サイエンティフィック・システム研究会
Scientific Systems

クラウド時代のデータ共有技術WG
成果報告書

2019 年 4 月 30 日

サイエンティフィック・システム研究会
クラウド時代のデータ共有技術WG

目　次

1.　はじめに(WG 活動概要) 1
1.1. WG 発足の背景 1
1.2. 活動方針 1
1.3. 活動内容 2
1.4. 活動期間 2
1.5. WG メンバー 2
1.6. 活動実績 2
1.7.　参考文献 3
2. クラウド時代のデータ共有技術の現状と問題 4
2.1. 昨今の要請と問題（概要説明） 4
2.2. クラウド技術の動向 4
2.2.1. クラウドのサービス～オンプレミスのサービス 5
2.3. 大規模データ・大量データを活用する分野における課題 5
2.3.1. クラウドサービスおよびオンプレミスシステムそれぞれの利点と欠点 6
2.3.2. オンプレミスとクラウドサービスの違い～システム管理者的視点より～ 7
2.3.3. ユーザがクラウドサービス利用時に期待する点 7
3.1. サービス動向 8
3.1.1. オンラインストレージ（ファイル保管・データ共有） 8
3.1.2. グループウェア（メール・ドキュメント共有・ファイル共有・スケジュール共有） 9
3.1.3. プロジェクトホスティングサービス 10
3.1.4. 認証サービス 10
3.2. ソフトウェア動向 10
3.2.1. ファイル共有 10
3.2.2. ファイルシステム 11
3.2.3. オブジェクトストレージ 12
3.2.4. ブロックチェーン 13
3.3. ハードウェア動向 13
3.3.1. ストレージ技術動向 13
3.3.2. 価格容量比率など 14
3.4. データ転送技術 14
3.4.1. ストレージゲートウェイ 14
3.4.2. 高速データ転送 14
3.5. 現在までの大規模・大量データの共有方法と分析 15
3.5.1. データを配布する側の事例　【A.1. 早戸委員】 15
【東京大学　宇宙線研究所 神岡】 15
3.5.2. データを配布する側の事例　【A.2. 藤田委員】 15
【宇宙航空研究開発機構 調布】 15
3.5.3. データを配布する側の事例　【A.3. 中村委員】 15
【高エネルギー加速器研究機構】 15
3.5.4. データを配布する側の事例　【A.4. 白崎委員】 15
【国立天文台】 15
3.5.5. データを受信する側の事例　【A.3. 中村委員】 15
【高エネルギー加速器研究機構】 15
3.5.6. ポリシーやガバナンス事例　【A.5. 渡邊委員】 16
【広島大学】 16

3.5.7. 各研究機関の分析 19
4. クラウド時代のデータ共有技術における課題(受容も含む) 22
4.1. データへのアクセス方法 22
4.1.1. 検討しておくべきデータアクセス性能に対する要請 22
4.1.2. アクセス方法に対する要請 22
4.2. データ移動量（オンプレミス→クラウド、クラウド→オンプレミス）に対する要請 23
4.2.1. 課金システム（データ総容量、アクセス頻度、アクセス制限）について 23
4.2.2. 調達フェーズと運用フェーズにおいて検討する必要のある事項 23
4.3. データ共有にかかるリスク対応 24
4.3.1. ガバナンス・コンプライアンス・国際法律 25
4.3.2. 連携する両者のポリシー整合性 25
4.3.3. 認証・証跡 25
4.3.4. データ保護・アクセス権限・特権 ID 25
4.4. データ共有時の利便性や管理の容易さを向上させる技術 26
4.4.1. データへのタグ付け（LOD、データジャケット、メタデータ）、VPX 26
5. データ共有技術におけるオンプレミス・クラウド利用に向けたチェックリスト 27
5.1. 共有データ種別・目的・方法 27
5.1.1. データ種別 27
5.1.2. データアクセス 27
5.1.3. 総データ量 28
5.1.4. データアクセス量 28
5.2. 可用性 28
5.2.1. サービス時間（計画的、事前通知、サービス終了リスク） 28
5.2.2. 稼働率・継続性 28
5.2.3. ディザスタリカバリ 28
5.3. システム管理・監視 28
5.3.1. リソース管理オペレーション 28
5.3.2. システム監視（ログへのアクセス） 28
5.4. 性能 29
5.4.1. スループット性能 29
5.4.2. レスポンス性能 29
5.4.3. 処理能力【件数】 29
5.5. データ管理 29
5.5.1. データ移行方法 29
5.5.1. バックアップ（容量、バックアップ時間、保存期間） 29
5.5.2. データ消去（ポータビリティ/委託） 30
5.5.3. 暗号化（データ漏えい対策、データ破損・改ざん検知） 30
5.6. セキュリティ 30
5.6.1. ユーザ管理・認証 30
5.6.2. 証跡・データ保護 31
5.7. クラウド・オンプレミス検討の方向性 31
6. 結び 33
A. 各委員のシステム紹介 34
A.1. 早戸委員 34
A.2. 藤田委員 47
A.2.1 概要 47
A.2.2 JAXA のデジタルデータ運用 49
A.2.3 データ管理の課題 50
A.3. 中村委員 51
A.4. 白崎委員 69

A.5. **渡邉委員** ····· 85
B. **クラウド時代のデータ共有技術紹介** ····· 97
B.1. **クラウドストレージ、ストレージソフトご紹介** ····· 97
B.2. **データ配布時の高負荷対策** ····· 105
B.3. **クラウドストレージ紹介** ····· 110
B.4. **クラウド時代のデータ共有技術** ····· 114
B.5. **「認証サービス」ご紹介** ····· 123
B.6. **ブロックチェーンで実現するデータ流通・利活用** ····· 133

表　目　次

表 1 一般的なクラウドの利点（AWS の場合） ····· 6
表 2　オンプレミスとクラウドの考慮点の比較 ····· 7
表 3　媒体の価格容量比率 ····· 14
表 4　法人文書重要度分類表 ····· 18
表 5　各研究機関のデータ分析 ····· 20
表 6　各研究機関のツール分析 ····· 21
表 7　クラウドに向かない状況一覧 ····· 32
表 8　オンプレミスやハイブリッドクラウドの選定基準一覧 ····· 32

図　目　次

図 1　データ容量に対するコスト ····· 9
図 2　データ転送量に対するコスト ····· 9
図 3　Erasure Coding の仕組み ····· 12
図 4　クラウドサービス利用基準 ····· 19

商標、登録商標について
・PRIMERGY、ETERNUS、FEFS は、富士通株式会社の登録商標です。
・その他一般に、会社名、機関名、製品名は各社や各機関の商標または登録商標です。
・会社名、システム名、製品名などには必ずしも商標表示（TM, ®）を付記しておりません。

1. はじめに(WG 活動概要)

WG 発足の背景や活動方針について以下に示す。

1.1. WG 発足の背景

クラウドサービスは、利用者が手元のコンピュータで利用していたストレージやソフトウェアを、ネットワーク経由で、利用者に提供するものである。利用者側は、最低限の環境（PC や携帯情報端末などのクライアント、その上で動く Web ブラウザーおよびインターネット接続環境など）を用意するだけで、どこからでも、また、どの端末からでも、さまざまなサービスを利用することができる。また、利用者は、クラウドサービス事業者が保有するコンピュータの処理能力を、柔軟に必要な分だけ利用することができる。

これまで、利用者はコンピュータのハードウェア、ソフトウェア、データなどを、自身で保有・管理し利用していた。しかしクラウドサービスを利用することで、機材の購入やシステムの構築、管理などに必要となるさまざまな手間や時間の削減が可能となり、業務の効率化やコストダウンを図ることができるようになる。

クラウドサービス（特に、以下の分類でいう IaaS）では、主に仮想化技術が使われている。仮想化技術とは、実際に存在する１台のコンピュータ上に、ソフトウェアの働きにより、何台もの仮想のコンピュータがあるかのような働きをさせることができる技術である。逆に複数台のコンピュータをあたかも 1 台であるかのように利用することもできる。この技術により、利用者から見て、インターネットの先にある自分が利用しているコンピュータの形態が実際にどうなっているのか見えづらいことを、図で雲のかたまりのように表現したことから、「cloud＝雲」という名称がついたと言われている。

クラウドサービスは、企業が情報資産を管理する手段として急速に普及している。また、個人が利用するインターネット上のさまざまなサービスも、利用者が意識するかどうかにかかわらず、クラウドサービス上で稼働するようになっている。

クラウドサービスを利用する場合には、データや認証情報がクラウドサービス事業者側のサーバに保管されること、また、データなどがインターネットを介してやりとりされることなどから、十分な情報セキュリティ対策が施されたクラウドサービスの選択が重要であるということを理解した上で利用することが大切である。

（出典：http://www.soumu.go.jp/main_sosiki/joho_tsusin/security/basic/service/13.html）

1.2. 活動方針

近年の共同研究は国際化が進み、大規模なデータを国内外で共有、相互利用している。これに伴い、大量のストレージやCPU資源が必要となったことから、中核となる研究機関のみで必要な計算機資源を保持し、ユーザに提供することは困難となり、複数の研究施設の持つ計算機資源を統合的に活用することが必要となってきた。このため、複数機関で大規模なデータの共有を効率的に行うことが一つの課題となっている。さらに近年では、公的資金を用いて行われる実験や研究などの成果を社会に還元することが必須となっており、一般公開する必要のあるデータも大規模となってきている。

これまでに、研究機関内など閉じた環境においてデータを共有する技術は確立されている。しかしながら、国内外に散らばる複数の研究機関が参画する場合のデータ共有に必要な技術（ソフトウェア等）は、未だ簡単に導入、運用できるものとはなっていない。近年特に問題となることが多いセキュリティ問題に対処することも必須であり、ネットワーク接続の制限やユーザの管理などにも配慮する必要がある。こういった様々な制約の中で、効率的かつ十分な性能を持ったデータ共有システムが求められている。

こういった現状を鑑み、本WGでは、いくつかの研究分野におけるデータ共有への要求を整理すると同時に、現在のデータ共有基盤技術を調査し、要求に適した技術を提示するとともに、既存技術における課題と必要となる技術開発の方向性について議論を行うことを目指す。

1.3. 活動内容

WG前半では、様々な分野におけるデータ共有の具体的な事例と要求、ならびに課題を集め、これらを整理するとともに、現在のデータ共有に関連する技術・規格や将来動向を調査した。

WG後半では、調査したデータ共有技術と、整理された要求や課題との関連を把握し、各事例に利用可能なデータ共有技術を検討するとともに、開発が必要となる技術について検討することを通し、今後のデータ共有の在り方について報告書にまとめた。

本WGのまとめが、国や分野を超えたデータ共有に関する検討の一助となることを期待する。

1.4. 活動期間

2017年1月～2019年4月

1.5. WG メンバー

		氏名	機関/所属(2018 年 12 月 31 日現在)
会員	担当幹事	井口　　寧	北陸先端科学技術大学院大学
	推進委員	○早戸　良成	東京大学宇宙線研究所
		藤田　直行	宇宙航空研究開発機構
		中村　智昭	高エネルギー加速器研究機構
		白崎　裕治	国立天文台
		渡邉　英伸	広島大学
賛助会員(富士通)	推進委員	○阿部　孝之	富士通(株)
		栄元　雅彦	富士通(株)
		坂口　吉生	富士通(株)
		甲斐　俊彦	富士通(株)
		宮本　巧輝	富士通(株)
		篠原　　誠	富士通(株)
		鈴木　康介	(株)富士通研究所
		浅野　典志	オブザーバ　富士通(株)
		荒井　亮太	オブサーバ　富士通(株)

○印：まとめ役

1.6. 活動実績

■第1回会合 ： 2017年2月1日(水) 富士通(株)本社
- WG活動計画(活動方針、活動の進め方/スケジュール感、成果)の確認
- 各委員の課題意識の共有と意見交換、活動で取り上げる課題の検討

■第2回会合 ： 2017年6月9日(金) 富士通(株)本社
- 各会員における運用や課題などの共有および議論
- データ共有に関するテクニカルなトレンドなどの紹介および議論

■第3回会合 ： 2017年10月6日(金) 富士通(株)本社
- 各機関におけるデータ分析・ツール分析の結果共有および議論
- WAN高速化技術の紹介および議論

■第4回会合 ： 2017年12月22日(金) 富士通(株)本社
- クラウドストレージゲートウェイ機能の紹介および議論
- クラウドサービス使用時に必要な対策(ユーザ認証・セキュリティ)の議論

・クラウドストレージの比較とデータ配布時の高負荷対策の議論
・各機関におけるデータ分析・ツール分析の結果確認と、まとめの方向性の検討

■第5回会合 ： 2018年4月3日(金) 富士通(株)本社
・ブロックチェーンを利用したデータ流通・利活用技術の紹介および議論
・成果報告書の作成方針、掲載項目、構成の検討

■第6回会合 ： 2018年7月17日(火) 東京大学宇宙線研究所(神岡)
・クラウドサービスとオンプレの比較の紹介および議論
・クラウドストレージゲートウェイの動向の紹介および議論
・富士通クラウドにおける認証サービスの紹介および議論
・ストレージの動向の紹介および議論
・クラウドストレージBoxの紹介および議論
・成果報告書の目次、掲載内容、執筆分担の検討

■第7回会合 ： 2018年10月3日(水) 富士通(株)本社
・成果報告書の目次、詳細内容、執筆分担、スケジュールの検討、確定

■第8回会合 ： 2018年12月14日(金) 富士通(株)北陸支社
・成果報告書の目次、詳細内容検討、執筆進捗確認、スケジュールの確定

1.7.　参考文献

今後のデータ共有技術のありかたを検討するにあたり、クラウドサービスの利用の参考となる資料を以下に示す。

題目　：　大学・研究機関のためのクラウドスタートアップガイド
更新　：　Ver.2.0 (2018/11/19)
　　国立情報学研究所クラウド支援室
URL　：　https://cloud.gakunin.jp/dist/pdf/startupguide-v2.pdf

2. クラウド時代のデータ共有技術の現状と問題

近年、大規模な事業を手掛ける企業においては、大規模なデータを保持、活用する技術が必要である。例えば、7500 万件の契約を持つ携帯事業者が、各契約あたり 5GB 程度のデータを保持するなら、全体で約 400PB のデータを扱うこととなる。また、1 万人の社員を持つ企業において、社員一人当たりが 1TB の領域を利用すれば、全体では 10PB が必要となる。ここまで大規模でなくとも、企業で扱われるデータは日々増加しており、実際に PB 級のデータ領域を要求する調達も少なからずある。一方で、企業内で大規模なデータを管理・利活用する技術を持つことは容易でない。このことは、総務省の調査からも見て取ることができる。2017 年度にクラウドサービスを一部でも利用している企業の割合は 56.9%となり、前年の 46.9%から大幅に増加している。これは、企業が大量のデータ管理や、その利用を、クラウドサービスの導入によって達成しようとしていることを示している。

そこで本 WG では、まず研究機関や大学、企業が必要とする大容量のデータ管理において、一般に提供されているクラウドサービスを調査、利用可能なケース、不適切なケースを検討した。

2.1. 昨今の要請と問題（概要説明）

利用されているクラウドサービスの内訳をみると、「ファイル保管・データ共有」（51.2%）が最も多く、次いで「サーバ利用」（47.6%）、「電子メール」（46.3%）の順となっている。また、クラウドサービスを利用している理由としては、「資産、保守体制を社内に持つ必要がないから」（45.2%）が最も多く、次いで「どこでもサービスを利用できるから」（34.8%）、「安定運用、可用性が高くなるから（アベイラビリティ）」（32.6%）が主なものとなっている。さらに、クラウドサービスを利用する企業のうち、「非常に効果があった」または「ある程度効果があった」として、効果を実感している企業の割合は 85.2%であり、多くの企業がクラウドサービスの利用による効果を実感している。中小企業庁が発表している『中小企業白書（2013 年版）』によると、企業がクラウドサービスを利用する理由として、「初期コストが安い」「導入までの期間が短い」「技術的な専門知識がなくても導入できる」をあげている。

（出典：http://www.chusho.meti.go.jp/pamflet/hakusyo/H25/PDF/0EHakusyo_part2_chap1_web.pdf）

一方で導入後の評価をみると、従業員数の数によりそのポイントが分かれている。従業員数の少ない企業ほど「初期コストが安い」という評価が低い。これは「IT 専任者もしくは、部門やガバナンスが整備されているかどうか」という点ではないかと推測される。 IT 専任者や部門が存在しない場合は「技術的な専門知識が不要(外部に委託)」という点がメリットになり得ると考えられる。逆にクラウドサービスを利用していない理由としては、「必要がない」（39.5%）が最も多く、次いで「情報漏えいなどセキュリティに不安がある」（38.1%）、「クラウドの導入に伴う既存システムの改修コストが大きい」（27.6%）が多くなっている。

（出典：http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/html/nd252140.html）

最近の国内や海外の状況はクラウドサービス利用をやめ、オンプレミスに戻すケースも見受けられている。システムをオンプレミスに戻す決定を行った理由について、セキュリティ、通信費用などのコストの問題、管理や信頼性に関する不満などが挙げられている。

データ共有という観点から興味深いのは、最もよく利用されているサービスがオンラインストレージであるということである。本サービスは、多くのクラウド業者が提供しており、主要なものとして Dropbox、Box、Google Drive、OneDrive、iCloud Drive、MEGA、Amazon Cloud Drive、Tera クラウド、SpiderOak、MediaFire などがある。一方で、業者(サービス)ごとに想定利用形態が異なり、これに伴って仕様や価格も大きく異なるため、選定においては注意深い検討が必要である。また、データ共有時には漏えいや改ざんなどを防ぐこと、すなわちセキュリティへの配慮は重要であり、この点での評価も必須となる。このように、組織の大きさにも影響されるクラウドサービスの利活用を見極めるためには、一般的なガイドラインの他にデータ共有技術におけるガイドラインも必要と考える。

このため本 WG ではデータ共有の観点でオンプレミス、クラウドの判断基準を議論し明らかにしたい。

2.2. クラウド技術の動向

一般的なクラウドサービスとオンプレミスサービスを列挙する。

2.2.1. クラウドのサービス～オンプレミスのサービス

クラウドサービスの分類としては、以下のカテゴリが存在しており、必要に応じてサービスに追加できる。比較対象も必要と考えオンプレミスにおいては、以下機能を必要に応じて追加できることを前提とした。

・SaaS(Software as a Service):クラウド内のサーバ上で実行されるアプリケーション（ソフトウェア)を提供
・PaaS(Platform as a Service)：クラウド内で動くアプリケーション（ソフトウェア）の実行環境を提供
・IaaS(Infrastructure as a Service):クラウド内の（多くの場合は仮想化された)計算機を提供
・DaaS(Desktop as a Service):ユーザが使用するデスクトップ画面を提供
・DSaaS(Data Storage as a Service) ストレージ容量およびアクセス・管理機能を提供
・DBaaS (DataBase as a Service):データベースのアクセスおよび管理機能をオンマンドに提供
・IDaaS(IDentity as a Service):アイデンティやそのアクセス権限管理の機能を提供

このように、多種多様なサービス形態がある中、サービス提供までの時間、Quality、価格（利用・管理コスト含む）の面と技術動向の面において情報が散見される状態である。

詳細なサービス一覧としては以下である。クラウドと異なり、オンプレミスは優先度の高いサービスを選定後にサービス提供する制約を前提とした。

・コンピューティング：仮想サーバ、コンテナ、Auto Scaling、ロードバランシング
・ストレージと配信：オブジェクトストレージ、ブロックストレージ、アーカイブストレージ、コンテンツ配信
・データベース：RDB、NOSQL、キャッシュ、データウェアハウス
・ネットワーク：インタコネクト、仮想ネットワーク、専用線、ロードバランシング
・セキュリティ・認証：ID 管理、アクセスコントロール、キーの管理
・分析：Hadoop、リアルタイム、機械学習、データウェアハウス、Data Pipelines
・アプリケーションサービス：キューイング、ワークフロー、ストリーミング、メール、通知、検索、API
・開発用ツール：ソースコード管理、デプロイ、継続的デリバー
・管理ツール：モニタリング、リソーステンプレート、使用量とリソースの監視、サービスカタログ

2.3. 大規模データ・大量データを活用する分野における課題

近年、スーパーコンピュータの演算性能の向上に伴って、高統計のシミュレーションデータの生成やデータ解析処理のスループットが向上したこと、ならびに、観測・測定機器の高精度化や多チャンネル化に伴ってデータの生成量は大幅に増加し、扱う必要のあるデータが大規模化した。一方で、利用する機器のコストは少数の組織や場合によっては一国でまかなえないものや国際的な共同研究が加速している状況から、大規模データ・多量データを有する研究機関は、組織を超えた共有がこれまで以上に重要となった。また、公的資金を用いた研究に対しては、そのコストを負担した社会への還元として、データの公開が求められるようになった。

一般社会においては、官公庁や企業がこれまで個別に収集・管理してきた様々なデータを統合的に扱うことで、生活の利便性向上なども含めた様々な問題の解決に役立つことが認知されるようになった。このため、様々な分野・領域でデータの集積と大規模化が急速に進んでいる。

これらは、計算機資源の共有・管理ならびに大規模データの保存・管理・共有において、クラウド技術の急速な発展が実現したことによって実現された。一方で、この過程において大規模なデータを取り扱うにあたっての様々な課題も認識されるようになった。特に大規模データの共有においては、データアクセスに求められる性能やアクセスポリシーが業務や事業ごとに異なるため、用いるべき技術も異なってくる。そのため、ユーザ、管理者、メーカ・ベンダ等の技術者が目的を共有し、それぞれの視点から課題を認識することで、初めて適切な技術選択や、開発すべき技術課題の明確化、そして課題への対応を実現することができる。本 WG の議論を通して、大規模データを共有したい場合にクラウドサービスを利用することが適切なのか、オンプレミスのシステムを構築する必要があるのかを判断するためには、どのようなことを考慮すべきか、最低限の基本的な情報を得ることもなかなか困難であるという共通認識となった。そこで、本まとめではクラウドサービスとオンプレミスそれぞれのもつ特性を整理し、技術、業務、調達などの観点から、クラウドとオンプレミスの選択に役立てることが可能となる資料とすることを目指すこととした。

2.3.1. クラウドサービスおよびオンプレミスシステムそれぞれの利点と欠点

ここでは、一般的なグラウドサービスとして AWS の利点をまとめた。

表 1 一般的なクラウドの利点（AWS の場合）

	自社データセンター	AWS クラウド
初期コスト	必要	不要
利用コスト	余剰リソースにもコスト発生	実際の利用分のみのコスト負担
災害対策コスト	高い （物理的に異なる場所のデータセンター向けにハードウェア発注、設定、運用を物理的に行う必要あり）	安い （稼働していない時間は無料）
インフラバージョンアップ/システム更改時のコスト	必要	不要
インフラ調達期間	数週間～数か月	数時間
データセンターのロケーション変更	長時間必要	短期間で可能
サーバスペック変更の容易性	時間とコストがかかる	余計なコストなしにすぐに対応可能
自社ネットワーク内での利用	可能(考慮不要)	可能
専用線による自社内サーバとの接続	可能	可能
第三者機関認証	自身で取得	AWS 側で取得済 （SOC 1/SSAE 16/ISAE 3402, SOC 2, ISO 27001, PCI DSS レベル 1, HIPAA, ITAR, FISMA Moderate, FIPS 140-2, CSA）

2.3.2. オンプレミスとクラウドサービスの違い～システム管理者的視点より～

オンプレミスとクラウドの考慮点の比較の結果を「表２　オンプレミスとクラウドの考慮点の比較」に示す。

表 2　オンプレミスとクラウドの考慮点の比較

	オンプレミス	クラウドサービス
調達	ハードウェア・ソフトウェア・サービス（メンテナンス）について仕様を策定する必要がある。	サービスに対しての仕様を策定すればよい。 （サービスとして提供されていない要請に対しては、別途開発等が必要なことはオンプレミスと同様である。）
調達物	インフラストラクチャ（ハードウェア・ソフトウェア）を調達する。設置期間なども必要である。	必要に応じてネットワーク回線を調達し、サービスを契約すればよい。
設計・開発	要求仕様に応じて柔軟な設計が可能、一方で、相応のコストが必要になる場合もある。	提供されているサービスを要求仕様にあわせて利用するための設計・開発は個別に必要である。
主要な経費	インフラストラクチャ費用 維持管理費用（電気、空調、ネットワーク回線、場所等）	サービス利用料（含データアクセスチャージ） ネットワーク回線費用
維持・保守費用	電気代、設置費用（空調システムなど）、場所利用料、保守費用などを別途計上しておく必要がある。	サービス料金に含まれる。
資産管理	インフラストラクチャの資産管理は必要である。	インフラストラクチャ資産管理は不要である。
セキュリティ	開発、運用時に考慮や対策が必要である。	サービスに含まれるものと含まれないものがあるため、注意が必要である。
そのほか	契約期間によるが、ハードウェア・ソフトウェアとも陳腐化しやすい。	細やかな要件やサービス事業者が保有するシステム性能を超えた要請には対応できない。 契約終了時に蓄積したデータをすべてとりだす場合は、必要なアクセス時間やアクセスチャージを事前に評価しておく必要がある。 利用する組織の予算上限を超えた場合、データへのアクセスが不可能となる可能性がある。

2.3.3. ユーザがクラウドサービス利用時に期待する点

特に会員の視点でクラウドサービスを利用することで効果を期待する事項として、以下のようなものがあった。

・研究・業務に必要な情報のやりとりを場所や時間に関係なくできること
・プロジェクト（組織横断）で研究・業務を自由に設計できること
・リージョンの壁が少なくなり、情報や研究・業務の格差が狭小化できること
・労働時間や場所やデバイスの制約が少なくなり、多様なワークスタイルを実現できること
・現場の技術・運用ノウハウを容易に継承できること
・必要なリソース（人・資金除く）をオンデマンドで利用できること
・法令もしくはカバナンスの観点の対策コストを抑制できること
・Auto Scaling に対応した仕組みであること
・投資に対して保証（犯罪・コスト削減）があること
・クラウドサービス事業者の倒産リスクに対して、対策を講じられること

3. 大規模・大量データの共有をささえる技術と事例

大規模・大量データの共有をささえる技術と事例について、以下に示す。

3.1. サービス動向

サービスの動向について、以下に示す。

3.1.1. オンラインストレージ（ファイル保管・データ共有）

近年オンラインストレージが急速に普及している。オンラインストレージとは、クラウド上に用意されたストレージを利用者に貸し出し、データのバックアップやファイル共有を行うサービスである。サービスの普及に伴い価格が下がったことで、サーバの運用コスト削減の目的や、台風や地震などの自然災害に対する BCP の施策として導入する個人や企業が増加している。

以下に、個人や企業で導入が進んでいるオンラインストレージの一例を紹介する。

(1) Dropbox

米 Dropbox 社によって運用されている世界最大級のオンラインストレージサービスで、全世界での利用者数は 5 億人以上である。ブラウザー経由で利用以外にも専用のクライアントソフトを導入することでデスクトップからエクスプローラ等で直接ファイルを操作することも可能になっている。また専用のフォルダに保存したデータが自動でクラウド上に保存されるなど使用者が特別な知識が無くても容易に使えるようになっている。

(2) OneDrive

米 Microsoft 社が提供するオンラインストレージサービスで同社の Office 365 の機能の 1 つとして提供されている。Office のデータを保存先として OneDrive を直接指定することができ、作成したデータファイルを直接クラウド上で保存でき便利になっている。全世界での利用者数は 2 億 5000 万人以上である。

(3) Box

米 Box 社によって提供されている企業向けに特化したオンラインストレージサービスで、利用者数は 4,100 万人以上ながら、導入企業数は 8 万 5000 社以上となっており、強固なセキュリティ面から Fortune 500 の約 7 割の企業が利用している。セキュリティ面以外の特長として SSO による外部サービス連携や 120 種類以上の拡張子に対応したプレビュー機能や各種 PC、モバイルデバイスから利用することができ Box とブラウザーがあればどこからでもファイルを閲覧可能になっている。

上記とは別に大手クラウド 4 社が提供しているクラウド上でのオンラインストレージのサービス料金をまとめてみた。サービスの料金は大きく、データ容量・データ転送量・リクエスト量から算出される。

(1) データ容量

データ保管量から算出される。料金は GB 単位で課金され、データが一定量を超えると単価が下がるタイプと固定のタイプがある。

(2) データ転送量

インターネットからクラウドへ送受信するデータ量から算出される。料金は GB 単位で課金され、データが一定量を超えると単価が下がっていく。

(3) リクエスト量

書き込みや読み込みなどストレージに対して行う操作回数から算出される。一定操作回数（1,000 回、10,000 回）単位で課金される。大きく分けると更新系と参照系で料金が異なり、10,000 回単位での価格は更新系：$0.05、参照系：$0.004 が目安になる。

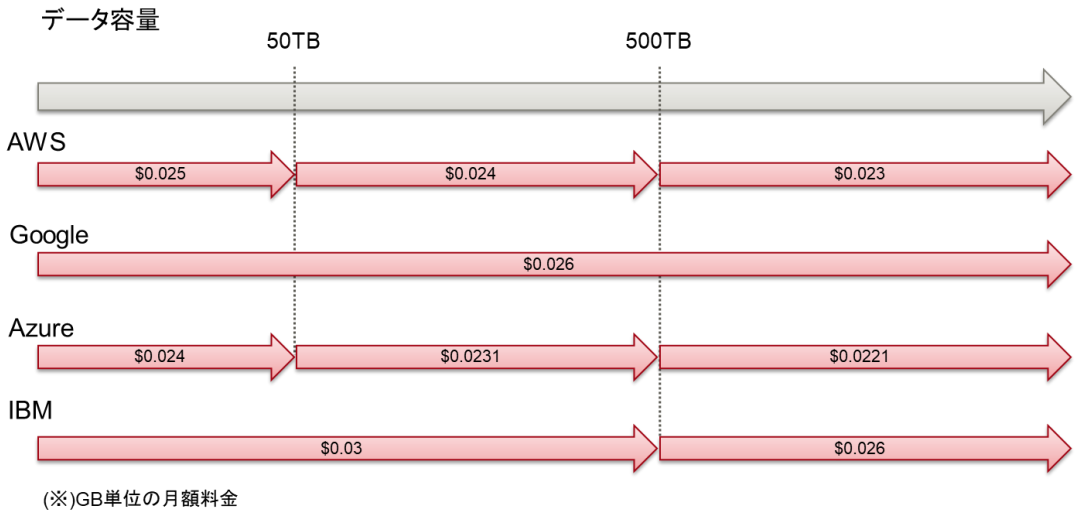


図 1　データ容量に対するコスト

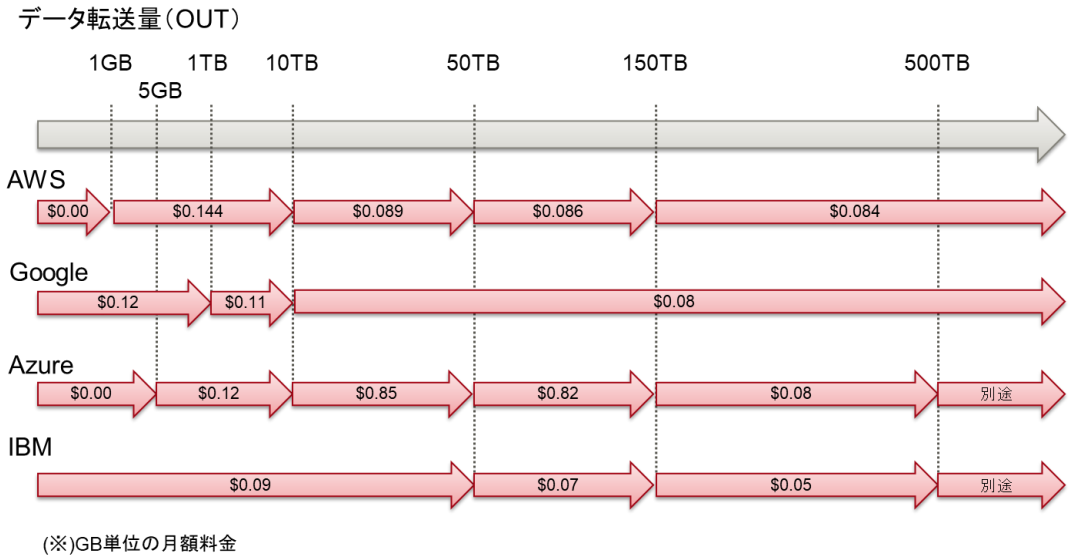


図 2　データ転送量に対するコスト

3.1.2. グループウェア（メール・ドキュメント共有・ファイル共有・スケジュール共有）

グループウェアは組織内の情報共有のためのシステムソフトウェアである。ネットワークに接続されたユーザ間での情報交換や共有、スケジュール管理などの業務に利用され、業務の効率化を目指すもので、最近はクラウドサービスとしての提供により導入・運用のハードルが下がってきている。

(1) Microsoft Office 365

Microsoft 社によって提供される SaaS 型クラウドサービスで、メールおよびスケジュール共有(Exchange Online)、ファイル共有(SharePoint Online)、オンライン会議(Skype for Business Online)などからなる。また、従来 PC で幅広く利用されてきた Microsoft Office アプリケーションのライセンスを月・年単位でのサブスクリプション契約ができる。

(2) G Suite

Google 社によって提供されるグループウェアで、メール(Gmail)、スケジュール共有(カレンダー)、ファイル共有(Google ドライブ)、オンライン会議(ハングアウト)からなる。スマートフォンなどから一般的に無償で個人利用されているサービスに対して機能や容量を拡張して提供している。

(3) Indico

CERN で開発された Web ベースの会議支援システムで、会議で使用する資料や参加者の管理、部屋の割り当てなど会議で必要なワークフローが整備されており、会議終了後にそのまま会議記録として保存して参照することができる。ソフトウェアは OSS として公開されており、100 箇所以上の研究機関に導入されている。

3.1.3. プロジェクトホスティングサービス

オープンソースソフトウェアの開発者向け機能を提供するホスティングサービスで、コードホスティングの他にソフトウェア開発に必要な、コードレビュー、障害管理システム、wiki 機能が備わっている。

(1) GitHub

GitHub 社が運営しているサービスで、存在するプロジェクトは 8000 万件以上。使用可能なバージョン管理システムは Git。2018 年 6 月に Microsoft に買収された。

(2) SourceForge

Slashdot Media 社が運営しているサービスで、存在するプロジェクトは 18 万。使用可能なバージョン管理システムは CVS や Subversion。

3.1.4. 認証サービス

オンプレミスのシステムとクラウドのサービスを併用する場合、利用者のアカウント情報の取り扱いが課題になる。システムやサービス毎に異なるアカウントを使用する場合、利用者視点で考えると別々のアカウントを使い分ける必要があり不便であるし、管理者視点で考えても別々にアカウント情報を管理する必要があり運用コストが増加する。認証サービスとは、認証統合の課題を解決し、利用者、管理者共に負担を軽減し、利便性を向上させるものである。

富士通では、Fujitsu Cloud Service for OSS の 1 機能として認証サービスを提供しており、その機能の特長は以下の通りである。

(1) 利便性の向上（シングルサインオン：SSO）

標準規約に従った認証方式(SAML、OpenID、Connect 1.0)の採用により、他のクラウドサービスと連携し、シングルサインオンシステムが構築できる。

(2) セキュリティリスクの軽減（多要素認証）

生体認証（手のひら静脈・指紋）の利用により、安全な認証システムの構築が可能になる。

(3) ID 管理連携（API を提供）

認証用の ID 情報に関して、ID 管理元からの API の利用により連携可能になる。

(4) 監査と認証のログ取得（API を提供）

セキュリティ監査に必要となるアクセスログ（追跡ログ）を API の利用により取得可能になる。

3.2. ソフトウェア動向

ソフトウェアの動向について、以下に示す。

3.2.1. ファイル共有

パブリッククラウドが提供するサービスに対するセキュリティなどのリスクを考慮して、プライベートクラウドやオンプレミスのシステム上にもオンラインストレージを構築可能である。

(1) ownCloud

Linux, Windows など複数の OS 上でクライアントが動作するファイル共有のオープンソースである。オンプレミスのサーバやクラウド上の仮想サーバで動作可能で、既存の Active Directory/LDAP と連携できるため、認証サーバで管理されている ID・パスワードで認証をかけることができる。また同期クライアントを利用することにより同期フォルダ内のファイルを自動で ownCloud 上にバックアップすることもできる。その他の機能としては、全文検索、ワークフロー、ファイル暗号化、監査ログなどが備わっている。

(2) Nextcloud

ownCloud の創始者 Frank Karlitscheck 氏と主要開発者が中心となって立ち上げた OSS で細かな違いはあるものの ownCloud の機能を追従して実装している。

(3) GitLab

Git ベースのソースコード管理機能、マージリクエスト、レビュー機能などを備えたオープンソースのソフトウェア開発支援環境で、「GitHub」のようなサービスをクローズ環境に独自構築できる。

3.2.2. ファイルシステム

本WGの会員のシステムでは得られたデータを外部の利用者向けに公開している。このため、外部利用者からのダウンロード要求で発生する大量のread要求を高速に処理できるファイルシステムが必要になる。単一サーバの場合、1サーバのハードウェア性能（ディスク、ネットワーク、CPU、メモリなど）で限界となるため、複数サーバに要求を並列に分散させる並列ファイルシステムが必要である。更に、1 つのデータを複数のサーバに分散して格納することで、特定のサーバへのアクセス集中を回避することも必要である。代表的な並列ファイルシステムにおいて大量のread要求を分散処理するための機能を以下に挙げる。

(1) Google File System

システム内に在る Master ノードがファイルデータを記録する Chunk サーバを管理し、クライアントは Master ノードに問い合わせ、自分がアクセスするファイルデータを記録している Chunk サーバを決める。ファイルデータは複数の Chunk サーバに複製されるため、クライアントはネットワーク上一番近い Chunk サーバからデータを read することができる。

(2) GlustreFS

「Replicated Volume」で構成された場合、複数のサーバの volume にファイルデータの複製を配置することができる。クライアントはDHT（Distributed Hash Table）を使ってどのサーバの volume に配置されているかが分かるため、サーバに問い合わせして早く応答が返ってきたサーバからファイルデータを read する。

(3) Lustre

Lustre 2.11 でサポートされたFLR1(File Level Replication phase 1)の機能でファイルデータの複製を複数の volume に配置することができる。クライアントはMDS(Meta Data Server)に問い合わせ、ネットワーク上近いサーバからファイルデータを read することができる。

上記の他に分散ファイルシステムとして広く使われているソフトウェアを以下に紹介する。

(4) OneFS

Dell EMC 社製ストレージ Isilon 上で動作するファイルシステム。従来のファイルシステム、ボリューム管理、RAID の 3 層を統合し、複数のストレージノードによる単一ファイルシステムを実現している。

拡張性に優れ、運用中にストレージノードを増設することでシームレスに容量を拡張でき、追加後はデータのリバランスを自動的に行う。これによりハードウェア更新を無停止で行える。

(5) BeeGFS

独 Fraunhofer 社が開発したオープンソースの並列分散ファイルシステムで、かつては FhGFS として知られていた。複数のサービスを同一サーバ上で動作させることが可能であり、オンデマンドで動的にファイルシステムを生成する BeeOND も提供している。

(6) SpectrumScale(旧 GPFS)

IBM 社によって提供される並列分散ファイルシステムで、Linux,AIX および Windows で使用することができる。高いスケラビリティを持つとともに、Hadoop や OpenStack への対応や拠点間の同期やキャッシングを行う Active File Management(AFM)機能などを持つ。

データアクセスを分散するためにデータを複製すると、ボリューム破壊等でデータが失われた場合でも正常なボリュームからデータを読み出すことができる。データの複製は冗長面では単純かつ強力な方法であるが、複製先のボリューム数を増やすと保管コストも単純に増える。このためデータを複製するのではなく、Erasure Coding を用いてデータに冗長を持たせているファイルシステムが広まってきている。

Erasure Coding では元のファイルから複数のデータブロックとパリティブロックを作成し、複数のボリュームに分散して格納する。実際にボリュームが故障した場合、残りの正常なデータブロックとパリティブロックから元データを復元することができる。先に書いた OneFS は Erasure Coding を実装している代表的なファイルシステムである。Erasure Coding は複製と比べて Write 時のコストが少ないのが特徴で、Write 性能を落とさずにシステムの可用性を高めることができる。

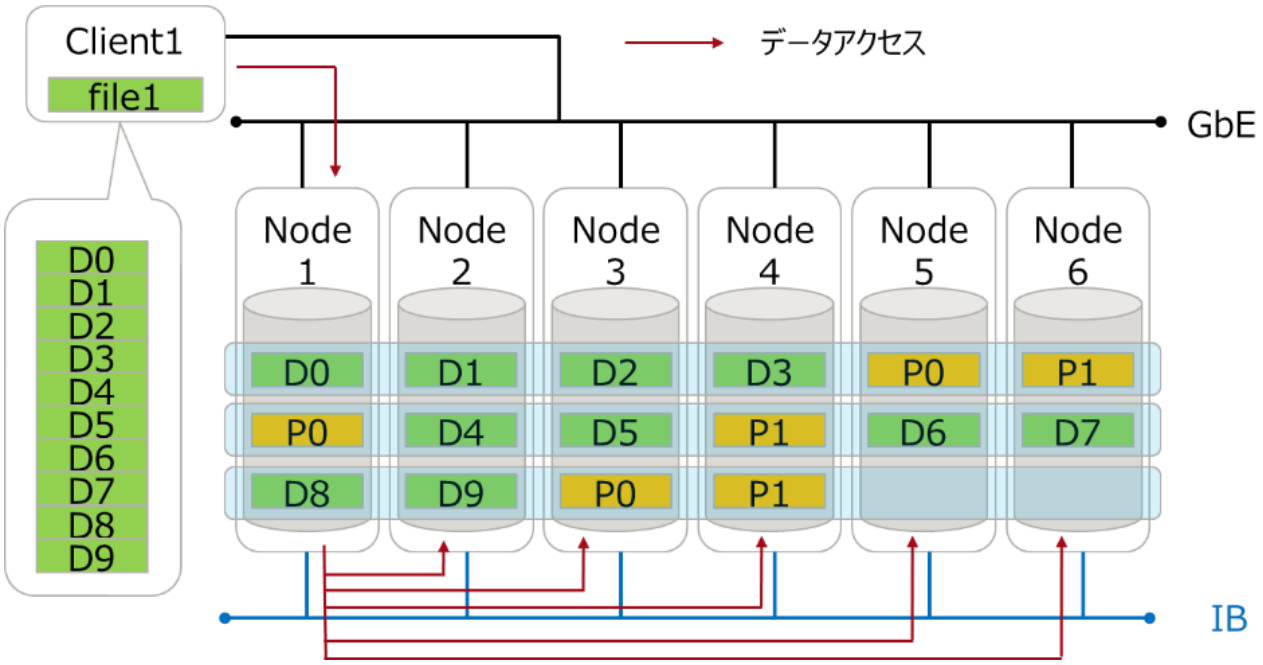


図 3　Erasure Coding の仕組み

今後、HPDA（High Performance Data Analytics）分野や AI 分野への HPC 技術を適用し、膨大なデータ解析が今後進んでいくと考えられる。このためクラウドとオンプレミスの長所を活かしたハイブリッドクラウド化が進んでいくと予想され、ストレージ関連各社はハイブリッドクラウド向けの製品を発表している。ハイブリッドクラウド向けのファイルシステムの例として以下を挙げる。

(1)Elastifile　Cloud　File System

Elastifile 社が販売しているハイブリッドクラウド用のファイルシステムで、ベアメタルサーバ、コンテナ、VM のどれからでも同じファイルを参照できる。拠点間でデータ共有する場合、参照先のデータ管理元から拠点のファイルシステム上に動的にデータを転送することにより利用者は特にデータ転送を意識することなくデータを参照することができる。

(2)Weka　IO

Weka.IO 社が販売しているファイルシステムでクラウドやオンプレミスの SDS(Software - Defined　Storage)上にあるデータをクライアントに近い SSD に自動転送して利用する。ファイルシステムは Linux コンテナ内で動作させ同一ノード内の他のプロセスから独立させている。VFS 用のドライバが利用でき、POSIX のインターフェイスを使用することもできる。

3.2.3.　オブジェクトストレージ

従来のファイルシステムはデータをディレクトリ構造で管理しているが、オブジェクトストレージはデータをフラットな空間にオブジェクトとして扱う。データは固有の ID を持ち API でアクセスされる。更新や上書きはできず、新規作成と削除となる。

スケラビリティに優れ PB 級に拡張可能で、無停止で可能である。主に Amazon S3 のようなオンラインストレージに利用されており、データ配信やアーカイブ、バックアップに適している。

(1) Ceph

オープンソースの分散オブジェクトストレージソフトウェアで、複数のアクセス方式に対応している。RADOSGW はオブジェクト単位のアクセスを提供し、RBD はブロック単位、Ceph　FS は POSIX 互換のファイル単位のアクセスを可能とする。

モニタと OSD の 2 種類の構成要素を持ち、モニタが Paxos アルゴリズムによりストレージノードへのオブジェクト配置を管理し、OSD がストレージノード上でのデータの読み書きや冗長化を行う。

(2) Scality RING

Scality 社が開発しているオブジェクトストレージソフトウェアで、データを格納するサーバをリング状

に構成し、Chordアルコリズムに基づいて実装している。コネクタサーバを用意することでNFSやCIFSの接続にも対応している。

また、複数のクラウドに跨るデータオブジェクトを一括管理するマルチクラウドデータコントローラーをオープンソースZenkoとして公開している

(3) minio

米Minio社が開発するオープンソースのオブジェクトストレージソフトウェアで、Go言語で実装されている。Dockerコンテナを起動させるだけでAmazon S3互換のサーバを簡単に構築することができる。

サーバは1台でも実行できるが、分散構成も取れるようになっていて、その場合オブジェクトは分散・複製されて保存される。Erasure CodingやBit rot detectionが実装されており、データ故障時にパリティを使ったデータの読出しや、データが読み出せなくなる前に他の場所へと書き込みなおすこともできる。

3.2.4. ブロックチェーン

近年、仮想通貨で使用されているブロックチェーンとは以下で定義されるものである。

(1) 「ビザンチン障害を含む不特定多数のノードを用い、時間の経過とともにその時点の合意が覆る確率が0へ収束するプロトコル、またはその実装をブロックチェーンと呼ぶ」

(2) 「電子署名とハッシュポインタを使用し、改竄検出が容易なデータ構造を持ち、且つ、該当データをネットワーク上に分散する多数のノードに保持させることで、高可用性及びデータ同一性等を実現する技術を広義のブロックチェーンと呼ぶ」

(出典：http://jba-web.jp)

ブロックチェーンは金融、医療、電力等様々な分野で適用されており、主にデータを安全に共有するという用途で多く使われている。データを安全に共有する技術が広まりデータ利活用が進むと、企業や団体が保有しているデータが有効活用され新たな価値が生み出されていくことに期待されている。

本WGではブロックチェーン技術を応用したデータ共有サービスとして富士通のVirtuora DXのサービスを取り上げた。

Virtuora DXは、ブロックチェーンの技術を拡張し、外部環境にデータを置くことなく企業間のデータ取引を可能にしている。データを共有したコンソーシアム内では、データの利用者はデータジャケットと言われるデータの内容、収集方法、共有条件といったデータの概要情報を参照してそのデータの利用価値を検索することができる。また、データジャケットをテキストマイニングすることにより人が認知できなかったデータのつながりを可視化する機能や、コンソーシアムのメンバー間でコミュニケーションが取れるトークルーム機能があり、メンバー間でのデータを活用した新たなサービスやビジネスの共創の促進への効果に期待している。

詳細は、【B.6.　ブロックチェーンで実現するデータ流通・利活用】を参照されたい。

3.3. ハードウェア動向

ハードウェアの動向を以下に示す。

3.3.1. ストレージ技術動向

近年のストレージ技術の動向としては2つあり、SCM(Storage Class Memory)とNVMe(Non-Volatile Memory Express)である。

SCMは、ノードのメインメモリとして使われているDRAMと、主流になりつつあるSSDで使用されているNAND Flashの中間に位置し、DRAMとNAND Flashの「応答性能」と「容量単価」のギャップを埋めるデバイスである。Intelのコンシューマ向け製品「Intel Optane Memory」に採用された「3D XPoint」や、EverspinのST-MRAMなどの高速・大容量の不揮発性メモリデバイスが登場している。従来の10倍から100倍の応答性を有しており、リアルタイム性が強く求められるような用途で広まっていくと思われる。

またNVMeはFlashなどの高速デバイスの性能を最大限に活かすためにマルチI/Oキュー採用やRDMA対応を行い、I/O設計を最適化している。これによりIOPS 40万（4Kランダムリード)、レイテンシ100マイクロ秒という高速なI/O性能でデバイスにアクセス可能になっている。

詳細は、【B. クラウド時代のデータ共有技術紹介】を参照されたい。

3.3.2. 価格容量比率など

本WG開催時点での、主要な各記録媒体の特長は以下の通りである。

表 3　媒体の価格容量比率

	SSD	HDD	光ディスク	テープ
媒体容量	30TB	12TB	3.6TB	12TB(非圧縮)
最大転送速度	>500 MB/s	200 MB/s	360 MB/s	360 MB/s
ランダムアクセス	○	○	△	×
マイグレーション周期	5年	5年	>50年	7～10年
データ保持期間	<10年	5～10年	>50年	30年
消費電力	小	大	小	小
メディア管理用空調	要	要	不要	要
TB単価(※)	2.5万円	4千円	8千円	2千円

(※) 調査時点

3.4. データ転送技術

データ転送技術の詳細について、以下に示す。

3.4.1. ストレージゲートウェイ

クラウドの台頭に伴い、既設のオンプレミス環境とクラウドを繋ぐ技術として「クラウドストレージゲートウェイ」が注目されている。オンプレミス側にクラウドストレージへの入り口を用意しオンプレミスからクラウドもしくはその逆のデータ移動をシームレスに行うもので、アプライアンスサーバのようなハードウェアとソフトウェアが一体化したような製品やWebブラウザー経由で利用できる製品などが存在する。現在の用途は「クラウドストレージへデータをバックアップする」、もしくは「クラウドストレージ上でデータを共有する」である。

富士通では既存のオンプレミス環境とクラウドとのデータ転送する製品として、Cloud Storage Gateway を販売している。製品の特長は以下の通り。

(1) 汎用的なNASインターフェイス

クラウドへの接続インターフェイスとして汎用的なNASインターフェース（NFS/SMB）でアクセスできるため、既存のバックアップソフトウェア等を継続して使うことができる。

(2) 重複排除および圧縮

バックアップ対象のデータ内で重複するデータを取り除き、データ量を削減する。さらに圧縮機能により、重複排除後のデータ自体を小さくする。

(3) データの暗号化及びセキュア通信

クラウドへ転送するデータはAES256方式にて自動で暗号化してデータを保存する。またクラウドとの接続はHTTPSによるセキュアな通信により第三者の覗き見を防止する。

3.4.2. 高速データ転送

企業WANの現状は以下のような理由で通信トラフィックが増加している。

(1) グローバル化に伴いファイル共有等で国内拠点と海外拠点間など長距離通信が増加

(2) SaaSやクラウドサービスの利用に伴い、拠点外への通信トラフィックが増加

(3) 災害対策として遠隔地のデータセンターにWAN経由でのデータバックアップ

通信トラフィックが増加することで離れた拠点間でデータ転送に時間が掛かるようになり、高速にデータ転送を行う技術が注目されている。

富士通では拠点間、拠点-クラウド間の帯域を有効活用し、スループットを向上させる WAN 高速化製品 Smart Communication Optimizer を販売している。製品の特長は以下の通り。

(1) 独自技術でデータ転送速度の大幅向上

富士通独自のトランスポート高速化技術（特許登録済）により、パケットロスやネットワーク遅延に強い。

(2) 簡単な導入

VMwareやKVMといった仮想化基板向けの仮想アプライアンスとして提供し、既存の仮想基板上にVMイメ

ージファイルをインポートするだけで簡単に利用できる。

(3) 導入効果の可視化

業務ごとや拠点ごとの往復遅延時間、パケットロス率、スループットを表示でき、製品導入後の効果を確認することが出来る。

3.5. 現在までの大規模・大量データの共有方法と分析

現在までの大規模・多量データの共有方法と分析の実例として、【A. 各委員のシステム紹介】に各委員のシステムや業務を紹介する。

3.5.1. データを配布する側の事例 【A.1. 早戸委員】

【東京大学 宇宙線研究所 神岡】

・一般向けデータ公開 (O(100GB))、web base、データベース
・一般向けデータ公開 (O(1PB))、データファイル
・共同研究者向け大規模データ共有

3.5.2. データを配布する側の事例 【A.2. 藤田委員】

【宇宙航空研究開発機構 調布】

・一般向けデータ公開 (O(100KB))、web や scp
・一般向けデータ公開 (O(～100KB))、ソースコード

3.5.3. データを配布する側の事例 【A.3. 中村委員】

【高エネルギー加速器研究機構】

・一般向けデータ公開 (O(100GB))、解析環境、解析ソフトウェア、ドキュメント
・一般向けデータ公開 (O(1PB))、データファイル
・共同研究者向け大規模データ共有 (O(100 PB))、実験データ、処理済みデータ

3.5.4. データを配布する側の事例 【A.4. 白崎委員】

【国立天文台】

・一般向けデータ公開 (O(1PB))、観測データファイル
・一般向けデータ公開 (O(10TB))、カタログデータベース
・国立天文台内におけるデータ共有 (O(1PB))

3.5.5. データを受信する側の事例 【A.3. 中村委員】

【高エネルギー加速器研究機構】

・共同研究者向け資料共有
・共同研究者向けコード共同開発
・今後の課題 (新しい認証・認可基盤の導入、データの恒久的な保存、データの一般公開)

限られた計算機資源により拡大する総データ量を処理していくために、既存のグリッドサイトだけでなく、HPC 計算センターや、クラウド計算機資源を積極的に利用できる環境が模索されている (eXtreme DataCloud [4])。これらの計算機資源と分散されたデータを、データレイクなどの新興技術を用いて統一的に取り扱うためには、現在グリッド基盤で用いられている X.509 証明書のみにもとづいた認証・認可基盤 (AAI) からの脱却が求められる。グリッド認証局の維持とユーザ管理はコストが高いだけでなく、エンドユーザーにとっては X.509 証明書の取り扱いの敷居が高いことも理由となっている。次世代の AAI としては、ユーザが所属する各組織の ID を連携させることによるシングルサインオンの実現 (Federated Identity Management for Research [5]) と、エンドユーザーの利便性を高めるための SAML 認証や OAuth を用いたトークンベースの AAI の採用が候補として検討されている (Authentication and Authorisation for Research and Collaboration [6])。

加速器実験で取得するデータ量は膨大なため、過去に行われた実験データの恒久的なアーカイブを続けるための様々な共同研究が、Data Preservation in High Energy Physics [7] の枠組みにより行われている。大型加速器の運転は電気料金などのコストが高く、同じ目的で再運転することは容易でないため、一度取得したデータを

将来にわたって再利用できることが望ましい。また、計算機器とその利用技術の進歩や、データ解析手法や理論物理の発展により、過去に取得したデータの再解析による新発見の可能性も期待されている。これらを実現するためには、大量データのアーカイブ技術だけでなく、バーチャルマシンやコンテナを用いたデータ読み出しのためのソフトウェアを含めた計算環境の保全なども必要となる。

加速器実験データの一般公開は、データ量の多さから、配布のために必要なネットワーク帯域幅と、データ処理のために必要な計算能力が膨大となるため、現在では新規参入の研究者をトレーニングする目的、学生の教育目的、アウトリーチ活動の目的に留まっているのが現状である（CERN Open Data [8]）。一方で昨今は、データ解析ソフトウェアなどの使い方を説明する文書を整備するだけでなく、Jupyter などのノートブックツールを用いて、データ解析環境と大量データへのアクセス方法を提供することにより、Web ベースの解析方法を通して、一般の人々が 1 PB 程度のデータにアクセスすることが可能となりつつある。今後は、その規模と簡便性を向上させていくことが課題となっている。

3.5.6. ポリシーやガバナンス事例　【A.5. 渡邉委員】

【広島大学】

・共同研究者向け資料共有

・共同研究者向けコード共同開発

クラウドサービスを利用する場合は、クラウド事業者との間に外部委託契約を締結する。このため、クラウド事業者の選定、サービス内容の確認、責任体制の構築等を慎重に行う必要がある。一方、選定基準や確認事項などについては学術機関共通の指標があるわけではなく利用する側で定めておかなければならない。本項ではクラウドサービス利用におけるガイドラインやチェックリストを先行して提供している広島大学の事例を紹介する。

広島大学では、大学が保有している情報（法人文書）の保存先としてクラウドサービスを利用する場合を想定した「広島大学クラウドサービス利用ガイドライン」（以下、ガイドラインという）を策定している。クラウドサービスを利用する際に広島大学として確認すべき項目の解説がされており、チェックリストにより広島大学が定める法人文書の重要度の応じたクラウドサービスの利用可否を容易に確認できる資料となっている。クラウドサービスの利用者は導入前にチェックリストの項目を確認し、その結果を指定された部署に提出する。また、利用者は 1 年を超えない期間ごとにその時点で最新のチェックリストの項目を再確認し提出する。これにより組織的なクラウドサービスの利用状況の把握することに加えて、情報セキュリティ関連の要求要件とクラウドサービスのセキュリティ対策状況の対応を定期的に見直す方針となっている。

チェックリストでは準備、検討、運用の 3 つのフェーズで確認すべき項目が定められている。以下にチェック項目の一覧を示す。

【準備】

- 利用に向けた準備（必須確認項目）
  - ➢ 取り扱う情報の確認
    - ✧ 情報の格付け
    - ✧ クラウドサービスの選択
  - ➢ 本学の組織・体制
    - ✧ クラウドサービス利用責任者
    - ✧ クラウドサービス利用担当者
  - ➢ 規則・契約
    - ✧ 規則との整合性
    - ✧ 契約の取り扱い

【検討】

- 利用範囲の明確化
  - サービスの品質
    - SLA
    - メンテナンス
    - 問い合わせ窓口・サポート体制
    - サービスの継続性
  - 機能とコスト
    - コンピューティング
    - ストレージ
    - ネットワーク
    - 管理機能
    - ライセンス
    - コスト
- クラウド事業者の選定
  - データセンター
    - データセンターの場所
    - 堅牢性
    - 機密性
  - クラウド事業者の信頼性
    - 経営状況の確認
    - 委託関係の確認
- 契約条件の確認
  - 責任範囲とペナルティ
    - 責任範囲の明確化
    - クラウド事業者のペナルティ
  - データセンターの所有権、返却・消去
    - データの所有権
    - データの返却
    - データの消去
  - 準拠法と管轄裁判所
    - 準拠法
    - 管轄裁判所

【運用】

- 運用体制の確認
  - システムの運用に関する項目
    - セキュリティ対策
    - ログの監視
  - データの管理に関する項目
    - 秘密鍵の管理
    - バックアップ
  - インシデントの管理に関する項目
    - インシデントの記録

広島大学では利用に向けた準備の項目を必須確認項目としている。ここではガイドラインの特徴である情報の格付けとクラウドサービスの選定ついて紹介する。詳細については、「広島大学クラウドサービス利用ガイドライン」を参照されたい。

- 情報の格付けとクラウドサービスの選定

広島大学の法人文書は、大学の業務に与える影響の大きさに応じて 4 段階の重要度に分類され、重要度ごとに管理方法やアクセス権限が定めてある「表 4　法人文書重要度分類表」。また、ガイドラインではクラウドサービスの利用時に満たすべき要件とクラウドサービスの主な機能から 4 段階の信頼度を定義し 4 段階の重要度と対応付けることで、法人文書の保存場所としてクラウドサービスの利用する際の基準を容易に確認できる工夫がされている「図 4　クラウドサービス利用基準」。

表 4　法人文書重要度分類表

区分		文書の種類	情報の類型（抜粋）
重要度Ⅳ	法人文書が流出(漏えい), 紛失, 改ざん等した場合，本学の業務等に深刻かつ重大な影響を及ぼすもの	特定の関係者以外に対し厳重に機密を保持すべきもの	○学籍簿，成績原簿に関する情報 ○指導要録の情報 ○医療に関する情報
重要度Ⅲ	法人文書が流出(漏えい), 紛失, 改ざん等した場合，本学の業務等に重大な影響を及ぼすもの	特定の職制, グループ又は部局等以外に対して機密を保持すべきもの	○学生指導に関する情報 ○卒業論文又は修士論文に関する情報 ○卒業後の就職・進路先情報 ○住所 ○氏名 ○生年月日 ○電話番号
重要度Ⅱ	法人文書が流出(漏えい), 紛失, 改ざん等した場合，本学の業務等に軽微な影響を及ぼすもの	公開を前提としていないもの	○学術講演会に関する情報 ○後援名義に関する情報 ○入試の実施計画に関する情報
重要度Ⅰ	法人文書が流出(漏えい), 紛失, 改ざん等した場合，本学の業務等にほとんど影響を及ぼさないもの	積極的な公開を前提としたもの	○公開講座等に関する情報 ○大学案内や広報誌に掲載している情報 ○募集要項に掲載している情報

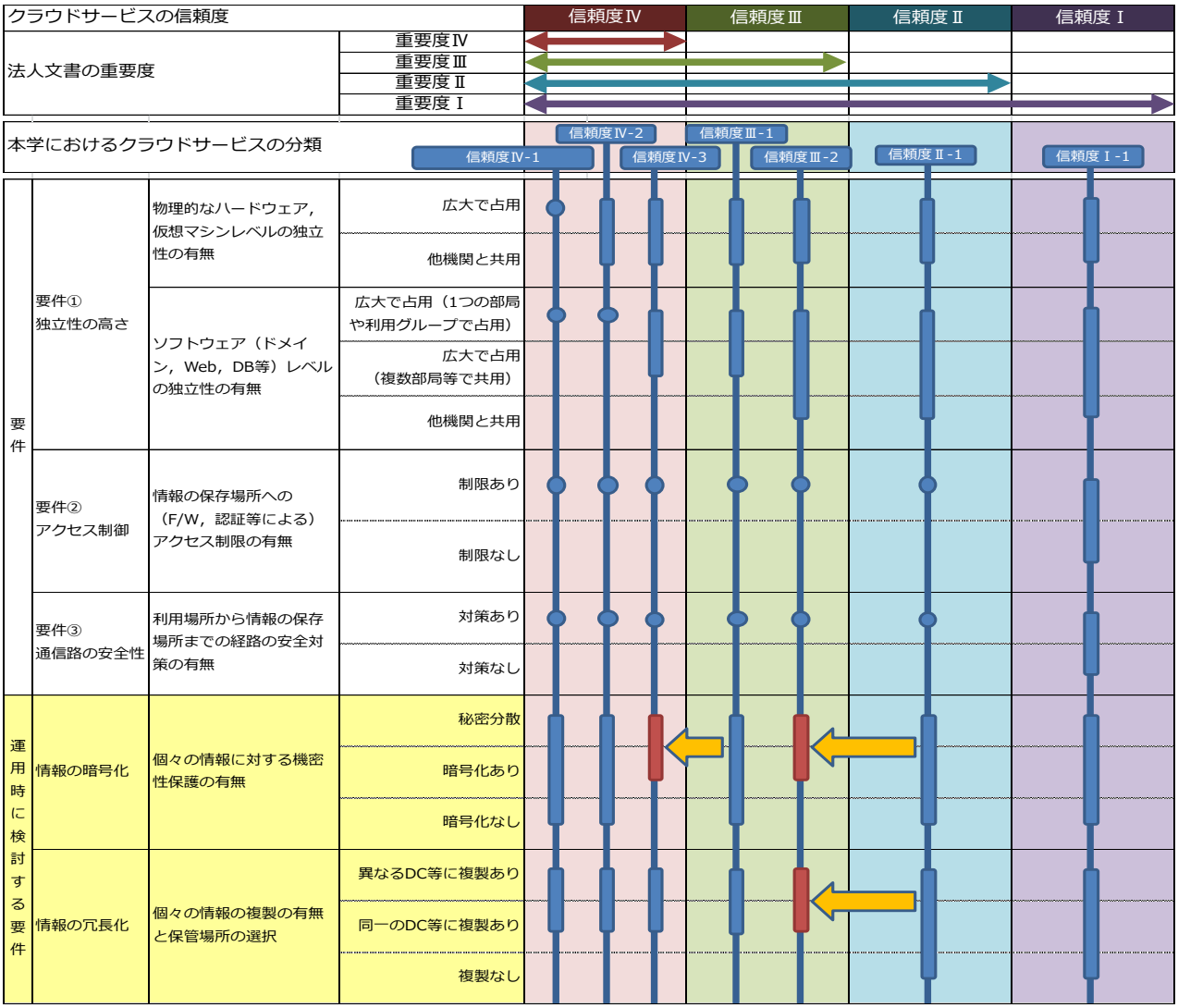


図 4　クラウドサービス利用基準

このように、データの保管場所としてクラウドサービスを契約する際は、事前に各学術機関で定めているデータに関する管理規則、個人情報や機密情報の取り扱いに関する規則に沿ってクラウドサービスを選定する必要がある。また、クラウドサービスの利用は業務の外部委託と同等と考えるべきであり、各学術機関で定めている業務請負に関する契約基準に準じていることが必要である。なお、クラウド事業者が個人情報等を別の委託先に渡すことがある場合は、取り扱いの承諾と基準を遵守する義務を委託先に負わせることも注意されたい。

3.5.7. 各研究機関の分析

当 WG では、現状を認知するために各研究機関のデータ分析とツール分析を実施した。

「表 5 各研究機関のデータ分析」では、各機関のデータの現状を、「表 6 各研究機関のツール分析」では現有技術を洗い出した。

表 5　各研究機関のデータ分析

会員様	共有したい相手	共有したいもの	単位	共有単位当たりのサイズ	総ファイル容量	Write･Read か Read Only	更新･アクセス頻度	現状の配布方法	一回の転送量	課題
北陸先端大	学内の他ユーザー 学外の共同研究者	データ，プログラム等	ファイル，フォルダ	1MB～数GB	300GB	Read/Write	随時	NextCloud	1MB～数GB	
東京大学 宇宙線研究所	共同研究者	データファイル	ファイル	0.1～2GB	O(1PB)			scp	１～１０００	情報の共有者と提供者間のメタデータ（ファイルの位置）管理(共有手法)が課題である。
	共同研究者	プログラムソース	ファイル	1～100kB	O(1TB)			svn., git etc.	１～１０００	
	共同研究者	プログラムバイナリ	パッケージ（VMイメージ？）	0.1～1GB	O(０．１PB)			http	O(1)	
	共同研究者	発表資料（ppt/pdf）	ファイル	0.1～100MB	O(１０TB)			http	O(1)	課題は、indicoは検索できないこと。 Docdbは会議には使いにくいことである。
	共同研究者	研究ノート	ファイル	0.1～100MB	O(１TB)				O(1)	
	関係する研究者（公開）	公表結果ファイル	ファイル	1～10MB	O(０．１TB)			http	O(1)	
	関係する研究者（公開）	一般公開用プログラムソース	ファイル	1～100kB	O(1TB)			svn, git etc.	１～１０００	完全公開であれば hepforgeサイトが存在する。
	関係する研究者（公開）	一般公開用バイナリ	パッケージ	0.1～1GB	O(０．１PB)			http	O(1)	
	関係する研究者（公開）	発表資料（ppt/pdf）	ファイル	0.1～100MB	O(１０TB)			http	O(1)	課題は、indicoは検索できないこと。 Docdbは会議には使いにくいことである。
宇宙航空 研究開発機構	共同研究者	データ	ファイル(セット)	100kB	O(1PB)	Read Only	不定期/季節変動あり	http、tar+scp	O(１００００)	課題はファイルの選別と圧縮に要する人の手間である。 情報の共有者と提供者間のメタデータ（ファイルの位置）管理(共有手法)が課題である。
	共同研究者	ソースコード	ファイルセット	～100KB?	～100KB?	Write･Read	毎日～毎週	Git	O(10)	セキュリティ(アクセスコントロール)が課題である。
	共同研究者	アプリケーションバイナリ	ファイル	?	?	Read Only	毎年	リムーバブルメディア	O(1)	ユーザー管理（輸出管理制限など）が課題である。
国立天文台	世界中の天文研究者	データ	テーブル形式データ	1kB～100GB	O(10TB)	Read Only	更新：年数回 アクセス頻度：～1000回/月	http	1kB～100GB	レコード数が膨大（＞１０億件）に推移した場合においても高速検索を実現したい。
	世界中の天文研究者	データ	観測データ	10MB～1TB	O(1PB)	Read Only	更新：ほぼ毎日 アクセス頻度：～1000回/月	http/ftp	1MB～1TB	自由に拡張可能なファイルシステムが必要である。 ディスクを追加するとデータが均等に自動分散されることが 理想である。 複数サーバーからの同時アクセスに対しても高いスループットを維持したい。
	天文台内部	データ	観測データ	10MB～1TB	O(1PB)	Read/Write	更新：ほぼ毎日 アクセス頻度：～数回/日	rsync/ftp など	1MB～1TB	国立天文台内の異なるプロジェクト間で利用可能な共通データ保存領域。 この保存領域にデータをコピーすることにより、そのデータを取得可能な URL が一意的に 定義される。それをキーとして天文台外にデータを配信できる仕組み。（DropBoxに類似）
高エネルギー 加速器研究機構	共同研究者	生データ	ファイル	0.1～10GB	10PB～100PB			gridftp, xrootd	データ取得期間は 継続的に転送	大規模実験で行われていることを列挙する。 ＊ グローバルな名前空間とレプリカの所在を示すファイルカタログをグループごとに持っている。 （ファイルの完全性、チェックサム、定義済データセット内のファイルが特定のサイトに揃っているかどうか、転送中かどうかの情報） ＊ ファイルとしてのメタデータだけでなく、ファイルの中身のサマリも付加(事象数など)し保存する。 課題として、ファイルはテープまたはディスクに存在する。テープに存在するデータの制御はディスクと異なりデータの取り出しに時間を要する。加えて、アクセス集中するとさらに時間を要する。 一方、全データは膨大なため、ディスクに保存できる期間は数か月程度が限界である。
	共同研究者	解析データ	ファイルまたはデータセット	5MBから100MB	O(10PB)			gridftp, xrootd, http	データセット単位O(1000)	
	共同研究者	ソフトウェア		O(10GB)	O(100TB)	read only	随時	CVMFS, VM, コンテナ		集中アクセス回避のために大規模なキャッシュが必要であること、データが大容量であるため転送中の同期の確認が必要であること、データサイズが大きくなると配布が難しくなっていくことが課題である。
広島大学										

表 6　各研究機関のツール分析

会員様	テクノロジー	目的	詳細	課題	期待する技術・トレンド
JAXAのソリューション	ファイルシステム/レプリカ機能	遠隔データ共有	Cloudコンピューティングのためのデータ共有ツールとして	I/O速度性能とのトレードオフであることが課題である。	
	VR/AR/MR	ディスカッション	離れた場所同士で同じシミュレーション結果を共有しながら議論を行う	報告ではない、議論を行える臨場感の実現することが課題である。	
	Web	データ配布	地球観測データや宇宙探査データを興味を持っている人 (場合によっては不特定多数)に配布する		Push型配布技術
加速器実験などで 利用されているツール	indico	会議支援ツール	会議の議事に従って発表資料などをおくことができる	ファイルのコンテンツを検索できないことが課題である。	
	docdb	ファイル共有	ファイルを共有する	アカウント管理（グループアカウントを用いて利用している模様）	
	hepforge	ソフトウェア共有・開発サイト	ソフトウェアの完全公開には利用可能		
	github	ソフトウェア共有・開発サイト	契約して実験グループとして利用することがある	アカウント管理が課題である。	
	website	一般公開		site管理の手間、CMSの場合はセキュリティが課題である。	
	plone	情報共有		アカウント管理、site管理の手間、セキュリティ、ソフトウェアの遅さが課題である。	
	gridftp	ファイル共有		アカウント管理、システム管理の手間である。 ソフトウェアが古い(globusはサポートを中止、CERNは当面サポートする)ことが課題である。	
	xrootd	ファイル配布		アカウント管理はグローバルな名前空間を提供しない限りエンドユーザーは使えない。 欧米サイトは日本から距離があるため、転送せずにダイレクトにアクセスする方法は使い物にならないことである。	
	fuse	(実行）ファイル共有	vmを一般公開して共有	完全に一般公開なので、アクセス制限は課題である。	
	CVMFS	解析環境			
	CERN VM	計算環境	ユーザ端末: Virtualbox, VMware, Hyper-V, Xen, KVM, VAGrant クラウド: OpenStack, Amazon, MS Azure, Google	サイズが大きくなりがち。 private/publicクラウド双方とも、deploymentツールと環境が必要である。	
	docker	計算環境	コンテナ	セキュリティに問題	
	singularity	計算環境	コンテナ（USのHPC分野で発展）	発展途上でノウハウが蓄積されてない。 最新Linux環境でなければ問題がある(setuidなど)	
	gitlab	ソフトウェア管理	大規模研究所では用意している	大規模研究所では用意している。 それらのアカウント管理、リモートアクセス管理が必要である。	
	Vac, Vcycle	計算環境	VM	主にUKで使われている。	
	BOINC	計算	ボランティアコンピューティング 大学等にある小規模クラスタを短い計算単位で利用	生成データの信頼性と管理が必要である。	
国立天文台で 利用しているツール	PostgreSQLなどの リレーショナル DB	データベース	カタログデータの格納、観測データのメタデータを格納	データベースの巨大化により検索時間が長くなる。 複雑な条件式の場合インデックスが利用できなく全件検索となることが課題である。	
	Tomcat などのWebコンテナと Struts などのWebアプリケーション フレームワーク	データ検索インターフェイス	ユーザがデータを検索したり取得したりするための Web インターフェイスを提供	様々なフレームワークがあり、どれがベストか選択が困難である。	
	Virtual Observatory アクセスツール	データ検索方法等の世界共通化	検索プロトコル、データフォーマット、ファイル共有方法の天文分野における世界標準	仕様が複雑かつ、曖昧な部分を含むため、仕様に合致していないサービスが多数存在する。	
	OpenLDAP	ユーザ認証	ユーザ認証を行うためのデータベース	シングルサインオンへの移行が必要である。	
利用が可能と 思われるツール	Gfarm	広域ファイルシステム		データ越境が課題である。	
	owncloud/NexTCloud	ファイル共有サービス		データ自動修復が必要である。	
	SharePoint	コンテンツ共有サービス		整合性判定（検出まで）が必要である。	
	学認	認証基盤	日本国内のみ		
	AWS、Azure		CPUの利用は行われている	データ越境が課題である。	
	高速ファイル転送ソフト		輸出規制	データ越境が課題である。	

4. クラウド時代のデータ共有技術における課題(受容も含む)

本WGの議論を通して、クラウド利用は、一般的にはコスト低減やインフラ調達期間の短縮などのメリットがあるという認識が得られた。一方で、特にセキュリティや事業者の継続性に関するリスク、ネットワーク帯域やアクセスチャージなどの付帯コスト、既存システムとの連携や障害対応などにおいて十分留意する必要があることも明らかになった。特に、本WGのメンバーが扱うような100億ファイルや10PBを超えるような大規模なデータ共有については、一般的に提供されているサービスとしては想定されていない点もあり、クラウドサービス提供者との綿密な調整が必要である。

また、現状のデータ共有の課題は二種類にわけられることがわかった。一つは、データへのアクセスに直接関係する課題、もう一つがデータへのアクセス許可など、認証やセキュリティに関連する課題である。以下では本WGで特に重要であると認識された課題について概説する。

4.1. データへのアクセス方法

クラウド利用を検討するにあたっては、まずデータへのアクセス単位やその方法と転送速度、反応速度などの性能をよく把握する必要がある。データファイルが大量数存在する場合や、ファイルが大量である場合、データベースのような細かいアクセスがある場合においても、要求性能が満たされ、かつ、利用量の上限がある程度予想できるケースはクラウドサービスも十分利用可能であると考えられる。一般にクラウドサービスはスケーリングしやすさが利点とされるが、規模拡大はサービス利用料の増加にもつながるため、適用事例がスケーリング可能かどうかについても事前の検討が必要である。事例によっては、オンプレミスのシステムを用いる方が、コストと性能のバランスをとりやすい場合もあるように見受けられた。

4.1.1. 検討しておくべきデータアクセス性能に対する要請

データへのアクセス単位としては、おおまかに、データベースソフトウェアからのアクセス（検索など）と、通常のソフトウェアからのファイル単位のアクセスに二分できると考えられる。また、ファイル単位の場合は、シーケンシャルアクセスとランダムアクセスに大別される。

以下、いくつかのパターンについて、検討しておくべき性能への要請を列記する。

- DBのようなアクセスの場合：時間当たりに処理するリクエスト数を見極める必要がある
- 大容量データファイル転送の場合：一回の要求に対し、「どの程度の時間が許容されるか」、「どの程度の転送速度が必要となるか」、「多重転送による高速化が必要か」、「帯域制限は必要か」、などを見極める必要がある。
- 多数のファイル転送の場合：転送開始までにどの程度の時間が許容されるのか、一回の要求で転送が必要となるファイルサイズ、ファイル数、そして転送完了までの許容時間を見極める必要がある。
- ランダムアクセスの場合：データファイルが利用可能になるまでの許容時間、アクセスパターン（単一ファイルへのアクセスが続くのか、複数のファイルへのランダムなアクセスなのか）、データ要求からデータが実際に届くまでのレイテンシはどの程度まで許容されるのかを見極める必要がある。ファイルの中に複数のオブジェクトを有している場合（見た目ファイルであるが中身はDBのようなアクセス）について、アクセスパターンが特殊である可能性があるので、注意が必要である、

4.1.2. アクセス方法に対する要請

・アクセスの手段（http/https/REST、Box、専用API）

用いるアクセスの手段を把握しておく必要がある。クラウドではhttp/https/RESTなどが主流である

- http/https/RESTの場合：認証・セキュリティ、アクセス履歴（証跡）などの情報について、利用者がどの程度管理、把握できるか事前に確認する必要がある
- Box/専用API（VOツール）の場合：データに適切なACLを付与するルールや運用(デフォルトの動作設定)などを検討する必要がある。
- オンプレミスでは、ファイルシステムの直接利用（マウント）やオブジェクトへの直接アクセス(オブジェクトストレージ)が主流である。
- ファイルシステムの直接利用の場合：CIFS、Lustre、SMB、NFSなどファイルアクセスとしてディレクトリ参照の重要性の確認が必要である
- オブジェクトの場合：cephなどのアクセス方法が必要か、上位にファイルシステムを介在させて利用するか

なども検討しておく必要がある。

4.2. データ移動量（オンプレミス→クラウド、クラウド→オンプレミス）に対する要請

予算の増加が許容される場合は、クラウドへの移動を検討することは比較的容易となる。年間の予算に上限が設定されているケースはデータ移動量を制御する仕組みも必要となるであろう。ローカル（オンプレミス）からクラウドへのデータ移動が安定していない場合にローカルのストレージがあふれてしまうような可能性についても考慮しておく必要があるだろう。

4.2.1. 課金システム（データ総容量、アクセス頻度、アクセス制限）について

・クラウド業者側でのサーバ設置場所によって使用料金が異なるケースがある。

国外の場合、サーバ設置場所によって、データ転送速度、レイテンシが異なる可能性もあるため、留意が必要であろう。（適用法令についても留意が必要。）

・課金の種類

- クラウドストレージのストレージ領域（データ保存領域）に対する課金
  - ディレクトリ単位やユーザ単位で上限設定ができるかどうかについても必要性を考慮しておく必要があると考えられる。
  - 特にユーザにストレージ領域への書き込みを許可する場合、蓄積を承認する仕組みも必要であろう。
  - 重複するデータを排除する機構が必要となる可能性もある。
- クラウドストレージの入出力のためのデータ転送(データ通信容量)に対する課金
  - データを蓄積するための課金はないことが多いが、確認が必要である。
  - データアクセス（読み出し）に対する課金がある場合が多く、アクセス量の見積が必要である。
  - データアクセス量などに対する制限が必要かどうか、検討が必要であろう。
  - アクセス制限においては、ユーザ認証との連携が必要となる可能性もある。
  - 大規模な読み出しを制限するなどの対策も必要であろう。
- レイテンシ要求（ｍｓ）
  - レイテンシ要請によりサービス価格が異なる場合があり、必要性能の事前検討が必要となる。
- アクセスコスト（高速回線代金)、ユーザ数、データ回収（システム移行）費用など
  - クラウドベンダまでの回線経費
  - クラウドベンダ内での接続について、接続速度によって料金が異なる場合があることに留意する。
  - アカウント（利用者）数で課金されるサービスの場合、利用者数の見積も必要となる。
  - 契約終了時、他システムにデータをコピーする必要があるかどうか、検討しておく必要がある。
  - 契約終了時に必要となるデータ転送量を用いて、必要となる経費を事前に見積もる必要がある。
  - 契約終了時に必要となるデータ転送時間を見積る必要がある。(必要な帯域の検討も必要であろう。)

4.2.2. 調達フェーズと運用フェーズにおいて検討する必要のある事項

・調達フェーズ

- クラウドサービス採用の利点
  - 特にストレージと同時にCPUなども同一のクラウドサービスプロバイダから提供を受ける場合、突発的な需要増に対応が容易となるため、クラウドを利活用することには大きなメリットがある。
  - 期間が短いプロジェクトにおいては、ほぼオンデマンドで調達できるため、無駄な時間を費やす必要がなくなる。
  - 小さい規模で開始し、将来的に規模を拡大する可能性があるプロジェクトにおいては、初期評価という観点でもクラウドサービスを用いて開始、クラウドサービスの利用を継続するか、オンプレミスに移行するかを検討する。このことが有用であろうと考えられる。
  - 既存のシステムからの移行の場合、各ベンダが提供しているサービス（ERP、CRM、勤怠など）も併用できる可能性があり、業務の統合が進む可能性がある。一方で、個人情報の保護などセキュリティ担保等の条件が満たされるかについては注意深い検討が必要である。また、それまでの業務プロセスに変更が必要となる可能性が高いため、これを受容できるかどうかについても検討が必要となる。

- クラウドサービス採用の留意点
  - 業務プロセスやシステムにおいて、ハード・ソフト（ライセンス契約含む）の規模を随時スケールさせることを前提にしなければ、クラウド本来の利点を活かしきることができない。これは、利用状況に応じた、比較的頻繁なサイジング再設計が必要となる可能性もある。再設計を適宜実施しない場合には、利点の一つであるコストの最適化につながらない。
  - クラウドストレージへのデータの入出力においては、とくに取り出しに課金がかかる場合があることに留意しておく必要がある。コールドストレージ（バックアップ）として利用する場合にはあまり問題にならないと思われるが、この場合でも、蓄積した全データを読み出す可能性がある場合には、読み出しに要する時間（ネットワーク帯域）と費用を見積もっておくことが重要である。
  - クラウドストレージからのデータ読み出しが大量かつ頻繁となる場合には、性能的に要請を満たすか、料金が想定を大幅に超えることがないかなどに留意する。また、想定外のアクセスがあった場合に追加負担が可能か、利用制限が可能かなどを事前に検討しておく必要がある。不必要、もしくは不正なアクセスによるコスト負担増加の可能性もあることにも留意が必要であろう。
  - クラウドストレージとの通信は回線帯域や利用状況に依存するため、性能評価時にはこれを考慮に入れておく必要がある。
- オンプレミスシステム採用の利点
  - オンプレミスシステムを採用する一番大きな利点はカスタマイズ性の高さであろう。データアクセスに特別な要請がある場合や、既設システムへの接続、稼働中のシステムからの継続性などにおいては、クラウドサービスへの移行は大きな負担となる可能性がある。
  - たとえば一定規模以上のストレージの場合は、トータルコストがクラウドサービスを上回らない場合も多いようである。
  - 一般的にデータ発生元、データ利用のリソースに近接してストレージを設置することができるため、レイテンシの短さが要請される場合や、高速アクセスが必要な用途に適している。
  - 性能改善のためシステム全体を通したチューニングなども可能である。
  - データの入出力に対して課金されることは通常ないため、アクセスが予想よりも多かった場合などにもコストの増加を気にする必要がない。
- オンプレミスシステム採用時の留意点
  - 調達に期間を要し、また、導入した機器を一定期間は利用しなければならないため、必要なリソースを精査し、調達時にシステム設計を入念に行うことが必須となる。
  - 調達した機器は、利用完了時までには時代遅れとなっている可能性があること、新たに出現した機材を追加導入することも難しいことを理解しておく必要がある。
  - ストレージの調達の場合、一定期間後に容量を追加する方がコストの削減になる場合もあるが、契約方法、期間によっては当初から一括で調達してもあまり変わらない場合もある。
  - 運用中の大幅なシステム構成の変更は現状では難しい。ただし、今後はプライベートクラウド環境をオンプレミスシステム上に構築することで、構成変更を容易にすることができるかもしれない。

・運用フェーズ

- クラウドサービス、オンプレミスシステムともに、稼働率や定常的な性能(通信・IOPS・ストリーム性能)の監視を行い、資産を十分に利活用する努力は必要である。特に、利用統計情報を適切に採取できない場合、無駄なコストが発生する可能性もあるため、注意が必要である。
- 契約順守状況の確認は重要である。オンプレミスやクラウド事業者の変更時などに注意が必要である。 監査手法の確立などの取り決めを事前に実施しない場合、オンプレミスやクラウド事業者が提供するサービスレベルの妥当性について、調達者は判断することが困難となる。

4.3. データ共有にかかるリスク対応

法令に準拠した信頼できるコンピューティング環境を構築するには、「コンプライアンス」「機密性」「安全性」「ガバナンス」「データのリスク管理」を少なくとも満たす必要があり、検討すべき項目を4点あげた。

4.3.1. ガバナンス・コンプライアンス・国際法律

・GDPR は遵守しないと罰金が科される可能性もある。データをクラウドサービスに移動することは、様々な地域へのレプリケーションの可能性があり、この GDPR の順守を考えなければならない。

・クラウドストレージに蓄えるデータがどのような種類のものか（個人情報などを含むのか、機密情報などを含むのか）、利用者の立場として貴重（失うことができないもの）か、再生成（再コピー）可能なものか、無期限で保存すべきものか、一定時間経過後に抹消すべきものか、管理責任があるものか、管理義務がないものか、などについて、すべてのデータについて、明確にラベリングすることが必要となる可能性もある。

・運用段階では、実際に提供されているサービスレベル品質や、事業者側のインシデント対応状況を利用者がモニタリングすることが重要である。利用者のデータがクラウド上の複数のサーバ上に保存されているケースでインシデントが発生した場合に、クラウドサービス提供者がどこまで調査を実施するか、情報提供を行うかなど、事前の確認が必須であろう。

4.3.2. 連携する両者のポリシー整合性

・データ提供元⇔データ活用先

- データがどこにあるのか、また、どのようにアクセスできるのか、という情報を共有することが必要となっている。ここには、アクセスへの許諾申請方法や具体的なアクセス方法などの情報が必要となる。また、検索機能を必要とする場合は、メタデータの生成も必要となる。
- 特にクラウドサービスを利用してデータを公開する場合、必要リソースやアクセスコストの評価のために、データへのアクセス頻度、アクセス量などといった情報だけでなく、データがどのように活用されているか、そのアクセスパターン（時期、時間帯）や用いられているアクセス方法などを把握しておく必要があり、これらを理解するための可視化も重要となっている。

・データ提供元⇔データ格納先

- 技術的なセキュリティが保証されない部分に関しては、契約約款や SLA の内容をクラウド事業者と詰めておく必要があるが、事業者によっては約款や SLA の変更に応じないケースもあるため、注意が必要である。
    - クラウドサービス導入が不可能である可能性もある。
    - これらを踏まえ契約関連で利用者が留意すべき事項の事例(広島大)にも示した。
- 利用状況は常時確認し、必要に応じて設定やリソースの契約を変更することが必須である。
    - 適切に行わない場合、データ提供元は予期しない金銭的負担を強いられる可能性がある。
    - 期間限定で公開をしていたサーバの利用終了手続きを怠ると、いつまでも課金される。
    - リソース利用量に制限をかけない場合、想定額以上の資金が必要となる可能性もある。

4.3.3. 認証・証跡

・認証・証跡

- 直接アクセスするユーザ、直接アクセスしたユーザからデータを参照する（プロジェクトとしてアクセスするケース）グループを識別できる仕組みが必要となる。
- 大学や研究機関のルールに即した使い方ができているか。調査して維持する仕組みの検討が必要である。
- 内部統制において、クラウド特性を考慮した監査基準が確立されているか確認することが必須である。
- 国際的なプロジェクトの場合、該否判定の書類審査などが存在する場合、本格利用までの準備期間が必要になることを念頭におく必要もある。

4.3.4. データ保護・アクセス権限・特権 ID

・データ保護

- データを削除した際に、各地に保存されたすべてのデータが削除されないと、データの不正利用につながる恐れや保護法を守れない恐れがある。
- 流出が問題となるデータの保存においては、クラウド上のデータが暗号化されているかどうかを確認する必要がある。
    - Dropbox などは自動的に暗号化している。
    - 暗号化されていないサービスの場合、事前に暗号化するなどの対策を利用者側で行うことも検討する必要がある。
    - 利用者が暗号化する場合は、暗号化キーの管理が課題となるであろう。KMIP（Key Management

Interoperability Protocol）のような暗号鍵管理プロトコルを用いて暗号化システム間を接続する可能性なども検討に値するであろう。

- 保存時だけでなく、データ転送時のセキュリティについても考える必要がある。
  - データ転送中のセキュリティは、クラウドサービスへの通信方法、ユーザの用意した回線、クラウド事業者のネットワークのセキュリティなどに依存する。
  - 保存中のデータ盗聴保護はデータの暗号化などによって実現することが可能となる。
- 保存したデータが正しく読み出せているか確認する仕組みが必要である。
  - データが破損していないか、改ざんされていないかなどをチェックする機構は一般的なサービスは持たないので、利用者側で必要性や実現手法の検討が必要であろう。

・アクセス権限・特権 ID

- 権限において、最重要データはオペレーションでも削除できない仕組みが必要（データ改ざん防止）
- 変更履歴の取得が可能であるかについても確認しておく必要がある。
- 運用者へのアクセス権限において、運用者の転職などを考慮したルール策定が必要（データ改ざん防止）

4.4. データ共有時の利便性や管理の容易さを向上させる技術

データの管理や利活用にあたっては、データの種別わけが重要である。この情報がファイルに付随していることで、正しい保存場所、保存方法を選択することが可能となるだけばかりなく、クラウドサービス利用の可否や、サービス選定時に求められるセキュリティ要件などを検討するときに重要な情報となる。

データの種別わけ自体は利用者に依存してしまうが、この種別わけの結果をデータとあわせて管理する方法については複数あり、実装例を以下に示す。

4.4.1. データへのタグ付け（LOD、データジャケット、メタデータ）、VPX

・データのラベリングについて、これを実現する技術としては以下のようなものがある。

- LOD（Linked Object Data）
- データジャケット
- IVOA(天文分野)
- VPX（ブロックチェーン技術の応用）
- ハッシュタグ

・メタデータとして必要となる可能性のある情報としては以下のようなものがある。

- データの生成日時
- データの生成場所
- データの種類
- データを所有者(権利者)、これを保証(認証)する組織（もしくは人）
- データの公開範囲（無制限、関係者のみ）
- データの有効期間
- データの再利用可否
- データの加工の可否
- データの再配布の可否
- データのチェックサム（破損・改竄確認）
- データの正当性保証情報（電子署名など）

5. データ共有技術におけるオンプレミス・クラウド利用に向けたチェックリスト

本WGにおいてシステム間の大規模な「データ共有」においてクラウドサービスの利用を検討するとき、各機関に共通して次のような懸念が存在することが確認された。

・第三者が提供するクラウドサービスに任せることやベンダロックインへの懸念
・実効帯域やストリーム性能やレスポンス性能への懸念
・不正アクセス・改ざん・情報漏えいなど情報セキュリティに関する懸念
・第三者のクラウドサービスのシステムや運用の可視化やポリシー定義などの責任分界点への懸念

これらの懸念はオンプレミスのシステムにおいても同様であり、事前に当事者間で基準について共通の認識を持つことが重要であるとの結論に至った。

また、クラウドサービスは個人から企業向けまで目的に応じたサービスが存在しており、利用を検討するときに留意すべき点を明確化しておく必要があるということも明らかになった。

一般的なクラウドサービスの導入にあたっては、国立情報学研究所が導入支援サービスを実施するにあたって作成した資料（https://cloud.gakunin.jp/ass/　で提供される）が大変有用である。一方で、データ共有という視点に限った場合には、この資料では若干使いにくいため、本WGにおける議論を踏まえ、チェックリストを作ることとした。

以下に掲載するチェックリストは、「データ共有」というカテゴリにおいて、オンプレミス・クラウドの適応を判断する情報システム担当者が、効率的かつ網羅的に検討を行うことができることを目指して作成した。
実際の導入にあたっては、前述の国情研の資料なども参照の上、ここに掲載されていない項目についても確認しておく必要がある点に留意していただきたい。

5.1. 共有データ種別・目的・方法

「データ共有」という観点から、比較的小さいサイズのデータを一般公開する場合、期間を設けて共有する場合、オフィス文書などを共有する場合は、クラウドサービスの利用が適している可能性が高い。一方で、特殊なアクセス方法等が必要な場合は、カスタマイズ性の観点からオンプレミスとせざるを得ないケースもある。さらに、データ総容量が1PBを超えるような場合、容量単価はオンプレミスシステムの方が安くなる可能性がある。また、クラウド上のデータの移行に時間やコストがかかる可能性があることもあり、オンプレミスでシステムを構築する方が適している可能性が高い。

5.1.1. データ種別

・ドキュメントファイル（文書、画像など）の共有では、専用のクラウドサービスが提供されており、アクセス手法（API）、アクセス制限、セキュリティなどについて要請を満たす場合は、クラウドサービスを用いたほうが利便性は高くなると考えられる。

- 失うことができない貴重なデータについては、クラウドサービスのみにデータを置くことは大きな危険が伴う。逆に、データセットのレプリカを置く場所として利用するのであれば、検討に値する。
- サービス契約完了時に消去可能な場合は、クラウドサービスに置くことも検討に値する。

5.1.2. データアクセス

・インターフェイス

- ベンダが提供するサービスをそのまま利用可能なケース（http/https、REST APIなど）や、試験的に情報共有するケースにおいては、クラウドサービス利用が検討に値する。
  ファイルシステムとしてマウントしたい場合、既存システムとの連携が必須となる場合などは、オンプレミスシステムの利用の方が適している可能性が高い。

5.1.3. 総データ量

・容量が大きい（数百 TB 以上）場合、オンプレミスシステムの方がコスト面で有利になる可能性が高い。
・ファイル数が多い（数億以上）場合などついても、オンプレミスシステムの方が有利になる可能性がある。
・契約期間内（５年程度）の容量増加と総容量が見通せる場合は、クラウドサービスの必要経費が正確に計算できるが、見通せない場合はオンプレミスとクラウドサービスのどちらが良いかの判断が難しくなる。

5.1.4. データアクセス量

・データへのアクセスに課金されないサービス（オフィス文書共有など）については、クラウドサービスを利用しやすいと考えられる。
・データへのアクセスに課金されるストレージサービスについては、アクセス量についても見積もりが必須となる。大量の読み出しアクセスが発生する可能性が高い場合は、オンプレミスシステムが無難であろう。

5.2. 可用性

データ共有を行うにあたり、可能な限り常時アクセスしたいという要望があると考えられる。また、運用システムの稼働が困難になった場合、どの程度の期間でこれを再開できるか、という観点からの検討も必要であると考えられる。これらについて、３点の確認項目にまとめた。

5.2.1. サービス時間（計画的、事前通知、サービス終了リスク）

・計画されたサービスの定期保守頻度および総時間（１年あたり）を確認する。
・計画停止において、事前停止通知日（期間）を確認する。
・サービス終了の事前通知期間を確認する。（サービス移行に要する時間の検討も別途必要である。）

5.2.2. 稼働率・継続性

・稼働率の定義と保証される条件を確認する。
稼働実績などは実稼働状況が確認できる無料オンラインサービス(CloudSquare Service Status)などを用いて調べることもできる。
・突然の停止時の措置（対応）を確認・検討する。
・重大障害時の代替方法を確認・検討する。

5.2.3. ディザスタリカバリ

・災害時のシステム復旧体制の有無と内容を確認する。
・別システムもしくはサービスに移行する必要が出た場合、運用ノウハウや資産を継承できるかを確認する。
・特に失うことのできないデータについては、多重化などが可能かについて確認する。

5.3. システム管理・監視

データ共有システムを利用するにあたって、クラウドサービスにおいては、一般的にサービス利用者が実施可能なオペレーションやアクセス可能なログが制限される。必要なオペレーションやログ監視が可能かどうかについて、２つの確認項目にまとめた。

5.3.1. リソース管理オペレーション

・サーバ、ブロックストレージ、ネットワークのリソース制御の API(CLI,SDK,Console など)仕組みについて確認する。

5.3.2. システム監視（ログへのアクセス）

・以下のログにアクセスすることが可能か確認する
- （データ）アクセスログ、改変ログ
- ロードアベレージ、パフォーマンスログ
- システム警告・エラーログ

5.3.3. システム運用

・セキュリティ確保のため、情報レベルなどを細かく規定した運用を実施したい場合は注意深い検討が必要である。オンプレミスが最初の候補になると考えられるが、最近は比較的機密度の高い情報についてもクラウドサービスを利用する広島大学などの事例もあり、不可能なわけではない。

5.4. 性能

近年、民間企業などにおいても、大容量、高スループットのストレージ・データ共有システムが必要となってきており、本WGのメンバーが求めるシステムに近い規模・性能のシステムをクラウドサービス上に構築するケースも出てきている。たとえば、スマートフォンからの利用を想定した場合、1ユーザあたり5GB程度の容量を7500万ユーザに提供すれば、総容量として400PBが必要となる。帯域としては10%のユーザが0.5MB/sの帯域を同時にアクセスすると仮定すれば4TB/s程度の帯域が必要となる。これらを実現できるクラウドサービスは、十分検討に値すると考えられる。性能については以下に３つの確認項目としてまとめた。

5.4.1. スループット性能

・実効帯域

- 回線帯域などを鑑みれば、10GB/s以上が必要であるならば、オンプレミスとする必要があるであろう。
- 逆に、100MB/sでよければ、クラウドサービスは十分検討に値すると思われる。

・更新増量

- 回線・転送速度を鑑みるに、平均で数TB/日で継続的にデータを蓄積する場合は、オンプレミスが適していると考えられる。
  逆に500GB/日程度以下であれば、クラウドサービスの利用が検討できると考えられる。

5.4.2. レスポンス性能

・RTT

- 10ms以下のレスポンスタイムが必須であるならば、オンプレミスを選択すべきと考えられる。
- これ以上であれば、クラウドサービスの利用が可能であろう。NII学術情報基盤オープンフォーラム2015の資料によれば、１か月平均では10～40msとのことである。

5.4.3. 処理能力【件数】

・Query数

- RTTとあわせて確認が必要である。
  クラウドサービスにおいても利用サーバ数を増やすことなどで処理件数を増やすことは可能であると考えられるが、コストにも影響するため、注意が必要である。

5.5. データ管理

クラウドサービスを利用する場合、データの移行や取り出しについて留意が必要である。また、データを消失させない、漏えいさせないといった観点からの検討も必要となる。

5.5.1. データ移行方法

・データ移行手法

- 現有データを新システムに移行するために必用な帯域、時間、コストの見積りを行うことは必須である。クラウドの場合、ネットワーク経由となるケースが多いが、テープやディスクなどの持ち込みなど、他の方法が提供されている場合もあるため、確認が必要である。

5.5.1. バックアップ（容量、バックアップ時間、保存期間）

・バックアップ方法

- オンプレミスシステムにマスターをおき、クラウドをバックアップとして利用するケース。
- クラウド上のシステムからオンプレミスに必要なバックアップをとるケース。こちらは、クラウド上の全フ

ファイルリスト取得が容易でない可能性などがあるため、事前の確認が重要。

- オンプレミスシステム内にスナップショット機能やテープ装置などバックアップ機能を持った機材を導入するケースは最も伝統的である。

・バックアップ時間

- 数十TB/月以下であれば、一般的なクラウドサービスでも転送速度的には問題ないと考えられる。
- バックアップデータ量が大量となる場合、読み出しコストがかかるクラウドサービスではコストが大きくなりすぎる可能性があるため、注意が必要。

・保存期間

- ディスクストレージはハードウェアの寿命と世代交代を考えると５年程度での入れ替えが必要となる。オンプレミスシステムでは、この程度以下の頻度で入れ替えが必要と考えられる。
  クラウドサービスでは、容量単価・アクセス単価は今後下がってゆくと考えられるため、契約期間や契約単価を随時見直す必要があることに留意し、これを考慮に入れた契約とする必要がある。

5.5.2. データ消去（ポータビリティ/委託）

・データ消去(データを破壊する)

- データ格納機材の物理的な破壊や目視を併用したRAID構成破壊など、完全なデータ消去が必要な場合、オンプレミスとする方が無難であろう。
- データの完全消去が必要な場合、たとえば米国政府はNational Industrial Security Program (NISP) の1995年版のオペレーティングマニュアルDoD 5220.22-Mなどに記されており、「全てのマッピング可能なセクタに何らかの文字で上書きを行った後、その補数の文字で上書きを行い、さらにランダムな文字コードで上書き処理を行う」という手法の実施を推奨している。こういったレベルの対応が可能かどうかについて、サービス提供者に確認する必要がある。
- データ消去を電子的に証明して、その証明書を電子的に発行する仕組みがある（日本ではCSAJなど）ため、これを利用することができるかについても検討が必要であろう。
  クラウドサービス提供者はデータ消去に関する規定を持っている場合が多い。一方で、この要件が必須となる場合は、これが順守されているかなどの精査も必要かと思われる。

5.5.3. 暗号化（データ漏えい対策、データ破損・改ざん検知）

・暗号化

- 記録時に暗号化を実施することで、漏えい対策になるほか、復号時にデータの破損や改ざん検知などを行うことも可能となる。
- オンプレミスで利用可能な技術は複数提供されているが、アクセス性能の劣化が許容範囲かどうかを確認する必要があるであろう。
- クラウドサービスで、クラウド上のストレージでデータを暗号化しているものもある。ただし、利用者はこれを認識しないで利用するものが大半である。
- クラウド上のストレージを利用する場合、データ転送時に暗号化、復号化を実施するストレージゲートウェイなどの機材の利用も検討する価値があるだろう。

5.6. セキュリティ

クラウドサービス提供者が提示するセキュリティの基準で要件が満たされる場合は、クラウドを優先的に検討してもよいであろう。一方で、セキュリティに対する要件が厳しい場合、複雑な場合は必然的にオンプレミスシステムの採用とならざるを得ないケースも残ってしまう。

5.6.1. ユーザ管理・認証

生涯ID、SSO、ID・パスワード自動再発行、プロジェクトをキーとした管理、LDAP連携、Active Directory連携 、デバイス認証、二要素認証、フェデレーション認証、アプリケーションのアサイン、一定期間アクセスがない場合のロック、スタッフからの流出を防ぐ機能などが必要と考えられる。

・ユーザ管理

- ローカルの認証に加えLDAP連携など、従来の管理方法を継続したいケースでは、クラウドサービスとの連携

がとれない場合もあるため、オンプレミスシステム利用が必須となってしまう可能性がある。

- クラウドサービスとの連携が可能な IDaaS(SKUID、１Password、RoboForm、Onelogin、CloudGate UNO、Azure Active Directory)を利活用できるケースもある。ただし、アカウント単位の課金となる可能性もあり、ユーザ数、コスト負担を行うものの確定などが必要となる。

5.6.2. 証跡・データ保護

・証跡の必要性

- 詳細な証跡監視・管理が必要でない場合は、クラウドサービスを利用しやすいであろう。
- 最近のクラウドサービスでは、多量、大規模な領域に対して「いつ、だれが（どこから、何をつかって)、だれと（どこから、何をつかって)、どのようなコンテンツ、何ＧＢまで」を共有したか長期間にわたり記録する仕組みを持つものもある。
- 逆に、米国防総省が規定したセキュリティ基準 TCSEC（Trusted Computer System Evaluation Criteria）のセキュリティ要件を満たした OS やファイルシステムの機能が必要なケースについては、クラウドサービスの利用は困難な可能性が高い。通信回線などの安全性担保なども含めて注意深い検討が必要となる。

・データ保護

- クラウドサービスにおいて予期しないデータ損失が発生した事例も過去に複数存在し、データのリカバリーはできていない。このため、データが失われる可能性を想定に入れた導入は重要であろう。
- クラウドサービスにおいて、レプリケーションのサービスを有している事業者も多いので、これらの併用検討は重要である。
- オンプレミスシステムにおいても、災害などでデータが失われてしまう可能性は十分にある。バックアップ先として、ストレージが地理的にも離れているクラウドサービスを利用することは、安全性担保、事業継続性担保という観点から有用である。
- 大規模ストレージシステムにおいては、いつのまにかデータの一部が破損する事故(Silent Data Corruption)の発生が問題となっている。オンプレミスシステムでは、ハードウェア機材の機能、OS（ファイルシステム）の機能で検出することも検討する必要があるだろう。クラウドサービスでは、どこまで保証されているか、一般的には不明のため、これらが問題となる場合は記録データのチェックサム（ダイジェスト）の取得・記録・比較を行うことも検討すべきであろう。

・改ざん防止（誰かが書き換えていないか)

- 改ざん防止対策が必須である場合、オンプレミスシステムにおいて、データ生成直後に改ざん防止のコード(電子透かし)など埋め込むなど、独自のガイドラインを構築する必要があるであろう。「Tripwire」や「AIDE」など改ざん防止手法を利用する方法、独自のデータフォーマットを用いてデータ改ざんを検出するなどの方法がある。
- クラウドサービスにおいては、経済産業省「クラウドサービス利用のための情報セキュリティマネジメントガイドライン」に準拠している WORM 機能を有するサービス（Amazon Glacier ボールトロックなど）を利用する必要がある。

5.7. クラウド・オンプレミス検討の方向性

データ共有というキーワードを主としてクラウド・オンプレミスの利点欠点などを議論した結果、用途や状況によって使い分けるハイブリッドも必要と考えた。各委員の業務の紹介や外的要因などを含めて議論する中で項目に対する重みも重要と認知された。選定の参考情報として以下に整理した。

選定するシステムにおいて内外の状況が【表 7　クラウドに向かない状況一覧】該当する場合は、ハイブリッドクラウドやオンプレミスの適用を主とした検討をお勧めする。該当しない場合は、クラウドファーストを検討できると考えている。

オンプレミスやハイブリッドクラウドの選定の議論の結果は、【表 8　オンプレミスやハイブリッドクラウドの選定基準一覧】にまとめた。

表 7　クラウドに向かない状況一覧

セキュリティに関する事項
データ公開を主とする短期間のプロジェクトがない。
運用やアプリケーションを外部委託しない。または、できない。
国際・国内法令・クラウドベンダ倒産に迅速に対応できない。
秘匿性の高いデータ・情報（国防・特許・研究途中）が存在し分割できない。
情報に関して統制・管理・相談できる部署が存在しない。
業務特性における人材に関する事項
クラウドに精通した人材を確保できない。または、不足している。
クラウドベンダのロックインを避けたい。
リソースの変動に対して判断する仕組みがない。変動を予測できない。
現有の ID 管理を刷新されると業務継続が困難である。
構築・運用コストに関する事項
学術ネットワークや既存専用線などの利活用ができない。
月次、年次の費用をオンプレミスと比較して低価格が必達である。
クラウドの制限（認証種類の制限や運用停止）に合わせて業務を変更できない。
月次の費用において大幅な変動が許容されない。
データ共有に関する事項
共有すべきデータの重要度、保持期間などラベリングの十分性を検証できていない。
データ転送または、アクセスにおいて、継続できることが業務遂行に重要である。
総容量 10PB 以上や 10 億ファイル以上のデータを共有したい。
データを消失しても、現実的な期間でのデータ再作成のスキームが確立されていない。

表 8　オンプレミスやハイブリッドクラウドの選定基準一覧

セキュリティに関する事項	
データのリスク分析や対応は検討済み	未検討または、今後の検討である
証跡が必要な業務であり、必要なデータを分類できている	分類できていない
秘匿性の高い業務（国防・特許・研究途中）と分割可能である	分割が難しい
上記の場合、オンプレミスと一部データのクラウド展開を検討可能	**上記の場合は、オンプレミスを推奨**
業務特性における人材に関する事項	
業務の変動を予測可能、または、傾向の変化が緩やかである	5 年前後でデータ移動が発生する
運用スキルを自組織に有したい	適切な判断のため、運用(ログ分析)のスキルを維持向上させたい
利用者が要望する検証環境が必要・存在（実験機連携など）する	現有の業務の踏襲が必須である
上記の場合、オンプレミスと一部データのクラウド展開を検討可能	**上記の場合は、オンプレミスを推奨**
構築・運用コストに関する事項	
データ課金対策としてデータ共有量を制御できる	データ共有量を制御できない
帯域課金対策としてアクセス量（スパイクピーク）が予測可能である	予測が困難である
統計情報（利用率、分布、通信、要求数、故障率）は、一部でよい	運用改善、次期機種選定などの比較として独自に必要である
上記の場合、オンプレミスと一部データのクラウド展開を検討可能	**上記の場合は、オンプレミスを推奨**
データ共有に関する事項	
データアクセス数やアクセス量が右肩上がりの業務である	数年、増加の傾きが急である(業務に変化あり)
データ転送の継続性は業務継続に直結する。(nnGB/秒の帯域が必須などの要件が明確)	データ生成・解析・共有までに大規模なデータ転送帯域が必要である
重要なデータは 2 重化以上など BCP 対策が確立(連携機関へのコピーやデータ転送のための帯域確保、再計算で生成など)されている	再現不可能や歴史的な非公開データが大量に存在する可能性が大きい
上記の場合、オンプレミスと一部データのクラウド展開を検討可能	**上記の場合は、オンプレミスを推奨**

6. 結び

近年、クラウドを用いたサービスは広く普及し、特に無料で利用可能なメールや動画配信サービス等は日常的に利用されるものとなった。また、ドキュメントやスプレッドシートの共有、共同作業ツールも非常に安価に利用できるようになり、オフィス系ソフトウェアや画像編集ソフトウェアについても、無期限に利用できるインストール型ソフトウェアの販売から、クラウドサービスの利用を含めたサブスクリプション型へと販売手法を変更している。これらのサービスは、ネットワークに接続され、ウェブブラウズが可能であれば、どこからでも利用可能であり、利便性が非常に高い。一方で、例えばメールサービスなどを組織が提供する場合、セキュリティ対策などが大きな負担となり、その運用コストも大きくなっているため、クラウドサービスと同様のレベルを維持することは大変難しくなっている。また、機能が不十分なサービスを提供したとしても、ユーザは利便性の高いクラウドサービスを独自に用いることもあり、管理ができない状況に陥ってしまう可能性も低くない。これらを勘案すると、サービスが一般的に提供されているメール、オフィスツール、比較的小さいドキュメントファイルの共有等については、適切なクラウドベンダ及びクラウドサービスを選択し、ユーザの利便性を高くすることが、セキュリティレベルを高めるうえでも有効であろうということが本 WG の議論から明らかになった。しかしながら、導入時の契約条項などには多くの留意点があるだけでなく、不用意な情報漏洩を起こさないためにも、適切な初期設定やユーザ権限の設定を行うことに、これまで以上に注意を払う必要があることも明らかとなった。また、一般にクラウドサービスは高い稼働率を維持しているが、大手クラウドサービスにおいて、半日程度ファイルへのアクセスが不可能になったという事例もあり、すべてを１社にゆだねることが適切かどうかも含め検討は必要であろう。ベンダ選定やサービス契約における留意点については、各種テンプレートの準備が進められており、これらを適宜利用することは有用である。推奨される各種設定などについての資料は充実していないが、今後徐々に整備されてゆくと予想される。

一方で、大規模なデータを幅広いユーザ間で共有する場合は、既存のクラウドサービスの適用が難しいケースがあることも認識された。数 PB を超えるような大規模なストレージを利用する場合、設置場所などに問題がなければコストメリットはあまり大きくない。また、データに頻繁にアクセスする場合は、通信における遅延など性能面でのオーバーヘッドがあるばかりでなく、データ読み出しに対する課金などにも考慮する必要がある。データ読み出し量の細かい調整、データ暗号化、データ破損に対する対策、さらにはデータへのアクセス認証方法などについても課題は多く残されていることも、本 WG の調査によって明らかになった。また、クラウド契約を解除する場合のデータの取り出しや移行に要する時間的、金銭的コストも無視できないため、具体的な利用方法について事前によく検討したうえでの選択が必要になる。これらを踏まえた場合、オンプレミスが有利となるケースも少なくないことも明らかになった。５章に、クラウドを利用したシステムを設計する場合に確認しておくべきチェック項目をまとめた。実際にオンプレミスとクラウドを取捨選択する場合には、その都度利用可能なサービスと要件について、比較・検討を続けていく必要がある。

今回の WG の議論を通してユーザ管理が、オンプレミスとクラウドサービスどちらについても大きな課題の一つとなることも明らかになった。複数の組織をまたいだコラボレーションの場合、ユーザの実在性検証、登録、権限付与・変更、削除などが管理上の大きな負担となる。認証については様々な標準化が進められているが、これらをいかに取り入れ、ファイル共有システムとしてうまく利用して行くかが、今後の大きな課題の一つとして認識された。

本まとめがシステム導入の検討の一助になれば幸いである。

A. 各委員のシステム紹介

各委員のシステムや業務を紹介する。

A.1. 早戸委員

【東京大学　宇宙線研究所 神岡】
データを配布する側の事例
・一般向けデータ公開（0(100GB)）、web base、データベース
・一般向けデータ公開（0(1PB)）、データファイル
・共同研究者向け大規模データ共有

検出器で取得した実験データを実際の物理解析に供するためには、まず一次処理（再構成）を行い、観測された粒子の生成(反応)位置、種別、運動量(エネルギー)などの情報に変換する。この変換されたデータについて、さらに解析したい物理事象の信号と、それ以外の事象（背景事象）をおおまかに区別することでさらにデータ量を減らし、実際にユーザが用いる解析用データファイル（DST）を生成する。 通常、DST は数百 GB～数 TB 程度となり、これを世界各国の研究者が利用する。これらのファイルは神岡にある 9PB に蓄積され、神岡施設での計算機資源に VPN でログインして直接扱うか、VPN 経由でファイルをコピー(scp)することで利用している。

検出器で観測されると期待される物理事象を、モンテカルロシミュレーションプログラムを用いて大量に生成する。モンテカルロシミュレーションプログラムは、検出器が生成するデータを模擬するものであり、これを実際の観測データと全く同じように処理することで、シミュレーションデータについても DST を生成する。統計のふらつきの問題を除くため、通常は実際のデータの数百倍程度の事象を生成する。例えば、大気ニュートリノ観測であれば、約 20 年分の観測データを解析するために 2000 年分相当（およそ 2000 万事象）程度の生成、処理を行う。

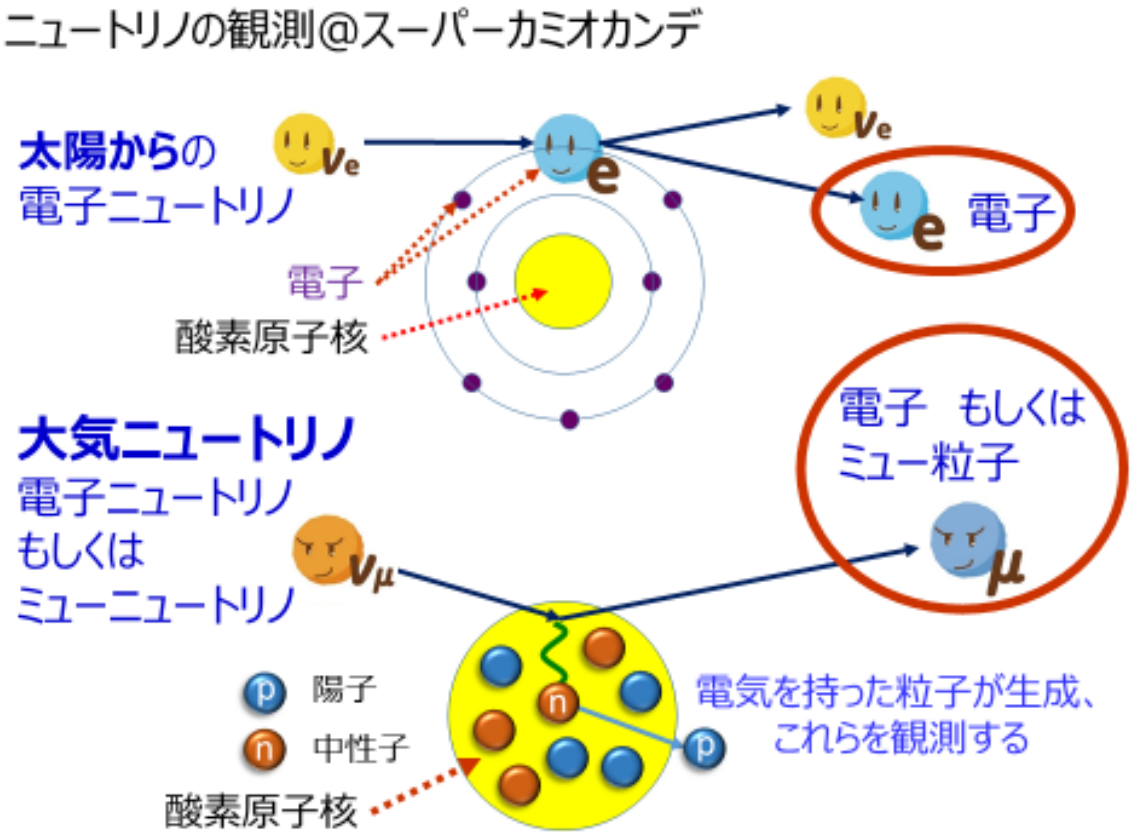
大気ニュートリノ事象のシミュレーションについては、その生成と高速な事象再構成は１事象１分程度以内で終了するため、神岡施設でも数週間程度でその生成を完了することができる。 一方、高精度な事象再構成ソフトウェアは、１事象５分程度を要するため、神岡施設の計算機資源では処理することができず、海外の複数の研究機関の計算機資源を用いて生成する必要がある。ここで生成したシミュレーションデータファイルは、総容量で数 TB を超えるが、これを神岡施設に転送、保管する必要があり、効率的かつセキュアな生成や転送が課題となっている。
事象シミュレーションに用いるプログラムや、事象再構成に用いるプログラム、入力に必要なプログラムは実験グループで開発したものとなり、すべてが一般公開されているわけではない。これを海外の研究機関やパブリッククラウドで実行するとき、各計算ノードに配布、展開する必要があるが、その配布方法には認証などの課題がある。
既に述べたように、ソフトウェアは世界各国の研究者が共同開発しており、これを効率的に行うため、svn や git などのソフトウェアレポジトリ（バージョン管理ソフトウェア）が用いられている。レポジトリそのものの管理だけでなく、ユーザの登録・削除、ならびにアクセス権の管理などが大きな課題となっている。

近年は、実験の結果だけでなく、そこで用いたデータなどもできるだけ一般に公開することが求められている。事象データそのものを公開することは、サイズ的な問題もさることながら、これを扱うには検出器の理解も必須となるため、あまり行われてはいない。
一方で、物理現象の解析に必要となるための解析済ファイルなどは公開されることが増えている。これらは通常数 GB～数百 GB となるが、その公開は機関の web サーバなどを用いて行われることが多い。読み出しについてはだれでも可能とする。
これらのデータは十分な速度で転送できる環境を整える必要があるばかりでなく、データの信頼性担保のため、正当な物であることを示すための技術も必要となる。さらには、事故もしくは故意の書き換えが発生しないようにするための高い管理技術やセキュリティ体制も求められる。

研究にあたっては、グループ内外で数多くの会議が開催される。そこで用いられた発表資料や文書などを共有するためのソフトウェアも必要となる。(indico、Docdb など)

論文やグループ内での報告書などを共同で編集することも必要となる。これも世界各国の研究者が効率的に作業できるような環境が必要となる（svn/Office365/GoogleDocs)。

コミュニケーションツールとして、従来はメールが用いられていたが、最近は Slack などのツールも利用されるようになっている。ここに蓄えられたコミュニケーションの記録などを適切に保管し、のちにアクセス可能な状態に保つことも重要となっている。



スーパーカミオカンデにおける素粒子観測

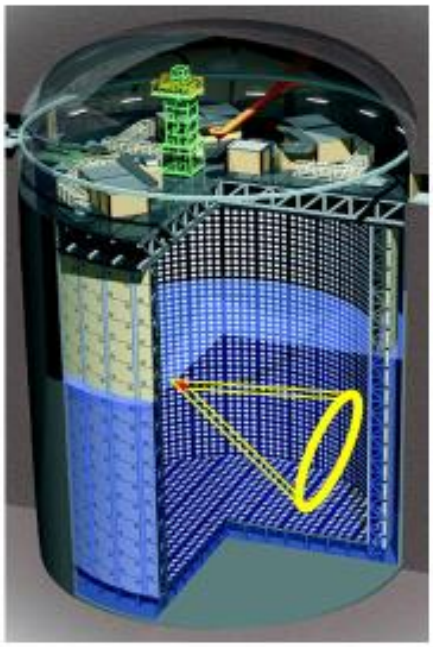
事象再構成の手法　チェレンコフ光の量と到達時間を利用

チェレンコフ光量
∝ 粒子運動量

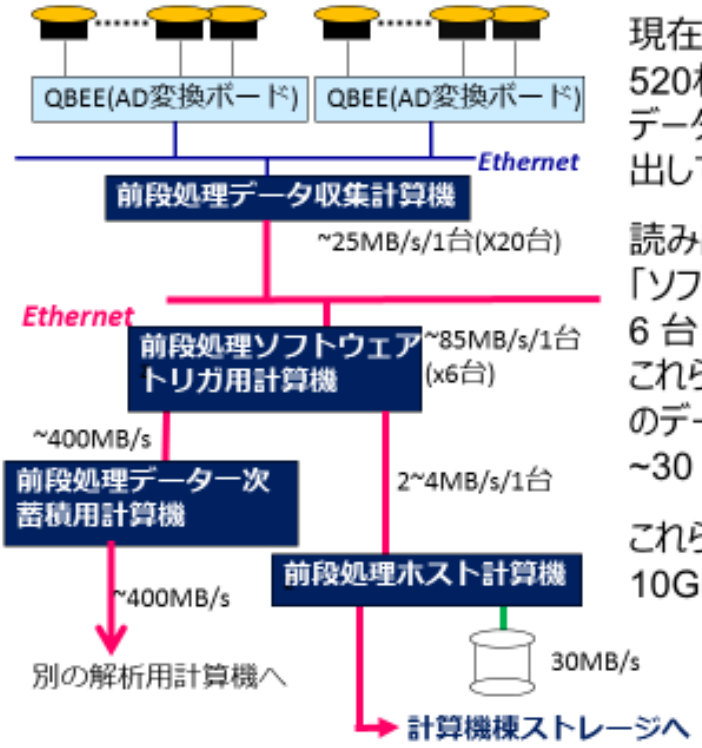
事象発生位置
~荷電粒子が走り出した点
→ 光子のPMTへの到達時間
リングパターン（広がり）

荷電粒子数（+γ）
→ チェレンコフリング数

荷電粒子種別
→ チェレンコフリング形状
μ・π・p　シャープなリング
e・γ　ぼやけたリング



スーパーカミオカンデのデータ収集システム

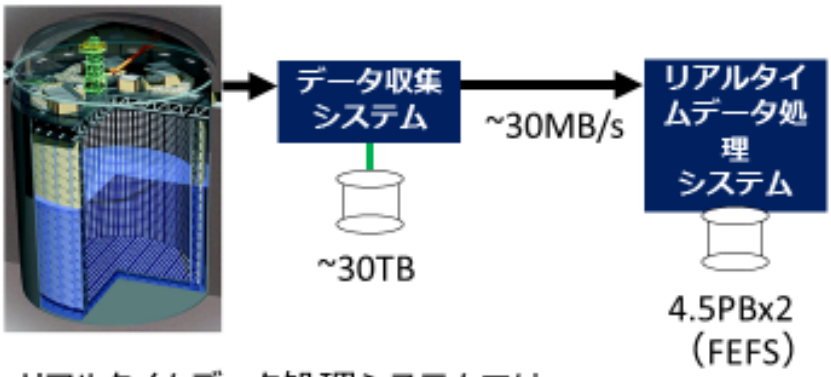


現在のデータ収集システムは、520枚のQBEEボードからのデータを、20台のPCで読み出しています。

読み出されたデータを処理する「ソフトウェアトリガー」計算機は6 台（48プロセス）。これらの計算機で約400MB/sのデータから、事象を選別、~30 MB/sに減らします。

これらの計算機は、GbEや10GbEで接続されています。

スーパーカミオカンデのデータ処理



リアルタイムデータ処理システムでは

１）データフォーマット変換

30MB/s → 16MB/s （圧縮込）

２）ノイズ除去

事象再構成したうえでノイズと思われる事象を捨てる

16MB/s → 2.5MB/s

を行い、一日あたり 200GB 程度の生データを保存している。

（永久保存）

スーパーカミオカンデのデータ解析

大気ニュートリノ＆陽子崩壊事象探索

大気ニュートリノ事象や陽子崩壊事象の探索は、

１）チェレンコフリング数～観測粒子数

２）粒子発生位置

３）各粒子の種別

４）各粒子のエネルギー、方向

を「再構成」することで行う。

基本解析：１事象～ ４０秒/core

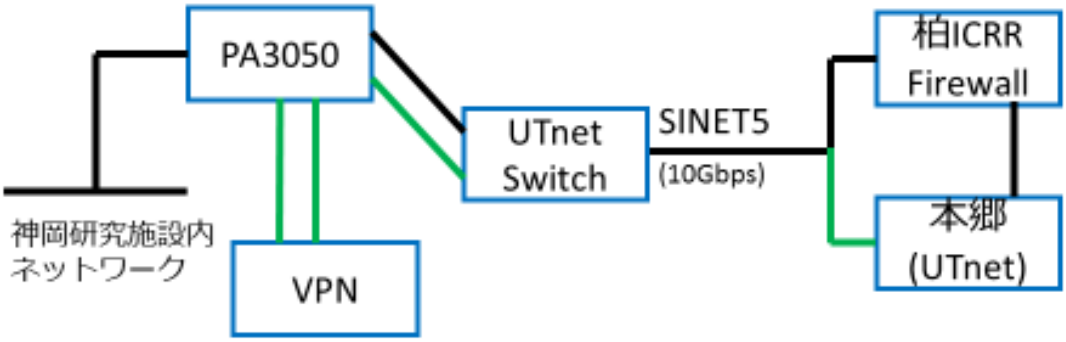
発展解析：１事象～３６０秒/core (Xeon 2.4GHz)

大気ニュートリノ事象 ~10事象/日

実データファイル ~2.5GB/年

解析結果ファイル ~80MB/年

神岡のネットワーク環境

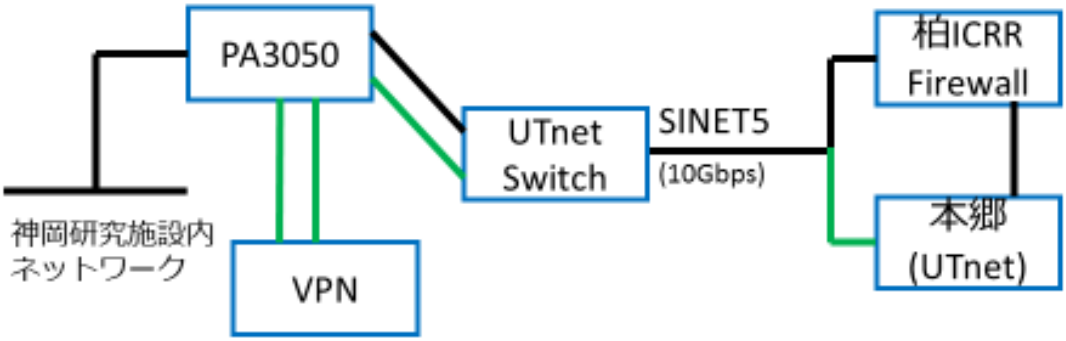


メインの経路は柏（宇宙線研究所）経由　1Gbps
う回路として、本郷経由　1Gbps　を利用している。

学外からのデータ/sshアクセスは基本的にVPN (AnyConnect) のみ。
特定ホストからの ssh を認める host も DMZ 内に存在するが、大容量
ディスクは直接見えないため、データコピーには不向き。

今後はシミュレーションデータ生成、解析のために　~200TB　程度の
データのやりとりを海外の研究施設と効率的に行える仕組みが必要。
（もってくるだけなら iRODS などを既に利用しているが、配る方が難しい）

神岡のネットワーク環境



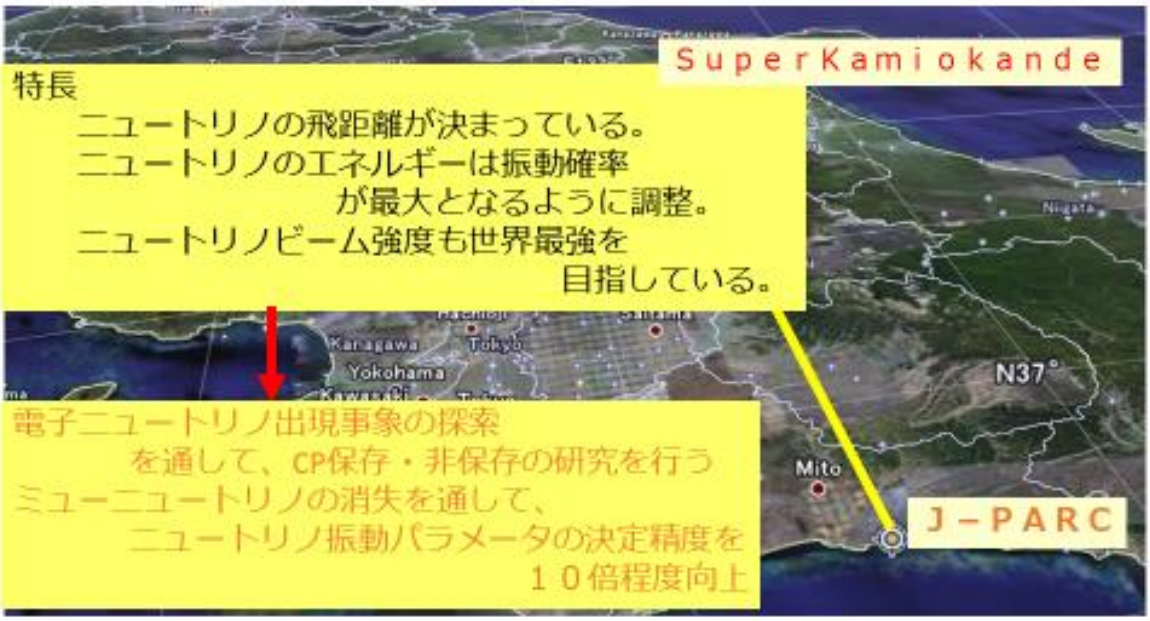
一般公開用データは、サイズが小さいものが多いため、現状では
web serverに置くことで対応している。

特定ユーザに公開するファイルについて owncloud も導入しているが
アカウント管理、公開範囲（権限）管理などが煩雑であることが課題。

一人のユーザが複数の実験グループに参加している
～特定の実験グループにだけで共有したいファイルも多い

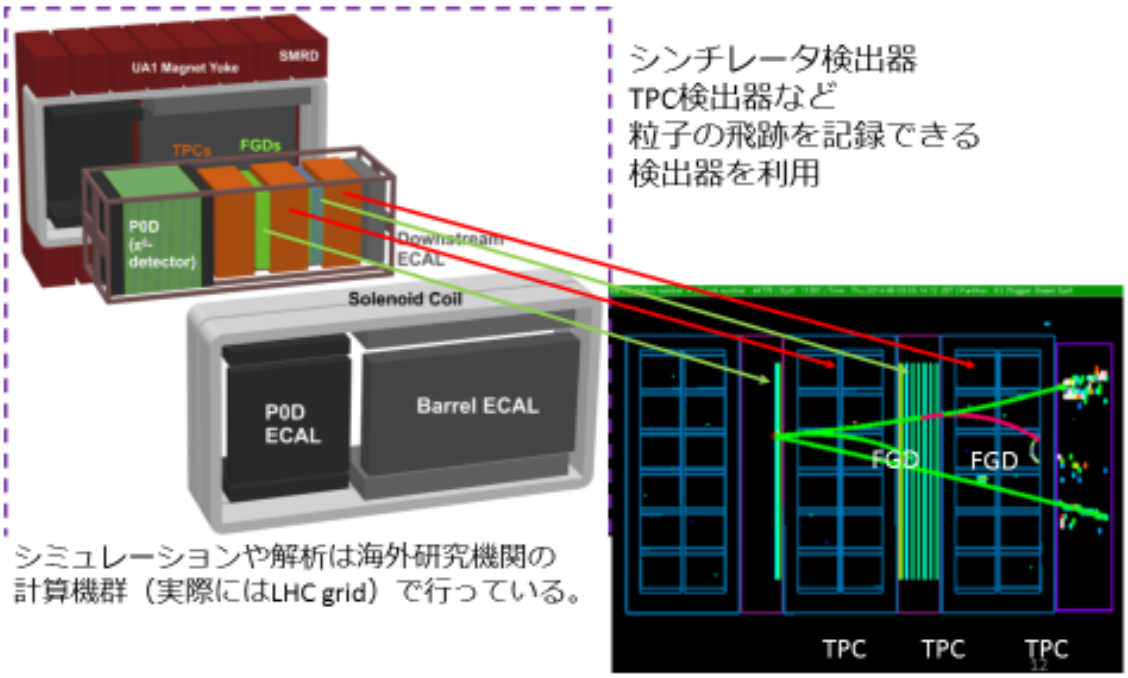
Ｔ２Ｋ実験

Ｊ－ＰＡＲＣ陽子加速器（東海村）を用いて
大強度のミューニュートリノビームを生成
２９５ｋｍはなれた、スーパーカミオカンデに打ち込む



Ｔ２Ｋ実験

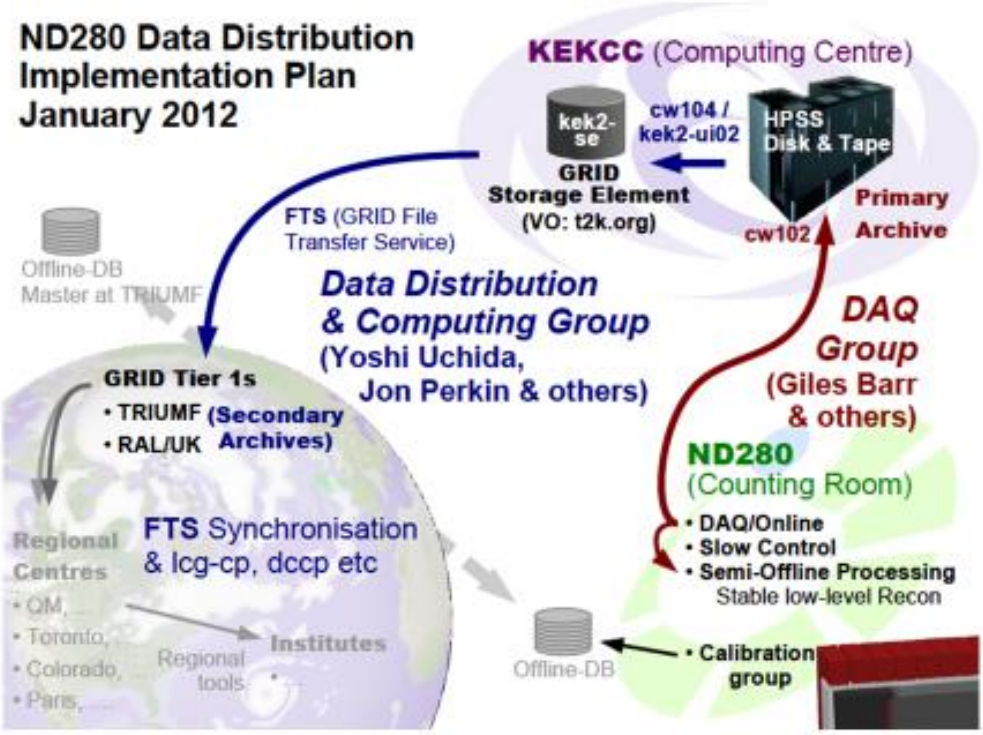
前置検出器～
J-PARC（東海村）に設置されたニュートリノ検出器



シンチレータ検出器
TPC検出器など
粒子の飛跡を記録できる
検出器を利用

シミュレーションや解析は海外研究機関の
計算機群（実際にはLHC grid）で行っている。

Ｔ２Ｋ実験



Ｔ２Ｋ実験

前置検出器・加速器のデータ

加速器（陽子）情報	~70TB (2010 ~ 2017)
前置検出器のデータ	~60TB (50GB/day)

シミュレーションデータ

ニュートリノビーム	~30TB（2010~2017)

実データ処理速度

前置検出器	5秒/1spill ~ 48hrs/1日分のdata

シミュレーション生成速度

ニュートリノビーム	1000 CPU hours (1 set)
前置検出器	現状 1 sec/CPU
	$2x10^9$ 事象程度を生成
	~ $2x10^9$ sec CPU ~ 20,000 days CPU
スーパーカミオカンデ	現状 360 sec/CPU
	10^6 事象程度を生成
	~ $4x10^8$ sec CPU ~ 4,500 days CPU

前置検出器のシミュレーションデータ生成はEU と カナダの GRID を利用

スーパーカミオカンデを使った実験での課題

シミュレーションを海外研究機関と
協調して行う場合のデータのやり取り
データサイズは数十ＴＢオーダ
少数の特定ユーザが行うことは可能

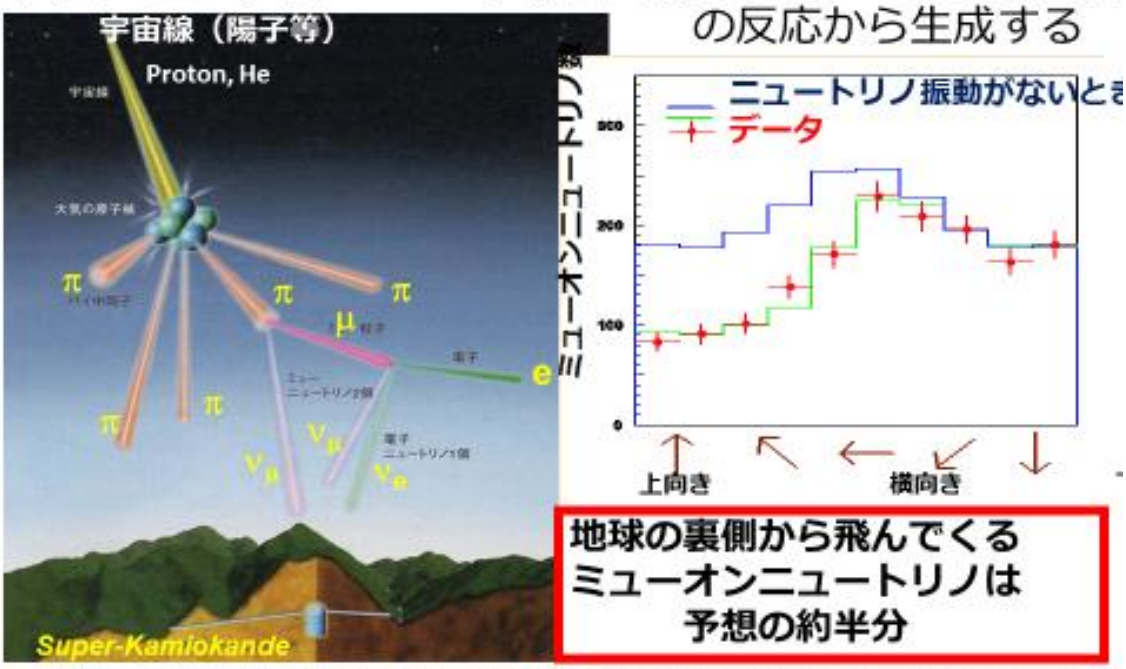
研究者間での解析データ共有
データサイズは～100GB
共同利用者は数百人、出入りも多い
～アカウント管理が大変

一般向けデータ公開
~100MB程度、web server で現状はなんとかなっている。
公開データが増える見込みがあるため、将来的に検討
が必要となってきそう。

ソフトウェア（ソースコード）共有
自前の svn サーバ
github (海外研究機関が契約中)

ドキュメント（プレゼンテーション資料、会議録など）
の共有方法も課題･･･

大気ニュートリノ　~　宇宙線（陽子など）と空気（窒素
の反応から生成する

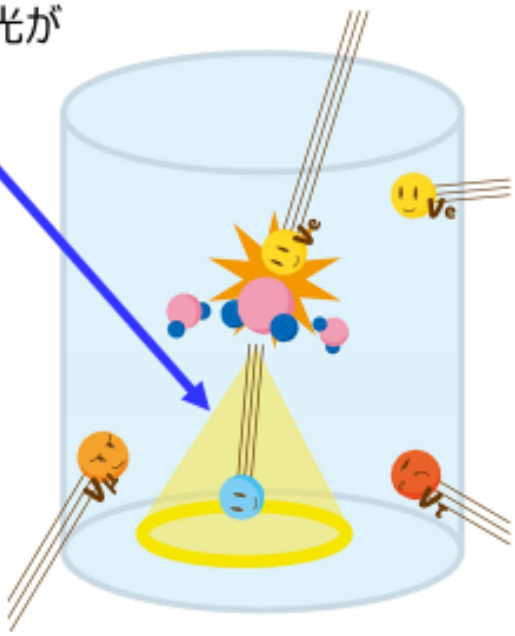


ニュートリノのエネルギーや方向と減り方の違いを詳細に調べる
⇒　ニュートリノの質量や混ざり方がわ

スーパーカミオカンデにおける素粒子観測

水中を電気をもった粒子が超高速でとぶと、
進行方向に円錐状のチェレンコフ光が
出ます。（本当は青い光）

実際に観測しているのは、
ニュートリノが弾き飛ばした電子や、
ニュートリノが原子核と反応して生成する
電子やμ粒子などの粒子です。

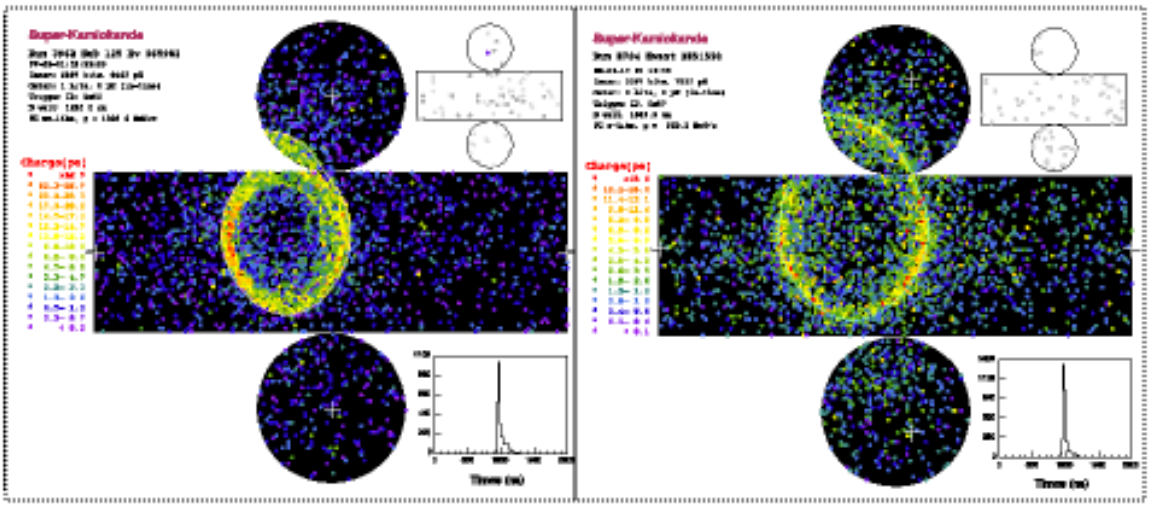


チェレンコフ光の量はわずかなので、
検出器の中は暗く保ちます。
超高感度の大型光センサー
（光電子増倍管）を水槽壁面
に大量につけることで検出します。

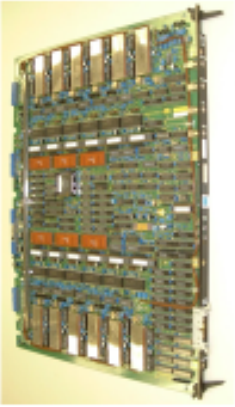
スーパーカミオカンデにおける素粒子観測

チェレンコフ光の分布（見え方）が
ミュー粒子と、電子で違う性質を利用して
ニュートリノの種類を判定しています。

ミュー粒子　　電子



スーパーカミオカンデ(初期)のデータ収集システム



各ATMボードから出力される「同時PMT信号検出数」の全ての和をとって、一定値を超えたときにデータを取得するようハードウェアで構成されていた。（アナログ信号）

データ収集用ボード120枚を1台のワークステーションで読み出していた。

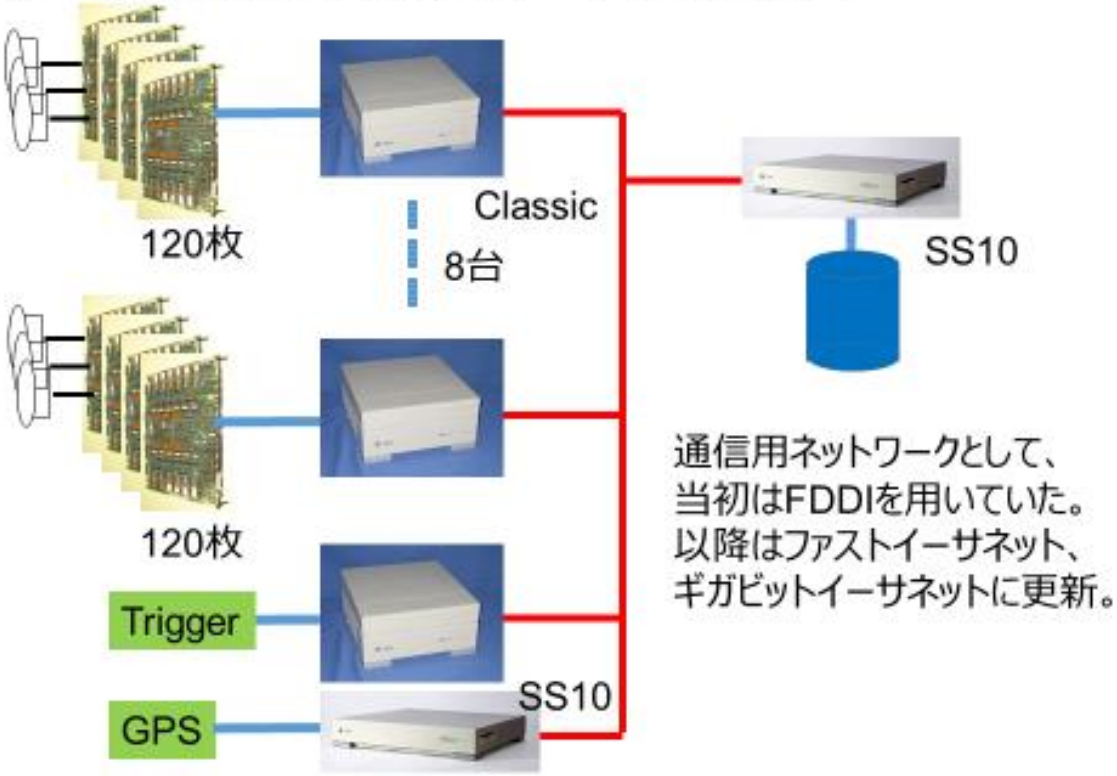
機材

1996年 ： Sun SPARC Classic (40MHz)
（後に SPARC Station 5, Ultra 5
などにアップグレードして、2012年まで稼動）

これ以外に、外水槽の光電子増倍管データを読み出す計算機、データ取得時の条件や前の事象との時間差の情報を読み出す計算機、GPS情報を読み出す計算機が用いられた。

これら読みだされたデータを一つのまとめるための「イベントビルダー」と呼ばれる計算機が存在。
こちらは、SPARC Station 10 や Ultra 10など性能の高いものを利用。

スーパーカミオカンデ(初期)のデータ収集システム



スーパーカミオカンデのデータ収集システム

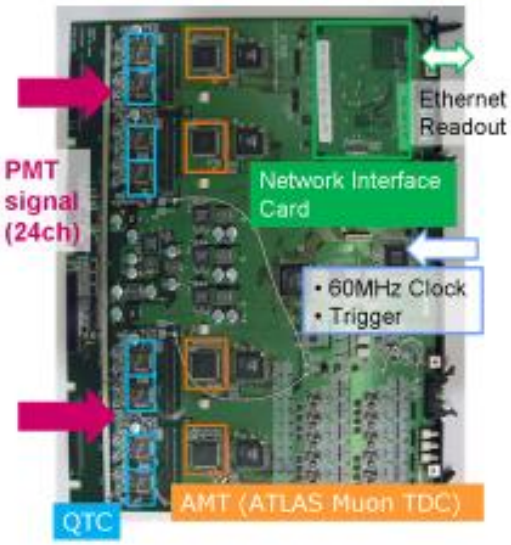
年数がたって、ATMボードは修理が困難に
より高性能な新システムの開発

電荷を計測するための「QTC」と呼ばれるASICから開発

QTC(CLC101)

各ボードで２４本のＰＭＴの信号を
読み出せるように変更（従来の倍）

各ボードがFast Ethernetを持ち、
データを直接計算機で読み出せる
ように変更。



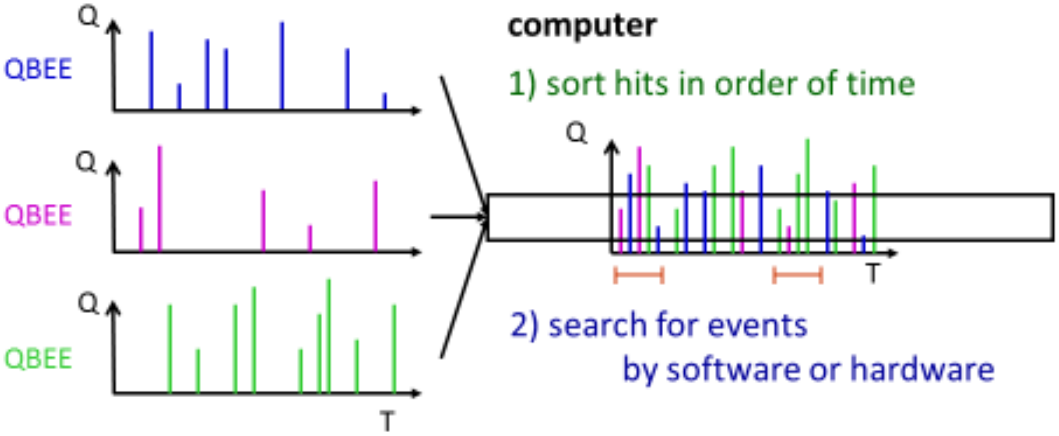
スーパーカミオカンデのデータ収集システム

データ収集のエネルギー閾値を大幅に下げる

T2K実験では、ビーム到着時の事象をバイアスなく残したい

旧システム：幅は1μs decay electron tagでは40μs程度
T2Kはビーム到着前後1ms

最初から全部のPMT hit を記録してしまい、ソフトで事象探索

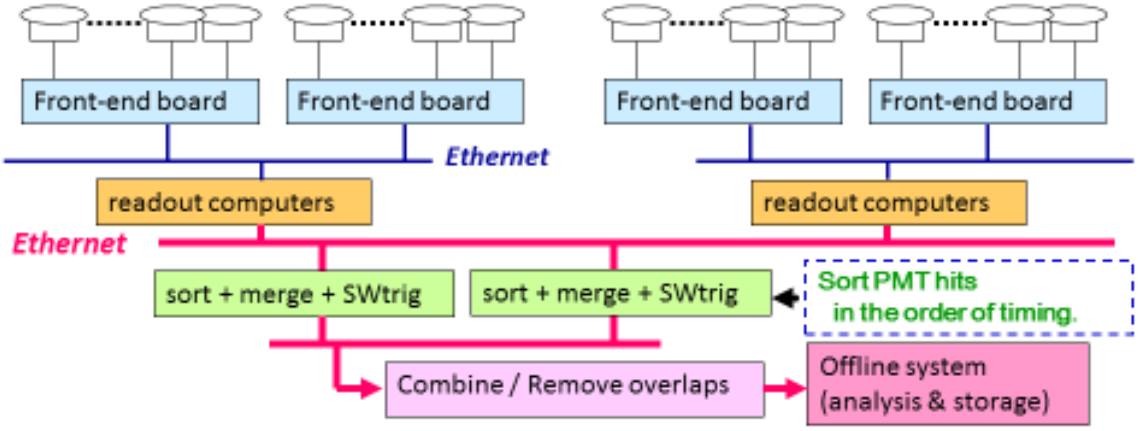


スーパーカミオカンデのデータ収集システム

1) 光センサーからの ~ 1/4 p.e. 以上の信号を常時全てデジタイズ
（電荷 & 時間）

2) 全ての 電荷と時間 のデータを PC から読み出す

3) ソフトウェアでトリガーをかける
時間順に全てのヒットを並べなおして
同時にPMTがなっていないかを探す。



ボードあたりのデータ量 ~1MB/sec * 約520枚 ~ 500MB/sec

スーパーカミオカンデのデータ収集システム用計算機の比較

1996~		2012~	
データ収集用		データ収集+事象選別用	
SPARC Station Classic x 9		IA32(64bit) server x 36	
CPU	**50MHz** (micro SPARC)	CPU	**2.27GHz x2** (Xeon L5640)
Memory	**64MB**	Memory	**24GB**
HDD	1.5GB	HDD	150GB
データ集積処理用		データ集積処理用	
SPARC Station 10		IA32(64bit) server	
CPU	**50MHz x 2** (Super SPARC)	CPU	**2.67GHz x2** (Xeon X5650)
Memory	**112MB**	Memory	**64GB**
HDD	**14GB**	HDD	**15TB**

スーパーカミオカンデのデータ収集システム

加速器からのタイミング信号を用いた
ニュートリノビーム到着時間によってデータを記録
（ビーム通過＝データ取得後、
数十ミリ～数秒程度以内にGPS情報が到着　）

二箇所のGPSの時間が「同じ」であることは、どう保証する？

コモンビューGPS受信機の利用（EW9955AR）
～　同一GPS衛星からのデータを用い、
日本標準時と比較することができる
（NICTが基準データを提供）
神岡と東海にコモンビュー受信機を設置、
それぞれのGPS受信機の時間差を
確認することが可能。

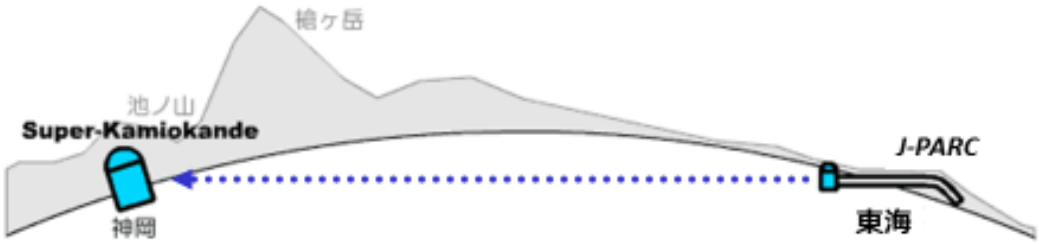
スーパーカミオカンデのデータ収集システム

事象の発生時間（絶対時間）　←　GPSを用いる。
~200ns程度の精度で決定可能

太陽ニュートリノ観測において、太陽方向を知る
超新星ニュートリノ到達の時刻を知る
長基線ニュートリノ実験（及び大気ニュートリノ観測）において、
加速器からのニュートリノによる事象かどうか同定する。

ニュートリノビーム射出時間を加速器側で記録

T2K実験では、ニュートリノビーム射出時間をすぐ神岡に転送、
SKにニュートリノが飛来する時間近傍のデータを全て記録
（SINET　L2VPNを利用、数十ms以内には転送できる）



A.2. 藤田委員

【宇宙航空研究開発機構 調布】
データを配布する側の事例
・一般向けデータ公開（0(100KB)）、web や scp
・一般向けデータ公開（0(～100KB)）、ソースコード

A.2.1 概要

宇宙航空研究開発機構(JAXA)調布では、スーパーコンピュータの利用において歴史的に、数値シミュレーションデータを大規模・大量データとして蓄積・利用してきた。図 A.2-1 に JAXA スパコンの歴史を示すが、1980 年代の開拓の時代から、現在は実用の時代へとその役割を変化させている。この変化に伴い、大規模なファイルに対する高速なシーケンシャルアクセス性能だけでなく、ものづくりの際の設計に不可欠なパラメトリックスタディ値計算でのデータアクセスパターンである小規模大量データへのアクセス性能も求められるようになった。2016 年の JSS2(JAXA Supercomputer System generation 2)の稼働を機に、衛星データの再処理も行うようになり(図 A.2-2)、小規模大量データの処理能力とデータ共有機能がより一層求められる時代になっている。

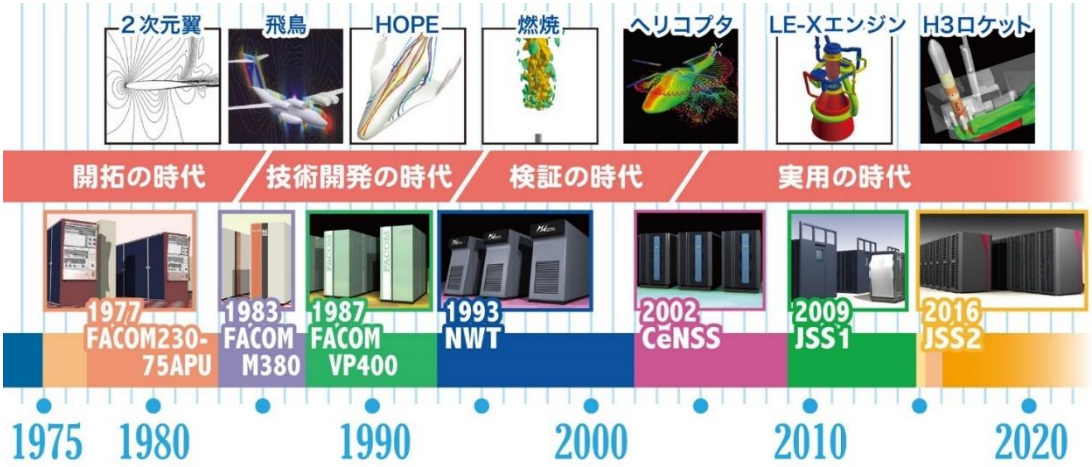


図 A.2-1　JAXA スパコンの歴史

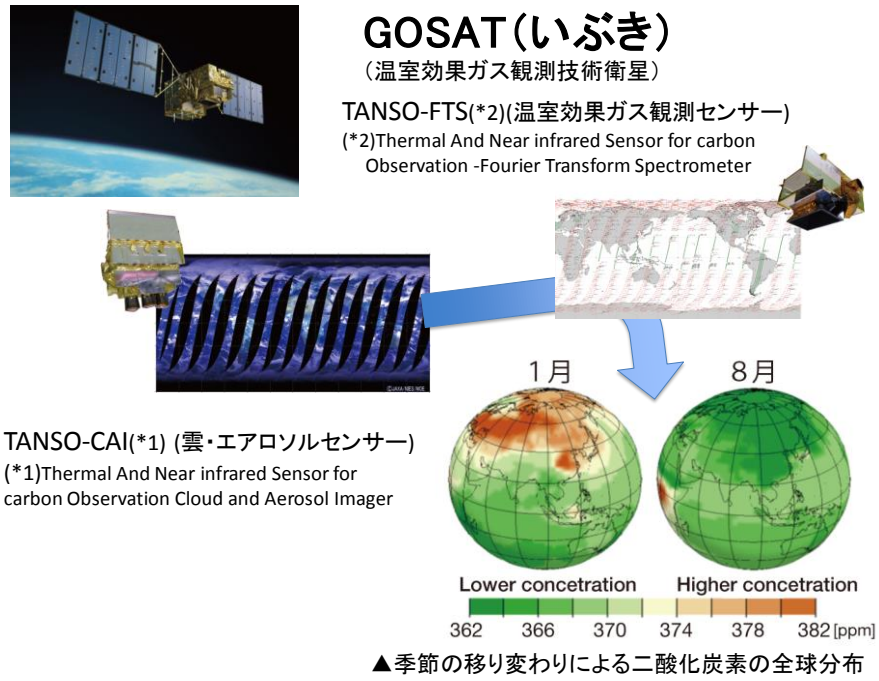


図 A.2-2　衛星データの再処理

図 A. 2-3 に、JSS2 のシステム構成図を示す。この中で、「ファイルシステム部(SORA-FS)」が 7PB の容量を持つ SRFS ファイルシステム、「アーカイバ部」が 60PB の容量をもつテープアーカイバであり、データ管理の中心的役割を担っている。

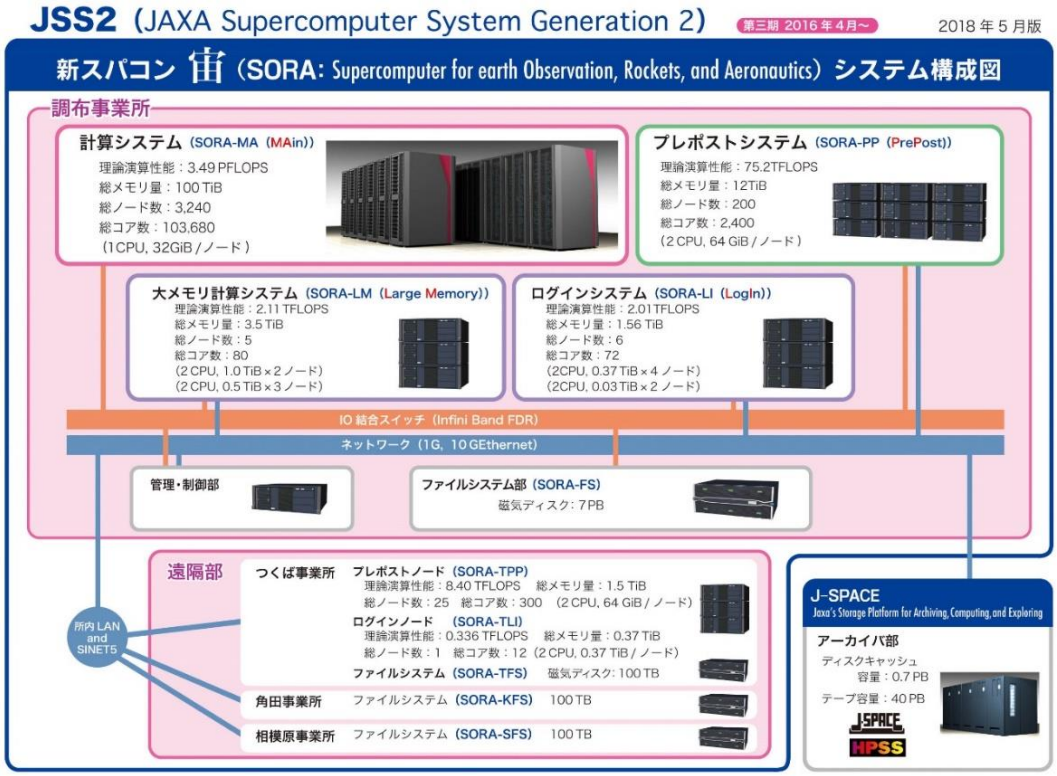


図 A. 2-3　JSS2 システム構成図

2016 年から 2018 年にかけての、アーカイバに蓄積されているデータのファイルサイズと数を図 A. 2-4 に、ファイルサイズ毎のデータ総量を図 A. 2-5 に示す。年を追うごとにファイル数や総容量は大きくなってきているが、サイズ 16MB 前後のファイル数が多いこと、総ファイル量はファイルサイズ 64GB 程度のものが多くを占めていることは、傾向として変わっていない。

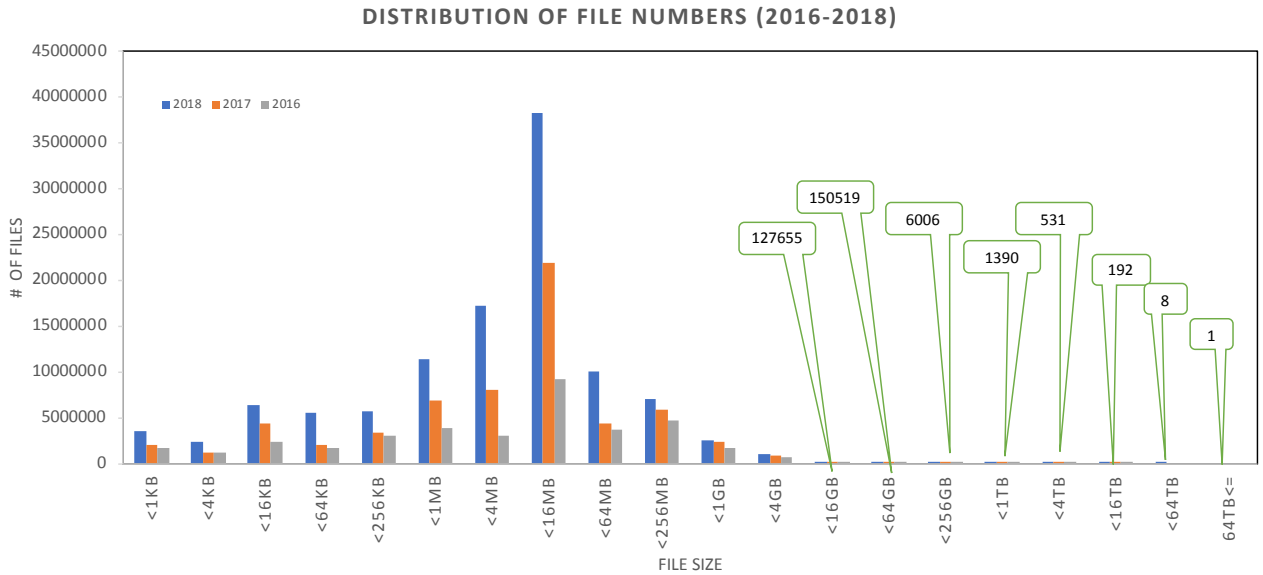


図 A. 2-4　アーカイバ内のファイルサイズと数

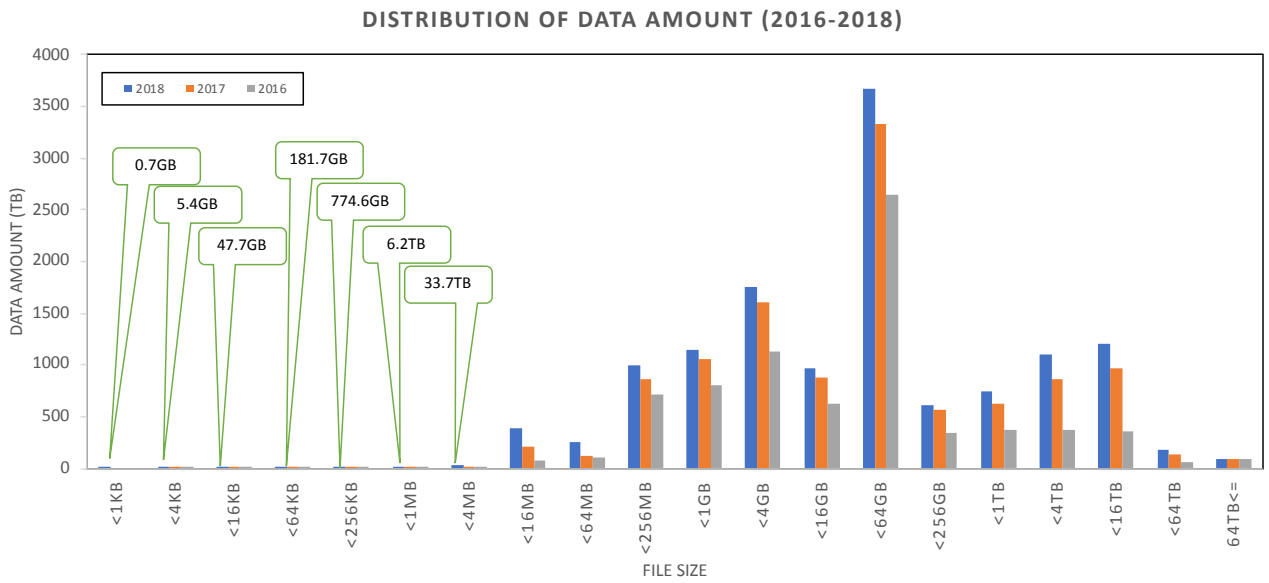


図 A. 2-5　ファイルサイズ毎のデータ総量

A. 2. 2 JAXA のデジタルデータ運用

前節では、スーパーコンピュータ(JSS2)でのデータ保管について紹介したが、本節では、JAXA 全体でのデジタルデータ運用に関して、その一部の活動を簡単に紹介しておく。

デジタルデータの保管場所（サーバ）は、データ種別毎には局所化しているが、JAXA 全体で見れば、多数の事業所にデータが分散して存在することになる。また、その総容量は PB のオーダーになっており、データ保持のための設備維持に多額の固定経費が必要である。例えば、次の通り。

数値シミュレーションデータ(7PB+60PB)・・・調布
衛星データ(5PB)　・・・・・・・・・・・・・・・筑波・鳩山
科学衛星データ(3. 3PB)　・・・・・・・・・・・相模原

また、データ種別単位に運用体制が存在し、データの健全な管理を行っているが、運用体制の集約による人的コスト集約(というか、そもそも人手不足の運用部門の立て直しの一手法としての集約)も、考える時期に来ている。

データにアクセスする研究者（クライアント）は、上記事業所だけでなく、図 A. 2-6(や、場合によっては図 A. 2-7 ないしは、NASA, ESA 等の協力機関)の場所からデータにアクセスする。

図 A. 2-6　JAXA 国内拠点

図 A. 2-7　JAXA 海外拠点

A.2.3 データ管理の課題

ここからは、JAXAにおけるデータ管理に関する課題を次の4つに分類し、順番に述べる。

(1)事業所を跨いだデータ管理
(2)データ公開の仕組み
(3)小サイズ大量ファイルのI/O性能
(4)ライフタイムの長いデータ管理ツール

(1)事業所を跨いだデータ管理

A.2.2.で述べた通り、JAXA全体を見渡した場合、多数の事業所にデータが保管され運用されている。これに起因する課題がいくつか存在する。

(1a)データ移動

データ処理に利用可能な計算機(CPU資源)が複数の事業所に存在する場合、計算機(CPU資源)の空き具合に応じて、処理させる計算機を切り替えることがある。この時、処理データをその計算機(CPU資源)のある事業所に移動する必要があるが、これにコスト(回線費、転送時間)がかかる。(そもそも、計算機／データの分散配置という設計が良くないのかもしれない。)

(1b)データ共有

データ移動は、基本的には人手を介して行われるが、これを自動化する意味で、データの事業所間での共有のニーズがある。過去に「SRFS on Ether」を制作し、事業所間を跨いだファイルシステムマウントの運用を行ったことがあるが、回線品質の悪さによる(と推測される)マウント切れが主たる理由で、現在では事業所間跨ぎのファイルシステムマウントの運用はしていない。「SRFS on Ether」から10年が経っているので、再挑戦する価値はあるかもしれない。

(2)データ公開の仕組み

公開範囲(ACL)のきめ細かい設定、特別なツールを使わない高速ダウンロード、データ処理計算機からデータ公開用サーバへのデータ転送というオーバーヘッドの存在等がデータ公開の際に問題になっている。

(3)小サイズ大量ファイルのI/O性能

データ共有に限った話ではないが、ファイルシステムは小サイズのファイル操作に対して性能を発揮できない。データ共有の仕組みでも同じことが言えるため、なにがしかの仕組みの導入により、高速に処理できるようにすることが重要である。

(4)ライフタイムの長いデータ管理ツール

ファイルシステムやアーカイブシステムが換装されても、ユーザからみたインターフェイスは変わらない、というデータ管理ツール(Put/Get、Archive/Stage)が欲しい。(ftp,scpが、ライフタイムが長いという意味で該当するのか。でも、速度面で不満がある、大量のファイルを管理するには工夫が必要なのも難点。)

以上

A.3. 中村委員

【高エネルギー加速器研究機構】
・一般向けデータ公開（O(100GB)）、解析環境、解析ソフトウェア、ドキュメント
・一般向けデータ公開（O(1PB)）、データファイル
・共同研究者向け大規模データ共有（O(100 PB)）、実験データ、処理済みデータ

<u>高エネルギー実験におけるデータ処理の現状</u>

現代の大規模粒子加速器を用いた高エネルギー素粒子・原子核実験は、加速器の高輝度化に伴うデータ取得レートの上昇と、測定器の大型化および検出器の高精度化によるビーム衝突事象ごとのデータサイズの増大により、取得する総データ量は拡大の一途を辿っている。実際に、欧州合同原子核研究機構（CERN）に設置されているLarge Hadron Collider（LHC）加速器を用いて行われるATLAS実験［1］で取得する生データ量は、2018年の時点で年間60 PBに達している。2026年からLHC加速器を増強して行われることが決定しているHigh Luminosity LHC（HL-LHC）加速器による実験では、総データ量が現状のLHC実験の10倍になることが見込まれている。また、2018年から高エネルギー加速器研究機構（KEK）で開始した、SuperKEKB加速器を用いたBelle II実験［2］では、数年間で100 PB程度の生データを蓄積することを予定している。

取得された生データは、検出器直近に配備するフロントエンド・エレクトロニクスにより即座にデジタル化され、事象選別された後にビーム衝突事象ごとに結合され、物理的に興味のある対象を含む事象のみが抽出され記録される。記録された生データは通常、加速器が設置されている研究所の計算機設備内のテープ・ライブラリに恒久的にアーカイブされると同時に、災害時に計算機設備が被害を受けた場合のデータ喪失を防ぐために、世界各国の研究所に配備されている計算機設備内のテープ・ライブラリに、国際学術ネットワークを通じてレプリカ・データが分散アーカイブされる。

生データへのアクセス頻度は高くはない。記録された生データは、テープ・ライブラリにアーカイブされる前に再構成し物理量に変換される。その後、対象とする物理解析に応じてデータを分離し、物理解析に使用するデータフォーマットへ圧縮と変換がなされる。生データへのアクセスは、検出器校正定数の更新や、再構成アルゴリズムに変更の必要性が生じた場合に限られる。一方で、物理解析用のデータは、世界各国の研究所と大学の計算機設備に分散され、繰り返し処理される。

物理解析用の実データは、検出器の構成が複雑であるため、実データと同じ統計量に相当するモンテカルロ・シミュレーションデータを最低限生成し、比較と評価がなされる。このシミュレーションデータの生成は主に、加速器研究所の外部サイトの計算機設備を用いて行われる。これら多数のサイト間での実データおよびシミュレーションデータの共有と計算機資源の相互利用は、Interoperable Global Trust Federation［3］のポリシーに従うグリッド認証局が発行する証明書を用いた認証・認可にもとづいて行われている。

<u>今後の課題（新しい認証・認可基盤の導入、データの恒久的な保存、データの一般公開）</u>

限られた計算機資源により拡大する総データ量を処理していくために、既存のグリッドサイトだけでなく、HPC計算センターや、クラウド計算機資源を積極的に利用できる環境が模索されている（eXtreme DataCloud［4］）。これらの計算機資源と分散されたデータを、データレイクなどの新興技術を用いて統一的に取り扱うためには、現在グリッド基盤で用いられているX.509証明書のみにもとづいた認証・認可基盤（AAI）からの脱却が求められる。グリッド認証局の維持とユーザ管理はコストが高いだけでなく、エンドユーザーにとってはX.509証明書の取り扱いの敷居が高いことも理由となっている。次世代のAAIとしては、ユーザが所属する各組織のIDを連携させることによるシングルサインオンの実現（Federated Identity Management for Research［5］）と、エンドユーザーの利便性を高めるためのSAML認証やOAuthを用いたトークンベースのAAIの採用が候補として検討されている（Authentication and Authorisation for Research and Collaboration［6］）。

加速器実験で取得するデータ量は膨大なため、過去に行われた実験データの恒久的なアーカイブを続けるための様々な共同研究が、Data Preservation in High Energy Physics［7］の枠組みにより行われている。大型加速器の運転は電気料金などのコストが高く、同じ目的で再運転することは容易でないため、一度取得したデータを将来にわたって再利用できることが望ましい。また、計算機器とその利用技術の進歩や、データ解析手法や理論物理の発展により、過去に取得したデータの再解析による新発見の可能性も期待されている。これらを実現するためには、大量データのアーカイブ技術だけでなく、バーチャルマシンやコンテナを用いたデータ読み出しのためのソフトウェアを含めた計算環境の保全なども必要となる。

加速器実験データの一般公開は、データ量の多さから、配布のために必要なネットワーク帯域幅と、データ処理のために必要な計算能力が膨大となるため、現在では新規参入の研究者をトレーニングする目的、学生の教育目的、アウトリーチ活動の目的に留まっているのが現状である（CERN Open Data［8］）。一方で昨今は、データ解

析ソフトウェアなどの使い方を説明する文書を整備するだけでなく、Jupyter などのノートブックツールを用いて、データ解析環境と大量データへのアクセス方法を提供することにより、Web ベースの解析方法を通して、一般の人々が 1 PB 程度のデータにアクセスすることが可能となりつつある。今後は、その規模と簡便性を向上させていくことが課題となっている。

参考文献

[1] ATLAS 実験, 素粒子の質量を生み出す起源とされる、ヒックス機構の証拠となるヒッグス粒子を発見した実験。この実験結果により、2013 年にヒッグス機構を理論的に提唱した、Fsrancois Englert 博士と Peter W. Higgs 博士がノーベル物理学賞を受賞した。
https://atlas.cern/

[2] Belle II 実験, 粒子と反粒子の対称性（CP 対称性）の破れを実験的に示した、KEKB-Belle 実験の測定結果を、高統計のデータを用いて精密に検証するための実験。CP 対称性の破れの起源を理論的に予言した、小林誠博士と益川敏英博士は、2008 年にノーベル物理学賞を受賞した。
https://www.belle2.org/

[3] Interoperable Global Trust Federation (IGTF)
https://www.igtf.net/

[4] eXtreme DataCloud (XDC)
http://www.extreme-datacloud.eu/

[5] Federated Identity Management for Research (FIM4R)
https://fim4r.org/about/

[6] Authentication and Authorisation for Research and Collaboration (AARC)
https://aarc-project.eu/

[7] Data Preservation in High Energy Physics (DPHEP)
http://hep-project-dphep-portal.web.cern.ch/

[8] CERN Open Data
http://opendata.cern.ch/

LHC実験とグリッドコンピューティング
現状と課題

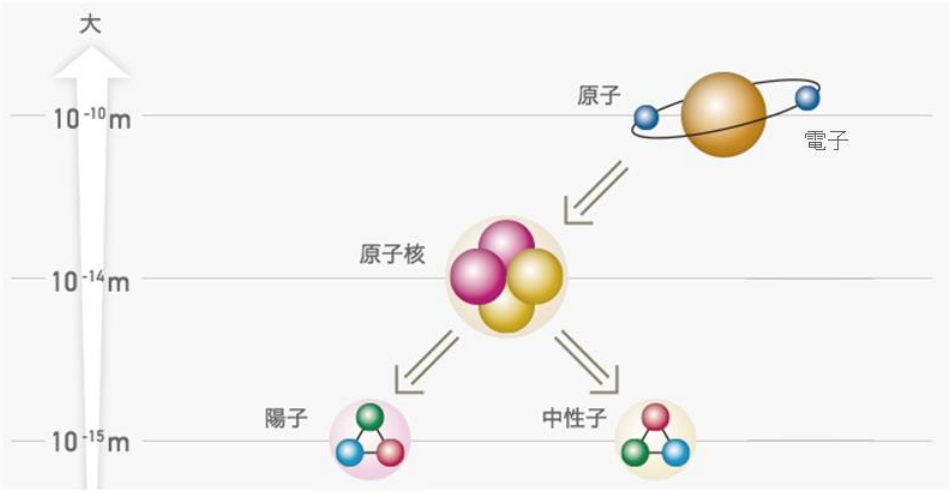
クラウド時代のデータ共有技術WG

中村智昭

KEK計算科学センター

KEK
Computing Research Center

原子・原子核



©ICEPP UTokyo

1987年に原子は物質の最小単位ではないことが分かった
1911年に原子核が発見

1919年に陽子，1932年に中性子が発見

素粒子標準理論

物質粒子

第1世代	第2世代	第3世代
u アップ	c チャーム	t トップ
d ダウン	s ストレンジ	b ボトム
ν_e eニュートリノ	ν_μ μニュートリノ	ν_τ τニュートリノ
e 電子	μ ミューオン	τ タウ

20世紀末までに発見済

力を伝える粒子

強い力	弱い力	電磁力
g グルーオン	W^+ W^- Wボゾン	γ 光子
	Z Zボゾン	

ヒッグス場に伴う粒子

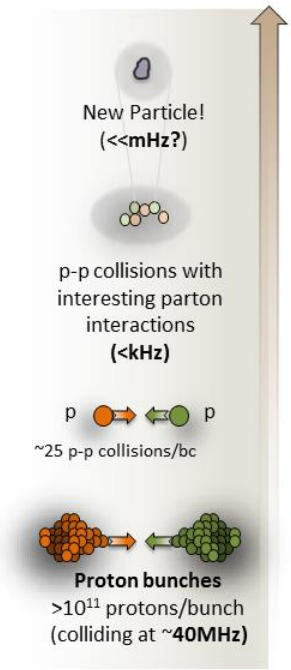
H ヒッグス粒子

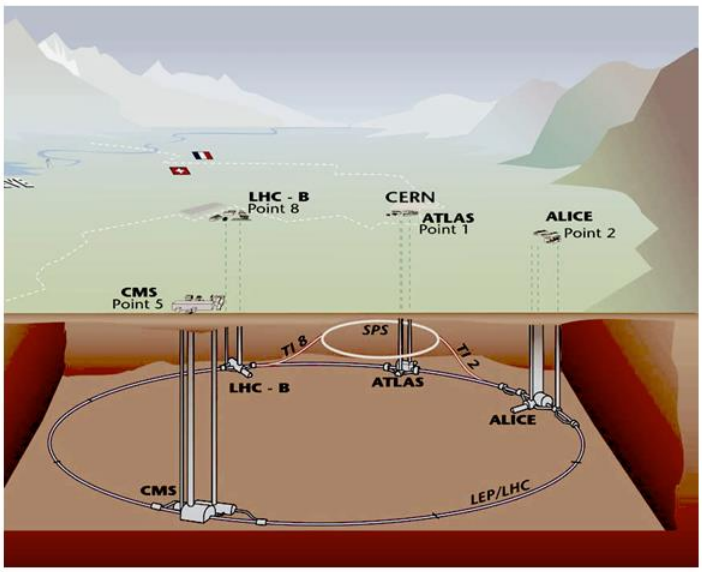
物質に質量を与えた起源とされる粒子
1964年に理論的に存在が予言

ヒッグス粒子の発見は
LHC実験の最も重要な目的

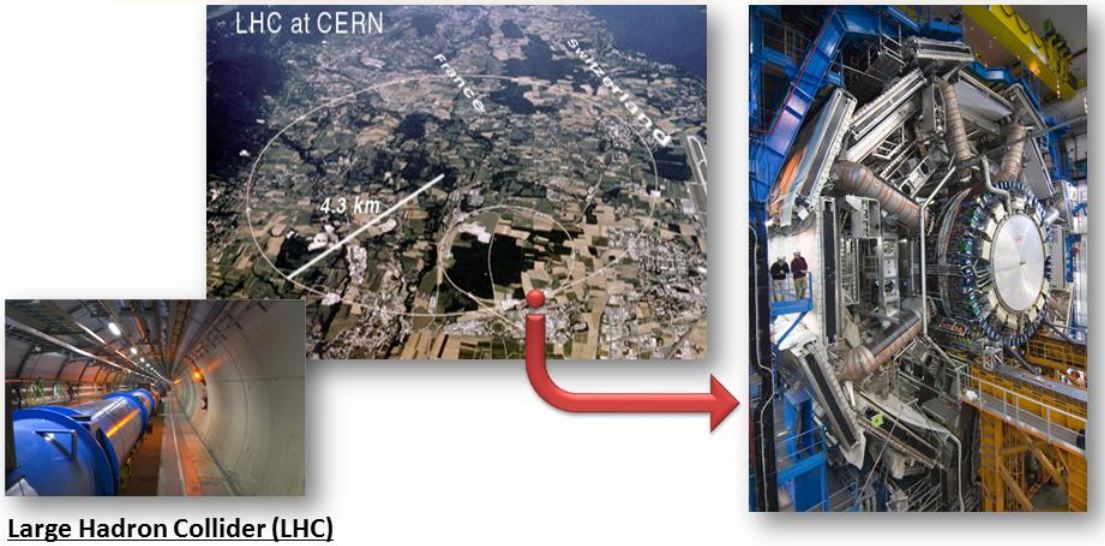
©ICEPP UTokyo

The Large Hadron Collider at CERN





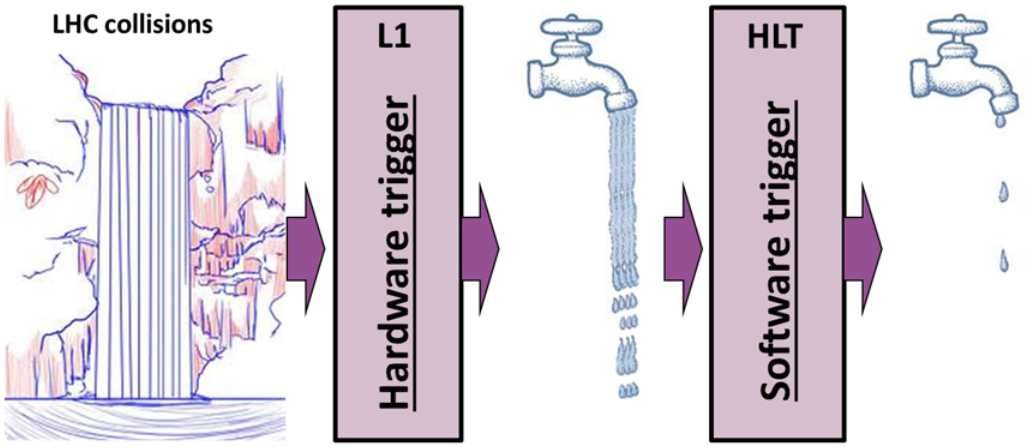
LHC-ATLAS実験



Large Hadron Collider (LHC)
ジュネーブ郊外
地下100m周長27kmのトンネル内に建設
重心系衝突エネルギー最大14TeV
ビームあたり362MJを貯め込む
大きさも衝突エネルギーも世界最大の加速器

ATLAS測定器
4箇所の衝突点に設置された4実験の1つ
~40か国から3,000人の研究者が参加
110Mチャンネルのセンサー読み出し
~1MB/事象, ~1kHz, 年間数10PB以上

新物理事象のトリガー



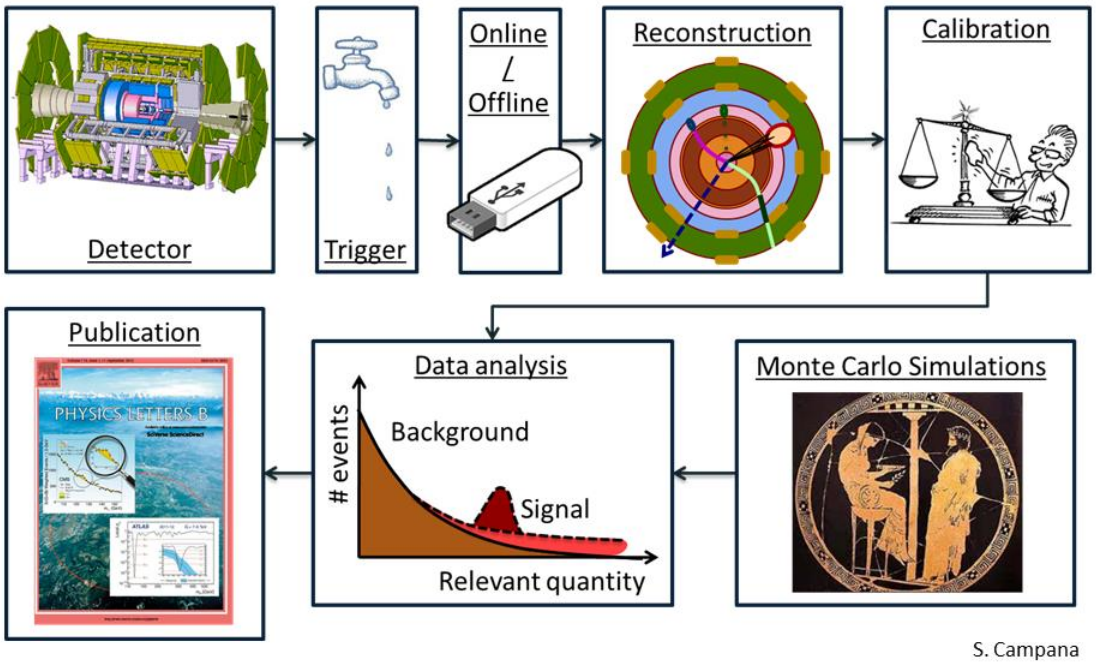
40MHz ~100kHz ~1kHz

エレクトロニクス

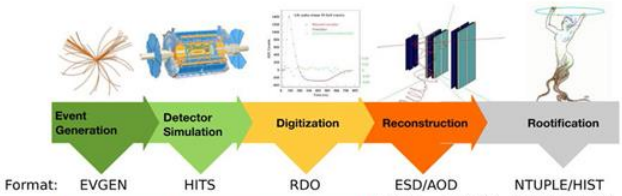
コンピュータクラスタ
~15k cores in 2015

S. Campana

データ処理の流れ



シミュレーションと物理解析



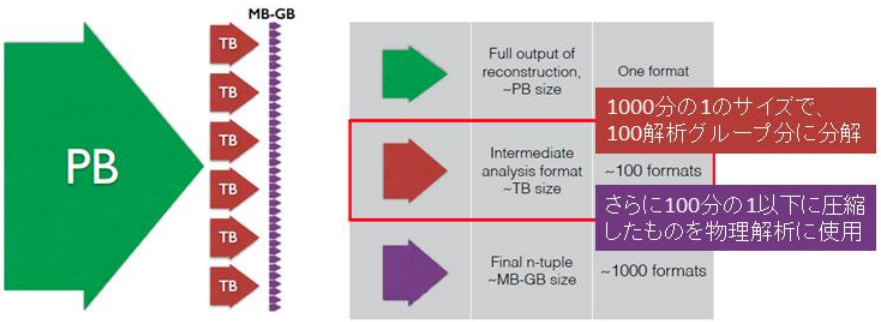
Gridをもちいて、実データ取得から数日後には物理解析が可能となるような解析環境を実現

onlineだけでなくofflineまで見通した計算機利用と資源配備の計画がないプロジェクトは破綻する

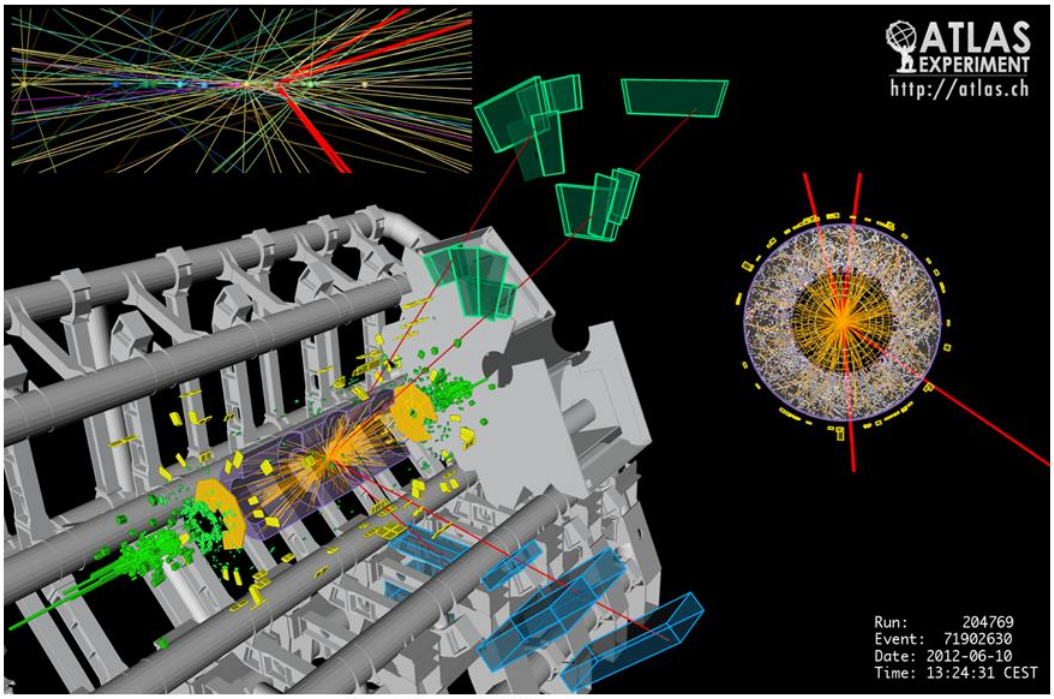
MCプロダクション　実データの数倍の統計量が必要
物理モデルの違い、検出器の状態変化に対応

実データ　衝突エネルギーの上昇、加速器の大強度化、検出器の高精度化
トリガーレートの上昇によるイベントサイズとデータ量の増加

データフォーマットを極力共通化し、組織化したデータアクセスを行うことにより、データ圧縮に要するI/Oを削減(トレインモデル)



2012年ヒッグス粒子の発見

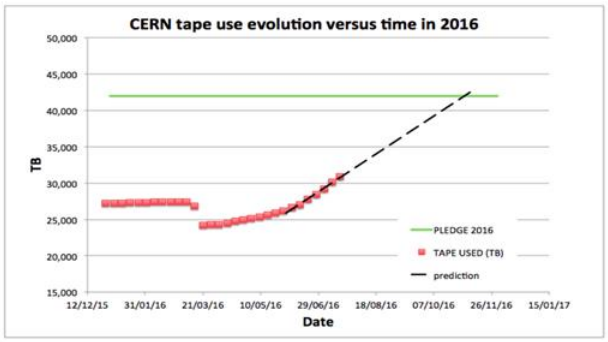


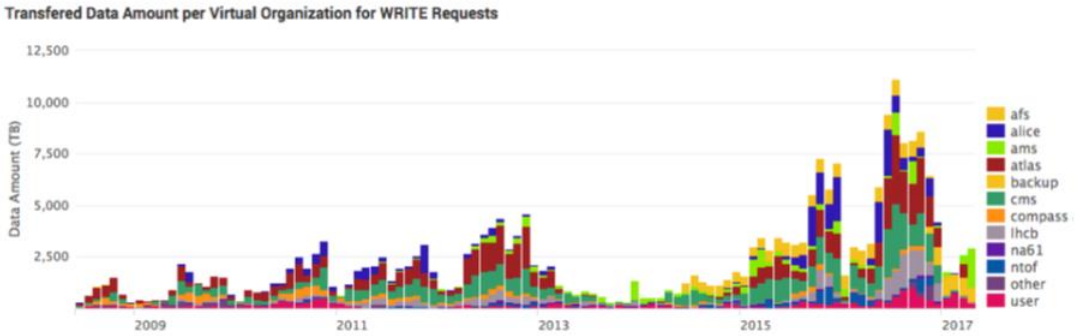
ヒッグス粒子の候補： H → ZZ → μμμμ

2013年ノーベル賞

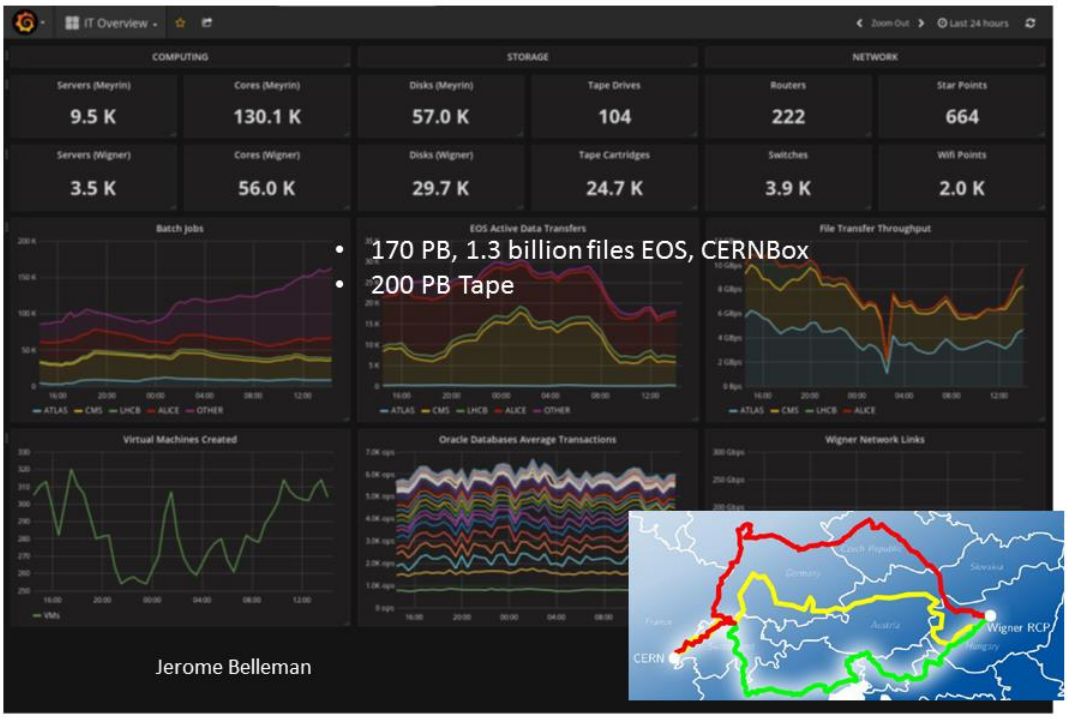
ヒッグス博士とアングレール博士 ©CERN

Raw data at CERN





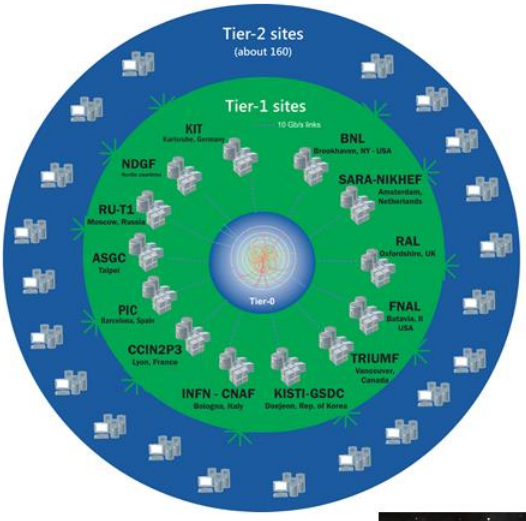
CERN計算機



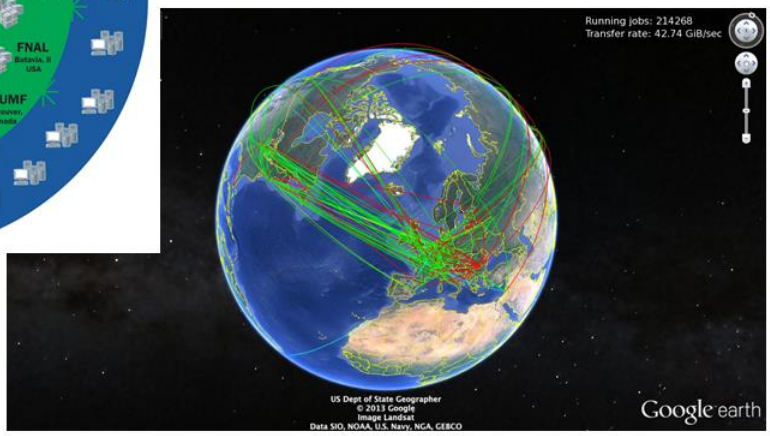
- 170 PB, 1.3 billion files EOS, CERNBox
- 200 PB Tape

Jerome Belleman

Worldwide LHC Computing Grid

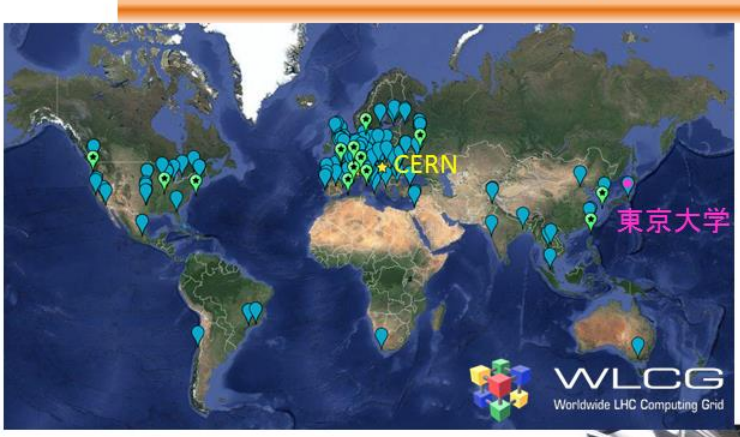


データアーカイブと再構成: CERN Tier-0
分散データアーカイブ: 13サイトのTier-1
シミュレーションと物理解析: 150サイト以上のTier-2



http://wlcg.web.cern.ch/

WLCG



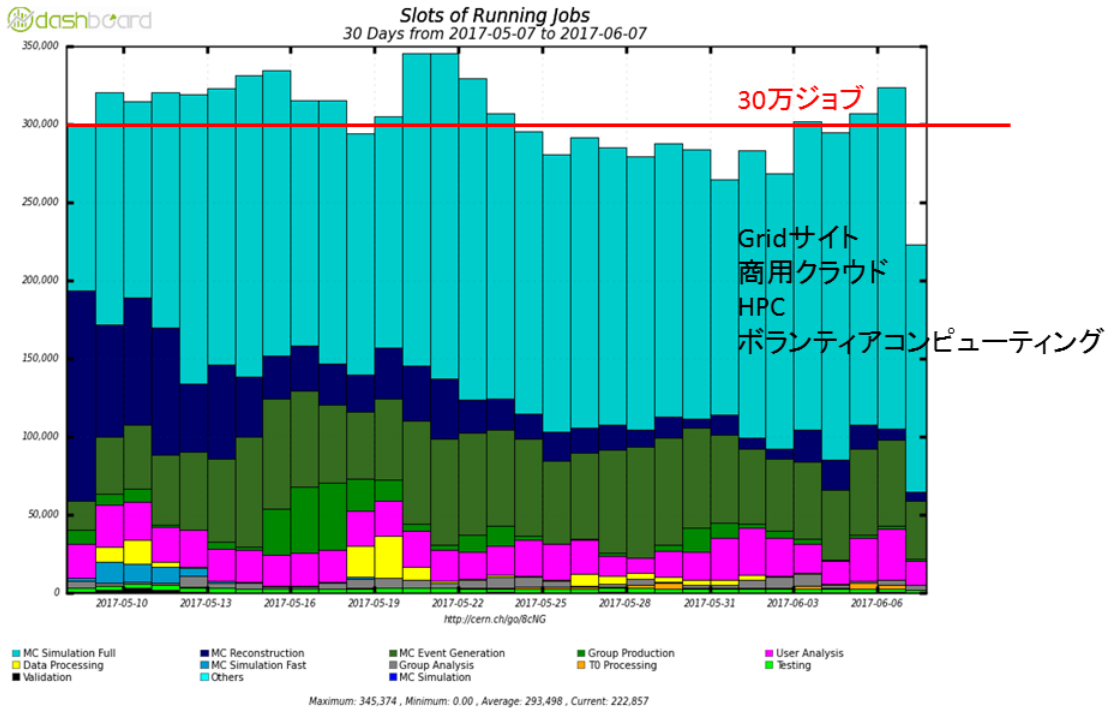
研究者は，自分のコンピューターからグリッドにアクセスすることにより，データや計算機の所在を意識することなく，即座に物理解析結果を得ることができる。

実験データの取得から物理解析結果を導くまで，パイプラインでの処理を可能とする仕組み。

40か国に点在する170の拠点が高帯域学術ネットワーク（10Gbps～100Gbps）で接続されている。
全体で約50万コアのCPU，約400PBのディスクストレージ，約600PBのテープアーカイブを用意している。

東京大学素粒子物理国際研究センターに配備するATLAS実験用のグリッド拠点
10k CPU core (サービス込み)，10 PB disk

ATLAS Grid Jobs（今週）



ストレージ

- AuthenticationとAuthorizationはX.509証明書（ホストと個人）とVOMS（Virtual Organization Membership Service）に基づくACL
- 中央のファイルカタログ（Run1の頃はLFC）とSRM（特にテープアーカイブ）
- データ転送はGridFTPかXRootDがメイン
- 転送制御はFTS3

dCache.org

StoRM

DPM
(Disk Pool Manager)

GT 6.0 GridFTP

XRootD

EOS

FTS3

CASTOR
CERN Advanced STORage manager

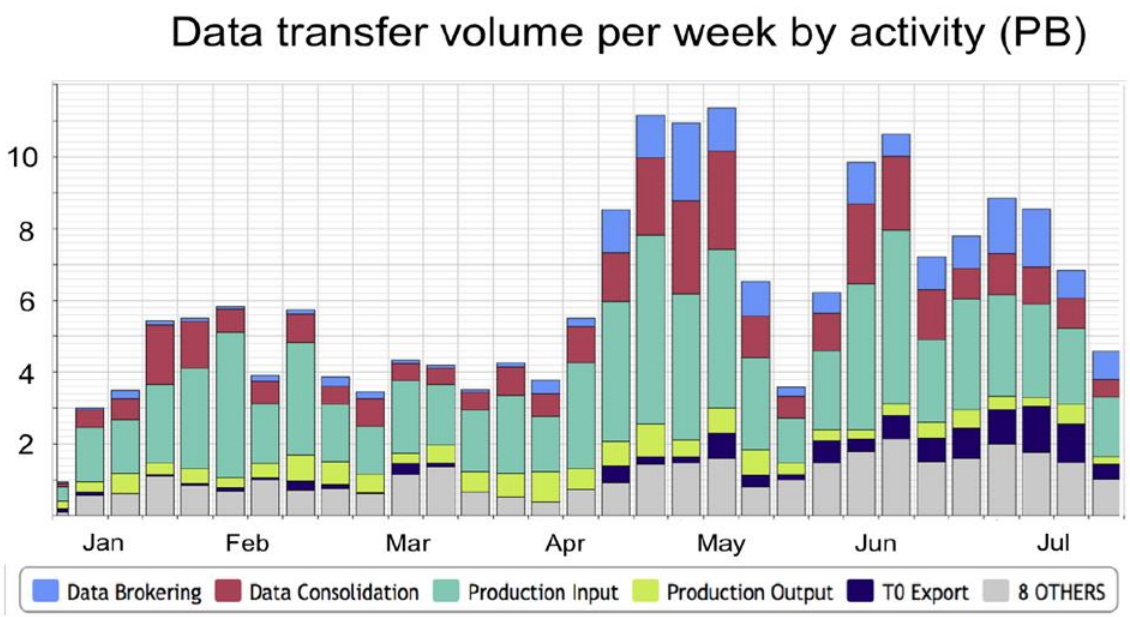
HPSS
High Performance Storage System

https://www.dcache.org/
http://lcgdm.web.cern.ch/dpm
https://italiangrid.github.io/storm/index.html
http://eos.readthedocs.io/en/latest/index.html
http://xrootd.org/
http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/
http://www.hpss-collaboration.org/
http://castor.web.cern.ch/
http://fts3-service.web.cern.ch/

マルチサイトでの転送

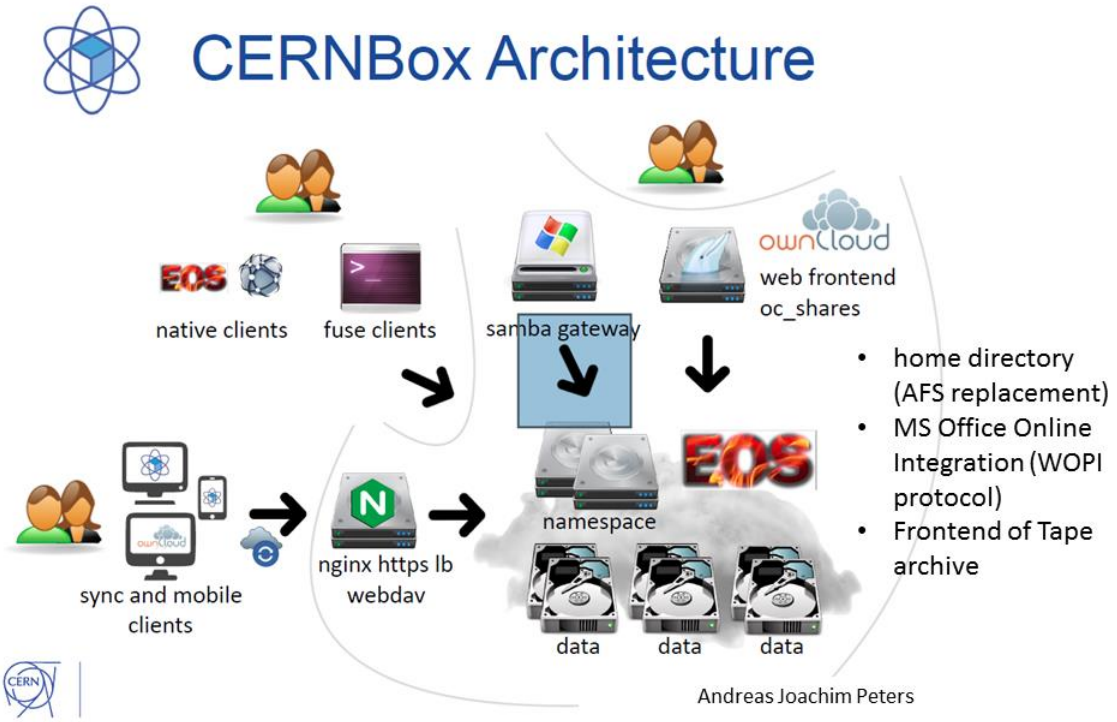
	TRANSFER-	STAGING-	DELETION-	CA+	CERN+	DE+	ES+	FR+	IT+	ND+	NL+	RU+	TW+	UK+	US+
TOTAL-	92 % 6 GB/s	95 % 232 MB/s	89 % 18 GB/s	92 % 375 MB/s	96 % 2 GB/s	93 % 556 MB/s	90 % 23 MB/s	91 % 767 MB/s	91 % 132 MB/s	93 % 329 MB/s	87 % 166 MB/s	81 % 247 MB/s	90 % 36 MB/s	93 % 726 MB/s	91 % 665 MB/s
CA+	98 % 983 MB/s	100 % 308 kB/s	100 % 497 MB/s	97 % 21 MB/s	100 % 219 MB/s	99 % 123 MB/s	97 % 5 MB/s	90 % 75 MB/s	98 % 10 MB/s	99 % 58 MB/s	95 % 7 MB/s	100 % 130 MB/s	98 % 2 MB/s	98 % 243 MB/s	99 % 89 MB/s
CERN+	99 % 2 GB/s	100 % 0 kB/s	100 % 15 MB/s	100 % 11 MB/s	100 % 2 GB/s	96 % 13 MB/s	99 % 1 MB/s	99 % 17 MB/s	100 % 5 MB/s	100 % 6 MB/s	98 % 1 MB/s	100 % 3 kB/s	100 % 57 kB/s	98 % 9 MB/s	100 % 23 MB/s
DE+	90 % 457 MB/s	95 % 441 kB/s	97 % 331 MB/s	83 % 17 MB/s	93 % 28 MB/s	94 % 222 MB/s	88 % 6 MB/s	91 % 16 MB/s	88 % 18 MB/s	93 % 38 MB/s	78 % 6 MB/s	54 % 630 kB/s	84 % 892 kB/s	92 % 50 MB/s	92 % 54 MB/s
ES+	93 % 260 MB/s	99 % 208 MB/s	93 % 8 GB/s	93 % 2 MB/s	94 % 3 MB/s	92 % 11 MB/s	97 % 329 kB/s	92 % 213 MB/s	88 % 1 MB/s	98 % 3 MB/s	92 % 5 MB/s	92 % 743 kB/s	97 % 118 kB/s	93 % 3 MB/s	97 % 18 MB/s
FR+	85 % 1 GB/s	100 % 805 kB/s	100 % 357 MB/s	92 % 245 MB/s	80 % 6 MB/s	90 % 120 MB/s	64 % 788 kB/s	79 % 296 MB/s	90 % 5 MB/s	88 % 151 MB/s	81 % 102 MB/s	69 % 110 MB/s	52 % 361 kB/s	89 % 236 MB/s	81 % 20 MB/s
IT+	91 % 95 MB/s	0 % 0 kB/s	92 % 504 MB/s	91 % 5 MB/s	87 % 15 MB/s	88 % 5 MB/s	99 % 200 kB/s	88 % 8 MB/s	89 % 6 MB/s	73 % 1 MB/s	96 % 4 MB/s	100 % 62 kB/s	99 % 292 kB/s	95 % 6 MB/s	95 % 45 MB/s
ND+	98 % 43 MB/s	100 % 0 kB/s	99 % 55 MB/s	100 % 4 MB/s	100 % 3 MB/s	95 % 5 MB/s	100 % 251 kB/s	99 % 7 MB/s	93 % 760 kB/s	100 % 63 kB/s	98 % 2 MB/s	100 % 2 MB/s	100 % 1 MB/s	98 % 12 MB/s	100 % 6 MB/s
NL+	95 % 90 MB/s	100 % 4 kB/s	95 % 32 MB/s	93 % 20 MB/s	99 % 19 MB/s	92 % 3 MB/s	96 % 883 kB/s	93 % 18 MB/s	85 % 5 MB/s	96 % 6 MB/s	89 % 282 kB/s	91 % 49 kB/s	89 % 106 kB/s	92 % 5 MB/s	95 % 12 MB/s
RU+	98 % 14 MB/s	100 % 956 kB/s	100 % 30 MB/s	99 % 118 kB/s	100 % 988 kB/s	99 % 672 kB/s	100 % 240 kB/s	97 % 4 MB/s	90 % 313 kB/s	100 % 721 kB/s	100 % 169 kB/s	100 % 10 kB/s	100 % 31 kB/s	97 % 3 MB/s	98 % 4 MB/s
TW+	95 % 28 MB/s	100 % 0 kB/s	100 % 22 MB/s	100 % 142 kB/s	100 % 3 MB/s	99 % 1 MB/s	92 % 41 kB/s	97 % 926 kB/s	99 % 118 kB/s	65 % 808 kB/s	99 % 131 kB/s	100 % 8 kB/s	91 % 5 MB/s	91 % 5 MB/s	100 % 13 MB/s
UK+	92 % 778 MB/s	100 % 0 kB/s	88 % 541 MB/s	92 % 32 MB/s	95 % 52 MB/s	93 % 33 MB/s	90 % 5 MB/s	93 % 93 MB/s	94 % 64 MB/s	93 % 40 MB/s	89 % 37 MB/s	71 % 1 MB/s	69 % 405 kB/s	90 % 114 MB/s	93 % 307 MB/s
US+	90 % 321 MB/s	44 % 22 MB/s	75 % 7 GB/s	91 % 20 MB/s	91 % 74 MB/s	92 % 19 MB/s	94 % 3 MB/s	88 % 19 MB/s	87 % 17 MB/s	94 % 23 MB/s	82 % 2 MB/s	80 % 3 MB/s	99 % 25 MB/s	92 % 42 MB/s	84 % 75 MB/s

ATLAS サイト間でのデータ転送（2016）



ATLAS, CMS, ALICE, LHCbを合わせると、定常的に40 GB/sぐらい

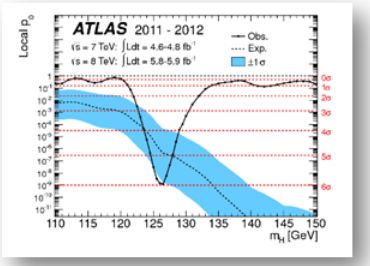
ownCloud and EOS at CERN



- home directory (AFS replacement)
- MS Office Online Integration (WOPI protocol)
- Frontend of Tape archive

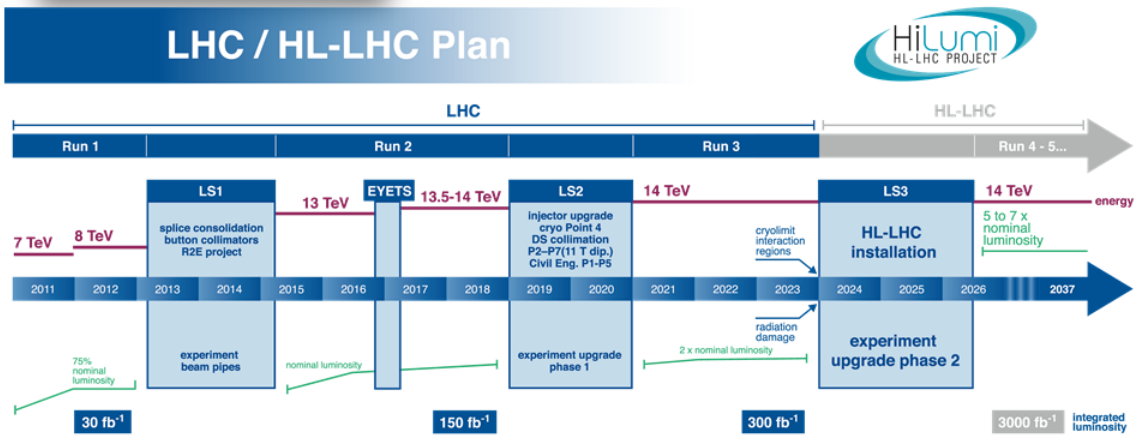
LHC実験の今後

LHC-Run1で発見されたヒッグス粒子

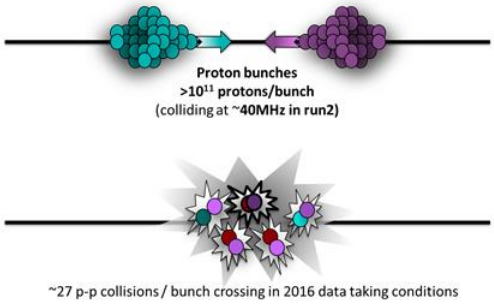


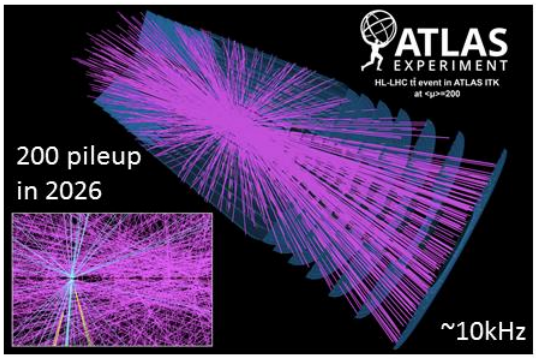
これから20年間でRun1の100倍の実験データを取得

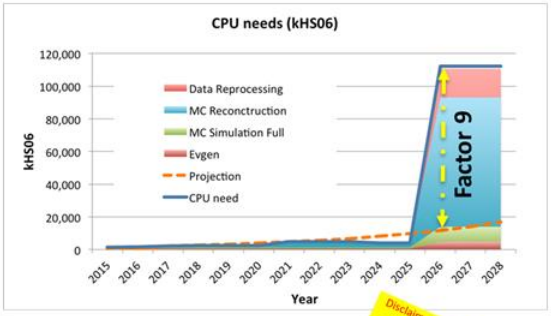
- ヒッグス粒子の詳細な性質を明らかにする
- 標準理論を超える物理(新粒子)を探索する

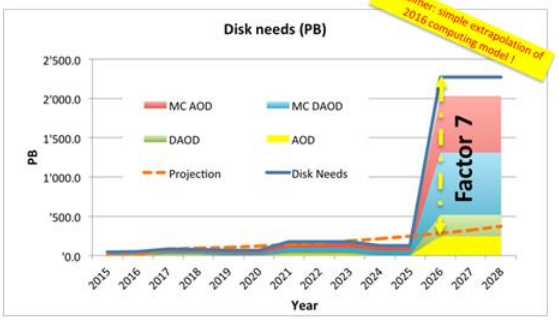


HL-LHC in 2026





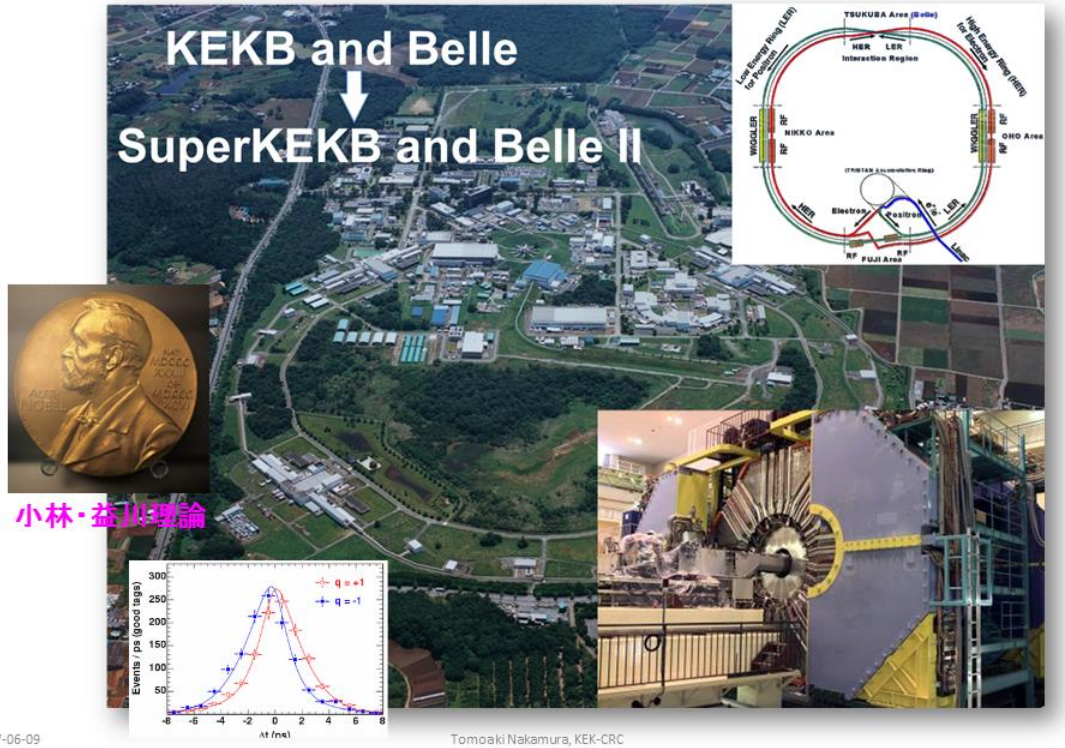




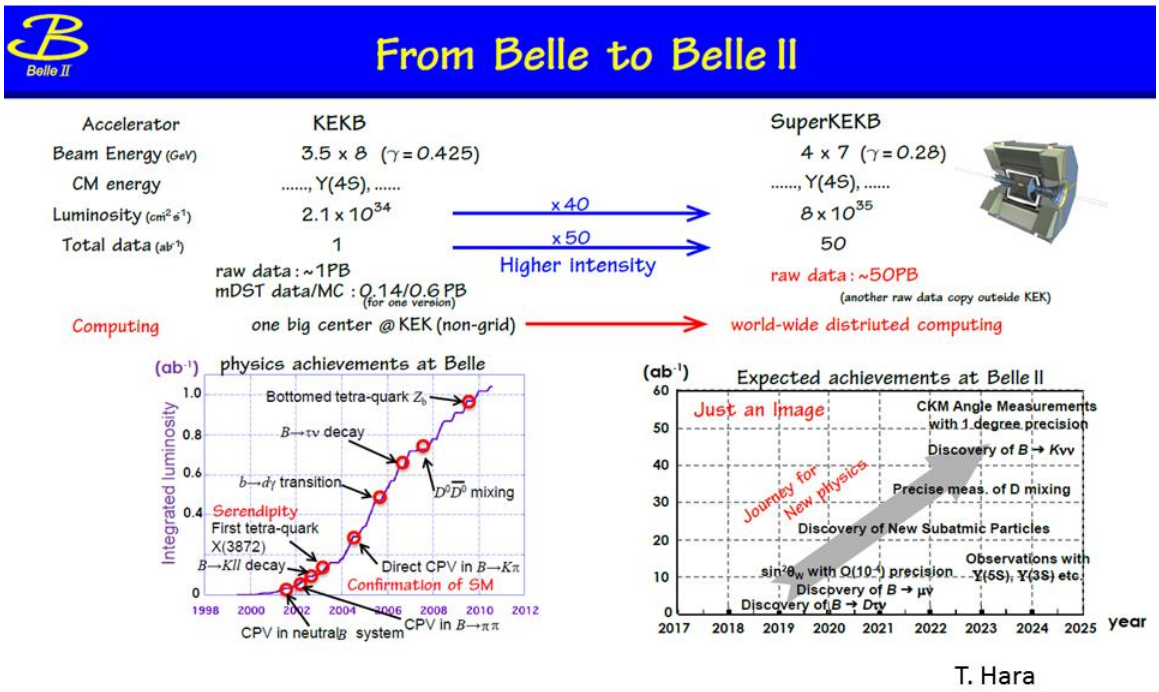
Disclaimer: simple extrapolation of 2016 computing model !

S. Campana

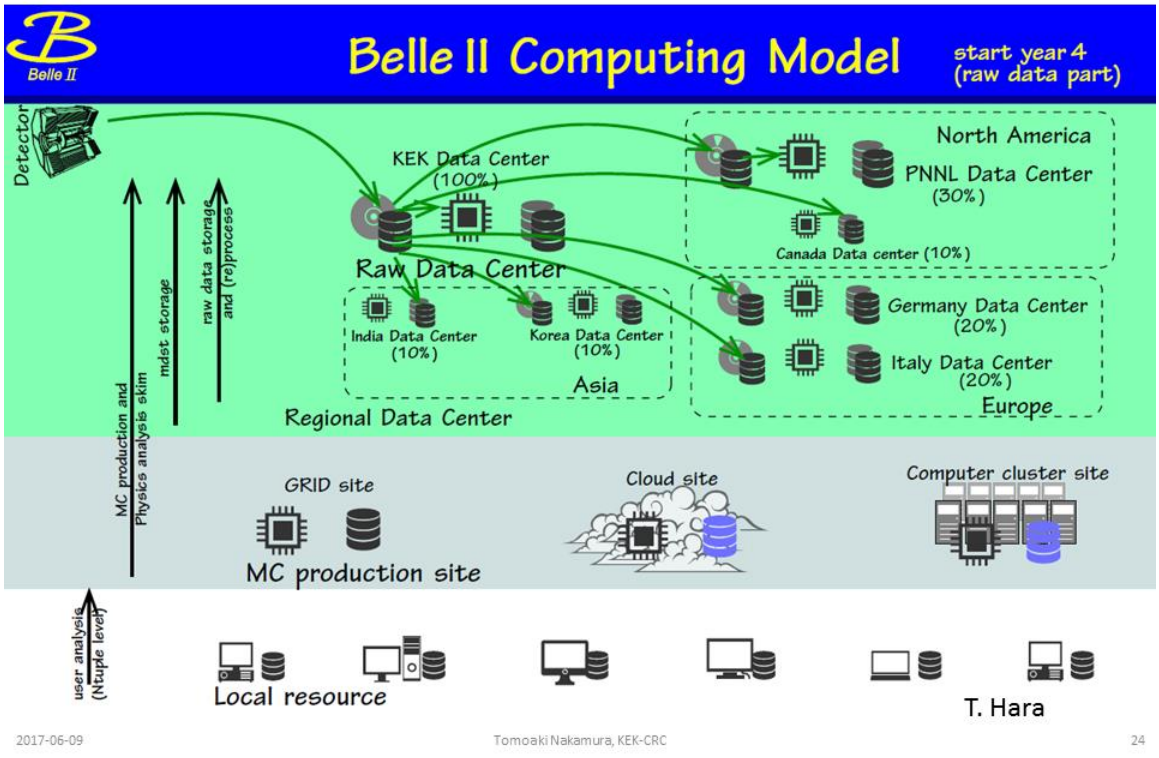
SuperKEKB/Belle II at KEK



50倍のデータ

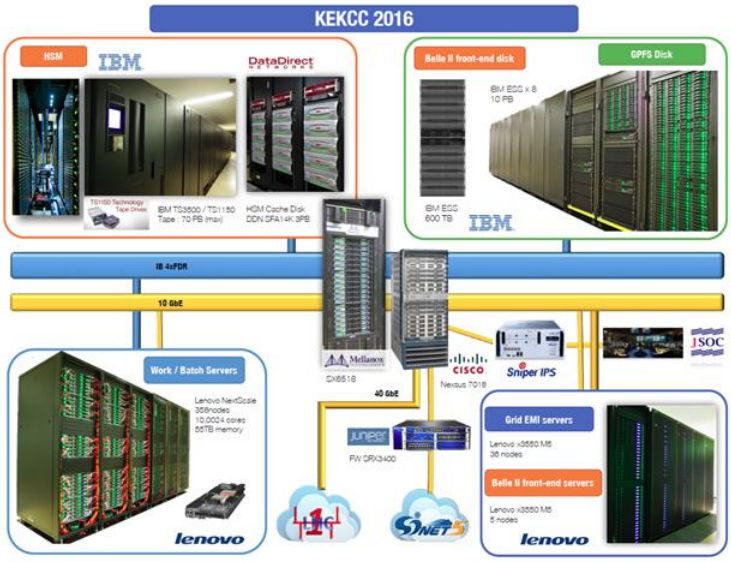


Belle II もGridを使う



KEK Central Computer System

System has already been in production mode since the last September.



SYSTEM RESOURCES

CPU : 10,024 cores
- Intel Xeon E5-2697v3 (2.6GHz, 14cores) x 2
  358 nodes
- 4GB/core (8,000 cores) /
  8GB/core (2,000 cores) (for app. use)
- 236 kHS06 / site

Disk : 10PB (GPFS) + 3PB (HSM cache)

Interconnect : IB 4xFDR

Tape : 70 PB (max cap.)
HSM data : 8.5 PB data, 170 M files,
5,000 tapes

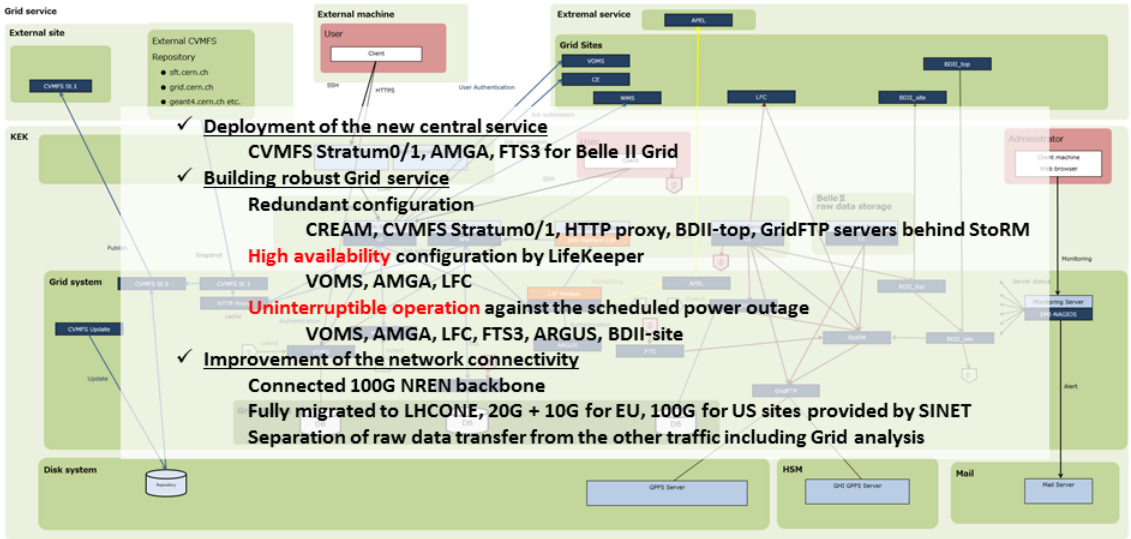
Total throughput : 100 GB/s (Disk, GPFS),
50 GB/s (HSM, GHI)

JOB scheduler : Platfrom LSF v9

K. Murakami

Grid system at KEK

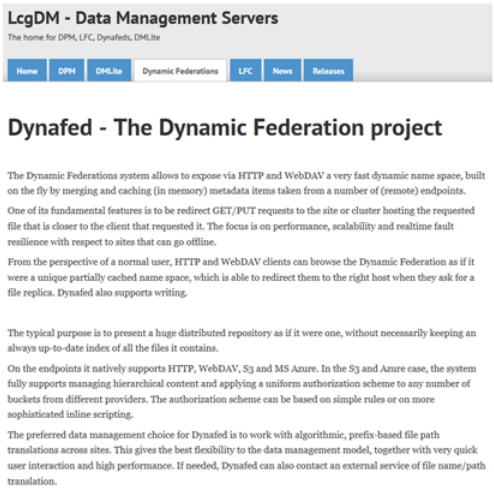
System has already been in production mode since the last September.



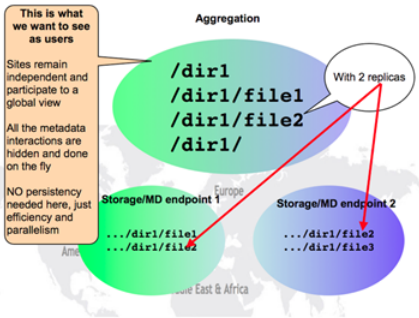
- ✓ **Deployment of the new central service**
  - CVMFS Stratum0/1, AMGA, FTS3 for Belle II Grid
- ✓ **Building robust Grid service**
  - Redundant configuration
    - CREAM, CVMFS Stratum0/1, HTTP proxy, BDII-top, GridFTP servers behind StoRM
  - High availability configuration by LifeKeeper
    - VOMS, AMGA, LFC
  - Uninterruptible operation against the scheduled power outage
    - VOMS, AMGA, LFC, FTS3, ARGUS, BDII-site
- ✓ **Improvement of the network connectivity**
  - Connected 100G NREN backbone
  - Fully migrated to LHCONE, 20G + 10G for EU, 100G for US sites provided by SINET
  - Separation of raw data transfer from the other traffic including Grid analysis

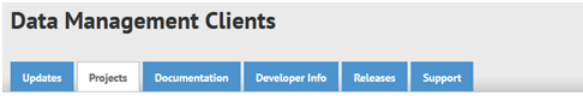
HTTPストレージ・フェデレーション

http://lcgdm.web.cern.ch/dynafed-dynamic-federation-project



ATLAS, Belle II ...





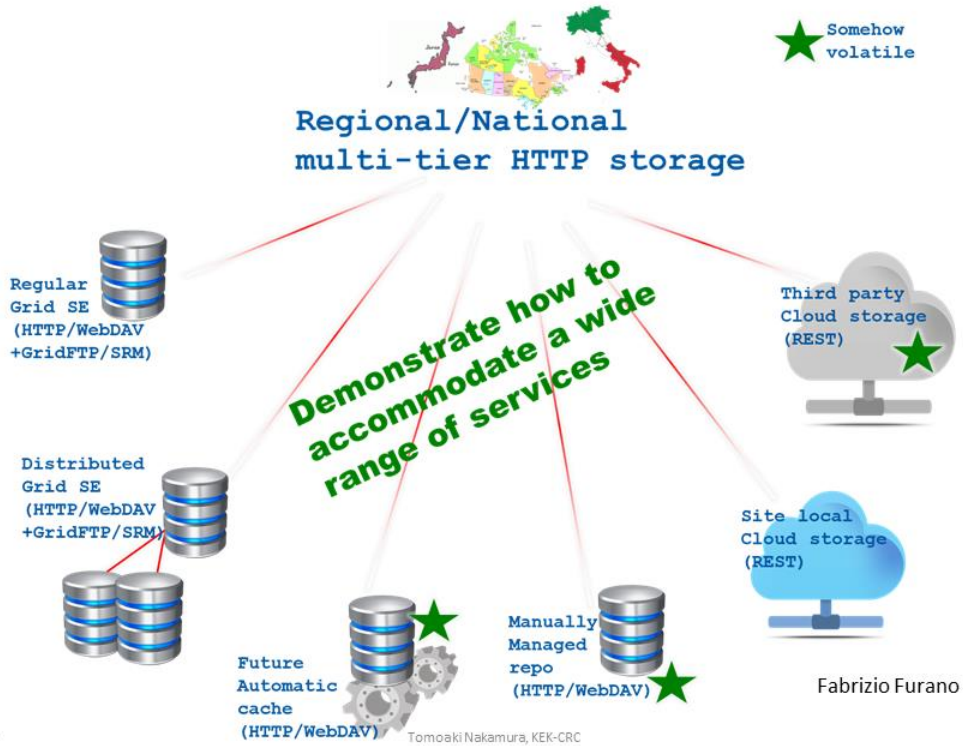
davix

The davix project aims to make file management over HTTP-based protocols simple. The focus is on **high-performance** remote I/O and data management of large collections of files. Currently, there is support for the WebDav, Amazon S3, Microsoft Azure, and HTTP protocols.

davix is multi-platform, open source and written in C++.

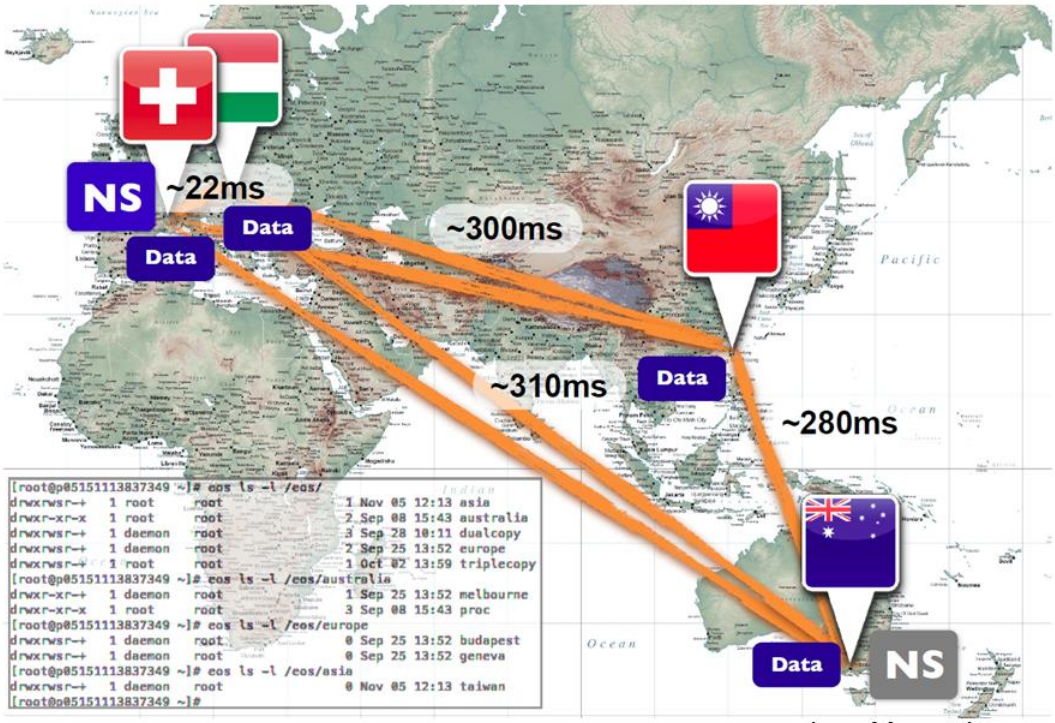
https://dmc.web.cern.ch/projects/davix/home

多種のストレージを統合



Fabrizio Furano

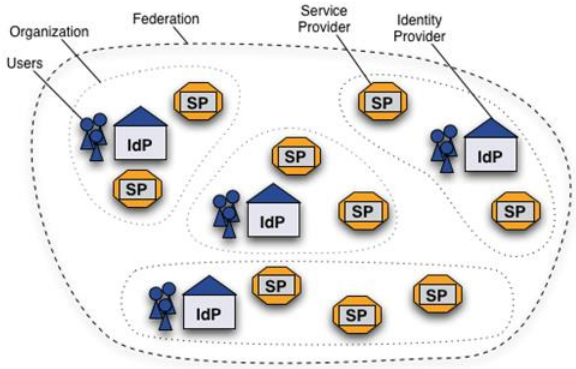
EOSストレージフェデレーション



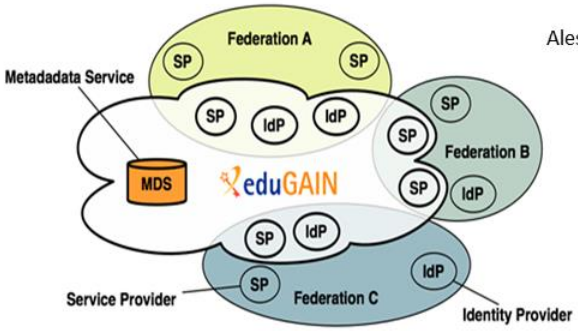
Luca Mascetti

Identity federation

ID federation



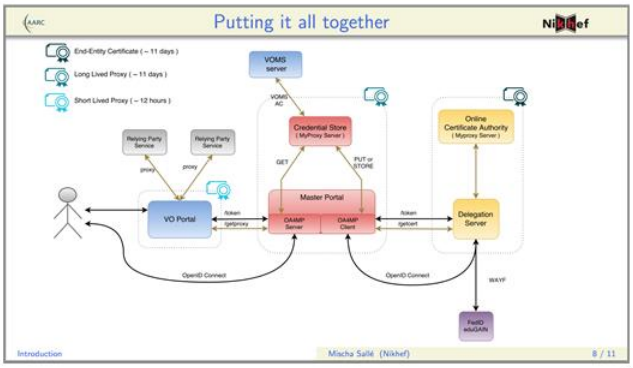
Inter-federation



Alessandra Scicchitano

WLCGプロトタイプ

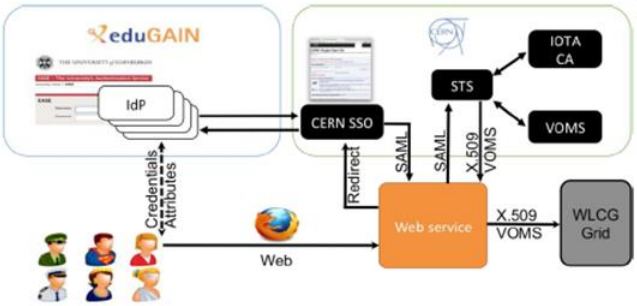
EGI



Security incident response...
Trust framework...

Hannah Short

WLCG



2017-06-09 Tomoaki Nakamura, KEK-CRC 31

A.4. 白崎委員

【国立天文台】
・一般向けデータ公開（O(1PB)）、観測データファイル
・一般向けデータ公開（O(10TB)）、カタログデータベース
・国立天文台内におけるデータ共有（O(1PB)）

<u>観測データ</u>

国立天文台では、すばる望遠鏡、ALMA 望遠鏡、野辺山電波望遠鏡、水沢 VERA、太陽望遠鏡等による観測によって取得されたデータの保管とインターネット経由による配信を行っている。

すばる望遠鏡によって取得されたデータはすべて天文分野における共通のフォーマットである FITS 形式で保管されている。他の観測装置の場合にはそれぞれの装置に固有のフォーマットで保管されているものと、FITS 形式で保管されているものとが共存している。

観測データのファイルあたりの容量はすばる望遠鏡の場合数十 MB 程度である。一方、ALMA 望遠鏡の場合は観測によりデータサイズはまちまちであり、数 10KB から 数 TB まで様々である。現時点（2018 年）における観測データの全容量は数 PB のオーダーである。

保管されているデータのメタデータはデータベース化されインターネット経由で検索可能であり、データ本体もダウンロード可能である。検索インターフェイスは Web ブラウザベースのほか、最近はヴァーチャル天文台仕様により世界中の天文データベースが利用可能となるアプリケーションを利用した検索も可能となっている。

今後想定される懸念事項としていは、ALMA の１ファイルあたりのデータ量が大きいため、ユーザがインターネット経由でダウンロードすることが困難な状況が増えていくであろうことである。また、従来使ってきた解析ソフトウェアの中にはそうした大容量データの解析には対応できていないものもあると考えられ、今後そうしたデータの解析にも耐えられるソフトウェアの開発が必要である。

<u>カタログデータ</u>

観測データから導き出される天体カタログデータの公開も行っている。

国立天文台が運用に関わる上記の望遠鏡以外にも、海外研究機関が運用している科学衛星 Gaia や JAXA の天文衛星 AKARI のデータより作成されたカタログも検索システムを構築し公開している。これらの全データ容量は数十 TB の規模である。

カタログデータは PostgreSQL 等のリレーショナルデータベースに入力され、テーブルあたりのカラム数は数十から数千個である。

Web ブラウザーを利用した検索が可能であり、CSV 形式の他、FITS 形式、VOTable 形式でダウロードできる。VOTable 形式はカタログデータや観測データのメタデータを交換するために定義されたデータ形式であり、XML で記述される。

観測データの場合と同様にヴァーチャル天文台に対応したアプリケーションによる検索も行われるようになってきている。また、単純にカタログファイル全体を Web ブラウザー等によりリンク先 URL からダウンロードすることも可能である。

カタログデータは観測データに比べて総容量は大きくないので、時間をかければダウンロードすることは可能である。しかしながら、データ公開開始直後は一度に大勢の研究者がデータダウンロードを行う傾向があり、ネットワーク回線への負荷がかかり過ぎることにより他のネットワークサービスへのアクセスもできない状況に陥ってしまうことが度々ある。こうした問題を回避できる仕組みが必要である。

<u>国立天文台内におけるデータ共有</u>

大量の観測データを利用者の計算機にコピーせずとも解析を行えるようにするため、また解析環境を研究者個人で用意するコストを抑えるため、共同利用計算機システムの運用を行っている。

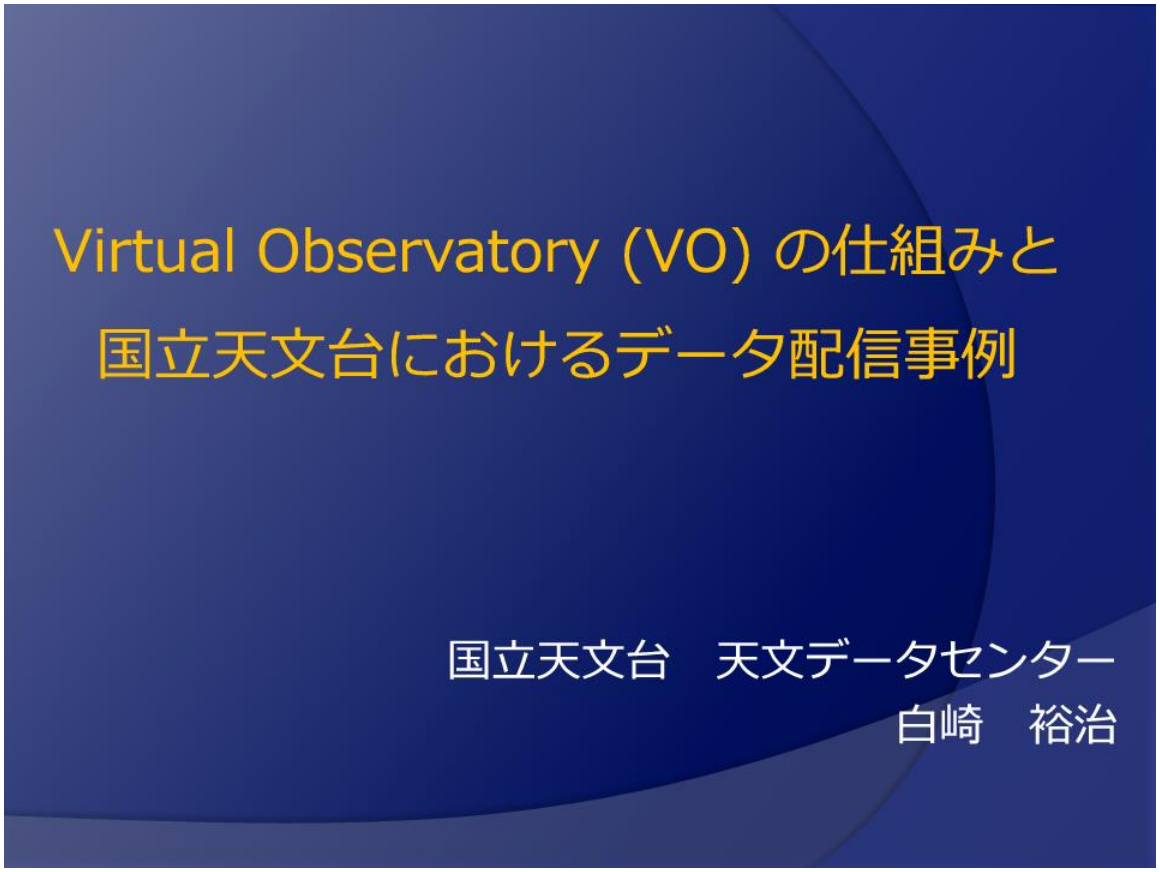
すばる望遠鏡や ALMA などのデータを解析したい研究者は、国立天文台三鷹キャンパスの共同利用計算機室からこの計算機システムにログインし利用することができる。また、遠隔から VPN によりログインして利用することも可能である。

この計算機システム上には数 PB 規模のデータ保管領域が用意されており、解析済みのデータを一定期間保存しておくことができる。解析対象となる観測データは、台外向けのデータ公開インターフェイスを利用してこの計

算機システム上にコピーしてから解析を行う。

共同利用者同士でのデータのやりとりにもこのデータ保管領域を利用することができる。もちろん、研究者個人の計算機に scp や rsync なででデータをエクスポートすることも可能である。

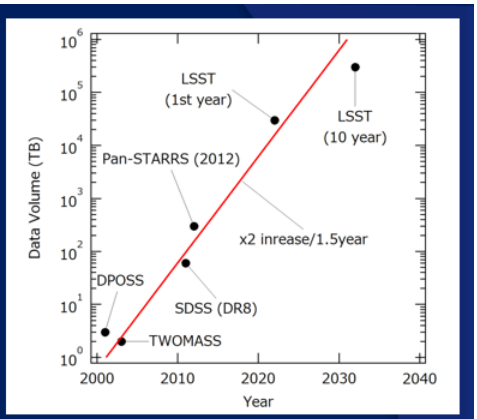
現在の課題としては、データ保管領域が多数のパーティションに分割されており、どのパーティションにデータを置くかは各利用者の判断に委ねられているため、特定のパーティションにデータが偏ってしまうことがあげられる。ほぼ100%まで利用されてしまったパーティションがある一方で、ほとんど利用されていないパーティションも存在している。現在分割されてしまっているパーティションを一つにまとめた上で、データ転送の帯域幅も十分ある大きな一つの保管領域として利用可能な仕組みが必要である。



内容

- 天文データの現状
  - データ利用における問題点
- VO の仕組みと利用方法
  - 天文データ共有の仕組み
  - VO サービス・デスクトップアプリの紹介
- 国立天文台におけるデータ配信
  - JVO, HSC, SMOKA

急増する天文データ



- 天文データは 1年半毎に倍増
  - CPU の計算性能は18カ月で2倍
  - I/O 性能の向上率は 10%/年
  - 並列 I/O・計算技術は必須
  - データ移動を極力避ける解析システム
- 望遠鏡・観測装置の大型化、高機能化
  - 高品質なデータ、取得のコストの増大 → 科学成果の最大化
  - 取得したデータを速やかに解析できる環境の構築
  - 研究者間で共有できる仕組みづくり

すばる望遠鏡

- 世界最大級の光赤外線望遠鏡
- アメリカ・ハワイ島　マウナ・ケア山山頂（標高4,205m）
- 口径 8.2m
- 共同利用装置
- 国立天文台が運用
- 年間 ～30TB の観測データを取得し占有期間が過ぎたデータを順次公開

ALMA 望遠鏡

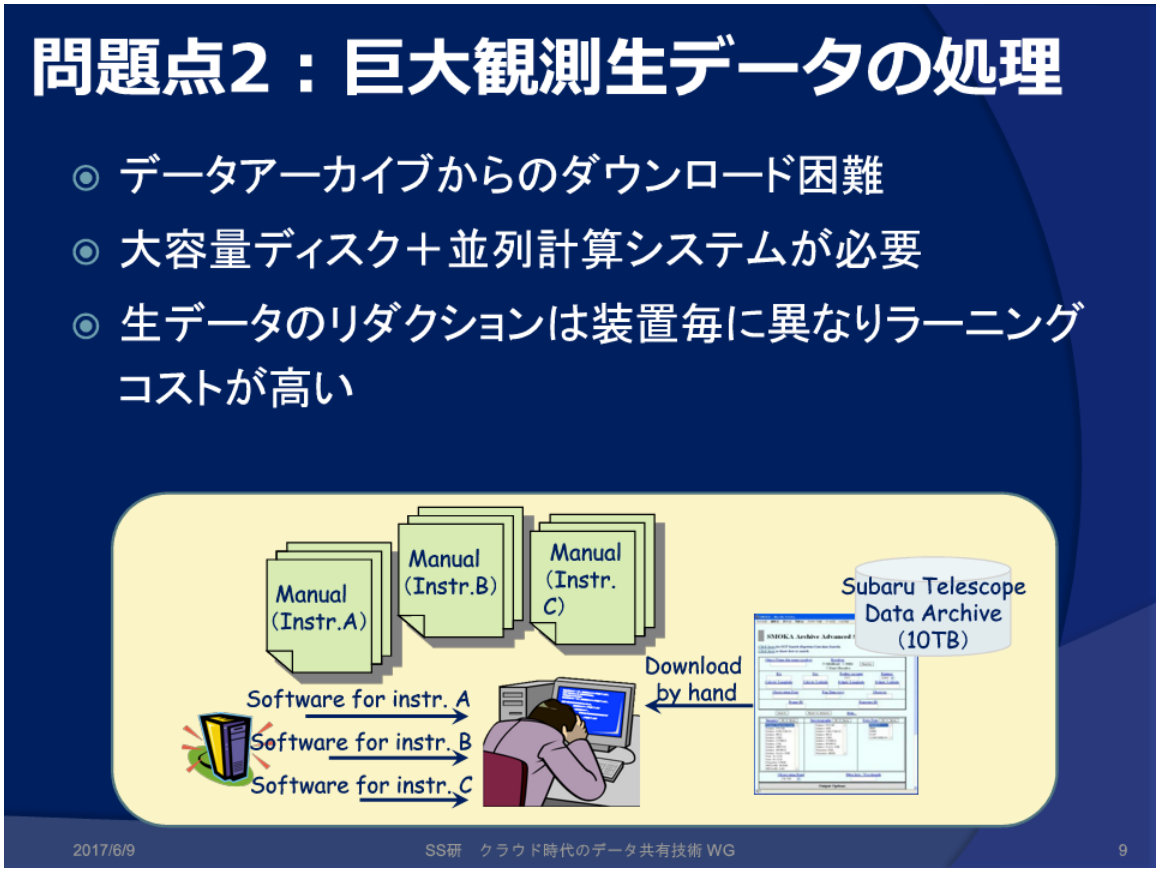
- 南米チリで観測を開始した電波望遠鏡
- 日・米・欧・台湾による国際プロジェクト
- 年間 200 TB のデータを生成

一つのデータファイルが1TB を超えることも

稼働中の主要な望遠鏡 (国内)
http://ja.wikipedia.org/wiki/ファイル:Spectre.svg を改変
ガンマ線
X 線
紫外線
可視光
赤外線
電波
すざく
すばる
野辺山45m電波望遠鏡
あかり
ASTE
MAXI
ひので
ALMA
ALMA/NAOJ
2017/6/9
SS研　クラウド時代のデータ共有技術 WG
6

稼働中の主要な望遠鏡 (海外)
http://ja.wikipedia.org/wiki/ファイル:Spectre.svg を改変
ガンマ線
X 線
紫外線
可視光
赤外線
電波
フェルミ
チャンドラ
ハッブル
ハーシェル宇宙望遠鏡
超大型干渉電波望遠鏡群(VLA)
GALEX
ニュートン
ジェミニ
大型ミリ波望遠鏡(LMT)
スピッツァー宇宙望遠鏡
スイフト
VLT
2017/6/9
SS研　クラウド時代のデータ共有技術 WG
7

問題点１：分散データ
アーカイブへの
アクセス
猫の手も借りたい
データベースアクセスインターフェイスの共通化
データ
取得を
自動化
サービスの
メタデータ DB
教育用教材としても利用可
2017/6/9
SS研　クラウド時代のデータ共有技術 WG

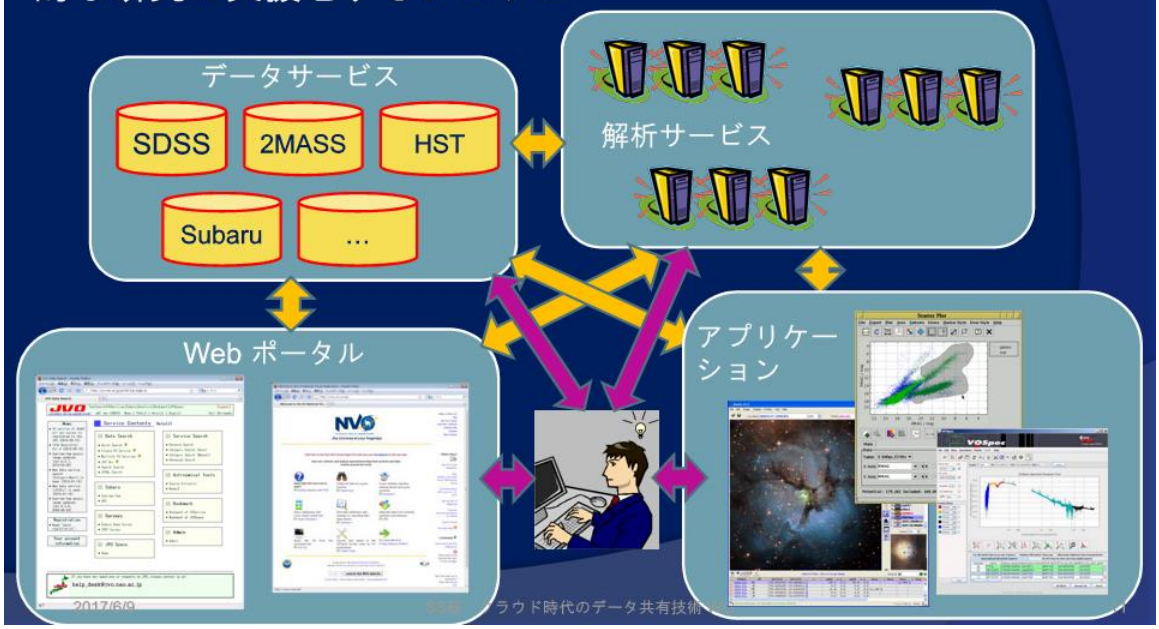

問題点2：巨大観測生データの処理
◉ データアーカイブからのダウンロード困難
◉ 大容量ディスク＋並列計算システムが必要
◉ 生データのリダクションは装置毎に異なりラーニングコストが高い
Manual (Instr.A)
Manual (Instr.B)
Manual (Instr. C)
Software for instr. A
Software for instr. B
Software for instr. C
Download by hand
Subaru Telescope Data Archive (10TB)
2017/6/9
SS研　クラウド時代のデータ共有技術 WG
9

解決策: データと計算資源の集約化

- 解析用計算機とデータアーカイブを国立天文台（三鷹）に集約
  - CPU: 632 core、メモリ: 6TB、解析用ストレージ: 1.2PB
  - データアーカイブ: 3PB
  - 解析用計算機とデータアーカイブは高速ネットワーク 10G で結合
  - 解析用計算機にリモートログインして解析実行
  - データ移動のコストを最小化
  - 並列計算環境も提供 → 大規模なデータ解析
  - 解析ソフトは必要なものをすべてインストール済み
  - データ解析講習会を開催

バーチャル天文台とは？

天文データベースの公開方式の国際標準を定義することにより、より高度なネットワーク経由でのデータ共有を可能にし、効率的な研究の支援をするシステム



International Virtual Observatory Alliance

- 2002年に結成
- 世界各国 21 VO プロジェクトが参加
- 天文データの共有をより効率的に行うための標準仕様策定団体
- VO標準に対応したアプリケーションの開発なども

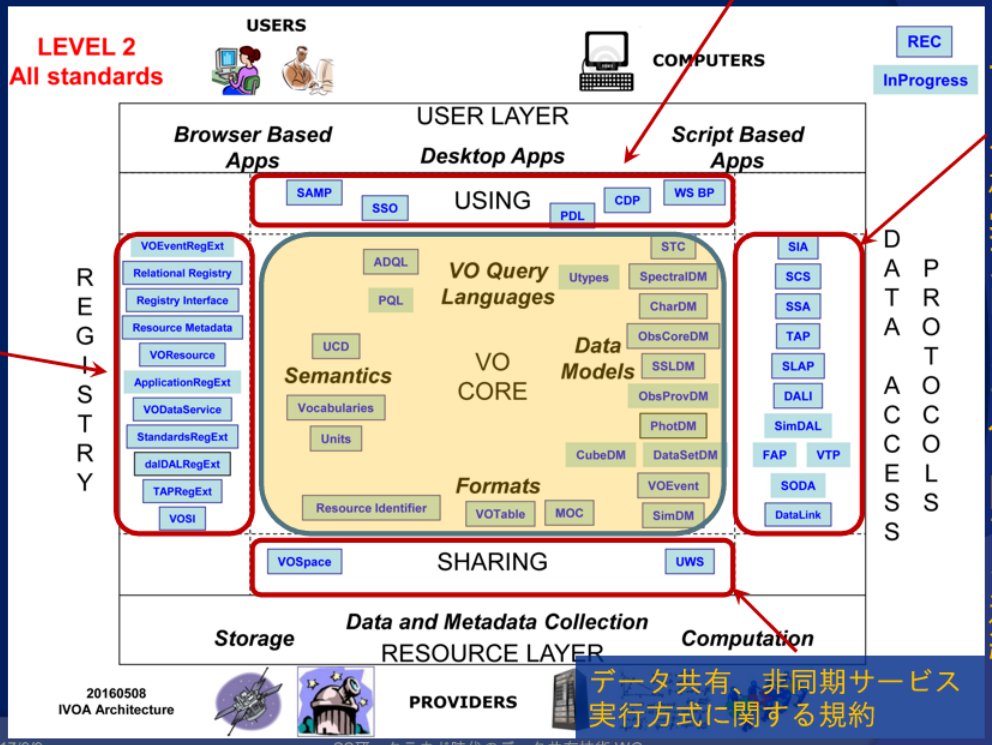
http://www.ivoa.net/

VO architecture

認証、アプリ間連携、サービス実行パラメータに関する規約

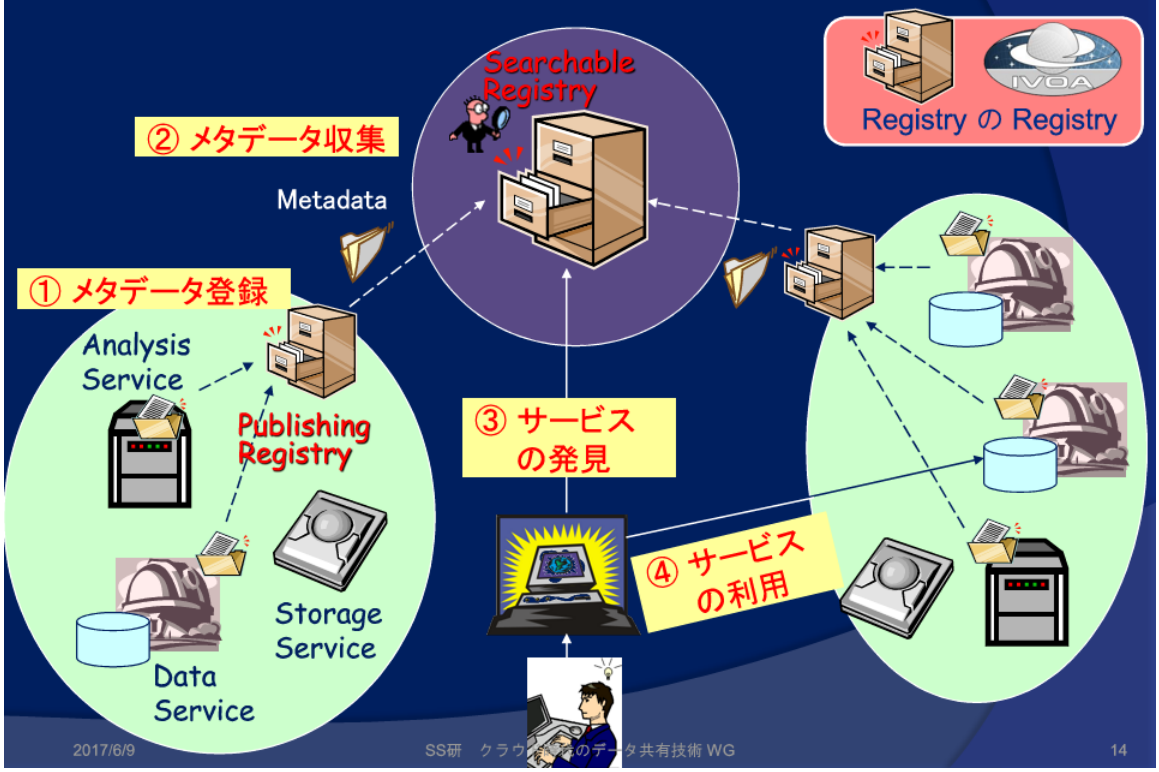
リソース検索プロトコルに関する規約

データ検索プロトコルに関する規約



データ共有、非同期サービス実行方式に関する規約

データサービスの公開と利用の仕組み



リソースメタデータのサンプル


```
- <vr:Resource xsi:type="vs11:CatalogService" created="2013-10-02T00:00:00.000Z" status="active"
  updated="2015-12-21T00:00:00.000Z">
    <title>Chandra X-ray Observatory Data Archive</title>
    <shortName xsi:type="xsd:string">CDA</shortName>
    <identifier xsi:type="xsd:anyURI">ivo://cxc.harvard.edu/cda.siap</identifier>
  - <curation>
      <publisher ivo-id="ivo://cxc.harvard.edu/cxo">Chandra X-ray Observatory</publisher>
    - <creator>
        <name>Chandra X-ray Center</name>
      </creator>
    - <contact>
        <name>Chandra Data Archive, Chandra X-ray Center</name>
        <address>60 Garden Street MS 67, Cambridge, MA 02138, USA</address>
        <email>archelp@head.cfa.harvard.edu</email>
      </contact>
    </curation>
  - <content>
      <subject>X-ray astronomy</subject>
      <subject>galaxies</subject>
      <subject>quasars</subject>
      <subject>stars</subject>
      <subject>X-ray</subject>
      <subject>photometry</subject>
      <subject>spectroscopy</subject>
      <subject>variability</subject>
    - <description>
        The Chandra X-ray Observatory is the U.S. follow-on to the Einstein Observatory. Chandra was formerly known as
        AXAF, the Advanced X-ray Astrophysics Facility, but renamed by NASA in December, 1998. Originally three
        instruments and a high-resolution mirror carried in one spacecraft, the project was reworked in 1992 and 1993. The
        Chandra spacecraft carries a high resolution mirror, two imaging detectors, and two sets of transmission gratings.
        Important Chandra features are: an order of magnitude improvement in spatial resolution, good sensitivity from 0.1 to
        10 keV, and the capability for high spectral resolution observations over most of this range.
      </description>
      <referenceURL>http://cxc.harvard.edu/cda/</referenceURL>
      <type>Archive</type>
      <contentLevel>Research</contentLevel>
      <contentLevel>University</contentLevel>
      <contentLevel>Community College</contentLevel>
    - <relationship>
        <relationshipType>derived-from</relationshipType>
        <relatedResource ivo-id="ivo://cxc.harvard.edu/cda">Chandra Data Archive</relatedResource>
      </relationship>
    </content>
```

```
  - <capability standardID="ivo://ivoa.net/std/SIA" xsi:type="sia10:SimpleImageAccess">
      <validationLevel>2</validationLevel>
    - <interface role="std" version="1.0" xsi:type="vs11:ParamHTTP">
        <accessURL use="base">http://cda.harvard.edu/cxcsiap/queryImages?</accessURL>
      </interface>
    </capability>
  - <coverage>
      <waveband>X-ray</waveband>
    </coverage>
  - <tableset>
      <schema/>
    </tableset>
  </vr:Resource>
```


タイトル

識別子

発行者情報

API 規約に関する情報

データの観測域や周波数域に関する情報

テーブルメタデータ関情報

コンテンツ内容に関する情報

位置指定によるカタログデータ検索例

http://cda.harvard.edu/cscvo/coneSearch?RA=53.117&DEC=-27.808&SR=0.01


```
−<VOTABLE version="1.2" xsi:schemaLocation="http://www.ivoa.net/xml/VOTable/v1.2 http://www.ivoa.net/xml/VOTable/v1.2">
 −<RESOURCE name="CSC">
  −<TABLE name="RA53.117,DEC=27.808,SR=0.01,VERB=2,TIME=2017/06/03_03:44:56">
     <DESCRIPTION> Chandra Source Catalog VO Cone Search Service </DESCRIPTION>
   −<GROUP utype="stc:AstroCoordSystem" ID="ICRS">
      <DESCRIPTION>ICRS Coordinate System</DESCRIPTION>
    −<GROUP utype="stc:AstroCoordSystem.TimeFrame">
       <PARAM name="TimeScale" datatype="char" arraysize="*" utype="stc:AstroCoordSystem.TimeFrame.TimeScale" value="TT"/>
       <PARAM name="TimeRefPosition" datatype="char" arraysize="*" utype="stc:AstroCoordSystem.TimeFrame.ReferencePosition" value="TOPOCENTER"/>
     </GROUP>
    −<GROUP utype="stc:AstroCoordSystem.SpaceFrame">
       <PARAM name="SpaceRefFrame" datatype="char" arraysize="*" utype="stc:AstroCoordSystem.SpaceFrame.SpaceRefFrame" value="ICRS"/>
       <PARAM name="SpaceRefPosition" datatype="char" arraysize="*" utype="stc:AstroCoordSystem.SpaceFrame.ReferencePosition" value="TOPOCENTER"/>
     </GROUP>
    −<GROUP utype="stc:AstroCoordSystem.SpectralFrame">
       <PARAM name="SpectralRefPosition" datatype="char" arraysize="*" utype="stc:AstroCoordSystem.SpectralFrame.ReferencePosition" value="TOPOCENTER"/>
     </GROUP>
   </GROUP>
   −<GROUP ID="ICRScoords" utype="stc:AstroCoords">
      <DESCRIPTION>ICRS coordinates</DESCRIPTION>
      <PARAM datatype="char" arraysize="*" ucd="pos.frame" name="cooframe" utype="stc:AstroCoords.coord_system_id" value="ICRS"/>
      <FIELDref ref="col1"/>
      <FIELDref ref="col2"/>
      <FIELDref ref="col3"/>
   </GROUP>
   −<FIELD ID="col0" arraysize="*" datatype="char" name="name" ucd="ID_MAIN" width="20">
    −<DESCRIPTION>
       Source name in the format 'CXO Jhhmmss.s +/- ddmmss'
     </DESCRIPTION>
   </FIELD>
   −<FIELD ID="col1" datatype="double" name="ra" precision="F5" ref="ICRScoords" ucd="POS_EQ_RA_MAIN" utype="stc:AstroCoords.Position2D.Value2.C1" width="9">
      <DESCRIPTION>Source position, ICRS right ascension</DESCRIPTION>
   </FIELD>
   −<FIELD ID="col2" datatype="double" name="dec" precision="F5" ref="ICRScoords" ucd="POS_EQ_DEC_MAIN" utype="stc:AstroCoords.Position2D.Value2.C2"
   width="9">
      <DESCRIPTION>Source position, ICRS declination</DESCRIPTION>
   </FIELD>
   −<FIELD ID="col3" datatype="double" name="err_ellipse_r0" precision="F2" ref="ICRScoords" ucd="phys.angSize.smaj;pos.errorEllipse" unit="arcsec"
   utype="stc:AstroCoords.Position2D.Error2.C1" width="7">
    −<DESCRIPTION>
       Major radius of the 95% confidence level error ellipse
     </DESCRIPTION>
   </FIELD>
   −<FIELD ID="col4" datatype="boolean" name="conf_flag" ucd="meta.code" width="5">
      <DESCRIPTION>Source regions overlap (source is confused)</DESCRIPTION>
   </FIELD>
```


位置指定によるカタログデータ検索例（続き）


```
   −<FIELD ID="col21" datatype="short" name="var_intra_index_w" ucd="src.var.index;em.X-ray" width="6">
    −<DESCRIPTION>
       Intra-observation Gregory-Loredo variability index in the range [0, 10] (highest value across all observations); HRC wide energy band
     </DESCRIPTION>
     <VALUES null="-32768"/>
   </FIELD>
   −<FIELD ID="col22" datatype="short" name="var_inter_index_b" ucd="src.var.index;em.X-ray" width="6">
    −<DESCRIPTION>
       Inter-observation variability index in the range [0, 10]; indicates whether the source region photon flux is constant between observations; ACIS broad energy band
     </DESCRIPTION>
     <VALUES null="-32768"/>
   </FIELD>
   −<FIELD ID="col23" datatype="short" name="var_inter_index_w" ucd="src.var.index;em.X-ray" width="6">
    −<DESCRIPTION>
       Inter-observation variability index in the range [0, 10]; indicates whether the source region photon flux is constant between observations; HRC wide energy band
     </DESCRIPTION>
     <VALUES null="-32768"/>
   </FIELD>
   −<DATA>
    −<TABLEDATA>
     −<TR>
        <TD>CXO J033226.5-274801</TD>
        <TD>53.110715 04</TD>
        <TD>-27.80050579</TD>
        <TD>0.751905</TD>
        <TD>F</TD>
        <TD>F</TD>
        <TD>F</TD>
        <TD>8.43114E-16</TD>
        <TD>5.77894E-16</TD>
        <TD>1.15579E-15</TD>
        <TD/>
        <TD/>
        <TD/>
        <TD>3.05034</TD>
        <TD>0.100104</TD>
        <TD>-0.0616492</TD>
        <TD>0.264752</TD>
        <TD>-0.427867</TD>
        <TD>-0.615706</TD>
        <TD>-0.228536</TD>
        <TD>8</TD>
        <TD>-32768</TD>
        <TD>-32768</TD>
        <TD>-32768</TD>
      </TR>
    </TABLEDATA>
   </DATA>
  </TABLE>
 </RESOURCE>
</VOTABLE>
```

ADQL によるカタログ検索例

ADQL: Astronomical Data Query Language. (天文版 SQL)

http://cda.cfa.harvard.edu/csctap/sync?REQUEST=doQuery&LANG=ADQL&QUERY=SELECT * FROM master_source WHERE ra BETWEEN 53 and 54 AND dec BETWEEN -28 and -27&FORMAT=VOTable

画像データ検索例

http://tgcat.mit.edu/tgSia.php?POS=53.4,-25.9&SIZE=1

```
-<VOTABLE version="1.1" xsi:noNamespaceSchemaLocation="http://www.ivoa.net/xml/VOTable/VOTable/v1.1">
  -<DESCRIPTION>
    TGCat VOTable service. For more information visit TGCat at http://tgcat.mit.edu or send inquiry to tgcat@space.mit.edu
  </DESCRIPTION>
  <COOSYS ID="J2000" system="ICRS" epoch="J2000" equinox="J2000"/>
  -<RESOURCE type="results" name="POS 53.4,-25.9 SIZE 1">
    <INFO name="QUERY_STATUS" value="OK"/>
    -<TABLE>
      -<DESCRIPTION>
        Table describing images for the query string POS 53.4,-25.9 SIZE 1 For questions/comments please see http://tgcat.mit.edu or write
      </DESCRIPTION>
      -<FIELD name="title" datatype="char" arraysize="*" ucd="VOX:Image_Title">
        <DESCRIPTION> Description of image </DESCRIPTION>
      </FIELD>
      -<FIELD name="link" datatype="char" arraysize="*" ucd="VOX:Image_AccessReference">
        <DESCRIPTION> Provides link to retrieve image </DESCRIPTION>
      </FIELD>
      -<FIELD name="obsid" datatype="char" arraysize="*" ucd="meta.id">
        <DESCRIPTION> Chandra Observation ID ( see cxc.harvard.edu ) </DESCRIPTION>
      </FIELD>
      -<FIELD name="ra" datatype="double" unit="degree" ucd="POS_EQ_RA_MAIN">
        <DESCRIPTION> Right Ascension </DESCRIPTION>
      </FIELD>
      -<FIELD name="dec" datatype="double" unit="degree" ucd="POS_EQ_DEC_MAIN">
        <DESCRIPTION> Declination </DESCRIPTION>
      </FIELD>
      -<FIELD name="format" datatype="char" arraysize="*" ucd="VOX:Image_Format">
        <DESCRIPTION> image format </DESCRIPTION>
      </FIELD>
      -<FIELD name="num_axes" datatype="int" ucd="VOX:Image_Naxes">
        <DESCRIPTION> Number of axes in image </DESCRIPTION>
      </FIELD>
      -<FIELD name="num_axis" datatype="int" arraysize="*" unit="pixel" ucd="VOX:Image_Naxis">
        <DESCRIPTION> size of each axis in image ( x,y,z ) </DESCRIPTION>
      </FIELD>
      -<FIELD name="scale" datatype="double" arraysize="*" unit="degree/pixel" ucd="VOX:Image_Scale">
        <DESCRIPTION> Scale of each axis in Deg/Pixel </DESCRIPTION>
      </FIELD>
      -<DATA>
```

```
-<TABLEDATA>
  -<TR>
    <TD> 33189 </TD>
    <TD> CXO J033200.3-274320 </TD>
    <TD> 3 </TD>
    <TD> 0 </TD>
    <TD> 3 </TD>
    <TD> 0 </TD>
    <TD> 6 </TD>
    <TD> 2 </TD>
    <TD> 6 </TD>
    <TD> 2 </TD>
    <TD></TD>
    <TD> -27.722388291465715 </TD>
    <TD> 90.00035095214844 </TD>
    <TD> 0.259531 </TD>
    <TD> 0.259531 </TD>
    <TD> 1.214E-14 </TD>
    <TD> 5.303E-15 </TD>
    <TD> 1.967E-15 </TD>
    <TD> 2.009E-15 </TD>
    <TD> 1.569E-15 </TD>
    <TD></TD>
    <TD> -0.0461659 </TD>
    <TD> -0.327494 </TD>
    <TD> -0.281328 </TD>
    <TD></TD>
    <TD> 53.00157039052561 </TD>
    <TD> 12.2483 </TD>
    <TD> 0 </TD>
    <TD> 0 </TD>
    <TD> 22 </TD>
    <TD> 0 </TD>
    <TD> 0 </TD>
    <TD> 0 </TD>
    <TD> 0.0 </TD>
    <TD> 0.0 </TD>
    <TD> 1960780.0 </TD>
    <TD> 1.24927E-14 </TD>
    <TD> 5.35621E-15 </TD>
    <TD> 2.00449E-15 </TD>
    <TD> 2.01351E-15 </TD>
    <TD> 1.09641E-15 </TD>
    <TD></TD>
    <TD> 1.29297E-14 </TD>
    <TD> 4.01756E-15 </TD>
```

```
-<TABLEDATA>
  -<TR>
    <TD> NGC 1360 [HRC-LETG] evt2 file </TD>
    -<TD>
      http://tgcat.mit.edu/tg3et.php?id=24638file=evt2.gz&src=vo-sia
    </TD>
    <TD> 608 </TD>
    <TD> 53.31104 </TD>
    <TD> -25.87168 </TD>
    <TD> image/fits </TD>
    <TD> 2 </TD>
    <TD> 65536 65536 </TD>
    <TD> .0000 66597223 </TD>
  </TR>
  -<TR>
    <TD> NGC 1360 [HRC-LETG] full field summary </TD>
    -<TD>
      http://tgcat.mit.edu/png/tgcid_2463_summary_im-b.png
    </TD>
    <TD> 608 </TD>
    <TD> 53.31104 </TD>
    <TD> -25.87168 </TD>
    <TD> image/png </TD>
    <TD> 2 </TD>
    <TD> 926 346 </TD>
    <TD> 0.0 </TD>
  </TR>
  -<TR>
    <TD> NGC 1360 [HRC-LETG] summary webpage </TD>
    <TD> http://tgcat.mit.edu/tgPlot.php?t=P&i=2463 </TD>
    <TD> 608 </TD>
    <TD> 53.31104 </TD>
    <TD> -25.87168 </TD>
    <TD> text/html </TD>
    <TD> 2 </TD>
    <TD> 0.0 </TD>
    <TD> 0.0 </TD>
  </TR>
</TABLEDATA>
</DATA>
</TABLE>
</RESOURCE>
</VOTABLE>
```

UCD属性値が
“VOX:Image_AccessReference” となっているカラムがデータ取得先 URL

VOSpace

リモートサーバー上のデータストレージを利用するための仕様

- ファイルのアップロード、ダウンロード、削除、コピー、移動
- ディレクトリ作成
- ファイルのメタデータ操作
- ファイルの URI リストの取得

ファイルの URI

- vos://nvo.caltech!vospace/myresult/siap-out-1.vot
- nvo.caltech!vospace は この VOSpace の Resource Identifier
  - ivo://nvo.caltech/vospace
  - "/" が "!" に置き換わる
- ivo://nvo.caltech/vospace のアクセス URL は Registry より取得
- どこの VOSpace のどのファイルかを一意に指定

VOSpace 動作例

例１. ファイルアップロード

1. http://endpoint/sync にアップロード先 URI を送信
2. 利用可能なプロトコル (通常 http/put)とファイルのアップロード先 URL を受信
3. 指定されたプロトコルで指定された URL にファイルを送信

例２. ファイルダウンロード

1. http://endpoint/sync にダウンローするファイルの URI を送信
2. 利用可能なプロトコル (通常 http/get) とファイルの取得先 URL を受信
3. 指定されたプロトコルで指定された URL からファイルを受信

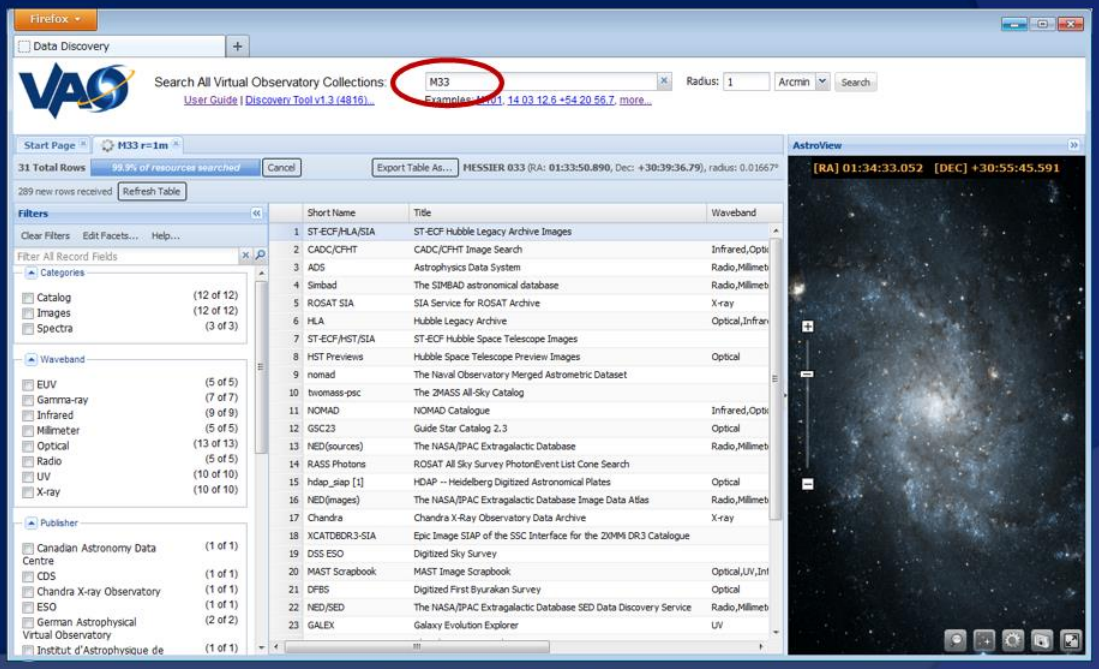
例３. 指定したディレクトリ配下のファイルリストを取得

1. http://endpoint/nodes/mydir に http/get でアクセス
2. mydir 配下のファイルのリストを受信

Datascope

http://www.usvao.org/science-tools-services/vao-tools-services-data-discovery-tool/

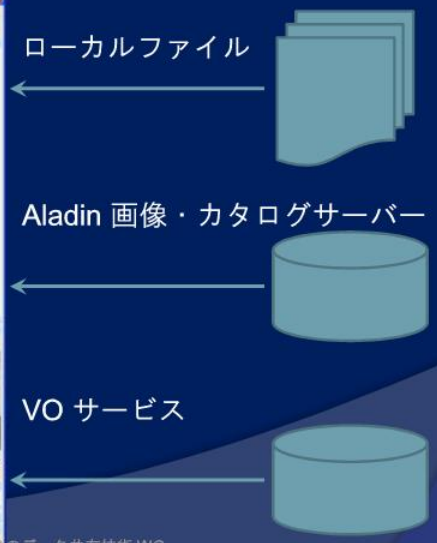
全VOサービスに対し、指定された領域のデータ検索を並列実行



Aladin

http://aladin.u-strasbg.fr/aladin.gml

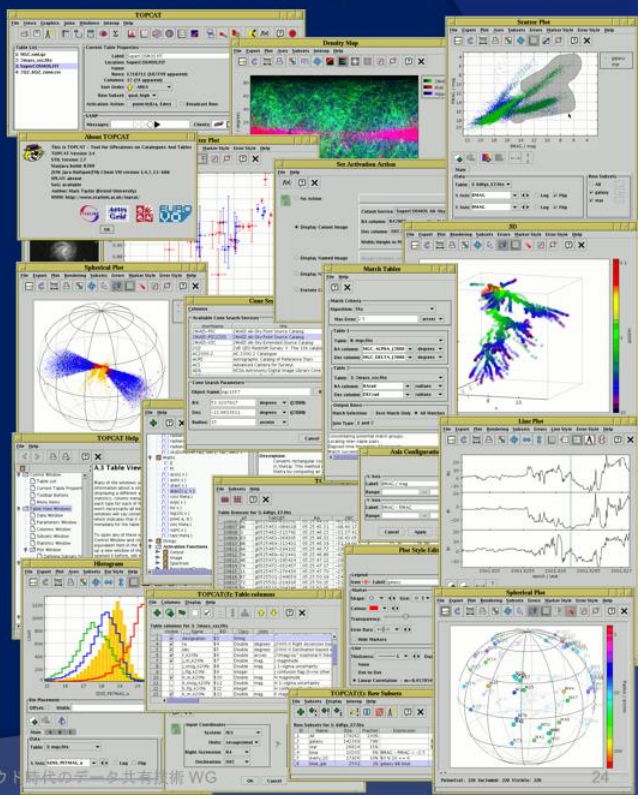
- フランス ストラスブルグ天文データセンター (CDS)
- 画像データ、カタログデータを取得し表示。



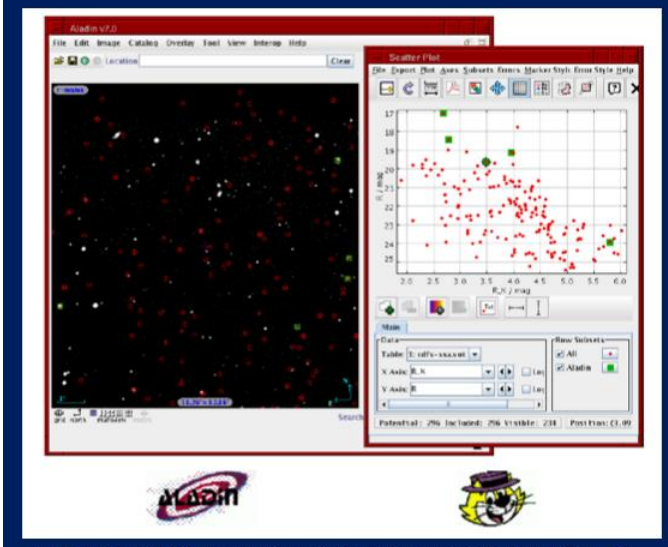
Topcat

http://www.starlink.ac.uk/topcat/

- イギリス Starlink プロジェクト (Mark Taylor)
- 様々な種類のプロット
- データはローカルファイルからロードする他、VOサービスからも取得可能。
- 複数カタログのクロスマッチ機能など。

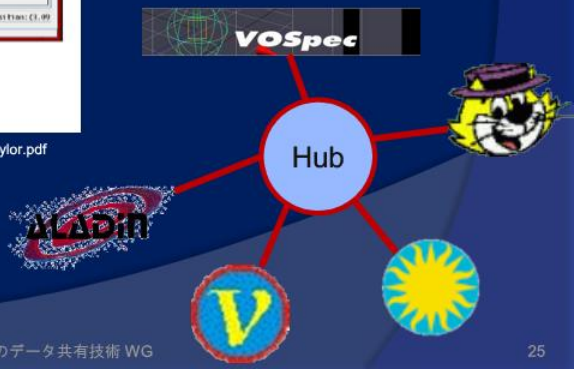


SAMP によるアプリケーション間連携



http://www.eso.org/sci/php/meetings/adass2011/Slides/PDF/All/ADASS_XXI_O26_Taylor.pdf

- Simple Application Messaging Protocol
- アプリケーション間の通信規約
- ハブを介してデータのやり取りやアクションの実行



Japanese VO (JVO) システム

JVO portal

- http://jvo.nao.ac.jp/portal
- VO 検索機能を利用した他サイトへの検索実行
- Subaru, ALMA などの処理済データの配信
- Subaru：10 TB, ALMA：15 TB (FITS データのみ)

VO インターフェイスによるデータ検索サービスの運用

- VO 対応アプリからの検索要求受付

HSC データ公開システム

http://hsc.mtk.nao.ac.jp/ssp/data-release/

現在公開中のデータセット

- ～100 平方度
- 五つの広帯域フィルター、四つの狭帯域フィルター
- 天体カタログ＋画像データ
- データ量　～100TB？
- ユーザ登録が必要

データ検索方法

- Web フォーム入力による検索
- Web フォームから SQL 検索
- コマンドラインから SQL 検索
- Web ブラウザでディレクトリ構造をたどってデータ取得

SMOKA システム

http://smoka.nao.ac.jp/

- すばる望遠鏡、岡山観測所、木曽天文台等で取得された全生データの公開
- 合計 120 TB
- Web フォームによる検索。
- データは ftp により取得。
- ユーザー登録が必要

国立天文台データアーカイブの現状の課題

- 各サブシステム毎にディスク領域を独立に確保
  - 同じデータが複数のサブシステムに重複して存在
  - ディスク資源の有効活用
- 複数のサブシステムで共通に利用されるデータは共通のデータ置き場に配置したい。
  - 各サブシステムとそのデータ置き場との間のデータ転送のバンド幅も十分なものが必要。
  - 一度に多数のプロセスが同時にデータを転送を行う場合がある。
- 必要な時に必要に応じてデータ領域の拡張を行いたい。
  - 現状は RAID 増設によるある新たな partition の追加。
  - partition があふれないよう監視する必要がある

A.5. 渡邉委員

【広島大学】
・共同研究者向け資料共有
・共同研究者向けコード共同開発

ポリシーやガバナンス事例

クラウドサービスを利用する場合、クラウド事業者との間で外部委託契約を行うことになるため、クラウド事業者の選定、サービス内容の確認、責任体制の構築等を慎重に行う必要がある。一方、選定基準や確認事項などについては学術機関共通の指標があるわけではなく利用する側で定めておかなければならない。本項ではクラウドサービス利用におけるガイドラインやチェックリストを先行して提供している広島大学の事例を紹介する。

広島大学では、大学が保有している情報（法人文書）の保存先としてクラウドサービスを利用する場合を想定した「広島大学クラウドサービス利用ガイドライン」（以下、ガイドラインという）を策定している。クラウドサービスを利用する際に広島大学として確認すべき項目の解説がされており、チェックリストにより広島大学が定める法人文書の重要度の応じたクラウドサービスの利用可否を容易に確認できる資料となっている。クラウドサービスの利用者は導入前にチェックリストの項目を確認し、その結果を指定された部署に提出する。また、利用者は1年を超えない期間ごとにその時点で最新のチェックリストの項目を再確認し提出する。これにより組織的なクラウドサービスの利用状況の把握することに加えて、情報セキュリティ関連の要求要件とクラウドサービスのセキュリティ対策状況の対応を定期的に見直す方針となっている。

SS研クラウド時代のデータ共有技術WG第2回会合@富士通本社, 2017/6/9

Gfarm事例を含めたセキュリティや関連ノウハウ・課題など

渡邉英伸

広島大学 情報メディア教育研究センター

広島大学

発表内容

広島大学

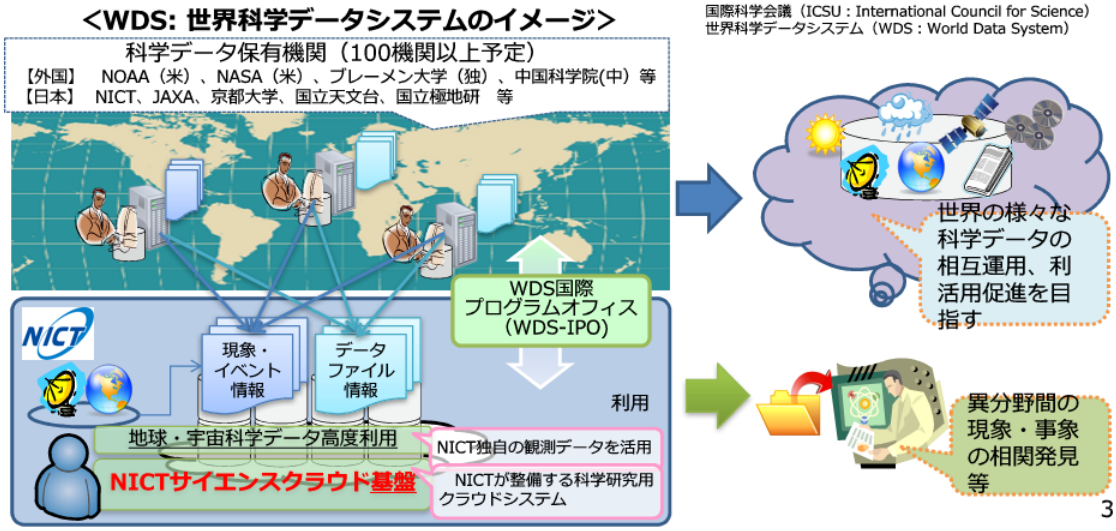
- Gfarm事例のセキュリティ
- クラウド時代の情報セキュリティ関連ノウハウ
- クラウド時代の情報セキュリティ課題
- 最後に

2

世界科学データシステム

広島大学

- 世界科学データシステム(WDS)の国際プログラムオフィス活動支援と国際連携
- グローバルオープンデータ利活用システムの実現を目指す研究開発
  - 科学データの分野横断検索など科学データ利活用支援に関する研究
  - 科学データの相互運用・一括情報管理など情報利活用基盤に関する研究



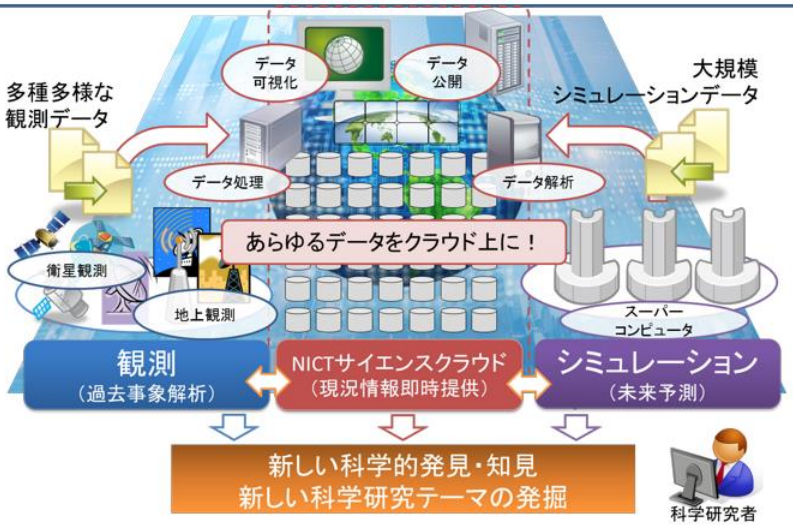
3

NICTサイエンスクラウド

広島大学

- **データ指向科学研究向けクラウドシステム**
  - 約550コア、約3.2PB容量の分散並列処理環境（※）
  - 約5億ファイルの蓄積（冗長化込）（※）
  - 約230アカウント発行（※）
  - 2010年よりテストベッド運用開始

※H27年1月時点

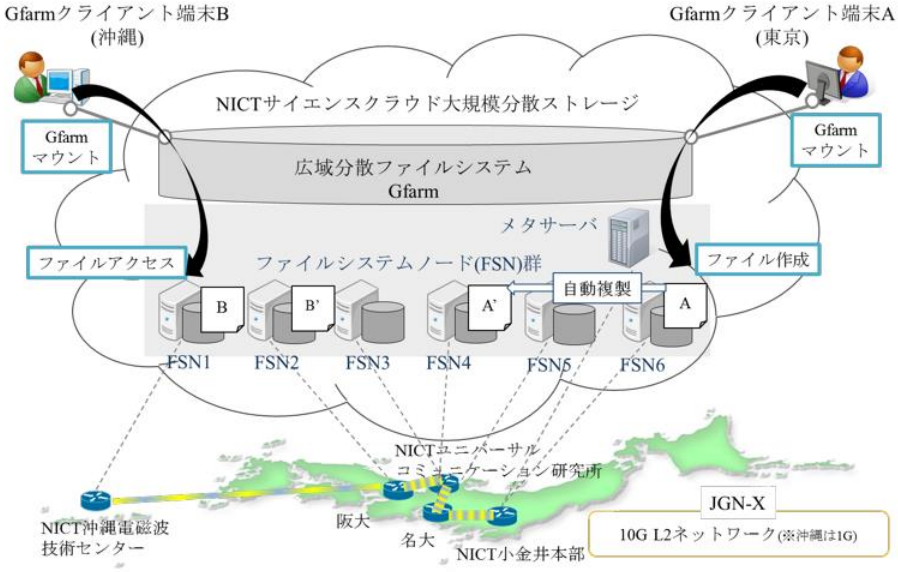


4

NICTサイエンスクラウド広域分散型ストレージ

広島大学

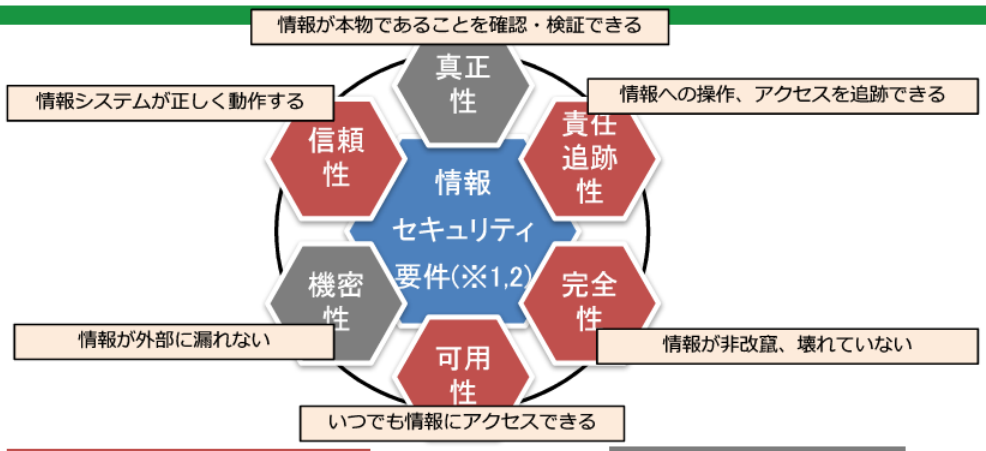
- **HPCIの共有ストレージに採用されているオープンソースの分散ファイルシステムGfarmで広域分散型ストレージを構築**
  - 東京、名古屋、京都、大阪、沖縄の5拠点に分散配置
  - JGN-Xの10GbEで接続



5

情報セキュリティ要件とH26年度Gfarmの対応状況

広島大学



Gfarmの対応済セキュリティ要件

- **信頼性**：徹底した機能試験によって信頼性を達成
- **可用性**：レプリケーション機能によりデータの冗長化は達成済、※2013年**SLA**(Service Level Agreement)評価手法の提案済
- **完全性**：2014年3月v2.5.8.5でファイルのチェックサム機能提供(それまではTCPチェックサムレベルでの担保のみ)
- **責任追跡性**：2012年v2.5でログフォーマット定義済、※ログを追跡する外部アプリケーションと連携可能

Gfarmのセキュリティ課題

- **真正性**：未実装
- **機密性**：ユーザオンデマンドのデータ暗号化が未実装
- 完全性：
- 責任追跡性

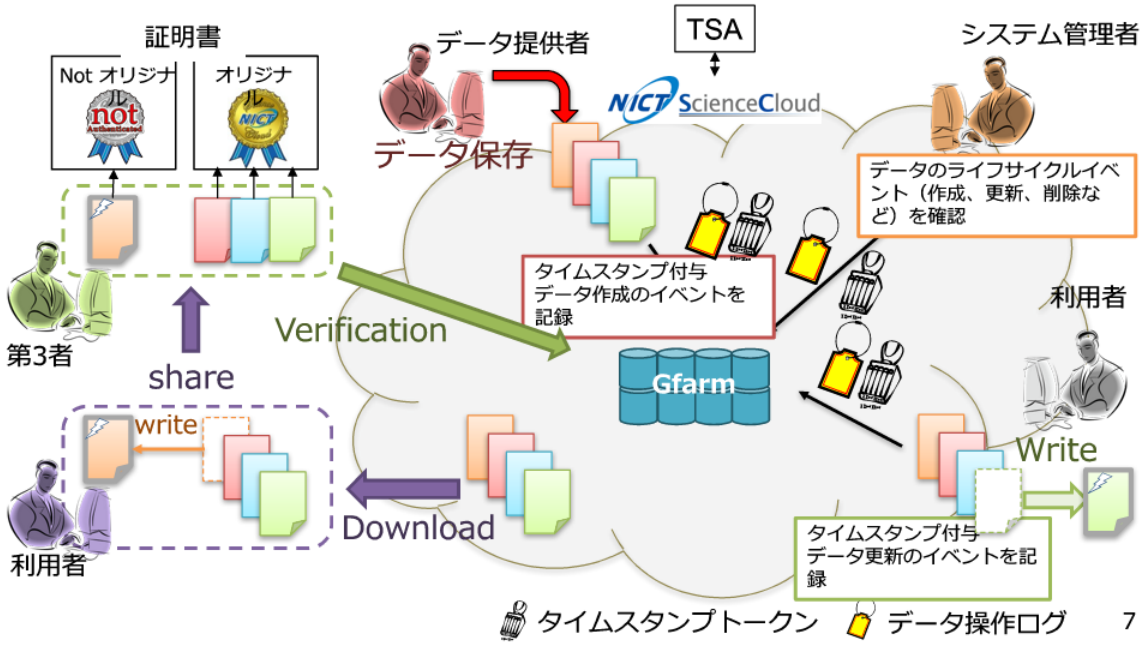
※1, JIS Q 27002 (ISO/IEC 27002)
※2, 経済産業省，高度情報化社会における情報システム・ソフトウェアの信頼性及びセキュリティに関する研究会，情報システム・ソフトウェアの信頼性及びセキュリティの取組強化に向けて～豊かで安全・安心な高度情報化社会に向けて～-中間報告書

6

エビデンスベースのデータ共有

広島大学

- **タイムスタンプ技術・ログトレース技術によってデータの完全性・真正性・責任追跡性を保証するデータ長期保全システムの構築中**



7

クラウド時代の
情報セキュリティ関連ノウハウ

広島大学

学術機関のクラウド活用を妨げる要因

広島大学

- **学術機関がクラウド化をしない1番の理由**
  – 情報セキュリティの不安　＞　コストの不安

数年間
変化なし

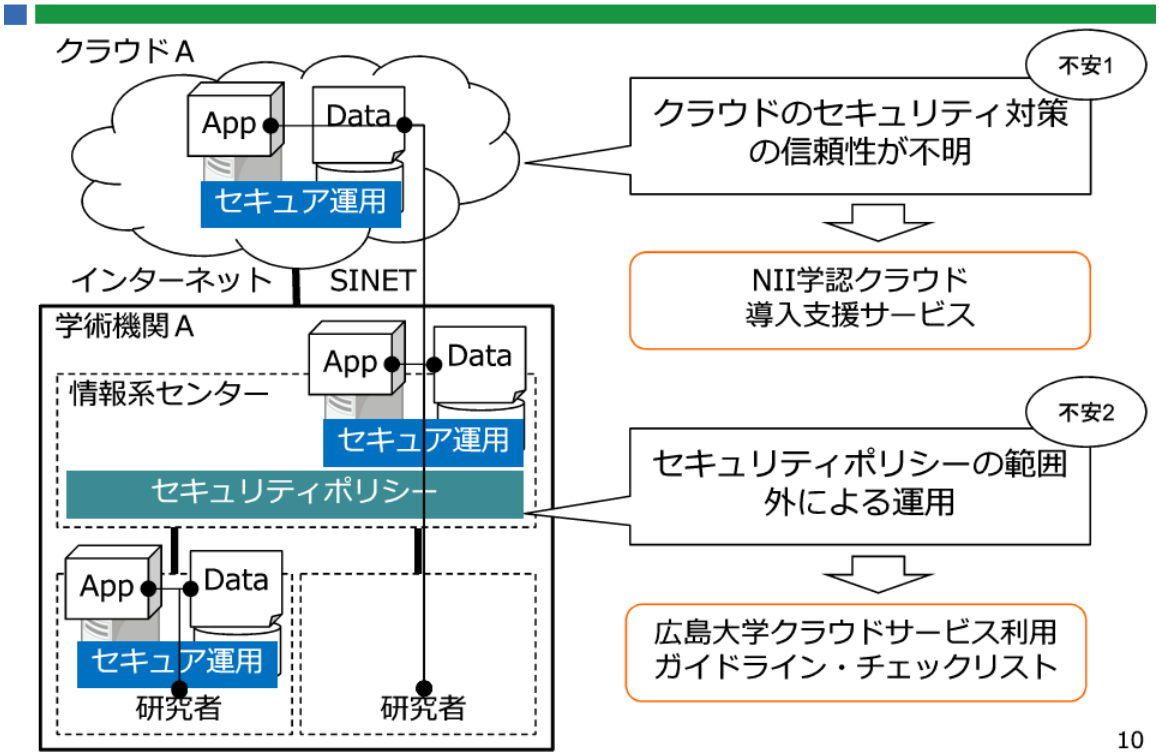
➢ パブリッククラウドサービスを利用しない、あるいは利用を妨げている理由	学術機関 (n=594)	企業 (n=722)
➢ 情報漏洩などセキュリティに不安がある	49.2	34.0
➢ クラウドの導入に伴う既存システムの改修コストが大きい	19.0	22.6
➢ ニーズに応じたアプリケーションのカスタマイズができない	19.0	12.8
➢ ネットワークの安定性に対する不安がある	17.8	15.2
➢ 法制度や所属機関の諸規則が整っていない	16.8	6.3
➢ メリットが分からない、判断できない	16.0	21.5
➢ 通信費用がかさむ	9.8	10.9
➢ クラウドの導入によって所属機関の諸規則との整合性に支障をきたす	8.9	6.2
➢ 必要がない	8.2	41.2
➢ その他	8.2	5.8

＜出展＞
コミュニティで紡ぐ次世代大学ICT環境としてのアカデミッククラウド（セキュリティ分野）：http://www.icer.kyushu-u.ac.jp/ac
平成24年通信利用動向調査（企業編）（総務省）：http://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html

9

クラウドセキュリティにおける現状

広島大学



10

NII学認クラウド導入支援サービス

広島大学

<出展>
吉田 浩，“大学・研究機関のクラウド利用拡大に向けたNIIの取組み － 導入検討から活用まで”，学等におけるクラウドサービス利用シンポジウム2017

11

NII学認クラウド導入支援サービス

広島大学

チェックリスト Ver.2.0

GakuNin Cloud

チェック項目（大項目）	詳細チェック項目数	主な詳細チェック項目
商品 / サービスの概要	4	タイトル、製品概要など
運用実績	2	契約法人数、運用年数
契約申込み	12	支払方法、ライセンス体系など
学認対応状況	2	Shibboleth利用可否、学認対応状況
信頼性	7	サービス稼働率の実績、計画停止の頻度など
サポート関連	5	サポート窓口、サポート回答時間など
ネットワーク・通信機能	9	SINET接続状況、通信の暗号化可否など
管理機能	11	稼働状況の一覧表示ツール、利用統計など
動作保証	3	利用可能OS、動作事例など
スケーラビリティ	6	資源の上限、作成可能なサーバ上限数など
データセンター	8	防犯設備、データの保存場所など
セキュリティ	10	セキュリティ対策、インシデント対応など
データ管理	10	データの多重化、ログなど
バックアップ	6	バックアップサービスの有無、リストアなど
クラウド事業者の信頼性	6	第三者認証、委託先での個人情報保護など
契約条件	6	責任範囲の明確化、損害賠償責任など
データの取り扱い	5	データの所有権 / 利用権、削除の方法など
データの引継ぎ	4	契約終了時の移行支援、イメージの移行性など

Copyright ©2017 National Institute of Informatics

<出展>
吉田 浩，“大学・研究機関のクラウド利用拡大に向けたNIIの取組み – 導入検討から活用まで”，学等におけるクラウドサービス利用シンポジウム2017

12

NII学認クラウド導入支援サービス

広島大学

現在の利用機関・プロバイダ

GakuNin Cloud

- 参加大学・研究機関(26)
  - 慶應義塾
  - 横浜国立大学
  - 室蘭工業大学　情報メディア教育センター
  - 東京有明医療大学
  - 鹿屋体育大学
  - 西日本工業大学
  - 佐賀大学
  - 京都大学
  - 名古屋大学
  - 群馬大学
  - 中京大学
  - 国立遺伝学研究所
  - 筑波大学
  - 九州大学
  - 中部大学
  - 宮崎大学
  - 広島大学
  - 千葉工業大学
  - 沖縄科学技術大学院大学
  - 岐阜経済大学
  - 東北大学
  - 静岡文化芸術大学　　他4機関
- クラウドプロバイダ(17)
  - 北海道大学情報基盤センター
  - 株式会社セールスフォース・ドットコム
  - 富士フイルムイメージングシステムズ株式会社
  - 株式会社セシオス
  - 伊藤忠テクノソリューションズ株式会社
  - 株式会社ミライコミュニケーションネットワーク
  - さくらインターネット株式会社
  - NTTコミュニケーションズ株式会社
  - クラウドエース株式会社
  - 日本マイクロソフト株式会社
  - 日本オラクル株式会社
  - GMOインターネット株式会社
  - 日本アイ・ビー・エム株式会社
  - ニフティ株式会社
  - 富士ゼロックス株式会社
  - アマゾン ウェブ サービス ジャパン株式会社
  - 株式会社ねこじゃらし

■ ：チェックリスト回答提供中
□ ：チェックリスト回答検証中(近日公開)

(2017年3月9日現在、順不同)

14

Copyright ©2017 National Institute of Informatics

<出展>
吉田 浩，“大学・研究機関のクラウド利用拡大に向けたNIIの取組み – 導入検討から活用まで”，学等におけるクラウドサービス利用シンポジウム2017

13

広島大学クラウドサービス利用ガイドライン・チェックリスト

広島大学

<出展>
西村 浩二，“ISMSクラウドセキュリティ認証の受審報告”，学等におけるクラウドサービス利用シンポジウム2017

14

ENISA：リスク項目(R1～R35)

広島大学

- **ポリシーと組織関連のリスク**

(R1)ロックイン
(R2)ガバナンスの喪失
(R3)コンプライアンスの課題
(R4)他の共同利用者の行為による信頼の喪失
(R5)クラウドサービスの終了または障害
(R6)クラウドプロバイダの買収
(R7)サプライチェーンにおける障害

- **技術関連のリスク**

(R8)リソースの枯渇(リソース割当の過不足)
(R9)隔離の失敗
(R10)クラウドプロバイダ従事者の不正－特権の悪用
(R11)管理用インターフェースの悪用(操作、インフラストラクチャアクセス)
(R12)データ転送途上における攻撃
(R13)データ漏えい(アップロード時、ダウンロード時、クラウド間転送)
(R14)セキュリティが確保されていない、または不完全なデータ削除
(R15)DDoS攻撃(分散サービス運用妨害攻撃)
(R16)EDoS攻撃(経済的な損失を狙ったサービス運用妨害攻撃)
(R17)暗号鍵の喪失
(R18)不正な探査またはスキャンの実施
(R19)サービスエンジンの侵害
(R20)利用者側の強化手順とクラウド環境との間に生じる矛盾

- **法的なリスク**

(R21)証拠提出命令と電子証拠開示
(R22)司法権の違いからくるリスク
(R23)データ保護に関するリスク
(R24)ライセンスに関するリスク

- **クラウドに特化していないリスク**

(R25)ネットワークの途絶
(R26)ネットワークの管理(ネットワークの混雑、接続ミス、適切でない使用)
(R27)ネットワークトラフィックの改変
(R28)特権の(勝手な)拡大
(R29)ソーシャルエンジニアリング攻撃(なりすまし)
(R30)運用ログの喪失または改ざん
(R31)セキュリティログの喪失または改ざん(フォレンジック捜査の操作)
(R32)バックアップの喪失、盗難
(R33)構内への無権限アクセス(装置その他の設備への物理的アクセスを含む)
(R34)コンピュータ設備の盗難
(R35)自然災害

<出展> *http://www.enisa.europa.eu/activities/risk- management/files/deliverables/cloud-computing-risk-assessment*
http://www.ipa.go.jp/security/publications/enisa/documents/Cloud%20 Computing%20Security%20Risk%20Assessment.pdf(日本語訳)

15

広島大学クラウドサービス利用ガイドライン・チェックリスト

広島大学

ガイドラインのチェック項目

広島大学

4. 利用に向けた準備（準備）
- 取り扱う情報の確認
  - 情報の格付け
  - クラウドサービスの選択
- 本学の組織・体制
  - クラウドサービス利用責任者
  - クラウドサービス利用担当者
- 規則・契約
  - 規則との整合性(3)
  - 契約の取扱い(2)

5. 利用範囲の明確化（検討）
- サービスの品質
  - SLA
  - メンテナンス
  - 問い合わせ窓口・サポート体制(2)
  - サービスの継続性
- 機能とコスト
  - コンピューティング
  - ストレージ
  - ネットワーク(3)
  - 管理機能
  - ライセンス(2)
  - コスト(2)

6. クラウド事業者の選定（検討）
- データセンター
  - データセンターの場所
  - 堅牢性
  - 機密性
- クラウド事業者の信頼性
  - 経営状況の確認
  - 委託関係の確認

7. 契約条件の確認（検討）
- 責任範囲とペナルティ
  - 責任範囲の明確化(2)
  - クラウド事業者のペナルティ
- データの所有権、返却・消去
  - データの所有権
  - データの返却(2)
  - データの消去(2)
- 準拠法と管轄裁判所
  - 準拠法
  - 管轄裁判所

8. 運用体制の確認（運用）
- システムの運用に関する項目
  - セキュリティ対策
  - ログの監視
- データの管理に関する項目
  - 秘密鍵の管理
  - バックアップ(2)
- インシデントの管理に関する項目
  - インシデントの記録(2)

2017/3/10　大学等におけるクラウドサービス利用シンポジウム2017

2017年3月に学術関連では日本で初めて、ISMSクラウドセキュリティ認証取得

<出展>
西村 浩二，“ISMSクラウドセキュリティ認証の受審報告”，学等におけるクラウドサービス利用シンポジウム2017

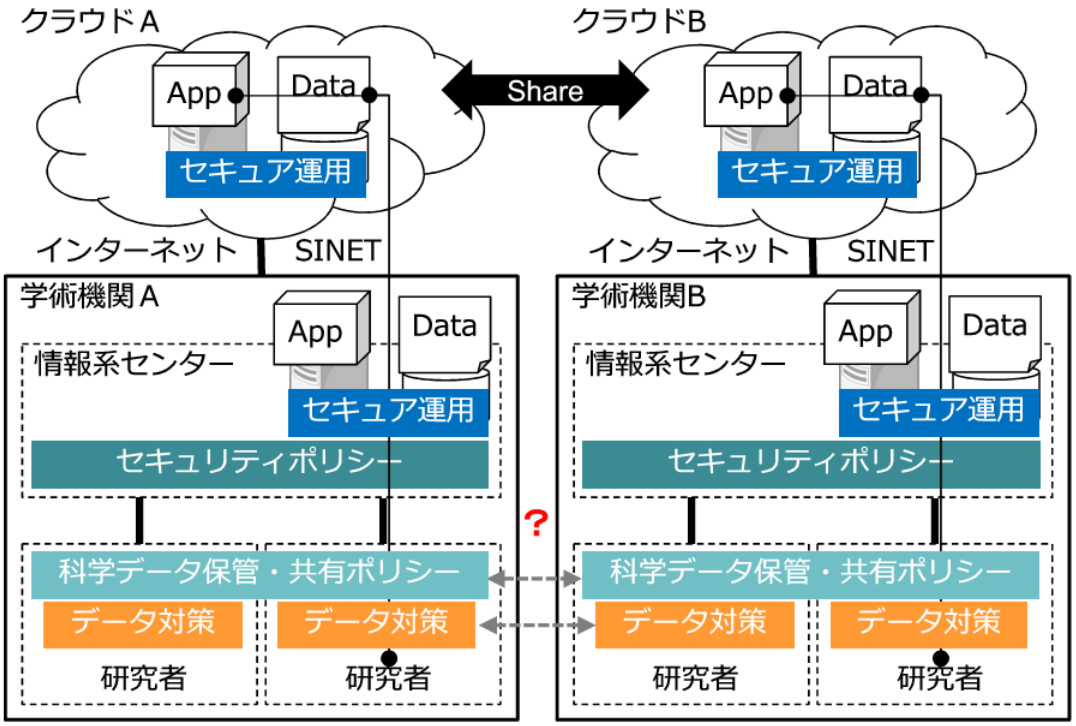
16

クラウド時代の情報セキュリティ課題

広島大学

マルチクラウドでのセキュリティ課題

広島大学



18

マルチクラウドでのセキュリティ課題

広島大学

レイヤ	課題		説明
Web	コンテンツ利用制御	マッシュアップ通信制御機構	同一生成元ポリシーの順守、コンポーネント間通信監視
		コンテンツ利用ポリシー	ポリシー宣言による開発者が意図しない動作の最小化
	アクセス制御の脆弱性		アクセス制御の不適切・不在、開発者側での対応
	クライアント側検証の脆弱性	エンドポイントでのリスク検出	信頼できないコード(Javascript等)を安全でない形で利用
		ミドルポイントでのリスク検出	非HTMLコンテンツをサンドボックスブラウザ上で実行
アプリケーション	セキュアなデータ保存	暗号化データの操作	プロキシ再暗合化技術、順同型暗号
		データの完全性	複数サーバによるストレージの正当性の確認
		データの完全性冗長性削除	冗長性削除技術を防ぐための暗号化
	認証と権限付与	認証フレームワーク	ロールベースアクセス制御、OpenID、x.509証明書
		クエリ制御	クエリの精査、セマンティックWeb言語
	ユーザ別のデータ隔離	プロセスの隔離	1つのコンポーネントが複数のデータを処理しない
		選択的複合技術	IDベース暗号、属性ベース暗号(ABE)
OS	特権の隔離		トロイの木馬、SELinux(Disableにされることが多い)
	カーネルの脆弱性		バグ、特権ユーザによる信頼できないソフトウェア実行
ハイパーバイザー	VMMのセキュリティ	脆弱性	VM上の悪意のあるコードによるVMM停止等
		VMベースのルートキット	ルートキット検出器による検出を回避
		透過性	リモートからのVMM存在の検出、VMMに対する攻撃
		プラットフォームの完全性	アプリケーション単位でのVM割当、vTPM、HIMA
	VMのセキュリティ	イントロスペクション	ハイパーバイザーからのゲストOSのデバッグ機能
		冗長性の削除	リソースの効率的活用の追求
HW&SWプリミティブ	プログラミングコード	マネージドランタイムセキュリティ	Java、.NETに対する信頼性の低下、JavaSnoop
		リソースの隔離	サンドボックスへのI/Oを通じたデータ混入
	ハードウェア処理	共有プロセッサリソース	ハイパースレッディング技術の脆弱性(サイドチャネル)

<出展>
高橋 他, “マルチテナントクラウドコンピューティング：セキュリティ上の課題とアプローチ”, 信学技報, ICSS2011-36, pp.25-32, 2011.

19

カントリーリスク

広島大学

- **米国愛国者法(Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act in 2001)**
  - テロリズム(201条)やコンピュータ関連の一定の犯罪(202条)について、通信傍受を行なう権利を認める
  - 一定の場合において、裁判所の命令なしに捜査官が電子メールやボイスメールを取得できる(209条)
  - 捜査官は、霊場の通知なく家宅など(データセンターを含む)を捜査できる(213条)
  - FBIは国際テロや秘密諜報活動の防止を目的とした捜査であることを明示し、金融機関や通信プロバイダの同意が得られれば、それらが有する顧客の個人情報を裁判所の関与なく得ることができる(505条)
- **EUデータ保護指令(Data Protection Directive)**
  - EU内での個人データの取り扱いに際するプライバシー権の保護が目的
  - EU域外の国に個人データを移転する場合、原則として、その国がEUデータ保護指令が要求する個人データ保護の水準を有していることを要求
  - 日本は、EUデータ保護指令上の個人データ保護の水準を有しているEU域外の国に含まれていない(例外規定はあるが注意が必要)
- **実例**
  - 2009年4月2日、FBIの捜査官が米国愛国者法に基づいてデータセンターを捜査し、2フロア分のサーバを押収したため、約50社の顧客に対するサービスが停止した

20

準拠法と管轄裁判所

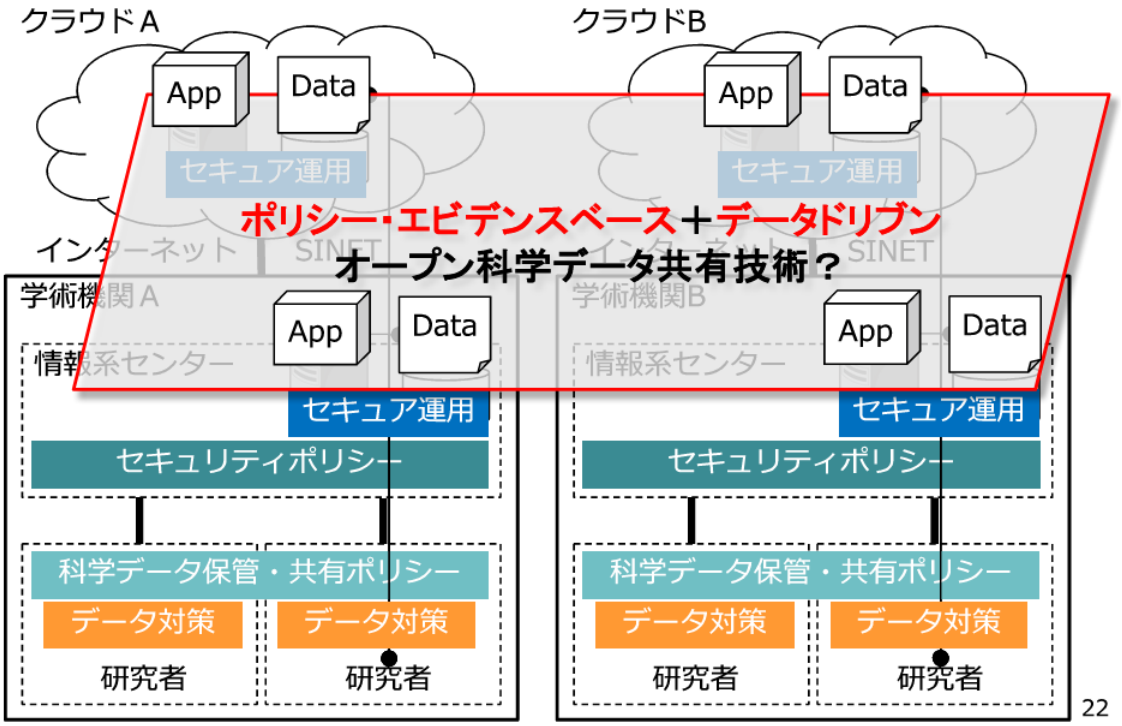
広島大学

- **Amazon Web Service**
  - AWSカスタマーアグリーメント
    - *http://aws.amazon.com/jp/agreement/*
  - 準拠法は「アメリカ合衆国ワシントン州法」
  - 管轄裁判所は「アメリカ合衆国ワシントン州キング郡に所在する州裁判所または連邦裁 判所」
- **Google Apps for Education**
  - Google Apps for Education 契約
    - *http://www.google.com/apps/intl/ja/terms/education_terms.html*
  - 準拠法は「アメリカ合衆国カリフォルニア州の州法」
  - 管轄裁判所は「アメリカ合衆国カリフォルニア州サンタクララ郡の裁判所」
- **Office 365 Education**
  - マイクロソフトオンラインサブスクリプション契約
    - *http://www.windowsazure.com/ja-jp/support/legal/subscription-agreement/*
  - 準拠法は「日本の法律」
  - 管轄裁判所は「東京地方裁判所」
- **Microsoft Azure**
  - マイクロソフトオンラインサブスクリプション契約
    - *http://www.windowsazure.com/ja-jp/support/legal/subscription-agreement/*
  - 準拠法は「日本の法律」
  - 管轄裁判所は「東京地方裁判所」

21

最後に

広島大学


クラウドA
クラウドB
App
Data
セキュア運用
ポリシー・エビデンスベース＋データドリブン
オープン科学データ共有技術？
インターネット
SINET
学術機関A
学術機関B
情報系センター
セキュリティポリシー
科学データ保管・共有ポリシー
データ対策
研究者
22

B. クラウド時代のデータ共有技術紹介

富士通からのデータ共有技術紹介を以下に示す。

B. 1. クラウドストレージ、ストレージソフトご紹介

FUJITSU
shaping tomorrow with you

SS研
クラウド時代のデータ共有WG
第4回会合
～クラウドストレージ、ストレージソフトご紹介～

2017年12月22日
富士通株式会社
次世代テクニカルコンピューティング開発本部
ソフトウェア開発統括部

Copyright 2017 FUJITSU LIMITED.

クラウドストレージ FUJITSU

- **特徴**
  - 負担・コスト軽減 ・・・ 管理者の負担や導入／運用コストを低減
  - 安全・安心 ・・・ 堅牢なセキュリティとデータ保全
  - 自由・柔軟 ・・・ 必要な時に必要なだけ
  - シームレス ・・・ いつでも、どこからでも
- **提供ベンダ**
  - Amazon AWS
  - Google Cloud Platform
  - Microsoft Azure
  - IBM Cloud
  - Alibaba Aliyunなど
- **動向**
  - 用途や価格体系、ベンダロックインやリスクの回避により、マルチクラウド利用が進行中
  - Vmware はvcloud air を売却、HPEもパブリッククラウドから撤退。

書き込み／読み出し

ログ保管・アーカイブ	★★★
コンテンツ保管・配信	★★★
オンラインストレージ	★★★
ファイルサーバー	★★
データベース	

更新

2 Copyright 2017 FUJITSU LIMITED.

クラウドストレージ比較

FUJITSU

■ クラウドストレージサービス比較

AWSが充実、各ベンダとも基本的なオブジェクトストレージサービスを提供。

	AWS	Google	Azure	IBM	Aliyun	Oracle
オブジェクトストレージ	○	○	○	○	○	○
アーカイブ	○	○	○	○		
ストレージゲートウェイ	○	○				
データ移行	○					
ファイルシステム	○	○	○	○		
ブロック	○			○		○

■ クラウドストレージサービス　価格比較(月あたり、GB/月の単価)

Capacity	AWS		Google		Azure		IBM		Aliyun	Oracle
	通常	アーカイブ	通常	アーカイブ	通常	アーカイブ	通常	アーカイブ	通常	通常
0 – 50TB	$0.023	$0.004	$0.023	Near $0.016 Cold $0.01	$0.023	**$0.002**	$0.03	$0.011	**$0.0197**	$0.0204
Next 450TB	$0.022				$0.022					$0.0201
Next 500TB	$0.021						$0.026			$0.0197
Next 4000TB					要相談					$0.0194
Over 5000TB										$0.019

Copyright 2017 FUJITSU LIMITED.

クラウドストレージ(AWS, Google)

FUJITSU

■ Amazon AWS

- https://aws.amazon.com/jp/products/storage/
- パブリッククラウドのリーダー的存在。IaaS, SaaSでシェア１位、IaaSは40%超。
- 幅広いユースケースにわたり、大量のサービスを提供。
  AWS Snowball, AWS Snowmobile といったデータ移行サービスも提供。
- Re:invent2017で新サービスを大量に発表。ストレージ観点では以下など。
  ・「Amazon S3 Select 」「Amazon Glacier Select」
  　データレイクの分析を容易にするため、必要なデータのみを取り出す仕掛け
  https://aws.amazon.com/jp/new/reinvent/

■ Google Cloud Platform

- https://cloud.google.com/storage/?hl=ja
- AI・機械学習を強みとして、AWS, MicroSoft Azureを追いかける状況。
  企業用途をアピールし、「エンタープライズ向けではない」という懸念払拭を図っている。
- ストレージ観点ではオブジェクトストレージ、FS、ブロックのラインナップに加えて、
  データ移行向けに「Cloud Storage Transfer Service」を提供。

Copyright 2017 FUJITSU LIMITED.

クラウドストレージ(Azure, IBM, Alibaba)

FUJITSU

Microsoft Azure

- https://docs.microsoft.com/ja-jp/azure/storage/
- IaaSで2位、SaaSで1位のシェア。豊富なMicrosoft製品をサービス化(office365など)。
  現在 40 のリージョンで、70 サービス・製品を展開。
- 小IO向けには Queue Storage を提供。最大64KBのメッセージを大量に格納可。
  Azure FilesではCIFSでの共有を提供。

IBM Cloud

- https://www.ibm.com/cloud-computing/jp-ja/products/storage/
- オブジェクトストレージ、FS、ブロックに加えて、コンテンツ配信ネットワーク(CDN)を提供。
  ネットワーク・トラフィックの混雑を回避し、遅延を削減。
  Akamaiの131カ国、約1,700のネットワークの存在と、IBMの19カ国、約60の
  クラウド・データセンターにおよぶグローバルなクラウド拠点を組み合わせ。
  http://www-03.ibm.com/press/jp/ja/pressrelease/53321.wss

Alibaba Aliyun

- https://jp.alibabacloud.com/product/oss
- IaaSで3位のシェア、中国のクラウド市場を支配。Softbankからも販売。
- オブジェクトストレージ、コンテンツ配信ネットワーク(CDN)、Table StoreというNoSQL データベースサービスを提供。

Copyright 2017 FUJITSU LIMITED.

クラウドストレージ(さくら、IDCF)

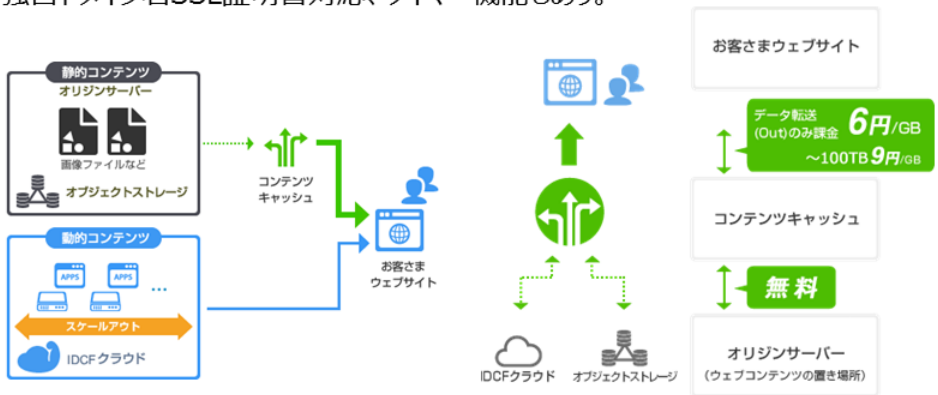
FUJITSU

さくら

- https://cloud.sakura.ad.jp/specification/object-storage/
- オブジェクトストレージの新規追加は停止中。
  データ転送の価格体系が従量課金なしの定額制あり。

IDCF

- https://www.idcf.jp/cloud/storage/
- コンテンツキャッシュはキャッシュサーバーによるアクセス負荷軽減。
  ワンタイムURL機能を使用すれば、ユーザーや期間を限定した配信が可能。
  独自ドメイン名SSL証明書対応、タイマー機能もあり。

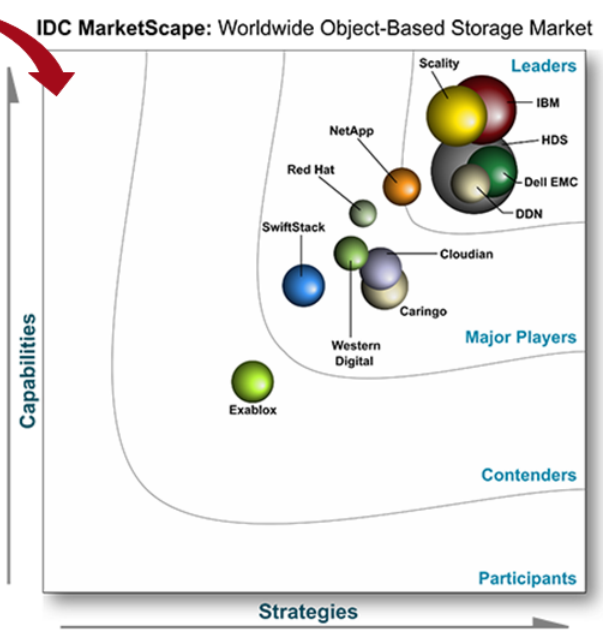


Copyright 2017 FUJITSU LIMITED.

オブジェクトストレージとファイルシステム

FUJITSU

	オブジェクトストレージ	ファイルシステム
単位	オブジェクト	ファイル
プロトコル	HTTP上のREST	NFS, CIFS
メタデータ	拡張可能	固定の属性
用途	コールドデータ クラウド	ファイル共有
構成	IDをフラットに管理	ディレクトリツリーで 階層管理
注意点	頻繁な更新は不向き 同一ファイルへの排他制御	システム外との共有
特徴	拡張性 分散アクセス レプリケーション 耐故障性	POSIX 成熟度



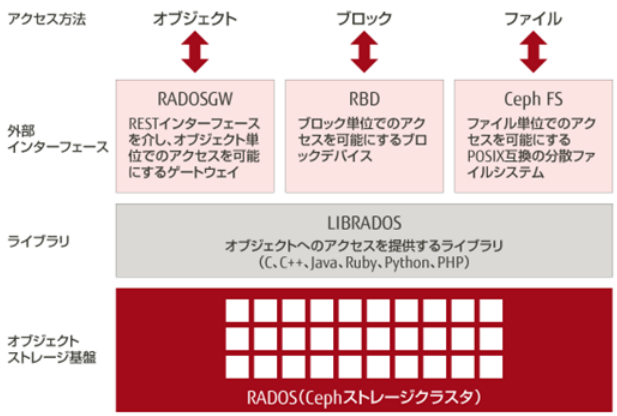
Source: IDC, 2016

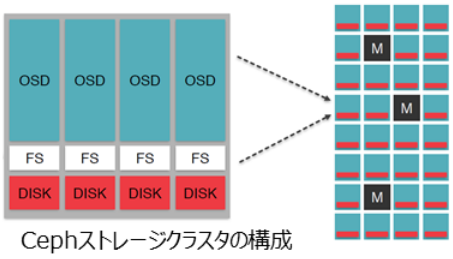
Copyright 2017 FUJITSU LIMITED.

オブジェクトストレージ(Ceph)

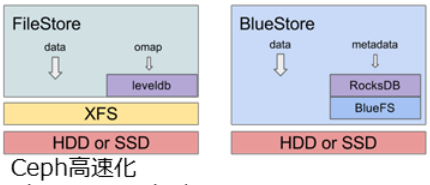
FUJITSU

- http://www.fujitsu.com/jp/products/computing/storage/lib-f/tech/beginner/ceph/
- 2007年から始まったオープンソース、LGPL v2.1。GitHubから入手可。
- Redhat(Redhat Ceph Storage), Suse(SUSE Enterprise Storage)が商用サポート。
- Reliable Autonomic Distributed Object Store (RADOS)による分散オブジェクトストレージ。CRUSHアルゴリズムにより大規模スケールアウト対応
- ストレージ、メタデータサービス、APIサービス、クラスター監視を提供する複数のノードによって構築。
- オブジェクトの他に、ブロックデバイス(RBD)とファイルデータ(CephFS)をサポート。





Cephストレージクラスタの構成



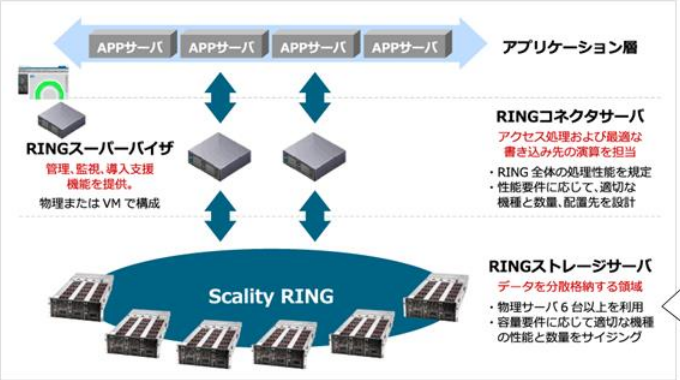
Ceph高速化
BlueStore=Block+NewStore

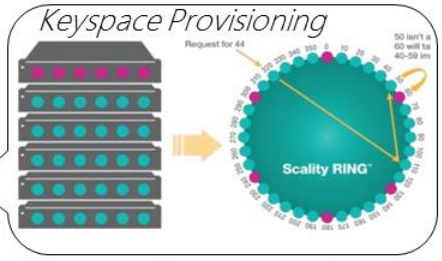
Copyright 2017 FUJITSU LIMITED.

オブジェクトストレージ(Scality)

FUJITSU

- **Scality Ring**
  - https://www.scality.com/products/ring/
  - RingはChordアルゴリズムを使用して分散管理
    Chordアルゴリズムは、キー空間(keyspace)をリング状に扱う。





- **Zenko**
  - https://www.zenko.io/
  - 複数のクラウドを同じAPIおよびレイヤで処理。
    フォーマットはMicrosoft Azure Blobストレージ、Amazon S3、
    Scality RINGおよびローカルDockerボリュームに対応。
  - CloudserverはS3 APIのオープンソース実装。Apache2.0。
    Docker イメージとして配布。

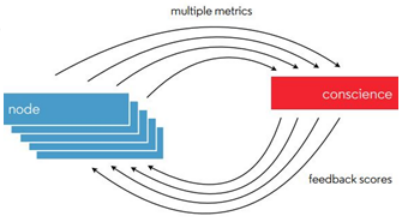
zenko.io

Copyright 2017 FUJITSU LIMITED.

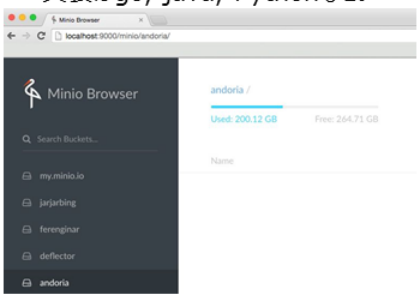
オブジェクトストレージ(OpenIO, minio)

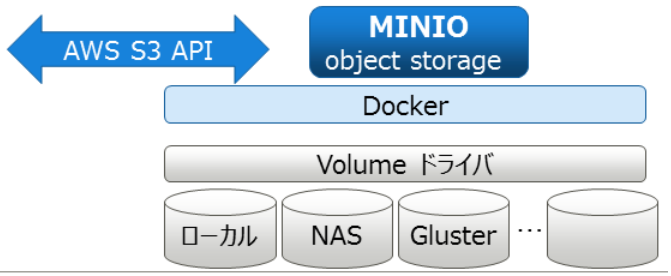
FUJITSU

- **OpenIO**
  - https://www.openio.io/product/product-overview
  - スケールアウト型オブジェクトストレージ、オープンソース。x86とARMをサポート。
  - 一部がGPLv3、その他はAffero General Public License v3。
  - Conscience - advanced write load-balancing
    = Service Discovery+ Load Balancing



- **minio**
  - https://github.com/minio/minio
  - Cloud Storage Server for Micro Services
  - Amazon S3 APIのオープンソース実装。X86とARMをサポート
  - 様々な領域を簡単にオブジェクトストレージ化でき、簡単起動。Docker起動がおすすめ
  - 実装はgo, java, Pythonなど。



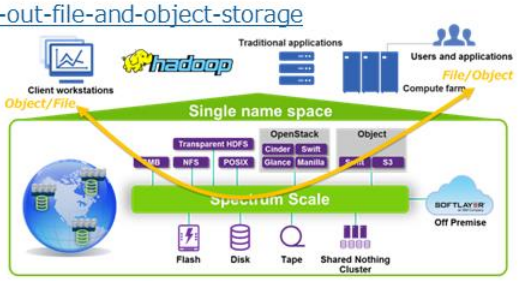


Copyright 2017 FUJITSU LIMITED.

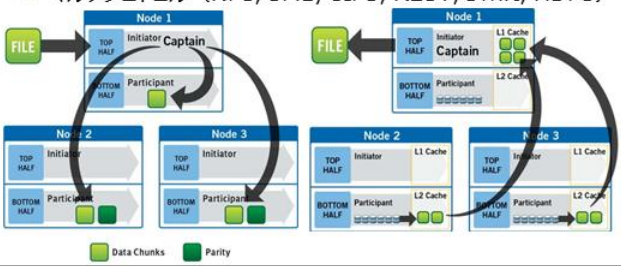
ファイルシステム(Spectrum Scale, OneFS)

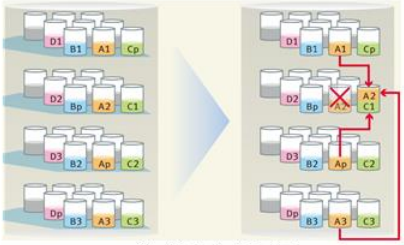
FUJITSU

- **IBM Spectrum Scale**
  - https://www.ibm.com/jp-ja/marketplace/scale-out-file-and-object-storage
  - 大規模分散共有ファイルシステム
  - Posix準拠でNFS, SMB, Swiftにも対応。
  - AIX, Linux, Windows Serverに対応
  - 最大256世代のsnapshotサポート
  - AFM(Active File Management)
    キャッシング機能を利用して遠隔アクセスを高速化



- **Dell EMC OneFS**
  - https://www.dellemc.com/ja-jp/storage/isilon/onefs-operating-system.htm
  - RAID不要なファイル保護、最大４ノード同時障害に対応
  - RAIDは領域単位なのに対し、バーチャルホットスペアで分割したファイルデータ単位にリビルド
  - オートバランスによる利用効率とパフォーマンスの最適化。手動でのデータ再配置が不要。
  - マルチプロトコル（NFS, SMB/CIFS, REST, Swift, HDFS）





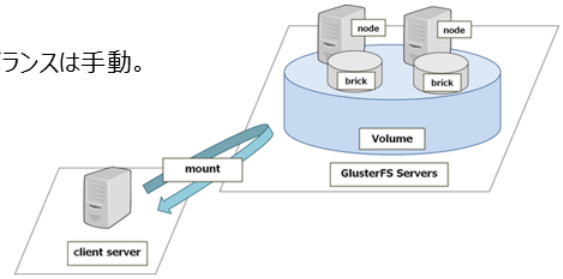
バーチャルホットスペア

Copyright 2017 FUJITSU LIMITED.

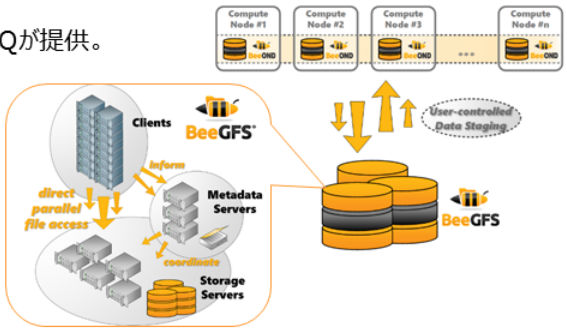
ファイルシステム（GlusterFS, BeeGFS）

FUJITSU

- **Gluster FS**
  - https://www.gluster.org/
  - オープンソース、FUSE実装でユーザ空間で動作。
  - 格納域であるbrickの動的追加・削除が可能。リバランスは手動。
  - NFS, REST APIに対応
  - ボリューム構成は以下に対応。組合せが可能。
    - ・Replicated Volume（レプリケーション）
    - ・Striped Volume（ストライプ）
    - ・Distirbuted Volume（分散配置）
    - ・Dispersed Volume（イレイジャーコーディング）



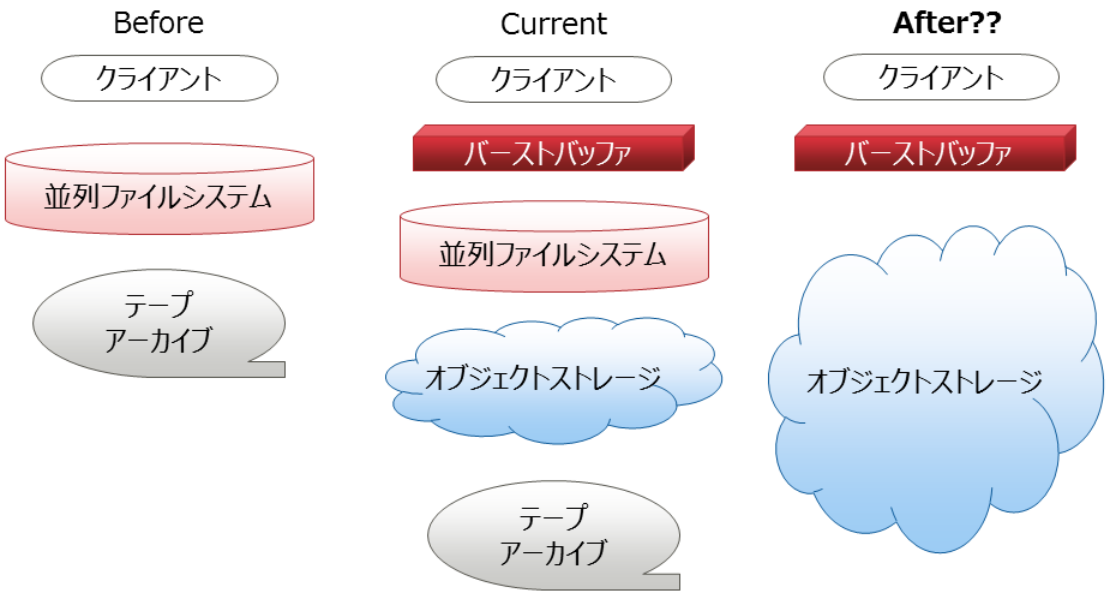
- **BeeGFS**
  - https://www.beegfs.io/content/
  - 2016年にオープンソース化、サポートはThinkParQが提供。
  - 構造はLustreに似ている。
    Management Service, Metadata Service, Storage Service, Clientで構成
  - BeeONDは計算ノード上のSSDを利用して、ジョブ実行に合せて動的にBeeGFSを生成可能



Copyright 2017 FUJITSU LIMITED.

ストレージスタックの変遷

FUJITSU



https://www.flashmemorysummit.com/English/Collaterals/Proceedings/2017/20170807_SNIA_T4_Grider.pdf

Copyright 2017 FUJITSU LIMITED.

データ管理技術

FUJITSU

- **NoSQL（格納、検索）**
  - RDBMS以外の大量データ、多様なデータを扱うデータベース。
  - 一貫性などを犠牲にする代わりに、柔軟性と拡張性を確保
  - 分散型でスケールアウトを特長とすることが多い。

データモデル	特長	ソフト
キー・バリュー	応答が速い	Amazon DynamoDB, Cassandra
カラム	大量集計、大量更新	Apache HBASE, Google BigTable
グラフ	関係性	Neo4j, TITAN
ドキュメント	スキーマレス	mongoDB, CouchDB

- **ブロックチェーン（証明、改ざん防止）**
  - 分散型台帳技術、ビットコインで有名。
    複数ノードで同一の記録を同期させて一つの台帳を維持。
  - 一定期間の取引をブロックとして台帳につなぐことで、全て記録。
  - ブロック作成は、取引データと直前のブロックのハッシュ値を持たせ、改ざんを実質不可能に。
  - P2Pによるメンバ間での直接取引が可能。
  - http://www.fujitsu.com/jp/services/infrastructure/network/blockchain/

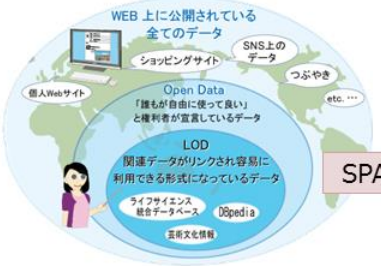
Copyright 2017 FUJITSU LIMITED.

データ管理技術

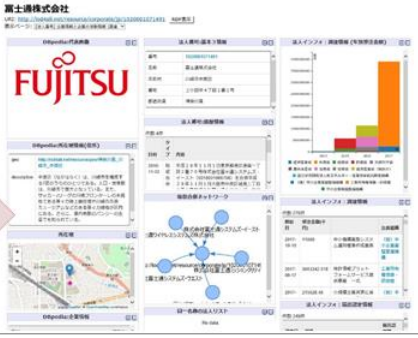
FUJITSU

- **ハッシュタグ（関連付け、検索）**
  - ＃付のキーワード。Twitterで使われ始めて10年。
    Twitter, Instagram, youtubeなどのSNSで幅広く使用。
  - 任意の文字列でタグ付けでき、情報のカテゴライズ、検索が可能。
  - 特徴がなく、役割を果たさない一定のハッシュタグは使用不可。

- **LOD:Linked Open Data（関連付け、検索）**
  - http://www.fujitsu.com/jp/group/labs/resources/tech/techguide/list/lod/
  - ウェブ上でデータを公開して共有するための方法。
  - 保有データを「Resource Description Framework（RDF）」形式に変換、LODにより結び付け。
  - クエリ言語SPARQL Protocol and RDF Query Language（SPARQL）にて検索
  - http://pr.fujitsu.com/jp/news/2017/06/15-1.html



SPARQLで抽出



Copyright 2017 FUJITSU LIMITED.

B. 2.　データ配布時の高負荷対策

FUJITSU
shaping tomorrow with you

SS研
クラウド時代のデータ共有WG
第5回会合
データ配布時の高負荷対策

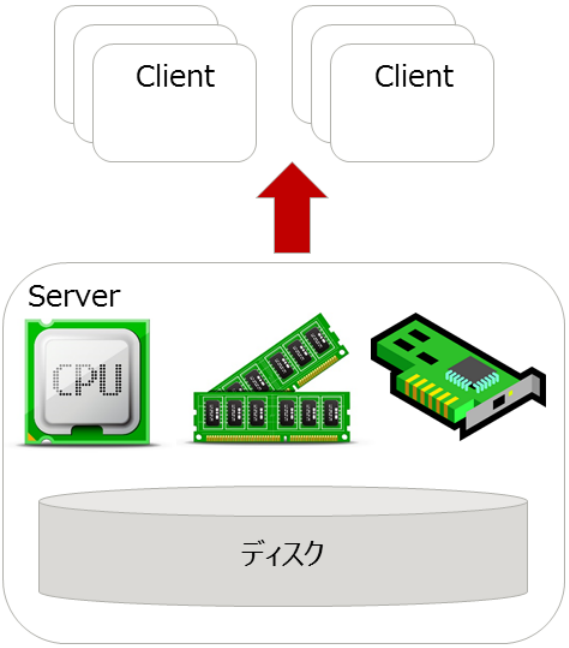
2017年12月22日
富士通株式会社
次世代テクニカルコンピューティング開発本部
ソフトウェア開発統括部

Copyright 2017 FUJITSU LIMITED

データ配布時のサーバ高負荷

FUJITSU

■ データ配布は大量のRead



■ Read処理時のボトルネック

■ ディスク帯域
- Server側のキャッシュで対応。

大量readでキャッシュのヒット率低下

■ ネットワーク帯域
- NICを増設して帯域を増やす。

CPUバスネックで効果が出ない場合も

■ CPU
- コア増やす or 高クロック化

ソフトの排他起因の場合は要仕様変更

サーバ1台の処理限界を
超えるとノードがダウンする

Copyright 2017 FUJITSU LIMITED

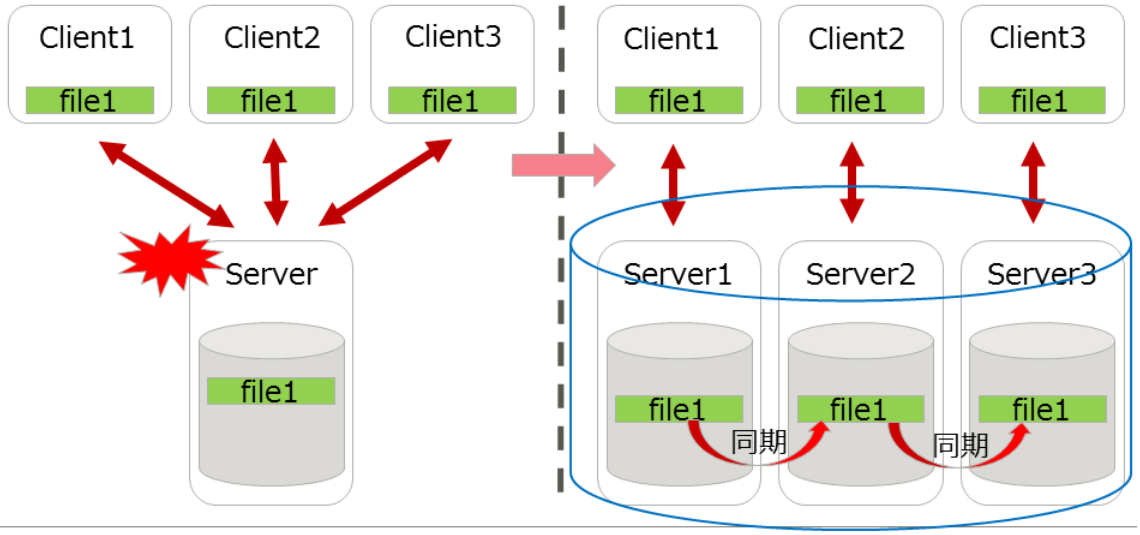
データRead時の高負荷対策（2）

FUJITSU

■ 複数サーバへ**データ複製**

■ 同1データへのアクセス集中してもハード処理限界。

→ 複数サーバに同1データを複製し、クライアントのアクセスを分散。



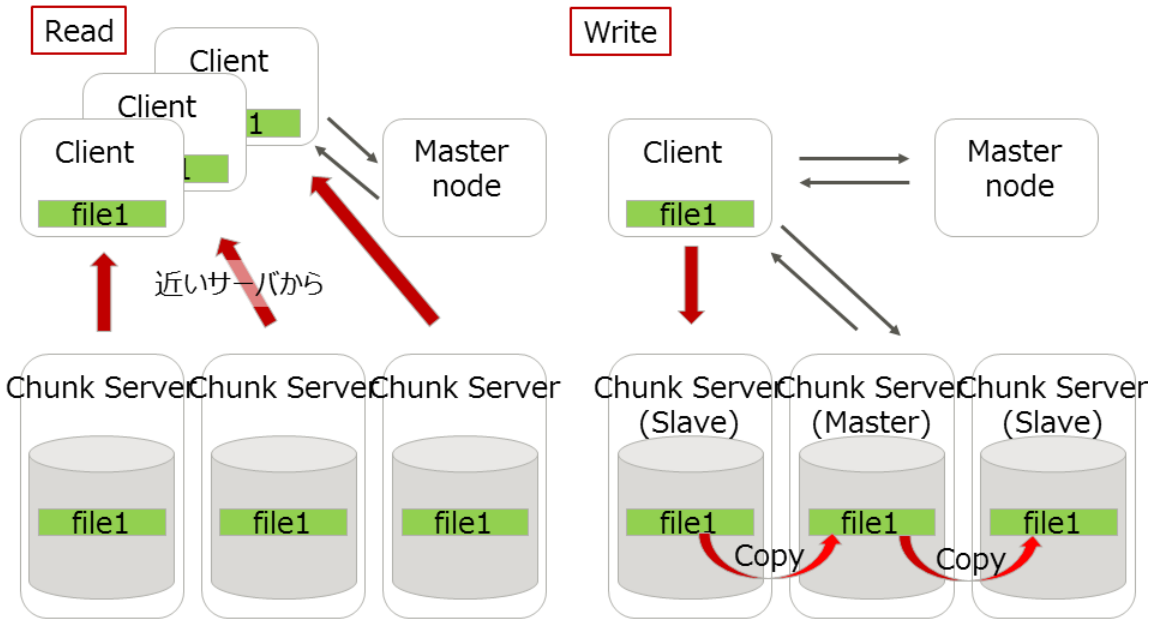
Copyright 2017 FUJITSU LIMITED

ファイルシステム事例（１）

FUJITSU

■ Google File System

マスターノードがファイル格納情報を管理



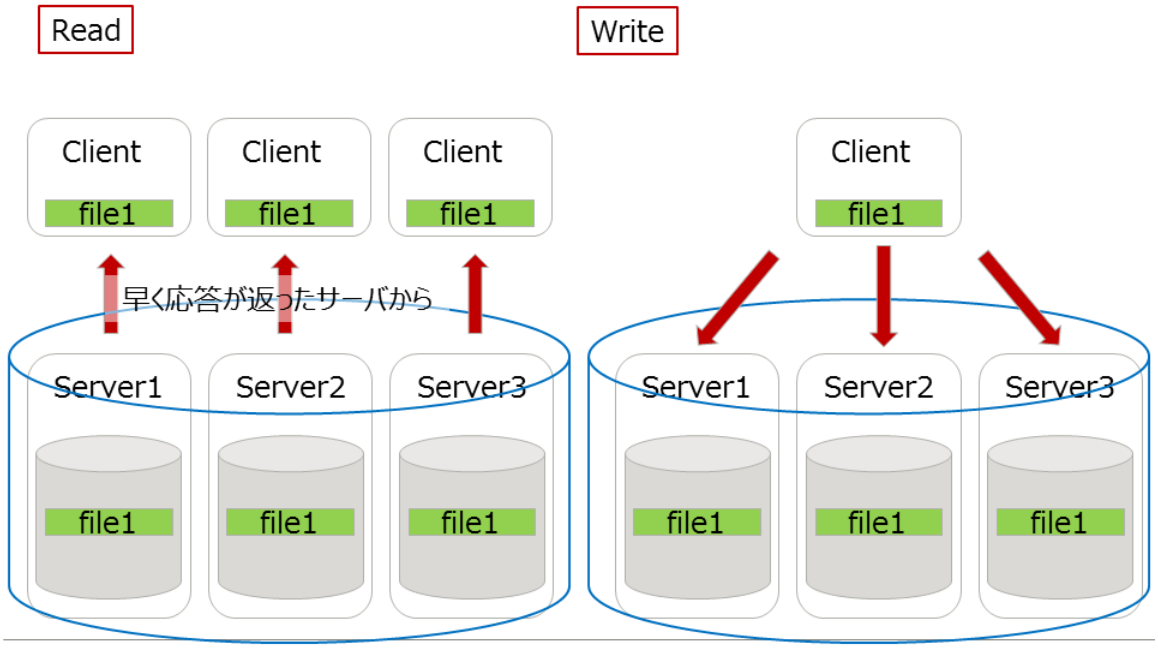
Copyright 2017 FUJITSU LIMITED

ファイルシステム事例（2）

FUJITSU

■ GlusterFS (Replicated Volume)

DHTでファイル格納情報を管理



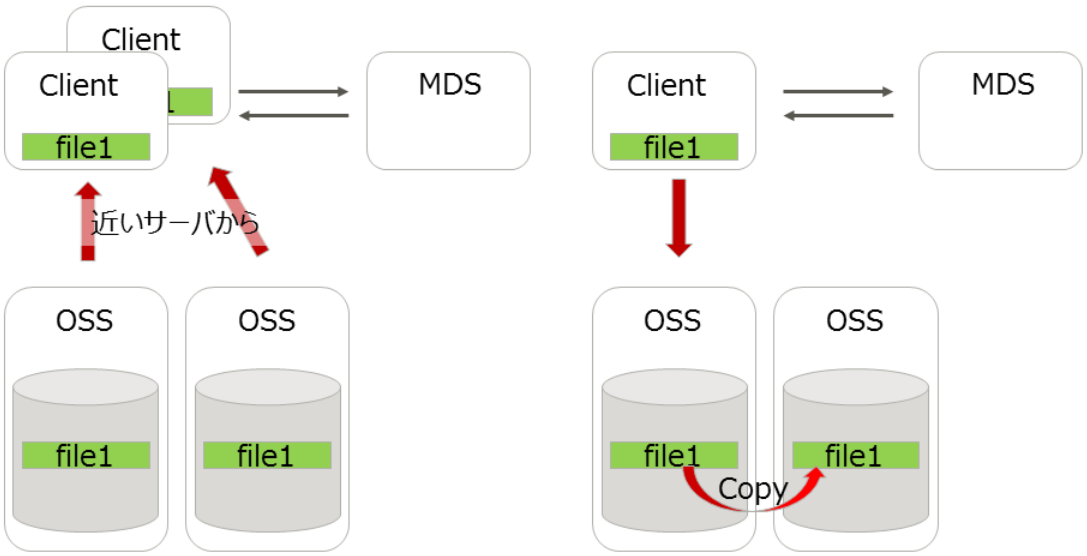
Copyright 2017 FUJITSU LIMITED

ファイルシステム事例（3）

FUJITSU

■ Lustre (File Level Replication：開発中)

MDSでファイル格納情報を管理



Copyright 2017 FUJITSU LIMITED

データ複製まとめ

FUJITSU

■ メリット
- ■ 特定のディスク装置に負荷が集中した場合に、負荷を分散できる。
- ■ ディスク故障時でもデータ消失を防ぐことが可能。

■ デメリット
- ■ Writeのコストが高い（一般的に性能低下）。
- ■ 物理ハードコスト（サーバ、ディスク装置）が増加。

扱うデータの性質によって使い分けが必要

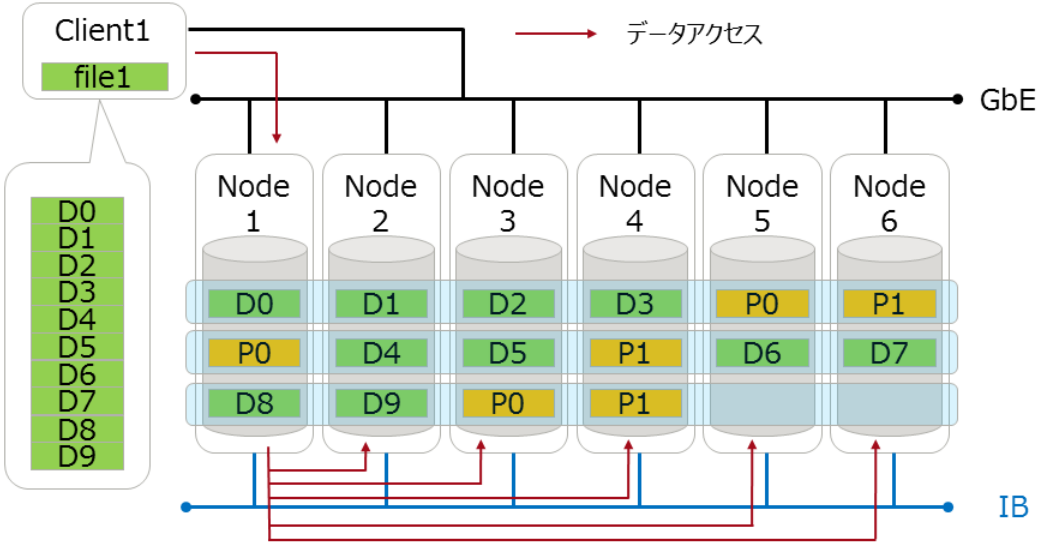
- ■ ほぼReadのみのデータ/観測データ　→ データ複製あり
- ■ Write中心のデータ/一時データ　→ データ複製なし

Copyright 2017 FUJITSU LIMITED

Erasure Cordingによるデータ保全（ご参考）

FUJITSU

■ EMC Isilon （OneFS）の実装例
- ■ パリティ付きでデータを複数ノード・ディスクに分散配置
  - ●ノード（ディスク）故障時でもデータの読出しが可能
  - ●データ複製と異なり、Write時のコストが少ない。



Copyright 2017 FUJITSU LIMITED

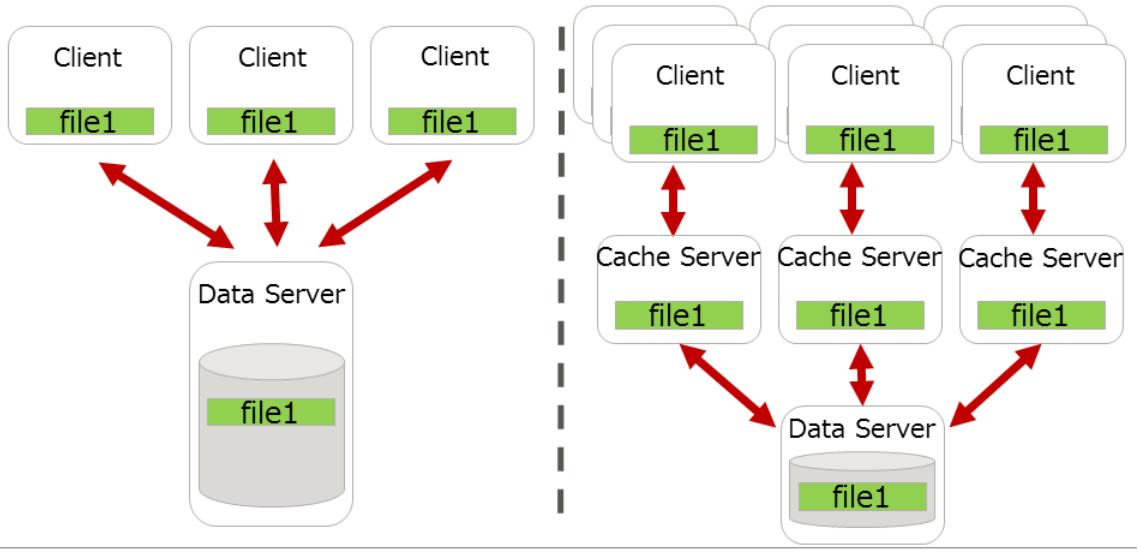
データRead時の高負荷対策（3）

FUJITSU

■ 複数サーバへ**データ複製**

■ 同1データへのアクセス集中してもハード処理限界。

→ クライアントとサーバの中間に中間ノードを配置し、アクセスを分散。



Copyright 2017 FUJITSU LIMITED

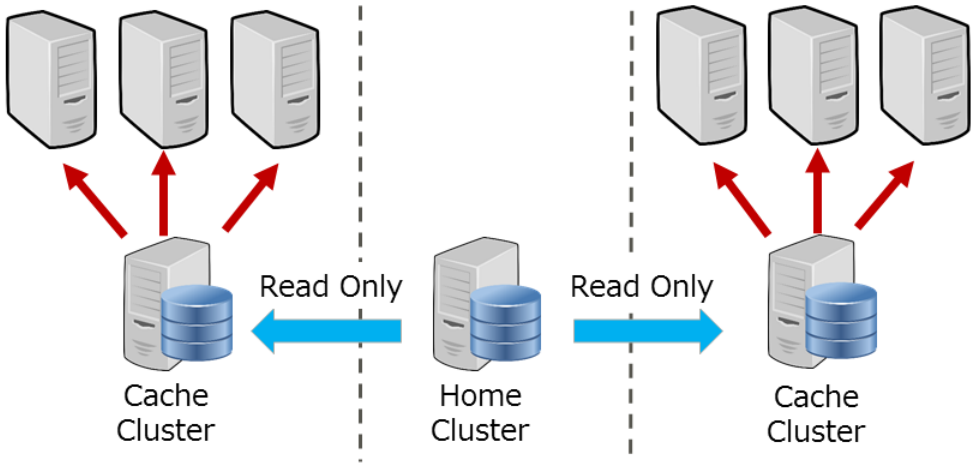
ファイルシステム事例（4）

FUJITSU

■ IBM AFM(Active File Management, Cache mode)

■ ストレージクラスタ間でデータを共有し、高速なファイルアクセスを実現。

■ 共有モードは他にも、Single Write、Local Updatesなど、Cache Cluster側での更新可能なモードも存在する。



Copyright 2017 FUJITSU LIMITED

B. 3.　クラウドストレージ紹介

FUJITSU
shaping tomorrow with you

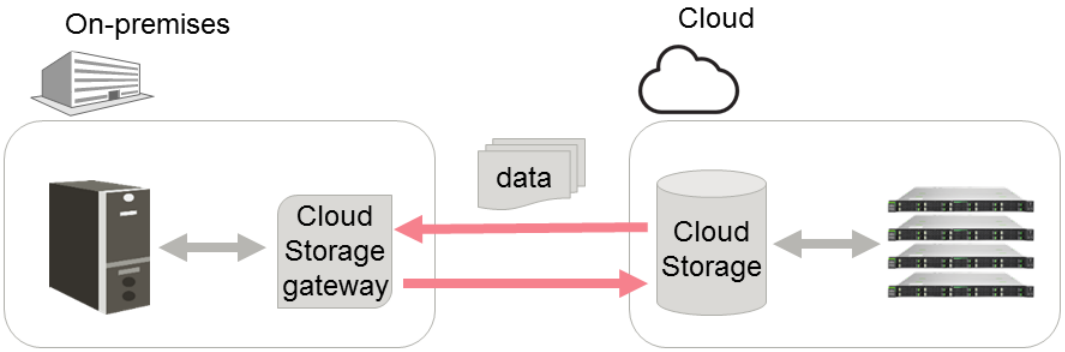
SS研
クラウド時代のデータ共有技術WG
第6回会合

富士通株式会社

Cloud Storage Gateway

FUJITSU

- オンプレミスとクラウドストレージ間でデータを送受信するサービスの総称。
  - タイプ①：クラウドストレージへデータをバックアップ。
    - 各拠点のデータをクラウド上に自動転送。有事の際にリストア。
  - タイプ②：クラウドストレージ上でデータを共有。
    - 業務データ自体をクラウドストレージ上で管理。クラウド上のサービスとも連携。



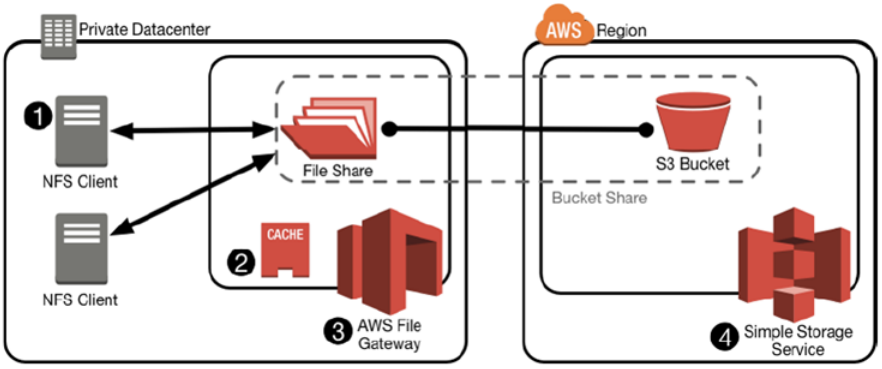
1

AWS Storage Gateway

FUJITSU

■ 特徴

- ■ クラウドストレージをNAS（NFS）、SAN（iSCSI）として利用可能。
- ■ Gateway上のストレージにキャッシュすることで高速アクセス可能。
- ■ キャッシュデータは非同期にクラウドストレージ上へ転送される。
- ■ **物理テープライブラリ装置の代替として仮想テープライブラリとして利用可能。**
  - 耐久性 99.999999999%のストレージに保管（料金:0.004USD/GB/Month)



参照：https://d1.awsstatic.com/whitepapers/Storage/aws-storage-gateway-file-gateway-for-hybrid-architectures.pdf

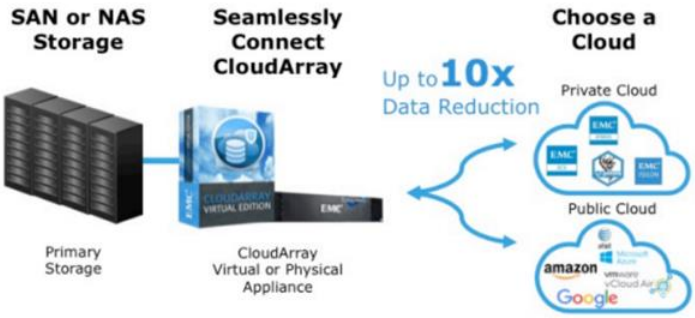
2

EMC CloudArray

FUJITSU

■ 特徴

- ■ クラウドストレージをNAS（CIFS/NFS）、SAN（iSCSI）として利用可能。
- ■ Gateway上のストレージにキャッシュすることで高速アクセス可能。
- ■ キャッシュデータは非同期にクラウドストレージ上に転送される。
- ■ **ECS（Elastic Cloud Storage）との組み合わせで遠地レプリケーションが可能。**



参照：https://www.emc.com/collateral/data-sheet/h13457-cloudarray-ds.pdf

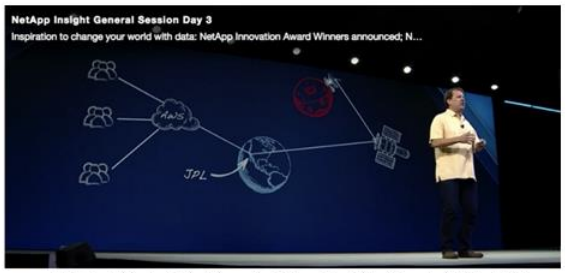
3

NetApp Cloud Sync

FUJITSU

■ 特徴

- ■ オンプレミスにあるサーバ上のデータをクラウドストレージへ転送。
- ■ 複数拠点間でデータ連携可能。
- ■ **SaaSとして提供。**
- ■ データバックアップならAltaVaultなど、用途毎に製品が分かれている。



NASAが火星探査で収集した画像データを研究所の
サーバからAWSへ転送し、解析サービスを利用。

参照：https://www.netapp.com/jp/communities/tech-ontap/jp-tot-201801-netapp-cloudsync.aspx

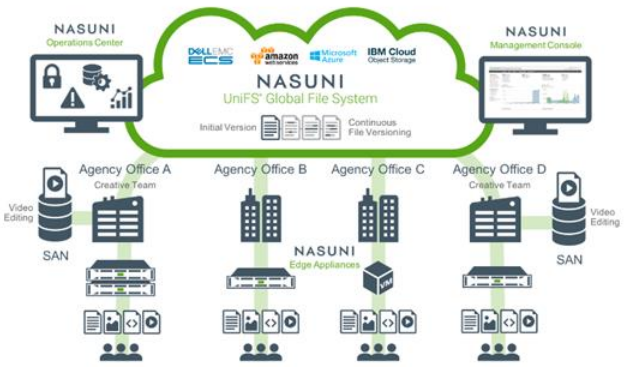
4

Nasuni UniFS + Edge Appliance

FUJITSU

■ 特徴

- ■ クラウドストレージをNAS（CIFS/NFS）、SAN（iSCSI）として利用可能。
- ■ **サーバ上のデータはWORM+差分で管理。**
- ■ **独自のロック機構で複数拠点間でデータ同時編集が可能。**
- ■ 中小規模向け（500人規模）



参照：https://www.nasuni.com/solutions/industries/media-advertising/

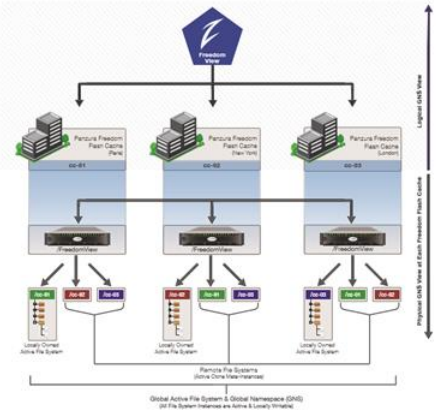
5

Panzura Cloud Storage Gateway

FUJITSU

■ 特徴

- ■ クラウドストレージをNAS（CIFS/NFS）として利用可能。
- ■ Gateway上のストレージにキャッシュすることで高速アクセス可能。
- ■ **独自のロック機構で複数拠点間でデータ同時編集が可能。**
- ■ 大規模向け（5000人超）



参照：http://go.panzura.com/rs/346-POI-893/images/Whitepaper-Global-File-Locking.pdf

各社サービス比較

FUJITSU

	AWS Storage Gateway	EMC CloudArray	NetApp Cloud Sync	Nasuni UniFS+App	Panzura Cloud Storage Gateway	Fujitsu Cloud Storage Gateway
複数拠点同時編集	No	No	-	Yes	Yes	No
データ重複排除/圧縮	No	Yes	（※）	Yes	Yes	Yes
データ暗号化（AES256）	Yes	Yes	（※）	Yes	Yes	Yes
通信経路暗号化	SSL	SSL	VPN/VPC	SSL	SSL	SSL
NFS support	Yes	Yes	Yes	Yes	Yes	Yes
SMB/CIFS support	No	Yes	Yes	Yes	Yes	Yes
iSCSI support	Yes	Yes	No	Yes	No	No
対応クラウド	Amazon S3	Amazon S3 等	Amazon S3	Amazon S3 他	Amazon S3 他	Amazon S3 他
Edge local cache	Yes	Yes	No	Yes	Yes	Yes

（※）NetApp AltaVaultで対応

B. 4.　クラウド時代のデータ共有技術

FUJITSU
shaping tomorrow with you

SS研
クラウド時代のデータ共有WG
第6回会合

2018年7月17日
富士通株式会社
次世代テクニカルコンピューティング開発本部
ソフトウェア開発統括部

Copyright 2018 FUJITSU LIMITED

クラウド時代のデータ共有技術

FUJITSU

■クラウドで提供されているサービス
～オンプレミスのサービスと比較した記述

- 大規模ストレージ系
- ドキュメント共有　Dropbox / box / OneDrive etc.
- 資料共有　（会議録、実験記録～ログブック）
- Evernote、googledocs、Office365
- Source code (development)
- Mail service (gmail)
- Database（データストレージ）

Copyright 2018 FUJITSU LIMITED

クラウド時代のデータ共有技術

FUJITSU

- クラウドとオンプレのサービス特徴比較
  - クラウドでは初期導入が安くて簡単。ただし、使っただけ費用がかかる。
  - オンプレではやりたいことができる。ただし、初期費用、維持管理の手間がかかる。

	オンプレ	クラウド
初期導入費用	高コスト	無料 or 低価格
運用費	保守費、電気代	使用分に課金
保守・資産管理	必要 バックアップやメンテナンス	不要
ライセンス	購入・管理が必要	込みの場合、不要 利用者が用意する場合も
構築期間	長期間(数か月) 調達含めると数年	アカウント登録後、即利用可
構成自由度	自由	制限付き
セキュリティ	自由に設定	用意されたプランから選択
性能	ハード含めたチューニング可	チューニングは限定的 買い足して、性能アップ
可用性・信頼性	より高いレベルで実現可能	用意された範囲で可能
前提知識	幅広く必要	利用したい機能に限定

Copyright 2018 FUJITSU LIMITED

ストレージ

FUJITSU

- クラウド
  - Amazon, Google, Azureは機能、価格ともにほぼ横並び。
  - 従来のPOSIXアクセス向けにはFUSEを使用
  - WasabiがS3準拠、低価格＆高速、ダウンロード固定料金を謳い登場。

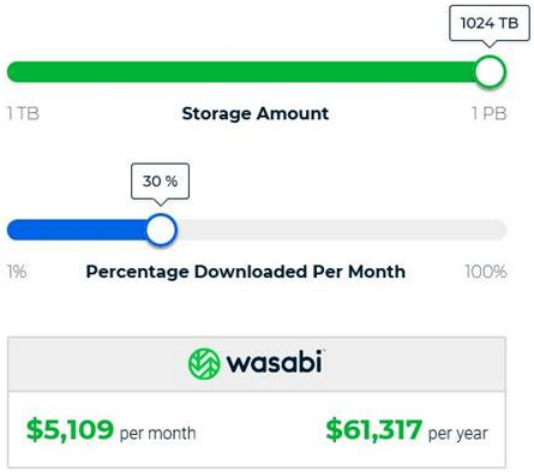
	Amazon S3	Google Cloud Storage	Microsoft Azure Storage	Wasabi S3-compatible Cloud Storage
価格	¥2.8/GB	¥2.58/GB	¥2.24/GB	¥0.55/GB
容量	無制限	無制限	無制限	
ファイル数	無制限	無制限	無制限	
ファイルサイズ	5TB	5TB	4.77TB	
レプリケーション	CRR	Regional Multi-Regional	ZRS GRS, RA-GRS	
バージョン管理	有り	有り	なし	
Web公開	IP制限、期限付が可能	IP制限、期限付きが可能	IP制限が可能	
マウント	FUSE(goofys)	FUSE(gcsfuse)	FUSE(blobfuse)	

Copyright 2018 FUJITSU LIMITED

クラウドストレージ

FUJITSU

- Wasabi
  - https://wasabi.com/s3-compatible-cloud-storage/
  - $.0049 per GB/month ($4.99 per TB/month)の低価格
  - Amazon S3準拠で５倍高速。465MB/5sec=92MB/s
  - ダウンロード、API呼び出しが無課金



Copyright 2018 FUJITSU LIMITED

ストレージサービス

FUJITSU

- クラウド
  - 企業利用が増え、セキュリティや共有の機能が充実
  - GoogleDriveはG Suite、OneDriveはOfficeとの親和性が高い
  - Boxは企業利用が増えている。強固なセキュリティなどが特徴。

	Google Drive	Microsoft OneDrive	Box	Dropbox (Business)	SugarSync
価格	¥1200/月	¥1360/月 office付	¥1800/月	¥2000/月	¥6160/月
容量制限	無制限	10GB	無制限	無制限	無制限
無料容量	15GB	5GB	10GB	10GB	–
オンライン編集	○	○	–	–	–
バージョン管理	○	○	○	○	○
セキュリティ管理	○	○	○	○	○
IRM Information Rights Management	○	○ Azure RMS	○	–	–
KMS Key Management Service	△ Clound KMS 一部国のみ	–	○ AWS KMS	–	–

Copyright 2018 FUJITSU LIMITED

ストレージ

FUJITSU

- オンプレ
  - オンラインストレージ(Owncloud, NextCloud)
  - オブジェクトストレージ（Ceph, OpenIO, Gluster）
  - NAS(NFS、FreeNAS)
  - ビッグデータ（Hadoop, Spark）
  - 並列分散FS（Lustre, BeeGFS, Spectrum Scale）
- オンラインストレージ
  - OwnCloud
    - https://owncloud.jp/
    - Dropboxライクな使い勝手、端末に依存しないファイルアクセスを提供。
      サーバ機能はLinuxOS、Webサーバ(Apache)、DB、PHPからなる。
  - NextCloud
    - https://nextcloud.com/
    - OwnCloud Enterprise相当をオープンソースとして提供。
    - 2016年、Owncloud創業者Frankさんが独立して設立。
      Owncloud Enterpriseの有償化が発端。
    - 最新版数は13.0.4(2018/6)。

Copyright 2018 FUJITSU LIMITED

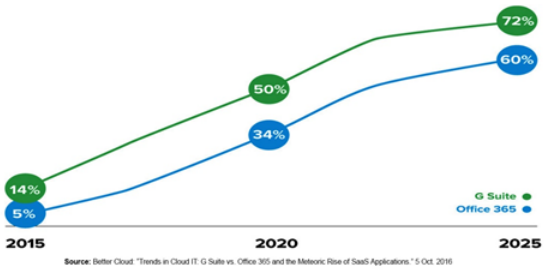
オフィス系

FUJITSU

- クラウド
  - Office365 と G suite の２択。違いは価格とMS Officeとの親和性。
  - Office365はOfficeライセンスをサブスクリプションで提供。

	プラン	月額／人	容量／人	備考
G Suite	Basic	¥600	30GB	
	Buisiness	¥1200	無制限	
	Enterprise	¥3000	無制限	DLP（Gmail）
Office 365	E1	¥870	1TB	Word, Excelは無し
	E3	¥2180	無制限	
	E5	¥3810	無制限	DLP（Exchange）

機能	G Suite	Office 365
メール	Gmail	Exchange Online
カレンダー	Google Calender	Outlook
ストレージ	Google Drive	OneDrive
メモ	Google Keep	OneNote
会議	Google ハングアウト	Skype for Buisiness
チャット・SNS	Google プラス	Skype for Buisiness
共有	Google サイト	Sharepoint Online



Copyright 2018 FUJITSU LIMITED

メール系

FUJITSU

- クラウド
  - GmailはAndroidユーザに、Outlook.comはWindowsユーザに向いている。
  - メール分類がGmailはラベルでタグ付け、Outlookは従来のフォルダ形式。

	Gmail	Outlook.com
無料容量	15GB（拡張は有料）	無制限
1通当たりの容量	25MB	25MB
分類	ラベル	フォルダ
プロトコル	POP/IMAP	POP/IMAP
アドレスのエイリアス	可能	可能

  - Amazon Simple Email Service
    - Amazonが提供するメール送信サービス。
    - バウンスはリージョン内の他ユーザとも共有。⇒携帯キャリアメールへのバウンス対応が面倒。
  - SendGrid
    - クラウドベースのSMTPプロバイダ
    - バウンスはユーザ単位の管理。
- オンプレ
  - Microsoft Exchange

Copyright 2018 FUJITSU LIMITED

ソース管理系

FUJITSU

- かつての Subversion(CVSなど)から 分散型である Git に人気がシフト
- クラウド
  - GitHub
    - https://github.com/
    - ソースコードのホスティングサービス。今年 6 月にMicrosoftが買収。
    - AWS, Google, MS Azure, IBMの各クラウド上にホスティング可能（Enterprise）
    - リポジトリ作成は無制限。プライベートの場合、Teamは$9/人月、Businessは $21/人月。
  - Gitlab
    - https://about.gitlab.com/
    - GitHub買収後にインポート数が10倍に激増。
    - プライベートリポジトリに関しても無償。
  - BitBucket
    - https://bitbucket.org/
    - 小規模向け5ユーザまでは無料
    - 中規模向けは$2/人月、大規模向けは$5/人月。

	GitHub	Gitlab	BitBucket
リポジトリ数	無限	無限	無限
サイズ	無限	1GB	2GB
人数	無限（公開）	無限	5人

Copyright 2018 FUJITSU LIMITED

メモ系, ELN系

FUJITSU

- メモ
  - 使い勝手はそれぞれ。既にGoogle, MSユーザなら、合わせて選ぶのがよい。
    Evernoteは2016年から無料ユーザの同時利用クライアントが2台に制限。

	Google Keep	OneNote	Evernote
ファイル添付	×	○	○
お気に入り	×	×	○
Undo/Redo	○	×	×
音声入力	○	×	×
タグ	○：ラベル	○：セクション	○：タグ
ノートブック	×	ノートブック	ノートブック
グループ管理	×	×	スタック

- ELN(電子実験ノート：Electronic Lab Notebook)
  - 医薬品、化学、生物、飲食料品などの業界で利用されている。
  - 用途に応じて、様々なベンダが展開中。
    参考: https://github.com/Labii/comparison-of-best-electronic-lab-notebook/blob/master/README.md

	オンプレ	クラウド	備考
Perkin Elmer	E-Notebook	Signals Notebook	FUJITSU Technical Computing Solution E-Notebook Enterprise
BIOVIA	BIOVIA Notebook	BIOVIA Notebook Cloud	
IDBS	E-WorkBook ELN	E-WorkBook connect	

Copyright 2018 FUJITSU LIMITED

データベース系

FUJITSU

- クラウド
  - Oracleを除き、RDBとNoSQLを品揃え。
  - AWS, Google, AzureはNoSQLで差異化、IBM, OracleはRDBで差異化。
    Fujitsu Cloud は Oracle と提携。

クラウド	RDB	NoSQL ドキュメント	NoSQL KVS	NoSQL グラフ	NoSQL カラム	Data warehouse
AWS	RDS (**Aurora**, MySQL, PostgreSQL,Oracle, MS SQLServer, MariaDB)	**DynamoDB** SimpleDB	ElastiCache (Redis, Memcached)	Neptune		Redshift
Google	SQL (MySQL, PostgreSQL) Spanner	**Datastore**	**Cloud Memorystore** (Redis)		Bigtable	BigQuery
Azure	SQL Database MySQL PostgreSQL	**Cosmos DB**	Azure Redis Cache			SQL Data warehouse
IBM	Db2 Hosted IBM Compose (PostgreSQL MySQL)	IBM Cloudant IBM Compose (MongoDB Elasticsearch RethinkDB)	IBM Cloudant IBM Compose (Redis, etcd)	IBM Graph		Db2 Warehouse
Oracle	**Oracle DB** MySQL	準備中				
Fujitsu	DB powered by Oracle® Cloud	NoSQLサービス (MongoDB)				

Copyright 2018 FUJITSU LIMITED

データ共有技術とその周辺技術　～現状と将来

FUJITSU

- ストレージシステムの動向（HDD/半導体ストレージの動向、容量と単価の見通し）
- 大規模ストレージの動向
  - 並列ファイルシステム（Lustre, FEFS, GPFS/HPSS）
  - オブジェクトストレージ（Ceph）
  - HDFS、Spark
  - 小中規模　NAS
  - クラウドストレージ
    - Dropbox/box/google drive/Onedrive　など
    - ファイルシステムとして利用できるもの
    - コールドストレージ
  - データコンテナ技術：オンプレミスからシームレスにつなぐ
    - コンテナイメージ、アプリイメージなど
    - （CERNVMFSなど）

Copyright 2018 FUJITSU LIMITED

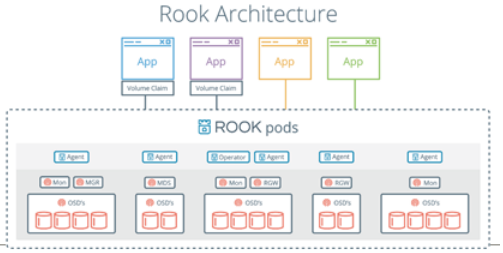
データ共有技術とその周辺技術　～現状と将来

FUJITSU

- データコンテナ技術：オンプレミスからシームレスにつなぐ
  - Kubernates(略記：k8s)
    - https://kubernetes.io/
    - コンテナ化したアプリケーションのデプロイ、スケーリング、および管理を行うための、オープンソースのコンテナオーケストレーションシステム。⇒**「コンテナのクラスタ化」**
    - google主導、CNCF(Cloud Native Computing Faundation)で開発。
    - Parsistent volumeは in-tree で下記表をサポート。

AWSElasticBlockStore	AzureFile	AzureDisk	CephFS
Cinder	FC	FlexVolume	Flocker
GCEPersistentDisk	Glusterfs	iSCSI	Quobyte
NFS	Ceph RBD	VsphereVolume	PortworxVolume
ScaleIO	StorageOS	Local	

    - Container Storage Interface (CSI)をα版としてサポート開始。
      ⇒3rdベンダのストレージサポートがk8s本体と切り離して開発可能に。
  - Rook
    - https://rook.io/
    - Kubernetesにストレージ機能を統合するソフトウェア
    - CNCFの15番目のプロジェクトとして採用
    - まずはCephが対象



Copyright 2018 FUJITSU LIMITED

データ共有技術とその周辺技術　～現状と将来

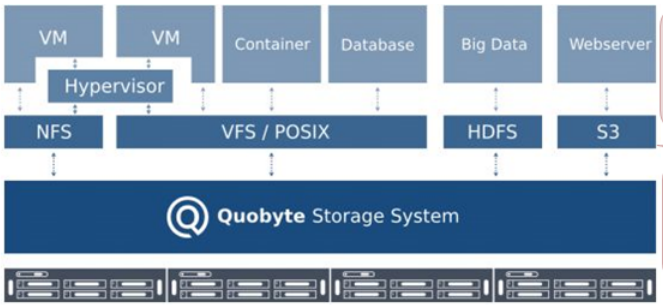
FUJITSU

■ 新興ファイルシステム

■ Quobyte Data Center File System

Quobyte

- https://www.quobyte.com/
- XtreemFSチームが2013年にQuobyteを立ち上げ
- 並列分散、POSIX準拠、マルチアクセス対応のUnified File System。
- CRC32, replication or erasure coding,
- 導入事例：Science and Technology Facilities Council's (STFC) JASMIN phase4
  42PBからスタートし、2022年には300PBの想定。
  - ・3 PB of Parallel File System storage (PFS)
  - ・30 PB of Scale-Out File System storage (SOF)
  - ・5 PB of Object Storage (OS)
  - ・0.5 PB of dedicated high-performance storage for scratch and home directory use.



Server: Min. 4, up to 10,000s
Capacity: Unlimited, max. 32PB per fle
File Systems: Unlimited
Files: Up to 100,000,000 per fle system

Native Client: Linux, Mac, Windows, HDFS
Proxies: NFSv3, NFSv4, SMB
Object: S3

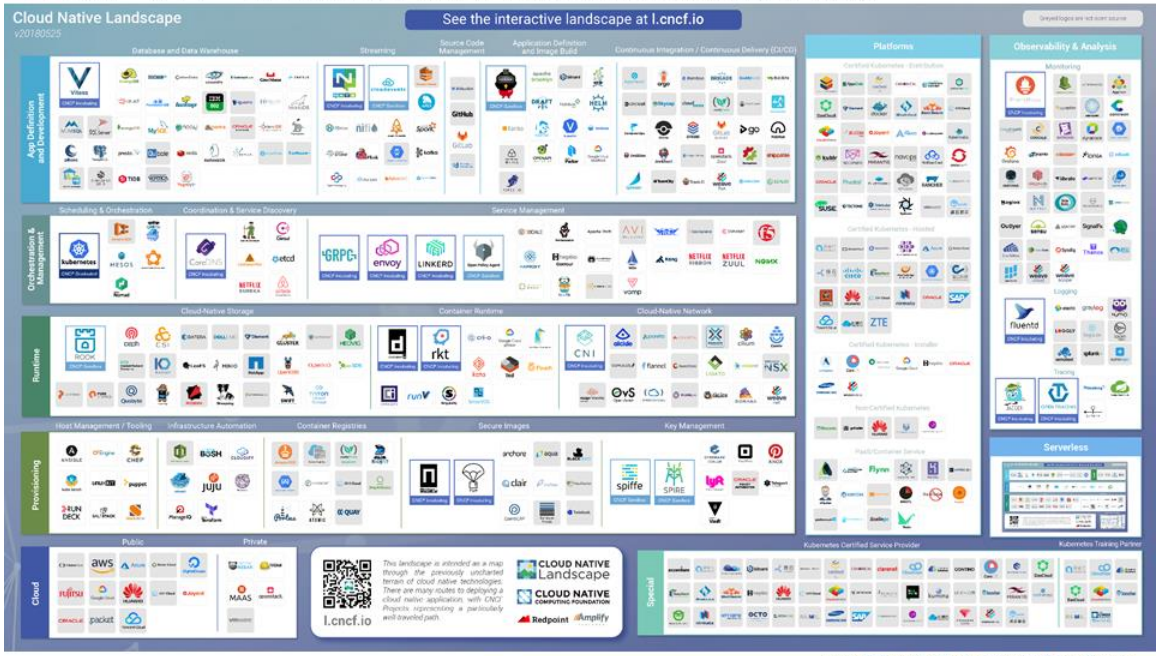
Copyright 2018 FUJITSU LIMITED

データ共有技術とその周辺技術　～現状と将来

FUJITSU

■ Cloud Nativeな技術はCNCF (Cloud Native Computing Foundation) が推進。

■ ストレージ、データ共有技術を含めて多数あり、継続的なウォッチが必要。



引用:https://github.com/cncf/landscape

Copyright 2018 FUJITSU LIMITED

データ共有技術とその周辺技術　～現状と将来

FUJITSU

■ FEFSの今後

■ 高速性の追求

- NVDIMMが出てきてハード性能が向上し、ソフトコストが見えてくる。
  今まではストレージが性能ボトルネックだったが、ソフトコストが課題に。
  ⇒ソフトコストをなるべく減らすため、技術動向調査と検討を実施予定。
  - メタ管理・排他ロックの見直し
    テープ同様にWriteOnceReadManyが増えてきているので、
    ReadOnly化含めた効率の良いファイルロック管理を検討。
  - サーバ・クライアント間のプロトコルコストを見直して、
    NVMe over Fabric や RDMAのアトミック命令の活用を検討。

■ 他ストレージとの連携強化

- 用途ごとに最適なストレージが分かれてきており、単一のファイルシステムだけではなく、
  それぞれのストレージとの連携が重要。
  ⇒以下などの連携を検討したい。
  - 高速化観点の階層ストレージ連携
  - バックアップ用途でのストレージゲートウェイ連携
  - AI観点でのコンテナ連携

Copyright 2018 FUJITSU LIMITED

B. 5.　「認証サービス」ご紹介

FUJITSU
shaping tomorrow with you

FUJITSU Cloud Service for OSS
「認証サービス」ご紹介

2018年6月
富士通株式会社

Version 1.20

目次

FUJITSU

- はじめに　～認証統合の要は「4つのA」～
  - 認証（Authentication）
  - 認可（Authorization）
  - 管理（Administration）
  - 監査と証跡（Audit & Audit Trail）
  - 認証サービスでの機能提供範囲
- 認証サービスとは
  - 認証統合の課題
  - サービス概要
- 認証サービスの特長
  - 利便性の向上（シングルサインオン（SSO））
  - セキュリティリスクの低減
- 機能概要
- 課金の考え方について
- 制限事項・注意事項

Copyright 2016-2018 FUJITSU LIMITED

はじめに　～認証統合の要は「4つのA」～

FUJITSU

認証統合は「4A」を軸にした検討が必要です

■認証(Authentication)
人間・機器・プログラム等のものが**想定通りのもの**であることを、他のものが**確認する**こと、またはその**機能**。

■認可(Authorization)
認証を受けた人間・機器・プログラム等のものに対して、機器・**プログラム・データに対するアクセスを許可する**こと、またはその**機能**。

4A

■管理(Administration)
認証や認可を行うための利用者情報、機器情報等やアクセス制御情報を保有し**利用・変更を可能にする**こと、またはその**機能**。

■監査＆監査証跡(Audit & Audit Trail)
機器・プログラム・データに対して**アクセスした記録や履歴**を保存し、そのアクセスの**正当性をチェック**や監査すること、またはその**機能**。

Copyright 2016-2018 FUJITSU LIMITED

認証（Authentication）

FUJITSU

人や機器の真正性を確認

■ 認証方式の識別特性とメリット・デメリット

識別特性	認証方式の具体例	メリット	デメリット
1．記憶	・文字列パスワード ・PIN(暗証番号)	○導入・運用コストが安価 ○わかりやすい	●紛失・盗難に気づきにくい ●簡単だと推測しやすい ●複雑だと覚えにくい（忘却しやすい）
2．所持品	・ICカード ・証明書 ・トークン ・機器に組み込まれたデータ	○悪用が困難 ○紛失・盗難に気づきやすい ○モノを持っている安心感	●導入・運用コストが高価 ●紛失・盗難しやすい ●壊れる
3．身体的特徴	・静脈 ・指紋 ・虹彩	○紛失・盗難されない ○なりすましが困難	●高価 ●プライバシー ●人によっては利用できない ●誤検知の場合もある ●ケガ等により損なわれる

Copyright 2016-2018 FUJITSU LIMITED

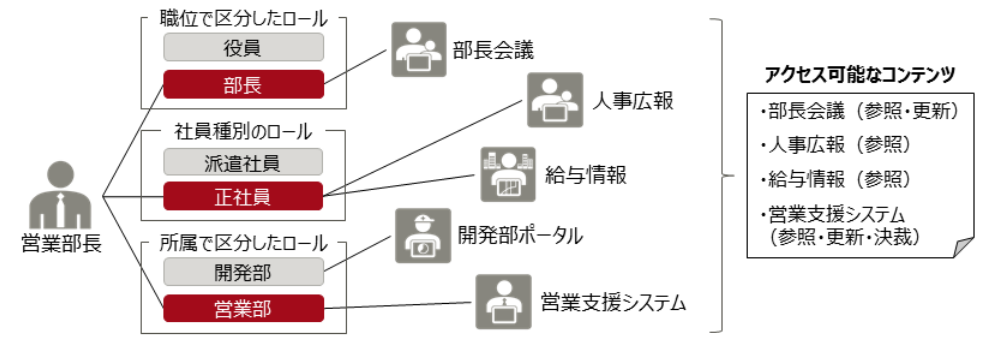
認可（Authorization）

FUJITSU

認証された人や機器にアクセス権限（ロール）を付与

■ ロール
職位、資格、所属する組織や職責を示すものです。
組織内における仕事や役割を表す用語を定義します。
ロールの定義にしたがって必要な認可が割り当てられます。

例　組織、所属を基準にロールを定義



Copyright 2016-2018 FUJITSU LIMITED

管理（Administration）

FUJITSU

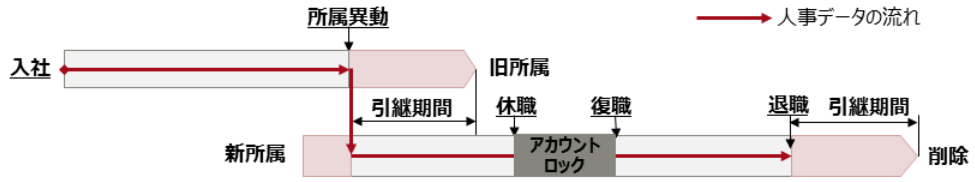
人や機器のライフサイクルを正しく管理

■ ライフサイクル管理
社員の入社、退職、所属異動、休職、復職など、アカウント情報の登録から削除までの管理機能です。

■ 非正規従業員の管理
人事情報に登録されない、現場裁量にて採用される非正規従業員などのアカウント情報管理も適切に運用できる必要があります。

■ 大規模異動処理への対応
国内企業に特有の年度初めの大規模人事異動などにおいて、アカウント情報を迅速に対応させる必要があります。

アカウント情報のライフサイクル管理の一例

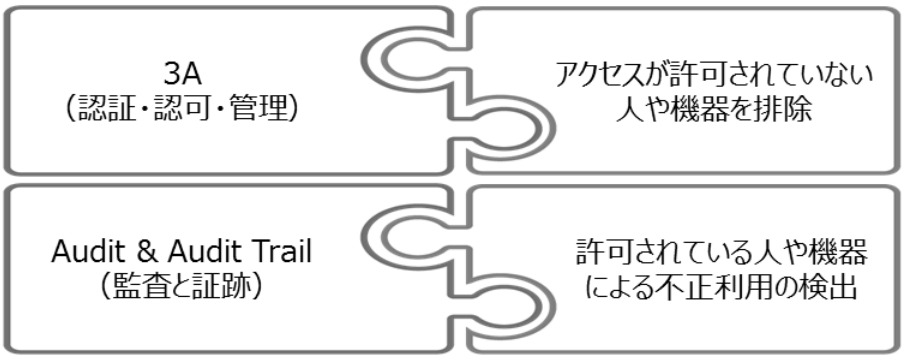


Copyright 2016-2018 FUJITSU LIMITED

監査と証跡（Audit & Audit Trail）

FUJITSU

「監査と証跡」により、3Aがより一層効果的に機能



記録される情報例

- ✓ いつ、だれが、だれの情報を操作したか
- ✓ いつ、だれが、どの端末にログインしたか
- ✓ いつ、だれが、どの端末でログインに失敗したか

Copyright 2016-2018 FUJITSU LIMITED

認証サービスでの機能提供範囲

FUJITSU

認証要素4つのうち、3要素を提供

■認証(Authentication)
- ・シングルサインオン
- ・ワンタイムパスワード
- ・生体認証

■認可(Authorization)

（システムに依存するため業務アプリ側での対応）

4A

■管理(Administration)
- ・ID管理機能連携（API提供）

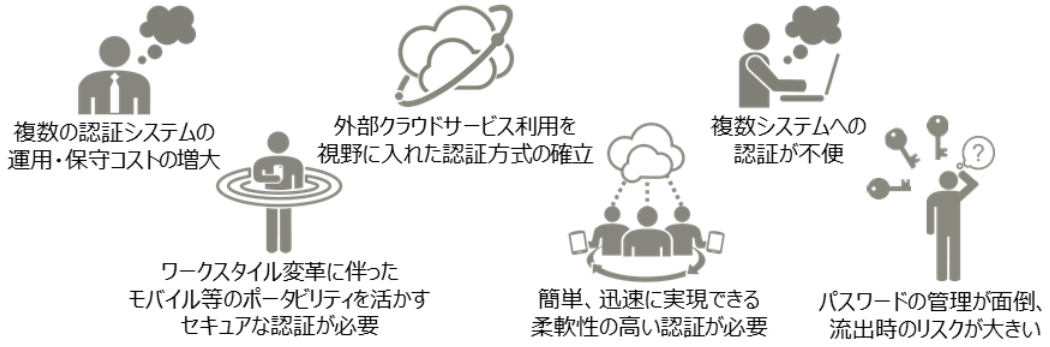
■監査＆監査証跡
(Audit & Audit Trail)
- ・監査ログ取得機能連携
（API提供）

Copyright 2016-2018 FUJITSU LIMITED

認証サービスとは –認証統合の課題-

FUJITSU

認証サービスの導入により、認証統合に関する様々な課題を解決します。



認証サービスで課題解決！

クラウド時代において、利便性を損なわずに本人と証明できる
パブリッククラウドサービスが必要となっている

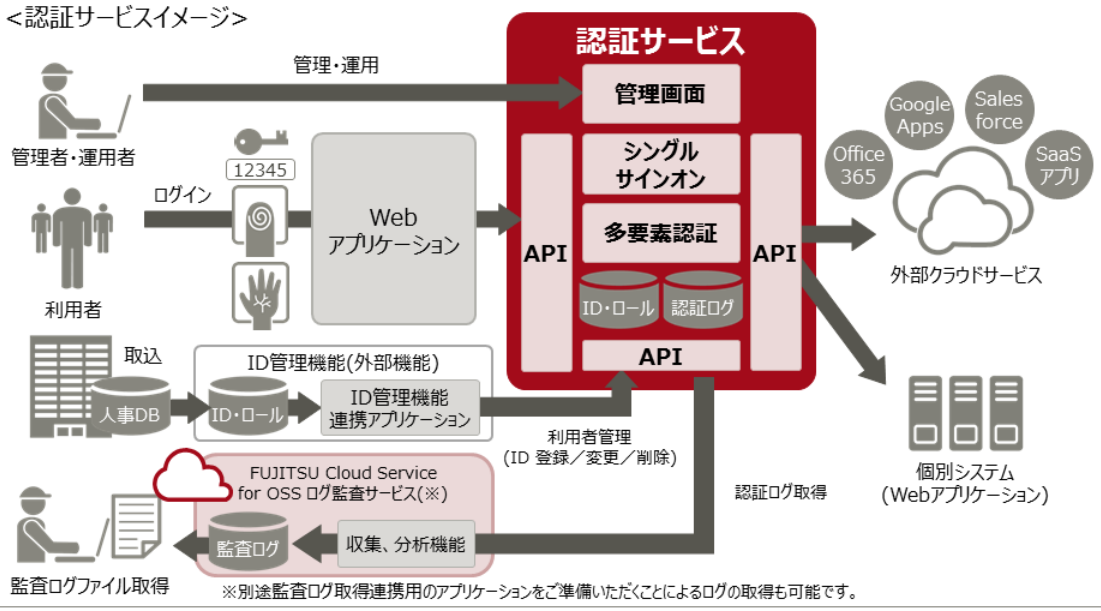
Copyright 2016-2018 FUJITSU LIMITED

認証サービスとは –サービス概要-

FUJITSU

Webアプリケーションに対して多様な認証機能をスピーディーに組み込みできるサービスです。本サービスの利用により、セキュリティ強化と利便性向上を実現します。

<認証サービスイメージ>



Copyright 2016-2018 FUJITSU LIMITED

認証サービスの特長

FUJITSU

利便性の向上（シングルサインオン（SSO））

標準規約に沿った認証方式(SAML,OpenID Connect 1.0)の採用により、他のクラウドサービスと連携、シングルサインオンシステムの構築が可能になります。

セキュリティリスクの低減（多要素認証）

富士通の生体認証（手のひら静脈・指紋）の利用により、安全な認証システムの構築が可能になります。

ID管理連携（APIを提供）

認証用のID情報に関して、ID管理元からのAPIの利用により連携可能になります。

監査と証跡のログ取得（APIを提供）

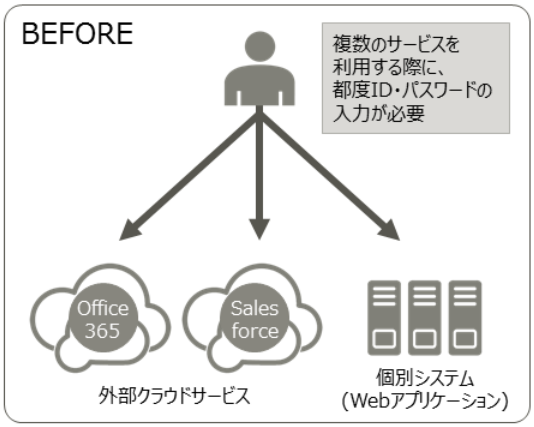
セキュリティ監査に必要となるアクセスログ（証跡ログ）をAPIの利用により取得可能になります。
ログ監査サービスと併せて利用することにより、外部機能を準備することなくアクセスログの取得が可能になります。

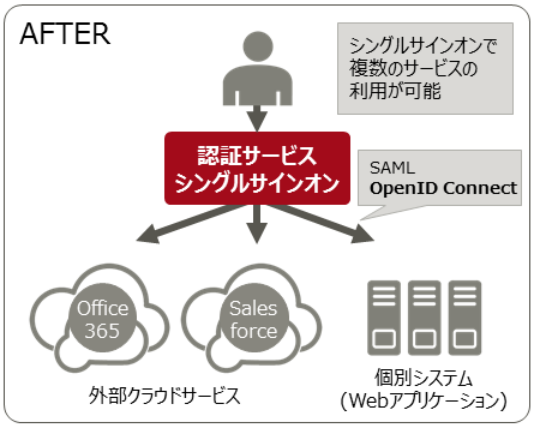
Copyright 2016-2018 FUJITSU LIMITED

利便性の向上（シングルサインオン（SSO））

FUJITSU

- 複数の業務サービス（Webアプリケーションや外部クラウドサービス）を利用する際に、都度ID・パスワードの入力を行うことなく1回の認証でアクセスすることができます。
- 標準規約に沿った認証方式(SAML,OpenID Connect 1.0)の採用により様々なクラウドサービスを一元してシングルサインオンシステムで利用が可能になります。
- ID／パスワードを一元管理することにより利便性が向上します。





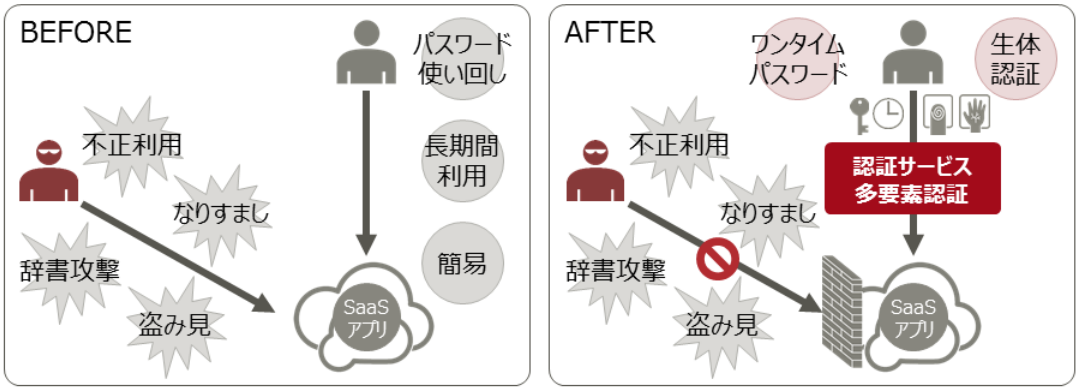
Copyright 2016-2018 FUJITSU LIMITED

セキュリティリスクの低減（多要素認証）

FUJITSU

「ワンタイムパスワード」や「生体認証」による認証を行うことにより、確実な本人特定で、ルール厳格化では防げない「なりすまし」、「不正利用」による情報漏えいのリスクを大幅に軽減することができます。

- ワンタイムパスワード
  1回限りの使い捨てパスワードを用いることにより、高い安全性を確保します。
- 生体認証
  生体情報を利用し、確実に本人であることを証明しアプリケーションにアクセスできます。



Copyright 2016-2018 FUJITSU LIMITED

機能概要

FUJITSU

■ 基本サービス

機 能		概 要
認証機能	シングルサインオン	標準規約に沿った認証方式(SAML,OpenID Connect 1.0)の採用により様々な業務サービスと連携することによりシングルサインオン環境を提供します。
	多要素認証	一般的なID、パスワードだけでなく、手のひら静脈認証、指紋認証、ワンタイムパスワードの多要素認証の提供によるセキュアな認証システムを提供します。

■ 管理画面

機 能	概 要
利用者管理	Webアプリケーションの利用者の管理を行います。運用テナントごとの最大ID数(申請ID数)、登録ID数、残ID数の確認や利用者の検索・追加・変更・削除が可能です。
各種設定	ログイン画面やセッションの設定、ドキュメント・ツールのダウンロード、認証連携を行うアプリケーションの登録を行います。
サービス内容	運用テナントの追加や利用状況の確認を行います。認証方式オプションの設定変更（ワンタイムパスワードオプション／生体認証オプションの設定／解除）が可能です。
認証設定	パスワードポリシーの設定、ADFS設定、IPアドレス制限の設定を行います。
個人設定	利用者のパスワード変更やタイムゾーンの設定を行います。ワンタイムパスワードや生体認証のオプションが設定されている利用者は、パスフレーズの確認や生体認証登録時のパスワードを設定することができます。

Copyright 2016-2018 FUJITSU LIMITED

機能概要

FUJITSU

■ REST API

機 能	概 要
認証	認証／ログアウト機能を提供します。
利用者管理	利用者情報について以下の機能を提供します。 ・利用者の取得 ・利用者の一覧取得 ・利用者の一括追加 ・利用者の変更 ・利用者の一括変更 ・利用者の一括削除 ・利用者の一括ロック解除
パスワードポリシー管理	パスワードポリシーの取得／変更機能を提供します。
IPアドレスによる認証制御	IPアドレスによる認証制御情報の取得／変更機能を提供します。
認証ログ取得	内部統制報告制度（J-SOX）、金融情報システムセンター（FISC）の監査に対応した認証ログを取得するAPIを提供します。 認証ログについて以下の機能を提供します。 ・認証結果ログ取得 ・認証サービスのREST API操作ログ取得

Copyright 2016-2018 FUJITSU LIMITED

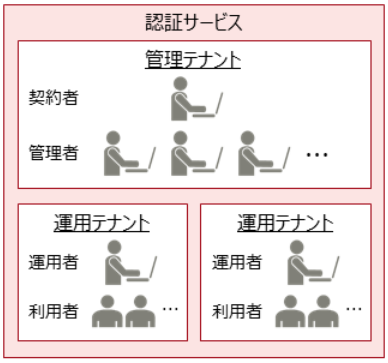
機能概要

FUJITSU

■ テナントの考え方

テナントとは、認証サービスを利用する単位です。本サービスでは、以下のテナント単位にアカウント（ID）の認証の設定を行い、管理・運用します。

名称	概 要
管理テナント	運用テナントの管理を行うためのテナントです。
運用テナント	Webアプリケーションや外部クラウドサービスとの認証連携を実現するためのテナントです。運用テナントごとに最大ID数を設定します。



■ アカウント（ID）

サービスのアカウントは、以下の4つに分類されます。

名称(権限)	概 要
契約者	FUJITSU Cloud Service for OSSと契約を結んでいるお客様です。認証サービスとして管理テナントを配備し、サービスを管理します。運用テナントの作成／運用者の作成、管理者の作成を行います。
管理者	管理テナントに所属し、テナントの管理者権限が設定されている担当者です。運用テナントを管理します。新たに管理者を作成することもできます。
運用者	運用テナントに所属し、テナントの運用者権限が設定されている担当者です。運用テナントに所属する利用者を作成し管理します。新たに運用者を作成することもできます。
利用者	Webアプリケーションや外部クラウドサービスの利用者です。運用テナントに所属します。

アカウント作成権限　○:作成可能／×:作成不可

		作成者自身のアカウント権限			
		契約者	管理者	運用者	利用者
作成対象	契約者	×	×	×	×
	管理者	○	○	×	×
	運用者	○	×	○	×
	利用者	×	×	○	×

Copyright 2016-2018 FUJITSU LIMITED

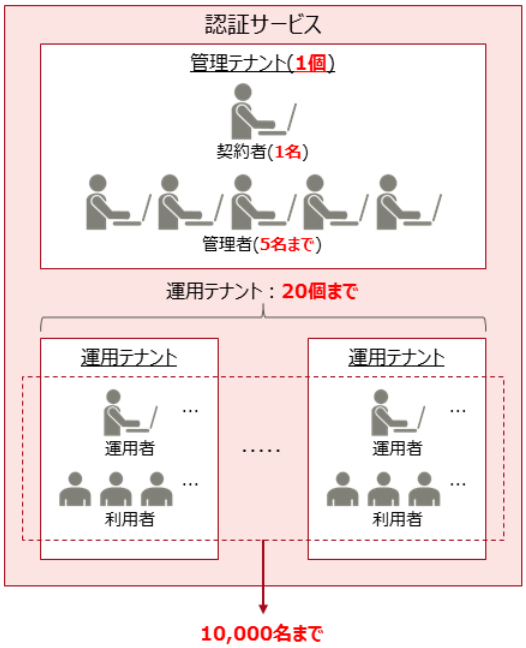
機能概要

FUJITSU

■ テナント数、アカウント(ID)数の考え方

名称	数	説明
管理テナント	1	認証サービスの利用を開始すると1つ作成されます。
運用テナント	～20	1つの管理テナントに対し20個まで作成可能です。
契約者	1	FUJITSU Cloud Service for OSSの契約者です。
管理者	～5	5名まで作成可能です。
運用者 利用者	※	管理テナントにひも付くすべての運用テナントの最大ID数（運用者＋利用者の上限数）を合算した値の上限は10,000です。

※10,000ID以上のご利用をご希望の際は、ヘルプデスクまでお問合せください。ヘルプデスクよりプラン変更方法をご案内いたします。詳細はFUJITSU Cloud Service公開ホームページのFAQをご確認ください。



Copyright 2016-2018 FUJITSU LIMITED

課金の考え方について

FUJITSU

■ 課金方法
申請ID数に応じた月額固定料金で課金されます。

■ 月額単価×申請ID数で課金されます。月額単価は申請ID数に応じて異なります。

- 申請ID数
  すべての運用テナントの「最大ID数」を合計した数です。
- 最大ID数
  お客様が運用テナントごとに任意に設定する項目です。
  該当の運用テナントに登録できる運用者ID＋利用者IDの上限数です。
- 申請ID数の集計の考え方
  ①日次で運用テナントごとに「最大ID数」を集計します。（集計時刻：UTC0時※）
  ②該当料金月の中で、もっとも大きな「最大ID数」が集計対象となります。
  毎月1日のUTC0時にすべての運用テナントの「最大ID数」を合計し、
  「申請ID数」とします。

※JST（日本時間）の9時

■ オプションを設定している運用テナントがある場合は別途オプション料金がかかります。
※ワンタイムパスワードオプションは無料です。

課金計算例は次ページをご参照ください。

Copyright 2016-2018 FUJITSU LIMITED

課金の考え方について

FUJITSU

■ 課金メニュー例

認証方式	1IDの単価（円）					
	～1,000ID	1,001～5,000ID	5,001～10,000ID	10,001～50,000ID	50,001～100,000ID	100,001ID～
基本料金	480	390	300	220	210	200
オプション料金(ワンタイムパスワード)	無料					
オプション料金(生体認証)	710	570	460	370	320	260

■ 課金計算例

- 登録ID数ではなく最大ID数を申請ID数とする
- 集計日前月の最大ID数に設定された値のうち、最大値を集計対象とする
- 日々の最大ID数は集計時刻(UTCの0時)時点の数を集計する

集計日

日付　10/10　10/11　10/12 ・・・ 11/01 0時

課金対象

テナント		ID数		
管理テナント		(課金計算対象外)	変更	
運用テナント1 生体認証オプション**なし**	**最大ID数**	**800**	**800→1500 →1000(最大値)**	**1000→500**
	登録ID数	0	300	400
運用テナント2 生体認証オプション**あり**	**最大ID数**	**800(最大値)**		
	登録ID数	700		

基本料金 申請ID数　：1800
オプション料金 申請ID数：800

【基本料金】
1,800ID×390円=702,000円

【オプション料金】
800ID×710円=568,000円

【合計】
1,270,000円

Copyright 2016-2018 FUJITSU LIMITED

制限事項・注意事項

FUJITSU

■ 本サービスを利用できるクライアント環境は以下のとおりです。

OS	Webブラウザ
Windows 7, 8.1, 10	Internet Explorer 11
iOS 10,11	Safari ※ただし、生体認証機能は未対応
Android 7,7.1	Google Chrome ※ただし、生体認証機能は未対応

※なお、動作保証がなされるOSは、ハードウェアの仕様に従います。

■ 本サービスの提供リージョンについては、FUJITSU Cloud Service公開ホームページのサービス仕様書をご参照ください。

Copyright 2016-2018 FUJITSU LIMITED

B. 6.　ブロックチェーンで実現するデータ流通・利活用

FUJITSU
shaping tomorrow with you

ブロックチェーンで実現するデータ流通・利活用 Virtuora DX

富士通株式会社

Copyright 2018 FUJITSU LIMITED

日本のデータ活用における現状

FUJITSU

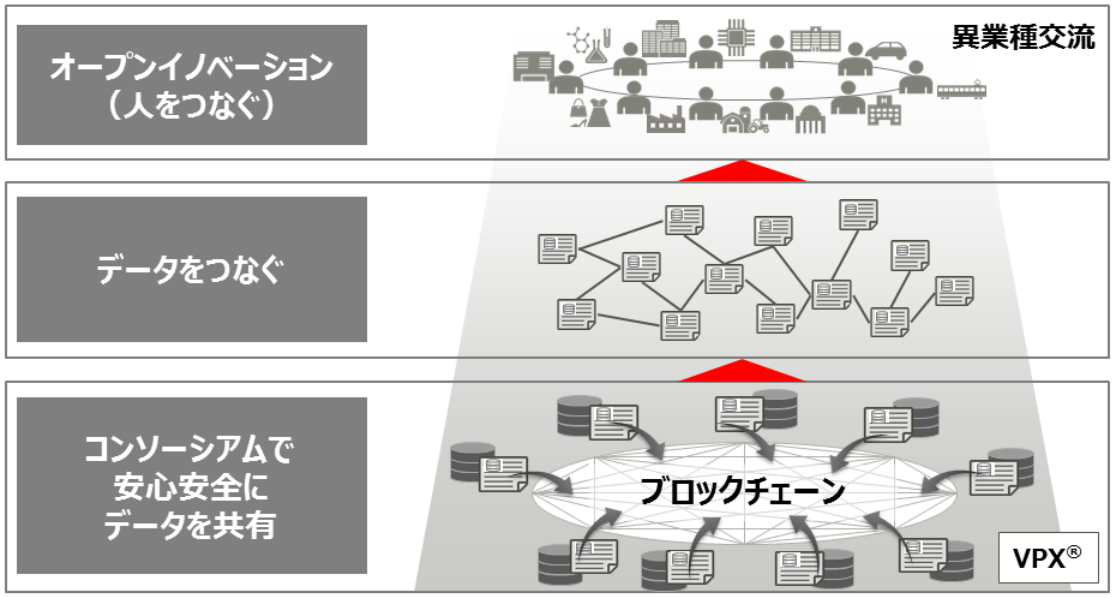
- 「データ立国ニッポン」、「データは21世紀の石油」
- 産業毎にデータ活用に取り組みながら業種を超えていく
- 日本は世界から見て、「デジタル時代に何もしない国」
- 各社、データサイエンティスト確保やオープンイノベーションに取り組み始める

Copyright 2018 FUJITSU LIMITED

Virtuora DX コンセプト

FUJITSU

組織が保有するデータを、業種・業界を跨いで安心安全に流通・利活用する「場」を提供し、Digital Co-Creation社会を実現



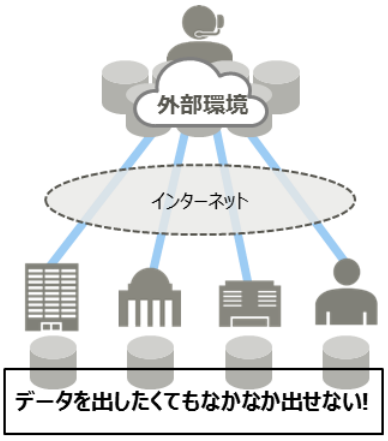
Copyright 2018 FUJITSU LIMITED

データ民主化を実現するキーコンセプト「分散型」

FUJITSU

従来のデータ共有方法
【中央集中管理型】

- 第三者機関にデータを託すことに対する不安
- データを移動するための作業・管理コスト
- 預かる側もシステム投資・運用負担、情報管理などのリスクを負う



ブロックチェーンを応用したデータ流通方法
【信頼でつながる分散共有型】

- データは手元で管理、信頼できる相手とだけ直接データを共有
- 分散台帳により、データ取り引きの履歴をトレース可能
- ブロックチェーンで処理の改ざんを防ぎ、不正データ流通を排除

データや取り引きのオーナーシップ	ピアツーピア直接取り引き	不正防止による安心・安全な環境

Copyright 2018 FUJITSU LIMITED

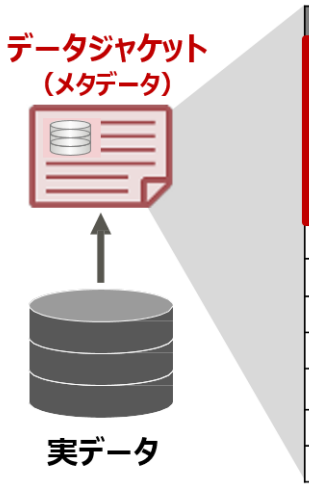
「データジャケット」でメタデータを人が理解できる形に変える

FUJITSU

- データはそのままでは人が理解しにくく、資産・利用価値の策定は難しい。
- 「データジャケット」によって、データの価値を人が理解できる形に記述

Creating a Market of Data
Data Jackets

データジャケットとは、
東京大学大学院工学系研究科システム創成学専攻大澤研究室が提唱した、**人が理解することを前提としたデータの概要情報(メタデータ)**。データの中身ではなく、概要情報(変数ラベル、保存形式、収集方法など)を共有することで、データの利用価値が多くの人に理解され、活用を促す記述形式。



項	記述項目
1	データのタイトル
2	データの概要説明
3	データ収集方法やコスト
4	データの共有条件
5	データが含む変数ラベル
6	データの所有者とその所在
7	データの種類（時系列、画像など）
8	データの保存形式（TXT、CSVなど）
9	分析・シミュレーションプロセス(データ分析を支える仮説・理論など)
10	分析・シミュレーションプロセスの結果
11	分析・シミュレーションプロセス以外に期待する分析
12	自由記述（データに関する補足事項）

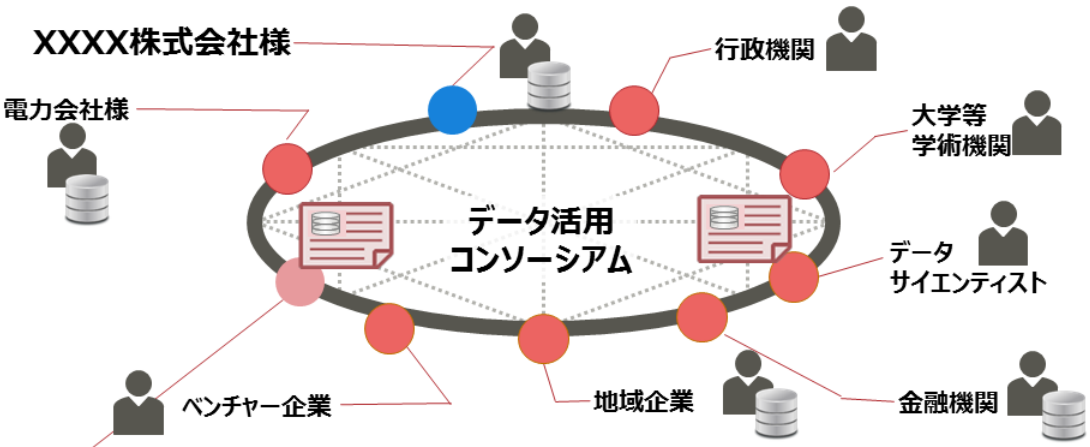
データの価値策定のための必須項目

Copyright 2018 FUJITSU LIMITED

信頼できる相手との「コンソーシアム」を組む

FUJITSU

- ブロックチェーンのアーキテクチャの名称としての「コンソーシアム」
- 共通の目的の元、業種を超えて信頼でつながる「コンソーシアム」



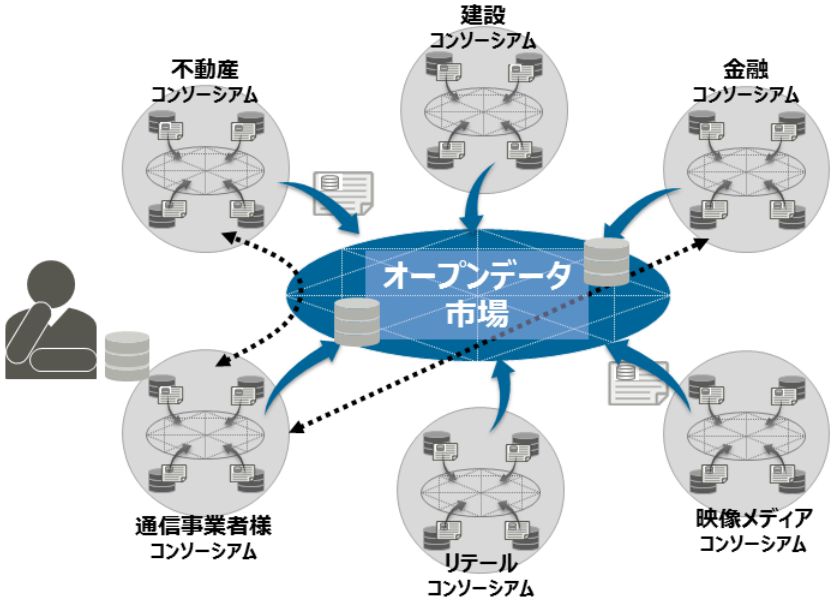
管理者	① コンソーシアムに、メンバーを招待する ② コンソーシアムメンバーに、データを持ち寄るよう呼びかける ③ サービスが適切に利用されているか監視する

Copyright 2018 FUJITSU LIMITED

「コンソーシアム」から「オープンデータ市場」へ

FUJITSU

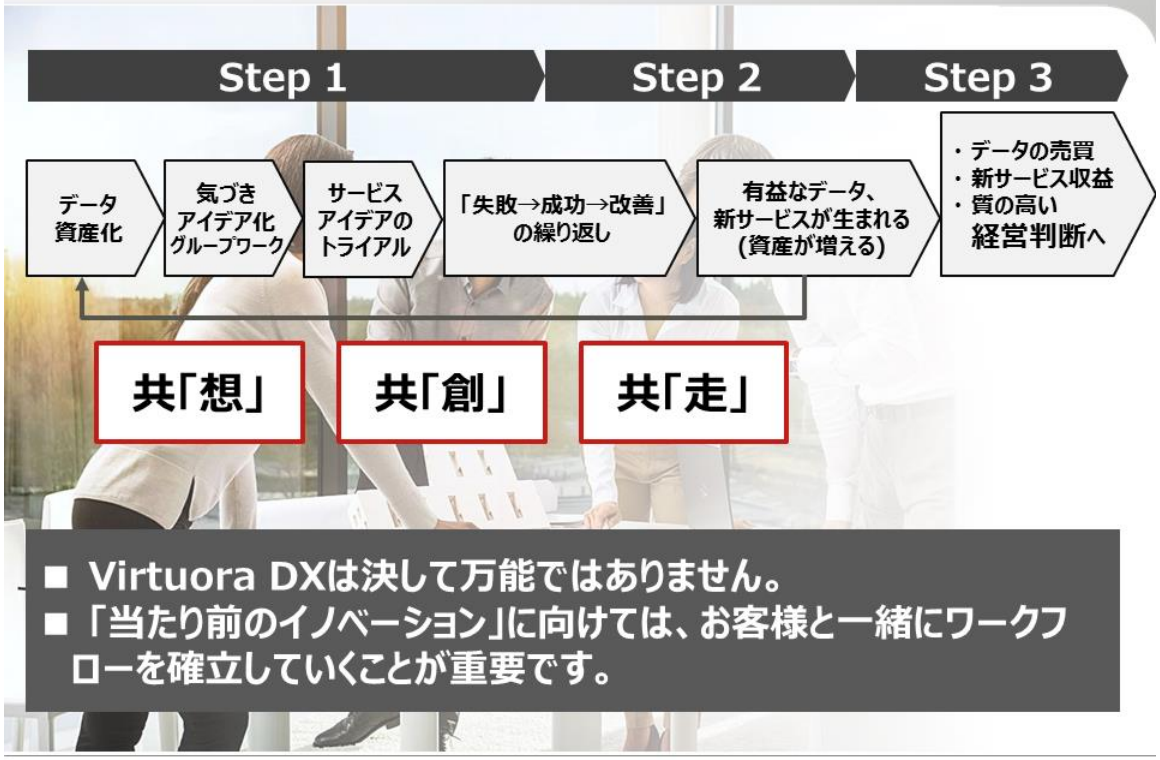
- Virtuora DXを使うほど、足りない(欲しい)データが見えてくる
- 「これなら外に出しても安全」「これなら売れるかも」と徐々に変化



Copyright 2018 FUJITSU LIMITED

新たなワークフローを共「ソウ」

FUJITSU



- Virtuora DXは決して万能ではありません。
- 「当たり前のイノベーション」に向けては、お客様と一緒にワークフローを確立していくことが重要です。

Copyright 2018 FUJITSU LIMITED

FUJITSU

ブロックチェーンの応用技術
富士通VPX®テクノロジー

Copyright 2018 FUJITSU LIMITED

富士通VPX®テクノロジー ～beyond VPN～

FUJITSU

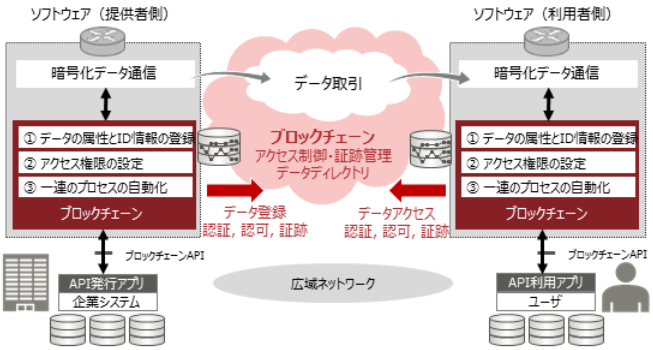
ブロックチェーンをネットワーク制御に応用した富士通独自の技術により、安全・安心なデータ流通ネットワークを実現
VPX®: Virtual Private digital eXchage

■ **データ属性とID情報の登録**
データの保管場所はそのままに、提供するデータの属性情報（データの種類や、そのデータに含まれる情報要素など）、およびデータの保管場所と紐づけたID情報をブロックチェーンの分散台帳に登録します。

■ **アクセス権限の設定**
データごとにアクセス権限を自由に設定できます。データ提供者は分散台帳を閲覧できる全参加者の中から、アクセス権限を付与した参加者にのみデータの情報を開示することが可能になります。

■ **データ取得プロセスの自動化**
データ利用者が公開されたID情報を使って実際のデータを取得する際に、アクセス権限認証、データ転送処理など一連のプロセスを自動化します。これにより、簡単かつ効率的なデータの受け渡しが可能になります。

安心	データ属性情報/IDのみ公開
安全	自由なアクセスポリシー設定
信頼	全ての処理の証跡を記録 不正取り引き排除(改ざん防止)
簡単	データ取り引きの自動化



富士通の独自技術VPX®(国内外特許出願済)によって Digital Co-Creation 社会の実現へ

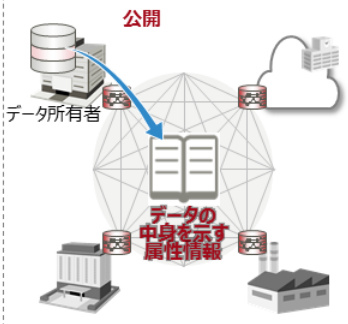
Copyright 2018 FUJITSU LIMITED

Virtuora DXがアシストするデータ流通

FUJITSU

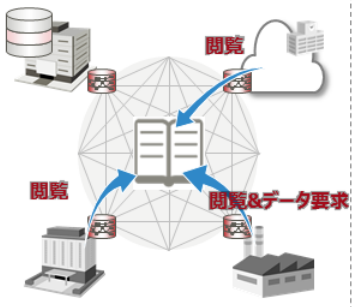
データの"ジャケット"を公開

所有者はデータの属性情報のみを
データジャケットとして公開。



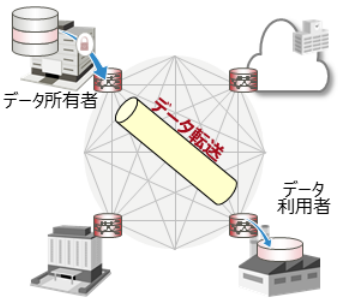
欲しいデータを見つけて交渉

データジャケットの閲覧により
データの所有者と利用者をマッチング



OKなら実データをやり取り

データの所有者と利用者間で合意の
後、直接データを転送するための制御



証跡管理 データの公開・発見・転送に係る情報を記録し、真正性を確認

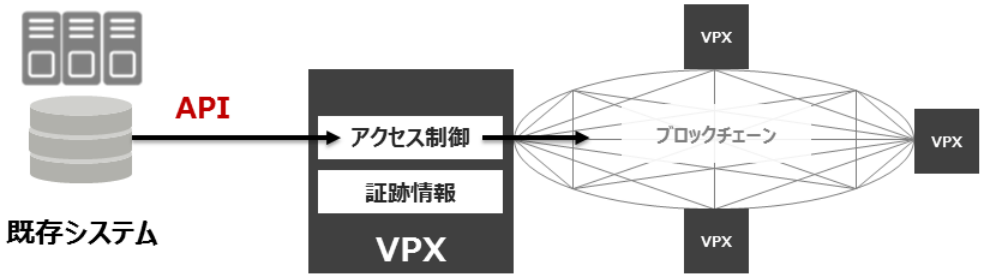
Copyright 2018 FUJITSU LIMITED

VPX®によるシステム連携

FUJITSU

■ 既存システムを変更することなく、安全なデータ流通が可能

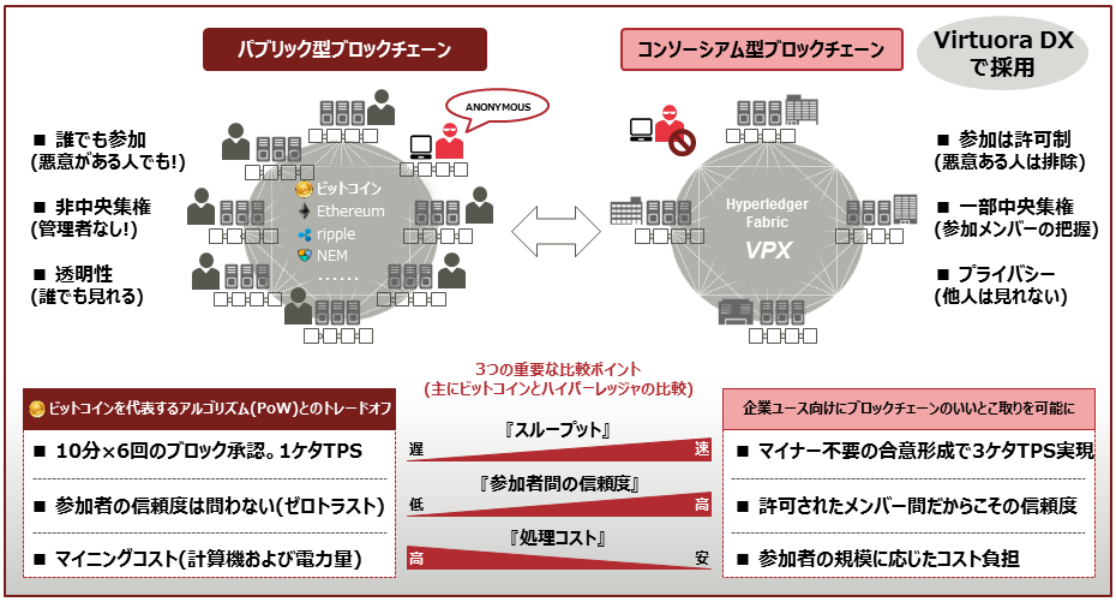
※ 既存システムのAPIをそのまま活用し、許可した相手のみアクセス可能



Copyright 2018 FUJITSU LIMITED

【ご参考】「ビットコインではないブロックチェーン」を使う意味

FUJITSU



ブロックチェーンの特性を適切に活かして実現できる「価値」(データ流通社会)を社会に提供

Copyright 2018 FUJITSU LIMITED

Virtuora DXのデモ画面

FUJITSU

- コンソーシアムメンバー間のアイデア創出活動を加速するコミュニケーション機能を統合します。
- 共有されたデータに関する詳細な条件確認や、新たなアイデアに関するディスカッションなど、組織や企業を超えたコミュニケーションをサポートします。

※本画面デザインは予告無く変更になる場合があります。

Copyright 2018 FUJITSU LIMITED

【発行者】サイエンティフィック・システム研究会
【編　集】サイエンティフィック・システム研究会
クラウド時代のデータ共有技術WG
【発行日】2019 年 4 月 30 日

◆ ◆ ◆

＜連絡先＞　サイエンティフィック・システム研究会 事務局
〒105-7123 東京都港区東新橋 1-5-2　汐留シティセンター
E-mail：ssken-office@ml.css.fujitsu.com
http://www.ssken.gr.jp/MAINSITE/

※著作権は各原稿の著者または所属機関に帰属します。無断転載を禁じます。