Authentication and Authorization Architecture in HPCI

38th EUGridPMA & IGTF All Hands meeting
September 19 2016
Geneva, Switzerland

Eisaku Sakane
sakane@nii.ac.jp
National Institute of Informatics
Japan

National Institute of Informatics

Table of Contents

- Overview of HPCI
- Authentication and Authorization Infrastructure in HPCI
- Topics

Overview

HPCI

- High Performance Computing Infrastructure
  - national project promoted by Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan
  - distributed computing infrastructure for high performance computing
    - “K computer”, supercomputers, high performance storage and high-speed network
  - Single Sign-On (Shibboleth and GSI), distributed file system (Gfarm), one-stop user support service (registration, helpdesk, etc)
  - first production level infrastructure for high performance computing in Japan
  - the production level operation started in Sep. 2012.

HPCI – System Providers

- **RIKEN Advanced Institute for Computational Science**
- **National Institute of Informatics**
- Information Initiative Center, **Hokkaido University**
- Cyberscience Center, **Tohoku University**
- Center for Computational Science, **University of Tsukuba**
- Information Technology Center, **The University of Tokyo**
- Global Scientific Information and Computing Center, **Tokyo Institute of Technology**
- Information Technology Center, **Nagoya University**
- Academic Center for Computing and Media Studies, **Kyoto University**
- Cybermedia Center, **Osaka University**
- Research Institute for Information Technology, **Kyushu University**
- Center for Earth Information Science and Technology, **Japan Agency for Marine-Earth Science and Technology**
- Center for Engineering and Technical Support, **The Institute of Statistical Mathematics**

National Institute of Informatics

Computational Resources (extracts)

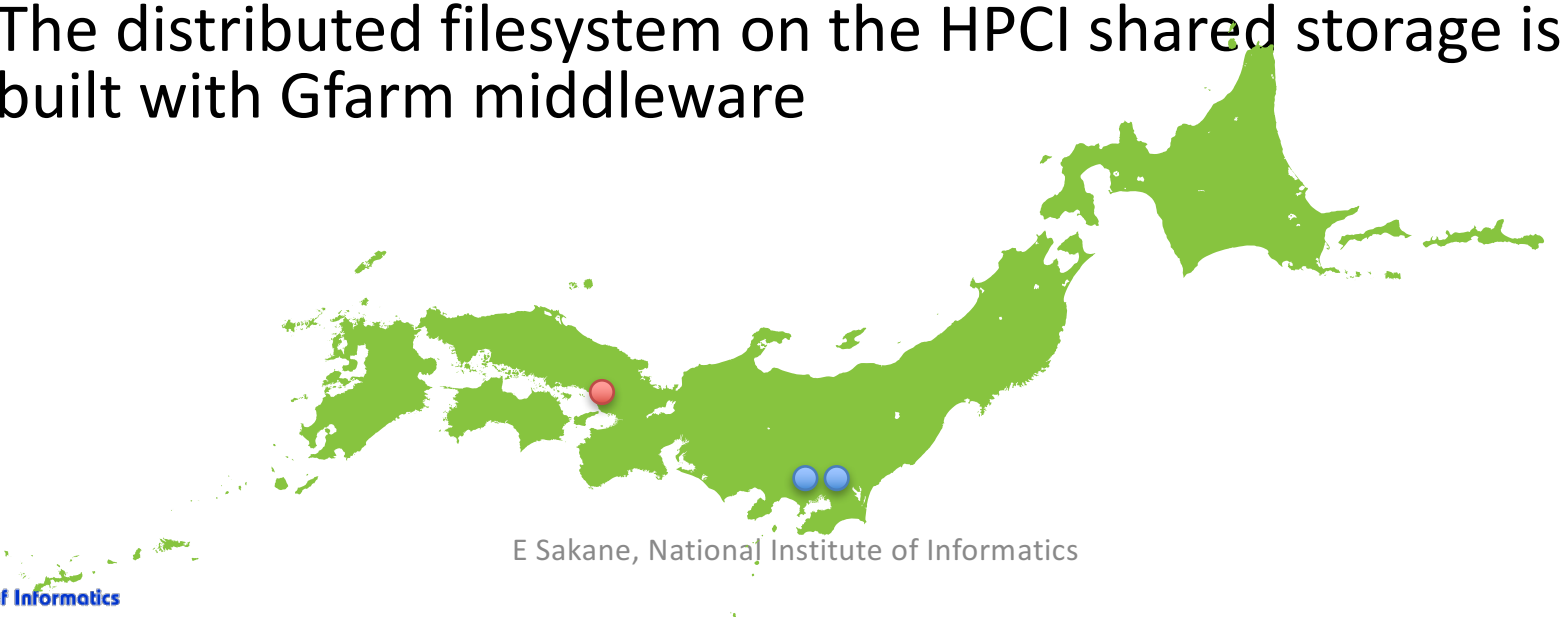
Provider	Resource Name	Amount of Available Resources
RIKEN AICS	K computer	82,944 nodes (10.62 PFLOPS), 30 PB
Hokkaido U	Super Computer HITACHI SR1600/M1	166 nodes (163 TFLOPS), 120 TB
Tohoku U	Supercomputer SX-ACE	2,560 nodes (707 TFLOPS), 4 PB
U of Tsukuba	COMA (PACS-IX)	90 nodes (230 TFLOPS), 300 TB
U of Tokyo	Supercomputer FX10	744 nodes (175.97 TFLOPS), 248 TB
TITECH	TSUBAME 2.5	Max. 420 nodes (64.3 TFLOPS), 1—10 TB per project
Nagoya U	FX100	2,880 nodes (3.2 PFLOPS)
Kyoto U	Supercomputer Cray XC30	32 nodes (33 TFLOPS, shared full year use)
Osaka U	Super Computer SX-ACE	1,024 nodes (282 TFLOPS), 200TB
Kyushu U	Super Computer Fujitsu PRIMEHPC FX10	384 nodes (90.8 TFLOPS)
JAMSTEC	Earth Simulator	5,120 nodes (1,310 TFLOPS), 4.7 PB
ISM	Data Assimilation Super Computer UV 2000	128 nodes (49 TFLOPS), 100 TB

As of FY2016

from http://www.hpci-office.jp/pages/e_h28_boshu_hpci_resource

HPCI – Shared Storage

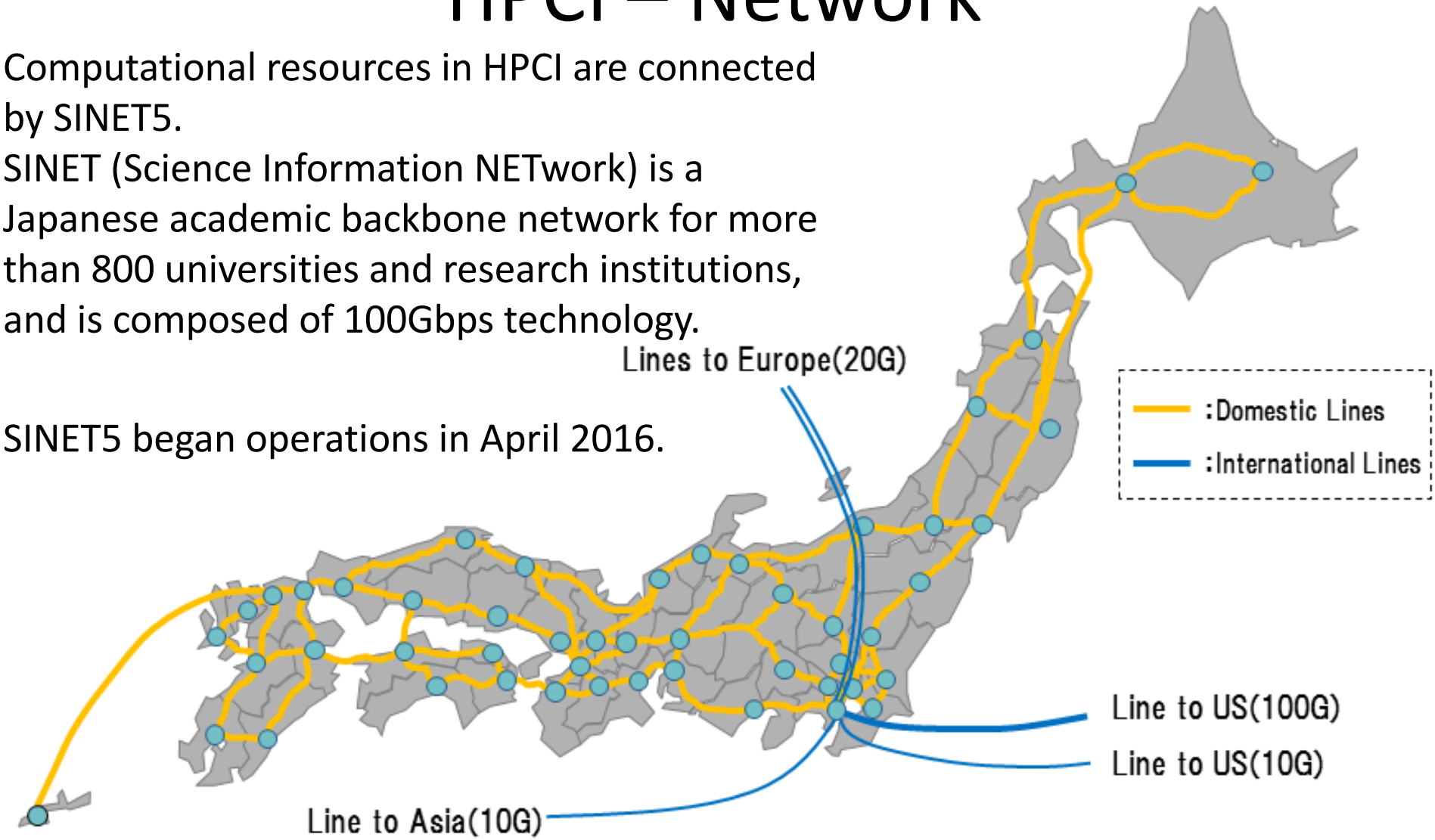
- Providers
  - Information Technology Center, The University of Tokyo (Eastern HUB)
  - RIKEN Advanced Institute for Computational Science (Western HUB)
  - Grobal Scientific Information and Computing Center, Tokyo Institute of Technology (TITECH HUB)
- Total size: 21.9PB
  - Eastern HUB: 11.5PB, Western HUB: 10PB, TITECH HUB: 0.4PB
- The distributed filesystem on the HPCI shared storage is built with Gfarm middleware



HPCI – Network

Computational resources in HPCI are connected by SINET5.
SINET (Science Information NETwork) is a Japanese academic backbone network for more than 800 universities and research institutions, and is composed of 100Gbps technology.

SINET5 began operations in April 2016.

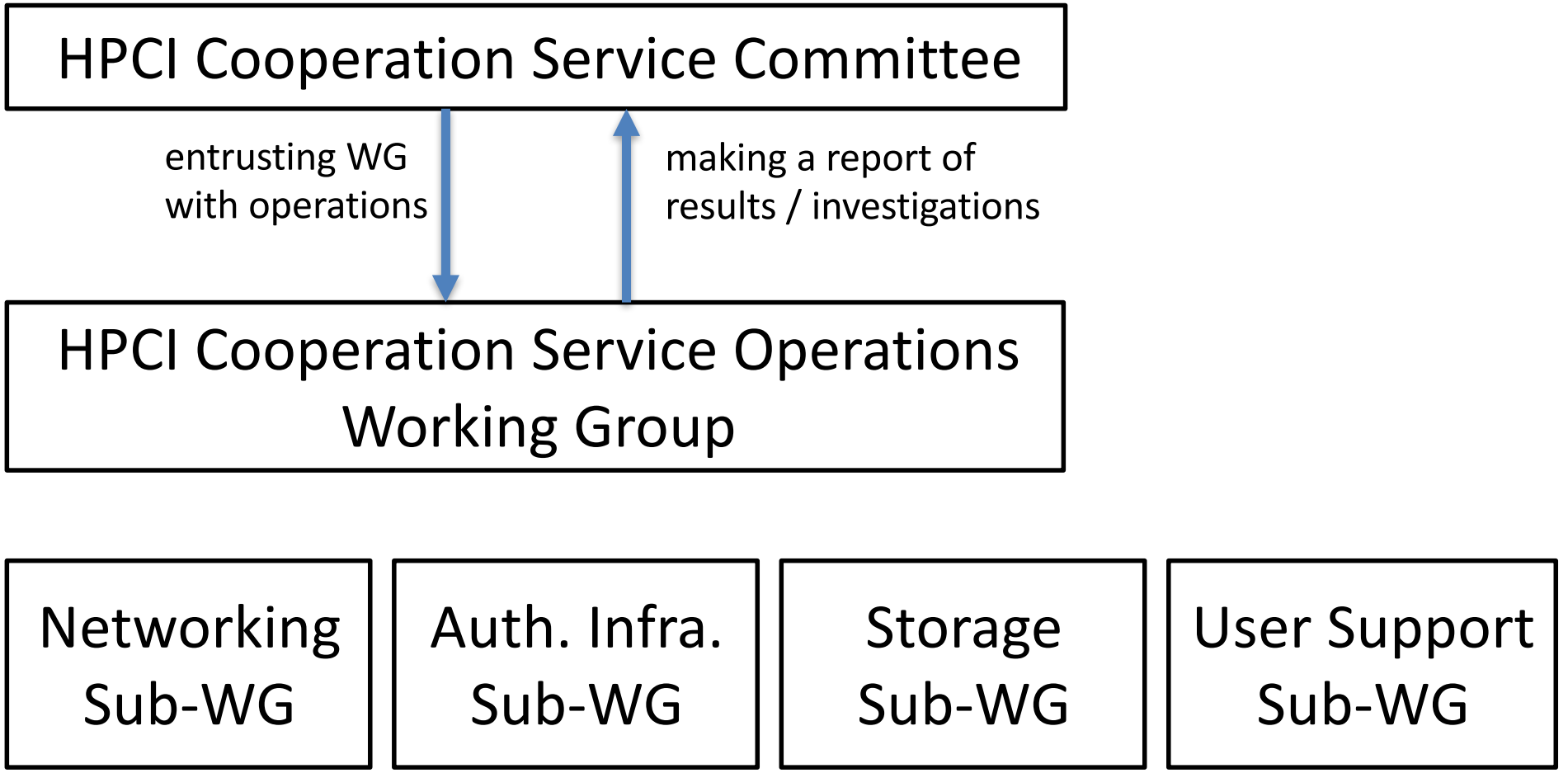


from https://www.sinet.ad.jp/en/aboutsinet-en

Guide to the HPCI System for Prospective Users

- Investigation of HPCI usage
  - Available project categories for proposals
    - Classification of projects and industrial use
  - Available hardware and software resources
- Application for project proposal
  - Acquisition of HPCI-ID
  - Submission of Application Form
  - Confirmation of awarding results
- Implementation of the project
  - Procedures after acceptance of the project proposal
- After finishing the project
  - Submission of User Report

Operating System

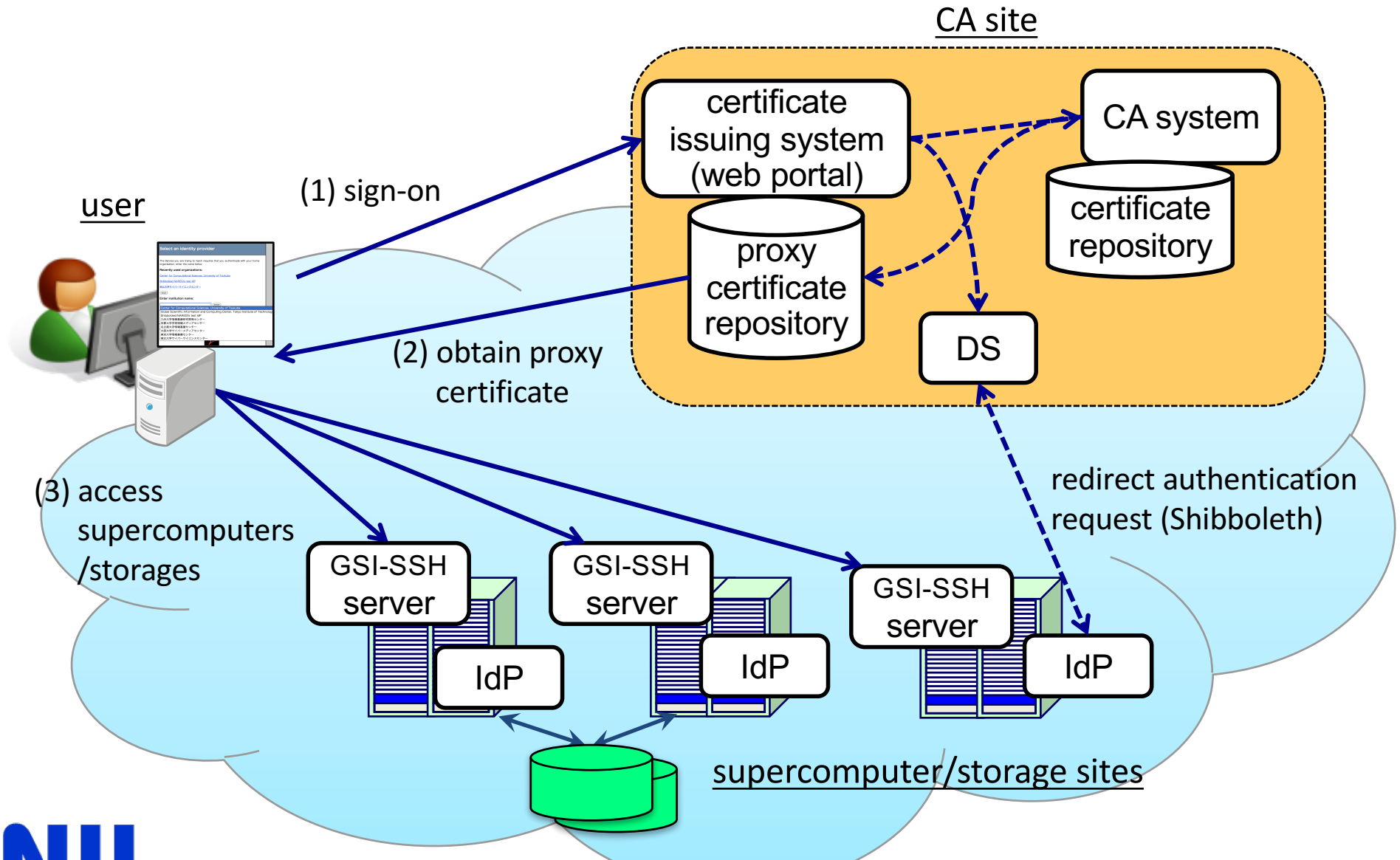


Authentication System

Authentication

- HPCI authentication system features include
  - **Single user ID and multiple accounts** called HPCI-ID, HPCI and local accounts
  - **A hierarchical initial identity vetting** system based on face-to-face meetings with photo-IDs
  - **Two kinds of credentials** for services in HPCI:
    - Shibboleth assertion for Web services: certificate issuance, CMS, etc.
    - GSI proxy certificate for access to supercomputers and shared storages

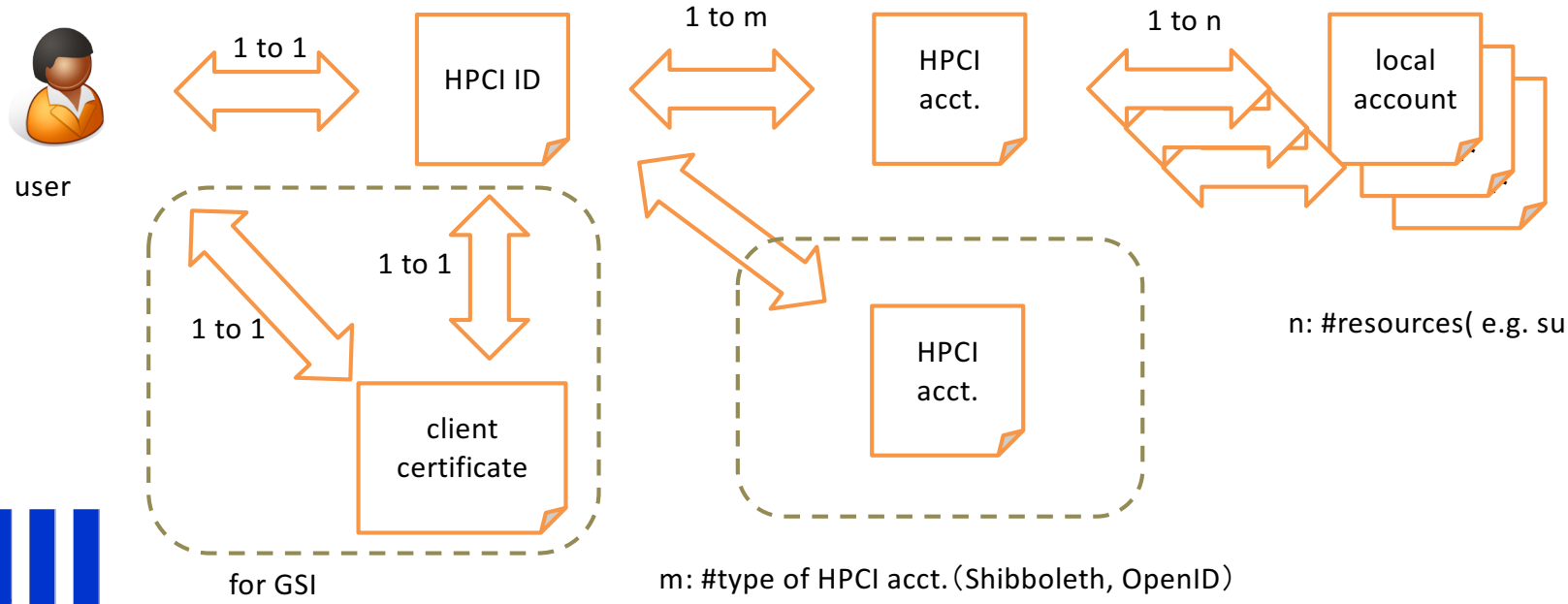
Single Sign-on in HPCI



National Institute of Informatics

ID／Accounts

ID/acct.	Issuer	#acct./user	
HPCI ID	HPCI operations office	1／1user	ID to distinguish a person applying HPCI
HPCI acct.	HPCI primary center (IdP)	m／1user	account for single sign-on HPCI resources
local acct.	Resource Provider	n／1user	local accounts (UNIX accounts) in resources (e.g. supercomputers)
client certificate	Certificate Authority	1／1user	client certificate used in GSI

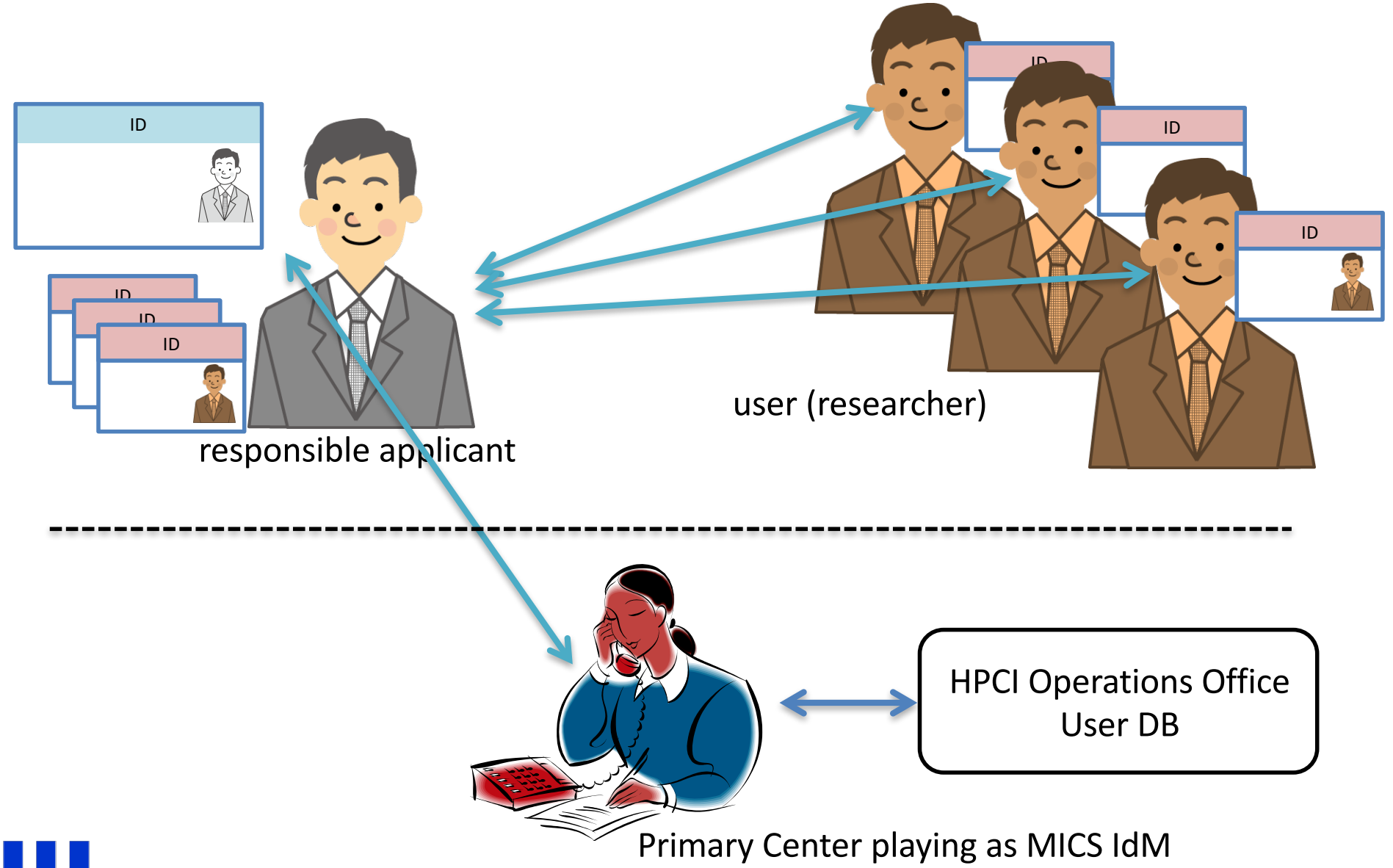


National Institute of Informatics

Procedures after acceptance of the proposal

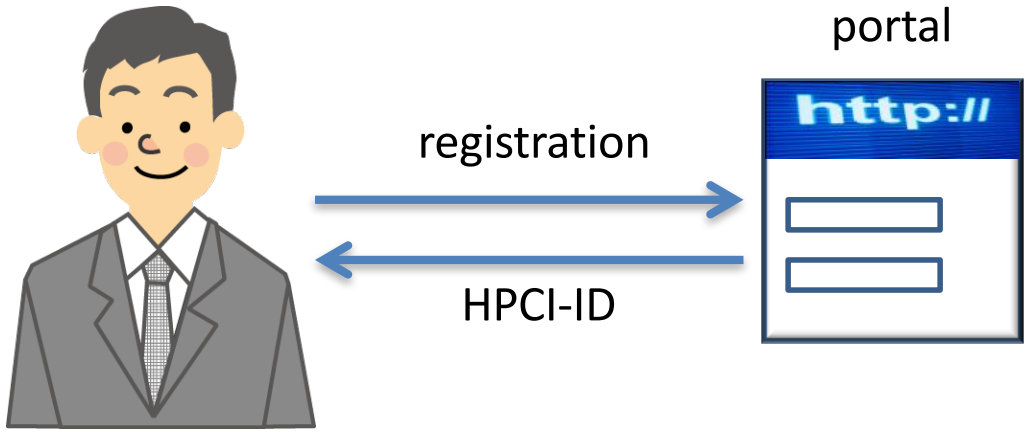
- Confirmation of the category and amount the allocated resource
- Face-to-face identity vetting
- Procedures for issuing HPCI account
- Receipt of the HPCI account
  - The HPCI account is issued by the primary center
- Receipt of the local account(s)
  - The local account is issued by each HPCI resource provider

Overview of initial vetting of identity in HPCI



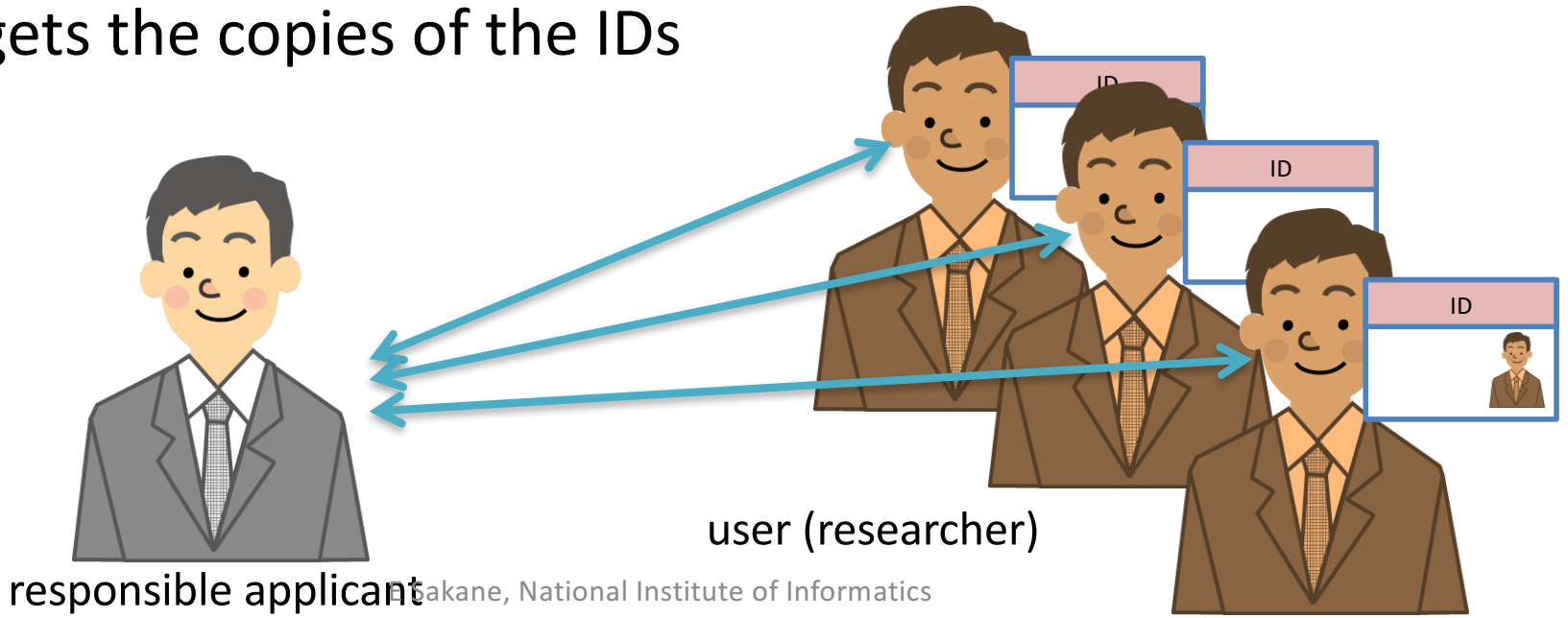
Initial vetting of identity in HPCI (1/3)

- First, each user registers information such as name and affiliation with the HPCI operations office via the HPCI portal, then gets an HPCI-ID.



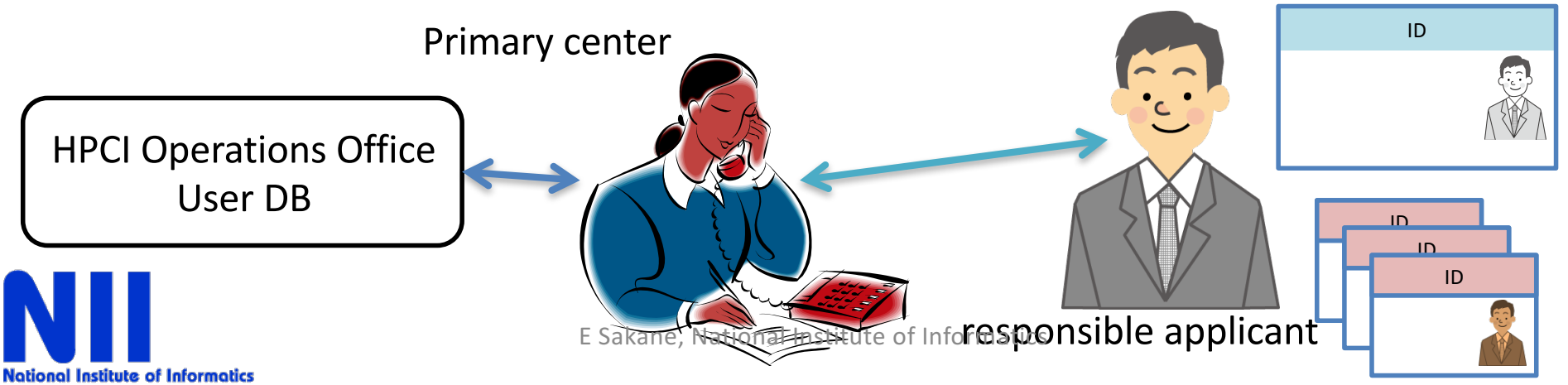
Initial vetting of identity in HPCI (2/3)

- Responsible applicant (project representative or the Deputy project representative)
  - vets the identity of all researchers in her/his project
    - based on a face-to-face meeting
    - confirmed via photo-identification and/or similar valid official documents
  - gets the copies of the IDs

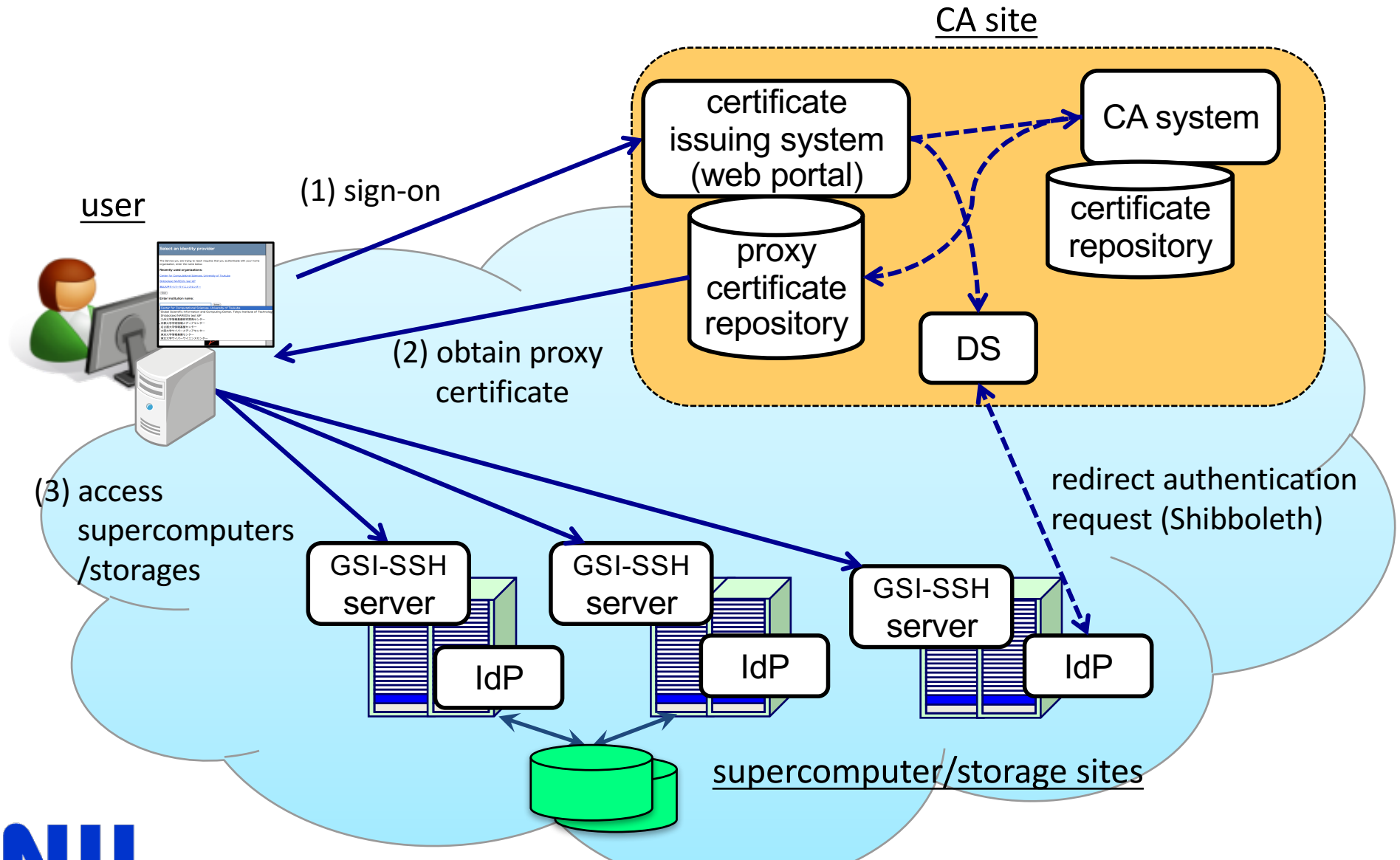


Initial vetting of identity in HPCI (3/3)

- An HPCI primary center
  - initially vets the identity of the responsible applicant
    - based on a face-to-face meeting
    - confirmed via photo-identification and/or similar valid official documents
  - receives the list of users (researchers) with the copies of IDs
  - confirms that each user in the list matches the copy of the ID.



Single Sign-on in HPCI



National Institute of Informatics

Software

role	system	software
Certificate Authority	CA system	NAREGI-CA
	certificate management	custom software
	certificate repository	MyProxy
	ID federation	Shibboleth
Portal （NII，supercomputer centers）	portal (cert. issuing system)	custom software
	Proxy certificate repository	MyProxy
	ID federation	Shibboleth
Identity Provider （supercomputer centers, AICS）	ID federation	Shibboleth
Resource Provider （supercomputer centers, AICS）	middleware to access resources	GSI-enabled SSH Gfarm

HPCI CA

- The production level operation started in Sep. 2012
  - approved by IGTF in Aug. 2014
  - implemented with NAREGI-CA software
- MICS based CA
  - based on MICS profile version 1.3
- Certificate statistics as of 2016/09/12
  #valid (#issued)
  - client certificates: 255 (1892)
  - host certificates: 183 (929)
    host: 78 (408)
    service (gfsd) : 105 (521)

PKI Participants in HPCI

- HPCI CA (NII)
  - issue certificates for access to resources
- HPCI Operations Office (RIST)
  - manage HPCI-ID and user information associated with the HPCI-ID
- IdP Operating Organizations (Supercomputing centers)
  - manage accounts for Shibboleth authentication
  - primary centers: Shibboleth IdPs
  - check centers: desks for initial vetting of identity
- HPCI system providers (Supercomputing centers)
  - GSI-enabled SSH servers that communicate with the CA to obtain a grid-mapfile created by the CA

PKI Participants in HPCI

- HPCI CA (NII)
  - issue certificates for access to resources

IdM org.

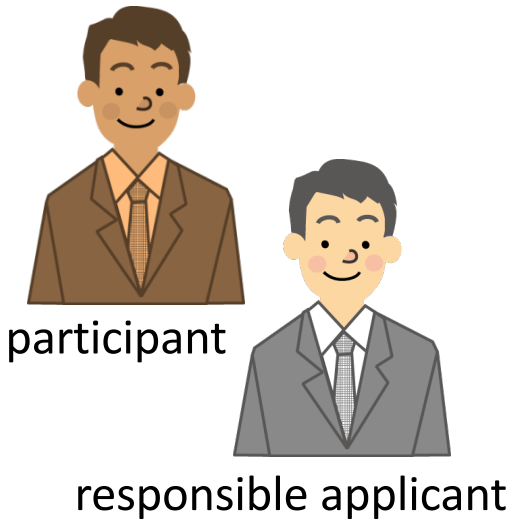
- HPCI Operations Office (RIST)
  - manage HPCI-ID and user information associated with the HPCI-ID
- IdP Operating Organizations (Supercomputing centers)
  - manage accounts for Shibboleth authentication
  - primary centers: Shibboleth IdPs
  - check centers: desks for initial vetting of identity
- HPCI system providers (Supercomputing centers)
  - GSI-enabled SSH servers that communicate with the CA to obtain a grid-mapfile created by the CA

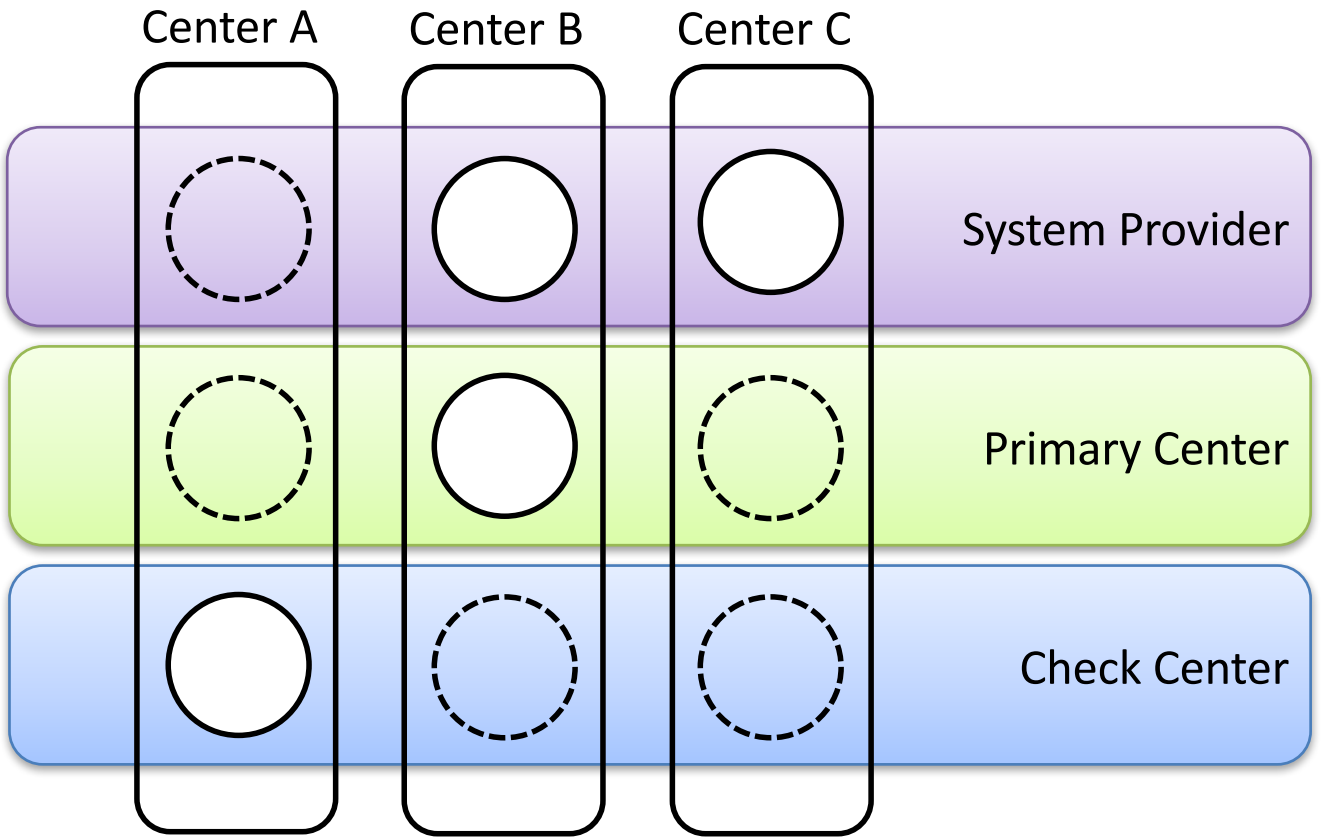
National Institute of Informatics

PKI Participants

HPCI Operations Office

HPCI CA





National Institute of Informatics

NAREGI-CA

- An open source software package for building a PKI: HPCI CA is built with NAREGI-CA
- Developed by NAREGI project (FY2003–2007)
- Originally designed to be suitable for grid computing environment. But this is not restricted to grid
- Currently maintained by NII in Japan
- Another protocol for interaction between CA components called “Lightweight Certificate Management Protocol (LCMP)”
  - IA-RA separated model (model A) is easily built
- HPCI CA is built with NAREGI-CA

GSI-SSHTerm

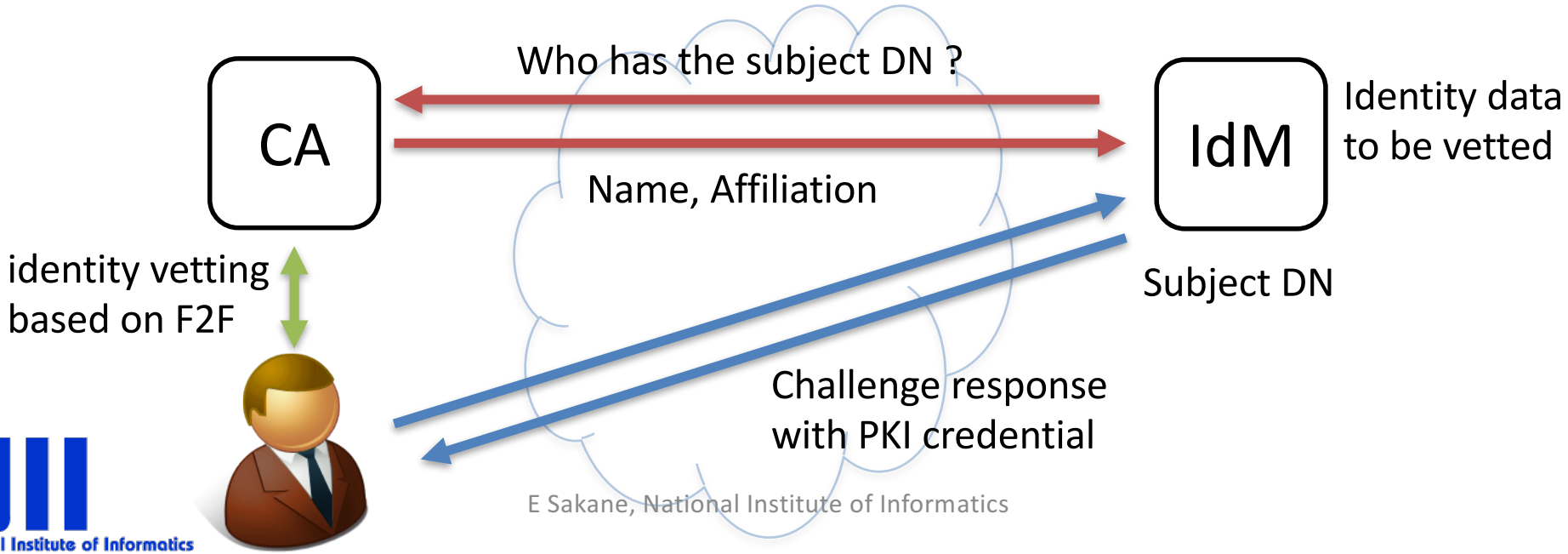
- HPCI recommends the GSI-SSHTerm as GSI-enabled SSH client software
- Current version: 0.91i-nii5
  - based on the development branch ‘jglobus2_branch’ of UK NGS
  - localized for HPCI
- improved to solve some problems:
  - fixed a handshake problem
  - added a heartbeat function to keep a session
- out-of-date cipher suite problem
  - GSI-SSHTerm does not support recent cipher suites such as aes128-ctr, hmac-sha2-256
  - We plan to add the following as supported ciphers: aes128-ctr, aes128-cbc, hmac-sha2-256, hmac-sha2-512

Current Issues

- Remote identity vetting of users
  - Anyone can apply a project proposal to the HPCI
    - researcher who do not live in Japan nor belong to organization located in Japan
  - How does the HPCI vet the identity of user who do not live in Japan ?
  - We need to establish methods of remote identity vetting of users.
- Possible methods
  - remote identity vetting with live video
  - using credential in cooperation with CAs

Using PKI Credential

- Basic Idea
  - challenge response between applicant and IdM
  - the IdM makes reference to the CA issuing the applicant's certificate for his/her identity data
  - After confirming correctly, the IdM consider the applicant's identity to be confirmed



Summary

- This talk presents an overview of HPCI and authentication and authorization system in HPCI
- Future plan
  - The HPCI second stage will start in FY2017
  - AAI in HPCI will be basically unchanged
  - We continue to work on development and management of AAI

http://www.hpci-office.jp