

# InfiniCortex

The challenge of building a global scale  
Supercomputer

Marek T. Michalewicz

Senior Director

A\*STAR Computational Resource Centre (A\*CRC)

Singapore

GPU Technology Workshop South East Asia 2014

July 10th 2014



# **Interplay of HPC, Storage, Networking and all components to deliver balanced national computational resources**

## **A\*STAR experience**

Marek T. Michalewicz

Senior Director

A\*CRC

# Necessary HPC Service Provisioning Components

- Computational hardware
- Software (middleware, compilers, libraries, application software, tools)
- Storage (backup, archive, on-line)
- Access and connectivity (network)
- User support and education – HPC expertise
- Visualisation

## Our Mission

Fostering world-class scientific research and talent  
for a vibrant knowledge-based Singapore

Chairman A\*STAR  
Managing Director A\*STAR

Science &  
Engineering  
Research Council

Joint Council  
Office

Biomedical  
Research  
Council

A\*STAR  
Graduate  
Academy

Corporate  
Group

## SERC

### A\*STAR Computational Resource Centre

Data Storage Institute (DSI)  
Institute of Chemical and Engineering Sciences (ICES)  
Institute of High Performance Computing (IHPC)  
Institute for Infocomm Research (I<sup>2</sup>R)  
Institute of Materials Research and Engineering (IMRE)  
Institute of Microelectronics (IME)  
National Metrology Centre (NMC)  
Singapore Institute of Manufacturing Technology (SIMTech)

## BMRC

Bioinformatics Institute (BII)  
Bioprocessing Technology Institute (BTI)  
Genome Institute of Singapore (GIS)  
Institute of Bioengineering and Nanotechnology (IBN)  
Institute for Medical Biology (IMB)  
Institute of Molecular and Cell Biology (IMCB)  
A\*STAR-Duke-NUS Graduate Medical School Neuroscience Research Partnership (NRP)  
Singapore Bioimaging Consortium (SBIC)  
Singapore Consortium of Cohort Studies (SCCS)  
Singapore Institute for Clinical Sciences (SICS)  
Singapore Immunology Network (SIgN)  
Singapore Stem Cell Consortium (SSCC)  
Experimental Therapeutics Centre (ETC)

### Commercialisation Outfits

- Exploit Technologies
- Experimental Therapeutics Centre
- Industry Development Group (BMRC)

### Other Scientific Services Outfits

- Biological Resource Centre
- National Breeding Centre
- Singapore Tissue Network
- National Metrology Centre
- Biopolis Shared Facilities

# A\*CRC Datacenter I



Level 17 at Fusionopolis

# A\*CRC Datacenter 2

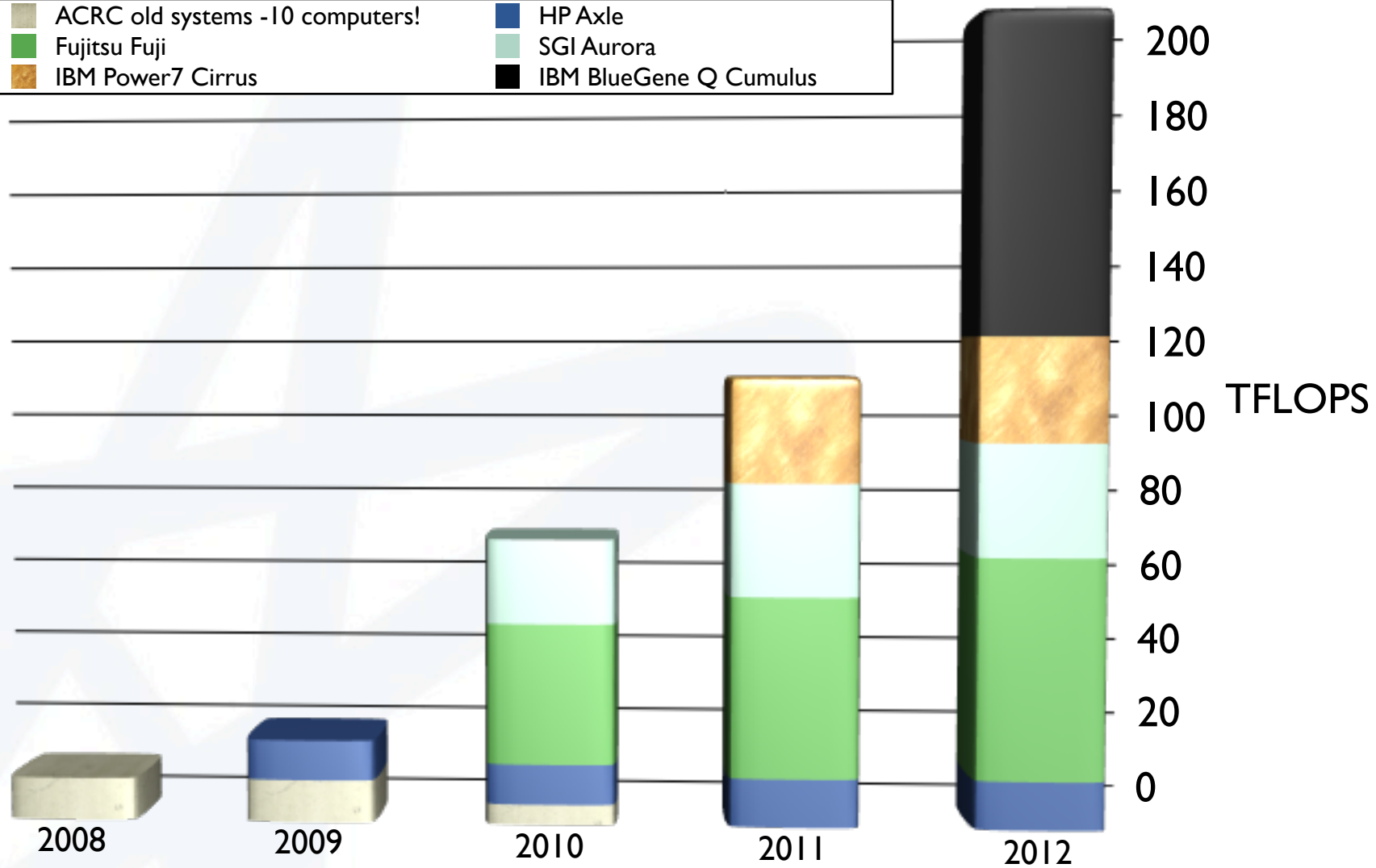
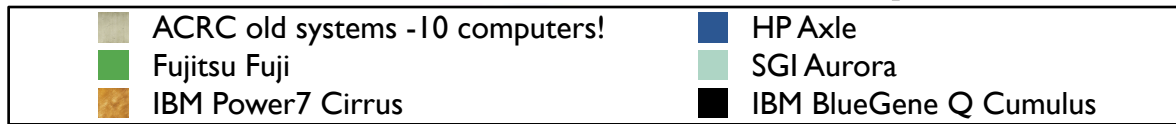


**Matrix Building at Biopolis**

# **Users Statistics**

- **~800 users**
- **5000+ A\*STAR staff**
- **all A\*STAR Institutes and Centres,**  
**but**
- **largest group from the Institute Of High  
Performance Computing**

# A\*STAR HPC Computational Power



2008

2009

2010

2011

2012



Agency for  
Science, Technology  
and Research

CREATING GROWTH. ENHANCING LIVES.



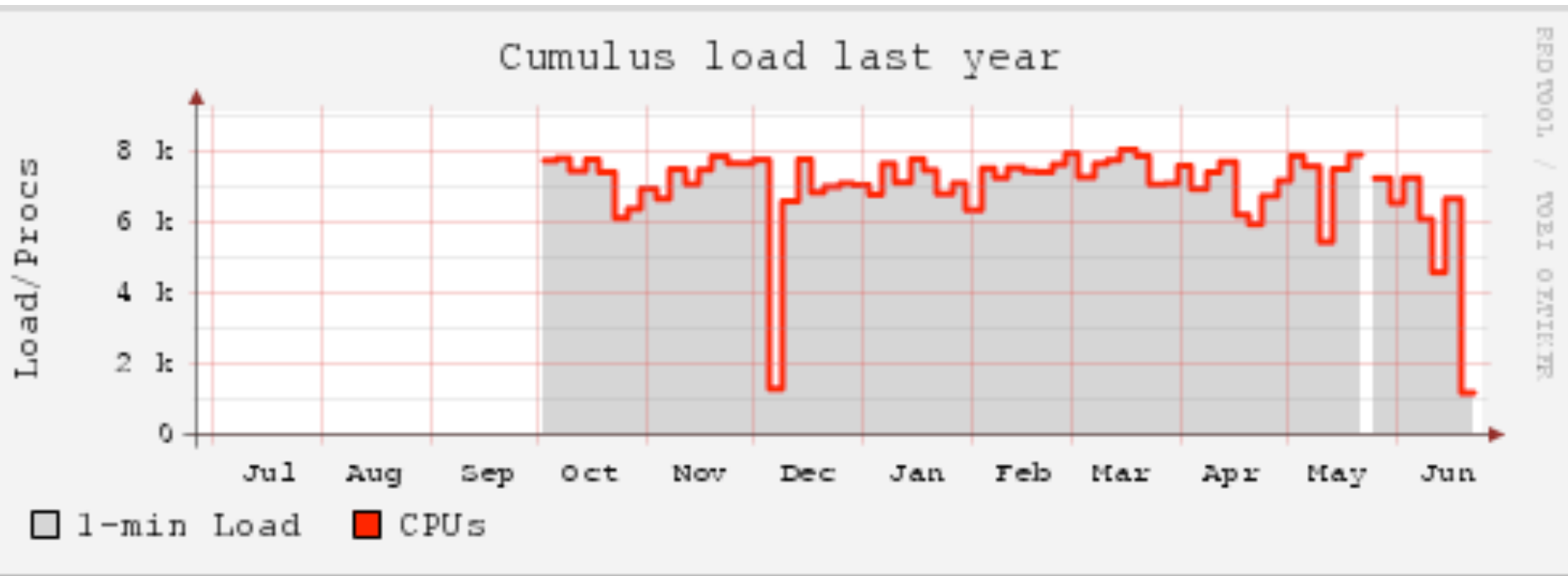
# The Cumulus System

- **IBM Bluegene/Q**

- Most powerful computer in Singapore
- 512 nodes
- PowerPC A2, 16 (18) cores per socket (node), 1.6GHz
- 1GB per core
- 5D torus
- Water-cooled
- 48kW (most energy-efficient system in the world)
- 100TB GPFS scratch disk



# System Utilisation, Oct 2013 to June 2014



# GPGPU in A\*<sup>\*</sup>CRC (I)

- First accelerated cluster was installed in 2011
  - a 12-node training server for the A\*<sup>\*</sup>CRC-IHPC-ADSC HPC Certificate Course
  - single Intel quad-core CPU (2.93GHz), 32GB RAM
  - 8 NVIDIA C2070 and 4 ATI 9700 GPU cards
  - Assembled by Jeremy Lim, Paul Hiew, Damien Leong and Stephen Wong

# GPGPU in A\*CRC (2)

- Training cluster turn into a production cluster in Q3 of 2011
  - NVIDIA accelerators only
  - Supported accelerated packages
    - ❖ AMBER
    - ❖ GROMACS
    - ❖ Quantum ESPRESSO
  - Administrated by Paul Hiew
- No popular scientific code support ATI accelerators
  - OpenCL is a high threshold for developers
  - OpenACC may provide an alternative solutions in future

# Our GPU Users

- Less than 10 users used our GPU servers
- most of the users are from the same group in A\*STAR Bioinformatic Institute
  - using AMBER for bio-molecular MD simulations
  - using single GPU for each job
  - demand from bio-molecular scientist is increasing
- One IHPC user used dual GPU for a special project using Quantum ESPRESSO

# Recent Development (I)

- In April, A\* CRC helped NUS and NTU students to prepare for Asian Supercomputing Challenge (ASCI4)
  - we provided a 4-node SGI cluster for training
    - ❖ each node has 2 Intel I2-core CPU and 128GB memory
    - ❖ 2 Xeon Phi accelerators
    - ❖ 9 K40 GPUs :
      - ★ 4 cards came with the SGI servers
      - ★ 5 cards borrowed from NVIDIA
  - K40 provides good HPL performance
  - Quantum ESPRESSO does not work on K40
  - NTU team won the 2nd place out of 82 international teams

Last week, I read in Lianhe Zaobao that a team of Nanyang Technological University students had come in second out of 82 international teams, at the world's largest supercomputer competition in Guangzhou. The teams had to develop the fastest and most optimised mini-supercomputer. The participants needed many skills: computer design, engineering and coding. (The Shanghai Jiaotong University team took the first prize.) (<http://on.fb.me/1nymr9z>)

Yesterday I learnt that Shannon Lee Xinjing, an National Junior College student, had won the Intel Foundation Young Scientist Award. Shannon developed a cheaper, eco-friendly way to make rechargeable zinc-air batteries more stable - by using brinjals! The rest of our team also did well. (<http://bit.ly/1pisSPr>)

I am happy to see our students excelling in the sciences, engineering and IT. These are fascinating and exciting subjects. They will also be valuable skills to have, in the high-tech world we live and work in. I will always be glad that I studied Maths and Computer Science at university, even though I did not continue with them afterwards. I hope more students will make these subjects their passion. - LHL

[Top - NTU team in Guangzhou (L-R): Frankie Wong, Qin Gemeng, Han Jianglei, Dhakshina Moorthy Pradeep, Lim Chun Leng / Photo Courtesy of NTU]

Bottom - S'pore team in LA (L-R): Mark Lim Kit (RI); Ong Jun Yi (NUS High); Vanessa Tan Yun Han (Dunman High); Du Yang Zhong Liang (NUS High); Shannon Lee Xinjing (NJC); Jeremy Goh Jin Quan (NUS High); Foo Cher Ying (RGS); Lim Ee Chuan Elliot (NUS High); Dousabel Tay May Yi (NUS High); Yu Shiyang (NUS High) / Photo Courtesy of MOE] - with Ong Jun Yi and 12 others.



# Singapore National Supercomputer Centre

**Proposal** (SG\$ **big**, three year funding, \*almost\* approved)

## **National Supercomputing Centre (NSCC)**

- Upgrade with new 1-2 Petaflop system to place Singapore in the league of developed countries with Petascale capacity (and consequently, capability)
- Centrally manage top end HPC resources and data storage across various locations hosted by A\*CRC, NUS and NTU
- Enable user access regardless of location

## **Science, Technology, Research and Education Network (STAREN)**

- Upgrade the backbone R&E network to 40 & 100 GE to provide easy access to the increased compute resources
- Linking A\*STAR, CREATE, NTU, NUS, SMU and SUTD
- **Upgrade international connectivity to US, Europe and Japan to 10GE (subject to agreements and co-funding from counterparts)**
- Enable Singapore organisations and industry to tap on full potential of NSCC



**A\*CRC is not actively involved in own research,  
nor product development,  
but:**

- we study state-of-the-art HPC technologies,
- we engage in forward trends discussions with vendors
- we observe and study best practices and trends
- we implement the best technological solutions that suit our users' needs.

# Short Review

2012.2

- Meeting in Beijing(A\*Star-LS)

2012.5

- Meeting in Beijing (A\*Star-LS)

2012.6

- Test 3A system(A\*Star)

2012.8

- Test 3B1000 system(A\*Star)

2012.10

- Workshop in Singapore. MOU(4 parties)

2013.1

- Test 3B1500 system (A\*Star)

2013.3

- Paul visit EXTOLL to work on PCI-E card

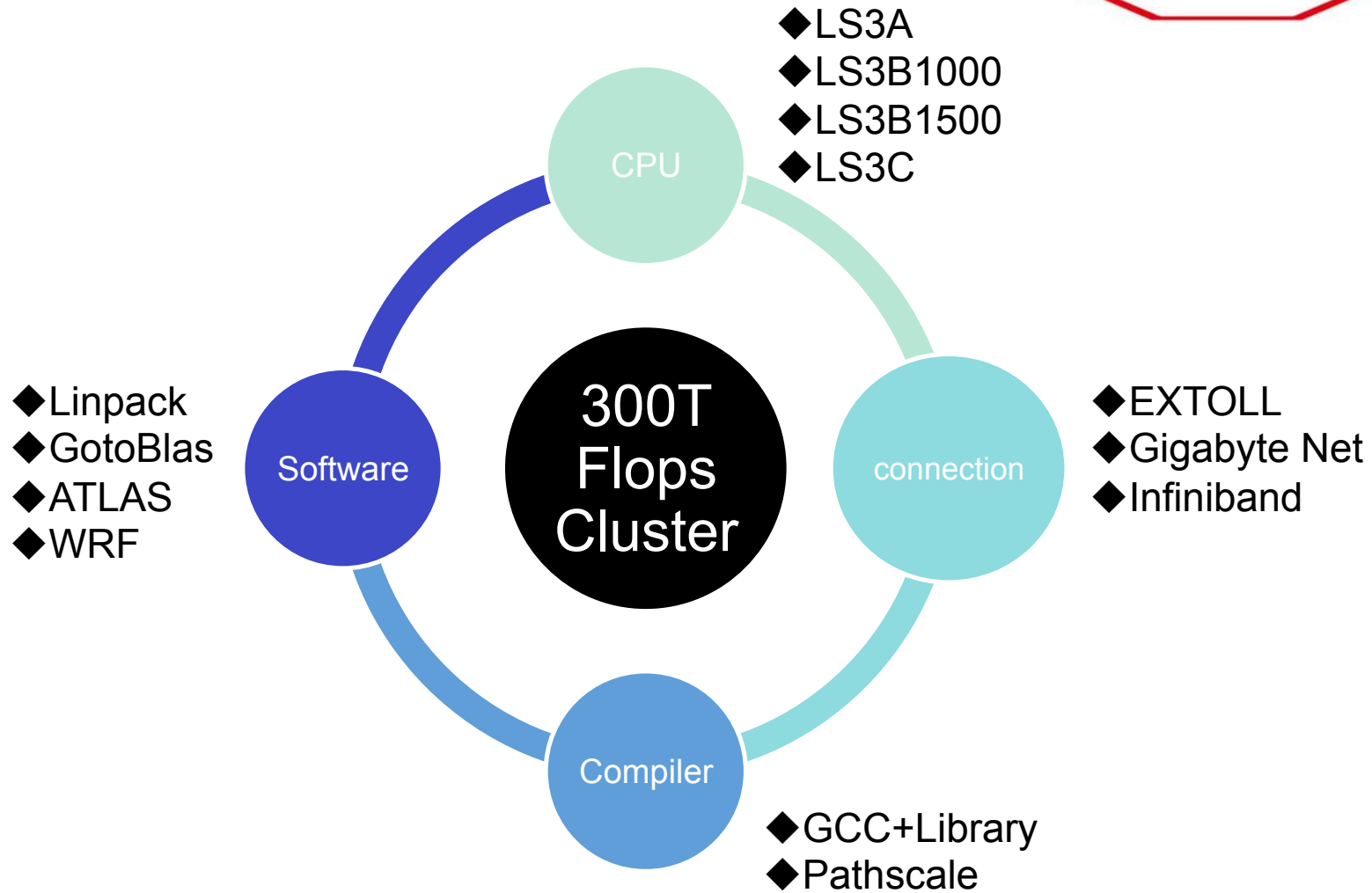
2013.4

- Workshop in Beijing(A\*Star-LS)

2013.4

- Paul visit Loongson solved PCI-E x16 problem

# Technical Map



# InfiniCortex

The challenge of building a global scale  
Supercomputer

Marek T. Michalewicz

Senior Director

A\*STAR Computational Resource Centre (A\*CRC)

Singapore

GPU Technology Workshop South East Asia 2014

July 10th 2014

# The GRAND Challenge

- 1. ACA 100 (Asia Connects America 100 Gbps, by November 2014)**
- 2. InfiniBand over trans-Pacific distance**
- 3. Galaxy of Supercomputers**

Partners:

xxxx Communications

Obsidian Strategics

A\*CRC Singapore

SingAREN

Japan: Tokyo Tech (Tsubame-KFC)

USA: Oak Ridge National Laboratory (Titan)

Taiwan

N-RENS and international counterparts



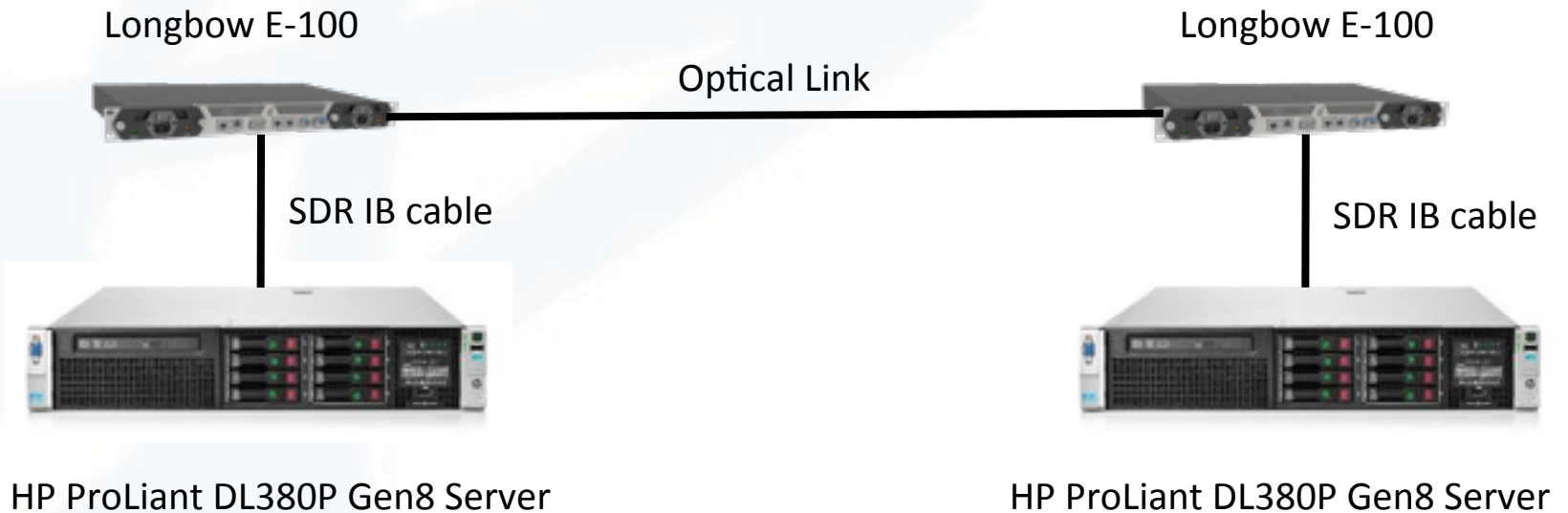
Computational  
Resource Centre



# Equipment Used

- Two HP ProLiant DL380P Gen8 servers
- Each server has
  - Two Xeon E5-2660 8-core CPUs @ 2.2 GHz
  - 128G RAM
  - DDR IB card (manually configured to SDR)
- Two Longbow E-100

# Equipment Setup





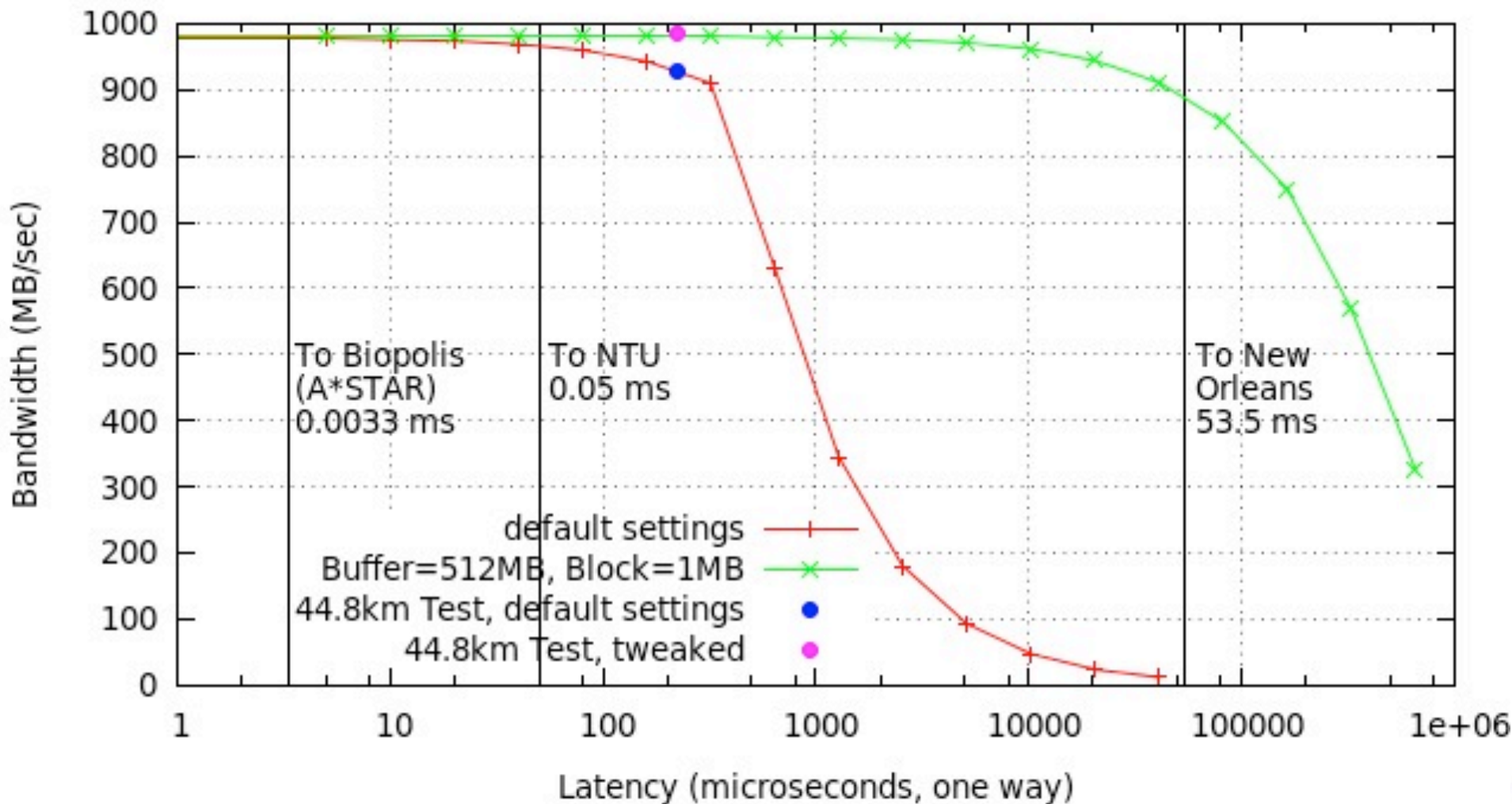
# A\*STAR @ Fusionopolis (FP) to Nanyang Technological University (NTU)



- Geographical distance: ~15 km
- Cable distance: ~ 22.4 km
- FP-NTU-FP loop cable distance : ~ 44.8 km

# dsync+ Results: Delay Simulated and Actual 44.8km Test

dsync+ Sustained Transfer Rate vs. Latency (Longbow E100 with 10Gbps SDR, no crypto)



# Galaxy of Supercomputers

Study of topologies

Yuefan Deng, A\*ARC & Stony Brook University  
Lukasz Orłowski, A\*ARC & Stony Brook University  
Marek Michalewicz, A\*ARC

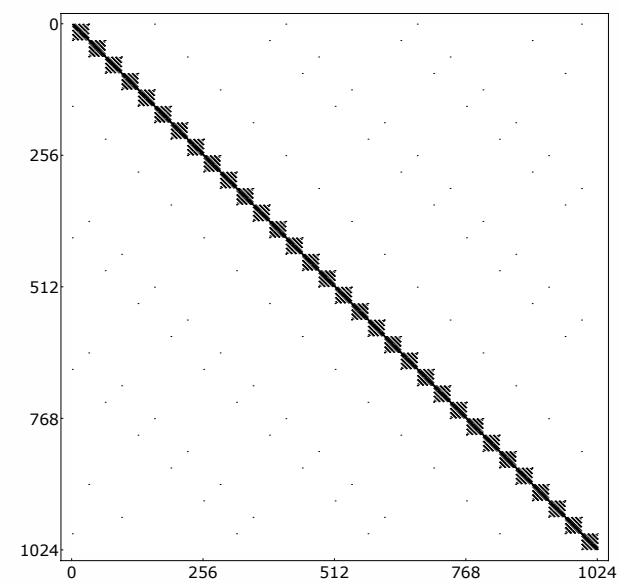
# Galaxy of Supercomputers

- Supercomputers located at different geolocations connected into a *Nodes of Super-Network (Super-Graph)*
- Supercomputers may have arbitrary interconnect topologies
- Galaxy is based on a topology with small diameter and lowest possible link number
- In terms of graph representation it is *embedding* of graphs representing Supercomputers' topologies into a graph representing the Galaxy topology
- **poster 9a “HPC in Asia” session**

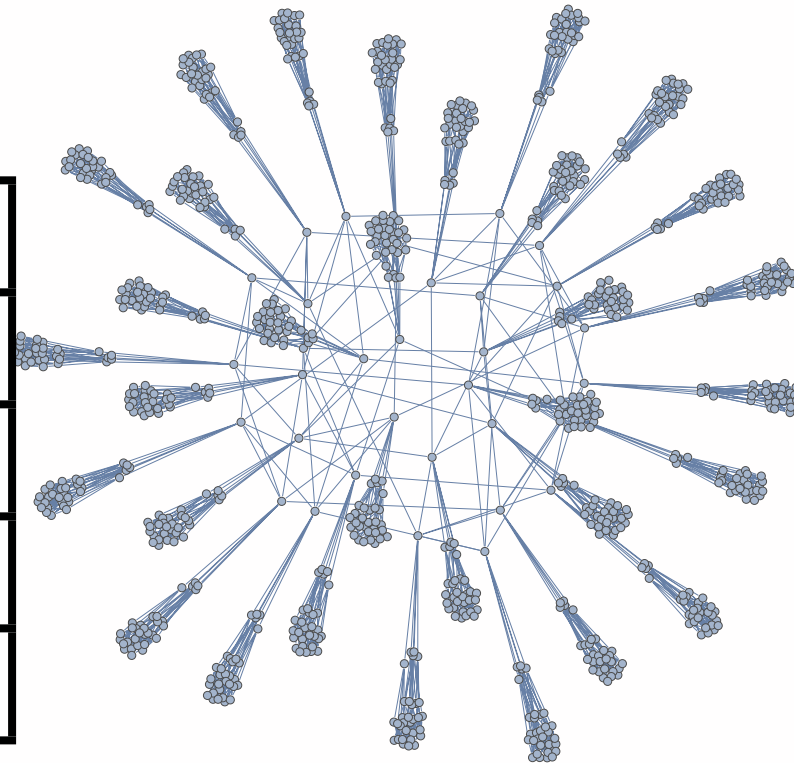


# 32k5 ⊗ 32k5

**Embedding** of a 5-connected graph on 32 nodes into itself proves to be comparable to TOFU or 5D torus with equal or similar number of nodes.

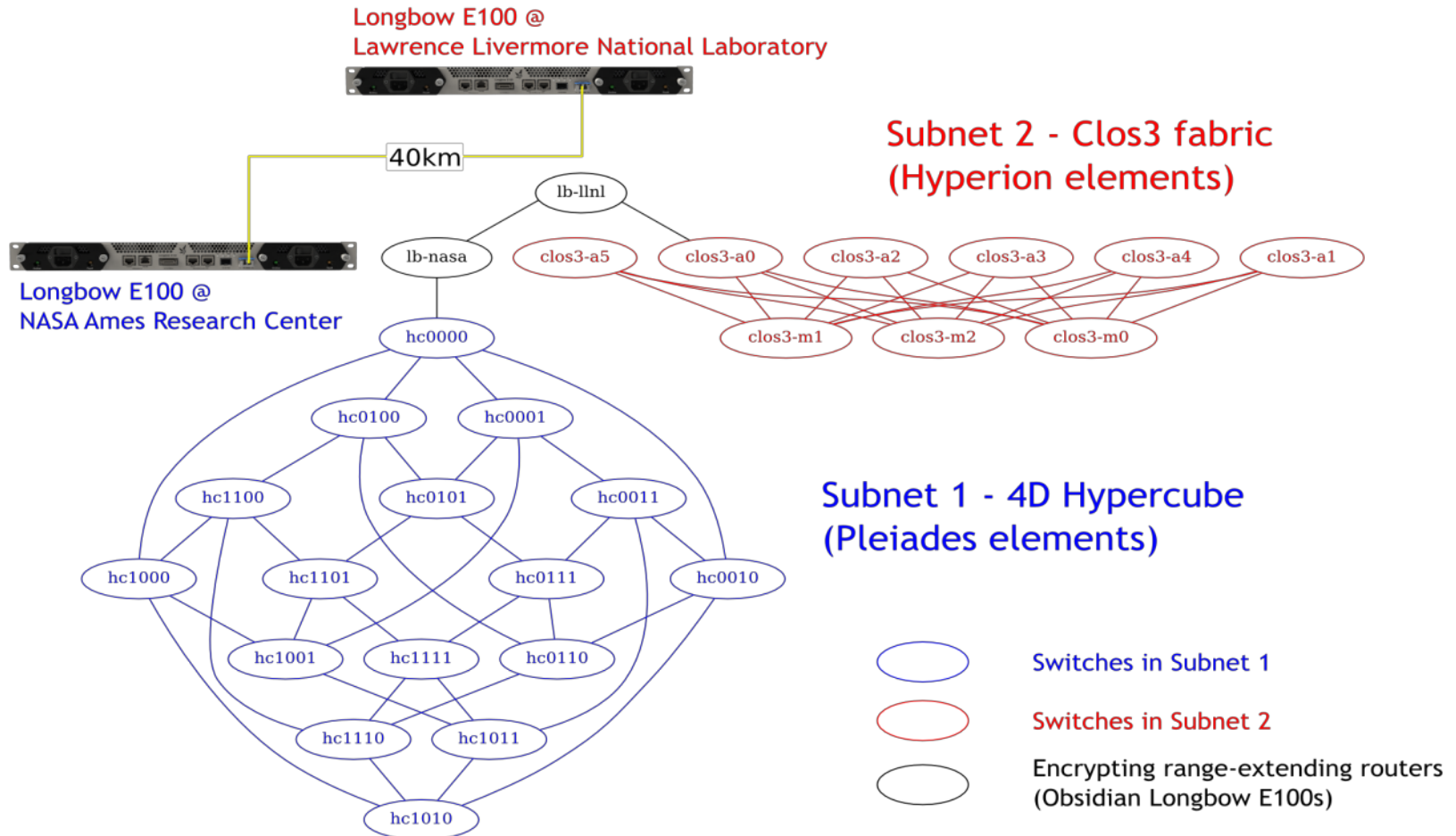


<i>Name of topology</i>	<i>Number of nodes</i>	<i>Number of link</i>	<i>Diameter</i>	<i>Mean path length</i>
<b>32k5 ⊗ 32k5</b>	<b>1024</b>	<b>2640</b>	<b>9</b>	<b>6</b>
<i>Tofu (6x5x3)</i>	<i>1080</i>	<i>5400</i>	<i>9</i>	<i>5</i>
<i>5D torus (4x4x4x4x4)</i>	<i>1024</i>	<i>5120</i>	<i>10</i>	<i>5</i>
<i>Tofu (4x4x8)</i>	<i>1536</i>	<i>7680</i>	<i>11</i>	<i>6</i>



# Galaxy of supercomputers: proof of concept

## System Software: BGFC - Next Generation InfiniBand Subnet Manager

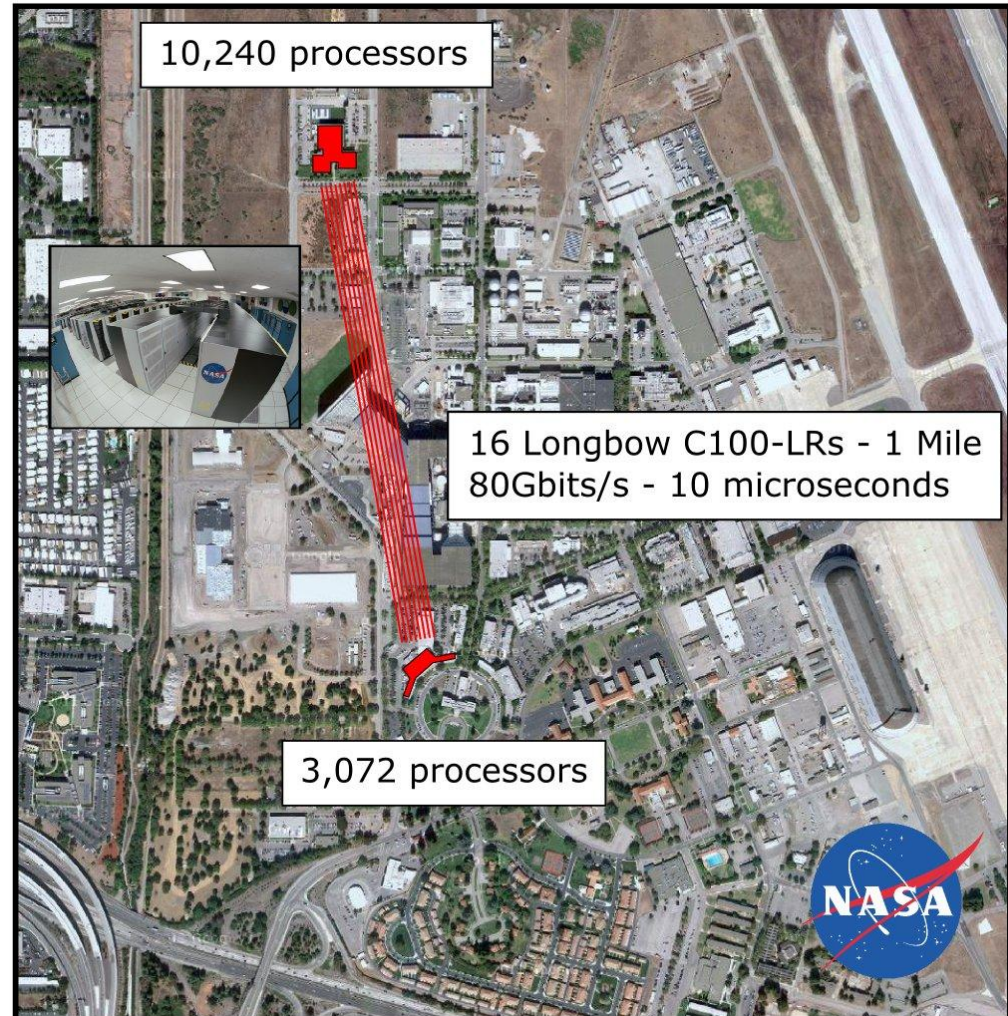


A 1Gbyte/second 40km routed and encrypted InfiniBand link was demonstrated at SC12, using BGFC to orchestrate the fabric comprising two subnets with very different internal topologies (“compound topology”).

# Galaxy of supercomputers: proof of concept

## Capability: Low Latency Server Aggregation

- NASA Ames purchased 16 Longbow C100 units expanding their flagship Itanium-based *Columbia* supercomputer to share jobs across one-mile of dark fiber to a second building.
- Expansion of supercomputers and data centers must contend with power and cooling constraints – these problems can often be resolved by Longbows.
- A similar model works for the linking of containerized data center pods in the field or within modular data centers.



# CONCLUSION:

## Invitation to participate in tackling InfiniCortex Challenge

1. trans-Pacific 100Gbps Research & Academic connectivity between Asia and USA  
**ACA 100** (by November 2014)

Building consortium of Asian partners to realise ACA 100

2. InfiniBand over trans-Pacific distance

3. demonstrating Galaxy of Supercomputers concept

Partners:

xxxx Communications

Obsidian Strategics

A\*CRC

SingAREN

N-RENS and international counterparts:

**YOUR ORGANISATION, YOUR COUNTRY**