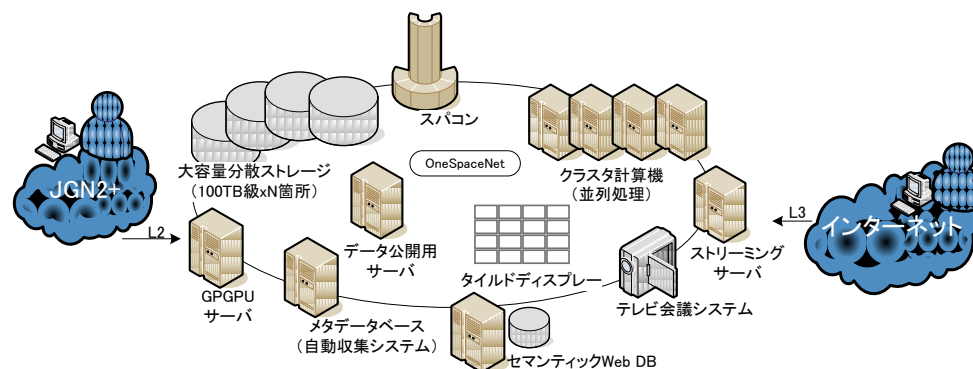


NICT サイエンスクラウド: OneSpaceNet クラウドストレージ

森川 靖大¹、山本 和憲¹、佐藤 建¹、井上 諭¹、坪内 健¹、
崔 超遠¹、亘 慎一¹、村田 健史¹、
NICT 宇宙環境計測グループ

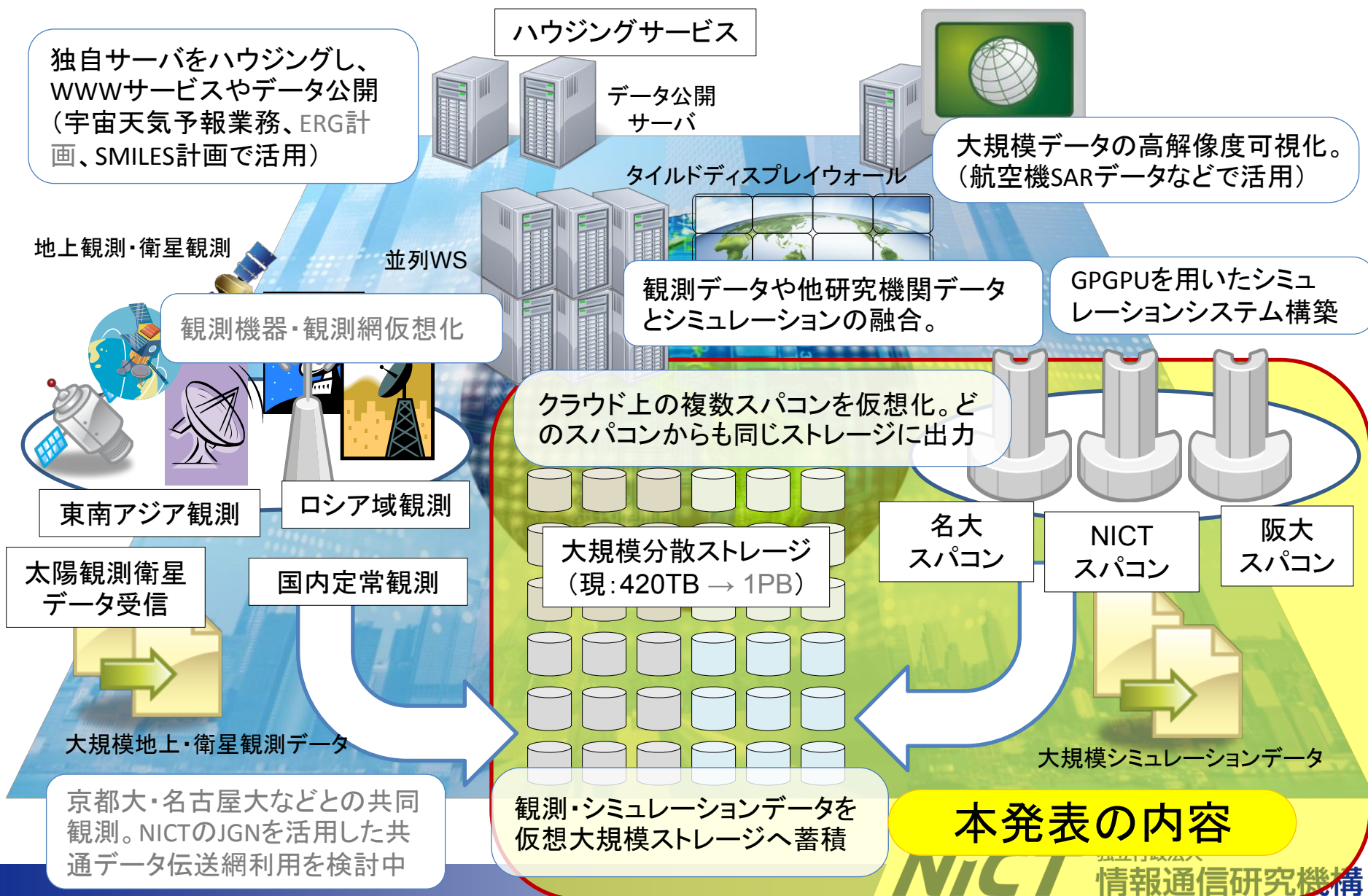
¹情報通信研究機構(NICT)



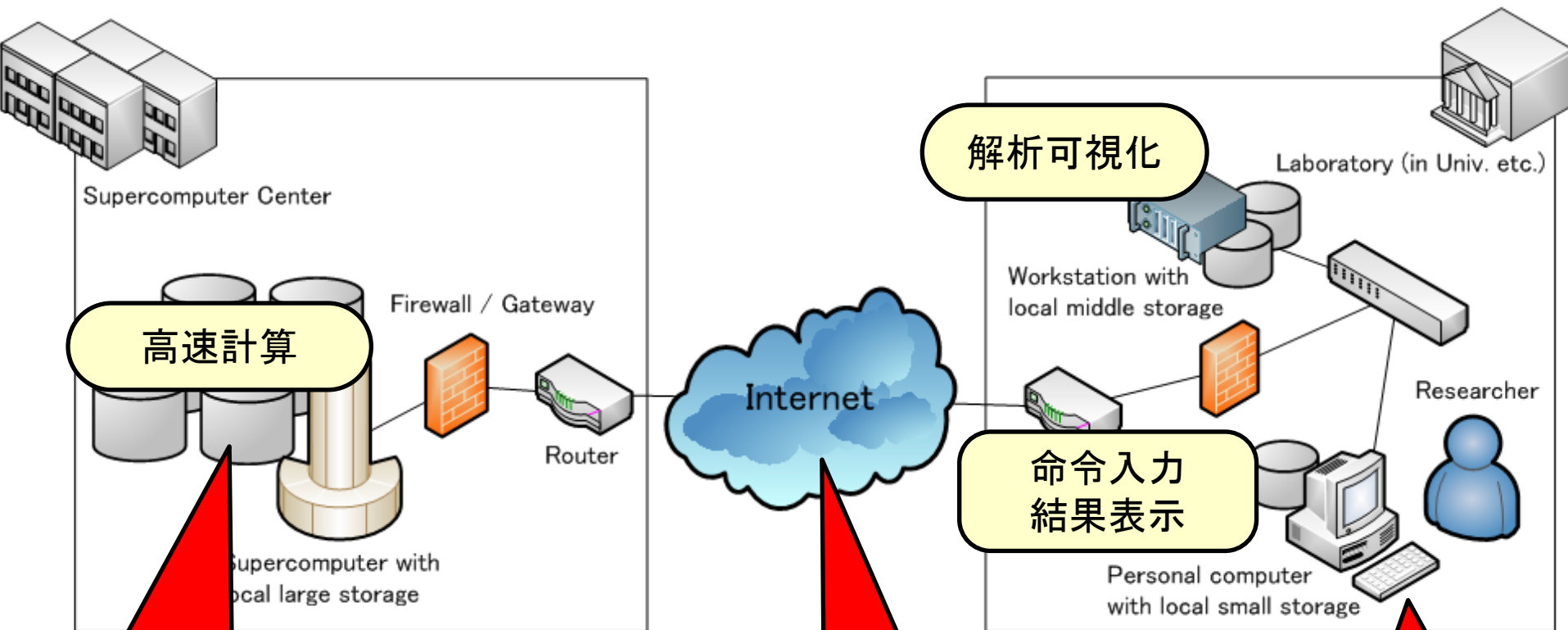
第3回データ科学ワークショップ

2010 年 11 月 25日(木) 17:00-17:30 @筑波大学計算科学研究センター・国際ワークショップ室

NICTサイエンスクラウド(現状・計画)



大規模データを扱う シミュレーション研究における問題



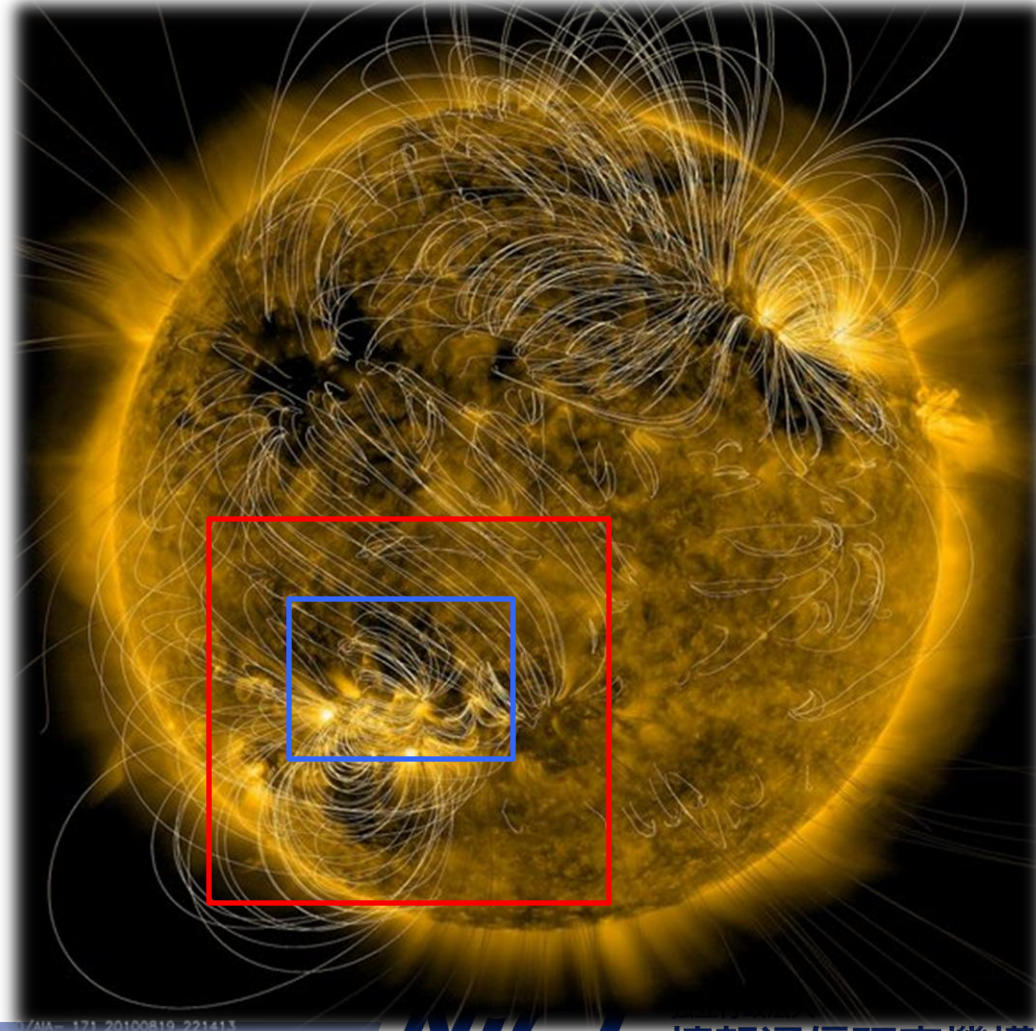
- 汎用的な解析可視化処理に不向き
- 計算速度に比してストレージ容量小 (個人割当: 数TB程度)

高遅延ネットワークを介したファイル転送 (数MB/sec 以下)

数百TB～数PBのストレージ管理の困難

太陽全球精密コロナ磁場の再現と太陽嵐のシミュレーション（井上 他）

- 「SDO」から得られる太陽全面の2次元の観測磁場データから、3次元磁場構造を数値的に外挿
- エネルギー蓄積過程を見るためには、太陽活動領域の部分に 256^3 の格子で、3hrs 程度の間隔のデータが必要。さらにCME（コロナ質量放出）の再現には 512^3 が必要
- 概算でトータル 7.7 TB（320GB/day）のデータがアウトプット



本研究の目標

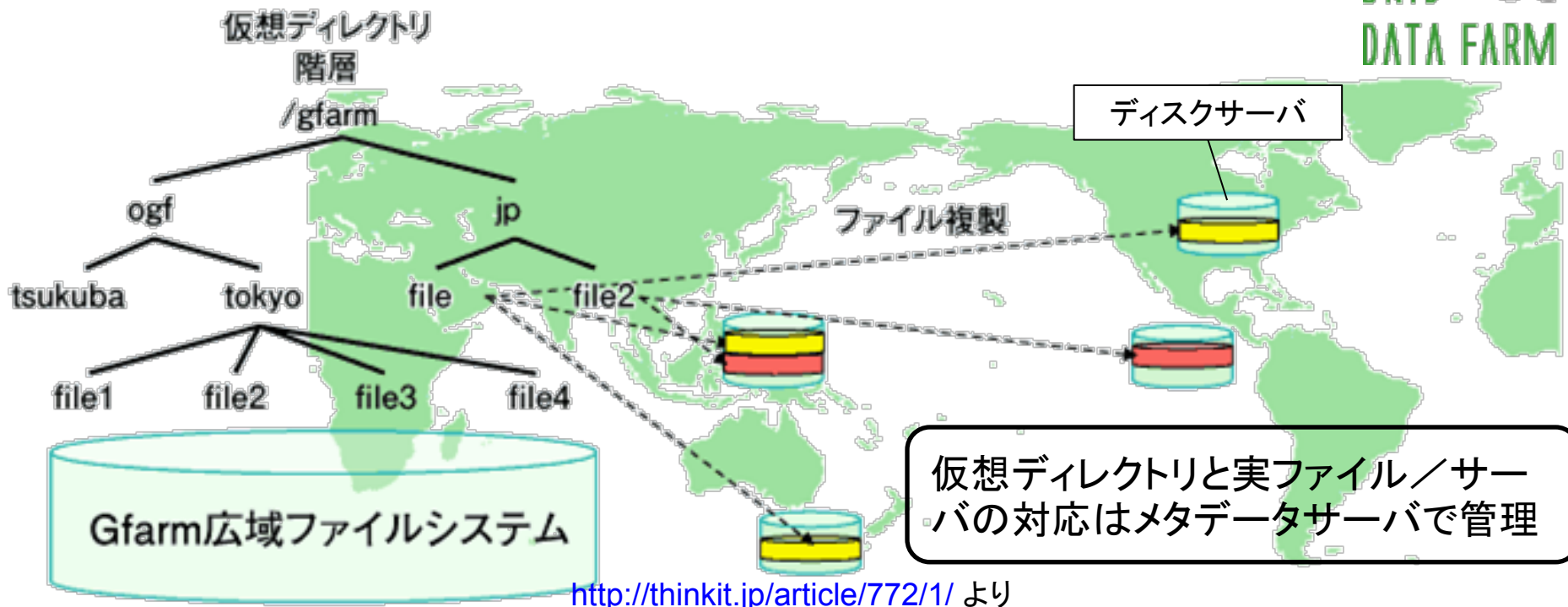
- 大規模ネットワーク分散ストレージの構築
 - 数百TB ～数PetaByte級のディスク容量
 - 従来の数倍～10 倍程度の機材コスト低減
 - 汎用のワークステーション・PCから解析可視化に利用可
- スパコンと解析可視化環境の融合
 - 計算結果を体感的に「即座に扱える」環境に
 - 太陽全球精密コロナ磁場の再現と太陽嵐のシミュレーション（井上 他）
- 利活用する基盤
 - Gfarm: ネットワーク分散ファイルシステム
 - JGN2plus: 全国規模の研究開発テストベッドネットワーク

Gfarm 広域ファイルシステム

(建部 他, 2007)

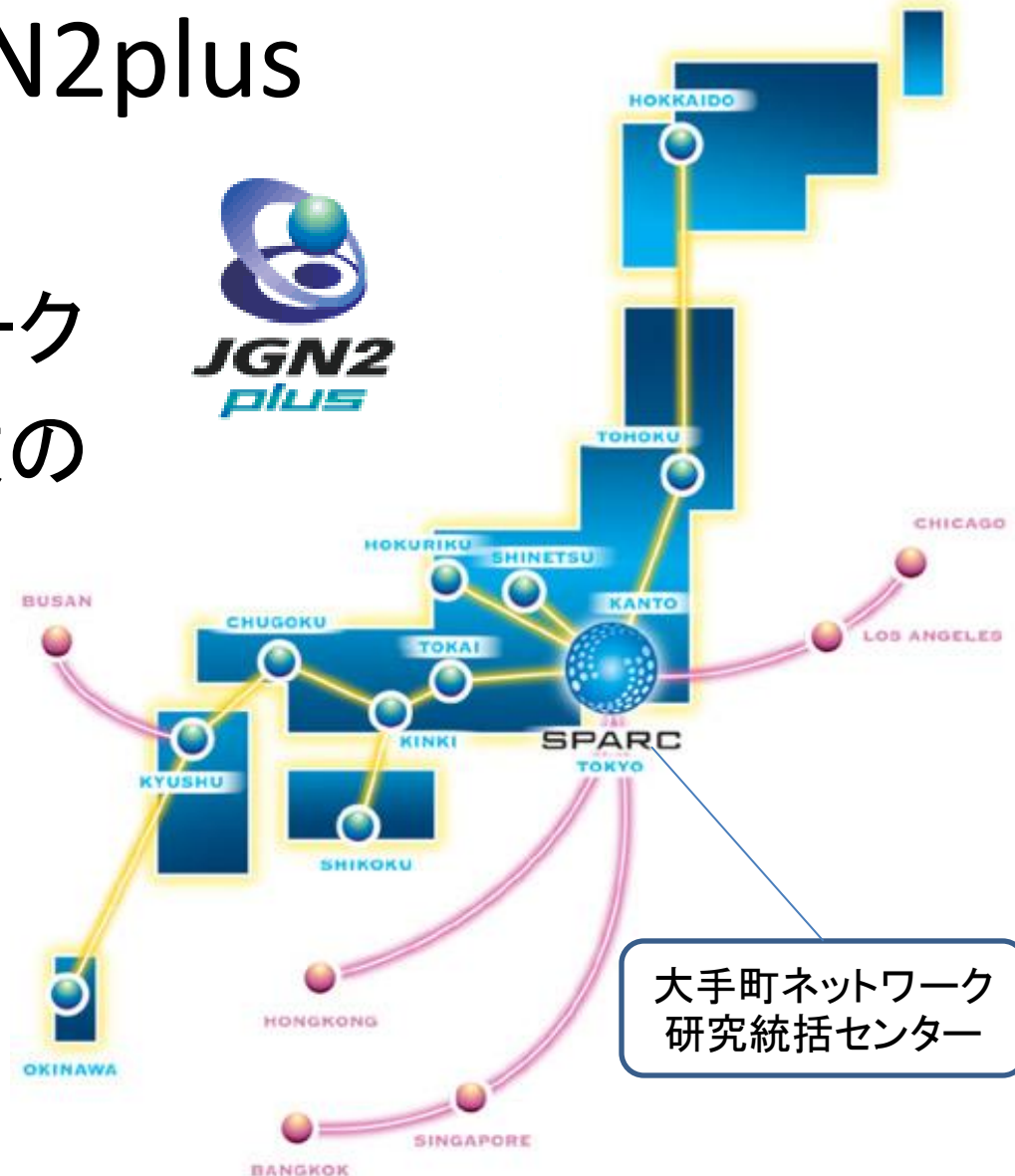
- 高速アクセス可能な広域の共有ファイルシステム
- 利用者はファイルの実際の格納場所を意識せず、仮想ディレクトリ階層へアクセス
- ファイル複製によるデータ冗長化と高速アクセス

GRID
DATA FARM



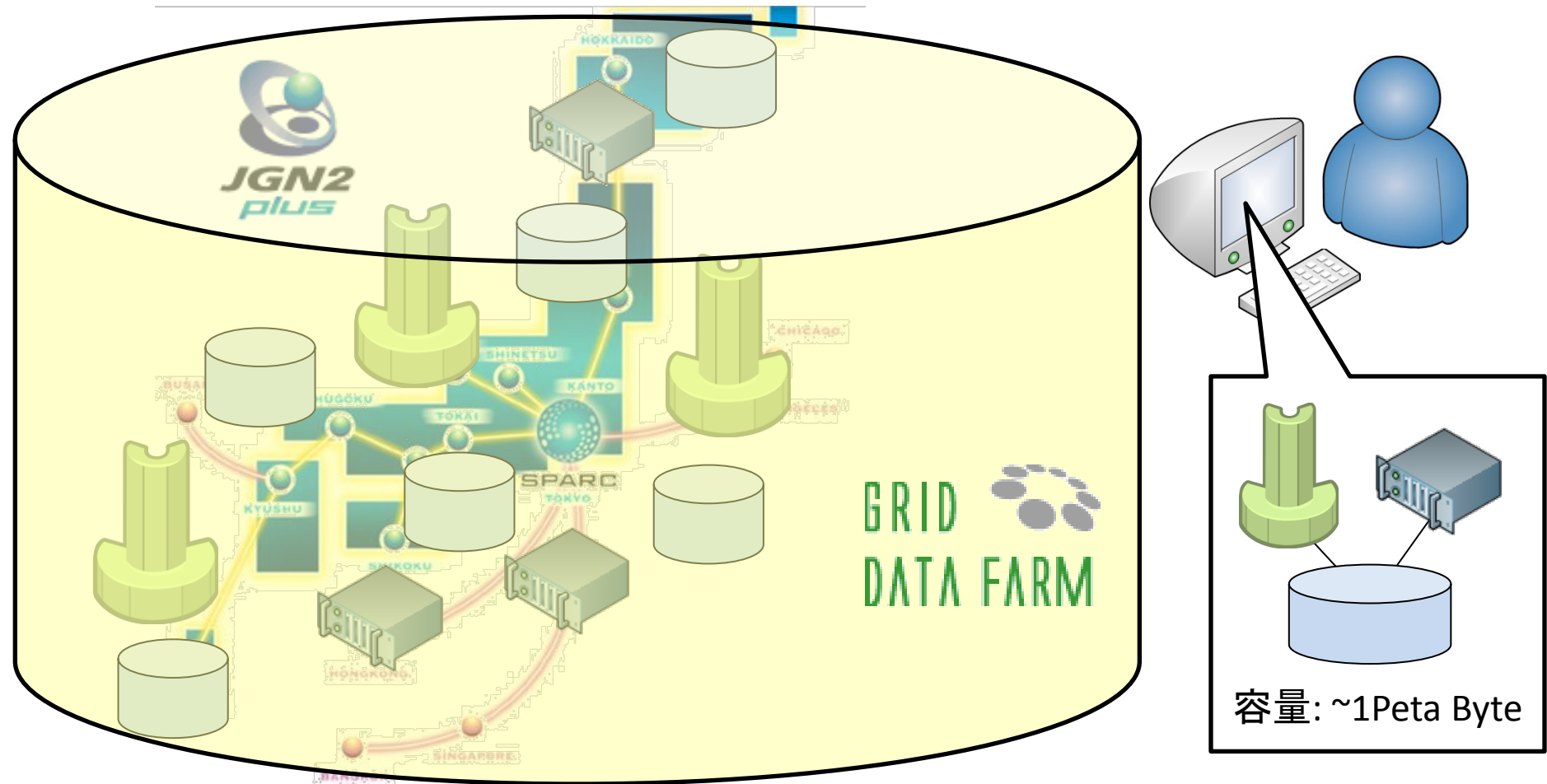
JGN2plus

- NICT による研究開発
テストベッドネットワーク
- NICT の各拠点と複数の
大学間を結ぶ
- ~10Gbps L2/L3
ネットワーク 提供



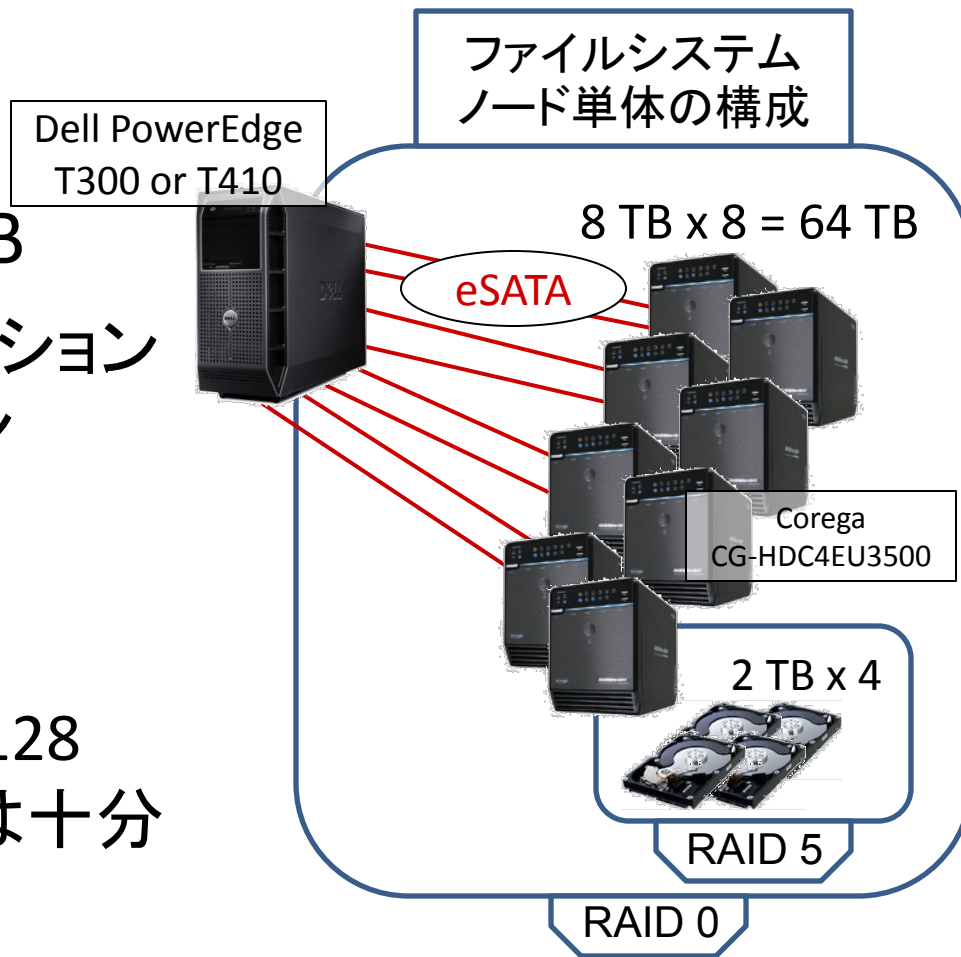
基本コンセプト

- JGN2plus にスパコン、解析ワークステーション、ファイルシステムノードを接続し、Gfarmによって大容量仮想ディスクを構成



廉価サーバ・ディスクによる機材費低減

- 市販の廉価サーバ・ディスクの利用
- 機材費：1～1.5万円/TB
 - 既存のメーカーソリューションの～10倍のコストダウン
- ローカル書込性能139 MB/sec
 - ネットワーク経由(最大128 MB/sec)で利用するには十分



東京・名古屋・大阪・沖縄への ディスクサーバ設置



NTT 小金井本部

NICT
計測技

【ストレージサーバ 2 台分】

- 物理/論理容量: 96 TB / 72 TB

【市販機材で構成】

- サイズ単価: ￥1.5万 / TB
- 書込性能: 139 MB/sec

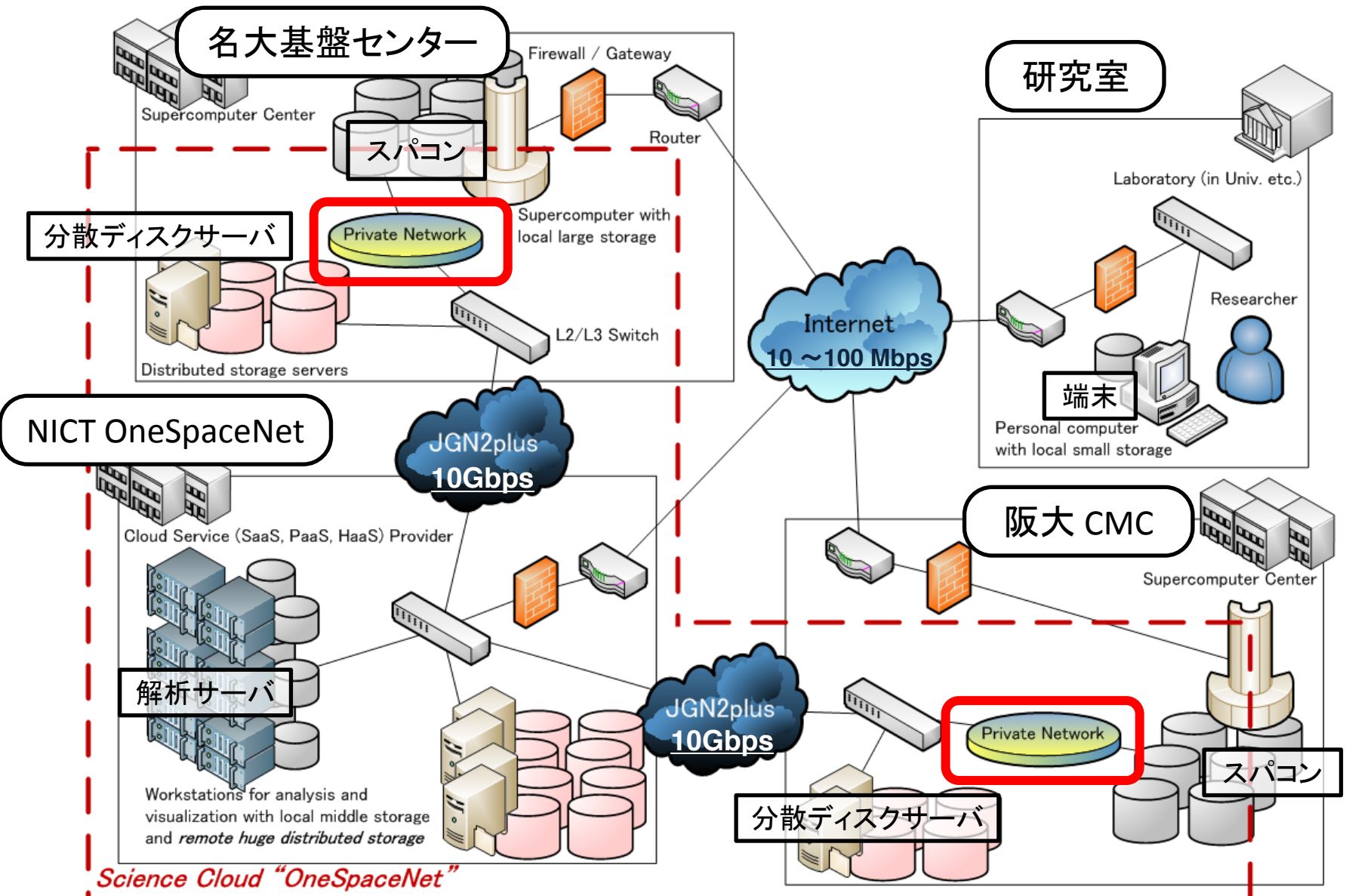
名古屋入子情報基盤センター

NICT

National Institute of Information and Communications Technology

独立行政法人
情報通信研究機構

スパコンと解析可視化環境とを繋ぐ、 ネットワーク・ストレージ設計



利用者(研究者)から見た 仮想研究室 “OneSpaceNet”

計算結果をスパコン付
属ストレージから仮想
ストレージへ自動転送

ジョブの投入、計算実行

“OneSpaceNet”
Supercomputer with
local huge storage

Virtual Huge Storage

gfarm2fs

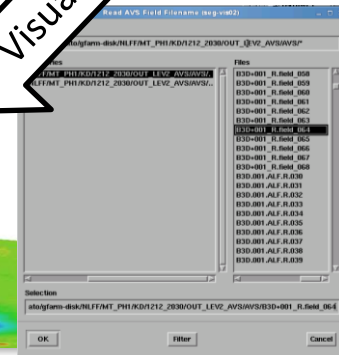
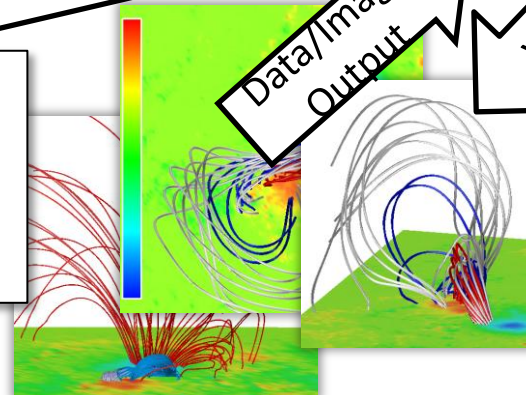
ストレージを「マウント」し
た汎用Linuxサーバから
AVS, IDL等の汎用アプリ
で解析可視化

Workstations for analysis and visualization
with local huge storage

```
morikawa@seg-vis02: /home/morikawa
[inosato@seg-vis02 ~]$ df -h -t fuse
Filesystem      Size  Used Avail Use% Mounted on
fuse            250T  207T   43T   83% /gfarm/inosato
[inosato@seg-vis02 ~]$ ls /gfarm/inosato/home/inosato/NLFF/M
T_PH1/KD/1212_2030/OUT_LEV2_AVS
ALPHA2D_NLFF      XRT_NLFF.031  XRT_NLFF.051
ALPHA_AVE_NLFF    XRT_NLFF.032  XRT_NLFF.052
ANGLE2D_NLFF      XRT_NLFF.033  XRT_NLFF.053
AVS               XRT_NLFF.034  XRT_NLFF.054
B3D_NLFF_1212_2030 XRT_NLFF.035  XRT_NLFF.055
BOTTOM_BX         XRT_NLFF.036  XRT_NLFF.056
BOTTOM_BY         XRT_NLFF.037  XRT_NLFF.057
```

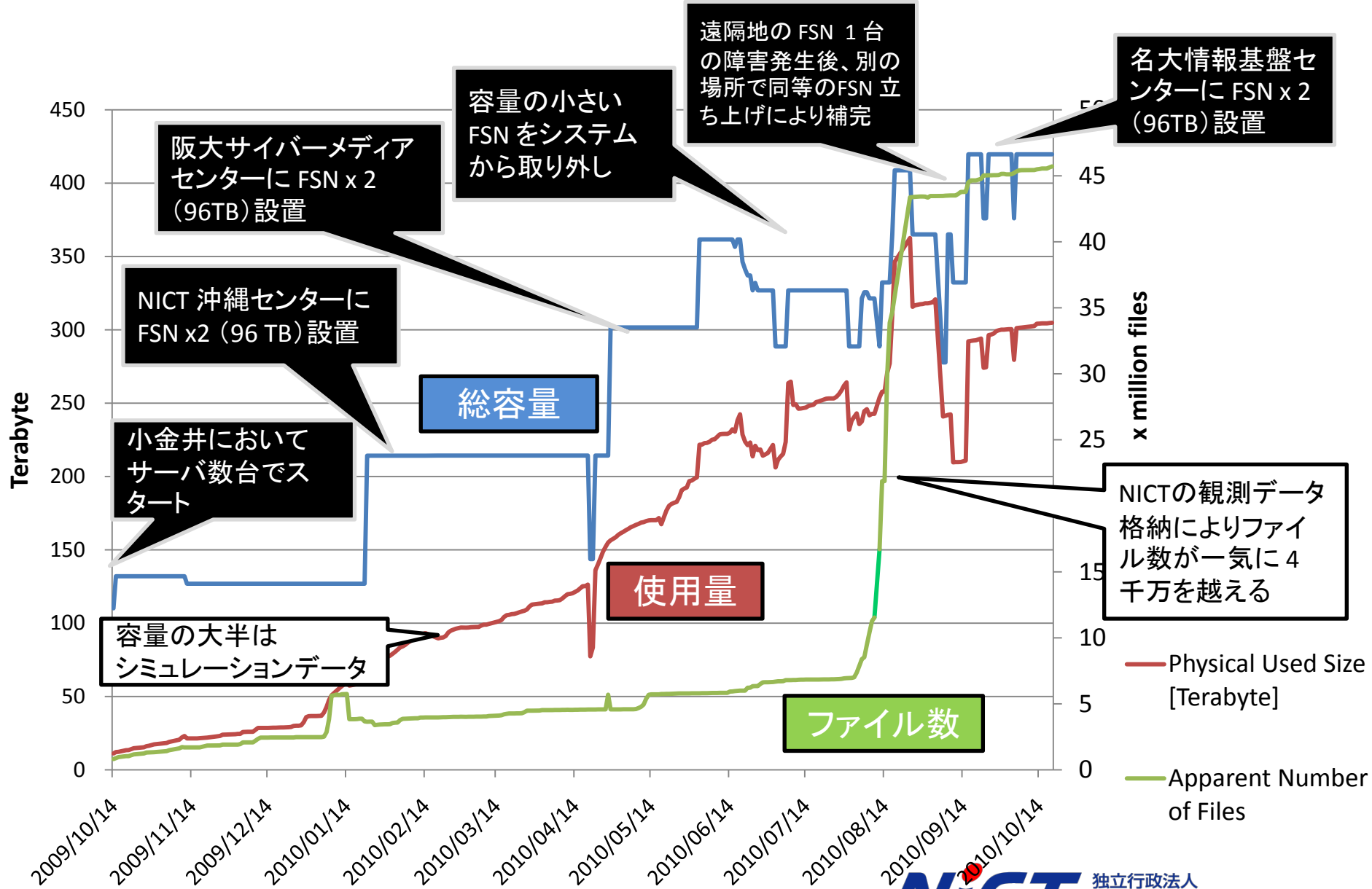
Data/Image
Output

Visualize



Researcher

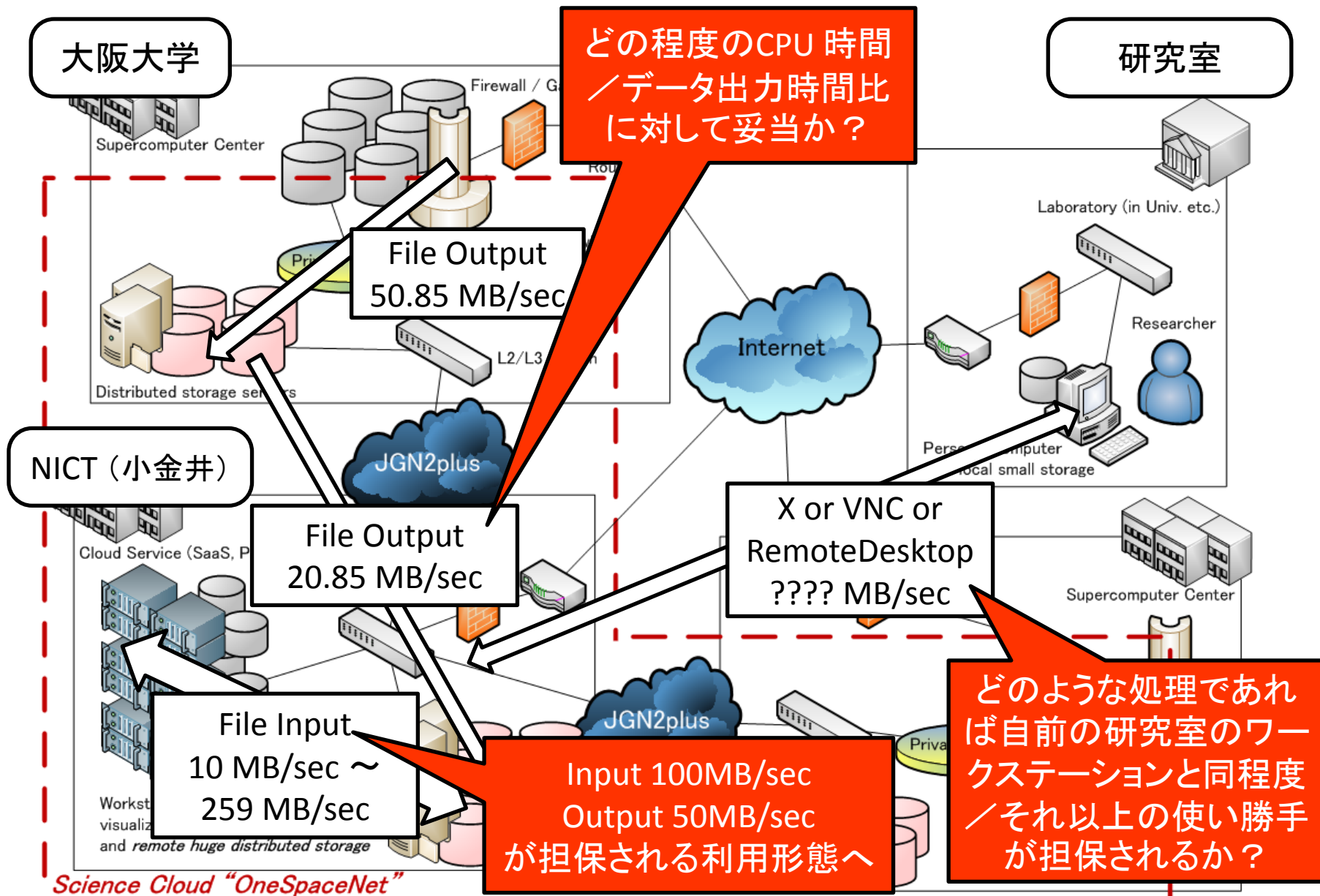
ストレージ使用量・ファイル数推移



利用実績・状況

運用期間	2009/08/14 ~ 2010/11/25					
使用/総容量 (物理容量換算)	305 / 420 TB					
ファイル数	45,711,900					
ユーザ使用量 (上位12ユーザ) ※ 複製分含まず	26T	fukxxxx	6.1T	gyoxxxx	3.5T	muraxxxx
	15T	tsuxxxx	5.8T	shimxxx	1.1T	inoxxxx
	9.3T	jyuxxxx	5.6T	shixxxx	790G	morxxxx
	6.3T	reaxxxx	5.1T	xsfxxxx	574G	hshixxxx
ファイルシステム ノード構成	1K-blocks	Used	Avail	Capacity	Host	
	35161993216	17407682956	17754310260	50%	seg-gfarm-n11.nict.go.jp	
	35161964544	2031048428	33130916116	6%	seg-gfarm-n40.nict.go.jp	
	35161964544	33787267488	1374697056	96%	seg-gfarm-n10.nict.go.jp	
	17580931072	17580931072	0	100%	seg-gfarm-n12.nict.go.jp	
	35161993216	34671799216	490194000	99%	seg-gfarm-n51.nict.go.jp	
	35161993216	32737894748	2424098468	93%	seg-gfarm-n50.nict.go.jp	
	35161993216	32941026628	2220966588	94%	seg-gfarm-n13.nict.go.jp	
	21488629724	21488561252	68472	100%	gst2.nict.go.jp	
	35161993216	32683036176	2478957040	93%	seg-gfarm-n14.nict.go.jp	
	35161964544	2542607172	32619357372	7%	seg-gfarm-n41.nict.go.jp	
	337946351580	245452786208	92493565372	73%		

通信速度の計測と今後の課題



まとめ

- NICT 宇宙天気クラウド “OneSpaceNet” における大規模ストレージ(420TB 程度)の構築と運用
 - JGN2plus により全国に展開
 - Gfarm v2 によりNW分散を意識しない利用形態
 - 汎用品の組合せにより機材コスト低減(1~1.5万円/TB 以下)
 - 今年度中に1PetaByte 近くまで増強の予定
- スパコンと解析可視化環境の融合へむけて
 - スパコンでの計算結果が「即座に」WSで解析可視化可能な環境の実現を目指す
 - 計算結果(NICT、阪大、名大スパコン)
→大規模ストレージへの転送システムの構築
 - 実証フェーズへ
 - 太陽全球精密コロナ磁場の再現と太陽嵐のシミュレーション(井上,他)
 - テストデータで、実測: 1.7 TB/day
 - 想定される 320GB/day はクリアできる見込