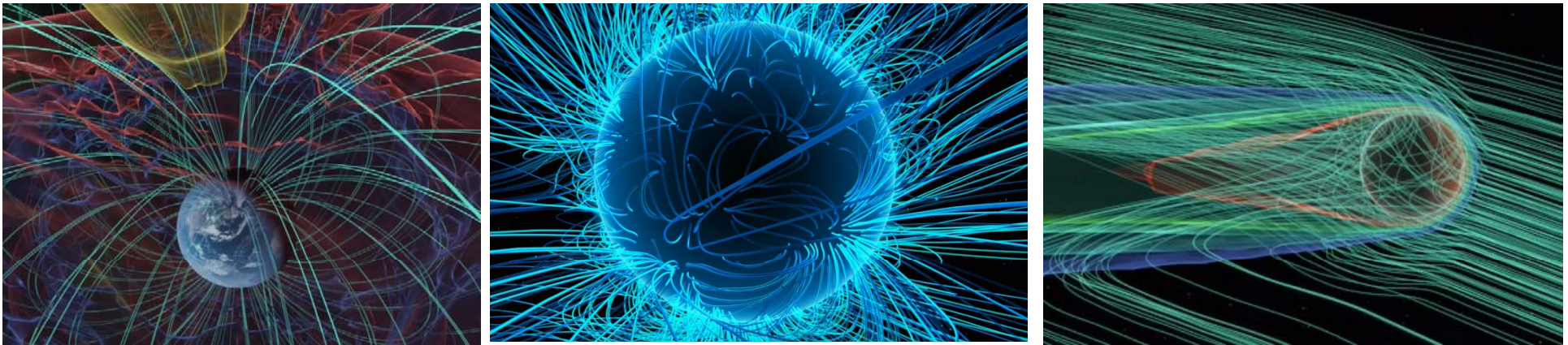


NICTサイエンスクラウド： 現状と将来計画



村田 健史

情報通信研究機構(NICT)

電磁波計測研究センター宇宙環境計測グループ

速報: WDS-IPOについて

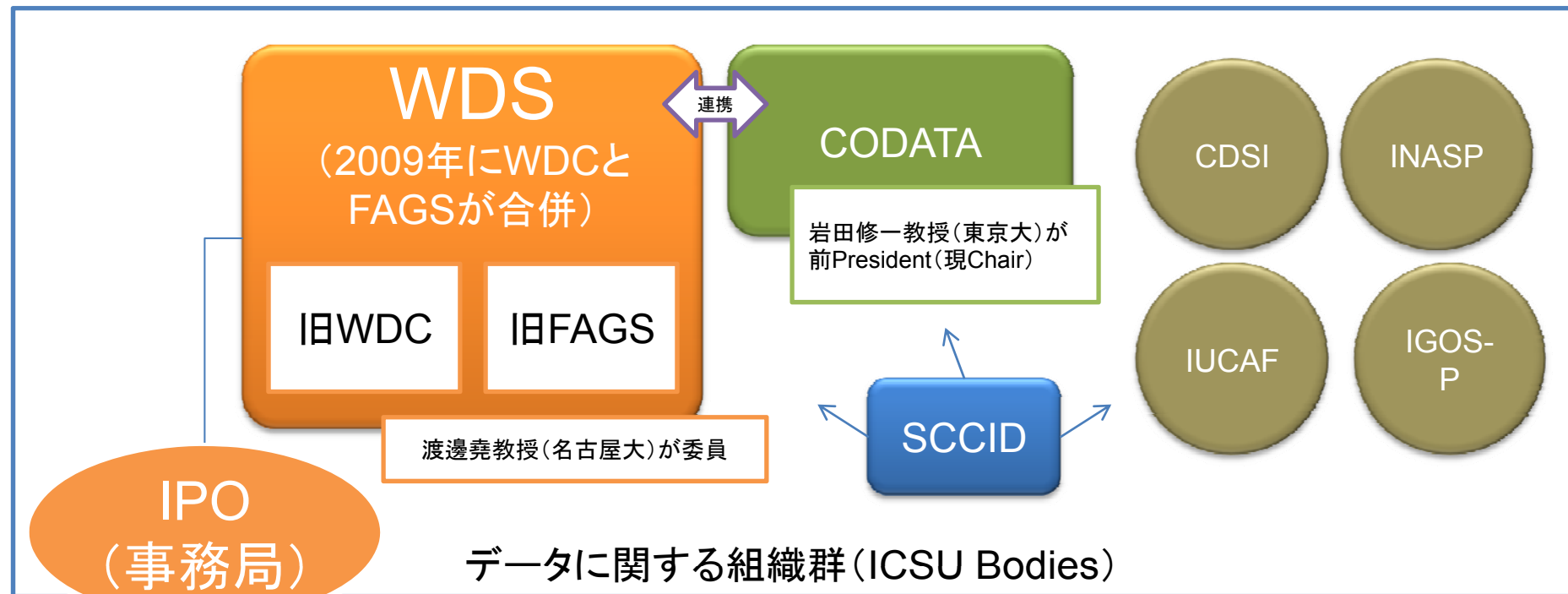
黒田玲子副会長(東京大)

ICSU (the International Council for Science: 国際科学会議)
1931年～

日本学術会議

金沢一郎会長

土居範久教授(慶応大)が現在のPAA Panelメンバー



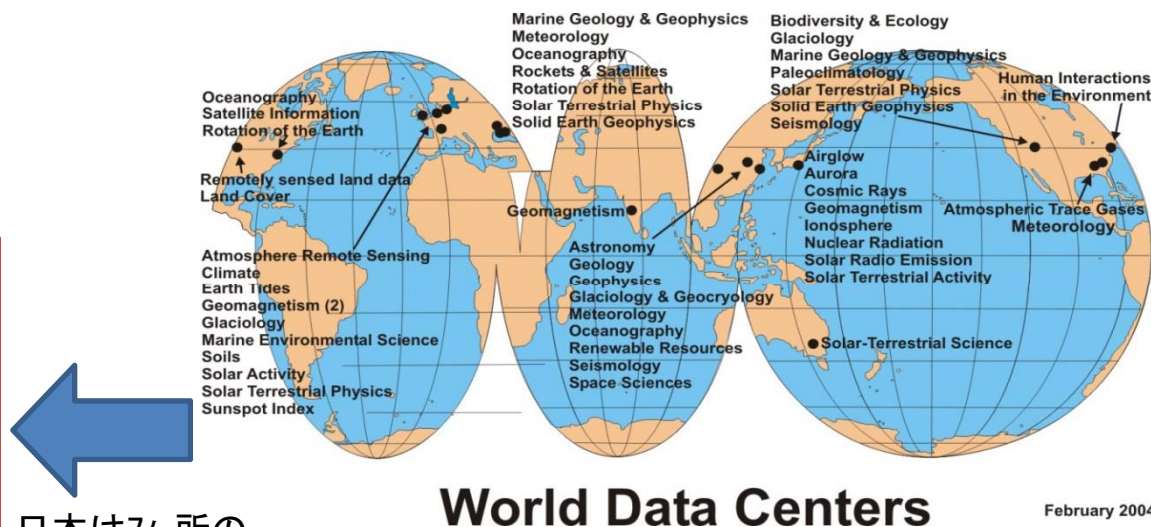


ICSU World Data Center (旧WDC)

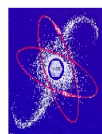


- 1957年に国際地球観測年(IGY)を期に設立(ICSU)
- 分野横断型国際データ組織の先駆
- 政治形態に捉われないデータの国際共有
- 品質管理されたデータの長期保全と提供
- 12ヶ国50ヶ所(主に天文・地球科学分野)
- **Full and Open Access** をモットー

分野	日本の所属機関
大気光	国立天文台
太陽電波	国立天文台
オーロラ	国立極地研究所
科学衛星	宇宙科学研究所
電離層	情報通信研究機構
宇宙線	名古屋大学太陽地球環境研究所
地磁気	京都大学地磁気世界資料解析センター



日本は7ヶ所の
WDCを保有



Federation of Astrophysical and Data Analysis Services (旧FAGS)

- 1957年、国際地球観測年(IGY)を期に設立(ICSU)
- 再解析データ、指数などの提供、研究プロジェクトの推進
- データセンター機能は持たない
- 13項目のサービス(天文、地球物理、電波科学分野)
- NICTの宇宙天気情報センターはInternational Space Environmental Service の一環

WDCとFAGSの問題点 → WDSへ

- 個々のデータセンターやサービスが個別に活動。
- 全体を束ねるシステムが無い。
- 活動方針の決定など、ガバナンスが機能していない。
- 天文・地球科学系分野に偏重

WDSを基盤とする世界データセンター計画 NICTサイエンスクラウド

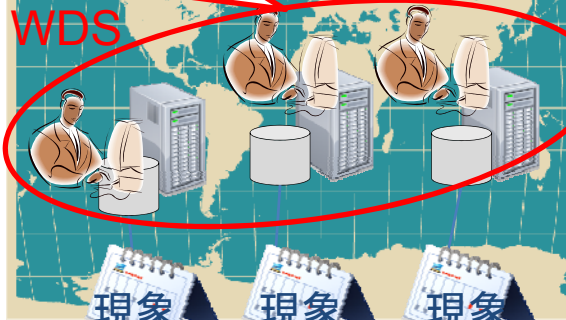
アメリカ海洋大気圏局 (NOAA)、NASA (GNSS)、Max-Planck研究所、パリ研究所、中国科学技術院 (CAS)、JAXA (日本)...

WDSの立候補機関 (世界26カ国・96組織)

USA	21
Germany	10
United Kingdom	10
France	7
Japan	7
China	5
Russia	5
Australia	4
Bergium	4
South Africa	3
Ukraine	3
Canada	2
Spain	2
その他13カ国	1

WDSプロジェクトの意義

- 巨大分散データベースとJGNxとの連携により、新の大規模クラウドデータセンタが実現。
- 最新の地球科学データを即座に利用できることにより、宇宙天気・宇宙科学スパコンシミュレーションが大きく進展する。
- けいはんな・神戸で行っているWeb検索やデータマイニング等の情報処理技術の活躍の場となる。



現象 現象 現象

Data Data Data

自動収集

NICTサイエンスクラウド

WDS-検索サービス

WDS-ファイルサービス

WDS-API

世界規模
①メタデータベース
②観測データベース
③イベントデータベース

10PB級分散
ストレージ

③イベント情報
日時・場所・減少

日時・場所を利用するすべてのアプリケーションからWDSイベント情報を検索・表示

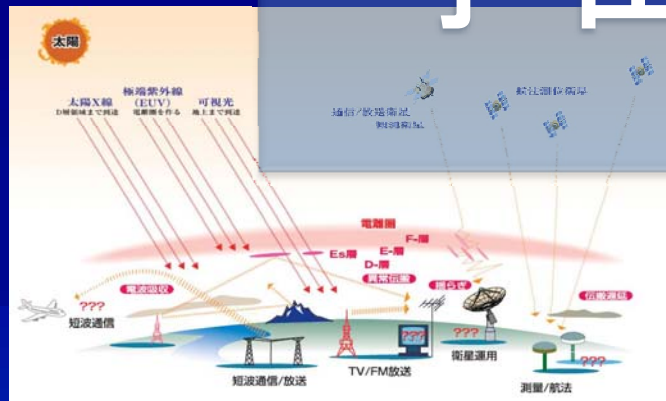
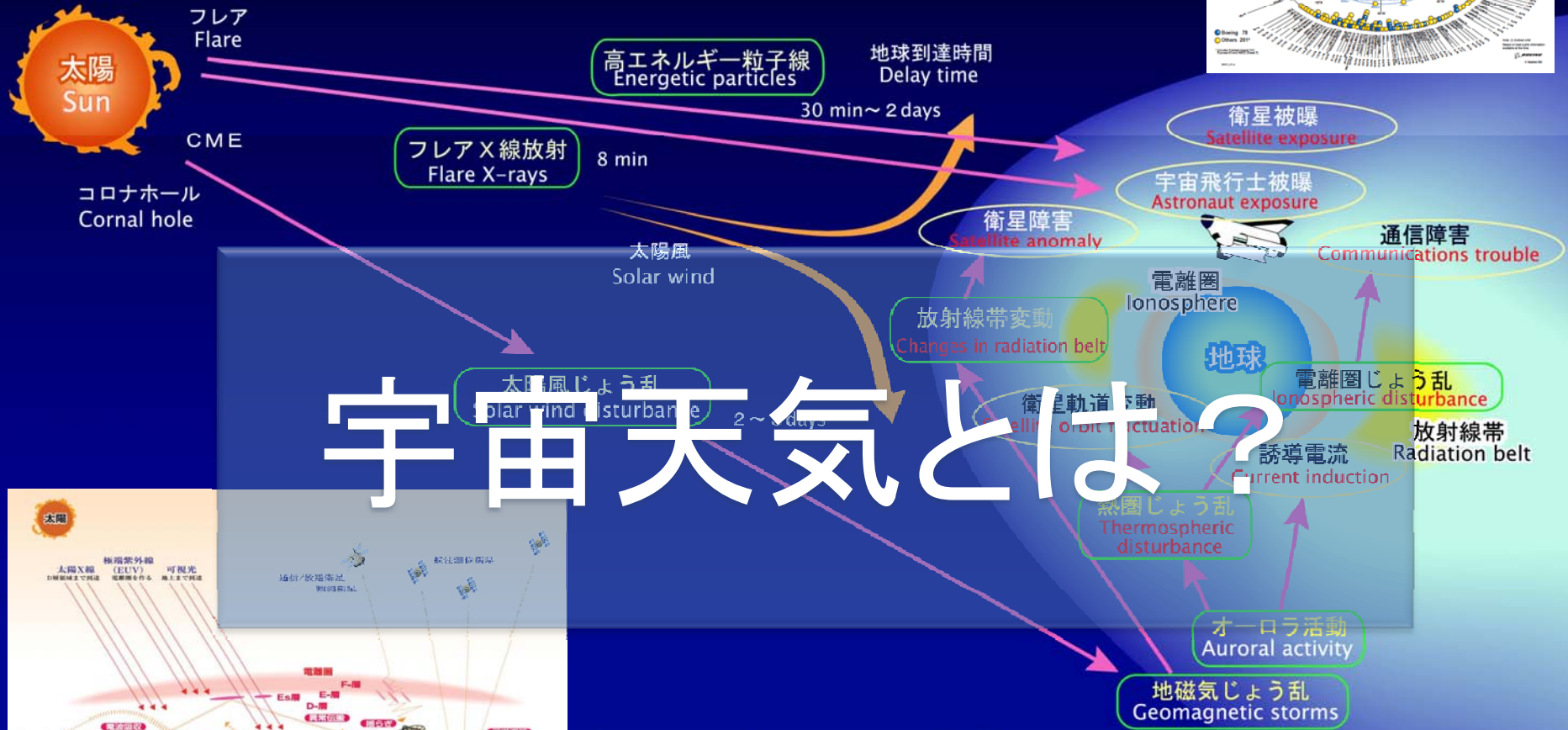
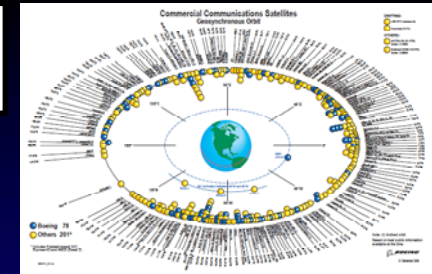
①メタ情報

世界のあらゆる学術データを検索・ダウンロード

②観測ファイル情報

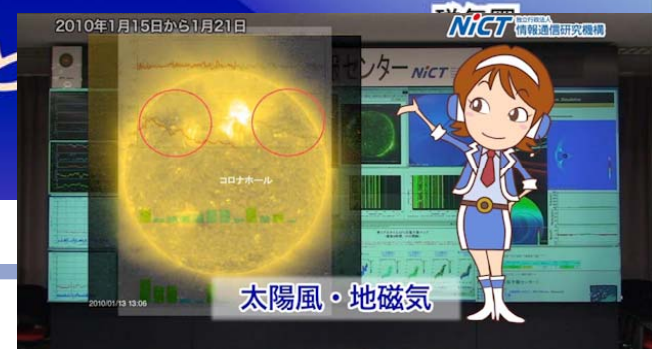
世界のあらゆる学術データを表示・解析・分析

静止軌道上の衛星
(300以上)



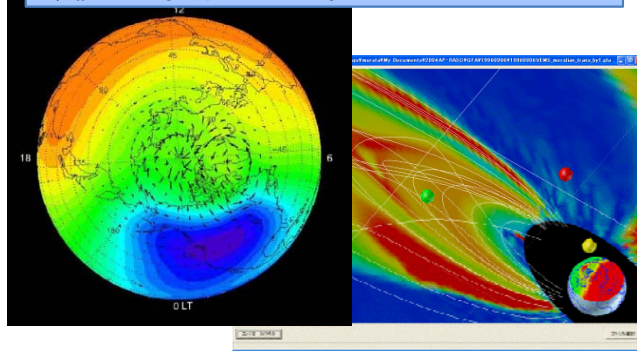
電離層(電離圏)の電波伝搬障害

Space weather: cause and effect
宇宙環境擾乱の発生と

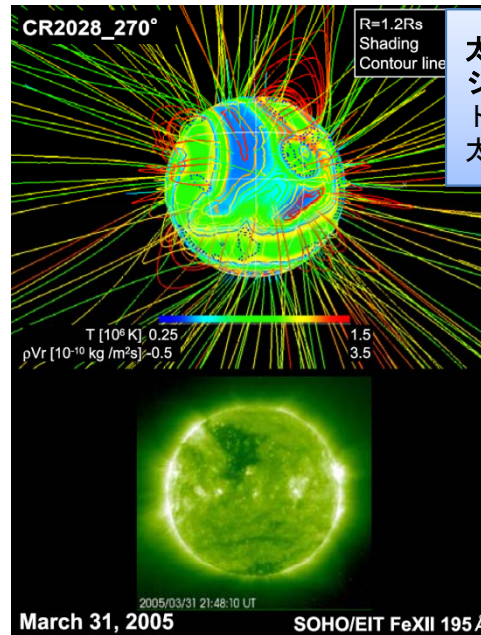
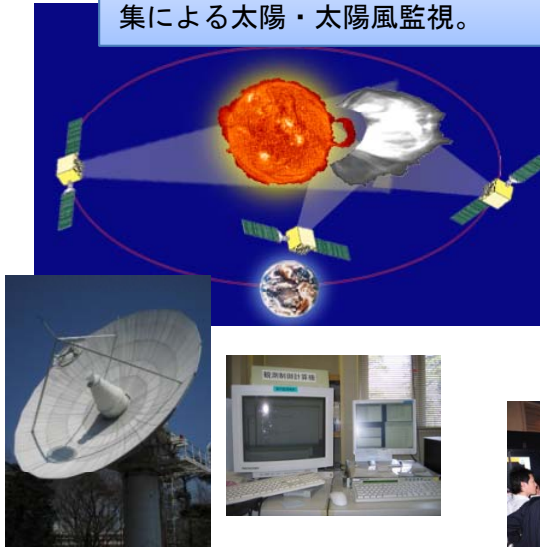


NICT research activities

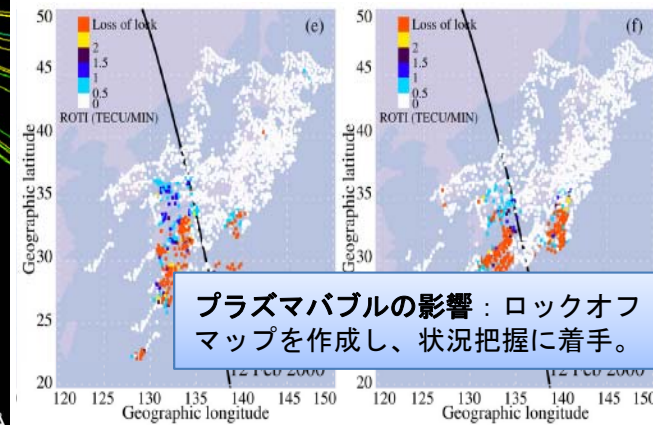
コンピュータシミュレーションによる磁気圏および電離圏電子密度変動の予測：太陽風の影響による磁気圏の変動や、下層大気の影響を含めた全球電離圏電子密度変動を数値的に予測している。



衛星による太陽・太陽風監視：二機のSTEREO衛星とACE衛星データの独自収集による太陽・太陽風監視。



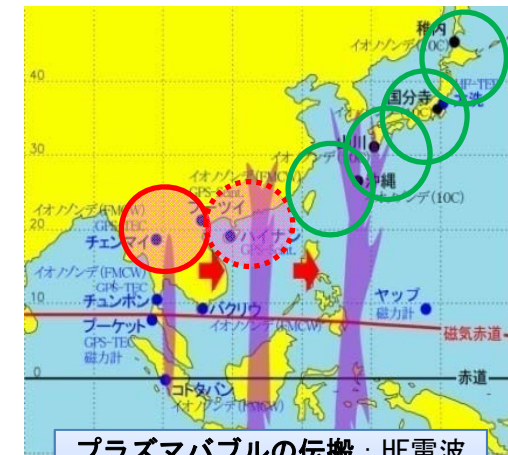
太陽磁場と太陽風のシミュレーション：磁気流体（MHD）コードにより太陽磁場と太陽からの太陽風のシミュレーション。



プラズマバブルの影響：ロックオフマップを作成し、状況把握に着手。



TWDによる大規模可視化
タイルディスプレイにより高解像度での3次元可視化を行う。



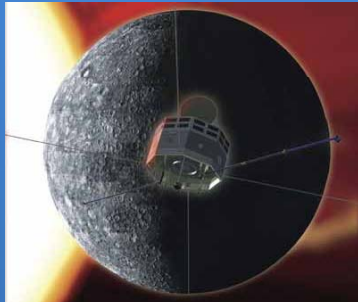
プラズマバブルの伝搬：HF電波やGPSを用いて限られた観測点から、およその伝搬状況を把握する手法を確立。

Mega-Science projects on Space Science(1)

Two huge data sets

Observation

Scientific Satellite Missions



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



<http://www.isas.jaxa.jp>

<http://www.isas.jaxa.jp>

~\$ 100
million

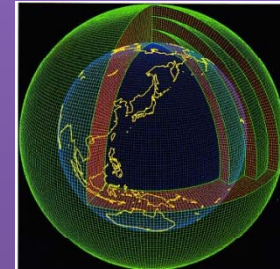


Simulation

Super Computer Projects



<https://aspara.asahi.com/blog/netnowplus/entry/uSreel7JOG>



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
billion

Peta-byte

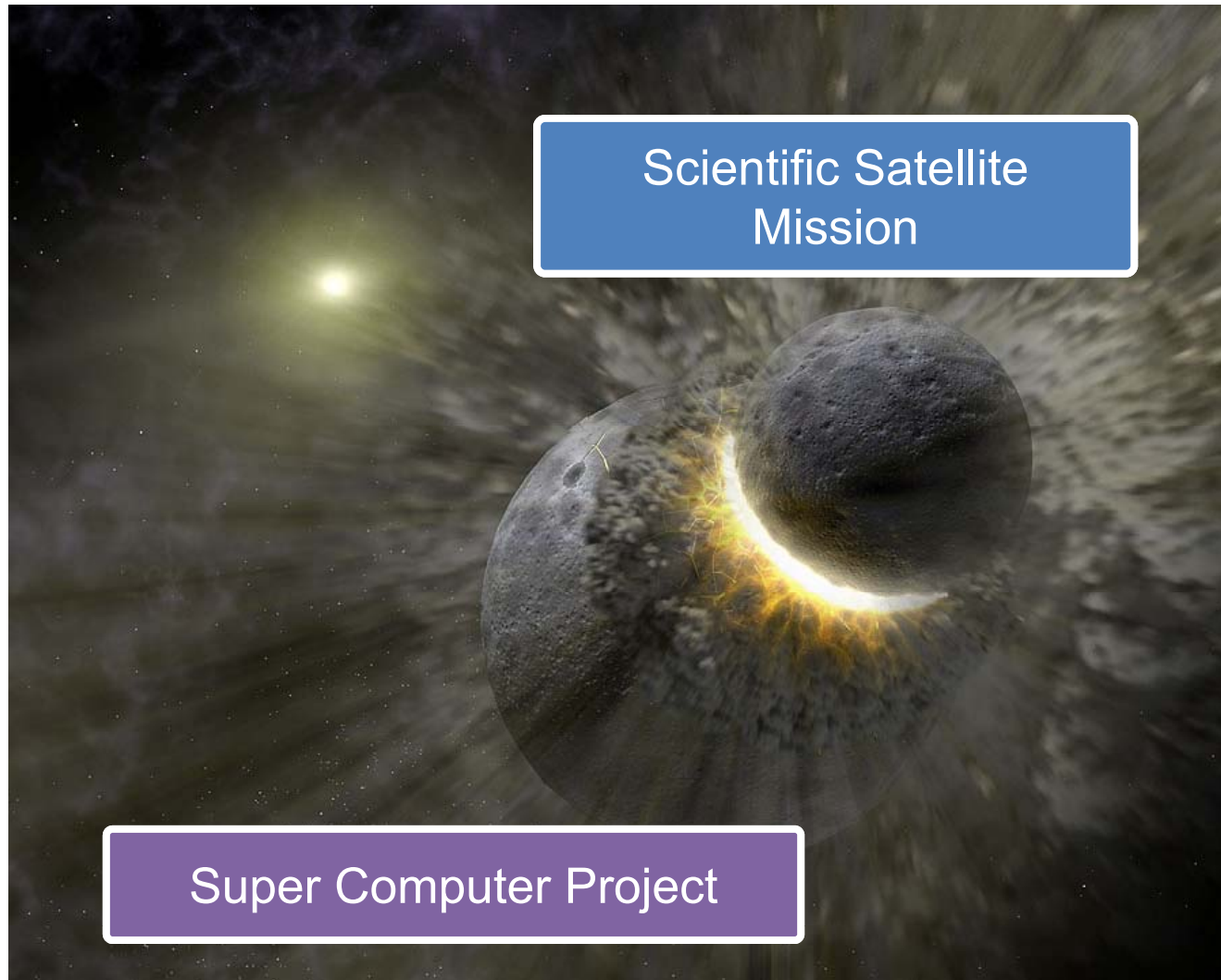
Peta-byte

Huge Data

どう解析？

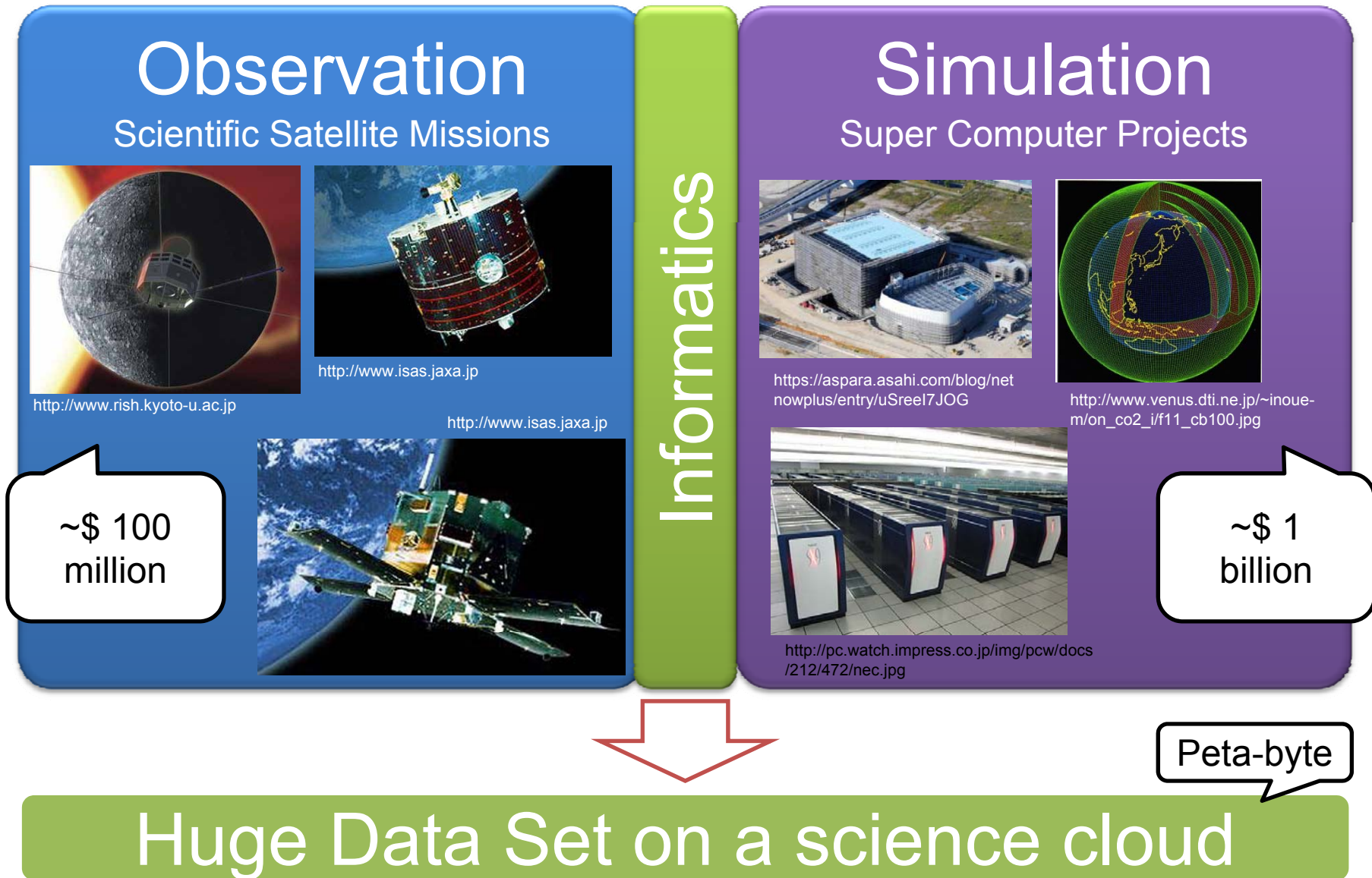
Huge Data

Collision of Mega-Sciences!

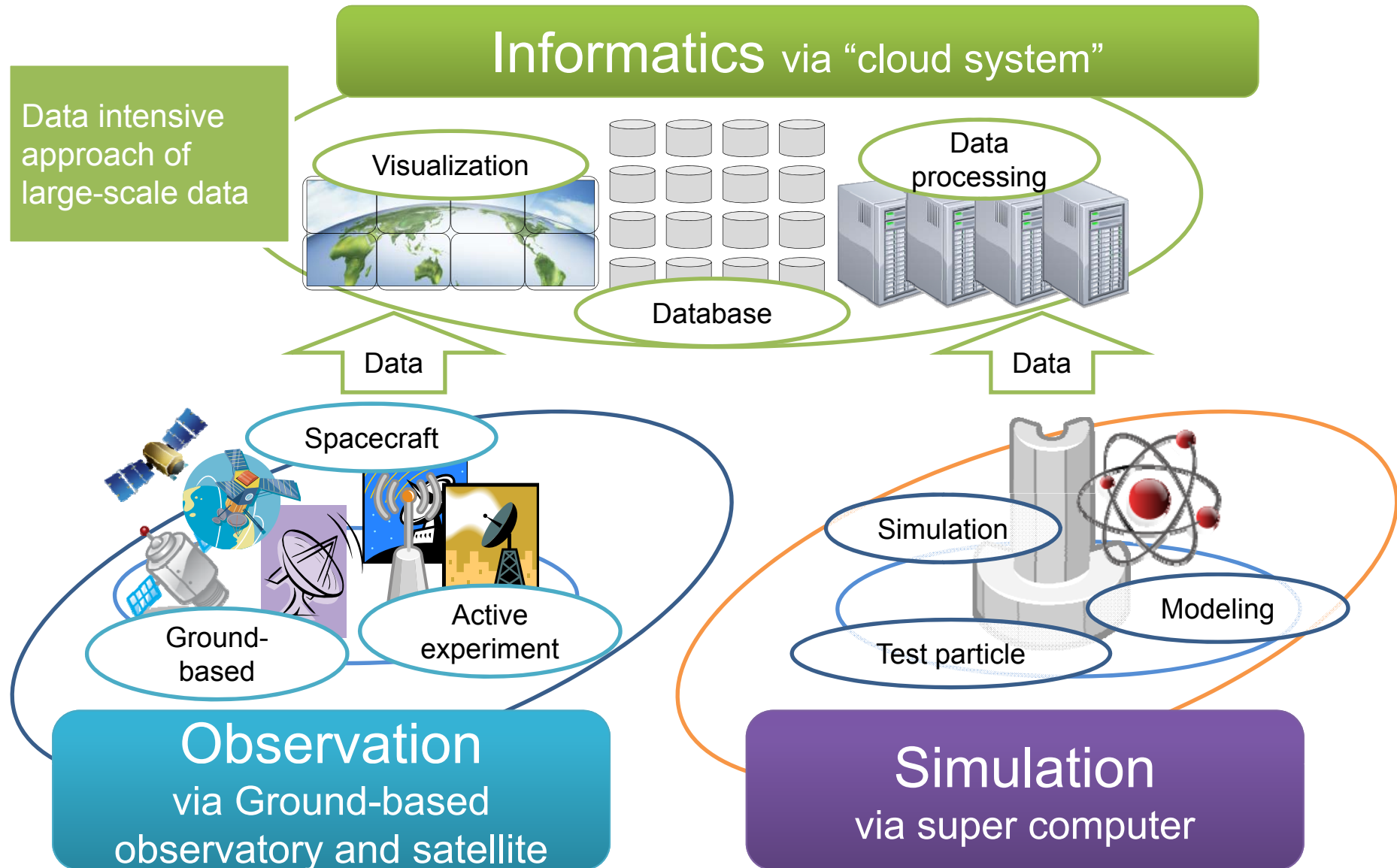


Mega-Science projects on Space Science(2)

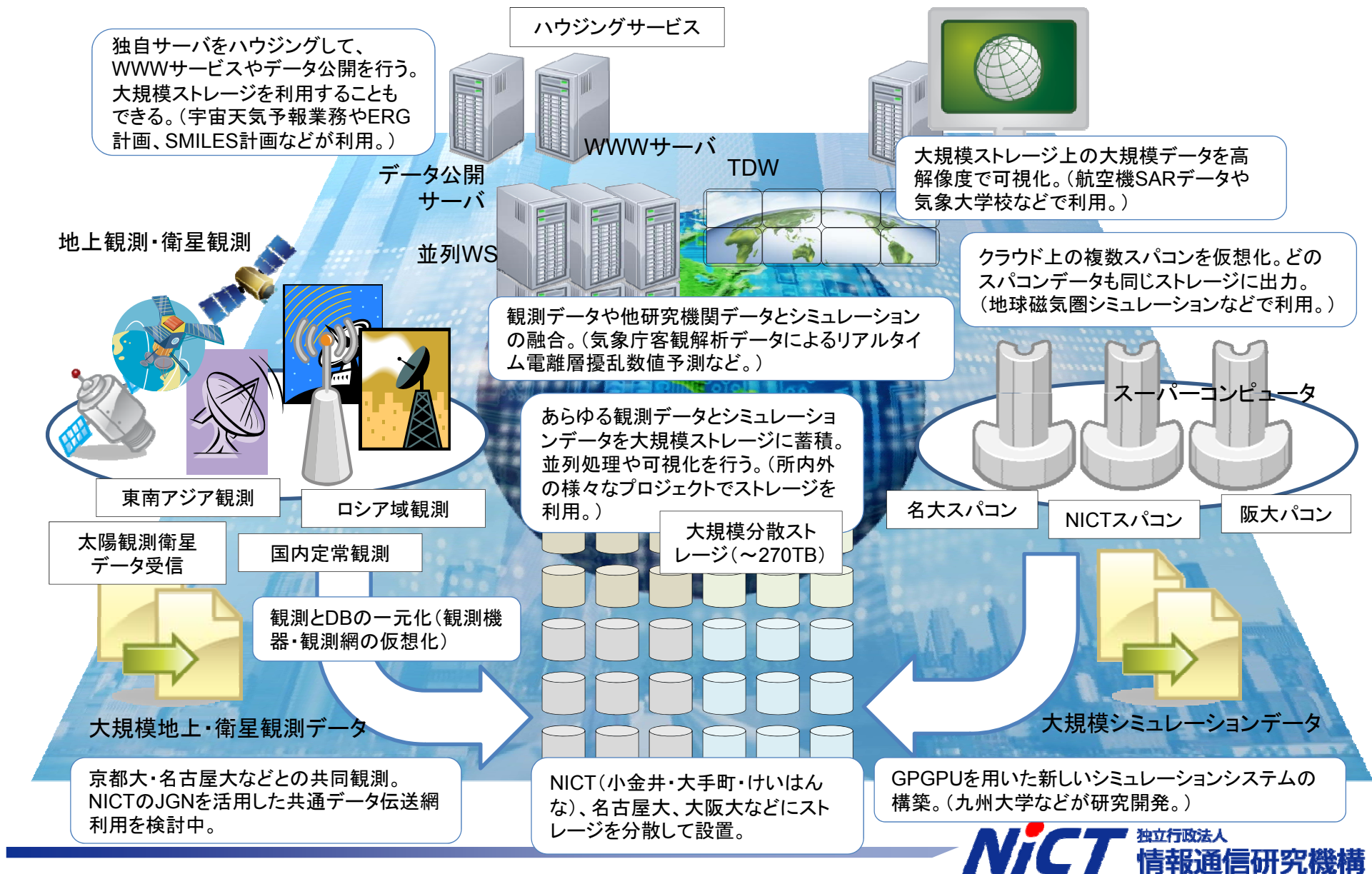
A roll of Informatics



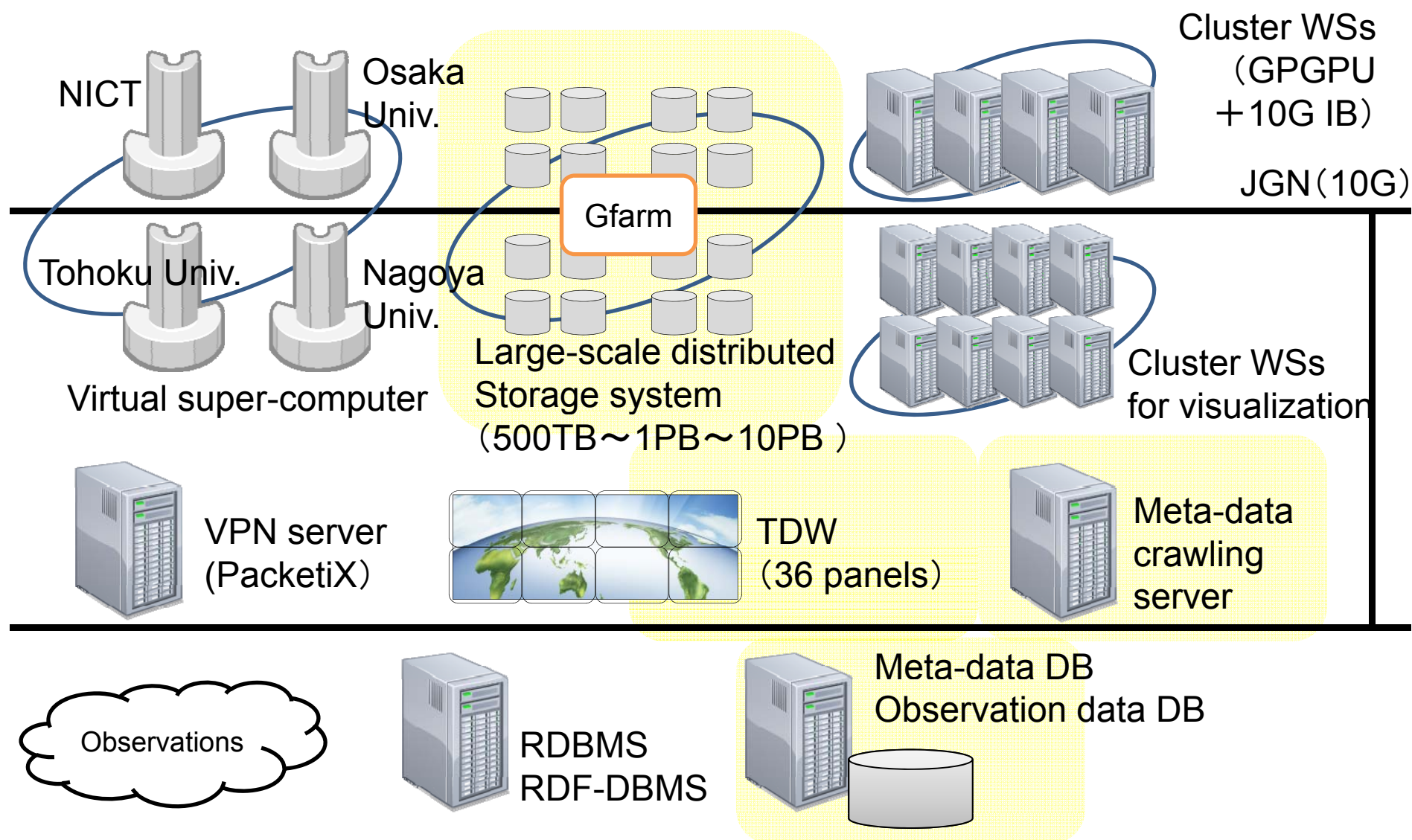
The third approach to Space Weather Research



NICTサイエンスクラウドのコンセプト

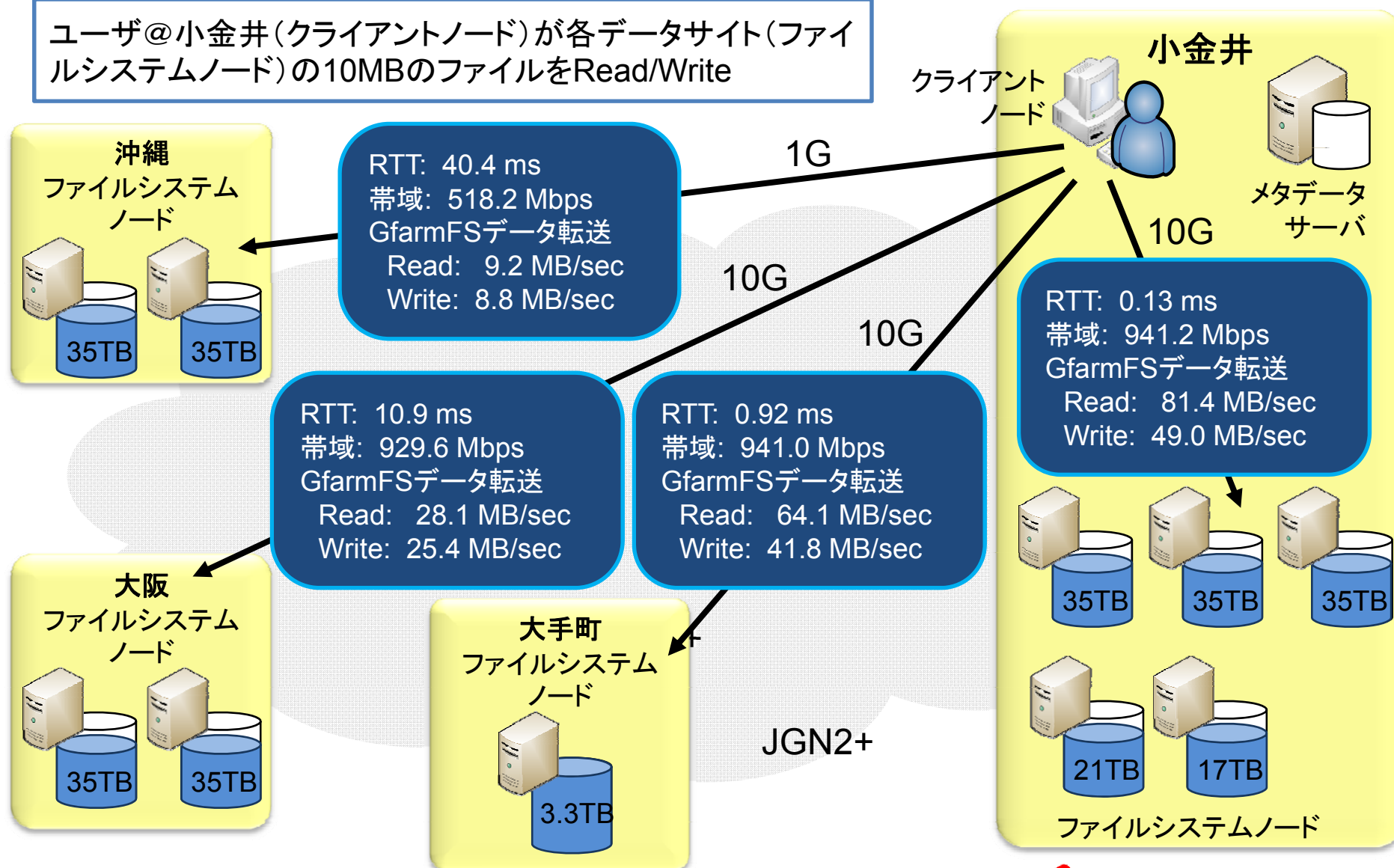


NICT science cloud (so far developed)



国内規模での分散ストレージ(Gfarm)のパフォーマンス

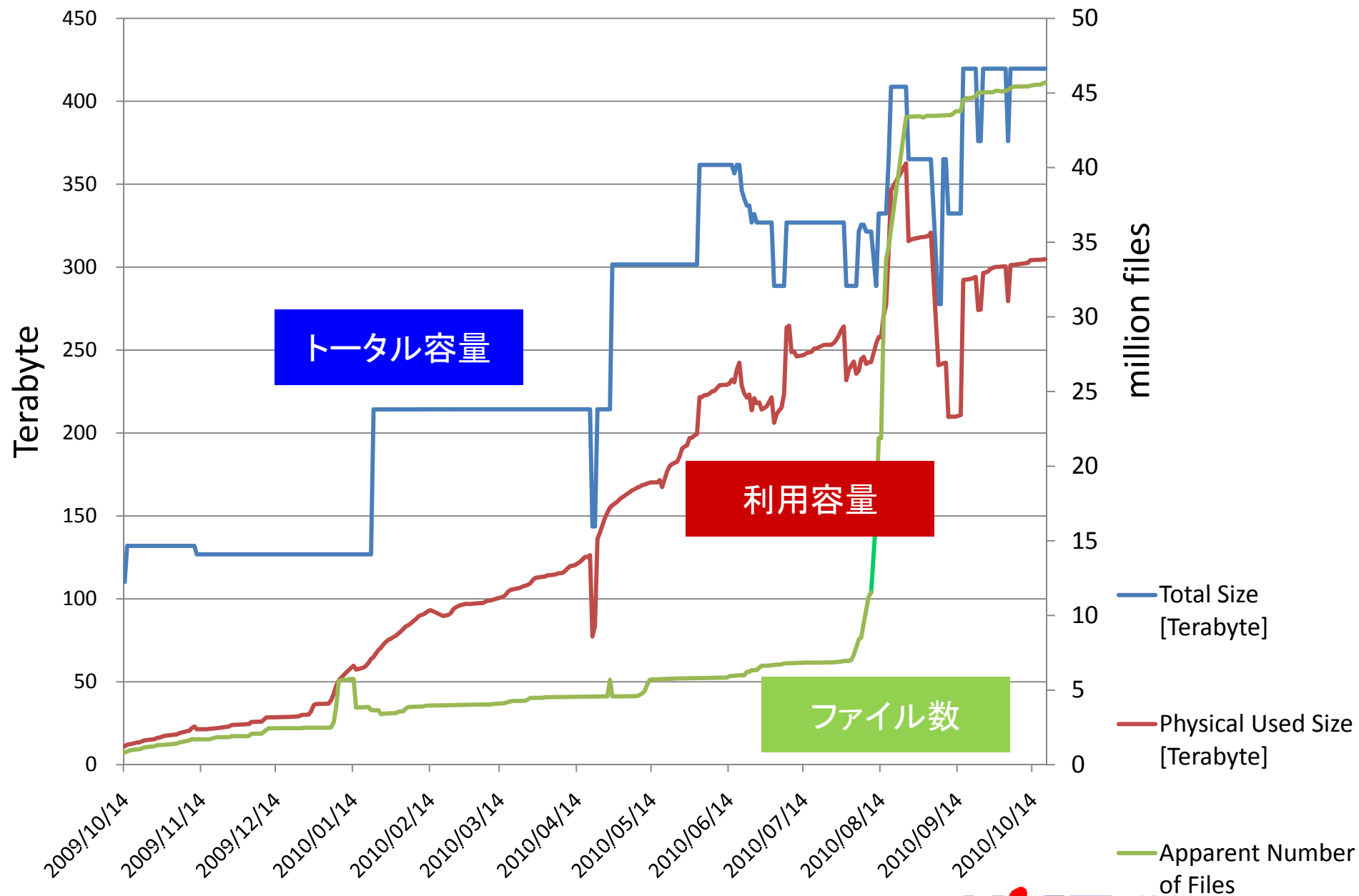
ユーザ@小金井(クライアントノード)が各データサイト(ファイルシステムノード)の10MBのファイルをRead/Write



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

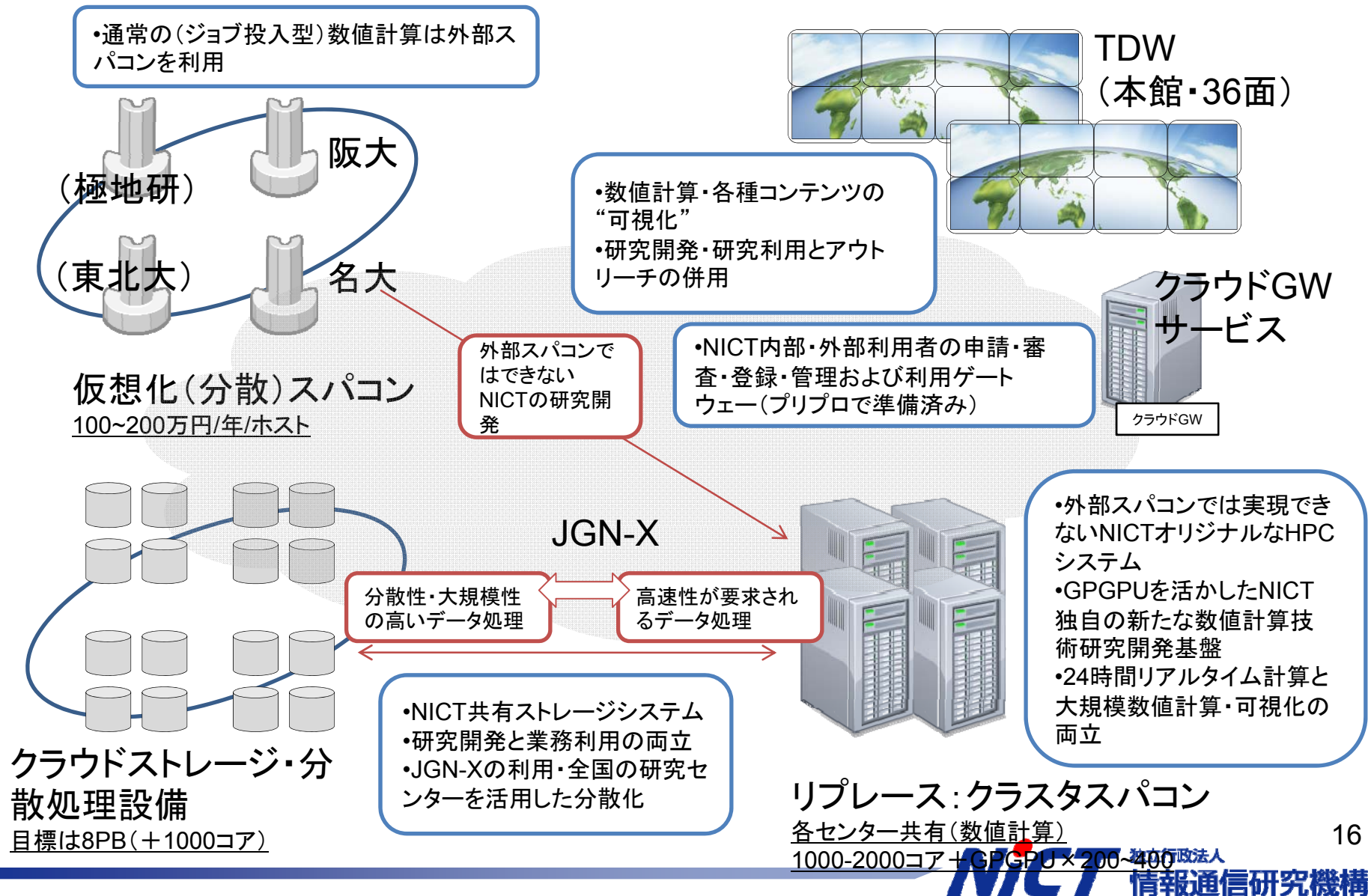
Gfarm storage

Operation	Form 2009/08/14 to 2010/10/20					
Used/Total	305 / 420 TB (-> 1PB in 2010)					
# of files	45,711,900					
Used area by users	26T 15T 9.3T 6.3T	fukxxxx tsuxxxx jyuxxxx reaxxxx	6.1T 5.8T 5.6T 5.1T	gyoxxxx shimxxx shixxxx xsfxxxx	3.5T 1.1T 790G 574G	muraxxxx inoxxxx morxxxx hshixxxx
File system	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%		



サイエンスクラウドによるNICT研究開発基盤

NICT共通研究基盤の充実(新しい研究成果・コスト削減)



附録: コンテナ型データセンター

IIJの実証実験の例



コンテナ本体は長さ約
8.6m、幅と高さがともに3m
くらい。



IIJの実証実験で使われているコンテナ型のデータセンターユニット。ラックやIT機器を搭載するITモジュールと空調モジュールで構成される。



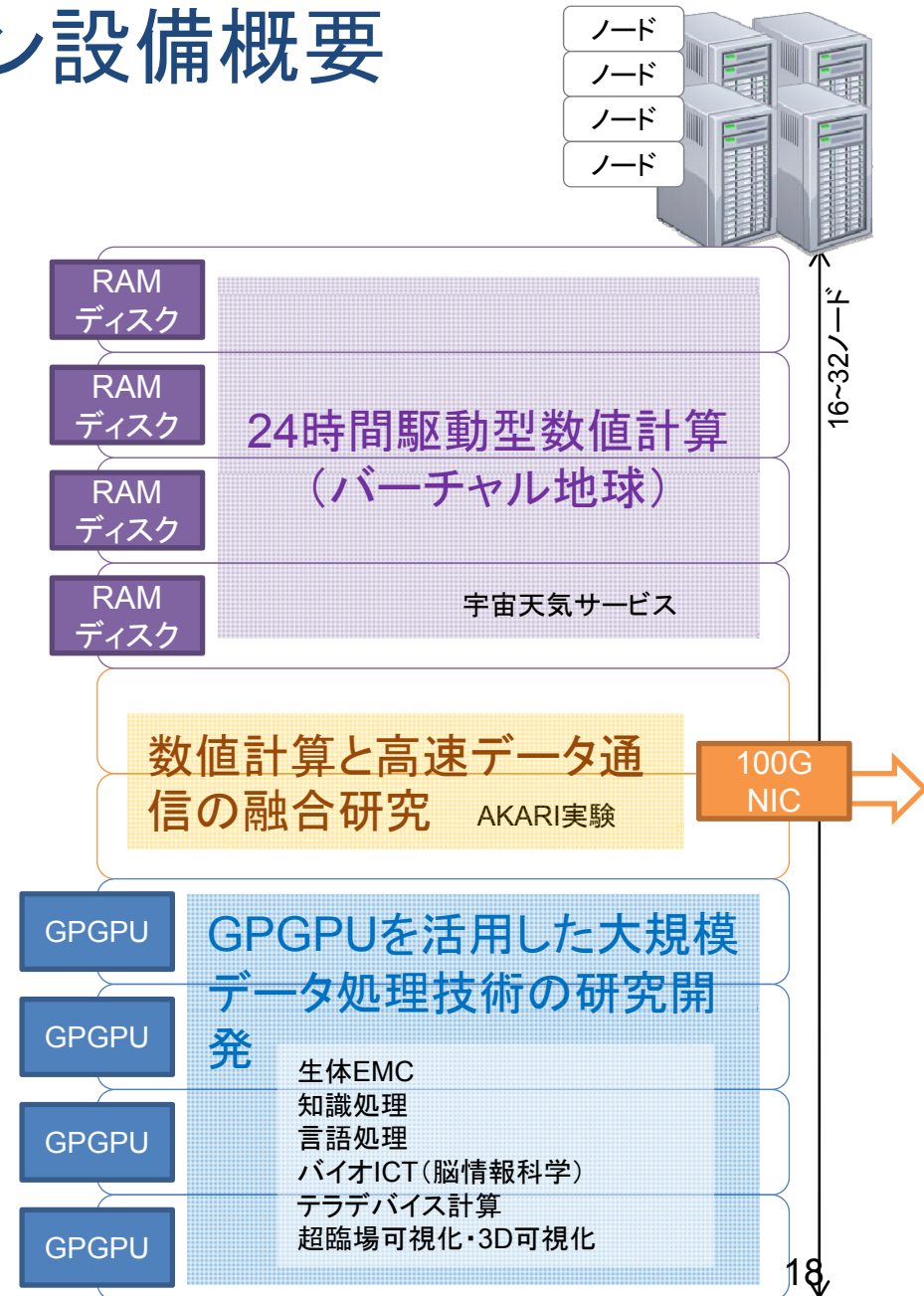
IT機器の背面の排気側にあたるホットアイル。



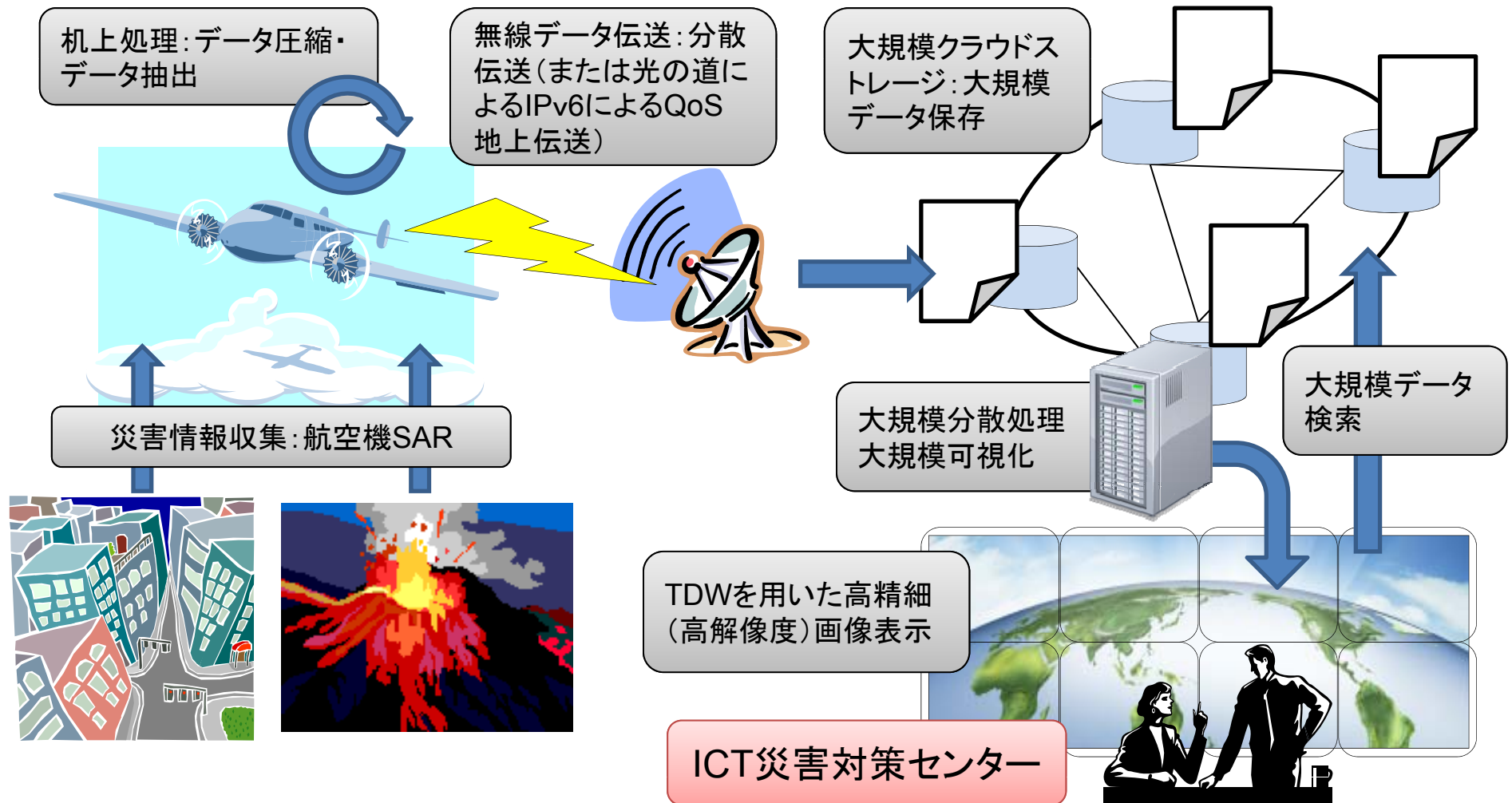
コンテナの後ろに接続されているのが空調モジュールと室外機。ここから外気を吸排気する。

クラウド型スパコン設備概要

- マルチノード構成
 - 16~32ノード構成
 - 用途によってノードのクラスタ化を変更
 - 所内の数値計算で共有利用
- クラウド型クラスタマシン
 - NICT内部(外部)研究者が自らハイスペッククラスタを購入・管理する必要がない
- 予定スペック
 - CPU: 256 (2048コア) ⇒ 20TFlops (理論値)
 - GPGPU: 400 ⇒ 200TFlops (理論値)
 - メモリ: 2TB
 - ネットワーク: 10G × ノード数 (最大640G)
- 特長(一般のスパコンとの差別化)
 - 理論性能の90%がGPGPU
 - 今後主流になるGPU計算技術開発
 - ノードごとの10Gネットワーク
 - 100G通信実験(AKARIプロジェクト)
 - 自動バックアップなど冗長性が高い24時間型シミュレーションシステムの研究開発
 - 気象庁などへのスピンアウト
 - ストレージレス
 - クラウドストレージの活用によるコストダウンと大規模データ処理
 - クラウド分散ストレージ・分散処理とのすみ分け
 - 大規模並列計算はクラウド分散ストレージシステム
 - 高速分散計算はクラウドスパコン
 - メモリのRAMディスク化
 - メモリのRAMディスク化によるデータファイルI/Oの高速化(他のスパコンシステムではできない独自性)

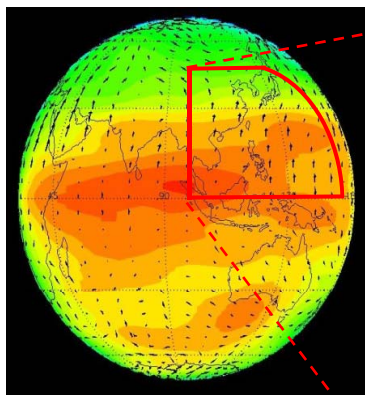


環境インフォマティクス 航空機SARによる被災地情報収集・対策システム



H.22年度シミュレーションPJ研究計画: 高度シミュレーション技法の検討

シミュレーション 大規模化



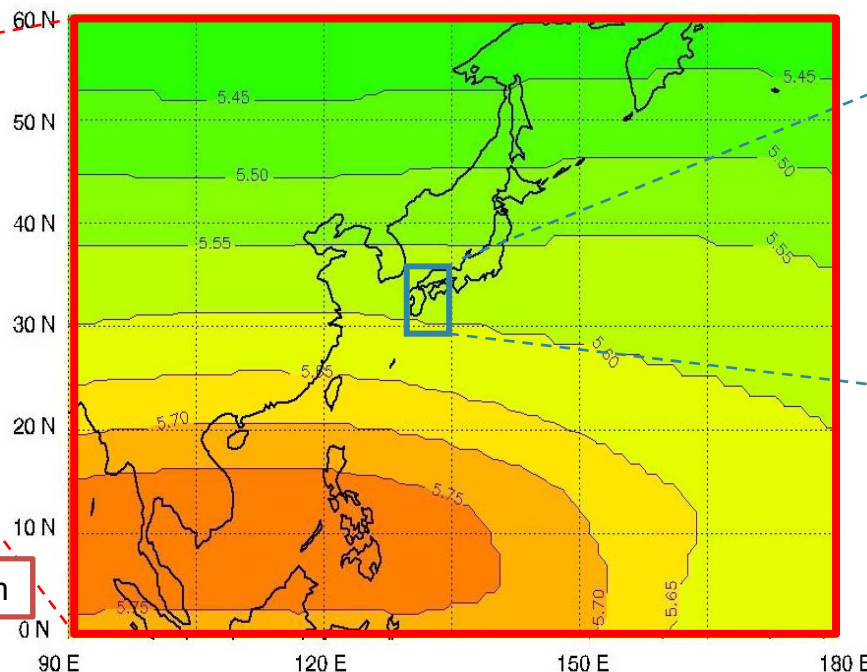
緯度1度×経度5度×高度10km

5000倍

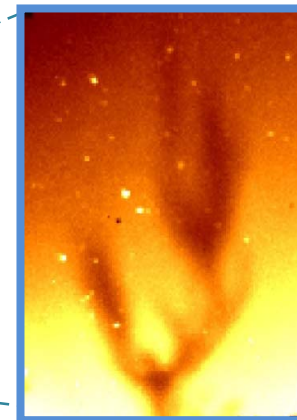
緯度0.2度×経度0.2度×高度5km

コード超並列化(GPGPU導入)
下層大気の効果(気象モデル)
大規模可視化システム開発

シミュレーション実用化の技術開発



光学観測



九州上空で光学観測されたプラズマバブル:GPS衛星測位誤差に最も影響が大きい電離圏電子密度の揺らぎ

シミュレーション評価モデル

シミュレーションによる定量的予測は困難

↓
擾乱度や推定値をシミュレーションデータから算出するスキームの開発

↓
4次元データ同化システムの検討

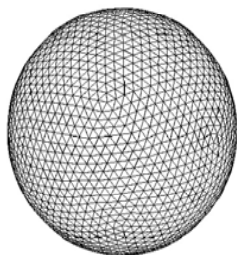
シミュレーションニーズ調査

太陽フレアモデル

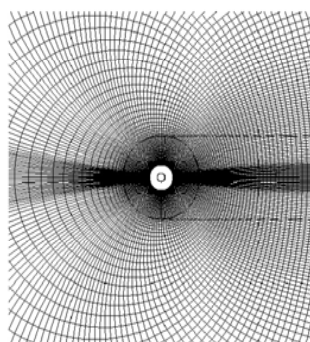
SEP(太陽高エネルギー粒子)は宇宙飛行体に最も影響が大きい→何をどこまで予測するかの調査

衛星帯電モデル

地球磁気圏モデルのリアルパラメータ化→SEPとの組み合わせで定量的な衛星帯電モデルの実現



非構造格子による大規模並列化



特殊格子モデル

サイエンスクラウドによる新しい研究手法の例

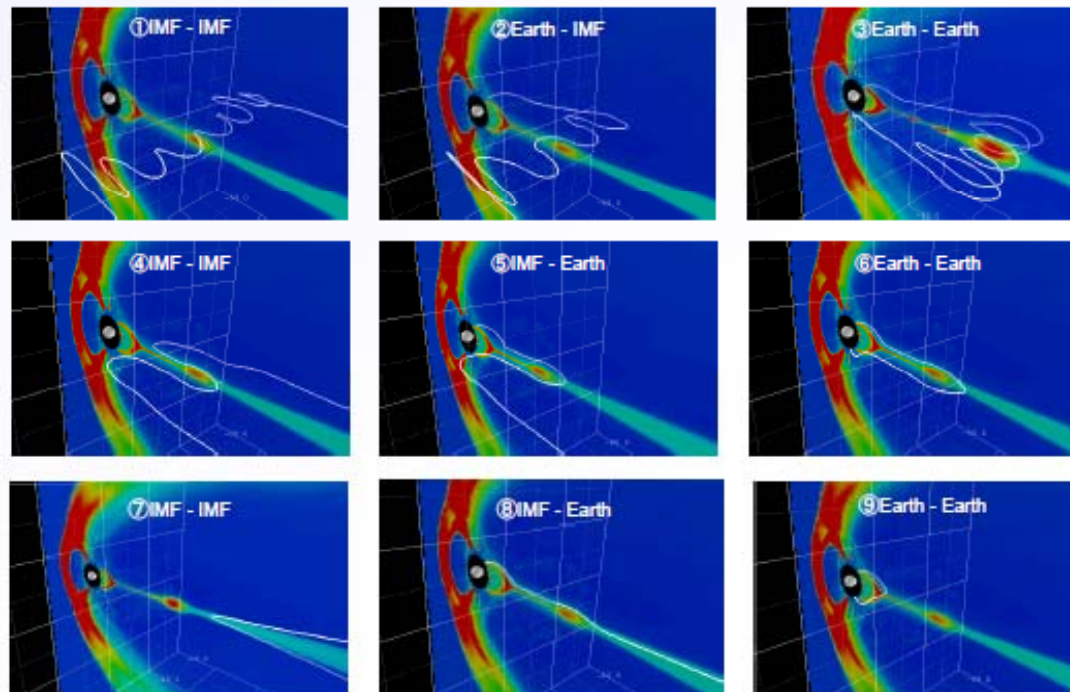
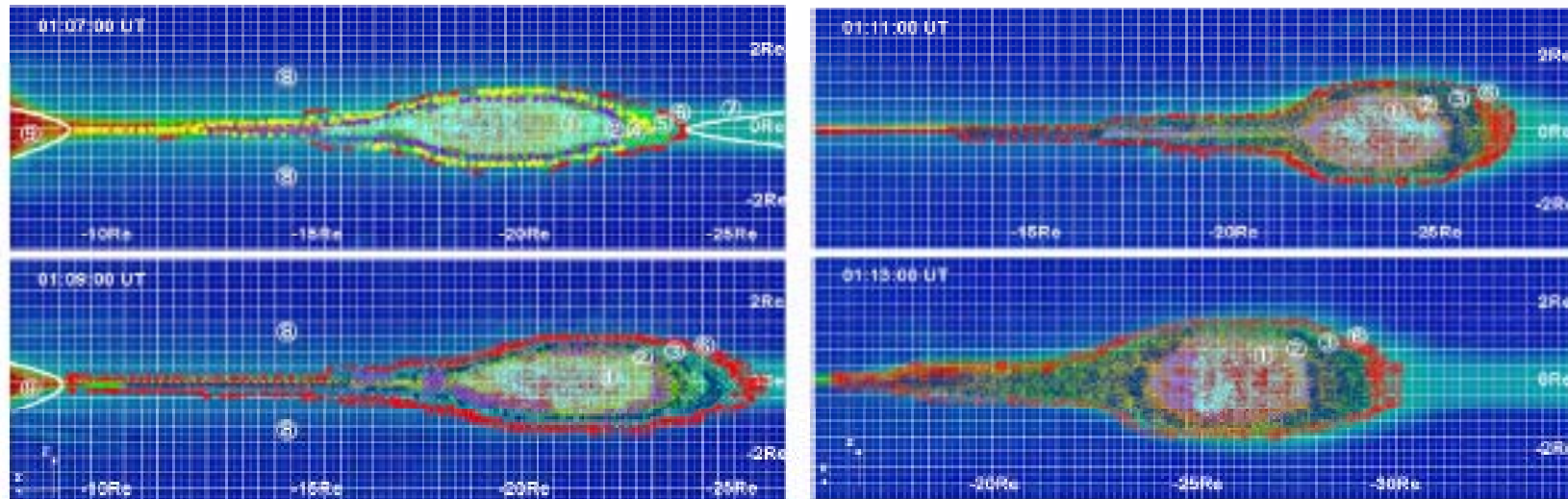
“Milk crown” visualization

- 超高時間分解能での3次元可視化環境構築
 - 3D時系列可視化
 - インタラクティブ解析
- 手法
 - シミュレーションの全時間ステップデータをディスクに保存
 - 本研究では90分=90000ステップ=7.2TB
 - データ保存のためにRAMディスクを利用⇒約10倍の高速化
 - 出力データはGfarm(分散ストレージ)で管理
 - バーチャルオーロラシステムにより可視化・解析
- 目的
 - 差分方程式にさかのぼり、シミュレーション空間内での物理素過程を解析



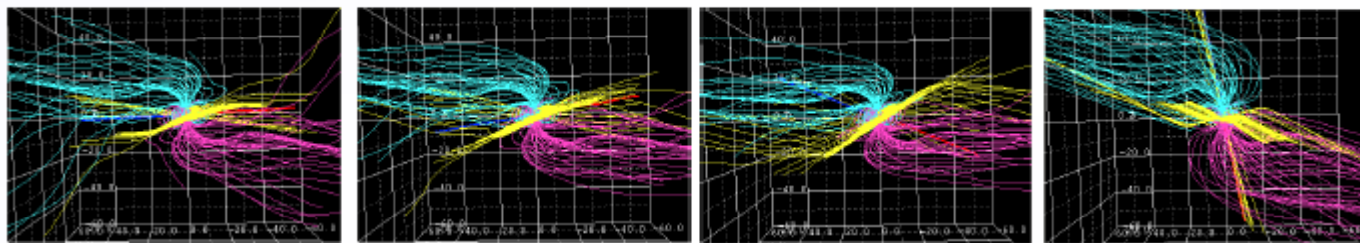
Milk crown可視化の例
Global MHDシミュレーションで
の磁力線追尾

①IMF-IMF、②IMF-Earth、③Earth-Earth (helical)
④IMF-IMF、⑤IMF-Earth、⑥Earth-Earth (not helical)



Motivation

Global MHDシミュレーションによる尾部プラズモイドの磁場構造分類
⇒プラズモイドのコアにIMFが“侵入”
⇒そのプロセスは？



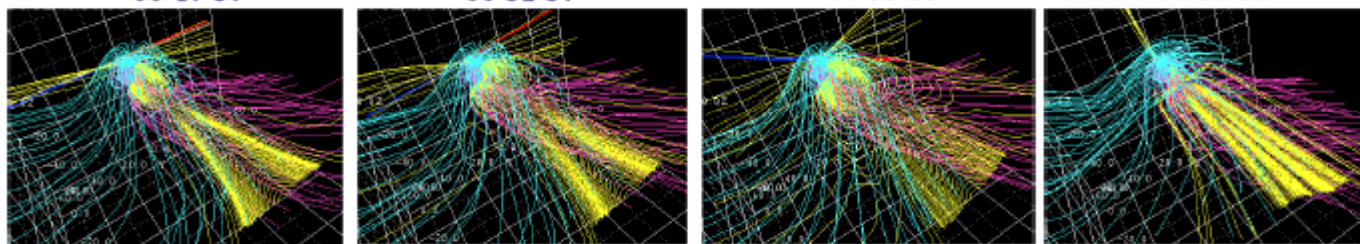
00:47:57

00:51:57

00:55:57

00:59:57

“前面”から

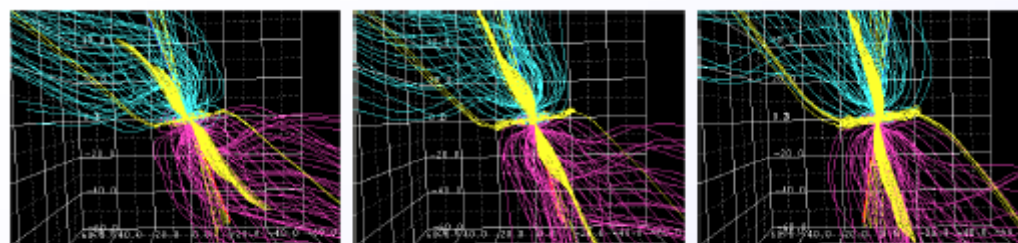


01:03:57

01:07:57

01:11:57

“後ろ斜め上”
から



01:03:57

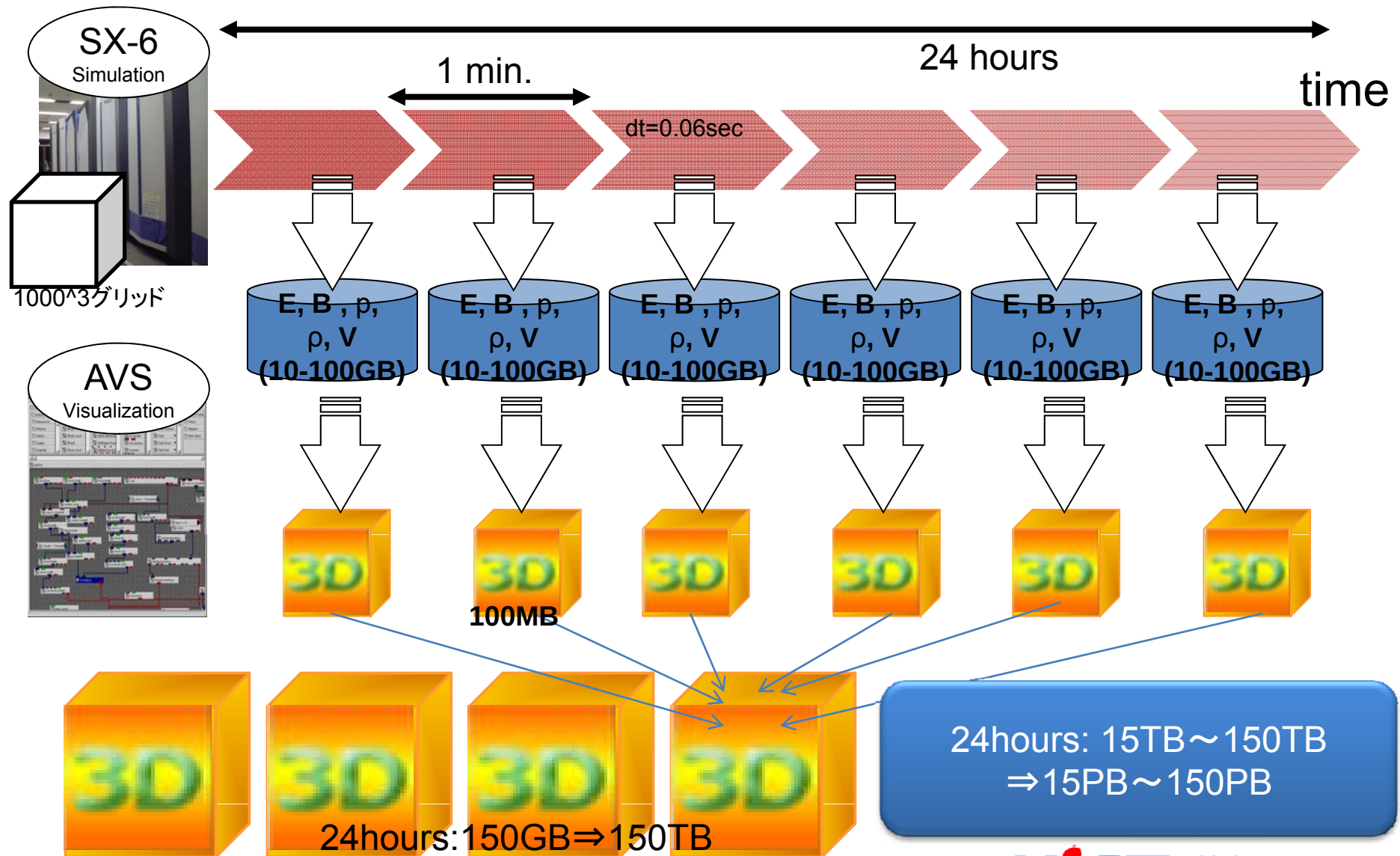
01:07:57

01:11:57

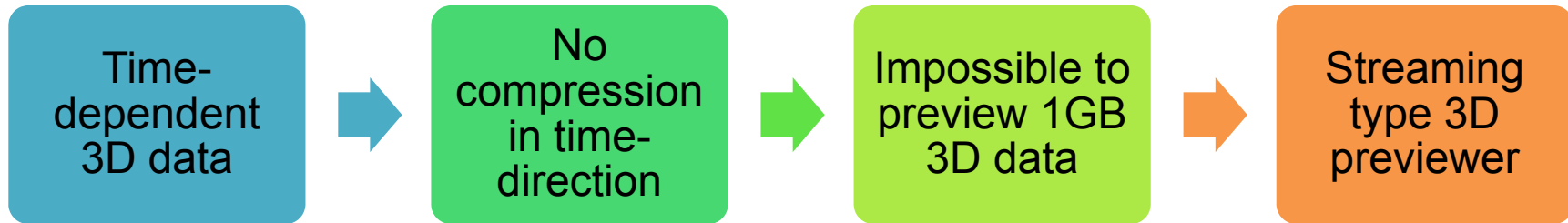
Global MHDシミュレ
ーションによる尾部
磁場のIMF応答

松岡他、2004

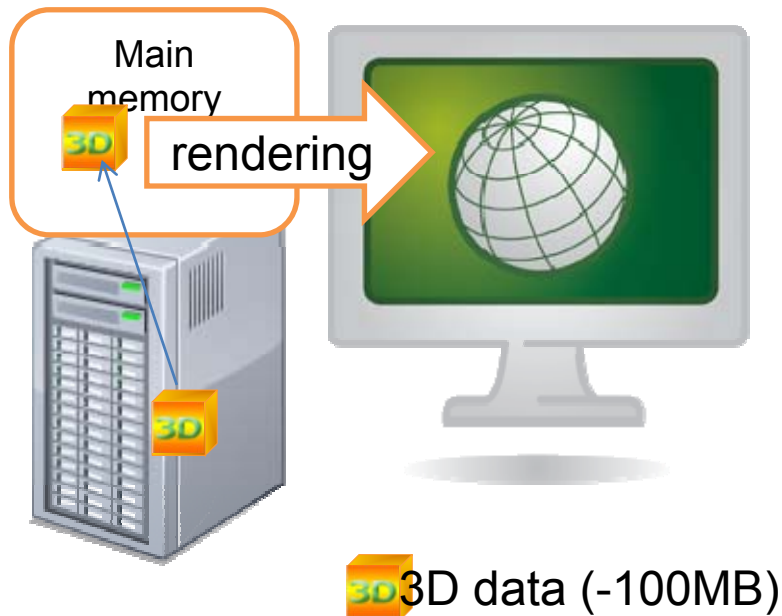
間抜けな国策？ 「なぜNo.1でないといけないんですか？」



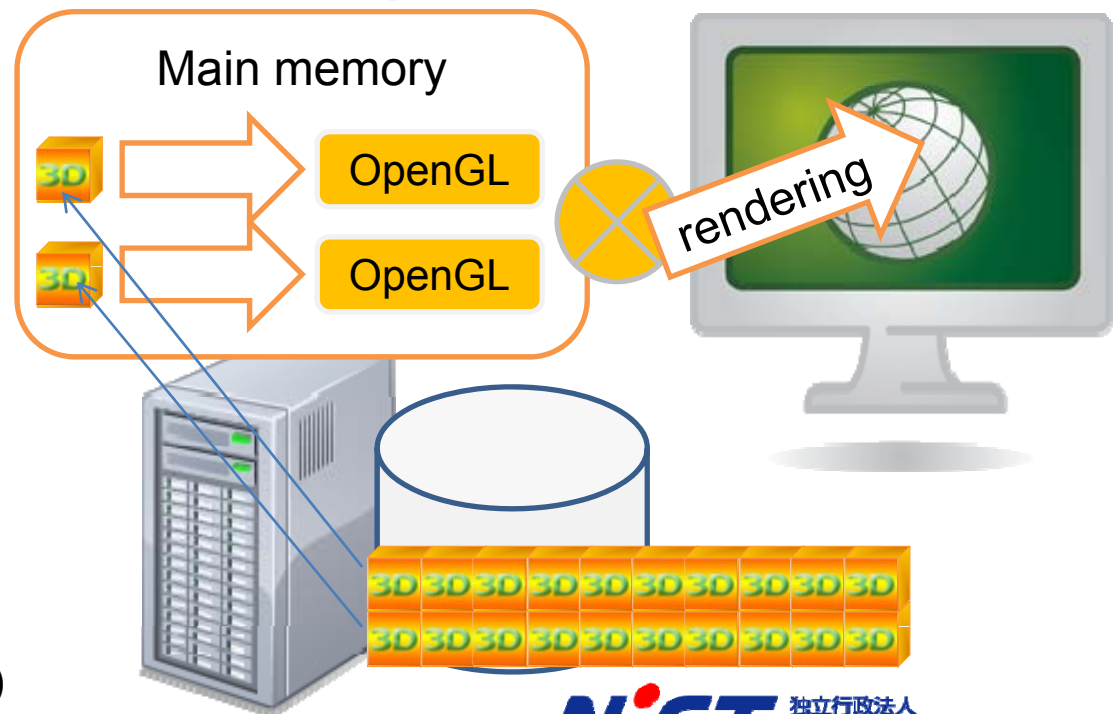
Large scale 3D data previewer



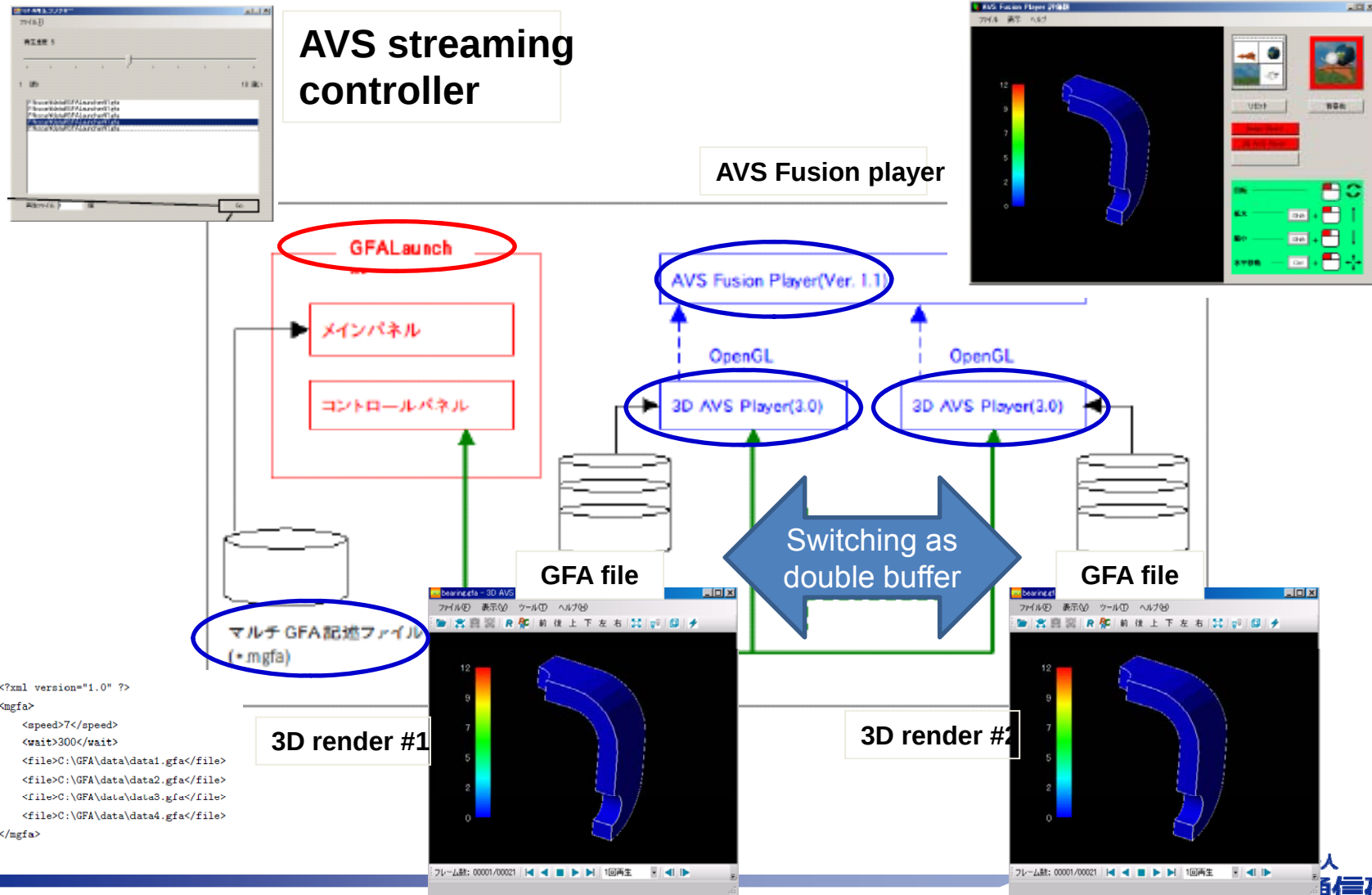
3D previewer (so far)



Our new previewer

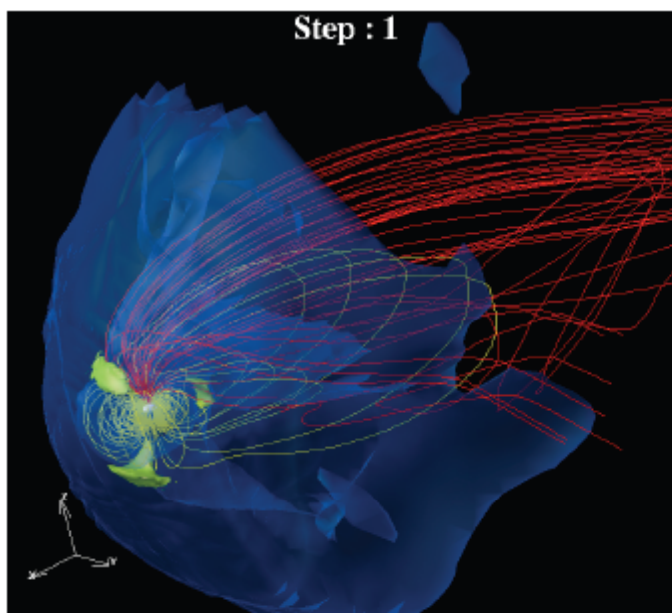


The longest 3D data preview in the world!



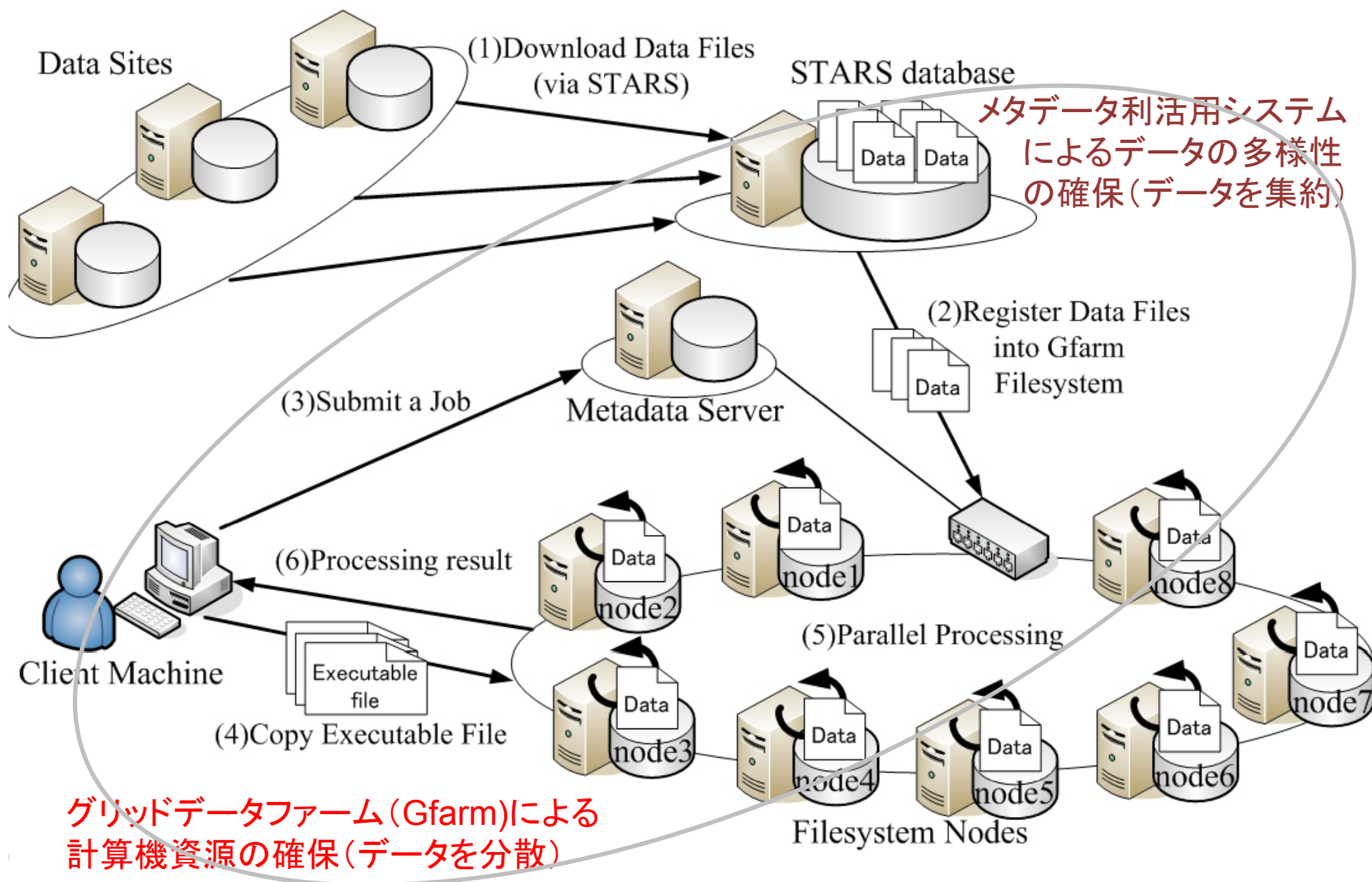
“Milk crown” viz. for SW simulation

- Simulation code
 - Global MHD simulation code (Tanaka & Fujita)
- Date & Time
 - 2003年11月12日23:30UT~(90分間=9万ファイル=7.2TB)
 - 可視化解像度 $dt=0.06$ (シミュレーションの $dt=0.02$)



- 空間グリッド: 88(経度)×120(緯度)×120(高さ)
 - 高さ方向のメッシュサイズは地球近傍(15Re以内)で0.25Re
- 時間ステップ: 0.02秒間隔...データは0.06秒間隔で保存(1000ファイル/分)
- 入力太陽風データ: 2003年11月12日23:30:00 – 13日01:00:00 UTのACEデータ(90分間=ファイル数9万)
- NICT SX-8Rで並列計算(node内8並列)
- 1ステップ辺りのデータ容量、約80MB
- 全データ容量、約7TB→gfarmストレージシステムに保存
- 等圧面や磁力線などの3D描画: virtual auroraシステム

パイプライン型データインテンシブ処理モデル



超高時間分解能可視化

- 可視化1
 - 確認:MHDシミュレーションで磁力線は本当にfrozen-inしているのか？(磁力線を追跡できるのか？)
- 可視化2
 - 長時間で可視化⇒どこでリコネクションが発生するか？(1分間可視化: $dt=60$ 秒)
- 可視化3
 - 短時間で可視化⇒リコネクションが起こる瞬間をとらえる！(4.2秒間可視化: $dt=0.06$ 秒)

超高時間分解能可視化

- 可視化1

- 確認:MHDシミュレーションで磁力線は本当にfrozen-inしているのか？(磁力線を追跡できるのか？)

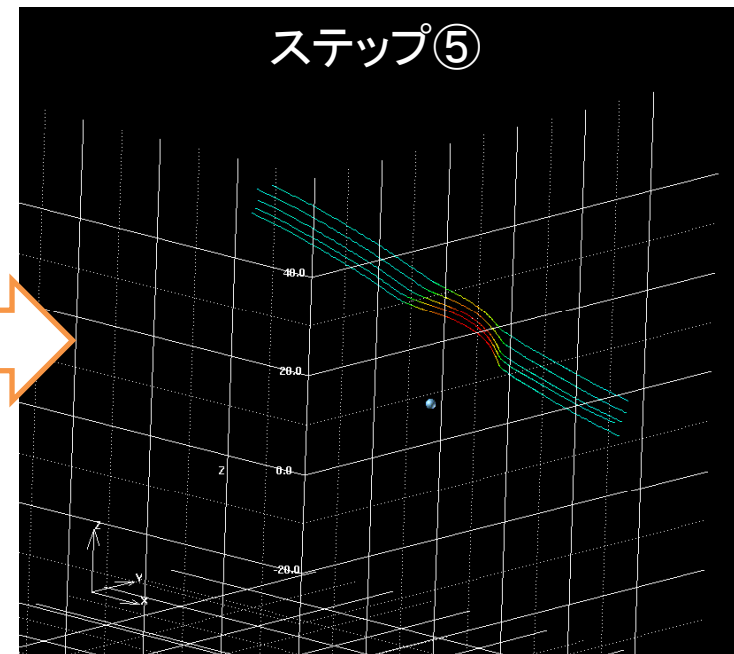
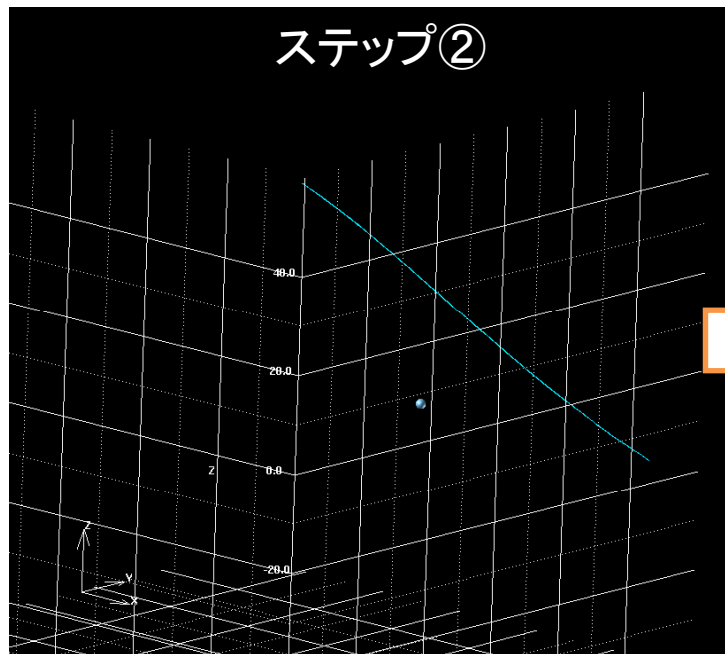
- 可視化2

- 長時間で可視化⇒どこでリコネクションが発生するか？(1分間可視化: $dt=60$ 秒)

- 可視化3

- 短時間で可視化⇒リコネクションが起こる瞬間をとらえる！(4.2秒間可視化: $dt=0.06$ 秒)

(可視化1)確認:MHDシミュレーションで磁力線は本当にfrozen-inしているのか？



①空間内の1点を選択⇒②その点を通る磁力線を3次元描画⇒③磁力線上の複数点を選択⇒④それらの点を時間方向に0.06秒進める(frozen-inを仮定)⇒⑤それらの点を通る磁力線を描画⇒④と⑤を繰り返す

磁力線はfrozen-inしていないように見える

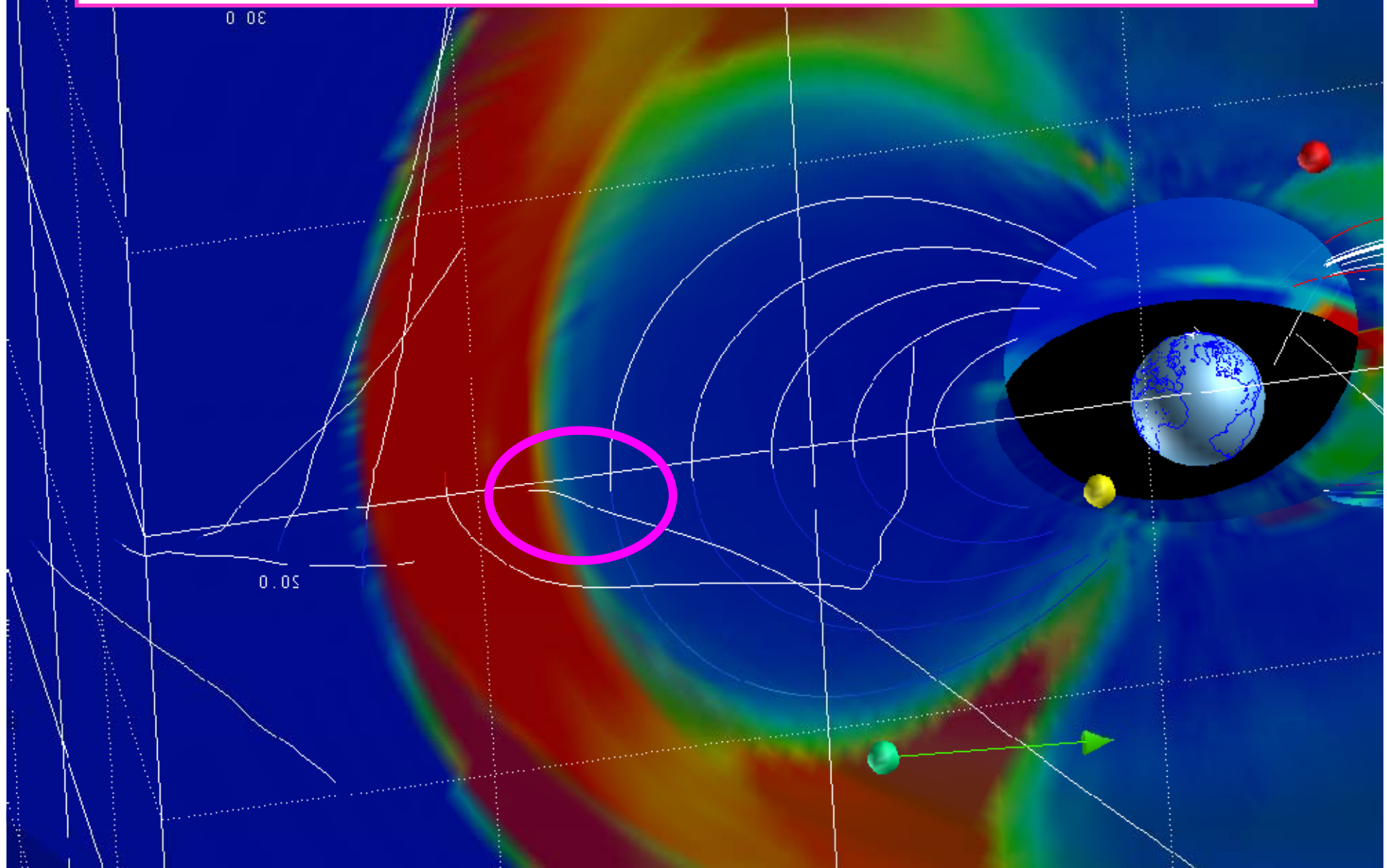
“磁束”可視化により誤差を吸収？



(避けえない)“数値誤差”が原因か？



Magnetic Reconnection?



超高時間分解能可視化

- 可視化1

- 確認:MHDシミュレーションで磁力線は本当にfrozen-inしているのか？(磁力線を追跡できるのか？)

- 可視化2

- 長時間可視化⇒どの時刻で(どの場所で)リコネクションが発生するか？(1分間可視化: $dt=60$ 秒)

- 可視化3

- 短時間可視化⇒リコネクションが起こる瞬間をとらえる！(4.2秒間可視化: $dt=0.06$ 秒)

超高時間分解能可視化

- 可視化1

- 確認:MHDシミュレーションで磁力線は本当にfrozen-inしているのか？(磁力線を追跡できるのか？)

- 可視化2

- 長時間で可視化⇒どこでリコネクションが発生するか？(1分間可視化: $dt=60$ 秒)

- 可視化3

- 短時間で可視化⇒リコネクションが起こる瞬間をとらえる！(4.2秒間可視化: $dt=0.06$ 秒)

Capture the
reconnection
moment !

Ultra-high time resolution visualization

by Tsubouchi, Tanaka, Fujita, Morikawa, Inoue, Yamaoto and
Shinagawa (NICT)

Data Sampling time

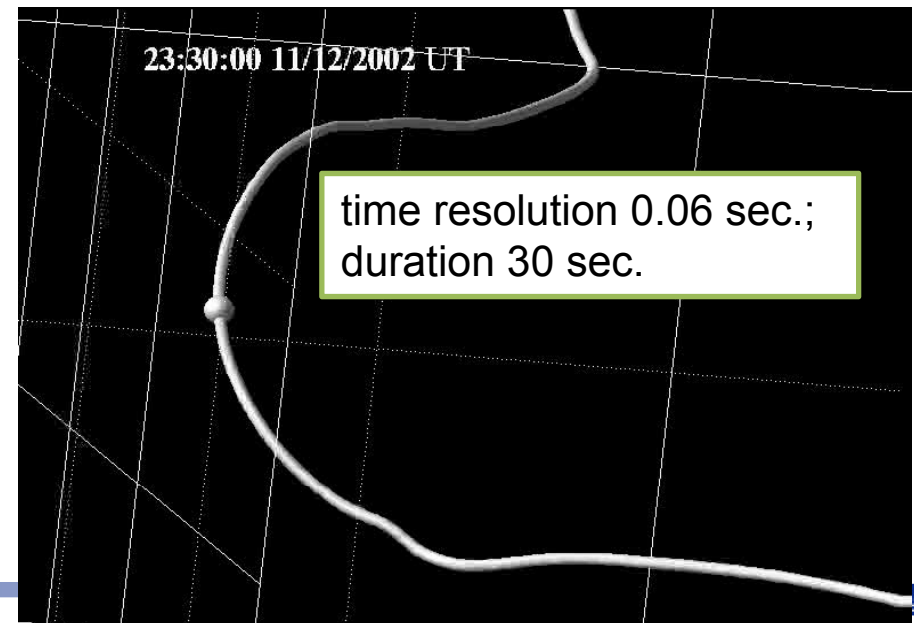
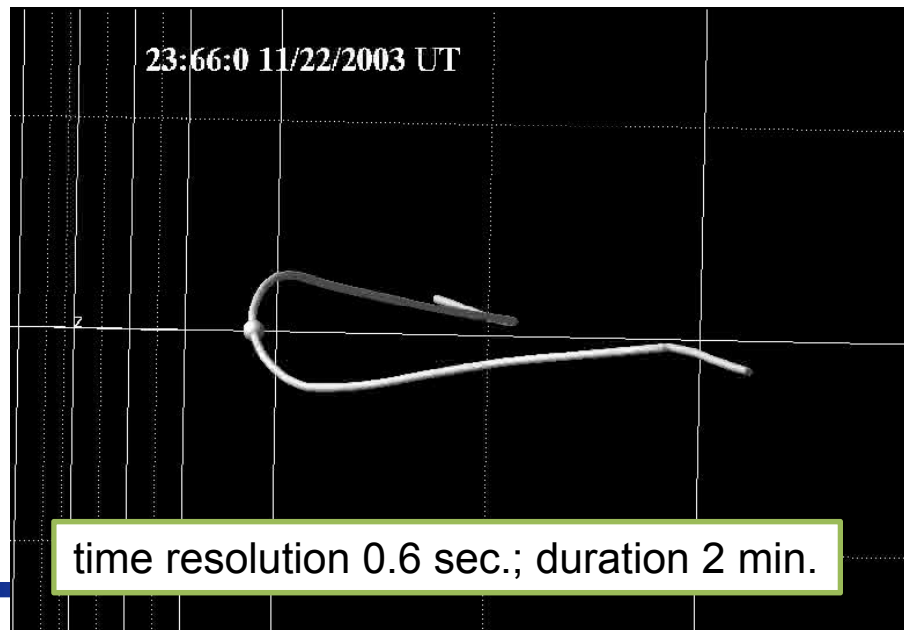
Method

A magnetic field line at an arbitrary time and position (a sphere in the panel) are chased on assumption of frozen-in theory. (Original Δt in MHD simulation is 0.02 sec.)

- time resolution: 60 sec. (not shown herein)
- frozen-in broken

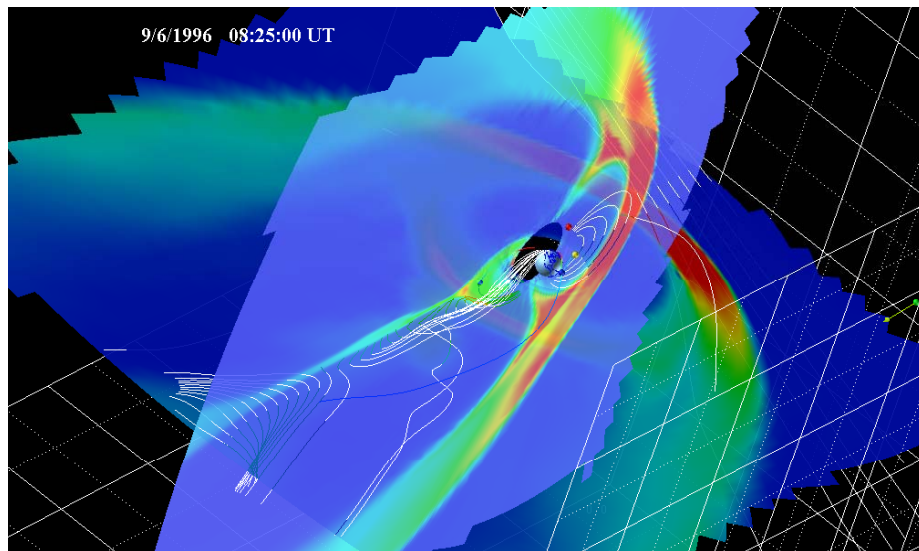
- time resolution: 0.6 sec. (7,000 files)
- frozen-in for one reconnection event

- time resolution: 0.06 sec. (70,000 files)
- reconnection captured!

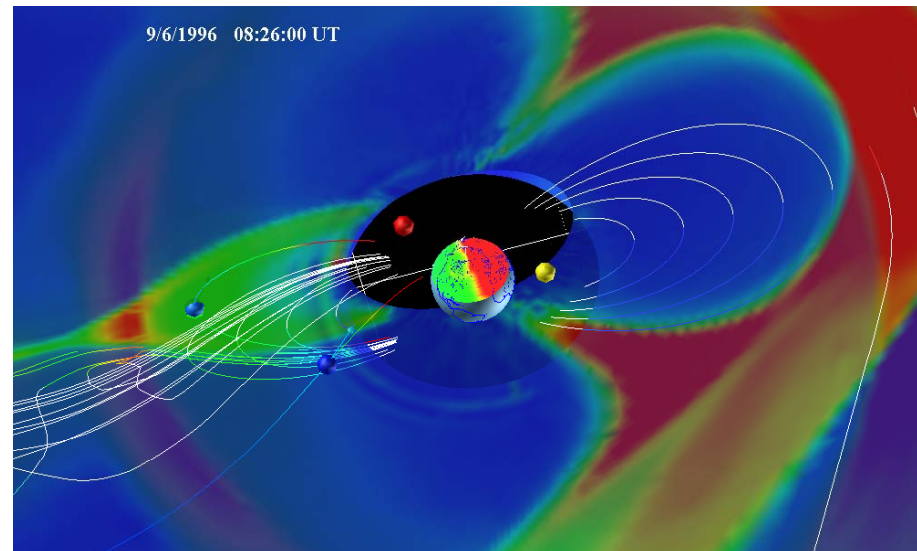


その他のサイエンスクラウドの可視化例①

マルチスケール可視化



Global view (Magnetosphere)



Earth view (aurora)

- aurora...from observation data (POLAR/VIS)
- Magnetosphere...from Global MHD simulation

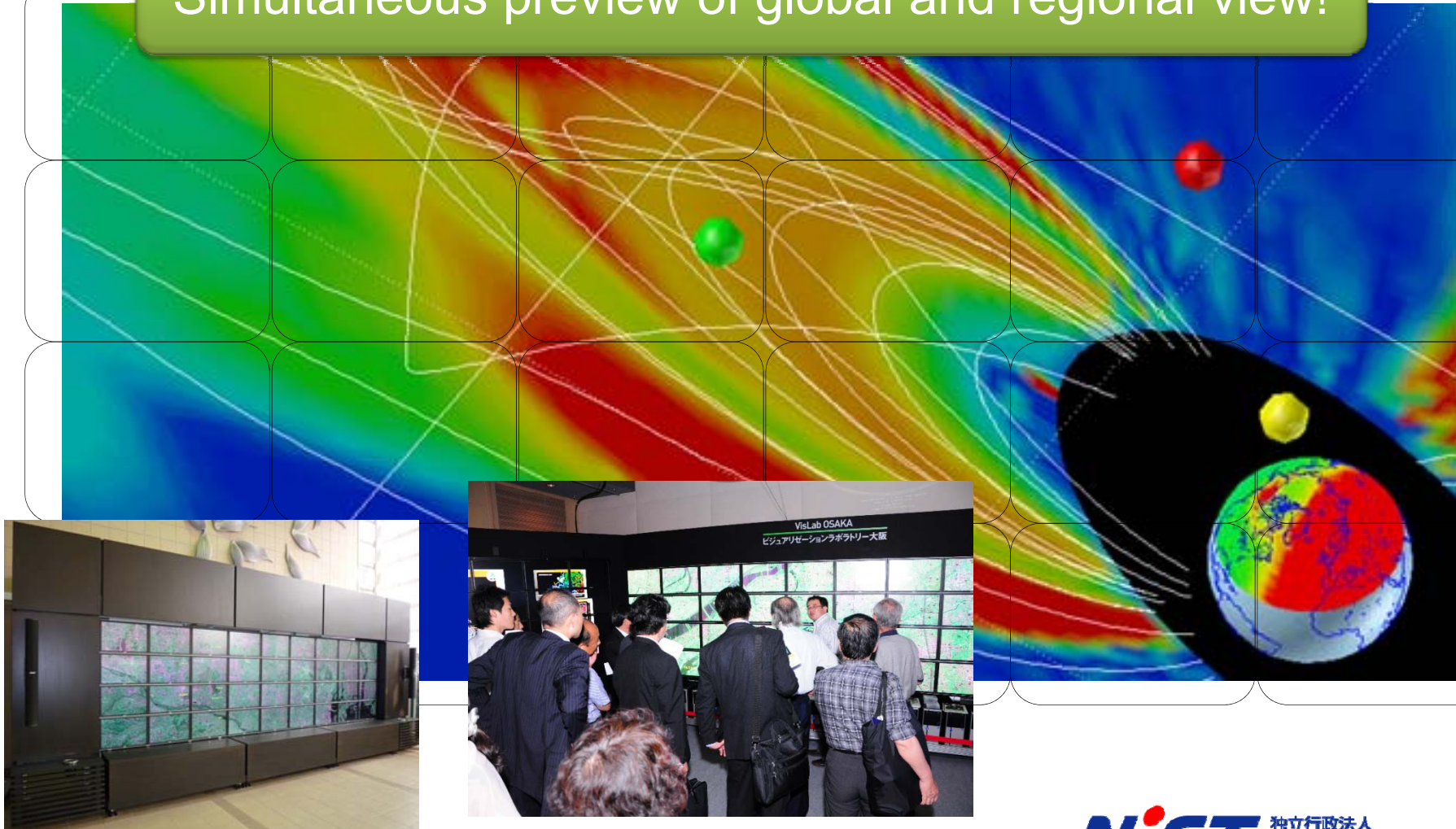
A Tool to investigate local/global and observational/simulational data

How can we take a look of both scales?

High-resolution Display Wall: TDW

マクロ⇄ミクロ

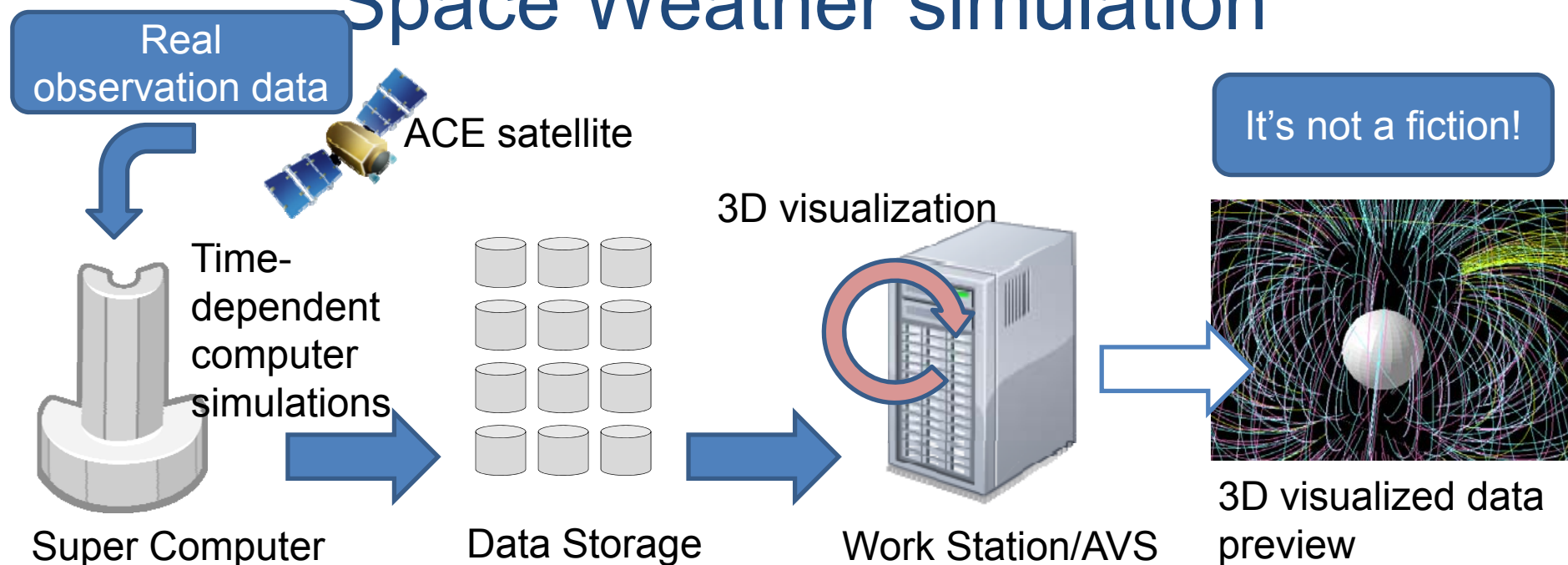
Simultaneous preview of global and regional view!



NICT TDW: 18000x4300 pixels

Collaborative Data analysis

Space Weather simulation



For easier understandings of space weather

More
beautiful!

More
reality!

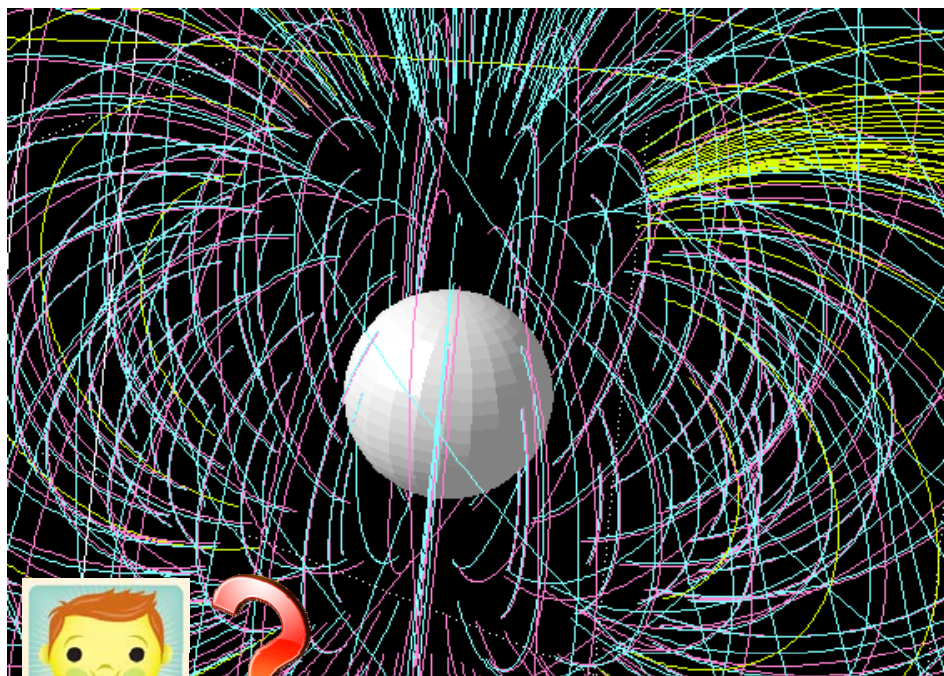
More
Story!



Which do you like?

3D visualization without story

- 4 independent computer simulations (SUN, Mercury, Venus and the Earth) in the solar-system

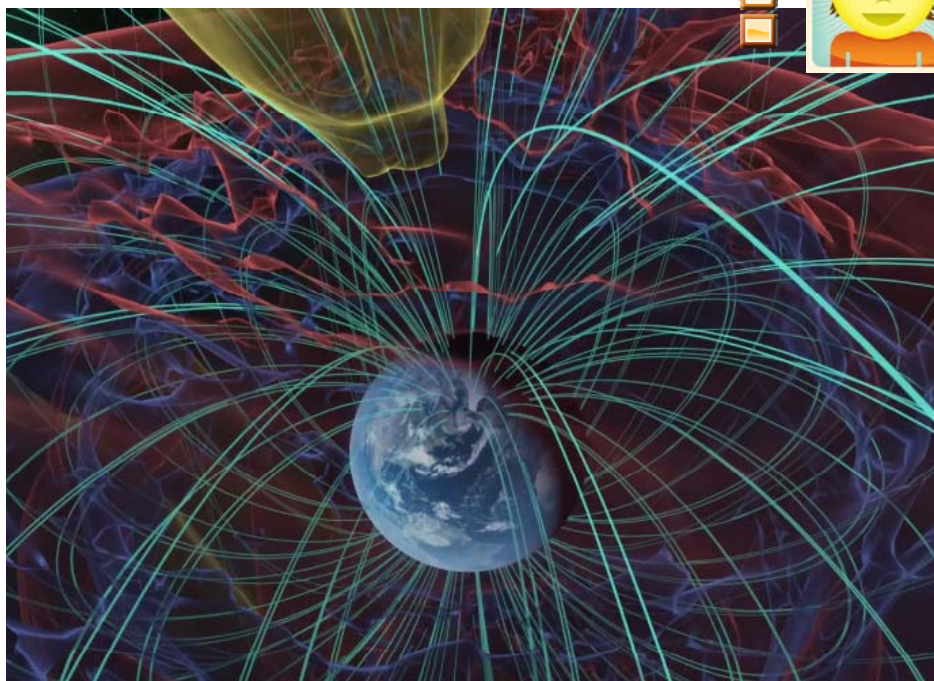
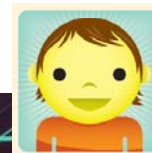


For research works

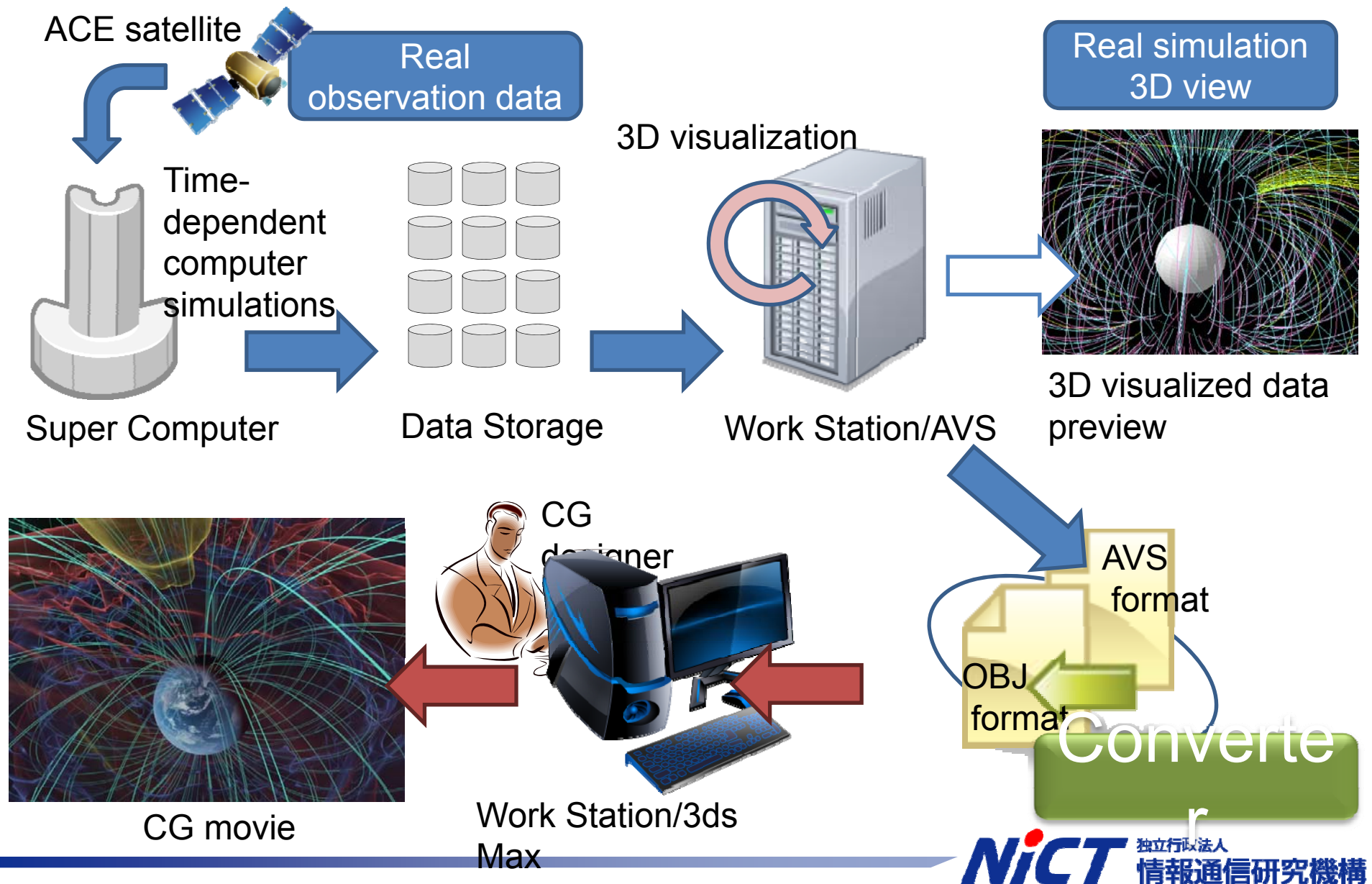
CG movie with stories

- CG movie for “Space weather” in the solar-system

For general use



A new technique to convert 3D viz to CG

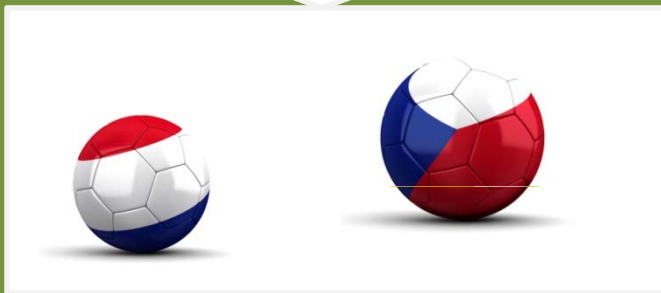


Data converter from AVS to CG

Labeling of objects
in different times



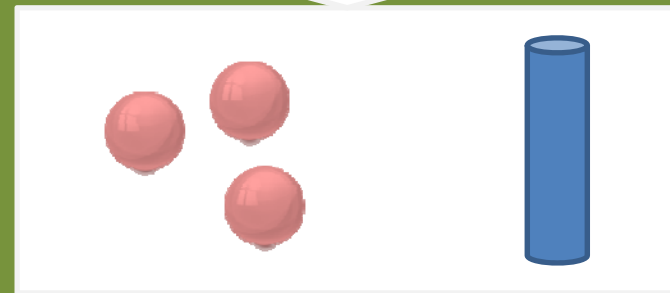
$t=t+dt$



Volume properties
to points and lies



Volume
property



未完成ですが...



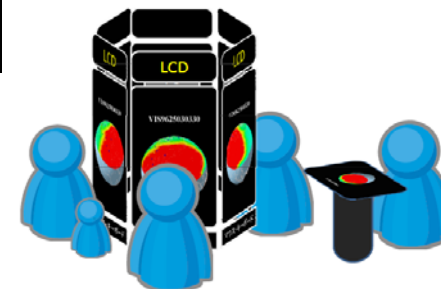
Vis Room@NICT on SW Cloud

- Contents
 - Space Weather simulation CG
 - 3D simulation for SW
 - Solar Observations
 - Google Earth



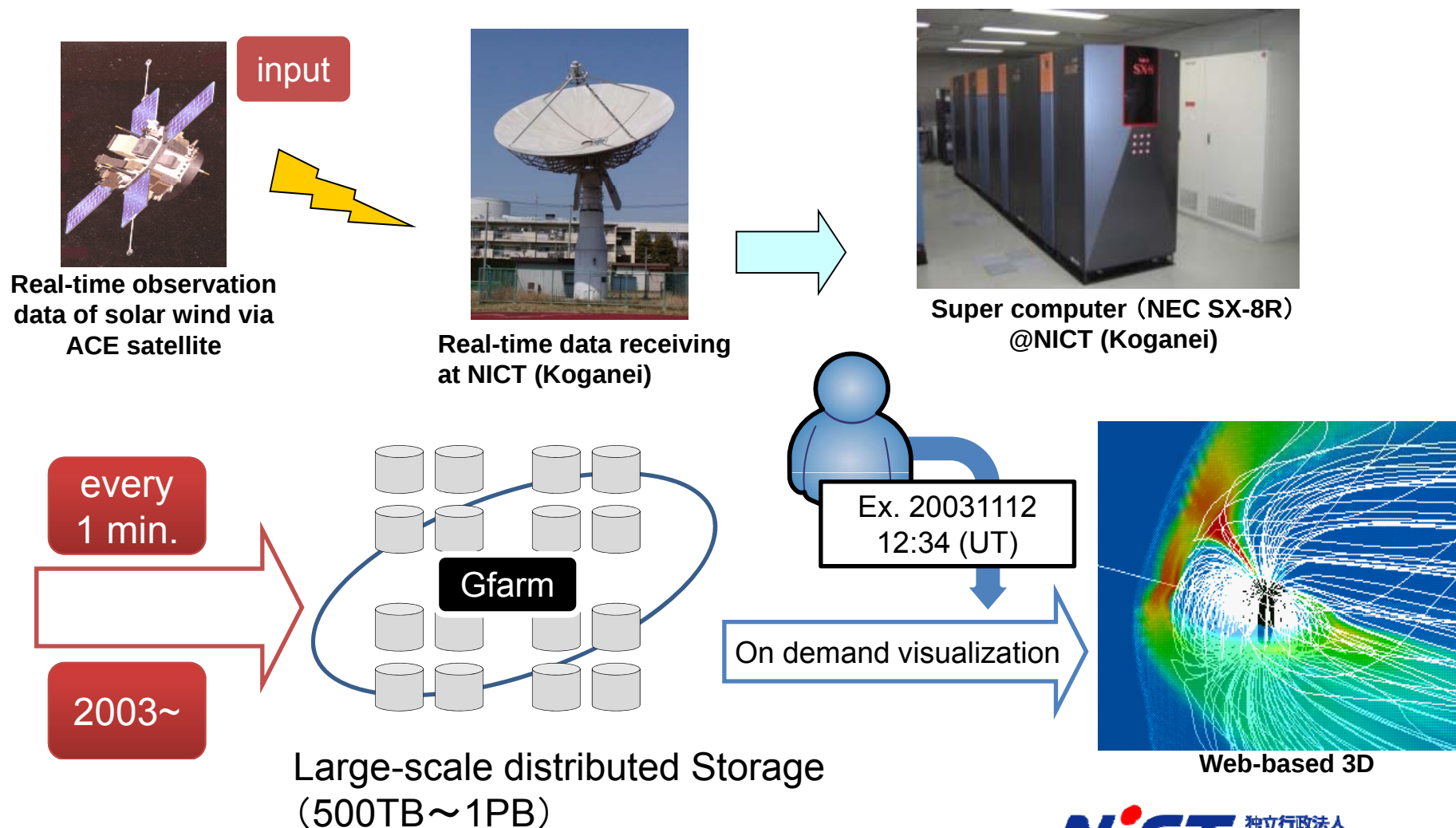
E-space@NICT (plan)

Tiled Display Wall@NICT



その他のサイエンスクラウドの可視化例②

オンデマンド3D可視化



NICTサイエンスクラウド ユーザ登録(2010年10月から開始)



<http://www.seg.nict.go.jp>

サイエンスクラウドユーザレベル

ゲストユーザ

- Webを利用する一般ユーザ(ユーザ登録なし)

Webユーザ

- Webを利用するユーザ(ユーザ登録あり)

Web研究利用者

- Webを利用する研究者(ユーザ登録あり)

まずはこのあたりから...

サイエンスクラウドリソース利用者

- 内部クラウドの基本リソース利用者
- (研究室でWS等の維持管理が困難な)WSやストレージの利用者を想定

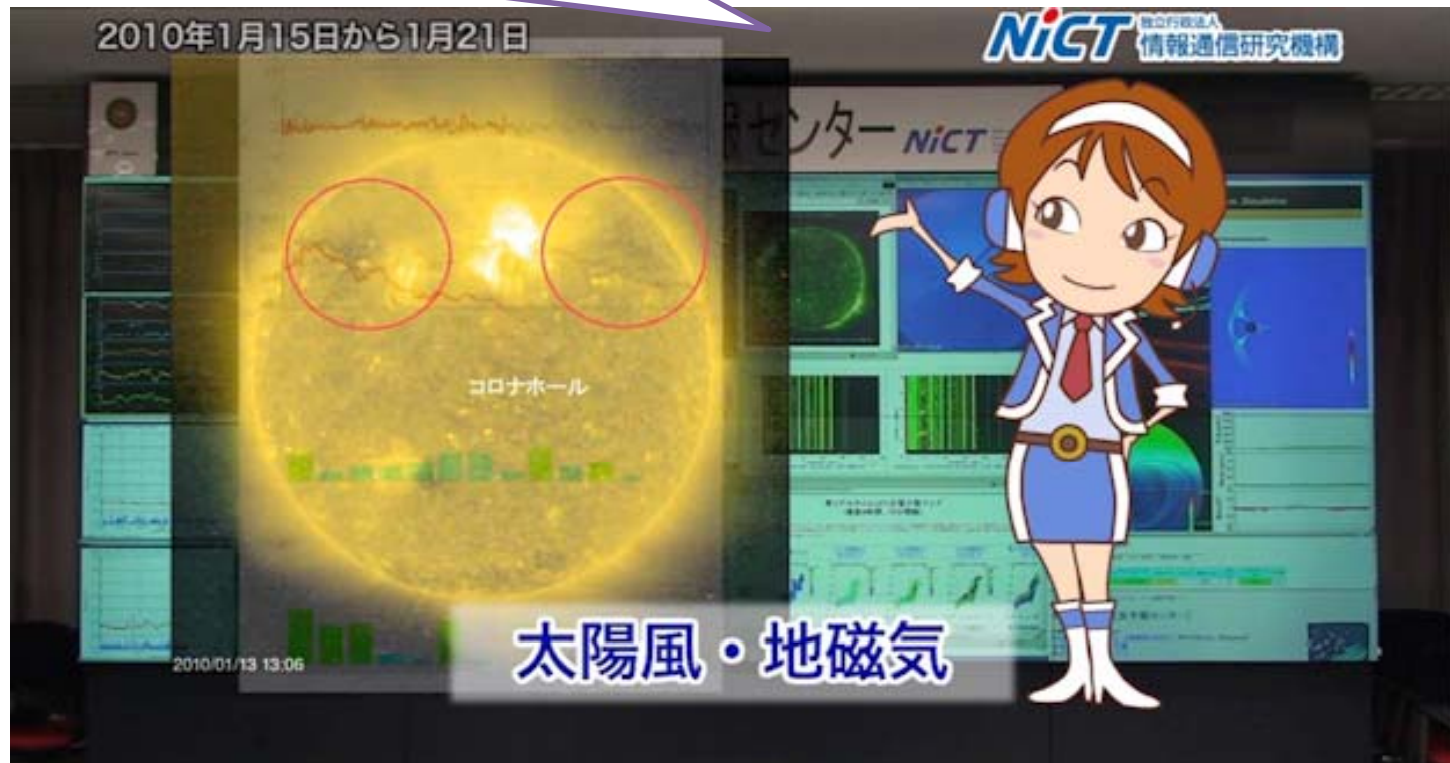
サイエンスクラウドミッション利用者

- すべてのサイエンスクラウドリソース(スパコン、ストレージ、解析サーバ...)の利用者
- NICTサイエンスクラウドを利用しなくてはできない研究(NICTとの共同研究)

まとめと今後の課題

- 2010年10月よりNICTサイエンスクラウドを公開
 - 次世代科学研究のためのクラウド環境
 - 利用者登録は5段階
- NICTサイエンスクラウドの応用例
 - 超高時間分解能可視化によるGlobal MHDシミュレーション磁力線追尾
 - “磁束”を追尾することで可視化誤差を緩和
 - リコネクションの“瞬間”を捉える事が出来た
 - 本技法を発展⇒IMFが尾部に輸送されるプロセスをすべて可視化する

Thank you for your attention!



Weekly Space Weather News
<http://www.seg.nict.go.jp/wsw>

本研究で利用したVirtual Auroraシステムは、統数研とNICTが開発したものです。