HPC with GPU

and its applications from Inspur

Haibo Xie, Ph.D

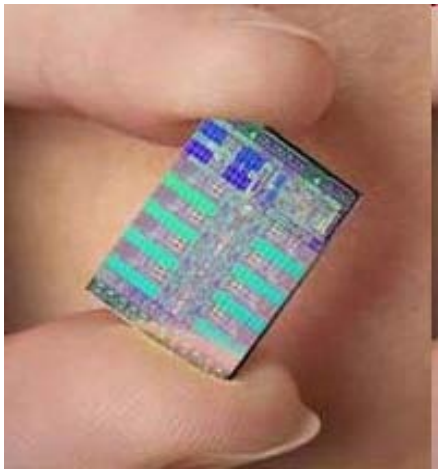
xiehb@inspur.com

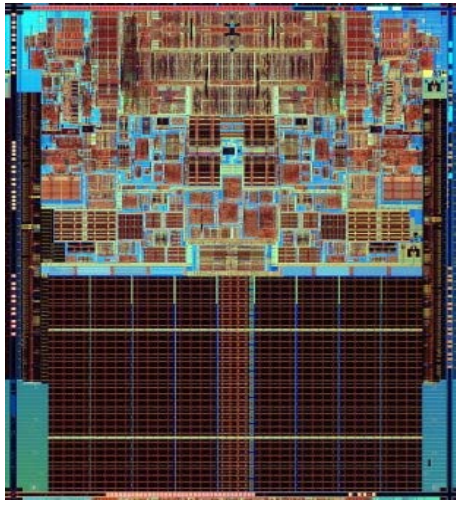
Agenda

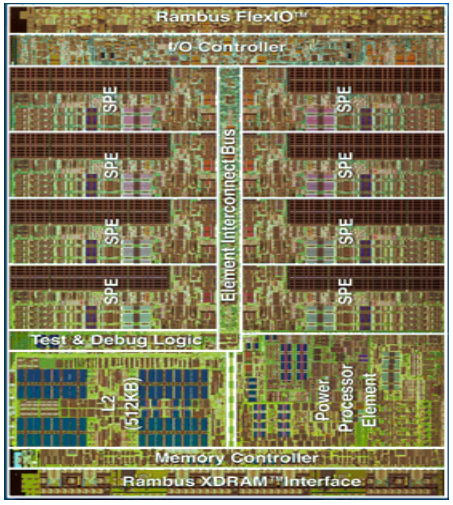
I. HPC with GPU

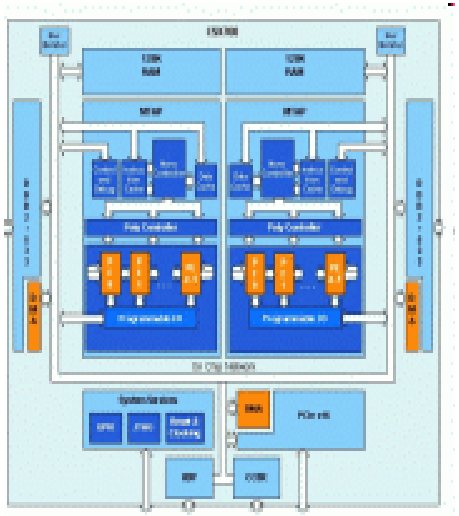
II. YITIAN solution and application

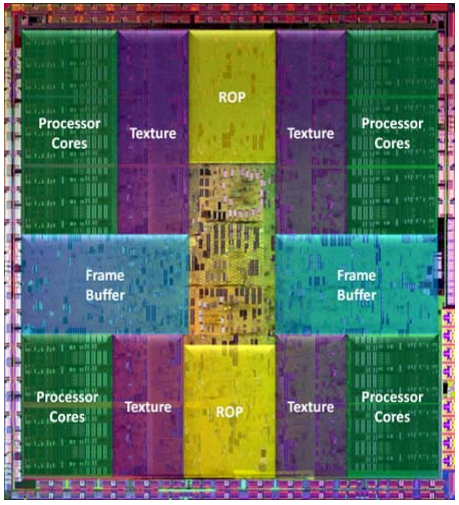
“New” Moore’s Law











inspur 浪潮

HPC？

- HPC stands for

  High performance Parallel Computing

  Heterogeneous Cluster Computing

  Hybrid/Cooperative Computing System

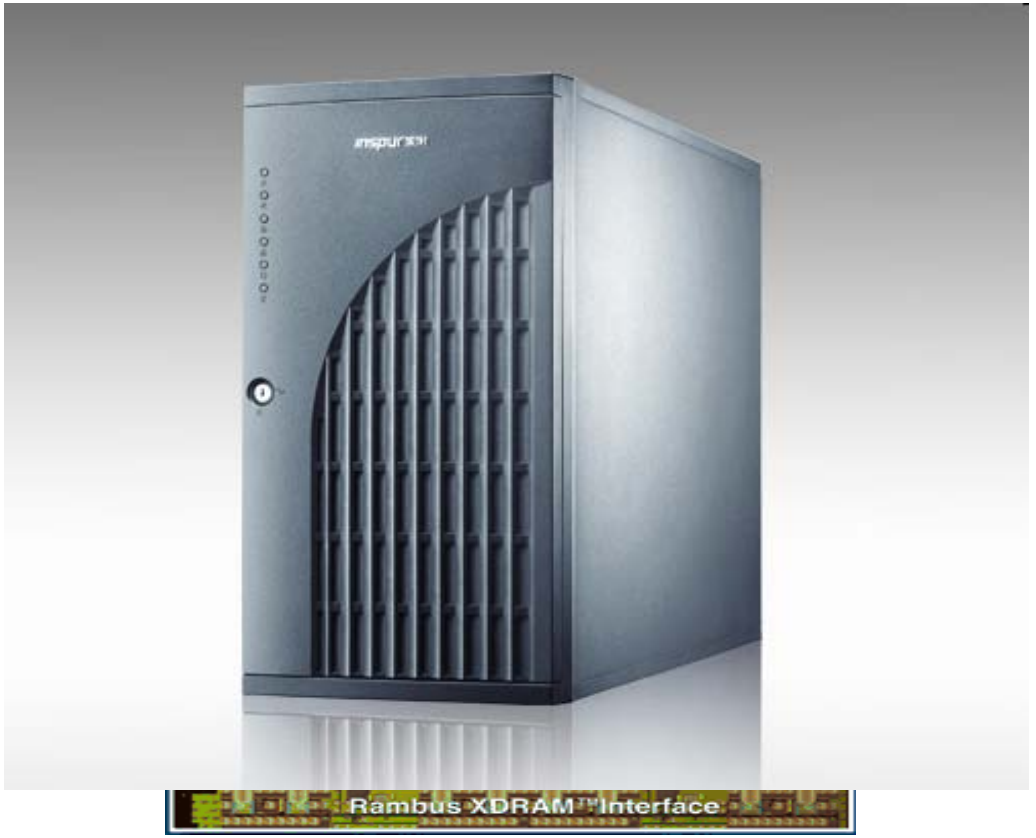
  Includes general and specialized processors/cores

  The future of HPC？

  TOP500：Roadrunner，TSUBAME，incoming supercomputer at ORNL

Enter Heterogeneous Parallel Computing

- Chip-level
- Desktop-level
- Cluster-level

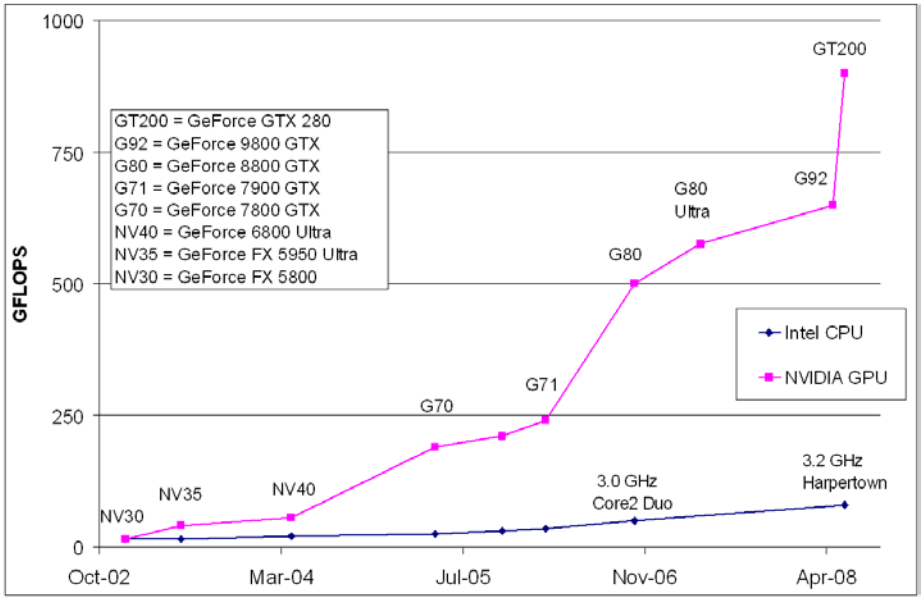


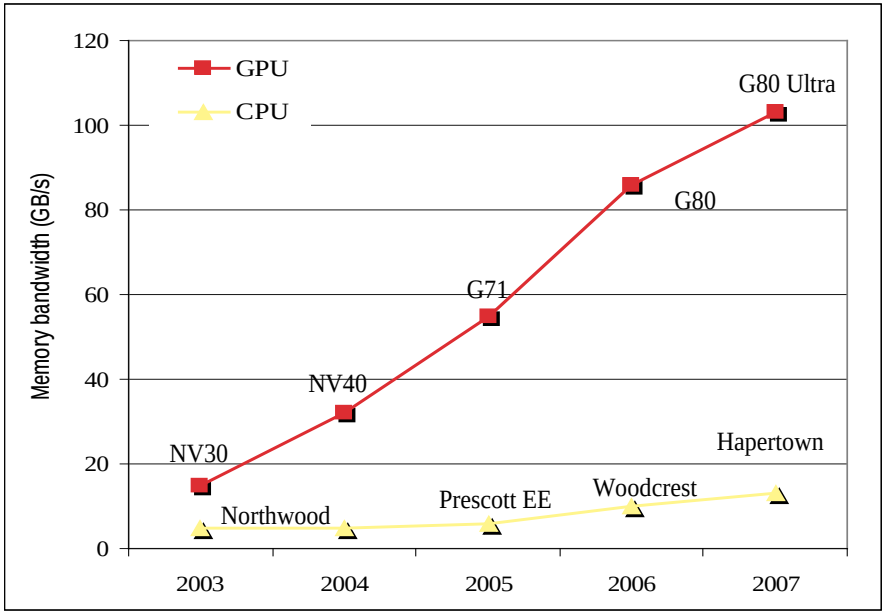
YITIAN with GPU HPC system

inspur 浪潮

Why GPU

- Many core with massive hardware thread
- Dedicated to computing
  - ✓ 1.3G * 240 * 3 ≈ 1TFLOPS peak （GT200）





inspur 浪潮

GPU for HPC, History, Opportunities and Challenges

GPGPU

- ✓ Initiative GPU for general purpose computing
- ✓ Terrible programmability
- ✓ Resident narrowly as desktop level application

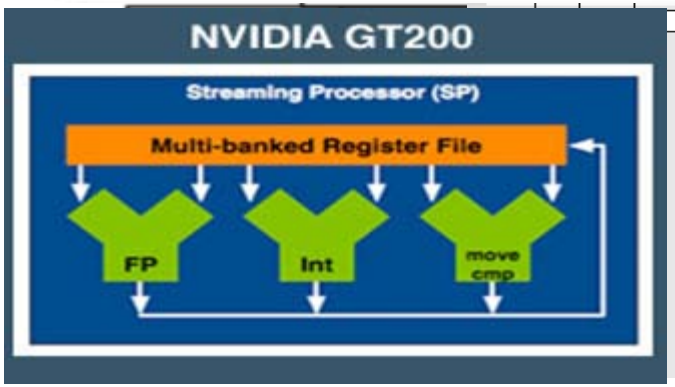
Opportunities

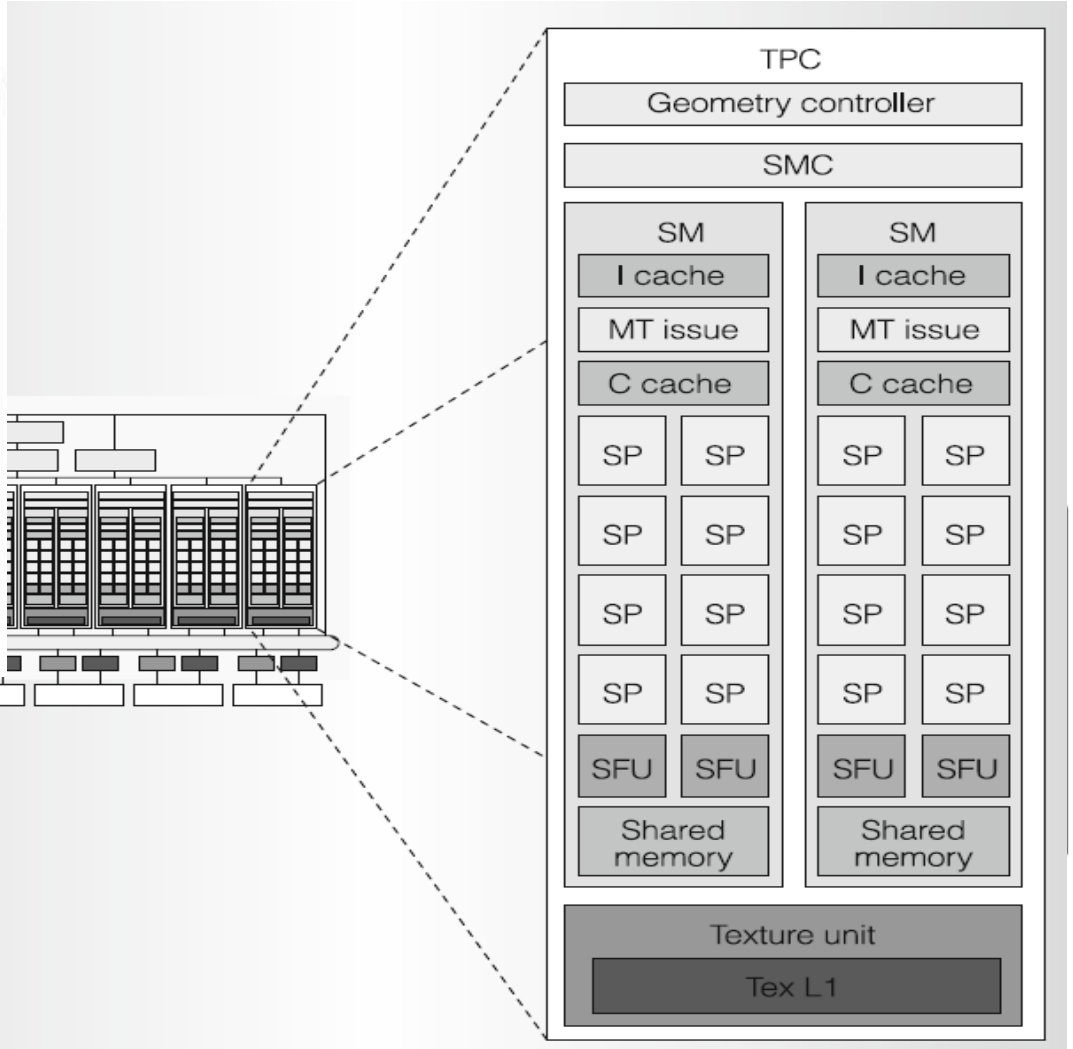
- ✓ Revolutionary architecture & programming model
  - Brook+? CUDA
- ✓ Ecosystem building-up is essential
  - Programmer wake up for chargeable lunch (Free lunch is over)
  - Low price, easy to get
  - Better perf/watt, perf/dollar
  - Tremendous R&D activities not only by academic, but only by commercial
- ✓ Targeting TOP500
  - Live product, rank 41, building by GSIC Center, Tokyo Institute of Technology

Challenges

- ✓ Programmability does matter!
- ✓ Mature ecosystem is essential
- ✓ Various architecture, lack of standard
  - openCL, way to fix that?

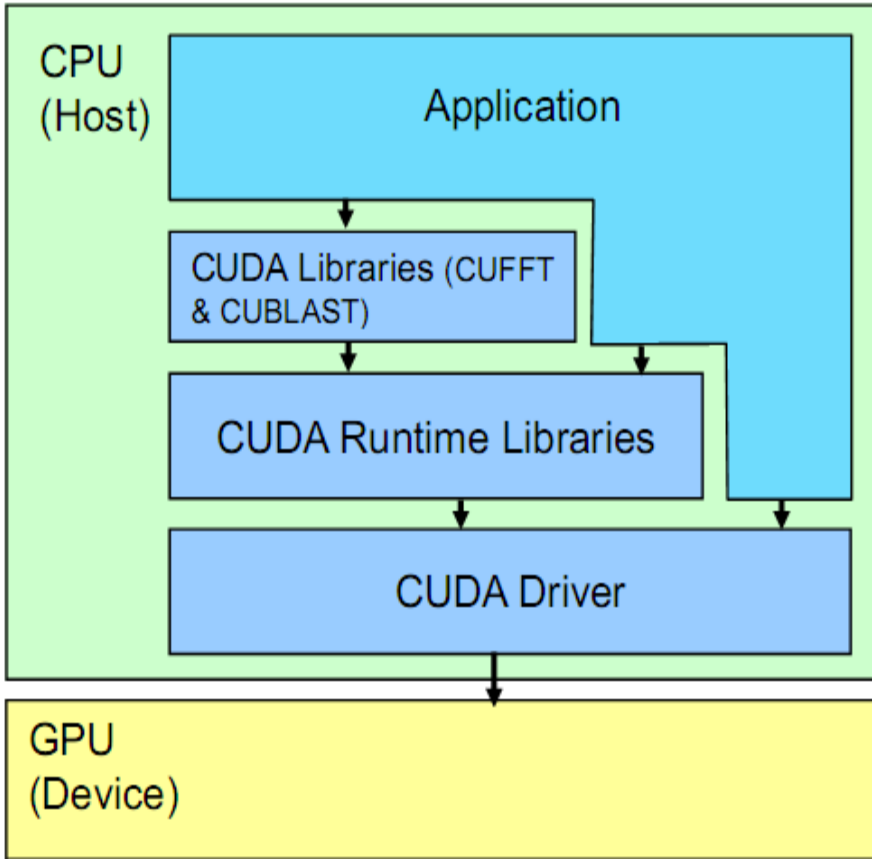
Hardware – GPU as computation device





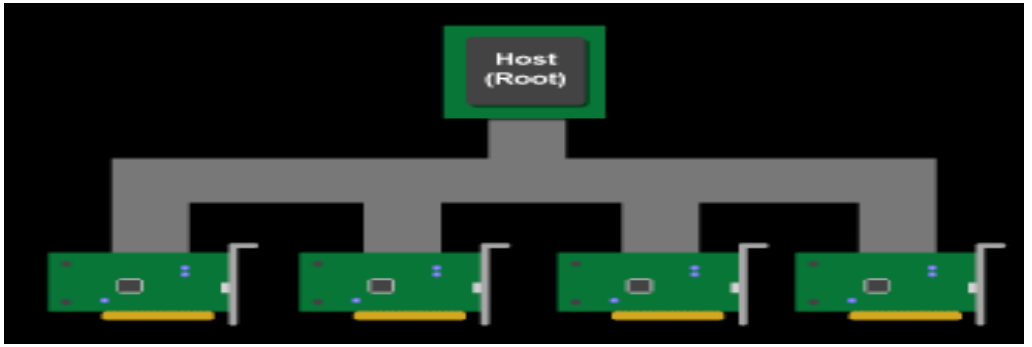
Programmability – Enter CUDA

- CUDA: Compute Unified Device Architecture
- Programming Model
  - ✓ Nvidia GPU based
  - ✓ Mapping between the hardware and the software
- SDK
  - ✓ Tool-chain （compiler, debugger, profiler）
  - ✓ Runtime
  - ✓ Language support
  - ✓ Library （BLAS, FFT, etc.）
- Thoughts about Heterogeneous/Hybrid computing
  - ✓ CPU+GPU
  - ✓ GPU as co-processor



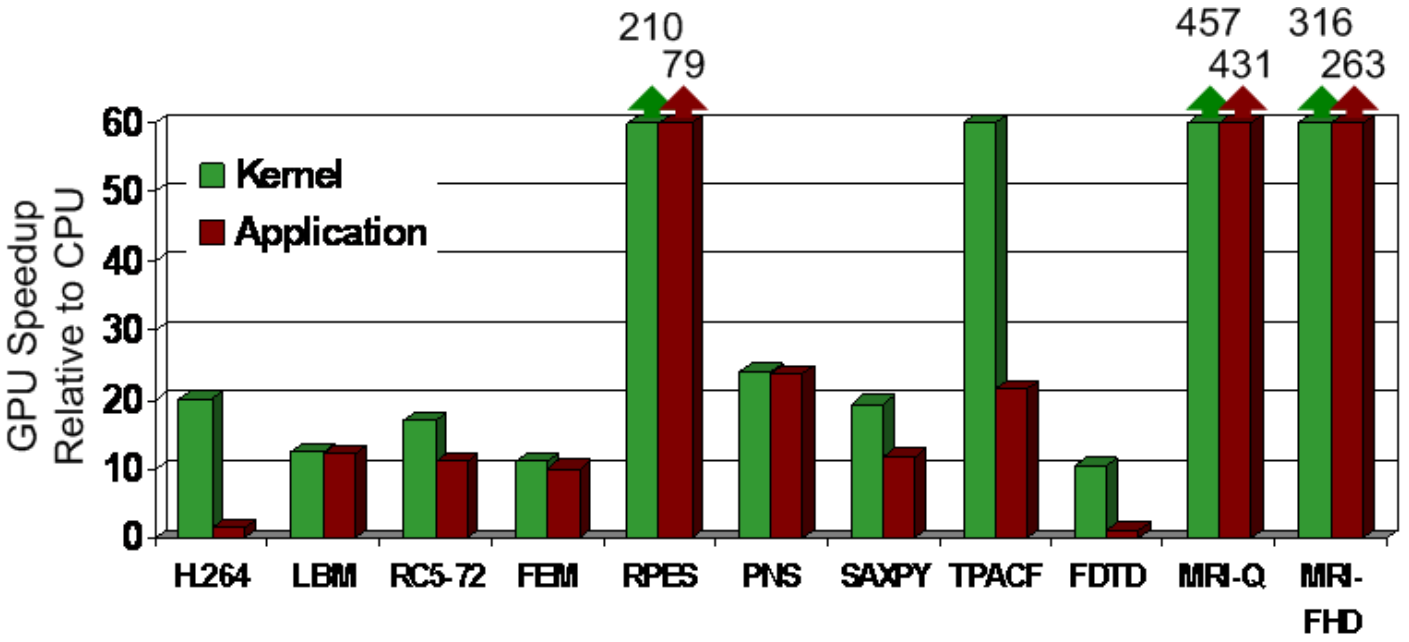
Heterogeneous Computing with CUDA

- CPU + GPU
  - ✓ MPI + CUDA?
  - ✓ OpenMP + CUDA?
  - ✓ Pthread + CUDA?



- Cooperative computing
  - ✓ CPU for general purpose workloads
  - ✓ GPU for offloaded specialized computational-sensitive workloads to GPU

CUDA application speedup



- GeForce 8800 GTX vs. 2.2GHz Opteron 248
- $10_{\times}$ speedup in a kernel is typical, as long as the kernel can occupy enough parallel threads
- $25_{\times}$ to $400_{\times}$ speedup if the function's data requirements and control flow suit the GPU and the application is optimized

Agenda

I. Introduction

II. YITIAN solution and application

“YITIAN” Solution

Personal HPC Solution

- ✓ “YITIAN” desktop supercomputing

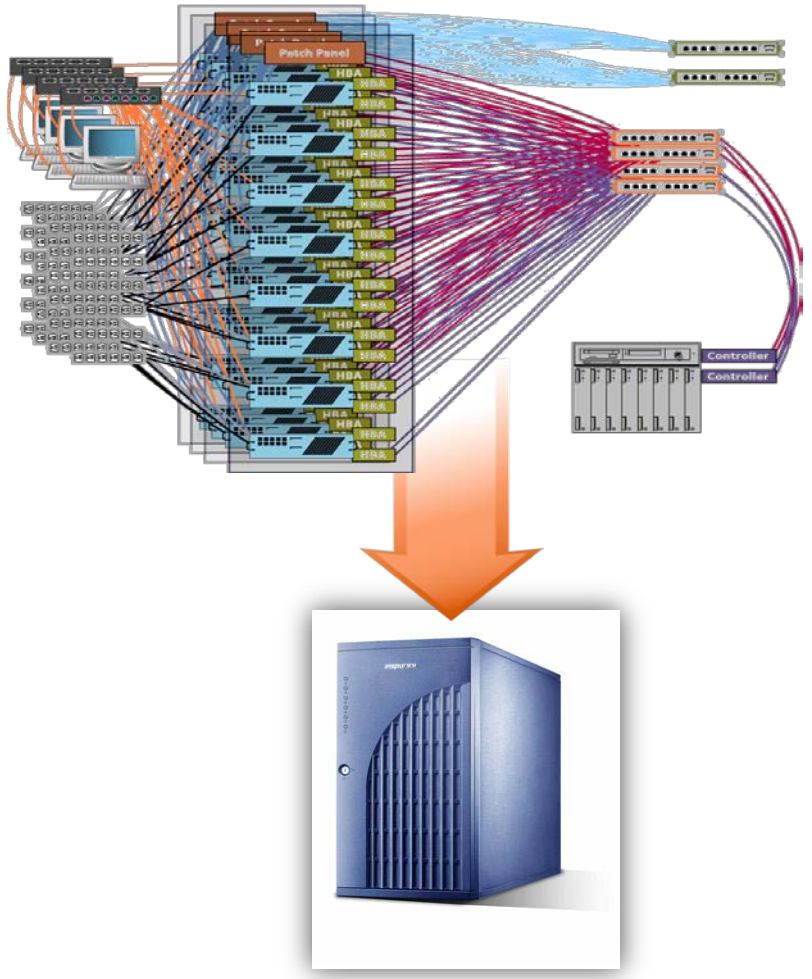
Industry Solution

- ✓ “TIANSUO 10000” with “YITIAN” Desktop HPC or “YITIAN” Tesla HPC cluster

inspur 浪潮

“YITIAN” desktop HPC

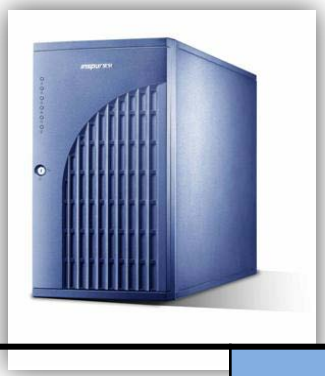
- **Personal HPC**
  Targeting personal HPC application
- **Cooperating computing**
  CPU + GPU Hybrid computing
- **Software stack**
  Mature ecosystem including R&D, software and hardware
- **Silencing and cooling**
  Emphasis on silencing and cooling



inspur 浪潮

More power, lower price – "YITIAN" desktop HPC

"YITIAN" personal HPC v.s. traditional cluster



VS

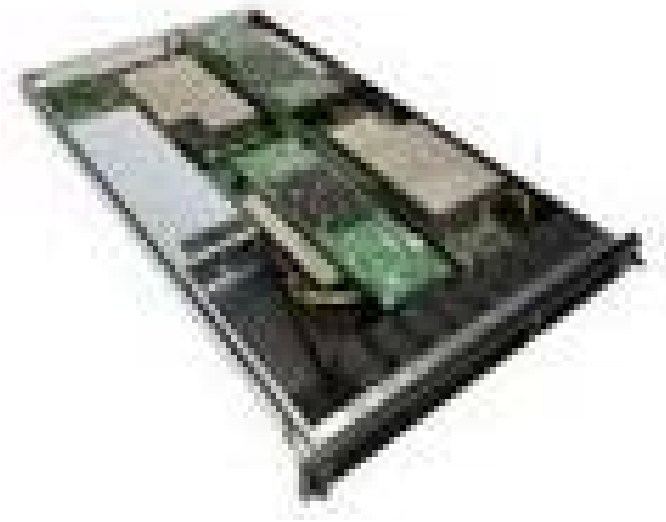
	"YITIAN" personal HPC	Traditional X86 Cluster
Performance	Tera-level	Tera-level
Size	Regular desktop	42U rack
Room requirement	None, regular office	Server room
Cooling requirement	None	Air conditioning
Power consumption	700W peak	7000w
Noise	22~45db	Above 60db
Price	RMB 50,000	RMB 500,000

inspur 浪潮

Need even more power? – "YITIAN" Tesla HPC cluster

"TIANSUO 10000"

- ✓ Flexible cluster system
- ✓ Consist of "YITIAN" Tesla HPC cluster
- ✓ Scalability with GPU acceleration



"YITIAN" Tesla HPC cluster

- ✓ S1070 with 4 C1060
- ✓ 4TFLOPS,16GB memory

Industry solution

- ✓ Supply more power as the foundation of your application

inspur 浪潮

“YITIAN” market

Market

Life science　　Graphics　Finance　　Engineering science

Computational Fluid Dynamic　Defense　Medical Imaging

Oil & Gas　Virtualization　Electronic design automation

Research、 Academic、 Chemicals & petroleum、 Finance

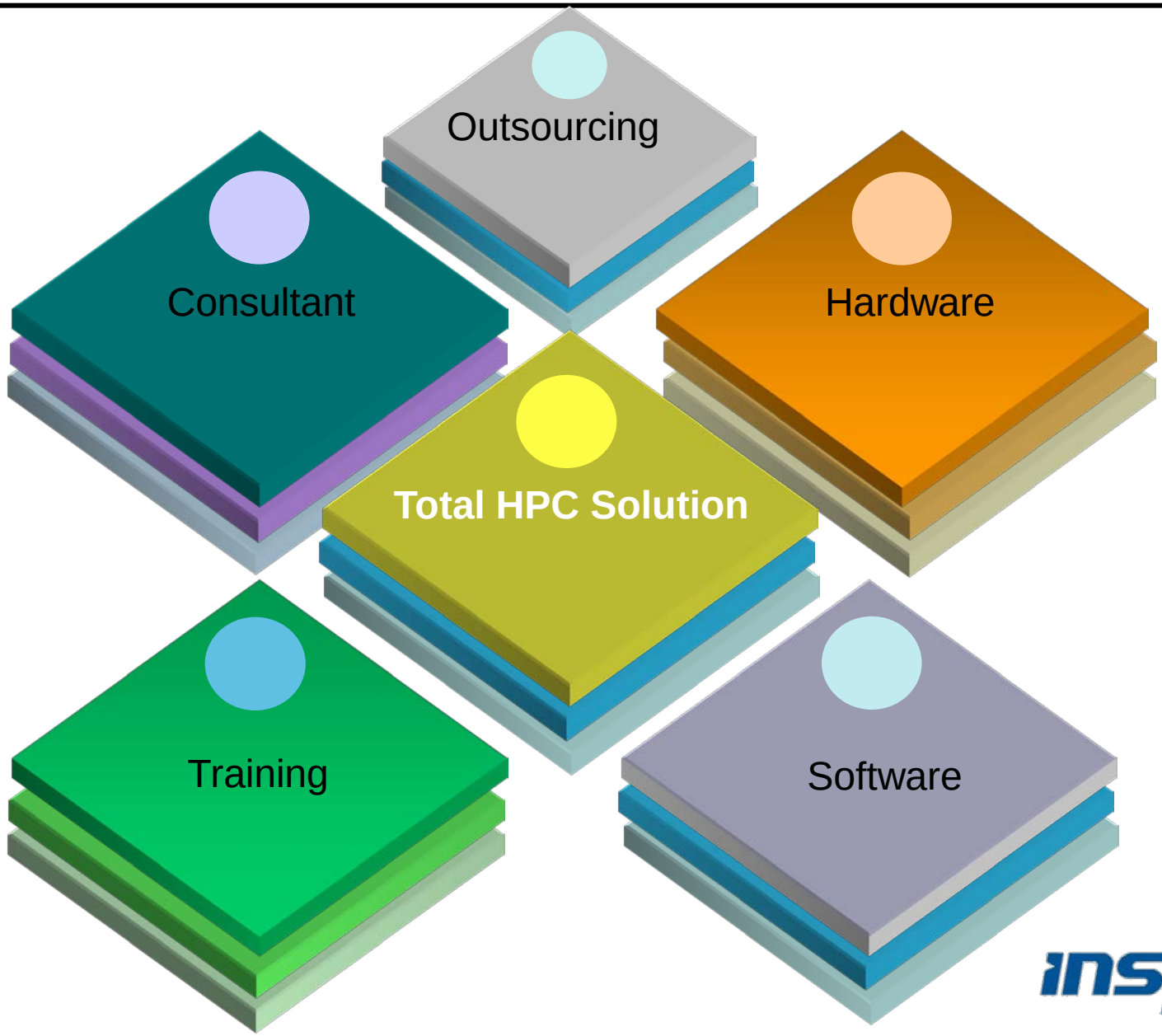
inspur 浪潮

Applications of "YITIAN"

Customer	Applications	Description
State key laboratory of information security, CAS	Information security, cryptography	Port application from CPU to GPU Typical 10x, up to 100x times speedup
Computer network information center, CAS	Network security	Multi-pattern matching algorithm for network security. 2.5 times speed up
State key laboratory of space weather, CAS	meteorology	N/A
National University of defense technology	Radar system and electromagnetic countermeasure	Application porting on-going
Beijing Institute of Genomics, CAS	Bio-information	BLASTN algorithm acceleration, Joint development with Inspur, paper published at HPCChina 2009
XI'AN jiaotong University	CFD	N/A
JINAN University	Bio-information	
54-th institute, CETC	N/A	

inspur 浪潮

Inspur offers



inspur 浪潮

Thanks !

inspur 浪潮