PRAGMA37
Resources and Cyber Learning Updates

Hsiu-Mei Chou

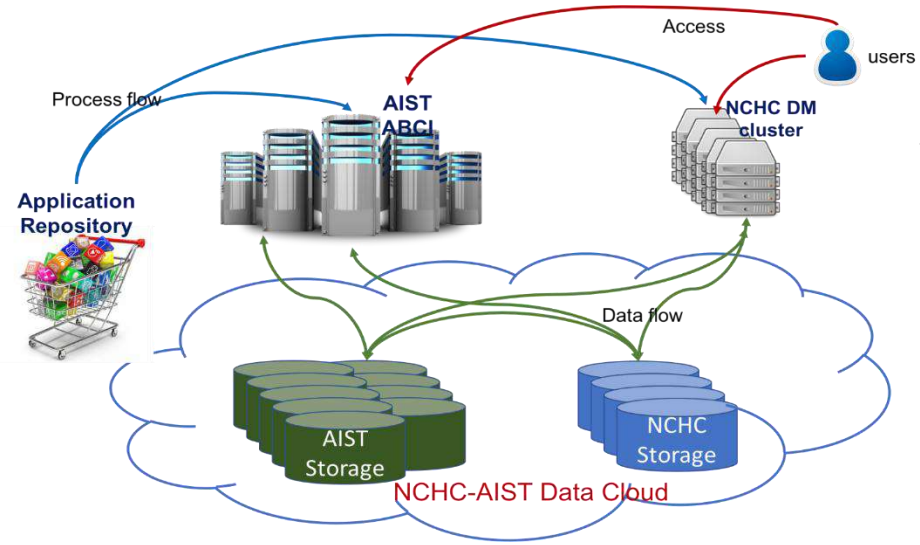
09/12/2019

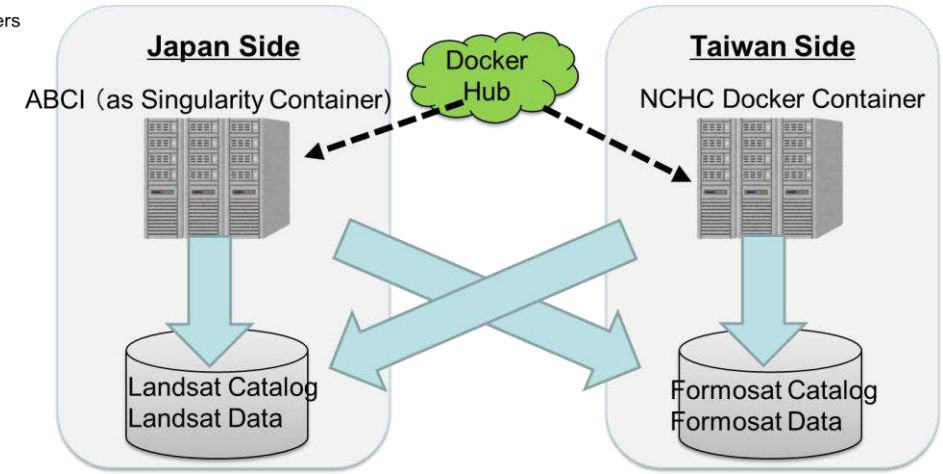
Focus Groups

- AI Focus Group
- DTN and Data Movement Focus Group
- Edge Computing Focus Group
- EDISON (Cyber Learning) Focus Group
- Distributable Lifemapper

AI Focus Group

- PRAGMA AI Platform:
  - Share knowledge (AI module) by container technologies.
  - Computing resources and data resources are also shared.
- Capabilities and status
  - Share trained model as well as non-trained model.
  - Access the data at each side.
  - POC and demonstration by AIST, NCHC and NSPO.
  - Expect to extend to the other collaborators.





Progress

- Had two meetings, one in Tsukuba, one in Taipei.
- Summarized the components of AI platform
  - Systems: ABCI, Taiwania2
  - Data: start by remote sensing data
  - Catalogue services: both AIST and NCHC are developing. planning cross reference.
  - Method to share AI modules: use container
  - Driving apps: solar panel detection, change detection, etc.
  - Visualization/HCI: integrate with SAGE2
  - Security: under discussion
- Give a demo this afternoon!

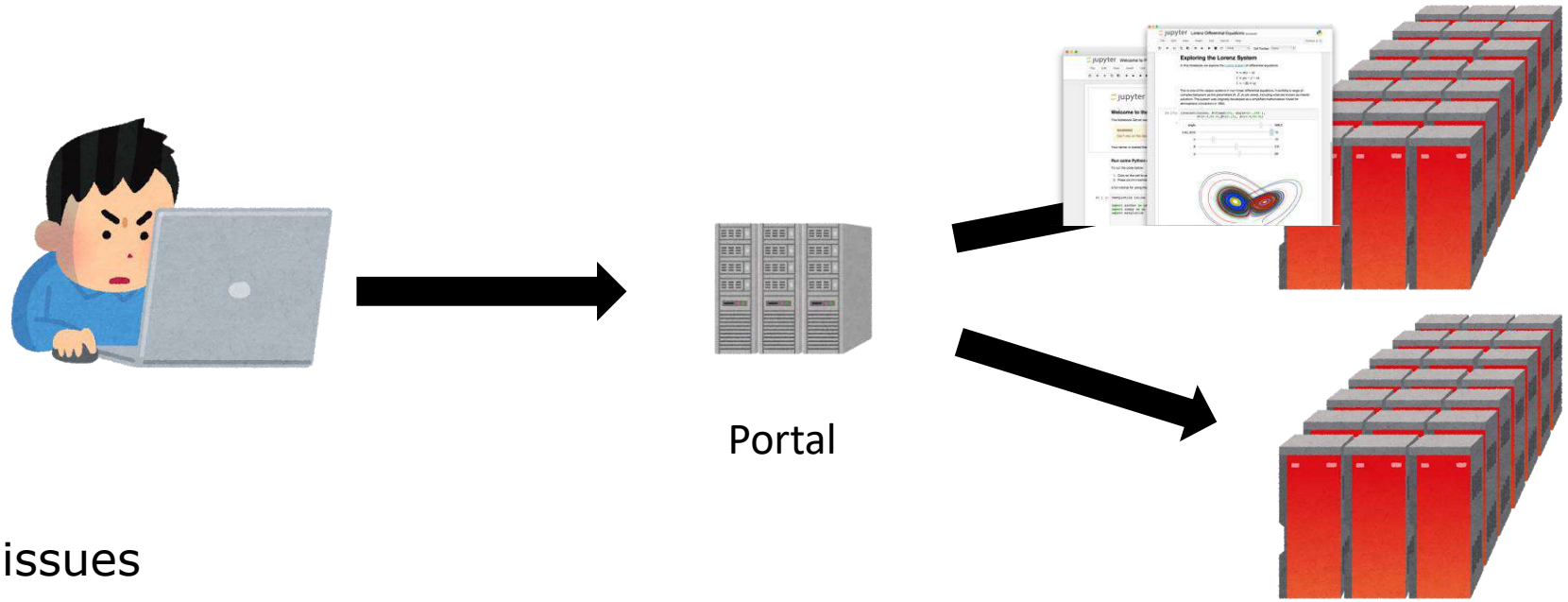
Demo overview:
Japan-Taiwan Data and AI Module Platform for Analyzing Remote Sensing data, Part 3

Hidemoto Nakada, Ryosuke Nakamura, Kyoung-Sook Kim, Jason Haga,
Yusuke Tanimura, Ryousei Takano, Yoshio Tanaka (**AIST**)
Hsiu-Mei Chou, Hsi-En Yu, Chun Hung Huang, Weicheng Huang (**NCHC**)
Bo Chen, Scarlet Peng (**NSPO**)

The Goal:

Provide easy to use interactive environment via poral : Jupyter notebook

- where the users explore data and execute AI models on the data.

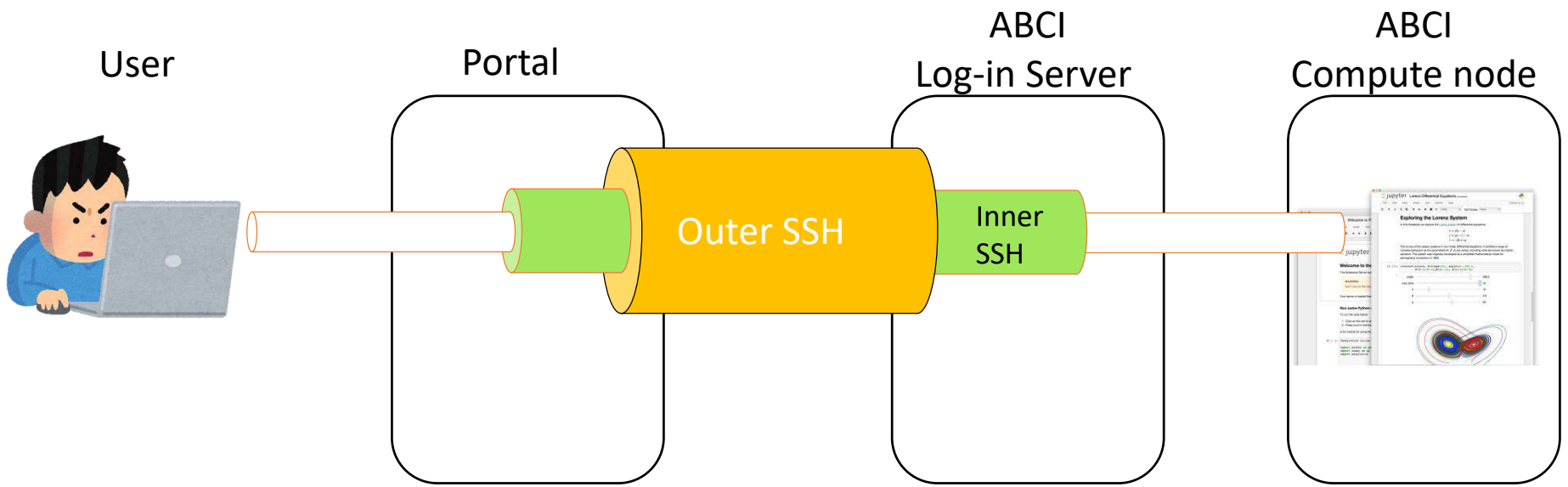


The issues

- How to allow access to the computers without installing the users private key on the portal
- How to provide connection to the computing nodes behind the firewall

The solution

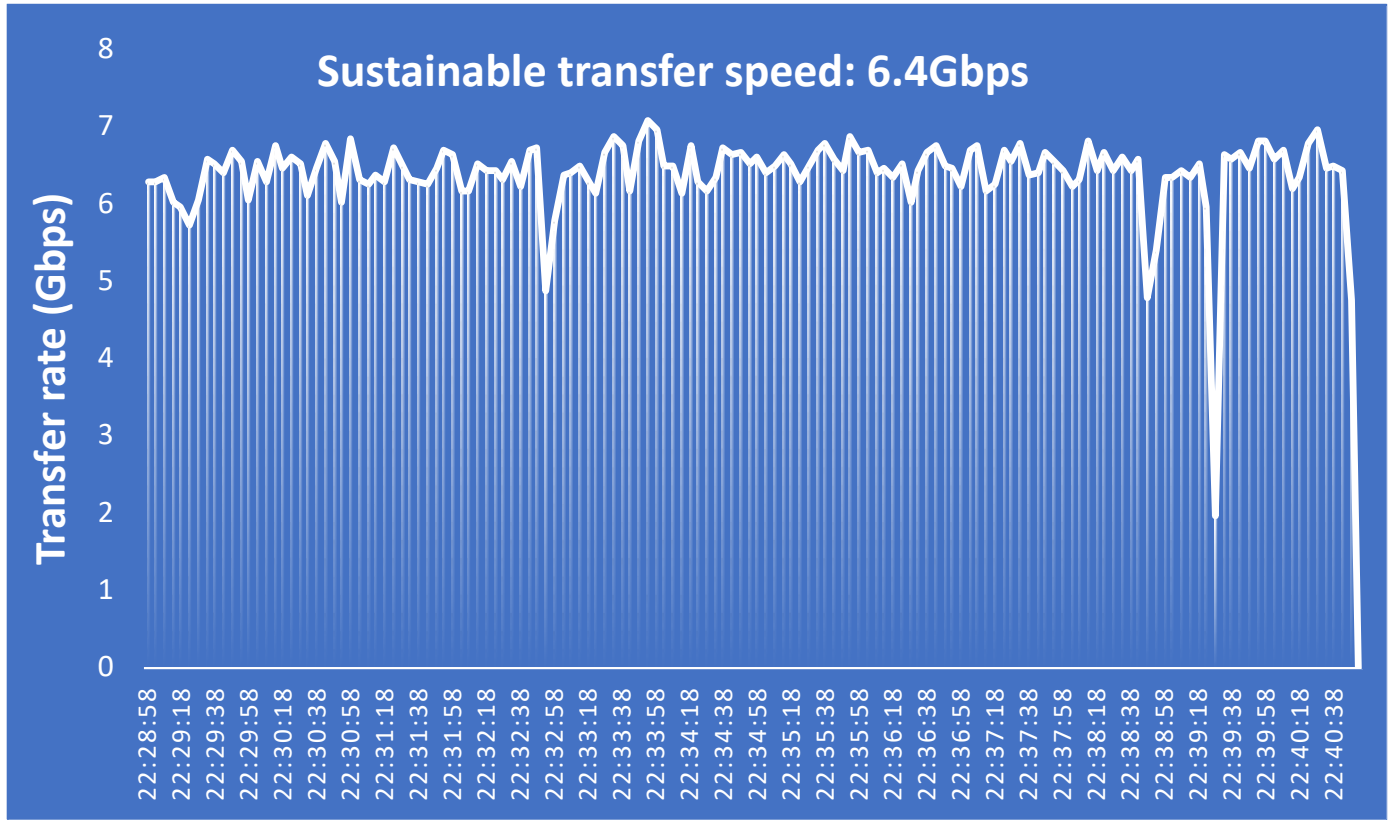
- Nested SSH tunneling to allow access to the supercomputer without installing the user's private key on the portal
- On-the-fly Port forwarding to allow connection to the compute node



DTN and Data Movement Focus Group

- Goal: We explore solutions to transfer big data (securely) among PRAGMA Clouds and share best practices.
- Possible solutions:
  - DTN (Data Transfer Node): i.e., FIONA@PRP
  - S3-based cloud storage
- Progress:
  - Set up DTN nodes on UCSD and AIST
  - Transfer zebrafish brain image data (0.5TB)

Disk-to-Disk data transfer experiment from UCSD to ABCI

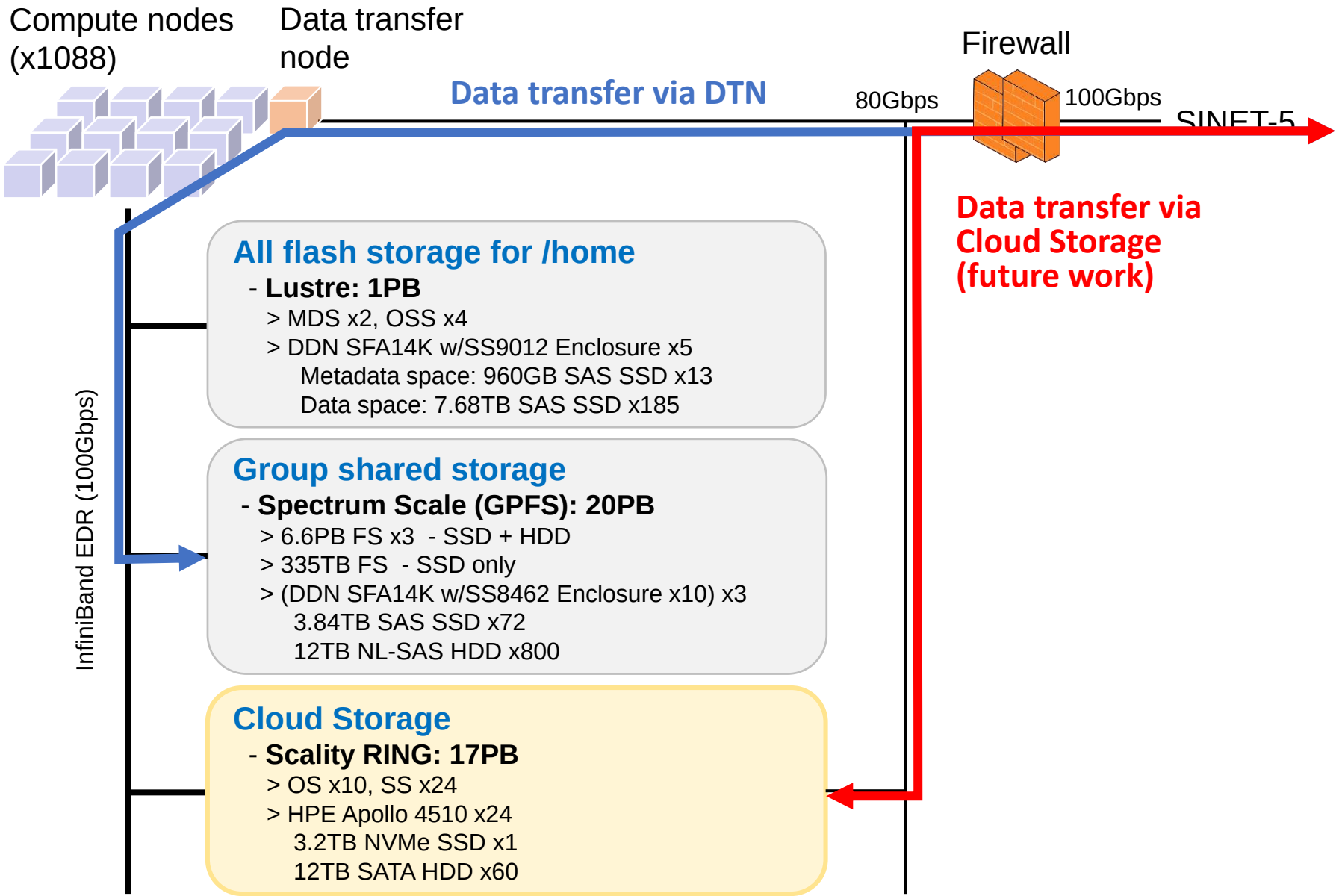


- Total data size: 574GB
- File size: 193MB
- Number of files: 2977
- Elapsed time: 12m10s

- Tool: FDT version 0.26.1

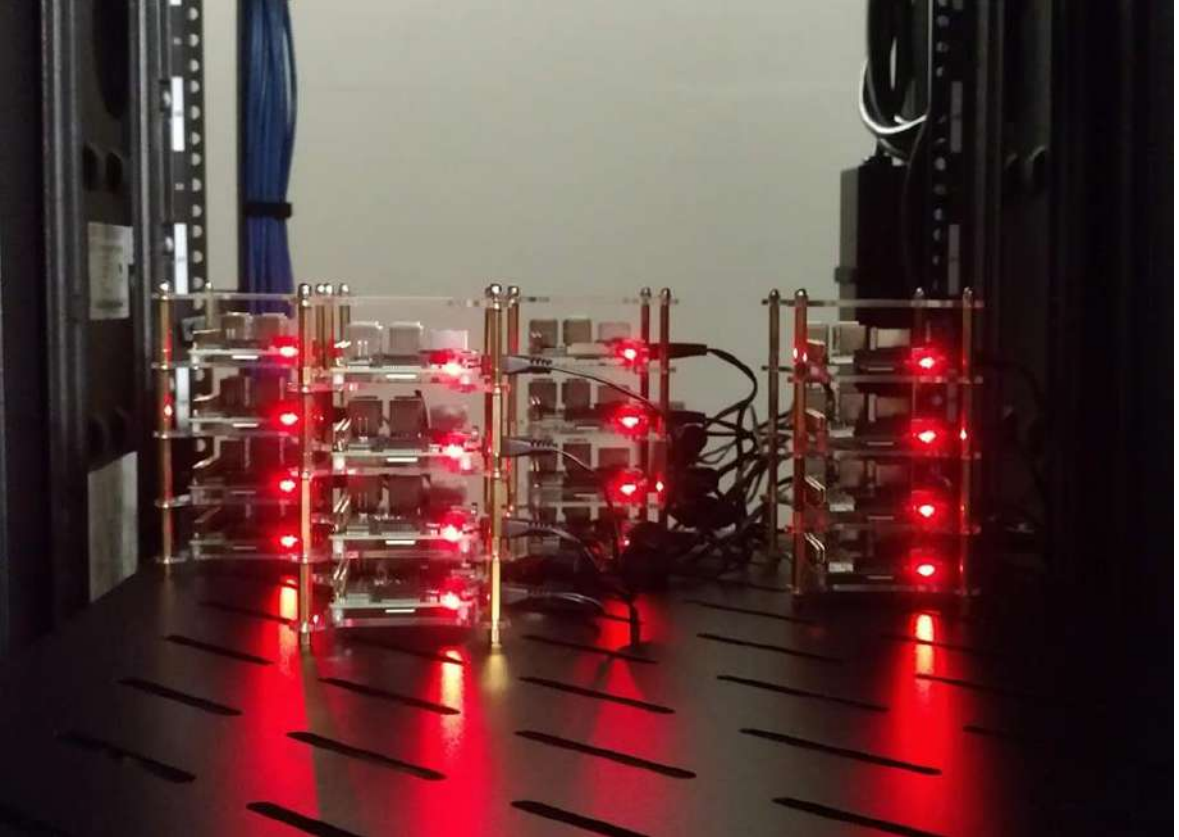
*) Zebrafish brain image raw data

ABCI Storage System



Distributed testbed for Edge computing

- Collaborators:
  - Renato Figueiredo, Vahid Daneshmand, Ken Subratie (UF)
  - Kohei Ichikawa, Keichi Takahashi, Kundjanasith Thonglek (NAIST)
- Deployed resources that can be used for edge computing research
- Initial setup:
  - Raspberry Pi devices at two institutions (UF, NAIST)
  - Local VMs for processing (in progress)
- Looking for more institutions to join our distributed cluster

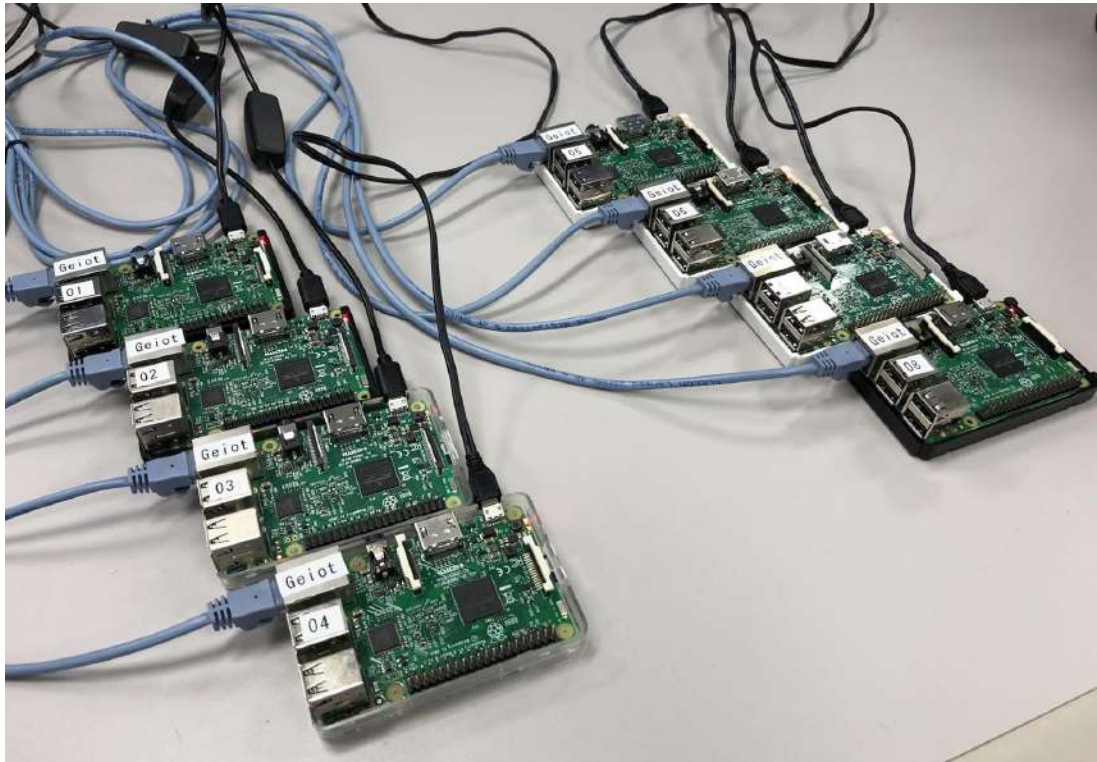


UF UNIVERSITY of FLORIDA

Institute of Science and Technology NAIST

UF - NAIST

Raspberry Pi Cluster



Cluster Specifications

- Cluster Size:
  - UF: 16 Raspberry Pi 3B+ Nodes
  - NAIST: 8 Raspberry Pi 3B+ Nodes
- Virtualization and orchestration:
  - Docker, Kubernetes
- Management + Master Node: VM @ UF
- Worker Nodes Specs:
  - SoC: Broadcom BCM2837B0 quad-core A53 (ARMv8) 64-bit @ 1.4GHz
  - GPU: Broadcom Videocore-IV
  - RAM: 1GB LPDDR2 SDRAM
  - Networking: Gigabit Ethernet (via USB channel), 2.4GHz and 5GHz 802.11b/g/n/ac Wi-Fi
  - OS: Raspbian
  - Overlay Network: IPOP VPN
  - Alternate Network Connectivity: autossh + Reverse SSH Tunneling

EDISON Focus Group

- Deploy and evaluate full EDISON deployment on PRAGMA Cloud
- Deploy EDISON on HKU infrastructure
  - introducing to HKU researchers to collect feedbacks for EDISON-AI
  - individual DEMOs for HKU researchers @PRAGMA37
- Federation authentication requirement for HKU researchers
  - not yet supported on the current version of EDISON-AI
- Resources
  - detailed discussion required @PRAGMA37
- KISTI/NCHC MoU
  - postponed

EDISON Focus Group

- Integration K8 into EDISON
- using PRP (K8 based) resources
  - internal group meetings in KISTI (Jul-Aug/2019, 2 times)
  1. integration between the EDISON scheduler and PRP container orchestration tools
  2. need to develop job wrapping libraries, i.e., wrapping interactive codes (Jupyter) to a binary for AI training

Distributable Lifemapper

- Kansas testing environment upgrade to Rocks 7
  - Success with Rocks 7 on virtual cluster host
  - Some problems with CentOS updates
  - Limited success with Rocks 7 on virtual clusters
- Lifemapper Server and Lifemapper Compute roll updates
  - Both have been successfully built and partially tested
  - Upgrade required updates to approximately 75 dependencies, some pre-built, some built from source
  - Rolls have not been fully tested because the testing environment is unstable