

**GPU** TECHNOLOGY  
CONFERENCE

# Advanced GPU Computing References with HP Workstations



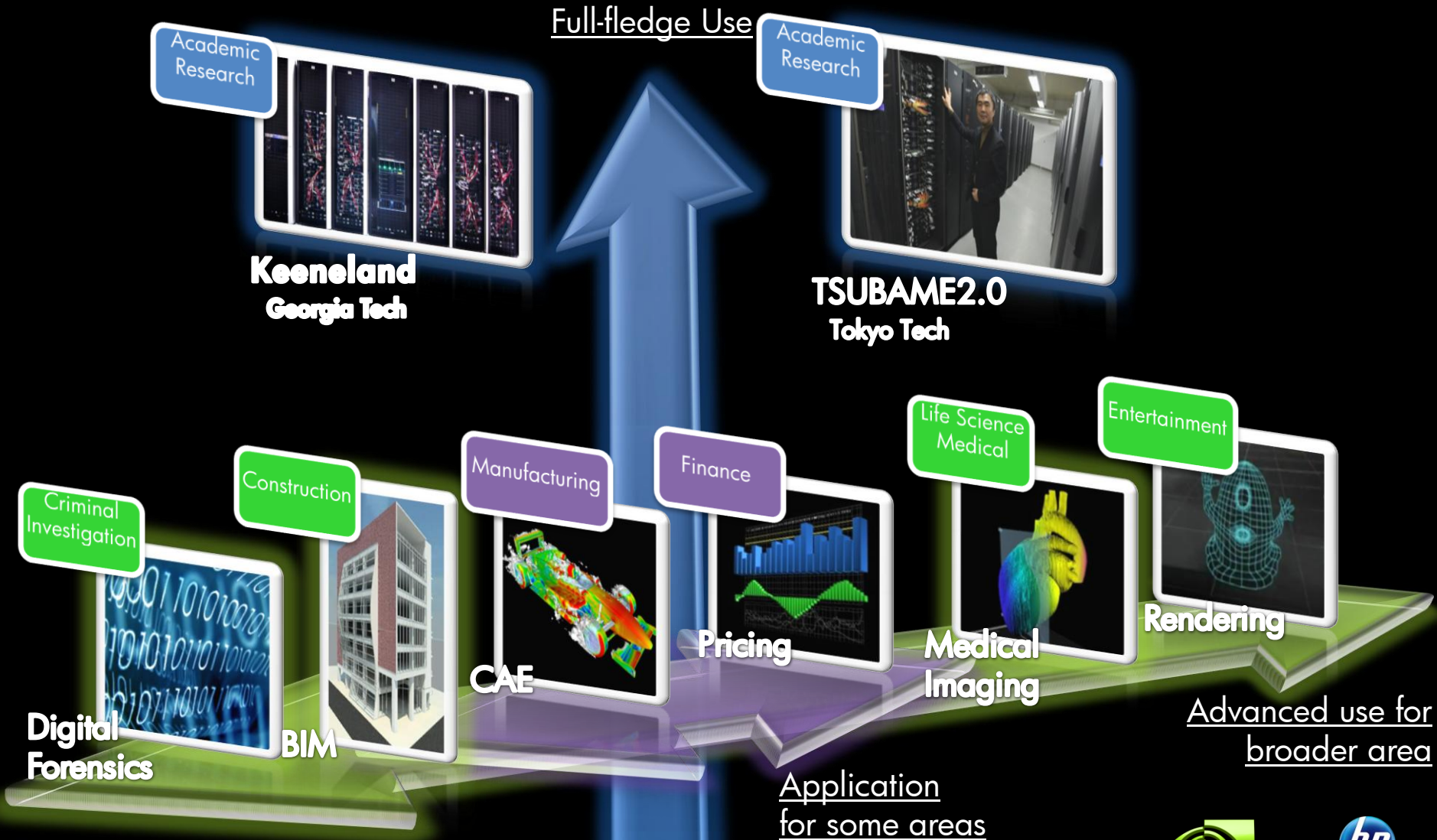
WORKSTATIONS

Woon Yung Chung 鄭運永  
Segment Marketing Manager  
HP Asia Pacific Workstations

Dec 14~ 15, 2011. Beijing



# Full-fledged Use Increase, Application Area Expand



## Square Enix

- SQUARE ENIX is one of the world largest game planning, development, publishing company in Japan.
- This company has successfully establishing the work flow to develop their HD quality interactive game with low cost. For this low cost with highest quality game development workflow, they are developing **GPU based Global Illumination renderer for HP Z800 with Tesla**. Mr. Eiji Fujii, Development Director and Mr. Shinji Ogaki, Senior Architect in R&D section will have a session in HP NVIDIA GPU Computing Day in Tokyo and talk about their success story with Z800 with Tesla.



# Digital Forensics - Criminal Investigation

Government

Contribution to  
Society



- Digital forensics for criminal investigation

- Encrypted and lost data analysis for decryption and recovery

- HP Workstation Z800 with dual Tesla C2050 provides extreme computing power to analyze huge data sets



# Customer in Digital Forensics

Contribution to  
Society

Korea case, 2009, phase 1

Customer: Public Sector

Application: digital forensics

Business need: to decrypt password encryption

HP compute nodes: 220 unit of Z800 cluster w/ Nehalem + Tesla → total 220 TeraFLOPS

- Compute node: Z800 w/ dual 2.8Ghz QC CPU / 6x1Gb RAM / Tesla C1060 / fx380

Reason HP selected:

- HP is #1 h/w vendor w/ 42% of sites in TOP500.org in Nov 2008 list. Competitor A is a distant 4th at 3.8% of sites.
- HP Workstation has lower noise, advance heat management, ease of maintenance of Z800, compared to white boxes vendors.
- Included a low end nVidia fx380 in addition to a Tesla C1060 for compute node monitoring purpose. This helped to increase total GPU throughput using fx380 CUDA computing power. Competitors tried to save cost by adding a non-CUDA-enabled graphic card for compute node monitoring and didn't provide adding CUDA-enabled graphic cards.
- Strong market share of HP workstation in AP. 60%~75% unit market share in Korea.
- HP Korea team provide top attention to the customer to bring early evaluation unit to customer as demo units, before product launch.



# Customer in Digital Forensics

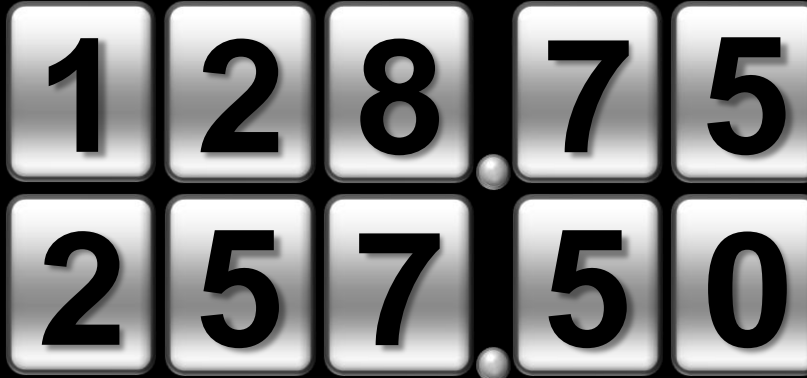
Contribution to Society

Korea case, 2010, phase 2

Customer:

Public Sector

GPU  
250set

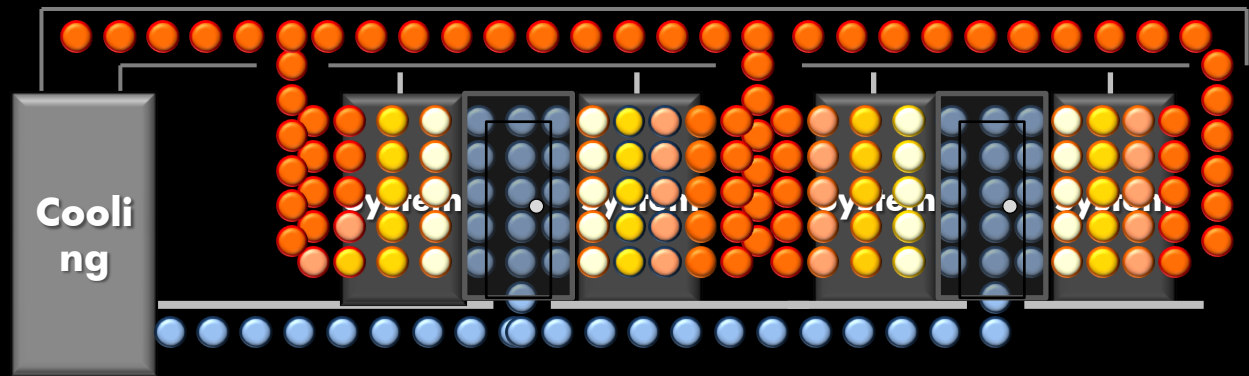
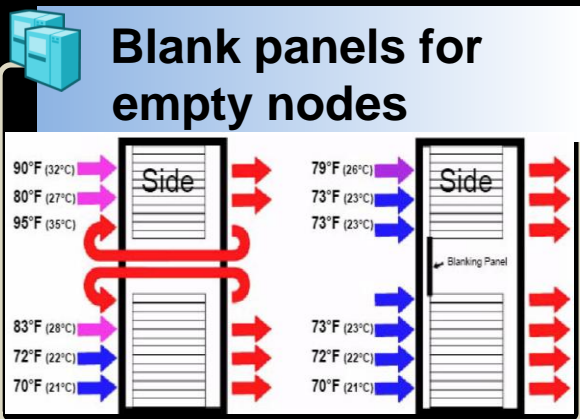


Double Precision  
515GFLOPS x 250

TFLOPS

Single Precision  
1030GFLOPS x 250

TFLOPS



Set up Cool Zone separation wall to maximize cooling efficiency

Blank Panel blocks the hot air circulation providing add'l cooling



# Integrated Monitoring

## Enables remote monitoring of compute nodes and racks



### Integrated Monitoring S/W, InnoSMMonitor

- Monitored targets

Compute Node	Rack
CPU, GPU temperature	Temperature
CPU, GPU Usage	Voltage, Current

- As customer adds second clusters of Z800 workstations, the monitoring s/w can be expanded to monitor overall integrated system.

### Integrated Monitoring S/W GUI

Customer is adding second cluster here.

The screenshot displays the InnoSMMonitor GUI. It features three main sections for Compute Node Groups: Group A, Group B, and Group C (additional). Each group contains a grid of CPU and GPU status indicators. A red dashed box highlights Group C, indicating it is a newly added cluster. On the right side, there are four monitoring panels: Overall Node Statistics, Individual Node Information, Rack Monitoring, and Power Wattage.

# Integrated Monitoring



## Added features of Integrated Monitoring S/W

The following features are added in 2010 to conveniently manage massive compute nodes. These s/w features flexibly manages any addition/changes to

**Per-Rack status mointo**

• Provides Intuitive Integrated View linked to physical location

**Integrated trouble**

• Realtime monitoring on troubled compute nodes. (Can rule out selectively specific troubled nodes in a job.)

• Easy to link with in-house code with API provided

**Note history mgmt**

• Per-node history mgmt of troubles, replacement records etc.  
• Enables browsing and mgmt of node history based on trouble-types

### Computing Node Group 1

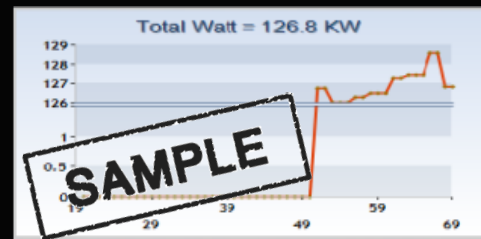


### Computing Node G1-107

Operated CN: 400  
CPU used: 354  
GPU used: 110  
CPU + GPU used: 96

CPU workload: 50 %  
GPU thermal 1: 38 °C  
GPU thermal 2: 12 °C

30.3	30.9	17.5	218.0	3815.0	01
30.6	34.9	18.1	224.0	4054.4	02
30.5	36.5	17.3	218.0	3815.0	07
30.6	34.9	18.1	224.0	4054.4	02
30.5	36.5	17.3	218.0	3815.0	07
30.6	34.9	18.1	224.0	4054.4	02
30.5	36.5	17.3	218.0	3815.0	07
30.6	34.9	18.1	224.0	4054.4	02
30.5	36.5	17.3	219.0	3861.7	07
30.6	34.9	18.1	224.0	4054.4	02
33.6	36.5	17.3	229.0	3961.7	03



# BIM (Building Information Modeling) - Construction

Reconstruction  
Assistance

Paper Less Studio Japan

3D building model visualization  
to improve quality and  
decrease environmental load

- To make construction schedule shorter and to achieve sustainable-design by several simulation
- NVIDIA and HP provides some HP Z800 Workstations with Tesla C2050 to accelerate Autodesk Revit as reconstruction assistance.

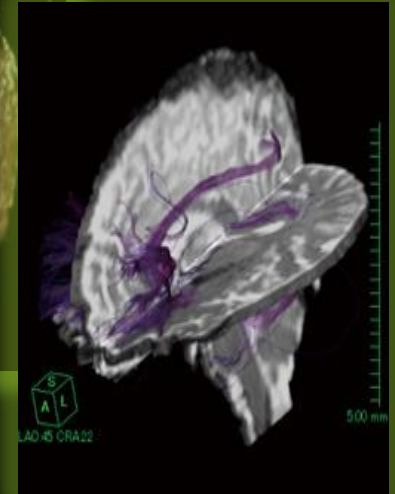


# Medical Imaging- Life Science/Medical AZE

Advanced  
Medical Tech



Highly advanced, high speed and sharpness  
medial imaging system



- Multiple analysis and evaluation  
to contribute comprehensible medical exam
- AZE's great application on HP Workstation  
with Tesla C2050 can achieve 20x rendering  
performance gain.



# HP Z WorkStations



	CPU	Graphics	GPU Compute
Z800	Up to dual 6-Core = 12	Up to 2 Quadro 6000	Up to 2 Tesla C2075
Z600	Up to dual 6-Core = 12	Up to 1 Quadro 5000 Up to 2 Quadro 2000	
Z400	Up to one 6-Core = 6	Up to 1 Quadro 5000 Up to 2 Quadro 2000	Up to 1 Tesla C2075
Z210	Up to one 4-Core = 4	Up to 1 Quadro 4000 (AMO)	
Z210 SFF	Up to one 4-Core = 4	Up to 1 Quadro 600	

Z400 – Tesla C2075 requires optional 600W power supply  
 Z800 – Dual Quadro 6000, or Tesla C2075 requires optional 1110W power supply

**Z800**  
Workstation

# Optimized, Separate Cooling System

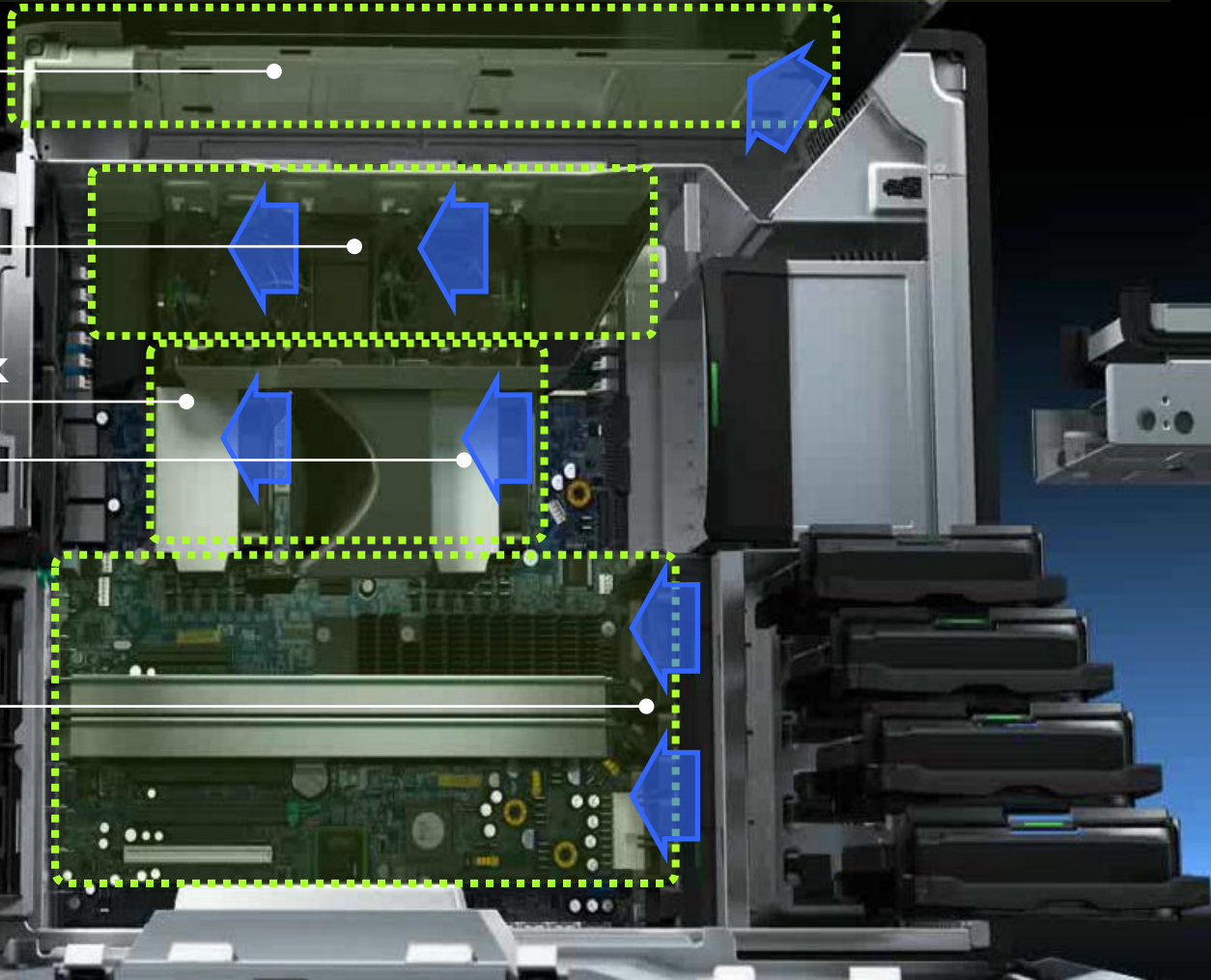
- Z800 supports the fastest Xeons at **3.46Ghz**/ 6 core or 3.60Ghz/ 4 core, also available with optional optimized cooling system.
- At the time, most 1U, 2U servers support only up to **2.93Ghz** Xeon 6 core CPUs, or 3.20Ghz/ 4 core due to limited cooling capability in 2010 Dec. **Now many servers support the same clock-cycle CPUs.**

**Power Supply**

**Fan for memory**

**CPU Active Heat sink**

**IO Device Fin  
For GPU**



# Z800 Workstation **Optional** Liquid Cooling Systems

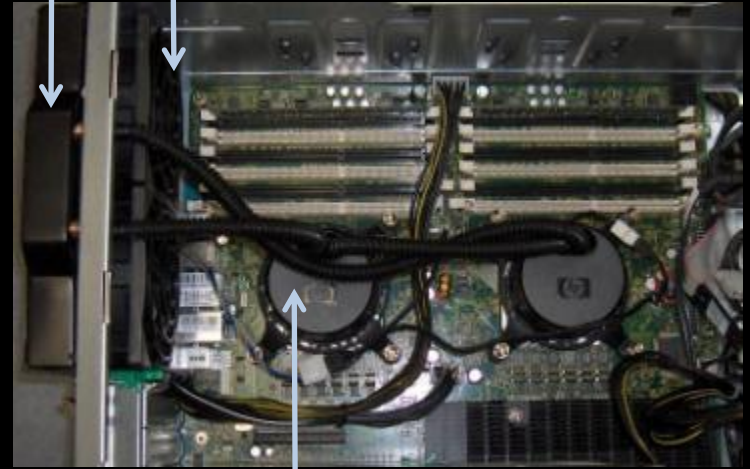
## HP Liquid Cooling Systems

- Because the CPU is usually the primary heat source in a system, that's where HP's liquid cooling solution is focused.
- HP's Liquid Cooling System (LCS) consists of a small reservoir, pump, and cold plate, as well as a radiator and fan.
- The LCS does not require servicing over the life of the workstation
  - Factory sealed, maintenance free
  - Minimizes noise-induced user fatigue and distraction
  - Enhances productivity
  - Available with HP Z400 and HP Z800 workstations

**Z800: Radiator mounted outside of chassis, effectively increasing cooling space.**

**Radiator**

**Cooling Fan**



**CPU cooling station with pump, reservoir and cold plate**

**Main Airflow Guide (under-side) with Liquid Cooling Duct Attached**



# OUR MOST **Mobile** Super Computing **MOBILE** WORKSTATION

New industrial design

Huron River architecture

Chassis redesigned for 16:9 panels

**NVIDIA Quadro 5010M** available

DreamColor display option, **1B colors**

DisplayPort 1.2 ready

Dual internal hard drives

7.9lbs (3.6kg) starting weight

720p HD camera

Redesigned keyboard

4 discrete function buttons

New travel batteries

Up to 4.5hrs battery life<sup>1</sup>



THE HP **ELITEBOOK**  
8760w

© Copyright 2011 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.



**NVIDIA**. WORKSTATIONS

**GPU** TECHNOLOGY  
CONFERENCE

# HP GPU Innovation with NVIDIA

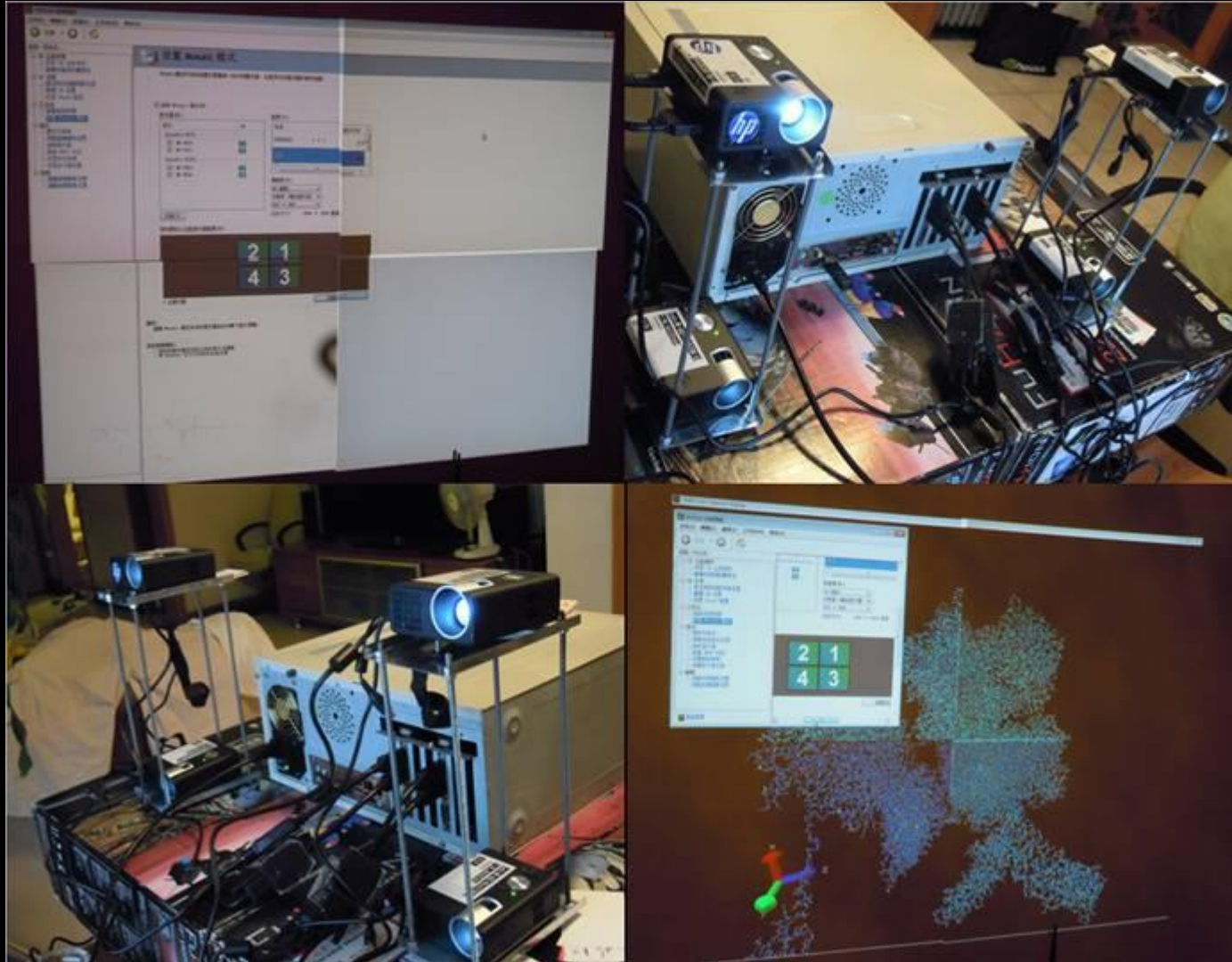


**nVIDIA**®

**Strong partnership to generate future innovation**

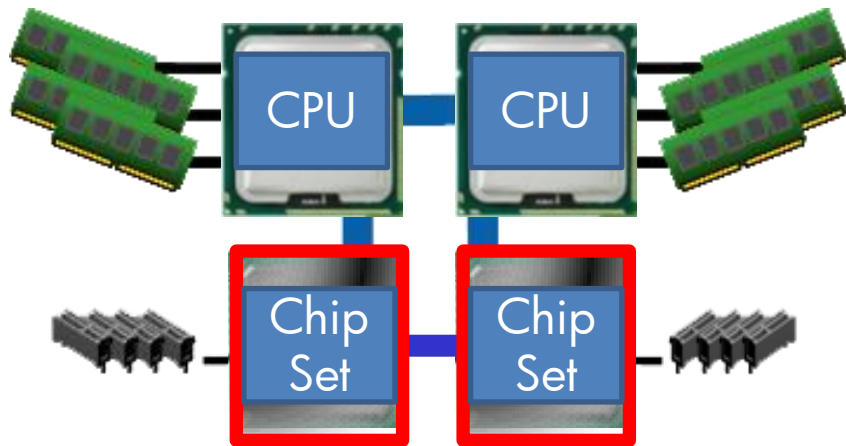
# Visit HP Booth !

Z800 Workstation with Tesla C2075 Tesla and QuadroPlex (Quadro 5000)



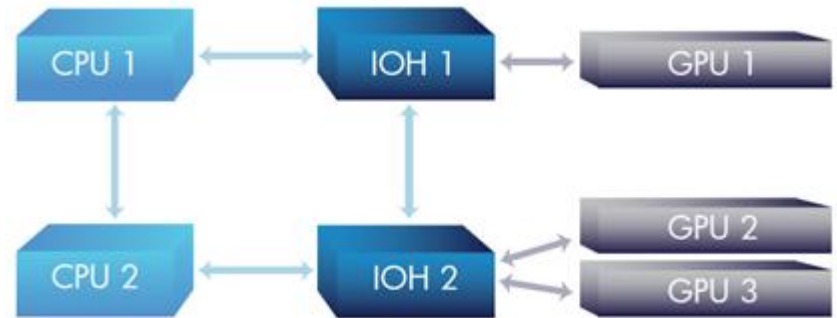
# Z800 Workstation Dual PCI-Express Gen2 x16

## Z800 Compute Node








- Z800 has two chipsets that can provide two full-speed PCI-E x16 slots. The two chipsets enable two Tesla cards to work at **full speed of PCI-E x16** bandwidth.

## Server based GPU Compute Node



- Some 2U server\* supports up to three M2070/M2090 Tesla cards. In maximum configurations, two of three GPUs needs to time-share the PCI-Express x16 channel.
- Data bandwidth per Tesla cards are shared btw the 2 Tesla cards, and is half as fast as Z800 case.
- Room temperature < 30° Celsius.
- Some 4U servers support up to 8 GPUs.

\* <http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA3-1993ENW.pdf>

For customers who use Windows HPC Server	Z800 as Compute Node	1U, 2U Server based Compute Node
GPU Card	<b>Tesla C series</b>	<b>Tesla M series</b>
Chipset	<b>2ea</b>	<b>2ea</b>
x16 PCI-E Gen2	 <b>Native 2 Slots</b>	<b>Native 2 Slots. Switching applied in 3 slots.</b>
Highest CPU bin	<b>Xeon X5690 – 3.46Ghz/6C/130W</b>	<b>Xeon X5690 – 3.46Ghz/6C/130W</b>
CPU Temperature Monitor	<b>○</b>	<b>○</b>
GPU Temperature Monitor	 <b>Windows: ○ Linux: ○</b>	<b>Windows: X Linux: ○.</b> (HP servers has sensors in PCI Express x16. IBM servers do not provide temperature sensors in PCI-Express slots.)
Supported OS	 <b>Win 7 (supports Win HPC Server cluster), Linux</b>	<b>Win 2008 HPC Server, Linux</b>
Cooling	 <b>Separate Cooling System for each component</b>	<b>Open structure</b>
Noise	 <b>&lt; 29dB</b>	<b>&gt; 50dB</b>