



SG Passion Made Possible

Data Mover Challenge The journey

Started in 2019

By

A/Prof. Lee Bu Sung, Francis
SingAREN

&

Director, International Network, NSCC(Singapore)



Objective

- Global research cooperation is driving the increasing demand for capabilities to effect the sharing of large amounts of research data, quickly and securely, between collaborating entities or computational resources.
- The *Data Mover Challenge* invites international teams to address the need for higher speeds and better coordination when moving large amounts of data between multiple sites with Data Transfer Nodes (DTNs), as well as to HPC centres around the world.

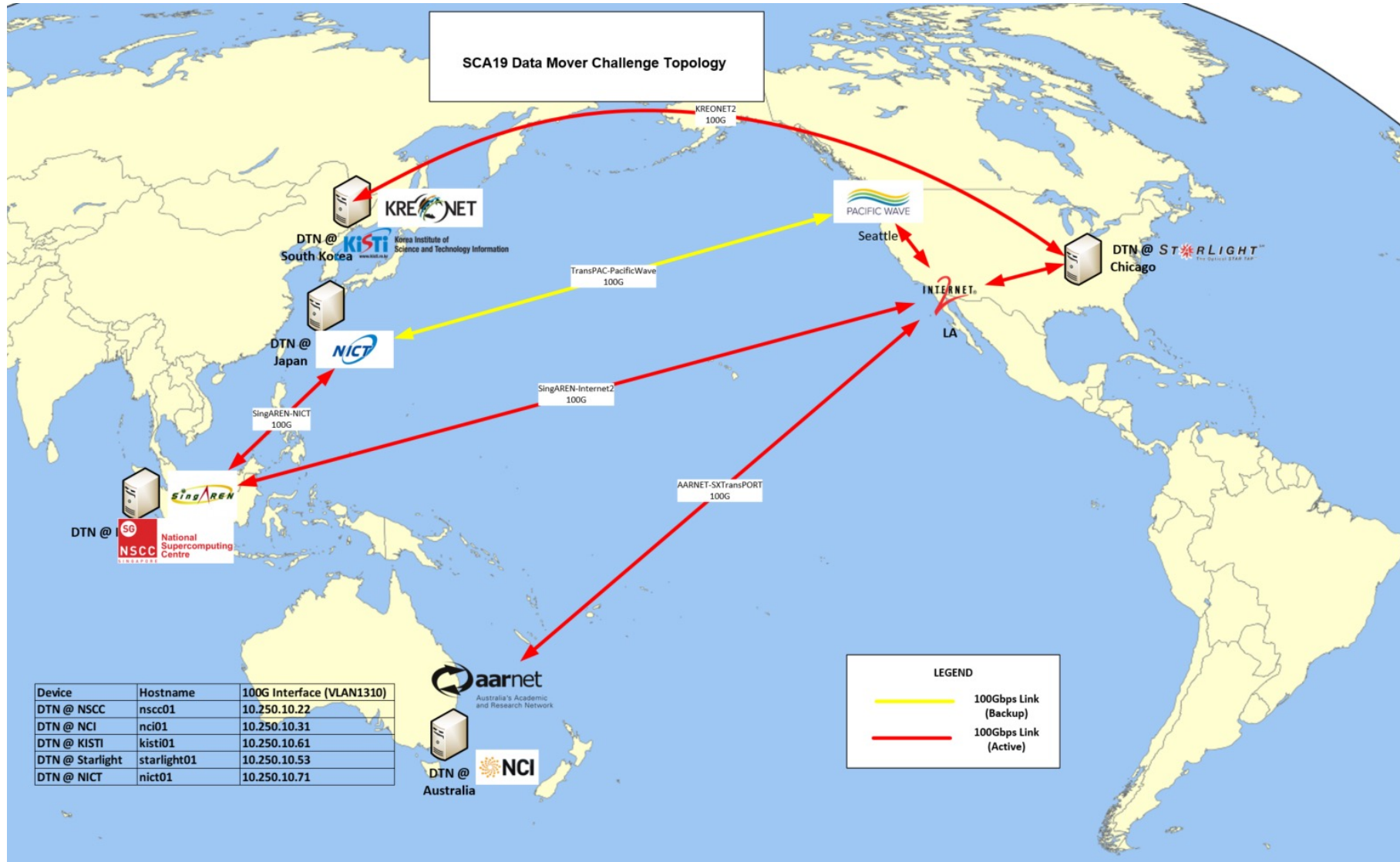
1st DMC 2019

“Move That Data” Challenge












- Inaugural **Data Mover Challenge (DMC)** organised by NSCC
- Period: Jan-March 2019
- Award ceremony 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



DMC Partners and Topology



DMC Partners

- AARNet 
 - Internet2 
 - KREONET 
 - KISTI 
 - NCI 
 - NICT 
 - PACIFIC WAVE 
 - SingAREN 
 - StarLight 
 - NSCC (Organiser) 
- 

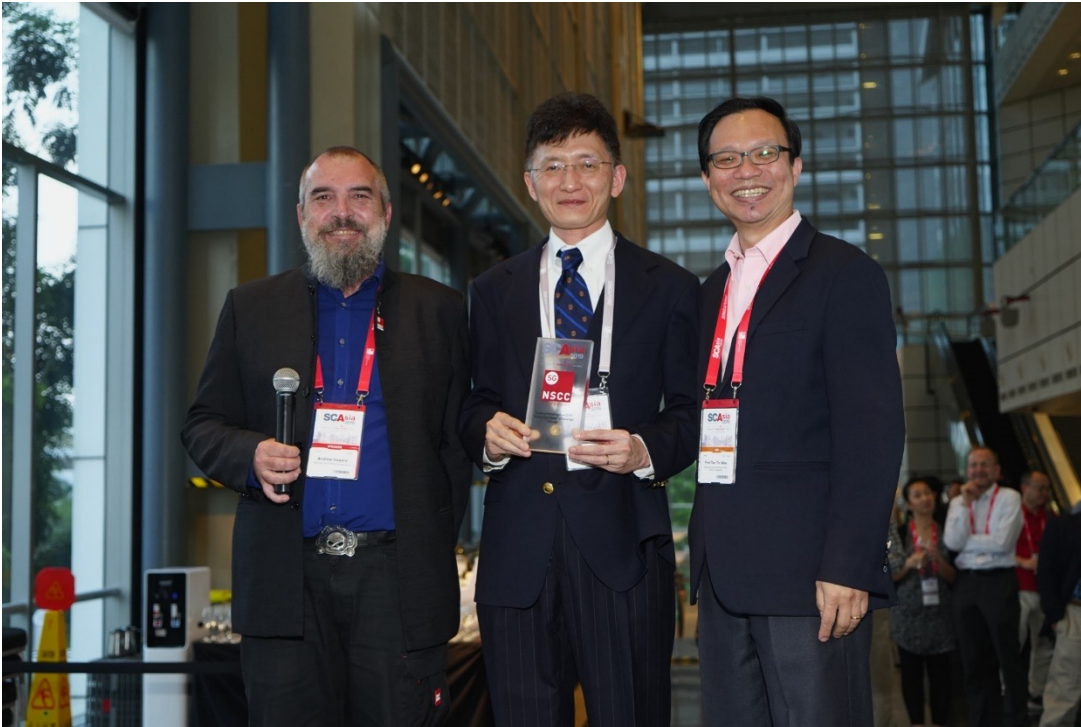
Judging Criteria

1. Error free data transfer performance of the data set
 - Memory to Memory Test for both directions
 - Disk to Disk Test for both directions
2. Technological Innovation base on solution submission and interview with judges
3. End-user Experience including user interface, ease of installation, licensing model, and integration with services such as access federation etc.

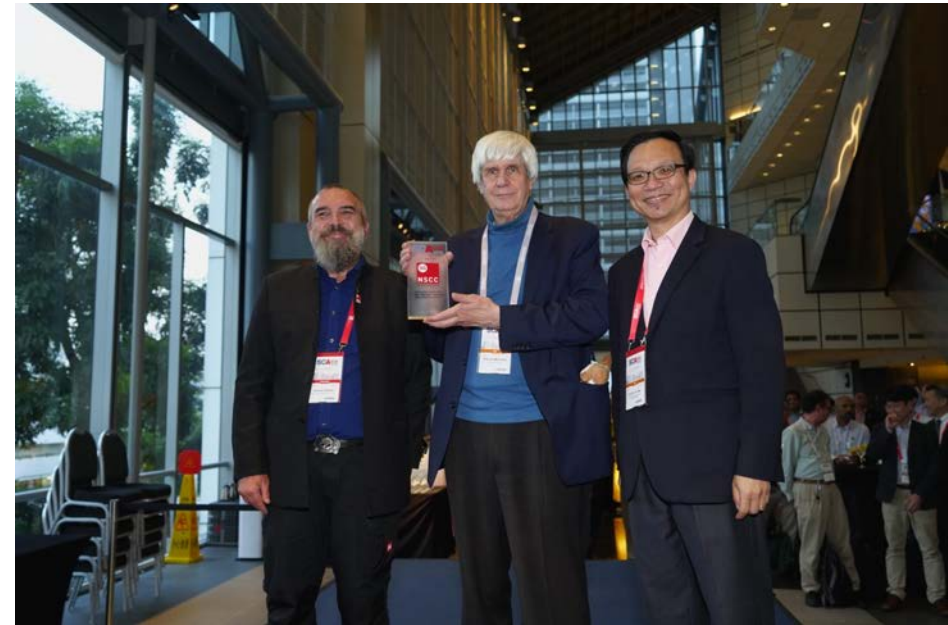
Participating Teams

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

Data Mover Challenge 2019 Winners



Overall winner: Zettar Inc



Most Innovative award: StarLight/iCAIR

2nd DMC

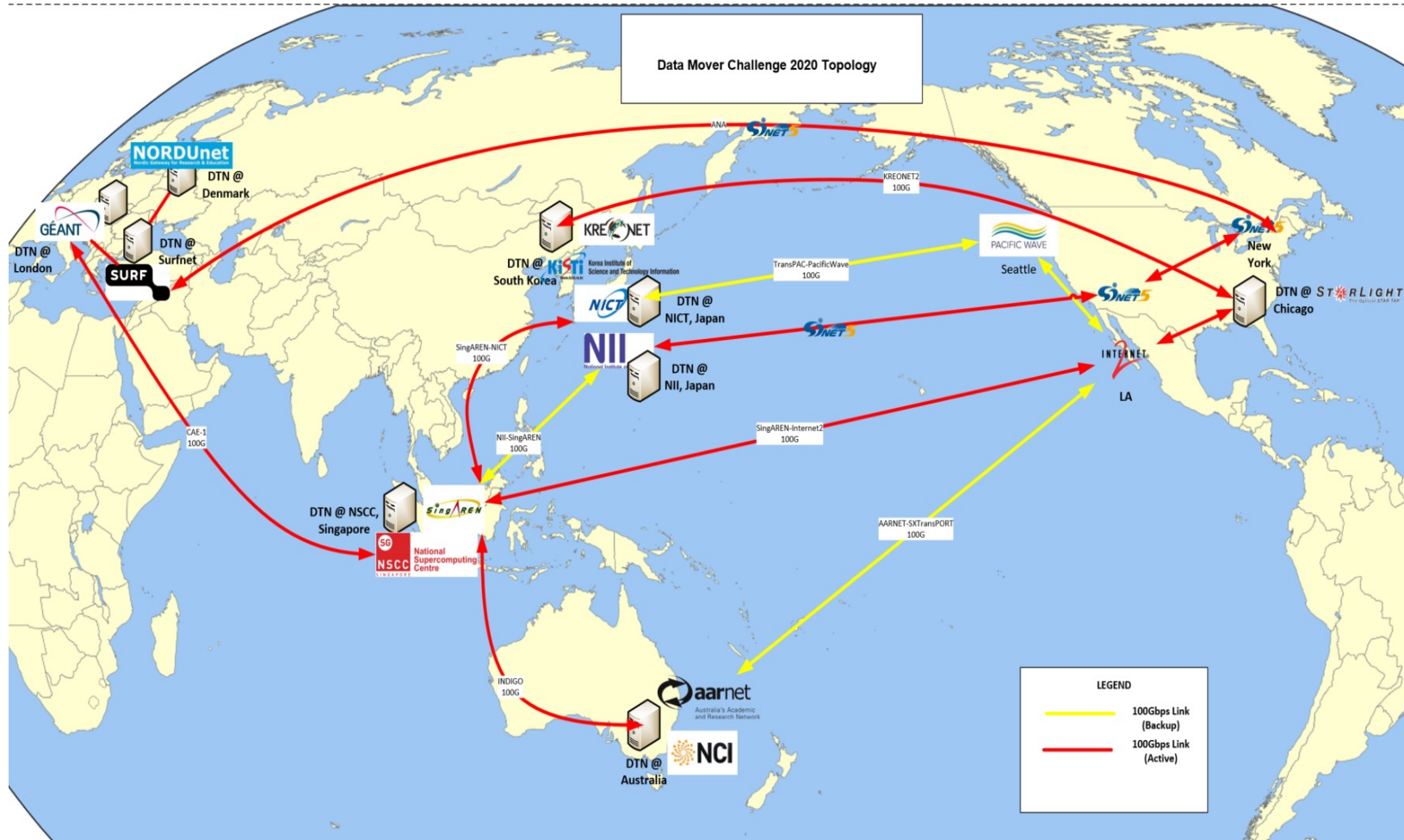
2020

“Data For Science” Challenge

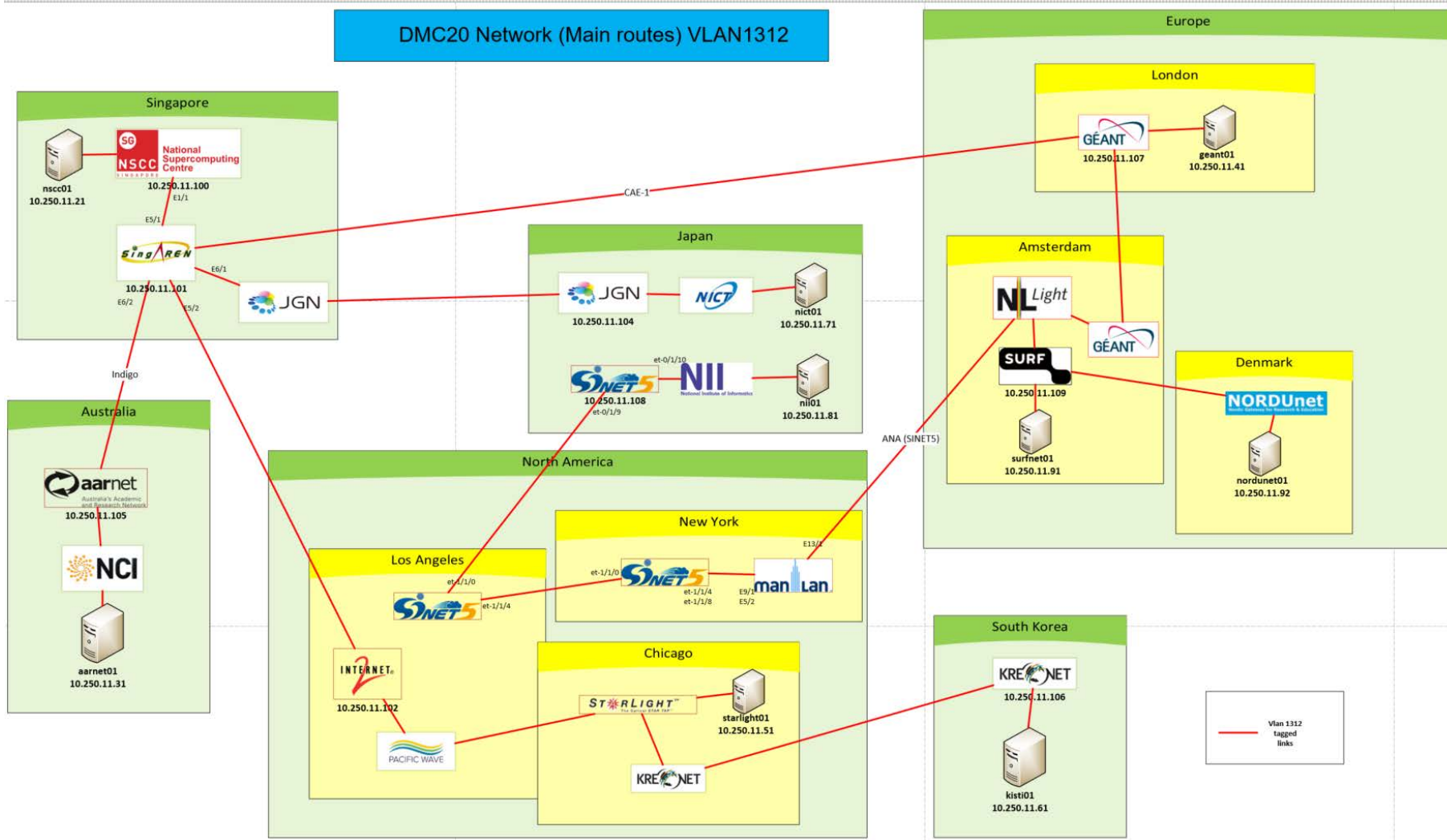
- **2nd iteration Data Mover Challenge (DMC)** organised by NSCC in collaboration with international R&E
- Period: Aug 2019 to Jan 2020
- Virtual Award ceremony held in conjunction with NSCC’s flagship conference - **SupercomputingAsia 2020 (SCA20)**
- Bring together experts from industry and academia
- To test their software across servers across the GLOBE connected by 100G international networks



DMC20 Partners and Topology



DMC20 Partners and Topology














DMC20 Partners

- AARNet  Australia's Academic and Research Network
- CENIC 
- GEANT 
- iCAIR 
- Internet2 
- Jisc 
- KISTI  Korea Institute of Science and Technology Information
- KREONET 
- NCI 
- NICT 



DMC20 Partners

- NII 
- SingAREN 
- NORDUnet 
- StarLight 
- PACIFIC WAVE 
- SURFnet 
- PACIFIC RESEARCH PLATFORM 
- TransPAC 
- SINET 
- NSCC (Organiser)  

Winning Teams at DMC 2020

Award	Organisation/Team
Most Systemic Award	Gauss Center for SuperComputing (GCS) e.V., Germany
Best Speed and Science Integration Award	International Center for Advanced Internet Research(iCAIR)/Starlight, USA
Experimental Excellence Award	Japan Aerospace Exploration Agency (JAXA)/National Institute of Information and Communication Technology (NICT), Japan
Most Innovative and Novelty Award	National Institute of Informatics (NII), Japan



Some of the participants, partners and judges of DMC20 met at the National Supercomputing Centre (NSCC) Singapore booth at SC19, Denver, USA for a Participants' Briefing Session to prepare for the challenge. Credit: NSCC Singapore



3rd DMC

2021

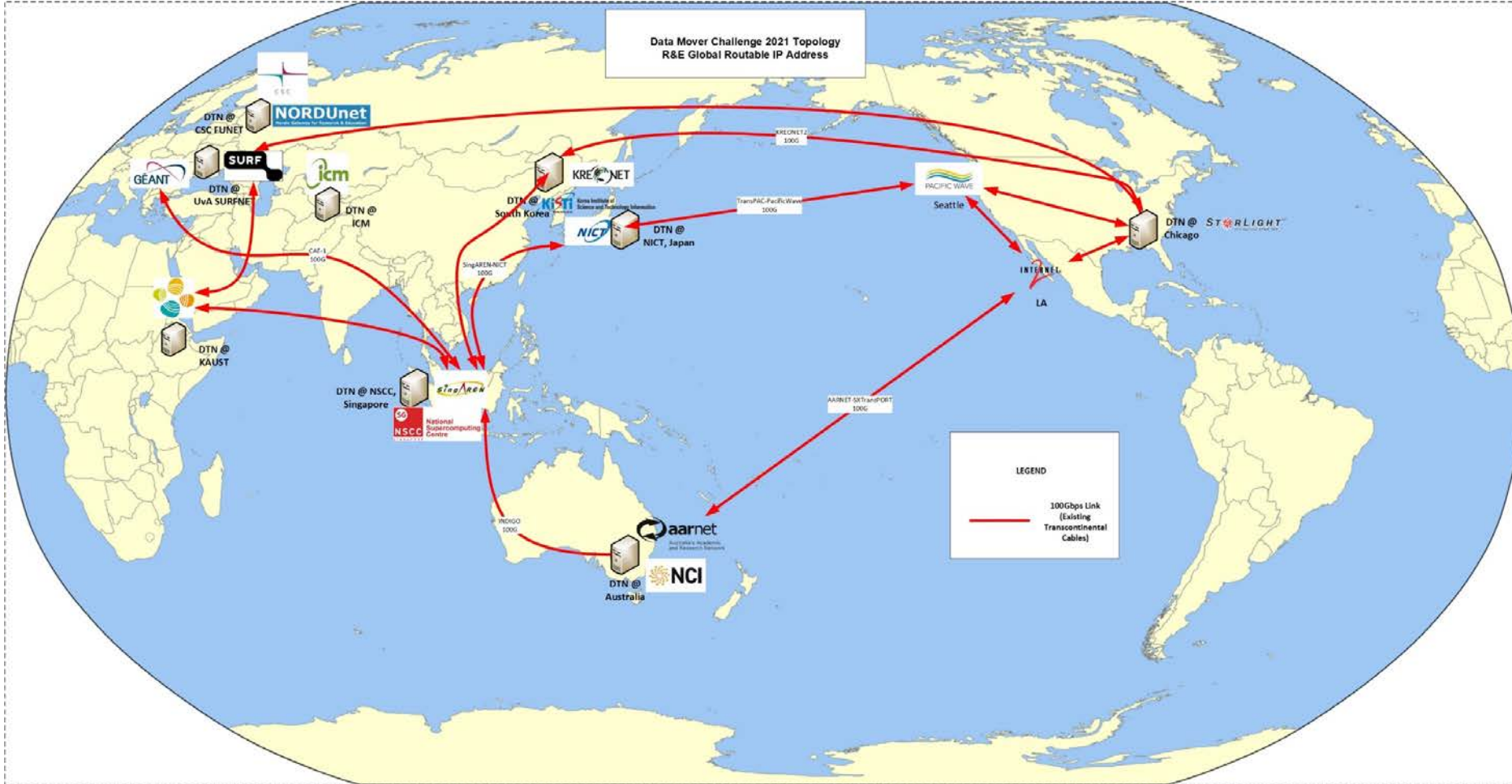
Data Mover Challenge (DMC) 2021/22

- Biennial **Data Mover Challenge (DMC)** organised by NSCC
- Duration: Aug-Nov 2021
- Award held in conjunction with NSCC's flagship conference – **SupercomputingAsia 2022**
- Bring together experts from industry and academia
- Provides a test bed for evaluating data transfer tools, techniques and infrastructure at a global scale
- To test their software by transferring a data set across Data Transfer Nodes (DTNs) set up in various countries connected by 100G international networks
- Reduce friction in large inter-facility data transfers

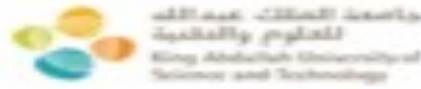
Data Mover Challenge (DMC) 2021/22

- Revamp the network connectivity from VLANs(used in DMC19 and DMC20) in favor of instrumenting and tuning NREN production networks services to support robust, production quality deployments of national, regional and international data transfers.
- Support for IPv6.
- Containerization of underlying toolkits to support future deployment as Research Platform services.
- DMC21 is also evolving to support a contemporary Research Platform deployment model to demonstrate flexible resource allocation and deployment

Data Mover Challenge (DMC) 2021/22



Supporting Partners:



Data Mover Challenge (DMC) 2021/22

- DTN Bandwidth, CPU, memory monitoring and statistics captured by netdata performance monitoring and displayed centrally on Prometheus server set up at SingAREN.



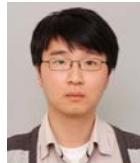
Winners DMC21

BEST LONG DISTANCE PERFORMANCE AND OVERALL WINNER

Ciena-iCair-UET



Gauravdeep Shami



Se-young Yu



Danial Ebling

MOST INNOVATIVE AND BEST IPV6 PERFORMANCE

Team MUSASHINO



Ken T. MURATA



Yasunori KAKIZAWA

Yuki MURAKAMI
Praphan PAVARANGKOON

Takamichi MIZUHARA
Ayahiro TAKAKI
Keichiro FUKAZAWA

MOST COMPLETE SOLUTION AND BEST SOFTWARE ARCHITECTURE

Arcitecta



BEST INTEGRATED SOFTWARE EXPERIENCE

Globus



Zhengchun Liu



Rajkumar Kettimuthu



Mike Link



Jack Kordas



Ian Foster

BEST VIRTUALISATION SUPPORT

Fast Is Good



Tecsun Yeep

Kenny Zou
Walken Huang
Gordon Shaw
Shuiwei Xiao

Weikang Gao
Jessica Lou
Tinahua Liu
Tingxin Huang



4th DMC 2023

DMC23 Network Partners



DMC23 DTN Partners



STARLIGHT™



ICT Solutions for Brilliant Minds



جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology



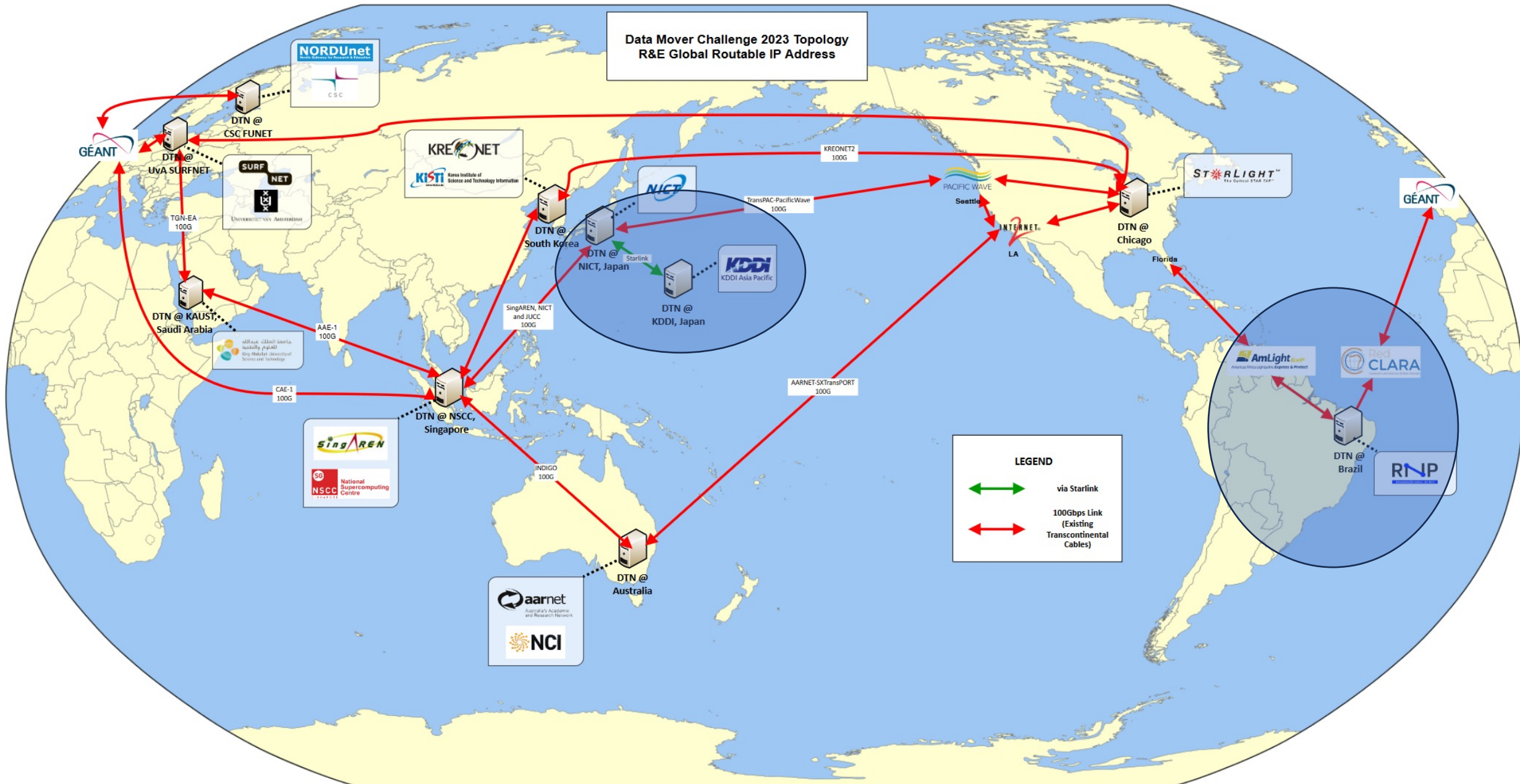
UNIVERSITEIT VAN AMSTERDAM



Korea Institute of
Science and Technology Information



Data Mover Challenge 2023 Topology
R&E Global Routable IP Address



LEGEND

- ↔ via Starlink
- ↔ 100Gbps Link (Existing Transcontinental Cables)

NORDUnet
 Nordic University Research & Education Network

KREONET
 Korea Institute of Science and Technology Information

NICT

KDDI
 KDDI Asia Pacific

SingAREN
 National Supercomputing Centre

aarnet
 Australian Academic and Research Network

NCI

STARLIGHT
 The Optical STAR TAP™

AmLight
 America's Fiber Optic Experts & Protect

CLARA

RNP

SURFNET
 SURFNET

UVA SURFNET

KAUST
 King Abdullah University of Science and Technology

South Korea

SingAREN, NICT and JUCC

DTN @ NSCC, Singapore

DTN @ Australia

INTERNET
 LA

PACIFIC WAVE
 Seattle

DTN @ Chicago
 Florida

DTN @ Brazil

GÉANT

GÉANT

TGN-EA
 100G

CAE-1
 100G

AAE-1
 100G

INDIGO
 100G

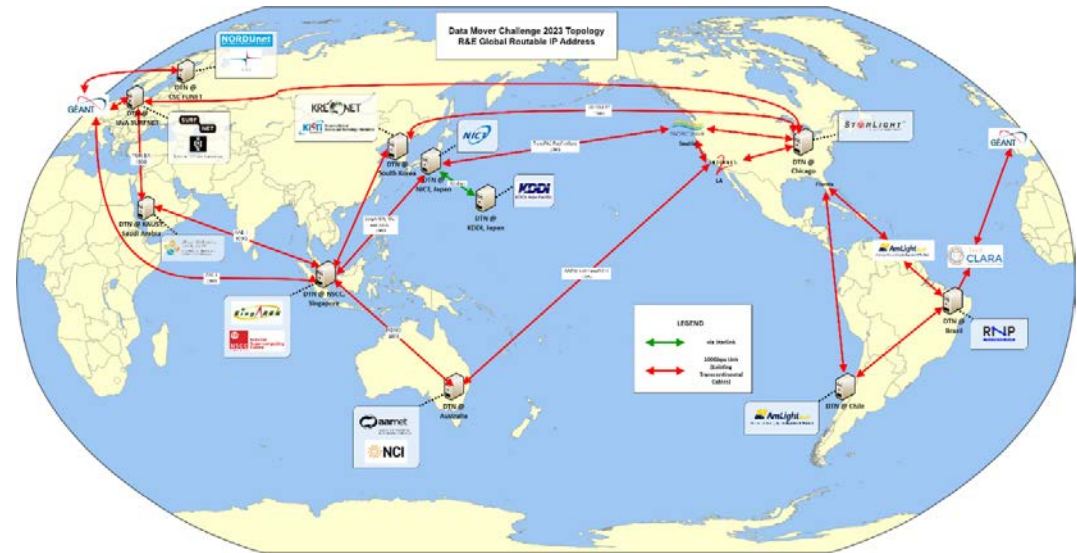
AARNET-SXTransport
 100G

TransPAC-PacificWave
 100G

KREONET2
 100G

Data Mover Challenge 2023

- Total of 7 teams from Asia, Europe, Oceania, USA.
- Competition started on 7 August 2023



Organisation	Solution Title:
Livewire	Big Data Through the Eye of a Needle
Raysync x Robust HPC Alliance	Raysync
Team Falcon	Falcon: Online High-Speed File Transfer Optimization
Team MUSASHINO	High-performance and Flexible Protocol and its application challenges high-speed file transfer in 100G on global LFNs
Just SSH!	mscp: Multi-threaded scp
Team CiTEO	DTN Optimization using Dynamic Network Estimation and DTN Impedance matching.
ADN	ADN: Application-Defined Networking for Efficient, Flexible Data Transport through Deep Infrastructure Visibility

NSCC to Amlight

```

1: [LOCALHOST]                pmtu 9000
1: _gateway                    0.413ms
1: _gateway                    0.246ms
2: et-6-2-41-nsc-sock1.singaren.net.sg 1.011ms
3: 192.168.203.11             184.711ms
4: 2000-i2-soeus.singaren.net.sg 187.096ms
5: fourhundredge-0-0-0-0.4079.core2.elpa.net.internet2.edu 231.652ms asymm 15
6: fourhundredge-0-0-0-22.4079.core1.elpa.net.internet2.edu 232.405ms asymm 14
7: fourhundredge-0-0-0-0.4079.core1.hous.net.internet2.edu 232.616ms asymm 13
8: fourhundredge-0-0-0-0.4079.core1.houh.net.internet2.edu 233.831ms asymm 12
9: fourhundredge-0-0-0-0.4079.core1.pens.net.internet2.edu 231.579ms asymm 11
10: fourhundredge-0-0-0-0.4079.core1.jack.net.internet2.edu 231.601ms
11: 198.71.45.187              237.723ms
12: 32.8.39.170.ampath.net     237.503ms
13: amlight01                  350.346ms reached

Resume: pmtu 9000 hops 13 back 13

```

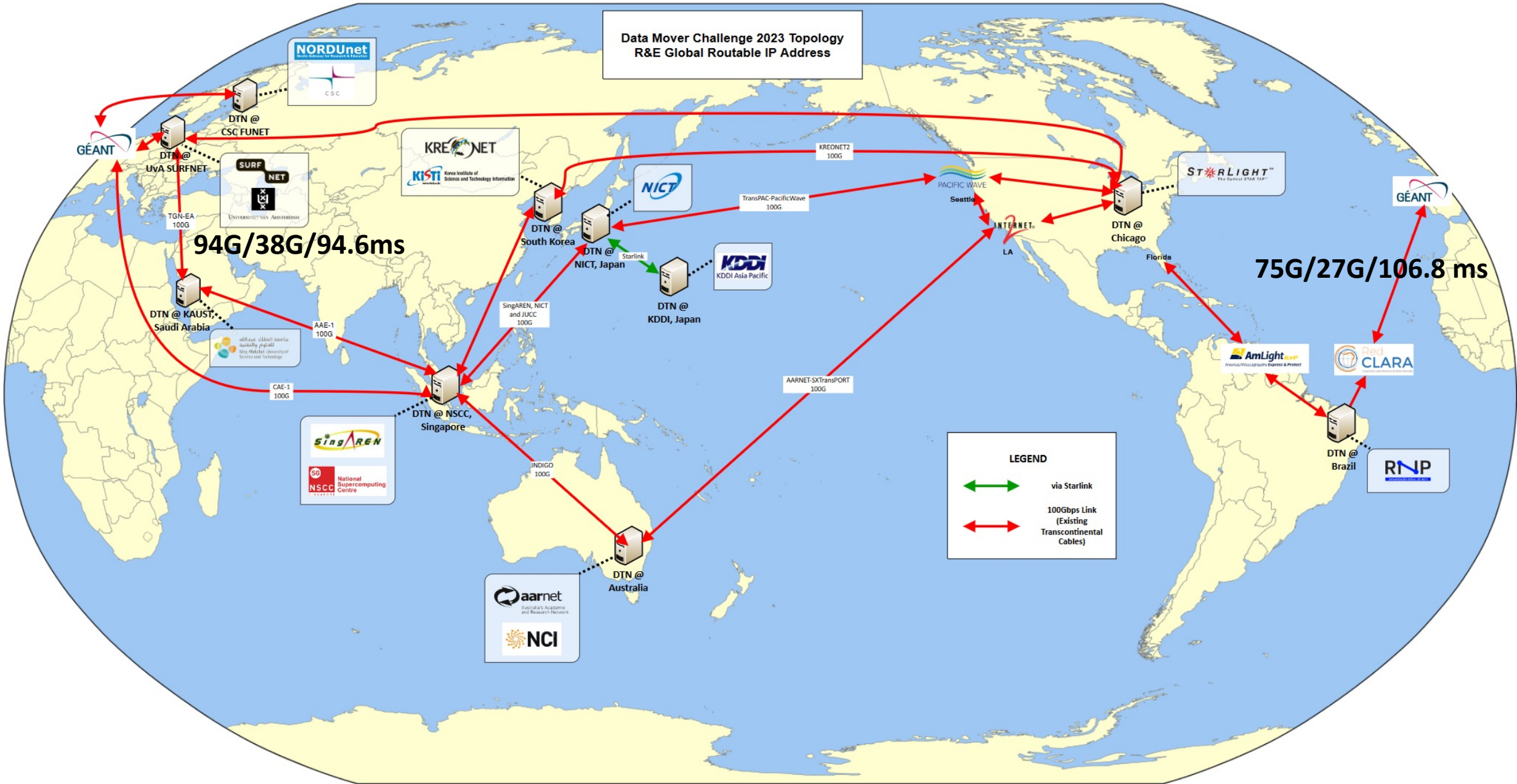


STARlink to NSCC(Singapore)

```
1: [LOCALHOST]                pmtu 1384
1: m7i-01-ge0-1-0-v21.jp.apan.net    0.267ms
1: m7i-01-ge0-1-0-v21.jp.apan.net    0.223ms
2: 10.0.0.1                        46.725ms
3: tpr5-xe0-0-0-v2.jp.apan.net       43.405ms
4: tyo-mx2010-et-3-1-0-v6.jp.apan.net 53.000ms asymm 3
5: 203.181.194.179                 98.163ms asymm 4
6: sg-mx480-1j-et-0-1-0-v2335.jp.apan.net 119.544ms asymm 5
7: singaren.jp.apan.net             130.170ms asymm 6
8: 41-nscs-soe1.singaren.net.sg      130.386ms asymm 7
9: nscs01                            130.117ms reached
Resume: pmtu 1384 hops 9 back 8
```

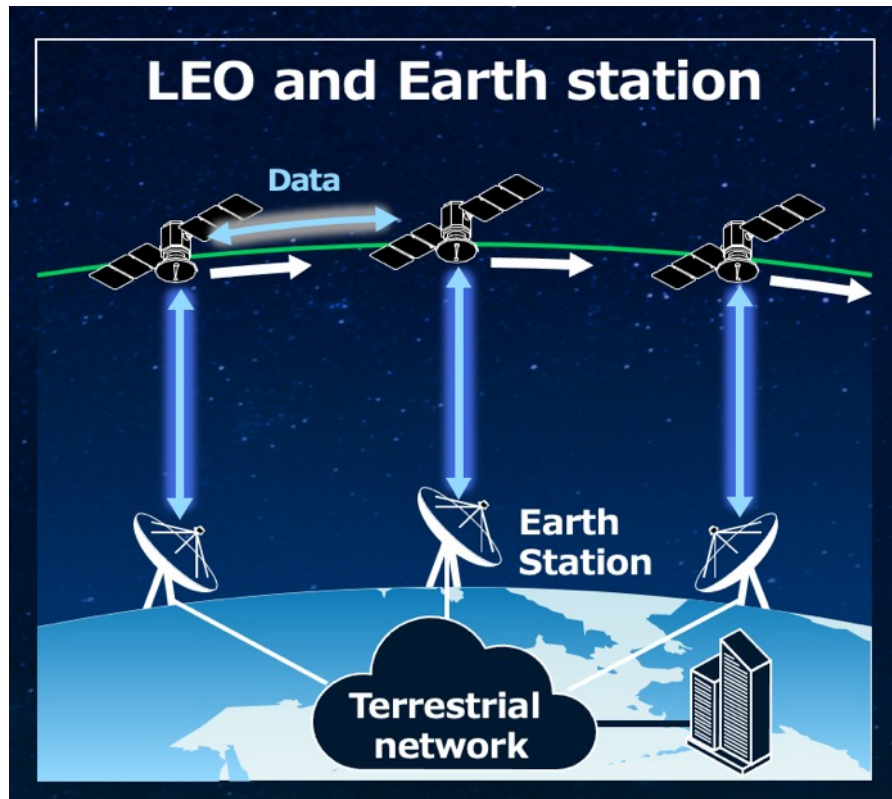


Data Mover Challenge 2023 Topology
R&E Global Routable IP Address



Starlink

- It is a Satellite broadband Internet service developed and operated by SpaceX
- Started in Northern America, Europe and Oceania since 2020 as a "beta" service. Available over 30 countries now.
- In 2022/Oct, service started in Japan (KDDI).
First release in the Asian region



Starlink satellites are positioned in LEO (low-Earth orbit) at an altitude of 550 km, so it can achieve significantly lower latency and higher transmission speeds for its end users.

Provide a Starlink to DMC23

- APAN-JP has sent a request to KDDI

KDDI

Solution Engineering & Operations Division



Solution ICT Engineering Department

- R&E NOC
- @Tokyo/Otemachi

Solution Operations Management Department

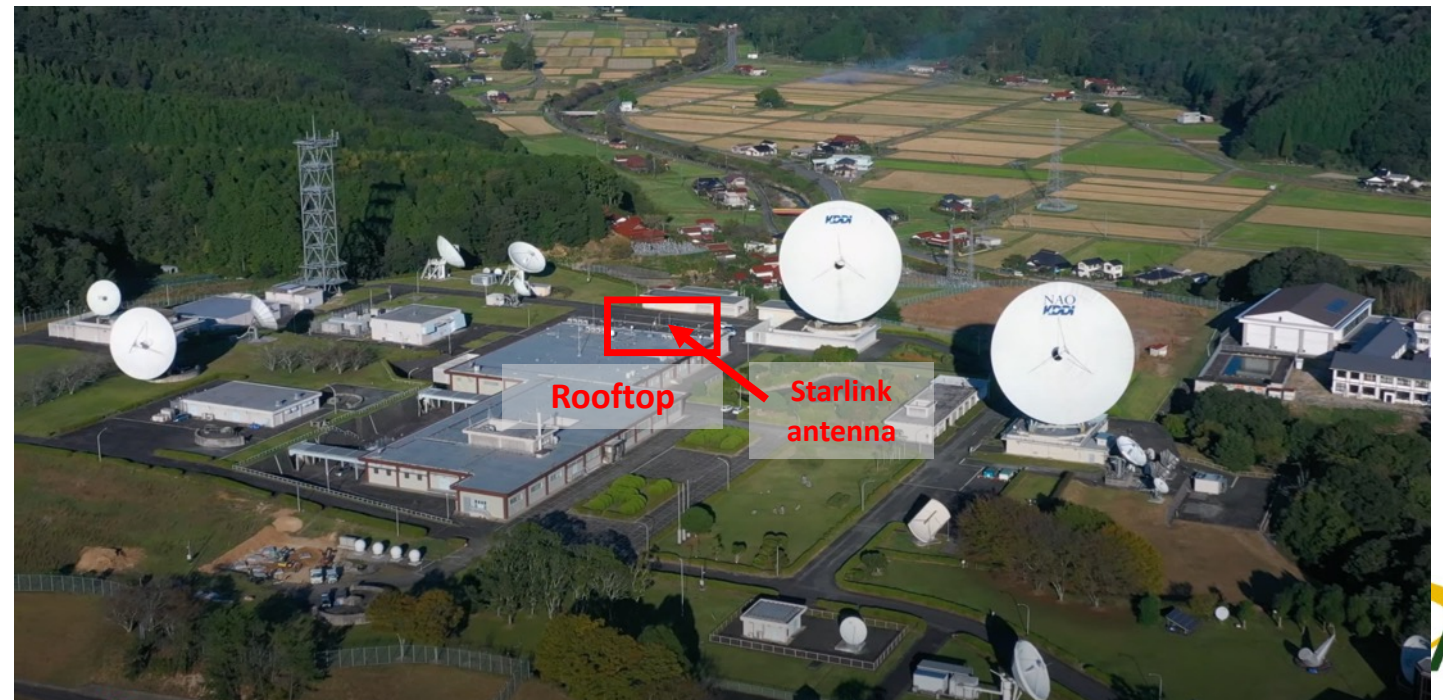
- Yamaguchi Satellite Communications Center)
- @Yamaguchi

User clients of Starlink

Type	Standard	High Performance
	 Circular	 Rectangular
Service	Residential	Business
Category	Phased array	
Size(cm)	Circular: 55 (diameter) Rectangular: 30x51	Panel size : 58x81
Weight(kg)	2.9kg (Antenna only)	6.9kg(Antenna only)
Angle	100°	140°
Wi-Fi	IEEE 802.11a/b/g/n/ac	
Ethernet port	Circular: One port Rectangular : Need adapter	Need adapter

KDDI Yamaguchi Satellite Communications Center

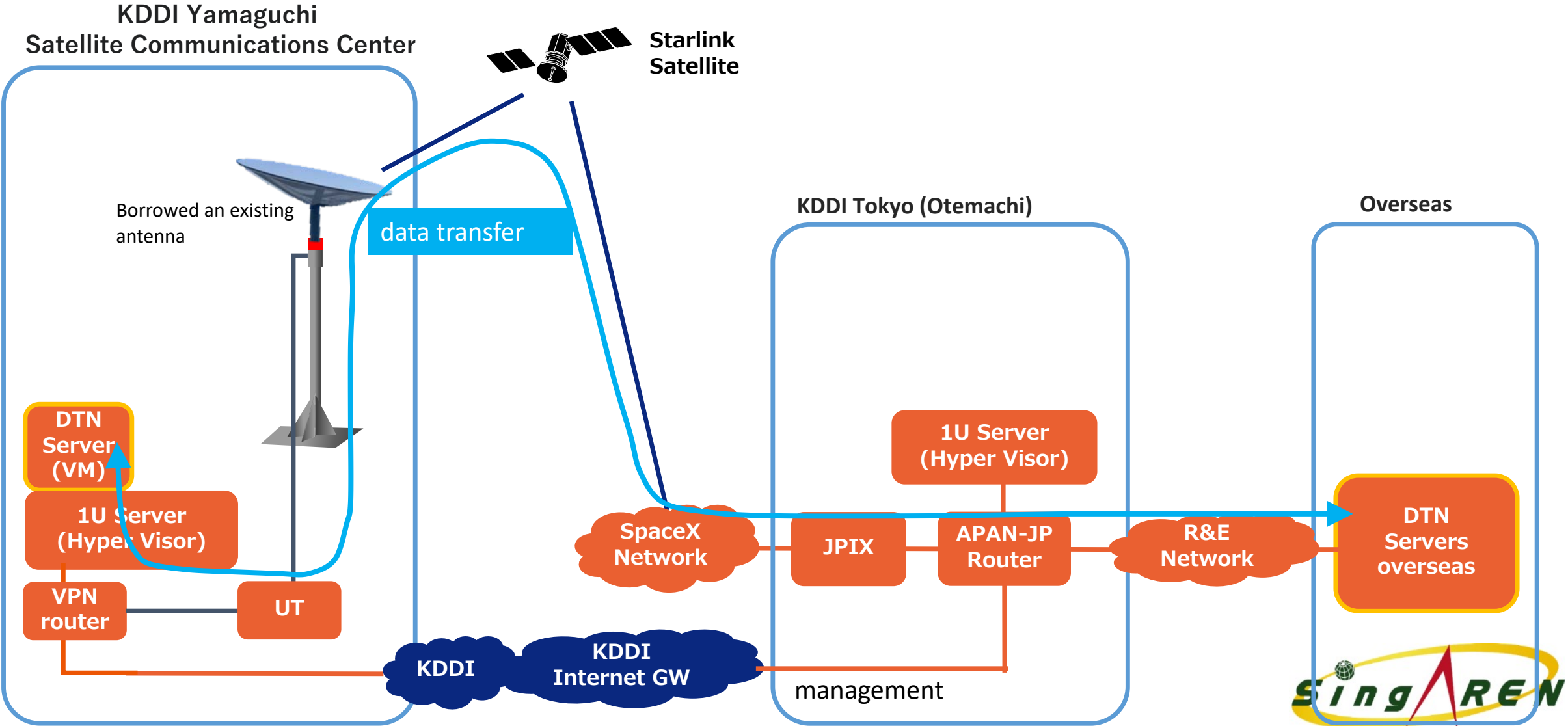
- KDDI opened satellite communication centers in Yamaguchi in 1969, respectively, to handle Japan's international satellite communications. It is one of the largest satellite communication facilities in Japan. There are currently 20 parabola antennas on the 160,000-square-meter site.



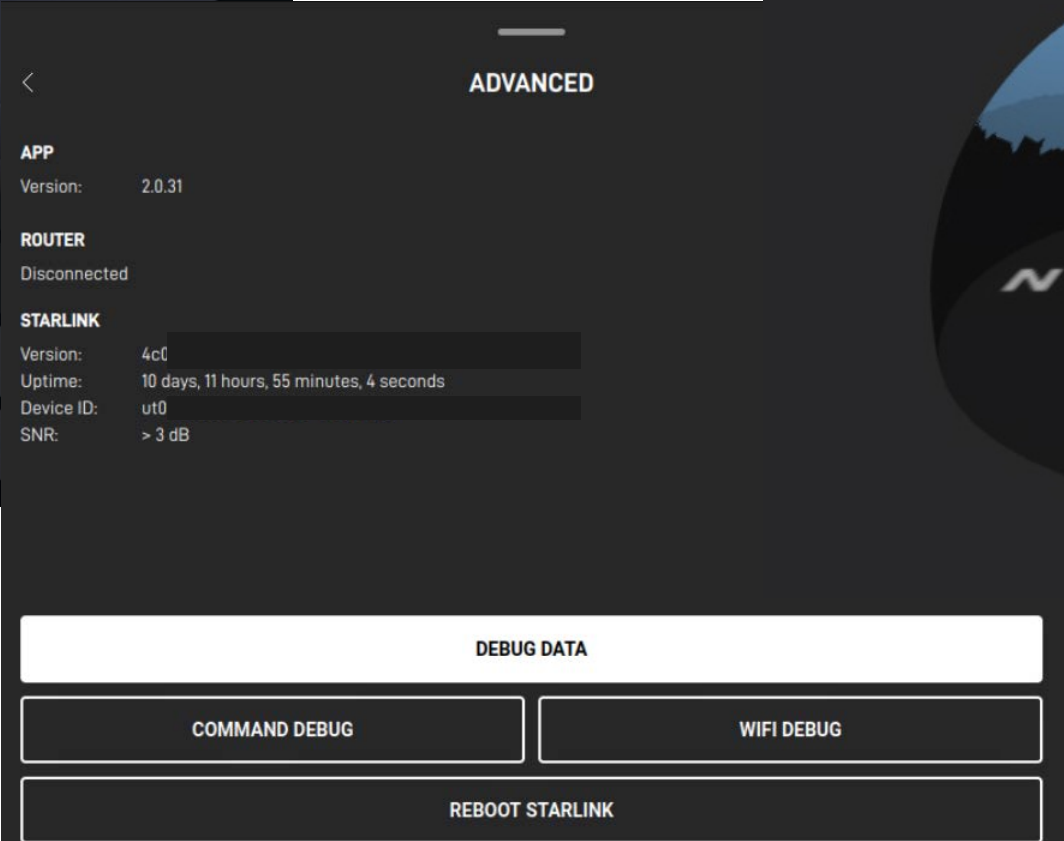
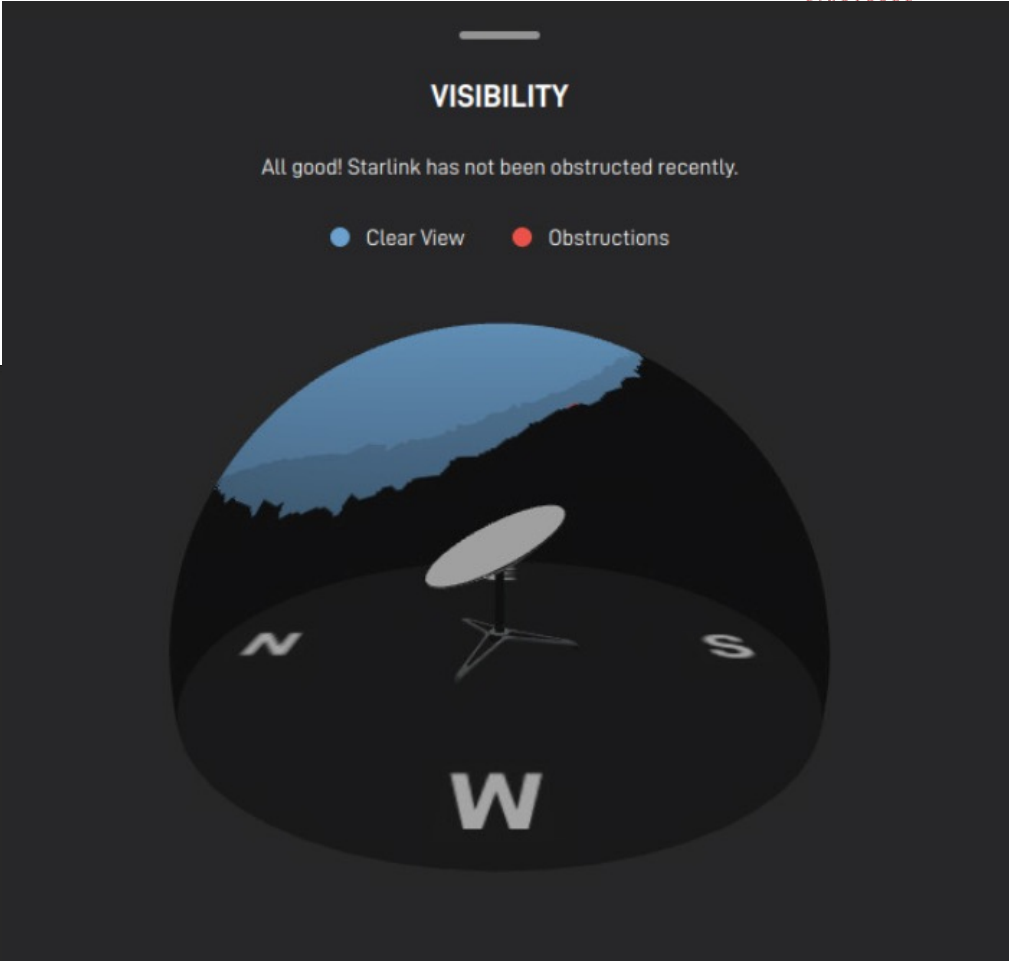
Starlink Antenna



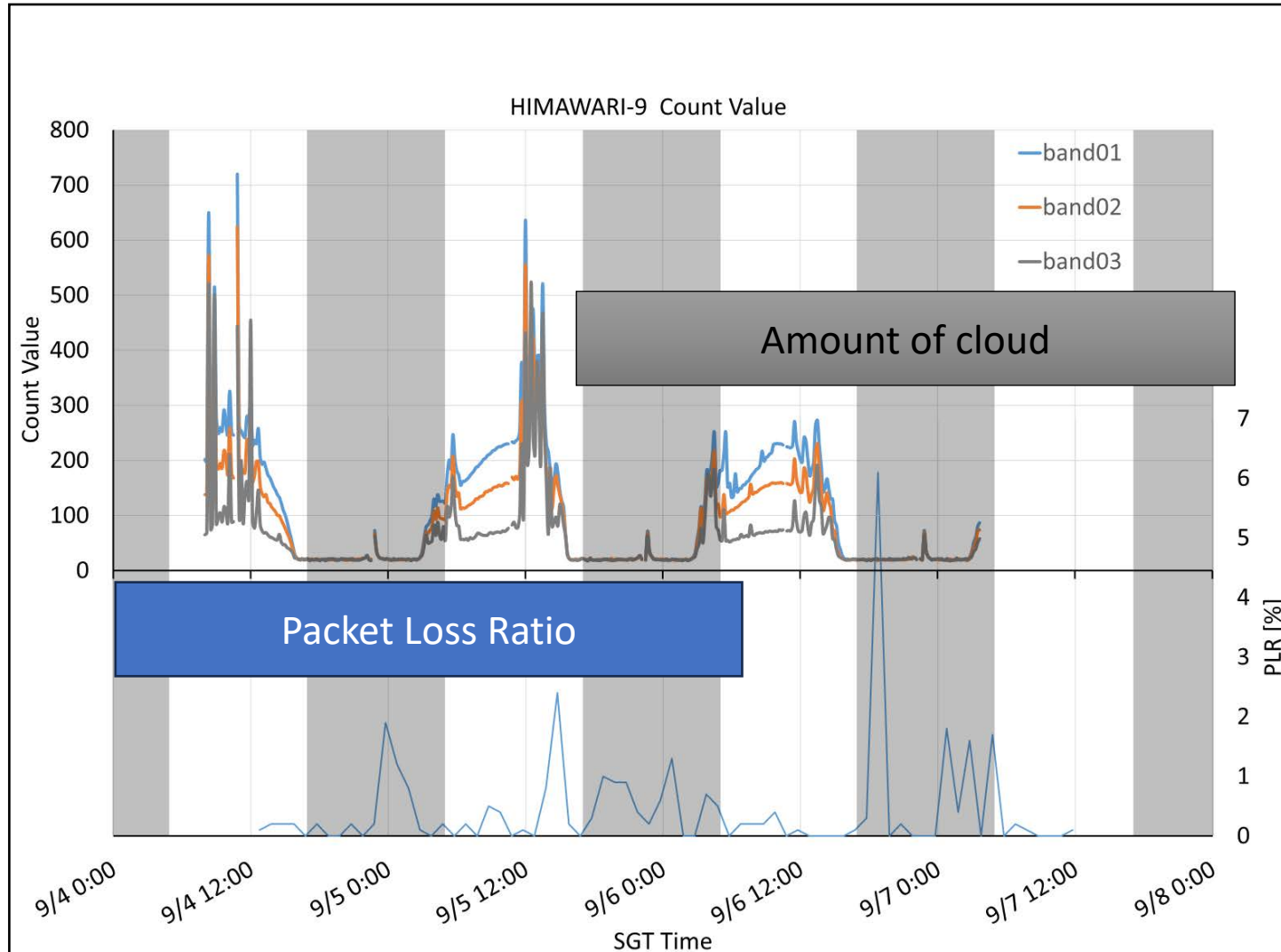
Image of connection from Yamaguchi to DMC



Screen of starlink

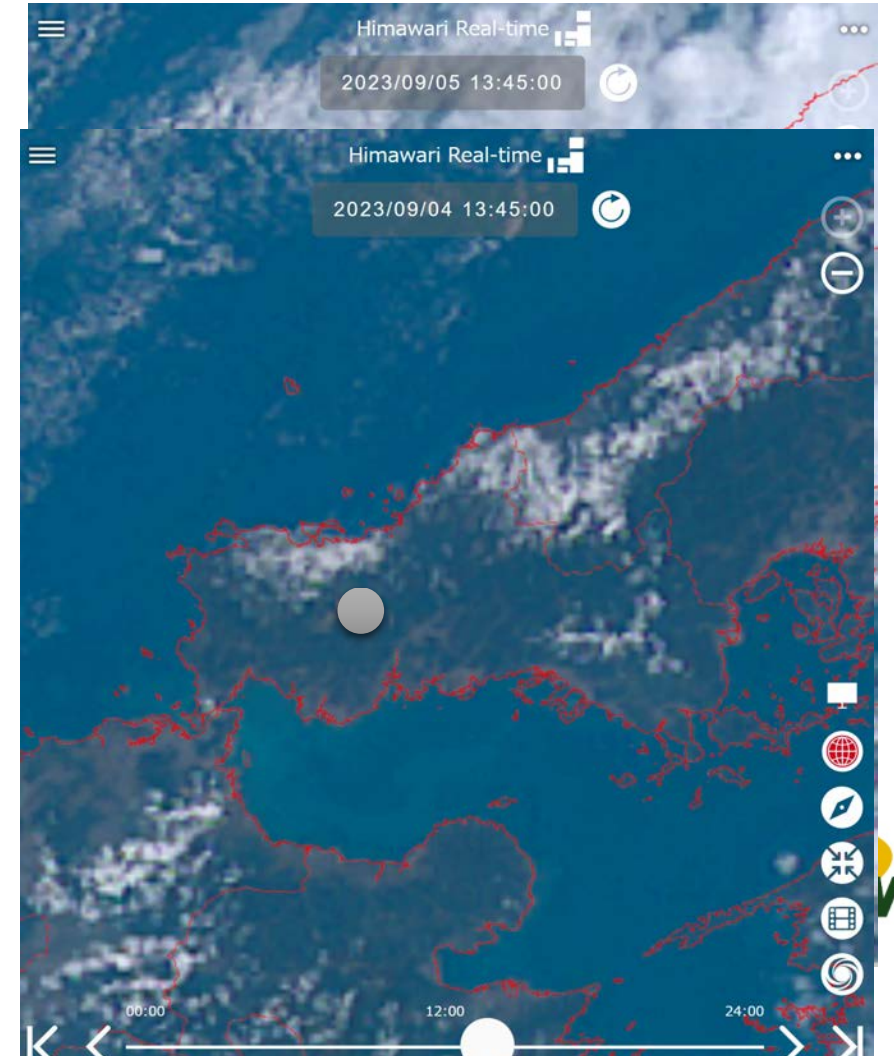


Weather@Yamaguchi and Network conditions

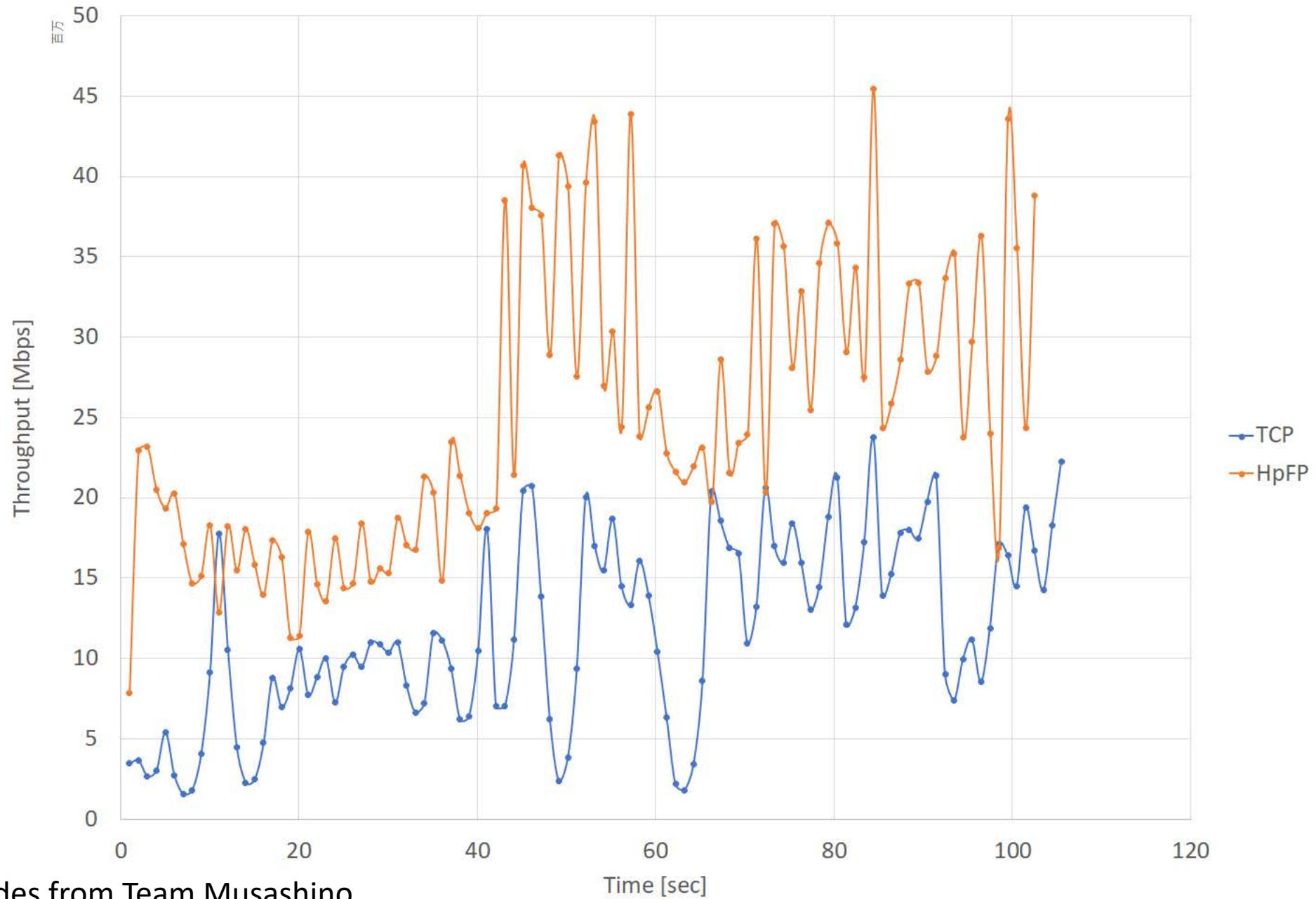


Slides from Team Musashino

Himawari real-time web
<https://himawari.asia>



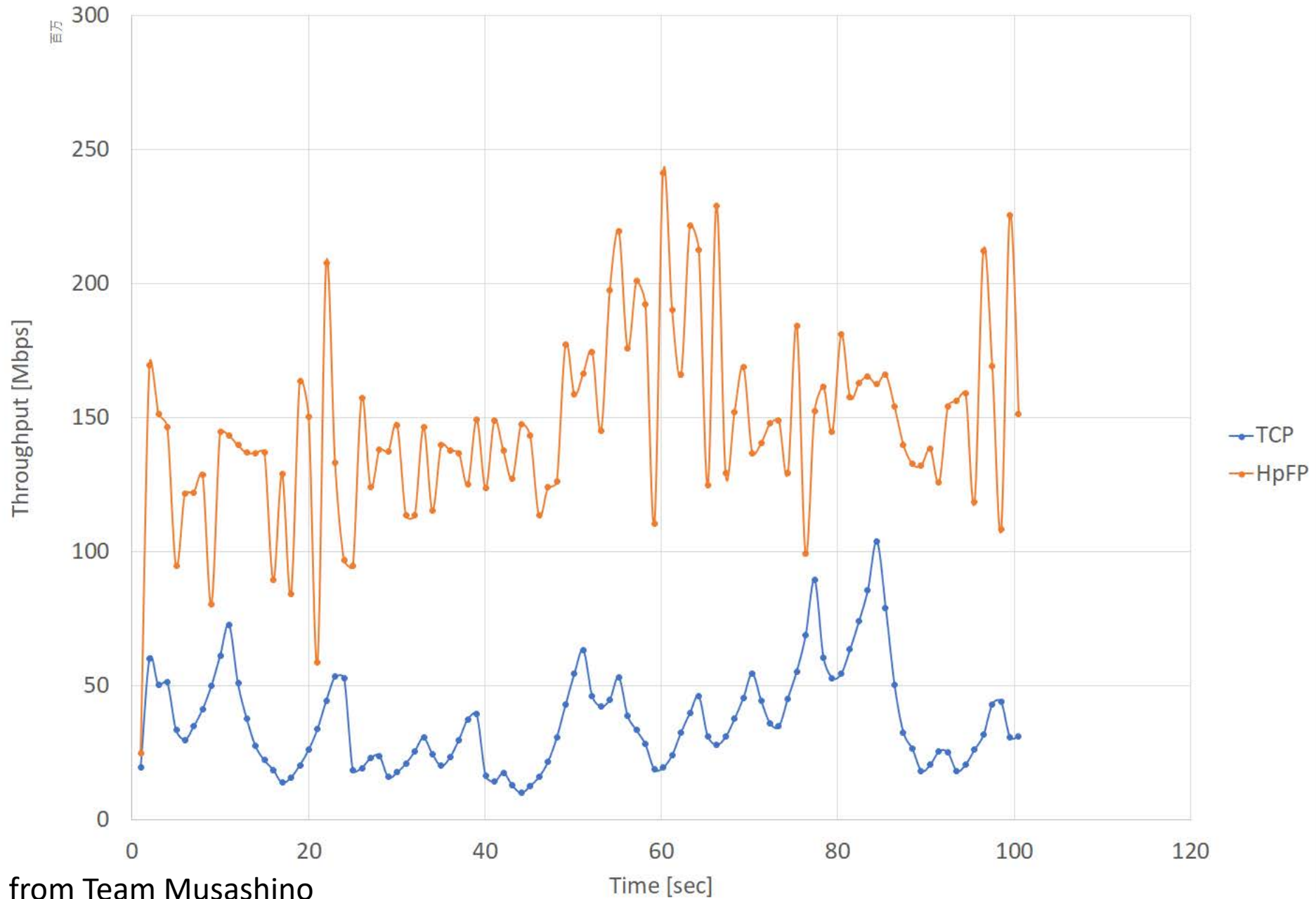
Starlink Upload Throughput



Slides from Team Musashino



Starlink Download Throughput



Slides from Team Musashino

Some comments from participants

- It was an opportunity for them to test their software in such a global and high-speed network.
- One team member doing his Phd found that it would be part of his thesis to test his optimised transport software in a global network.
- There was protocol that is well suited for the STARlink, which has high packet loss.
- Simplicity of mscp to installed.
- Use of AI techniques to optimised the parameters, e.g. # of workers, block size, etc..



SupercomputingAsia 2024

19 – 22 February 2024, ICC Sydney



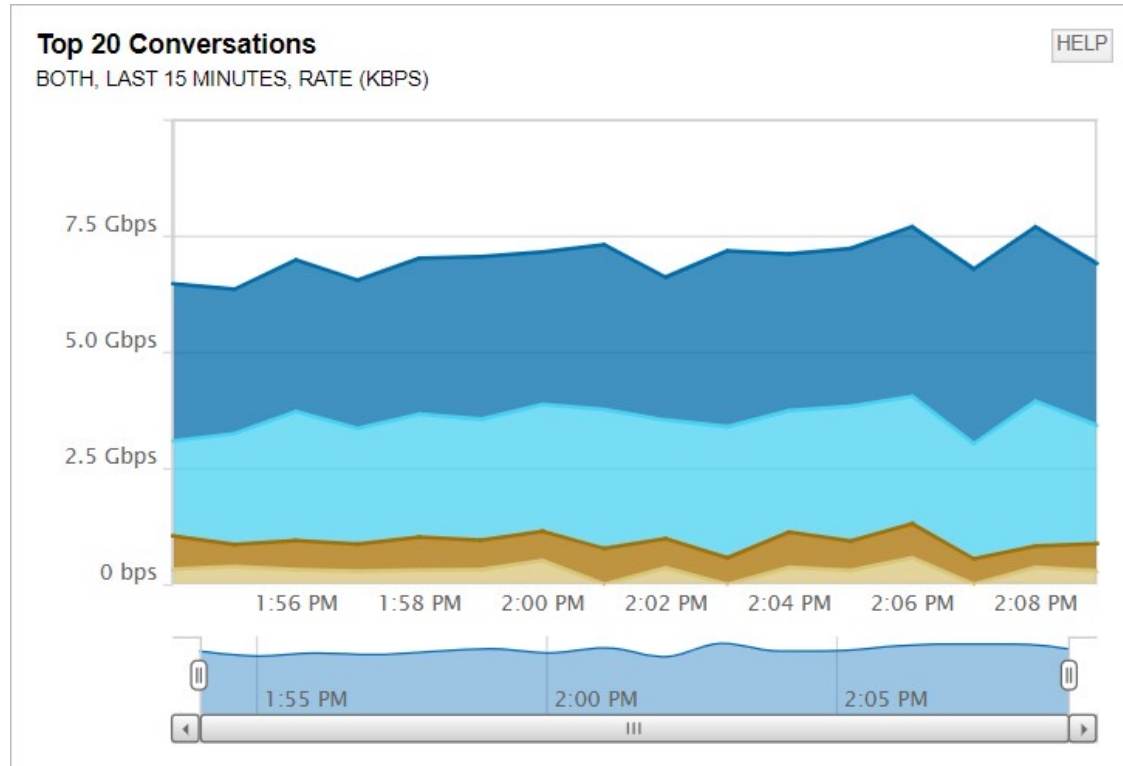
	DMC19	DMC20	DMC21	DMC23
Region	Asia-Pacific & USA	Asia-Pacific, US & Europe	Asia-Pacific, US, Europe, Saudi Arabia	Asia-Pacific, US, Europe, Saudi Arabia, South America
Network Architecture	L2-VLAN	L2-VLAN	L3_L2 only between AARnet and NICT	L3 and STARlink
IP version	IPv4	IPv4	IPv4 & IPv6	IPv4 & IPv6
DTN OS	CentOS 7.5	CentOS 7.6	CentOS 7.9 Ubuntu 18.04 LTS	Ubuntu 20.0.6 LTS
Container engine	Singularity v2.6	Singularity v3.2	Singularity v3.7 Rootless Docker v20.10.08	Singularity v3.7.3 Rootless Docker v24.0.6
Test Scenario	Two-way transfer	One-way simultaneous,	One-way simultaneous IPv4 & IPv6	One-way simultaneous and STARlink IPv4 and IPv6
Monitoring Solution	MRTG	MRTG, NetData + Grafana	MRTG, PerfSONAR + MaDDASH, Prometheus Node	MRTG, PerfSONAR + MaDDASH, Prometheus Node

Acknowledgement

- Jian Ma, Kodai Motohashi, Hirotaka Sato (KDDI) for their support to set-up the Starlink.
- Network and DTN partners for the support for the Data Mover challenge
- Participants of DMCs

Thank You

High Speed Data Transfer



Gadi supercomputer
NCI(AUS)



ASPIRE 2A
Singapore

