

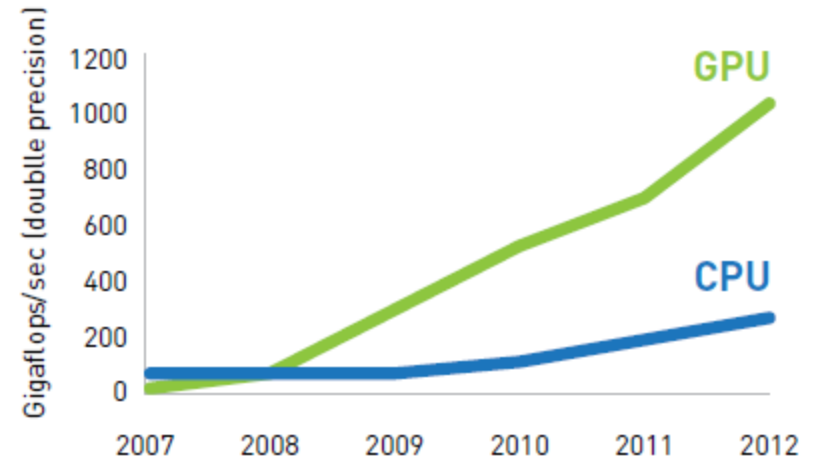
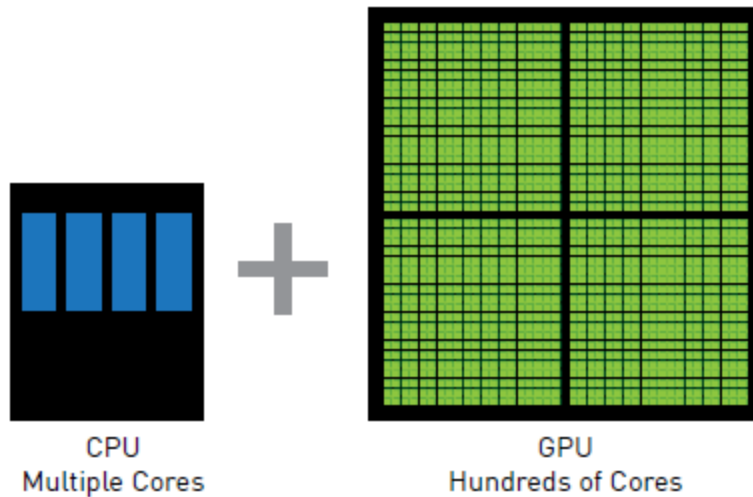


GPU, GPU cluster
Peta-Flop-in-a-Cabinet

Lehoczki Gábor – Silicon Computers Kft.

CPU + ?

- FP Coprocessor
- FPGA
- GPU graphics card as coprocessor



General Purpose GPUs (GPGPU)

- Volume product – cheap
- PCI-Express – easy to deploy
- OpenCL / CUDA / Firestream
 - Easy to program

GPU – nVidia Tesla

- nVidia Tesla C2050, M2050, and S2050
 - 448 CUDA core
 - 512 GFLOPS
 - 3 GB ram
 - \$3.000



GPU - AMD ATI FireStream (Radion)

Specifications	AMD FireStream™ 9350	AMD FireStream™ 9370
Number of GPUs	1	1
Memory Capacity	2GB DDR5	4 GB DDR5
Double Precision Floating Point	400 GFLOPS	528 GFLOPS
Single Precision Floating Point	2.0 TFLOPS	2.64 TFLOPS
TDP	175W	225W
Stream Cores	1440	1600
SIMD Processors	18	20
Core Clock Frequency	700 MHz	825 MHz
System Interface	PCIe 2.1	PCIe 2.1
Memory Bandwidth	128 GB / S	147 GB / S
Memory Clock Frequency	1.0 GHz	1.2 GHz
Dimensions	4.376" x 9.5"; Single slot	4.376" x 10.5"; Dual slot
Aux. Power Connector	6-pin (2x3)	8-pin (2x4)
Thermal solution	Passive heat-sink	Passive heat-sink
Display output	1 DP	1 DP

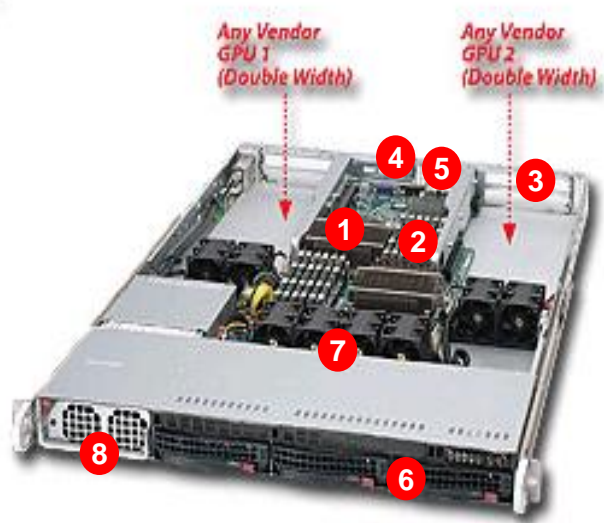
Desktop workstation with GPUs

- Big Tower House
- Large PSU
- Many PCI slots
- Extra cooling



- © Jurek Zoltán, Tóth Gyula MTA SZFKI

1U rack mount server for 2 GPUs



NVIDIA compute GPU support:

- S2050, S2070, M2050, M2070, M2070Q, M2090

NVIDIA graphics GPU support:

- Quadro FX 1800, 3800, 4800, 5800
- Quadro 2000, 4000, 5000, 6000

1	Processor Support Dual Nehalem-EP Xeon CPU, 5500 series
2	Memory Capacity 12 DIMM, DDR3 1333/1066/800 MHz
3	Expansion Slots 2 x PCI-e X16 Gen 2.0 (Full height, double width) - Optional riser cards to split one or both x16 slots in two x 8 slots - 1 x PCI-e X4 (Low-profile)
4	I/O ports 1 x VGA, 1 x COM, 2 x Gbit LAN, 2 x USB 2.0 ports
5	System management On board BMC (Baseboard Management Controllers) supports IPMI2.0
6	Drive Bays 3x hot-swap 3.5" SATA HDD
7	System Cooling 4 fans w/ Optimal Fan Speed Control
8	Power Supply 1400W High-efficiency power supply

SGI Altix XE500

3U, 2 socket rack mount server

Processor Support	Dual Nehalem-EP Xeon CPU, 5500 series
Memory Capacity	18 DIMM, DDR3 1333/1066/800 MHz
Expansion Slots	2 x PCI-e X16 Gen 2 (full height) 4 x PCI-e X8 Gen 2 (full-height)
I/O ports	1 x VGA, 2 x COM, 2 x Gbit LAN, 6 x USB 2.0 ports, 2 x PS/2 ports, 1 x Mic, 5 x Audio
System management	On board BMC supports IPMI2.0
Drive Bays	8 x 3.5" SAS/SATA HDD, RAID 0, 1, 5, 6, 10
System Cooling	5 x fans w/ Optimal Fan Speed Control
Power Supply	2 x 1200W High-efficiency redundant power supplies



NVIDIA compute GPU support:

- S2050, S2070, C2050, C2070

NVIDIA graphics GPU support:

- Quadro FX 1800, 3800, 4800, 5800
- Quadro 2000, 4000, 5000, 6000

Altix UV10

4U, four socket rack mount server

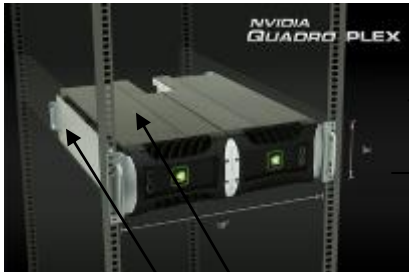
- 4 Intel Xeon E7 family (“Westmere EX”)
- 1 TB Memory
- **10 PCI Slots**
- 8 disc drive + DVD
- Redundant Power & Cooling
- 4U rack mount



Altix UV 100/1000

large scale SSI (smp) server with many GPUs

NVIDIA® Tesla™ or Quadro® Plex Enclosures Cable Directly to UV External PCIe Mezzanine Riser

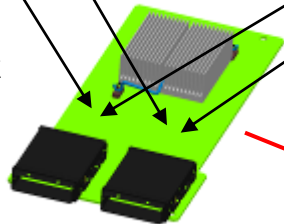


3U

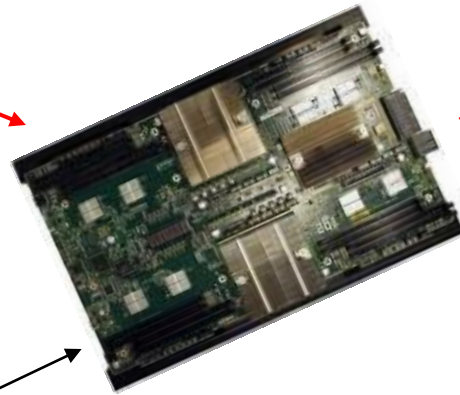
NVIDIA Tesla unit =
2 or 4 GPU, Two x16 links



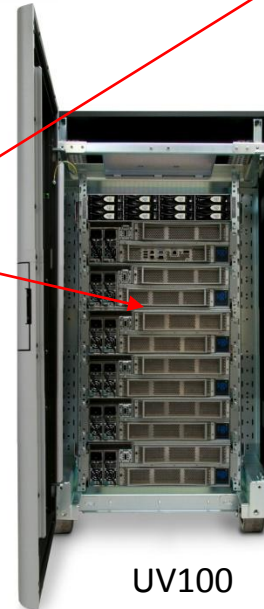
NVidia Quadro Plex
unit = 2 GPU +
Gsync, one X16 link
(2 units shown)



**Altix UV 100/1000
(2) PCIe Gen2 x16 Cable**



Each UV 100 or UV 1000 blade can connect up to one
NVIDIA Tesla or Quadro Plex enclosures



UV100



UV1000

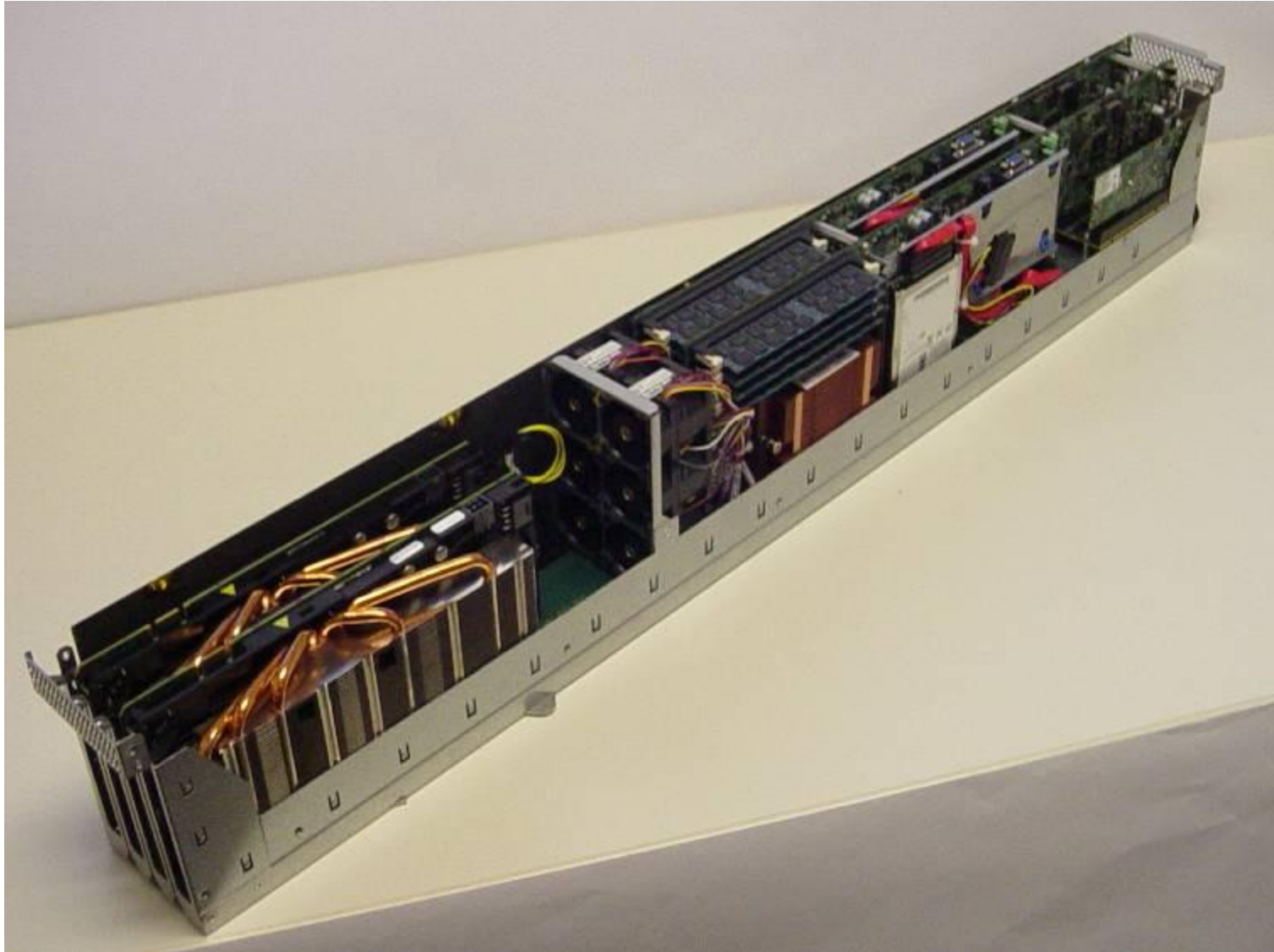
GPU cluster - STIX™ Architecture

- Design starts with the PCIe x 16 slot and...
- Wraps enough of a motherboard around it to give the accelerator access to:
 - More memory
 - Storage
 - I/O

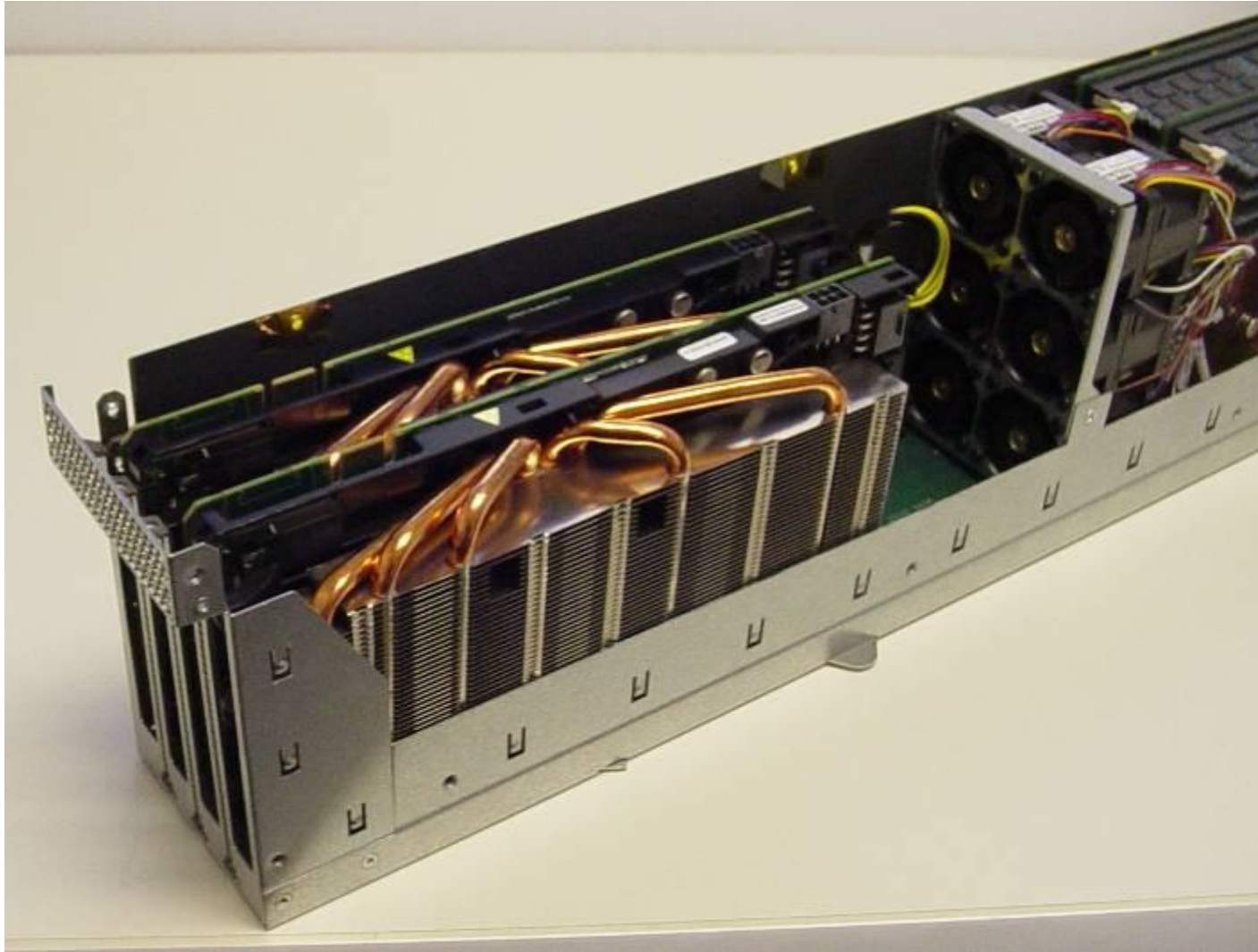
STIX™ Architecture

- Each Stick has two ‘slices’
 - A double-wide accelerator + motherboard + optional Infiniband card
 - ‘mirror’ images of each other
- Motherboard
 - 1S AMD Opteron 4100
 - 4 DIMM slots, DDR3, 1333 MHz
 - 2 x SATA2 drives (250GB, 500GB, 1 TB) supported
 - BMC
 - 2 x GigE
 - USB

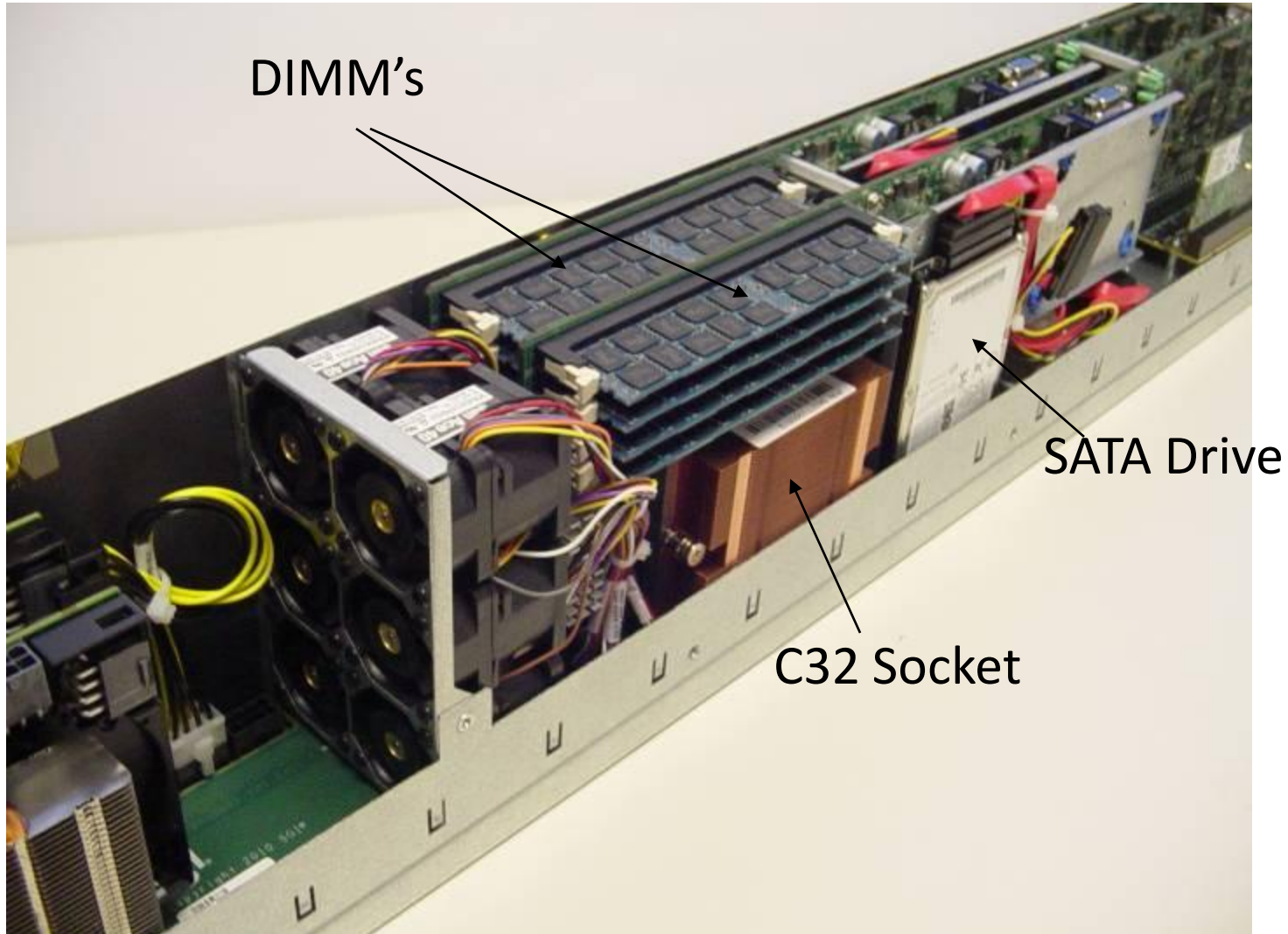
The Stick!



Accelerator Card Detail (NVIDIA M2050's)



Motherboard Details



Front Stick Detail



PCIe Infrastructure Capabilities

- Full Bandwidth PCIe gen 2 x 16 slot for the accelerator
 - Full 16 lanes
 - No PCIe switches for multiplexing
 - Up to 300W total power to include all current accelerator cards, plus *future—flexible and future-proofed!*
- PCIe gen 2 x 8 slot for networking
 - Single or dual-port QDR Infiniband HBA support for optimized cluster integration

Enclosure

- 2U
- 3 bays for sticks



SGI Prism XL Rack = 584 TFLOP SP



126x ATI Radeon HD5970 = 584 TFLOP SP

126x ATI Radeon HD9370 = 332 TFLOP SP

252x ATI FireStream 9350 = 504 TFLOP SP




Fully Populated SGI Prism™ XL Rack

42U

Switches & headnode in additional I/O rack

SGI Accelerator Execution Environment (AEE)

- AEE provides an integrated execution and development environment for different PCIe based hardware accelerators
- Highly scalable to petascale deployments

Accelerators	 NVIDIA NVIDIA GPUs	 AMD The future is fusion AMD GPUs	 TILERA Tilera many-core SoC CPUs
Development Software	Third-party tools from the leading vendors: NVIDIA, AMD, Tilera, Allinea, CAPS Enterprise, Portland Group, Rogue Wave		
System Management	Accelerator aware management using SGI Management Center-Premium Edition		
Job Scheduling	Accelerator aware workload scheduling using Altair PBS Professional		
OS	Red Hat RHEL 5.5, CentOS 5.5		

Customers/prospects/comparison

- P....d
470x Nvidia M2070
GPU: 40TFLOPS for \$2M
 - NIIF-Debreceni Egyetem
256x Xeon X5680 (3,3 GHz, 6c)
CPU: 18TFLOPS for \$1,5M
- ↔
- NIIF
GPU extensions of the new HPC
centers
2-8 GPU