



# Integration, Development and Results of the 500 TeraFlop Heterogeneous Cluster (*Condor*)

September 2011

Mark Barnell  
Air Force Research Laboratory

*Integrity ★ Service ★ Excellence*



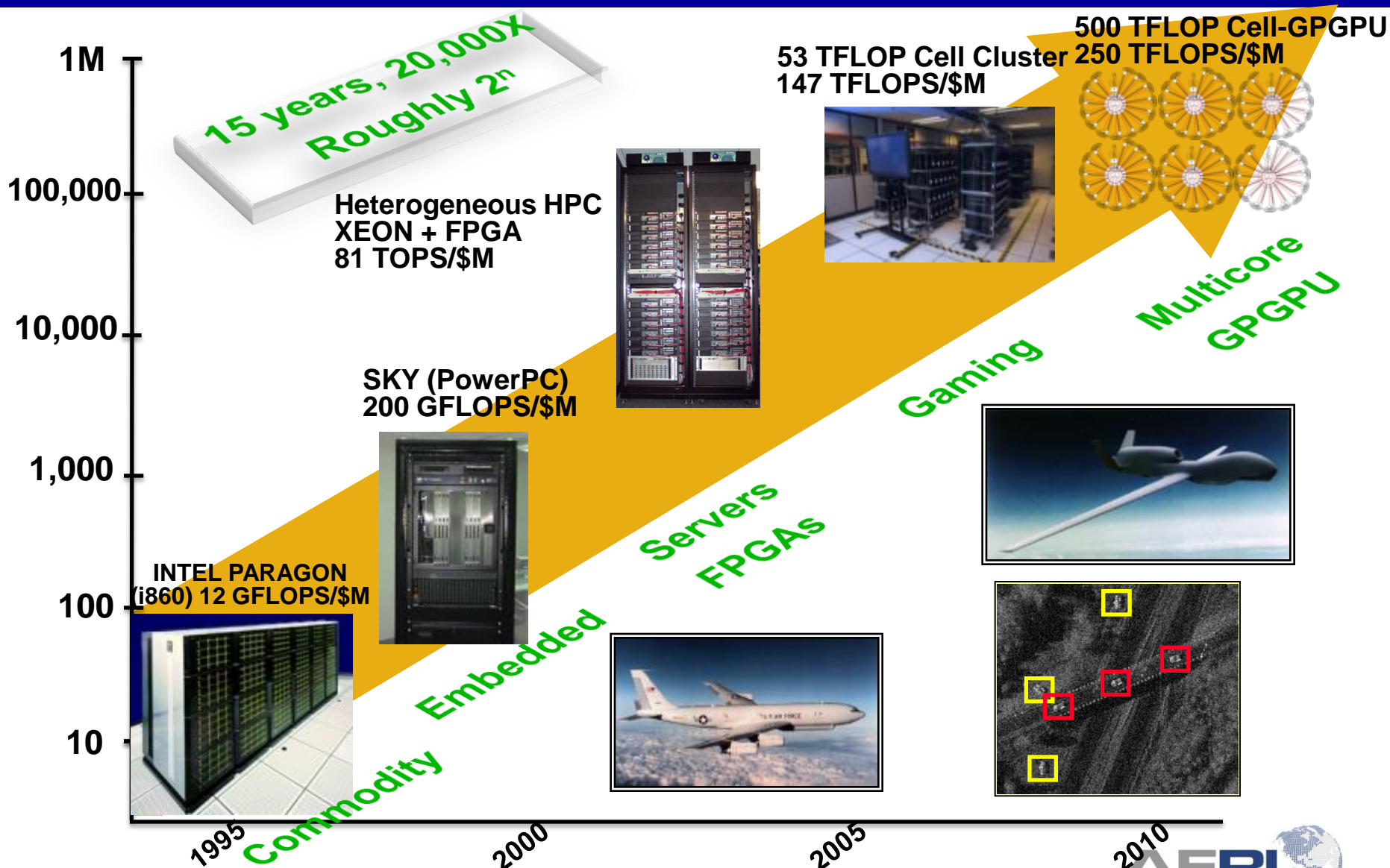
# Agenda



- **Mission**
- **RI HPC-ARC & HPC Systems**
- **Condor Cluster**
- **Success and Results**
- **Future Work**
- **Conclusions**



# Exponentially Improving Price-Performance Measured by AFRL-Rome HPCs





# Mission



- **Objective:** To support CS&E R&D along with HPC to the Field experiments by providing interactive access to hardware, software and user services with special attention to applications and missions supporting C4ISR.
- **Technical Mission:** Provide classical and *unique, real-time, interactive* HPC resources to the AF and DoD R&D community.



# HPC Facility Resources



**Cell BE Cluster** *53 TFLOPS Peak Performance*

**EMULAB** *Network Emulation Testbed*

**Legend:**

- HPC SDREN Assets April 2011
- HPC Assets on HPC DREN Network

**HORUS** *22TFLOPS TTCP Field Experiments*

**CONDOR CLUSTER** *Urban Surveillance Cognitive Computing Quantum Computing*  
 Funding: \$2M HPCMP DHPI

**Online: Nov 2010**



# HPC Facility Resources GPGPU Clusters



**ATI Cluster** 32 TFLOPS ATI  
FirePro 8800

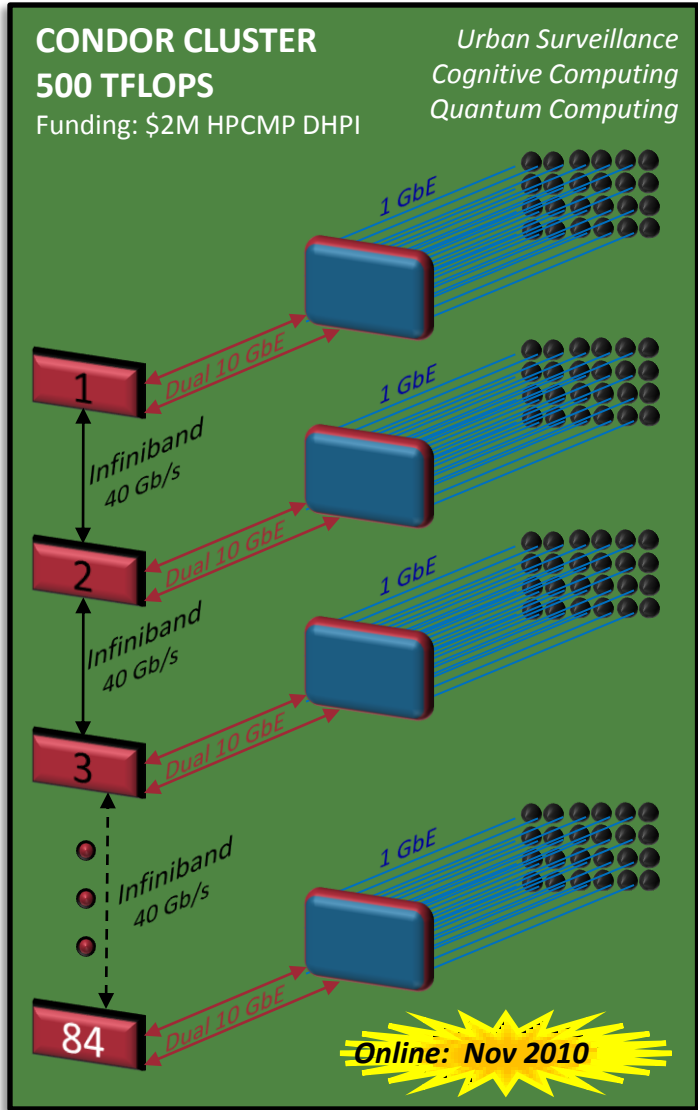


**Online: Jan 2011**

**Legend:**

HPC GPGPU Assets  
on DREN Network

**HORUS** 22TFLOPS TTCP  
Field Experiments

- Upgrade all Nvidia GPGPUs to C2050s & C2070s Tesla cards August 2011
- 200 Kepler cards ~400K will have a 3x improvement (3Tflop DP) 220W
  - Condor one of the greenest HPC in the world (1.25 Gflop/W DP&SP)
- Redistribute 100 Fermi Tesla cards to other HPC and research sites
  - NE Univ, UMASS, ARSC, MIT/LL, AFRL/RI, AFRL/RV





# The Condor Cluster



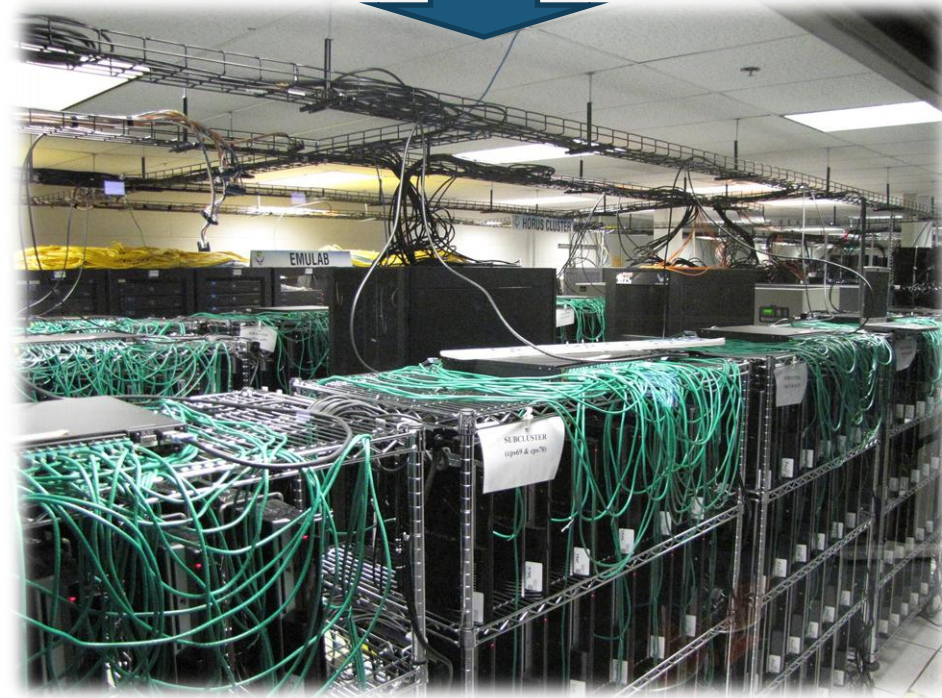
**FY10 DHPI**

Key design considerations: **Price/performance & Performance/Watt**



## 1716 SONY Playstation3s

- STI Cell Broadband Engine
  - PowerPC PPE
  - 6 SPEs
  - 256 MB RAM



## 84 head nodes

- 6 gateway access points
- 78 compute nodes
  - Intel Xeon X5650 dual-socket hexa-core
  - (2) NVIDIA Tesla GPGPUs
    - 49 nodes – (98) C2050
    - 15 nodes – (30) C2070
    - 14 nodes – (28) C1060
  - 24 GB RAM (\*48GB)

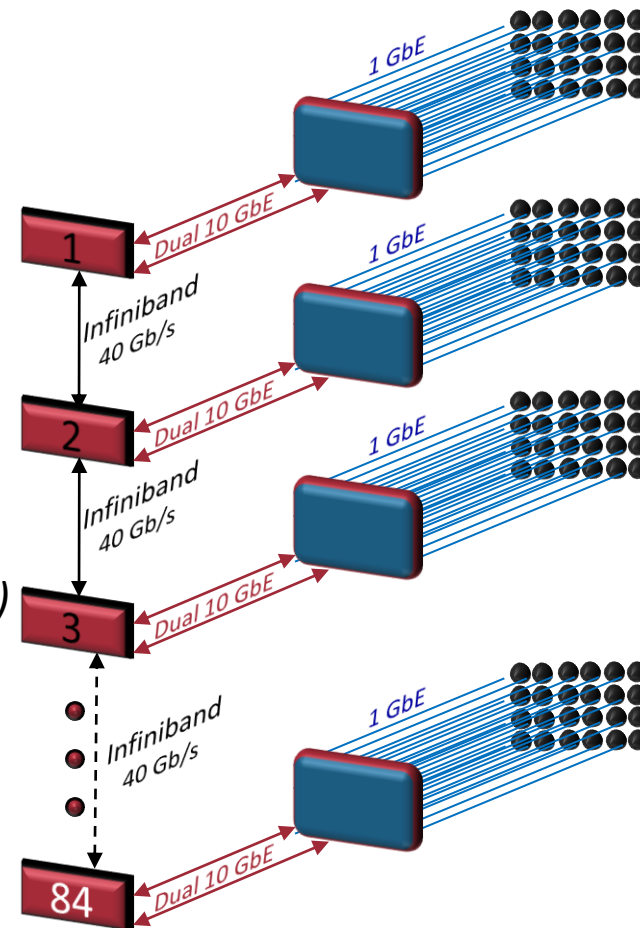


# Condor Cluster (500 Tflops)



Online: November 2010

- 263 Tflops from 1,716 PS3s
  - 153 GFLOPS/PS3
  - 78 subclusters of 22 PS3s
- 230 Tflops from server nodes
  - 84 sever nodes (Intel Westmere 5650 dual socket Hexa (12 cores))
  - Dual GPGPUs in 78 server nodes
  - Horus Cluster (~22 Tflops 10 nodes, 18 C2050s)
- Cost: Approx. \$2M



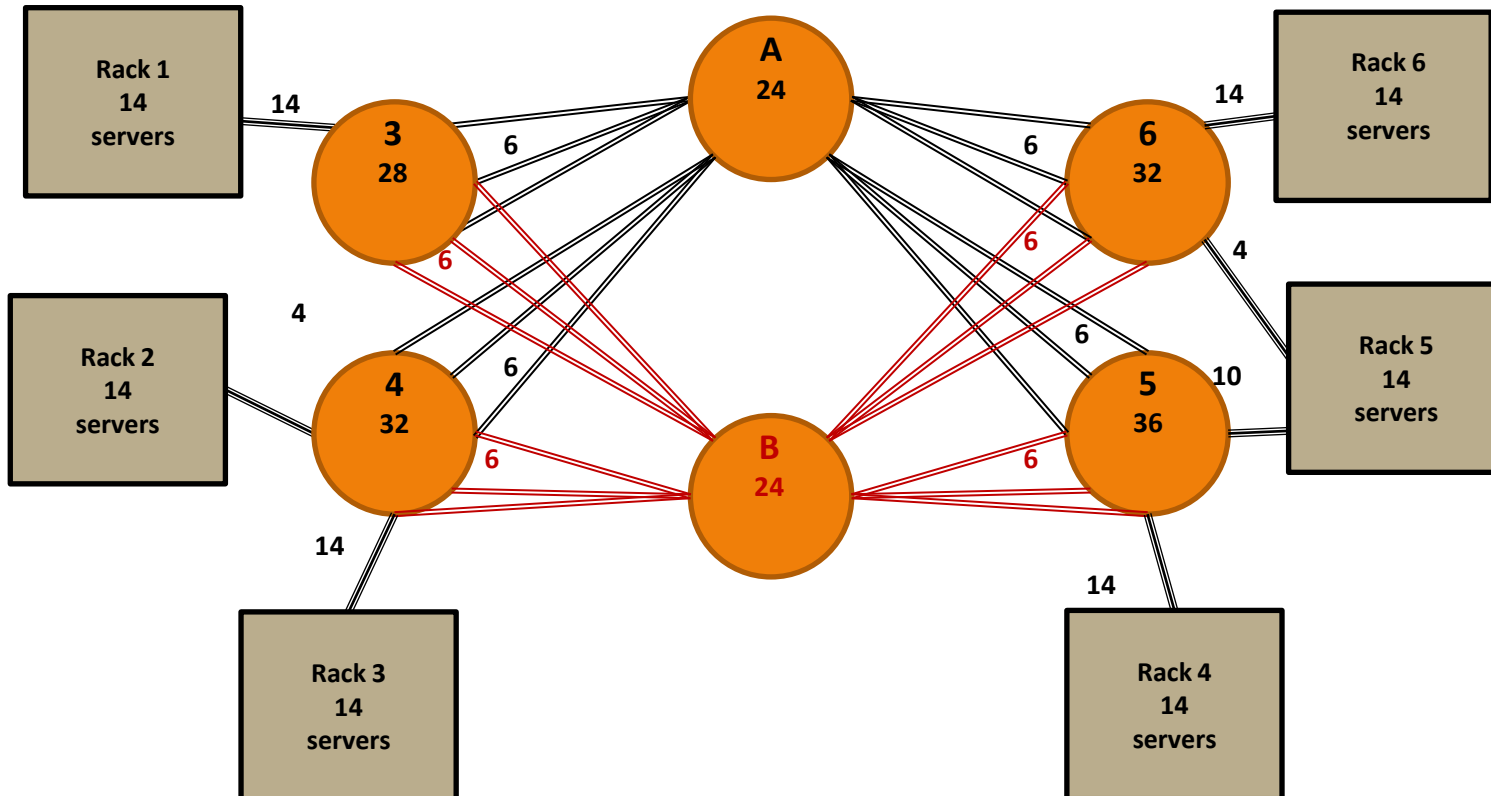
Sustained Throughput Benchmarks YTD: Xeon X5650: 16.8 Tflops, Cell 171.6 Tflops, C2050 : 60.2 Tflops, C1060: 23 Tflops....**CONDOR TOTAL 271.6 Tflops**



# Condor Cluster Networks

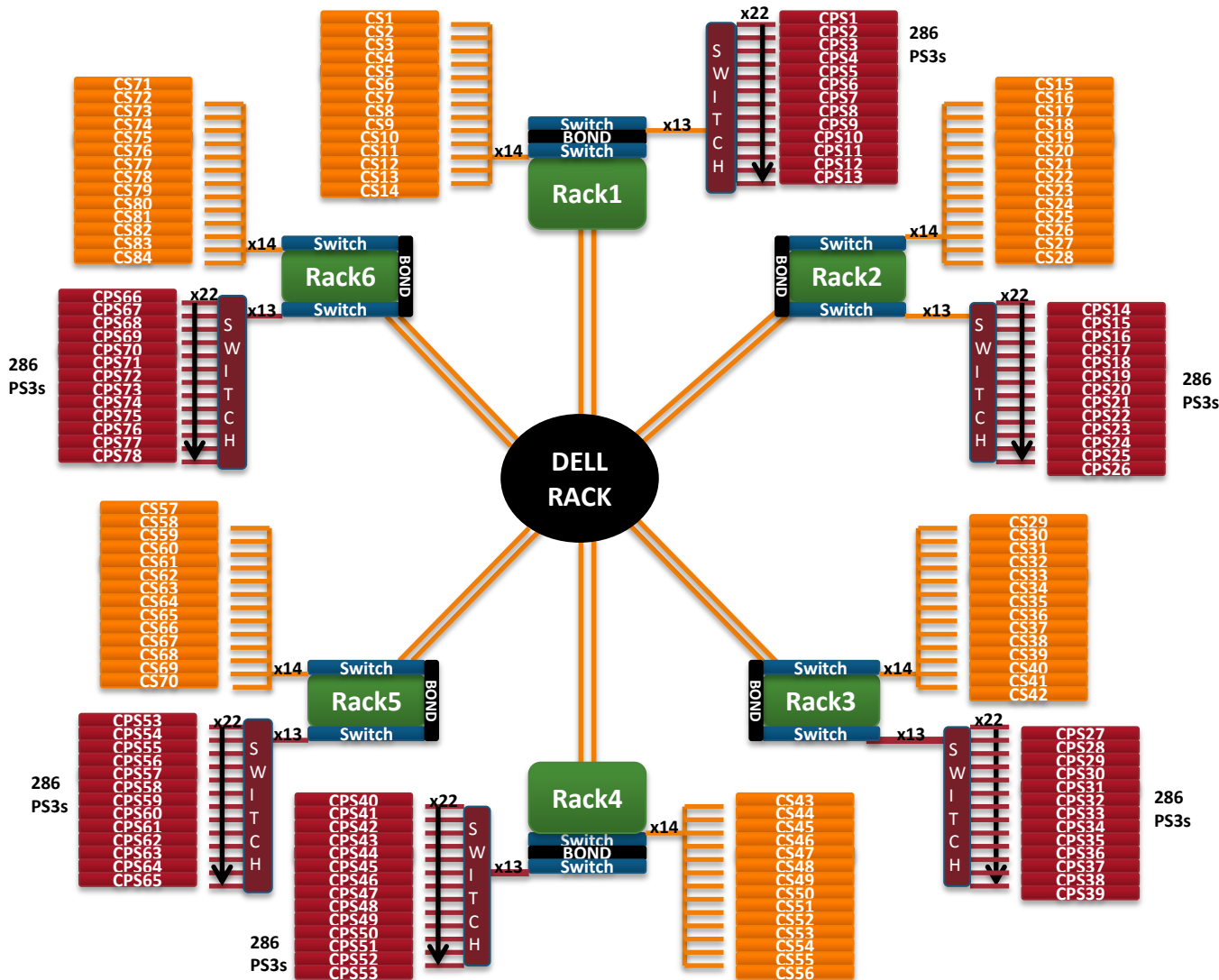


**Infiniband Mesh Non-Blocking 20Gb/s  
(5) 12200 & (1) 12300 Qlogic 40Gb/s Infiniband (36 port) switches**



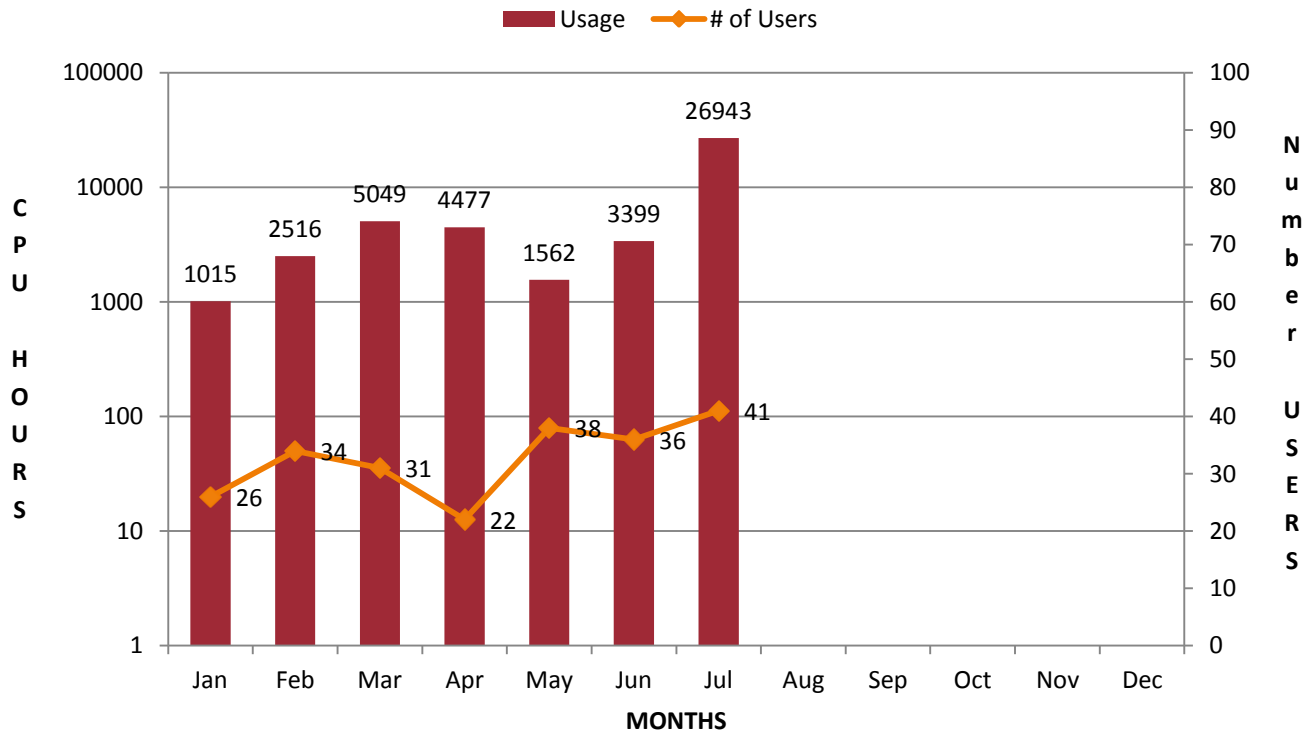


# 10GbE Star-Hub





# Condor System CPU Hours



Max Interactive CPU Hours:  $78 \text{ Nodes} * (12+2\text{GPUs}) * 12 \text{ (Hrs)} * 21 \text{ (days)} = 380016 \text{ Hours}$   
 JULY 7%

Total Interactive Login Time:  $84 \text{ Nodes} * 24 \text{ Hrs} * 21(\text{days}) = 42336 \text{ Hours}$   
 15,445 Login Hours JULY 36.5%





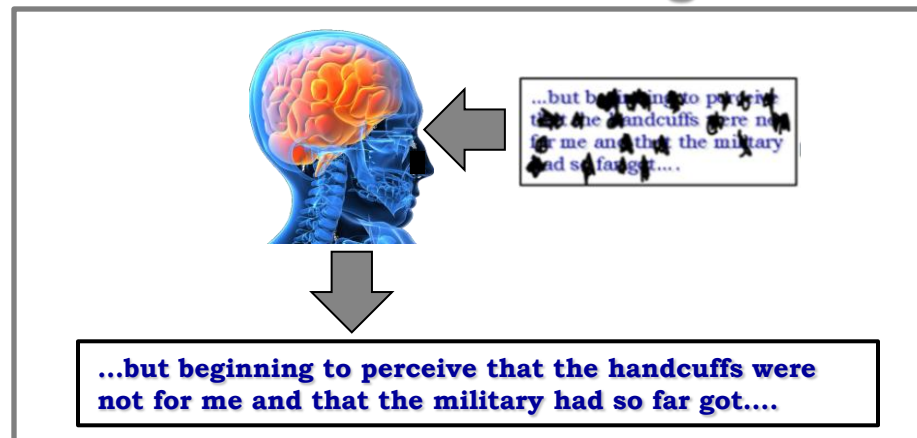
# Solving Demanding, Real-Time Military Problems



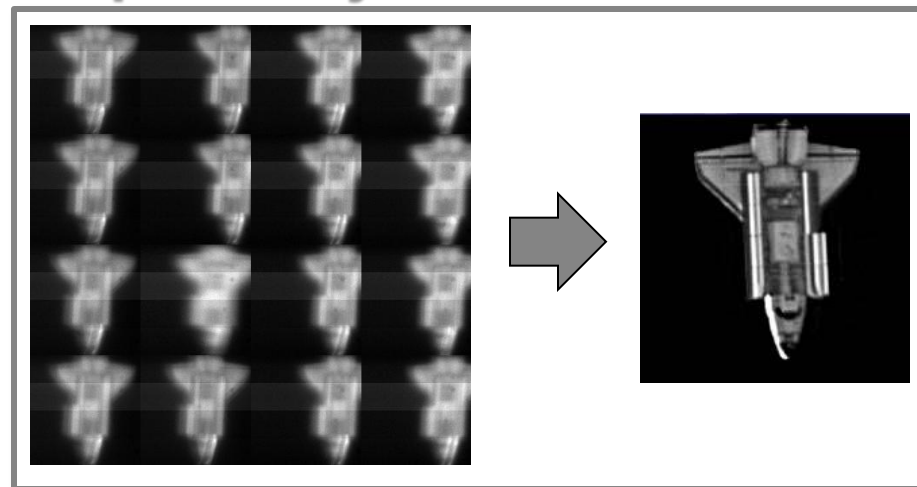
## Radar processing for high resolution images



## Occluded text recognition



## Space object identification

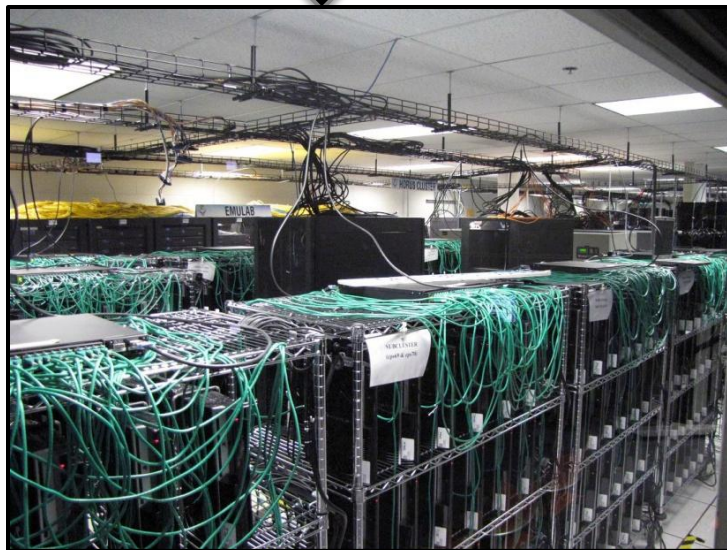




# RADAR Data Processing for High Resolution Images



Radar processing for high resolution images in real-time





# Optical Text Recognition Processing Performance



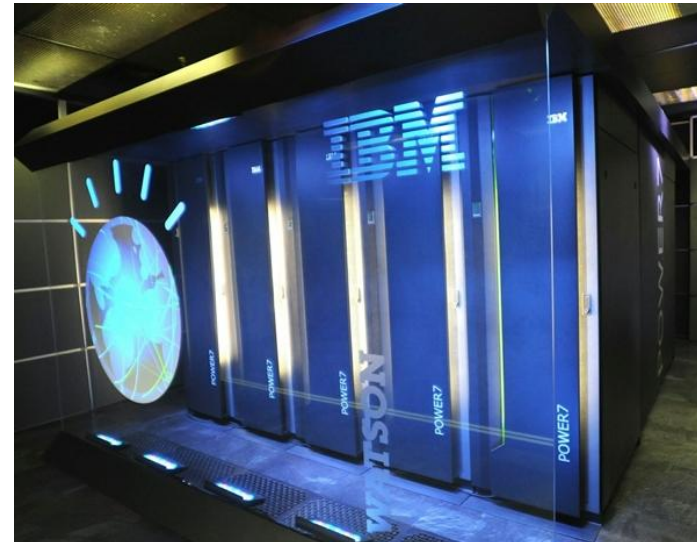
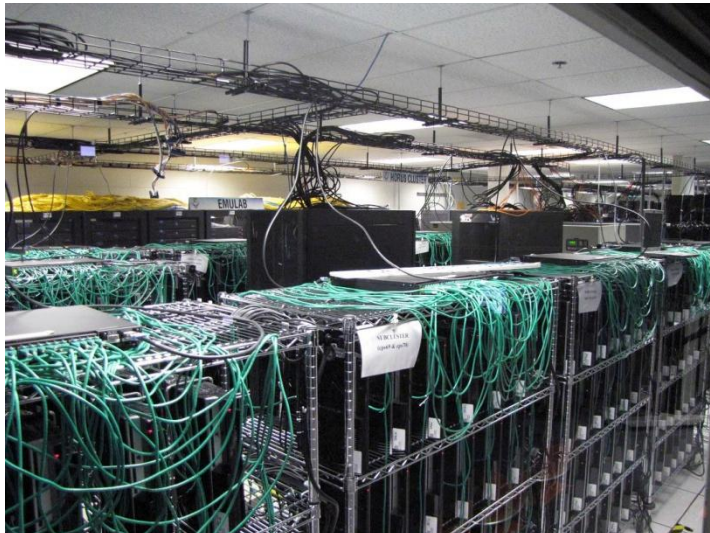
- **Computing resources involved in this run**
  - 4 Condor servers – 32 Intel Xeon processor cores
  - 88 PlayStation 3's – 616 IBM Cell-BE processor cores

Total X86 Cores Running	Total Cell-BE Cores Running
32/936	616/12012
Total Number of Pages Processed	
21	
Pages per Second	Sentences per Second
3.9	139.4
Words per Second	Characters per Second
985.6	6711.0

- 40 Condor servers – 320 Intel Xeon processor cores
- 880 PS3s – 6160 IBM Cell-BE Processor cores (21 pages/sec)



# Condor vs. Watson



- 69940 processor cores, 2.8TB memory
- **Typical application:** Occluded Text Recognition
- **Core function:** Determine the missing information in a damaged text image

- 2880 processor cores, 16TB memory
- **Typical application:** Jeopardy
- **Core function:** Determine the missing information in a Jeopardy clue

Alice was beginning to get very tired of sitting by her sister on the bank



Alice was beginning to get very tired of sitting by her sister on the bank

THIS IS THE MOST USED VOWEL SOUND IN THE ENGLISH LANGUAGE



E

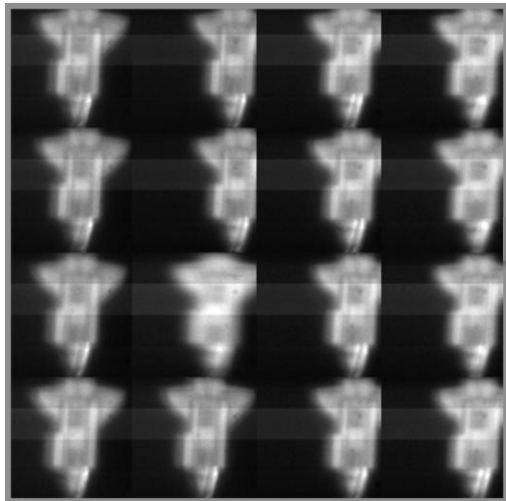




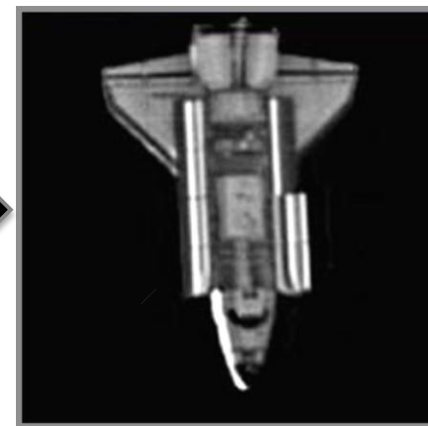
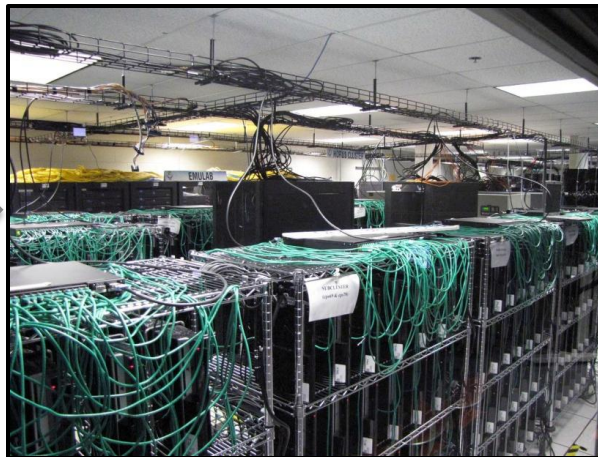
# Space Object Identification



Combining frames to create high quality  
images in real-time



*Low resolution frames*



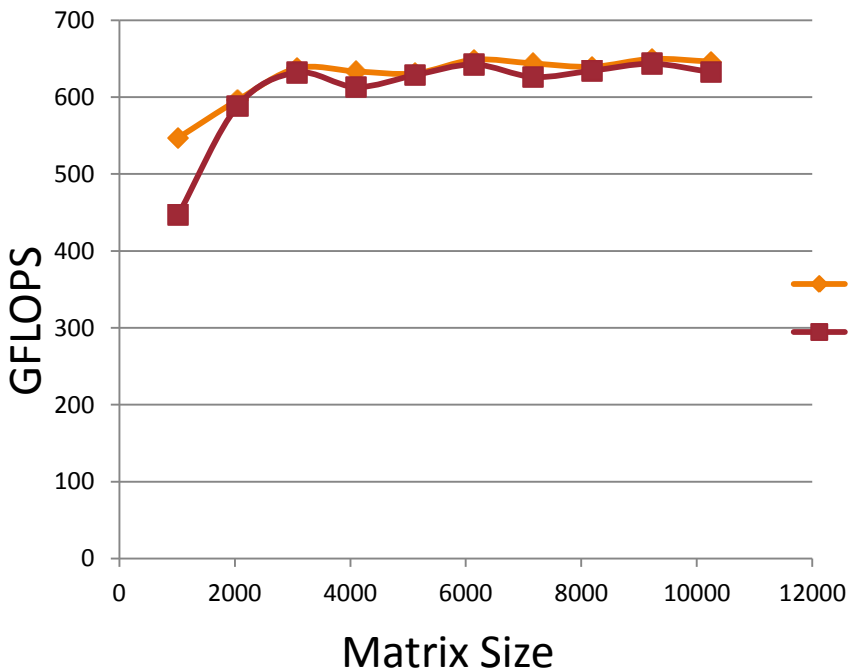
*High resolution image*



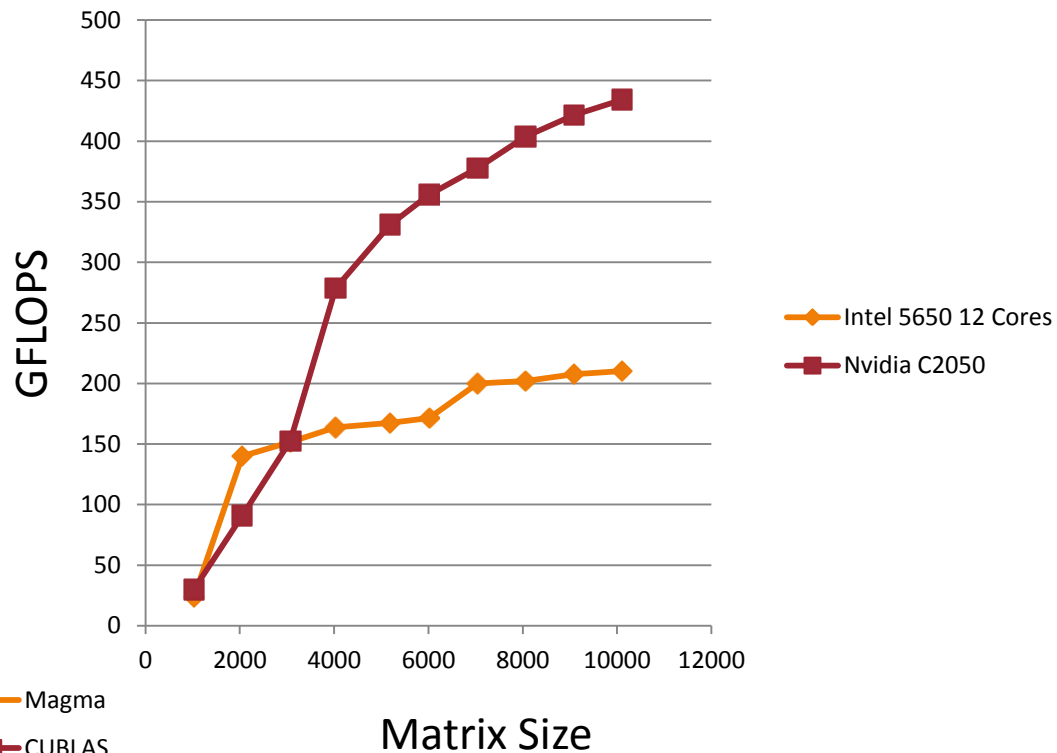
# Matrix Multiply



### Matrix-matrix multiplication test C2050 (MAGMA vs CUBLAS)



### MAGMA-only, one-sided matrix factorization

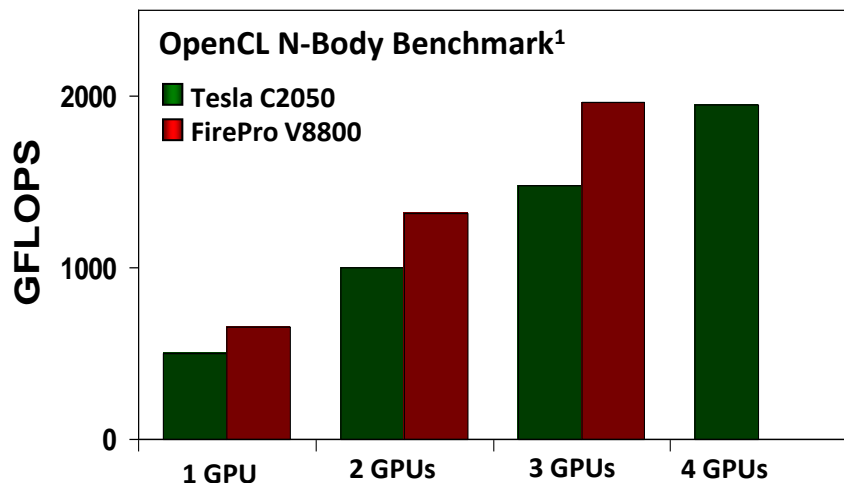




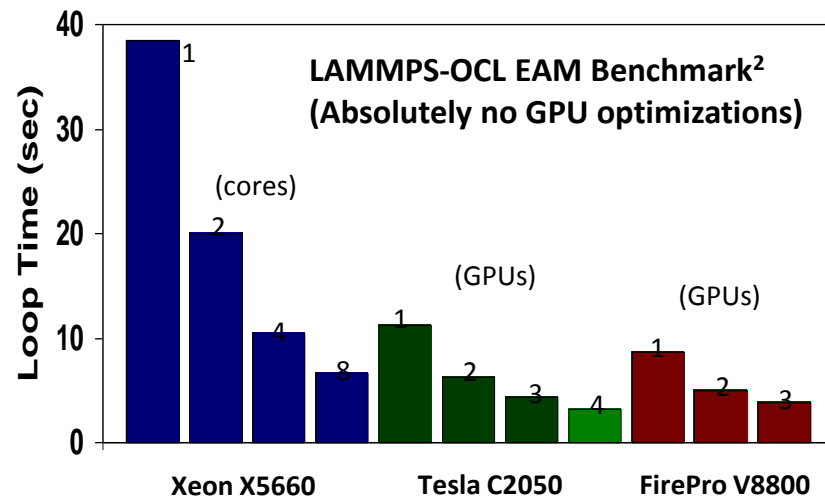
# LAMMPS on GPUs

Brown Deer  
Technology

- **Condor/Firebird provides access to next-generation hybrid CPU/GPU architectures**
  - Critical for understanding the capability and operation prior to larger deployments
  - Opportunity to study non-traditional applications of HPC, e.g., C4I applications
- **CPU/GPU compute nodes provides significant raw computing power**
  - OCL N-Body benchmark with 768K particles sustained performance ~ 2 TFLOPS using 4 Tesla C2050s or 3 FireBird V8800s
- **Production chemistry code (LAMMPS) shows speedup with minimal effort**
  - Original CPU code ported to OpenCL with limited source code modifications
  - Exact double-precision algorithm runs on Nvidia and AMD nodes
  - Overall platform capability increased by 2x (2.8x) *without any GPU optimization*



<sup>1</sup>MPI-modified BDT N-Body benchmark distributed with COPRTHR 1.1



<sup>2</sup>LAMMPS-OCL is a modified version of the LAMMPS molecular dynamics code ported to OpenCL by Brown Deer Technology





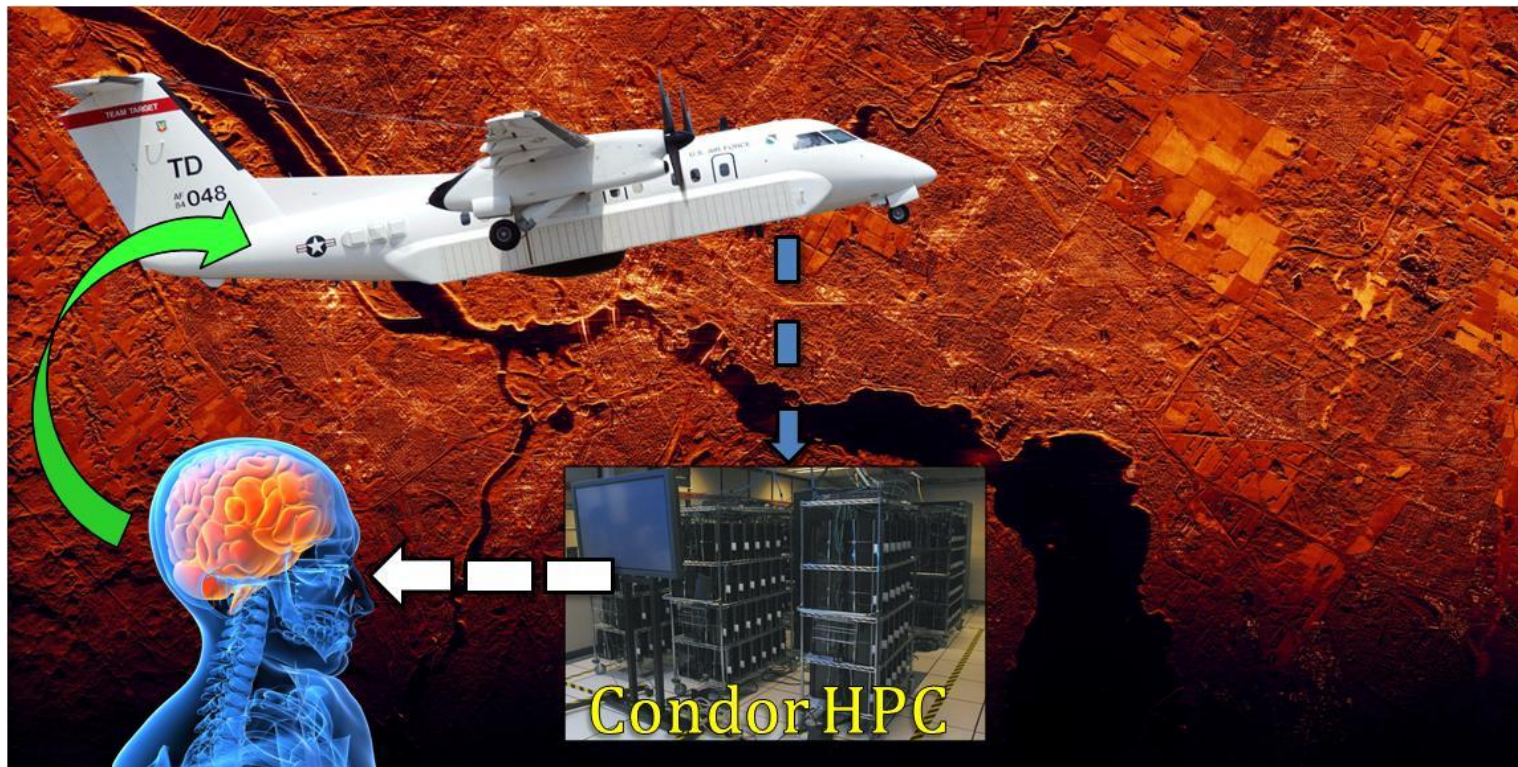
# Future Work



- **Improved OTR applications**
  - Multiple languages
- **Space Situation Awareness**
  - Heterogeneous algorithms
- **Persistent Wide Area Surveillance**



# Autonomous Sensing in Persistent Wide-Area Surveillance



## • Cross-TD effort

- Investigate scalable, real-time and autonomous sensing technologies
- Develop a neuromorphic computing architecture for synthetic aperture radar (SAR) imagery information exploitation
- Provide critical wide-area persistent surveillance capabilities including motion detection, object recognition, areas-of-interest identification and predictive sensing



# Conclusions



- Valuable resource to support entire AFRL/RI, AFRL and tri-service RDT&E community.
- Leading large GPGPU development and benchmarking tests.
- This investment is leveraged by many (150+) users
- Technical benefits – Faster, higher fidelity problem solution; multiple parallel solutions, heterogeneous application development



# Questions?

