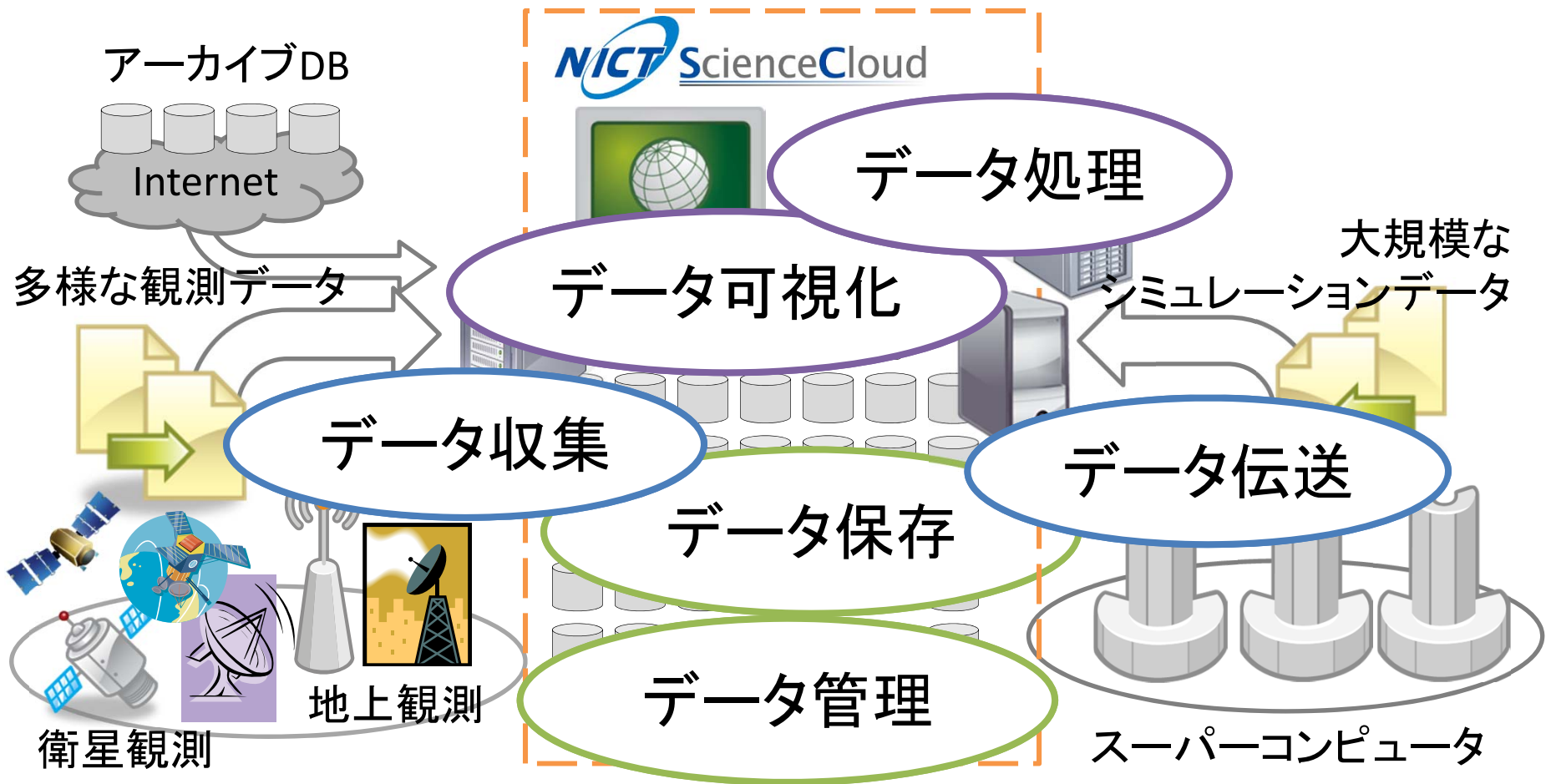


NICTサイエンスクラウドの 運用とデータ処理技術報告

村田 健史・渡邊 英伸・長屋 嘉明
情報通信研究機構

20150220 平成26年度名古屋大学太陽地球環境研究所研究集会

NICTサイエンスクラウド システムコンセプト



NICTサイエンスクラウドの基盤技術開発

基盤技術(i) データ収集・伝送

グローバルデータ収集

- 世界中の観測拠点を監視する広域観測網監視システム
- インターネット上の科学データを自動収集・処理
- 異分野データ・ソーシャルデータを融合表示

遠隔高速ストレージ

- 遠隔地からのクラウドストレージの高速I/O技術

基盤技術(ii) データ保存・管理

クラウドストレージ トレーサビリティ

- 広域分散ストレージのトレーサビリティシステム・タイムスタンプシステム開発と試験運用

セキュアWeb開発手順

- 研究者がWebアプリケーションをセキュアに開発する手順の確立
- H.25年度に6例のWebアプリで有効性を実証・運用

基盤技術(iii) データ処理・可視化

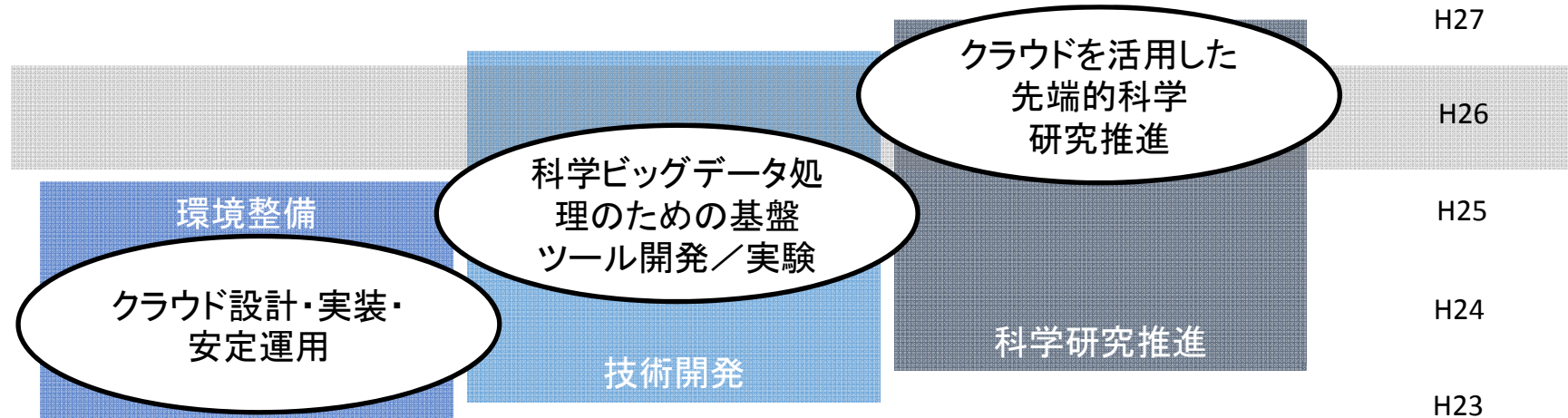
ビッグデータ並列処理

- 分散ストレージと分散処理の連携によるビッグデータ高速処理技術開発
- ビッグデータ処理システム開発開始・3次元レーダデータリアルタイム処理

ビッグデータ可視化

- 宇宙天気・気象レーダの3次元可視化システム

個々の基盤技術をマッシュアップして初めて一つのシステム(アプリ)として機能する



申請プロジェクト一覧(H.25年度)

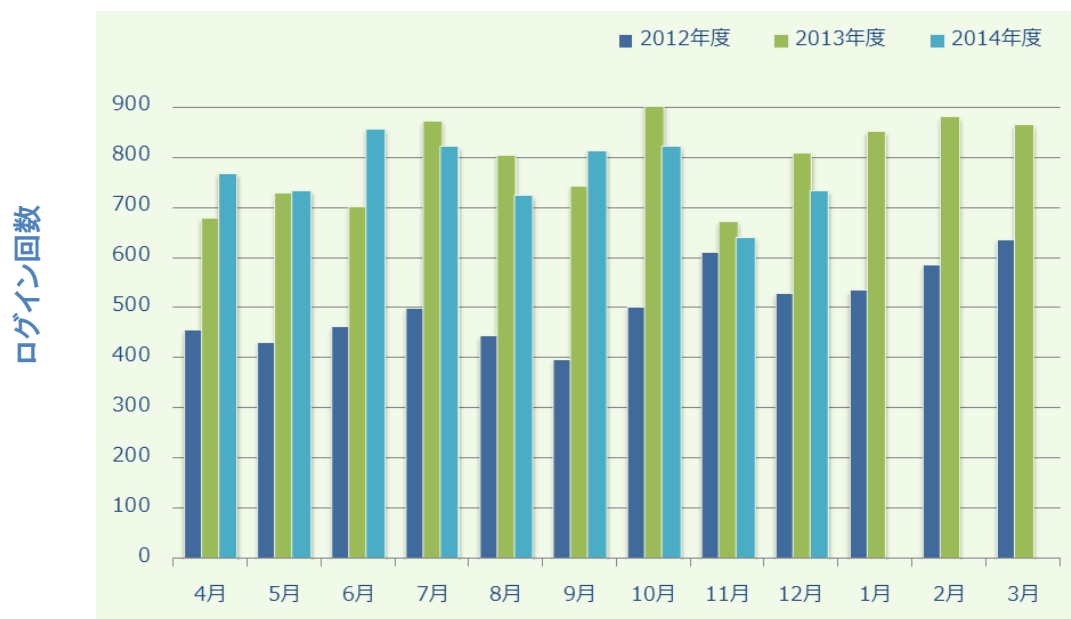
NICTプロジェクト比率 18/30

PJ 番号	プロジェクト名称	研究代表者	代表者所属	分担者数 (NCT外)
1	GNSS全電子数計測を利用した電離圏の研究	津川 卓也	NICT宇宙環境インフォマティクス研究室	20(19)
2	東南アジア域低緯度電離圏観測(SEALION)	津川 卓也	NICT宇宙環境インフォマティクス研究室	4(0)
3	短波到来方向探査装置を利用した電離圏の研究	津川 卓也	NICT宇宙環境インフォマティクス研究室	3(1)
4	宇宙天気シミュレーション	品川 裕之	NICT宇宙環境インフォマティクス研究室	16(10)
5	次世代宇宙天気情報処理の研究	亘 慎一	NICT宇宙環境インフォマティクス研究室	2(0)
6	太陽圏モデリング	亘 慎一	NICT宇宙環境インフォマティクス研究室	14(10)
7	大気圏・電離圏長期シミュレーションデータの解析	陣 英克	NICT宇宙環境インフォマティクス研究室	3(2)
8	太陽風-磁気圏電離圏システム相互作用	久保田 康文	NICT宇宙環境インフォマティクス研究室	2(2)
9	ジオスペース・放射線帯予測	長妻 努	NICT宇宙環境インフォマティクス研究室	5(1)
10	南極観測	長妻 努	NICT宇宙環境インフォマティクス研究室	11(5)
11	Integrated Satellite Observaion SIMulator for a Coherent Doppler Lidar (ISOSIM-L)による衛星搭載ドップラーライダーのフィジビリティスタディ	石井 昌憲	NICTセンシング基盤研究室	5(4)
12	SMILES/GOSAT (プロジェクト申請未提出)	笠井 康子	NICTセンシング基盤研究室	45(37)
13	フェーズドアレイ気象レーダのデータ利用システム(気象レーダの3次元視覚化)	佐藤 晋介	NICTセンシングシステム研究室	9(7)
14	生体電磁環境プロジェクト	渡辺 聡一	NICT電磁環境研究室	4(0)
15	太陽圏の巨視的構造とダイナミックスの研究	鷺見 治一	アラバマ大学	3(0)
16	科学衛星搭載プラズマ波動観測器で得られた波形データの特徴解析	笠原 禎也	金沢大学総合メディア基盤センター	2(2)
17	惑星間空間磁場北向き時の磁気圏電離圏対流機構の解明	渡辺 正和	九州大学理学研究院地球惑星科学部門	5(3)
18	GNSS可降水量データベース	藤田 実季子	独立行政法人海洋研究開発機構	1(1)
19	静止軌道衛星帯電プラズマ環境の解析および予測の研究	中村 雅夫	大阪府立大学	1(1)
20	地球磁気圏の形状と自由エネルギーに関する初期研究:大規模3次元電磁流体計算と観測を比較する方法の確立	齋藤 実穂	名古屋大学	1(1)
21	気象分野におけるビッグデータ利活用技術の研究	大野 智生	気象庁観測部気象衛星課	2(2)
22	バーチャルオーロラツールを活用したデジタル磁気嵐現象の研究	海老原 祐輔	京都大学生存圏研究所	2(1)
23	SS-MIX標準ストレージのNoSQL実装と並列分散処理の検証	木村 映善	愛媛大学医学部	1(0)
24	NICTサイエンスクラウドを用いたゲノムデータ管理基盤に関する研究開発	原田 憲治	(株)カイ研究開発部	3(0)
25	社会インフラのメンテナンスに資するシミュレーションとセンシングデータの解析	中畑 和之	愛媛大学大学院理工学研究科	1(1)
26	Global MHDシミュレーションの大規模可視化によるプラズマダイナミクス	深沢 圭一郎	九州大学	3(3)
27	SALMON (プロジェクト申請未提出)	村山 泰啓	NICT統合データシステム研究開発室	未
28	時系列データ表示アプリケーション(STARS touch)の開発※	村田 健史	NICT統合データシステム研究開発室	5(5)
29	NICTサイエンスクラウドセキュリティ技術開発※	渡邊 英伸	NICT統合データシステム研究開発室	5(5)
30	NICTサイエンスクラウド高速データ転送表示技術開発※	渡邊 英伸	NICT統合データシステム研究開発室	6(6)
合計		30		184 (129)

※は技術開発プロジェクト。NICT外部利用者が多いプロジェクトはバーチャルラボとしての利用であると予想される。

アクセス数(～H.25年度)

サイエンスクラウド利用者数 (H.24年度～H.26年度)



- ログイン回数は、サイエンスクラウドユーザが1ヶ月の間にゲートウェイサーバへログインした日数を示す。(1日に複数回ログインした場合は1回としてカウントする。)

サイエンスクラウドログイン数 TOP20 (H.25年4月～H.26年3月)

No.	PJ番号	所属	ログイン回数
1	12	NICTセンシング基盤研究室	332
2	4, 6, 15, 22	九州大学	321
3	6	Kyung Hee University (米国)	299
4	4, 5, 8, 9	NICT宇宙環境インフォマティクス研究室	281
5	4, 7	NICT宇宙環境インフォマティクス研究室	280
6	4, 7	九州大学	275
7	1, 2	NICT宇宙環境インフォマティクス研究室	261
8	2, 3, 5, 9, 10	NICT宇宙環境インフォマティクス研究室	238
9	4, 8, 20	気象大学校	211
10	10	NICT宇宙環境インフォマティクス研究室	204
11	4, 6	Kyung Hee University (米国)	163
12	12	NICTセンシング基盤研究室	154
13	2, 9, 10	NICT宇宙環境インフォマティクス研究室	135
14	26	NICT統合データシステム研究開発室	128
15	12	大阪府立大学	123
16	17	九州大学	114
17	4	NICT宇宙環境インフォマティクス研究室	112
18	12	東京工業大学大学院	107
19	4, 6, 15	NICT宇宙環境インフォマティクス研究室	98
20	6, 15	Alabama University (米国)	97

全ユーザログイン回数合計: 9385

成果リスト(学術論文・その他)

	学術論文(査読付き)	その他論文
H.24年度	25件	3件
H.25年度	56件	7件

サイエンスクラウド運用の現状:プロジェクト一覧(2012-)

#PJ	Category	Zone	Tool	Web
#1(21)	O	CEAH	N	●
#2(5)	O	C-A-	W	●
#3(4)	O	C-A-	W	●
#4(17)	S	C-AH	V	●
#5(3)	S	C-A-	P	●
#6(15)	O	C---	-	
#7(4)	S	C-A-	-	●
#8(3)	S	C-A-	V	
#9(6)	OS	CEA-	N	●
#10(12)	O	---H	W	
#11(6)	S	CEAH	-	
#12(46)	O	C---	T	
#13(10)	O	C---	WT	
#14(5)	H	CE--	-	
#15(4)	S	C---	-	

PJ	Category	Zone	Tool	Web
#16(3)	O	C---	T	
#17(6)	S	C---	V	
#18(2)	O	C---	-	
#19(2)	S	C---	-	
#20(2)	S	C---	-	
#21(3)	O	C---	-	
#22(3)	S	C---	-	
#23(2)	H	CE--	-	
#24(4)	H	C---	-	
#25(2)	S	C---	-	
#26(4)	S	CE--	VP	
#27(1)	O	-E-H	-	
#28(6)	I	CEA-	NWP	●
#29(6)	I	C---	-	●
#30(7)	I	CE--	P	

(191)

Category: Observation, Simulation, Human, Informatics
太字は代表者がNICT所属

Space Science (18)

Earth Science (5)

Human Science (3)

Informatics (3)

NICTサイエンスクラウド 科学研究利用形態

サイエンスクラウド上で独自プラットフォームと独自アプリを構築

サイエンスクラウドのプラットフォーム上に独自アプリを構築

サイエンスクラウドが提供するアプリケーションを活用



独自アプリ

独自アプリ

SaaS

独自プラットフォーム

PaaS

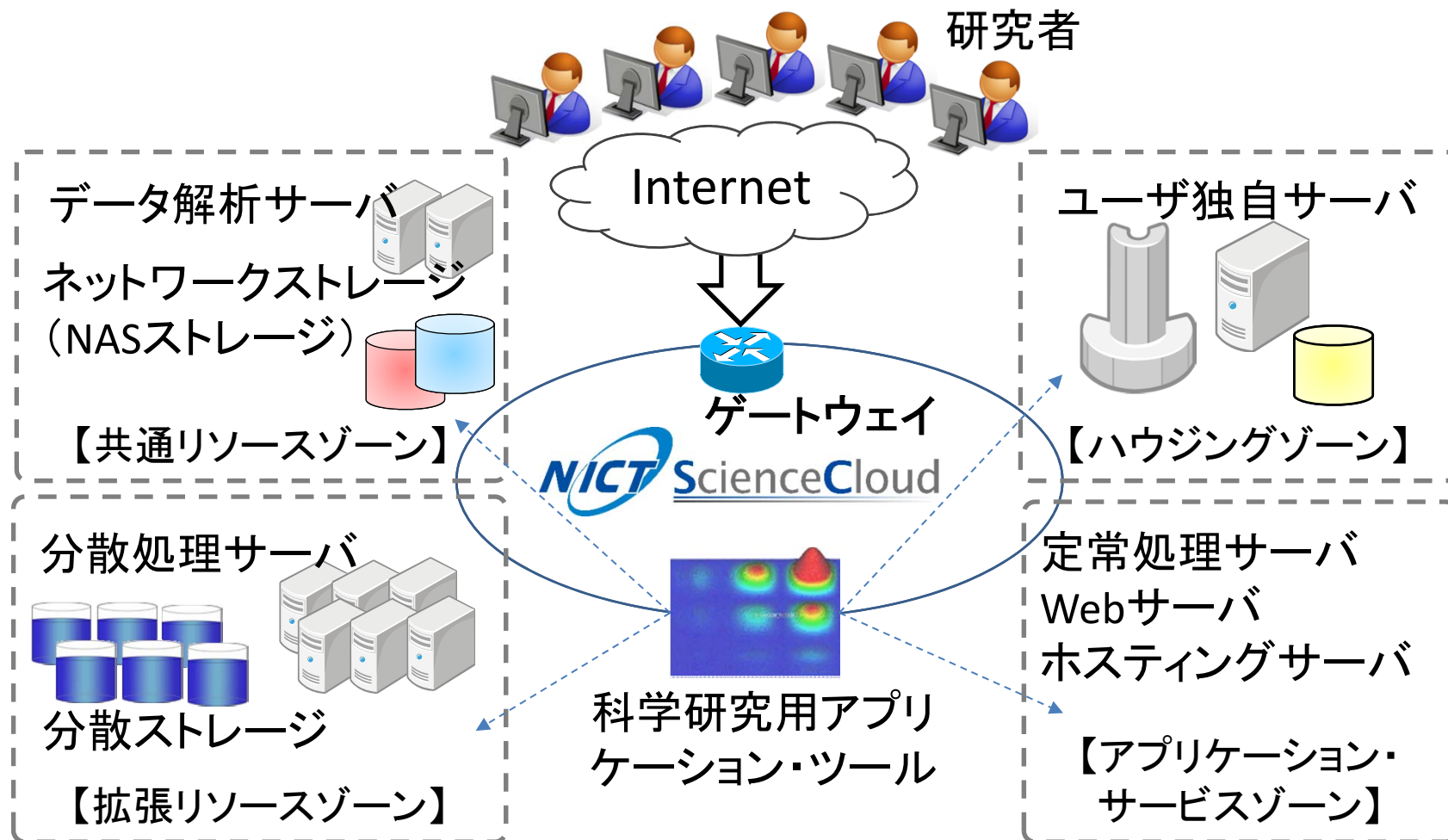
PaaS

IaaS

IaaS

IaaS

NICTサイエンスクラウド リソース・サービス



NICTサイエンスクラウドオリジナルアプリ・ツール

目的	アプリ名・ツール名	概要
データ収集	NICTY/DLA	インターネットで公開されている科学データをクロールするツール。メタ情報自動収集ツール(NICTY)とデータファイルダウンロードエージェント(DLA)から構成。
データ収集	WONM(Wide-area Network Monitoring)システム	広域観測網の観測所・観測拠点の観測システムを監視し、データ転送を自動的に行うツール。サーバツールとクライアントツールから構成されるが、クライアントツールをあらかじめセットアップした小型アプライアンスサーバを利用できる。
データ伝送	遠隔高速ストレージシステム(High-bandwidth Virtual Remote Storage System)	分散ファイルシステム(Gfarm)を仮想ストレージとして、遠隔地から高速データファイルの読み込み・書き出しを行うツール。クライアントサーバにセットアップすることで、APIとして利用できる。
データ管理	WSDBank(World Science Data Bank)	サイエンスクラウドのストレージ(NAS, 分散ファイルシステム)上のデータファイルにアクセスするためのWebアプリケーション。
データ管理	Gfarmトレーサビリティ	分散ファイルシステム(Gfarm)上のデータファイルの履歴をファイル単位(インスタンス単位)で追跡するツール。管理者用。
データ処理	Pwrake	複数の計算ノードでデータファイルを分散処理するための並列処理用タスクスケジューラ。NFSでもGfarmでも利用できるが、Gfarmと組み合わせローカルファイルに優先的にアクセスすることでI/O高速化を実現するアフィニティスケジューリングが可能となる。
データ処理	Torque/Maui	クラスタ計算環境で並列処理に適したタスクスケジューラ。リソース・マネージャ(Torque)とスケジューラ(Maui)から構成される。
データ可視化	バーチャルオーロラツール	Global MHDシミュレーションデータを可視化するツール。AVS Express/Devにより実装。
データ可視化	STICKER	フェーズドアレイ気象レーダデータとTwitter等のソーシャルデータを融合表示するWebアプリケーション。NICT情報利活用研で開発。
データ可視化	VDVGE	JAMSTECによるGoogle Earth用ボリュウムデータ表示ツール。
データ可視化	STARStouch	異分野字形例つデータ融合表示ツール(Webアプリ)。GEOTAIL衛星版(公開済み)、ひまわり衛星データ版(近日公開)、読売新聞版(SOMATO)、フェーズドアレイレーダ版などを開発。

サイエンスクラウド利用状況(H.25年度)

プロジェクト名称	研究代表者	研究分野	クラウドアプリケーション・サービス			クラウド拡張リソース		
			NICTY	WONM	V_AURORA	定常処理サーバ	Webサーバ	Pwrake T/M
GNSS全電子数計測を利用した電離圏の研究	津川 卓也	O	○	—	—	○	○	—
東南アジア域低緯度電離圏観測(SEALION)	津川 卓也	O	—	○	—	—	—	—
短波到来方向探査装置を利用した電離圏の研究	津川 卓也	O	—	○	—	—	—	—
宇宙天気シミュレーション	品川 裕之	S	—	—	○	—	—	—
次世代宇宙天気情報処理の研究	亙 慎一	T	—	—	—	—	—	P
太陽圏モデリング	亙 慎一	O	—	—	—	—	—	—
大気圏・電離圏長期シミュレーションデータの解析	陣 英克	S	—	—	—	—	—	—
太陽風-磁気圏電離圏システム相互作用	久保田 康文	S	—	—	○	—	—	—
ジオスペース・放射線帯予測	長妻 努	OS	○	—	—	○	○	—
南極観測	長妻 努	O	—	○	—	—	—	—
Integrated Satellite Observaion SIMulator for a Coherent Doppler Lidar (ISOSIM-L)による衛星搭載ドップラーライダーのフィジビリティスタディ	石井 昌憲	S	—	—	—	○	—	T/M
SMILES/GOSAT (プロジェクト申請未提出)	笠井 康子	O	—	—	—	—	△	T/M
フェーズドアレイ気象レーダのデータ利用システム(気象レーダの3次元視覚化)	佐藤 晋介	O	—	○	—	—	—	P
生体電磁環境プロジェクト	渡辺 聡一	+	—	—	—	—	—	—
太陽圏の巨視的構造とダイナミクスの研究	鷲見 治一	S	—	—	—	—	—	—
科学衛星搭載プラズマ波動観測器で得られた波形データの特徴解析	笠原 禎也	O	—	—	—	—	—	P
惑星間空間磁場北向き時の磁気圏電離圏対流機構の解明	渡辺 正和	S	—	—	○	—	—	—
GNSS可降水量データベース	藤田 実季子	O	—	—	—	—	—	—
静止軌道衛星帯電プラズマ環境の解析および予測の研究	中村 雅夫	S	—	—	—	—	—	—
地球磁気圏の形状と自由エネルギーに関する初期研究:大規模3次元電磁流体計算と観測を比較する方法の確立	齋藤 実穂	S	—	—	—	—	—	—
気象分野におけるビッグデータ活用技術の研究	大野 智生	O	—	—	—	—	—	—
バーチャルオーロラツールを活用したデジタル磁気嵐現象の研究	海老原 祐輔	S	—	—	—	—	—	—
SS-MIX標準ストレージのNoSQL実装と並列分散処理の検証	木村 映善	+	—	—	—	—	—	—
NICTサイエンスクラウドを用いたゲノムデータ管理基盤に関する研究開発	原田 憲治	+	—	—	—	—	—	—
社会インフラのメンテナンスに資するシミュレーションとセンシングデータの解析	中畑 和之	S	—	—	—	—	—	—
Global MHDシミュレーションの大規模可視化によるプラズマダイナミクス	深沢 圭一郎	S	—	—	○	—	—	P
SALMON (プロジェクト申請未提出)	村山 泰啓	O	—	—	—	—	—	—
時系列データ表示アプリケーション(STARS touch)の開発	村田 健史	T	○	○	—	—	△	P
NICTサイエンスクラウドセキュリティ技術開発	渡邊 英伸	T	—	—	—	—	○	—
NICTサイエンスクラウド高速データ転送表示技術開発	渡邊 英伸	T	—	—	—	—	—	P

Simulation (S)、Observation (O)、Technical (T)、Others (+)

T/M: Torque/Maui

H.25年度サイエンスクラウドによる予算削減効果(試算)

ハードウェア購入費用は含まず

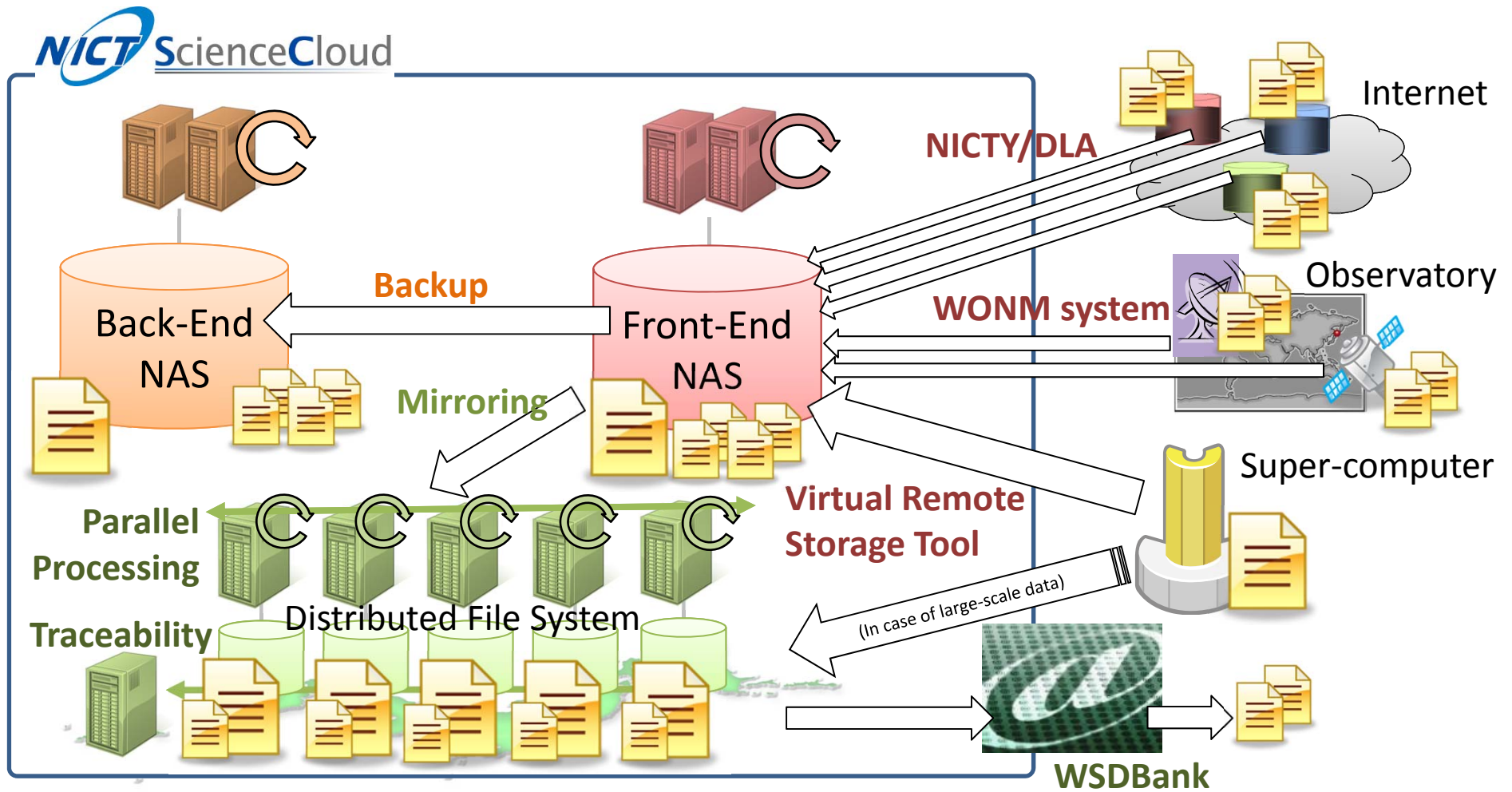
	クラウド運用支援(派遣)	ネットワーク・システムセキュリティ構築・運用支援	Webアプリ開発・公開・運用	システム構築・運用支援(共通リソース)	システム構築・運用支援(アプリケーションツール)	システム構築・運用支援(拡張リソース)	サーバ保守(けいはんな含)	商用アプリケーションライセンス	FW/IP S保守	計

2.12倍

H.25年度サイエンスクラウドによる予算削減効果(試算):作業内容一覧

クラウド運用支援(派遣)	<ul style="list-style-type: none"> サイエンスクラウドユーザ対応、データ整理 可視化ルーム環境整備
ネットワーク・システムセキュリティ構築・運用支援	<ul style="list-style-type: none"> システム設計支援(業務基盤) システム全体セキュリティ管理支援 ネットワーク機器監視・管理支援 ネットワーク監視(監視ツールによる)
Webアプリ開発・公開・運用	<ul style="list-style-type: none"> Webアプリケーション開発 Webアプリケーション開発コンサルティング Webアプリケーション公開支援
システム構築・運用支援 (共通リソース)	<ul style="list-style-type: none"> システム設計・構築・運用(共通リソース) ユーザ管理、セキュリティ管理、ネットワーク管理 システム管理・運用手順(書)策定
システム構築・運用支援 (アプリケーションツール)	<ul style="list-style-type: none"> ツール・アプリ運用、保守 <ul style="list-style-type: none"> ・NICTY(N:975) ・WONM(W:715) ・WSDBank(B:5,850) ・Torque/Maui(T:1,300) ・Gfarm/Pwrake(P:2,340) ・V_AURORA(V:65)
システム構築・運用支援 (拡張リソース)	<ul style="list-style-type: none"> システム設計・構築・運用(拡張リソースおよび独自システム) プロジェクトコンサルタント・Webアプリケーション構築支援 (年度による変動がある)
サーバ保守(けいはんな含)	<ul style="list-style-type: none"> PCサーバ/RHEL/NW機器/JP1/報告(P:2,124) Gfarm用ディスクアレイ/HDD交換(Dg:2,123) バックアップ用ディスクアレイ/HDD交換(Db:2,123)
商用アプリケーションライセンス	<ul style="list-style-type: none"> IDL(I:200) AVS(A:606) XenServer(X:94)
FW/IPS保守	<ul style="list-style-type: none"> 小金井NW機器

科学データ収集・転送・管理・保存・処理
 Science Data File Crawling/Transfer, Preservation/Management and Processing



NASストレージ(フロントエンド)

フロントエンドNAS、バックエンドNAS 利用者 TOP10 <ファイル数>

No	プロジェクト名(代表者、実担当)※1	PJ番号	所属	フロントエンド ファイル数	バックエンド ファイル数	Gfarmファイ ル数(論理)	備考
1	GPS(津川)①	1	NICT	121,572,578	121,572,578	2,482	
2	SEALION(津川)⑥	2	NICT	32,984,392	32,984,392	92,783	
3	STARS(村田)②	28	NICT	5,004,053	5,004,053	12,554,814	
4	フェーズドアレイ(佐藤)④	13	NICT	201,415	2,092,463	52,639	バックエンドにファイルを格納
5	国内電離圏観測(長妻、加藤)⑥	9	NICT	1,839,518	1,839,518	1,839,518	
6	南極観測(長妻、北内)⑥	10	NICT	1,143,327	1,143,327	0	
7	南極観測(長妻、久保田実)⑥	10	NICT	602,733	602,733	602,733	
8	次世代ドップラーCOBRA(佐藤)⑥	—	NICT	0	201,415	201,415	バックエンドにファイルを格納
9	Global MHD(深沢)⑧	26	九州大学	79,104	79,104	614,841	Gfarm/Pwrake処理
10	太陽系モデリング(久保)⑥	6	NICT	74,650	74,650	105,803	

<ディスク使用量>

No	プロジェクト名(代表者、実担当)※1	P番号	所属	フロントエンド ディスク使用量	バックエンド ディスク使用量	Gfarmディスク 使用量(論理)	備考
1	フェーズドアレイ(佐藤)④	13	NICT	1.1TB	127TB	1.9TB	バックエンドにファイルを格納
2	GPS(津川)①	1	NICT	51TB	51TB	1.3GB	
3	Global MHD(深沢)⑧	26	九州大学	16TB	16TB	58TB	Gfarm/Pwrake処理
4	STARS(村田)②	28	NICT	2.4TB	2.4TB	9.9TB	
5	SEALION(津川)⑥	2	NICT	2.2TB	2.2TB	28GB	
6	磁気圏シミュレーション(久保田)⑦	8	NICT	1.5TB	1.5TB	100TB	Gfarm/Pwrake処理
7	次世代ドップラー(佐藤)⑥	13	NICT	0B	1.1TB	1.1TB	バックエンドにファイルを格納
8	電離圏シミュレーション(陣)⑫	7	NICT	900GB	900GB	2.2TB	
9	太陽圏観測(亘、石橋)⑥	6	NICT	780GB	780GB	153GB	
10	国内電離圏観測(長妻、加藤)⑥	9	NICT	650GB	650GB	646GB	

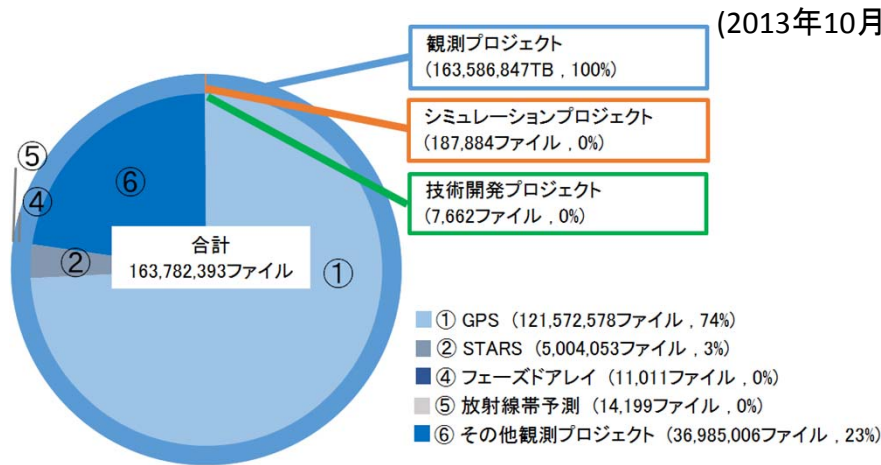
※1 右の○番号は次ページの円グラフの番号を示す。赤文字はバックアップが完了していることを示す。

WONMシステム利用状況(2013年10月時点)

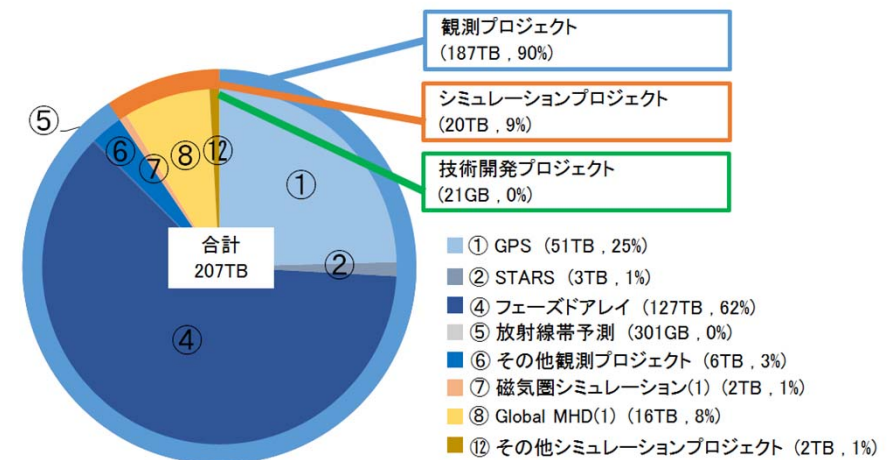
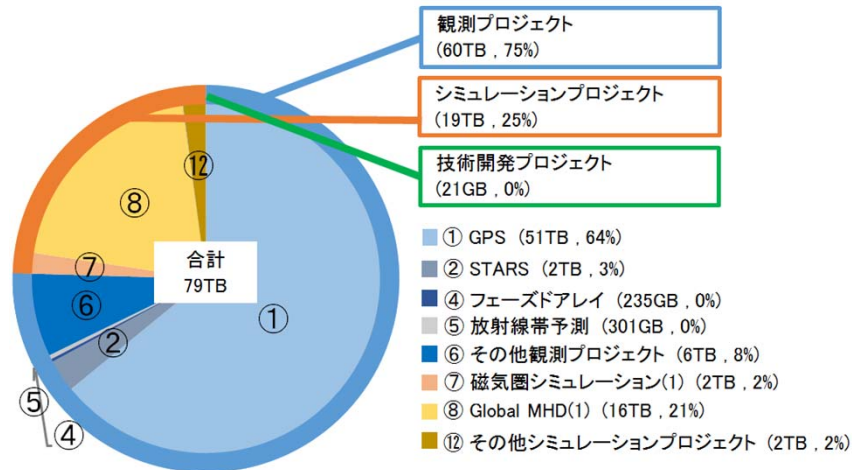
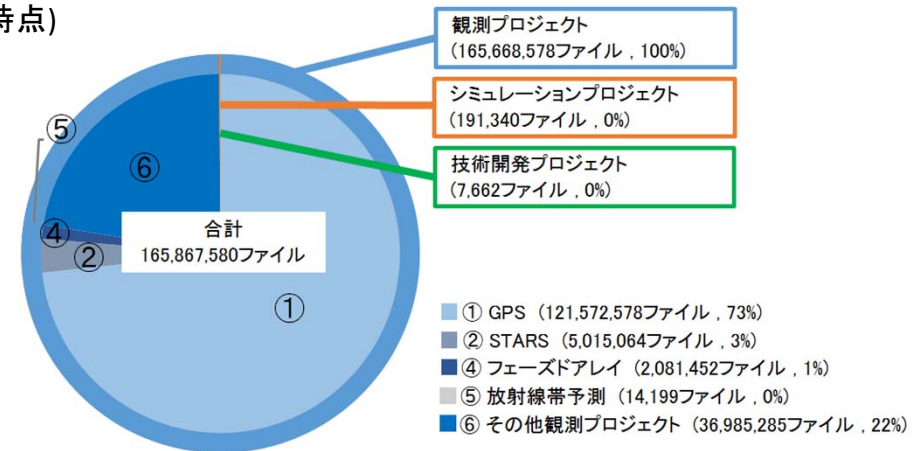
プロジェクト	プロジェクト担当者	拠点	サーバ数	備考
Observation Network	宇宙環境インフォマティクス研究室 山本和憲(kaz-y@ml.nict.go.jp)	小金井	1	テスト用
サイエンスクラウドサーバ 管理	NICTサイエンスクラウド事務局 (osn-system@ml.nict.go.jp)	小金井	96	
		沖縄	3	
		大阪	2	
		名古屋	2	
		けいはんな	15	
Observation Network Test		小金井	3	テスト用
HIRAS	宇宙環境インフォマティクス研究室 山本和憲(kaz-y@ml.nict.go.jp)	—	0	
ISD-J		山川	1	
HF-TEP		—	0	
Magnetometer		—	0	
		チェンマイ	3	
	バンコク	1		
	チュンポン	3		
	プーケット	3		
	コタバン	1		
	バクリウ	3		
セブ	3			
HF Radar	宇宙環境インフォマティクス研究室 長妻努(tnagatsu@nict.go.jp)	キングサーモン	1	
Syowa Station	宇宙環境インフォマティクス研究室 長妻努(tnagatsu@nict.go.jp)	昭和基地	2	
		稚内	1	
		サロベツ	1	
Ishii Lab.	仙台高専 石井誠四郎(ishii@sendai-nct.ac.jp)	仙台	1	
Phased Array	センシングシステム研究室 佐藤晋介	小金井	1	
合計			18	147

NASストレージ(フロントエンド・バックエンド)

フロントエンドNASのプロジェクトごとの ファイル数(上)とディスク使用量(下)

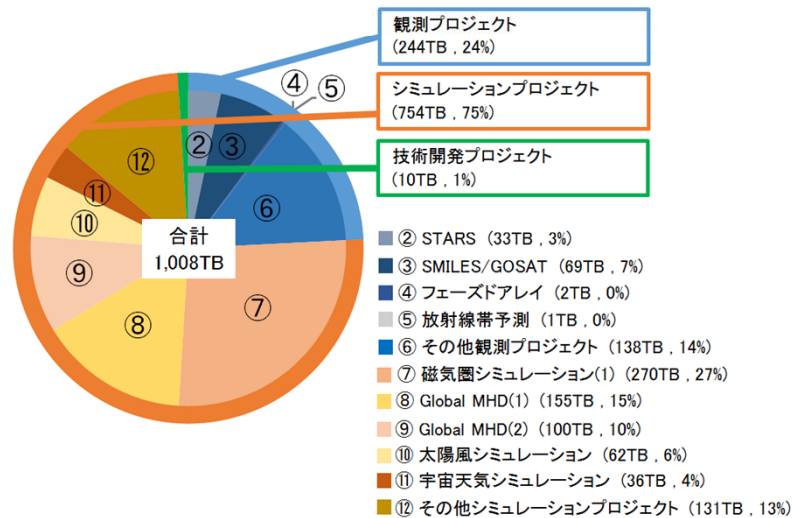
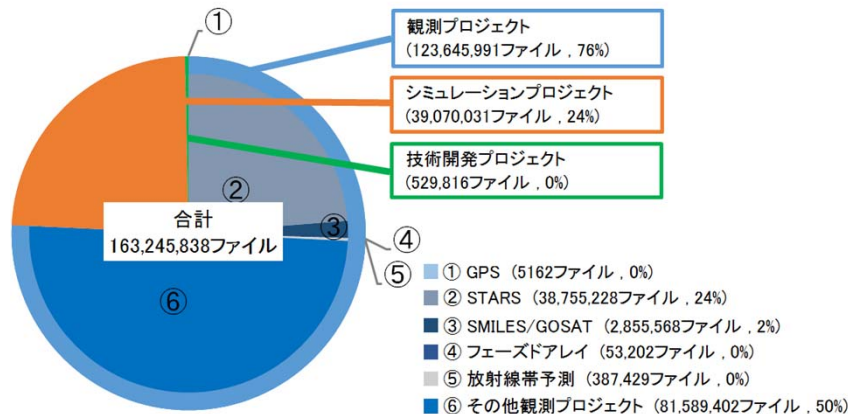


バックエンドNASのプロジェクトごとの ファイル数(上)とディスク使用量(下)



Gfarmストレージ

Gfarmストレージのプロジェクトごとのファイル数(上)とディスク使用量(下)



Gfarmストレージ 利用者 TOP10

<ファイル数>

No	プロジェクト名(代表者、実担当)※1	P番号	所属	論理ファイル数	物理ファイル数
1	STARS(村田)②	28	NICT	12,554,814	38,754,293
2	STARS(村田、戸田)⑥	28	NICT	9,637,425	67,045,227
3	太陽風シミュレーション(坪内)⑩	-	東京大学	3,885,733	11,122,014
4	宇宙天気シミュレーション(品川、島津)⑫	4	NICT	2,606,391	8,951,153
5	宇宙天気予報(亘)⑥	5	NICT	2,525,909	5,551,580
6	国内電離圏観測(長妻、加藤)⑥	9	NICT	1,838,440	5,571,436
7	宇宙天気シミュレーション(品川、田)⑫	4	NICT	1,706,773	4,991,702
8	SMILES/GOSAT(笠井)③	12	NICT	1,418,171	2,851,682
9	宇宙天気シミュレーション(藤田)⑪	4	気象大学校	1,003,717	3,549,663
10	磁気圏シミュレーション(久保田)⑦	8	NICT	978,337	2,592,825

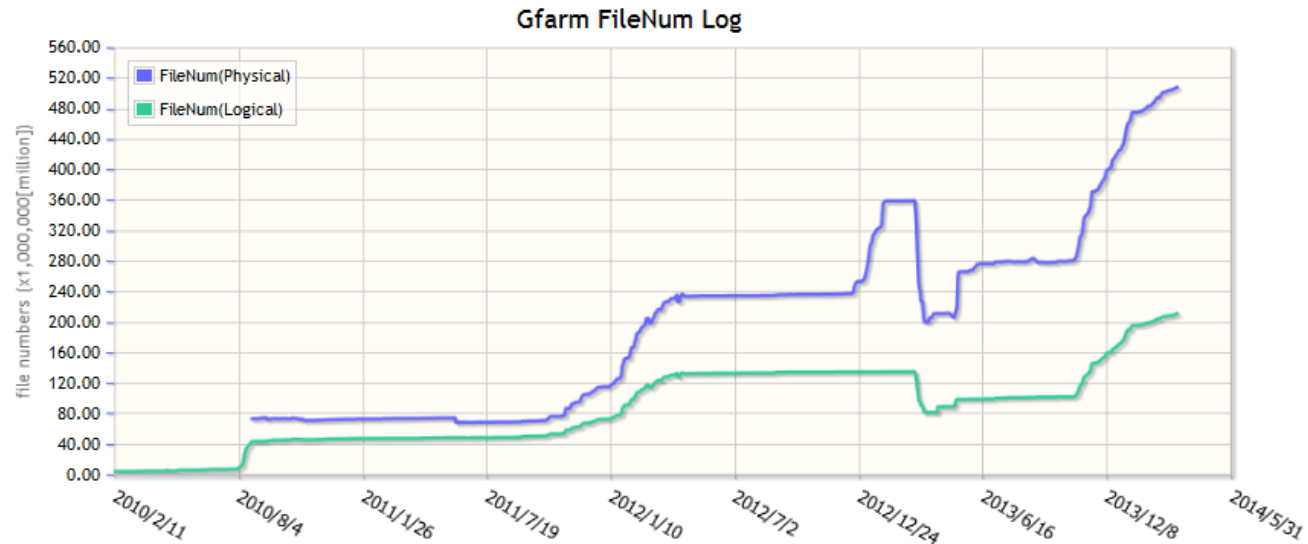
<ディスク使用量>

No	プロジェクト名(代表者、実担当)※1	P番号	所属	論理ディスク使用量	物理ディスク使用量
1	磁気圏シミュレーション(久保田)⑦	8	NICT	100TB	270TB
2	Global MHD(深沢)⑧	26	九州大学	58TB	155TB
3	Global MHD(深沢、郭)⑨	26	NICT	42TB	100TB
4	PiSAR(上本)⑥	-	NICT	29TB	74TB
5	SMILES/GOSAT(笠井)③	12	NICT	35TB	69TB
6	太陽風シミュレーション(坪内)⑩	-	東京大学	21TB	62TB
7	宇宙天気シミュレーション(藤田)⑪	4	気象大学校	14TB	36TB
8	宇宙天気シミュレーション(品川、田)⑫	4	NICT	10TB	31TB
9	STARS(村田)②	28	NICT	9.9TB	33TB
10	STARS(村田、戸田)⑥	28	NICT	7.4TB	51TB

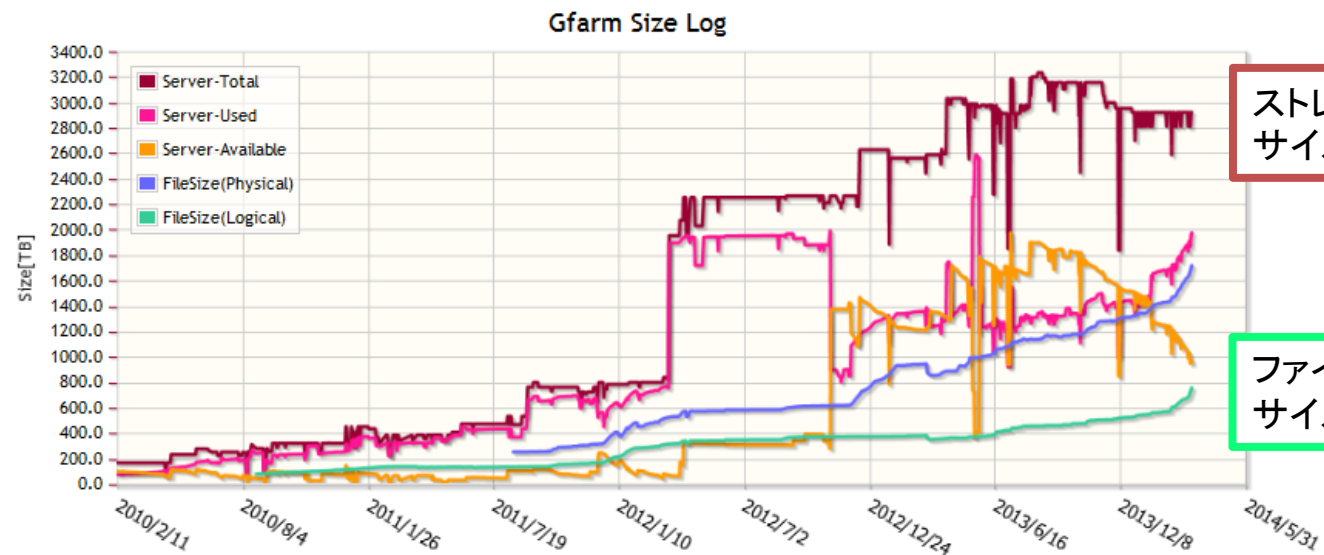
※1 右の番号は円グラフのプロジェクト番号を示す。

分散ストレージシステム (Gfarm) 履歴 (2010年2月～2014年3月)

- 管理ファイル数
(論理・物理)



- 参考
- ストレージサイズ
 - 管理ファイルサイズ
(論理・物理)

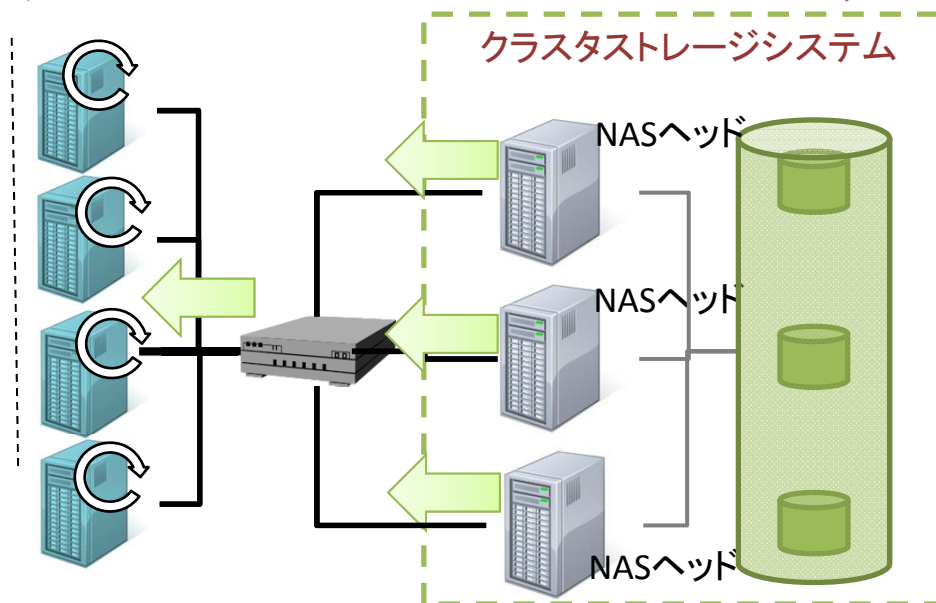
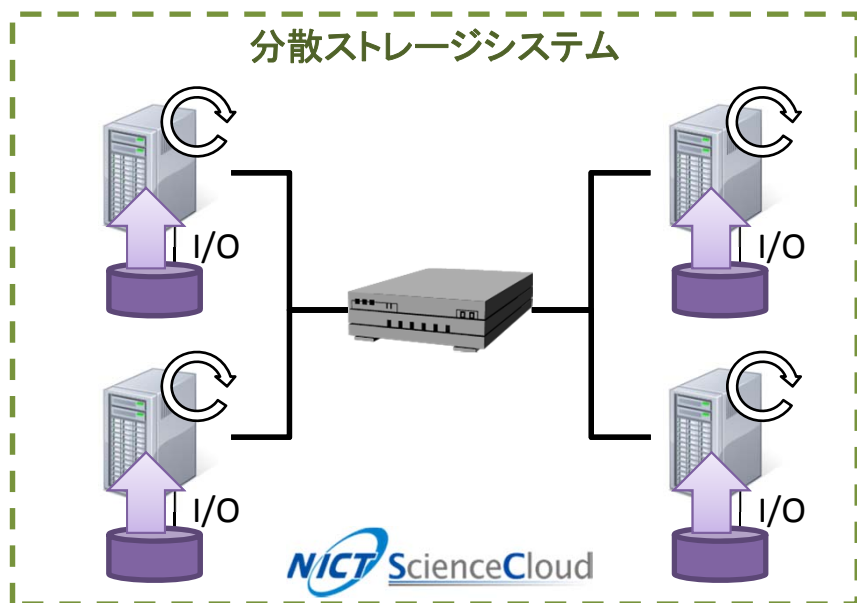


ストレージ
サイズ

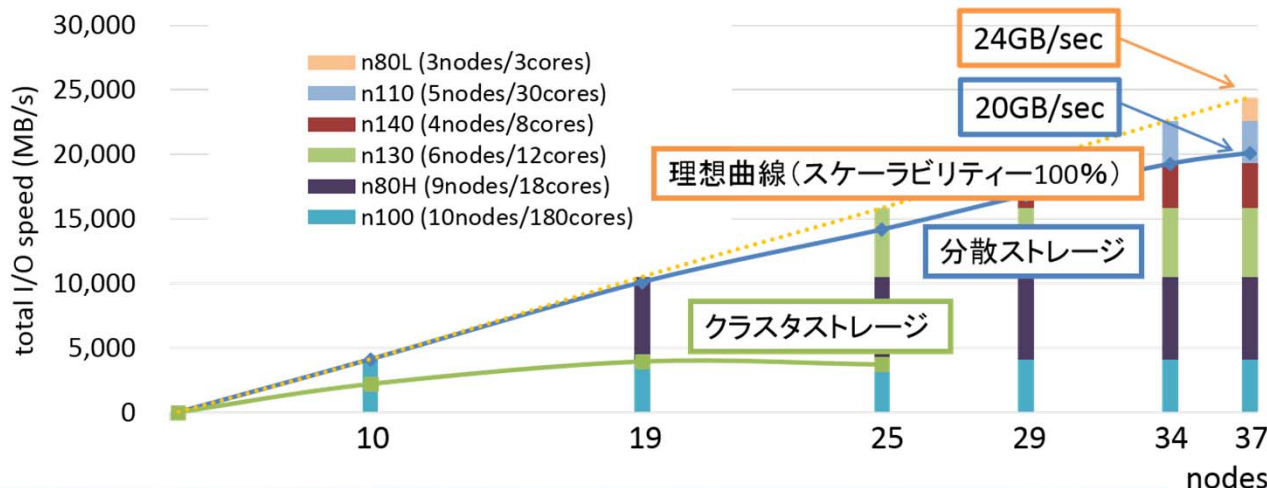
ファイル
サイズ

分散ストレージの高速データI/O性能評価(クラスタストレージとの比較)

比較対象(クラスタストレージシステムを使ったシステム例)



クラスタストレージと分散ストレージのI/Oのスケラビリティ比較(クラスタストレージシステムとしてGPFSを利用)



- 広域クラウド内に37台のファイルサーバを設置
- クラスタストレージはネットワークがボトルネックになる
- 分散ストレージは80%以上のスケラビリティを達成

NICT開発技術 信頼性機能(タイムスタンプ)

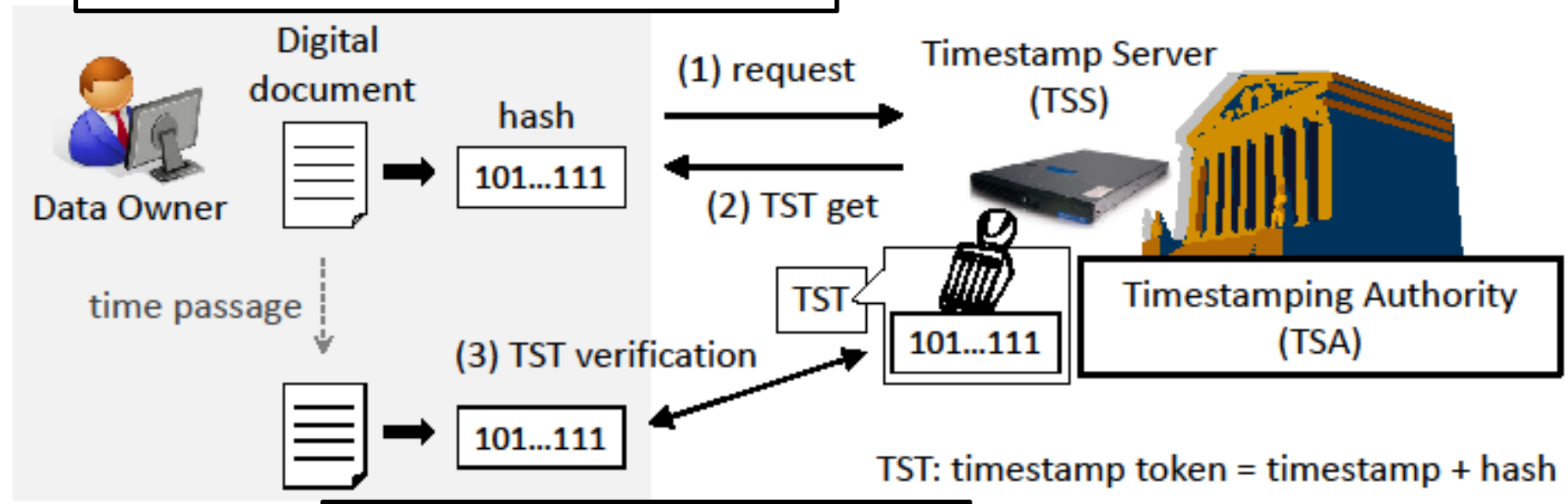
背景

- 時刻認証局(TSA)は日本では4社のみ(アマノ、セイコーソリューションズ、NTTデータ、北海道総合通信網)
- 現状では各社とも自社のTSAサーバでのタイムスタンプしか認めていないためクラウドの高速・大容量のデータベースには未対応
- 今後はこのようなクラウドに直結(内包)タイムスタンプ方式の実用化を視野に入れたビジネスモデルが必須である

サイエンスクラウドの成果

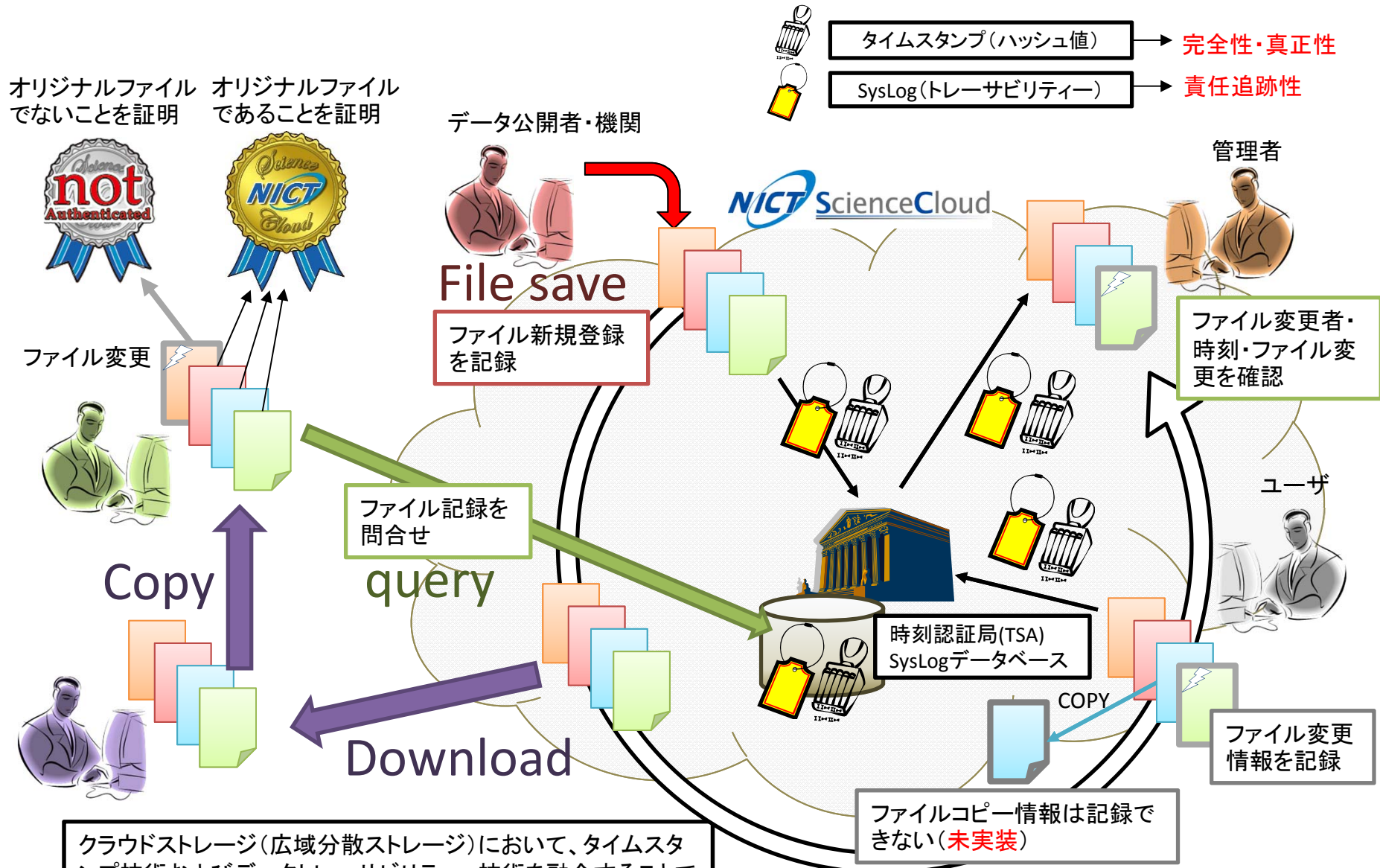
- 分散ストレージシステムとタイムスタンプ・サービスの協調機能によりデータトレーサビリティを実現
- 重要なデータの完全性と真正性を担保

ファイルをストレージに保存するだけで(ユーザは意識せずに)タイムスタンプ付与が可能となった。



ファイルあたりタイムスタンプ付与(90ms)、タイムスタンプ検証(40ms)の現実性の高い性能を達成。

NICT開発技術 総合的信頼性認証機能

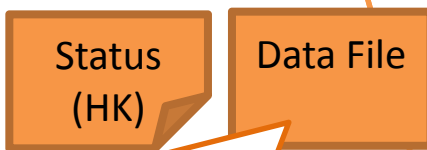
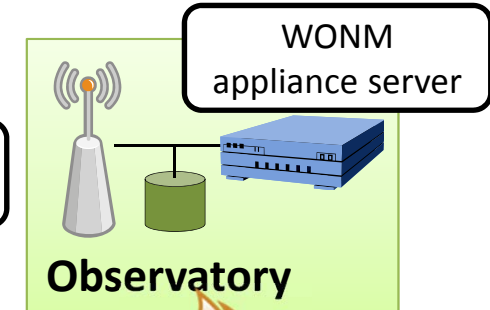
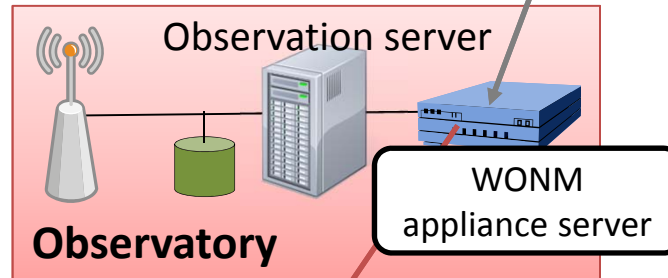
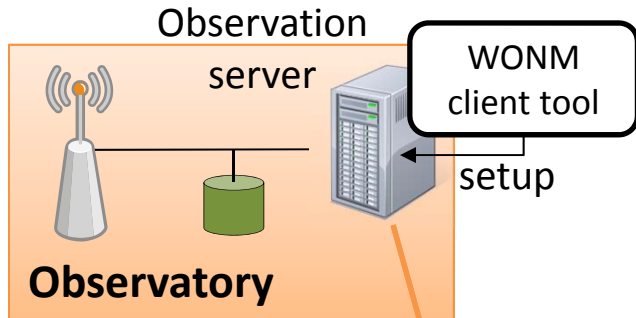


クラウドストレージ(広域分散ストレージ)において、タイムスタンプ技術およびデータトレーサビリティ技術を融合することで、データ完全性・真正性・責任追跡性認証システムを実現。

広域観測網監視システム Wide-area Observation Monitoring System

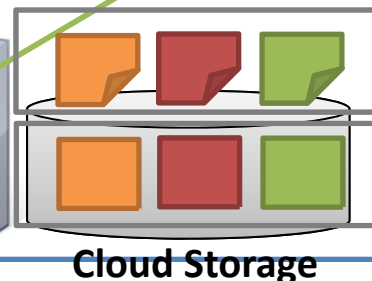
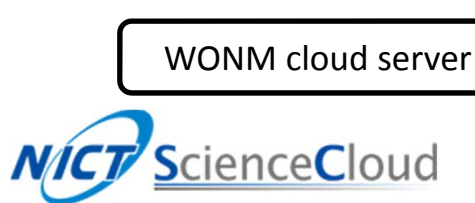


- ① Monitoring Server
- High tolerance
 - Automatic recovery
 - Redundancy



- ② Data Transfer
- Data Transfer Retry
 - High performance network band
 - Data consistency

- ③ Easy & Integrated Management
- Warning System
 - Monitoring System

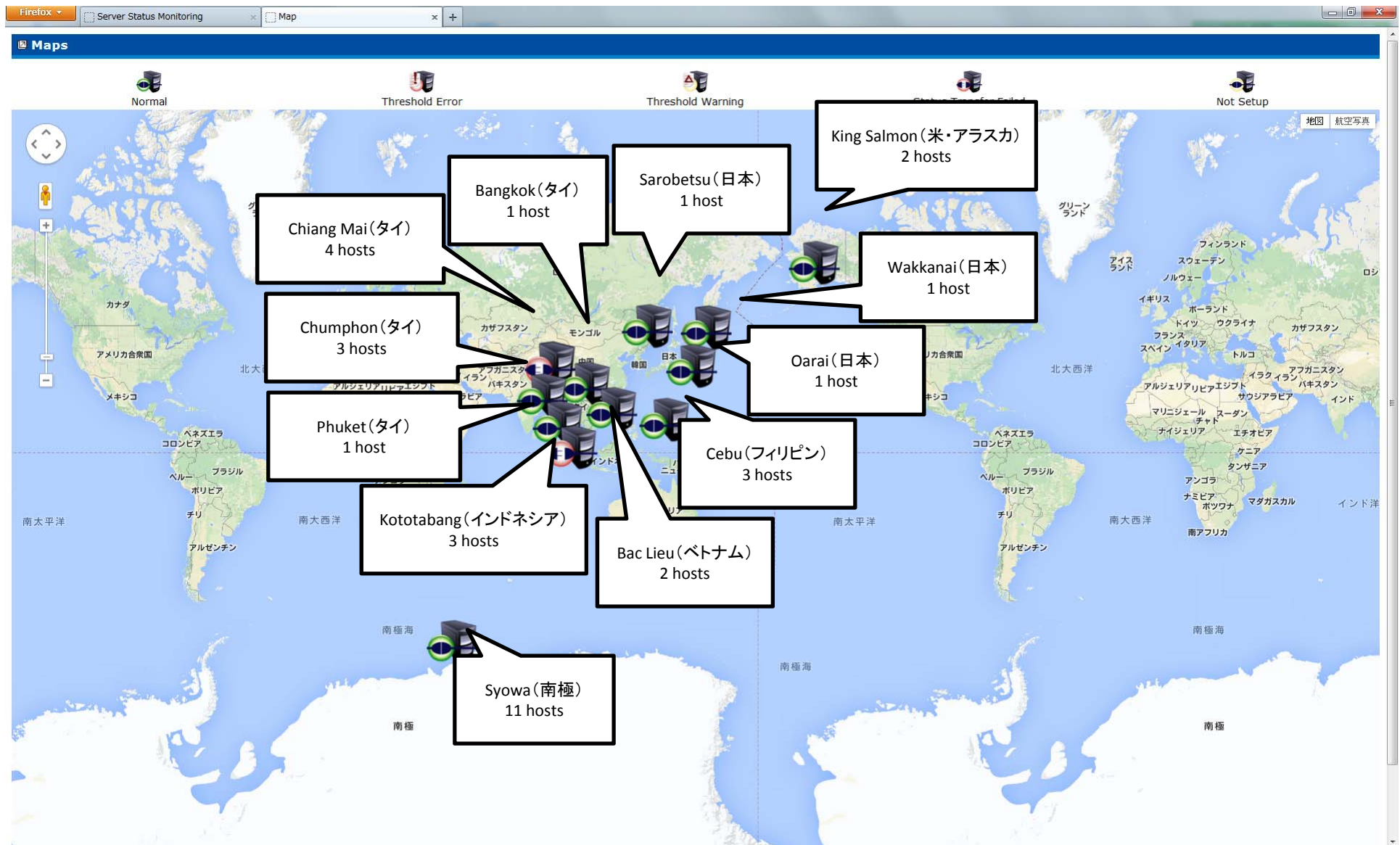


WONM Web

Data Analysis



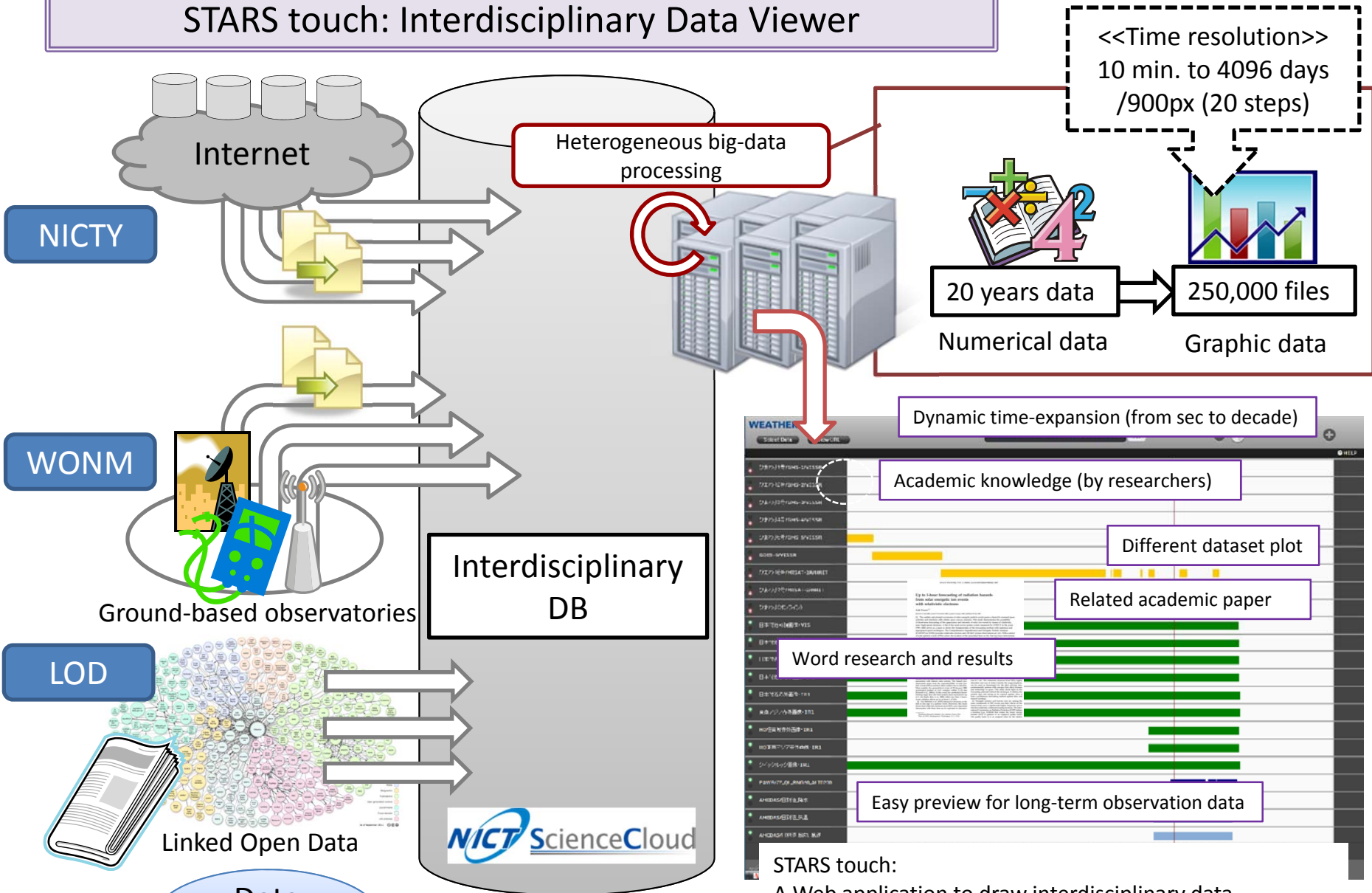
WONMシステムによる観測所管理(宇宙環境のみ: ~H.26年度)



データ収集実績(～2014年2月)

Data	PJ#	Organization	Tool	Number of collected files	Total data size
GNSS observation (宇)	1	UNAVCO, GSI	NICTY	23,506,753	9.6TB
SEALION (セ)	2	NICT	WONM	60,787	1.6TB
GOSAT satellite (セ)	12	JAXA, ECMWF	*	3,737,123	58.4TB
Phased-array meteorological data (セ)	13	Osaka Univ., NICT	*	2,358,677	217.1TB
SMILES (セ)	12	JAXA	*	564,253	1.4TB
Antarctica (ionosphere observation) (宇)	10	NICT	WONM	280,859	424GB
Geo-magnetic data (宇)	9	NICT, Kyoto Univ.	NICTY	277	1.5TB
GMS satellite	21	JMA	NICTY	26,868	26.8GB
GEOTAIL satellite	28	Kyoto Univ.	NICTY	29,840	204GB
GOES satellite (宇)	9	NOAA	NICTY	1,638	59MB
ACE satellite (宇)	9	NOAA	NICTY	759	203MB
KODAMA satellite (宇)	9	JAXA	NICTY	3,890	1.8GB
KAGUYA satellite	16	Kanazawa Univ.	*	3,227	221GB
Space Weather real-time simulation (宇)	4	NICT	NICTY	63	3.5TB

分野横断型時系列データプレビューア STARS touch: Interdisciplinary Data Viewer



背景: STP分野のデータ表示ツール

CDAWeb

- NASA/GSFCで開発
- Webアプリ

DARTS

- 宇宙科学研究所で開発
- Webアプリ

TDAS/UDAS

- UCLA/名古屋大学 (IUGONET)で開発
- IDLアプリ (Web版もある)

STARS

- 愛媛大学で開発
- Windowsアプリ

- OSが制限されるとユーザが固定する (広く活用されない)
- Webアプリはデータ解析機能追加が困難
- 有償アプリを基盤とすると利用者が制限される
- ほとんどのアプリが時刻・データを指定してサブミットすることでデータ表示⇒データの連続閲覧が面倒

◆基本操作① データ存在表示◆

Webアプリケーション起動時の表示画面。

データごとにデータファイルが存在する期間をバー表示する。複数衛星データ比較でデータ存在期間を確認する場合などで有効。データ存在バー上または時間軸上でマウスでスライドすることで表示時間を連続的に変更することができる。

The screenshot displays the GEOTAIL SATELLITE web application interface. The browser address bar shows 'geotail.nict.go.jp'. The main content area features a 'Select Data' button and a list of satellite data sources: GEOTAIL/PWI/SFA, GEOTAIL/MGF, GEOTAIL/LEPSDB, GEOTAIL/PWI/WFC, and GEOTAIL/ORBIT. Each source has a corresponding blue bar representing data availability over time. A time axis at the bottom shows dates from 1997/4/1 to 1997/8/1. A mouse cursor is positioned over the time axis, with a callout box labeled '時間移動' (Time Movement). Another callout box labeled '観測データの有無を一目で認識' (Recognize data presence at a glance) points to the blue bars. The interface also includes a 'ZOOM' slider and a 'HELP' button.

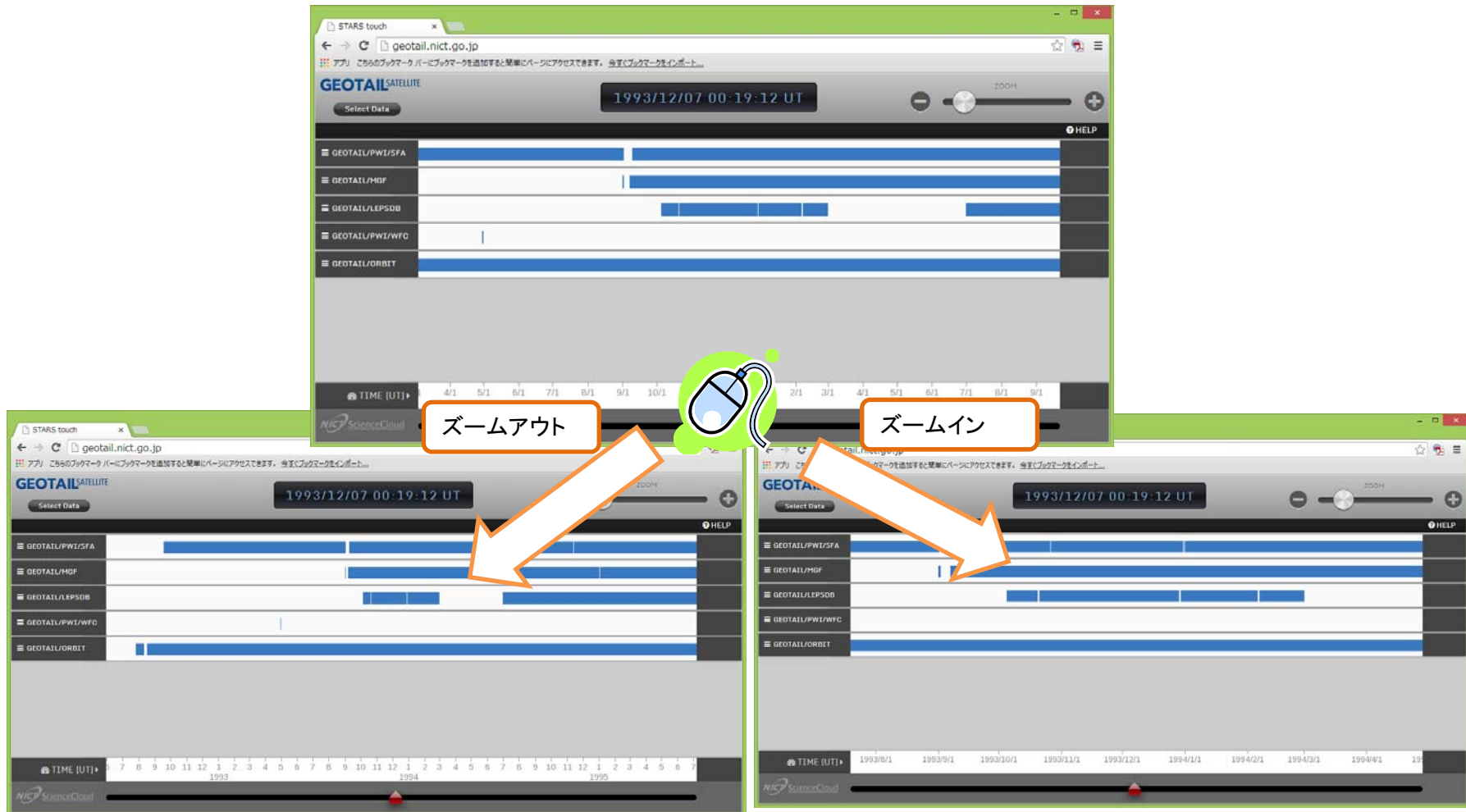
◆基本操作② データ時刻表示(スライダ)◆

時刻スライダにより時刻を指定する(表示している時間帯の任意の時刻を指定することができる)。時刻スライダが表示する時刻は上部の時刻窓に表示される。



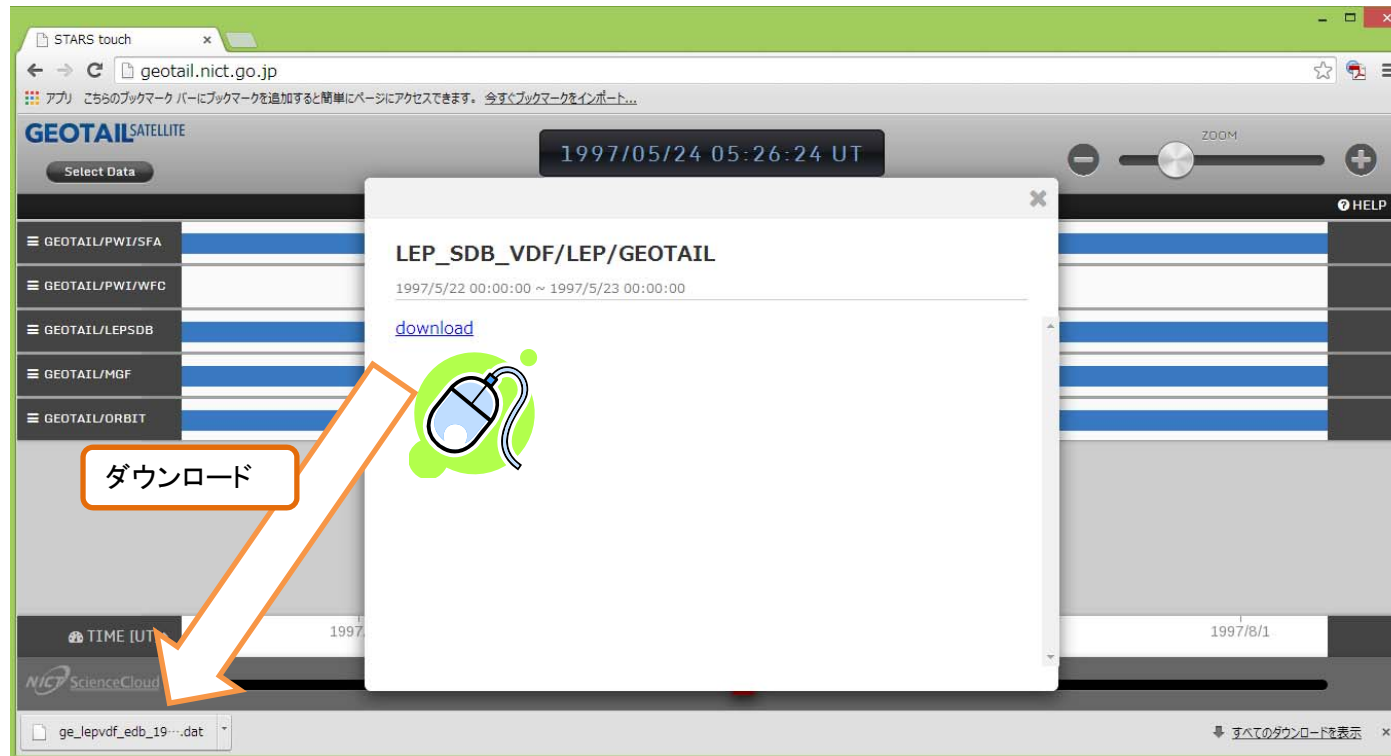
◆基本操作③ 時間幅変更◆

時間軸上でマウスダブルクリックにより表示時間を連続的に変更することができる。
左ダブルクリックで時間ズームイン、右ダブルクリックで時間ズームアウトする。



◆基本操作④ データダウンロード◆

データバー上でマウス左ダブルクリックで、該当する任意時刻のデータファイルをダウンロードする。データによりフィルダウンロードできるデータとできないデータがある。(ダウンロード機能は、次期バージョンでUIが変更になる予定である。)



STARStouchアプリ対象データ(予定)

科学データ(STARStouch)

- [GEOTAIL衛星版\(ver1.0公開\)](#)
- STPグローバル研究版
 - [地磁気データの融合表示\(京都大学他\)](#)
 - オーロラ地上観測画像(昭和基地)
 - 太陽風モニタリング衛星(ACE衛星)
 - 太陽極端紫外観測(SOHO衛星画像など)
- ひまわり衛星・気象データ版

社会データ(SOMATO)

- **新聞記事版(Ver1.0をH.27年度中に公開)**
 - 読売新聞社記事をSTARStouchに表示(アプリ名:SOMATO)
 - 特許申請の関係で詳細は省略

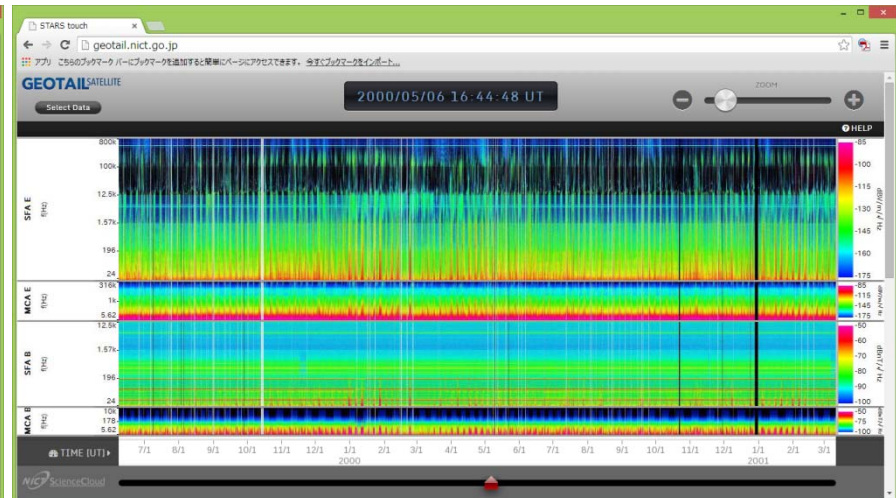
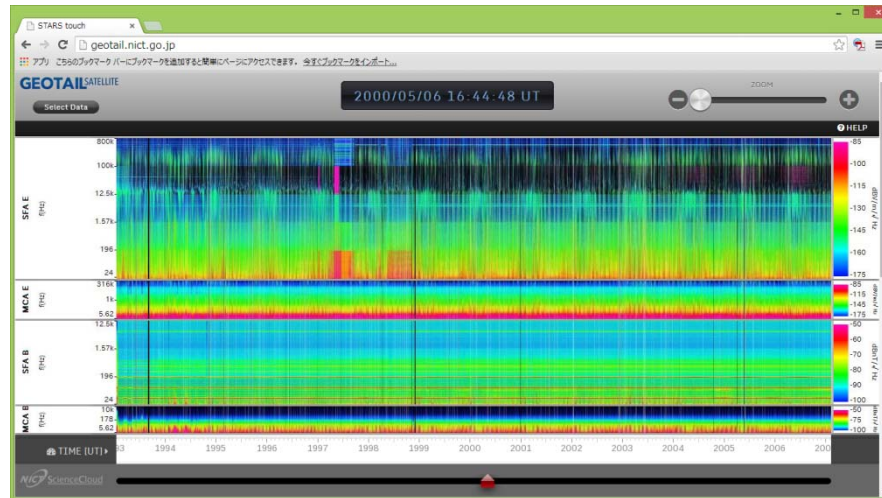
◆時間表示機能① 時間スライド表示◆



◆時間表示機能② ズームイン・ズームアウト◆

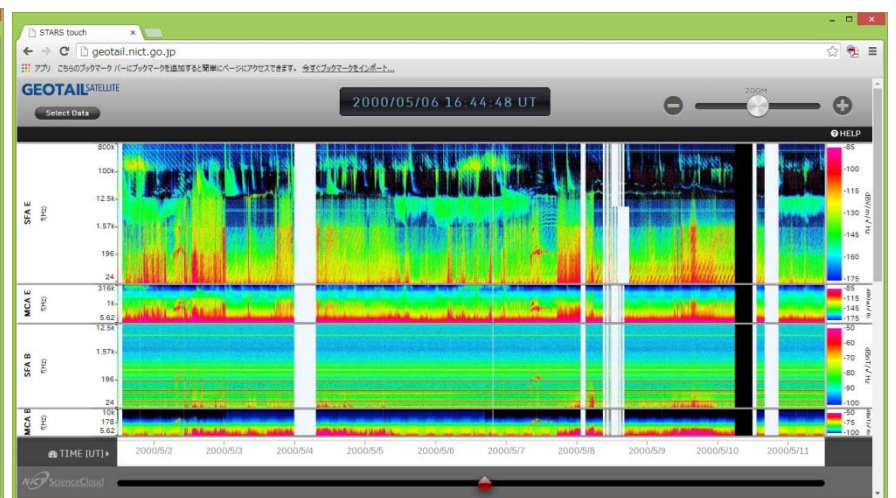
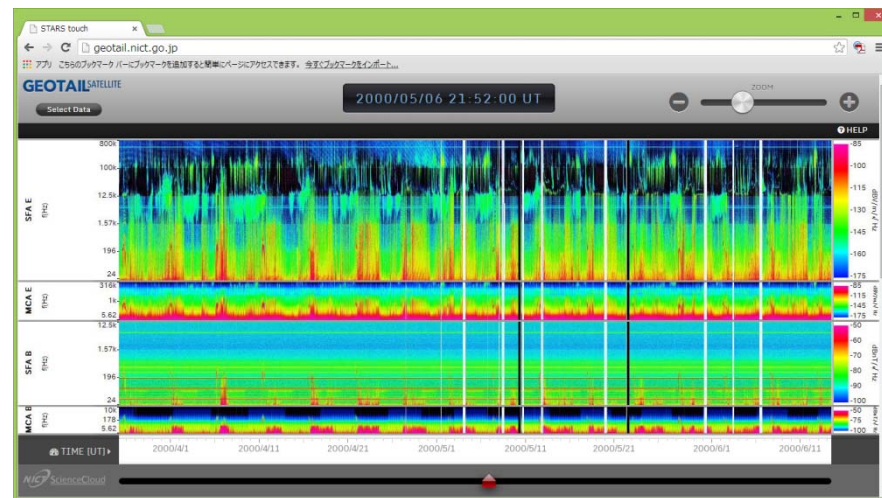
4096日(11.2年)

512日(1.4年)



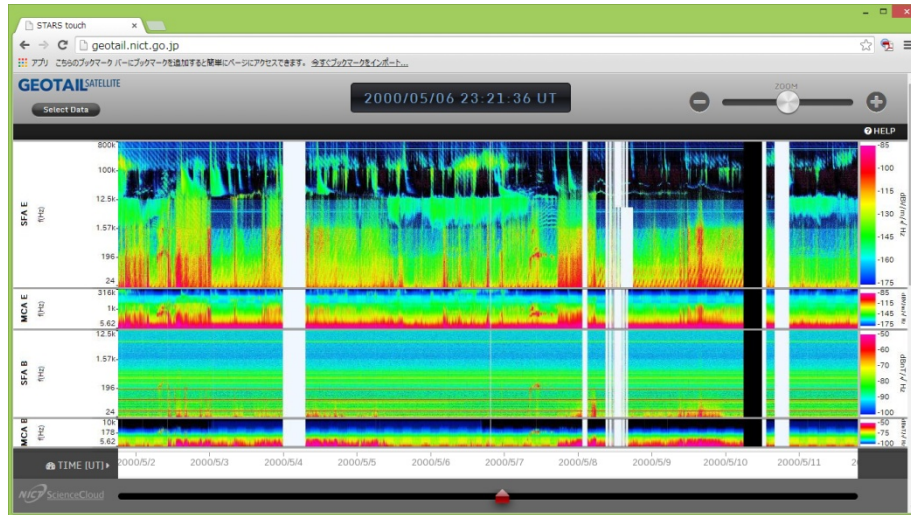
64日

8日

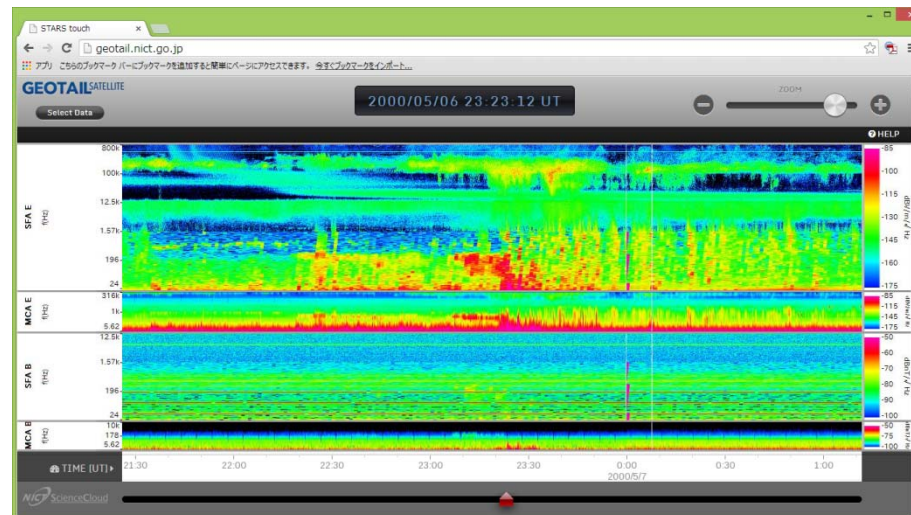


◆時間表示機能② ズームイン・ズームアウト◆

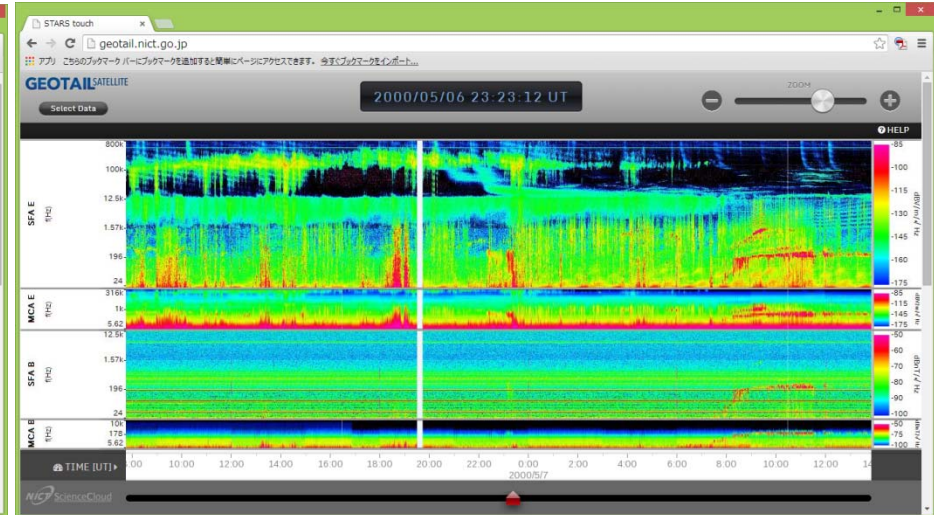
8日



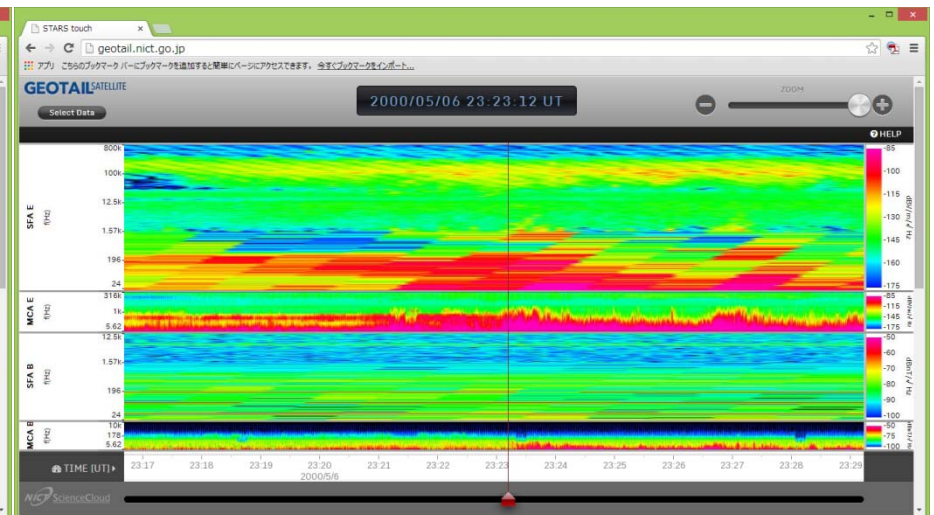
3時間



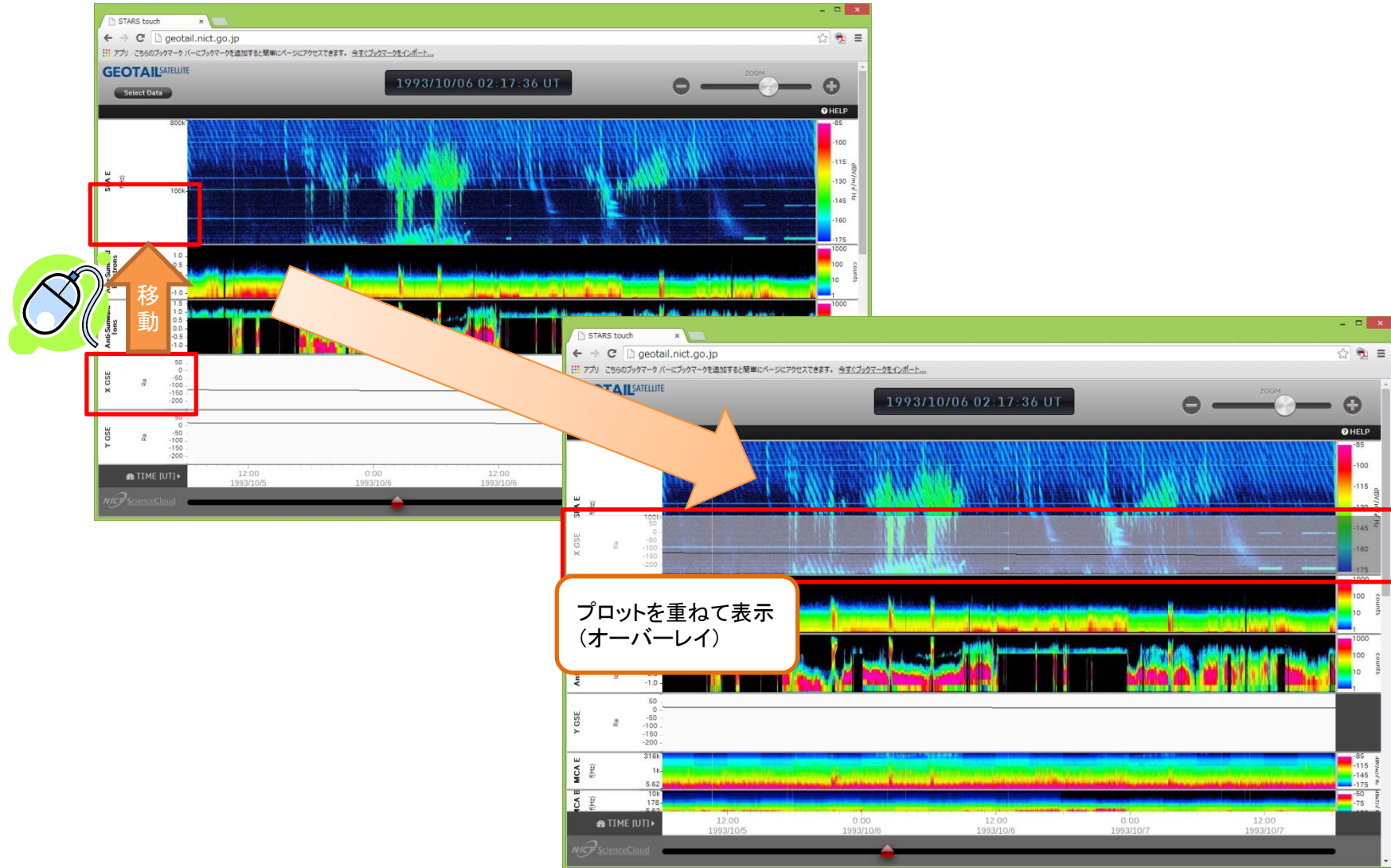
1日



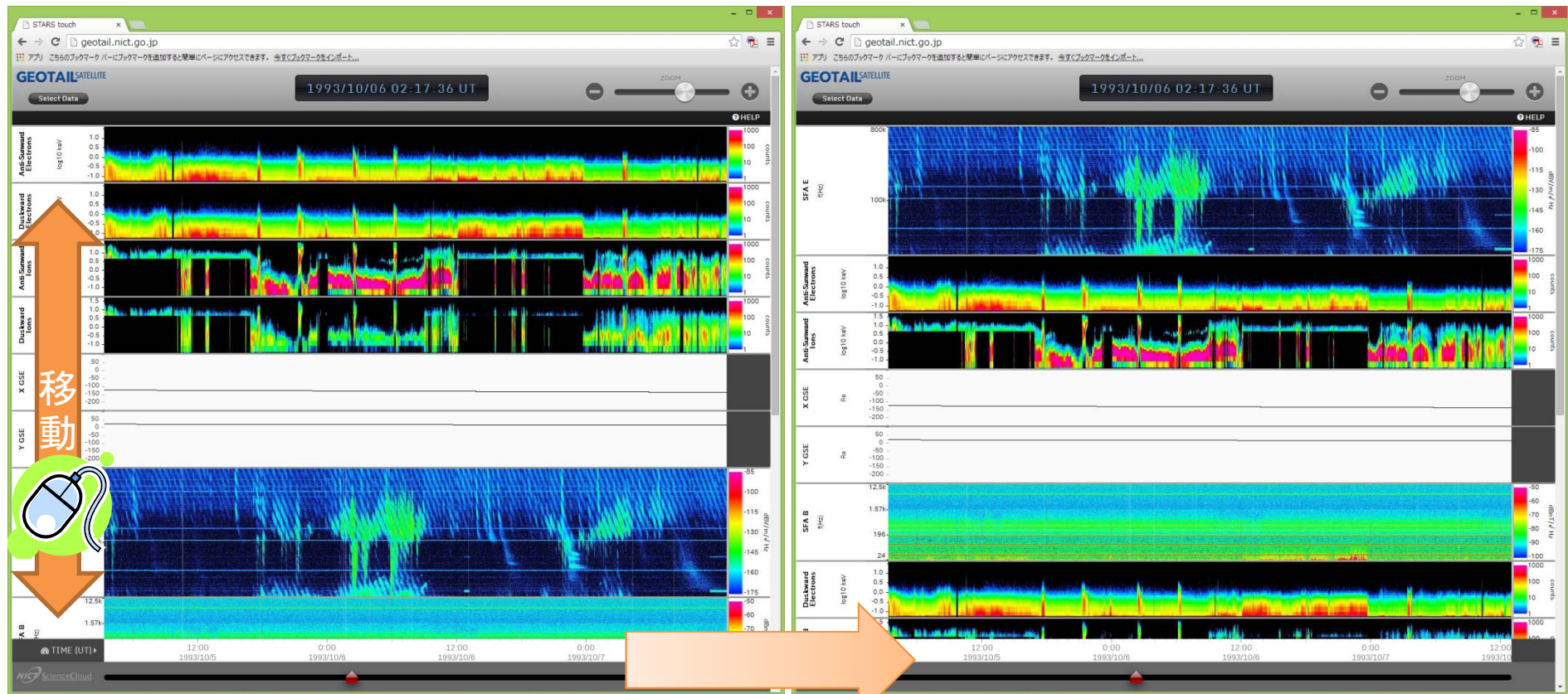
10分



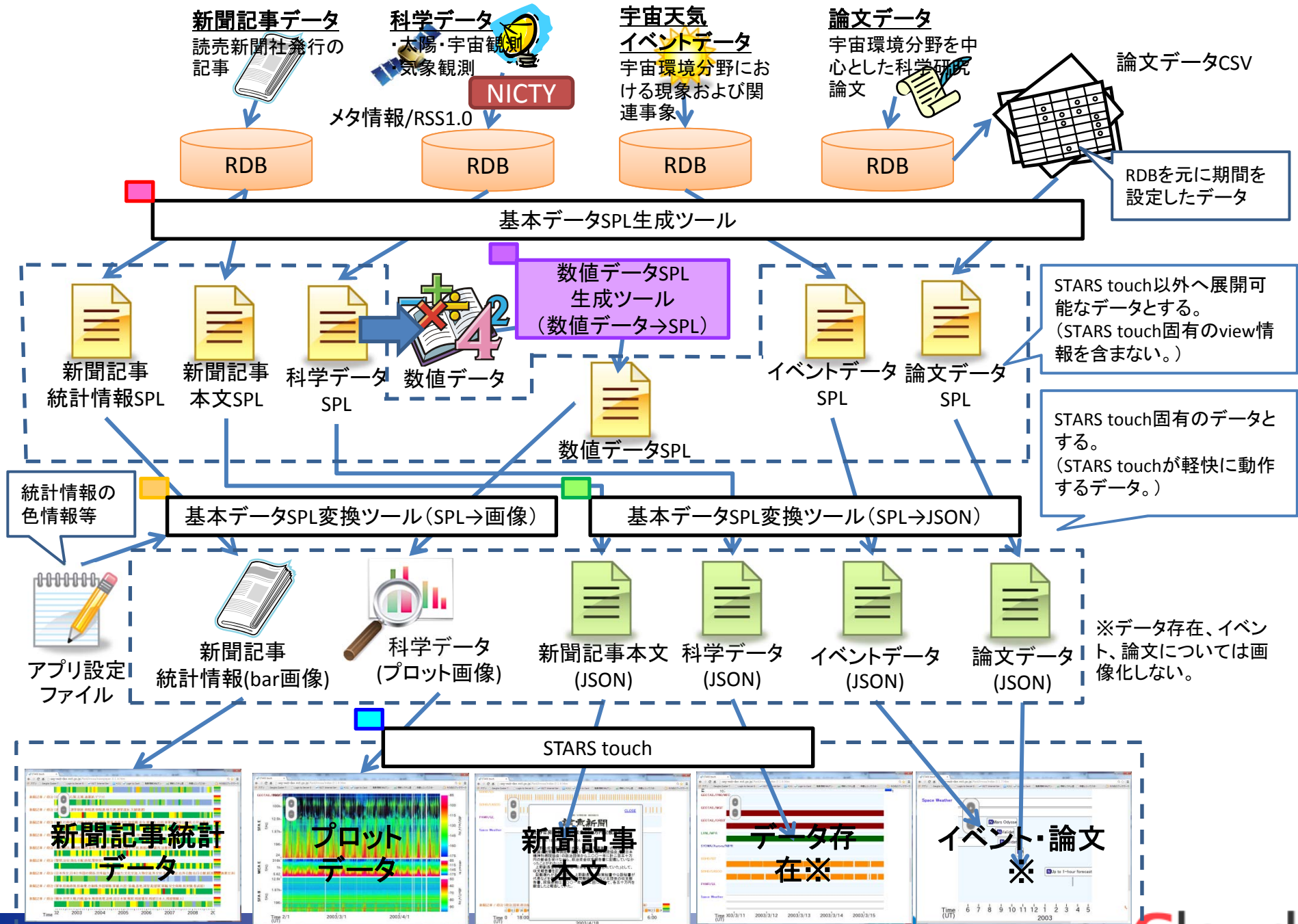
◆データオーバーレイ表示機能◆



◆表示データレイアウトカスタマイズ機能◆

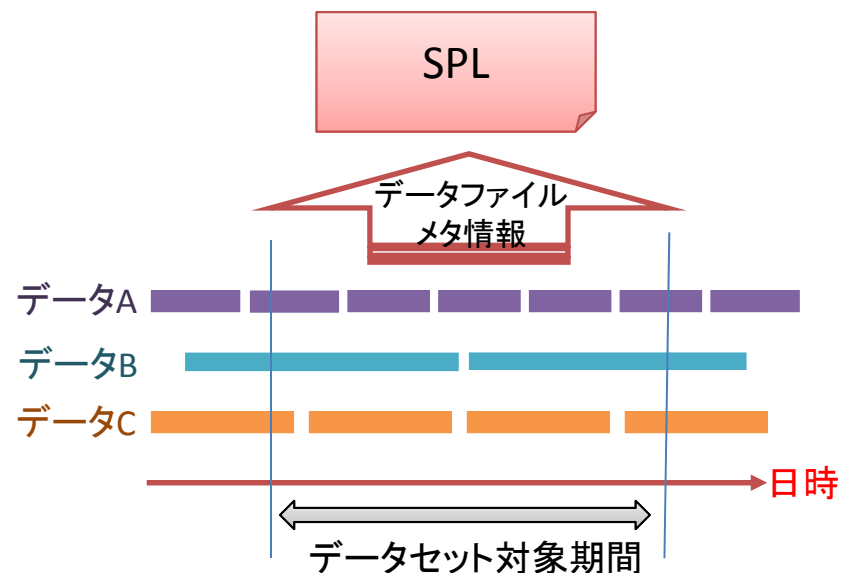


利用者が見たいデータに組み替えることが可能



SPLについて

- SPL(Stars Project List)とは、データファイル群(データセット)のメタ情報(データ種別、データファイル名、データ期間など)をXML形式で定義した情報
- SPLを利用することで効率的なデータ解析が可能
- SPLはSTARSメタデータベースより検索、取得することもできる



SPL(データファイルの例)

```
▼<StarsProjectList>
  ▼<StarsEventList st="2449231.5" et="2449232.5">
    <User id="10001" name="Guest"/>
    ▼<Mission id="20000" name="GEOTAIL">
      ▼<Team id="20200" name="MGF/GEOTAIL">
        ▼<Data id="20209" name="GE_EDAB3SEC" label="GEOTAIL/MGF">
          <Component name="D20209" type="StarsFiles"/>
          <File id="355520487771606823" name="ge_edb3sec_mgf_19930913_v01.cdf" st="2449243.5" et="2449244.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930913_v01.cdf"/>
          <File id="355520487771606826" name="ge_edb3sec_mgf_19930916_v01.cdf" st="2449246.5" et="2449247.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930916_v01.cdf"/>
          <File id="355520487771606829" name="ge_edb3sec_mgf_19930919_v01.cdf" st="2449249.5" et="2449250.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930919_v01.cdf"/>
          <File id="355520487771606832" name="ge_edb3sec_mgf_19930922_v01.cdf" st="2449252.5" et="2449253.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930922_v01.cdf"/>
          <File id="355520487771606835" name="ge_edb3sec_mgf_19930925_v01.cdf" st="2449255.5" et="2449256.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930925_v01.cdf"/>
          <File id="355520487771606838" name="ge_edb3sec_mgf_19930928_v01.cdf" st="2449258.5" et="2449259.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19930928_v01.cdf"/>
          <File id="355520487771606846" name="ge_edb3sec_mgf_19931006_v01.cdf" st="2449266.5" et="2449267.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19931006_v01.cdf"/>
          <File id="355520487771606849" name="ge_edb3sec_mgf_19931009_v01.cdf" st="2449269.5" et="2449270.5"
          url="ftp://cdaweb.gsfc.nasa.gov/pub/data/geotail/edb3sec_mgf/1993/ge_edb3sec_mgf_19931009_v01.cdf"/>
        </Data>
      </Team>
    </Mission>
  </StarsEventList>
</StarsProjectList>
```

データ名

観測期間(ユリウス日/UT)

ダウンロードURL

SPL(数値データの例)

```
<StarsProjectList>
  <StarsEventList st="2450204.5" et="2450235.5">
    <User id="20009" name="smallsat"/>
    <Mission id="20000" name="GEOTAIL">
      <Team id="20200" name="MGF/GEOTAIL">
        <Data id="20209" name="GE_EDAB3SEC" timeresolution="0000">
          <Component name="BGSE" type="StarsScalars" unit="nT" scalemin="-100" scalemax="100" validmin="-100" validmax="100">
            <Element index="0" label="edb Bx (GSE)"/>
            <Element index="1" label="edb By (GSE)"/>
            <Element index="2" label="edb Bz (GSE)"/>
            <Element index="3" label="eda Bx (GSE)"/>
            <Element index="4" label="eda By (GSE)"/>
            <Element index="5" label="eda Bz (GSE)"/>
          </Component>
          <Epoch juliustime="2450204.50003714">
            <Value index="0">4.8000001907348633</Value>
            <Value index="1">-2.0999999046325664</Value>
            <Value index="2">1.9600000381469727</Value>
            <Value index="3">4.8000001907348633</Value>
            <Value index="4">-2.0999999046325664</Value>
            <Value index="5">1.9600000381469727</Value>
          </Epoch>
          <Epoch juliustime="2450204.50007257">
            <Value index="0">4.8400001525878906</Value>
            <Value index="1">-2.130000114440918</Value>
            <Value index="2">2.0199999809265137</Value>
            <Value index="3">4.8400001525878906</Value>
            <Value index="4">-2.130000114440918</Value>
            <Value index="5">2.0199999809265137</Value>
          </Epoch>
          <Epoch juliustime="2450204.500108">
            <Value index="0">4.78000020980835</Value>
            <Value index="1">-2.130000114440918</Value>
            <Value index="2">2.1500000953674316</Value>
            <Value index="3">4.78000020980835</Value>
            <Value index="4">-2.130000114440918</Value>
            <Value index="5">2.1500000953674316</Value>
          </Epoch>
        </Data>
      </Team>
    </Mission>
  </StarsEventList>
</StarsProjectList>
```

unit : 単位
scalemin/scalemax : プロット縦軸最小/最大値
validmin/validmax : データ最小/最大値

数値要素(プロット単位)
index : 要素番号
label : 要素名

ある日時での数値データ
juliustime : 日時(ユリウス日/UT)

数値データ
index : 要素番号(index@Elementに対応)

まとめ

- (いよいよ)科学ビッグデータ時代の到来
 - 今後はTB～PBスケールのデータセットが増加
 - 大量データをどのように効率的に管理・保存できるか
 - 大量のデータから必要なデータをどれだけ迅速に探せるか
 - データサイズだけではなくデータ種類の増加
 - 様々なデータに対して、出来る限りコストを掛けず公開する手法は？
 - 様々な研究分野においてクラウド(的な考え方)を活用できる分野が一步進むことができるでしょう