

ISGC 2017 (Taipei, Taiwan)

“e-Science in Singapore - HPC without Borders”

(9th March 2017)

Dr John Kan
Chief Information Officer, A*STAR
Deputy Chairman, A*STAR Computational Resource Centre
Chief Data-Analytics Officer, A*STAR
President, SingAREN Network
Board Member, APAN Committee



Super Computing - Key to Economic Growth and Prosperity

- Cloud Computing, IoT, data mining, data analytics, supercomputing, are now perceived as essential to the wealth, security, growth, and economic prosperity of nations.
- Super Computing simulation and analysis have become a source of discovery. Extraordinary amounts of data are now collected and revolutionize storage. Genomics and personalized medicine as well as Cloud Computing require ultra high security and reliability, fast data replication and disaster recovery.
- Exascale computing is seen as the next Frontier with the USA, Europe, Japan and China allocating each in excess of one billion US\$ to be the first to reach this milestone by the early 2020's.
- We invite ISGC members and the larger R&E community in the region to interconnect and collaborate paving the road to a distributed exascale computing, and reap the benefits of this new phase of economic and human development.

Singapore



Physical Land Area: 718 sq km

Population: 5.5 million

3.9 million (70.8%)
Singapore Residents

Literacy Rate: 96.7% (Aged 15 & above)

76.4% of those aged 25-34 years
have tertiary qualifications

2014 GDP S\$390.1bn (US\$307.9bn)

Real Growth : 2.9%

Per Capita GDP: S\$71,318
(US\$56,284)

Independent since 9 August 1965

Sources: Singapore in Brief 2015, Department of Statistics

Agency for Science, Technology and Research

>1,700

INDUSTRY PROJECTS A YEAR
5 INDUSTRY PROJECTS A DAY

>2,800

PAPERS PUBLISHED A YEAR

>200

LICENSES A YEAR
4 LICENSES A WEEK

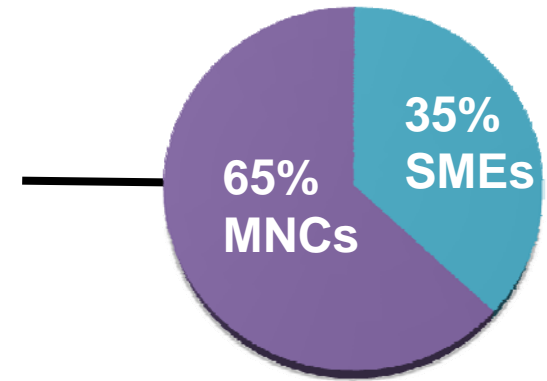
1

RSE SPUN OUT TO INDUSTRY A DAY*

** average number of Research Scientists and Engineers (RSE) per working day in a calendar year*

>240

PATENTS FILED A YEAR



>13

START-UPS A YEAR

A*STAR Computational Resource Centre

Scope:

- serve A*STAR users
- exploration of novel technologies and architectures

Many unique systems:

- Cumulus (IBM), Cirrus (IBM), Aurora (SGI), Fuji (Fujitsu), Axle (HP), Cray

Large range of architectures:

- Intel x86, IBM Power, NVIDIA Tesla, AMD FirePro, Intel Xeon Phi, Micron Automata

Involvement in exploration projects:

- InfiniCortex – InfiniBand ring around the world
- Intel Center of Excellence
- IBM OpenPower Foundation member (*4 x IBM Power 8 servers*)
- Micron Automata Center of Excellence (*2 x Automata boards*)


National Supercomputing Centre (NSCC)



~1 PFLOP System

- **1,288 nodes** (dual socket, 12 cores/CPU E5-2690v3)
- **128 GB DDR4 RAM/ node**
- **10 Large memory nodes** (1x6TB, 4x2TB, 5x1TB)


FUJITSU



13PB Storage

- **HSM Tiered, 3 Tiers**
- **I/O 500 GB/s flash burst buffer**
- **10x Infinite Memory Engines (IME)**


DDN



EDR Interconnect

- **EDR (100Gbps) Fat Tree within cluster**
- **InfiniBand connection to remote login nodes at stakeholder campuses (NUS/NTU/GIS) at 40/80/500 Gbps throughput**

Mellanox TECHNOLOGIES



Sample of NSCC applications

Numerical Ocean Basin

TECHNOLOGY CENTRE FOR OFFSHORE AND MARINE, SINGAPORE (TCOMS)

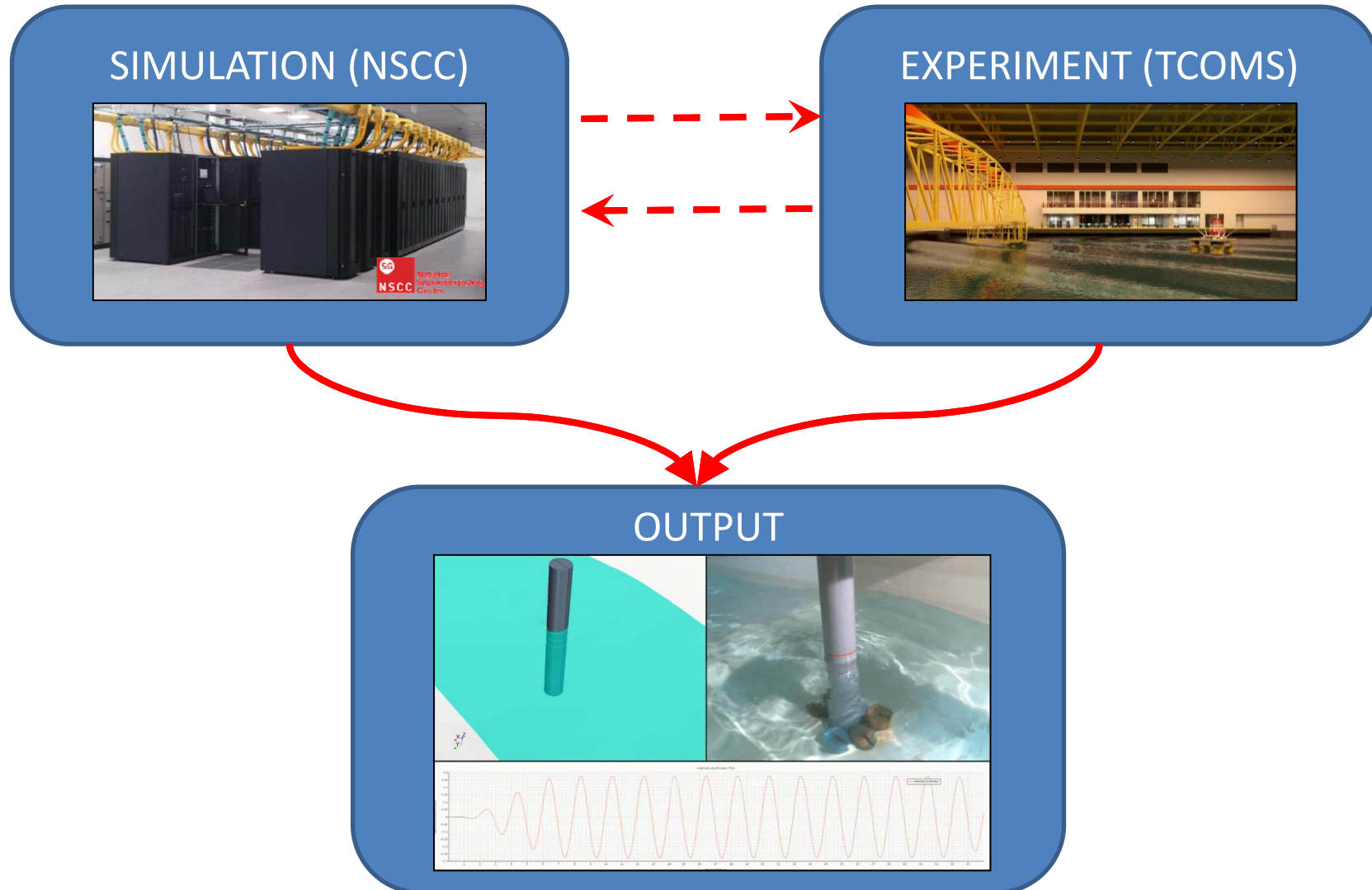
A state-of-the-art Ocean Basin facility is being developed by TCOMS@NUS in Singapore, comprising a deep tank equipped with wave and current generation systems to simulate ocean environment. NSCC partners TCOMS to provide the computational power required for the numerical modeling of waves and current flows of test models and experiments for marine and offshore structures.



Collaborative Project

Technology Centre for Offshore and Marine Singapore (TCOMS)

COUPLED EXPERIMENTAL - NUMERICAL WAVE FLUME TEST



NSCC applications

GenomeAsia 100K Consortium

GENOME@NANYANG TECHNOLOGICAL UNIVERSITY



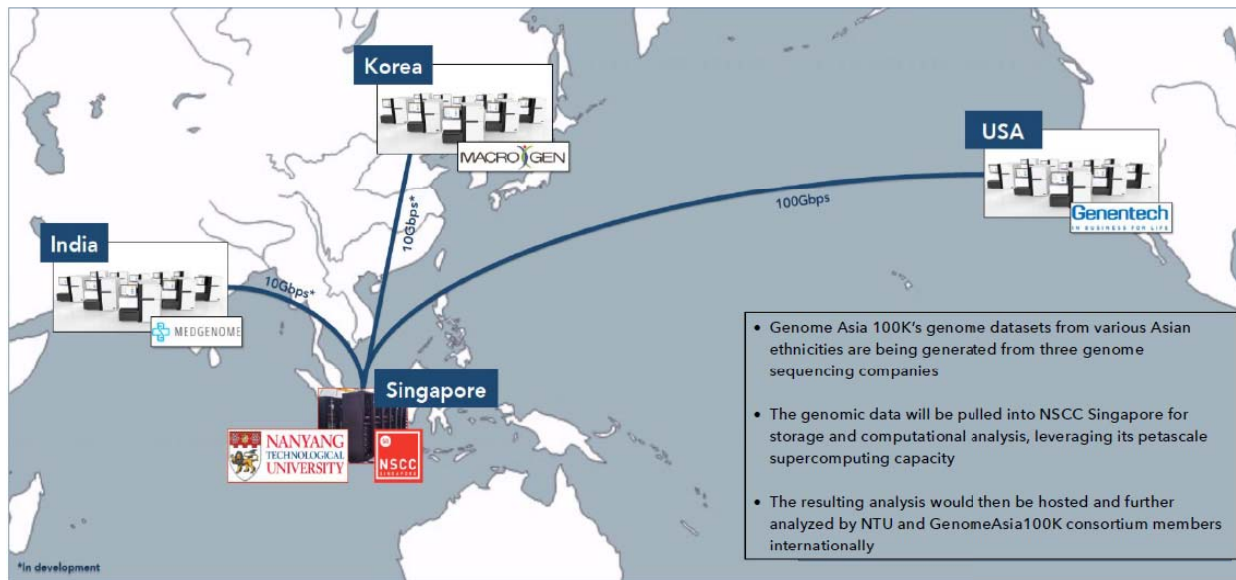
GenomeAsia 100K

The GenomeAsia 100K initiative aims to sequence 100,000 genomes from North, South and East Asia populations, with the goal of accelerating precision medicine and clinical application for Asian patients by leveraging new information and understanding from the collected genomics data.

Collaborating with NTU acting as host to the initiative to utilise NSCC's infrastructure to undertake the computational analysis of massive datasets for accelerating downstream analysis and experiments leading to new knowledge and insights in Genome science.

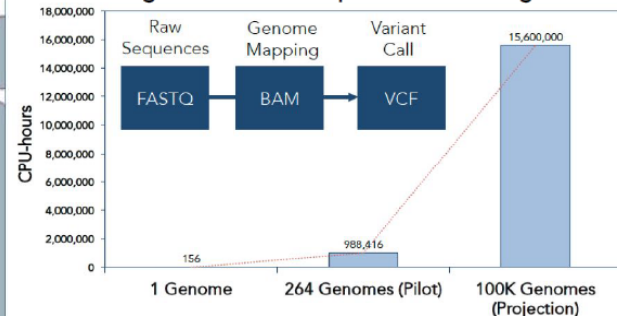
Collaborative Project (NTU-NSCC-GenomeAsia 100K consortium)

- **Aim:** To sequence 100,000 genomes from various South, North & East Asia populations
- **Goal:** To accelerate precision medicine and clinical applications for Asian patients
- **NTU** acts as the host and **NSCC** provides the HPC resources
- **68TB** of genome data from USA and Korea aggregated and hosted on NSCC platform



Pilot Benchmark/ Projections:

Processing benchmark with pilot data of 264 genomes



- **Using NSCC HPC resources, total compute time to process genome data for 100,000 genomes will only take up 1.1% of the 4-year project, thus providing more time for novel downstream analysis and experiments**

NSCC applications

National Precision Medicine Initiative

Led by Prof. Patrick Tan (Deputy Director, Biomedical Research Council of A*STAR and NSCC Steering Committee member) the NPMI will be a key partner and user of NSCC's supercomputing, advanced networking, visualisation and data wormhole highspeed connectivity to globally located healthcare information resources.

Some new Developments

Platform for Deep Learning

STRATEGIC PARTNERSHIP WITH NVIDIA:

Towards a Deep Learning platform leveraging on Theano, Caffe, TensorFlow and Torch on our 128 Tesla K40 GPU nodes.

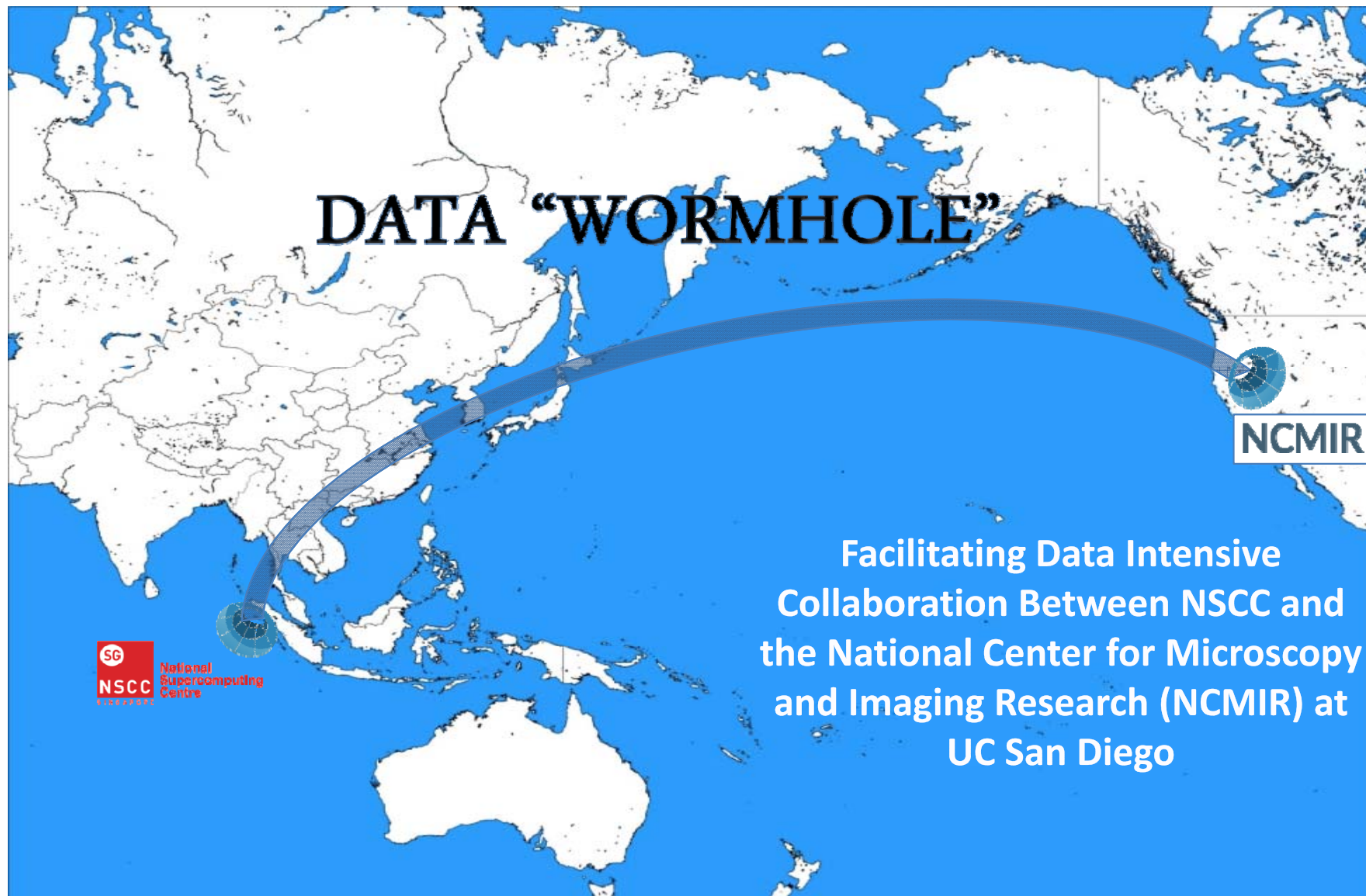


Data Wormhole



High bandwidth, low latency, transcontinental data transfer gateway in collaboration with International Centre of Advanced Internet Research (iCAIR), University of California San Diego (UCSD) to support our participation in the Global Research Platform, an extension of the Pacific Research Platform.

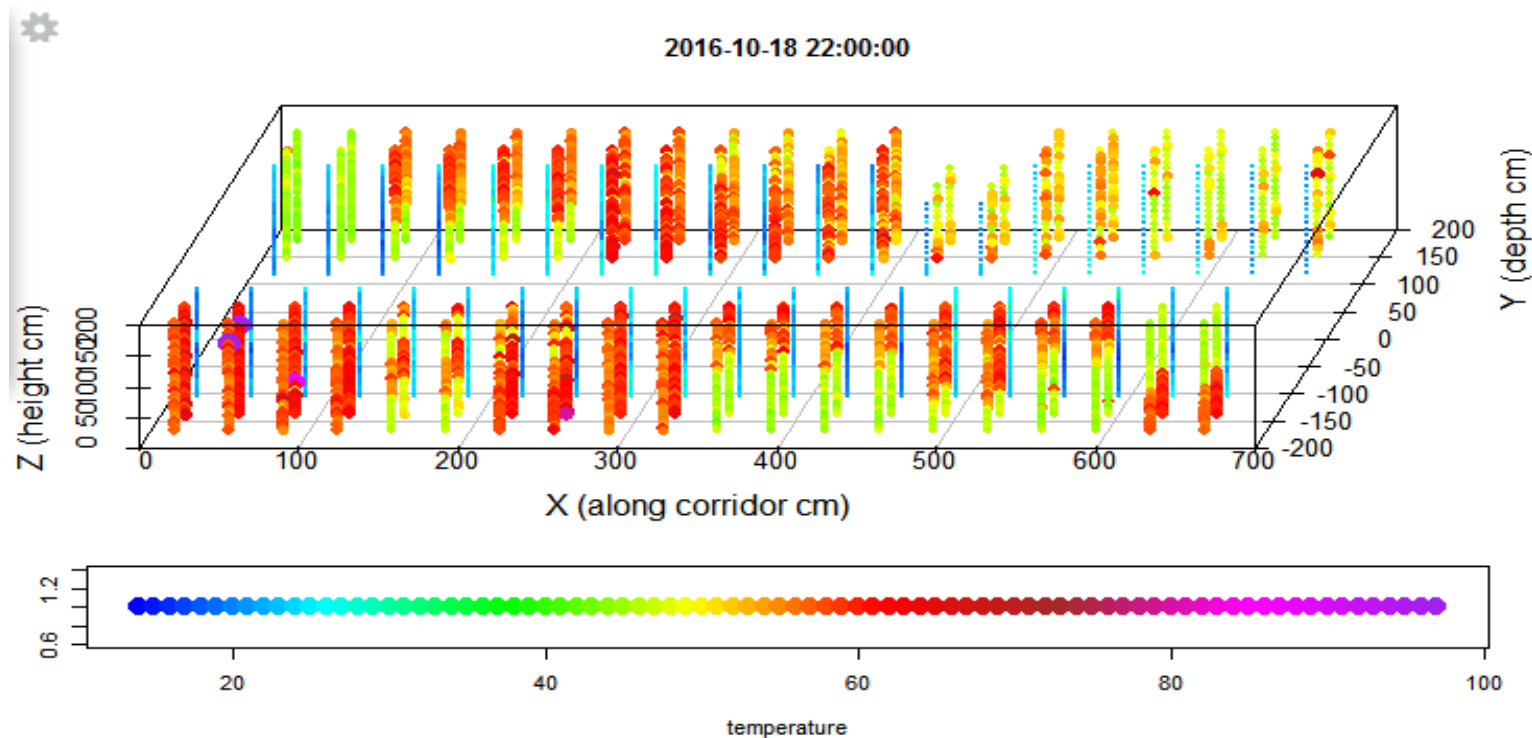
Collaborative Project (Data “Wormhole”)



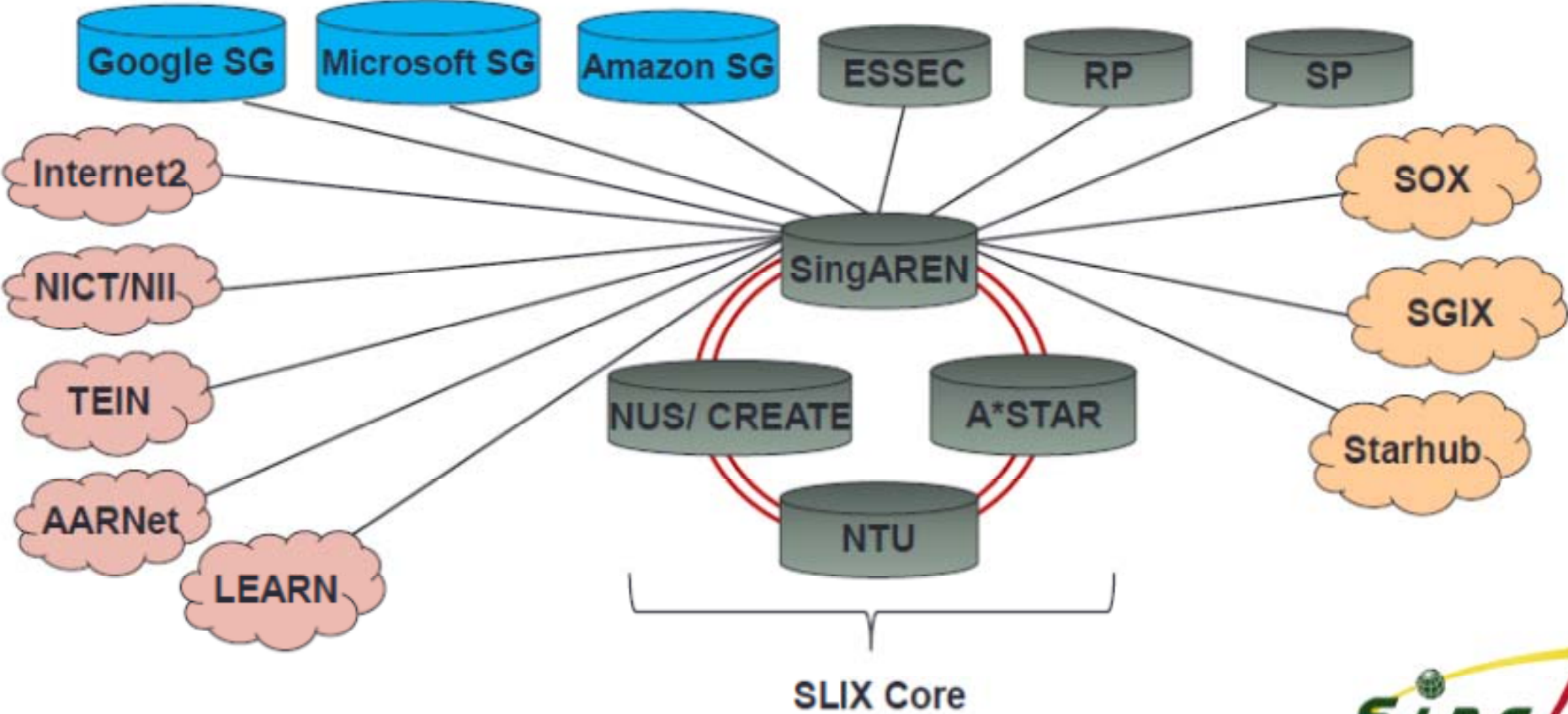
Collaborative Project

IMDA-NTU-NSCC (Green Data Centre)

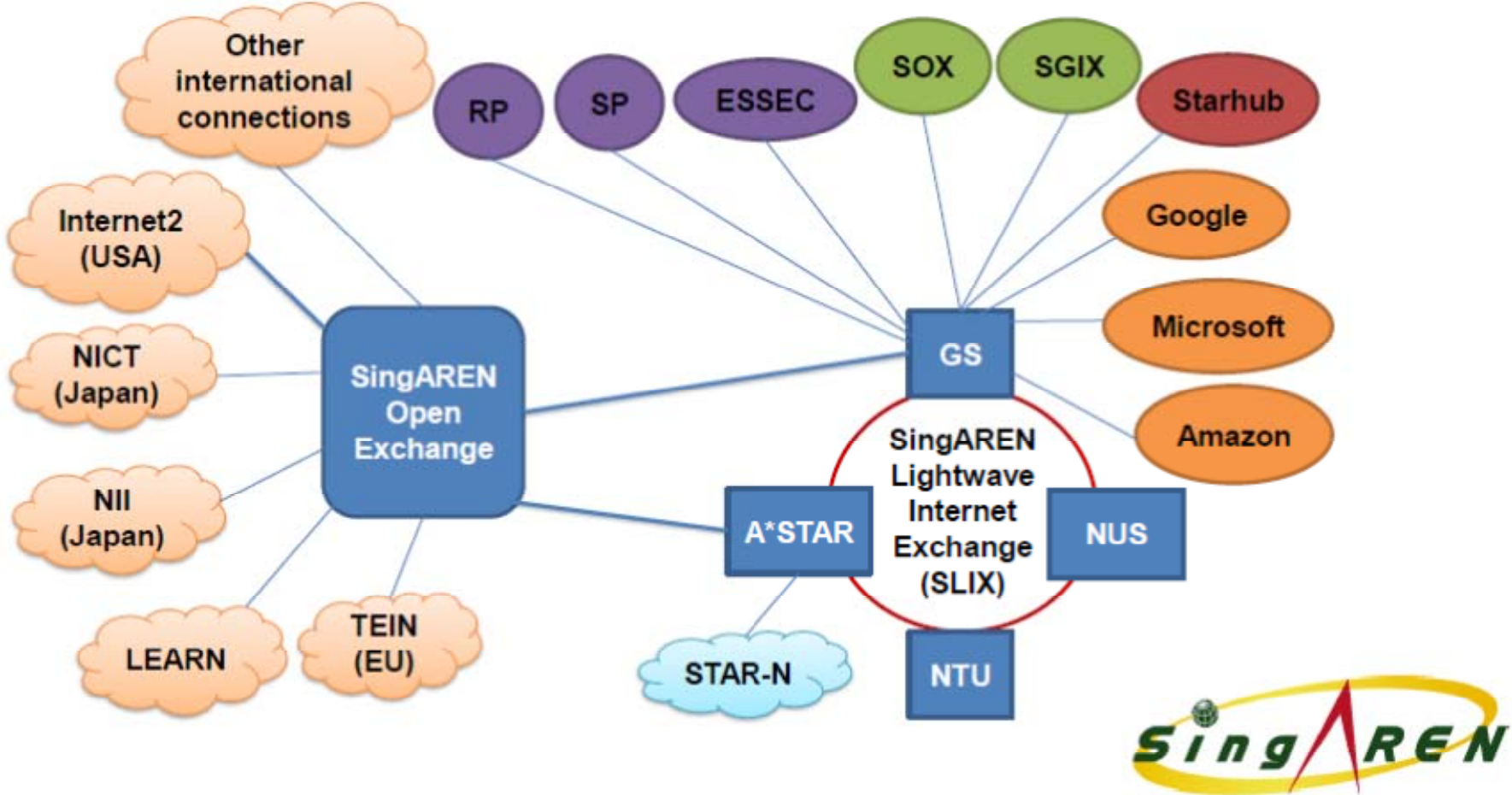
- Development of an interactive 3D visualization of the thermal profile of high performance server clusters in NSCC
- Ability to monitor the thermal dynamics of servers and spot outliers or anomalies
- Eventual aim is to improve the energy efficiency at NSCC's data centre



SLIX: SingaREN Lightwave Internet Exchange

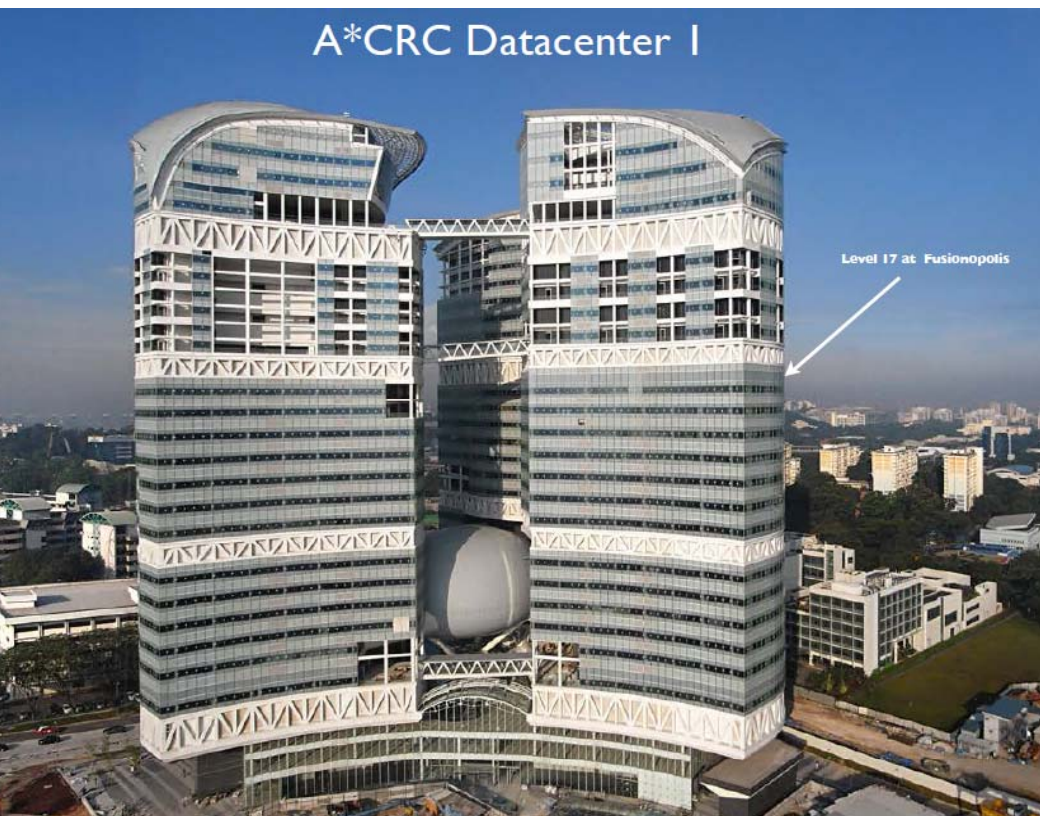


SingAREN Open Exchange



'Maximize my effective throughput between my storage and my compute'

Fusionopolis and Biopolis are 2km apart



Tests started with Mellanox METRO-X early 2013. Today the sites are connected with nx40gbps connections running native InfiniBand and reaching approx. 98.4% of maximum theoretical possible throughput. Dark fibre running at 400Gbps, soon 1Tbps using Infinera connects the two sites.

RDMA - InfiniBand – RoCE: Going the distance

- Remote Direct Memory Access (RDMA) is the core concept for efficient operation between the multiple cores of a cluster.
- Ethernet, InfiniBand and proprietary standards are used for communication between cores. InfiniBand has become the protocol of choice in recent years providing the best throughput at the lowest latency and runs on more than 50% of the Top100 supercomputers. The Ethernet camp has developed RoCE (RDMA over converged Ethernet) in response.
- A long distance version of InfiniBand was developed under the impetus of the US military. The long distance implementation of RoCE is a possible alternative.

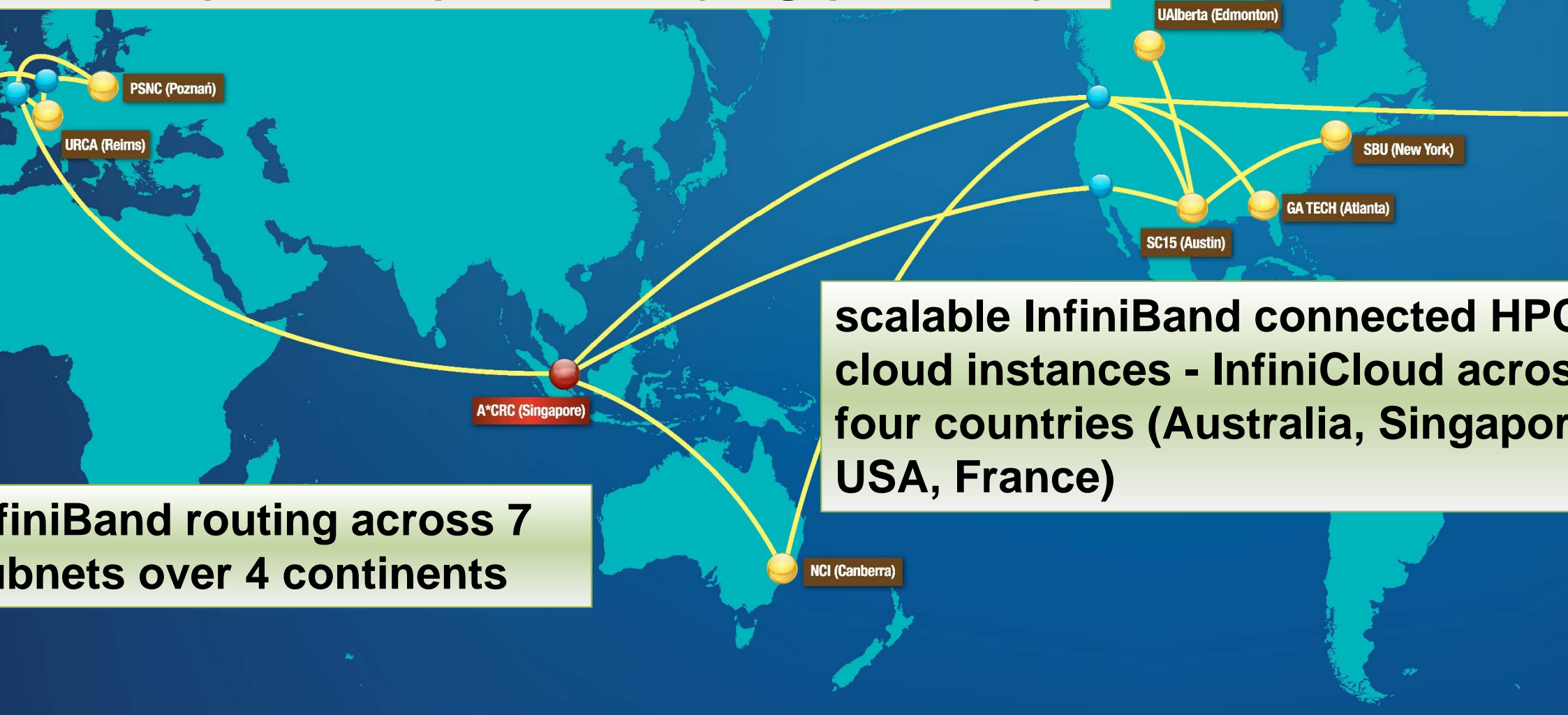


Singapore's approach to exascale: InfiniCortex

- A geographically dispersed constellation of compute, storage and associated power needs, working as one; not grid, not cloud.
- Successfully demonstrated at SC14, SC15 and SC16 as well as an InfiniCloud approach at ISC16.
- The five elements needed to succeed are being lined up:
 - Supercomputer interconnect topology based on graph theory work done by Y Deng, M. Michalewicz and L. Orlowski.
 - Availability of high speed uncongested bandwidth
 - Long distance InfiniBand or RoCE to increase effective throughput over any given link and InfiniBand or RoCE routing.
 - Application layer: from simple file transfers to complex workflows with Oakridge developed ADIOS and multi-scale models.
 - Partnerships: need to reach a critical mass to go to phase II and eventually to production mode.

InfiniCortex at SC 2016

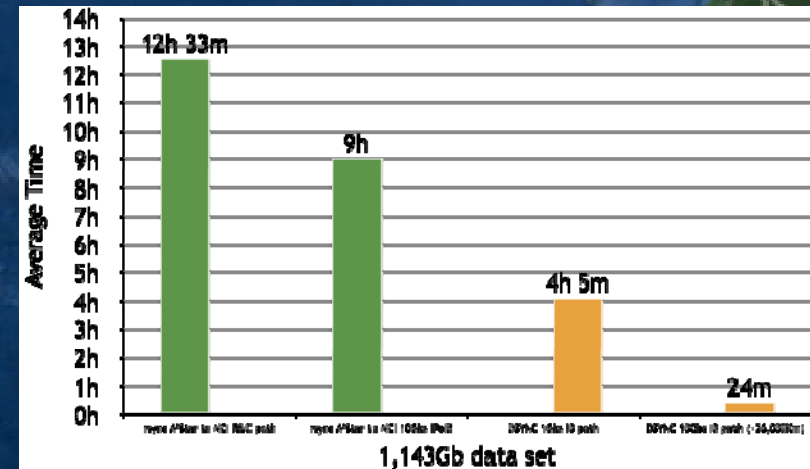
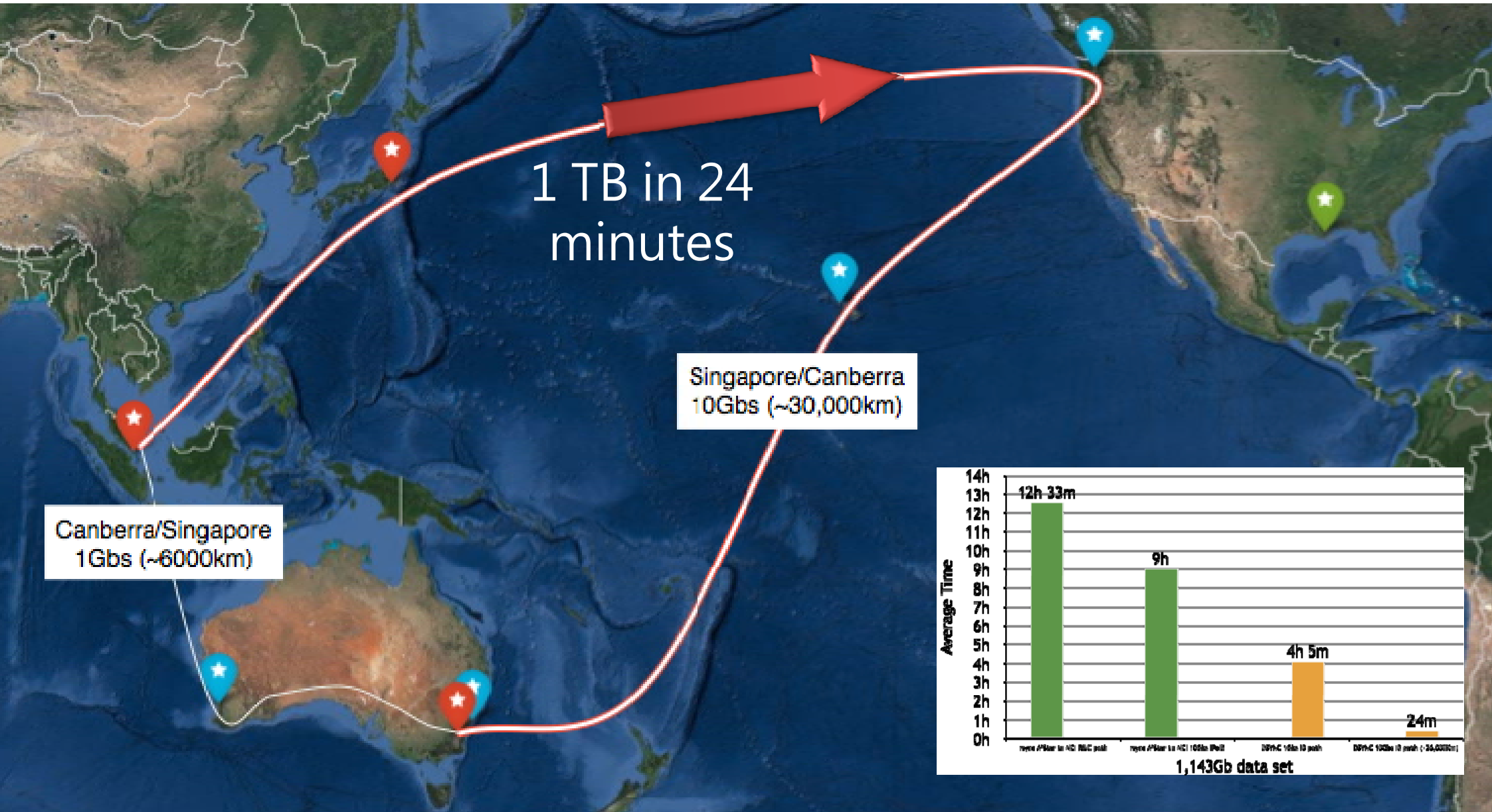
largest spanning InfiniBand network: ring-around-the-world with up to 100Gbps sections (Singapore-USA)



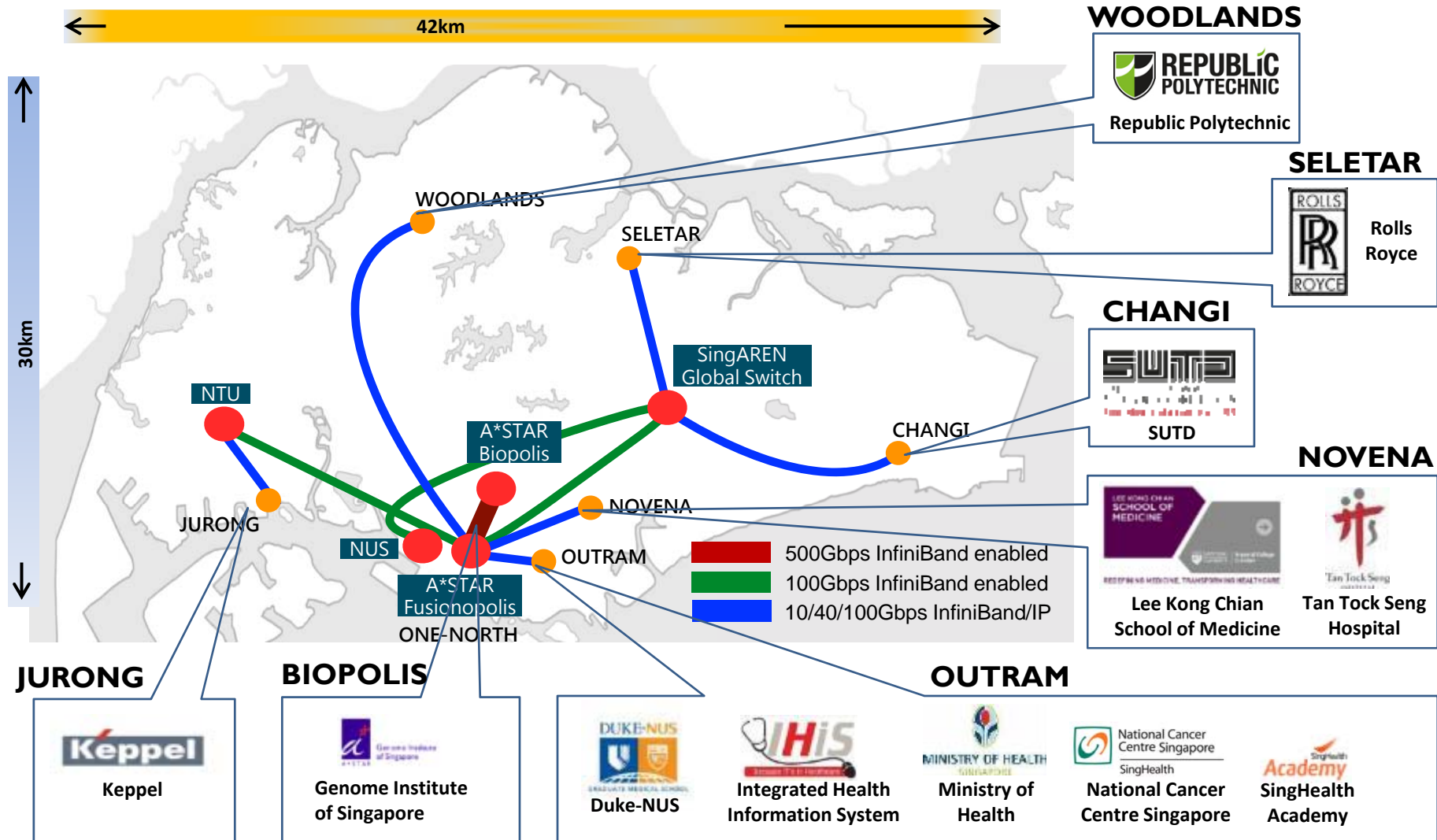
InfiniCloud at ISC 2016 Frankfurt, Germany



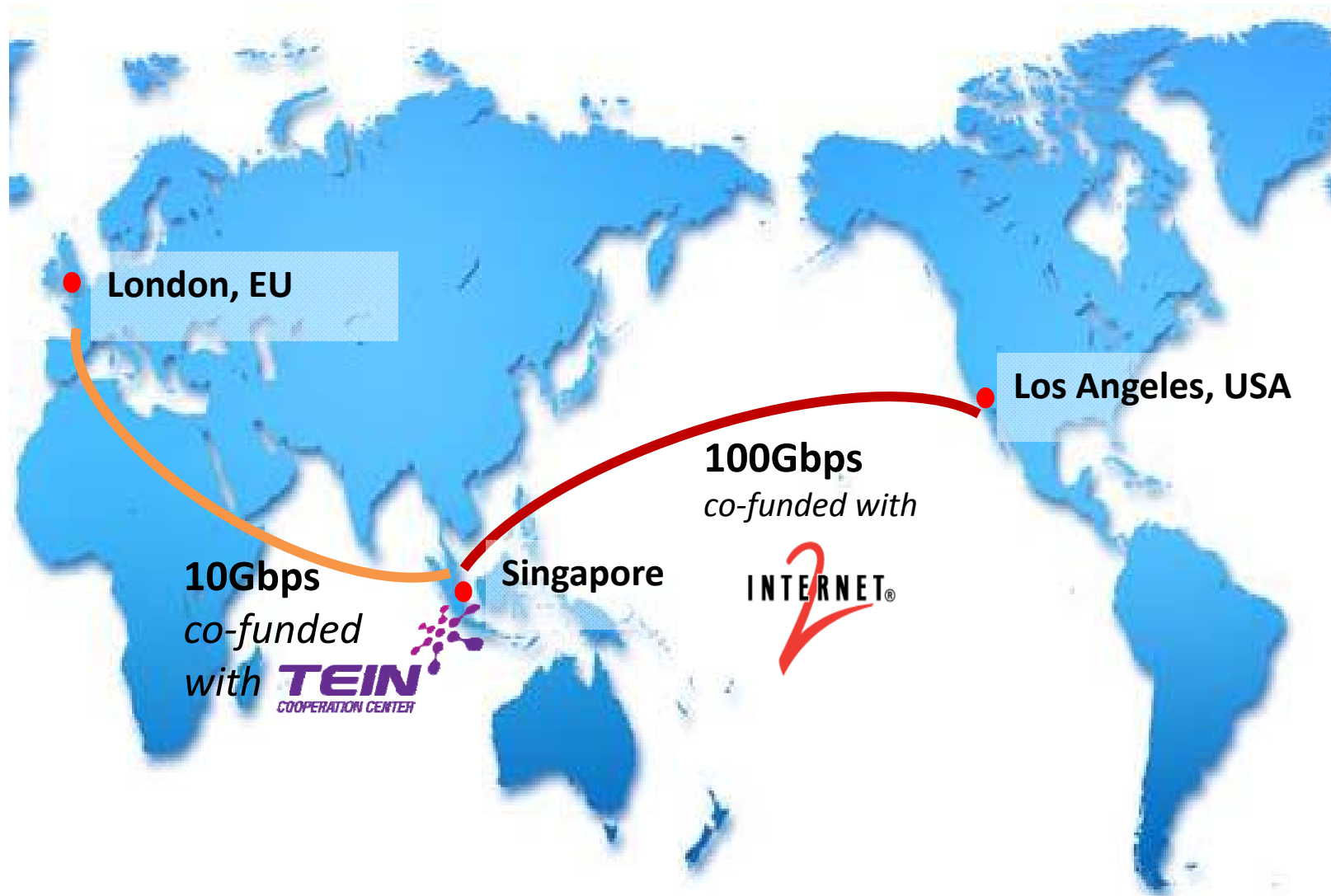
Data Transfer Test (dsync)



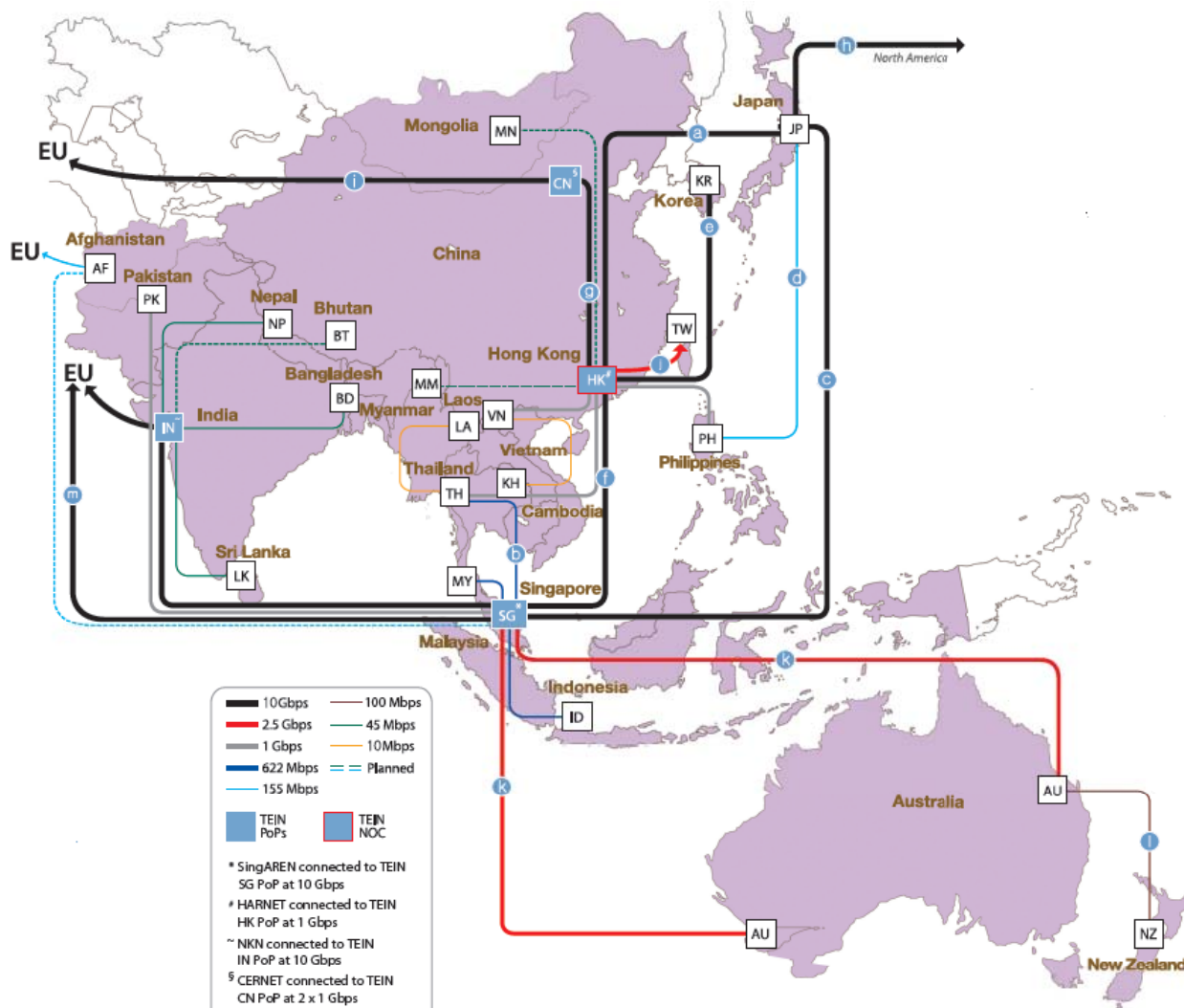
S'pore InfiniBand Connectivity and Fabric



Singapore Intercontinental Connectivity



Regional Network Connectivity (via SingAREN)



10 Gbps 100 Mbps
 2.5 Gbps 45 Mbps
 1 Gbps 10 Mbps
 622 Mbps Planned
 155 Mbps

TEIN PoPs TEIN NOC

- * SingAREN connected to TEIN SG PoP at 10 Gbps
- # HARNET connected to TEIN HK PoP at 1 Gbps
- ~ NKN connected to TEIN IN PoP at 10 Gbps
- § CERNET connected to TEIN CN PoP at 2 x 1 Gbps
- ¶ ThaiREN Connected to TEIN SG PoP at 600 Mbps
- ¶ GLORIAD-KR connected to TEI HK PoP at 1 Gbps
- ¶ Internet2 connected to TEIN SG PoP at 1 Gbps

TEIN Project Partners

AF Afghanistan	ID Indonesia	NZ New Zealand
AU Australia	JP Japan	PK Pakistan
BD Bangladesh	KR Korea	PH Philippines
BT Bhutan	LA Laos	SG Singapore
KH Cambodia	MM Myanmar	LK Sri Lanka
CN China	MN Mongolia	TH Thailand
HK Hong Kong	MY Malaysia	TW Taiwan
IN India	NP Nepal	VN Vietnam

The following links are fully financed/co-financed by the link owners whose support is gratefully acknowledged

a NICT National Institute of Information and Communications, Japan	g CERNET China Education and Research Network, China
b NICT National Institute of Information and Communications, Japan	h TrinityPAC co-funded by Japan and the USA
c NII National Institute of Informatics, Japan	i ORIENT plus co-funded by China and EU
d MAFFIN Ministry of Agriculture, Forestry and Fisheries Research Network, Japan	j ASGC Academia Sinica Grid Computing, Republic of Chinese Taipei
e NIA National Information Society Agency, South Korea	k aarnet Australia, Academic and Research Network, Australia
f NIA National Information Society Agency, South Korea	l REANVZ Research and Education Advanced Network New Zealand
	m National Supercomputing Centre, Singapore

As of August 2016

TEIN CC
 COOPERATION CENTER
www.teincc.org



TEIN is co-funded by the European Commission through the Directorate-General for Development and Cooperation-EuropeAid



Com Reso

Strategic Projects

Domain Areas



LIFE SCIENCES

Accelerate biomedical discoveries through **high performance applications in genomics**, thus improving the effectiveness of clinical treatments and personalised medicine.

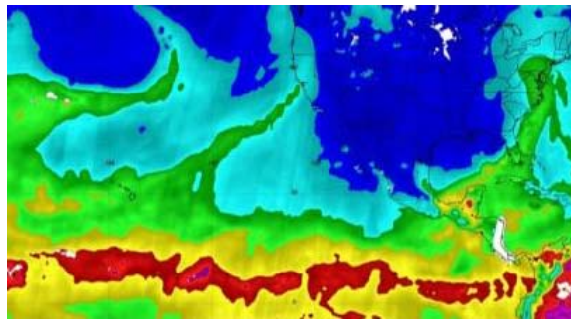


[Image courtesy of insideHPC]

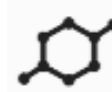


CLIMATE MODELLING

Contribute to atmospheric science and **improves the accuracy of weather forecasts** by broadening the range of parameters included in the simulations.

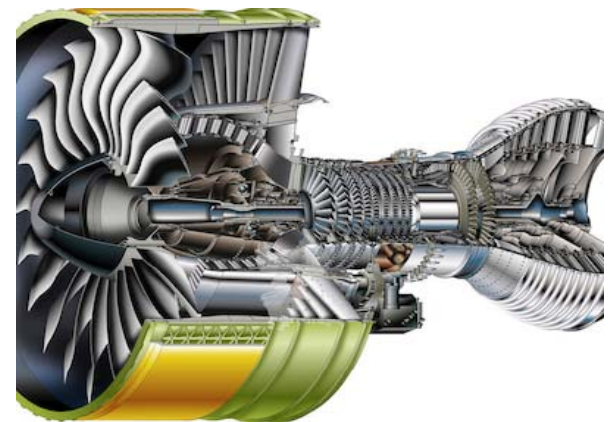


[Image courtesy of NASA]



MANUFACTURING

Enhance modeling, simulation and analysis to **speed up the design cycle for a faster time-to-market** for new and advanced products.



[Image courtesy of EnterpriseTech & Airbus]

Domain Areas



COMPUTATIONAL FINANCE

Perform high performance **computational modelling of market conditions, pricing model, risk models**, and contingencies to allow financial institutions to **accurately meet real-time goals**.



[Image courtesy of MIR Labs]

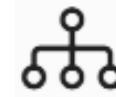


DIGITAL MEDIA PRODUCTION

Accelerate **rendering with high realism**, reduces time to market for producers and increases the quality of production for users.



[Image courtesy of Omens Studios]



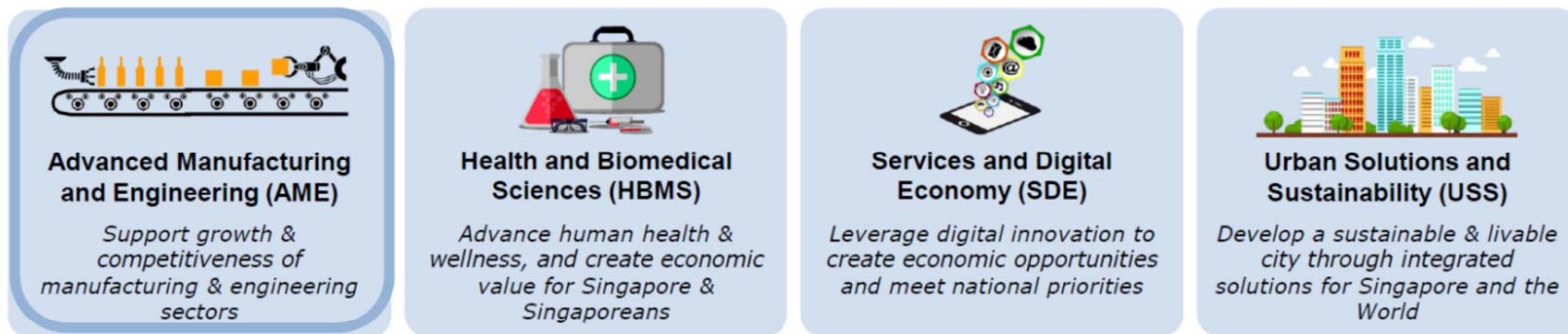
DATA CENTRE & NETWORKING

Offer an unprecedented high performance network testbed coupled with **high performance data analytics for quasi-real-time intrusion detection** and cybersecurity optimisation

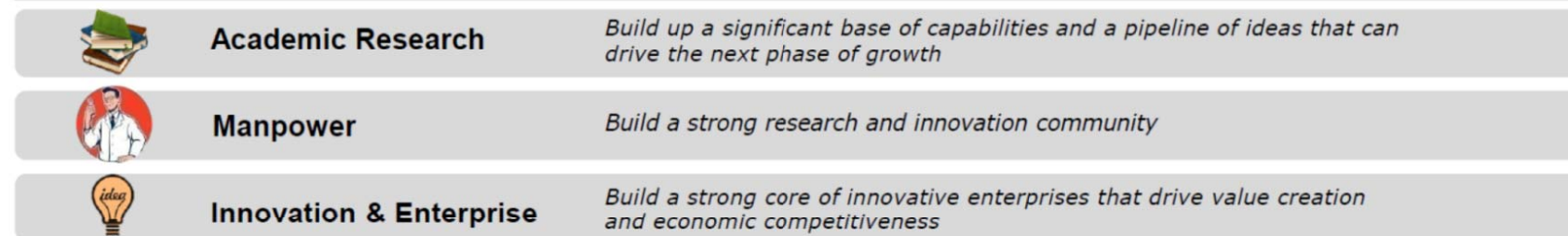


Strategic Focus

- To provide HPC resources for National R&D projects (RIE2020 / Science & Technology Research - \$19B)
- To identify within these domains, those which HPC can have maximum impact



CROSS CUTTING PROGRAMMES (HORIZONTALS)



Outreach & Manpower Capability Development

Outreach

Conduct Regular Training Workshops

Workshops

- Introductory
- Optimisation Techniques
- Parallel Profiling & Debugging
- Advanced Job Mgmt.

A*STAR



NTU



NUS



Local & Overseas Conference Participation



1 – 5 Aug 2016
Hong Kong



21–23 Sep 2016
Singapore



18 May 2016
Singapore



13 – 18 Nov 2016
Salt Lake City, Utah, USA



20 – 22 Jun 2016
Frankfurt, Germany



4 – 5 Oct 2016
Melbourne, Australia

Roadshows / Collaborations / Industrial Engagement

MOU with Industry (Mar 2016)



Industrial/Government Engagements



International Collaborations

- International collaborations with Supercomputing Centres (RIST, SDSC, ICAIRS, HKU, Compute Canada, ANU/NCI, PAWSEY)

**Research Organisation for
Information Science & Technology
(20 Jun 2016)**



with Dr. Motoi Okuda
(Deputy Head of Kobe Centre, RIST)

**iCAIRs
(20 June 2016)**



with Joe Membretti

**University of San Diego (UCSD)
(20 Jun 2016)**



with Prof. Mark Ellisman

**HKU
(3 Aug 2016)**



with Prof. Andy Hor

Thank you!