宇宙科学と大規模可視化
サイエンスクラウドで初めて見えた現象

世界No.1になった「京」コンピュータは一度の計算でPB（ペタバイト）データを計算します。これらのデータを安全に保存する方法はあるのか、大規模データをどのように処理するか、そこから得られる新しい科学は何かなど、NICTサイエンスクラウドが目指すテーマとその成果についてご紹介します。

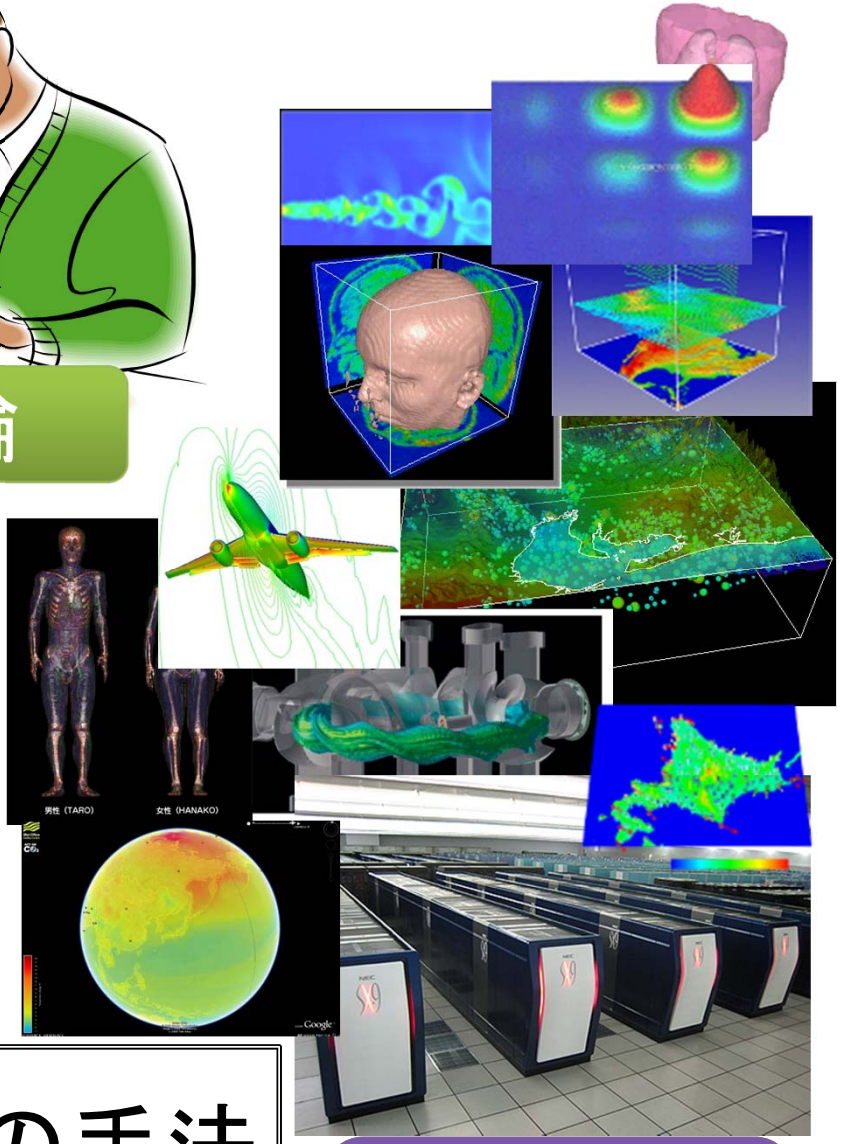
平成24年12月3日
情報通信研究機構
統合データシステム研究開発室
村田 健史（たけし）

①理論

②観測・
実験

科学研究の手法

主要な3つの方法

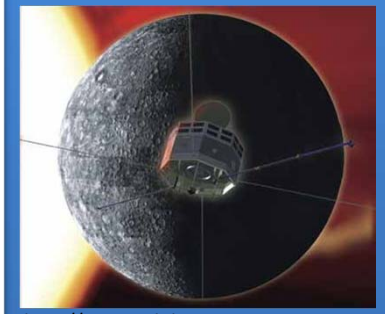


③シミュ
レーション

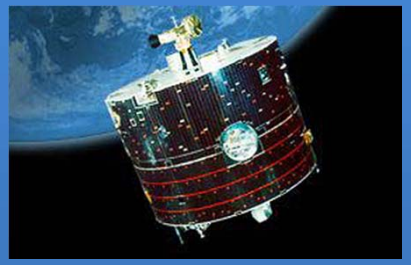
宇宙科学のメガサイエンスプロジェクト(1)

衛星観測と数値シミュレーション

科学衛星観測
Scientific Satellite Missions



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



http://www.isas.jaxa.jp

http://www.isas.jaxa.jp

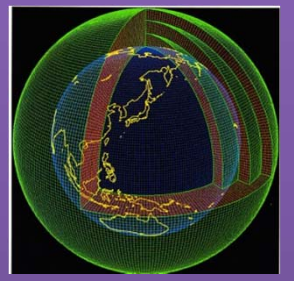
100億円

ペタバイト

ビッグデータ

数値シミュレーション
Super Computer Projects

https://aspara.asahi.com/blog/netnowplus/entry/uSreeI7JOG



http://www.venus.dti.ne.jp/~inoue-m/on_co2_i/f11_cb100.jpg

http://pc.watch.impress.co.jp/img/pcw/docs/212/472/nec.jpg

1,000億円

ペタバイト

ビッグデータ

別々に研究が進められている

メガサイエンスの衝突

http://cosmos21.exblog.jp/

第四の科学研究手法（Jim Gray）

http://research.microsoft.com/en-us/collaboration/fourthparadigm

Microsoft Research

Search Microsoft Research

Videos Projects Publications People Downloads

Home Our Research Connections Careers Hub

About Us | Research in Action | Opportunities | Research Accelerators

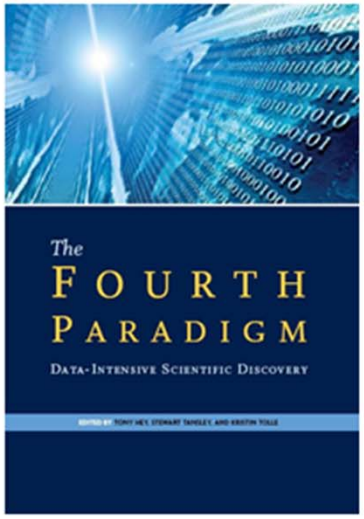
Tell us what you think.
Take our 6-question site survey.

> Microsoft Research Connections > The Fourth Paradigm: Data-Intensive Scientific Discovery

The Fourth Paradigm: Data-Intensive Scientific Discovery

データ指向型科学

Presenting the first broad look at the rapidly emerging field of data -intensive science



Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets.

The speed at which any given scientific discipline advances will depend on how well its researchers collaborate with one another, and with technologists, in areas of eScience such as databases, workflow management, visualization, and cloud computing technologies.

In *The Fourth Paradigm: Data-Intensive Scientific Discovery*, the collection of essays expands on the vision of pioneering computer scientist Jim Gray for a new, fourth paradigm of discovery based on data-intensive science and offers insights into how it can be fully realized.

Download

- Full text, low resolution (6 MB)
- Full text, high resolution (93 MB)
- By chapter and essay

Purchase from Amazon.com

- Paperback
- Kindle version

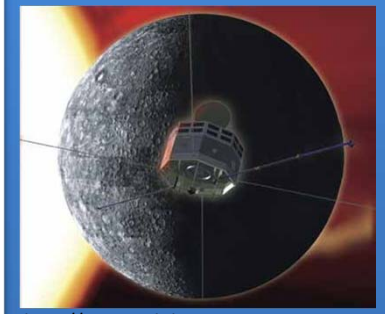
In the News

- Sailing on an Ocean of 0s and 1s (*Science* Magazine)
- A Deluge of Data Shapes a New Era in Computing (*New York Times*)

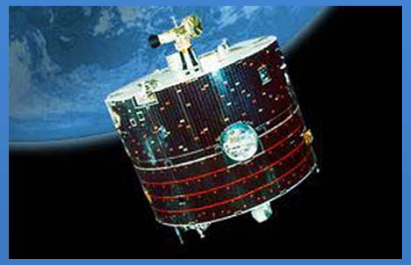
宇宙科学のメガサイエンスプロジェクト(2)

インフォマティクスの役割

科学衛星観測
Scientific Satellite Missions



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



http://www.isas.jaxa.jp

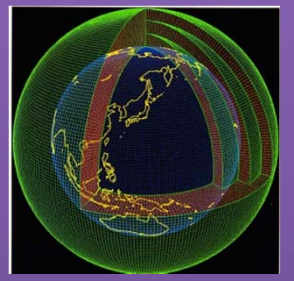
http://www.isas.jaxa.jp

100億円

インフォマティクス

数値シミュレーション
Super Computer Projects

https://aspara.asahi.com/blog/netnowplus/entry/uSreeI7JOG



http://www.venus.dti.ne.jp/~inoue-m/on_co2_i/f11_cb100.jpg

http://pc.watch.impress.co.jp/img/pcw/docs/212/472/nec.jpg

1,000億円

データ指向型科学

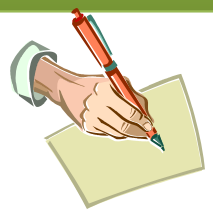
ペタバイト

ビッグデータ

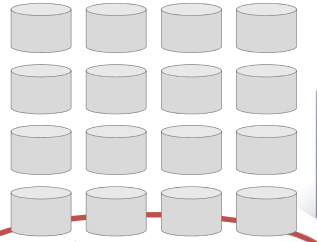
科学を行う方法（宇宙科学の場合）

NICT ScienceCloud

①理論⇒
紙と鉛筆

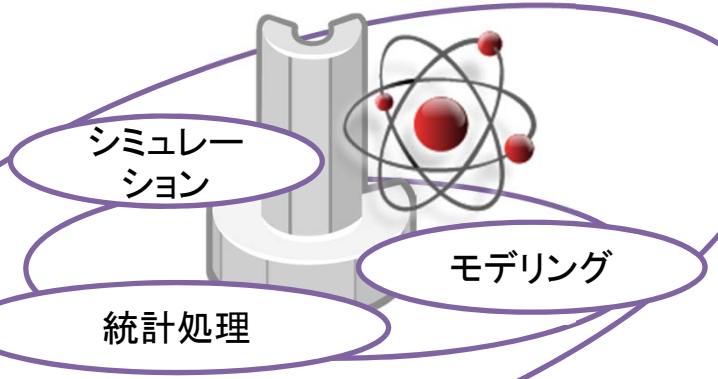


④データ指向型研究（インフォマティクス）⇒クラウド



Data

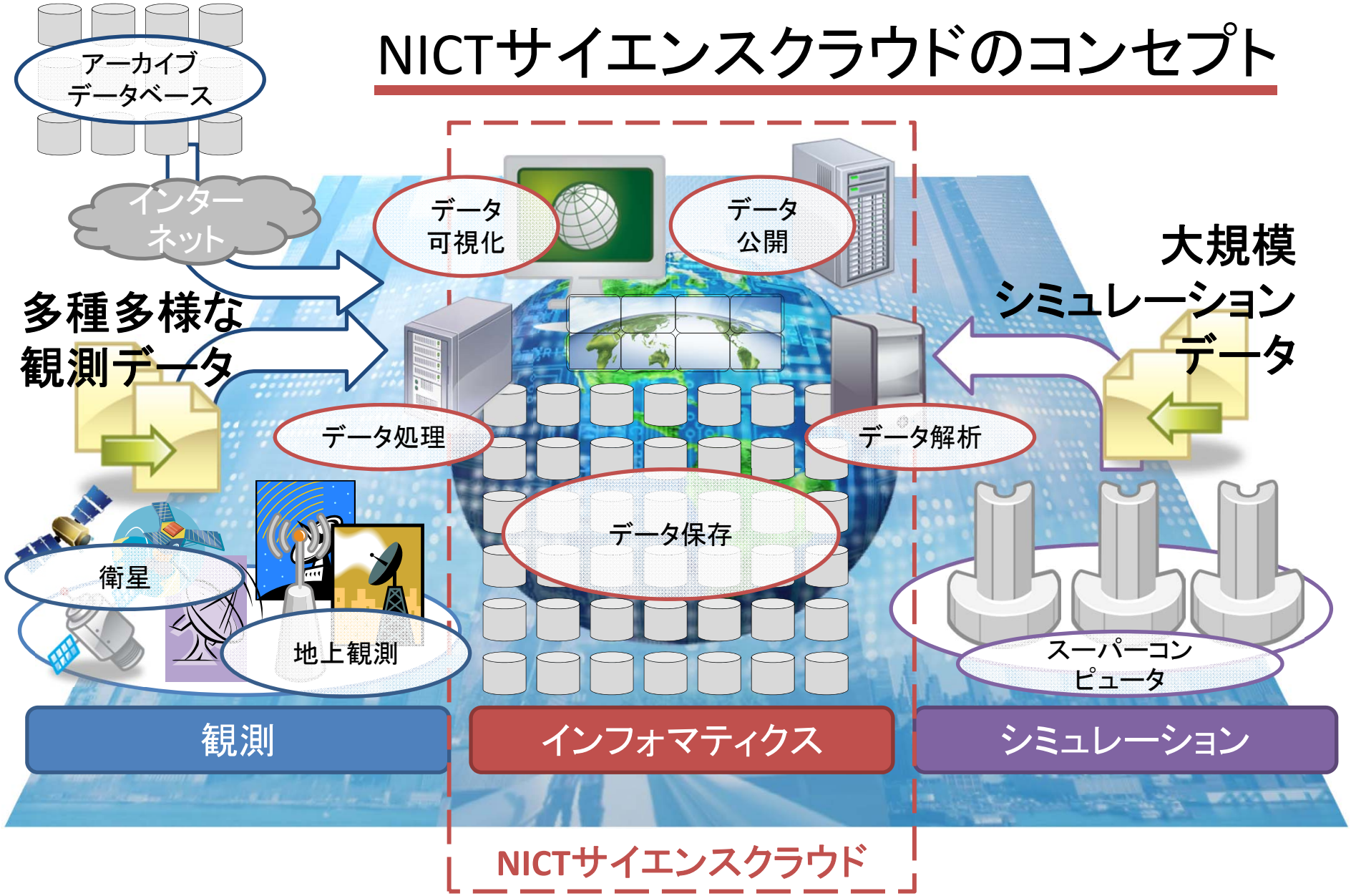
Data



②観測⇒科学衛星・地上観測

③コンピュータシミュレー
ション⇒スパコン

NICTサイエンスクラウドのコンセプト

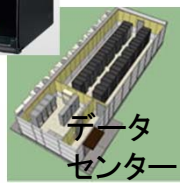


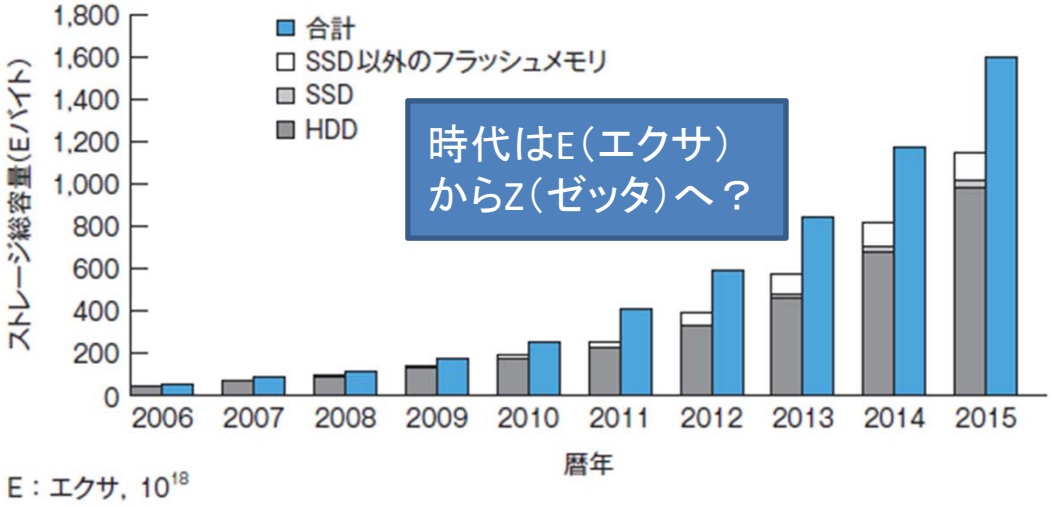
あらゆるデータをクラウド上に！

世界の情報ストレージ総容量/増加する科学データ

世界のストレージ動向

- MB（メガバイト）
- GB（ギガバイト）
- TB（テラバイト）
- PB（ペタバイト）
- EB（エクサバイト）

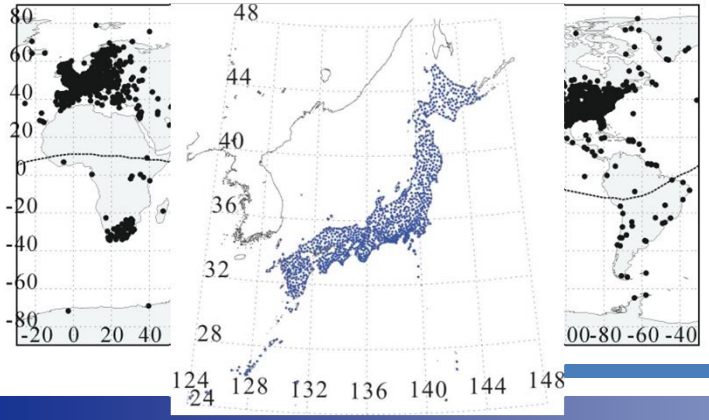


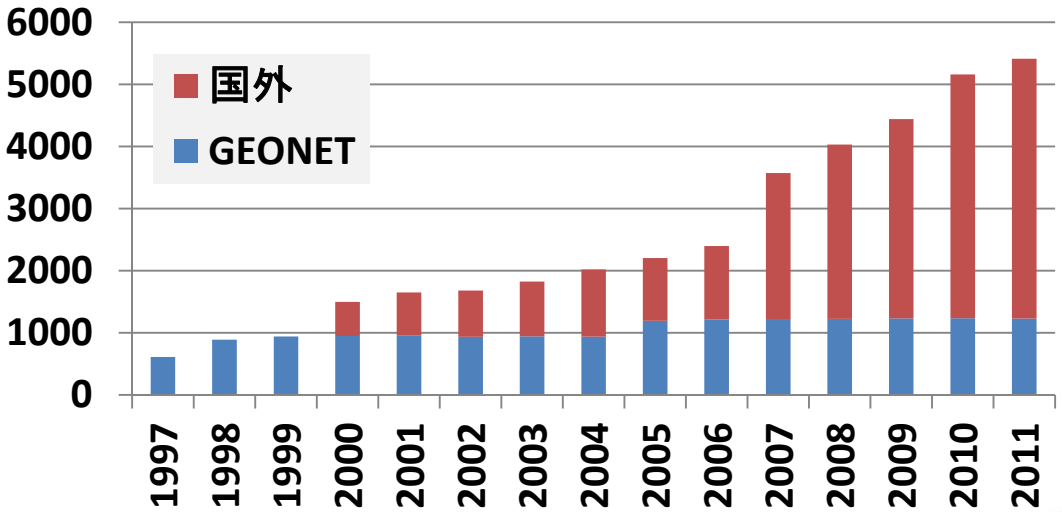


http://www.toshiba.co.jp/tech/review/2011/08/66_08pdf/b02.pdf

科学研究分野の例

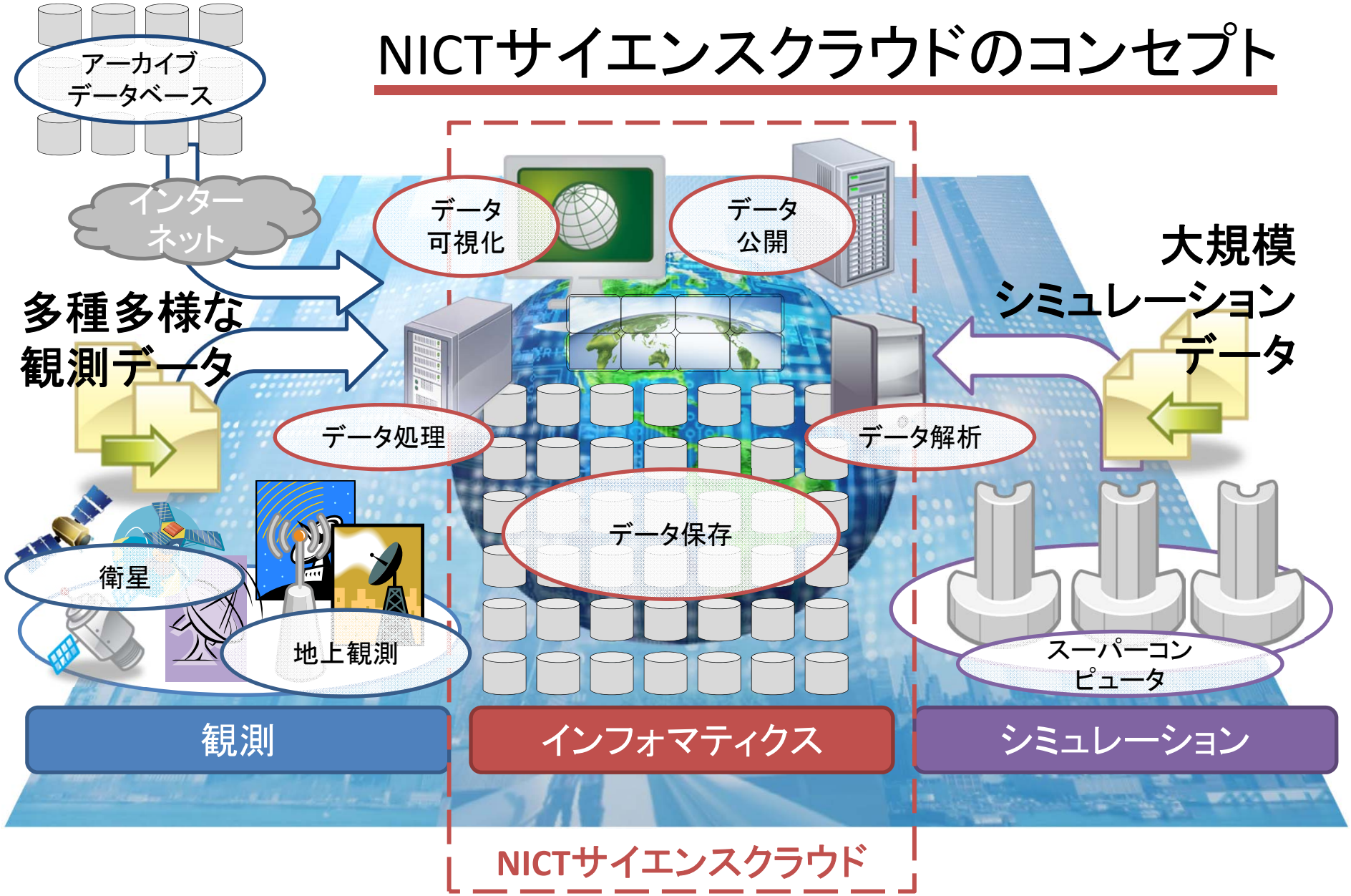
GPS公開データ





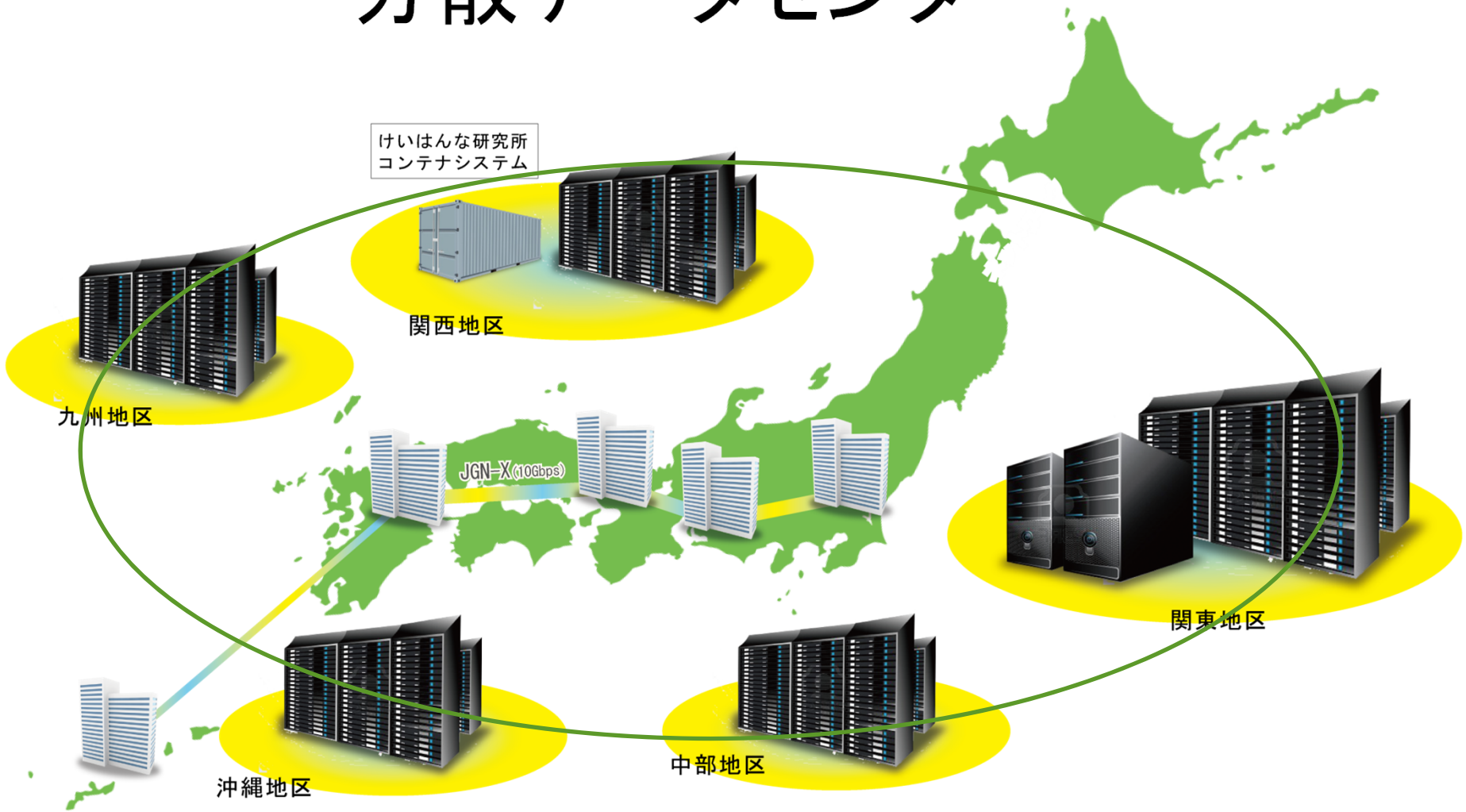
津川（NICT）らによる

NICTサイエンスクラウドのコンセプト

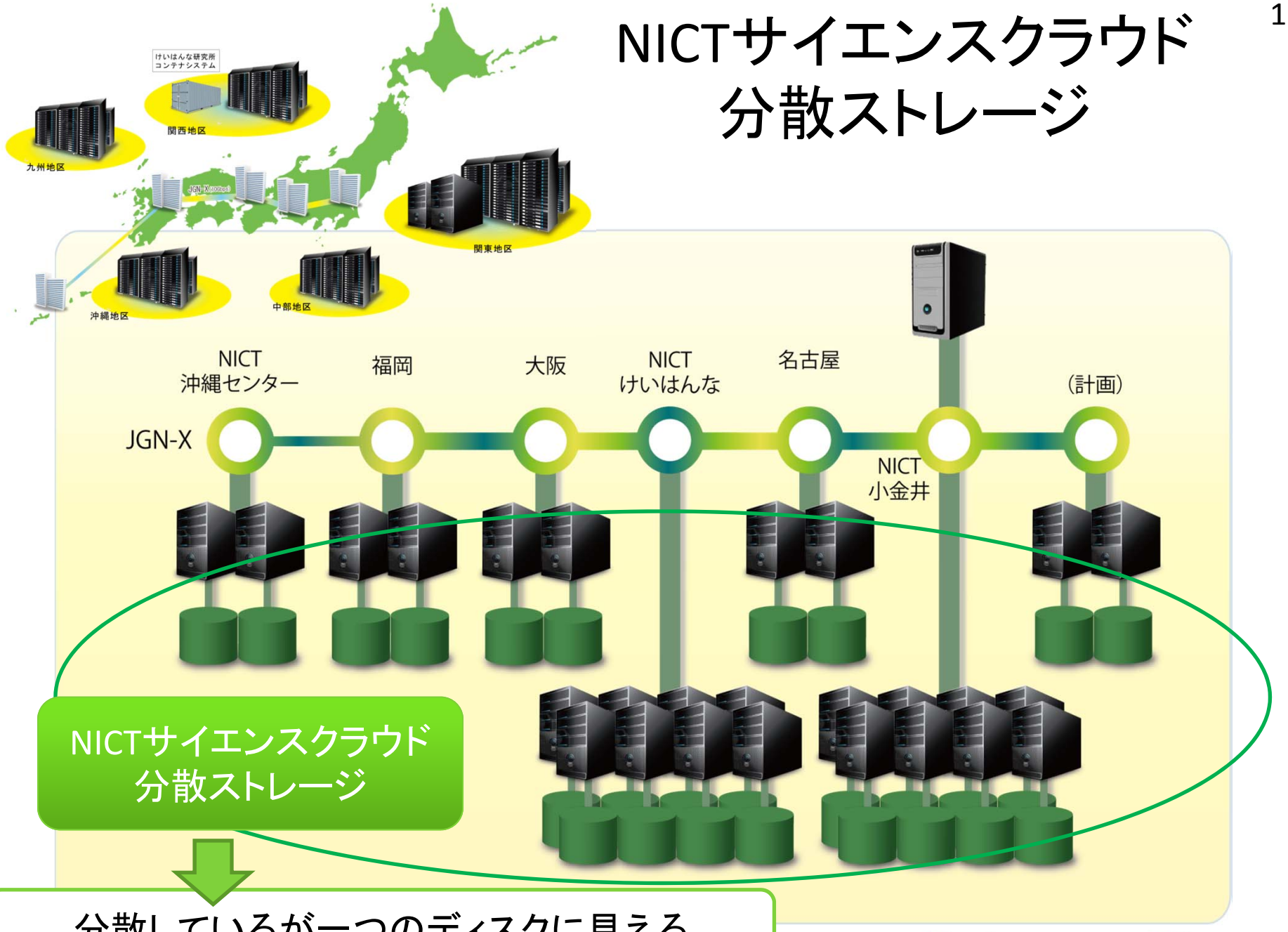


あらゆるデータをクラウド上に！

NICTサイエンスクラウド
分散データセンター



NICTサイエンスクラウド 分散ストレージ



科学を行う方法（宇宙科学の場合）

ビッグデータ

④データ指向型研究（インフォマティクス）⇒クラウド

①理論⇒紙と鉛筆

データ指向型科学を実現するには…

どうやってデータを集めるか？
どうやってデータを保存するか？
どうやってデータを読み出すか？
どうやってデータを処理（解析）するか？
どうやってデータを公開するか？

②観測⇒科学衛星・地上観測

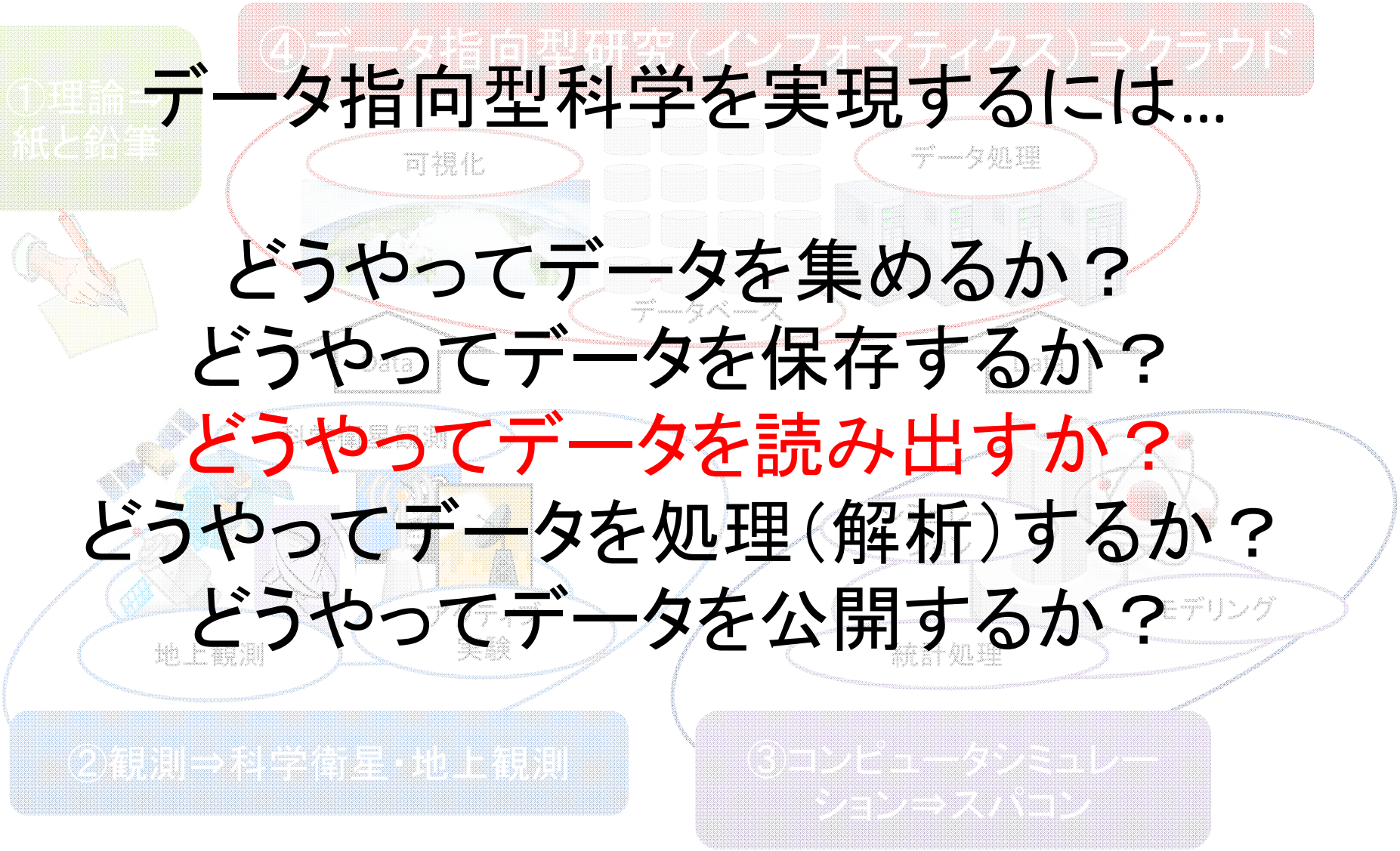
③コンピュータシミュレーション⇒スパコン

科学を行う方法（宇宙科学の場合）

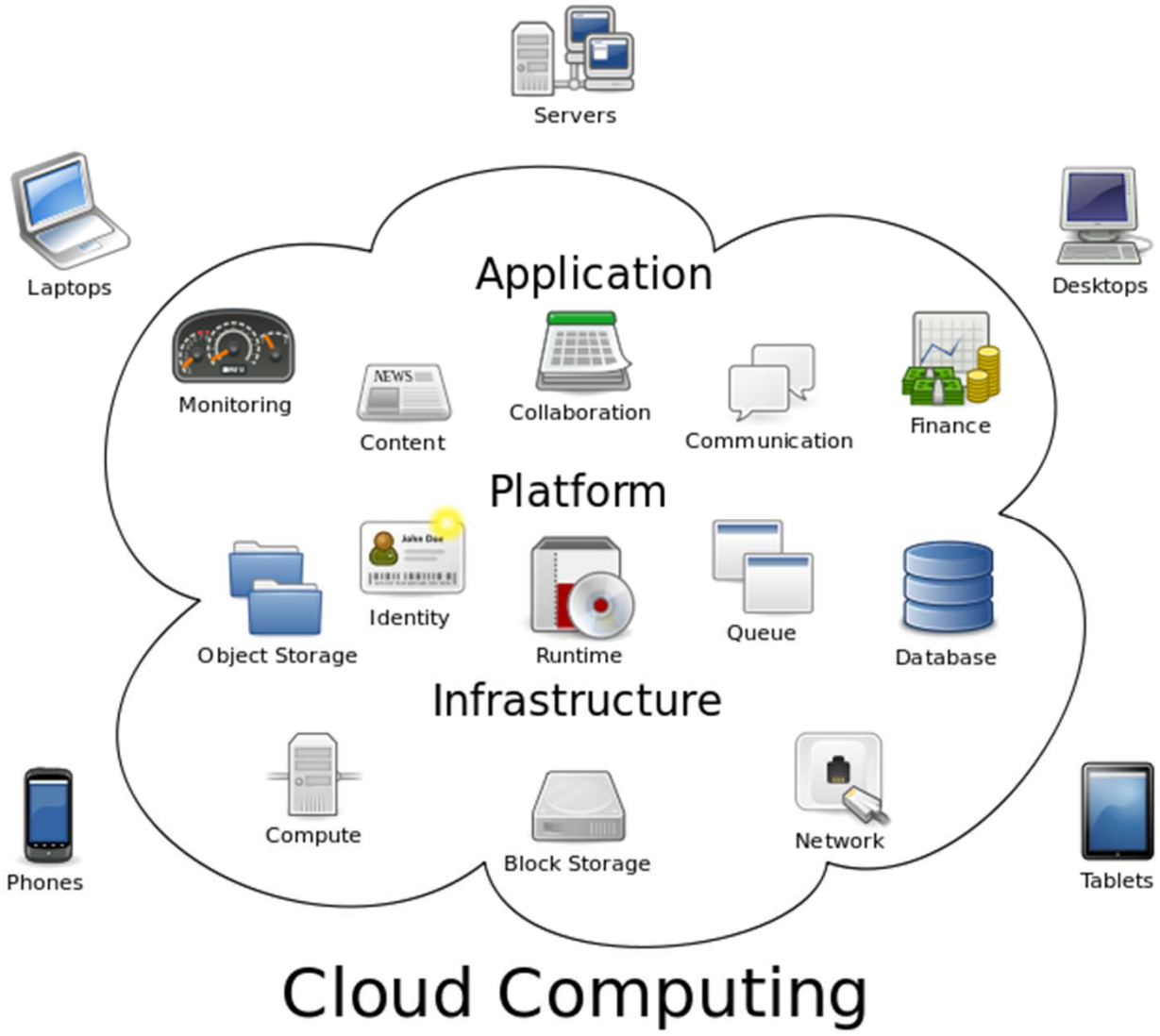
ビッグデータ

データ指向型科学を実現するには…

どうやってデータを集めるか？
どうやってデータを保存するか？
どうやってデータを読み出すか？
どうやってデータを処理（解析）するか？
どうやってデータを公開するか？



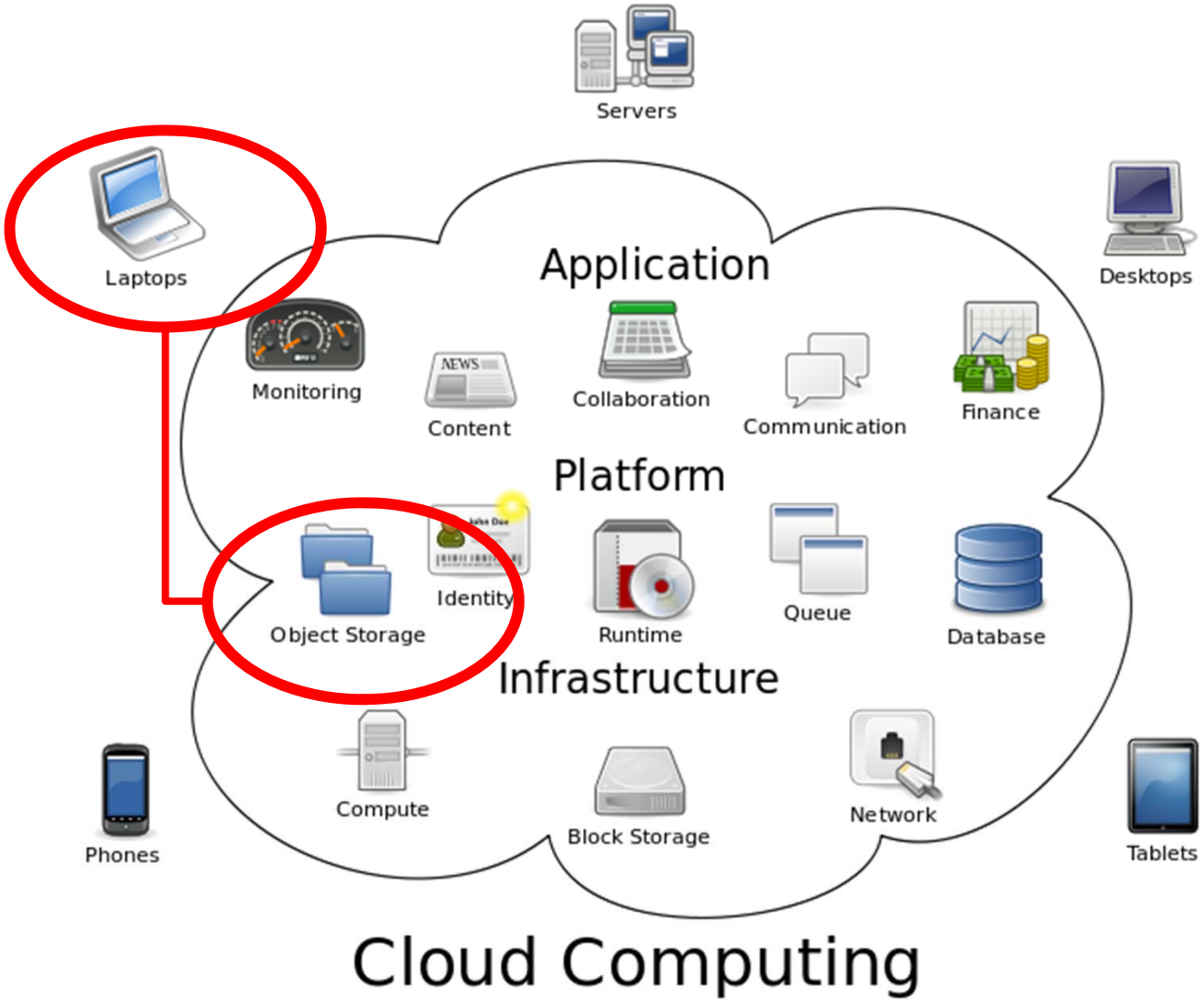
クラウドとは？（Wikipediaより）



クラウドコンピューティング（英: cloud computing）とは、ネットワーク、特にインターネットをベースとしたコンピュータの利用形態である。ユーザーはコンピュータ処理をネットワーク経由で、サービスとして利用する。

http://ja.wikipedia.org/wiki/%E3%83%95%E3%82%A1%E3%82%A4%E3%83%AB:Cloud_computing.svg

NICTサイエンスクラウド利用例

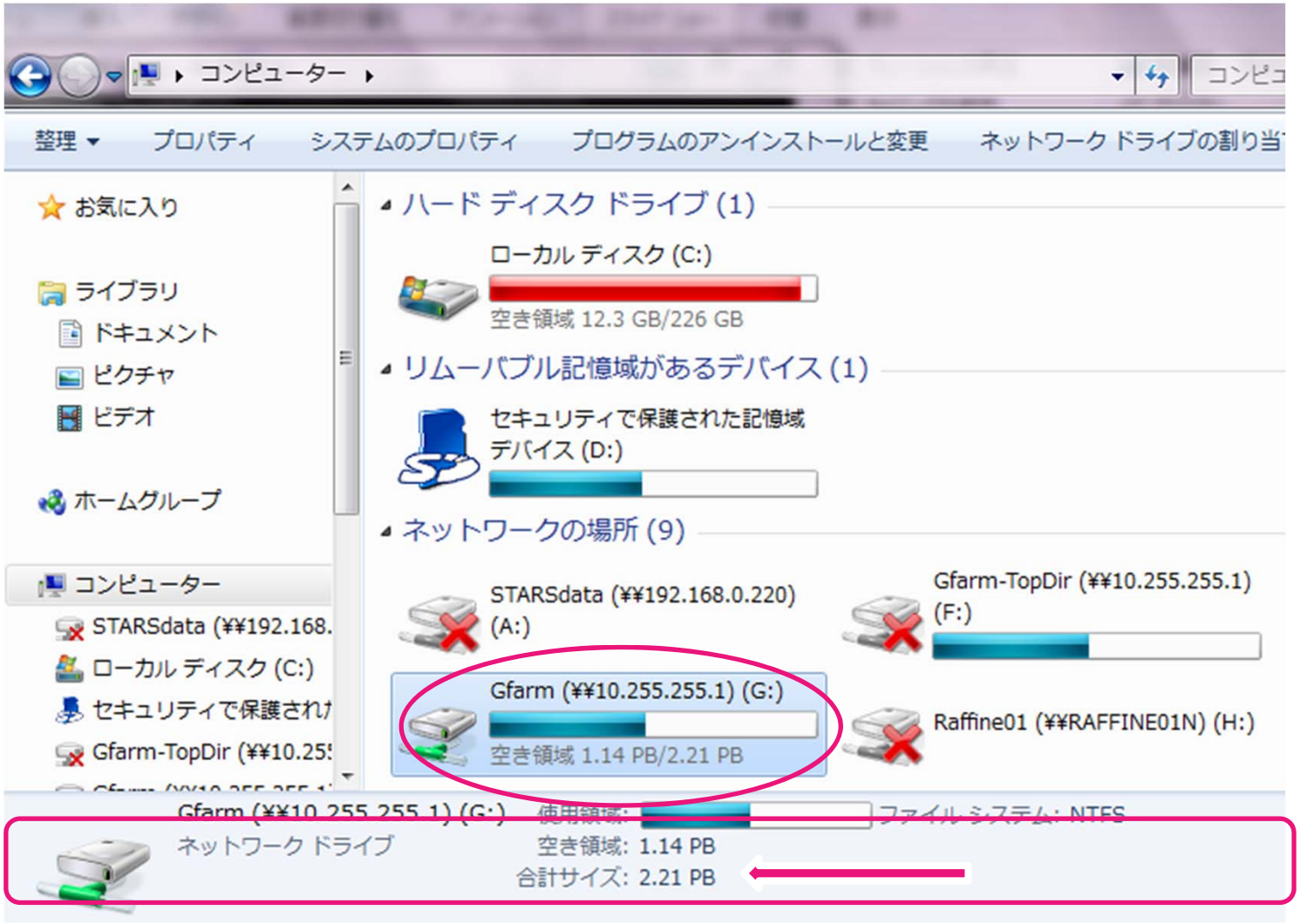


Laptopパソコンからクラウドストレージ（NICTサイエンスクラウド）を利用

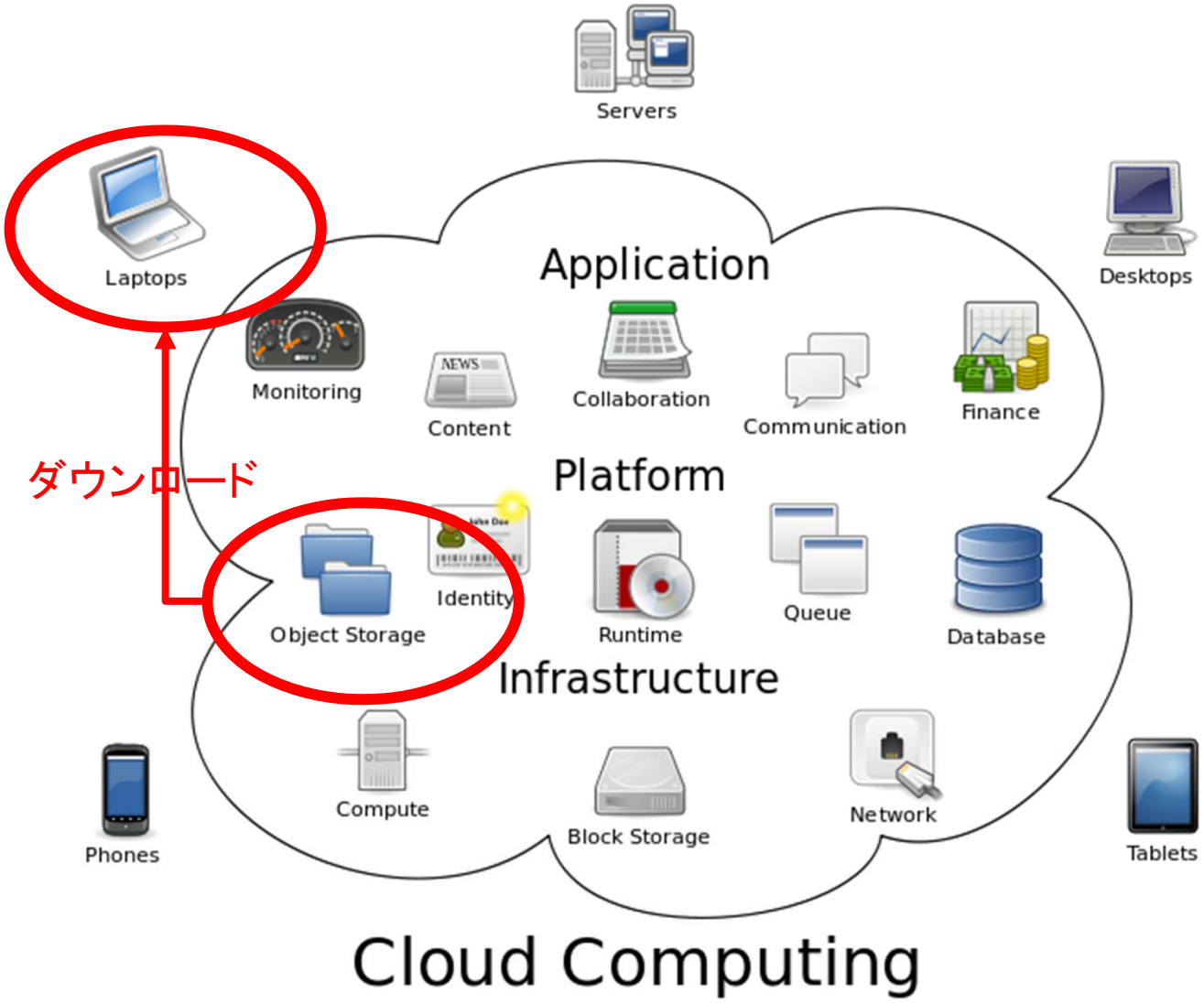
ストレージサイズが問題ではない！

http://ja.wikipedia.org/wiki/%E3%83%95%E3%82%A1%E3%82%A4%E3%83%AB:Cloud_computing.svg

私のラップトップPC


コンピューター
整理
プロパティ
システムのプロパティ
プログラムのアンインストールと変更
ネットワーク ドライブの割り当
お気に入り
ライブラリ
ドキュメント
ピクチャ
ビデオ
ホームグループ
コンピューター
STARSdata (¥¥192.168.
ローカル ディスク (C:)
Gfarm-TopDir (¥¥10.25
ハード ディスク ドライブ (1)
ローカル ディスク (C:)
空き領域 12.3 GB/226 GB
リムーバブル記憶域があるデバイス (1)
セキュリティで保護された記憶域
デバイス (D:)
ネットワークの場所 (9)
STARSdata (¥¥192.168.0.220)
(A:)
Gfarm-TopDir (¥¥10.255.255.1)
(F:)
Gfarm (¥¥10.255.255.1) (G:)
空き領域 1.14 PB/2.21 PB
Raffine01 (¥¥RAFFINE01N) (H:)
Gfarm (¥¥10.255.255.1) (G:)
ファイル システム: NTFS
ネットワーク ドライブ
空き領域: 1.14 PB
合計サイズ: 2.21 PB

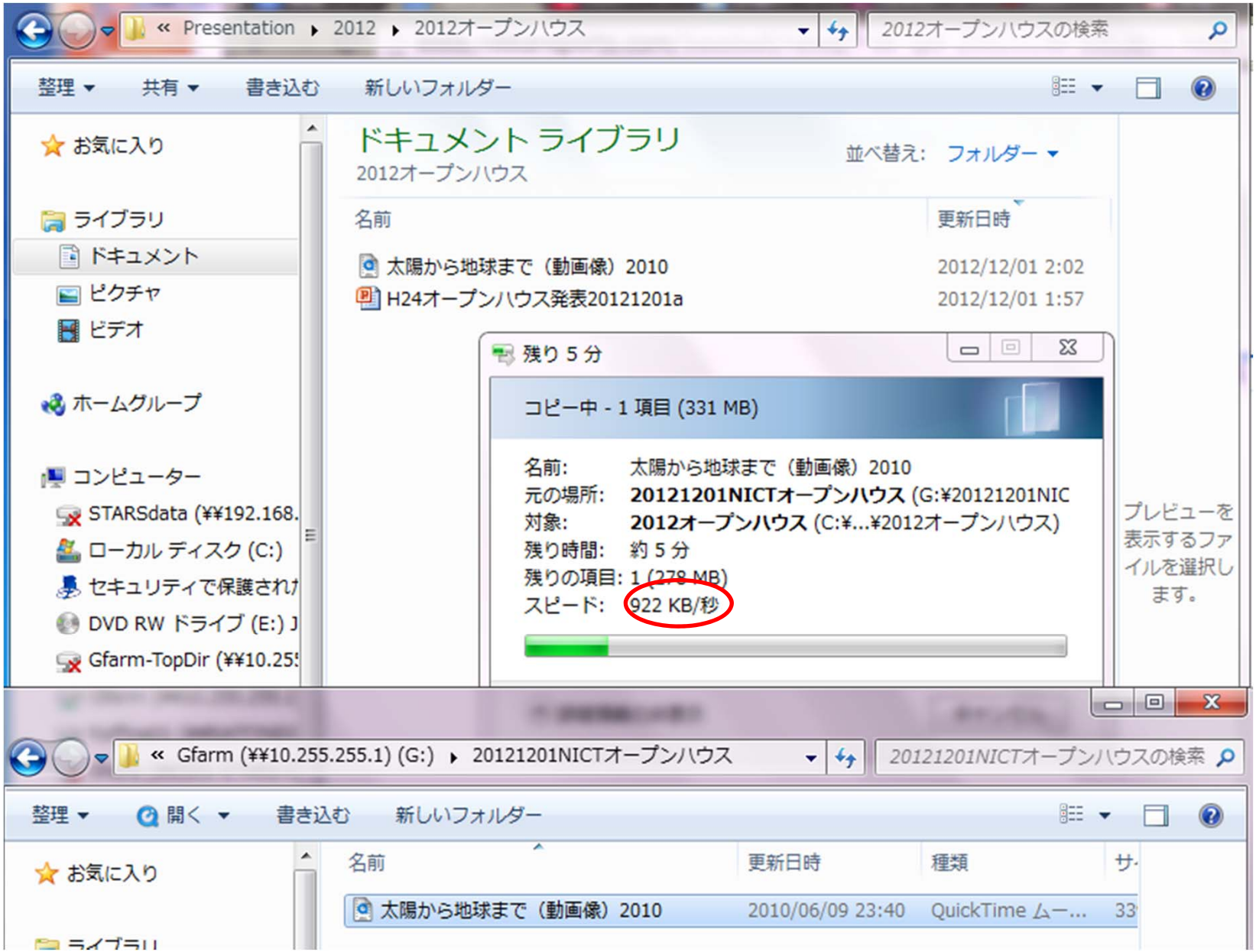
NICTサイエンスクラウド利用例



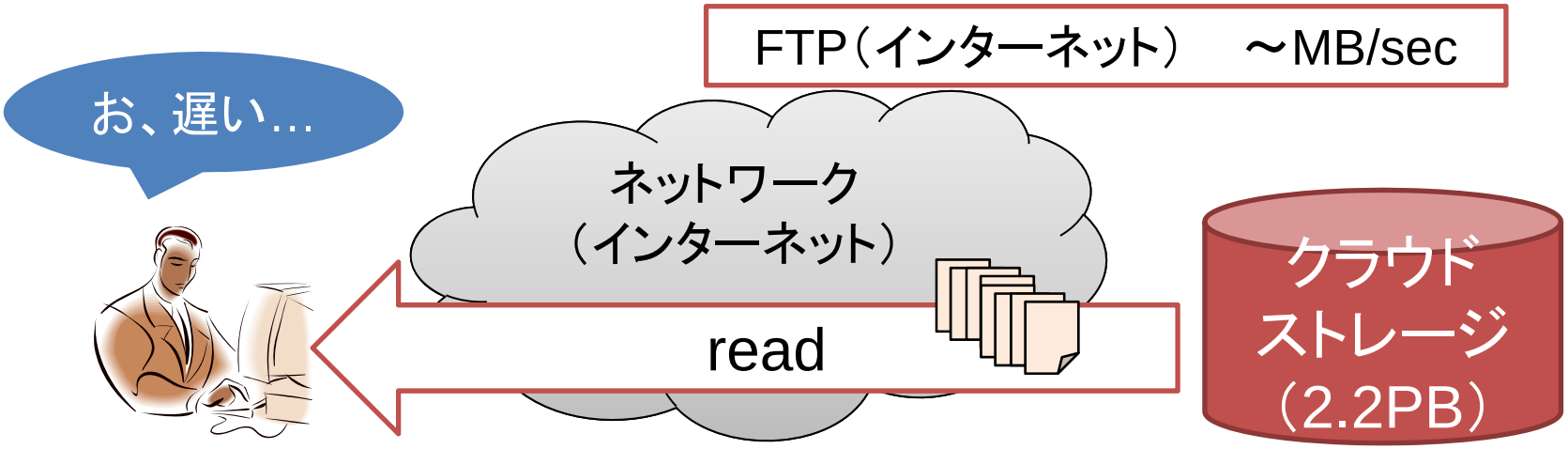
Laptopパソコンからクラウドストレージ（NICTサイエンスクラウド）を利用

動画像（330MBをクラウドストレージからダウンロード）

http://ja.wikipedia.org/wiki/%E3%83%95%E3%82%A1%E3%82%A4%E3%83%AB:Cloud_computing.svg


« Presentation ▸ 2012 ▸ 2012オープンハウス
2012オープンハウスの検索
整理
共有
書き込む
新しいフォルダー
お気に入り
ライブラリ
ドキュメント
ピクチャ
ビデオ
ホームグループ
コンピューター
STARSdata (¥¥192.168.
ローカル ディスク (C:)
セキュリティで保護された
DVD RW ドライブ (E:) J
Gfarm-TopDir (¥¥10.25
ドキュメント ライブラリ
2012オープンハウス
並べ替え: フォルダー
名前
更新日時
太陽から地球まで（動画像）2010
2012/12/01 2:02
H24オープンハウス発表20121201a
2012/12/01 1:57
残り 5 分
コピー中 - 1 項目 (331 MB)
名前: 太陽から地球まで（動画像）2010
元の場所: 20121201NICTオープンハウス (G:¥20121201NIC
対象: 2012オープンハウス (C:¥...¥2012オープンハウス)
残り時間: 約 5 分
残りの項目: 1 (278 MB)
スピード: 922 KB/秒
プレビューを表示するファイルを選択します。
« Gfarm (¥¥10.255.255.1) (G:) ▸ 20121201NICTオープンハウス
20121201NICTオープンハウスの検索
整理
開く
書き込む
新しいフォルダー
お気に入り
名前
更新日時
種類
太陽から地球まで（動画像）2010
2010/06/09 23:40
QuickTime ムー...

クラウドストレージを科学研究で使うと…



一般的なインターネットが1MB/secと仮定として…

全てのデータファイルを
ダウンロード（2.2PB）

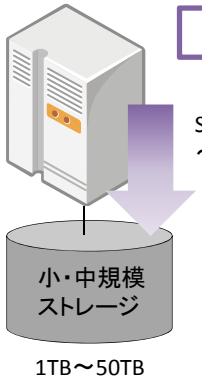
2,200,000,000秒 =70年 ！

NICT ScienceCloud

クラウドストレージを高速に利用したい

データ書き込み方法

ローカルディスク書き込み



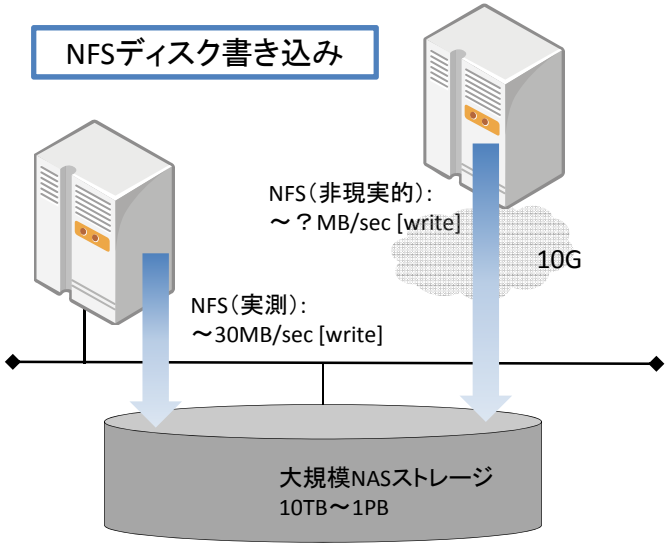
SATA3（実測）:
～400MB/sec [write]

SATA3の登場で、I/Oは400MB/secまで高速化された。ただし、ローカルディスクサイズにはサイズの限界があるため、100TB以上のデータは書き込めない。また、保存ファイルの冗長化にも不安が残る。

データ書き込み時間の比較（100TB）

- ローカルディスク書き込み時間
  - 100TBのローカルディスクは（原則）存在しない
  - 250000秒=約3日
- NFS（NAS）ディスク書き込み時間
  - 約3000000秒=約38日
- 仮想分散ディスク書き込み
  - バンド幅（10G）の活用を仮定
  - 10000秒=3時間弱

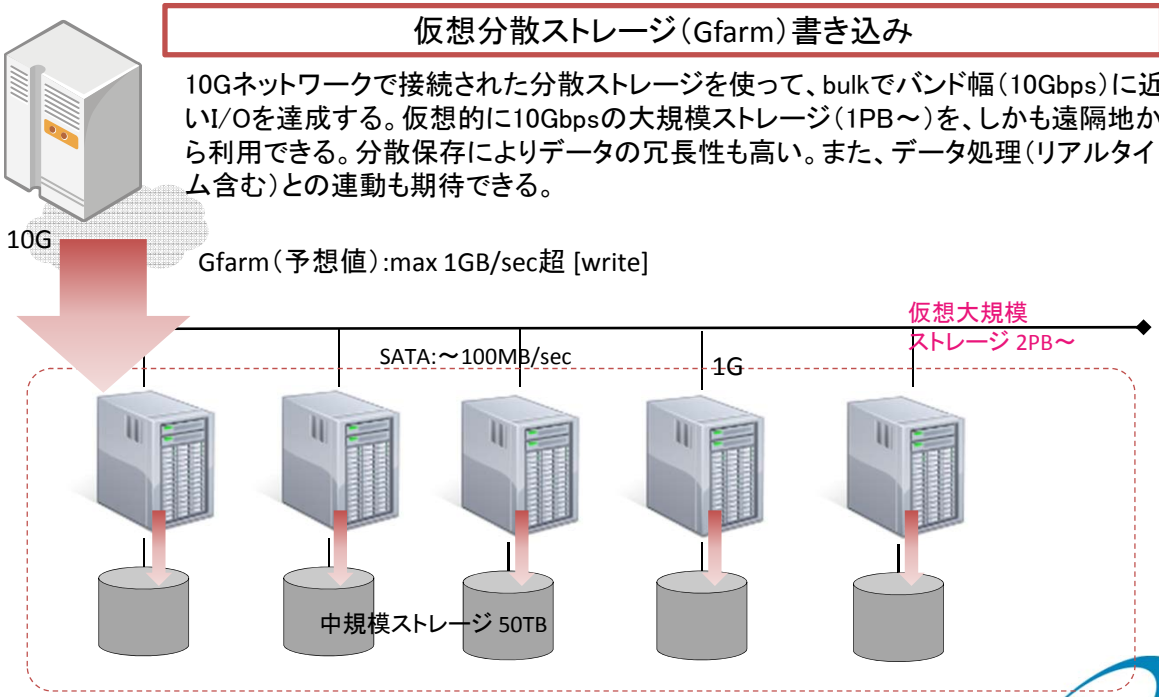
NFSディスク書き込み



NFSの高速化は容易ではなく、高速ネットワークを使っても50MB/secを超えるのは難しい。また、長距離ネットワーク越えのディスクマウントは現実的ではない。

仮想分散ストレージ（Gfarm）書き込み

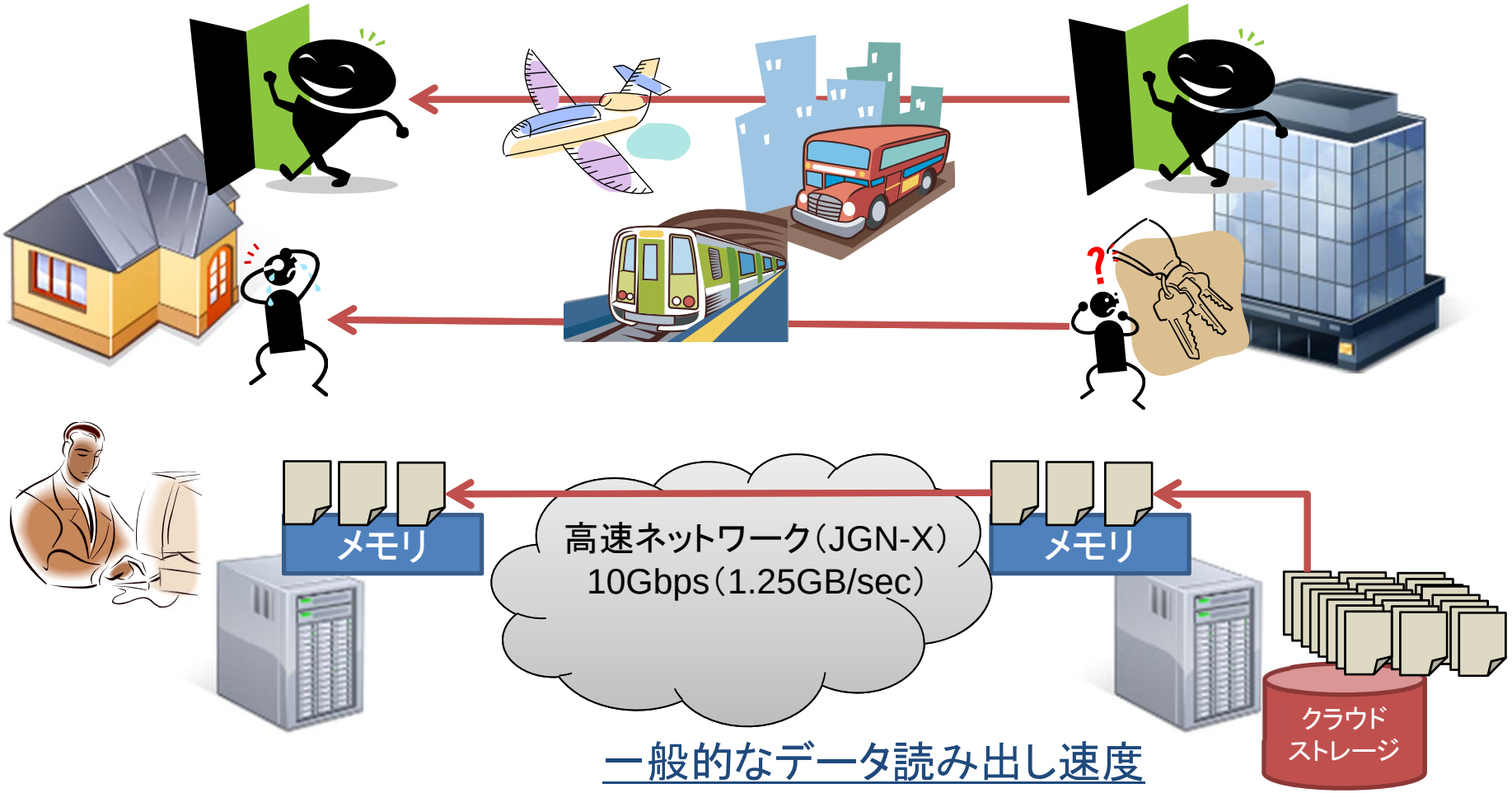
10Gネットワークで接続された分散ストレージを使って、bulkでバンド幅（10Gbps）に近いI/Oを達成する。仮想的に10Gbpsの大規模ストレージ（1PB～）を、しかも遠隔地から利用できる。分散保存によりデータの冗長性も高い。また、データ処理（リアルタイム含む）との連動も期待できる。



阪大レーダ：約30MB/sec　宇宙天気シミュレーション（書き込み）：40MB/sec以上

クラウドストレージの高速化

遅いのはネットワークか？



一般的なデータ読み出し速度

- NAS　400Mbps（50MB/sec)
- SSD disk　1.8Gbps（225MB/sec）
- （SATA3/RAID5　3.8Gbps）
- （RAM disk　3.6Gbps）

NICT ScienceCloud

日米間高速ストレージ実験①（2012年11月）

PC@アメリカ

目標
10Gpbsでデータを読む！

「2000個（200GB）のファイルをください。」

日米間10Gbpsの専用回線

NICT
サイエンスクラウド

指示

制御コンピュータ

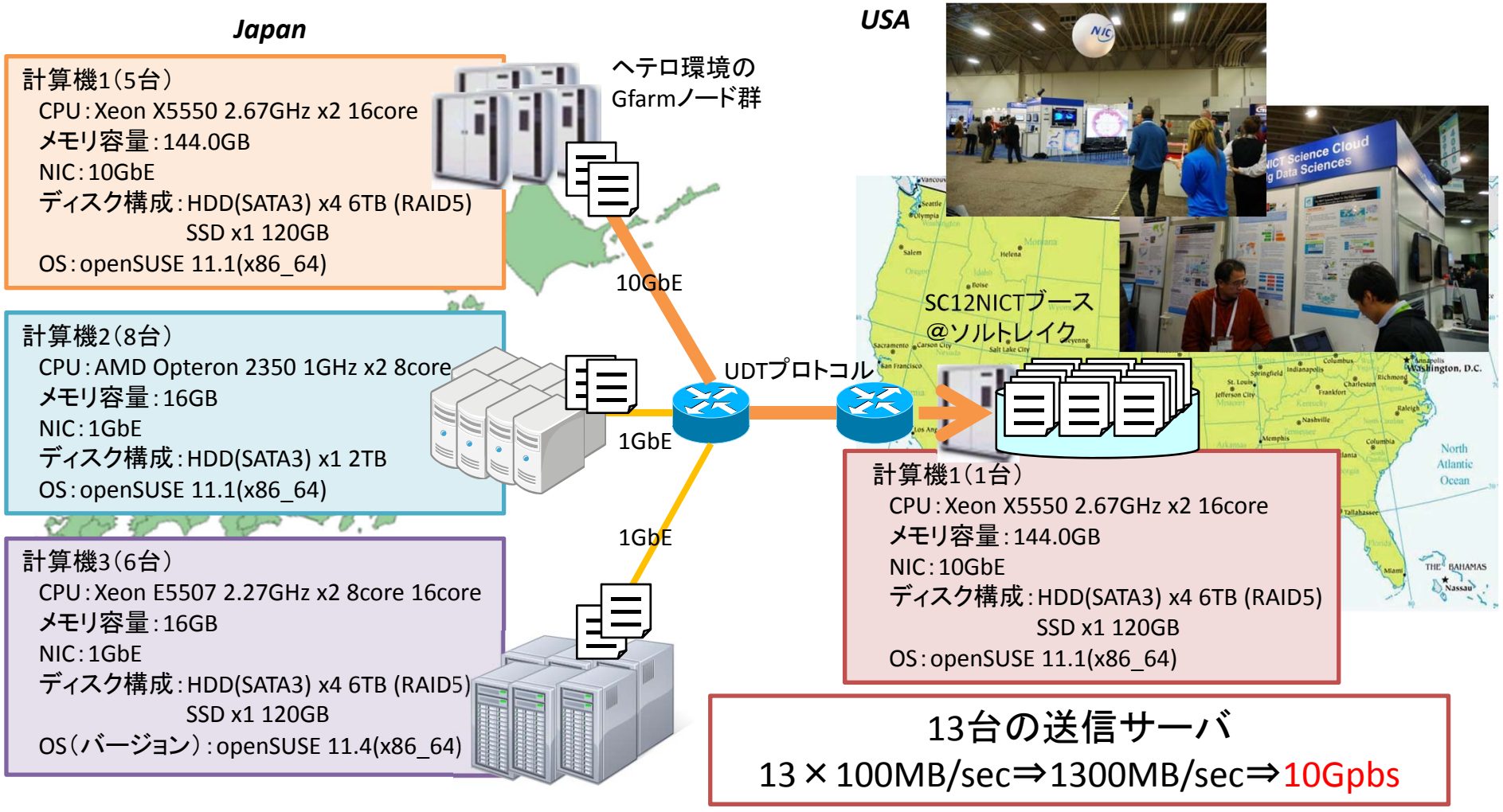
13台の送信サーバ: 13×100MB/sec⇒1300MB/sec⇒約10Gpbs

NICT ScienceCloud

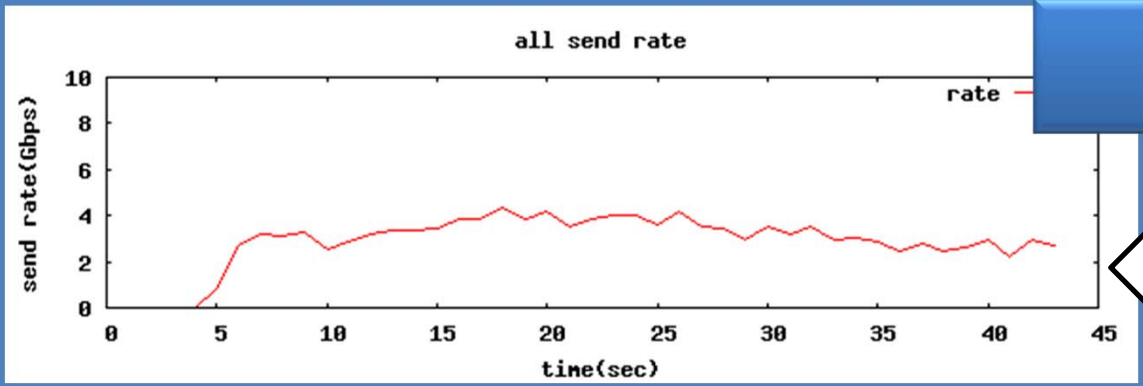
日米間高速ストレージ実験②（2012年11月）

NICT ScienceCloud

- 遠隔地からクラウドストーレジを自分のストレージのように仮想利用
- 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の複数ファイルを1台のクライアント計算機＠米国から読み出し



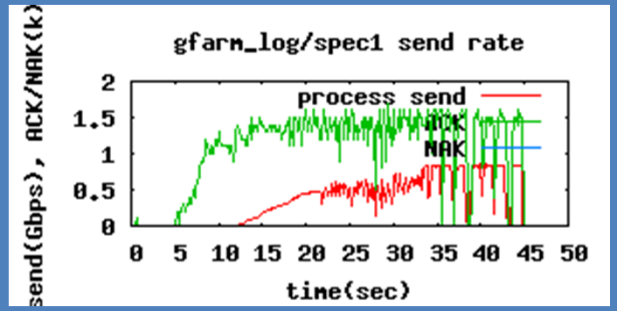
8セッションの場合

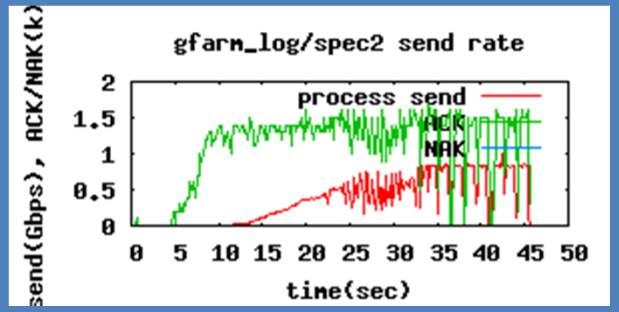


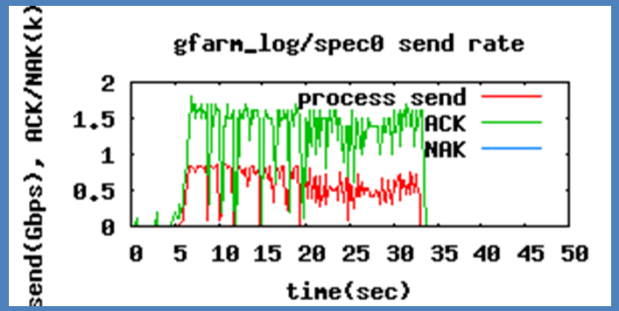
Local Disk (Read)

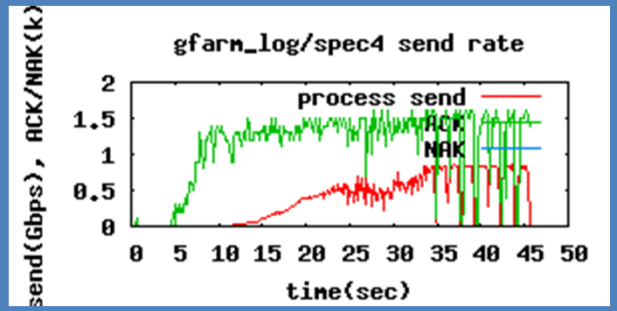
- SSD disk -> 1.8Gbps
- SATA3/RAID5 -> 3.8Gbps
- RAM disk -> 3.6Gbps

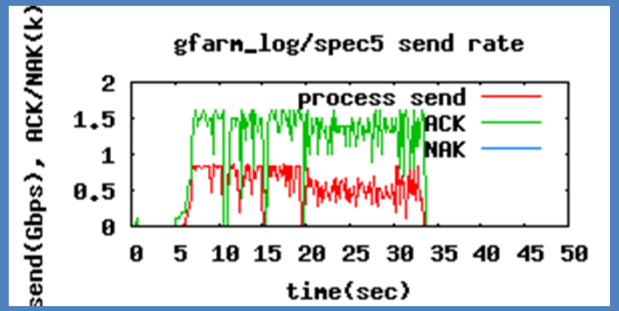
- 125MB/150 files
- 8 sessions
- read

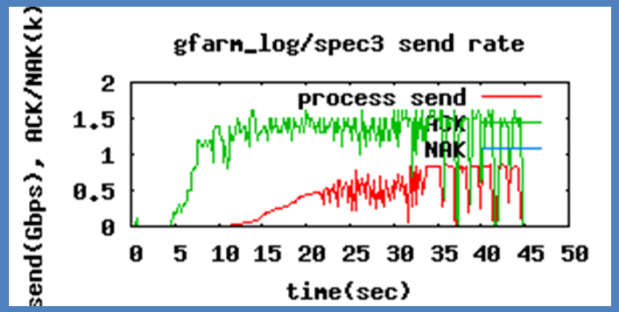


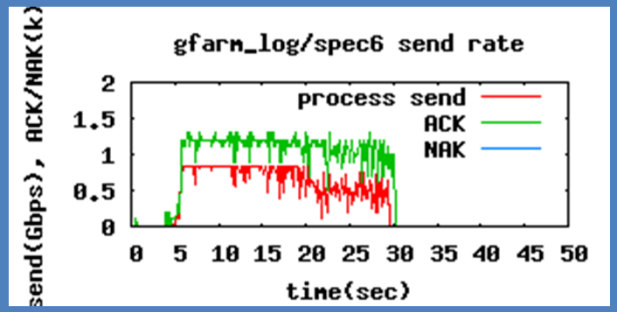


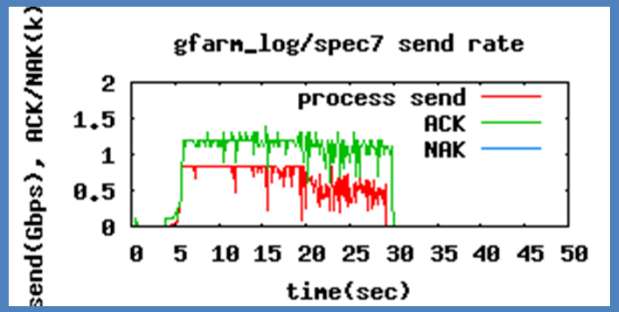












日米間高速ストレージ実験結果（2012年11月）

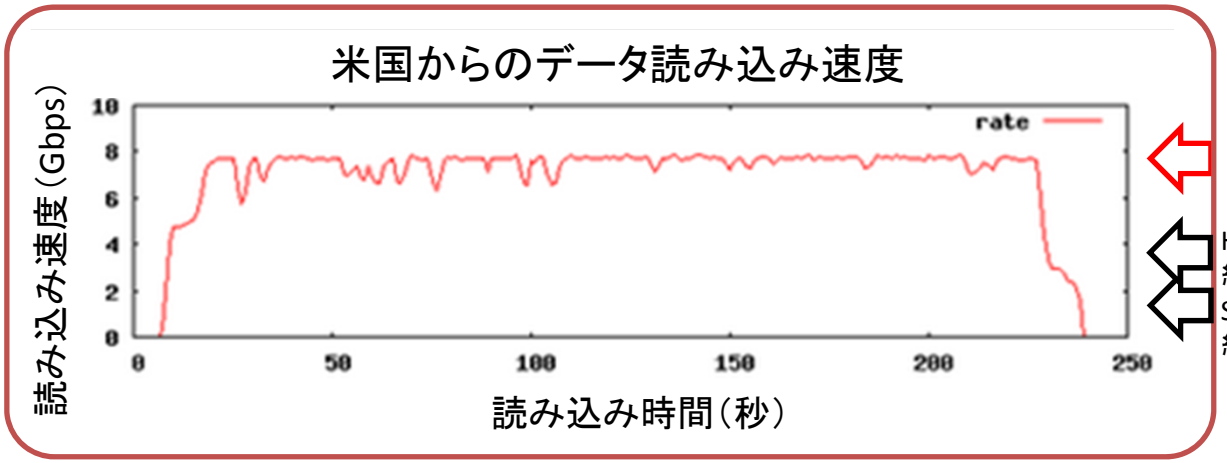
NICT ScienceCloud

6.95Gbps（平均速度）

※参考：単体ディスクI/O（Read）性能

SSD：約1.76Gbps

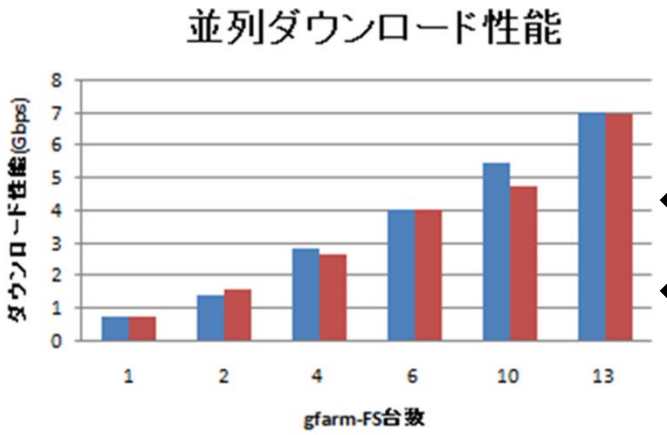
HDD（SATA3・RAID5）：約3.78Gbps



約7Gbps

HDD（SATA3,RAID5）：
約3.78Gbps
SSD：
約1.76Gbps

208GB（DVD約44枚分）のデータファイルを日米間で240秒（4分）で伝送した。

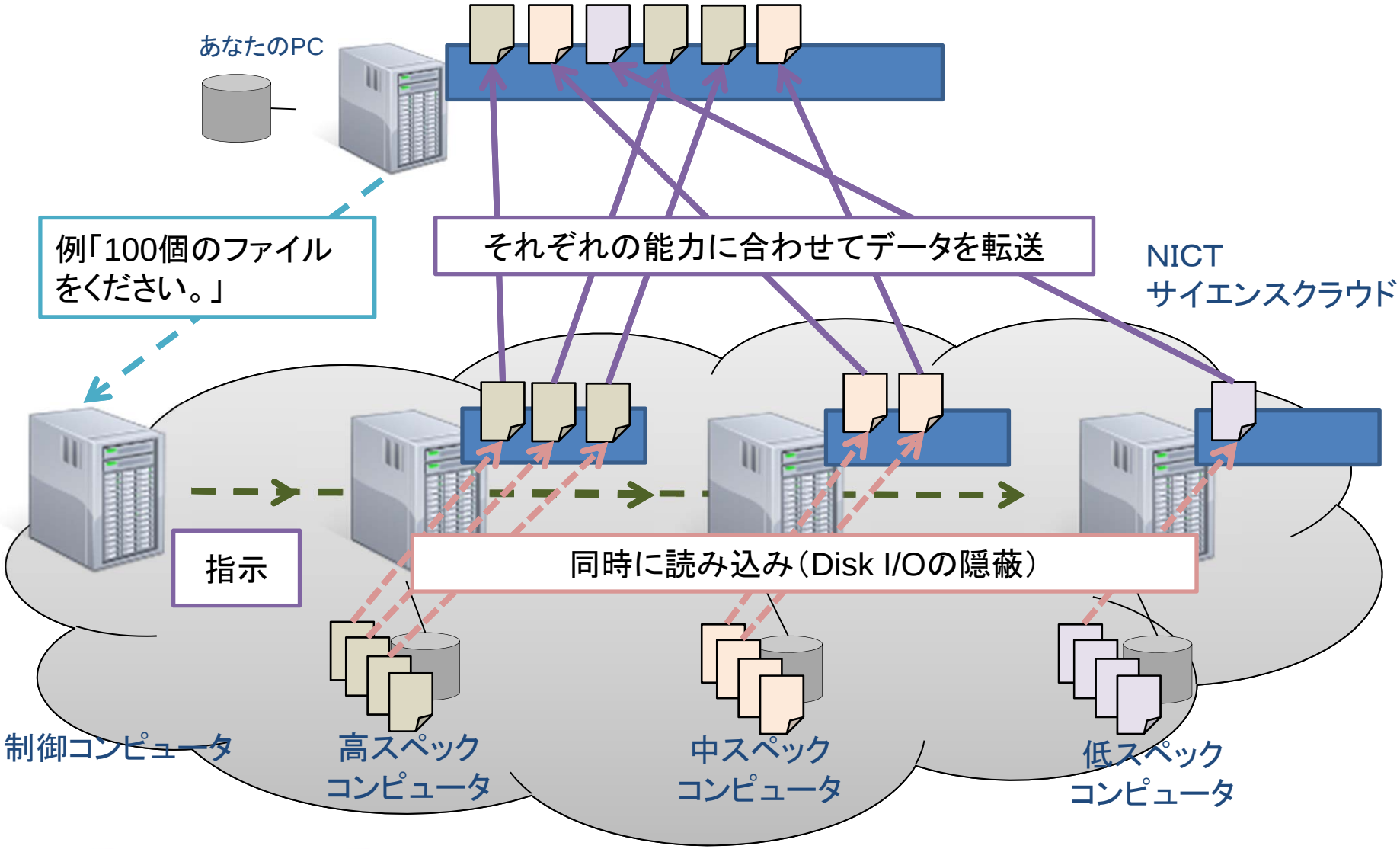


HDD（SATA3,RAID5）：
約3.78Gbps

SSD：
約1.76Gbps

今後の課題：クラウドストレージの高速化

並列データ伝送技術



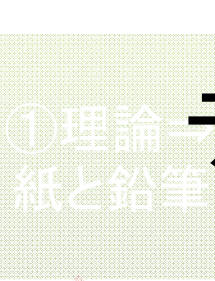
NICT ScienceCloud

科学を行う方法（宇宙科学の場合）

ビッグデータ

データ指向型科学を実現するには…

どうやってデータを集めるか？
どうやってデータを保存するか？
どうやってデータを読み出すか？
どうやってデータを処理（解析）するか？
どうやってデータを公開するか？



①理論⇒紙と鉛筆

④データ指向型研究（インフォマティクス）⇒クラウド

②観測⇒科学衛星・地上観測

③コンピュータシミュレーション⇒スパコン

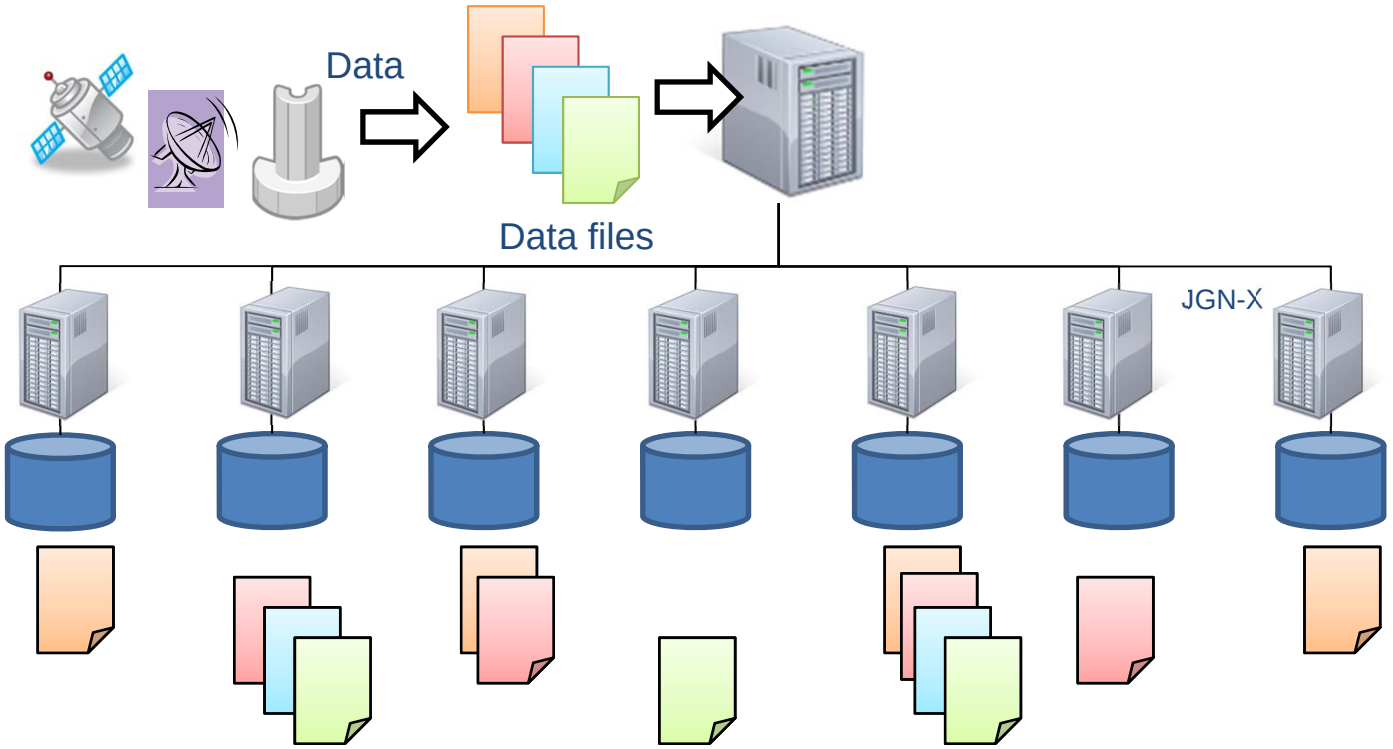
A Virtual Remote Disk for Remote Visualization of Computer Simulations

Ken. T. Murata, K. Yamamoto, H. Watanabe,
T. Kurosawa, E. Kimura, S. Watari, M. Ishii,
KDDI team, JGN-X team, and H. Tatebe

National Institute of Information and Communications Technology
Applied Electromagnetic Research Institute
Space Weather and Environment Informatics Laboratory
4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan
TEL: +81-42-327-7931 FAX: +81-42-327-6978
E-mail: SciCloud-office@ml.nict.go.jp

NICT ScienceCloud

どうやってデータを処理(解析)するか？①

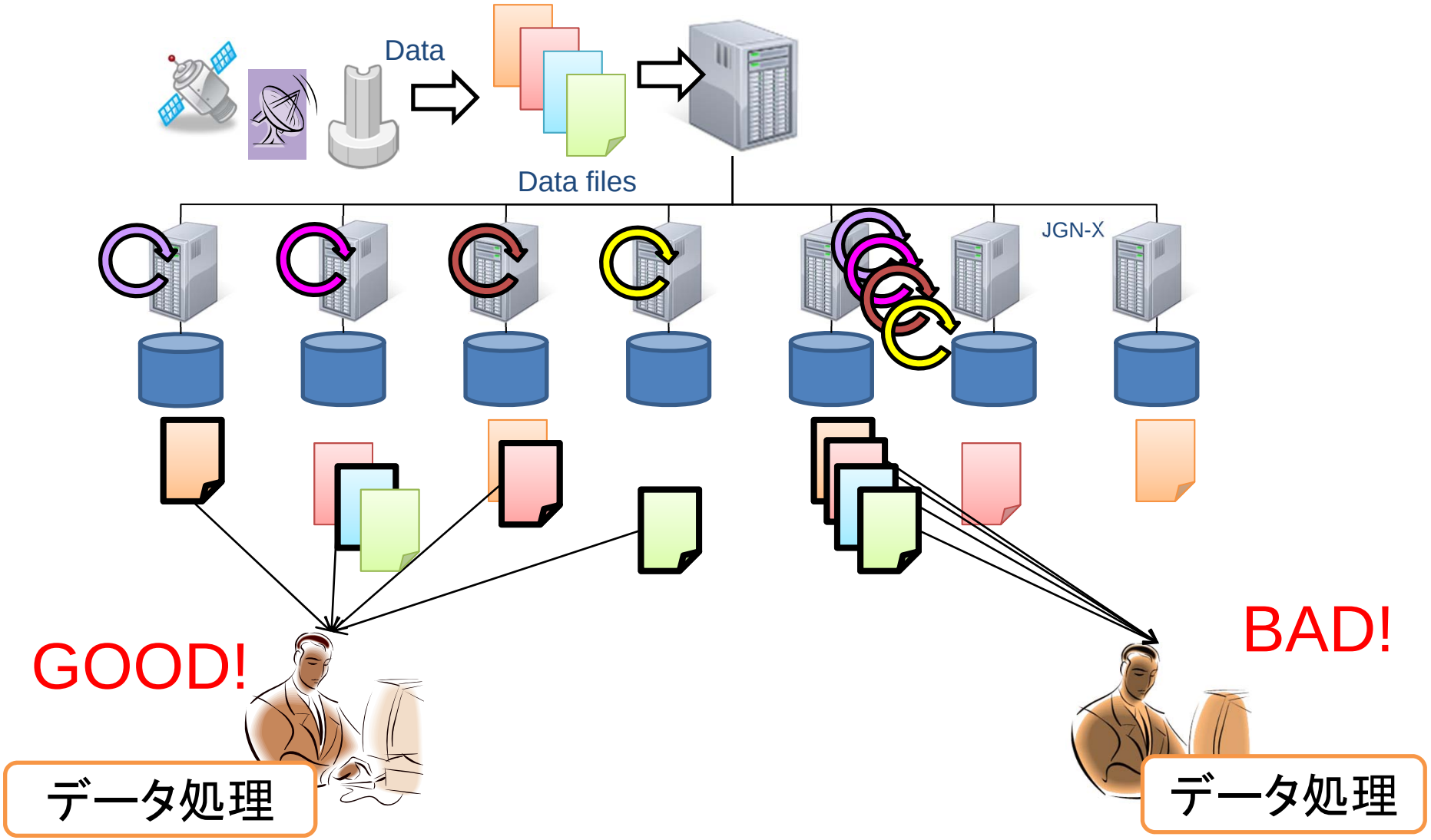


4つのファイルはそれぞれ3か所に保存される。 ← バックアッププレス

サーバは、①ファイル管理サーバ(ファイルシステムノード)機能と②ファイル処理サーバ(クライアントノード)を兼ねる。 ← ディスクI/Oの隠蔽

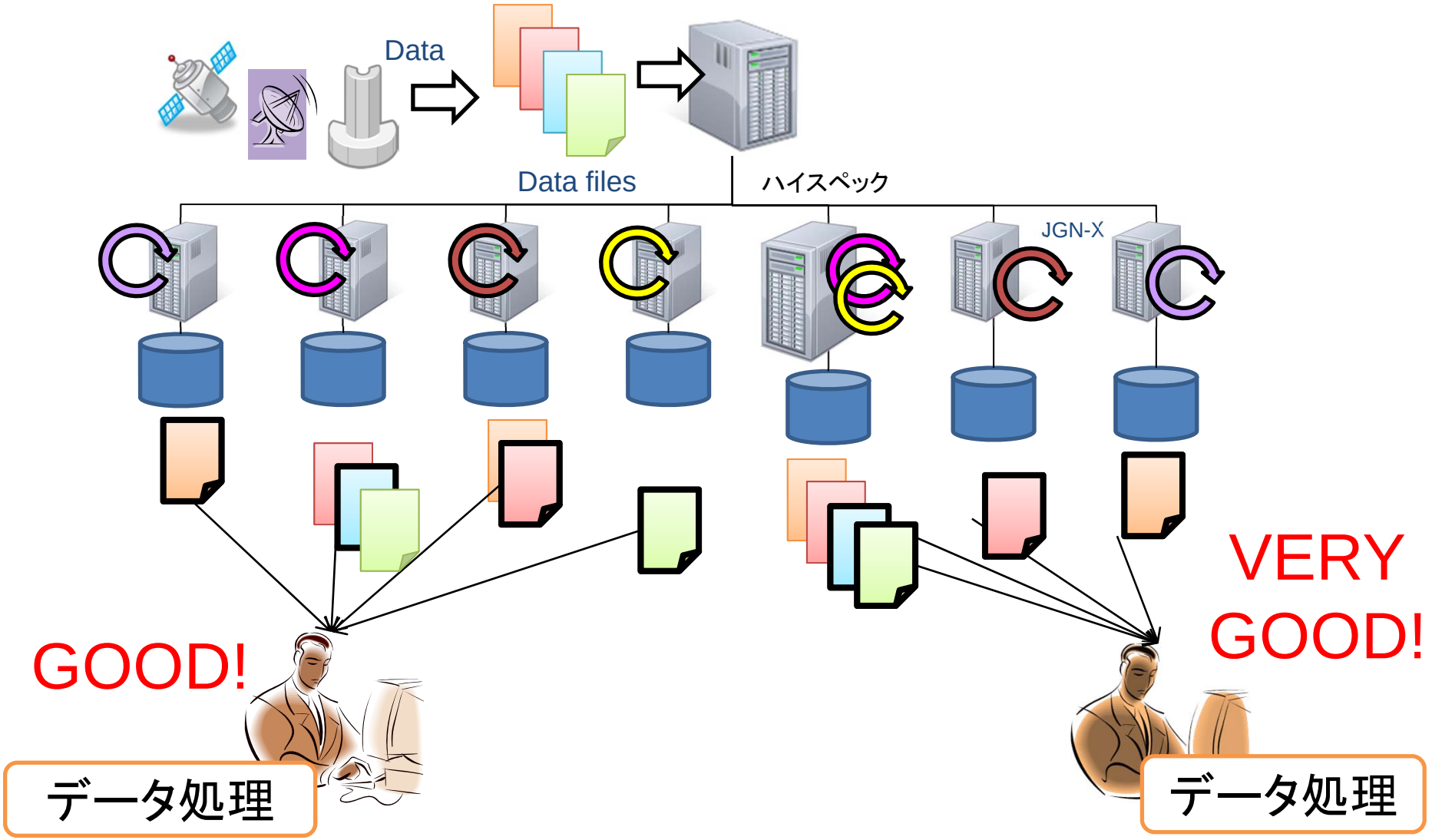
NICT ScienceCloud

どうやってデータを処理(解析)するか？②



NICT ScienceCloud

どうやってデータを処理(解析)するか？③



NICT ScienceCloud

日米間高速データ処理実験①（2012年11月）

PC@アメリカ

目標
10Gpbsでデータを読む！

「2000個（200GB）のファイルをください。」

日米間10Gbpsの専用回線

NICT
サイエンスクラウド

指示

制御コンピュータ

高スペック
コンピュータ（5台）

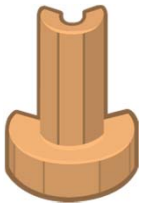
中スペック
コンピュータ（6台）

低スペック
コンピュータ（8台）

NICT ScienceCloud

19台の処理・送信サーバ

Typical Workflow of Visualization (post processing)



Super Computer

Computer Simulation

Numerical Data (n-HDF5)

- Binary
- Unreadable
- Large size

Visualization

Visualized Object Data (o-HDF5)

- Binary
- Unreadable, but Visible
- Small size

Rendering

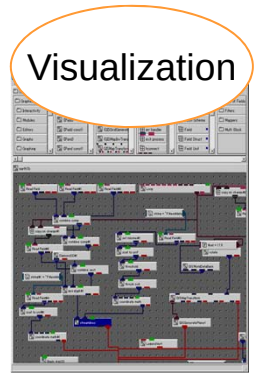
Display (eg., TDW, 3D)

NICT ScienceCloud

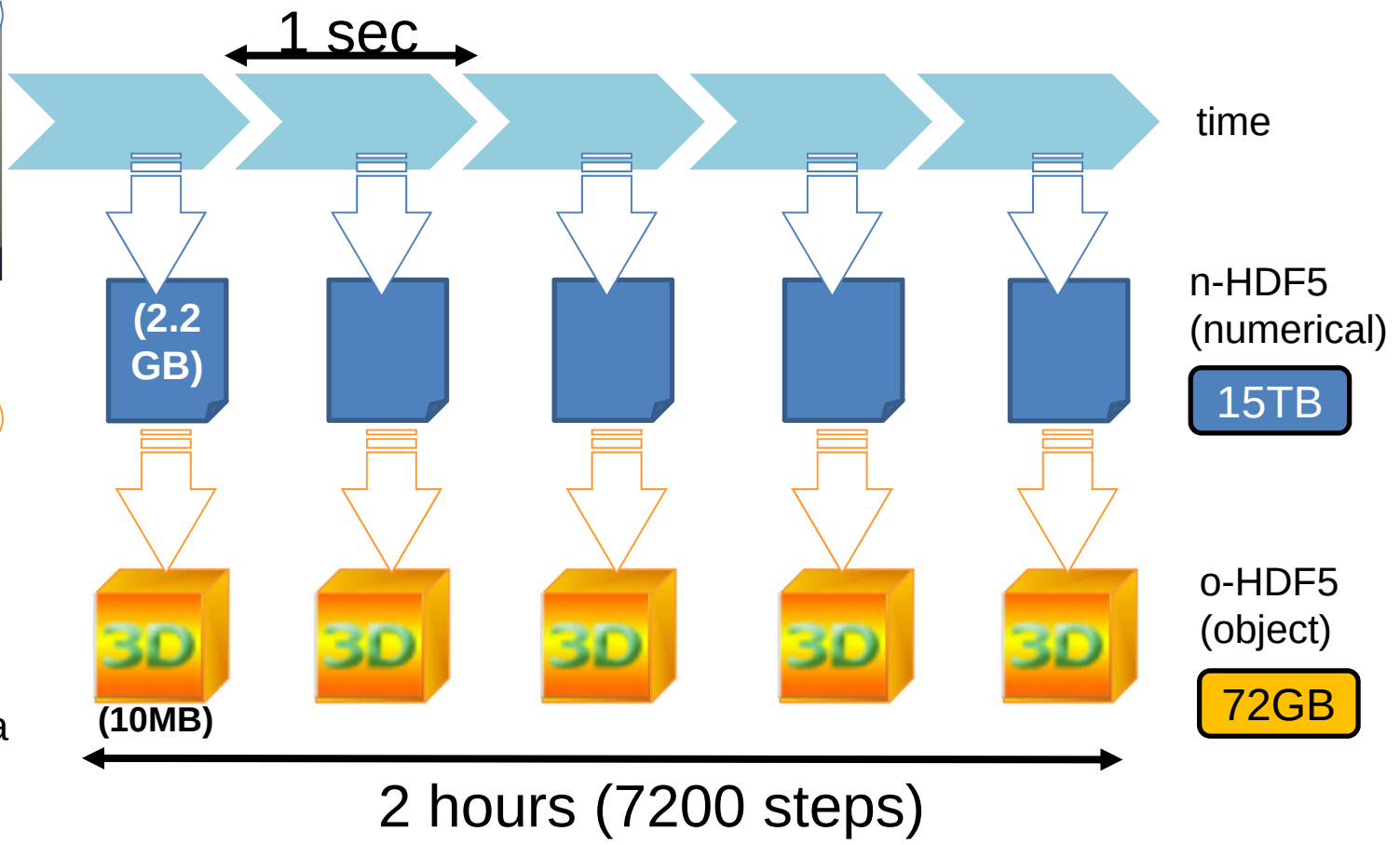
A simulation data size (example)

・Spatial and time resolution
450(x) × 300(y) × 300(z) –uniform grid (dx=0.2Re, dt=0.5sec)

(NICT SX-8)



Virtual Aurora tool



NICT ScienceCloud

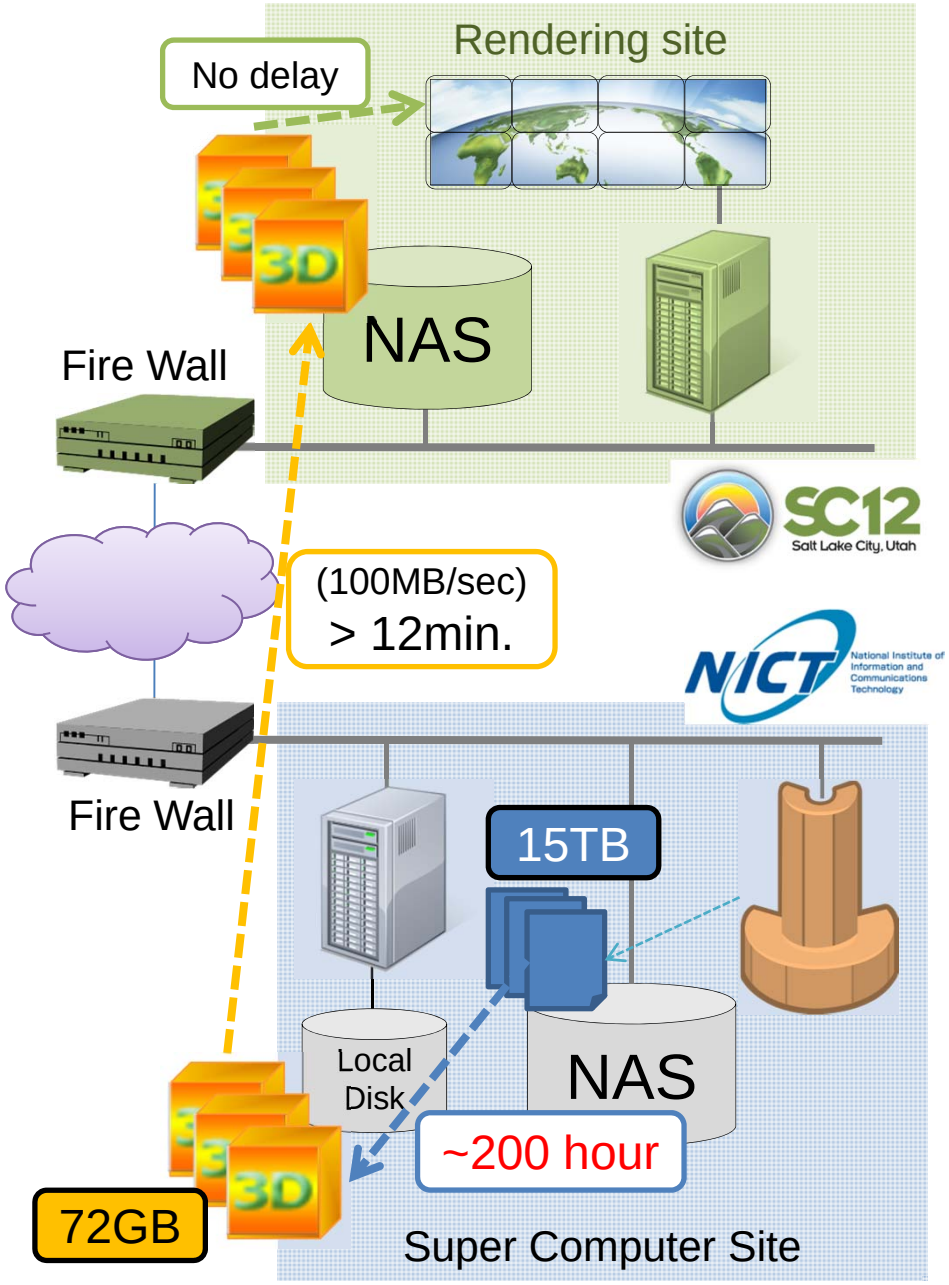
Typical Data Visualization System for Super Computer

- Rendering Site
  - The server has two functions: receiving the transferred data (in object HDF5 format) and rendering on the display.

Numerical data (n-HDF5)
2.2GB/file, file # 7200
(total 15TB)

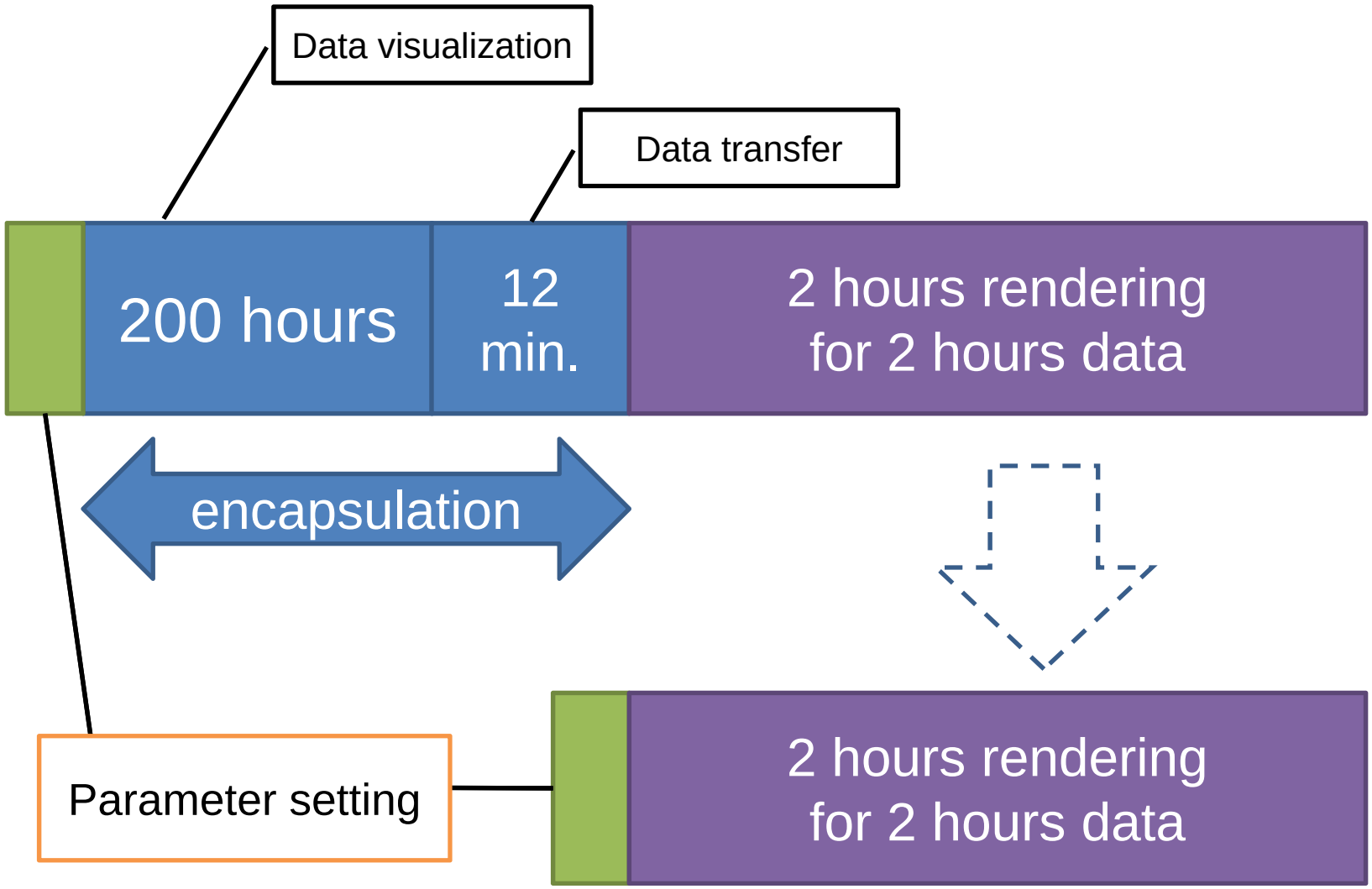
Visualized data (o-HDF5)
10MB/file, file # 7200
(total 72GB)

- Super Computer Site
  - Numerical simulation data files are storage on storage(s) in SC site. The storage is either local disk or NAS.
  - The server has two functions: processing data files on the storage and transfer the data to the rendering site.



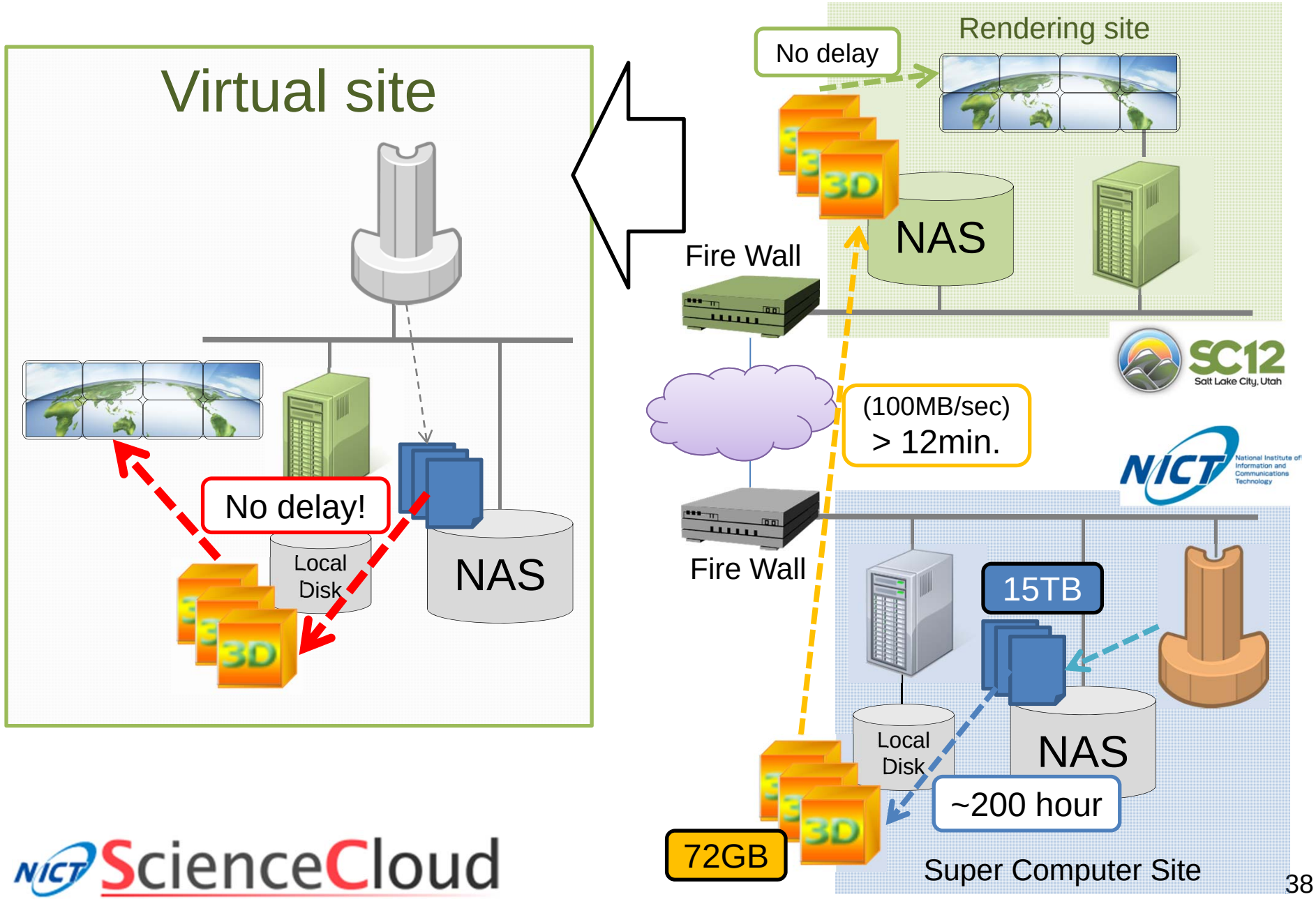
NICT ScienceCloud

Examination: data size and transfer time
In case of a 2 hour global MHD simulation



NICT ScienceCloud

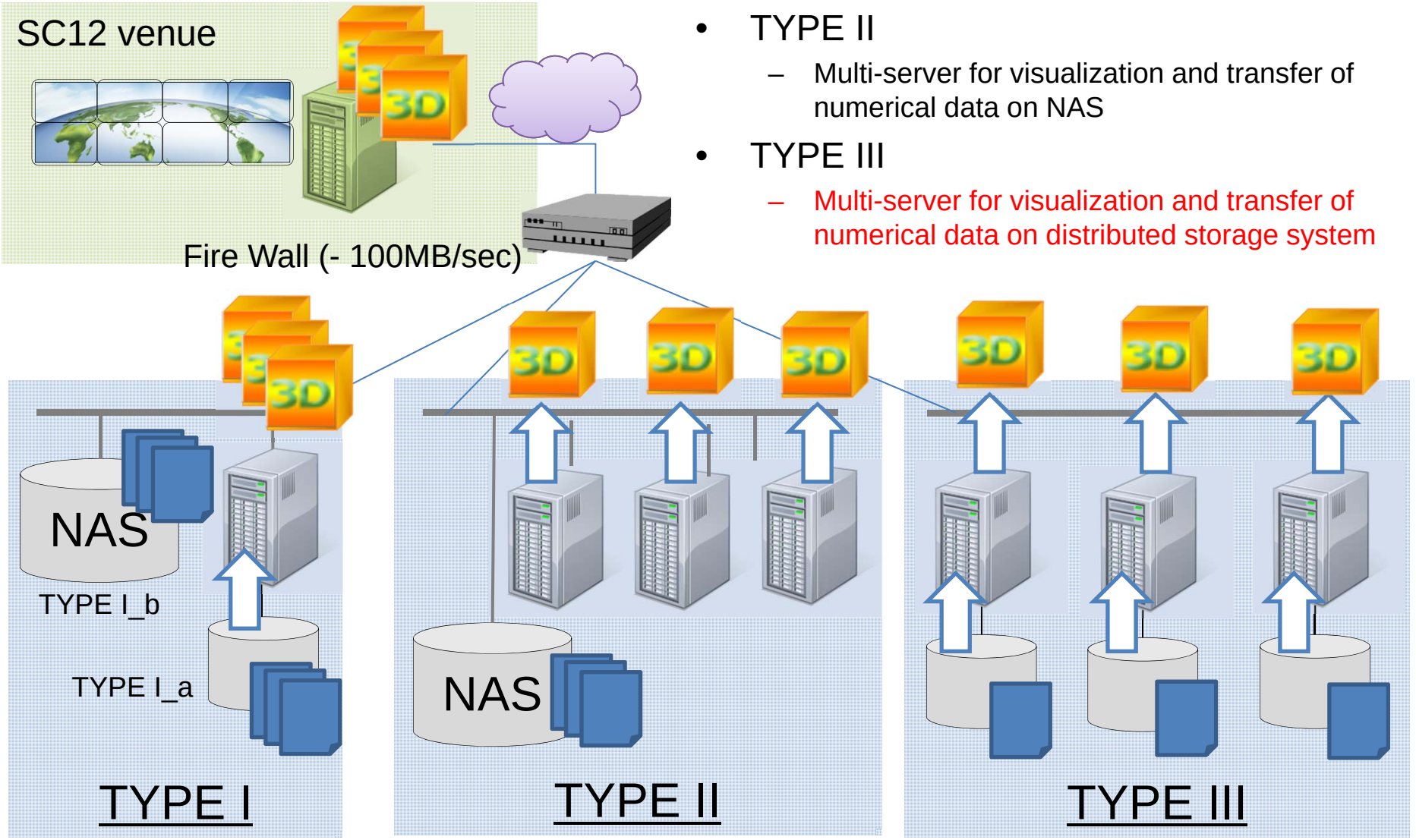
Typical Data Visualization System for Super Computer



NICT ScienceCloud

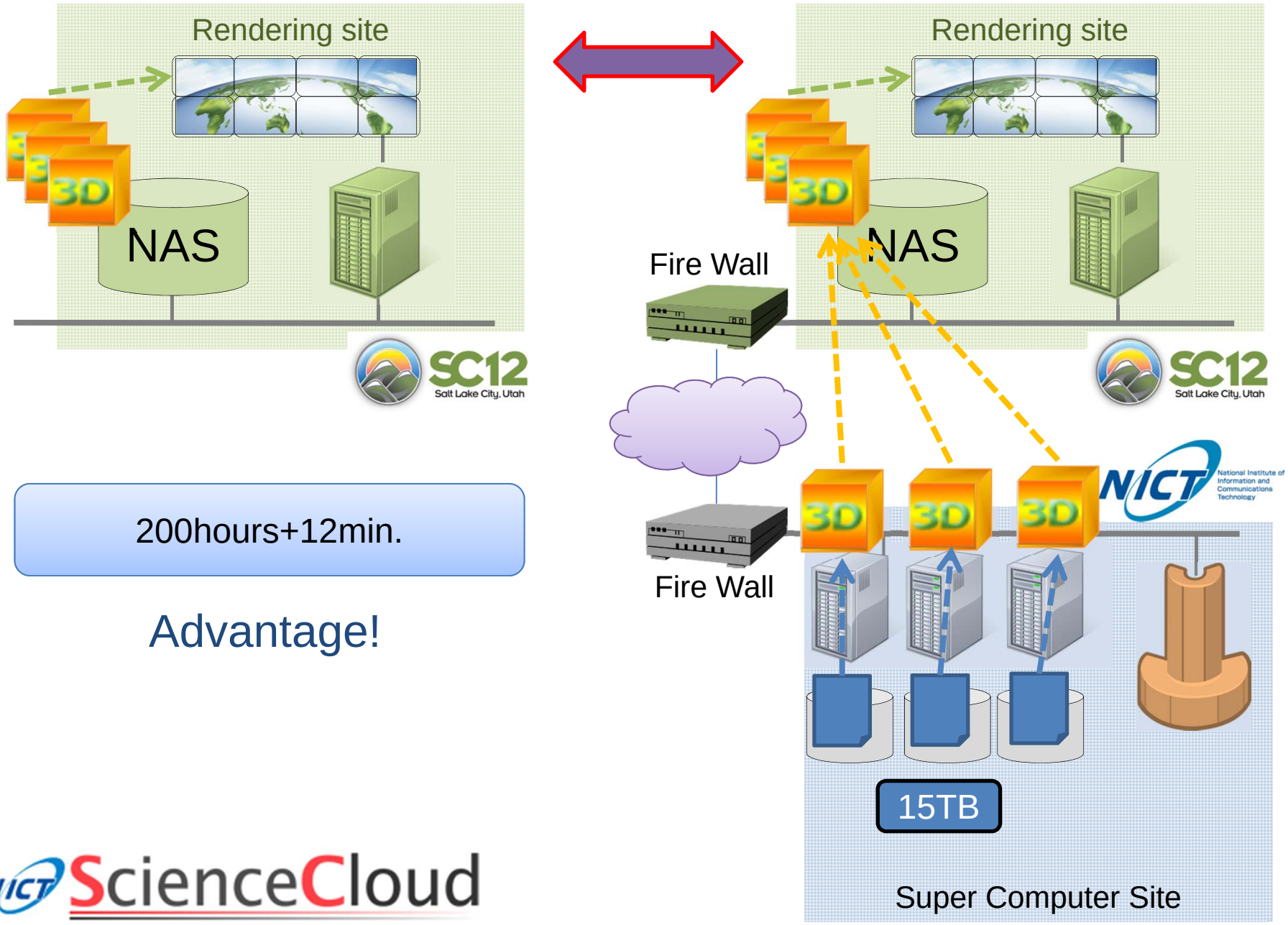
Systems for real-time visualization

- TYPE I
  - Single-server for visualization and transfer of numerical data on NAS/local disk
- TYPE II
  - Multi-server for visualization and transfer of numerical data on NAS
- TYPE III
  - Multi-server for visualization and transfer of numerical data on distributed storage system



NICT ScienceCloud

Competition!

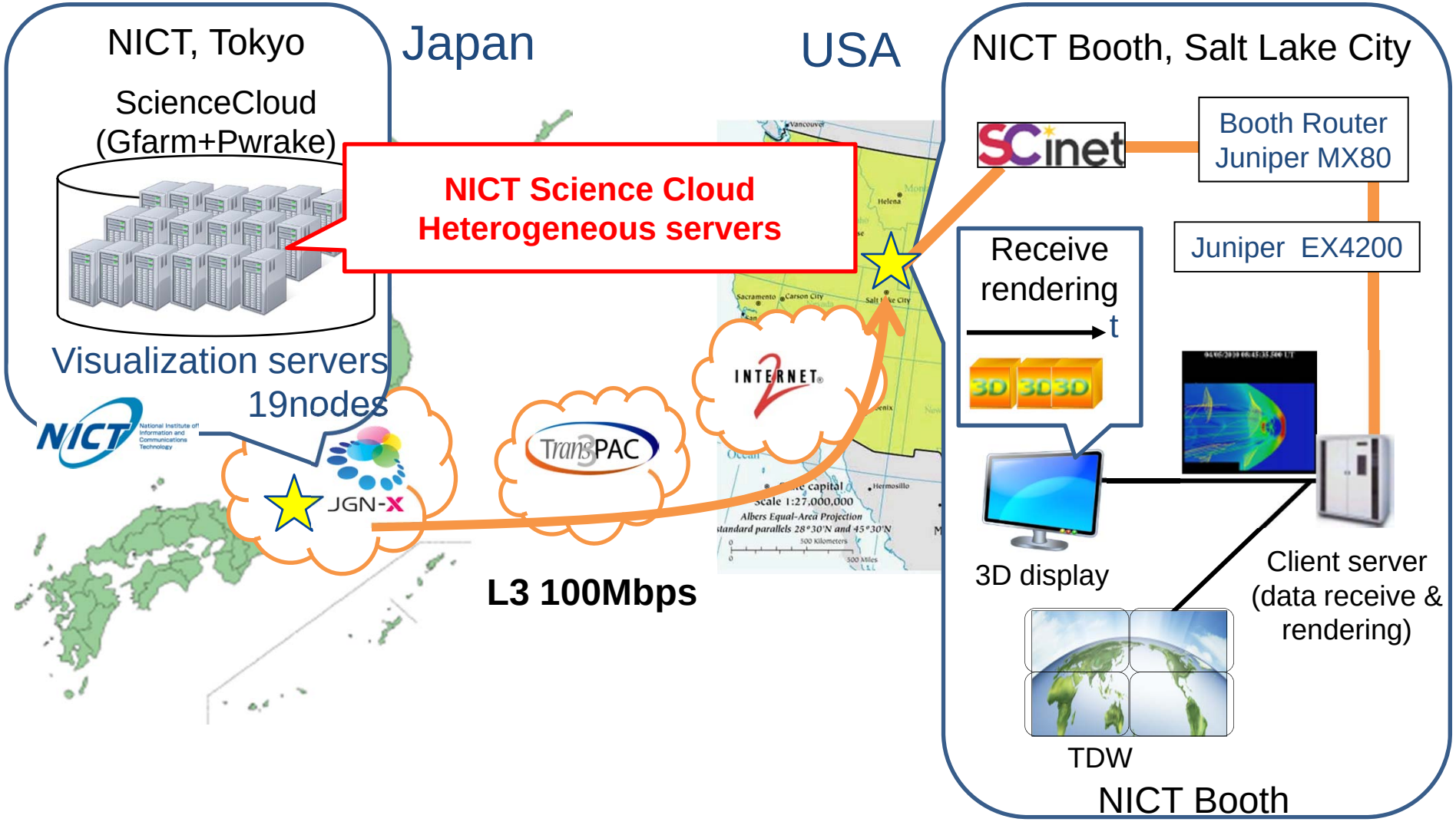


200hours+12min.

Advantage!

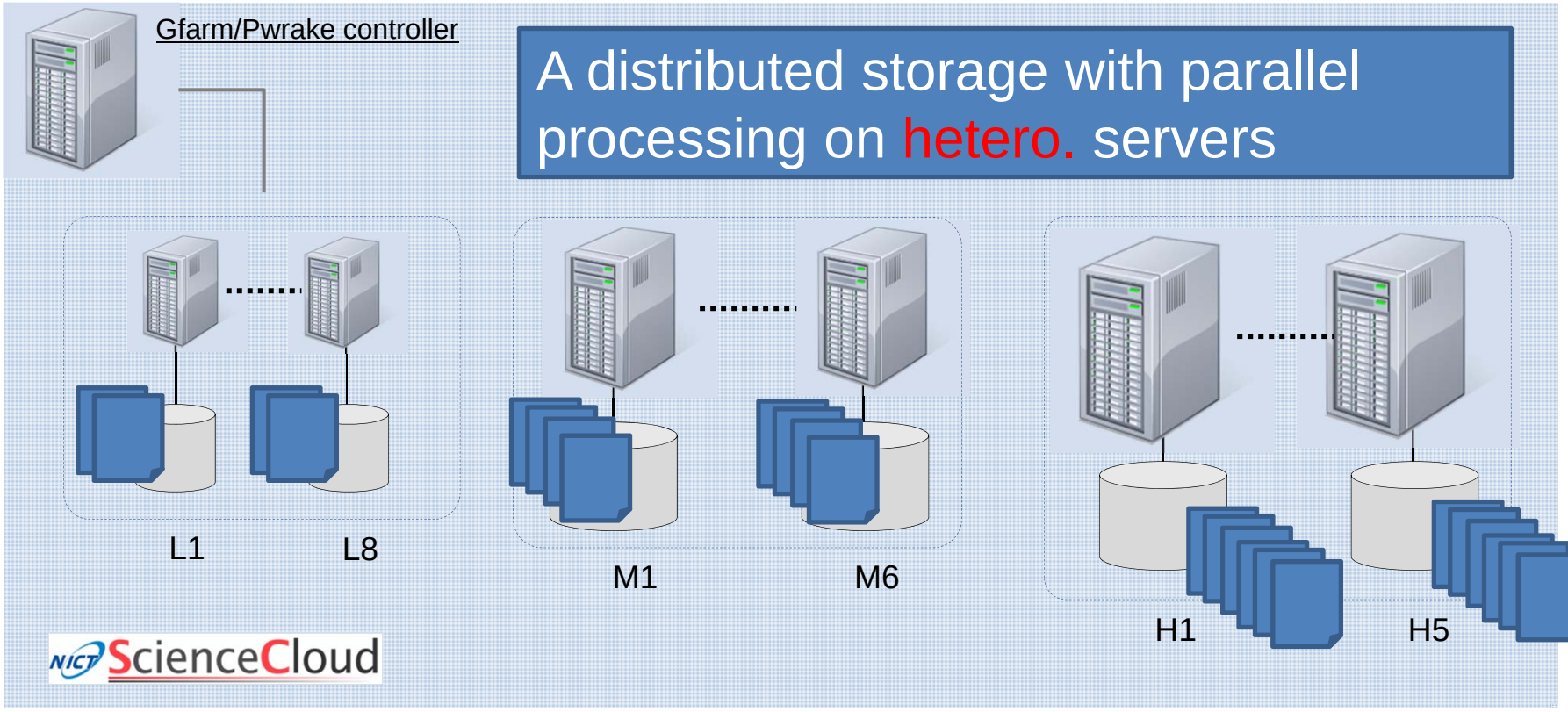
NICT ScienceCloud

Real-time visualization over international network



NICT ScienceCloud

Heterogeneous environment at processing site



Low Spec machine (#8)

- CPU #: 8 cores
- CPU: AMD Opteron Processor 2350
- Main memory: 16 GB
- OS: openSUSE 11.1 (x86_64)
- HDD: SATA 2 x1
  Read: 127 MB/sec Write: 96MB/sec
- NIC: 1GbE

Middle Spec machine(#6)

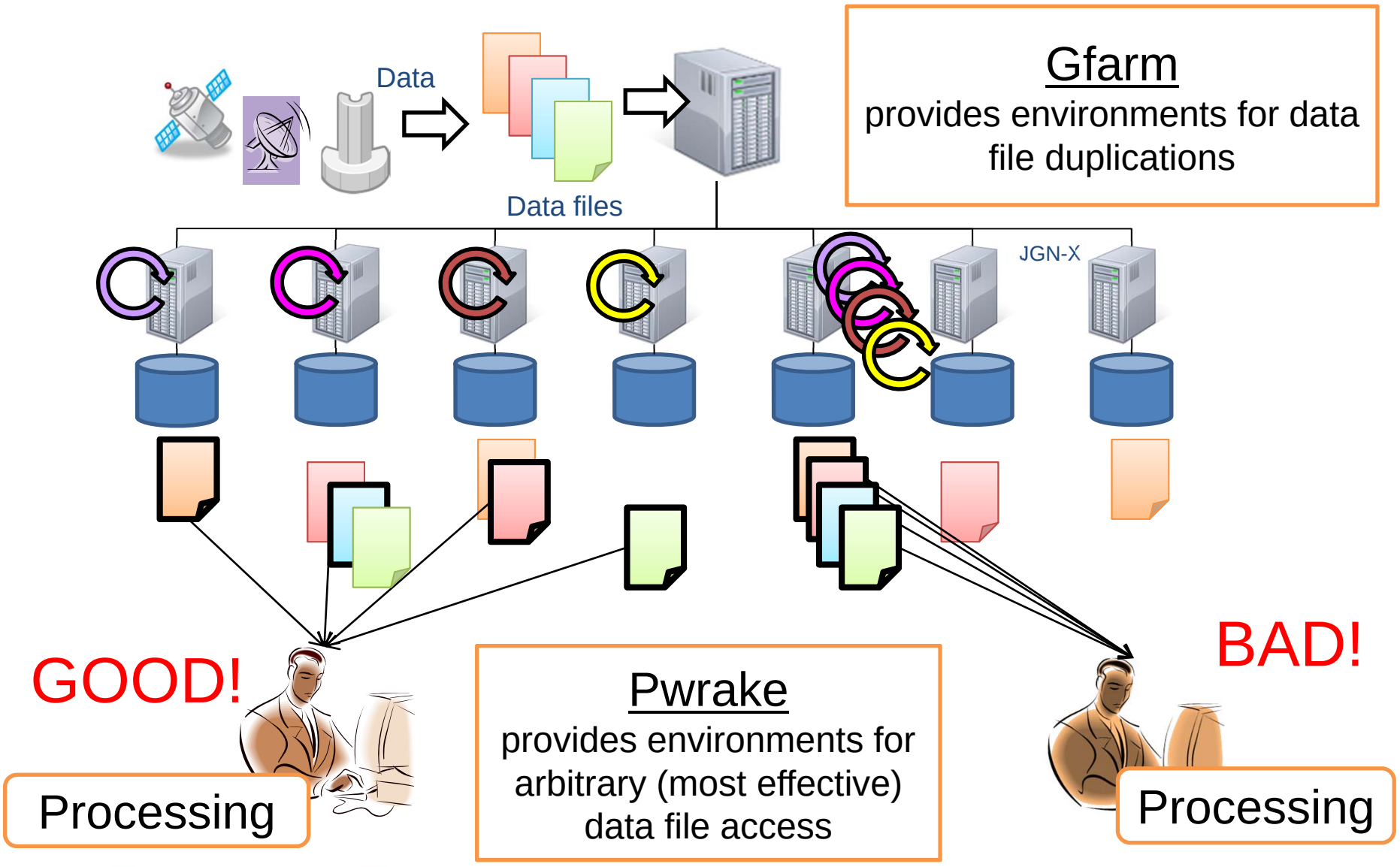
- CPU #: 8 cores
- CPU: Intel Xeon E5507@2.27GHz
- Main memory: 16 GB
- OS: openSUSE 11.4 (x86_64)
- HDD: SATA 3 x4 (RAID5)
  Read: 371 MB/sec Write: 137MB/sec
- NIC: 1GbE

High Spec machine(#5)

- CPU #: 8 cores
- CPU: Intel Xeon X5550@2.67GHz
- Main memory: 144GB
- OS: openSUSE 11.1 (x86_64)
- HDD: SATA 3 x4 (RAID5)
  Read: 371 MB/sec Write: 137MB/sec
- NIC: 10GbE

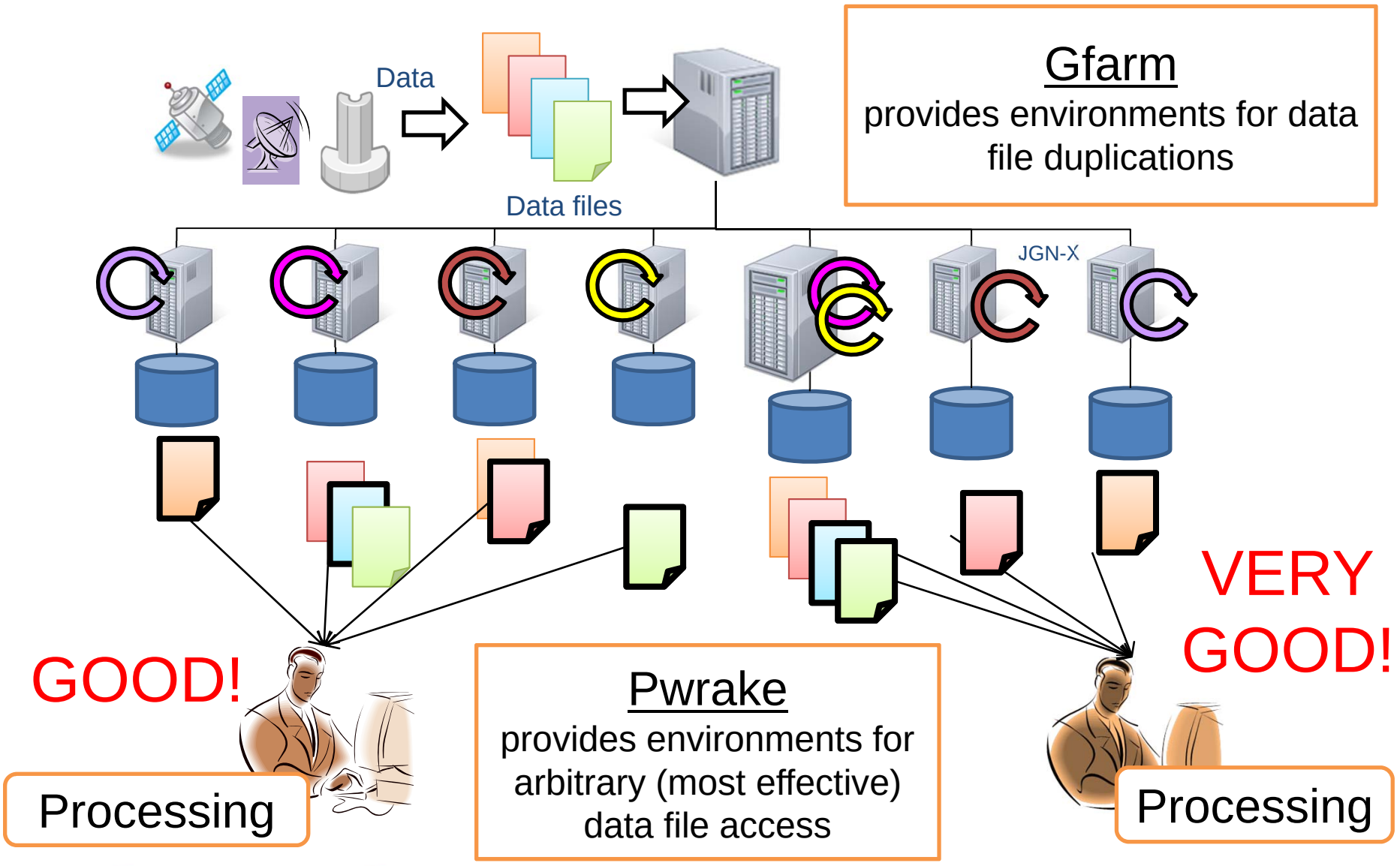
NICT ScienceCloud

Gfarm (Grid Datafarm)/Pwrake (1)



NICT ScienceCloud

Gfarm (Grid Datafarm)/Pwrake (2)



日米間高速データ処理実験②（2012年11月）

PC@アメリカ

目標
10Gpbsでデータを読む！

「2000個（200GB）のファイルをください。」

日米間10Gbpsの専用回線

NICT
サイエンスクラウド

指示

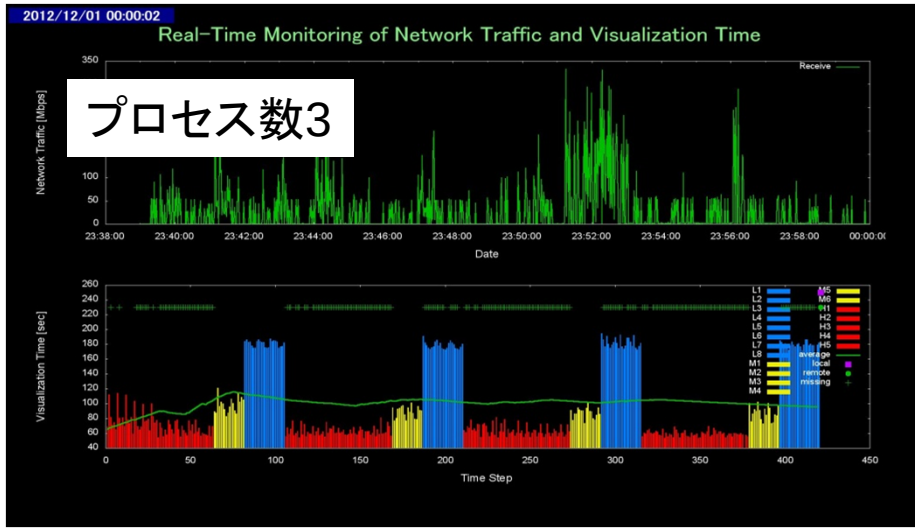
制御コンピュータ

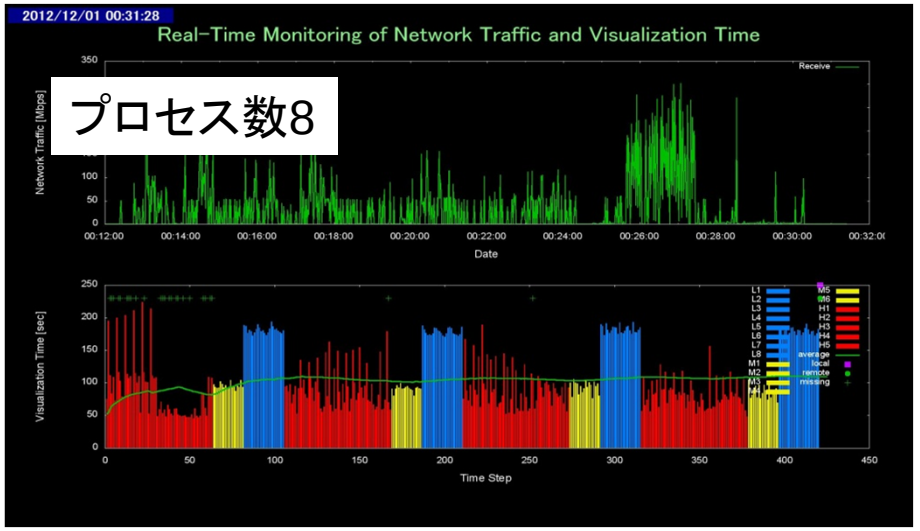
高スペック
コンピュータ（5台）
8コア・16GBメモリ
3プロセス（固定）

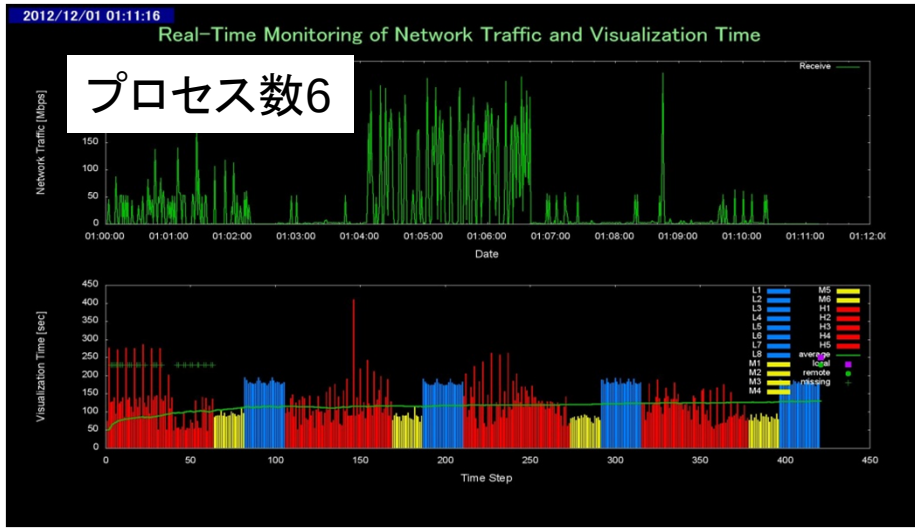
中スペック
コンピュータ（6台）
8コア・16GBメモリ
3プロセス（固定）

低スペック
コンピュータ（8台）
8コア・144GBメモリ
3～8プロセス

プロセス数3/6/8の場合の比較







- 時刻ステップごとのデータ処理時間（縦軸は固定していません）
- 青（L）・黄（M）・赤（H）
- プロセス数が増加⇒ディスクアクセス集中が発生

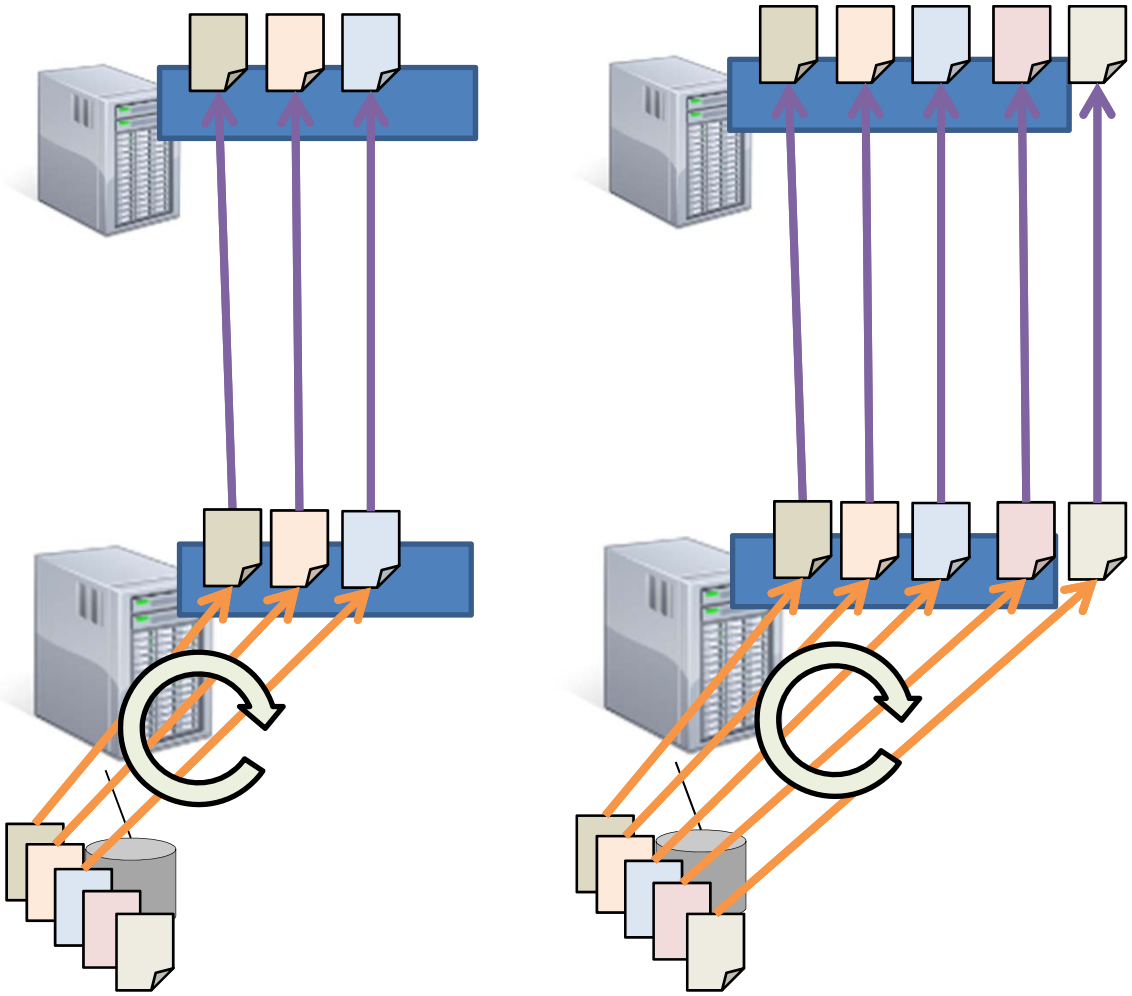
- プロセス数6を超えるとディスクアクセス集中が発生

NICT ScienceCloud

プロセス数の設定

3プロセス並列

5プロセス並列



データ総量の増加
（メモリ容量超過）

データ通信トラフィックの増大

処理の並列化

ディスクアクセスの集中

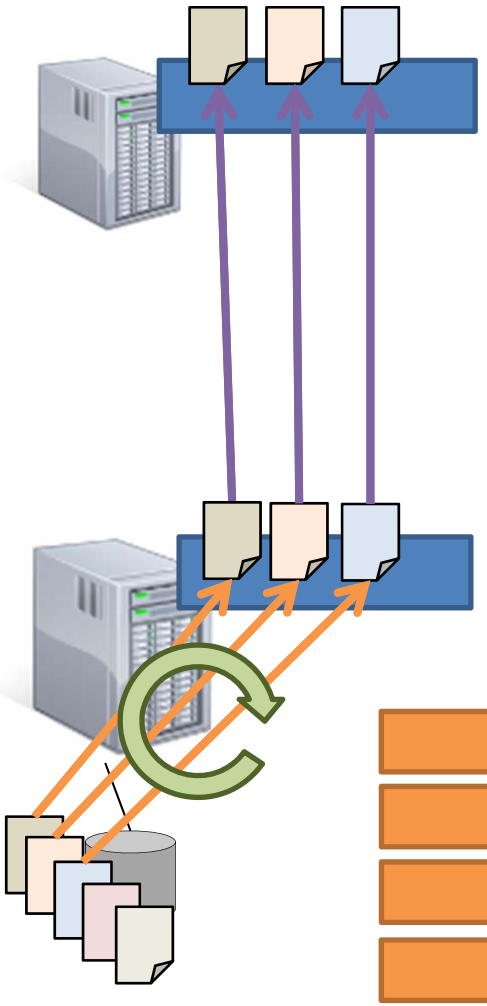
5プロセスが有利

5プロセスが不利

http://www.hitachi.co.jp/products/it/storage-solutions/techsupport/basicknowledge/raid.html

プロセス数の最適化

3プロセス並列の場合



3プロセス並列

I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間

処理速度向上

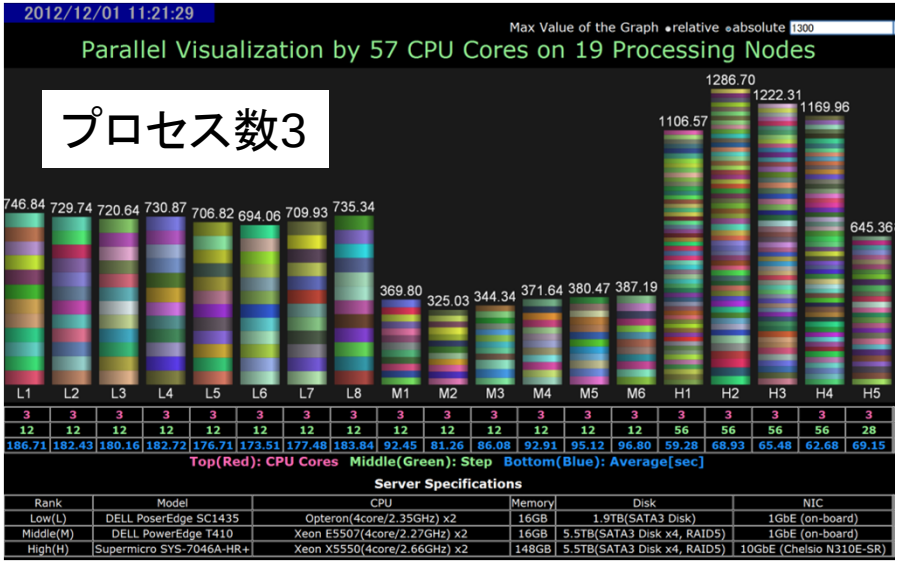
5プロセス並列

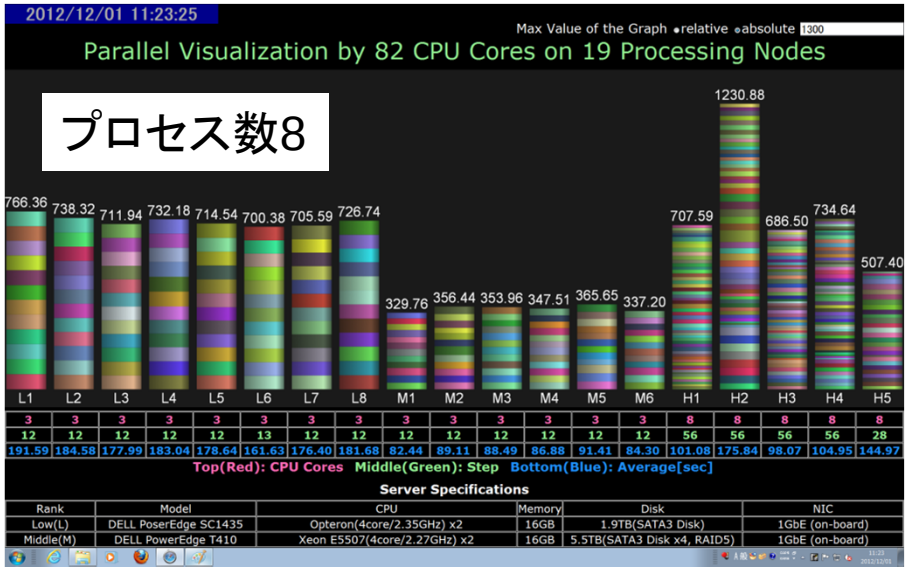
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間

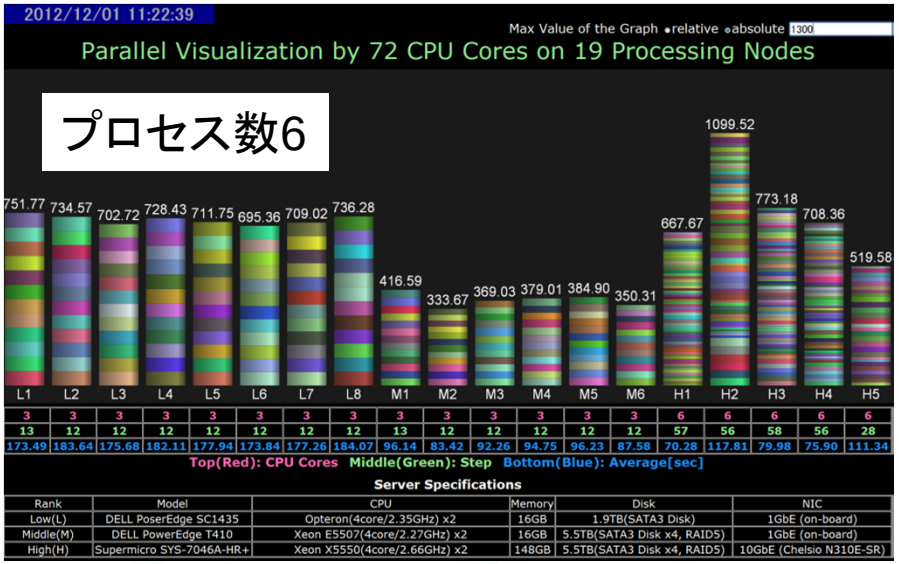
処理速度低下

I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間
I/O時間	処理時間	通信時間

プロセス数3/6/8の場合の比較：負荷バランス



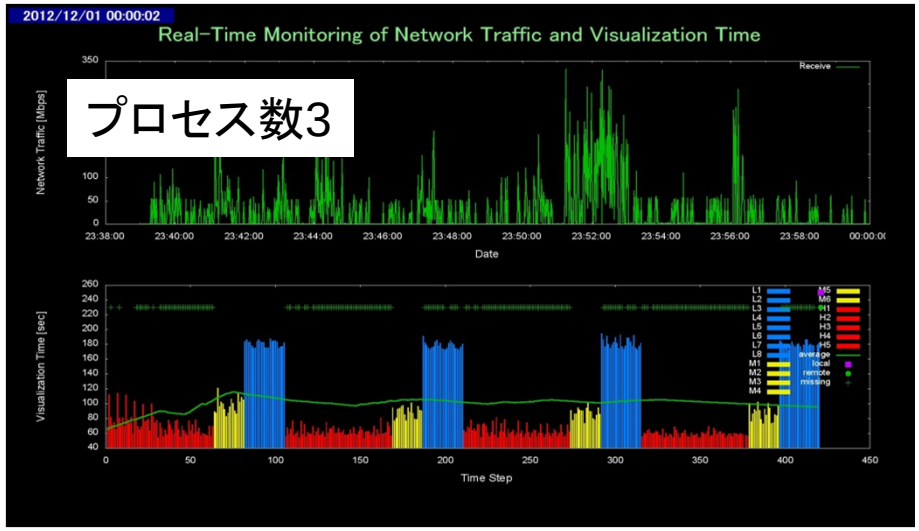


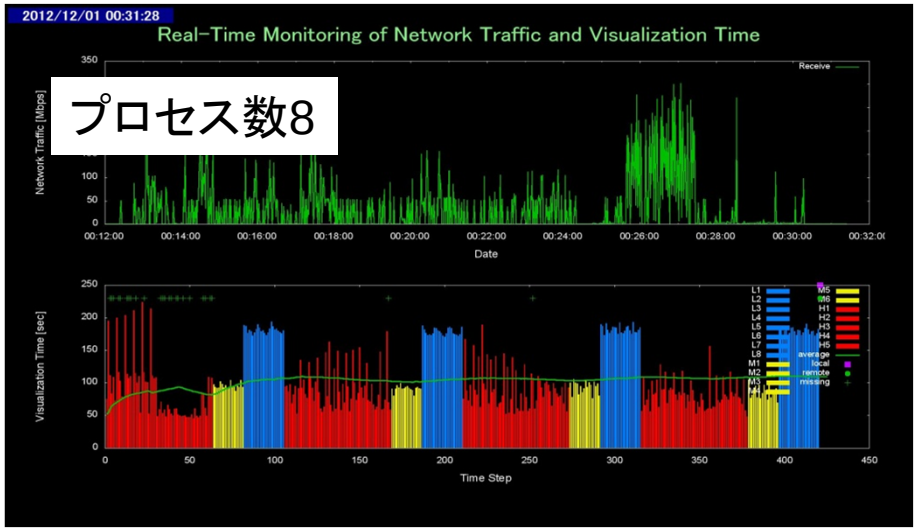


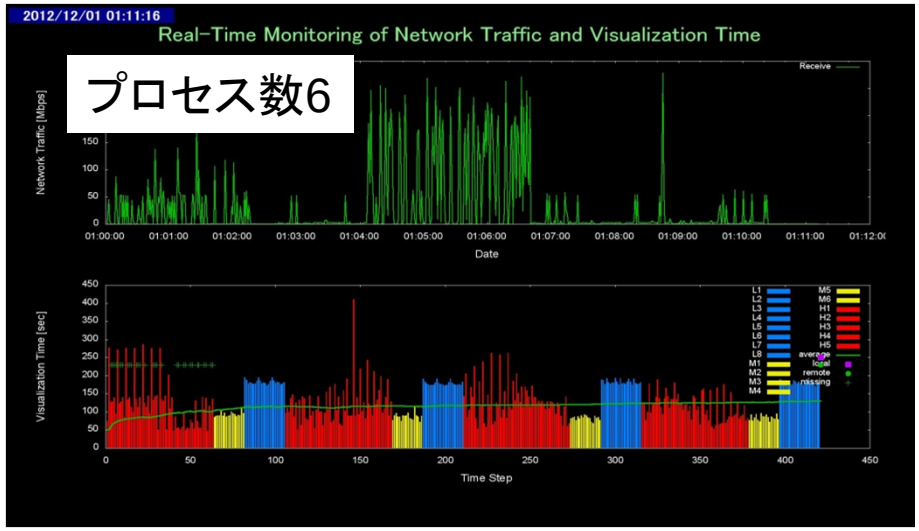
- プロセス数6を超えるとディスクアクセス集中が発生
- スケーラビリティーとディスク集中のバランスがよいのは6プロセス処理の場合

NICT ScienceCloud

プロセス数3/6/8の場合の比較：ステップごとのデータ処理時間





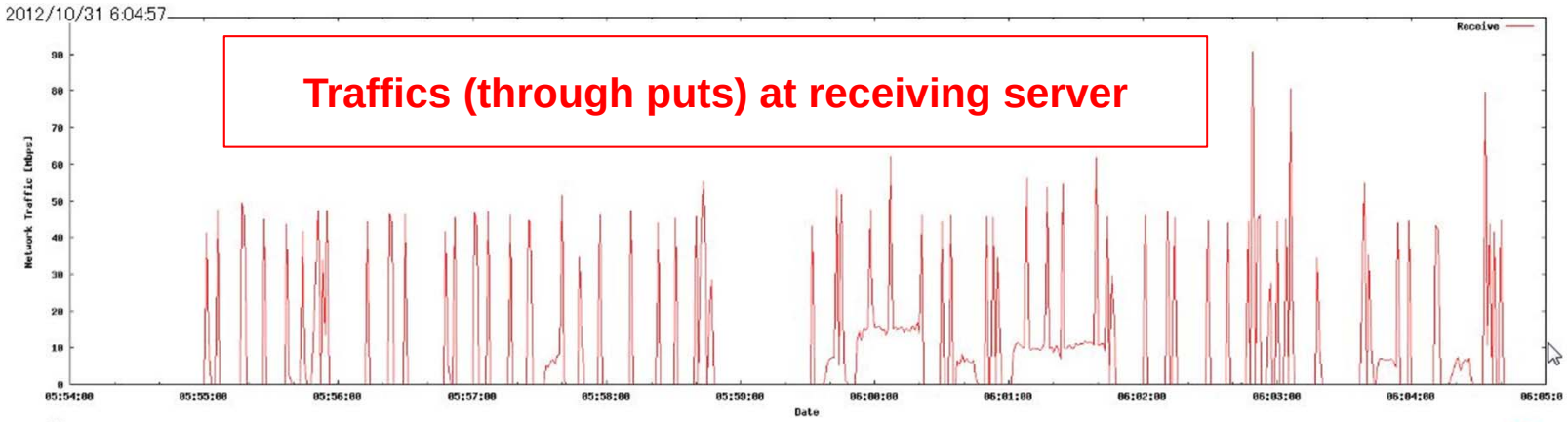


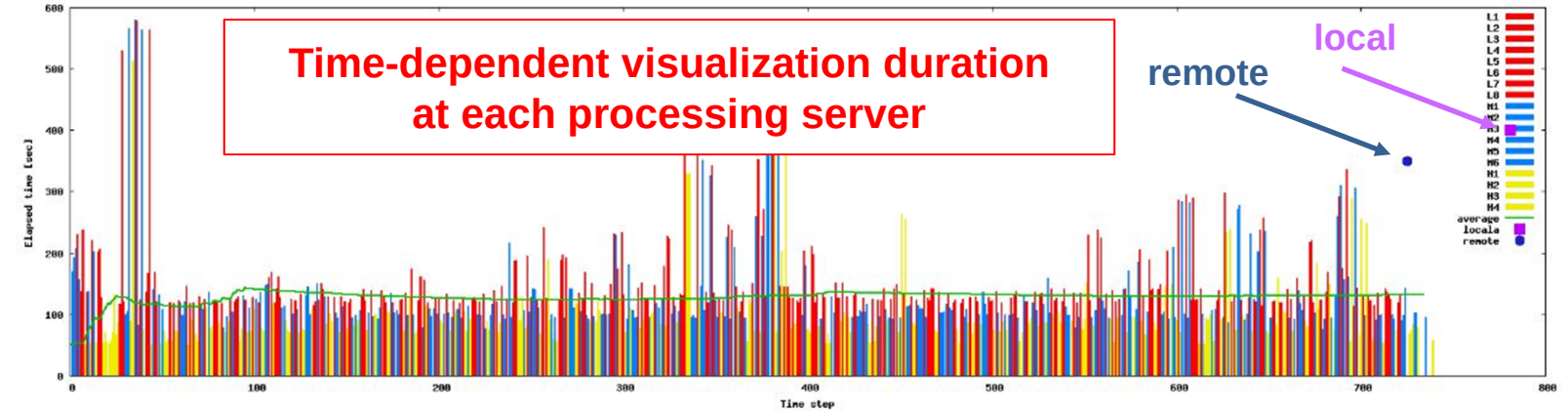
- 時刻ステップごとのデータ処理時間（縦軸は固定していません）
- 青（L）・黄（M）・赤（H）
- プロセス数が増加⇒ディスクアクセス集中が発生

- プロセス数6を超えるとディスクアクセス集中が発生

NICT ScienceCloud

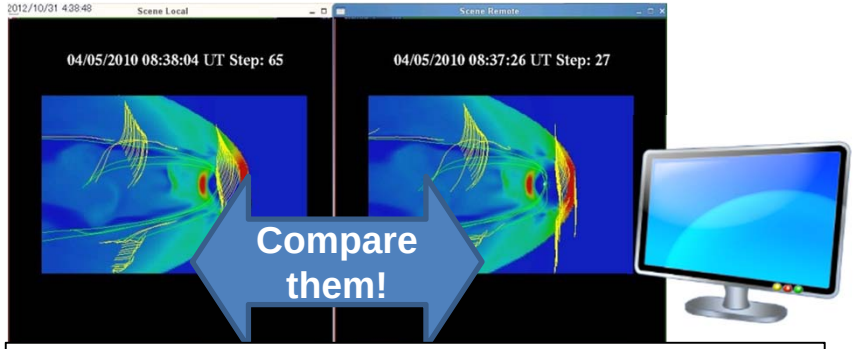
（参考）FIFO型の場合





Visualization time

Time step

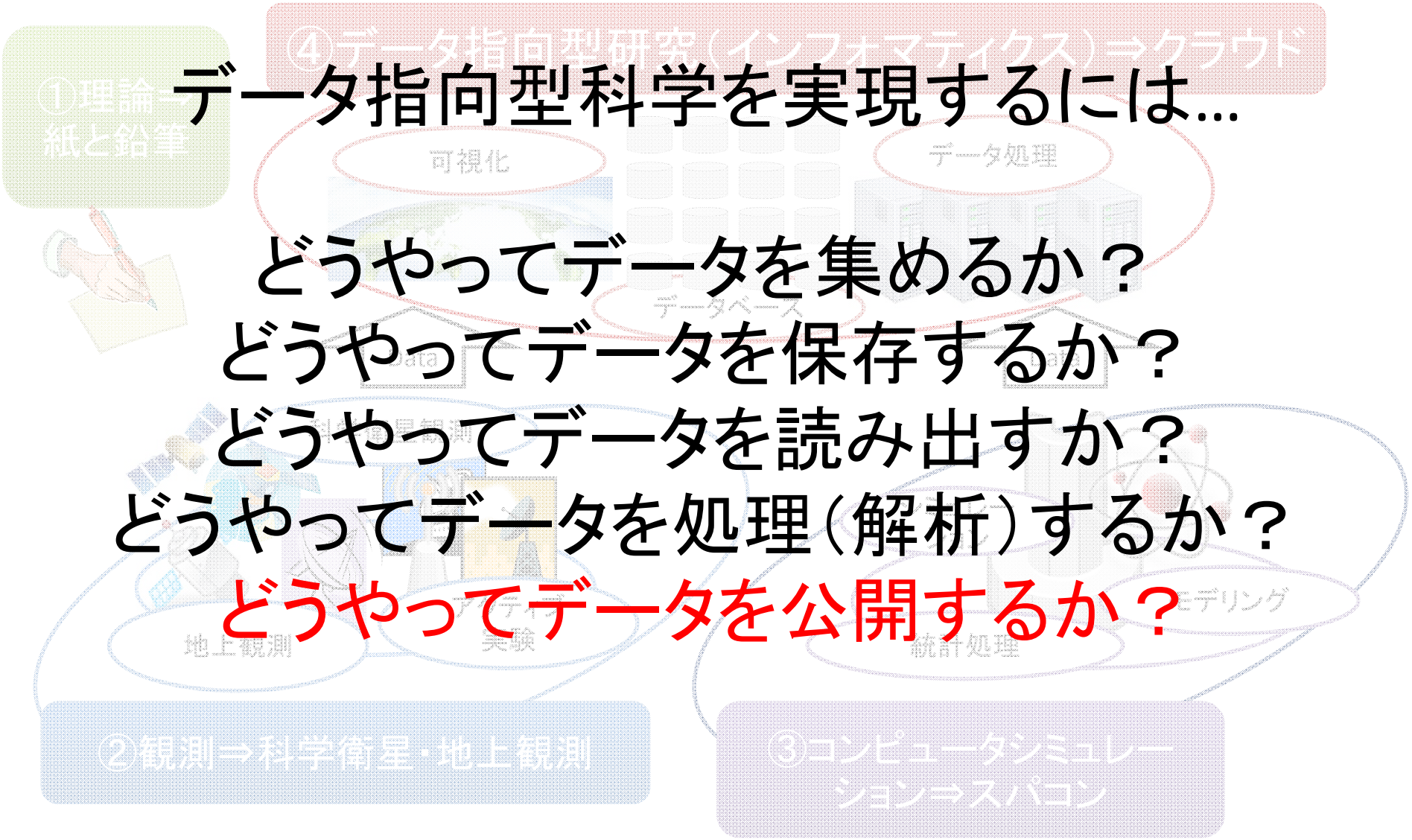


3D rendering at the client server

NICT ScienceCloud

科学を行う方法（宇宙科学の場合）

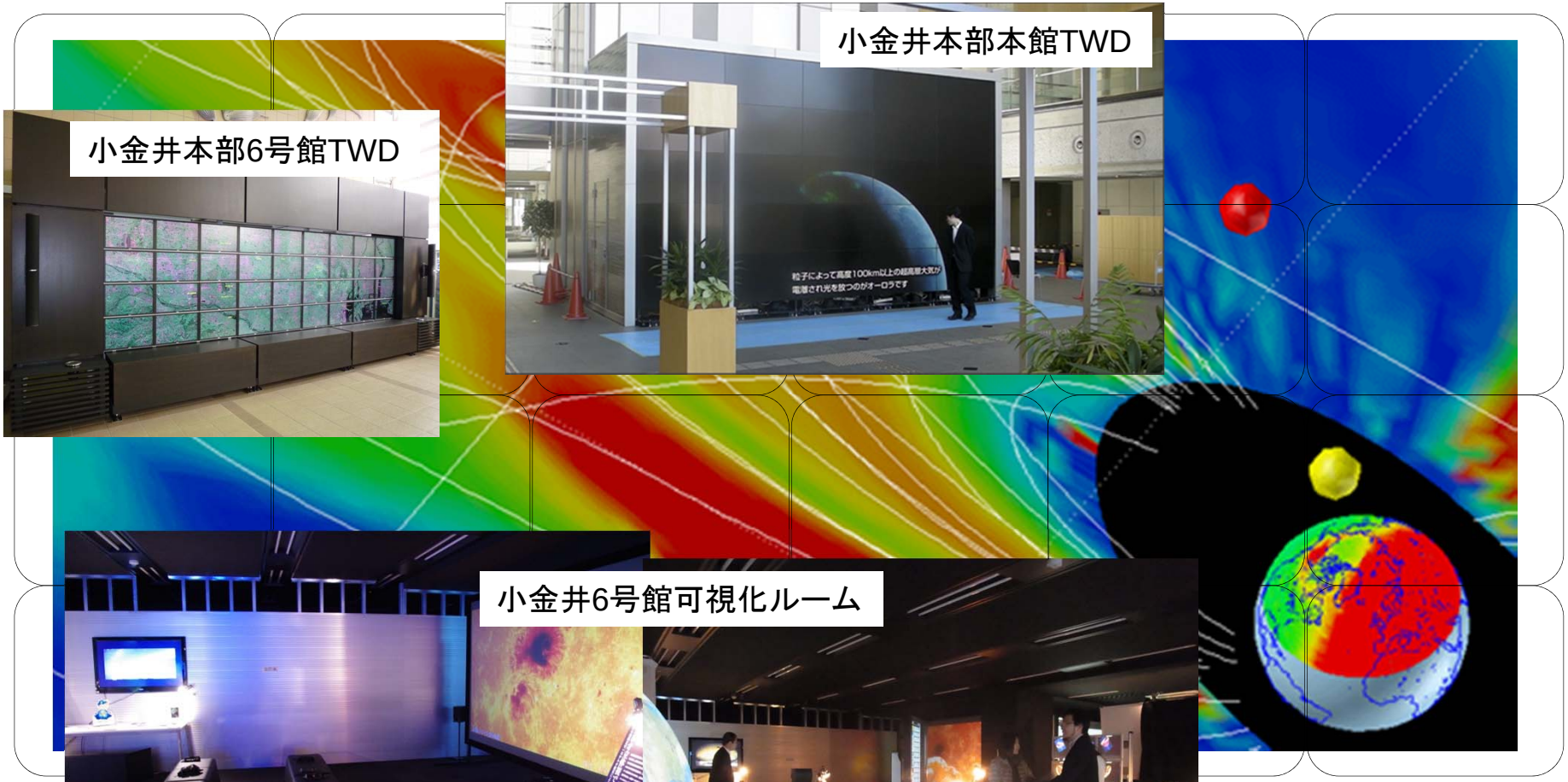
ビッグデータ



データ指向型科学を実現するには…

どうやってデータを集めるか？
どうやってデータを保存するか？
どうやってデータを読み出すか？
どうやってデータを処理（解析）するか？
どうやってデータを公開するか？

どうやってデータを公開するか？



NICT ScienceCloud

Visualization of global MHD simulation with extremely high time resolution

Y. Kubota and space weather team

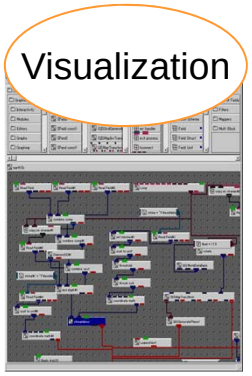
National Institute of Information and Communications Technology
Applied Electromagnetic Research Institute
Space Weather and Environment Informatics Laboratory
4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan
TEL: +81-42-327-7931 FAX: +81-42-327-6978
E-mail: SciCloud-office@ml.nict.go.jp

NICT ScienceCloud

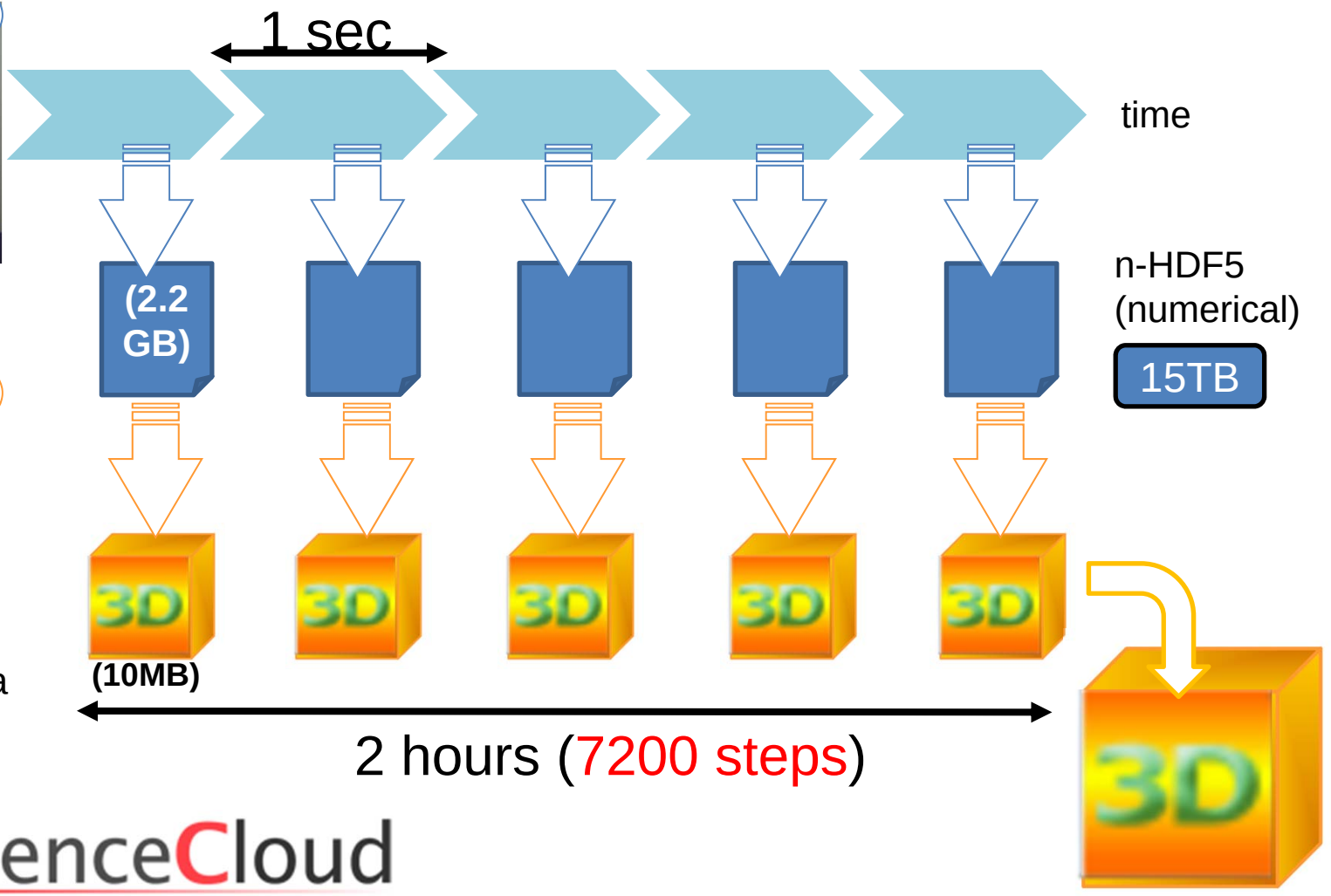
Simulation data size (typical)

・Spatial and time resolution
450(x) × 300(y) × 300(z) –uniform grid (dx=0.2Re, dt=0.5sec)

(NICT SX-8)

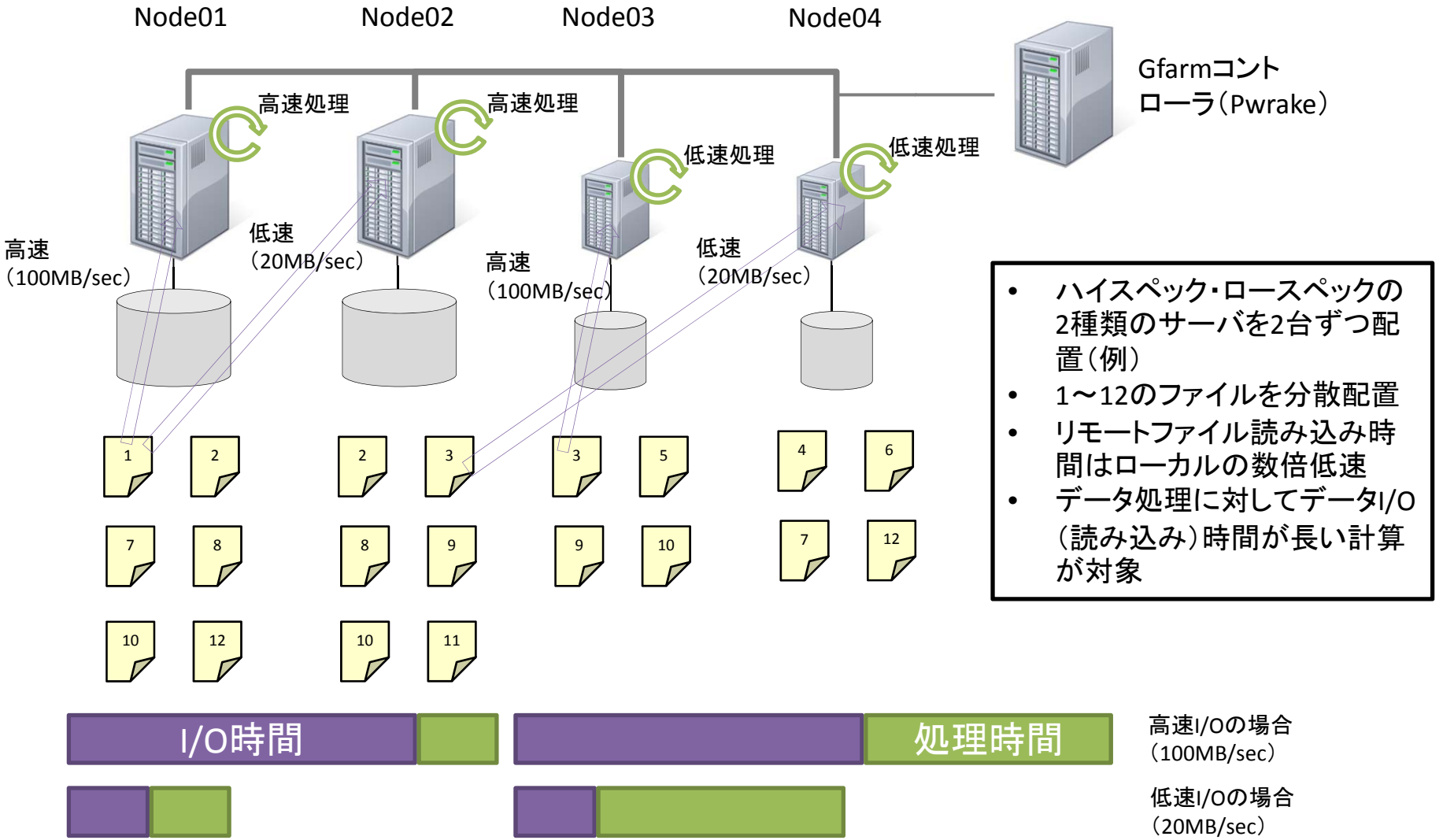


Virtual Aurora tool



NICT ScienceCloud

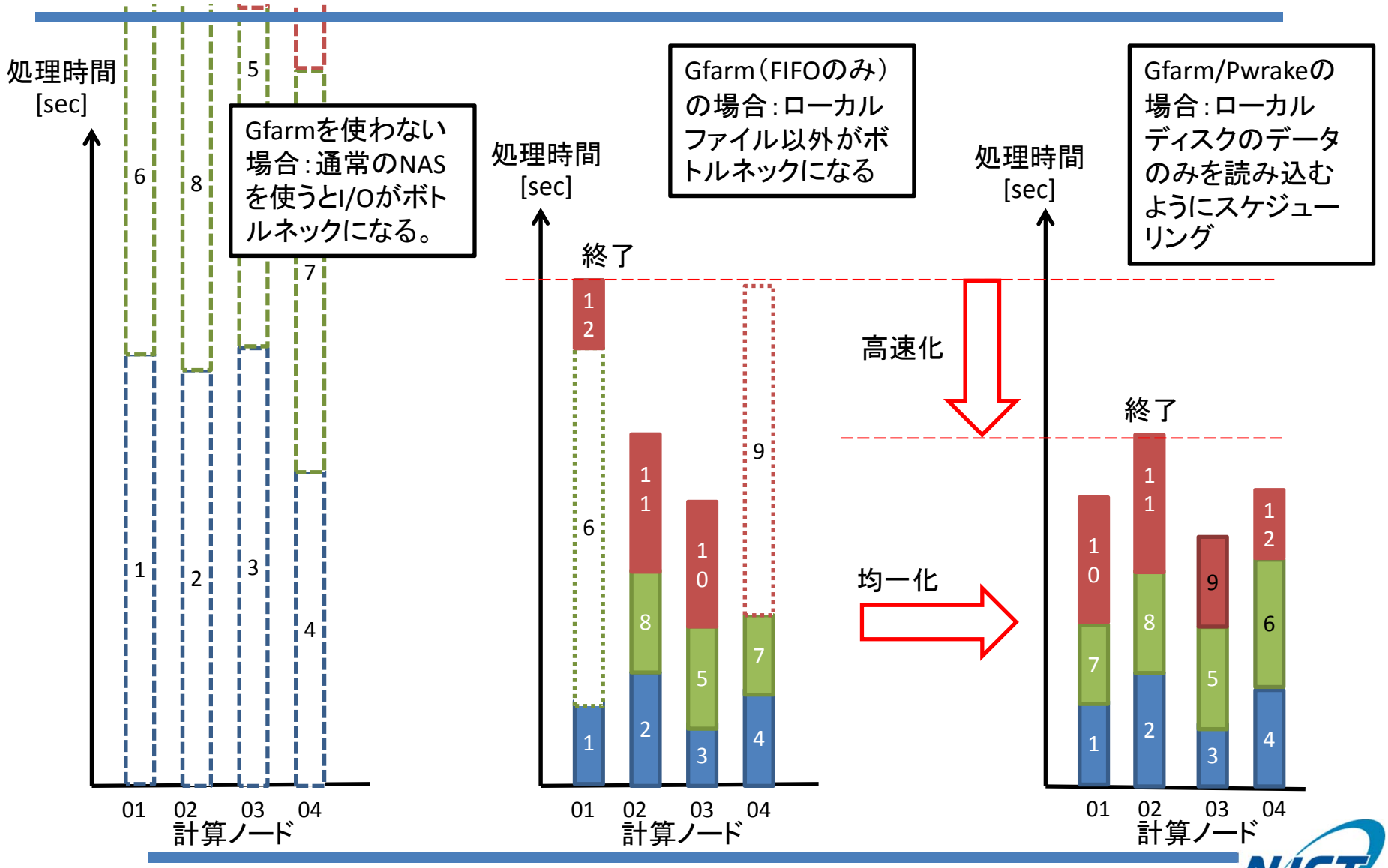
Gfarm/Pwrakeによる並列処理（モデル例）



- ハイスペック・ロースペックの2種類のサーバを2台ずつ配置（例）
- 1～12のファイルを分散配置
- リモートファイル読み込み時間はローカルの数倍低速
- データ処理に対してデータI/O（読み込み）時間が長い計算が対象

Gfarm/Pwrakeによるノード処理最適化

FIFO型と非均一分散型の組み合わせ例



Distributed Storage/ Parallel Visualization of 7200 time step data（data size: 15TB）

Traditional Method (36 days)

Data Read Time (I/O Time)
20MB/sec -> 1.5 million sec. -> 18 days

Data Processing Time
14400 steps x Tracing time (10 sec.) -> 14 thousands sec. -> 2 days

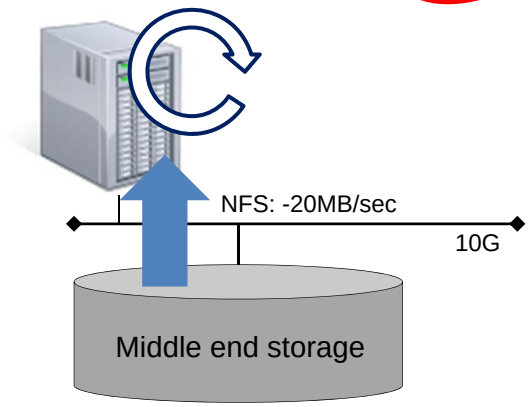
Visualization Time
14400 steps x120sec -> 16 days

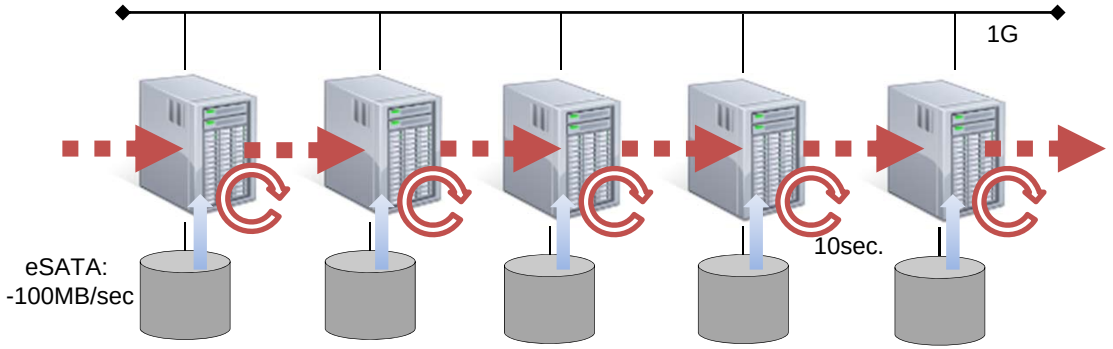
Parallel Visualization (2 days)

Data Read Time (I/O Time)
100MB/sec x 200 cores -> 1500 sec.
-> 30 min.

Data Processing Time
14400 steps x Chasing time (10 sec.) -> 14 thousands sec. -> 2 days

Visualization Time
14400 steps x 120sec./200 cores -> 4 hours





Disk: Low end storage ($160/TB)

CPU: 200 cores

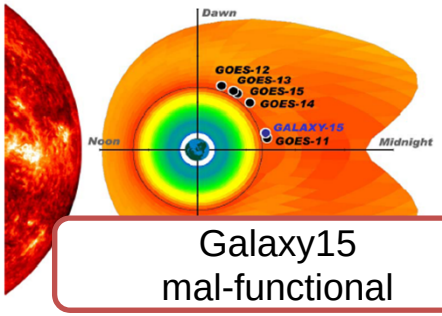
NICT ScienceCloud

“Crown-milk view” of reconnections at the dayside magnetosphere

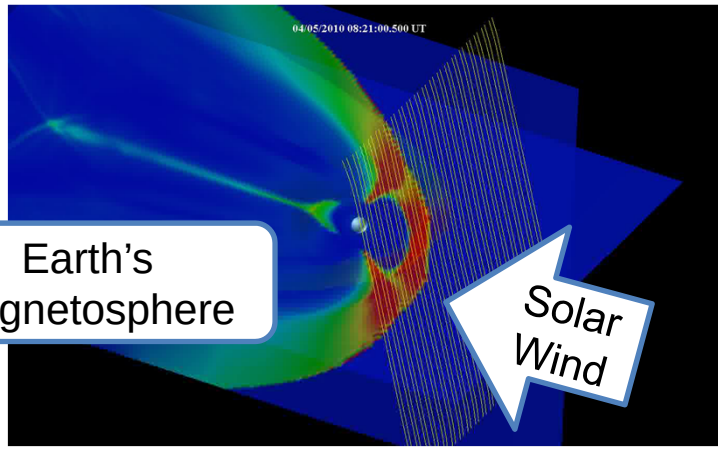
The first successful visualization of solar-wind transfer into magnetosphere

Galaxy 15 (133 W) Anomaly 09:48 UT

Satellite Locations

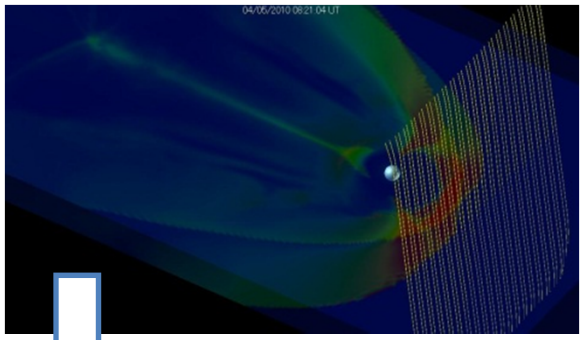


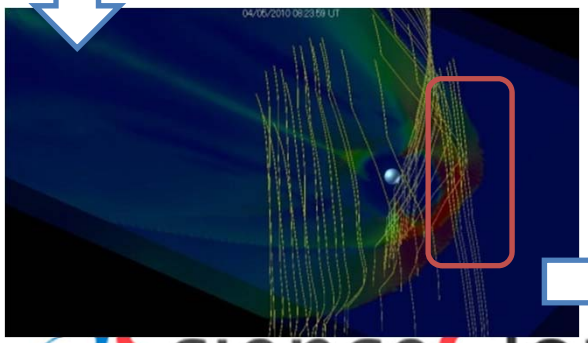
Galaxy15 mal-functional

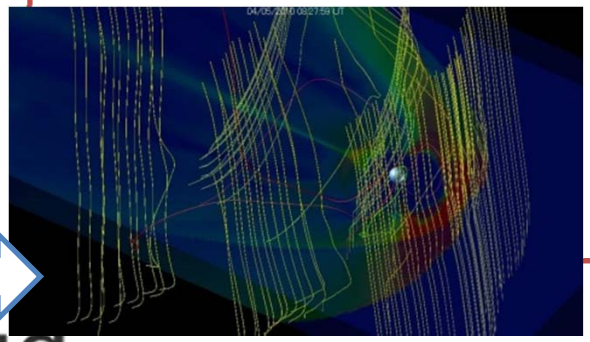


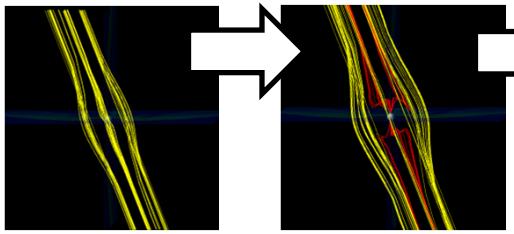
Open Question

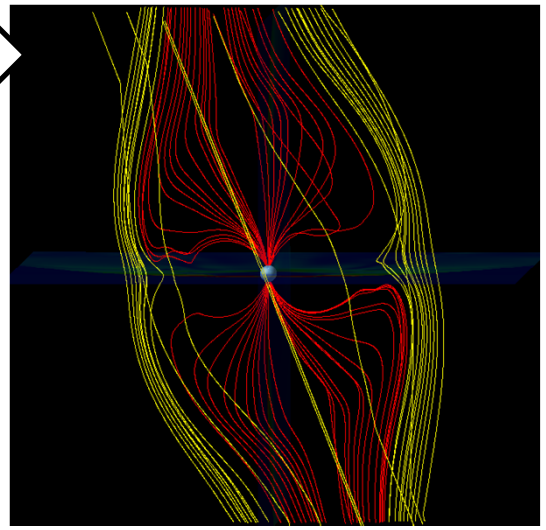
How does the solar-wind penetrate into Magnetosphere?



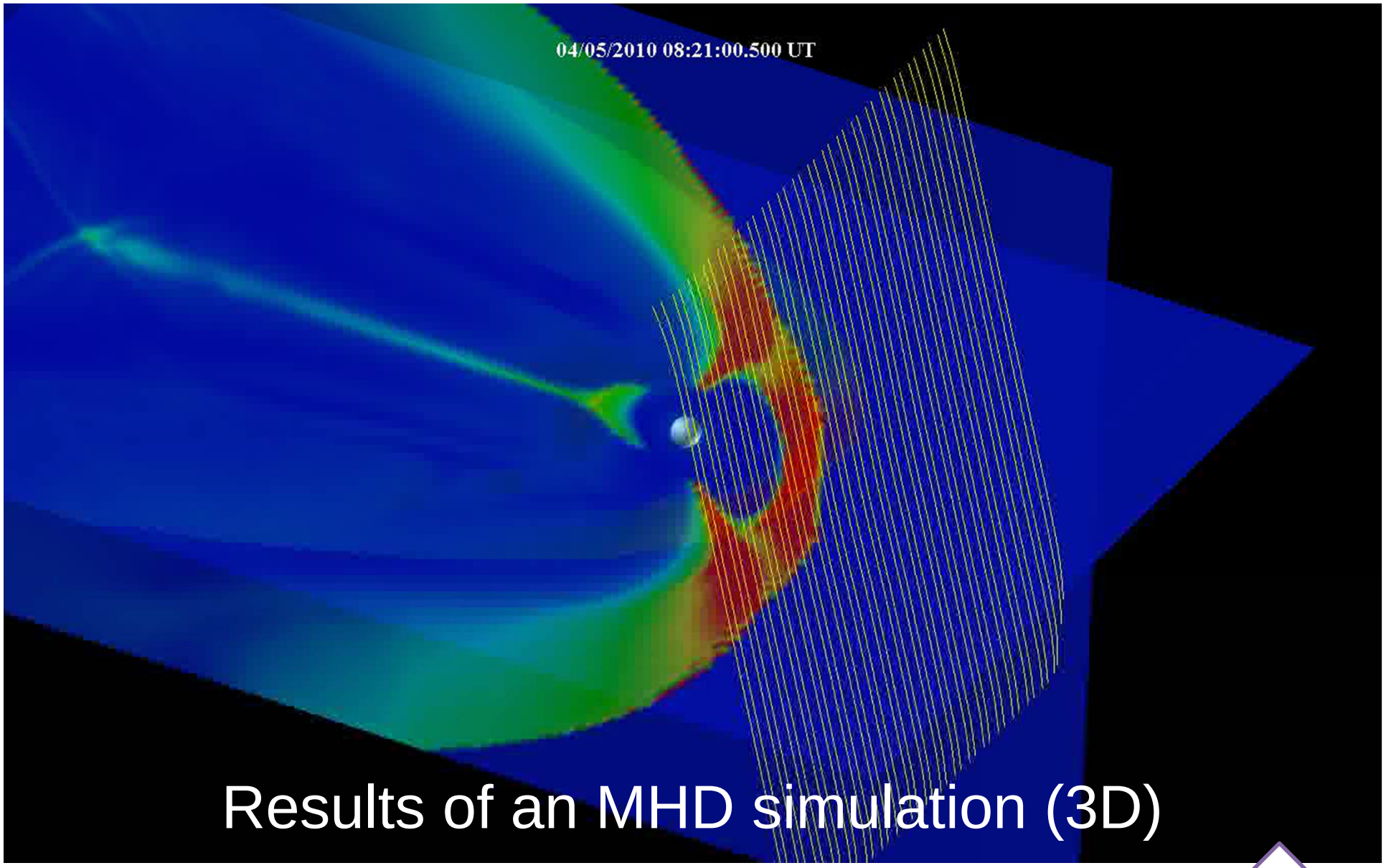








NICT ScienceCloud



Results of an MHD simulation (3D)

Global MHD simulation (1 hour=7200 steps)

NICT ScienceCloud

64bit 3D player developed by NICT

科学を行う方法(宇宙科学の場合)

ビッグデータ

データ指向型科学を実現するには…

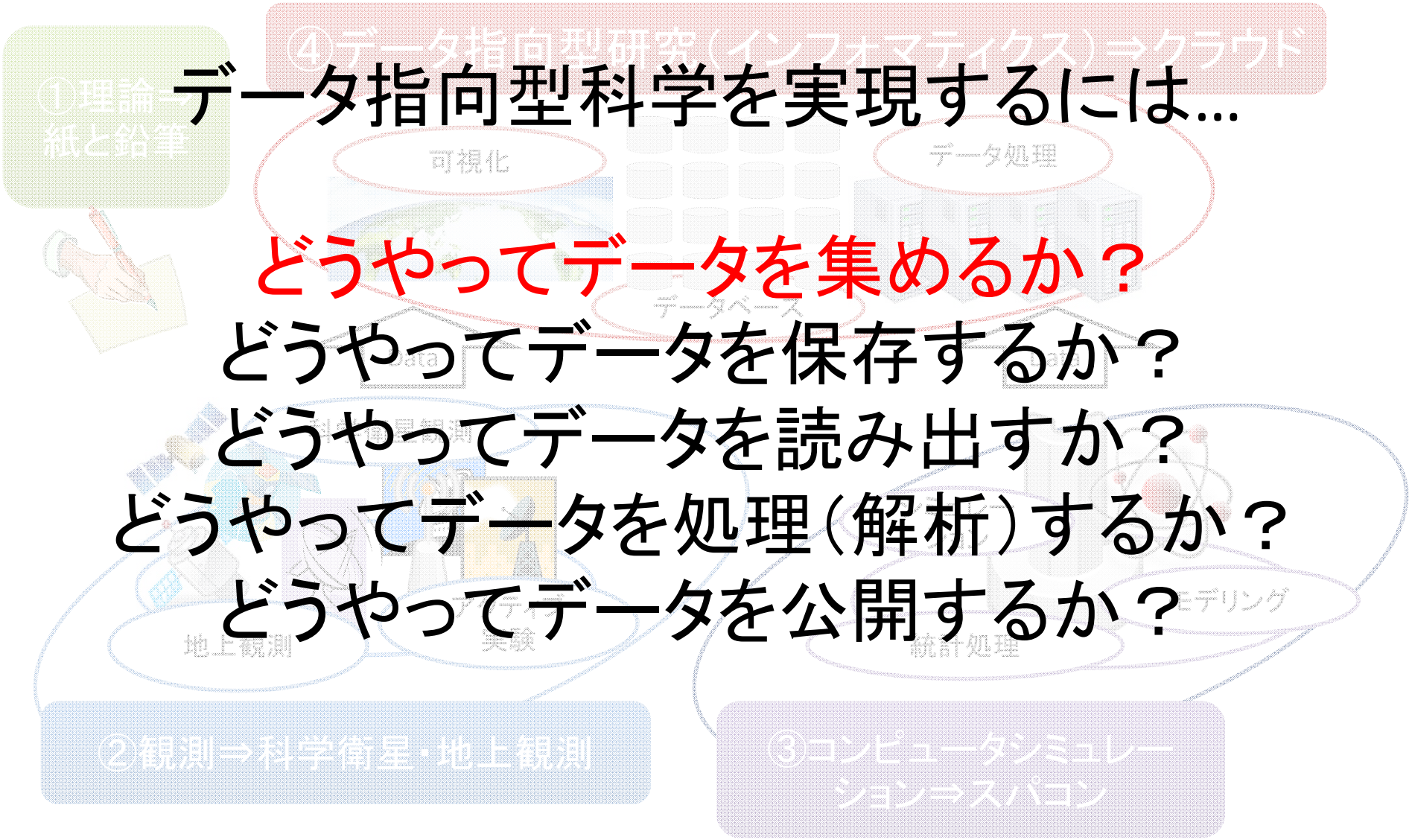
どうやってデータを集めるか？

どうやってデータを保存するか？

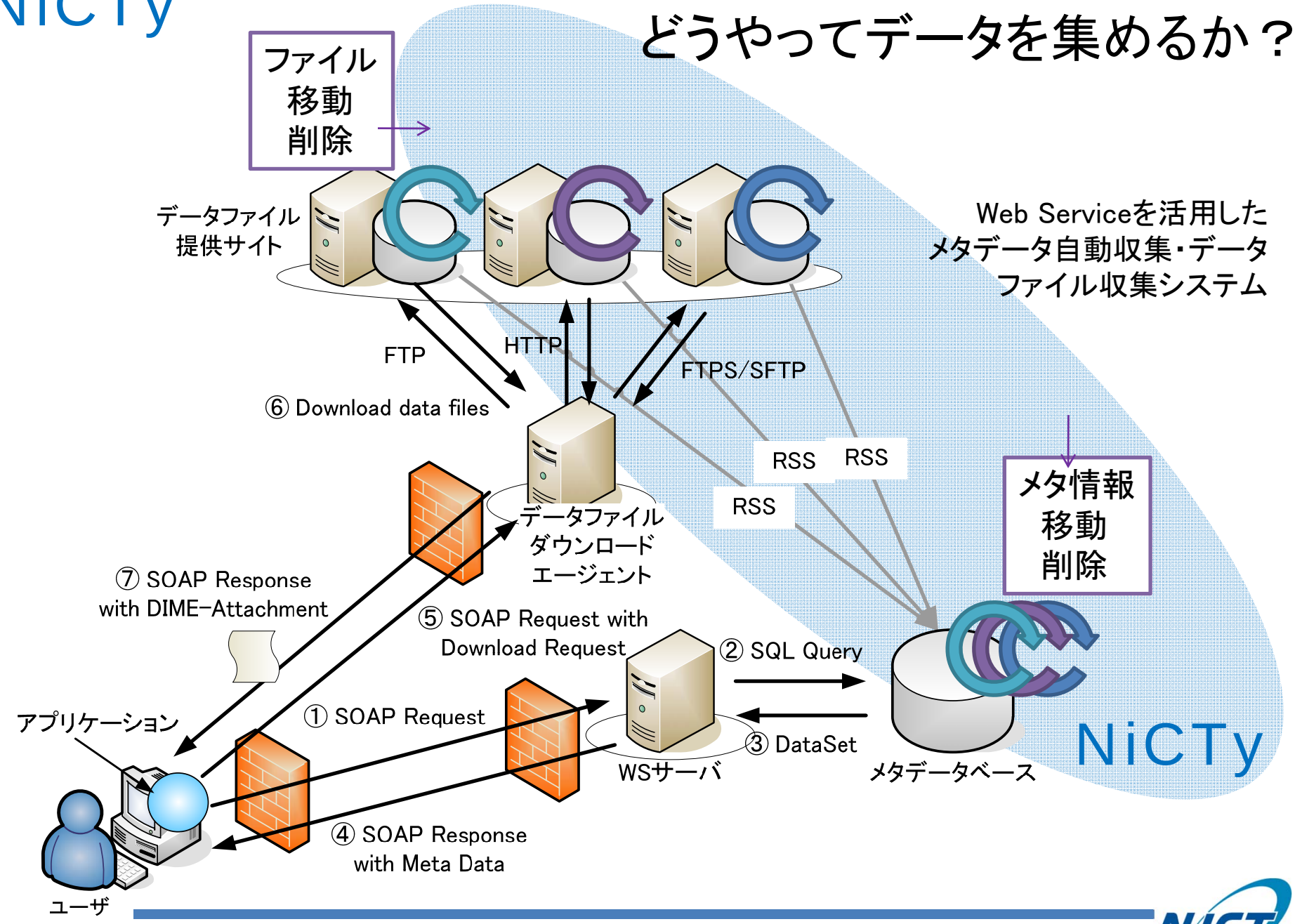
どうやってデータを読み出すか？

どうやってデータを処理(解析)するか？

どうやってデータを公開するか？



どうやってデータを集めるか？



科学を行う方法（宇宙科学の場合）

ビッグデータ

データ指向型科学を実現するには…

どうやってデータを集めるか？

どうやってデータを保存するか？

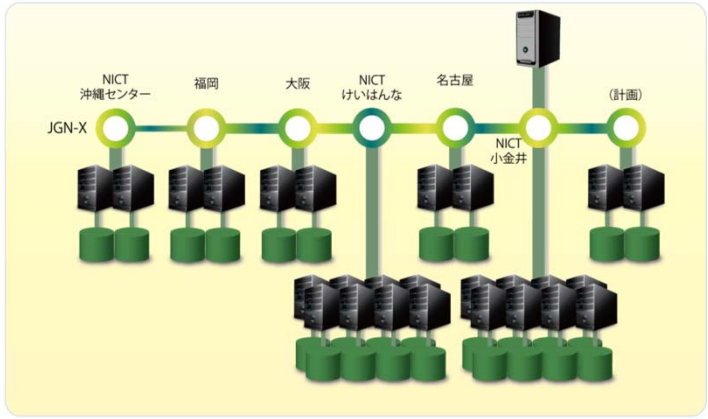
どうやってデータを読み出すか？

どうやってデータを処理（解析）するか？

どうやってデータを公開するか？

どうやってデータを保存するか？

NICTサイエンスクラウド内のデータ保存



冗長性

データの
ダウンロード

データの
保存

データの
ダウンロード

データの
保存

データの保存

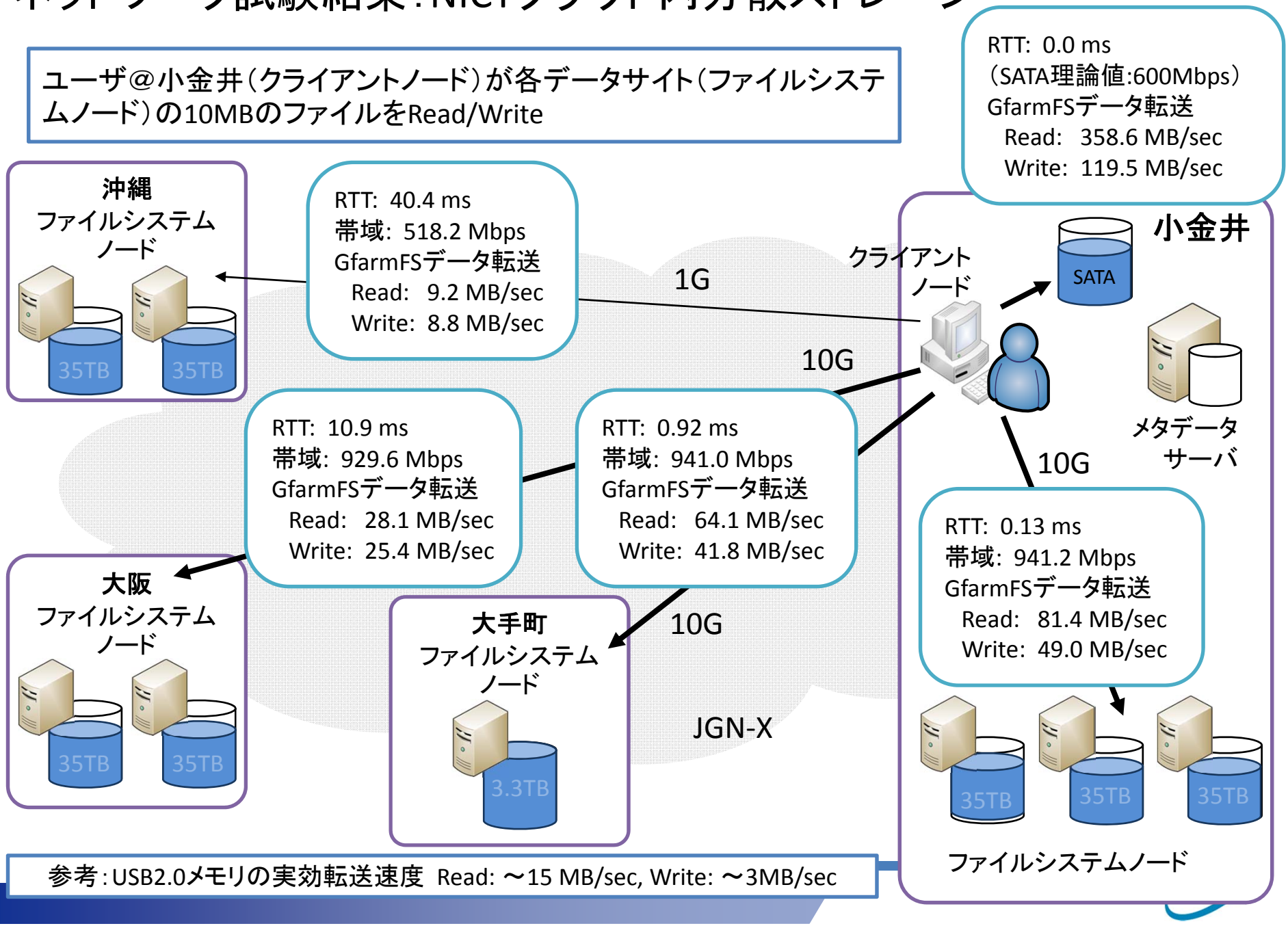
データの
アップロード

データの
ダウンロード

データの保存

ネットワーク試験結果：NICTクラウド内分散ストレージ

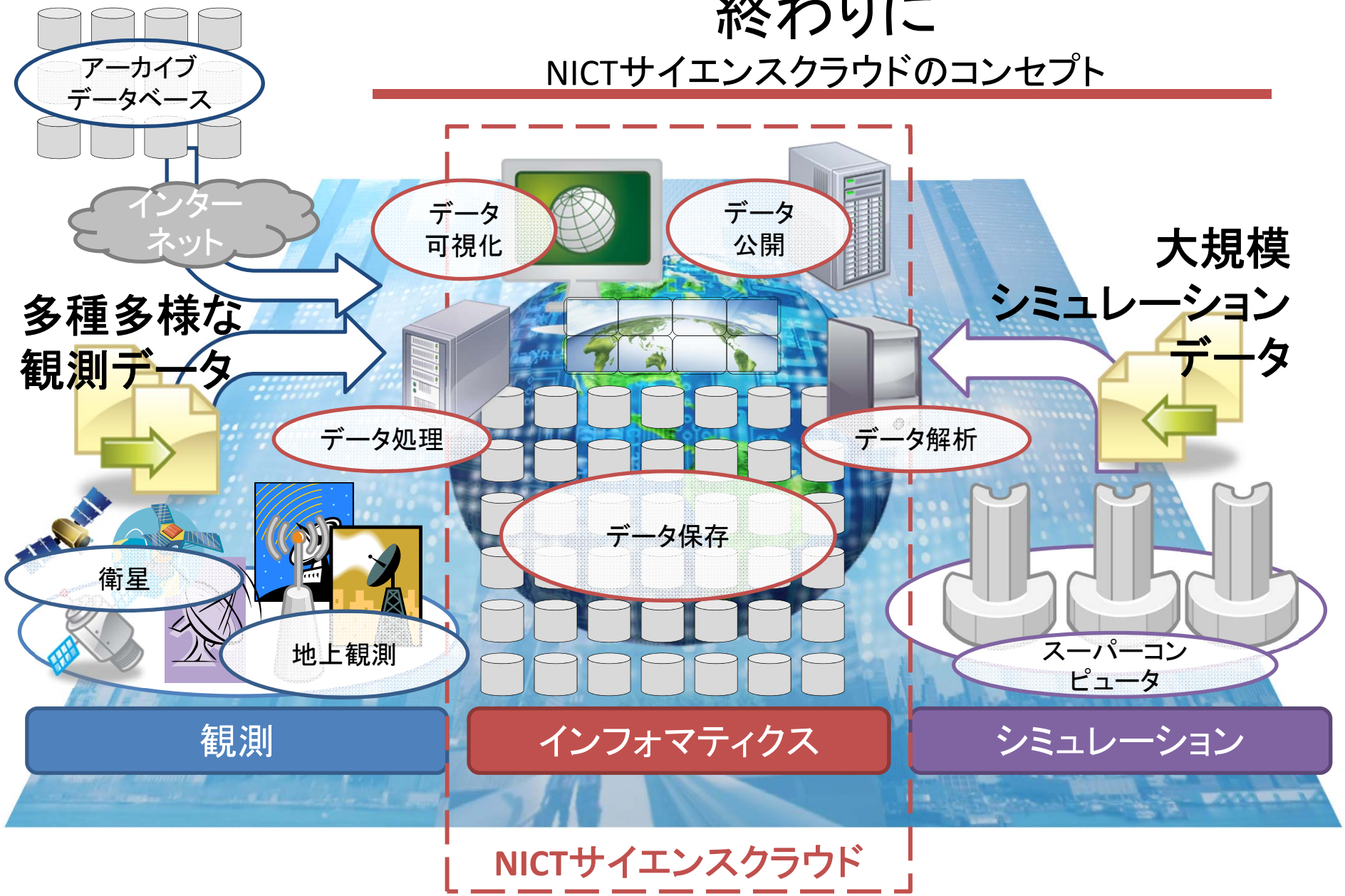
ユーザ@小金井（クライアントノード）が各データサイト（ファイルシステムノード）の10MBのファイルをRead/Write



参考：USB2.0メモリの実効転送速度 Read: ～15 MB/sec, Write: ～3MB/sec

終わりに

NICTサイエンスクラウドのコンセプト



クラウドは総合技術！

ご清聴ありがとうございました!

NICT ScienceCloud