2011年12月21日

NIIのHPCIへの取り組み

合田 憲人
国立情報学研究所

平成23年第2回学術情報基盤オープンフォーラム

Kento Aida, National Institute of Informatics

HPCI

- HPCI
  - 京コンピュータと 国内のスーパーコンピュータや大規模ストレージを連携して利用するための革新的ハイパ フォーマンス・コンピューティング・インフラ
- HPCI基本仕様の策定（2010～）
  - 5サブWG：ネットワーク，認証，ユーザ管理支援，共用ストレージ,先端ソフトウェア運用
  - 北大，東北大，筑波大，東大，東工大，名大，京大，阪大，九大，理研，NII

本講演の目的

NIIのHPCIへの取り組みとして，HPCI 基本仕様として検討されているネットワーク基盤，認証基盤に関する報告

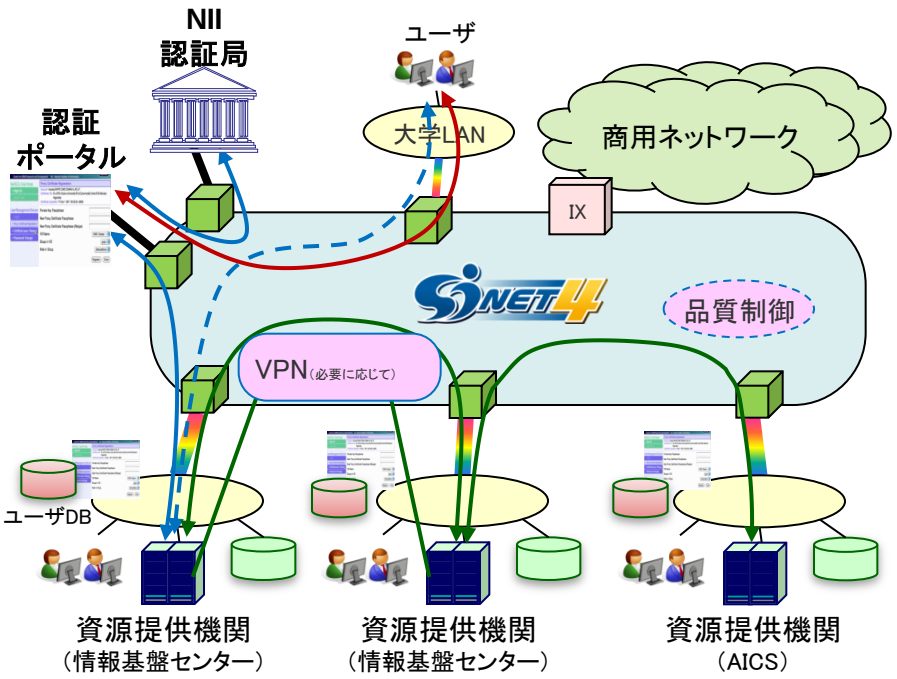
HPCIに対する取り組み

■HPCI基本仕様

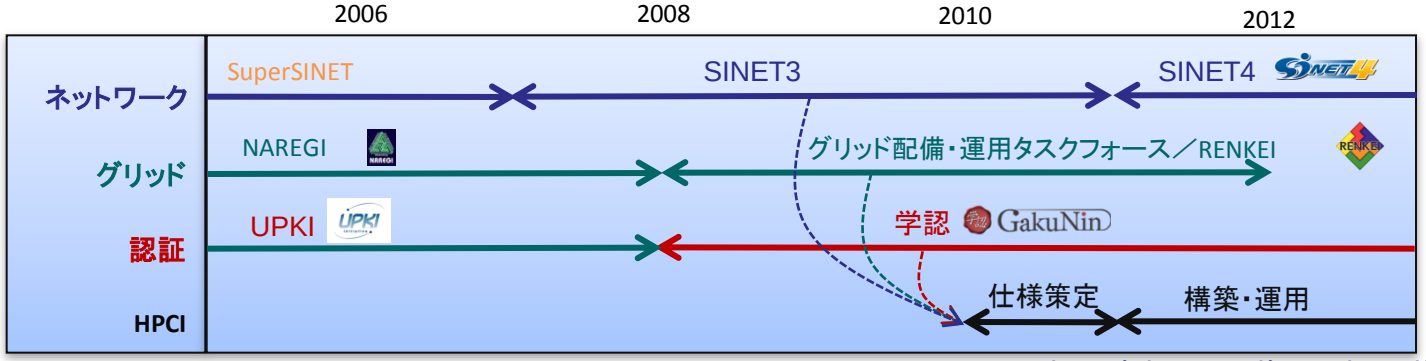
ネットワーク（NII）	資源提供機関群を接続するためのネットワーク基盤をSINET4を活用して構築・運用する. ✓ HPCIのニーズに対応してネットワーク増強. ✓ 商用ネットワークとの接続も確保.
認証基盤（NII）	資源提供機関群のリソース（計算機・ストレージ等）にユーザがシングルサインオンして利用するための認証基盤を構築・運用する. ✓ NIIが中核拠点となり，情報基盤センターと連携して認証基盤を運用.
共有ストレージ（情報基盤センター，理研）	資源提供機関群の計算機からアクセス可能な高性能な共有ストレージを構築・運用する.
先端ソフトウェア運用基盤（情報基盤センター）	HPCI上での将来の利用が想定される先端的なソフトウェア運用実験を行うための実証基盤を構築・運用する.
ユーザ管理支援（情報基盤センター，HPCI事務局）	HPCIの課題申請，利用申請，障害対応，ヘルプデスク業務等のユーザ管理支援業務を実施する.

■ネットワーク・認証基盤サービス

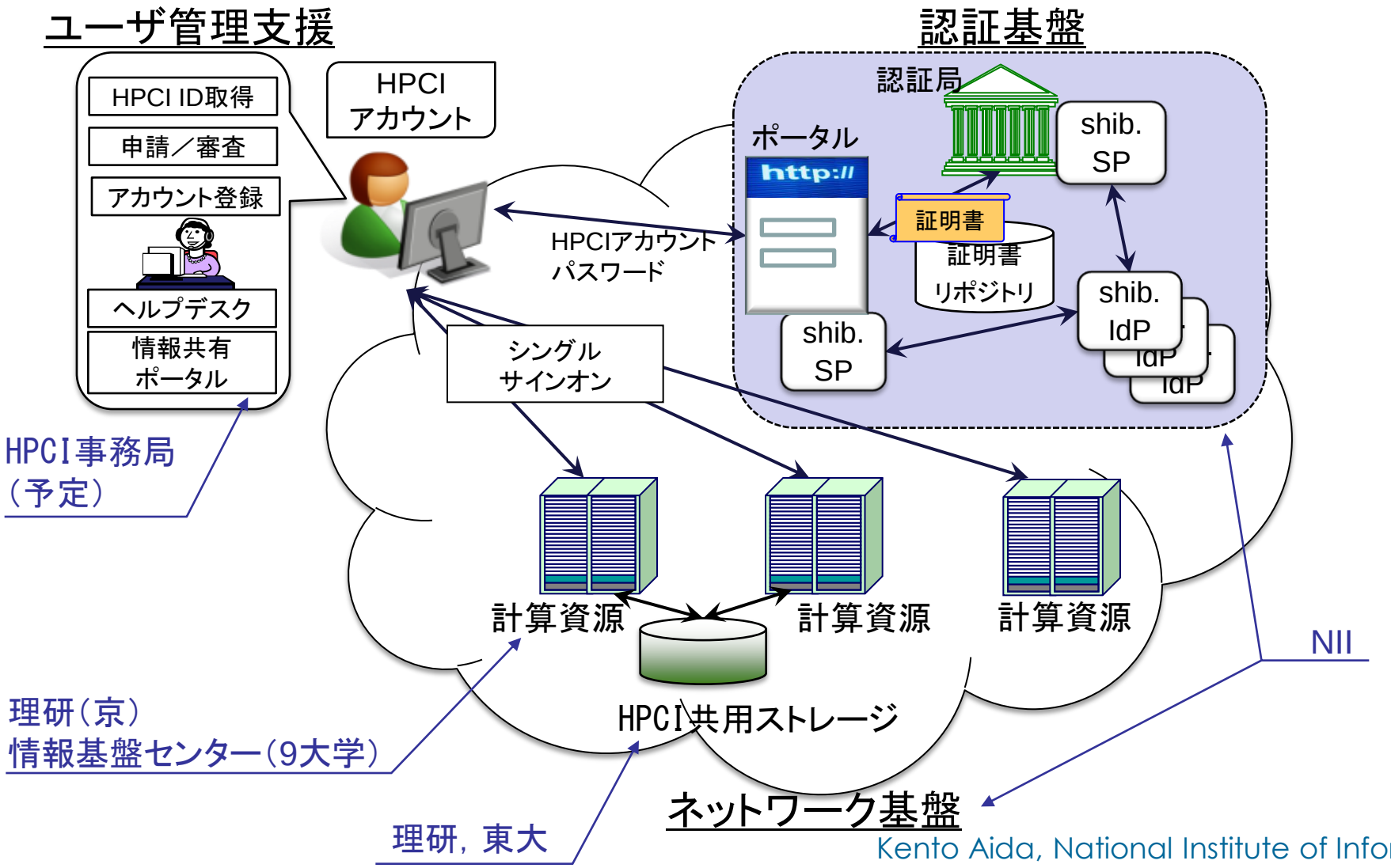
- ✓ 認証ポータルにサインオンすることにより，資源提供機関群の計算機やストレージ利用が可能（1つのアカウントで全ての資源を利用可能）.
- ✓ SINET4を活用した高速・高品質ファイル転送が可能.



■ロードマップ（NIIのこれまでの取り組み）



インフラ概要（2012年11月スタート時）



AICS and Supercomputer Centers in Japanese Universities

AICS, RIKEN：
K computer (10 Pflops, 4PB)
Available in 2012

Hokkaido Univ.：
SR16000 M1 (172TFlops, 22TB)
Cloud system (43TFlops, 14TB)

Kyoto Univ.
T2K Open Supercomputer
(61.2 Tflops, 13 TB)

Tohoku Univ.：
NEC SX-9(29.4Tflops, 18TB)
NEC Express5800 (1.74Tflops, 3TB)

Osaka Univ.：
SX-9 (16Tflops, 10TB)
SX-8R (5.3Tflops, 3.3TB)
PCCluster (23.3Tflops, 2.9TB)

Univ. of Tsukuba：
T2K Open Supercomputer
95.4Tflops, 20TB

Univ. of Tokyo：
T2K Open Supercomputer
(140 Tflops, 31.25TB)

A 1 Pflops machine without accelerator will be installed by the end of 2011

Kyushu Univ.：
PC Cluster (55Tflops, 18.8TB)
SR16000 L2 (25.3Tflops, 5.5TB)
PC Cluster (18.4Tflops, 3TB)

Nagoya Univ.：
FX1(30.72Tflops, 24TB)
HX600(25.6Tflops, 10TB)
M9000(3.84Tflops, 3TB)

Tokyo Institute of Technology：
Tsubame 2
(2.4 Pflops, 100TB)

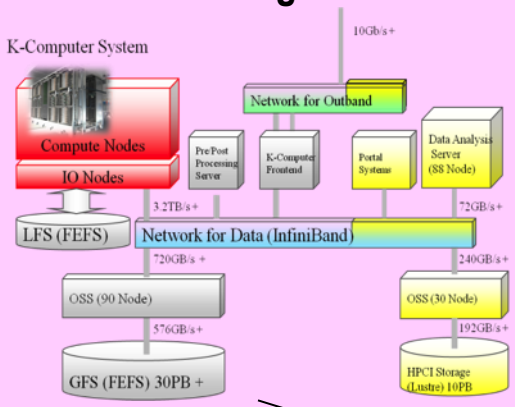
source: Y. Ishikawa, Univ. of Tokyo

Storage

HPCI WEST HUB

AICS, RIKEN

- 10 PB+ storage

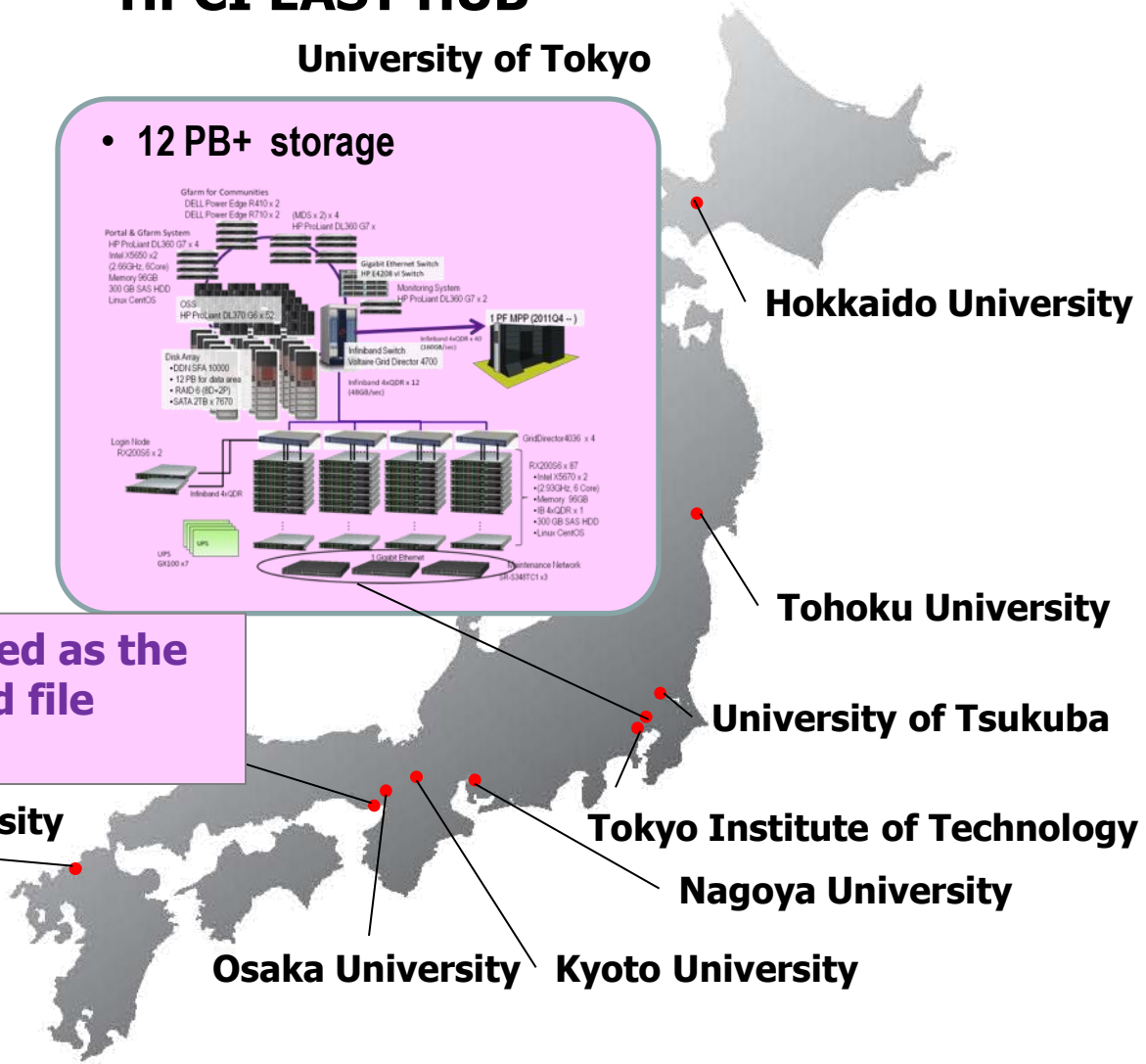


HPCI EAST HUB

University of Tokyo

- 12 PB+ storage

Gfarm2 is used as the global shared file system



Hokkaido University

Tohoku University

University of Tsukuba

Tokyo Institute of Technology

Nagoya University

Kyoto University

Osaka University

Kyushu University

source: Y. Ishikawa, Univ. of Tokyo

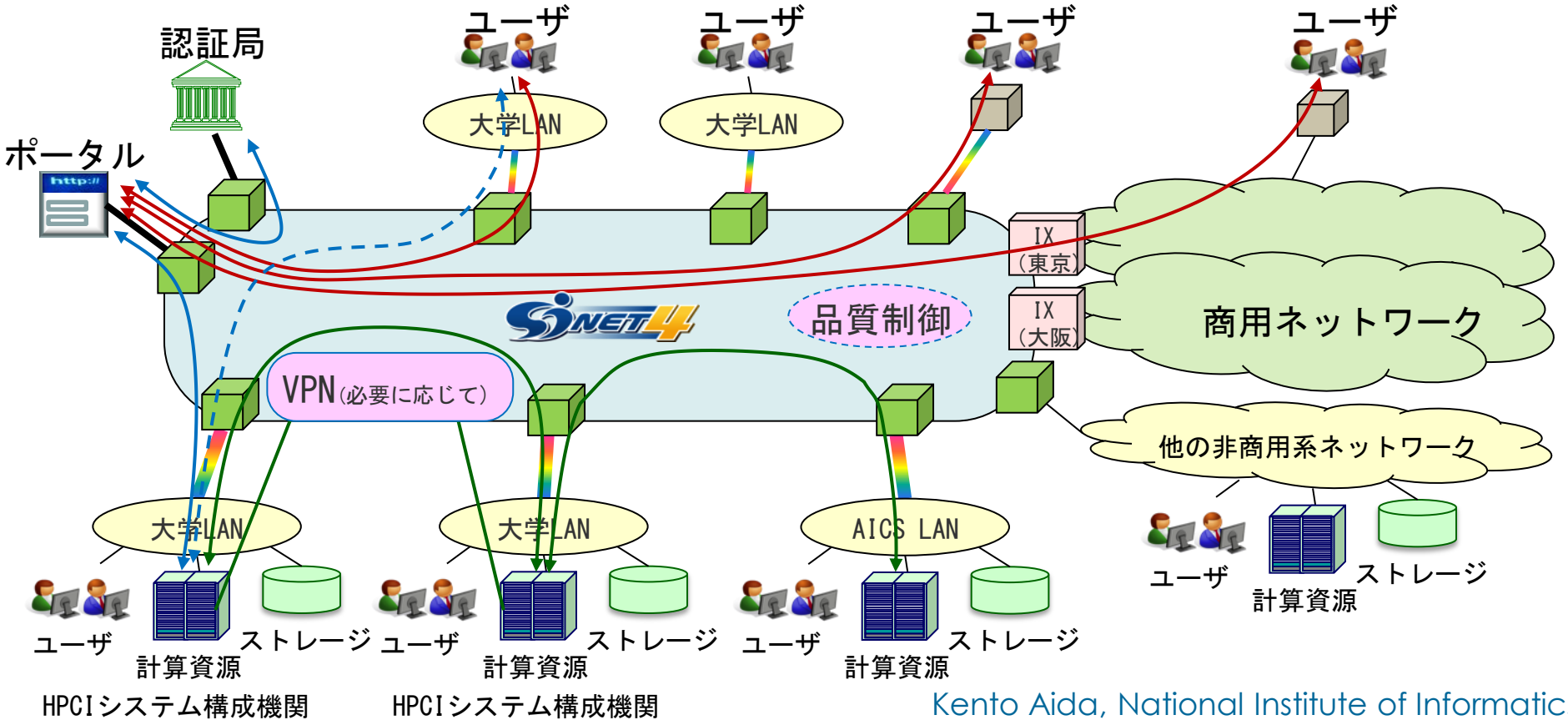
ネットワーク基盤

ネットワーク基盤の課題

- 幅広いユーザに対応した資源への制限のないアクセス
  - SINET4 のみを用いてアクセス
  - 商用ネットワーク等を経由してSINET4に接続してアクセス
  - 指定された拠点に出向いてアクセス
- 高速ネットワーク環境
  - 十分なバックボーン帯域
  - 通信品質制御
- セキュリティ
  - インターネット or VPN?
  - ファイアウォール

ネットワーク基盤の設計

- SINET4の利用
  - 商用ネットワーク接続:
    東京134(30Gbps), 大阪22(11Gbps)
  - 東京ー大阪80Gbps(H23予定)
  - L3VPN, L2VPN/VPLS, QoS, リソースオンデマンド等



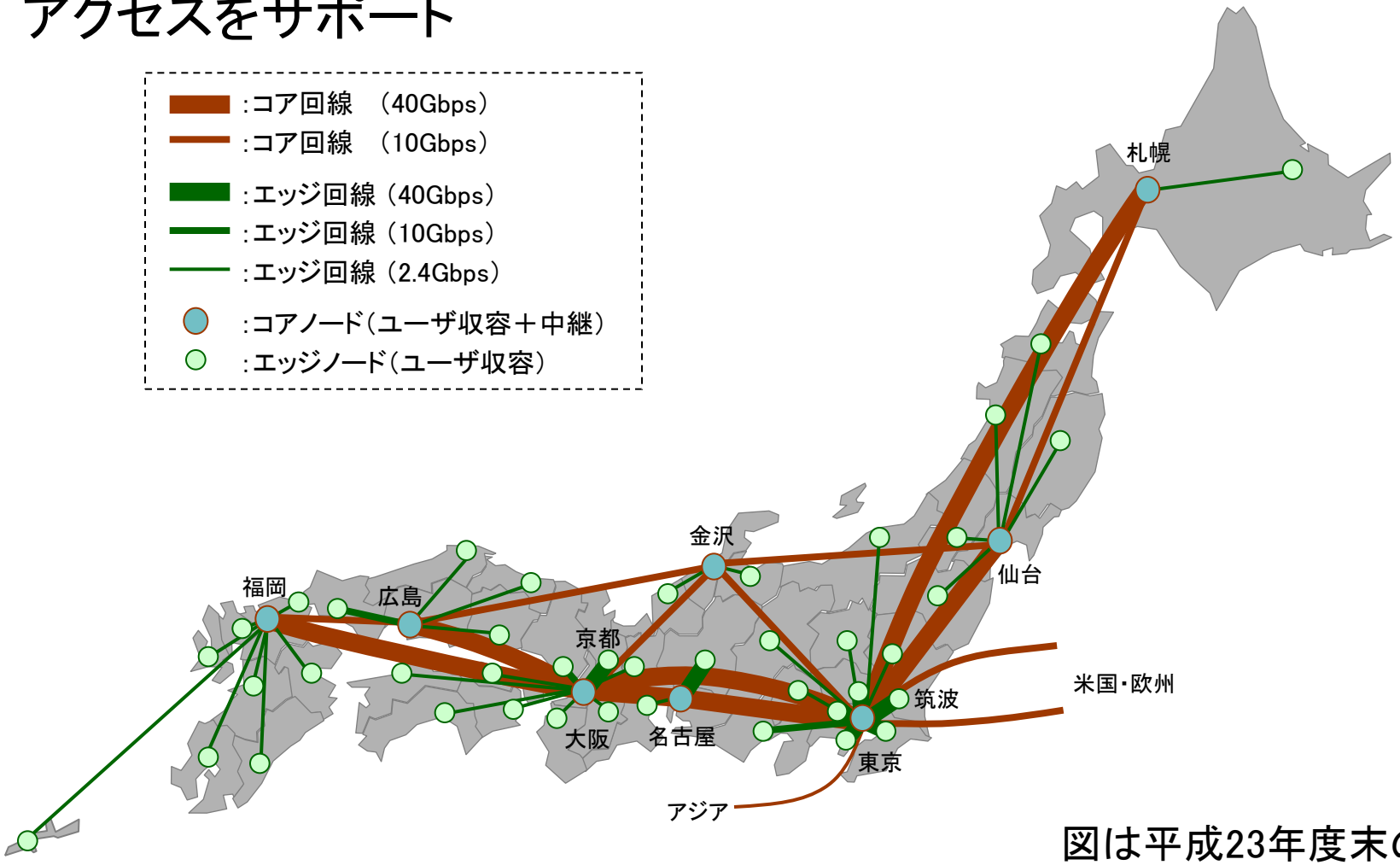
所要帯域の分析

松／竹／梅：データを6分／1時間／10時間で転送

No	機関名	アンケートからの抜粋部分	換算帯域（Gbps）			
			対地	梅	竹	松
1	機関A	AICS～機関A： 数TB 東京大学～機関A： 数10GB	AICS～機関A（3TB） 東京大学～機関A（30GB）	0.67 0.067	**6.7** **0.67**	67 6.7
2	機関B	AICS～機関B： 1Gbps 研究所Y～機関B： 10Gbps 研究所Z-機関B： 10Gbps	AICS-機関B 研究所Y-機関B 研究所Z-機関B	1 - -	**-** **10** **10**	- - -
3	機関C	AICS～機関C： 1Gbps 機関C（～東京DC）： 10Gbps	AICS～機関C 機関C（-東京DC）	1 -	**-** **10**	- -
4	機関D					
5	機関E	AICS～利用者ローカル環境： 10TB/1日 AICS～拠点間： 100GB	AICS～神戸DC AICS～神戸DC	2.2 0.022	**-** **0.22**	- 2.2
6	機関F	機関F（～横浜DC）： 10Gbps	機関F（～横浜DC）	-	**10**	-
7	機関G	AICS～機関G： 20TB	AICS～機関G	4.4	**44**	440
8	機関H	AICS～機関H： 10GB/s	AICS～機関H	-	**8**	80
9	機関I	機関I（～名古屋DC）： 100Gbps	機関I（～名古屋DC）	-	**10**	100
10	機関J	機関J（～札幌DC）： 2.4Gbps	機関J（～札幌DC）	-	**2.4**	24
11	機関K	AICS～機関H～機関D/機関K： 10GB/s	AICS～機関H～機関D/機関K	-	**8**	80
12	機関L	AICS～機関L： 128GB/日	AICS～機関L	0.028	**0.28**	2.8
13	機関M	AICS～機関M： 50TB/4日以上	AICS～機関M（12.5TB/1日）	2.75	**27.5**	275
14	機関N	AICS～機関N： 10GB AICS～機関N： 1TB/1日	AICS～機関N AICS～機関N	0.002 0.22	**0.022** **2.2**	0.22 22
15	機関O	AICS～機関O： 10Gbps以上	AICS～機関O	-	**10**	-
16	機関P	AICS～機関P：10GB	AICS～機関P	0.002	**0.022**	0.22
17	機関Q	AICS-機関Q: 1TB/1時間	AICS-機関Q	0.22	**2.2**	22

SINET4の構成

■ 各県配備のSINETノードと北海道－九州を縦断する40Gbpsベースの超高速・高信頼バックボーンによりHPCIの計算資源へのアクセスをサポート



図は平成23年度末の構成

セキュリティ

- 基本はインターネット接続
  - 必要な通信ポートの到達性、そのほかはHPCIシステム構成機関のポリシに依存する
- 先端ソフトウェア実験向けに専用の論理網も検討

HPCIシステム構成機関における利用通信ポート

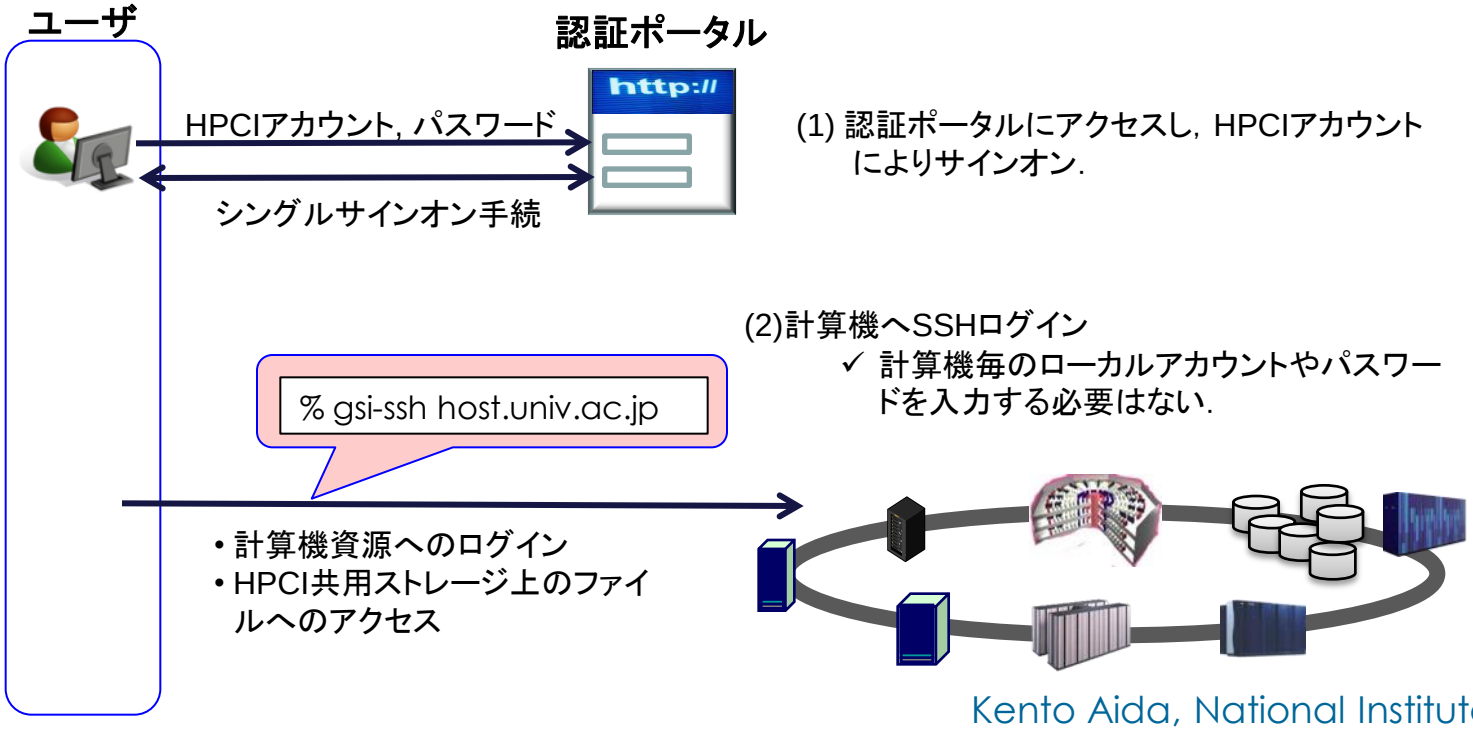
	ポート	プロトコル	用途	機能
1	22/tcp	GSI-SSH	GSI-SSHサーバ, VMホスティングサーバ	認証 先端ソフト基盤
2	80/tcp	http	Shibboleth IdP, 認証ポータル, VMホスティングサーバ	認証 先端ソフト基盤
3	443/tcp	https	Shibboleth IdP, 認証ポータル, VMホスティングサーバ	認証 先端ソフト基盤
4	600/tcp, udp	gfsd	Gfarm	ストレージ
5	601/tcp	gfmd	Gfarm	ストレージ
6	7512/tcp	MyProxy	Proxy証明書リポジトリ	認証
7	11412/tcp	airad *1	証明書管理システム	認証

*1 outbound

認証基盤

認証基盤の課題

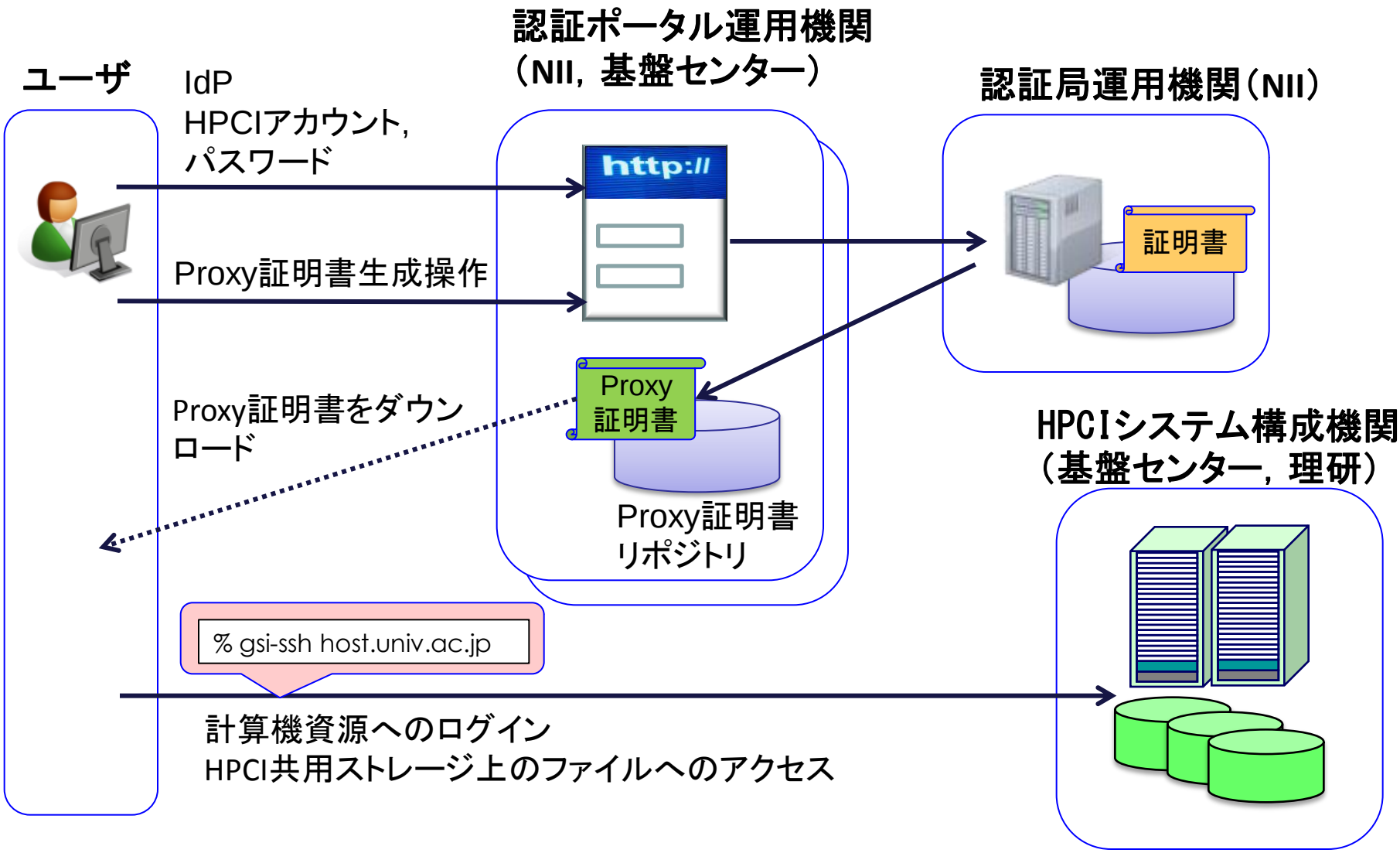
- シングルサインオンの実現
  - 1回のサインオンで全ての計算機やストレージにアクセス可能
- 既存技術の検討
  - 様々な認証方式を検討，運用実績を考慮
- アカウント管理
  - 統一管理 or 分散管理？



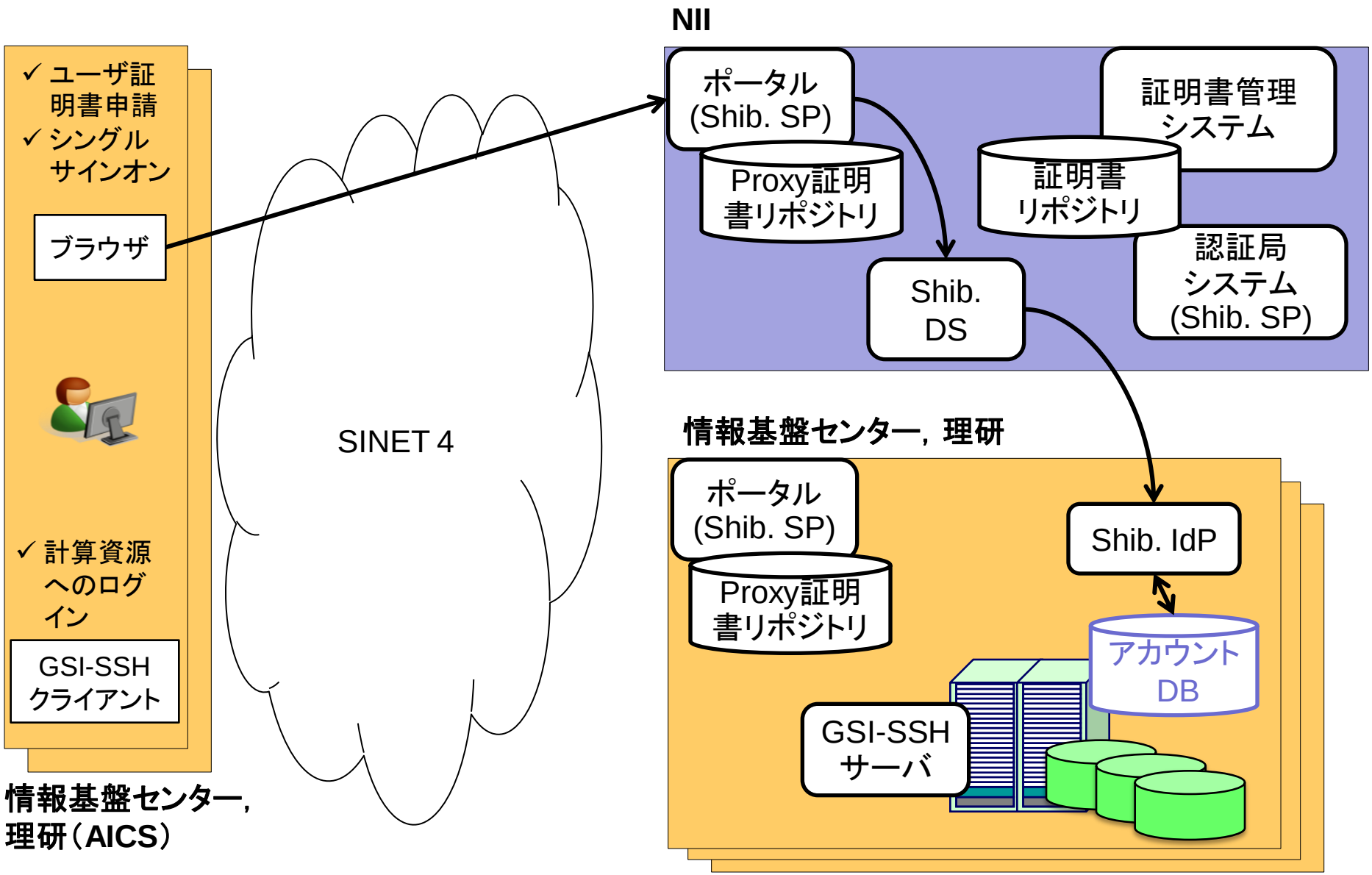
Shibboleth + GSI

- HPCIアカウント
  - Shibbolethを利用
    - ✓ 分散したユーザ管理システムを連携させる技術
  - HPCIシステム構成機関で運用されるユーザDBを連携
    - ✓ アカウント管理を分散.
    - ✓ ユーザはどこかのHPCIシステム構成機関にHPCIアカウントを持てばよい.
- シングルサインオン
  - Grid Security Infrastructure (GSI)を利用
    - ✓ PKIに基づくシングルサインオンを実現する技術
    - ✓ 従来のパスワード認証に安全性高.
    - ✓ グリッド技術におけるデファクトスタンダード
    - ✓ Proxy証明書の生成と委譲
  - 認可処理
    - ✓ 個人のユーザ証明書と計算機資源のローカルアカウントをマッピング.
    - ✓ HPCIシステム構成機関の運用方針に従った認可設定が可能.

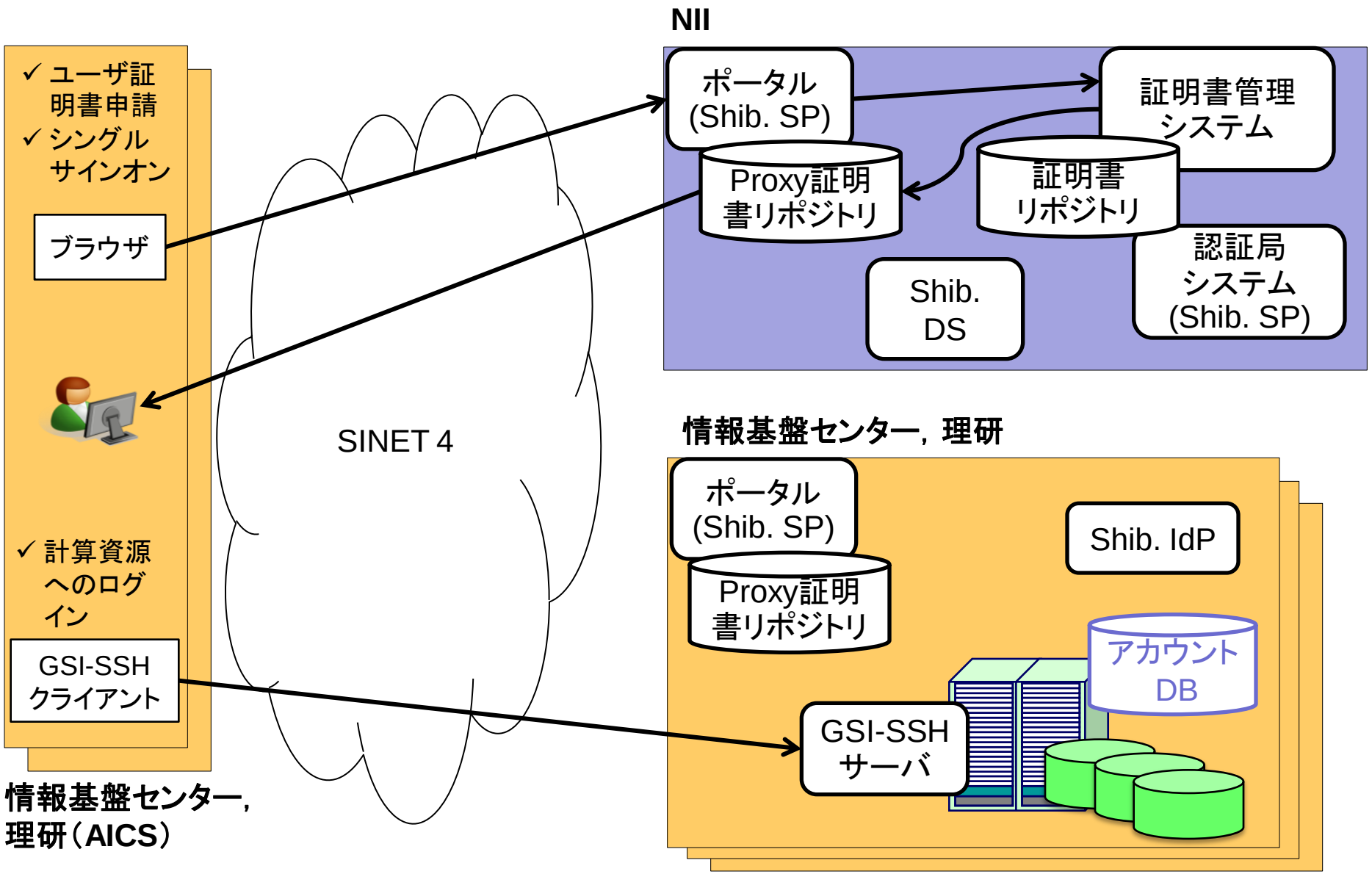
シングルサインオンの流れ



認証基盤アーキテクチャ



認証基盤アーキテクチャ（続）



認証基盤ソフトウェア

ソフトウェア仕様

機関	システム	ソフトウェア
認証局運用機関 （NII）	認証局システム	NAREGI-CA
	証明書管理	開発中
	証明書リポジトリ	MyProxy
	認証連携	Shibboleth
認証ポータル運用機関 （NII，情報基盤センターの一部）	認証ポータル	開発中
	Proxy証明書リポジトリ	MyProxy
	認証連携	Shibboleth
HPCIアカウントIdP運用機関 （情報基盤センター，理研）	認証連携	Shibboleth
HPCIシステム構成機関 （情報基盤センター，理研）	GSI認証機能付ミドルウェア	GSI-SSH

まとめ

- NIIのHPCIへの取り組みとして，HPCI 基本仕様として検討されているネットワーク基盤，認証基盤について報告
- 現状と今後の予定
  - ネットワーク
    - ✓ H23年度よりSINET4が稼働.
    - ✓ バックボーンネットワークの帯域増強，京コンピュータの接続方法について検討中.
  - 認証基盤
    - ✓ 9大学情報基盤センター，理研，NIIによるプロトタイプソフトウェアを用いた試験実施中.
    - ✓ H23年度中に試験運用開始予定.