2019 Japan-US Network Opportunity 2 (JUNO2)

Project Title: *Disaster-Resiliency Strategies for Next-Generation Metro Optical Networks*

ATN: 1818972

Speakers:
Dr. Yoshinari Awaji (NICT)
Dr. Massimo Tornatore (UCDavis)

October 11, 2019
Principal Investigator Meeting @ Chicago (IL)

UCDAVIS

NICT

Project Team

- **PIs**
  - Massimo Tornatore; University of California, Davis; PI
  - Biswanath Mukherjee; University of California, Davis; Co-PI
  - Yoshinari Awaji; NICT (Japan); Team leader (PI Japan side)

- **@NICT**
  - Sugang Xu; NICT (Japan); Collaborator
  - Yusuke Hirota; NICT (Japan); Collaborator
  - Masaki Shiraiwa; NICT (Japan); Collaborator

- **@UCD**
  - Dr. Sifat Ferdousi; University of California, Davis; PDF
  - Andrea Marotta; University of L'Aquila, Italy; Visiting Student
  - Giap Le; University of California, Davis; GSR

- **International Collaborators**
  - Prof. Francesco Musumeci; Politecnico di Milano, Italy
  - Prof. Gustavo Bittencourt; University of Bahia, Brasil

UCDAVIS

NICT

Two in-person meetings

- Davis, Feb 27th to Mar. 1st
  - Pre-OFC visit by Dr. Awaji and Dr. Hirota
- Sendai, July 4th and 5th
  - Pre-OECC visit by Dr. Tornatore
- Several other interactions via email and telco

Massimo Tornatore
@MaxTornatore

Two intense days discussing disaster resiliency for next generation optical metro networks in our JUNO2 #NSF project . Thanks to Dr. Yoshinari Awaji and Dr. Yusuke Hirota from #NICT for visiting us in #UCDavis. @dmedhi

Massimo Tornatore
@MaxTornatore

Visiting #NICT in Sendai, Japan, to discuss disaster resiliency in next generation optical networks for our #JUNO2 #NSF project. Inter-carrier recovery, #MachineLearning for failure management, and several other research directions can be explored over NICT's #OPCI testbed

1:14 AM - 5 Jul 2019

Motivation

A new generation of optical metro networks is needed to turn the vision of "Smart Cities" into reality

- From a **rigid ring-based aggregation infrastructure** to a **composite network-and-computing ecosystem** to support critical 5G services (e.g., autonomous driving)

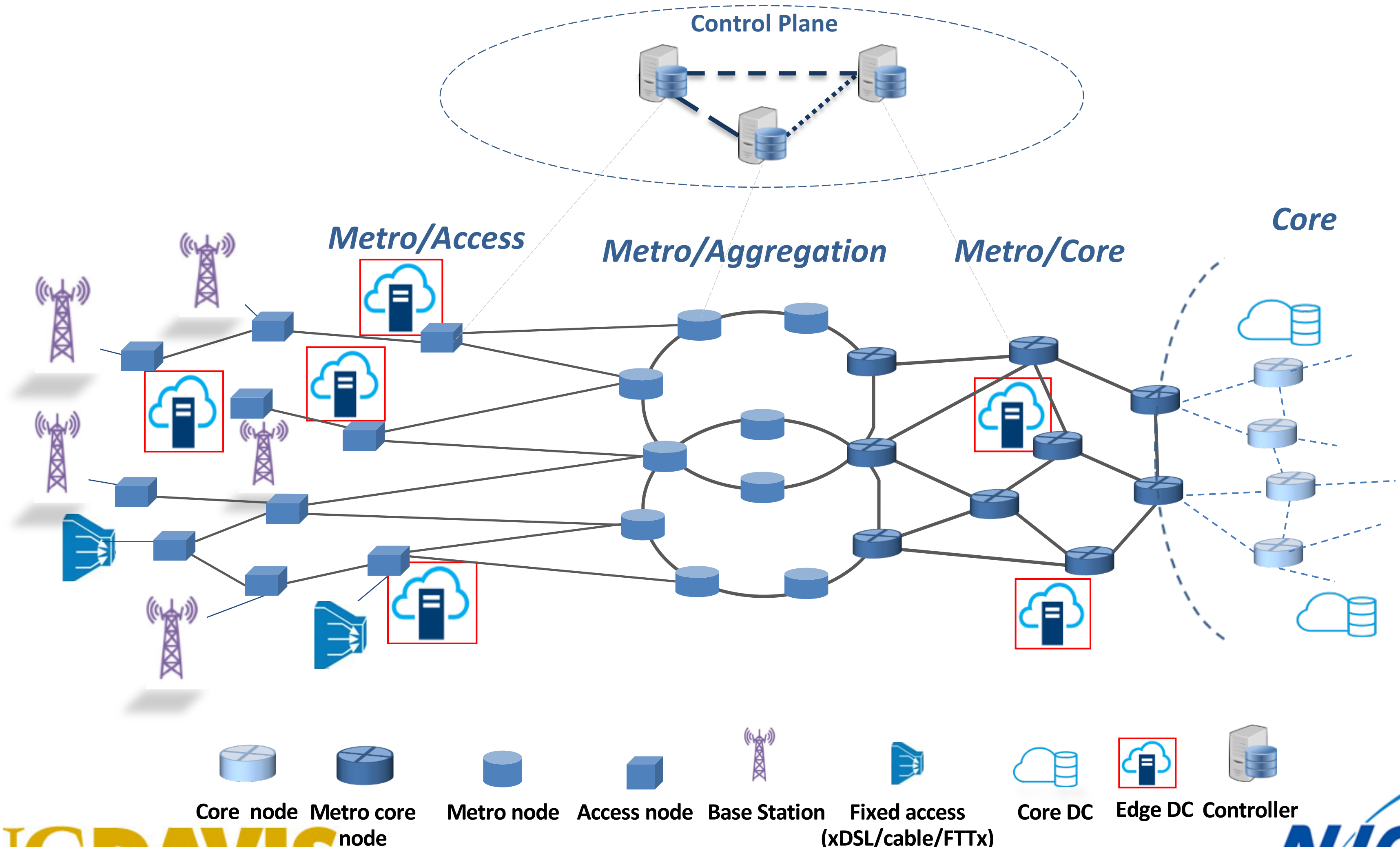
- Several technical enablers:
  - Increased reconfigurability enabled by **SDN**
  - **Integration of optical and wireless** access networks
  - Metro nodes becoming edge data centers (**edge computing**)
  - **Network slicing** to logically partition network, computing, and storage resources
  - …

UCDAVIS

NICT

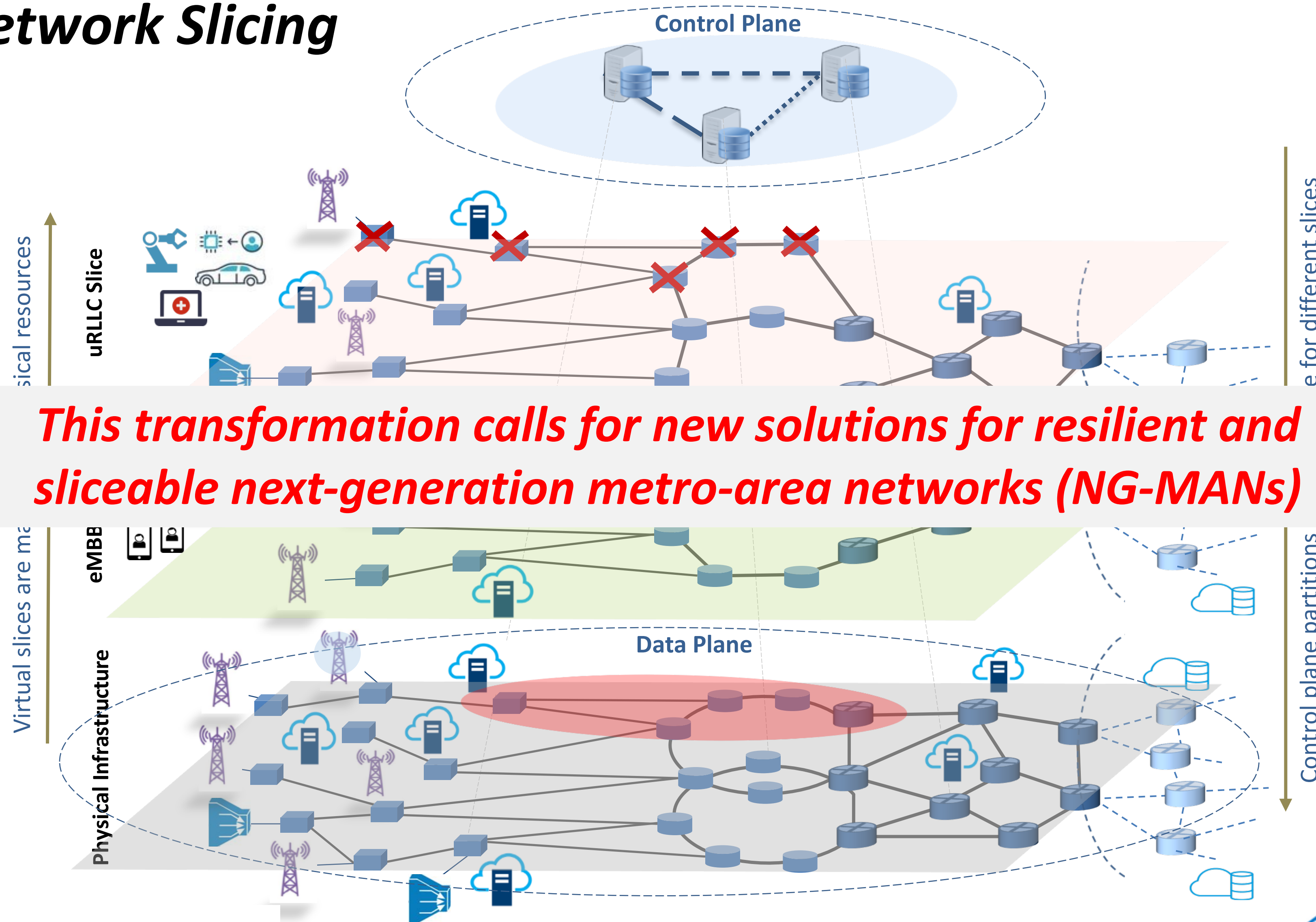
Evolution of Metro Access Networks (I)

Edge computing and SDN



Evolution of Metro Access Networks (II)

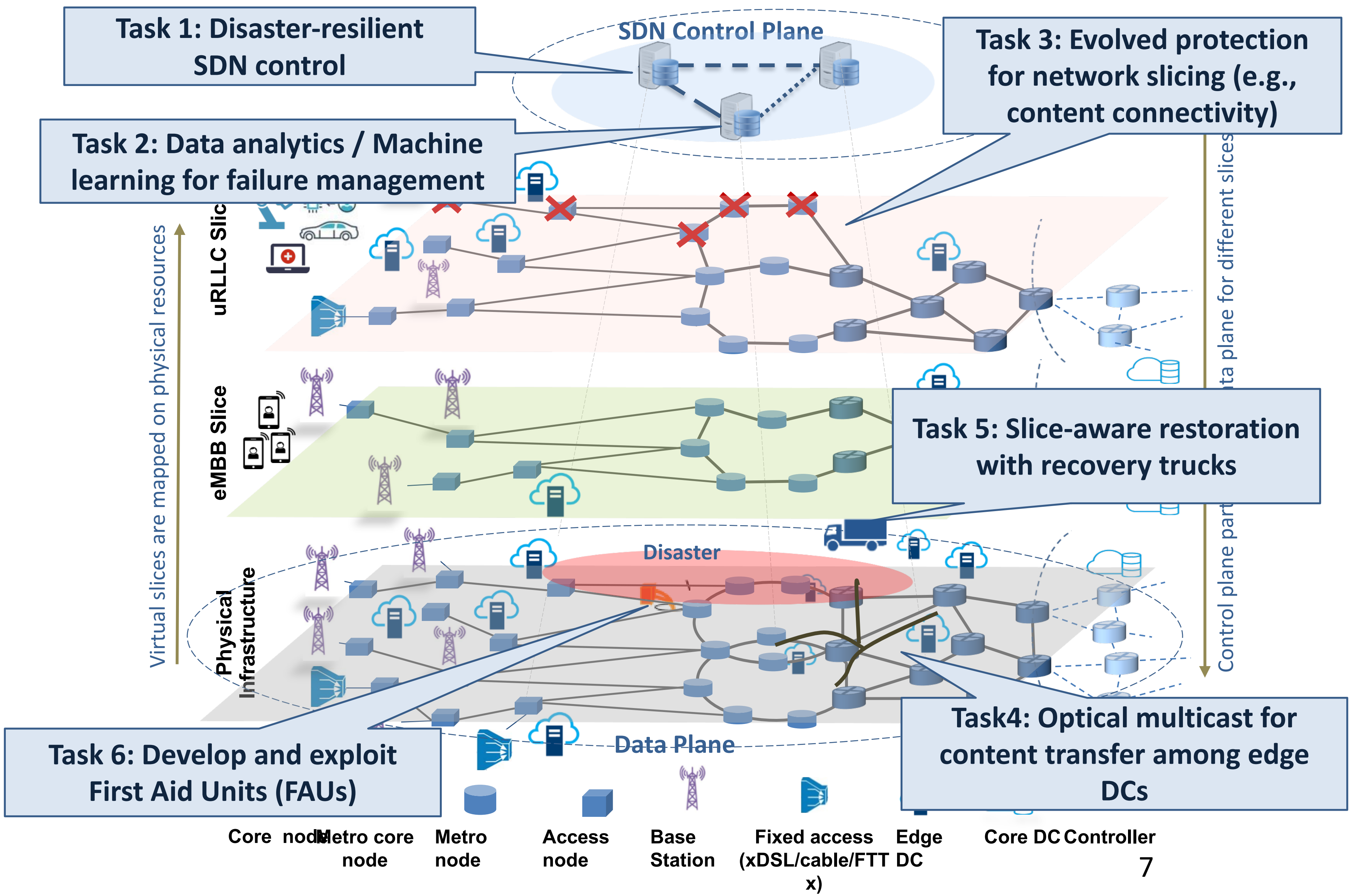
Network Slicing



This transformation calls for new solutions for resilient and sliceable next-generation metro-area networks (NG-MANs)

Disaster affecting underlying NG-MAN

Project Overview



Overall Topics

	Disaster-resilient NG-MANs		
Phase	Pre-disaster phase		Post-disaster phase
Theme	Theme 1: Prevention Disaster-resilient design of the SDN control plane	Theme 2: Preparedness Protection strategies for sliceable NG-MANs	Theme 3: Fast recovery Recovery strategies with portable devices in NG-MANs
UC Davis	Task 1 Disaster-resilient SDN controller placement	Task 3 Disaster-resilient content-connected slice mapping	Task 5 Slice-aware restoration with recovery trucks for NG-MAN
NICT	Task 2 SDN modules for network slicing and data analytics	Task 4 Optical multicast for effective content transfer among edge DCs	Task 6 Develop and exploit First Aid Unit (FAU)
Key technologies	SDN/NFV, Machine learning	Flexible grid & multicast transmission	Novel portable/movable disaster recovery systems

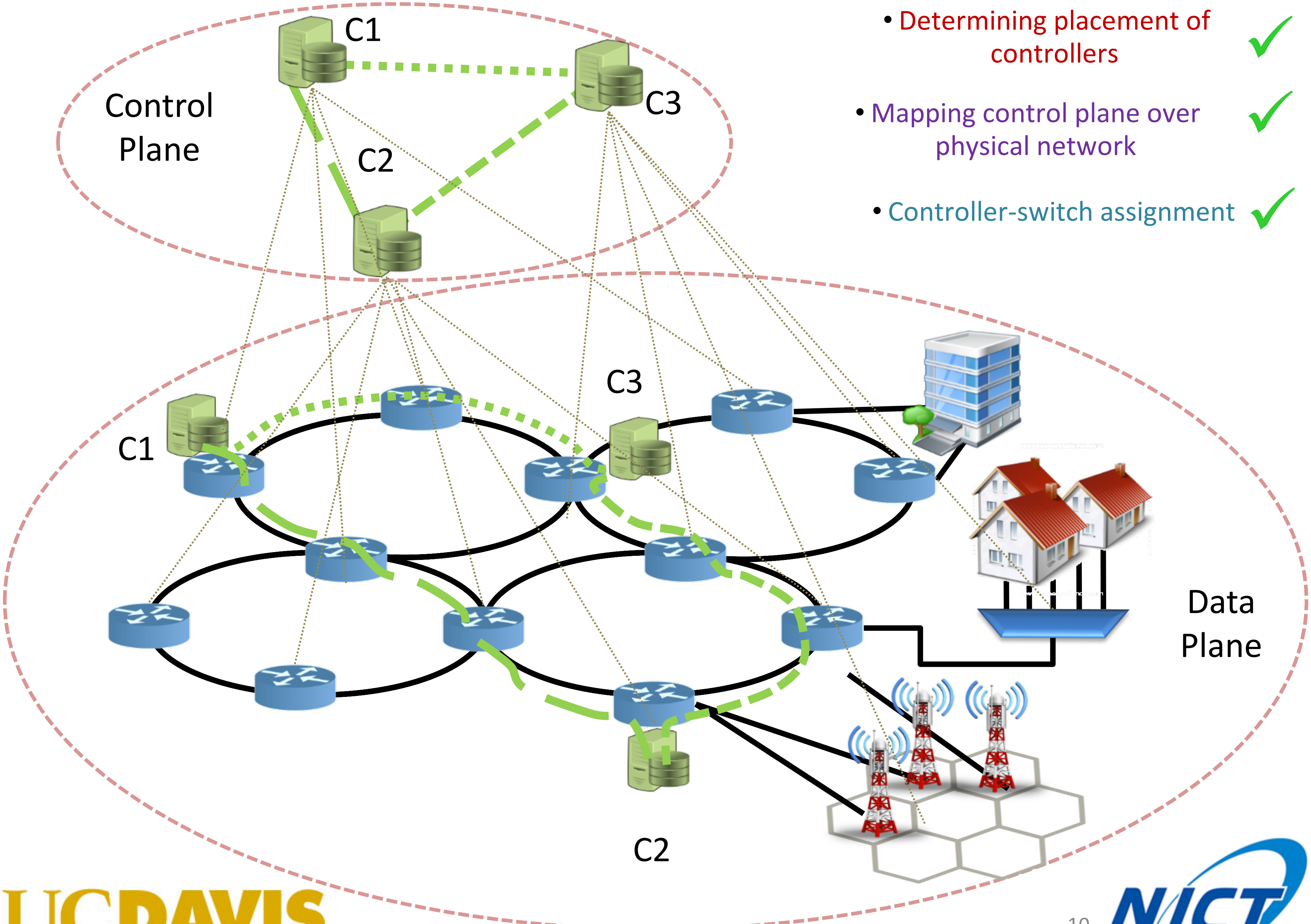
UCDAVIS

NICT

Task 1
Disaster-resilient SDN controller placement

- Mostly concluded activity
  - S. Savas, M. Tornatore, F. Dikbiyik, A. Yayimli, C. Martel, and Biswanath Mukherjee, "RASCAR: Recovery-Aware Switch-Controller Assignment and Routing in SDN," in *IEEE Transactions on Systems and Network Management*, vol. 15, no. 4, pp. 1222-1234, Dec. 2018
  - R. Lourenco, S. Savas, M. Tornatore, B. Mukherjee, "Robust Hierarchical Control Plane for Transport Software-Defined Networks", in *Optical Switching and Networking*, Vol. 30, pp. 10-22, Nov. 2018

- Determining placement of controllers ✓
- Mapping control plane over physical network ✓
- Controller-switch assignment ✓

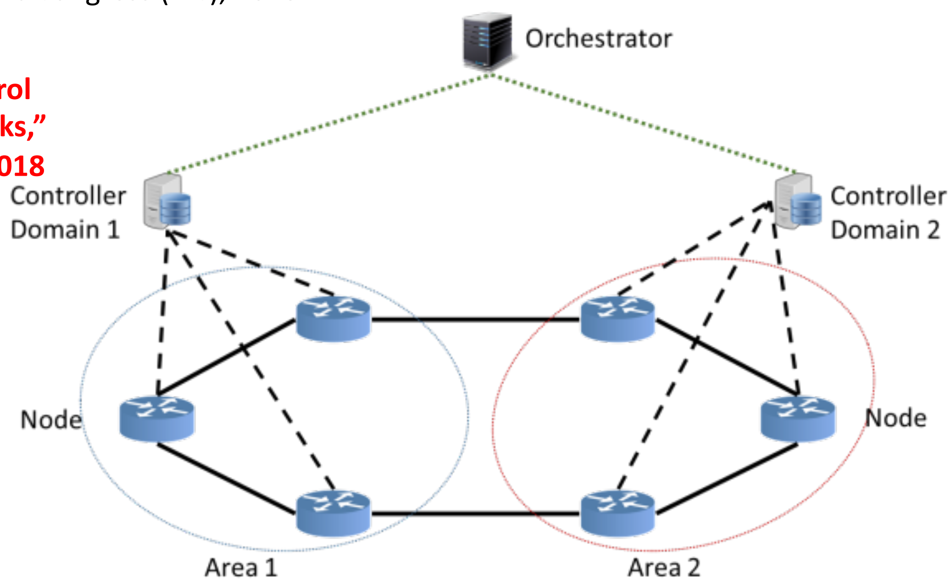


UCDAVIS

NICT

Our contribution on fault-tolerant controller placement (CP)

- **[Pre-planned controller replicas]** F.J. Ros, P.M. Ruiz, Five nines of southbound reliability in software-defined networks, in: Proc. of the ACM HotSDN, New York, NY, USA, 2014 & (B. Killi, et al., Capacitated next controller placement in software defined networks, IEEE Trans. Netw. Service Manage. 14 (3) (2017) 514–527)
- **[Path diversity]** F. Müller, et al., Survivor: an enhanced controller placement strategy for improving SDN survivability, in: Proc. of the IEEE GLOBECOM, 2014
- **[Disaster awareness]** S. Savas, et al., "Disaster-resilient control plane design and mapping in Software-Defined Networks," *In Proc. of HPSR,* Budapest, Hungary, July 2015
- **[Malicious Attacks]** D. Santos, A. de Sousa, C.M. Machuca, Robust SDN controller placement to malicious node attacks, in: Proc. of IEEE DRCN, 2018.
- **[Several types of failures]** D. Hock, et al., Pareto-optimal resilient controller placement in SDN-based core networks, in: Proc. of the 25th International Teletraffic Congress (ITC), 2013.
- **[T-SDN (Orchestrator+Controllers)]**
  - **R. Lourenço, et. al, "Robust hierarchical control plane for transport software-defined networks," Optical Switching and Networking, vol. 30, 2018**



- Ok controller has been properly placed…

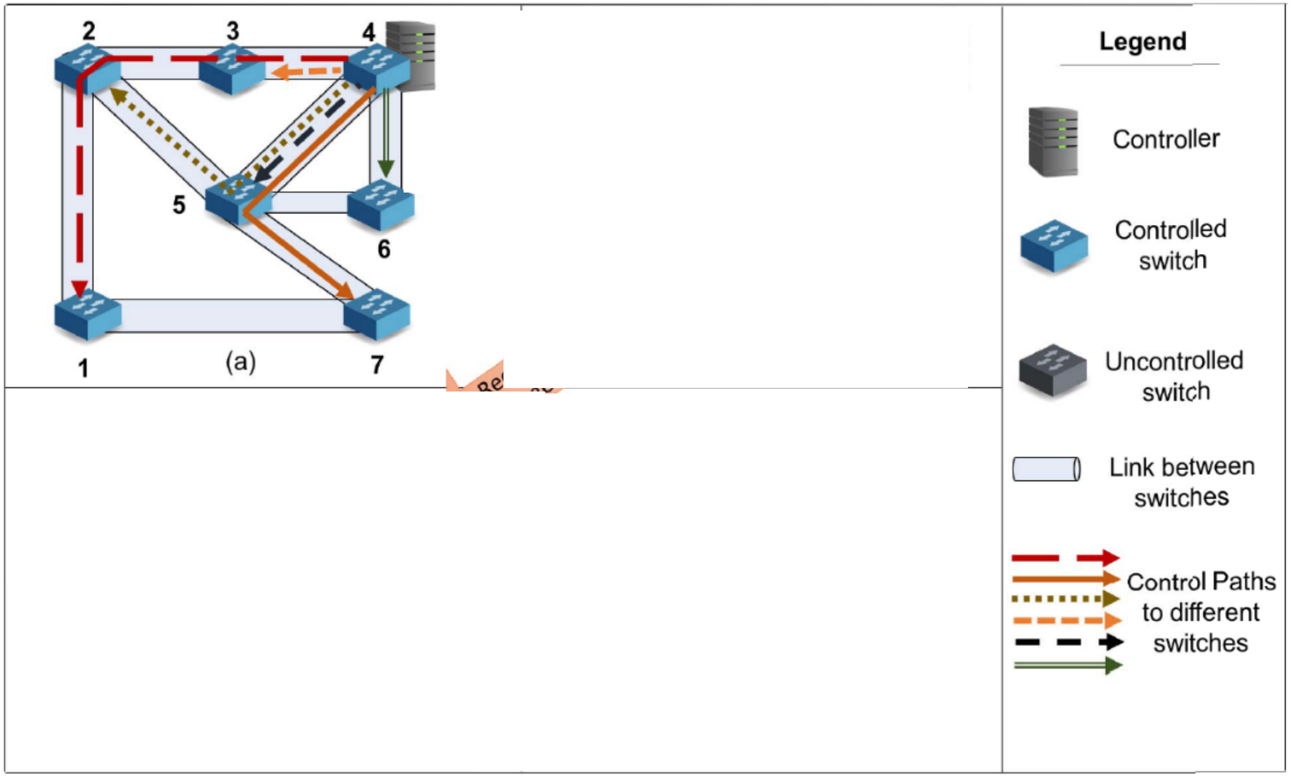
- .. still *how we interconnect controllers to switches* is crucial to minimize recovery time!

S. Savas et. al, "RASCAR Recovery-Aware Switch-Controller Assignment and Routing in SDN", *IEEE Transactions on Network and Service Management*, vol. 15, no. 4, pp. 1222-1234, Dec. 2018

Looking at "controller to switch" paths!

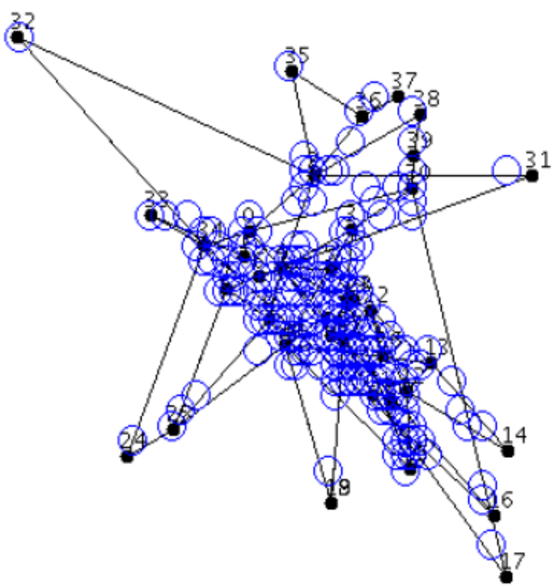
Multi-stage recovery

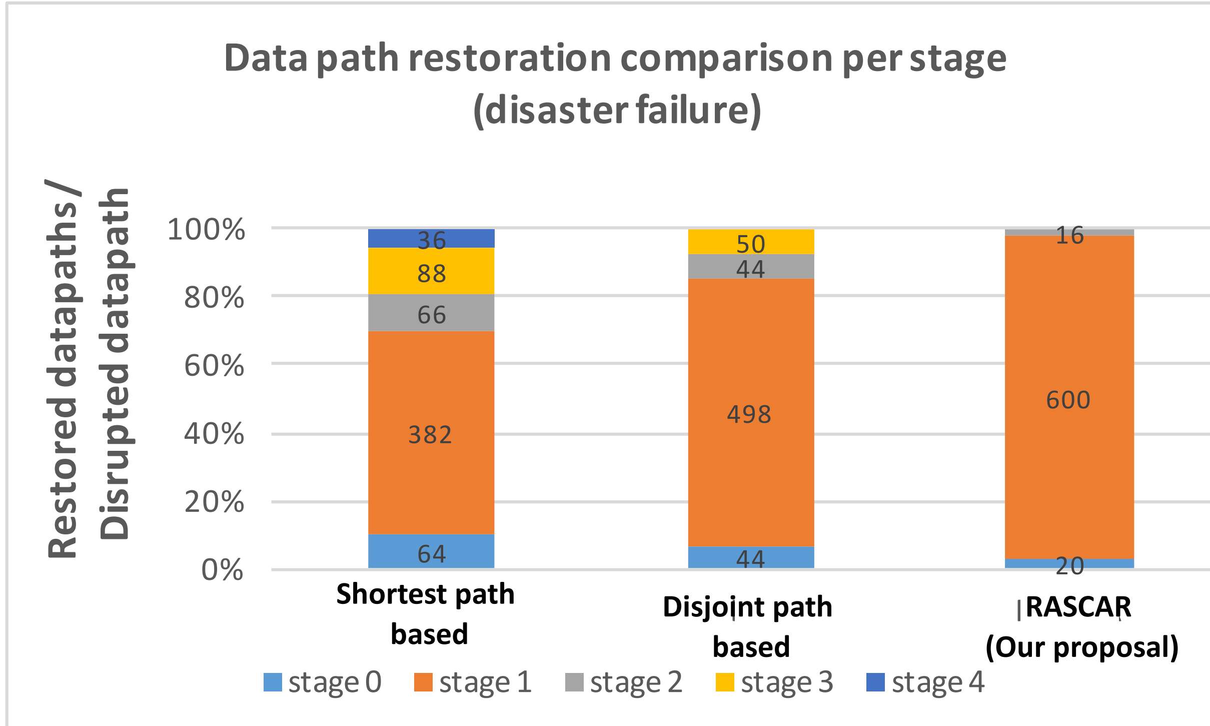
- Even a single failure affects multiple switch-to-controller **control paths**
- When switches lose control paths, they become "**uncontrolled**":
  - route traffic using old flow entries
  - cannot exchange control messages (e.g., flow setup request, flow installation)
  - cannot be used for data path restoration



Results: recovery speed vs. cost

- **GEANT Network**
- **All possible disasters with radius = 100km**

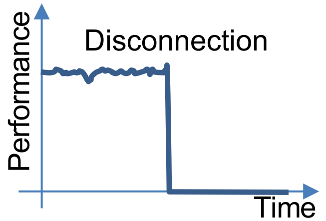


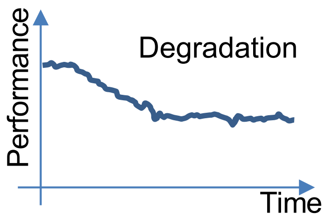


<u>Note: what about additional cost?</u>

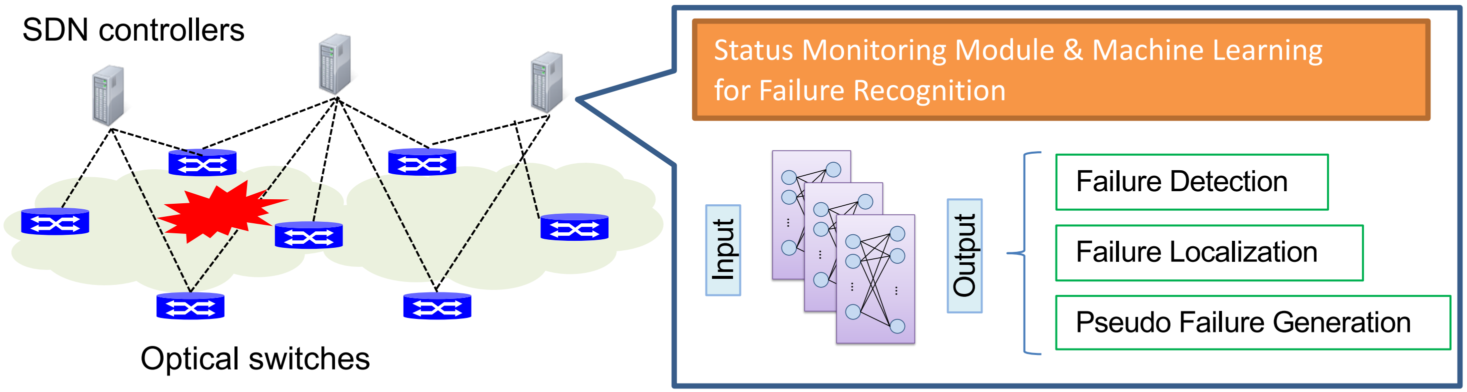
Less than 1% additional resource consumption.
Only control paths become longer!

Task 2: SDN Modules for Network Slicing and Data Analytics





SDN to support advanced network functionalities



- ***SDN-Based monitoring for Detecting Failures***: Failure detection mechanisms that make use of the global network view provided by SDN to promptly correlate disaster failures.
  - Detection mechanism for disconnected links and failed nodes
  - Machine Learning based failure recognition mechanisms for incomplete network connectivity

- Purpose:
  - With the distributed SDN controllers, perform the resilient slice management, and, failure detection and analytics ➡ Support advanced network functionalities, quick response to failures suppressing performance degradation

UCDAVIS

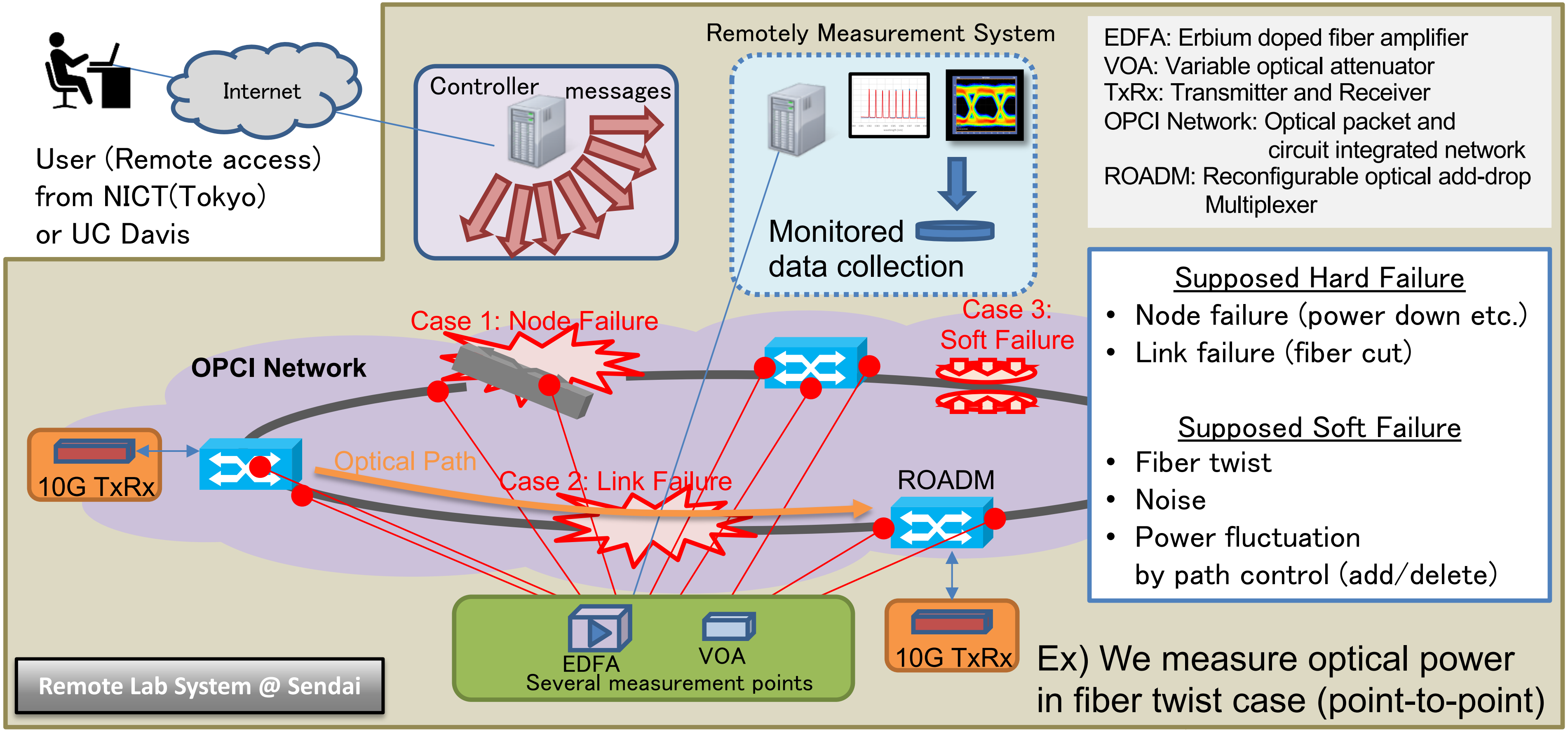
NICT

Outline of 1st year's research in Task 2

- Issues (3 years)
  - Create dataset for learning including failure cases
  - Develop machine learning algorithms
  - Establish failure recognition scheme in SDN control
  - Cooperative operation between multiple SDN controllers

- 1st years achievements
  - Preliminary experiments to obtain dataset.
    - Only optical power
    - Limited failure events
    - Manually setup for every events
  - In addition, we discussed a closer collaboration research.

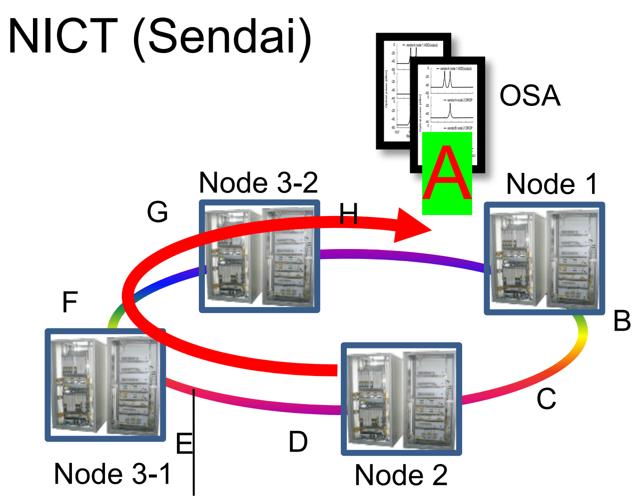
Development of Sendai Testbed for Advanced Network Functionality



- In Task 2, we obtained several data of optical power (1 sec interval).
  - Not enough data to learning and recognize failures.
- Next step: We develop CLI-based multipoint measurement system for collecting large dataset and utilize the dataset.

Collaborative Research

- Real data acquisition at Sendai testbed.
  - Locally setup experiments at NICT side
  - Data acquisition in local (NICT side)
  - Remote control of several equipment and data acquisition (UC Davis side)
  - Evaluation and discussion

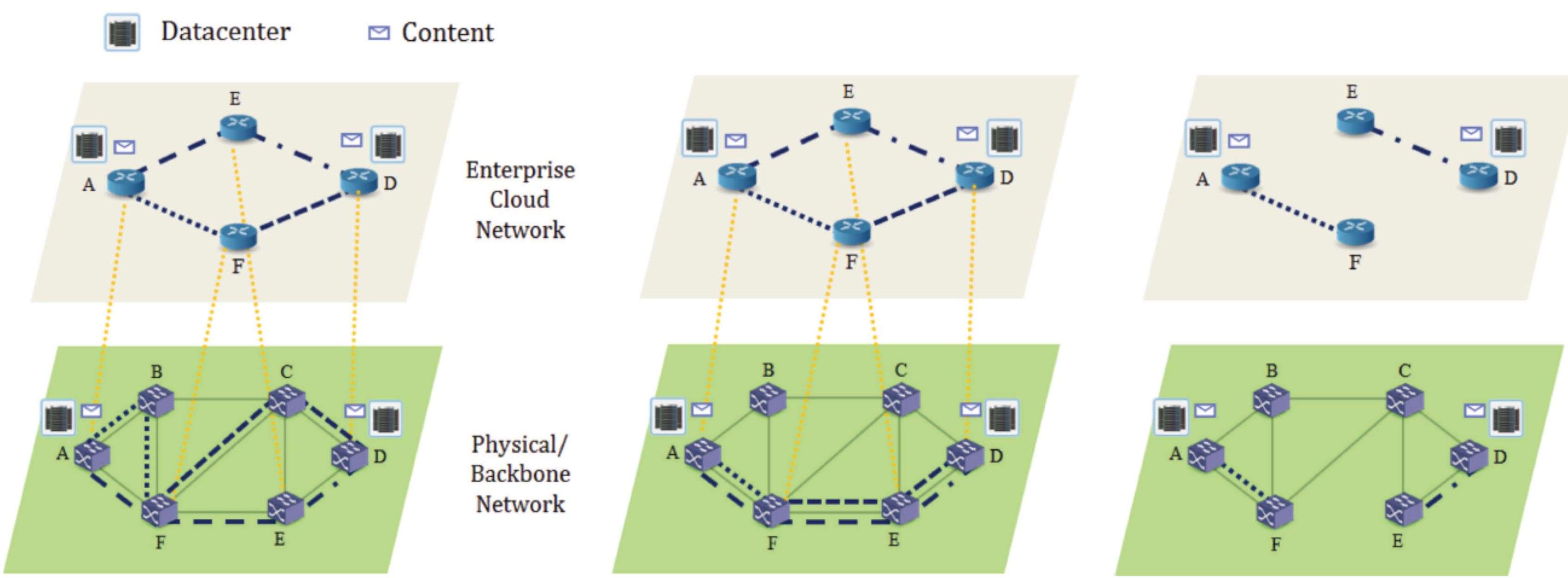


Develop failure recognition algorithm in both side and discuss in detail for better algorithm.

Task 3: "Content Connectivity" for Disaster Resiliency

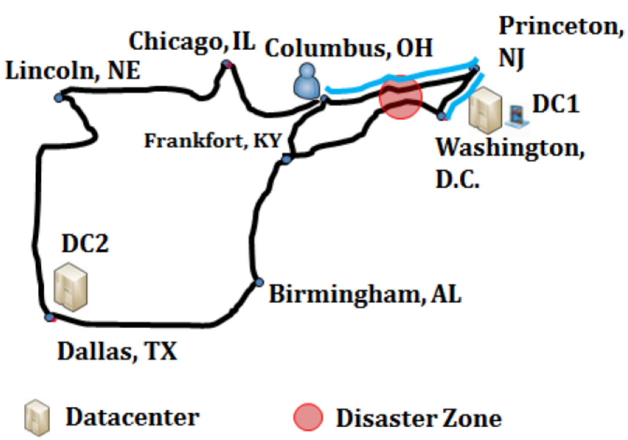
Network Connectivity - **ensure reachability of any pair of nodes in the network (end-to-end) in case of failure**

Content Connectivity - **ensure reachability of content from any point of a network (end-to-content) in case of failure**

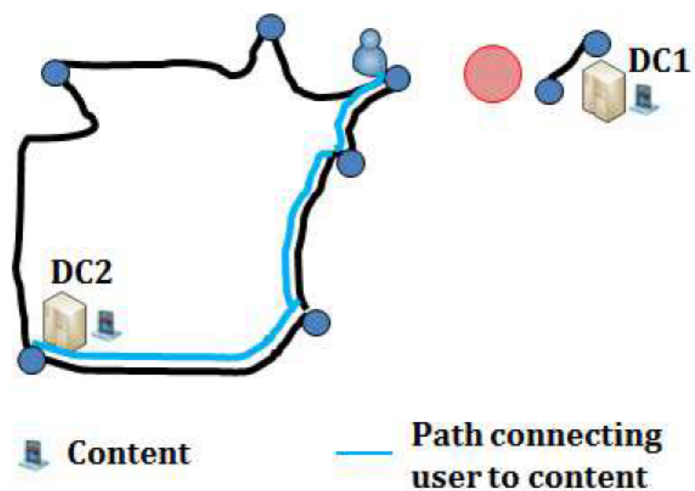


Content Connectivity: Why it is important?

- Content connectivity:
  - **costs less** than network connectivity
  - might be **feasible** when network connectivity is not



Network connectivity not possible



Content connectivity possible

[3] M. Farhan Habib, M. Tornatore and B. Mukherjee, "Fault-tolerant virtual network mapping to provide Content Connectivity in optical networks," in *Proc. IEEE/OSA Opt. Fiber Comm. Conf. (OFC'13)*, Mar. 2013.

Current activity: generalization of content connectivity modeling against n link failures

We developed an Integer Linear Program

Given:

- ✓ Physical topology: $G_P(V_P,\ E_P)$
- ✓ Logical topology: $G_L(V_L,\ E_L)$
- ✓ Set of data center locations: D
- ✓ Number of wavelengths per fiber: W

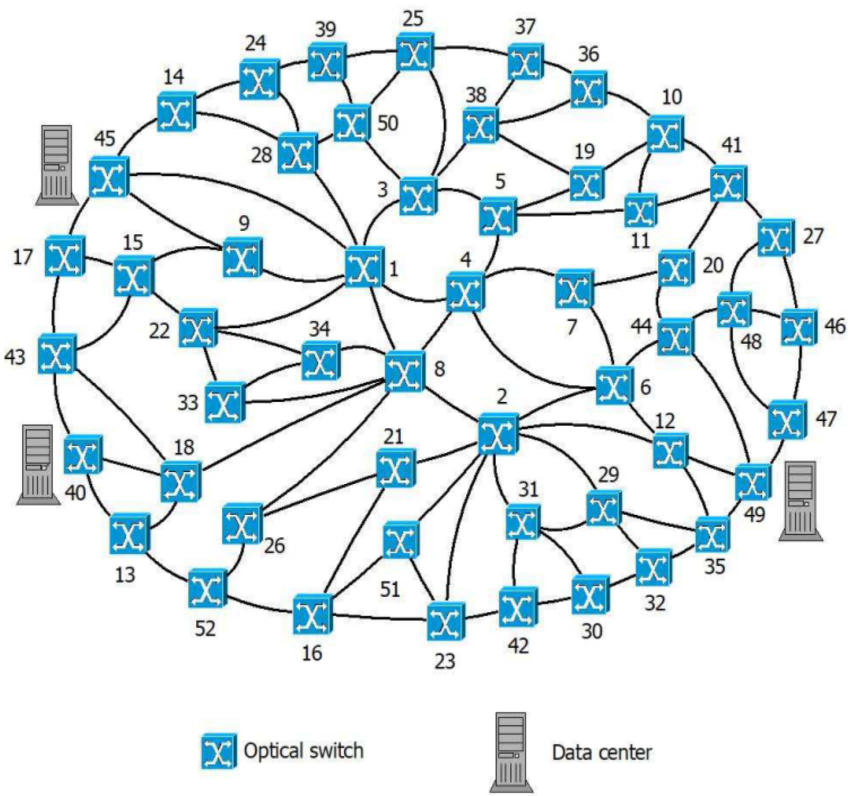
DONE

Outputs:

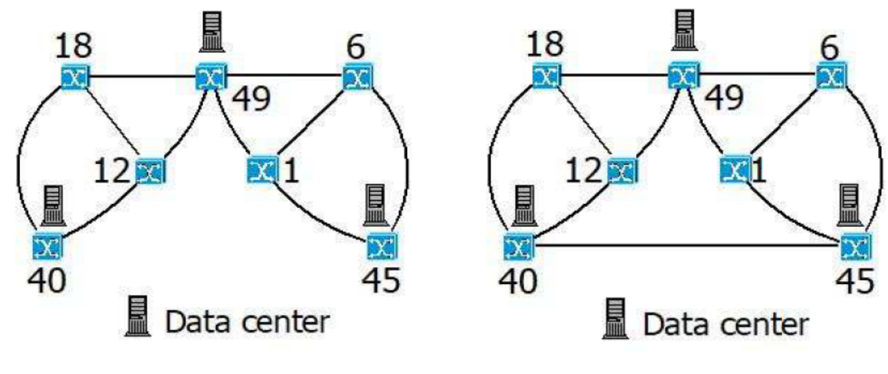
- ✓ Content-connected mapping of a virtual network over physical network against n−physical-link failures (previous work only up to n=2)
- ✓ Minimize total wavelength channel number

G. Le, A Marotta, S. Xu, Y. Hirota, Y. Awaji, M. Tornatore, B. Mukherjee, "Logical Network Mapping With Content Connectivity Against Multiple Link Failures in Optical Metro Networks", IEEE ANTS 2019

Illustrative Numerical Examples



Physical Network:
Modified Telecom Italia Network
(52 nodes, 98 links)

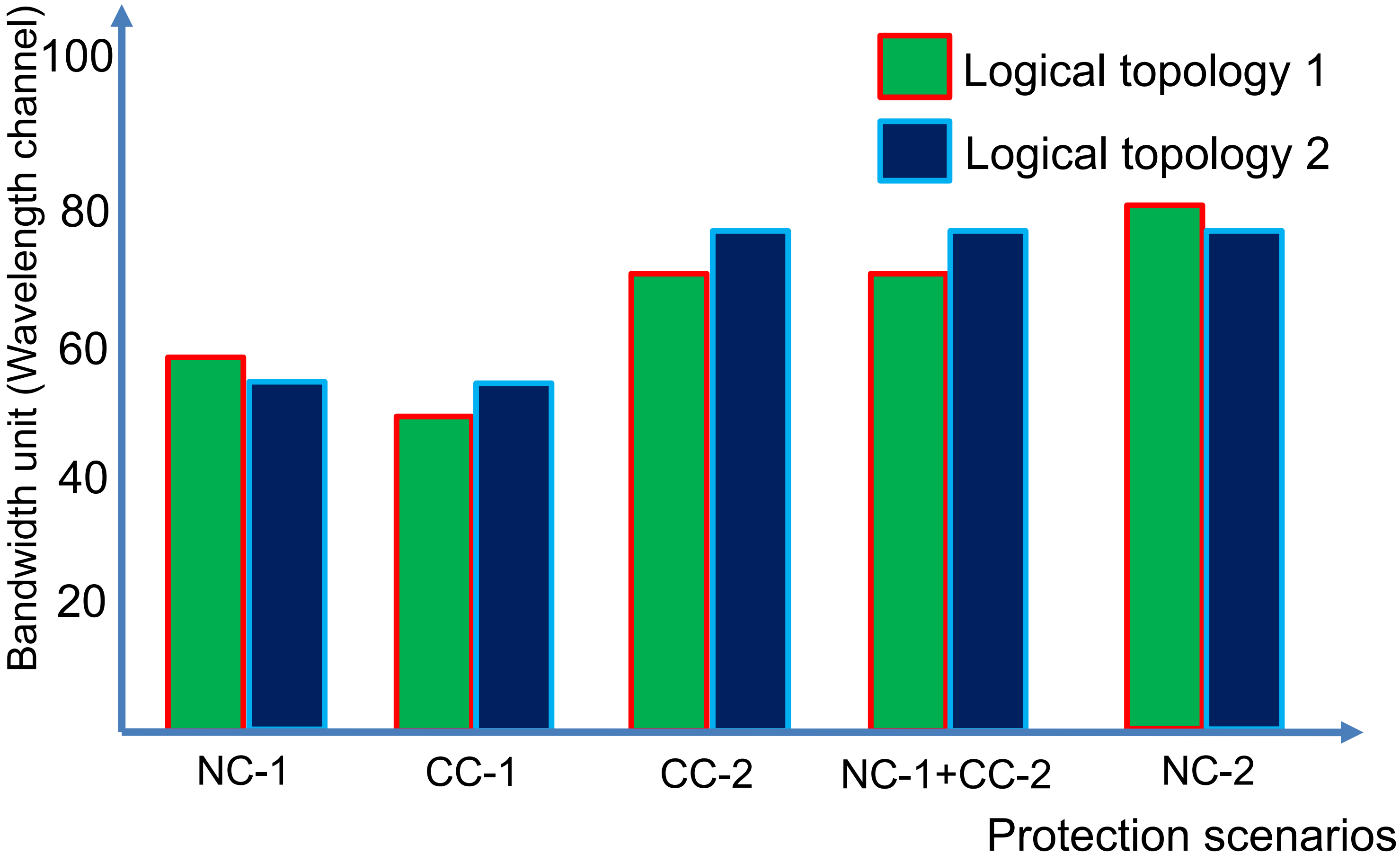


a) Logical topology 1

b) Logical topology 2

Virtual Networks
(7 nodes, 10 links (a), and 11 links (b)

Illustrative Numerical Examples



Our new approach is more generic (arbitrary n) and more scalable than those in [10], [11]

[10] M. F. Habib, M. Tornatore, and B. Mukherjee, "Fault-tolerant virtual network mapping to provide Content Connectivity in optical networks," *Proceedings of OFC*, Mar. 2013.

[11] A. Hmaity, F. Musumeci, and M. Tornatore, "Survivable virtual network mapping to provide content connectivity against double-link failures," *Proceedings of DRCN*, Mar. 2016.

Content Connectivity [CC]: In-Progress Research

- ✓ Generalize conclusions:
  - ✓ When NC > CC and when NC = CC?
  - ✓ When NC not possible but CC possible?

- ✓ Main question: to ensure a given survivability level to a 5G service using a slice, shall I add more data centers or more logical connections?

Task 4: Optical Multicast for Effective Content Transfer among Edge DCs

"Content connectivity" can be realized only if relevant data is replicated in several edge DCs

- *Constant and intensive synchronization and backup procedures among edge DCs via optical multicasting (Verification on OPCI network testbed)*

- ***Selection of multicast tree leaves for disaster resiliency*** - How do we determine which nodes receive the synchronization data of the multicast tree?
- ***Slicing and multicasting*** - How to slice network resources for multicast transmissions?
- ***Verification of optical multicasting with SDN control in Sendai testbed*** - SDN controllers need to monitor transmitted data and to construct multicasting slices considering the locations of edge DCs.

OPCI network testbed @Sendai NICT branch

UCDAVIS

NICT

Outline of 1st year's research in Task 4

- In Task 4, we study how to deliver large size data for enhancing content connectivity.
- 3 types of transmissions for data transfer to multiple receivers.

	Required bandwidth	Elapsed time	Switching load
Serial unicast	Low	Very long	Not heavy
Parallel unicast	Very large	Short	Moderate
Optical multicast	Low	Short	Not heavy

* To duplicate packets electrically is not suitable for Large data transfer.

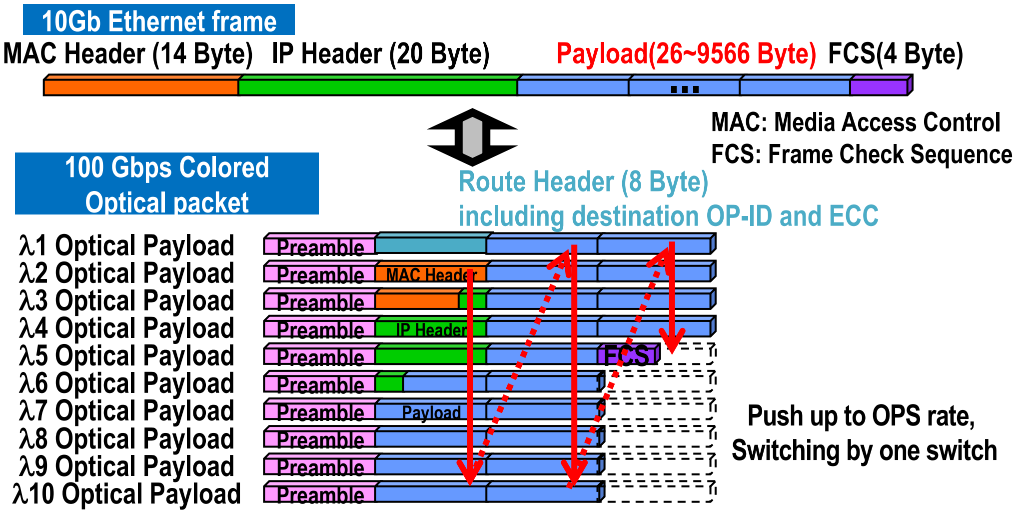
- Optical path multicast is difficult because …
  - Wavelength continuity constraint is difficult in case of multicast against unicast.
  - One 10G Tx cannot execute handshake to multiple 10G Rxs simultaneously.

- Proposal: Optical level multicast packet transmission
  – We firstly developed OPCI node and demonstrate its functionality.

UCDAVIS

NICT

Experimental Setup



- Demonstration on optical level multicast
  - Multicast 1: Background traffic
  - Multicast 2: Video Streaming
  - Multicast 3: Large File Transfer

Optical Multicast 1: Background Traffic

Demo movie

Optical Multicast 2: Video Streaming

Demo movie

Task 5: Slice-Aware Service Restoration with Recovery Trucks for NG-MANs (I)

Some examples of deployable recovery units (e.g., recovery trucks, FAUs)?

- In t e, recovery both repa *f/service* whi ing on (unl covery)
  - *How*
  - *Wh*
  - *Wh*

"Slice-aware" routing and deployment strategy to minimize downtime penalty and ensure fast restoration of important slices

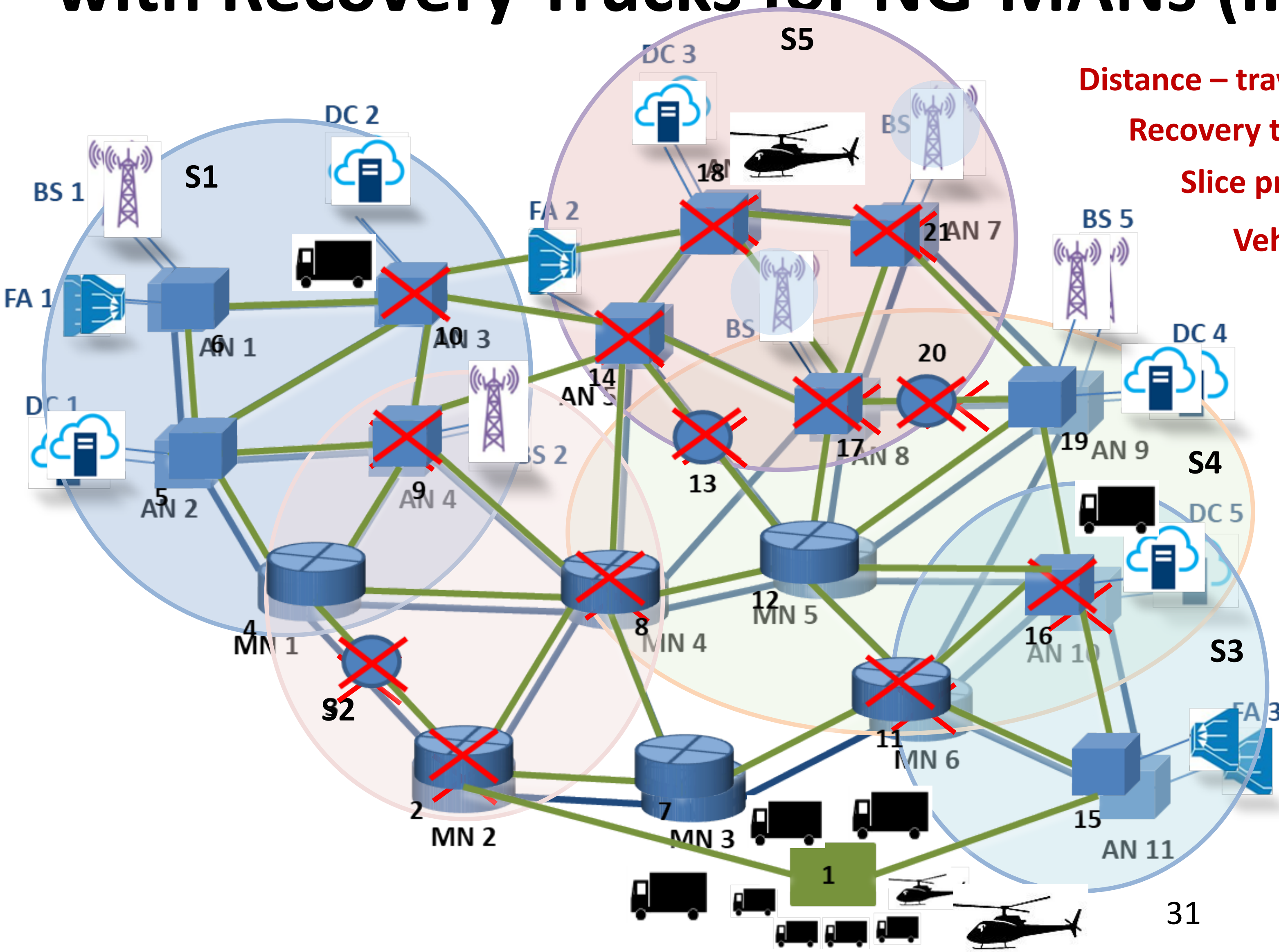
Task 5: Slice-Aware Service Restoration with Recovery Trucks for NG-MANs (II)

Distance – travel time

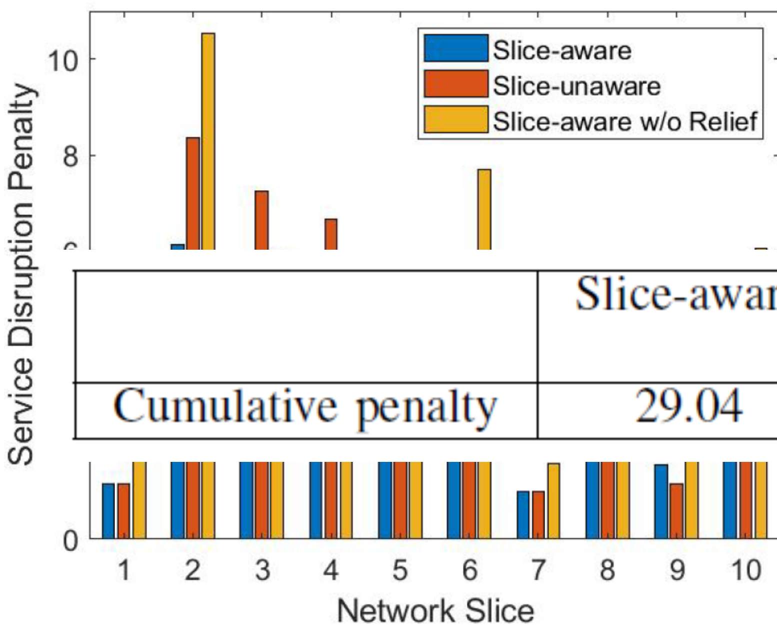
Recovery time

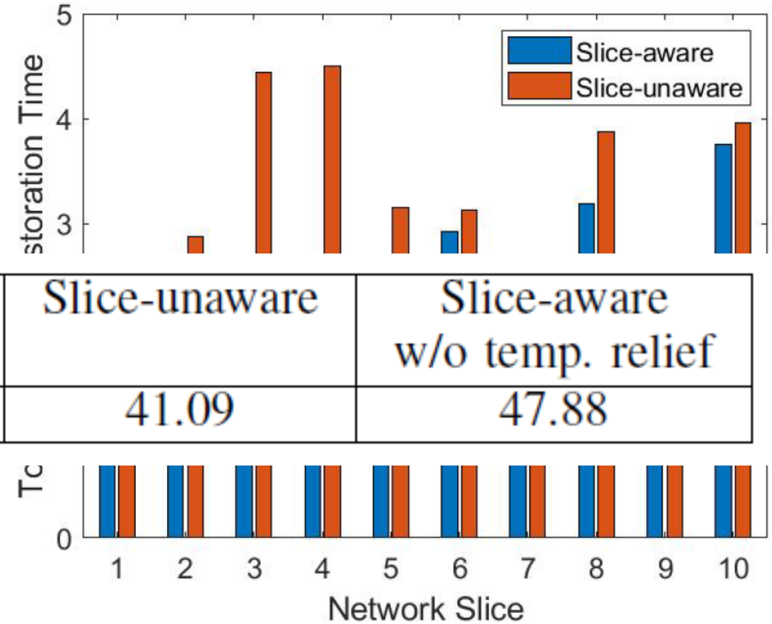
Slice priority

Vehicle type



Task 5: Slice-Aware Service Restoration with Recovery Trucks for NG-MANs (II)





	Slice-aware	Slice-unaware	Slice-aware w/o temp. relief
Cumulative penalty	29.04	41.09	47.88

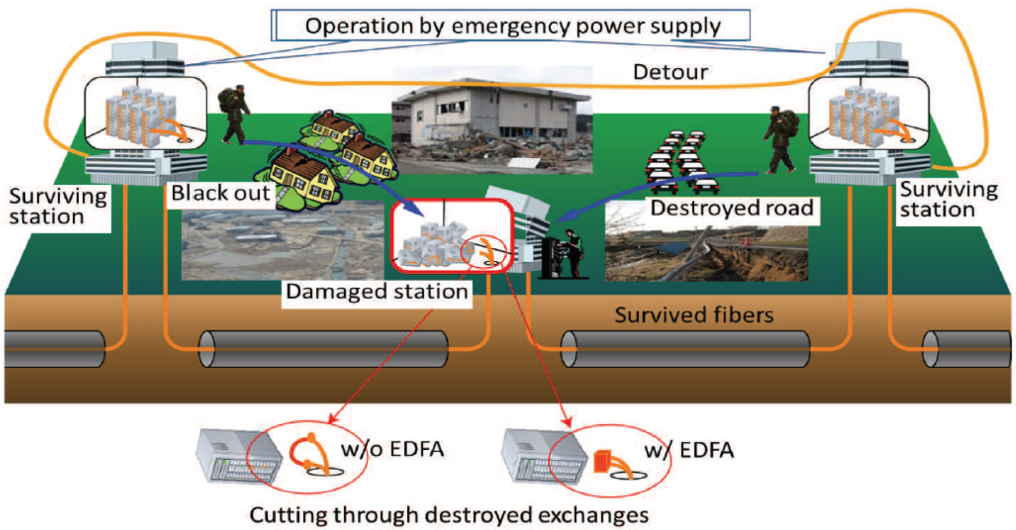
60% reduction in service disruption penalty

46% service-restoration time savings

S. Ferdousi, M. Tornatore, S. Xu, Y. Awaji, and B. Mukherjee, "Slice-Aware Service Restoration with Recovery Trucks for Optical Metro-Access Networks," accepted to IEEE GlobeCom conference 2019

Task 6: Develop and exploit First Aid Unit (FAU)

Interconnection of surviving facilities and FAU to restore network connectivity

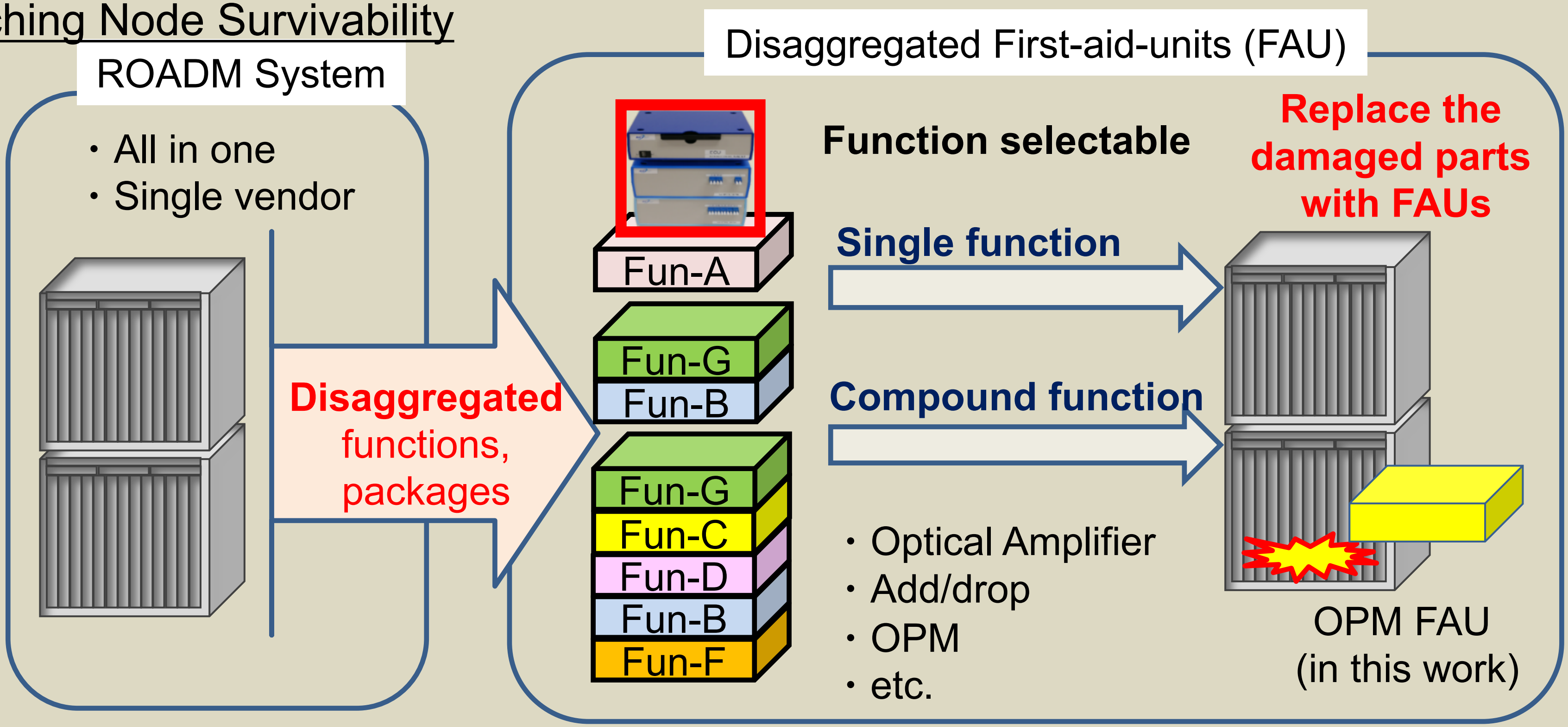


- *Impairment information collection utilizing monitoring systems*
- *Impairment-aware emergency optical network planning*
- *Development of supportive hardware and control system*

- *Impairment information collection*
  - Employ emergency optical performance monitoring (OPM) subsystem in FAU
  - Create emergency telemetry platform using outside surviving network resources (e.g., wireless/Internet)
- *Impairment-aware emergency optical network planning*
- *FAU upgrade*

Emergency First-aid-Unit (FAU)-based OPM Quick Recreation and Robust Telemetry in Disaster Recovery

Switching Node Survivability



Control Plane Survivability

- When disaster occurs, capacity in control plane also decreases or control plane may be heavily congested.

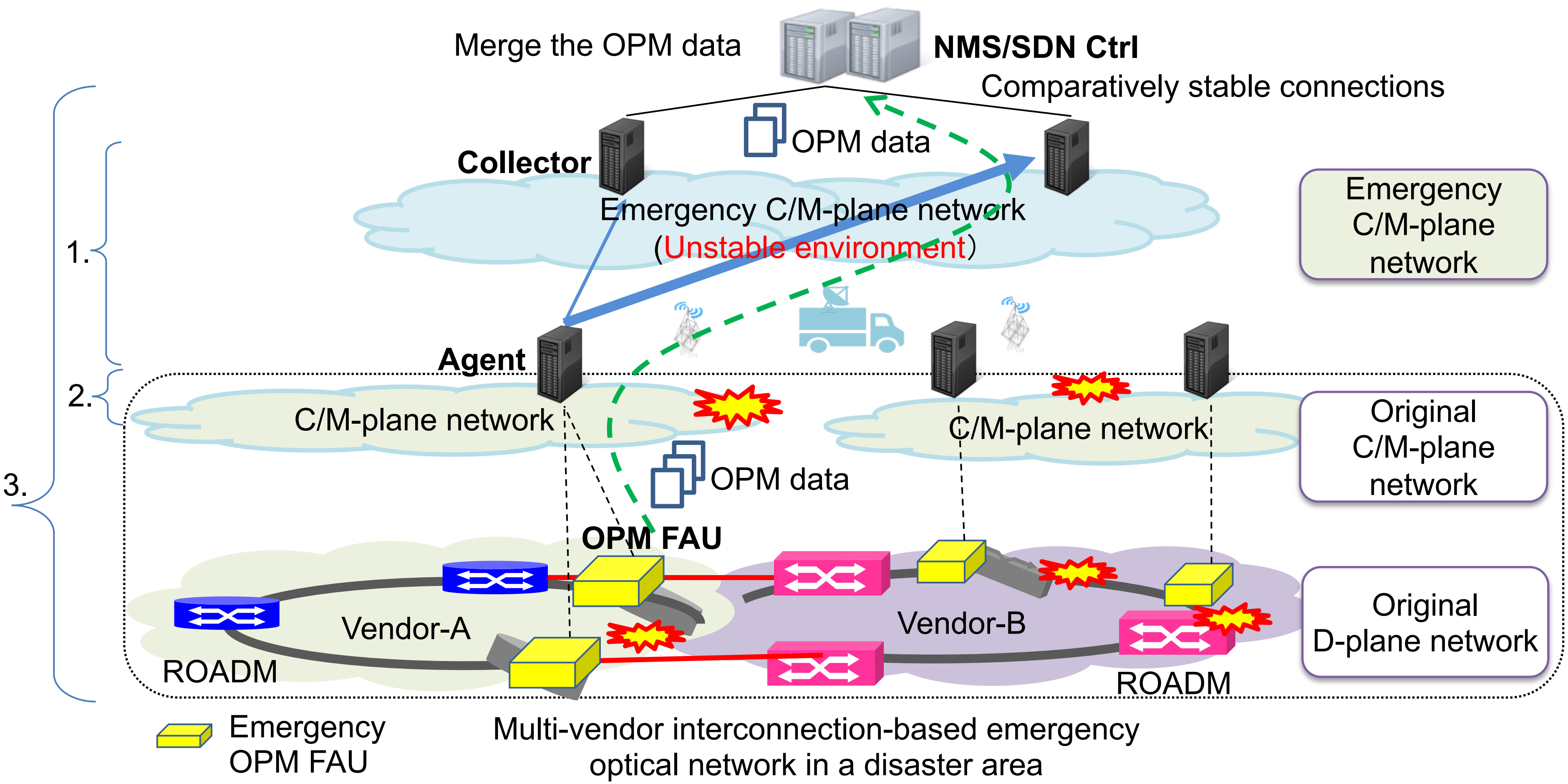
Target in this work

- Recreation of the lost capabilities after disasters as early as possible
  - Optical performance monitoring (OPM)
  - Telemetry

Proposal for OPM Recreation and Open and Robust Telemetry

Main Issues and Proposal (including the following three new capabilities)

1. Robust Telemetry Protocol between agents and collectors
2. OPM Data Analysis and Triage in Agent
3. Multi-vendor OPM FAUs integration via OpenConfig

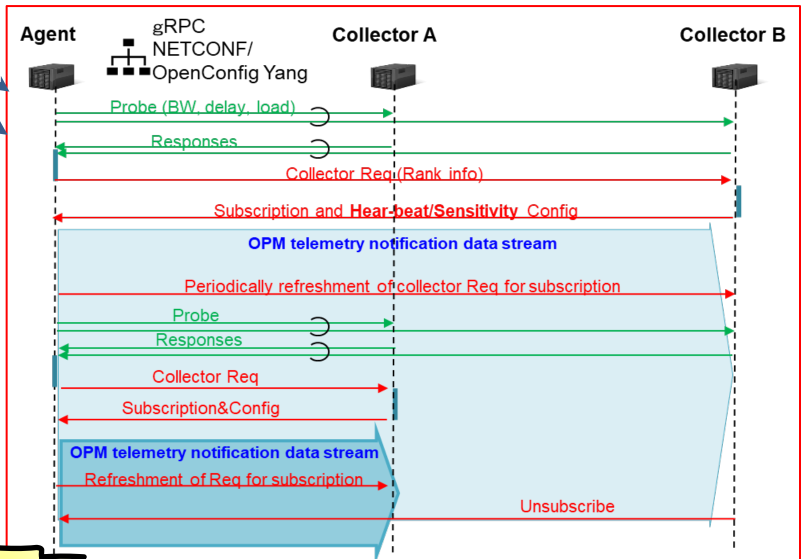


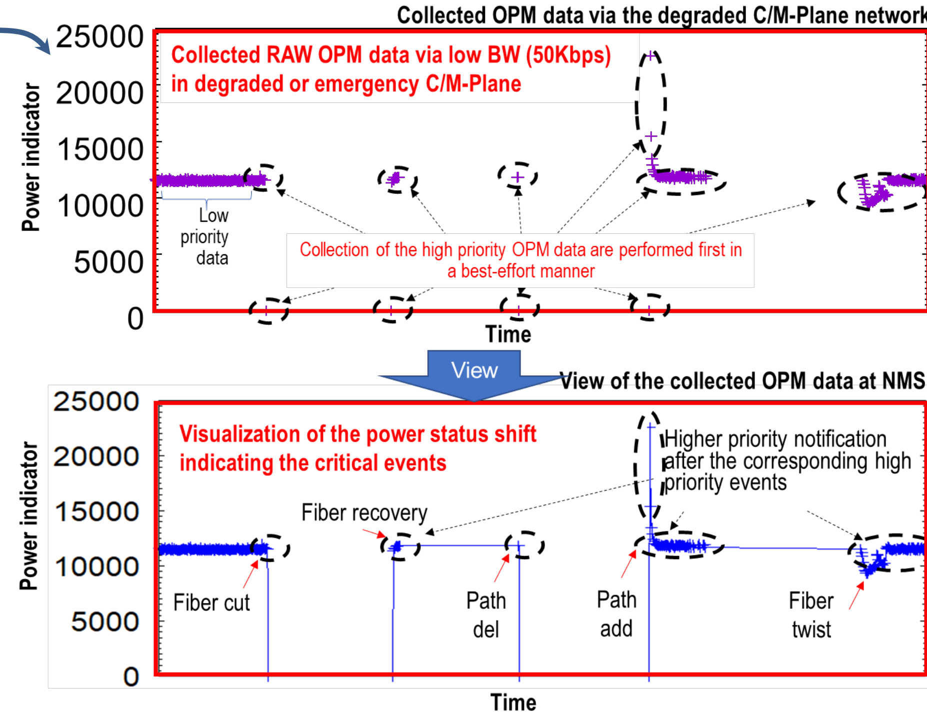
Multi-vendor interconnection-based emergency optical network in a disaster area

Results and Future Works

Main Issues and Proposal (including the following three new capabilities)

1. Robust Telemetry Protocol between agents and collectors
2. OPM Data Analysis and Triage in Agent
3. Multi-vendor OPM FAUs integration via OpenConfig





This study is presented in ECOC2019 as HIGHLY SCORED paper.

S. Xu et al., "**Quick OPM recreation and robust telemetry in emergency optical networks for early disaster recovery**", ECOC, M.1.E.2, Sep. 2019.

Next Step:

- R&D of the Robust-telemetry in Open OPM telemetry systems introducing the active OPM telemetry capability to quickly grasp the status of network in disaster recovery
- Modeling study for the optimal placement of the emergency FAU resources.

Recapping

- 2 in-person visits
- 6 publications (3 journals, 3 conferences)
  - S. Savas, M. Tornatore, F. Dikbiyik, A. Yayimli, C. Martel, and Biswanath Mukherjee, "RASCAR: Recovery-Aware Switch-Controller Assignment and Routing in SDN," in IEEE Transactions on Systems and Network Management, vol. 15, no. 4, pp. 1222-1234, Dec. 2018.
  - R. Lourenco, S. Savas, M. Tornatore, B. Mukherjee, "Robust Hierarchical Control Plane for Transport Software-Defined Networks", in Optical Switching and Networking, Vol. 30, pp. 10-22, Nov. 2018
  - F. Musumeci, C. Rottondi, G. Corani, S. Shahkarami, F. Cugini, and M. Tornatore, “A Tutorial on Machine Learning for Failure Management in Optical Networks,” in IEEE/OSA Journal of Lighwave Technology, vol. 37, no. 16, pp. 4125-4139, 15 Aug.15, 2019
  - S. Ferdousi, M. Tornatore, S. Xu, Y. Awaji, and B. Mukherjee, "Slice-Aware Service Restoration with Recovery Trucks for Optical Metro-Access Networks," accepted to IEEE GlobeCom conference 2019
  - S. Xu, Y. Hirota, M. Shiraiwa, M. Tornatore, Y. Awaji, N. Wada, B. Mukherjee, "Quick OPM Recreation And Robust Telemetry In Emergency Optical Networks For Early Disaster Recovery", accepted to European Conference on Optical Communication (ECOC) 2019
    - Highly scored, invited for a full extension in JLT
  - G. Le, A Marotta, S. Xu, Y. Hirota, Y. Awaji, M. Tornatore, B. Mukherjee, “Logical Network Mapping With Content Connectivity Against Multiple Link Failures in Optical Metro Networks”, accepted for publication, IEEE ANTS 2019
- + 3 in preparation
- 1 keynote (DRCN 2019) / 1 tutorial (OECC) /1 survey (JLT)

Thanks for your attention!

Disaster-Resiliency Strategies
for Next-Generation Metro Optical Networks
ATN: 1818972

National Science Foundation
WHERE DISCOVERIES BEGIN

On-going activities

- **Task 2**
  - Published a survey on failure management in optical networks (JLT)
  - Experimenting ML-based failure management using data from Sendai testbed
- **Task 3**
  - Evolution of work on content connectivity (joint paper @ ANTS19)
  - Reliable network slicing (to be submitted to OFC 2020, not covered)
- **Task 4**
  - Alert-based data evacuation algs. (to be submitted to a journal, not covered)
  - Experimental activities on optical multicast trees
- **Task 5**
  - Slice-aware restoration with recovery trucks (joint paper @ GB19)
- **Task 6**
  - OPM recreation and robust telemetry (joint paper @ ECOC19)