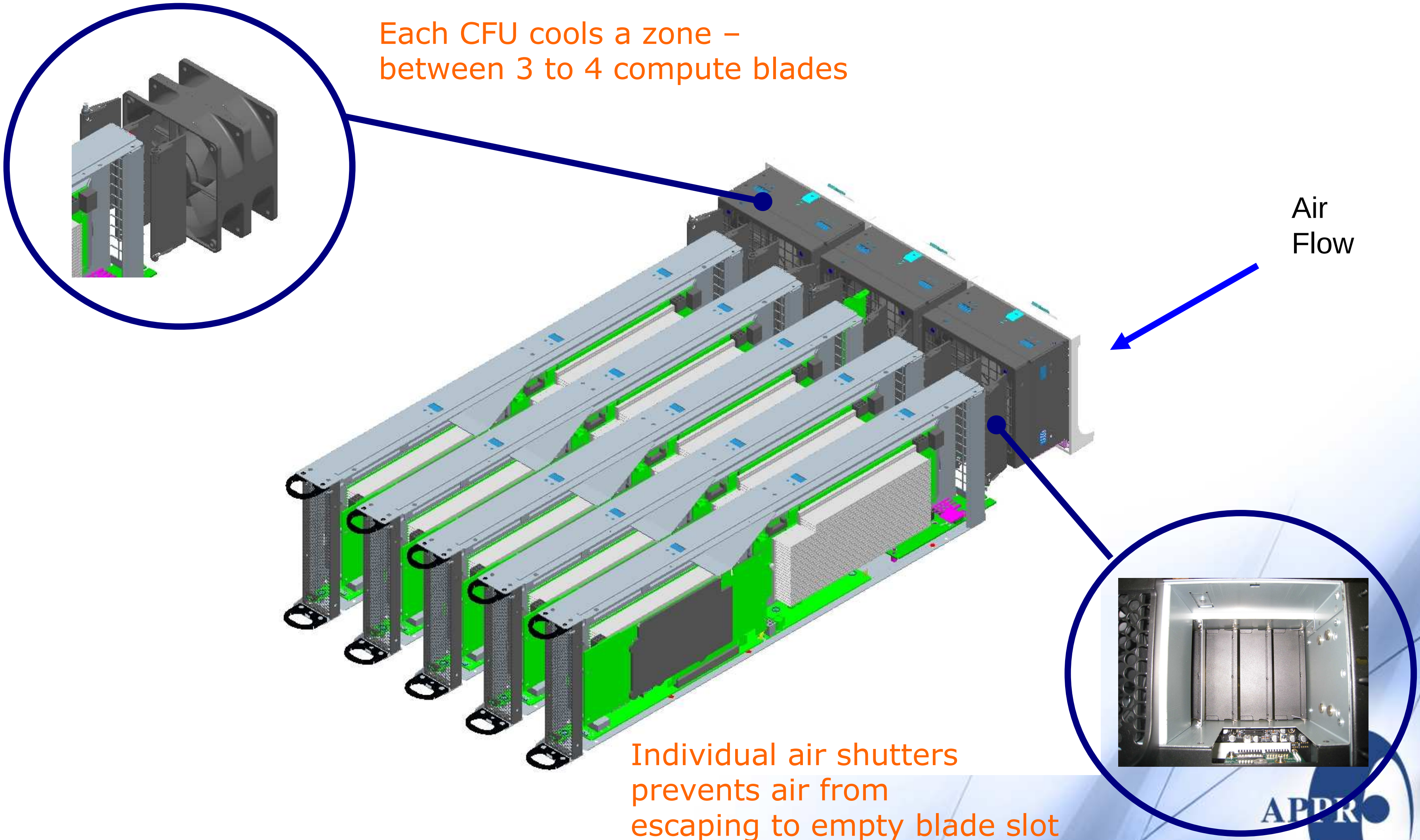
GreenBlade



- 3x Cooling Fan Unit (CFU)
- Each CFU has two, redundant 120mm fans
- CFU LED: Green for normal and Red for Warning
- With Chassis Manager, fan speed can be controlled
- Customizable IDLE/NORMAL/HIGH fan operation settings

System Focus

:: Cooling



Each CFU cools a zone – between 3 to 4 compute blades

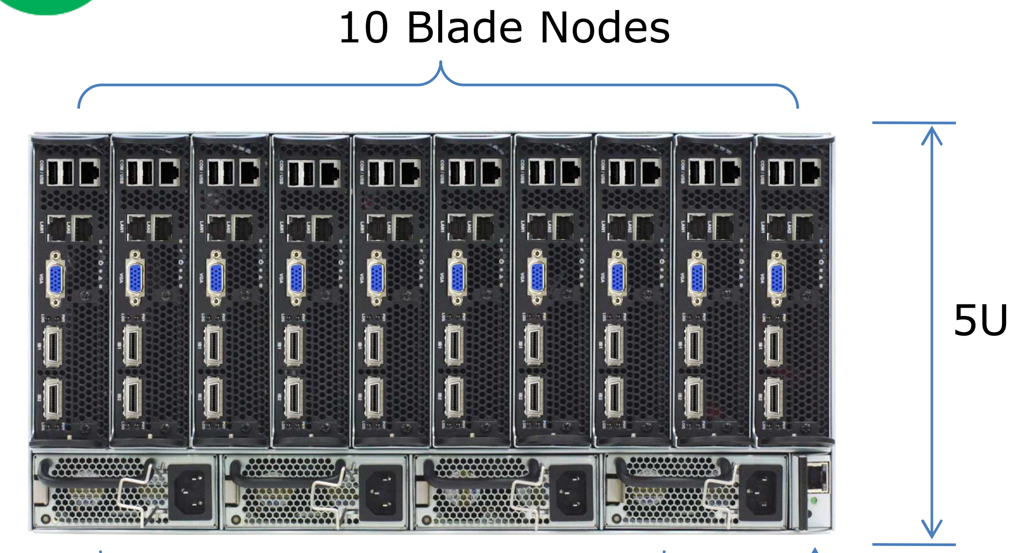
Air Flow

Individual air shutters prevents air from escaping to empty blade slot

System Focus

:: Power Distribution

GreenBlade



10 Blade Nodes

5U

4 Pluggable PSU (1625W/PS)
(N+1 or 2+2 Redundant configuration)

Chassis Manager

- Up to 4 power supplies in N+1 configuration
- Over 90% efficient power supplies
- Standard redundant configuration is 2+1 (~300W/node allocation)
- 3+1 redundant configuration can support rich system configuration (~460W/node allocation)

System Focus

:: Advantages of Shared Infrastructure

10x 1U Servers

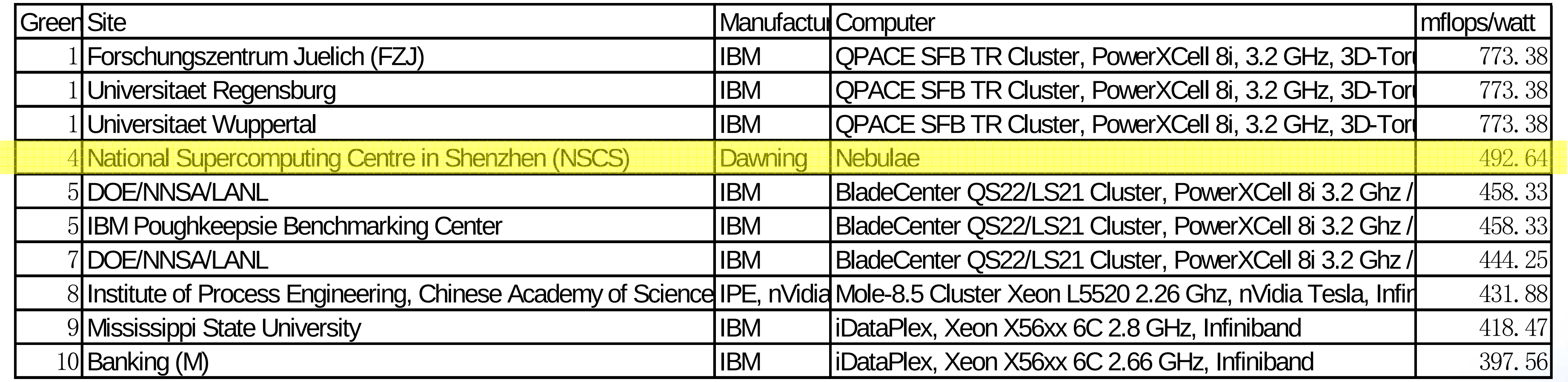
VS

GreenBlade

- @ 60W avg. saving/server, one GB chassis = 600W power savings
- 600W savings translate to 2 additional servers
- Go from (12x10) 120 fans to 6 fans
- Go from 10 power supplies to 2 power supplies (non-redundant)
- Go from 20 power supplies to 3 or 4 power supplies (redundant)

System Focus

:: Green500

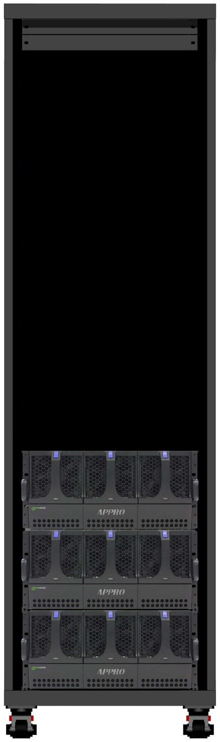


Green	Site	Manufactu	Computer	mflops/watt
1	Forschungszentrum Juelich (FZJ)	IBM	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Tor	773.38
1	Universitaet Regensburg	IBM	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Tor	773.38
1	Universitaet Wuppertal	IBM	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Tor	773.38
4	National Supercomputing Centre in Shenzhen (NSCS)	Dawning	Nebulae	492.64
5	DOE/NNSA/LANL	IBM	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz /	458.33
5	IBM Poughkeepsie Benchmarking Center	IBM	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz /	458.33
7	DOE/NNSA/LANL	IBM	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz /	444.25
8	Institute of Process Engineering, Chinese Academy of Science	IPE, nVidia	Mole-8.5 Cluster Xeon L5520 2.26 Ghz, nVidia Tesla, Infir	431.88
9	Mississippi State University	IBM	iDataPlex, Xeon X56xx 6C 2.8 GHz, Infiniband	418.47
10	Banking (M)	IBM	iDataPlex, Xeon X56xx 6C 2.66 GHz, Infiniband	397.56

- The Edge Cluster is a GreenBlade-based hybrid cluster with nVIDIA Fermi GPUs and Intel Westmere hosts
- Achieved over 100TF on Linpack this month ~ 667MFLOPs/Watt
- @ 667MFLOPs/W – would be ranked #4 in June 2010 list

System Focus

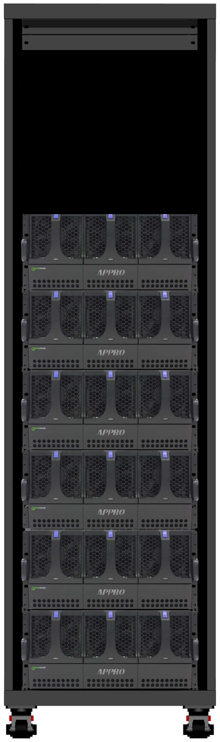
:: 10.5kW/rack Configuration



1.8kW savings

System Focus

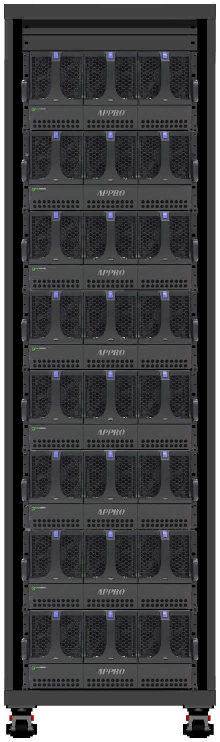
:: 21kW/rack Configuration



3.6kW savings

System Focus

:: 28.5kW Configuration



4.8kW savings

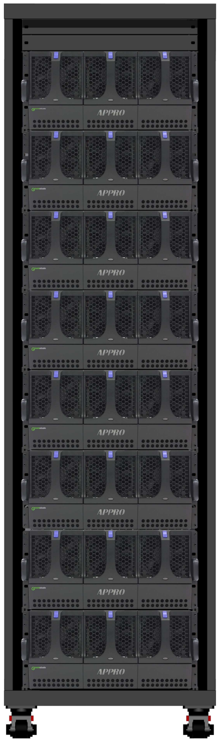
Supercomputer Solutions

System Focus

:: Rich Infrastructure Eco-system

APC
by Schneider Electric

EMERSON
Network Power



Raritan.
Know more. Manage smarter.

Server Technology
Solutions for the Data Center Equipment Cabinet

VETTE CORP®

RITTAL

wright · line
outfitting technology-intensive environments

Looking into the future…

:: GreenBlade2

- **Features & Benefits over Gen1**
  - Additional cost reduction for even better price/performance without feature/quality sacrifices
  - Target 30% energy reduction
    - More efficient air flow
    - More efficient system components (MB, VRM..)
    - More efficient power supplies
    - More efficient cooling
  - More Intelligent Chassis Manager
    - Integrated DCM(Data Center Manager) for better real-time power monitoring, management & policy engine
    - Better algorithms for more dynamic fan speed control
  - Always looking for customer feedback to improve the product

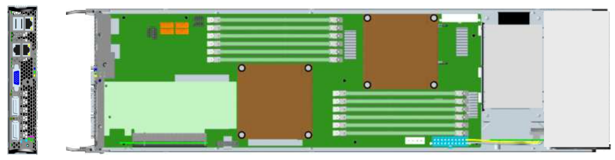
Do More with Less with Appro

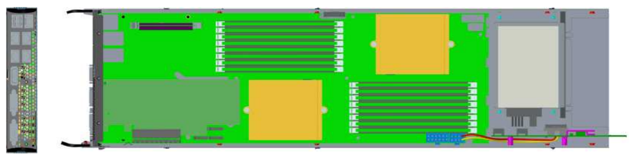
APPRO

Supercomputer Solutions

Product Focus: GreenBlade™

:: Compute Blades

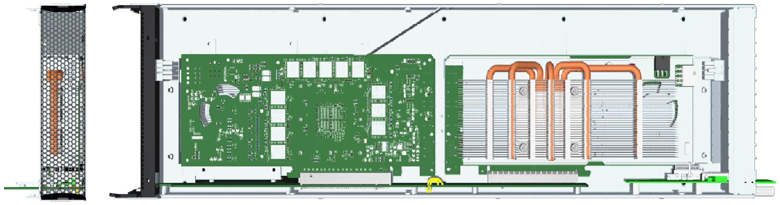




- Host System is hot-swappable
- Large memory footprint – up to 96GB
- Internal storage flexibility - supports up to 2x 2.5” disks – SATA HDDs or SSDs
- Integrated dual port Gigabit ethernet
- Integrated QDR Infiniband (option)
- Integrated IPMI 2.0 Remote Management

Product Focus: GreenBlade™

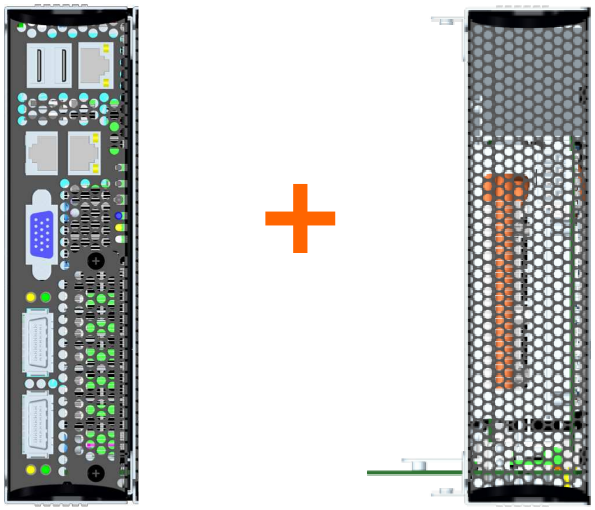
:: GPU expansion Blade



- **Features & Benefits**
  - GPU expansion Blade is hot-swappable
  - Supports two nVIDIA "Fermi" GPUs – M2050 or M2070
  - PCIe Gen2 x16 interface for maximum bandwidth
  - Flexibility to support either AMD or Intel hosts
  - Intelligent power control – GPU Blade can be independently powered down to save overall system power

Product Focus: GreenBlade™

:: GPU Compute Blades



GPU Compute Blade

Compute Host Blade

GPU Expansion Blade

- Direct PCIe bus slot to slot connection
- No need for external PCIe cables
- Host/GPU Pair is a single GPU blade system
- Each GPU system is hot-swappable & easily serviceable
- All monitoring sensors and data are integrated between host and GPU
- Host and GPU module can be upgraded independently