NICTサイエンスクラウド:
広域分散型クラウドの基本性能と有効性の検証

村田健史・渡邉英伸（NICT）・鵜川健太郎・村永和哉・鈴木　豊（SEC社）
・黒澤　隆（日立）・木村映善（愛媛大）・建部修見（筑波大）・山本和憲
・久保田康文・笠井康子（NICT）・荻野竜樹（名大）

http://sc-web.nict.go.jp/

NICT ScienceCloud

ビッグデータ：なぜサイエンスでは難しい？

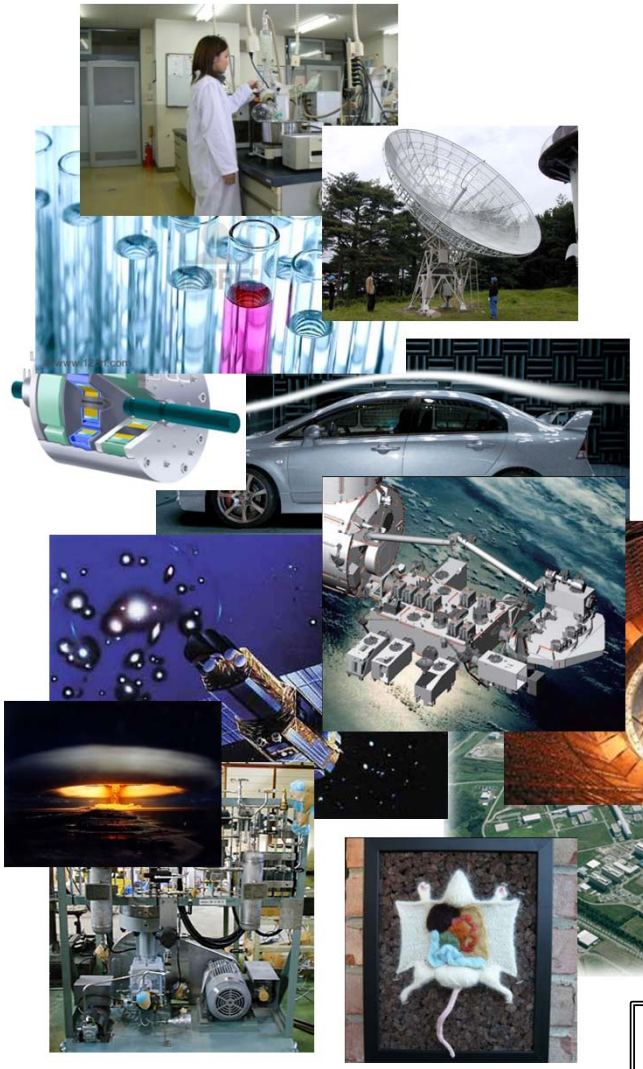
一般・民間・行政

- データ
  - バーバルデータ・分析結果・画像データ…
- 大規模データ
  - より幅広い情報（収集・分析・提供）
- データ処理
  - 機械学習・統計解析・パターンマッチング…
- サービス・アプリケーション
  - より便利なサービス（生活／行政・アミューズメント・医療・経済／ビジネス／金融…）
  - 便利なサービスは普及する

アカデミア・科学研究

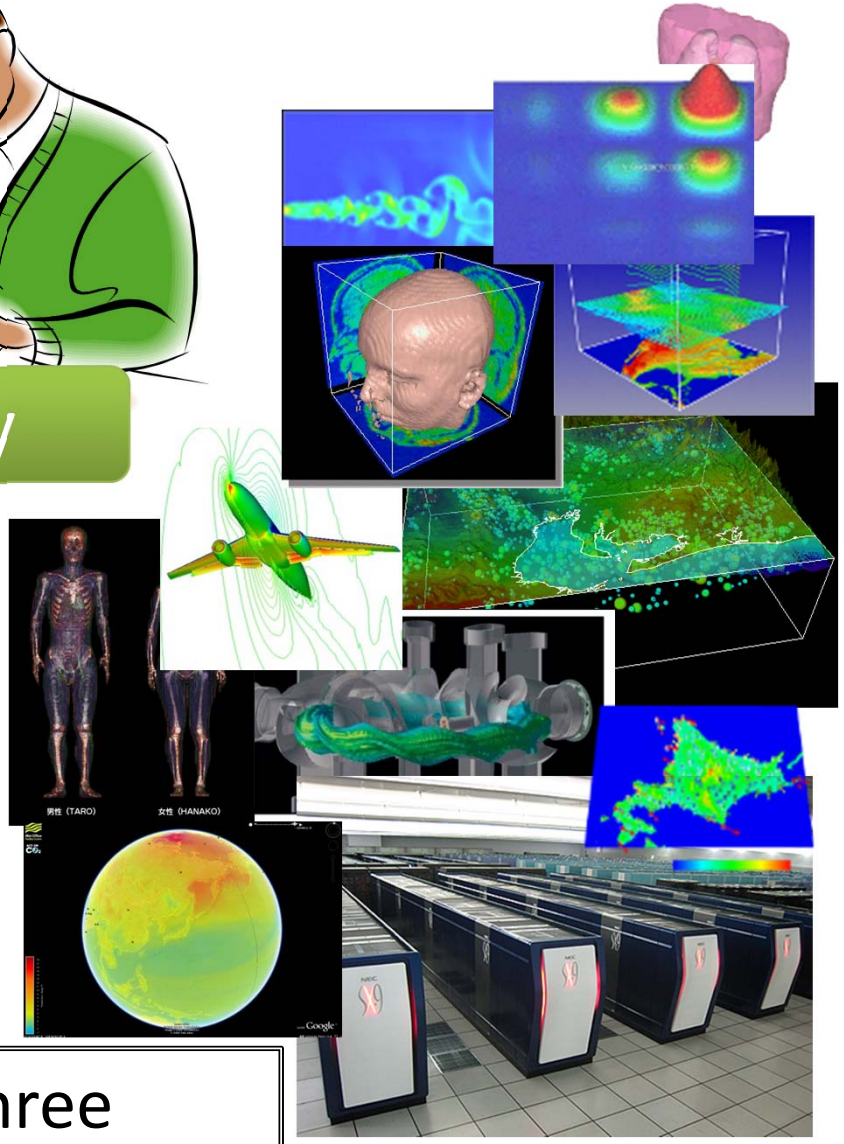
- データ
  - 数値（バイナリ）データが多い
- 大規模データ
  - 実験結果・数値計算結果・公開科学データ・共同研究データ
- データ処理
  - テーマごとに特殊な目的（汎用性低い）
- サービス・アプリケーション
  - 新たな科学的発見が可能なアプリケーション（ビッグデータ処理による）
  - 研究成果が達成できる方法が普及する

Theory



Observation
Experiment

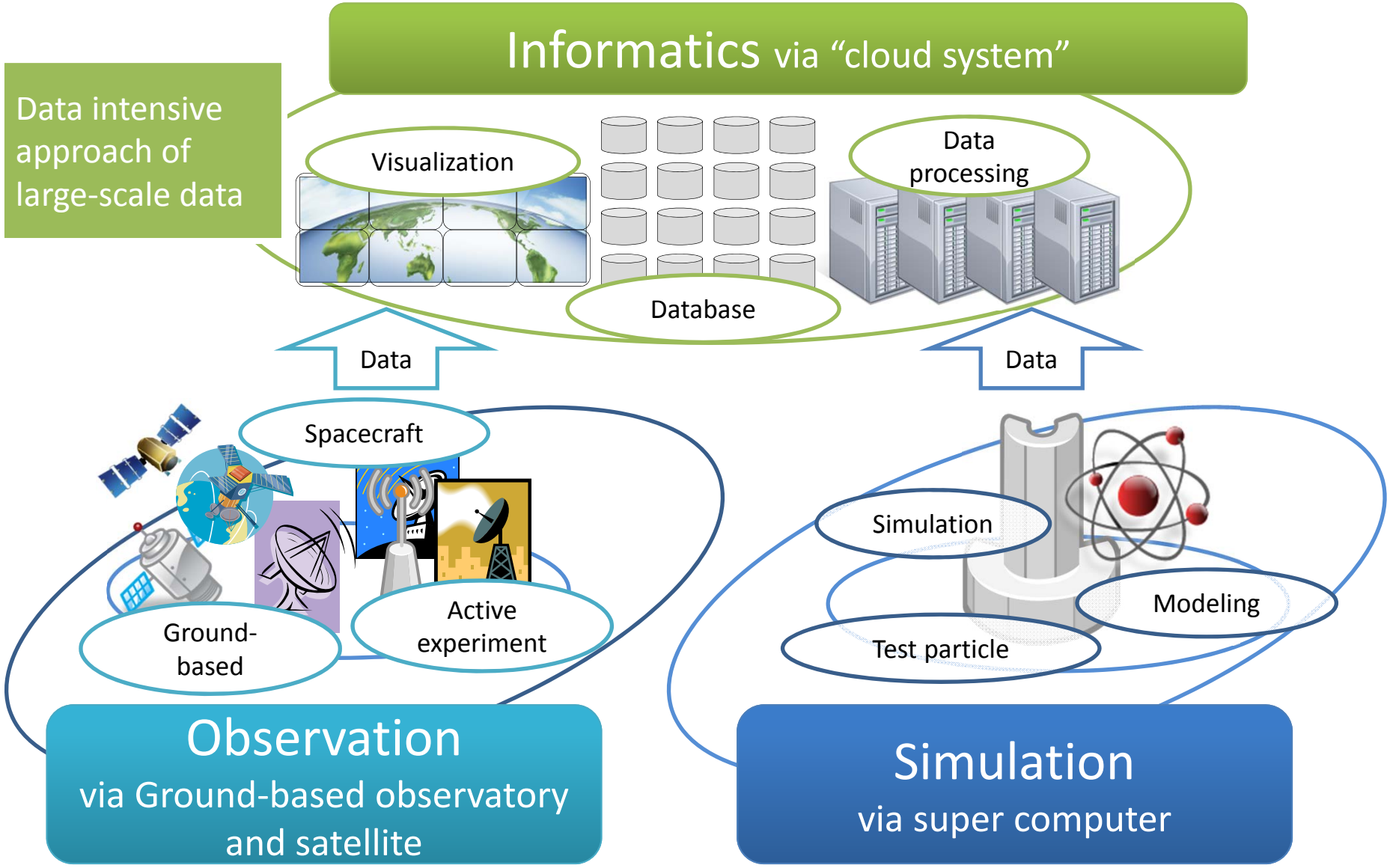
Major Three
Methodologies
For Science



Simulation

NICT ScienceCloud

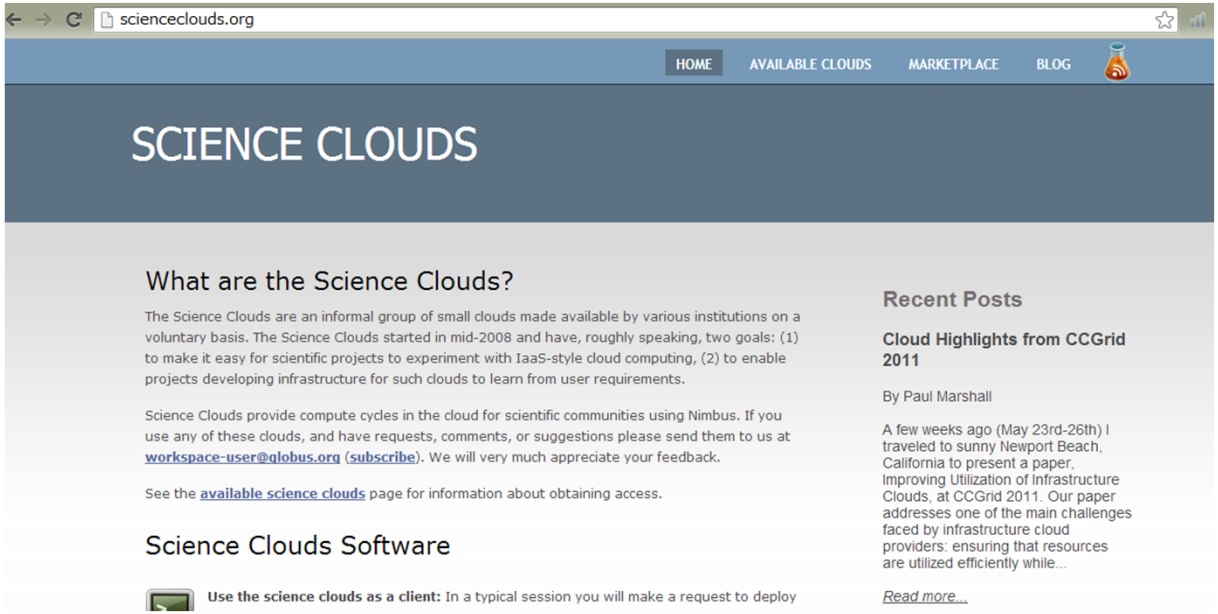
“Science Cloud”; A facility for the 4th methodology



Science Clouds in USA/EU

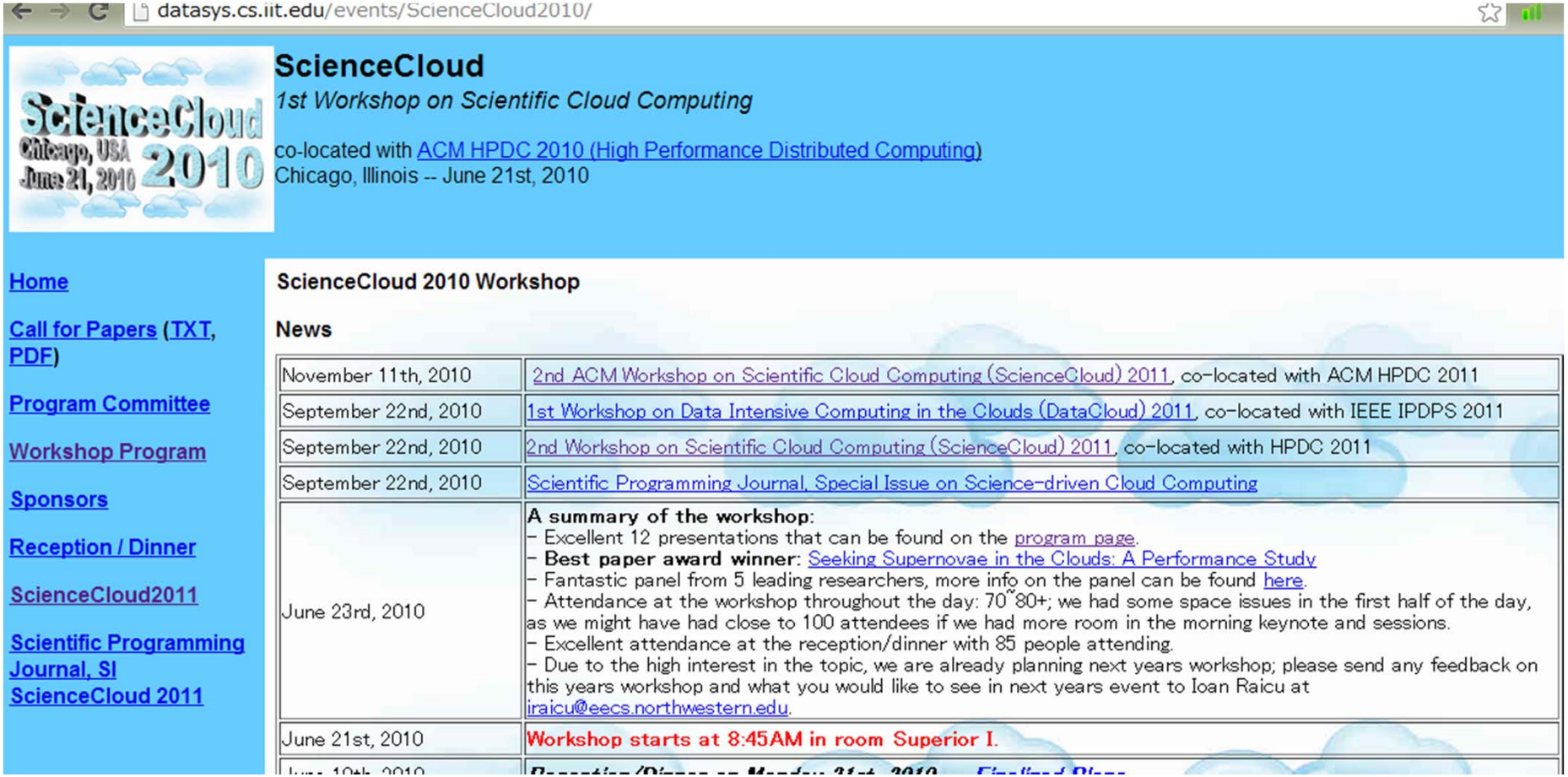
Country/ Organization	Cloud Name/ Group Name	URL	Notes
CERN, CNR-IREA	HELIX NEBULA	http://helix-nebula.eu/	To create a multi-tenant Open Market Place for Science
INRIA Rennes (France)	Centre de recherché commn	http://www.inria.fr/centre/rennes	Scalable Storage for Sharing Application Data on Azure Clouds: TomusBlobs
	The Venus-C project	http://venuscfeedback.codeplex.com/	
	SCIDP-ES	http://www.scidip-es.eu/	Long-term data preservation for Earth Science (Cloud?)
(Lavanya Ramakrishnan)	FRIEDA	http://frieda.lbl.gov/	Flexible Robust Intelligent Elastic Data Management in Cloud Environments
	The Open Science Data Cloud (OSDC)		A petabyte-scale science cloud managed and operated by the Open Cloud Consortium (OCC)
	“Science Clouds”	http://scienceclouds.org/	Science Clouds Blog

Science Cloud International Trend(1)
Illinois University http://scienceclouds.org/



- "Science Cloud" started around 2008.
- The major objectives are:
  - Development of Scientific Experiments via IaaS-type Cloud Environment
  - Development of Infrastructures Based on Users' Requirements

Science Cloud International Trend (2)
First Science Cloud Workshop@Chicago (2010)



datasys.cs.iit.edu/events/ScienceCloud2010/

ScienceCloud

1st Workshop on Scientific Cloud Computing

co-located with ACM HPDC 2010 (High Performance Distributed Computing)
Chicago, Illinois -- June 21st, 2010

- Home
- Call for Papers (TXT, PDF)
- Program Committee
- Workshop Program
- Sponsors
- Reception / Dinner
- ScienceCloud2011
- Scientific Programming Journal, SI ScienceCloud 2011

ScienceCloud 2010 Workshop

News

November 11th, 2010	2nd ACM Workshop on Scientific Cloud Computing (ScienceCloud) 2011, co-located with ACM HPDC 2011
September 22nd, 2010	1st Workshop on Data Intensive Computing in the Clouds (DataCloud) 2011, co-located with IEEE IPDPS 2011
September 22nd, 2010	2nd Workshop on Scientific Cloud Computing (ScienceCloud) 2011, co-located with HPDC 2011
September 22nd, 2010	Scientific Programming Journal, Special Issue on Science-driven Cloud Computing
June 23rd, 2010	**A summary of the workshop:** - Excellent 12 presentations that can be found on the program page. - **Best paper award winner:** Seeking Supernovae in the Clouds: A Performance Study - Fantastic panel from 5 leading researchers, more info on the panel can be found here. - Attendance at the workshop throughout the day: 70~80+; we had some space issues in the first half of the day, as we might have had close to 100 attendees if we had more room in the morning keynote and sessions. - Excellent attendance at the reception/dinner with 85 people attending. - Due to the high interest in the topic, we are already planning next years workshop; please send any feedback on this years workshop and what you would like to see in next years event to Ioan Raicu at iraicu@eecs.northwestern.edu.
June 21st, 2010	**Workshop starts at 8:45AM in room Superior I.**

7

Science Cloud International Trend (3)
The third Science Cloud Workshop@Chicago (2012)

The fourth Science Cloud Workshop @NY/USA (2013)

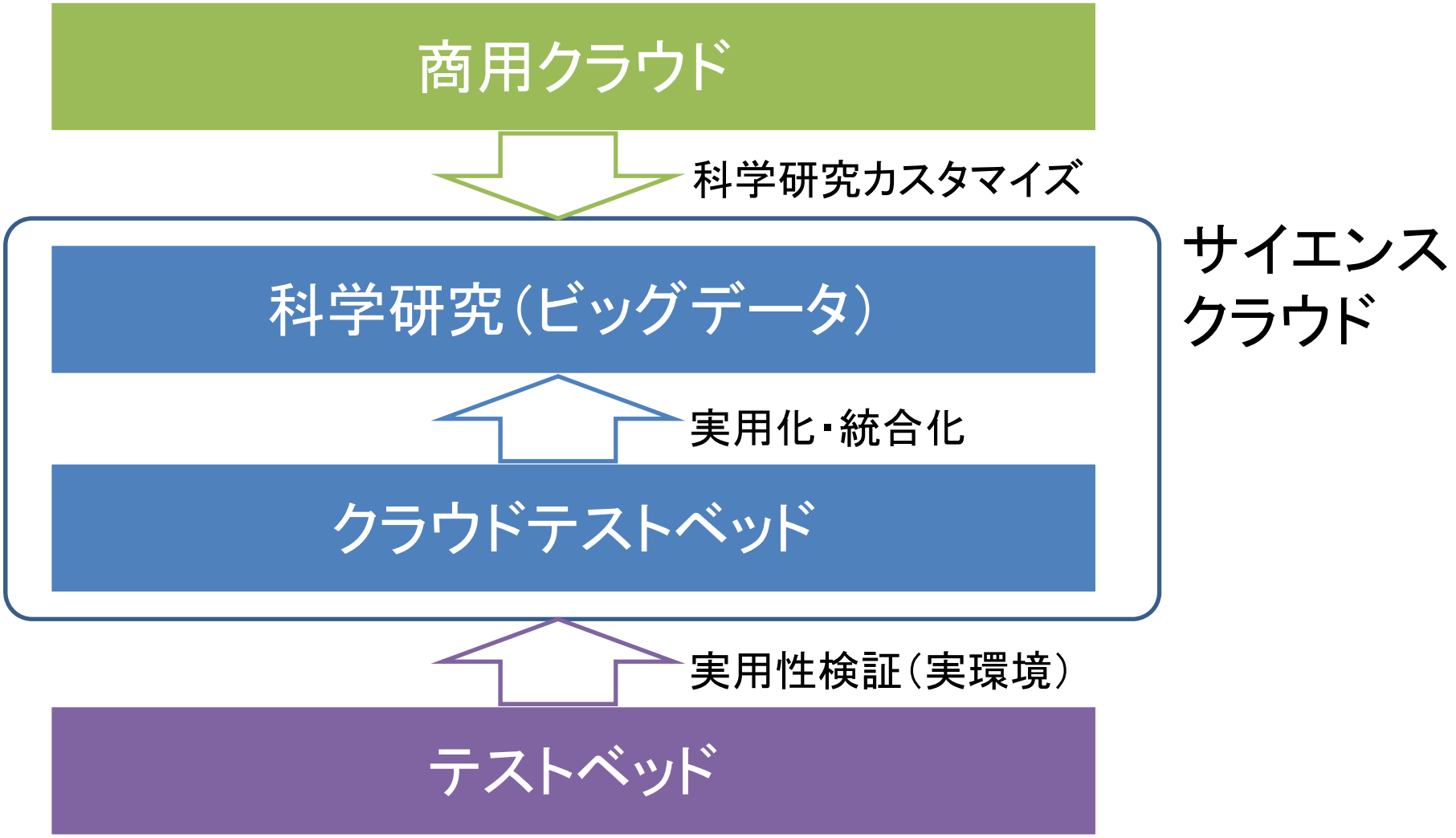
Science Cloud International Trend (4)
Topics of Interest（The third Workshop）

- Scientific application cases studies on cloud infrastructure
- Performance evaluation of cloud environments and technologies
- Fault tolerance and reliability in cloud system
- Data-intensive workloads and tool optimizations
- Use of programming models such as Map-Reduce and its implementations
- Storage cloud architectures
- I/O and Data management in the cloud
- Workflow and resource management in the cloud
- Use of cloud technologies (e.g. NoSQL databases, etc) for scientific applications
- Data streaming and dynamic computations in clouds
- Application of cloud concepts in HPC environments
- High performance parallel file systems and interconnects in virtual environments
- Research and best practices in Cloud security

サイエンスクラウドは商用クラウド・民間クラウドと何が違うのか？

サイエンスクラウドをビッグデータ・データ指向型科学研究にどう活用するか？

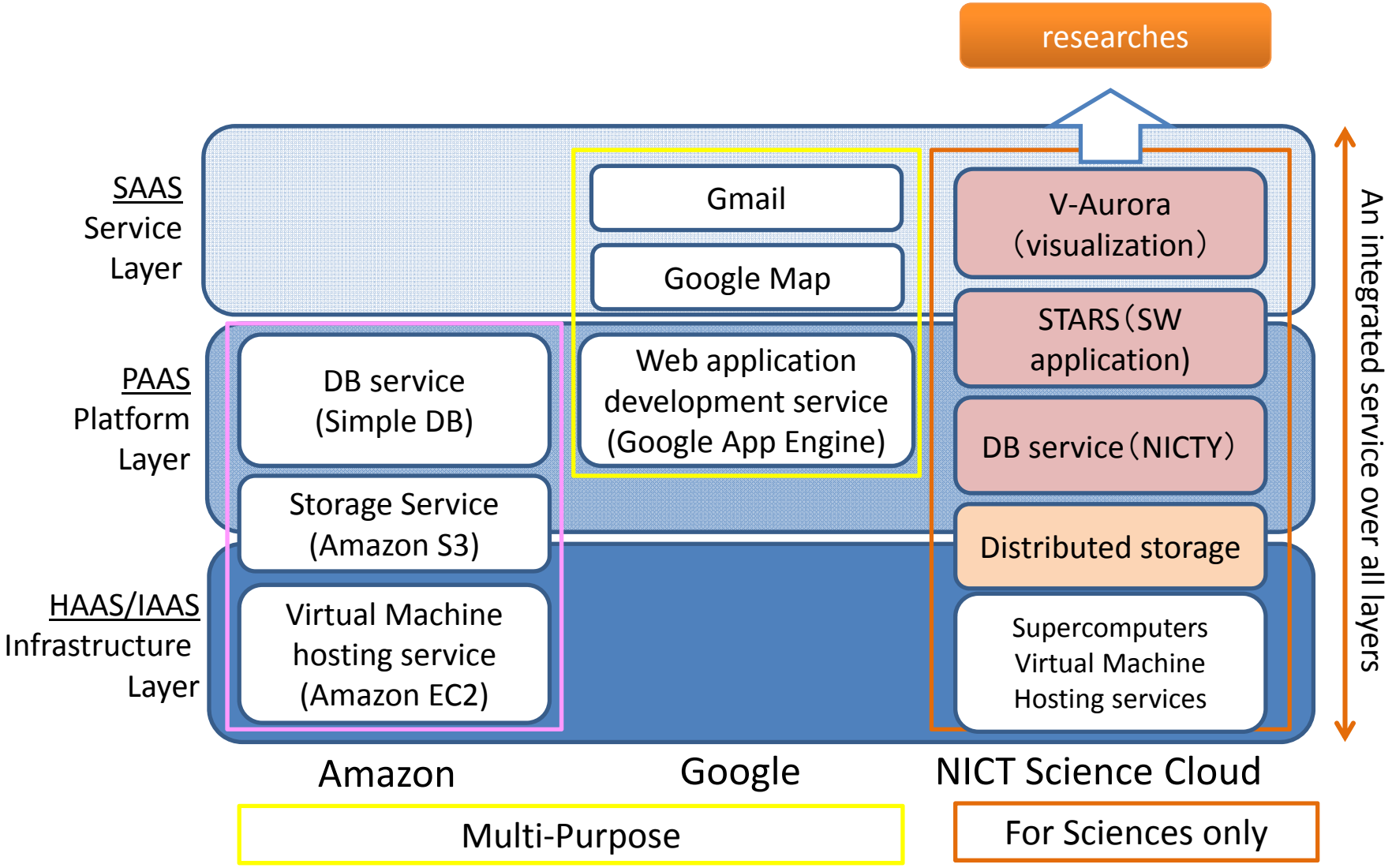
サイエンスクラウドの意味



NICTサイエンスクラウドの特長

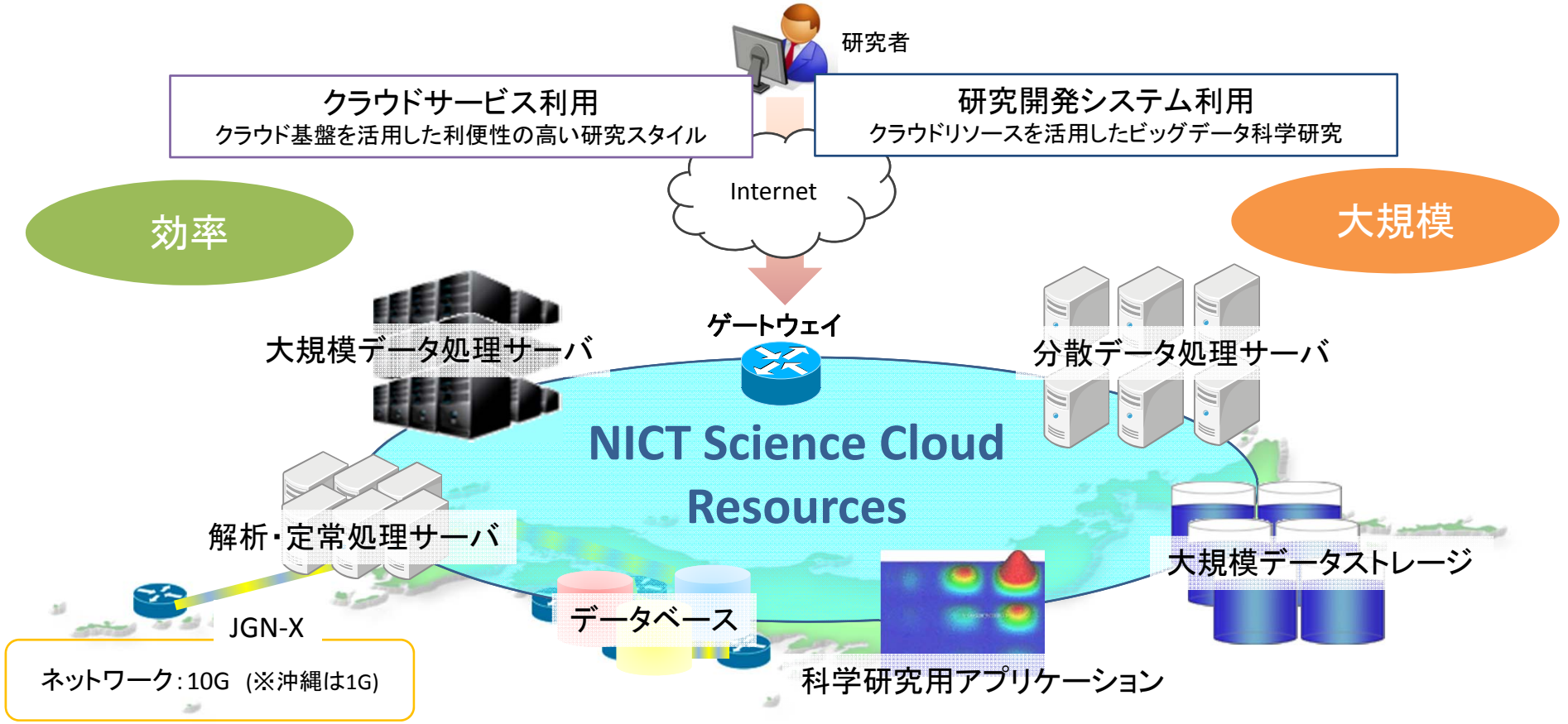
	民間クラウド	NICTサイエンスクラウド	独自研究環境
計算機環境	▲仮想化環境の提供が主体	◎科学研究に特化した様々な計算機環境	◎科学研究に特化した計算機環境
ストレージ環境	▲規模により課金（一般には数TB程度）	◎研究テーマに応じて配分（数100TBも可能）	○一般には100TB程度まで
運用	◎クラウド業者が運用	◎NICTサイエンスクラウドが運用	▲独自運用
セキュリティー対策	◎安心・安全な環境を提供	◎安心・安全な環境を提供（NICT規定に準ずる）	▲独自にセキュリティー対策
研究環境	▲さまざまな利用制約	○共同研究的利用	◎自由な環境を構築

NICT Science Cloud and commercial clouds



NICT ScienceCloud

サイエンスクラウド利活用概要



解析・定常処理サーバ
【台数】5台
【スペック】Opteron/ Xeon（8～12コア）、メモリ32GB～128GB
【保存場所】東京（京都にマイグレーション）

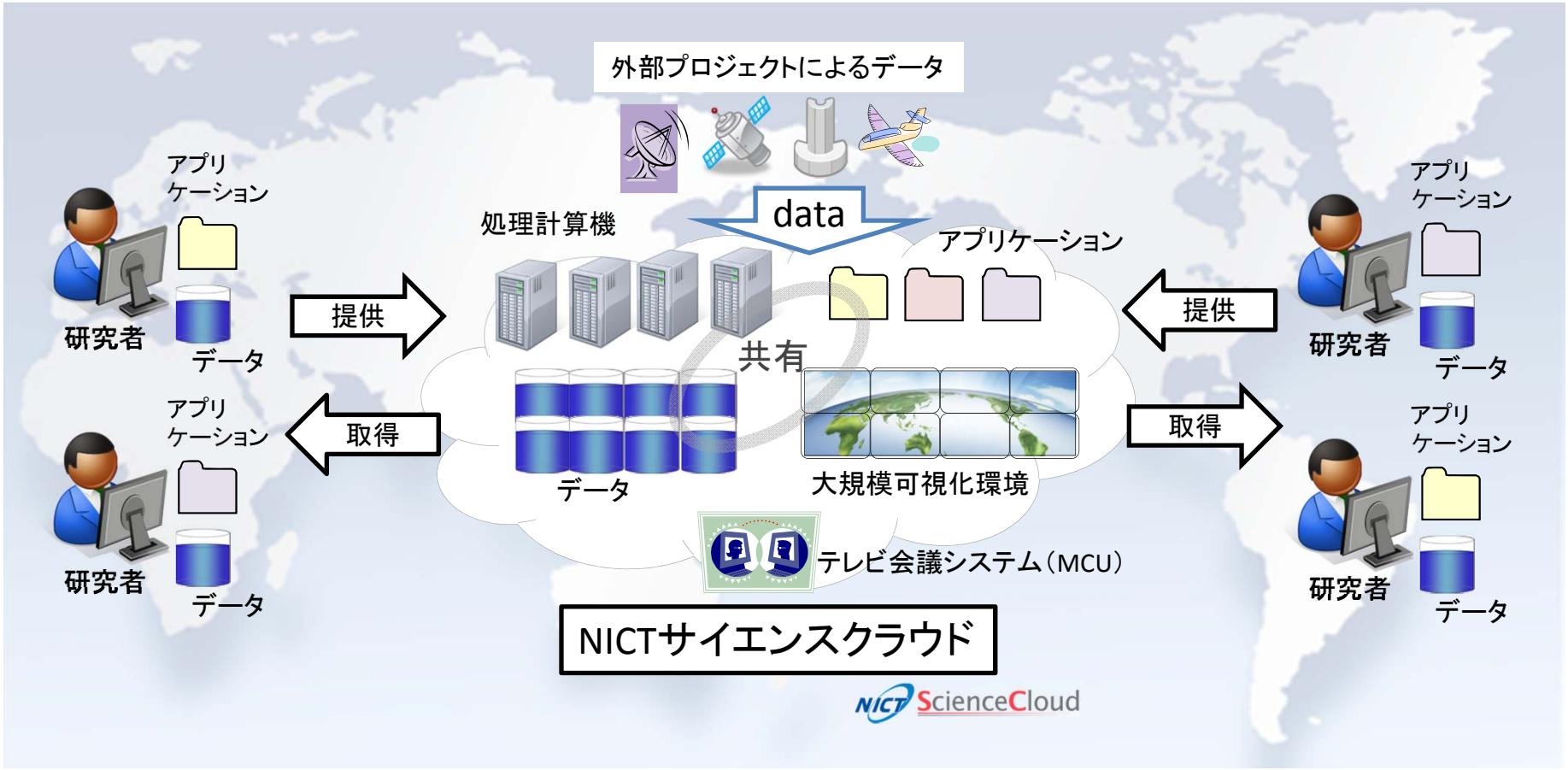
データ保存ストレージ（NAS）
【総容量】150TB
【保存場所】東京（京都にミラーリング）

大規模データ処理サーバ*1
【CPU】合計307コア*2、1.83TFLOPS*3
【RAM】平均71GB/node
【HDD】1.4PB

分散データ処理サーバ*1
【総容量】1.8PB*4
【ディスク規格】SATA3*5、300MB/s（read）*3
【保存場所】東京、京都、沖縄

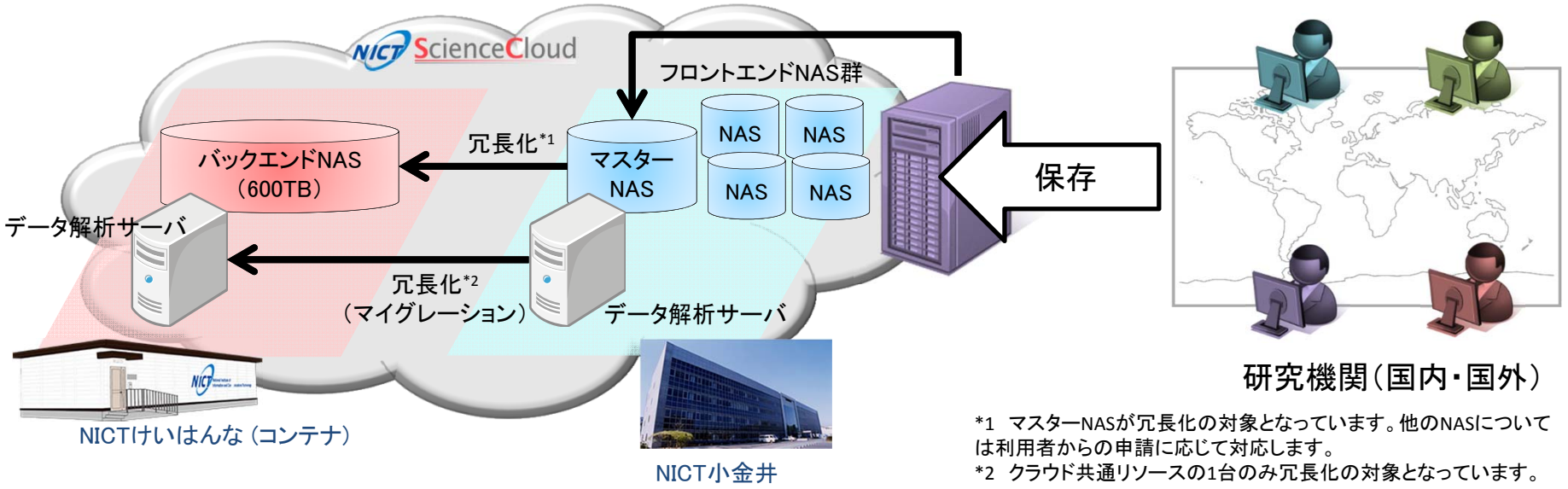
仮想ラボラトリーによる協調的研究開発

- NICTサイエンスクラウド上に、研究者個人や研究グループの協調的研究環境（バーチャルラボ）を構築し、目的に沿った研究を進めることができます。
- 国内外に分散した研究者が、計算機リソースやデータ、プログラムを提供・交換・共有・再利用・協調的解析する場合などに有効です。
- 共有するプログラム、データを研究者サイトにダウンロードして解析することもできます。

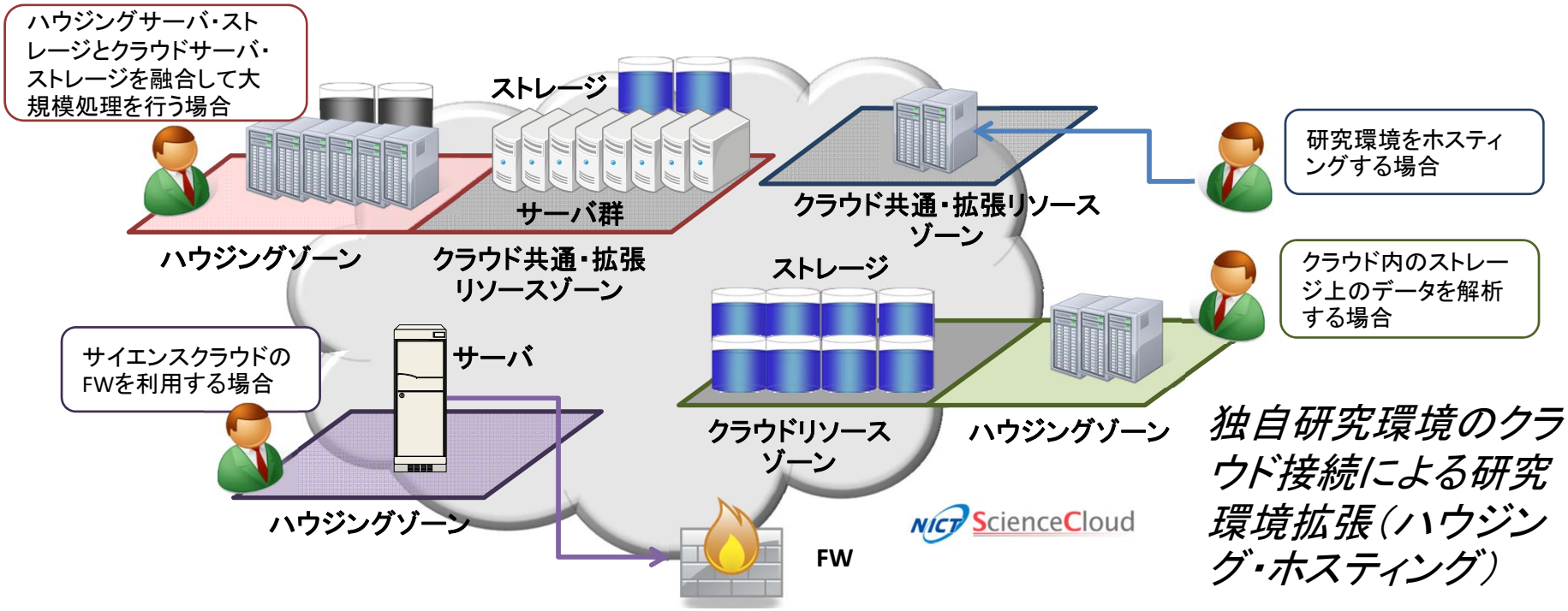


NICT ScienceCloud

冗長化環境の利活用

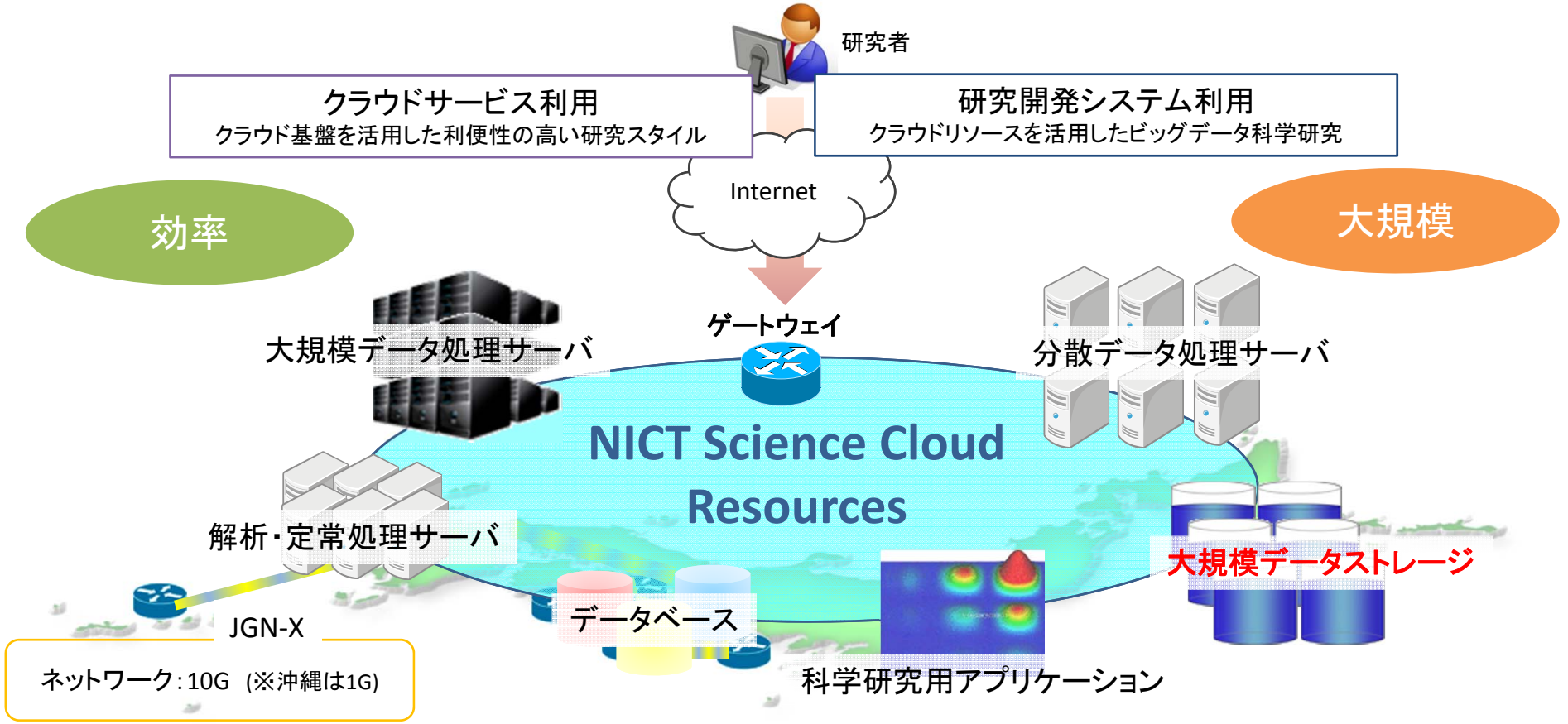


*1 マスターNASが冗長化の対象となっています。他のNASについては利用者からの申請に応じて対応します。
*2 クラウド共通リソースの1台のみ冗長化の対象となっています。



独自研究環境のクラウド接続による研究環境拡張（ハウジング・ホスティング）

サイエンスクラウド利活用概要



解析・定常処理サーバ
【台数】5台
【スペック】Opteron/ Xeon（8～12コア）、メモリ32GB～128GB
【保存場所】東京（京都にマイグレーション）

データ保存ストレージ（NAS）
【総容量】150TB
【保存場所】東京（京都にミラーリング）

大規模データ処理サーバ*1
【CPU】合計307コア*2、1.83TFLOPS*3
【RAM】平均71GB/node
【HDD】1.4PB

分散データ処理サーバ*1
【総容量】1.8PB*4
【ディスク規格】SATA3*5、300MB/s（read）*3
【保存場所】東京、京都、沖縄

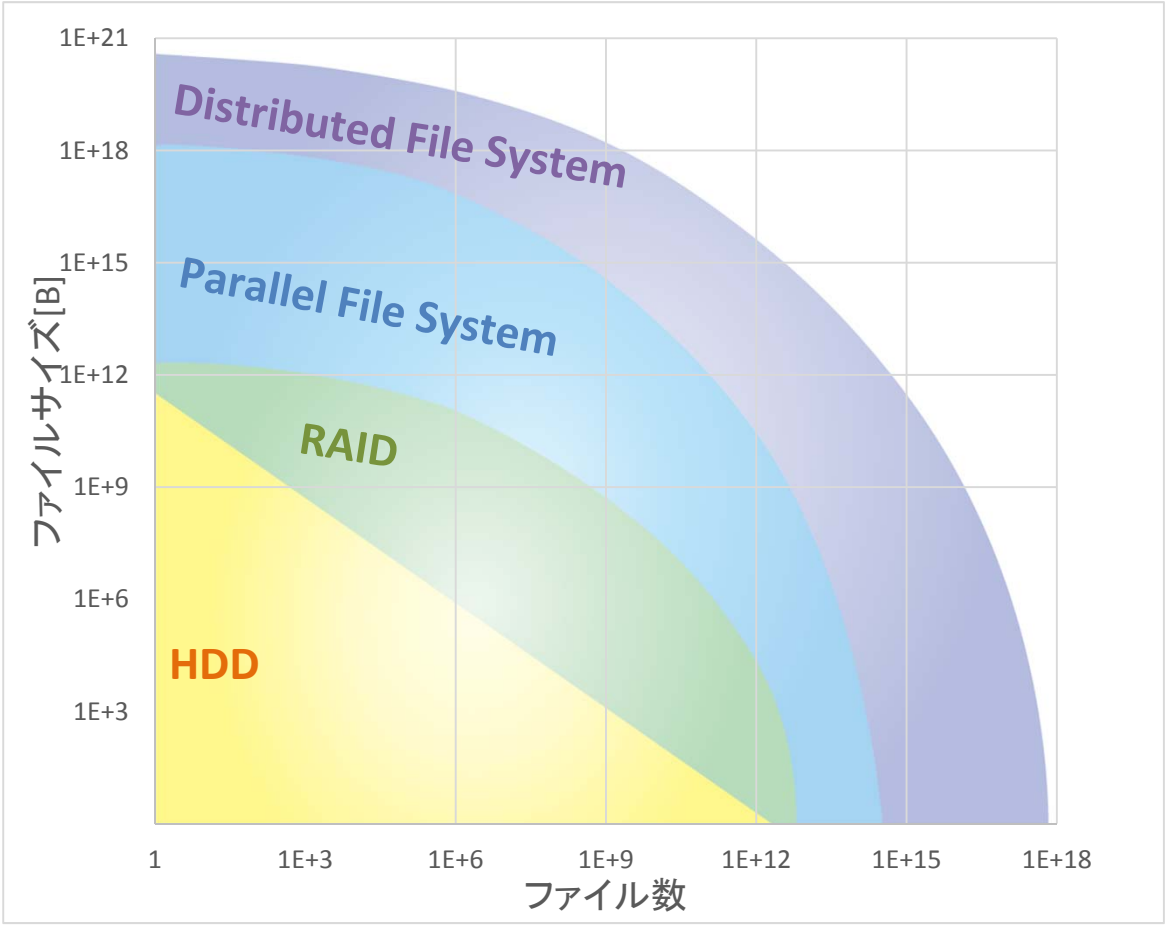
データ格納技術におけるファイルサイズとファイル数の関係

分散ファイルシステム

GoogleFS（Google）
Hadoop FS（Yahoo!）
Gfarm（筑波大）

並列ファイルシステム

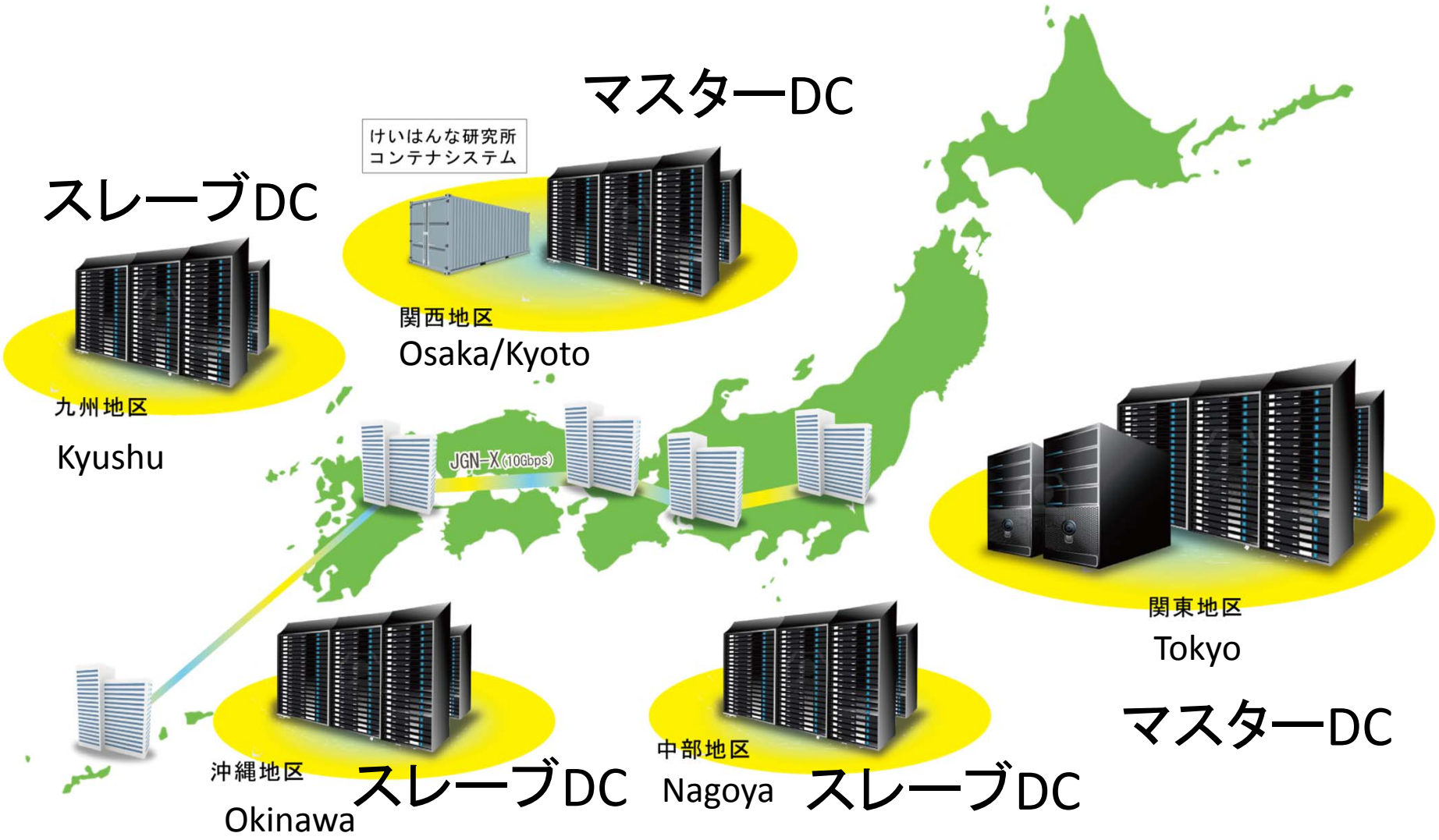
LusterFS（Sun Microsystems）
⇒（例）TSUBAME（1.5PB/60GB/s）
GPFS（IBM）
PVFS2（OSS）



小西史一、情報処理、pp.845-852、2009-09-15から転載

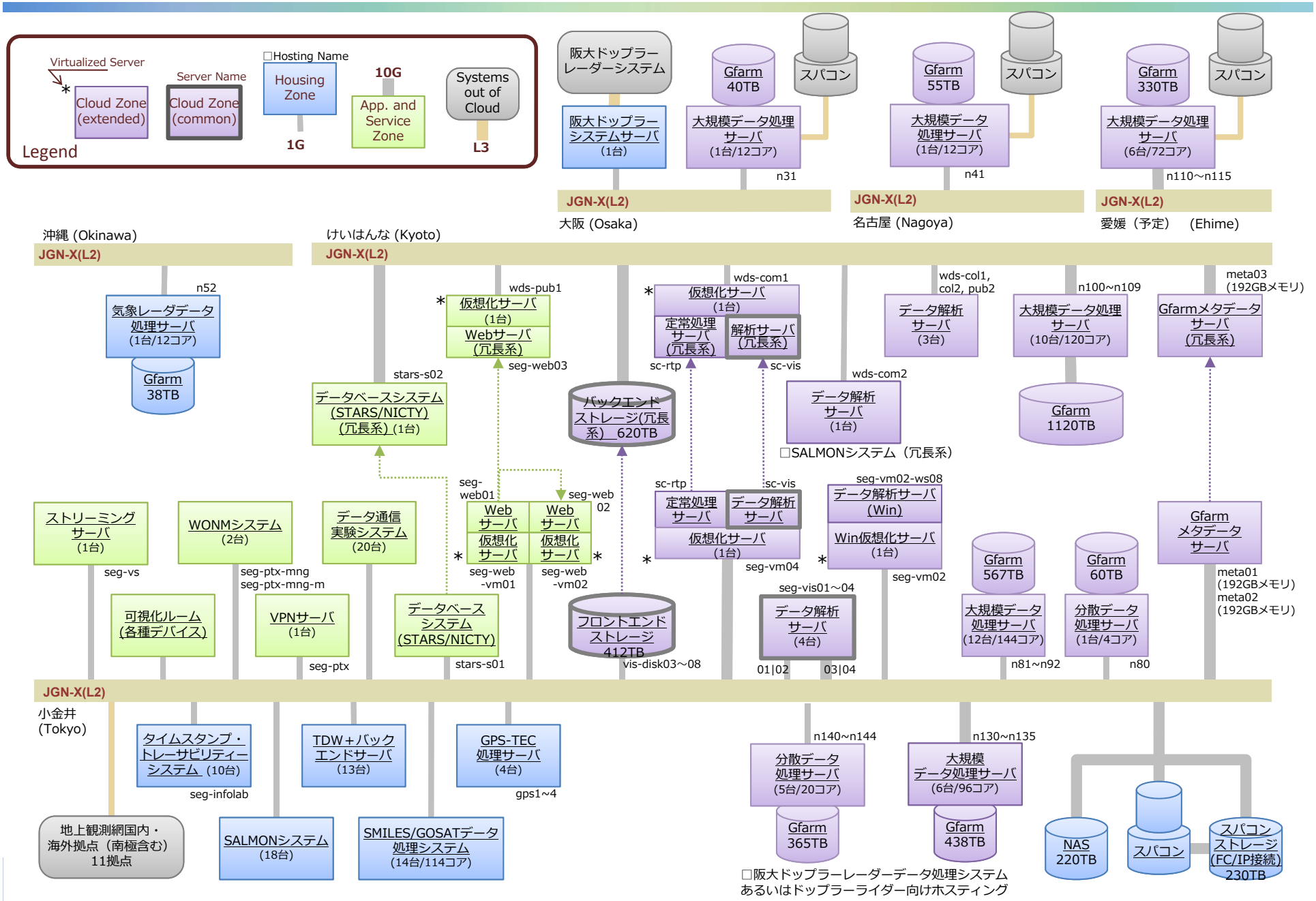
NICT ScienceCloud

NICT Science Cloud (Distributed System)



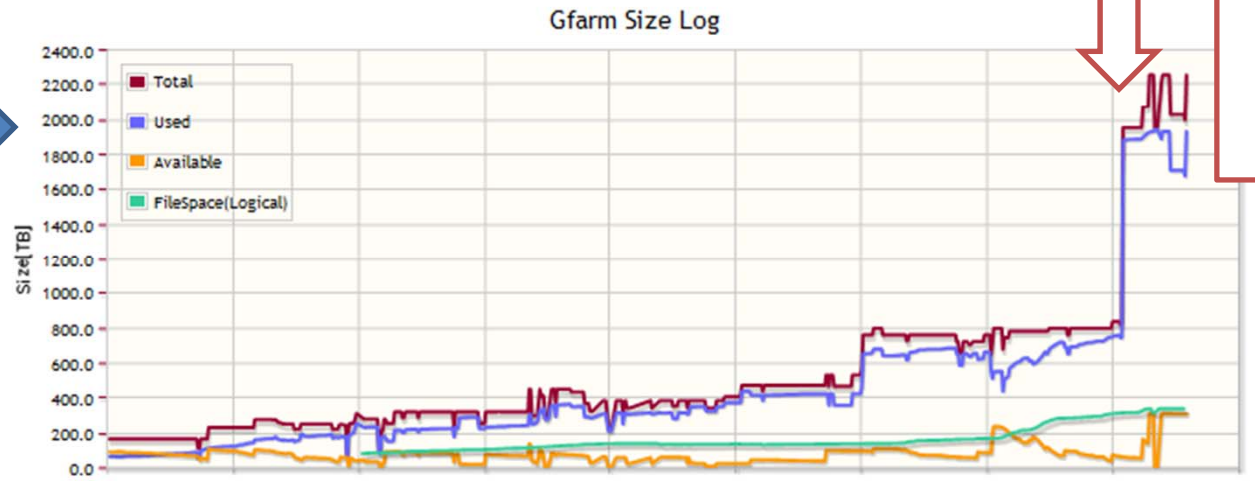
NICT ScienceCloud

System Configuration and Deployment of NICT Science Cloud (in Japanese) 2013



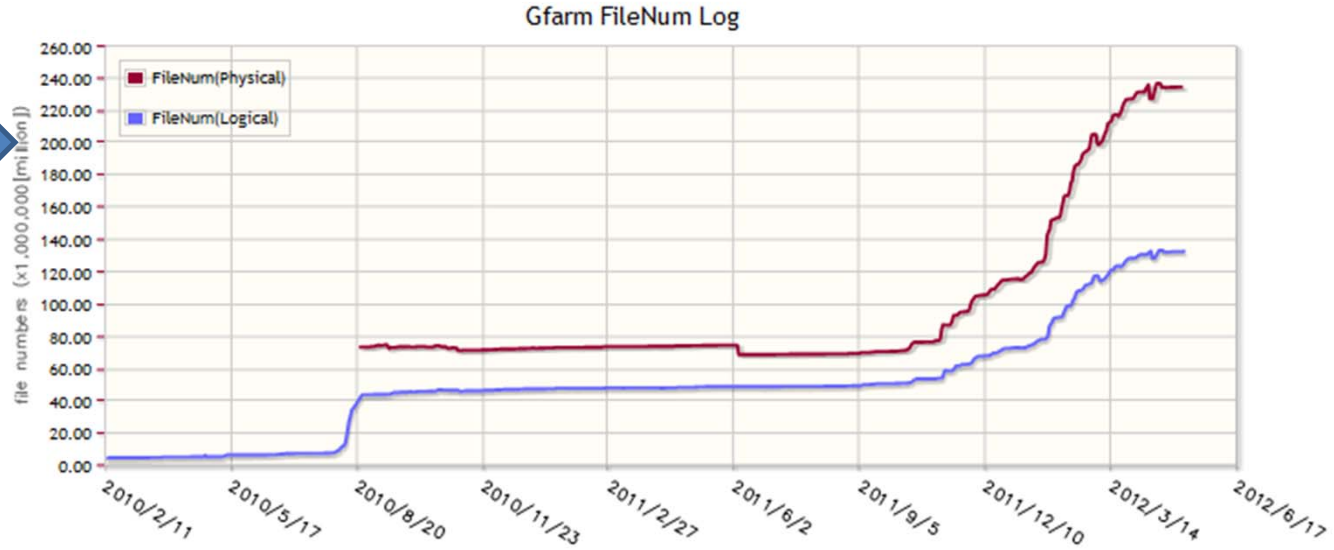
NICT Science Cloud Storage

2PB



NICT Container System

200M files



NICT ScienceCloud

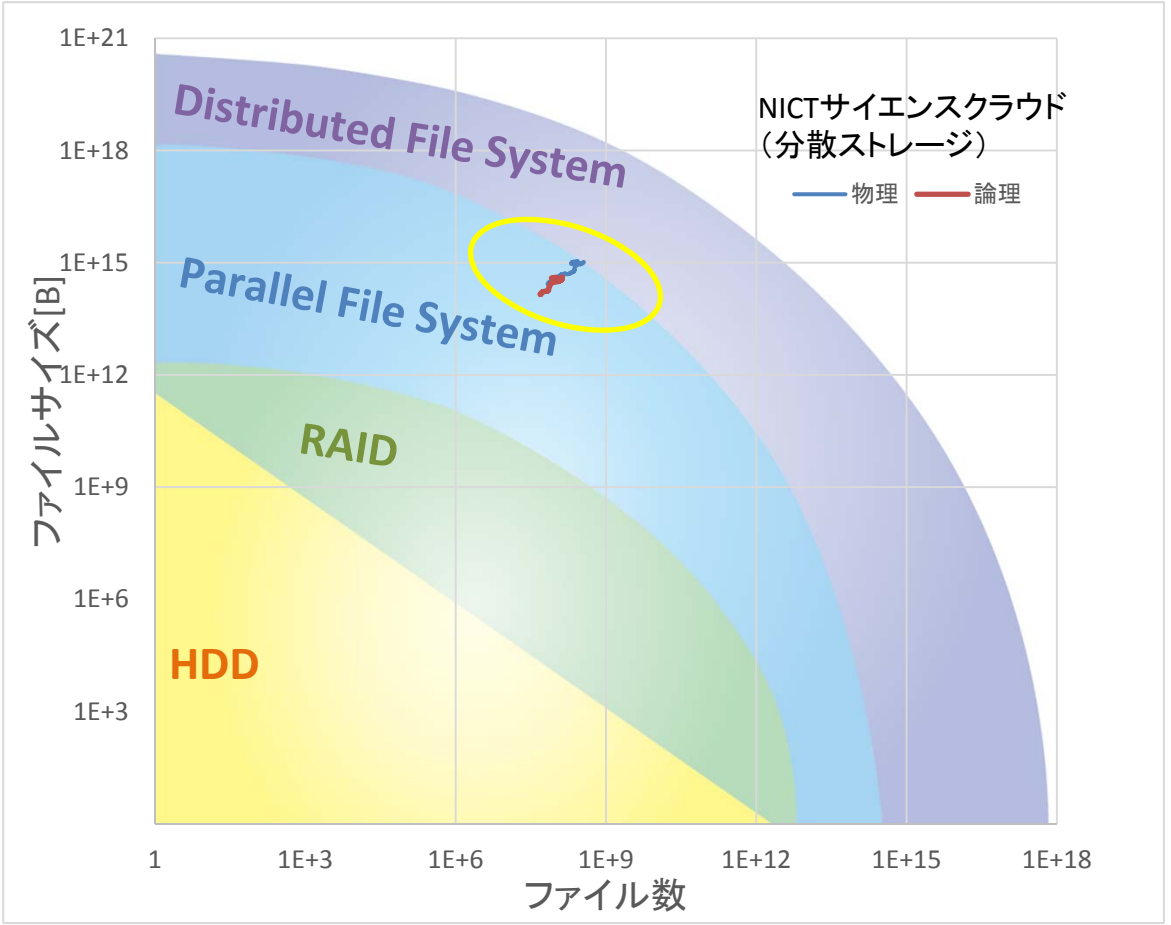
データ格納技術におけるファイルサイズとファイル数の関係

分散ファイルシステム

GoogleFS（Google）
Hadoop FS（Yahoo!）
Gfarm（筑波大）

並列ファイルシステム

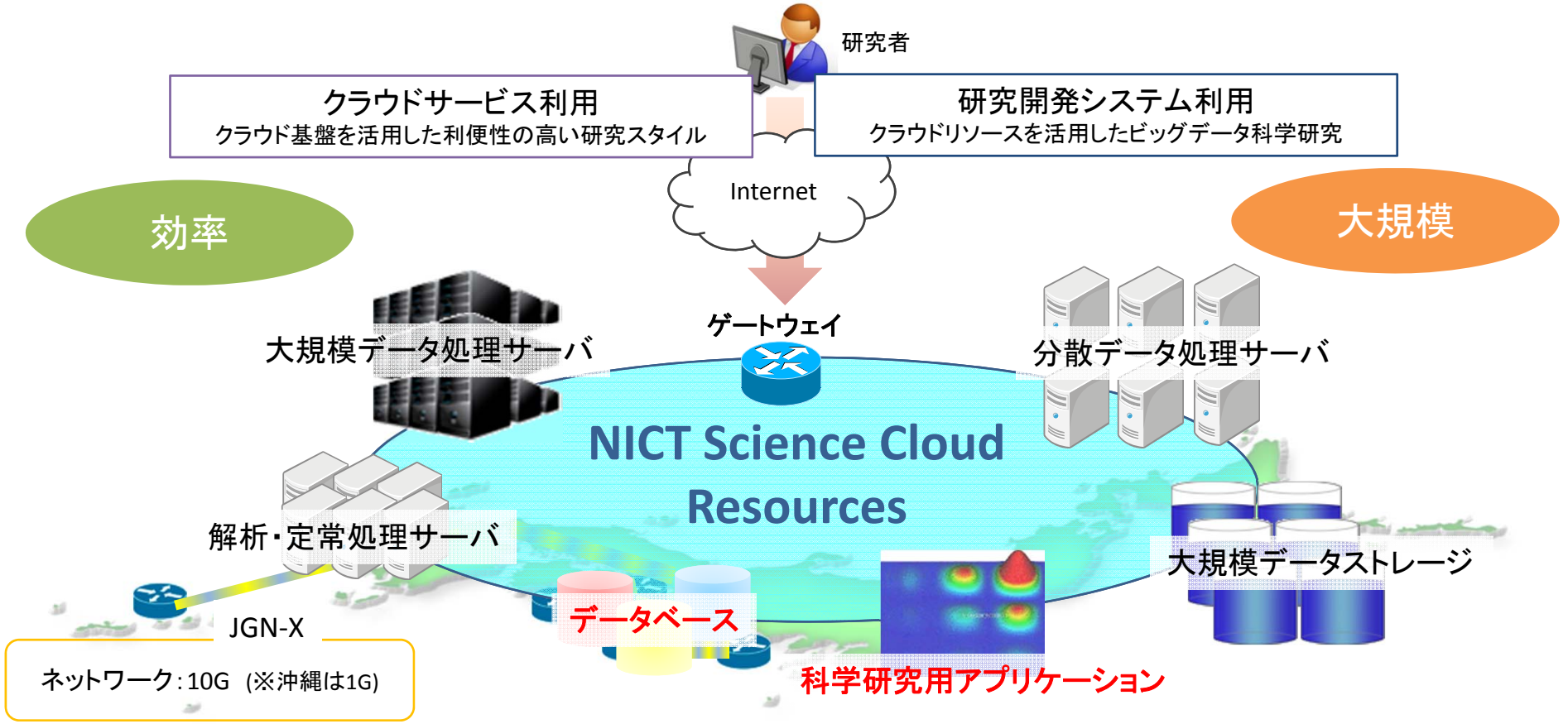
LusterFS（Sun Microsystems）
⇒（例）TSUBAME（1.5PB/60GB/s）
GPFS（IBM）
PVFS2（OSS）



小西史一、情報処理、pp.845-852、2009-09-15にデータを追加

NICT ScienceCloud

サイエンスクラウド利活用概要



解析・定常処理サーバ
【台数】5台
【スペック】Opteron/ Xeon（8～12コア）、メモリ32GB～128GB
【保存場所】東京（京都にマイグレーション）

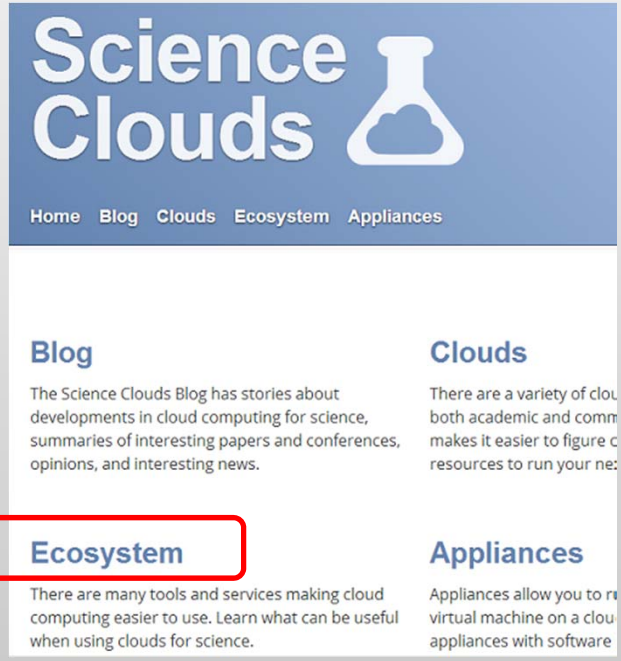
データ保存ストレージ（NAS）
【総容量】150TB
【保存場所】東京（京都にミラーリング）

大規模データ処理サーバ*1
【CPU】合計307コア*2、1.83TFLOPS*3
【RAM】平均71GB/node
【HDD】1.4PB

分散データ処理サーバ*1
【総容量】1.8PB*4
【ディスク規格】SATA3*5、300MB/s（read）*3
【保存場所】東京、京都、沖縄

ECO system in Science Cloud

Cloud ecosystem is a term used to describe the complex system of interdependent components that work together to enable cloud services.



There are many tools and services making cloud computing easier to use.

http://scienceclouds.org/

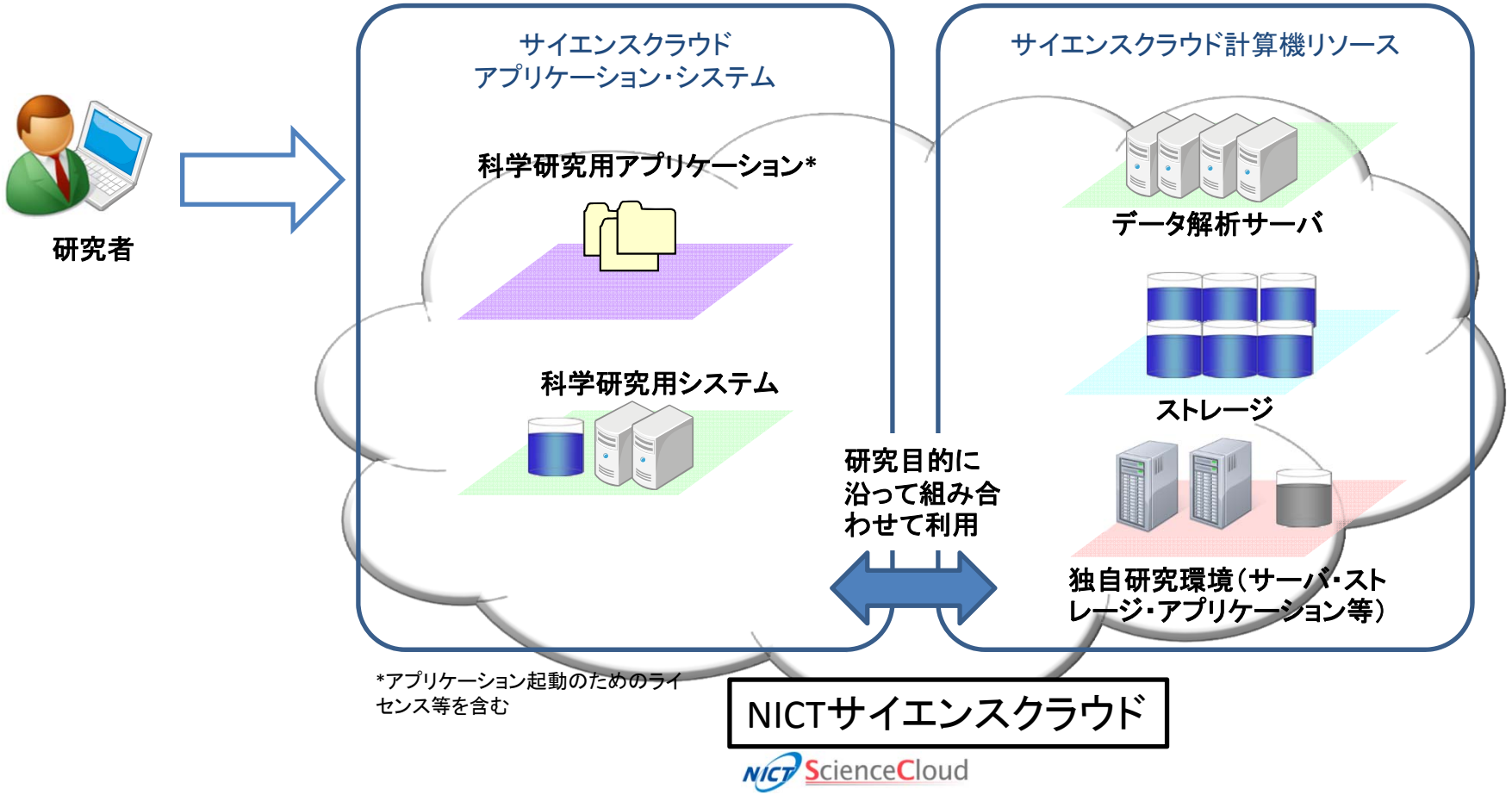
HELIX NEBULA
THESCIENCECLOUD

http://www.helix-nebula.eu/

- NICT Science Cloud
  - ECO system in Cloud for Scientific Researches

NICT ScienceCloud

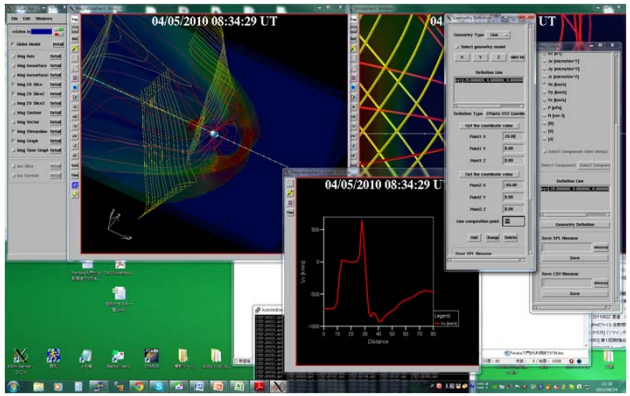
Eco systems in the NICT Science Cloud



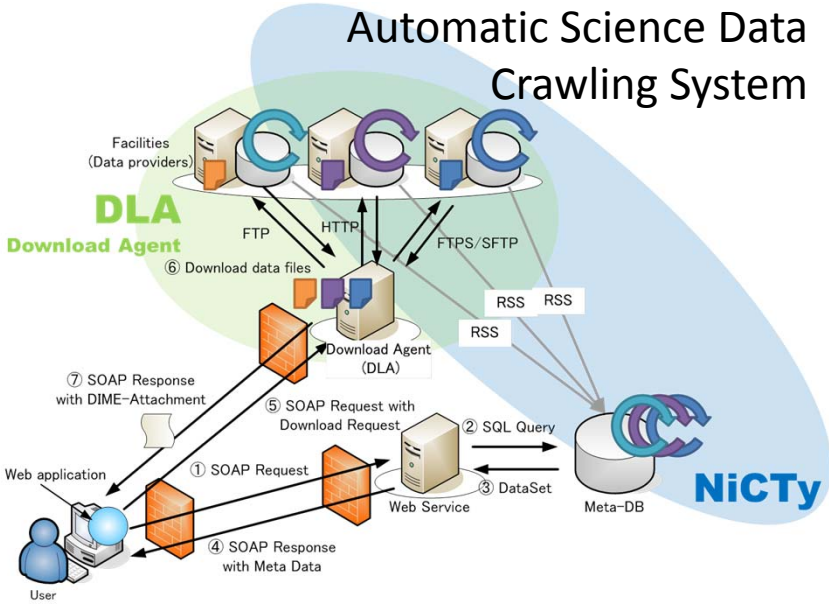
NICT ScienceCloud

Eco systems in the NICT Science Cloud

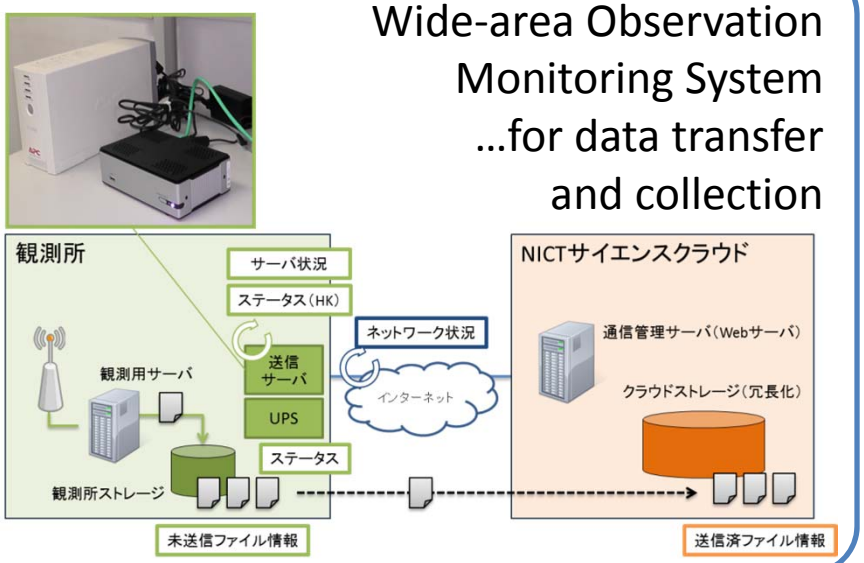
3D Visualization Tool



Automatic Science Data Crawling System



Wide-area Observation Monitoring System …for data transfer and collection



World Science Data Bank (Web Application)

…for reliability and traceability

NICT ScienceCloud

Ionospheric Disturbance via High-Resolution GPS-TEC

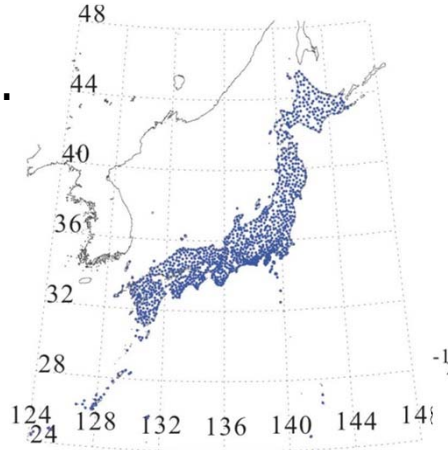
Region
# of GPS Rec.

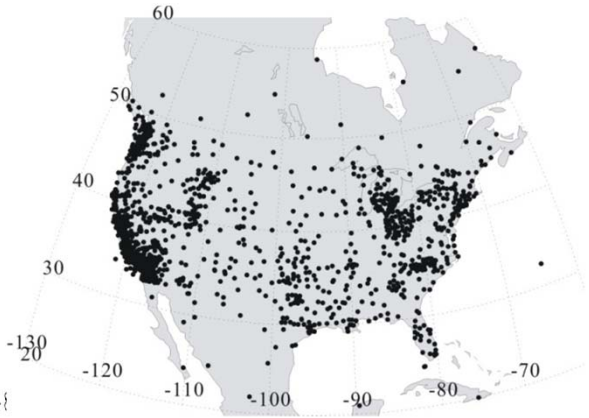
~1,200 receivers

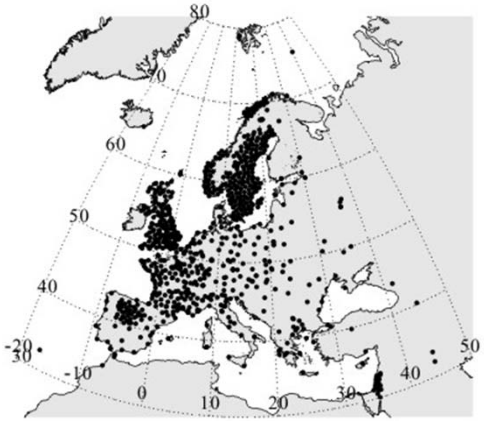
~1,600 receivers

~830 receivers

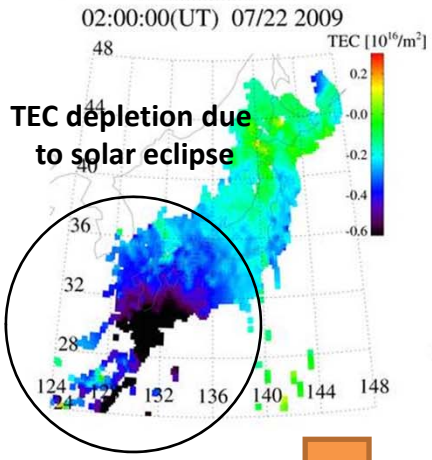
GPS Receivers

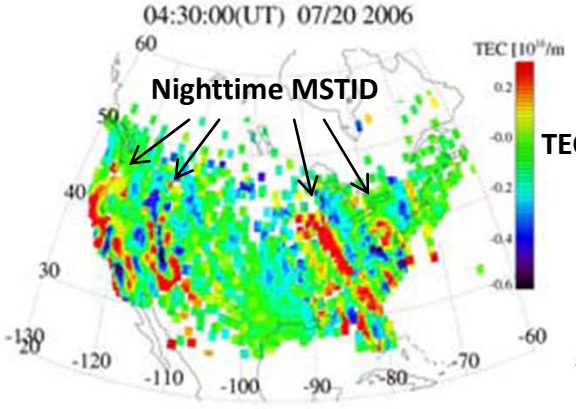


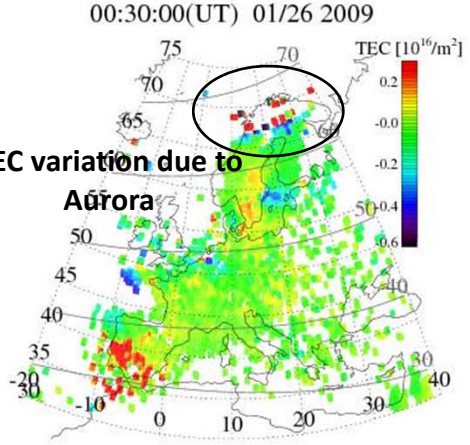




Detrended TEC Map (30sec, 0.15°x0.15°)



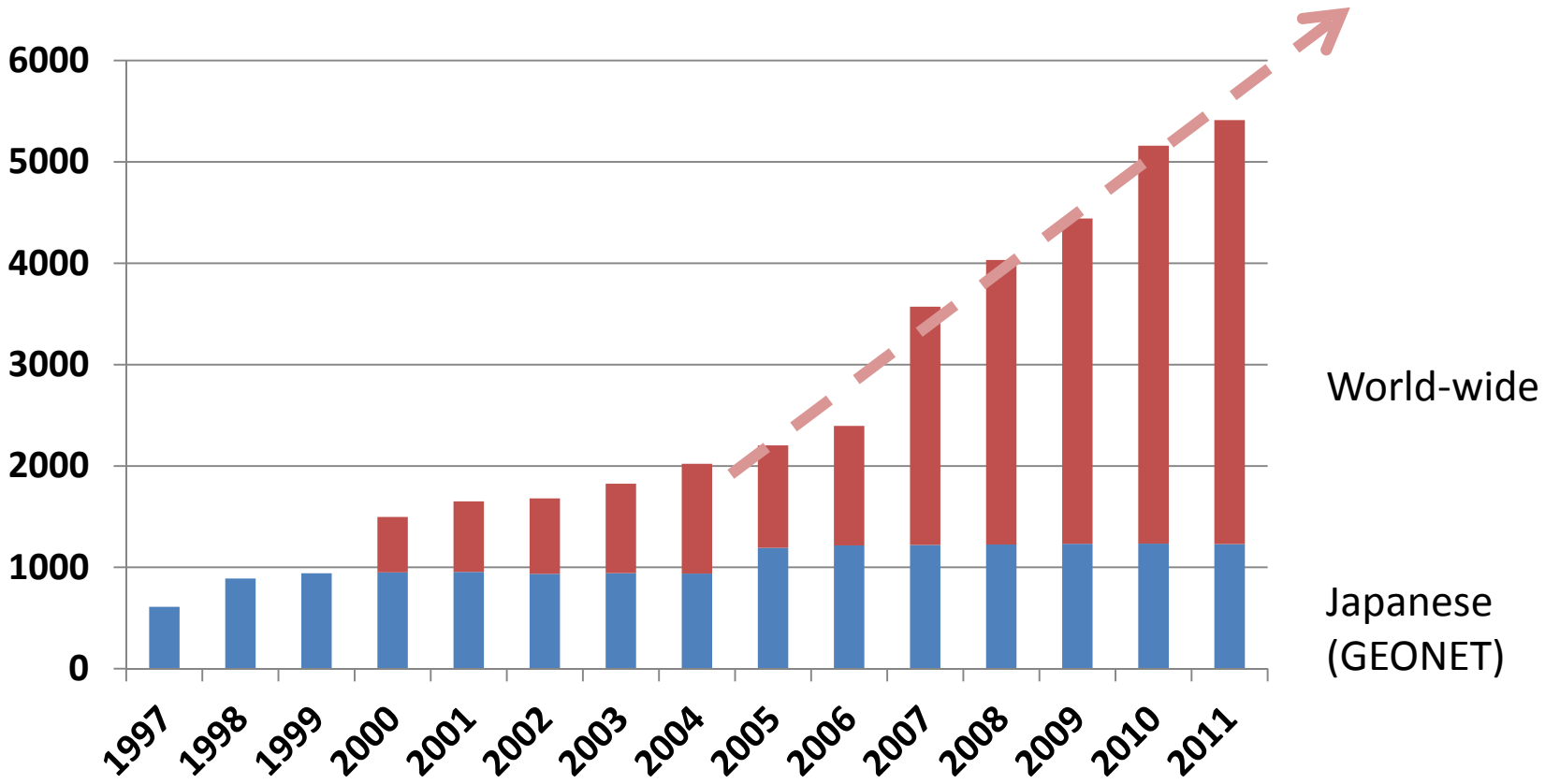




Asia-Oceania GPS-TEC map…

Tusgawa et al. (NICT/Japan)

Trend of the number of GSP receivers

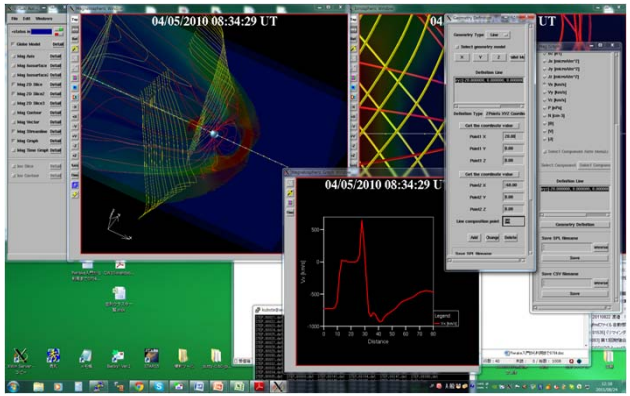


- The annual trend of the number of GSP receivers (world-wide since 2000 and domestic (Japan) since 1997) .

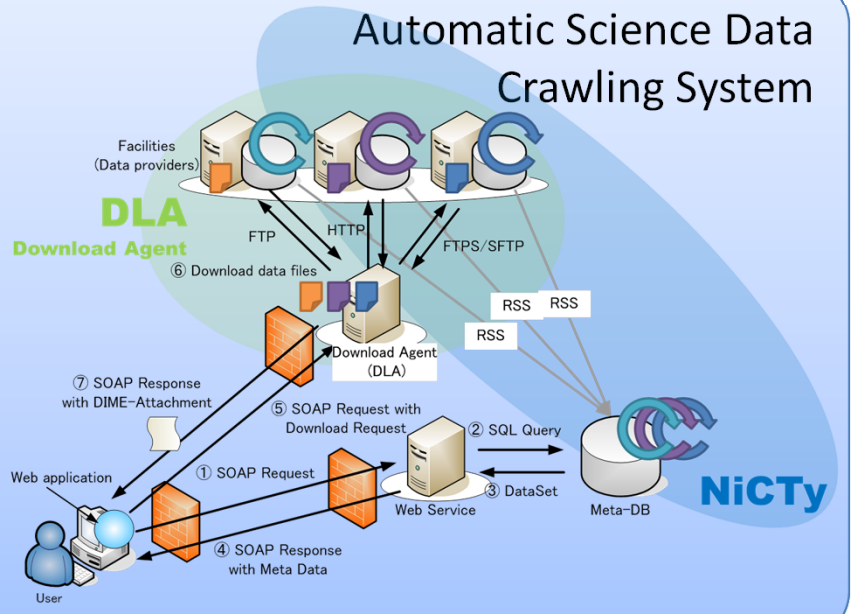
Tusgawa et al. (NICT/Japan)

Eco systems in the NICT Science Cloud

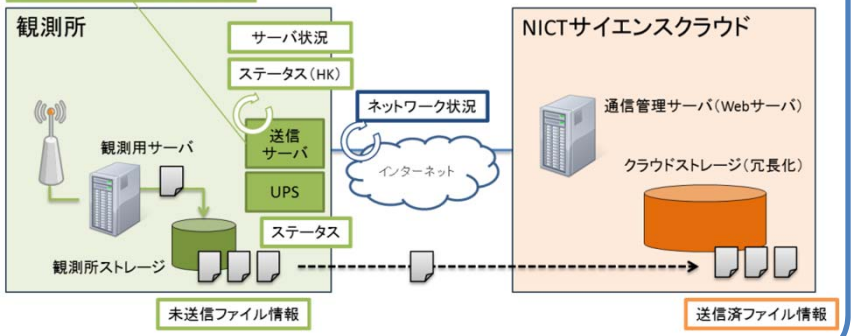
3D Visualization Tool



Automatic Science Data Crawling System



Wide-area Observation Monitoring System

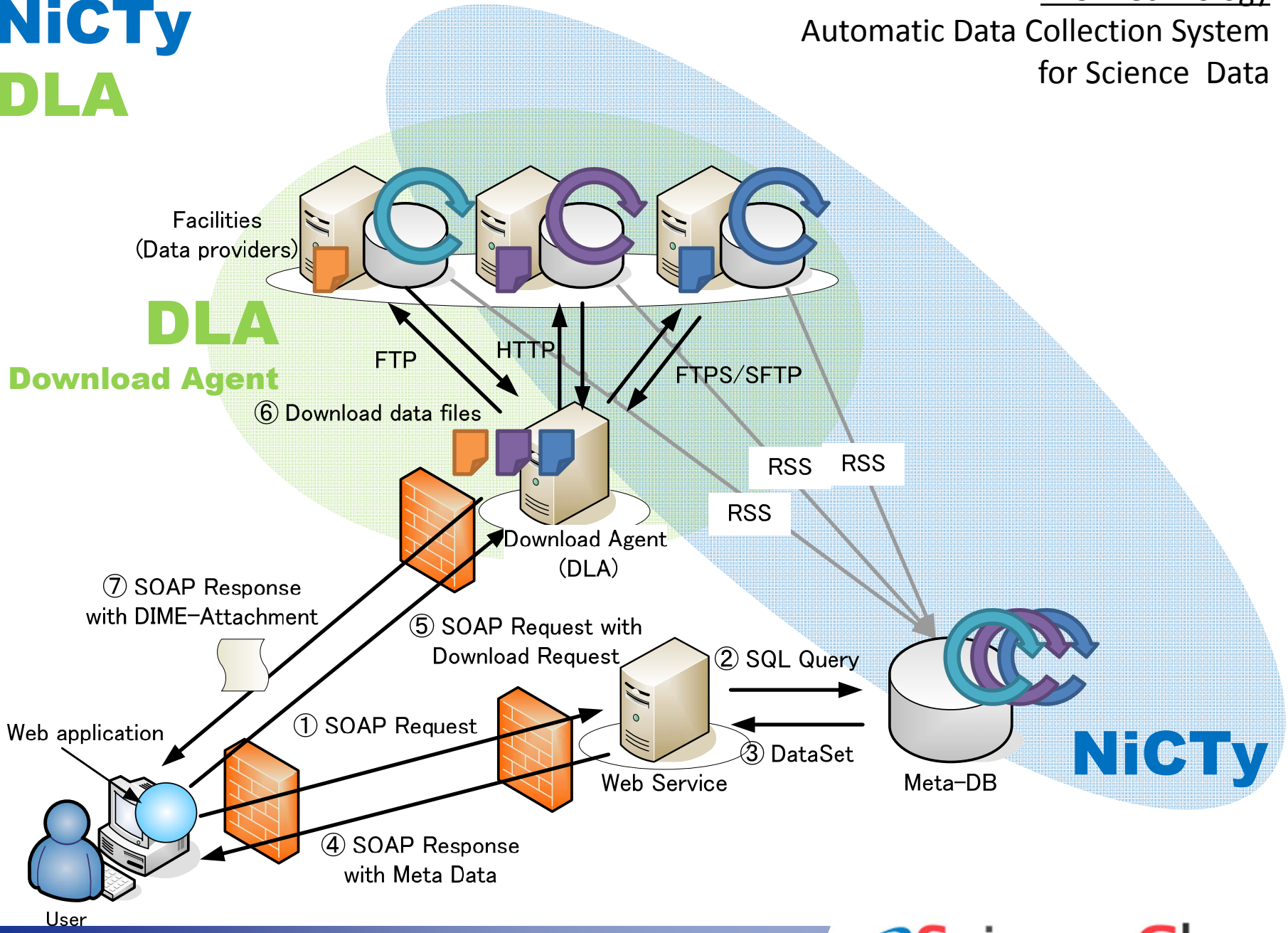


World Science Data Bank (Web Application)

NICT ScienceCloud

NiCTy DLA

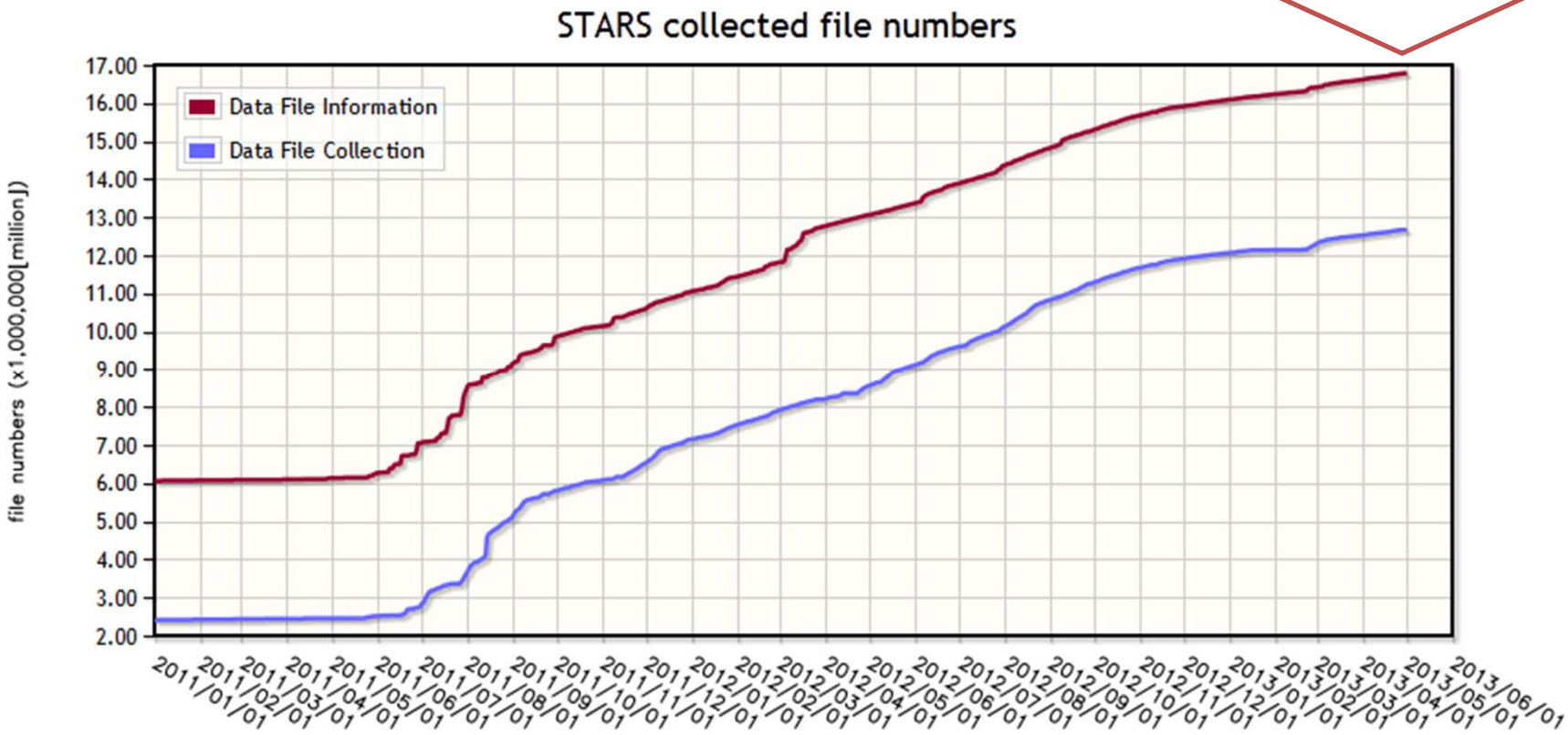
NICT Technology
Automatic Data Collection System
for Science Data



NICT ScienceCloud

Automatic Data Collection

More than 17 million files

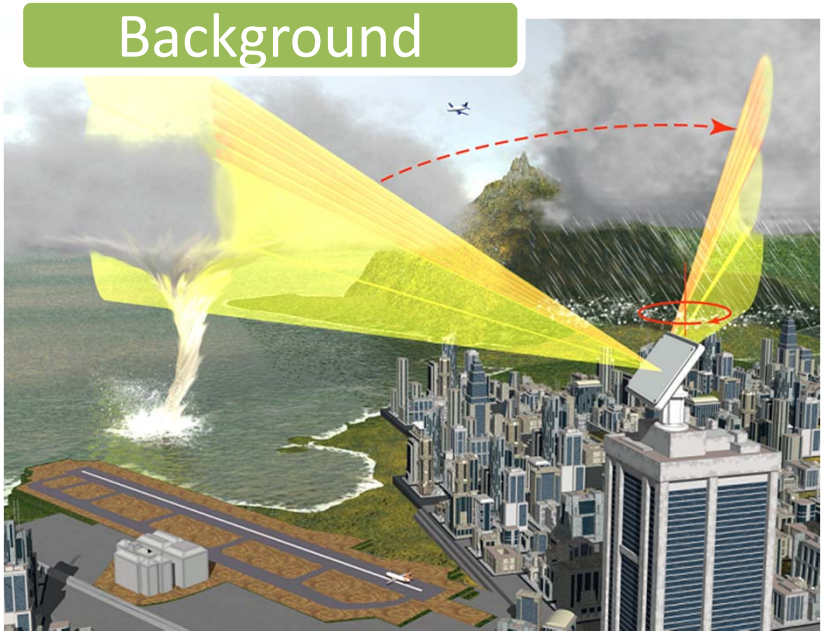


NICT ScienceCloud

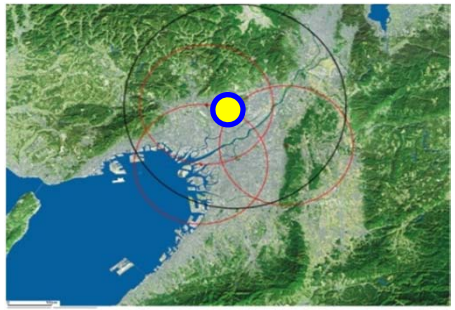
3D remote sensing via phased-array radar (meteorological radar located at Osaka Univ.)

Shinsuke Sato
et al. (NICT, Japan)

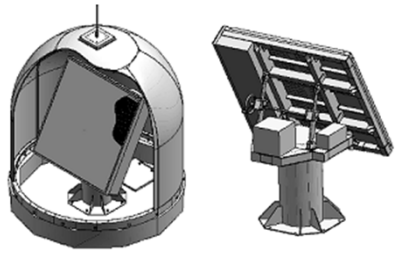
Background



An next generational phased-array radar with as high spatial resolution as 100m with temporal duration of 10 seconds can detect 3D profiles (structures) of localized rain and tornado/flurry



Located at Osaka University



Outlooks of the 3D phased-array radar which enables high-speed scanning

Objectives

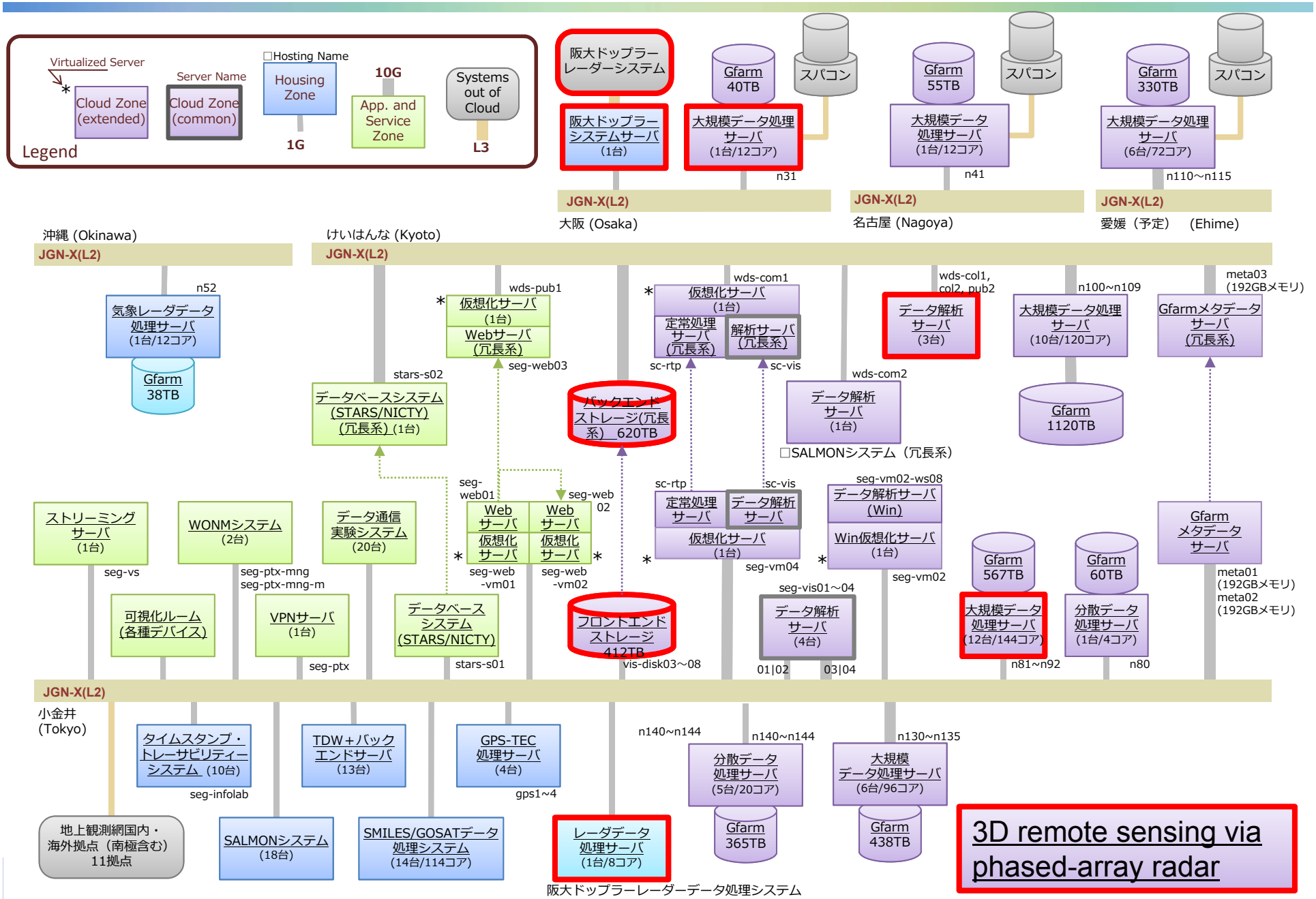
Meteorological radar so far = 2D

3D meteorological Radar

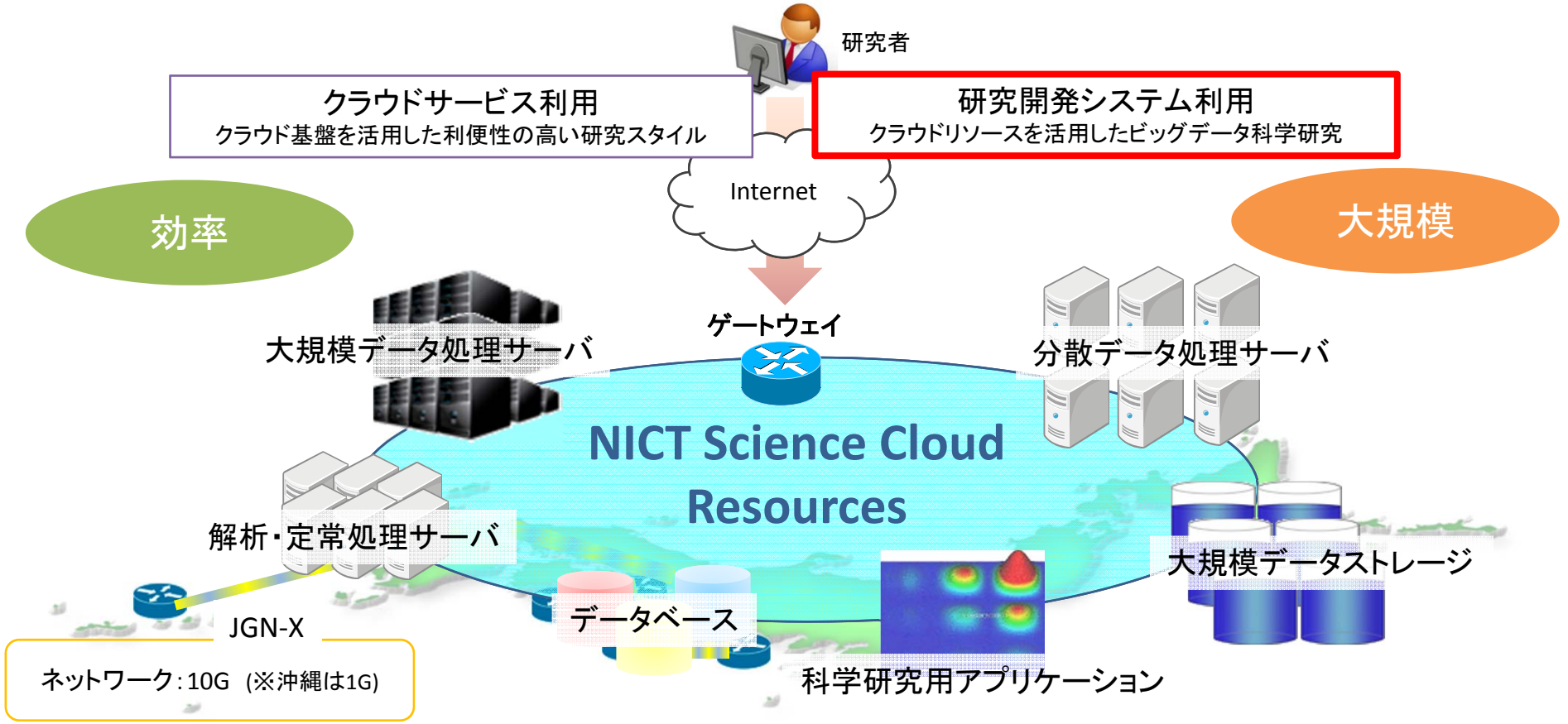
- Data generation rate: 200-500MB/30sec.
- Archived data:200-500TB/year
- International data analysis team (collaborative works)
- 3D radar visualization within 1min. after observation.


An Integrated System using NICT Science Cloud
Cloud Resource
App. and Service
Housing
*under development
3D observation
L0 to L1/L2
conversion
3D Phased-Array Data System (Sato, NICT)
Real-time 3D visualization
(within 1 min.)
Data transfer
server
Gateway
LDAP
NICT Science Cloud Viz Gallery
High-speed
data transfer
technique*
NICT ScienceCloud
JGN-X
Heterogeneous Many Core Servers
Heterogeneous
Parallel Data
Processing
Technology
Phased-array
Radar
Remote
Observation
data transfer
system
1GB/sec
30MB/sec
Concurrent Data Processing
Automatic Data Backup
Data Processing
Server
Cloud Storage
Data Backup
Virus Check
Large-scale storage (NAS)
Virtual Laboratory
(international)
Application
(commercial/NICT
original)
WSDBank
Web app.
Automatic data collection
(crawling)
Wide-area data management*
Time Stamp technique*
Traceability*
Wide-area distributed
storage
Internet
== Flexible ==
Customization and Construction of
Original System for Your Science
== Technical ==
Special Applications, Tools and
Systems for Sciences
== Economical ==
Commission of Basic Operations
and maintenance of Cloud

System Configuration and Deployment of NICT Science Cloud (in Japanese) 2013



サイエンスクラウド利活用概要



解析・定常処理サーバ
【台数】5台
【スペック】Opteron/ Xeon（8～12コア）、メモリ32GB～128GB
【保存場所】東京（京都にマイグレーション）

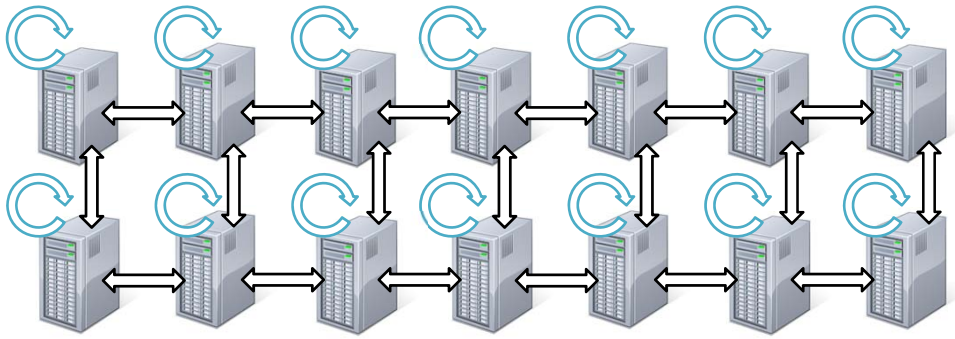
データ保存ストレージ（NAS）
【総容量】150TB
【保存場所】東京（京都にミラーリング）

大規模データ処理サーバ*1
【CPU】合計307コア*2、1.83TFLOPS*3
【RAM】平均71GB/node
【HDD】1.4PB

分散データ処理サーバ*1
【総容量】1.8PB*4
【ディスク規格】SATA3*5、300MB/s（read）*3
【保存場所】東京、京都、沖縄

HPCとMTC/DIC

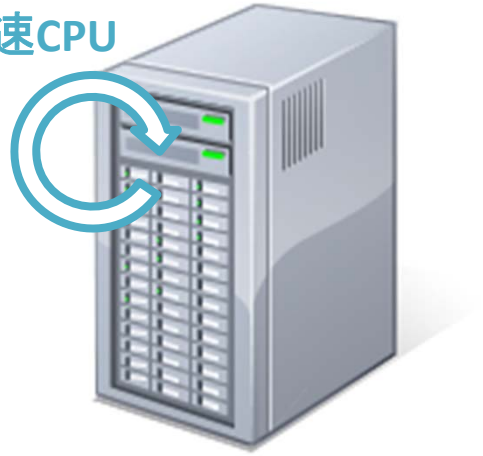
これまで：HPC



CPU高速化

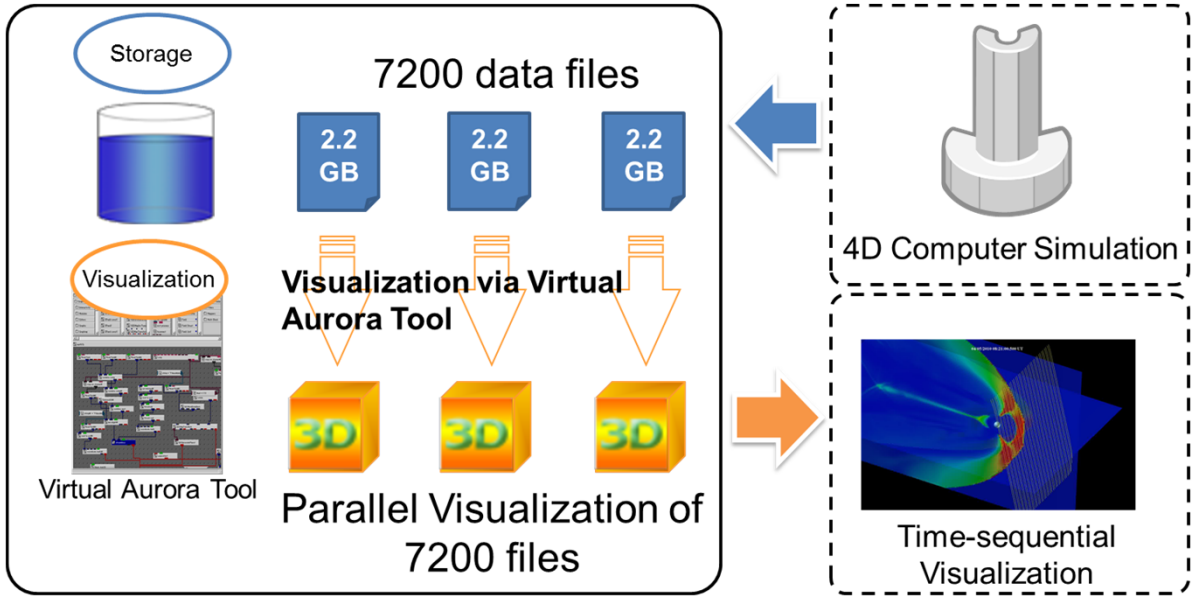
実現する仮想化コンピュータ

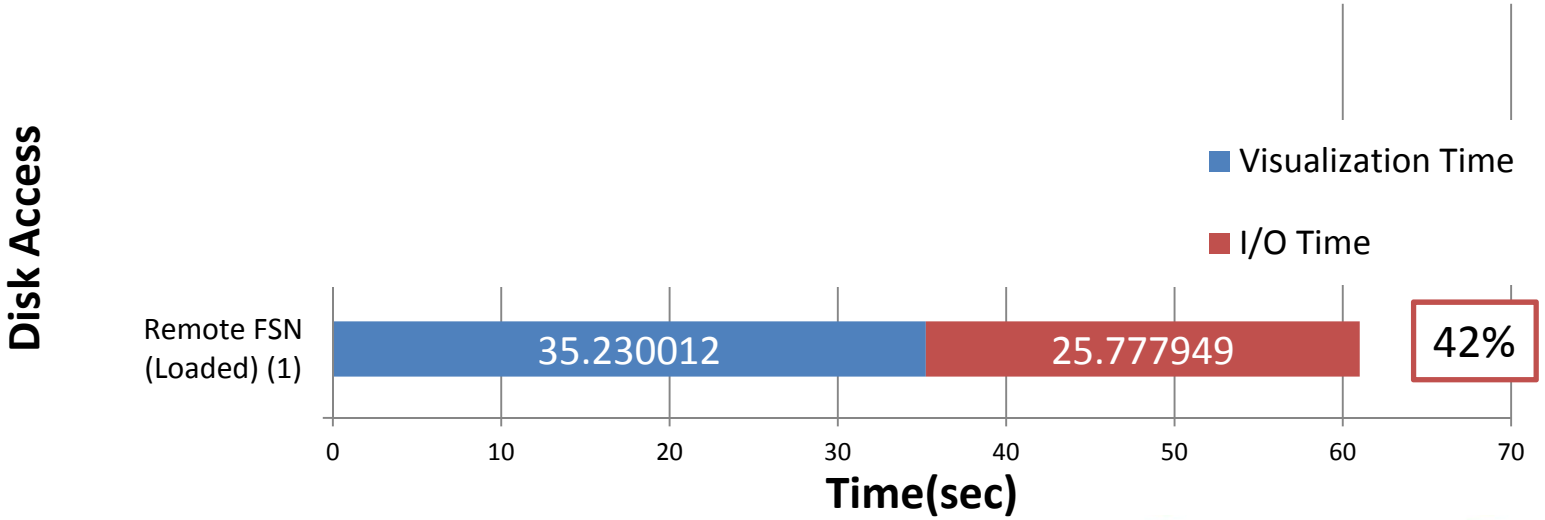
仮想高速CPU



HPC: High Performance Computing

数値シミュレーションの例

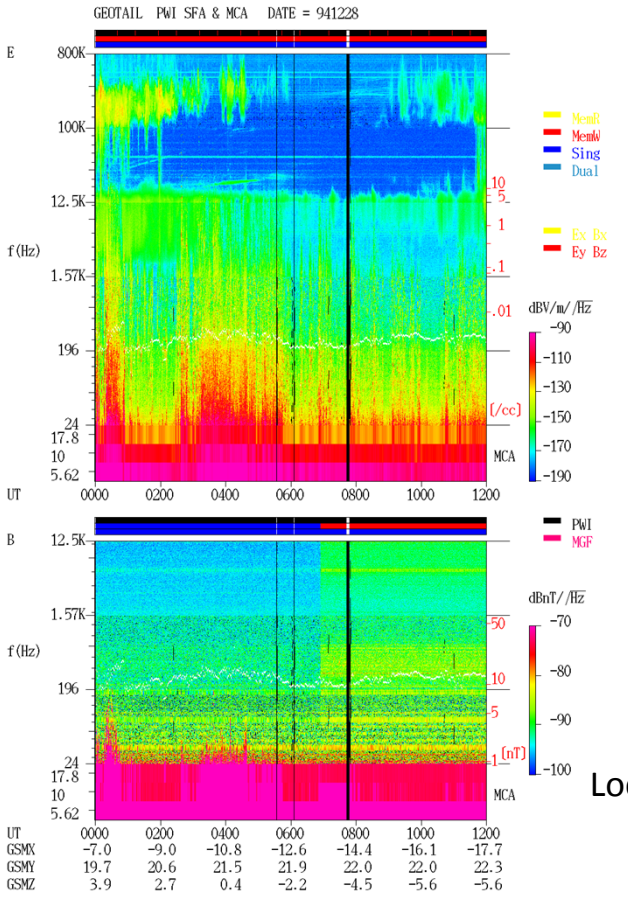




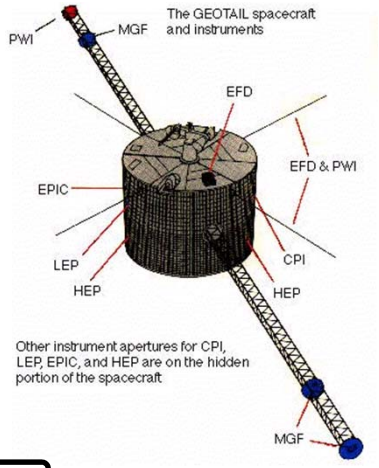
NICT ScienceCloud

科学衛星観測の例

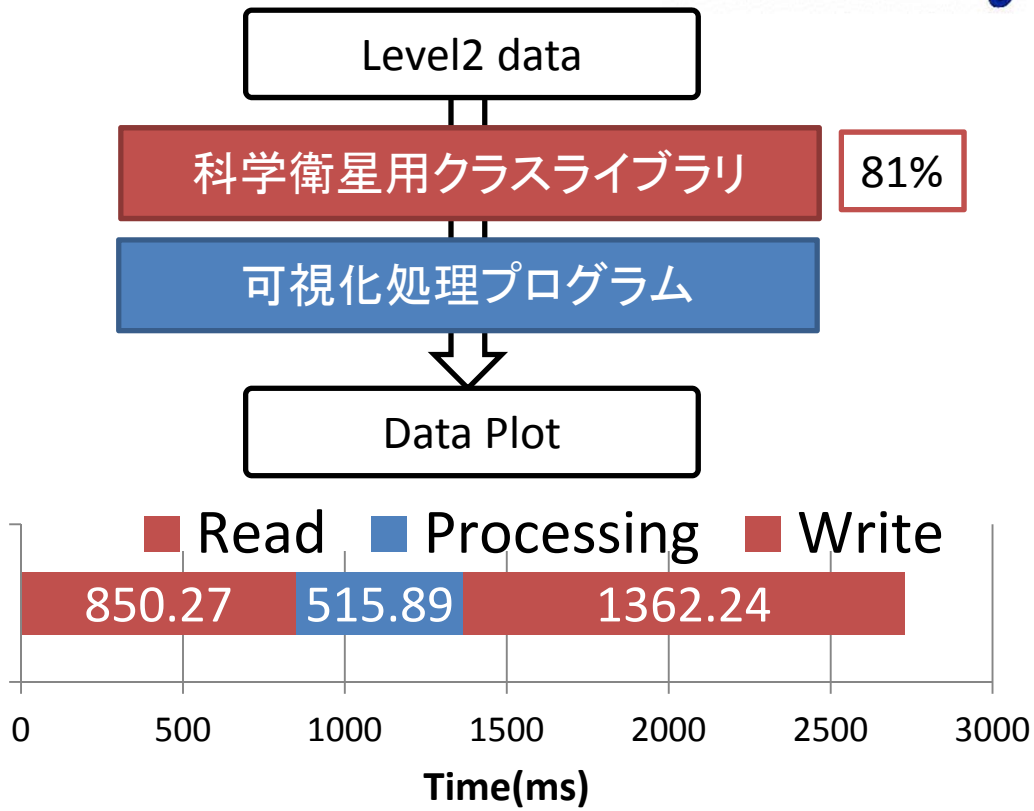
プラズマ波動
（ダイナミックスペクトル）



http://www.rish.kyoto-u.ac.jp/gtlpwi/



GEOTAIL衛星
地球磁気圏観測衛星
1992年に打ち上げ（現在も運用中）

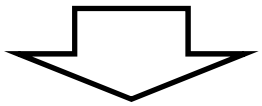


NICT ScienceCloud

HPC, MTC, and DIC

High Performance Computing (HPC)
heavily focusing on compute-intensive applications

- Computing Oriented
- The third paradigm
  - Computational Science
- Infrastructure
  - Super Computer
- Parallelization
  - CPU

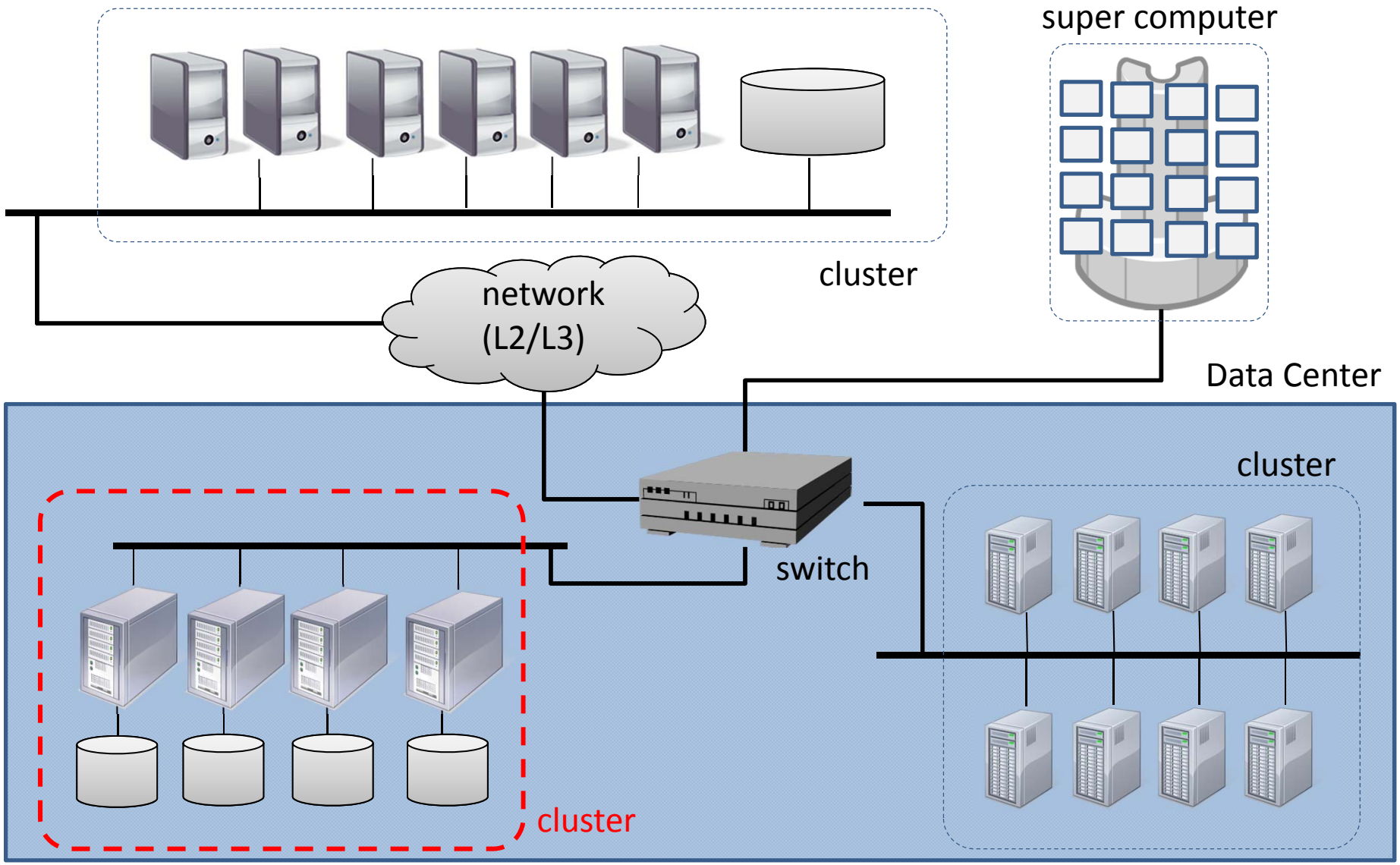


Many-Task Computing (MTC)
focusing on using many resources over short periods of time

Data Intensive Computing (DIC)
focusing on data distribution, data-parallel execution, and harnessing data locality by scheduling of computations close to the data

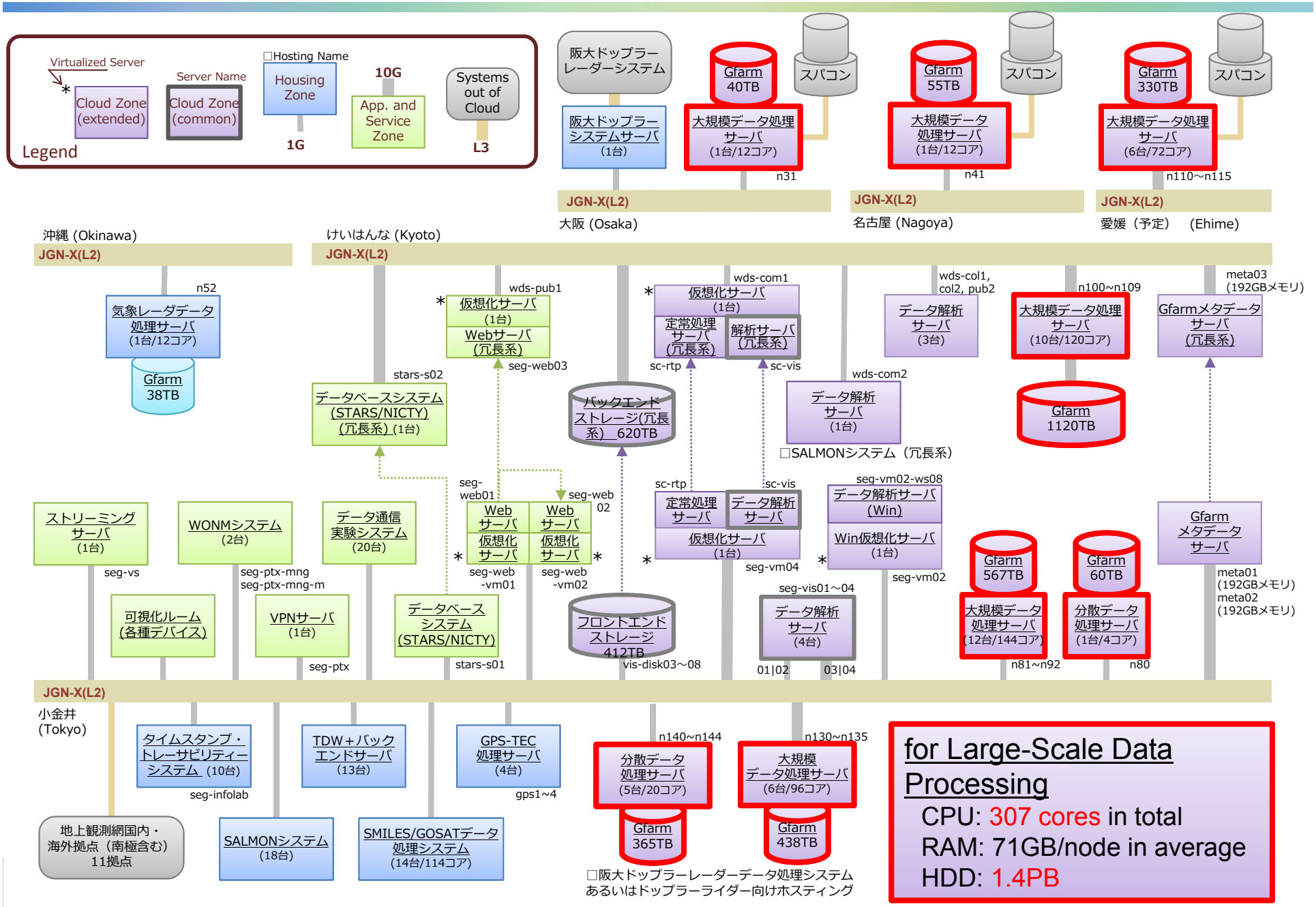
- Data Oriented
- The forth paradigm
  - Data-intensive Science
- Infrastructure
  - Science Cloud
- Parallelization
  - CPU, I/O, network…

Scalable Clusters in Science Cloud



NICT ScienceCloud

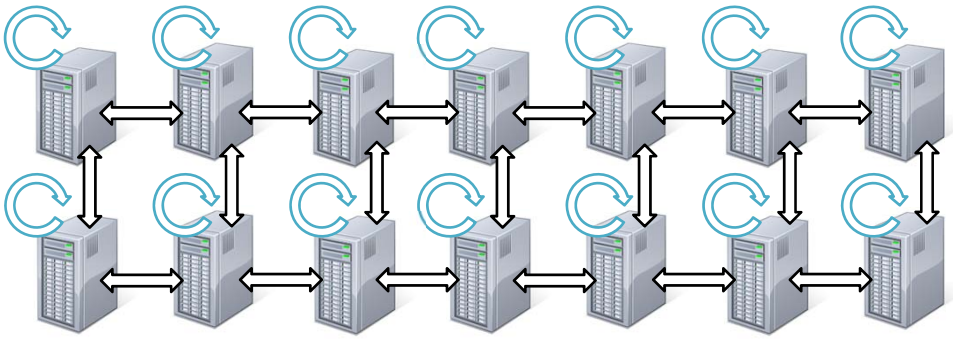
System Configuration and Deployment of NICT Science Cloud (in Japanese) 2013



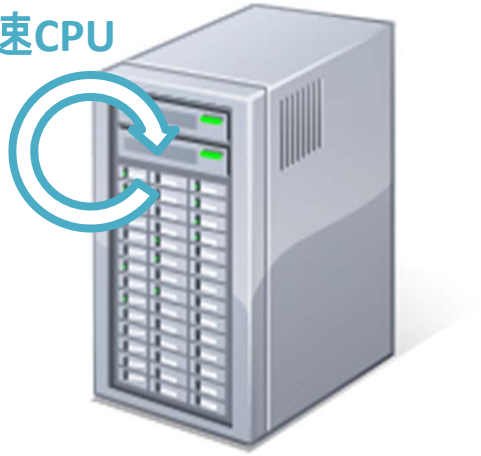
HPCとMTC/DIC

実現する仮想化コンピュータ

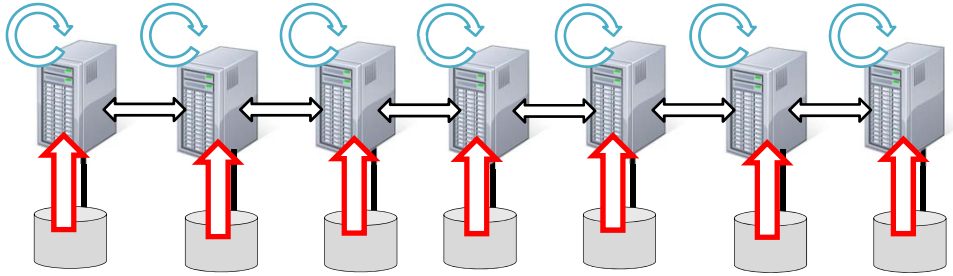
これまで：HPC



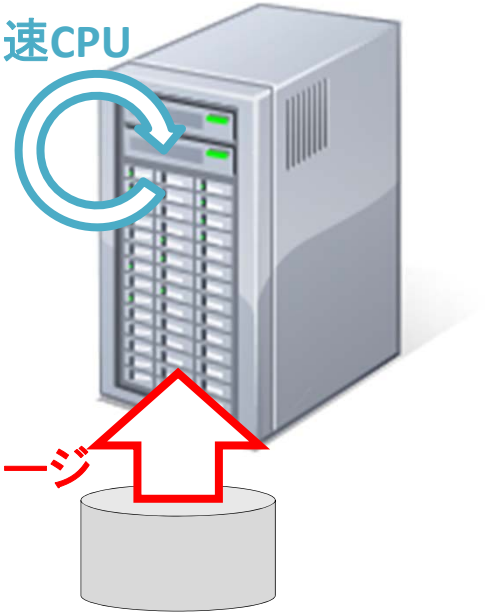
CPU高速化



本研究：HPC＋高速I/O=MTC（DIC）



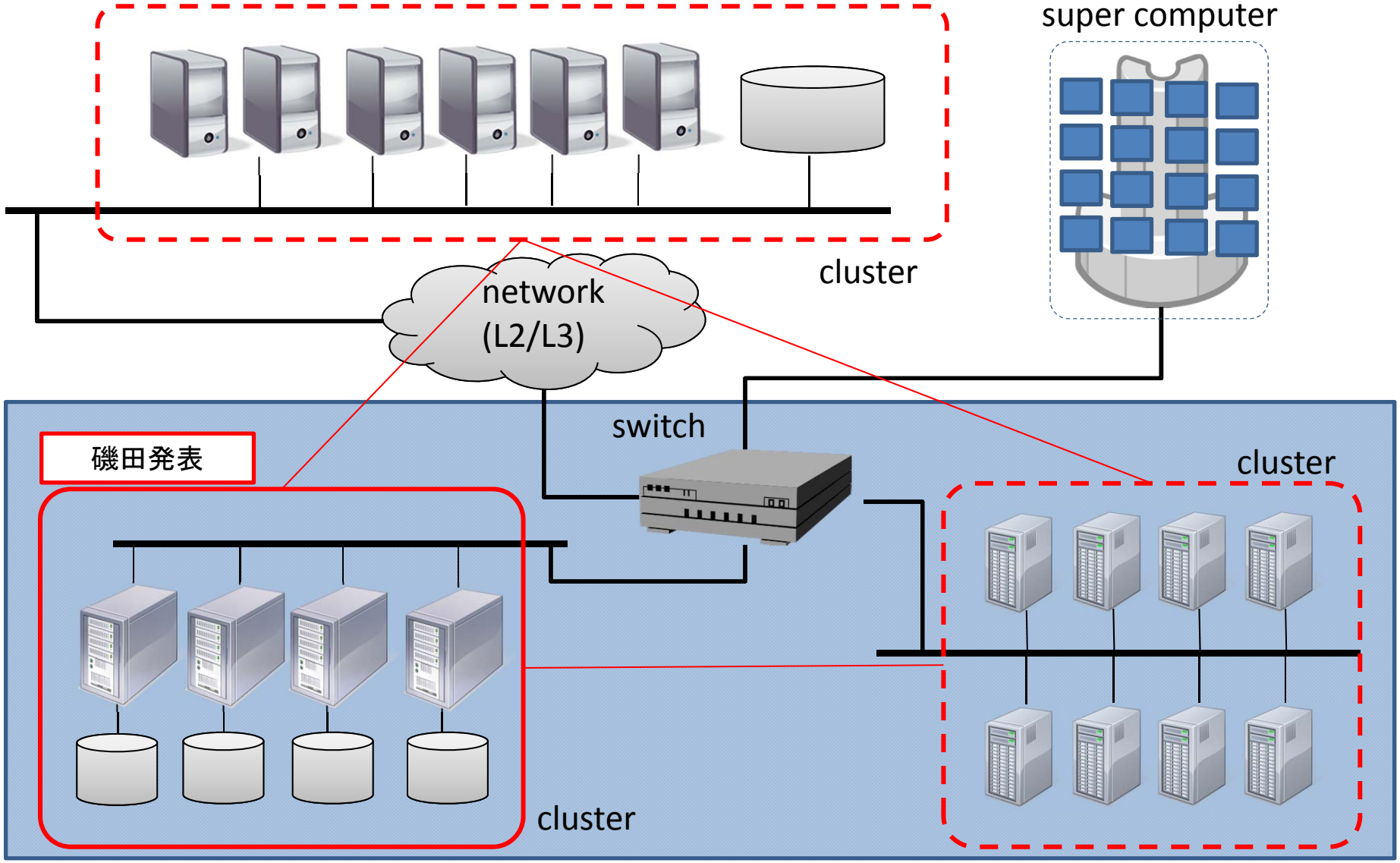
CPU高速化
I/O高速化



HPC: High Performance Computing　MTC: Many-Task Computing

NICT ScienceCloud

NEXT STEP : マルチクラスタ計算技術

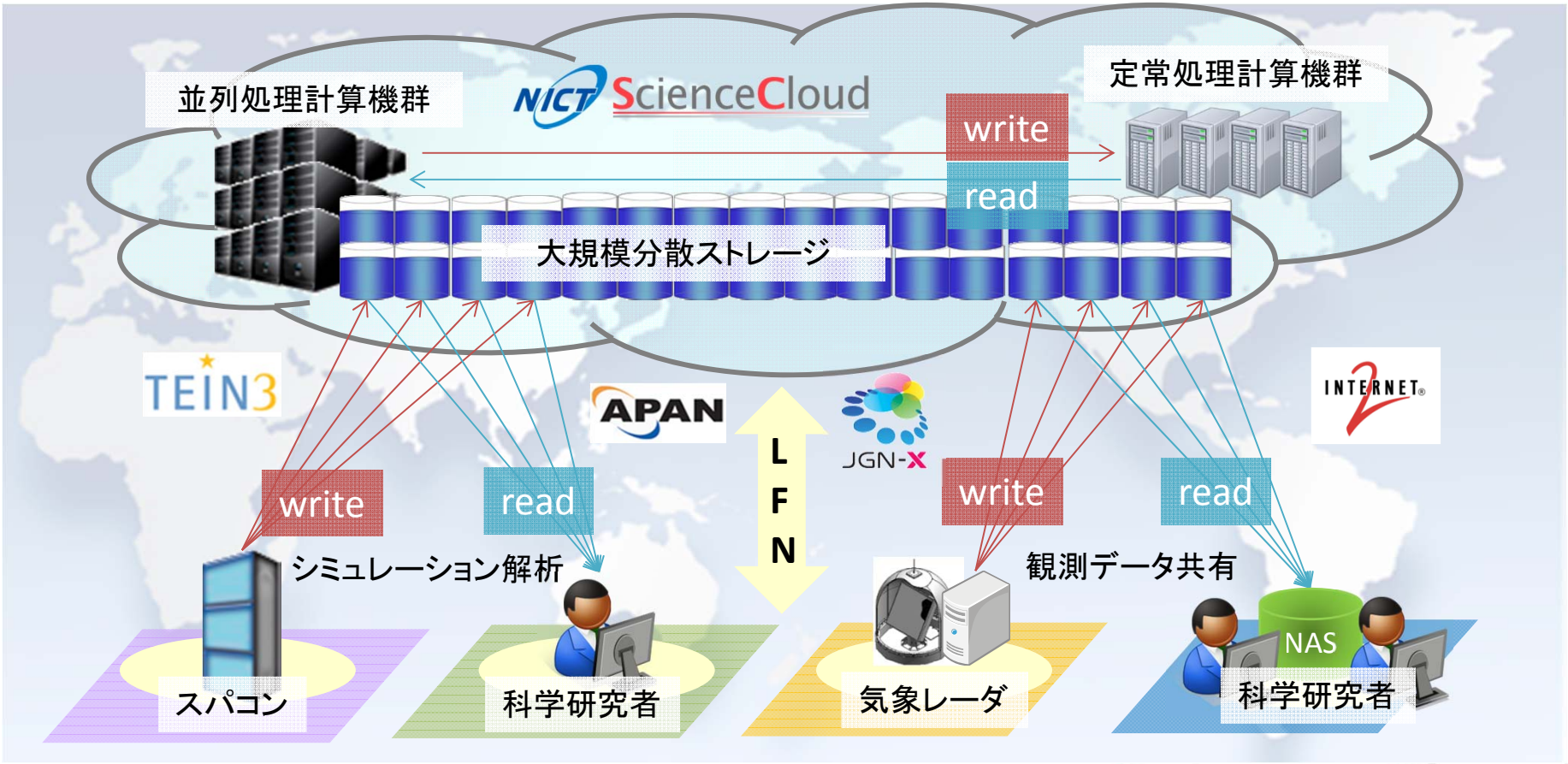


NICT ScienceCloud

ビッグデータ解析のための基盤技術：大規模データ伝送

- 課題：長距離高帯域ネットワーク(LFN: Long Fat Network)のデータ移行
  - ✓ TCPはLFN上では高いスループットを維持しがたい

➢ 目的：LFN上で一対多のクライアント・サーバ型の高速データ移行（ダウンロード・アップロード）が可能なアプリケーションを実現



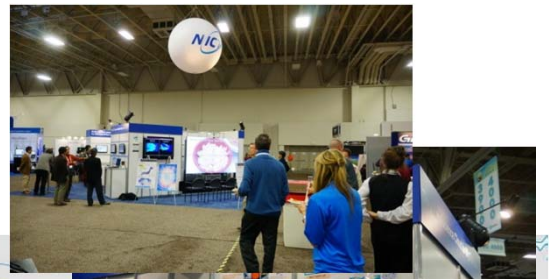
参考：S. Floyd et.al, HighSpeed TCP for Large Congestion Windows

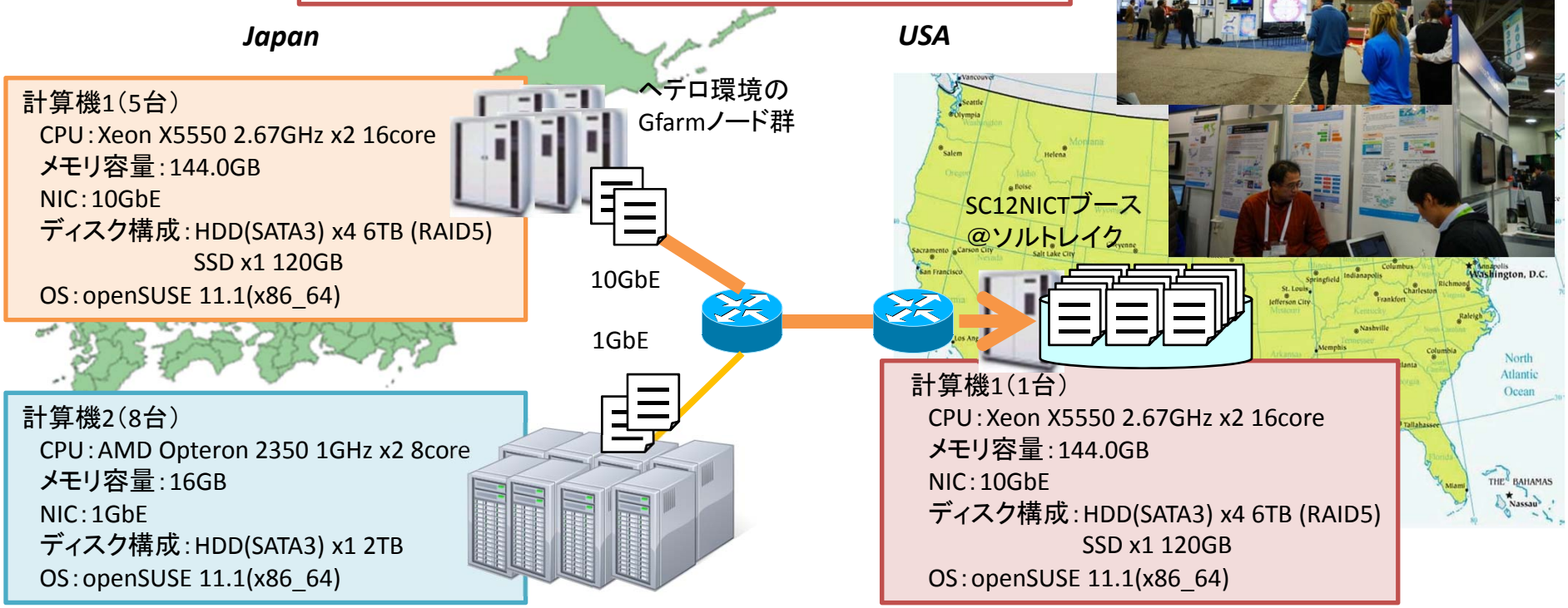
NICT ScienceCloud

日米間並列ファイル転送実験（並列ファイルダウンロード）

- NICTサイエンスクラウド上の13台 のファイルノード上に配置した125MBの512個のファイルを1台のクライアント計算機＠米国から連続読み出し
  - ✓ 小金井(東京)・ソルトレイク（米国）のRTTは143ミリ秒
  - ✓ タイムスケジュールで実験ネットワークを占有利用

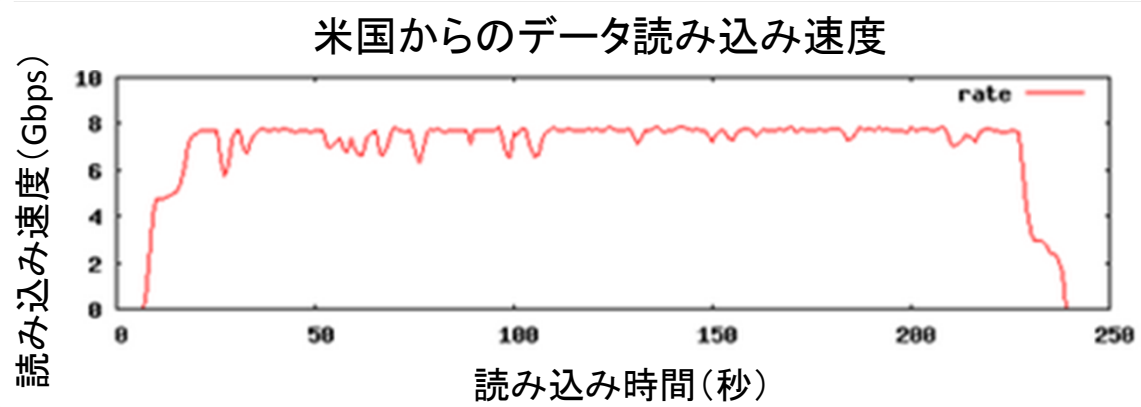
13台の送信サーバ（=13セッション）
13×100MB/sec⇒1300MB/sec⇒10Gpbs

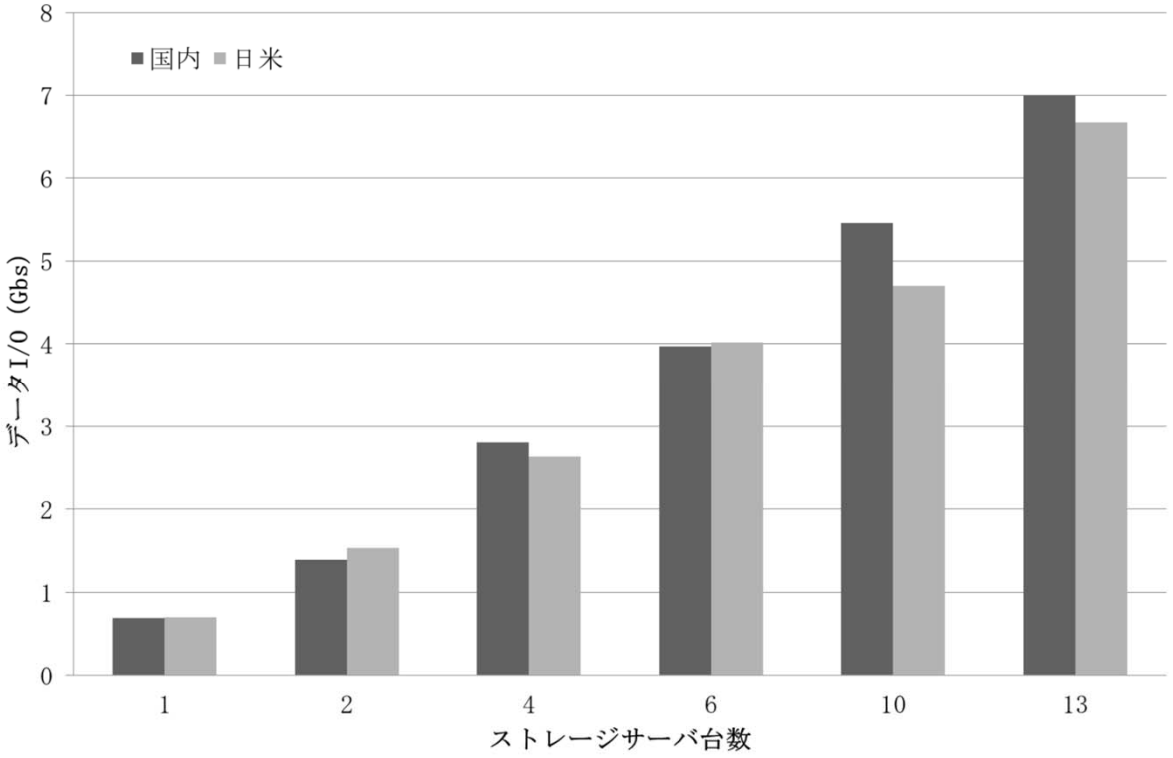




実験結果　データI/O（Read）

- 6.95Gbps（平均速度）
- 参考：単体ディスクI/O（Read）性能
  - SSD：約1.76Gbps
  - HDD（SATA3・RAID5）：約3.78Gbps
- マルチプロセス・マルチスレッドによる並列ファイル転送を実施
  - 13セッションを4プロセスで並列転送
  - マルチスレッドで性能低下した場合の影響範囲を調整

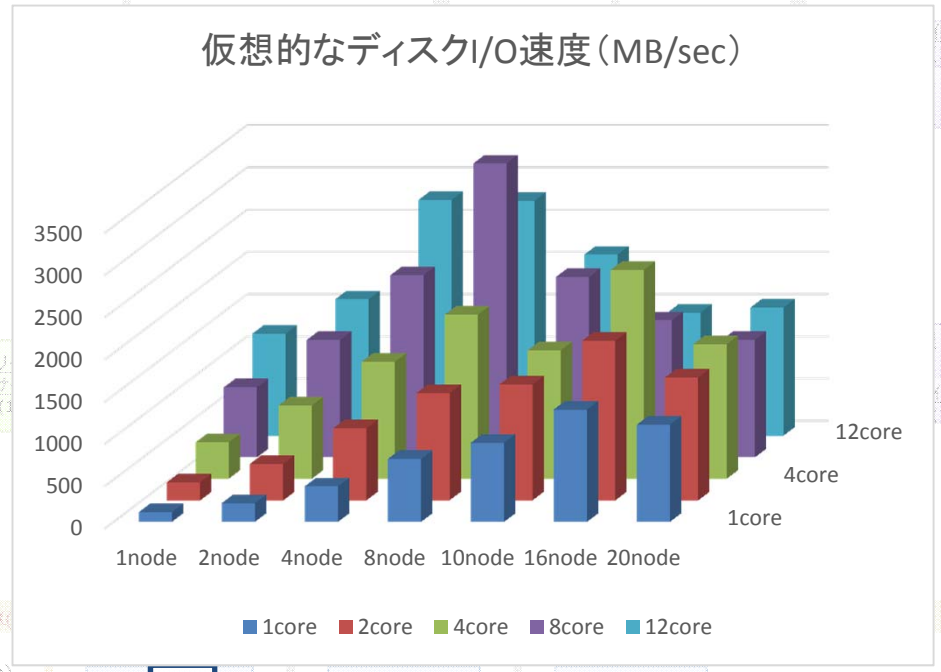




NICT ScienceCloud

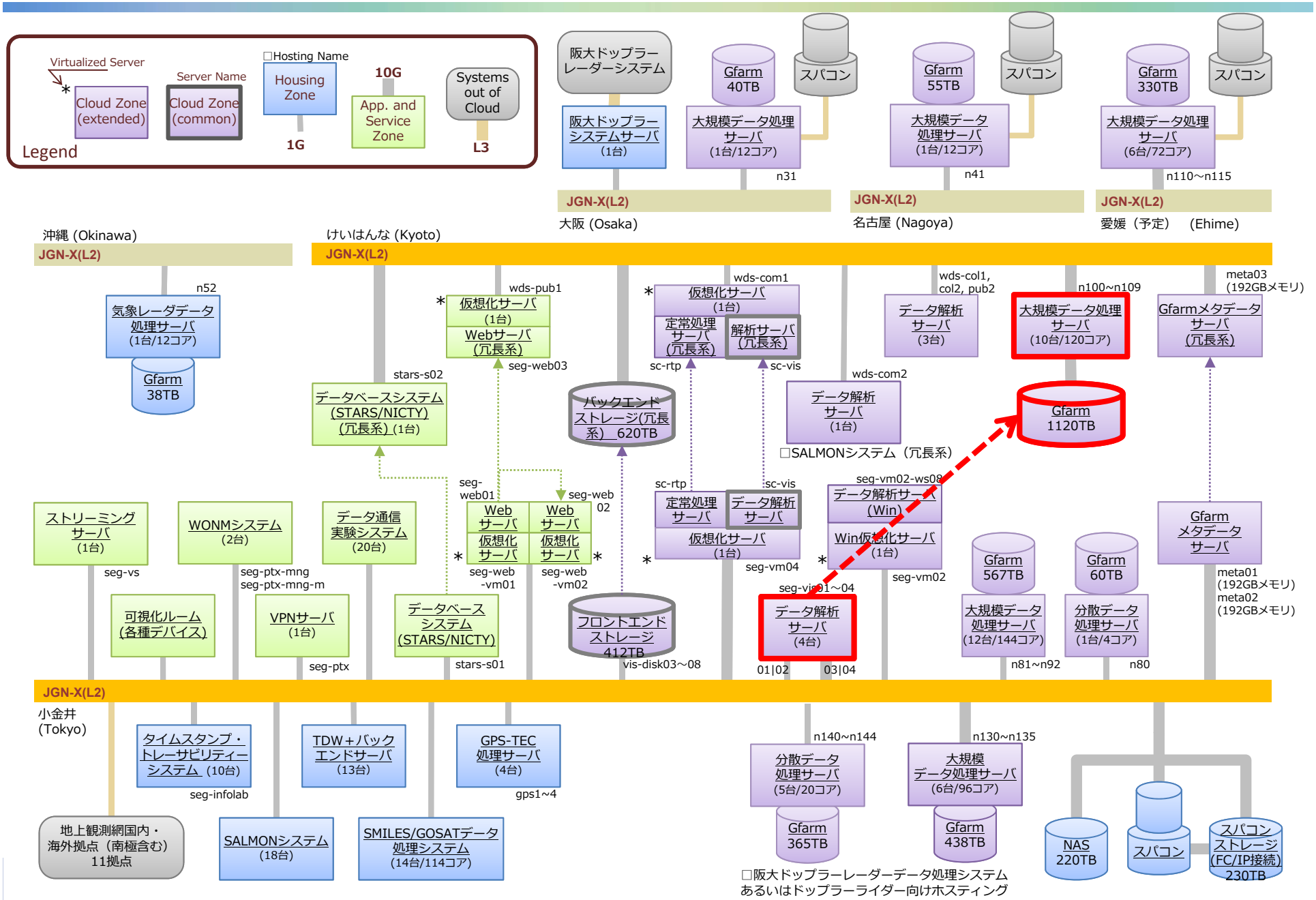
仮想高速ストレージ実験

仮想ファイルI/O速度（20ノード）

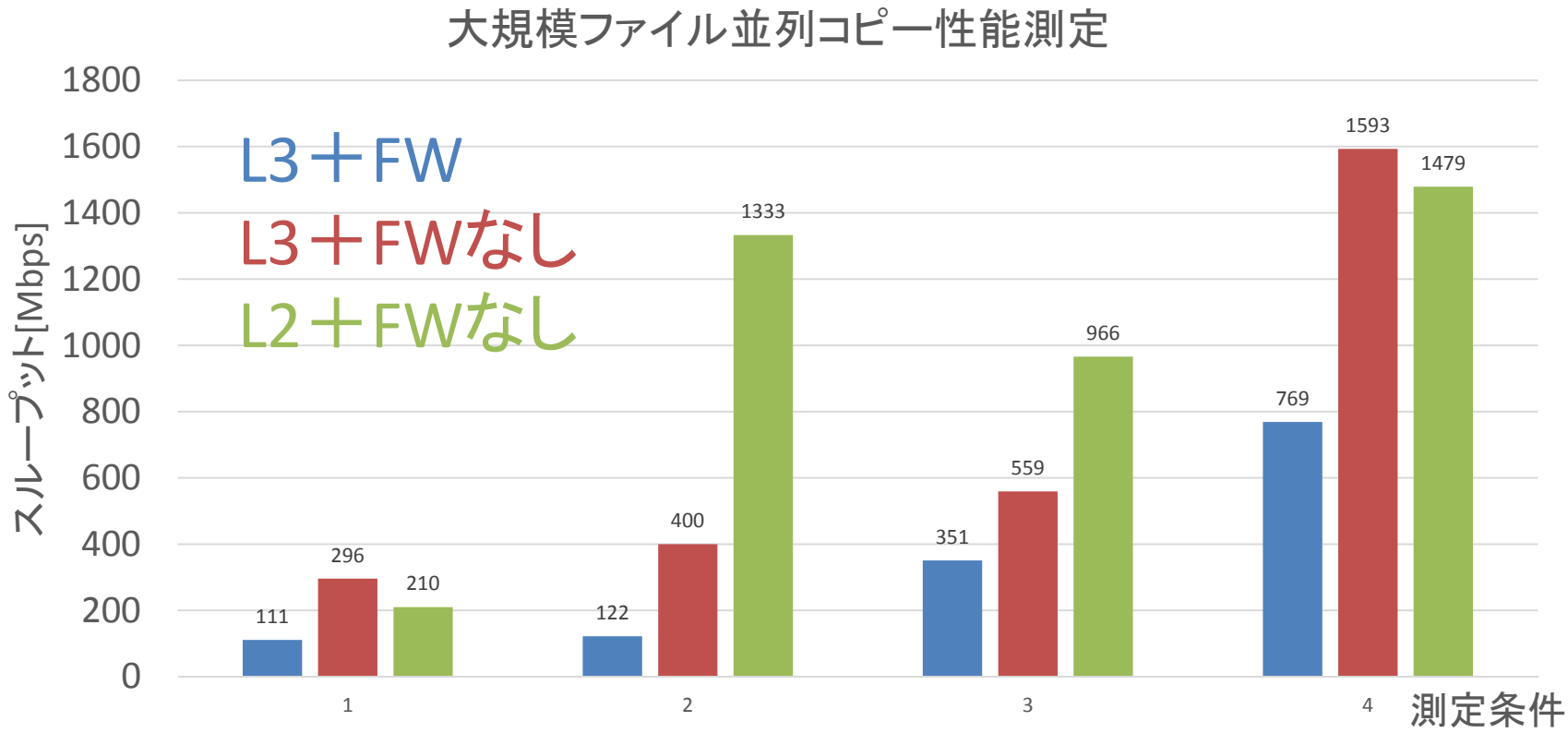


ローカルファイルI/O（約110MB/s）の
約30倍（3.5GB/s）を実現

System Configuration and Deployment of NICT Science Cloud (in Japanese) 2013



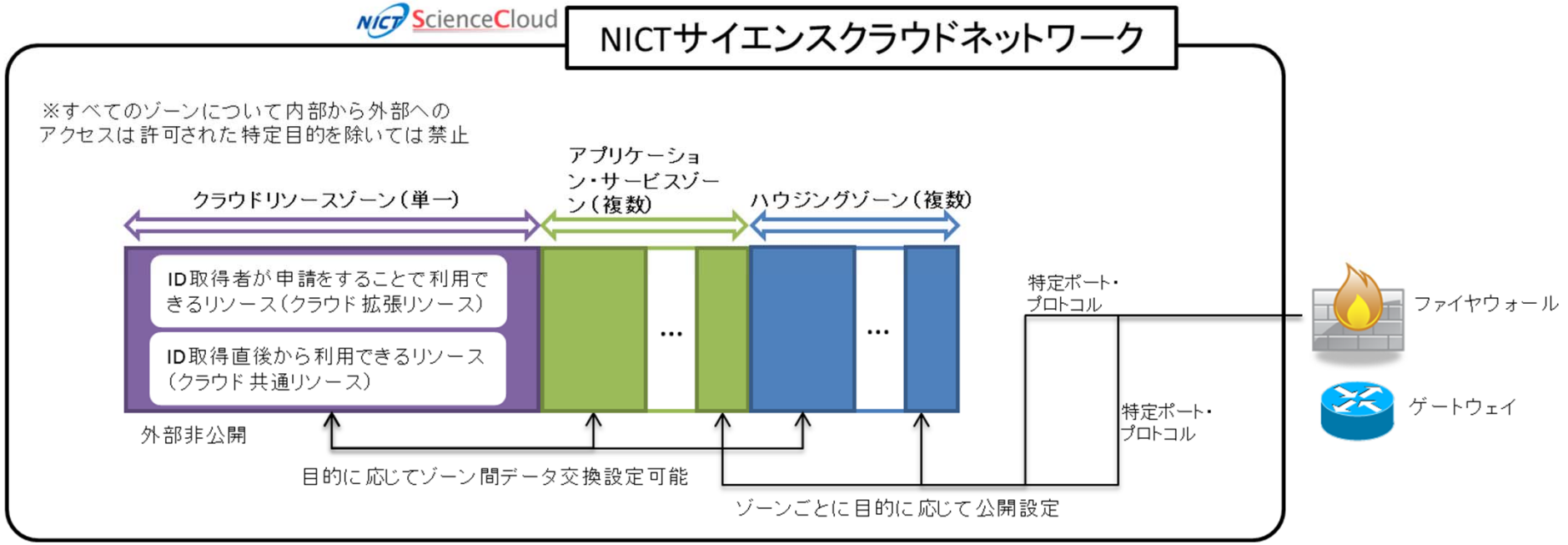
クラウド内・DC間データファイルコピー



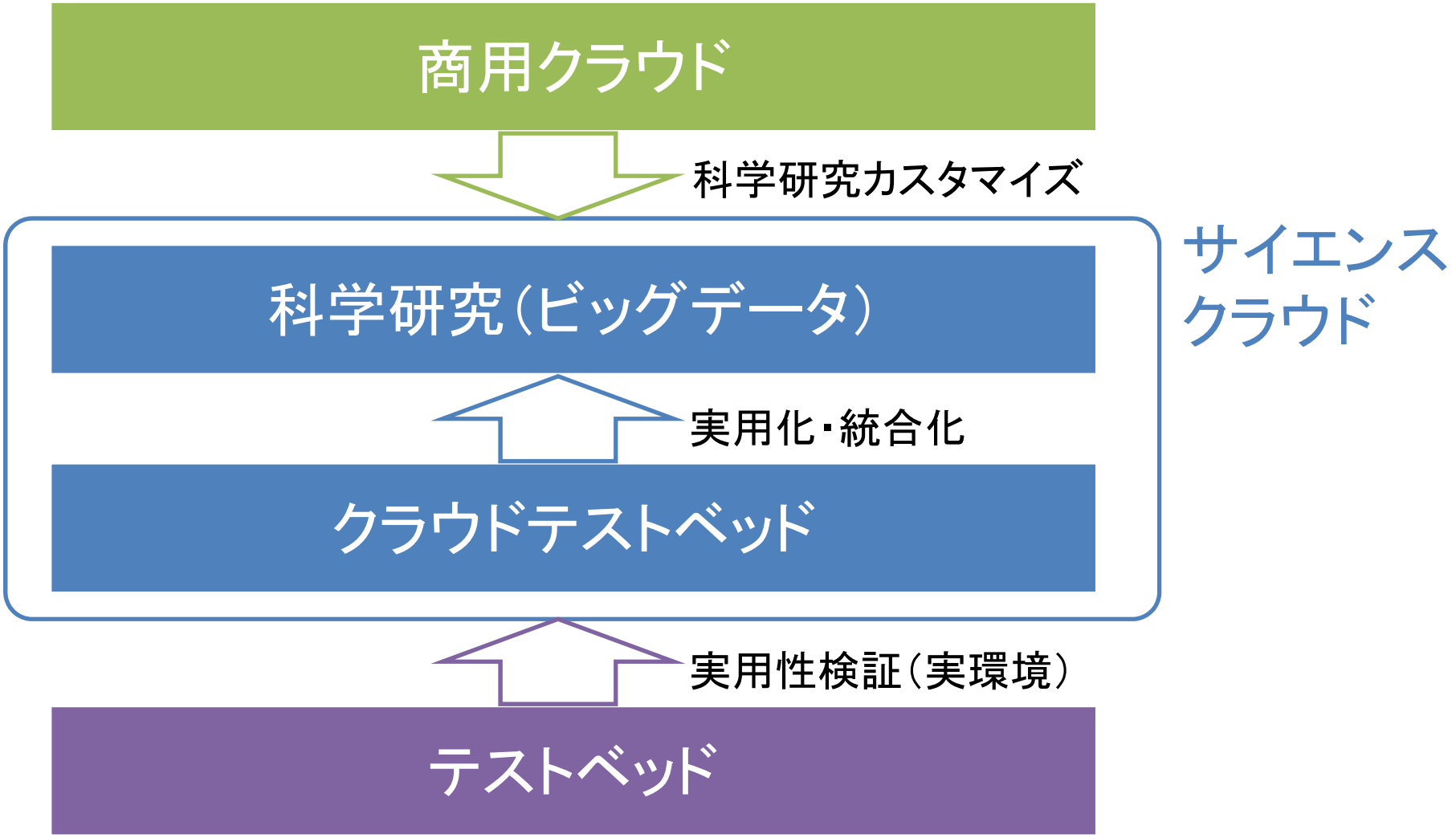
ファイルサイズ	1MB	1MB	100MB	100MB
並列数	10	100	10	100

NICT ScienceCloud

NICTサイエンスクラウドネットワークセグメント



サイエンスクラウドの意味



まとめ

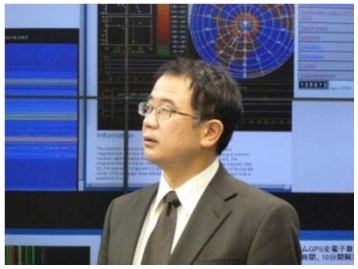
- データ指向型科学
  - 第4のパラダイム（研究基盤＝サイエンスクラウド）
- NICTサイエンスクラウド（2010～）
  - 効率的で大規模な科学研究基盤
  - 民間クラウドとの違い
    - 科学研究環境（アプリケーション）の利活用
    - 科学研究用ビッグデータ処理の柔軟性

http://sc-web.nict.go.jp/

Thank you for your attention!

“Cloud makes nothing new.” Prof. Hideo Miyahara
(ex-president of NICT)

“Cloud makes nothing new,
but makes it possible.” Ken T. Murata



New President of NICT:
Prof. Masao Sakauchi