InfiniCortex: Present and Future

Invited Paper

Marek T. Michalewicz^{1*}
Jonathan Low^{4*}
Gabriel Noaje^{7*}
Jakub Chrzęszczak^{10***}

Tan Geok Lian^{2*}
David Southwell^{5**}
Dominic Chien^{8*}
Andrew Howard^{11***}
Liou Sing-Wu^{13*}

Lim Seng^{3*}
Jason Gunthorpe^{6**}
Yves Poppe^{9*}
Tan Tin Wee^{12*}

ABSTRACT

Commencing in June 2014, A*STAR Computational Resource Centre (A*CRC) team in Singapore, together with dozens of partners world-wide, have been building the InfiniCortex. Four concepts are integrated together to realise InfiniCortex:

- i) High bandwidth (~ 10 to 100Gbps) intercontinental connectivity between four continents: Asia, North America, Australia and Europe;
- ii) InfiniBand extension technology supporting transcontinental distances using Obsidian's Longbow range extenders;
- iii) Connecting separate InfiniBand sub-nets with different

¹Corresponding Author: michalewicz@acrc.a-star.edu.sg

²tanl@acrc.a-star.edu.sg

³lims@acrc.a-star.edu.sg

⁴jonathan-low@acrc.a-star.edu.sg

⁵dsouthwell@obsidianstrategies.com

⁶jgunthorpe@obsidianresearch.com

⁷noajeg@acrc.a-star.edu.sg

⁸chiensh@acrc.a-star.edu.sg

⁹ywpoppe@acrc.a-star.edu.sg

¹⁰jakub.chrzesczyk@anu.edu.au

¹¹andrew.howard@anu.edu.au

¹²tantw@acrc.a-star.edu.sg

¹³liousw@acrc.a-star.edu.sg

*A*STAR Computational Resource Centre

1 Fusionopolis Way, #17-01 Connexis, 138632 Singapore

**Obsidian Strategies Inc

4445 Calgary Trail NW #120

Edmonton, AB T6H 5R7, Canada

***The Australian National University, Canberra, ACT 0200, Australia

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored.

CF'16 May 16-19, 2016, Como, Italy

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4128-8/16/05.

DOI: <http://dx.doi.org/10.1145/2903150.2912887>

This work is licensed under a Creative Commons Attribution International 4.0 License.

net topologies to create a single computational resource: Galaxy of Supercomputers [10]

iv) Running workflows and applications on such a distributed computational infrastructure.

We have successfully demonstrated InfiniCortex prototypes at SC14 and SC15 conferences. The infrastructure comprised of computing resources residing at multiple locations in Singapore, Japan, Australia, USA, Canada, France and Poland. Various concurrent applications, including workflows, I/O heavy applications enabled with ADIOS system, Extempore real-time interactive applications, and in-situ real-time visualisations were demonstrated.

In this paper we briefly report on basic ideas behind InfiniCortex construct, our recent successes and some ideas about further growth and extension of this project.

CCS Concepts

•Computer systems organization → Distributed architectures; •Networks → Very long-range networks; Cloud computing; •Software and its engineering → Ultra-large-scale systems; Peer-to-peer architectures;

Keywords

InfiniCortex, Exascale computing, globally distributed concurrent supercomputer, InfiniCloud, InfiniBand, RDMA, ADIOS, in-situ visualisation, workflows, Extempore, data transfer, Square Kilometer Array, SKA

1. INFINICORTEX:

THE WORLD'S FIRST GLOBALLY DISTRIBUTED AND CONCURRENT GALAXY OF SUPERCOMPUTERS

1.1 Why InfiniCortex is very different from GRID, Cloud and Internet connected resources?

1.1.1 RDMA - the great differentiator

RDMA (Remote Direct Memory Access) is a hardware acceleration technology that enables direct zero-copy data movement between user-space virtual memory regions residing on remote servers linked by RDMA-capable fabrics.

Extensive offloading ensures that CPU loading remains minimal and since OS buffers are avoided memory system loading is reduced to just the direct read and write between user-space source and target regions, minimising end to end latency and latency jitter. CPU intensive protocol stacks present computational limits on data transfer rates whereas RDMA-based transfers are limited only by host memory bandwidth which is readily scalable; single very high capacity flows can both originate from and arrive onto single server nodes. The open standard InfiniBand Architecture is currently the dominant RDMA-capable interconnect used by more than 50% of technical computing clusters and supercomputers. Routinely used for inter-process communication within LAN clusters and for Storage Area Networks (SAN), InfiniCortex leverages InfiniBand range extenders to capture the benefits of RDMA transport across global WAN circuits.

1.1.2 Highly efficient bandwidth utilisation

The challenge of efficiently driving high bandwidth*delay product WAN links using lossy protocols such as TCP/IP is well documented. Even across 'straight shot' uncontested point to point connections the packet discard mechanism for handling transient congestion conditions can trigger exaggerated and performance-crippling cycles of retransmission and insertion rate throttling. TCP/IP was designed for dramatically smaller flows and shorter distances than the links called for by InfiniCortex. Range extenders such as Obsidian's Longbow E100s maintain InfiniBand's lossless flow control semantics across WAN segments. InfiniBand's lossless flow control scheme functions by suspending instantaneous transmission onto a segment whenever the hardware-maintained link-layer state machine signals the possibility of buffer overrun at the receiver. Unlike TCP/IP's optimistic heuristics, InfiniBand's precisely calculated approach avoids routine retransmission and throttling cycles due to congestion, because overflow conditions are simply never allowed to develop. Lossless WAN link characteristics combine with the advantages of RDMA to deliver near ideal WAN link behaviours - sustained maximum throughput flows between user-space applications with minimal latency and virtually zero jitter at high bit rates across global distances.

1.1.3 Hardware encryption

InfiniCortex connects numerous globally distributed HPC resources owned and operated by different agencies and countries, and routinely carries data demanding protection from unauthorised access. The Longbow E100 devices provide integrated cryptography hardware to meet these requirements. In-line SHA-1 and advanced key management provides continuous end-point authentication to ensure each party's identity has not been compromised by a man-in-the-middle agent. WAN encapsulated data is encrypted in flight using AES-192-GCM (Galois Counter Mode - a variant of the AES standard intended for efficient implementation in silicon), providing better than financial services grade protection. The implementation has no measureable effect on throughput performance and adds a negligible 100ns to the device's port-to-port latency. Since these cryptographic functions are integrated into the hardware data path of the range extenders they do not load the sender or receiver CPUs. In contrast, implementing encrypted TCP/IP connections server-side often has a further serious impact on

link throughput, latency and jitter.

1.1.4 InfiniBand routing

It is possible to implement a network such as InfiniCortex as a single InfiniBand subnet, however as a matter of practicality it is much more desirable to segment the network into multiple interconnected InfiniBand subnets. A multi-subnet fabric offers administrative control and sovereignty in each region as well as fault isolation, performance scaling and virtually unlimited aggregate node-count scaling, breaching the 48k node limit of an InfiniBand subnet's LID address space. InfiniCortex is currently implemented as a compound fabric combining seven subnets, each implementing various internal topologies distributed across the globe. Multi-subnet InfiniBand fabrics are implemented with Obsidian Crossbow R400-6 devices (six-port QDR speed native InfiniBand routers). Each subnet hosts a cluster ($n \geq 1$) of Bowman Global Fabric Controller (BGFC) nodes that also serves as InfiniBand subnet manager. BGFC has a number of unique features to equip it for this role, including a deterministic LFT generator based on idealised geometry solutions that are proveably incapable of deadlock, a mechanism to aggregate arbitrary combinations of dissimilar cluster topologies, n-way active:active implementation for fault tolerance and performance scaling and persistent configuration databases for reproducible networks and immediate detection of discovered topology deviations, such as bad cables.

1.2 First global scale proof-of-concept of InfiniCortex infrastructure

A*CRC team headed by Dr Marek Michalewicz conceived and has been leading the InfiniCortex project. Since June 2014, A*CRC has assembled a collaboration of several dozen partnering institutions, and, by November 2014, demonstrated the first Trans-Pacific 100Gbps InfiniBand-connected concurrent supercomputer composed of HPC resources on three continents: Australia, Asia and North America; in four countries: Australia, Japan, USA and Singapore; with participation of seven Universities: Tokyo Institute of Technology, Japan; ANU, Australia; Georgia Institute of Technology, Rutgers, Stony Brook University, University of Tennessee from USA; the NTU; and three research Organisations: A*CRC, Oak Ridge National Laboratory, USA and National Computing Infrastructure, Canberra, Australia. A*CRC was also supported by seven research network organisations and thirteen commercial partners and sponsors [8].

The two main commercial partners were Obsidian Strategies, that provided InfiniBand range extenders and routers, and Tata Communications that provided 100Gbps trans-Pacific connectivity.

At Supercomputing 2014 conference in New Orleans the team demonstrated transmission of data at speeds of up to 100Gbits/s directly from the A*CRC datacentre to the A*CRC booth at the exhibition floor of the convention centre [8]. This throughput was a record-breaker as it delivered data 100 times faster than the research link between Singapore and USA in use at the time. It was also a record speed of transmission over a single-wave between Asia and North America.

The demonstration merged four separately important and interesting concepts integrated for the first time together to realise InfiniCortex:

i) High bandwidth intercontinental connectivity between Asia

and the USA;

- ii) InfiniBand technology over trans-continental distances using Obsidian Strategies range extenders;
- iii) Connecting separate InfiniBand sub-nets with different sub-net topologies to create a single computational resource: Galaxy of Supercomputers
- iv) Running workflows and applications on such a distributed computational infrastructure, especially using ADIOS I/O framework.

InfiniCortex project was one of the 15 best projects selected for the special 'Emerging Technologies' track of the conference [8] and was also featured throughout SC14 at the A*CRC booth.

The demonstration consisted of several real use cases, applications and ADIOS enabled workflows:

1. Demonstration of remote parallel file transfer using dsyn+ [3] and Lustre file system on a remote computer resources (Tokyo Tech);

2. InGeneoS: Intercontinental Genetic sequencing over trans-Pacific networks and Supercomputers. The project consisted of several parts and demonstrated several unique achievements:

i) ability to create a genuine 'HPC Cloud' instances of arbitrary size, and interconnected with InfiniBand (InfiniCloud) [5, 7]

ii) sequence alignment workflows for input files of up to 1.2TB and processed either at NCI InfiniCloud instances, or on A*CRC's large SMP Aurora computer with 12TB shared memory. It took 12.5 hours to transfer a 1.2TB file over 26,000km using TCP/IP but with long-range InfiniBand that time was reduced more than 30 times to 28 minutes. This demonstrated that hospital generated genomic sequence data could be sent and processed on a remote supercomputing resources almost anywhere, and would return diagnostic quality results to the doctor's tablet, at the patient's bedside, in less than 30 minutes! (A*CRC and NCI, Australia);

3. Near Real Time (NRT) Plasma Disruption Detection in Nuclear Fusion research reactors (part of EPSI project) (ORNL, ADIOS Group, and Princeton Plasma Research Laboratory);

4. Microscopy Image Analysis in Cancer detection with real time feature detection in images of up-to 120k×120k pixels (Stony Brook University and ORNL);

5. Nanotechnology and design of materials: Molecular Dynamics using LAMMPS with staging over distributed resources (Georgia Tech and ORNL);

6. Synthetic tests of very fast file transfer using dsync+ reaching the physical limit of the network (A*CRC);

7. Sub-net detection using Crossbow InfiniBand routers: successful detection of sub-nets and traffic control (A*CRC and Obsidian Strategies).

2. INFINICORTEX DEMONSTRATION AT SUPERCOMPUTING 2015 [9]

At SC15 expo floor in Austin, TX, A*CRC presented the largest InfiniCortex infrastructure to date (Figure 1) Over one and half days, at our expo booth, we've built a substantial installation of necessary hardware consisting of one rack with 10 HP servers, 14 Longbows E100, one Crossbow R400 router, one Longbow C400 and one rack of DDN Lustre storage. This InfiniBand network run on top of the

Figure 1: Geographical distribution of InfiniCortex 2015

Figure 2: Seven separate sub-nets realised during SC15 demo

dark-fibre network infrastructure prepared by A*CRC Network in collaboration with almost 20 network organisations from around the world.

For the first time the InfiniBand network consisted of seven *separate* Infiniband subnets (Figure 2):

- A*CRC - Singapore and Stony Brook University - USA
- SC15 show floor - USA
- Poznań Supercomputing and Networking Centre - Poland
- University of Reims Champagne-Ardenne - France
- Georgia Tech - USA (Clos3 networked cluster)
- National Computational Infrastructure - Australia
- University of Alberta and Obsidian Research - Canada

2.1 Performance metrics

To meet the challenge issued by SCinet to demonstrate peak bandwidth usage of the 100Gbps network link between Singapore and Austin, a sustained 80Gbps of bandwidth traffic was utilised, peaking to 140Gbps with help from Tokyo University.

Each of the ten Longbows range extenders on the 100Gbps link was able to utilise maximum bandwidth of 10Gbps in

both sending and receiving directions without the need for rate limiting.

The Longbow range extender connected to Stony Brook University was limited to a total of 6Gbps bandwidth, either 6Gbps uni-directionally or 3Gbps bi-directionally. This correlates to the bandwidth availability paid for by the team at Stony Brook, which is 5Gbps of bandwidth with a allowable burst-buffer of 1Gbps.

Using a Longbow on the 100Gbps link, a peak speed of 818MB/s was observed during a dsync+ data transfer exercise. As a comparison, a 1GB file transfer using Aspera software achieved a transfer rate of 150MB/s (1201476Kbps).

2.2 Partners

InfiniCortex partners involved in the SC2015 InfiniCortex project were:

- National Computational Infrastructure (NCI) - Australia
- Poznań Supercomputing and Network Centre (PSNC) - Poland
- University of Reims Champagne-Ardenne (URCA) - France
- Georgia Tech Institute of Technology - USA
- Stony Brook University - USA
- University of Alberta - Canada

Each partner brought in specific expertise and own codes and applications which were run on top of the ring-around-the-world InfiniBand network.

2.3 Network Architecture

The InfiniCortex is applied as an overlay application on the vast network infrastructure built for this purpose. To support it, the underlay backbone network infrastructure has to be built towards the SC15 booth. Collaborators have worked closely with us to build the network connectivity towards the convention venue. We have commissioned network connectivity for the following endpoints:

- Singapore towards SC15 venue Trans-Pacific link at 100 Gigabits/sec. This capacity involves deep collaboration with many partners as no one has full ownership of the capacity at any point. This was achieved through:
- Poznań, Poland (PSNC) Trans-European and Trans-Atlantic link towards SC15 venue at 30 Gigabits/sec.
- Existing 10 Gigabits/sec Trans-Pacific connectivity to ESnet.
- Singapore towards GÉANT in London via the Trans-Eurasia Information Network (TEIN) at 10 Gigabits/sec. This is the highest bandwidth capacity ever commissioned between A*STAR Singapore and Europe for Research & Education purpose.
- 200 Gigabits/sec links into A*STAR booth at SC15 which allows InfiniCortex to connect to the outside world.

We have installed the latest equipment from leading network solution providers like Ciena and Arista to make the global infrastructure possible. We have installed the Arista 7504E with 30 Terabits/sec switching capacity, and also demonstrated SingAREN's capability to transport more than

100 Gigabits/sec aggregated data to A*STAR datacentre at A*STAR Fusionopolis. This is in addition to A*STAR's existing infrastructure consisting of equipment from Cisco, Ciena etc.

The global network architecture for SC15 have a very high capacity and coverage which rivals that of a telecom carrier in scale, and made possible by many sponsoring and collaborating industry leaders.

2.4 Demonstrations with NCI Australia, University of Reims and Stony Brook University [6]

A*STAR staff in collaboration with partners in Australia, Europe and the USA used the connectivity provided by the InfiniCortex fabric to build the first native InfiniBand High Performance Cloud - InfiniCloud 2.0 [6].

InfiniCloud 2.0 was built using a customized OpenStack [4] installation running on top of CentOS7 [2] operating system. It combines computational resources available across four continents (Asia, Australia, North America and Europe) into one system with a single interface (dashboard, CLI and API) - which can be used as a single, globally distributed supercomputer. InfiniCloud 2.0 users have the freedom of utilizing computational and storage resources spanning across four continents according to their needs and requirements - and the ability to either distribute processing evenly across the sites (demo: Geopipeline), or to request a specific locations (demo: Extempore).

High bandwidth, native InfiniBand fabric gives users freedom to either utilize high-speed, efficient data transfers using Obsidian Strategies dsync+ [3] or to access data in-place, using a specifically tuned instance of Fraunhofer Institute BeeGFS parallel file system [1].

2.4.1 Genomics Geopipeline

An intercontinental genetic sequencing pipeline by Prof. Kenneth Ban from National University of Singapore was run over InfiniCloud infrastructure, utilizing ElastiCluster tool to create a virtual cluster made of instances of arbitrary size, on demand and interconnected with InfiniBand. The input files were sourced from the remotely mounted parallel file system hosted in Singapore and the computational HPC Cloud resources were created instantly across four continents. Compute nodes 1 and 4 were located at NCI in Australia, node 2 was at A*STAR, Singapore, nodes 3 and 5 were at Stony Brook University in New York, node 6 at University of Reims, France and the dashboard was accessible from the show floor in Austin. The embarrassingly parallel nature of the codes made it possible to transparently scale the cluster beyond the size of locally available pool of resources at each and every site.

2.4.2 Plasma Physics PIC modelling with real-time interactive model control with Extempore

Yet completely different demonstration was a Plasma Physics Particle-in-Cell (PIC) code of Victor Decyk, Henry Gardener and Ben Swift, run and augmented with Extempore language code created by Andrew Sorensen, allowing real-time, interactive steering of the physics simulations - given the nature of communications-intensive, latency sensitive code, this process was carefully distributed over only two continents: the PIC code run in Australia with results displayed and interactive code modifications and steering

performed in Austin, USA. The PIC simulation was additionally embellished with Extempore created 'sonification' of the physics process, adding musical effects to a plasma physics simulations!

The above demonstrations showed the potential of InfiniCortex to solve the difficult computational problems by joining supercomputing resources available across physically distant sites on the global scale, allowing users to scale beyond the resources available locally, creating a 'Galaxy of Supercomputers' [10]. A video recording of the demo is archived and is available from A*CRC.

2.5 Other InfiniCortex demonstrations at SC15

2.5.1 Demonstrations with Georgia Tech

Particle tracking using data stored at A*CRC resources in Singapore. The applications were compiled and executed on A*CRC resources at the SC15 booth. The computation and visualisation was performed and a video recording of the demo is archived and available from A*CRC.

2.5.2 Demonstrations with University of Alberta

At it's expo booth, Obsidian Research conducted the demonstration with the University of Alberta of efficiently migrating huge genomic datasets across global distances, leveraging the global InfiniBand network, with InfiniBand routing in place and linked to the A*CRC booth at SC15.

2.5.3 Demonstrations with Oak Ridge National Labs

XGC is a gyrokinetic Particle-in-Cell code designed to model the development of an edge pedestal in the radial density and temperature profiles of tokamak fusion plasmas. The code also computes scrape-off and wall loss physics. In XGC, the gyrokinetic plasma ion and electron guiding centers evolve in time within a five-dimensional phase space using a realistic magnetic equilibrium and limiter geometry.

The code is extremely computationally intense and runs on entire Titan system in ORNL. The cost of running the code per day is \$300,000.

2.5.4 Shared Network Architecture for University of Tokyo/National Institute of Information and Communications Technology, Japan

Not only did the network infrastructure support InfiniCortex, we also provided the network elements to support our partner for SC15: University of Tokyo and NICT for their demonstration. Professor Kei Hiraki, University of Tokyo, led a demonstration to display their research on 100Gbps Long Fat Networks with High Delay Product. A virtualized network routing environment on this same infrastructure was built based on their needs and the demonstration was performed successfully from SC15 convention floor. Due to high bandwidth needs from InfiniCortex and University of Tokyo, we agreed on time-sharing the capacity based on goodwill arrangement.

During the exhibition, we were seeing upwards of 80 Gigabits/sec of usage by InfiniCortex and over 60 Gigabits/sec from University of Tokyo. In total, we were peaking at 140 Gigabits/sec in bandwidth on the network architecture.

3. BENEFITS OF INFINICORTX PROJECT

As a direct result of our InfiniCortex demonstrations the following events occurred over the last two years:

1. Internet2 has last year signed off a deal with SingAREN for 50%-50% bandwidth sharing across Pacific for the first 100Gbps link - this translates to tremendous savings for Singapore and immediately provides 10x increase in research bandwidth available between the two countries and positions Singapore in easy reach for Big Data sharing with the USA,
2. TEIN*CC has signed off a similar 50%-50% bandwidth sharing deal with SingAREN for a 10Gbps link with Europe (landing in London).

Both of these two agreements were accelerated by and were a direct result of our InfiniCortex demonstrations.

InfiniCortex is not only about data transfer (or BIG data transfer) - it is about creating a Global Supercomputer (or Galaxy of Supercomputers). We have last year created seven separately routed sub-nets using InfiniBand - the first ever. We have also demonstrated true instances of HPC Cloud which spanned four continents - never done before! With extended range InfiniBand we can utilise close to 100% of available bandwidth, while with traditional Internet (TCP/IP) the bandwidth utilisation is about 15-20%. This translates into huge savings in infrastructure cost and utilisation.

The entire test-bed we've created is extremely important to commercial cloud services, hyper-scale datacenter providers and carriers. Imagine datacentres in Norway or Iceland, where electricity is cheap and abundant. If they are going to be commercially successful, they must receive and send data in huge volumes and very efficiently - but what technology is best for this purpose?

The largest experimental and observational facility to be created soon, the Square Kilometer Array (SKA) in Australia and South Africa, will be producing data at a rate of >1 PBytes a second - how will the observational data be transmitted to the rest of the world for processing and analysis?

We hope InfiniCortex, with close to 100% bandwidth utilisation and hard-wired encryption will be one of the most attractive and effective solutions.

3.1 Research collaborations resulting from InfiniCortex project

A*CRC has already signed Memoranda of Understanding (MoU) with:

- Reims University (France),
- Poznań Supercomputing and Networking Centre (Poland),
- Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw (Poland),
- Stony Brook University (USA),
- George Washington University (USA),
- Indian Institute of Science Education and Research (IISER) in Pune, (India).

Establishing high bandwidth permanent networks over global distances combined with superior computing facilities lead to multitude of new, joint research collaborations. A*CRC began new collaborations with Princeton Plasma Research Laboratory, National Computing Infrastructure and the Australian National University, University of Alberta (Canada), Georgia Tech (Atlanta, USA), Oak Ridge National Laboratories (DoE, USA), Tokyo Tech (Japan), *etc.* - the unprecedented number of new, highest quality research collaborations established in a very short time - largely due to the value of InfiniCortex concept and infrastructure we bring to the collaborations.

We have received millions of dollars worth of support, services and equipment as donations from dozens of commercial sponsors - the top tier league of technological corporations - just to take part in this project.

3.2 Further plans for InfiniCortex

3.2.1 Singapore Nation-Wide InfiniBand fabric

The newly created Singapore National Supercomputer Centre (NSCC) and A*CRC are working very closely on implementing a nation-wide InfiniBand fabric to interconnect NSCC with National University of Singapore, Nanyang Technological University, Singapore University of Technology and Design, National University Hospital and other stakeholders (Figure 3). The NSCC datacenter site, which is co-located in the same building as A*CRC datacenter, and which hosts a new Peta-FLOPS supercomputer with state-of-the-art 13 Peta-Bytes of HPC storage will have InfiniBand connectivity to login nodes at all these sites. The InfiniBand network will further expand to industrial sites in Singapore (Seletar, Changi Business Park, Jurong, *etc.*). This will provide high throughput, low latency direct connection to the supercomputing facilities in Fusionopolis. This infrastructure will provide Singapore's researchers an unparalleled 'at-your-fingertip' easy access to HPC resources. This initiative to provide easy access to the HPC resources will have positive impact on the research and economy of the nation. The network architecture and the technology used to implement it in this country-wide operational, production infrastructure has been prototyped and tested through InfiniCortex project.

3.2.2 Plans for InfiniBand connected Data Transfer Nodes (DTN) at Point-Of-Presence (POP) in London and USA

NSCC and A*CRC are also looking into implementing Data Transfer Nodes to be placed in Europe and US POPs, running InfiniBand across the transpacific distances. The work is progressing and we are also talking to Pacific Research Platform which has plans of creating a 'Global Research Platform' for big scientific data exchange. We are also exploring various new technologies to implement the most suitable DTN solutions.

3.2.3 Further geographical extension of the InfiniCortex network and new partnerships

This year we are going to connect with several Indian institutes of higher education, and strengthen our links with Australia. We also plan to extend InfiniCortex to South America, Africa and Arabian Peninsula - subject to available resources and partnerships in those locations.

Figure 3: Singapore InfiniBand research network providing access to NSCC Supercomputing facilities. (Green - existing connections; Red - 500Gbps (5x100) single pair of fibre Infinera CloudX connection; Blue - proposed connections to R&D, industrial and hospital campuses)

A*CRC staff have been working for 12-16 hours a day for months on end to accomplish these feats. Communication and discussions with partners at all geographical locations forced us to have frequent conference calls at weird late hours. And something extraordinary: the ring-around-the-world InfiniBand connectivity has been created by A*CRC Network Group which consists of *only two staff members!* A*CRC Network Group staff has accomplished this with tremendous amount of dedication and extra long hours of work, and they deserve the highest recognition. I must put on record an outstanding performance by Ms Tan Geok Lian, The Network Group Manager, and Mr Lim Sen, our Network Engineer.

An efficient system provisioning mechanisms is an inevitable part for the functional Galaxy of Supercomputers. In order to meet this requirement we have been working with the engineers from Bright Computing to implement and evaluate the mechanisms for remote system provisioning, and we expect to demonstrate this functionality in SC16.

Currently A*CRC is partnering with researchers from Ludwig-Maximilians-University, Munich and Leibniz Supercomputing Centre, Garching, Germany in a project proposal entitled 'Singaporean-German Urgent computing platform based on Global distance in infiband netwOrks (SUGO)' to explore InfiniCortex infrastructure for a dedicated, reliable and robust delivery of 'urgent' computing in situations such as flush-flooding or natural disasters.

These are just few examples of what new goals we have in the short time-frame this year, but it is our imagination that limits us or provides us with further ideas.

3.3 Publicity, Promotion and Recognition of InfiniCortex

We have already presented InfiniCortex at several international conferences, each time with noticeable impact:

1. Supercomputing 2014 (Nov, New Orleans, USA) - we have had about 300 visitors in our booth and we were invited to present at a special, top ranking session: 'Emerging Technologies'
2. Supercomputing 2015 (Austin, Texas, USA) - 200+ visitors stopped at our booth, but this time InfiniCor-

tex information and banners were displayed at eight different booths by our partners GEANT, Compute Canada, Obsidian Strategics, Poznań Supercomputing and Networking Centre, *etc.*

3. International Supercomputing Conference, Frankfurt, 2015
4. TERENA conferences: 2014 (Dublin, Ireland) and 2015 (Porto, Portugal) - special sessions dedicated to InfiniCortex
5. APAN 38, 39, 40 - local SE Asia networking conferences - with special sessions and presentations of InfiniCortex
6. Internet 2 regional work-group meeting Abu Dhabi, 2014
7. ANKABUT conference 2015
8. ASREN conferences 2014 and 2015
9. HPC workshop at University of Reims, France, 2015
10. ISGC 2014, Taiwan
11. Big Data and Exascale Computing Workshops in Spain and at ISC15 in Germany
12. Big Data Forum, China, Nov 2015
13. 12th Annual OpenFabrics Alliance Workshop, Monterey, USA, April 2016

The term 'InfiniCortex', which the author coined, has been referred hundreds of times by various on-line reports and sites.

The A*CRC group which was responsible for the InfiniCortex project was awarded the following honours in 2015:

- A*STAR Awards 2015: STAR Innovation Award for InfiniCortex (Galaxy of Supercomputers) Project
- Innovative Project GOLD Award: for InfiniCortex Project, Ministry of Trade and Industry, Singapore 2015
- Technology Leadership Award: Agency for Science, Technology and Research; FutureGov, Singapore awards
- PS21 ExCEL Awards Most Innovative Project/Policy (Merit) - Public Service Awards Singapore 2015

4. CONCLUSIONS

InfiniCortex has girded the globe. This is an early time in the development of this new type of concurrent computing infrastructure. This year, in 2016, we plan to strengthen our links with Australia by increasing the bandwidth, opening new paths and increasing the range of applications - especially in the context of fast and efficient observational data transfer from the Square Kilometer Array.

We also plan to open-up new connection with India and several new terminal points in the USA and other geographies. We also plan to grow tentacles of InfiniCortex into embedded devices and sensors which are going to be components of Internet of Things of the future.

5. ACKNOWLEDGMENTS

A*CRC team wishes to thank all our collaborators at home and around the world: Kenneth Ban, Vincenzo Capone, Francis Bu Sung Lee, Lawrence W.C. Wong, Scott Klasky, Yuefan Deng, Firat Coskun, Matthew Wolf, Manish Parashar,

Tahsin Kurc, C.S. Choong-Seock Chang, Artur Binczewski, Michal Kulczewski, Piotr Kopta, Krzysztof Kurowski, Arnaud Renard, Michael Krajecki, Satoshi Matsuoka, Shin'ichi Muira, Nils Otto vor dem gentschen Felde, Cerlane Siew Hoon Leong, Torsten Wilde, Dieter Kranzlmuller, Andrew Sorensen, Henry Gardener, Ben Swift, Viktor Decyk, Pieter Tsiknas, well as all contributors and supporters from Obsidian Strategics, SingAREN, NTU, NUS, ESnet, Internet2, SCinet, TEIN*CC, NCI, AARNET, GEANT, PSNC, Pioneer, RENATER, Universite de Reims, ORNL, Rutgers University, Georgia Tech, Stony Brook University, University of Alberta, CANARIE, Cybera, Compute Canada, Tata Communications, Data Direct Networks, Arista Networks, Ciena, Juniper, Intel, Mellanox, Nvidia, SGI, HPE, Aspera, Altair, Bright Cluster Computing, ByteScale, BlueTel, Pacific Wave, Century Link, 3D Networks, and others.

6. REFERENCES

- [1] www.beegfs.com.
- [2] www.centos.org.
- [3] www.obsidianresearch.com/products/software/dsync+/index.html.
- [4] www.openstack.org.
- [5] K. H. K. Ban, J. Chrzęszczczyk, A. Howard, D. Li, and T. W. Tan. InfiniCloud: Leveraging the Global InfiniCortex Fabric and OpenStack Cloud for Borderless High Performance Computing of Genomic Data. *Supercomputing Frontiers and Innovations*, 2(3):14–27, July-Sept 2015.
- [6] J. Chrzęszczczyk, A. Howard, A. Chrzęszczczyk, J. Low, K. Ban, T. W. Tan, B. Swift, and P. Davis. InfiniCloud 2.0: Distributing High Performance Computing across continents. In *Supercomputing Frontiers 2016, Singapore*, March 2016.
- [7] J. Low, J. Chrzęszczczyk, A. Howard, and A. Chrzęszczczyk. Performance Assessment of InfiniBand HPC Cloud Instances on Intel Haswell and Intel Sandy Bridge Architectures. *Supercomputing Frontiers and Innovations*, 2(3):28–40, July-Sept 2015.
- [8] M. Michalewicz, D. Southwell, T. W. Tan, Y. Poppe, S. Klasky, Y. Deng, M. Wolf, M. Parashar, T. Kurc, C. C.-S. Chang, S. Matsuoka, S. Muira, J. Chrzęszczczyk, and A. Howard. InfiniCortex: concurrent supercomputing across the globe utilising trans-continental InfiniBand and Galaxy of Supercomputers. *Supercomputing 2014: The International Conference for High Performance Computing, Networking, Storage and Analysis, At New Orleans, LA, USA*, November 2014.
- [9] G. Noaje, M. T. Michalewicz, G. L. Tan, S. Lim, Y. Poppe, T. W. Tan, D. Southwell, J. Low, D. Chien, S.-W. Liou, J. Chrzęszczczyk, A. Howard, and J. Gunthorpe. Around The Globe Towards Exascale: InfiniCortex Past and Present. In *Supercomputing Frontiers 2016, Singapore*, March 2016.
- [10] L. Orłowski, Y. Deng, and M. Michalewicz. Galaxies of Supercomputers and their underlying interconnect topologies hierarchies. *International Supercomputer Conference, Leipzig, Germany*, June 2014.