A Proposal of On-demand Staging leveraging Job Management System and Software Defined Networking

Susumu Date[1], Takashi Yoshikawa[1,2], Yasuhiro Watashiba[2], Yoshiyuki Kido[1], Shinji Shimojo[1], Masahiko Takahashi[3], Masaki Kan[3], Masaki Muraki[4]

Osaka U [1], NAIST [2], NEC [3], TIS [4]

INTRODUCTION OF CMC

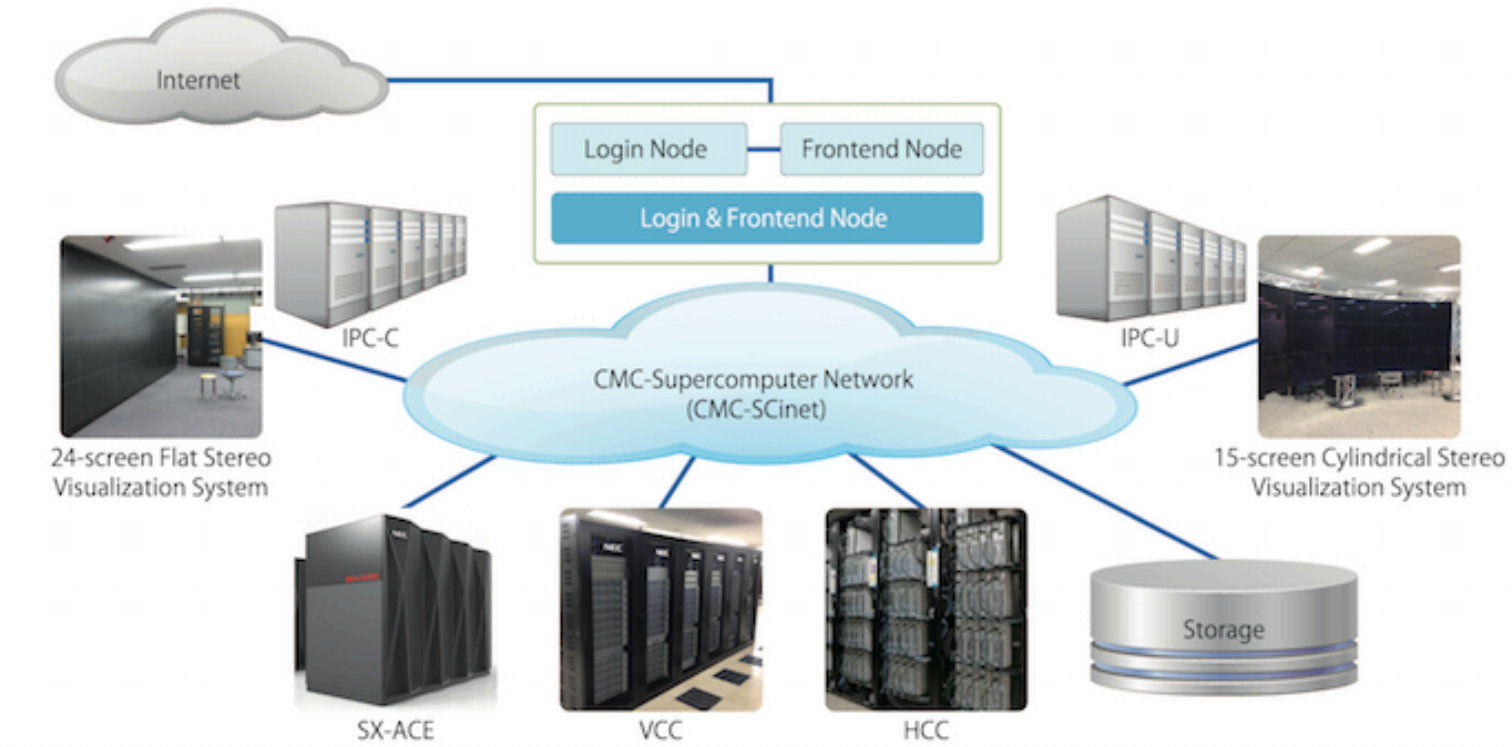
Cybermedia Center

- a computing center at Osaka University
  - *http://www.cmc.osaka-u.ac.jp/*
- has a responsibility of providing a powerful high-performance computing environment for university researchers across Japan as a national joint-use facilities.
- Is in charge of the campus IT infrastructure and for promoting its effective use.

CMC's Suita main building

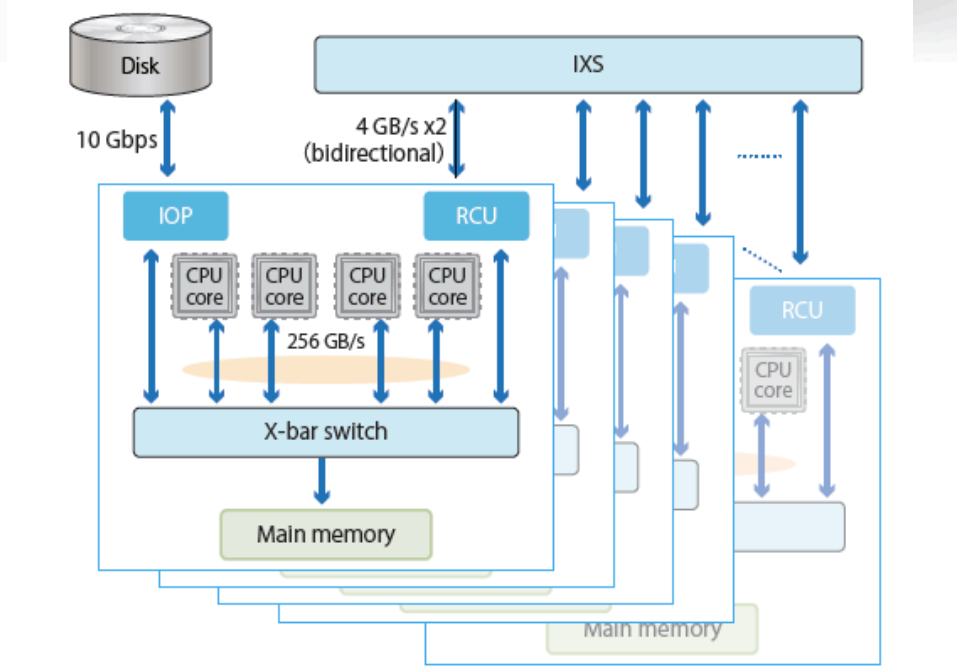
IT Core Annex: new datacenter building

High Performance Computing and Visualization Environment at the CMC



- Large-scale computing system
  - Vector-type: SX-ACE
  - Scalar-type: HCC and VCC
- Large-scale Visualization system

SX-ACE: computer service since Dec. 2014



Type: Vector

OS: Super-UX

# of nodes: 1,536

# of cores: 6,144

Total memory: 96 TB

Peak performance: 423 TFlops

Each node has

- 4-core multi-core vector CPU (256 GFlops)
- 64 GB memory

IXS (Internode crossbar switch) connects 512 nodes and allows 4GB x 2 communication

VCC (PC cluster system for large-scale visualization)

Type: Scalar

OS: Linux

# of nodes: 65

# of cores: 1300

Total memory: 4,160 TB

Peak performance: 26.0 TFlops

Accelerator: NIVIDA Tesla K20 x 59

- A cluster system composed of 62 computing nodes, each of which has 2 Intel Xeon E5-2670v2 and 64GB main memory
- **Dynamically Reconfigurable Cluster System** using ExpEther, system virtualization technology from NEC

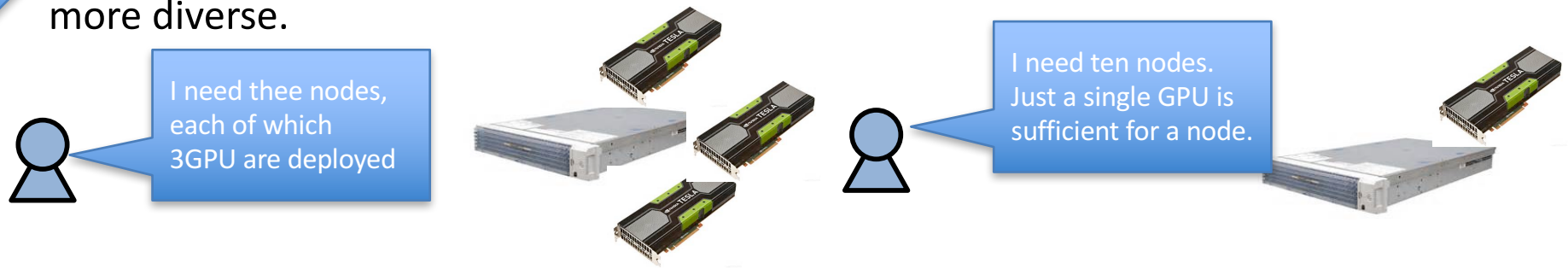
Not limited to visualization purpose!

RESEARCH PART

Research Background(1)

- Necessity and importance of supercomputing is rising for every field of science
  - Development of scientific data measurement devices
  - Proliferation of computation results due to processor's technological advancement
  - Technological advancement in networking area
  - Expectation to big data analysis, deep learning and so forth
- High-performance computing environments have been diversified.
  - The number of processor cores: 6, 8, 12, 16…
  - Many options of accelerators: GPU, FPGA, Phi

The combination of these choices have been making users' computing environment more diverse.



Supercomputing centers have a resource flexibility to accommodate a diversity and heterogeneity of requests pertaining to computing resources.

Research Background(1)

- Necessity and importance of supercomputing is rising for every field of science
  - Development of scientific data measurement devices
  - Proliferation of computation results due to processor's technological advancement
  - Technological advancement in networking area
  - Expectation to big data analysis, deep learning and so forth
- High-performance computing environments have been diversified.
  - The number of processor cores: 6, 8, 12, 16…
  - Many options of accelerators: GPU, FPGA, Phi

The combination of these choices have been making users' computing environment more diverse.

I need ten nodes.

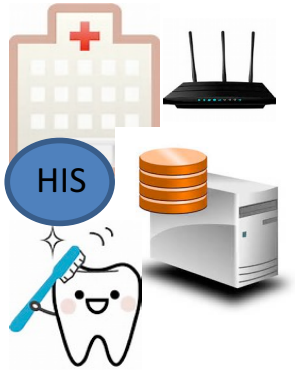
How can we satisfy these diverse computational requirements and requests from users on finite computing resources (budget)?

Supercomputing centers have a resource flexibility to accommodate a diversity and heterogeneity of requests pertaining to computing resources.

Research Background(2)

- Large amount of privacy-rich and confidential data which scientists and researchers want to analyze with supercomputer systems exists even in our campus.

Osaka University Dental Hospital Case



We would like to use super computer systems for deep learning towards severity assessment of patients' cases, learning analysis. … But data security prevents …although our university has supercomputer systems..

Security policy..
At least we would like to minimize the exposure of our privacy-rich and confidential data to other people even in the campus.

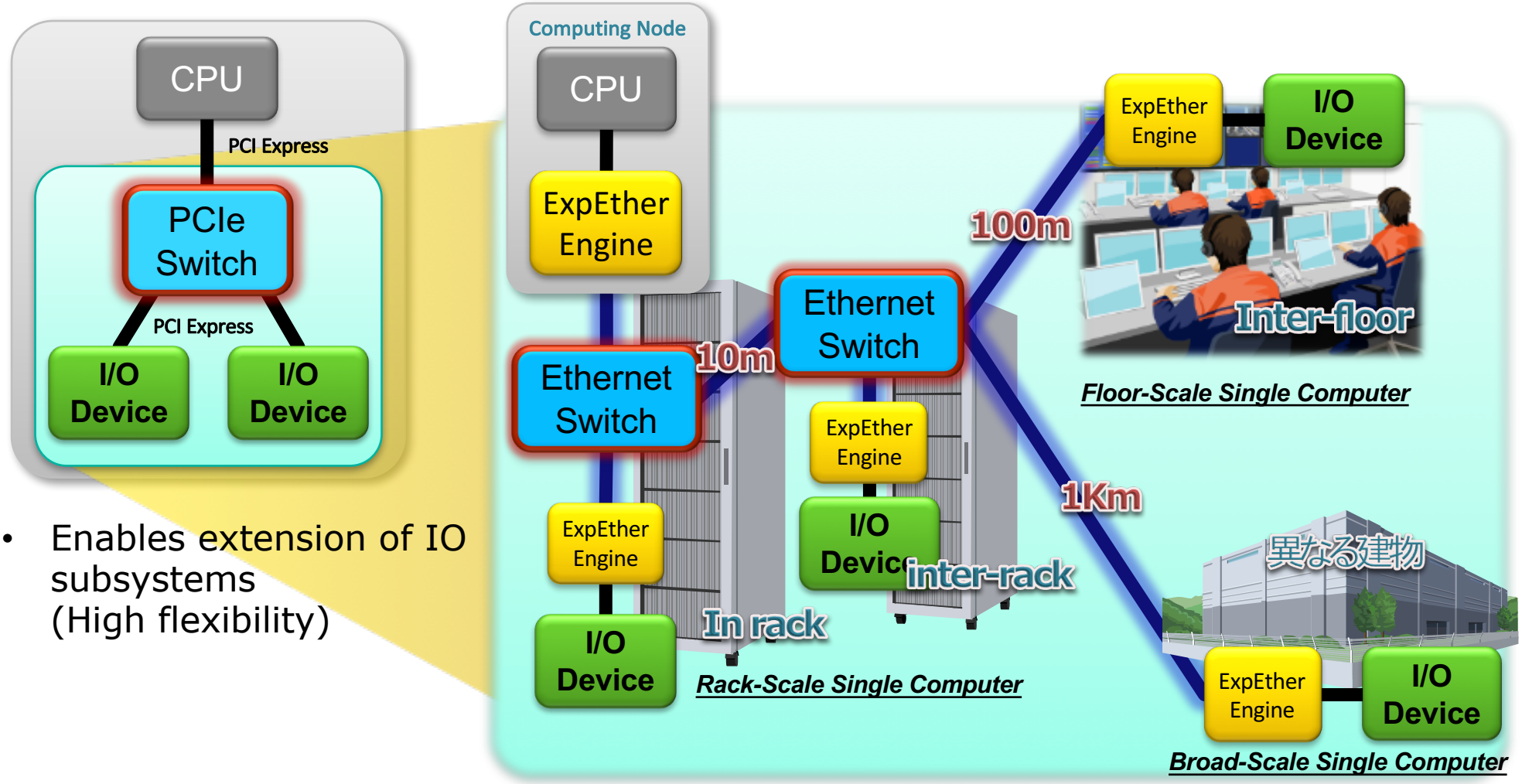
Is there any technological solution that allows us to safely transfer data from/to supercomputer systems?

Is there any technological solution that allows us to place our data only during our computation?

A Solution for heterogeneity and diversity of computing

- Flexible management of our scalar-typed computer system with ExpEther-based Resource pool in response to users' requests

- The world's first technology that extends PCI Express onto Ethernet (L2 layer)
  - Enables "spacious single computer" using networks"



- Enables extension of IO subsystems (High flexibility)

A Solution for heterogeneity and diversity of computing

- Flexible management and administration of our scalar-typed computer system in response to users' requests

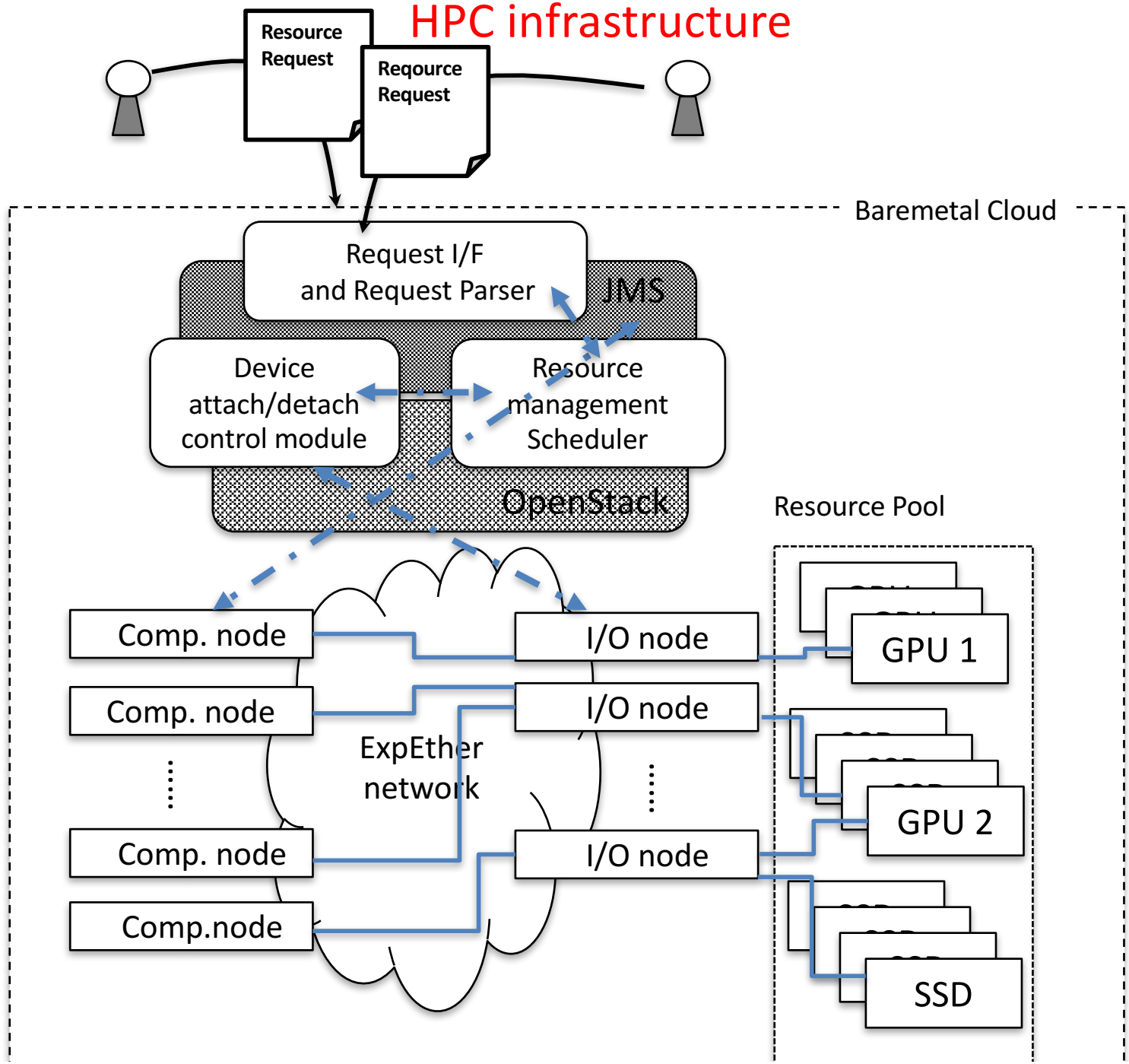
PROBLEM

Administrator has to manually set up user-desired configuration using ExpEther server based on user requests.

IDEA

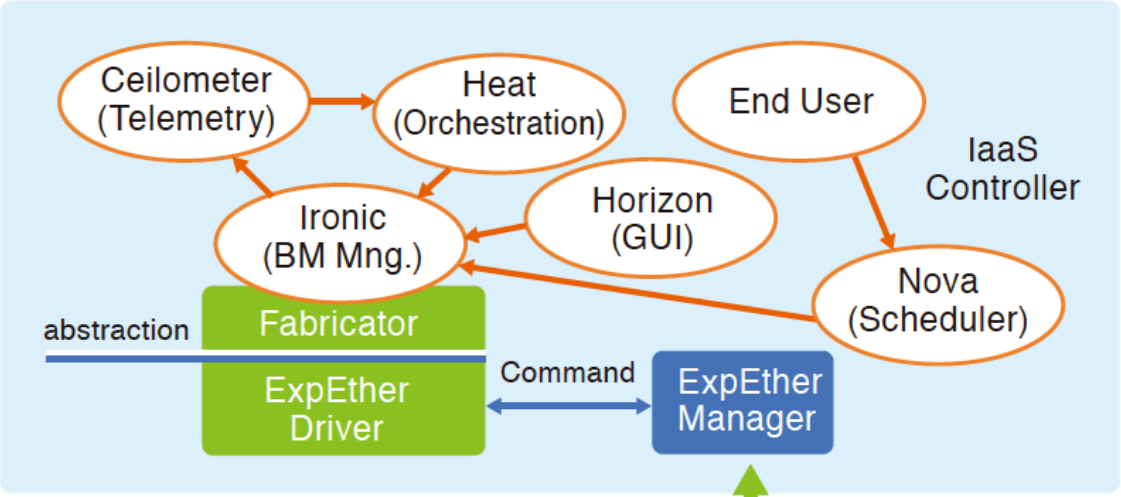
Can we directly conveys user's requests to computing resources to ExpEther?

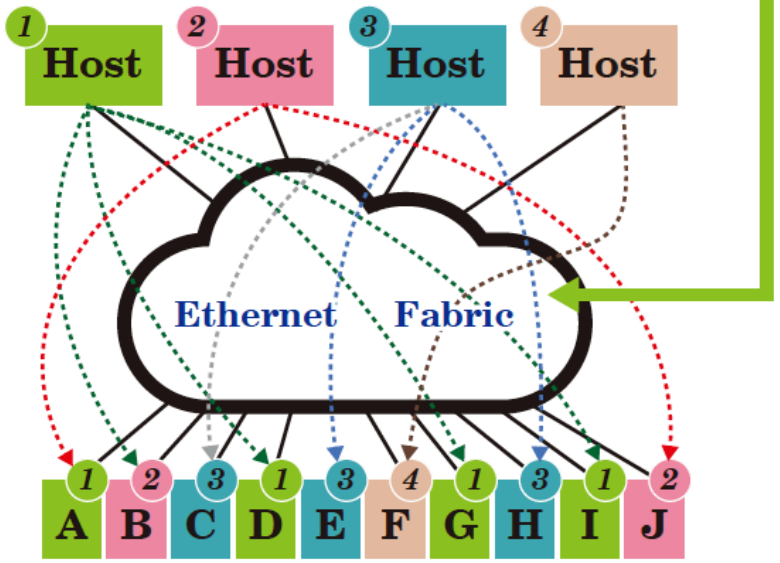
Basic idea: dynamic reconfigurable HPC infrastructure



Prototype implementation
SC2016 demonstration in collaboration with NEC laboratory

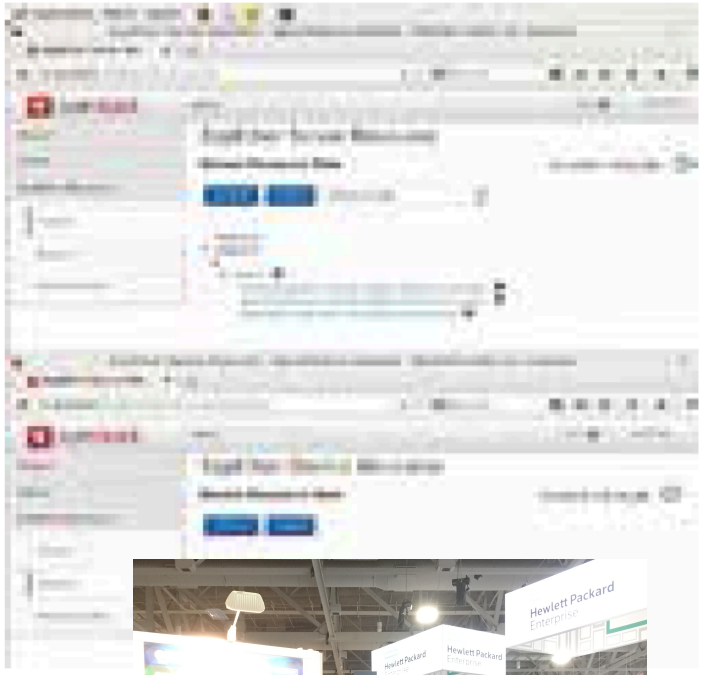
Modify Ironic to control each device





Key technologies
- Openstack(Ironic)
- Exp-Ether

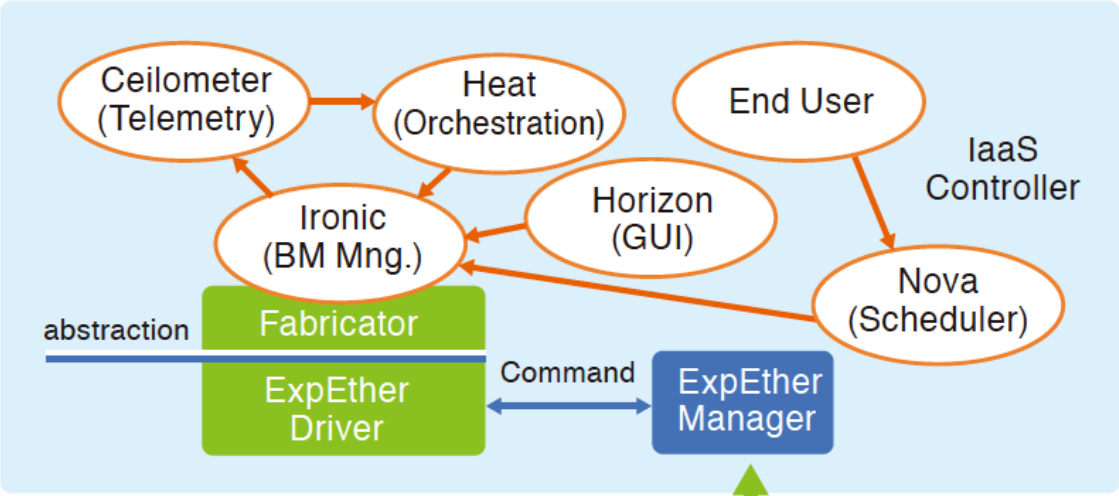
GUI on modified Horizon



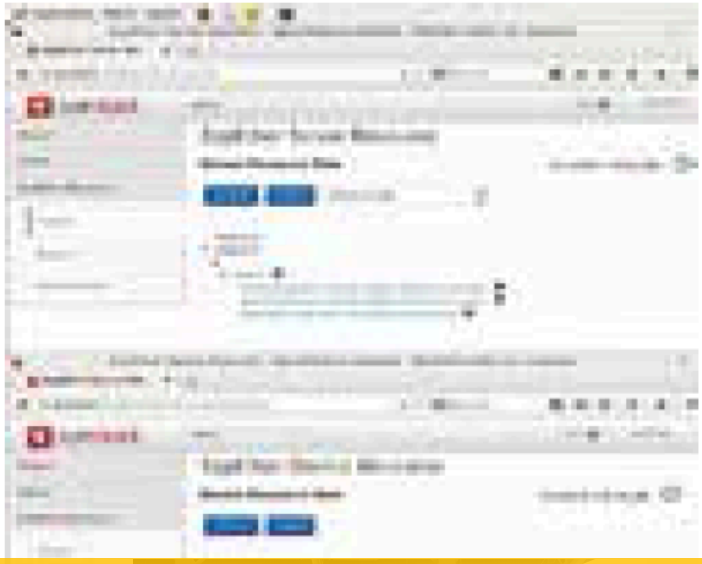
Guest Prof. Takashi Yoshikawa (Osaka U/NEC)

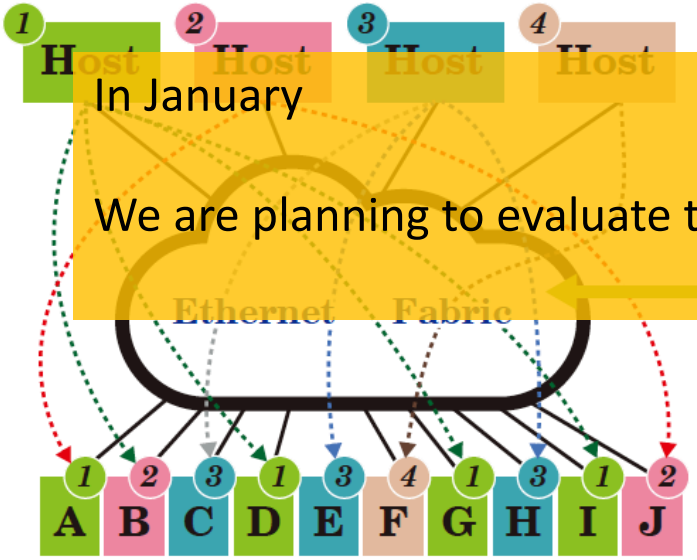
Prototype implementation
SC2016 demonstration in collaboration with NEC laboratory

Modify Ironic to control each device



GUI on modified Horizon

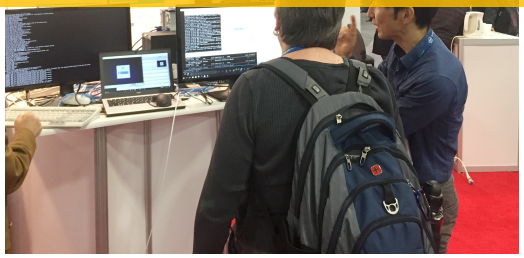




In January

We are planning to evaluate the prototype system on VCC.

- Openstack(Ironic)
- Exp-Ether

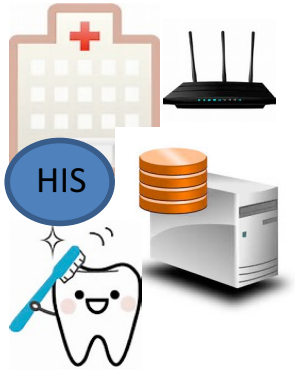


Guest Prof. Takashi Yoshikawa (Osaka U/NEC)

Research Background(2)

- Large amount of privacy-rich and confidential data which scientists and researchers want to analyze with supercomputer systems exists even in our campus.

Osaka University Dental Hospital Case



We would like to use super computer systems for deep learning towards severity assessment of patients' cases, learning analysis. … But data security prevents …although our university has supercomputer systems..

Security policy..
At least we would like to minimize the exposure of our privacy-rich and confidential data to other people even in the campus.

Is there any technological solution that allows us to safely transfer data from/to supercomputer systems?

Is there any technological solution that allows us to place our data only during our computation?

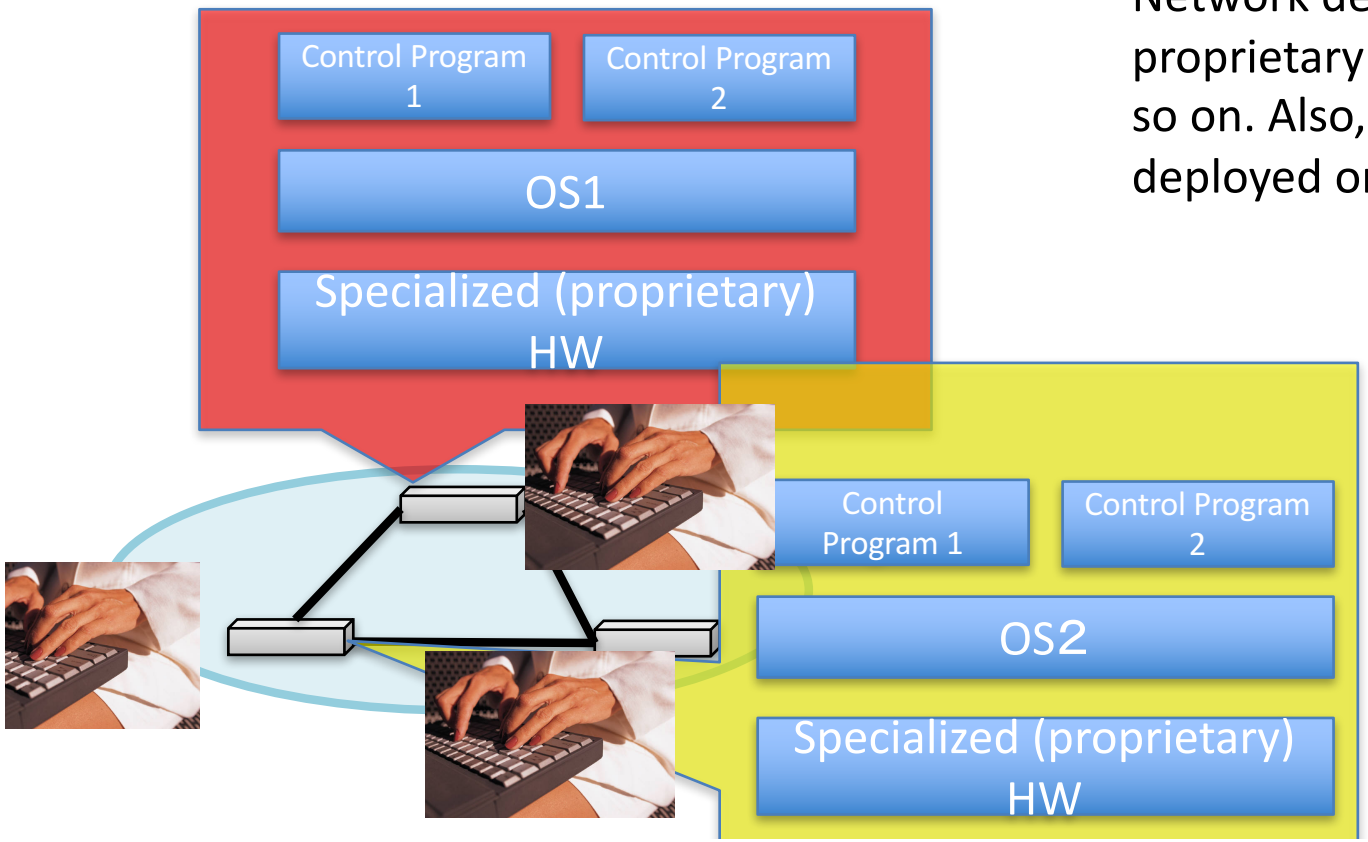
A solution : smart campus for supercomputing (planning)

- Leveraging the dynamic network programmability brought by Software Defined Networking for staging-in/out data from/to our supercomputer systems.

SDN: A new concept of network architecture

Network functions are divided to “control plane (control)” and “data plane (packet forwarding)”.

Traditional networking



Network devices are developed with proprietary OS, specialized hardware and so on. Also, control program for routing is deployed on them.

- Administrators need to configure network devices through the use of different suites of APIs.

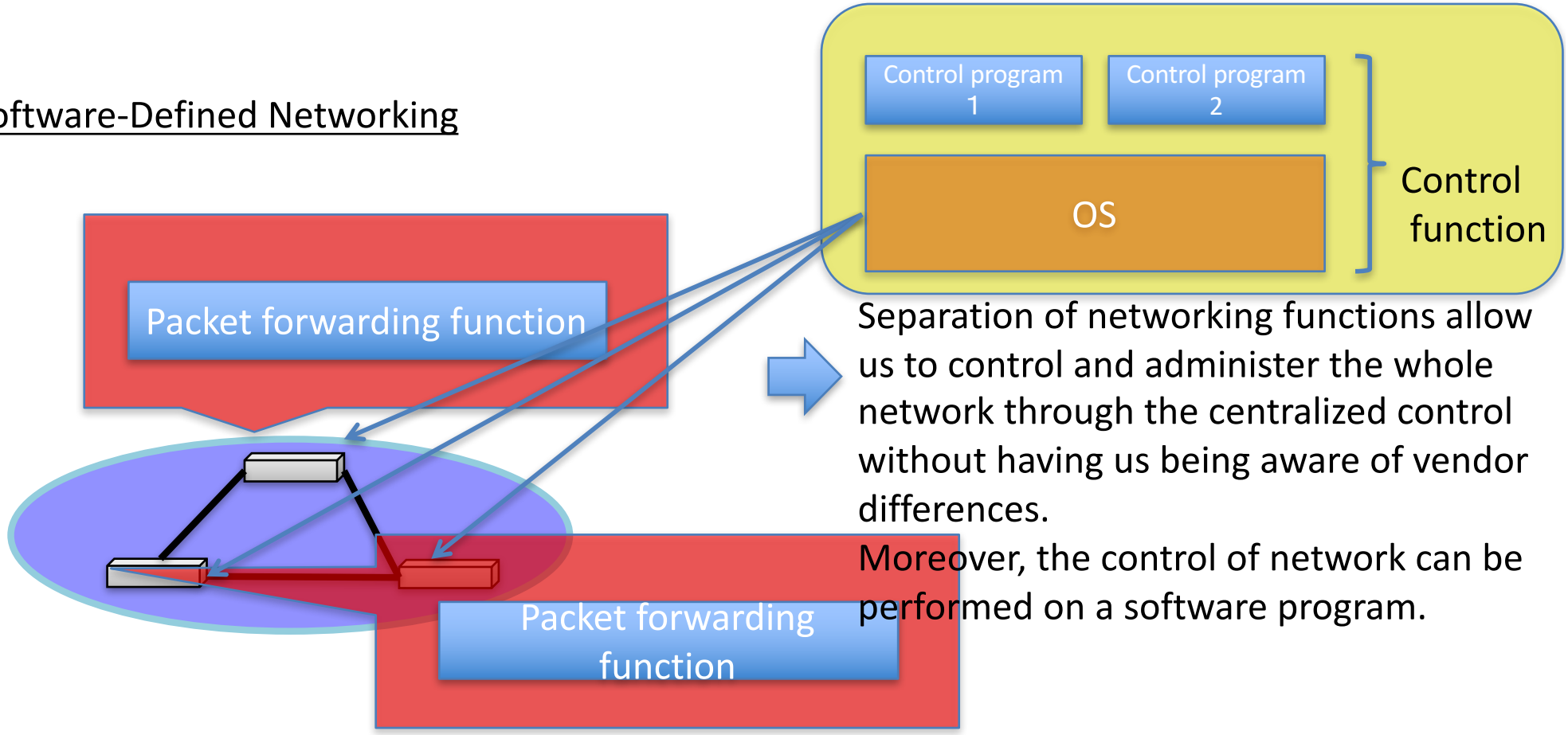
A solution : smart campus for supercomputing (planning)

- Leveraging the dynamic network programmability brought by Software Defined Networking for staging-in/out data from/to our supercomputer systems.

SDN: A new concept of network architecture

Network functions are divided to “control plane (control)” and “data plane (packet forwarding)”.

Software-Defined Networking

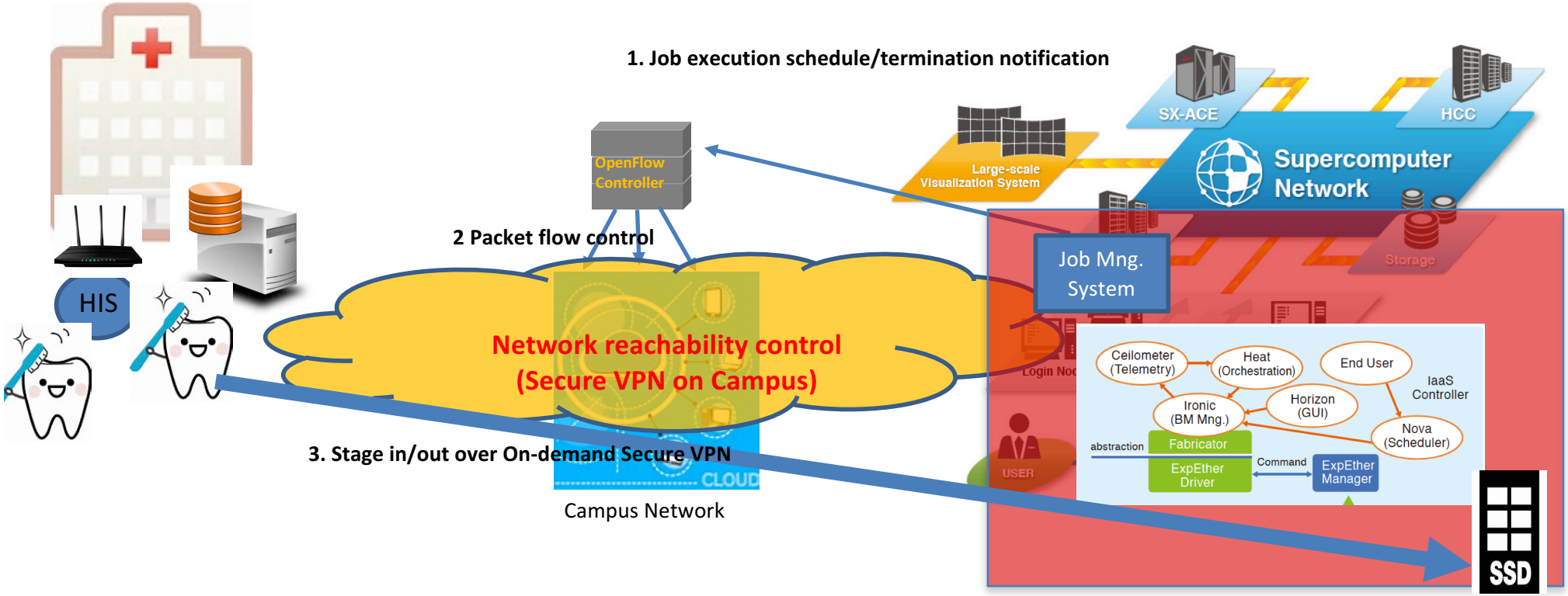


Separation of networking functions allow us to control and administer the whole network through the centralized control without having us being aware of vendor differences.

Moreover, the control of network can be performed on a software program.

A solution : smart campus for supercomputing (planning)

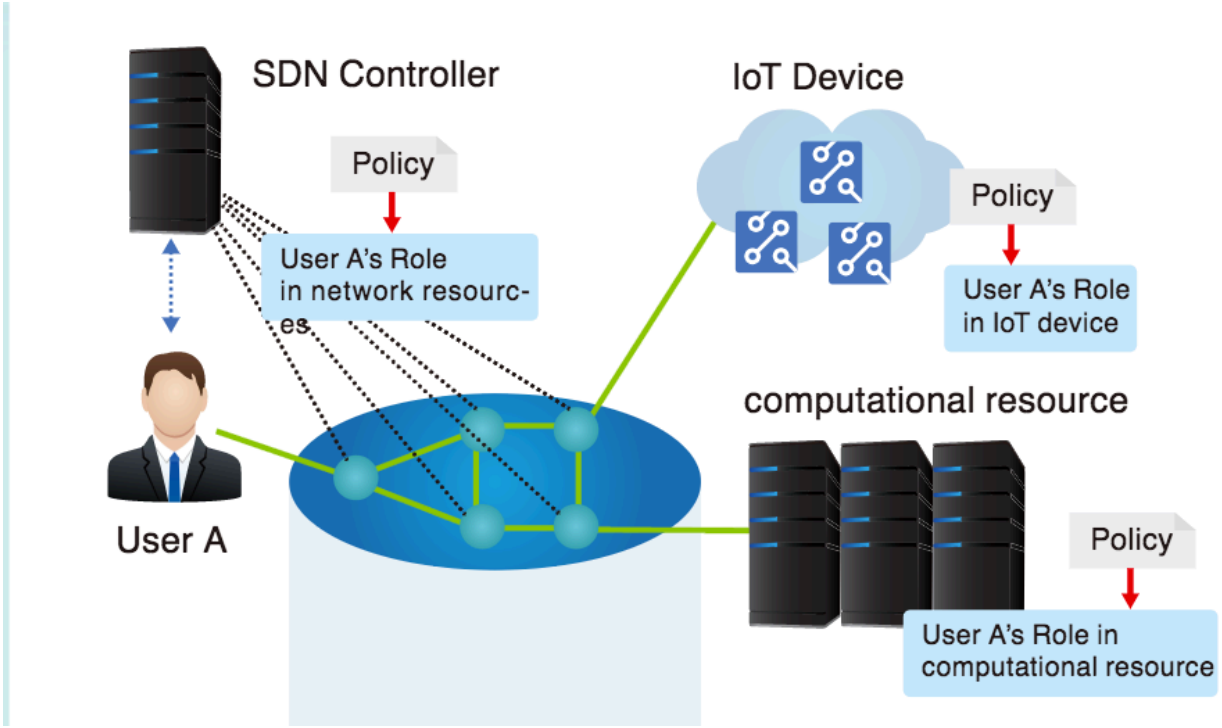
- Leveraging the dynamic network programmability brought by Software Defined Networking for staging-in/out data from/to our supercomputer systems.



- Local SSD is dynamically attached/detached through ExpEther based on user requirements.
- By taking advantage of network programmability, network reachability is assured only during data movement before/after computation.

Current Progress

- FlowSieve: RBAC-based access control to network resources
  - When a user attempt to access network, authentication is conducted.
  - Each of network resources have a security policy and assigns a role to the authenticated user based on its security policy (usually by the owner of network resources)
  - Based on the assigned role to the user, packet flows are restricted and controlled.



Takuya Yamada, Keichi Takahashi, Masaya Muraki, Susumu Date, Shinji Shimojo, *"Network Access Control Towards Fully-controlled Cloud Infrastructure",* CloudCom2016, Ph.D Consortium, Dec. 2016.

Summary

- A reconfigurable HPC infrastructure is presented against the recent heterogeneity and diversity of computing requests.
- Also, we showed an idea (proposal) of on-demand staging leveraging job management system and software defined networking.
- Currently, we have been working on flexible supercomputing solutions through the collaboration with NEC and TIS.