National Supercomputing Centre

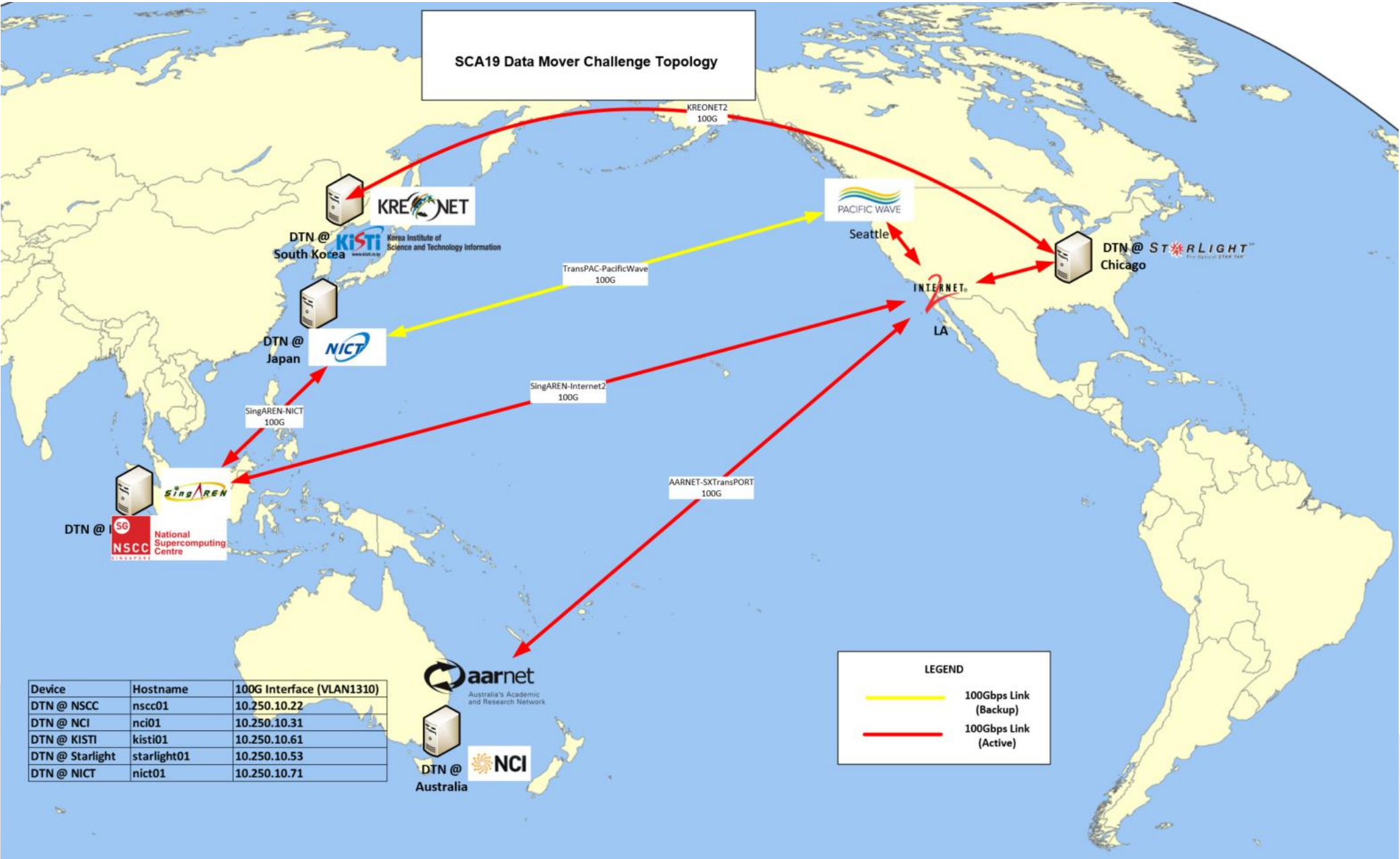
Data Mover Challenge 2019/20

Alvin Chiam
Senior Specialist II (Network Engineer)

Data Mover Challenge 2019

- Inaugural **Data Mover Challenge 2019 (DMC19)** organised by NSCC
- Held in conjunction with NSCC's flagship conference - **SupercomputingAsia 2019 (SCA19)**
- Bring together experts from industry and academia
- To test their software across servers in various countries connected by 100G international networks
- To share knowledge and solutions in data transfers over long distance, high latency international links

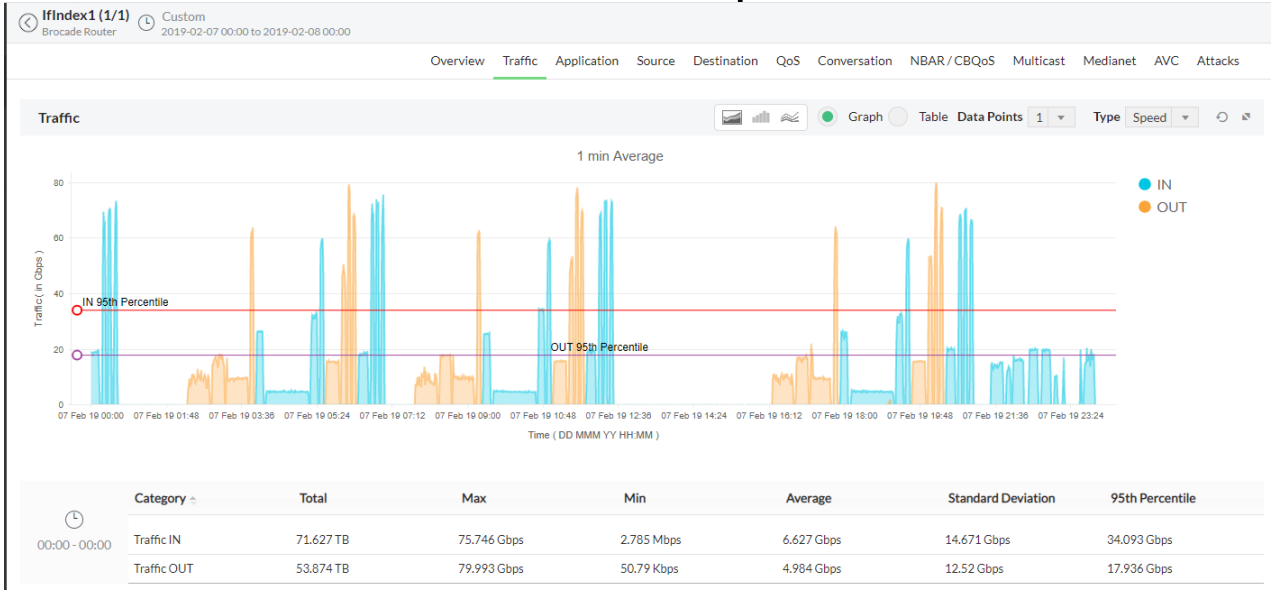
DMC19 Partners and Topology



Device	Hostname	100G Interface (VLAN1310)
DTN @ NSCC	nscc01	10.250.10.22
DTN @ NCI	nci01	10.250.10.31
DTN @ KISTI	kisti01	10.250.10.61
DTN @ Starlight	starlight01	10.250.10.53
DTN @ NICT	nict01	10.250.10.71

Partner DTN Setup

- Basic OS and Software
  - **Centos7.5, Iperf3, Singularity**
- OS and 100G NIC tuning according to ESnet FasterData Guide
- /DMC directory created on each DTN with subdirectory
  - /test - a READ/WRITE directory as the destination the data set transfer
  - /data - a READONLY directory as the source of the data set
- Bandwidth utilization and statistics captured at NSCC and SingAREN



DMC19 Competition Scope

- 2TB data set consisting of Genome and Satellite mixed data type and sizes
- DTNs 2-way Data Transfer scenarios

  nscc01 <-> starlight01

  nscc01 <-> nict01

  nscc01 <-> nci01

  nscc01 <-> kisti01

Rules and Conditions

- Participants will log in with non-root user accounts
- Data transfer tools deployed only on **Singularity Containers** running on the DTNs
- Limited BIOS and administrator/root level access by request basis only
- Changes implemented by local partner systems administrator only

DMC19 Participating Teams

Organisation	Solution Title:
SEAIP/NCHC+Starlight	SEAIP DTN-as-a-Service
Fermilab	BigData Express
Zettar Inc.	Zettar zx hyperscale data distribution software platform
The University of Tokyo	Secure Data Reservoir
Argonne National Laboratory	Using GridFTP and Globus Online for Large Data Transfers
iCAIR/Northwestern University	STARLIGHT DTN-as-a-Service for Intensive Science
JAXA & Fujitsu	Smart Communication Optimizer

DMC19 Winners

1. **Overall Winner Award**

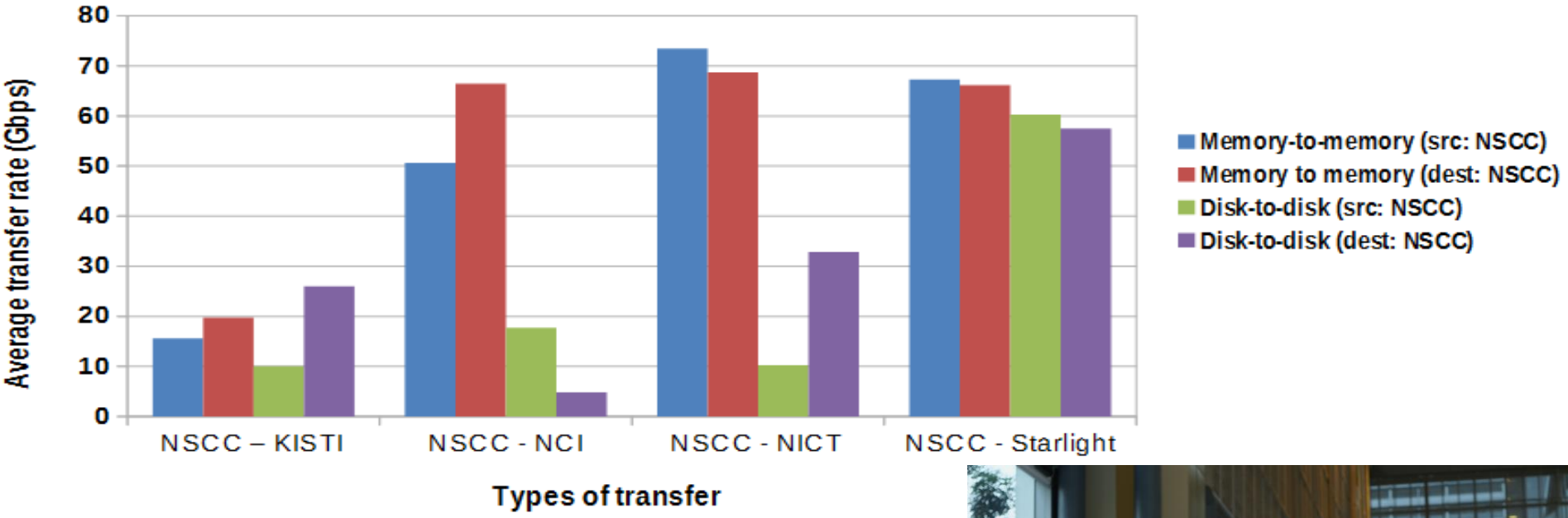
Zettar Inc. They achieve an amazing 68 Gbps Disk-to-Disk transfer rate between Chicago and NSCC (Singapore).

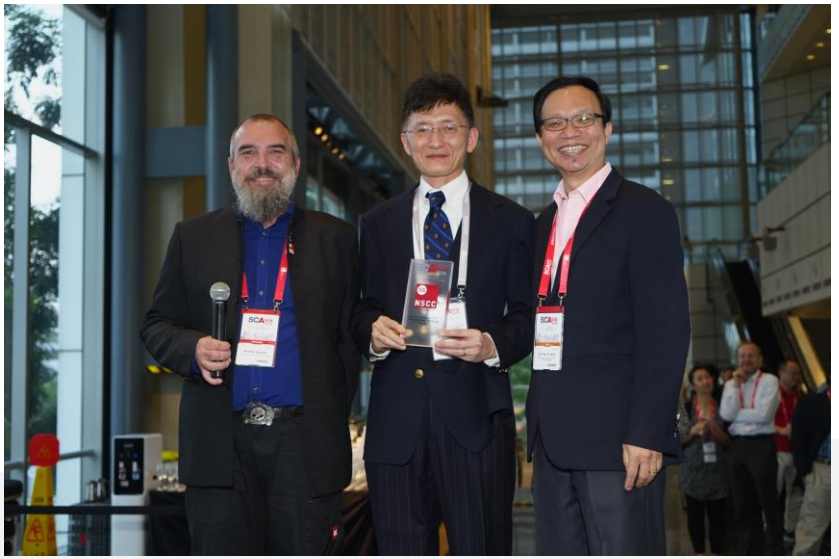
2. **Most Innovative Award**

StarLight/iCAIR (International Center for Advanced Internet Research) team. They impressed the judges with their flexible framework to support large data transfers.

SCA19 Data Mover Challenger Zettar Inc. Results

2-way transfers. Both memory-to-memory and disk-to-disk (in Gbps)



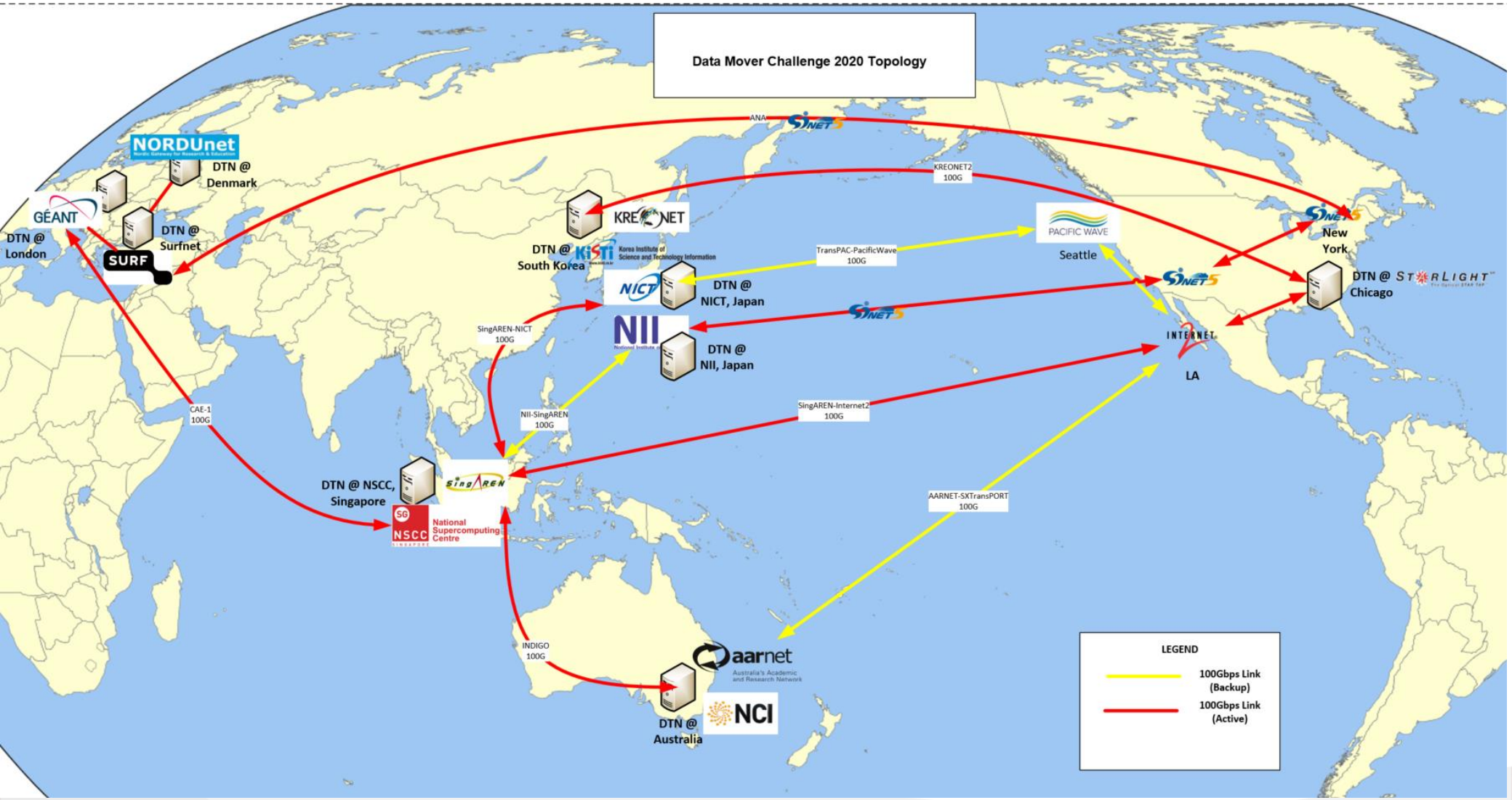


Data Mover Challenge 2020

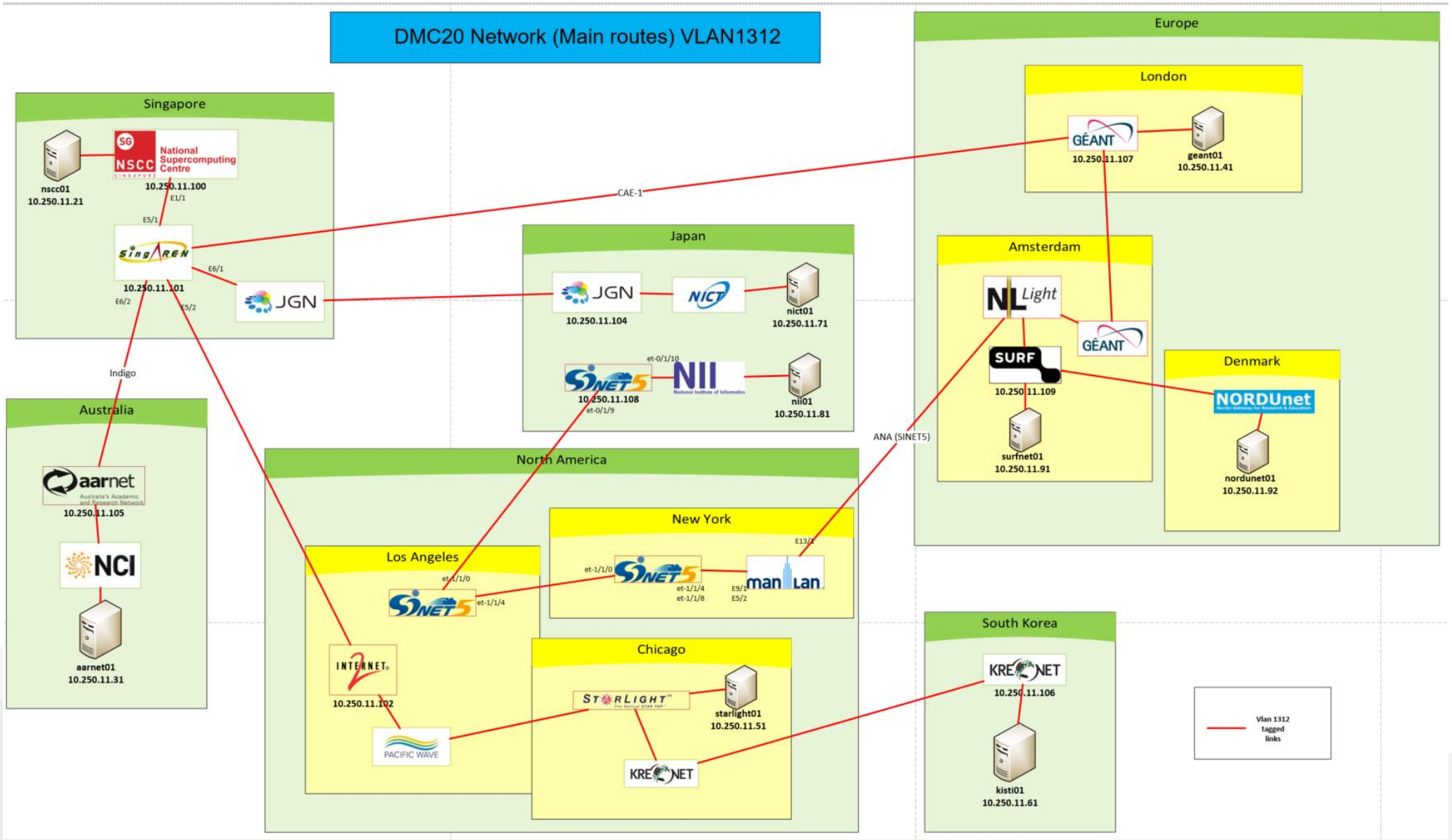
"Data For Science"

- 2nd iteration **Data Mover Challenge (DMC)** https://www.sc-asia.org/dmc20/ organised by NSCC Singapore
- Organised by NSCC Singapore in collaboration with International R&E partners
- Held in conjunction with **SupercomputingAsia 2020 (SCA20)** https://www.sc-asia.org/
- A platform to bring together experts from industry and academia
- To test their data transfer software across DTNs set up in various countries connected by 100G international networks

DMC20 Partners and Topology



DMC20 Partners and Topology



DMC20 Partners

- AARNet 

- CENIC 

- GEANT 

- iCAIR 

- Internet2 

- Jisc 

- KISTI 

- KREONET 

- NCI 

- NICT 

DMC20 Partners

- NII 

- NORDUnet 

- PACIFIC WAVE 

- PACIFIC RESEARCH PLATFORM 

- SINET 

- SingAREN 

- StarLight 

- SURFnet 

- TransPAC 

- NSCC (Organiser) 

Data Mover Challenge 2020

- Running from Aug 2019 – Jan 2020 (excluding Oct and Nov 2019)
- Each Team given 5 days
  - **Monday– Wednesday: Software Setup**
  - **Thursday – Friday: Running and demonstration of Software to judges**
- Participating Teams invited to SCA20 for award presentation and solution showcase

Important Dates

1 JUN 2019

Launch of DMC20

1 JUN – 31 JUL 2019

Submission period of entries for DMC20

31 JUL 2019

Close of entry submissions for DMC20

7 AUG 2019

Announcement of DMC20 Shortlisted Participants

26 AUG 2019 – 23 JAN 2020

Start - Challenge to begin

23 JAN 2020

End Challenge to conclude

24 FEB – 27 FEB 2020

SupercomputingAsia 2020 (SCA20) - Gathering the Best of HPC in Asia!

26 FEB 2020

DMC20 Award Ceremony - Announcement of Winner(s)

DMC20 Competition Scope

- About 2TB data set consisting of Genome and Satellite mixed data type and sizes

Memory-to-Memory and Disk-to-Disk Data Transfer scenarios:

- 1. surfnet01 -> simultaneous transfer to nscc01 (via CAE-1-SingAREN) and aarnet01 (via CAE-1-Indigo)
- 2. aarnet01 -> geant01 (via Indigo-CAE-1-GEANT)
- 3. nscc01 -> starlight01 (via SingAREN-I2-PacificWave-StarLight)
- 4. nii01 -> nordunet01 (via SINET-US-NetherLight-Nordunet)
- 5. nict01 -> kisti01 (via NICT/JGN-SingAREN-Internet2-StarLight-Kreonet

Judging Criteria

1. Error free data transfer performance (average transfer rate and time taken) of the data set:

- Memory to Memory Test **(Weightage 10%)**.
- Disk to Disk Test **(Weightage 50%)**.

2. Technological Innovation base on solution submission and interview with judges **(Weightage 20%)**.
    - Disk Management (management of files for efficient transfer)
    - Access Management (integration with federated services and single sign-on)
    - Protocol adaptability
    - OS scheduling
    - Others
3. End-user Experience including user interface, ease of installation, licensing model, and integration with services such as access federation etc. **(Weightage 20%)**.

DMC20 Competition Teams

Organisation	Country	Solution Title:
Visuality Systems	Israel	NQ line of products - SMB implementation
Gauss Centre for Supercomputing (GCS) e.V.	Germany	Unicore UFTP
JAXA: Japan Aerospace Exploration Agency, NICT: National Information of Communications and Technology	Japan	High-speed satellite data transferring over the world from JAXA/NICT via HpFP protocol
National Institute of Informatics	Japan	MMCFTP
National Center for High-performance Computing/NARLabs	Taiwan, U.S.A, Thailand, Japan, Indonesia, Vietnam	SEAIP DTN-as-a-Service
iCAIR/Starlight	United States	StarLight DTN-as-a-Service
Compute Canada/Calcul Canada	Canada	DTN as a service

For more information

www.sc-asia.org/dmc20

National Supercomputing Centre

NSCC's Flagship Annual HPC Conference

SupercomputingAsia

SCAsia 2020 Supercomputing

Gathering the **Best of HPC** in Asia

24 – 27 February 2020
SINGAPORE

- **SupercomputingAsia (SCA)** – Annual conference that hosts fringe and co-located events to strengthen local, regional and global cooperation in HPC technology.
- Insights from thought leaders in academia and industry on latest and emerging trends in HPC.
- Networking opportunities with public and private HPC community in Asia.

www.sc-asia.org

SCA19 Highlights

726 Number of Delegates

26 Countries from Asia, Europe and America

44 Exhibitors from public and private sector

12 Tracks related to HPC trends now and for the future

Expert Tutorials & User Group Sessions **5**

#exascale #quantum
#AI #hyperscale

Thank You