

06/08/2012

DataDirectTM
N E T W O R K S
I N F O R M A T I O N I N M O T I O N TM



Distributed Hyperscale Collaborative Storage

Web Object Scaler(WOS) Family 製品紹介
第一回IPABセミナー

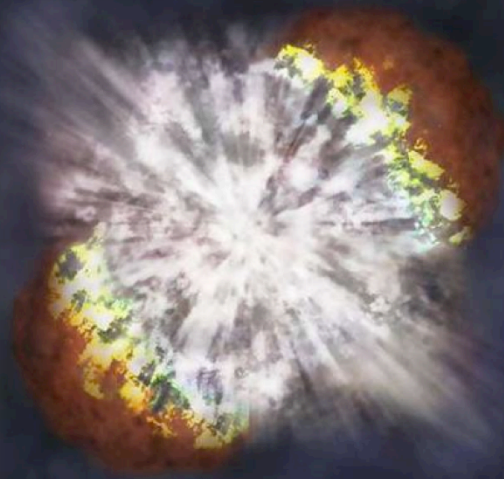
DataDirect Networks Japan, Inc.

<http://www.ddn.com/products/web-object-scaler-wos>

Nobu Hashizume

nhashizume@ddn.com

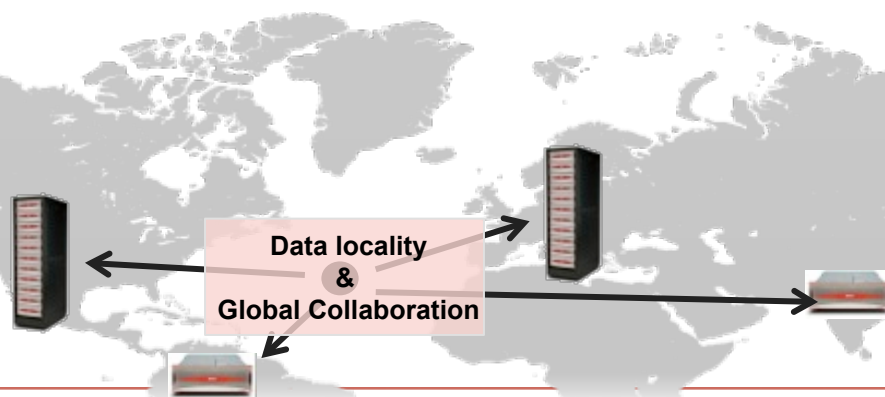
BIG DATA



Hyper-scalable • Distributed • Collaborative

DDN Web Object Scaler

- ▶ Big Data環境のためのコンテンツ・ストレージ・ビルディングブロック
 - 業界最高水準のスケールアウト・オブジェクトストレージ・アプライアンス
 - 前例のない高性能・高効率製品
 - Built-in シングルネームスペースとグローバル・コンテンツ・ディストリビューション
- ▶ 協調環境に最適化
 - 地理的要因を考慮したアクセス・レイテンシの最適化
 - “Just-in-time”プロビジョニング
- ▶ 業界最高水準のTCO削減率
 - “Near ZERO”管理機能
 - 自動“Continuity of Operations”アーキテクチャ



WOS 利用形態



WOSアクセスゲートウェイ

- NFSプロトコル
- アクセスコントロール
- 複数サイト上の複数ゲートウェイによるスケーラビリティ
- HA & DR保護機能

クラウドプラットフォーム

- S3-Compatible & WebDAV APIs
- Multi-tenancy
- Reporting & Billing
- Remote storage, file sharing, and backup agents

ネイティブオブジェクト

- C++, Python, Java, PHP, HTTP REST インタフェース
- PUT, GET, DELETE オブジェクト、RESERVE ObjectID, etc

• WOS Core

- ネイティブオブジェクトアクセス
- ポリシベース・ジオレプリケーション
- ObjectAssure

• WOS Access NFS Gateway

- スケーラブルな複数ゲートウェイ
- DR保護機能とHAフェイルオーバー
- リモートサイト間のデータベース同期機能
- ローカルRead/Writeキャッシュ
- WOSへのLANもしくはWANアクセスを提供

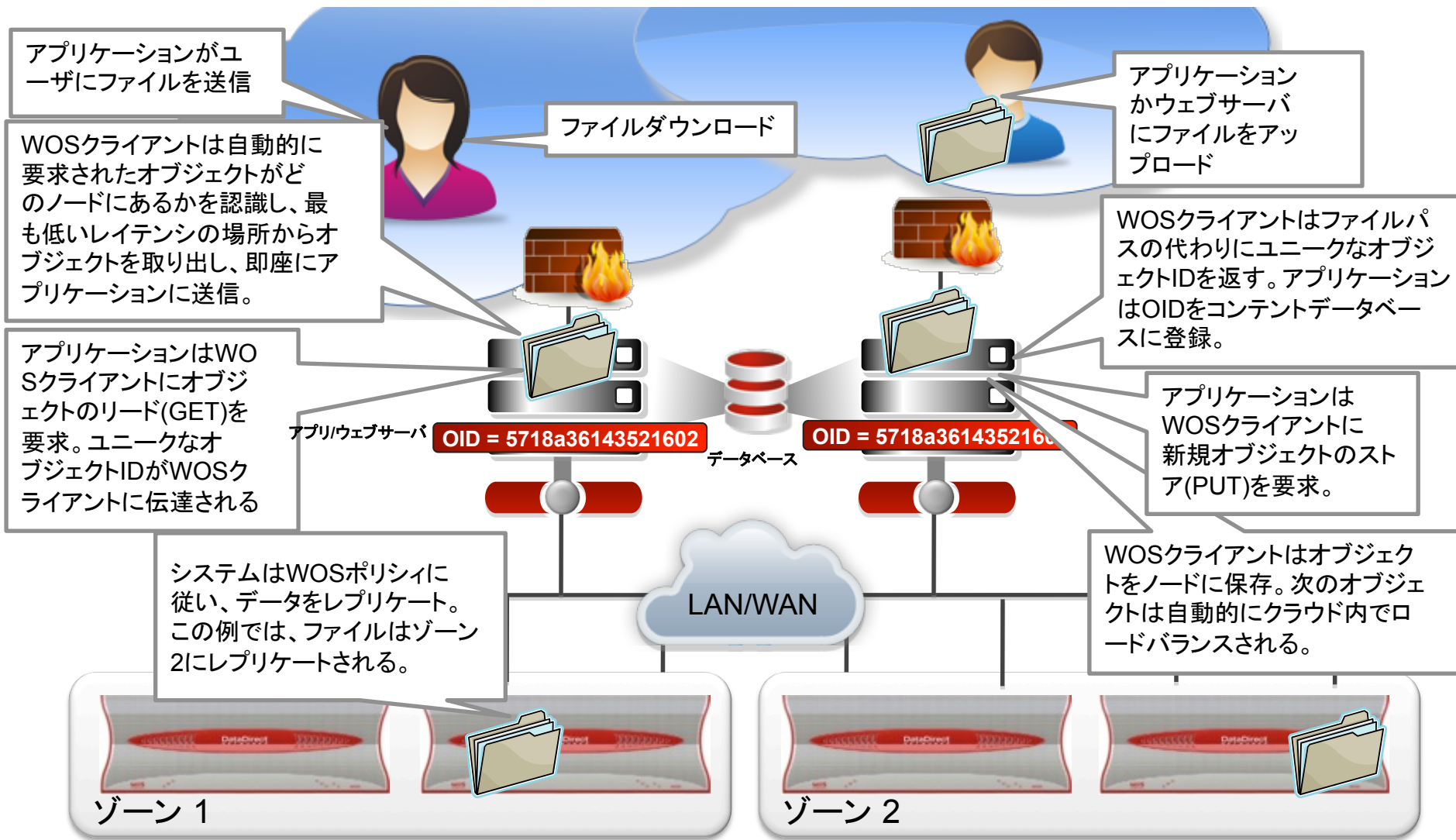
• WOS Cloud

- クラウドサービスプロバイダ、プライベートクラウドが対象
- 僅かな投資でS3アプリケーションからWOSストレージを利用可能
- マルチテナンシィ、従量課金、テナント毎のレポートをサポート



Web Object Scaler Building Block

WOSコンセプト



WOSCoreビルディングブロック

WOS 6000 - 4U high density 60-drive

WOS 1600 - 3U high-performance 16-drive

特徴

- 業界最高水準のDDNハイパフォーマンスストレージプラットフォーム上に構築
- ノード毎に4x GbEか2x 10GbE/IBのネットワークコネクション
- 業界最高の密度とスケーラビリティ
 - 1.98PB / ラック、クラスタ全体で23PB
 - 660 ドライブ / ラック
 - 220億オブジェクト / ラック、1280億オブジェクト / クラスタ
 - 512バイトから500GBまでのどのサイズでも99%の実効容量を実現
- リニアなクラスタスケーラビリティ
- 低レイテンシ
 - 1MB以下のオブジェクトの read/writeは1回のディスクI/Oで実行

WOS 6000



4U, 60-drive WOS
Node (SAS/SATA)



2PB / 11 Units
per Rack

WOS 1600



3U, 16-drive WOS
Node (SAS/SATA)



544TB / 15 Units
per Rack

WOSレプリケーション

- ▶ ポリシー単位で同期/非同期レプリケーションを選択
 - オブジェクトの種類をベースにし定義
- ▶ 同期レプリケーションはアプリケーションに”write”成功を伝達する前に、リモートサイトへの複製完了を保証
- ▶ 性能向上のためにはリモートWOSノードへの非同期レプリケーションが有効
 - データロスの最小化/回避のために、データのコピーを、”PUT”操作がOIDを返す前に必ず二つ”local”(最低レイテンシな) ノード上に作成
 - レプリケーションはバックグラウンドで実行され、レプリケーション成功後ローカルに作成されたデータロス回避用のコピーは削除
 - 大きなファイル、大きなデータセットに特に有効

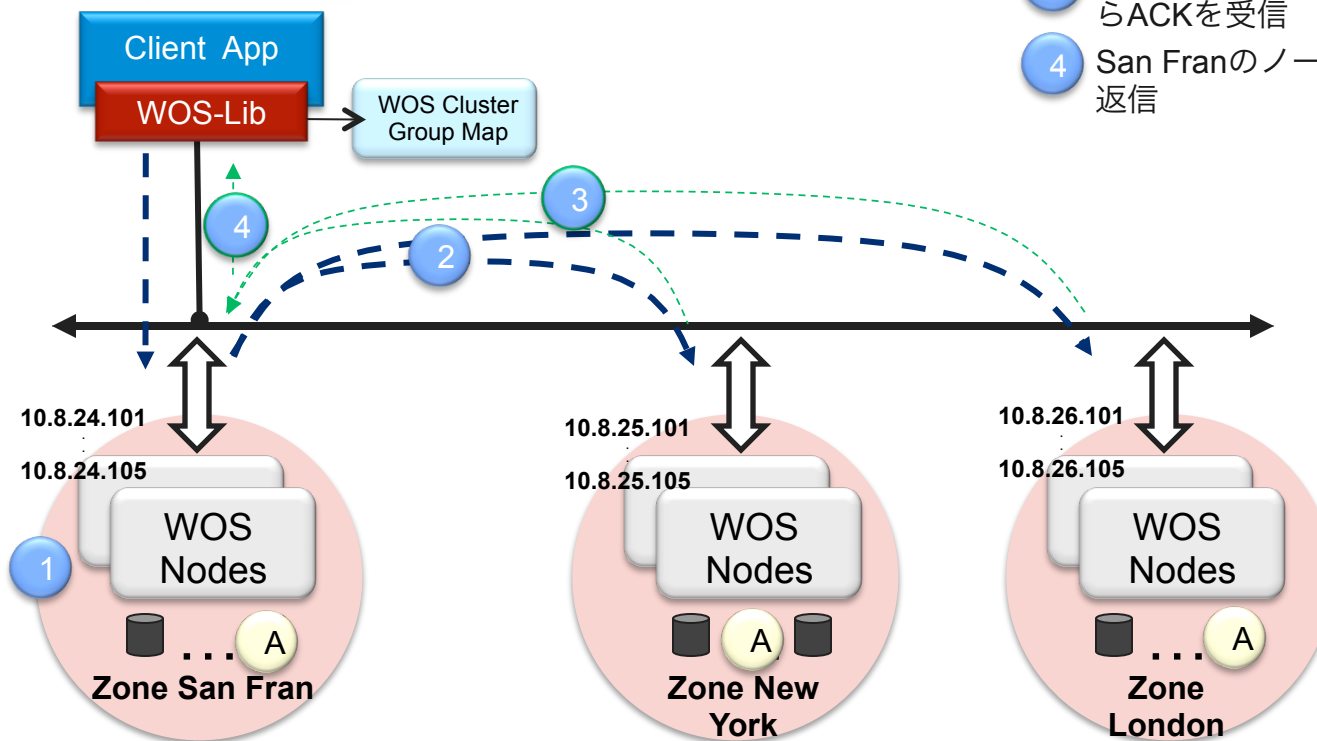


WOS 同期PUT操作



同期レプリケーション PUT操作

- 1 WOSLibは“最近接”なノードを選択
- 2 “San Fran”ゾーンのノードはオブジェクトをディスクに書き込み、ポリシーに従ってNew Yorkとロンドンに複製を作成
- 3 San FranのノードがNew YorkとLondonゾーンからACKを受信
- 4 San FranのノードがアプリケーションにOIDを返信



WOS 非同期PUT操作

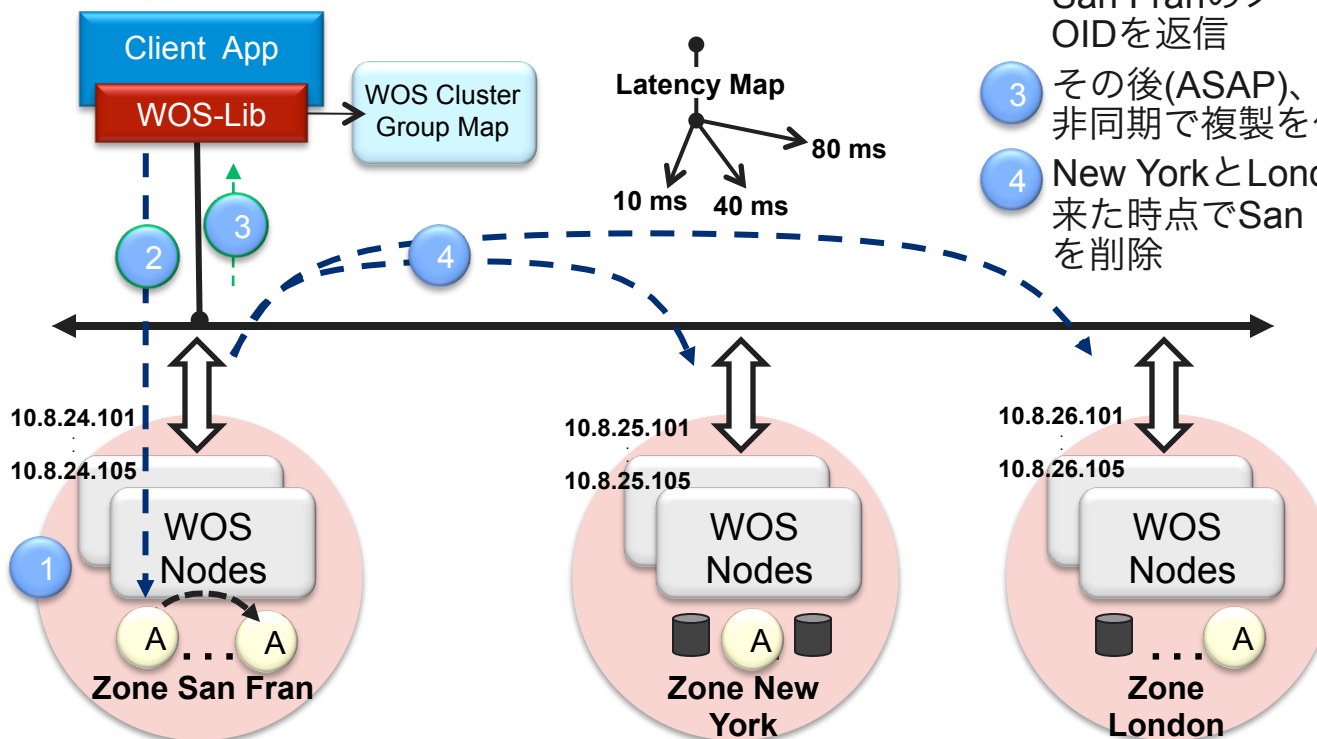
Acme WOS 1

- San Fran
- New York
- London
- Tokyo

Policies	
Policy Name (ID)	Zone Replication
Policy Name: UnitedStates	
<input type="button" value="Create Policy"/>	
<input type="button" value="Cancel"/>	
Zone	Replica Count
San Fran	<input type="text" value="1"/>
New York	<input type="text" value="1"/>
London	<input type="text" value="1"/>
Tokyo	<input type="text" value="0"/>

非同期レプリケーションPUT操作

- 1 WOSLibは“最近接”なノードを選択
- 2 “San Fran”ゾーンのノードはノード内の異なるディスクにオブジェクトのコピーを二つ作成
- 3 その後(ASAP)、New YorkとLondonゾーンに非同期で複製を作成
- 4 New YorkとLondonゾーンからACKが返って来た時点でSan Franゾーンの余分なコピーを削除

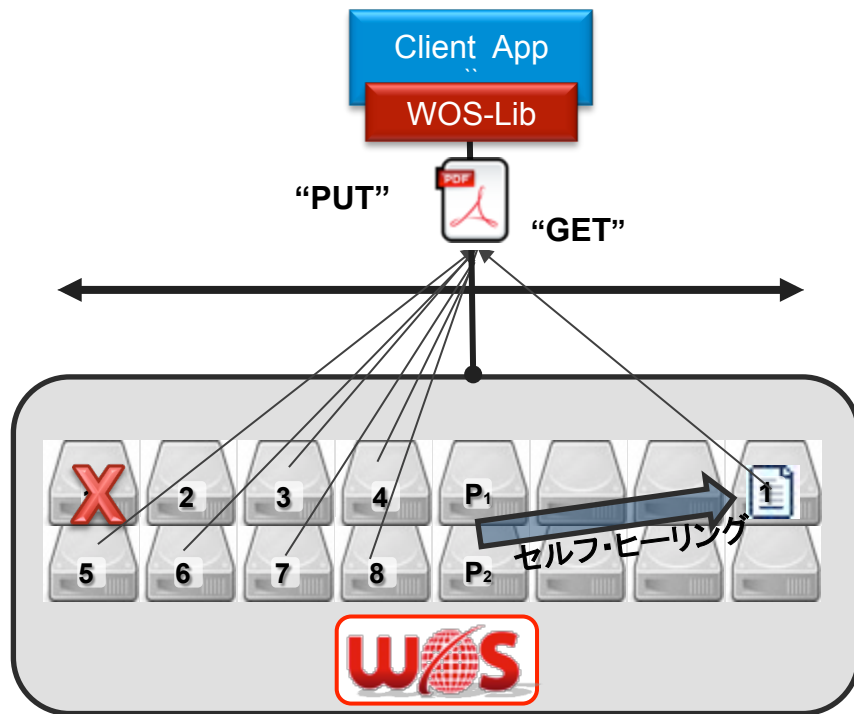


ObjectAssure : シングルコピー・データ保護

- ▶ ObjectAssure : 非クラスタ化データ保護をベースとしたerasure-code
 - erasure-codeはオブジェクトをフラグメントに分け、それを異なるディスク上に配置することにより、冗長性を確保
 - データはどのフラグメントの組み合わせからも復旧可能
 - ObjectAssureはハイパースケール、高性能クラウドストレージ向けの初めてのerasure-code保護機能
- ▶ ObjectAssureでは、ディスク2台同時故障まではデータロス、データ可用性を失わずに動作可能
 - レプリケーションコスト抜きのデータ保護
- ▶ シングルコピー環境を提供
 - 実効容量の拡張とストレージ全体コストの更なる圧縮を実現



ObjectAssure : シングルコピー・データ保護



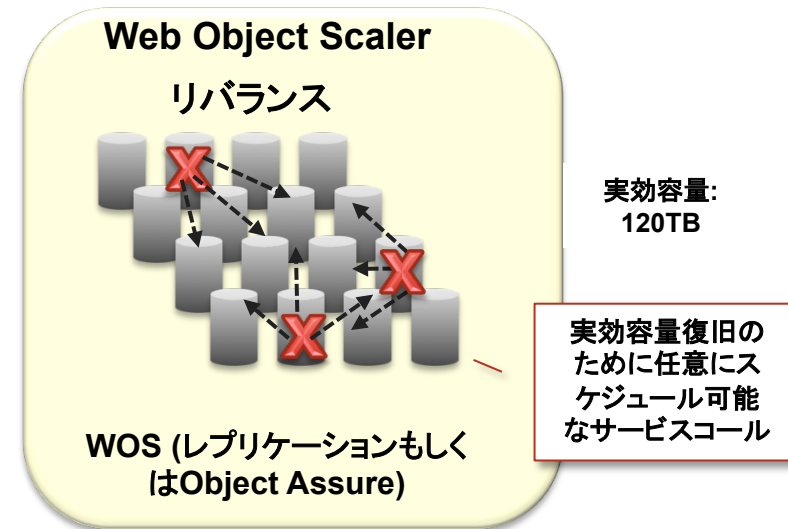
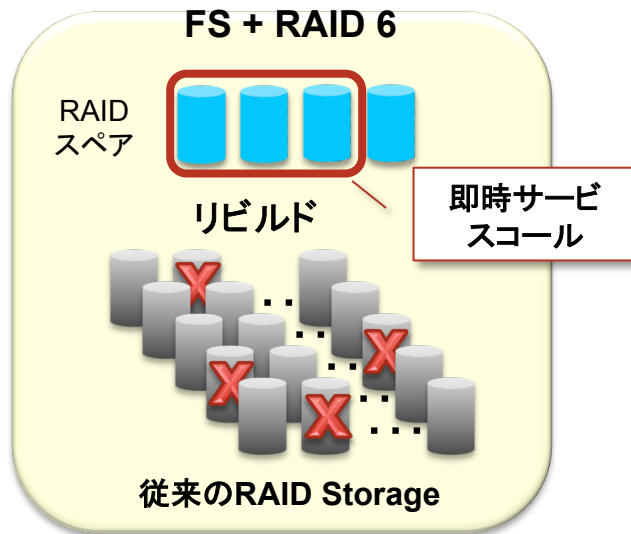
オペレーション

- ▶ Object Assure(OA)はWOS-Lib、REST API双方で利用可能
- ▶ OAは単一WOSノード上で機能する
- ▶ OAはWOSポリシーでシングルレプリカを指定することによって機能する
- ▶ WOSクラスタ上ではOAとレプリケーションの両方の機能が共存可能
- ▶ OAはWOSノード毎に同時発生した複数ディスクエラーを検知可能、且つ同時発生した二つのディスクエラーを復旧可能

ObjectAssure(OA)プロセスフロー

1. PUT操作において、OAはオブジェクトを8ドライブに分割し、更に2パリティドライブを生成(8+2)
2. ドライブ障害またはオブジェクトが破損した場合、WOSはパリティドライブを用いてオブジェクトを別ドライブ上に再構築する(セルフ・ヒーリング)。障害ドライブ上のデータはノード上の複数ディスクに分散される
3. WOS OAは必要であれば、リード(GET)中にデータ訂正を行う

インテリジェントデータ保護 RAIDリビルド vs WOSリバランス



RAIDドライブ再構築

- ▶ 無駄な容量 – スペアドライブが必要
- ▶ 長時間の再構築 – 故障ディスクの使用済み容量に関わらず全てのドライブに渡る再構築が必須
- ▶ データロスリスクが高い – スペアディスクが無い場合、再構築されない
- ▶ サポートコストの増大 – スペアディスクの補充のために即時サービスコールが必要
- ▶ 書き込み性能の低下 – RAIDは書き込み性能が低下、小さいファイルの書き込みでは顕著

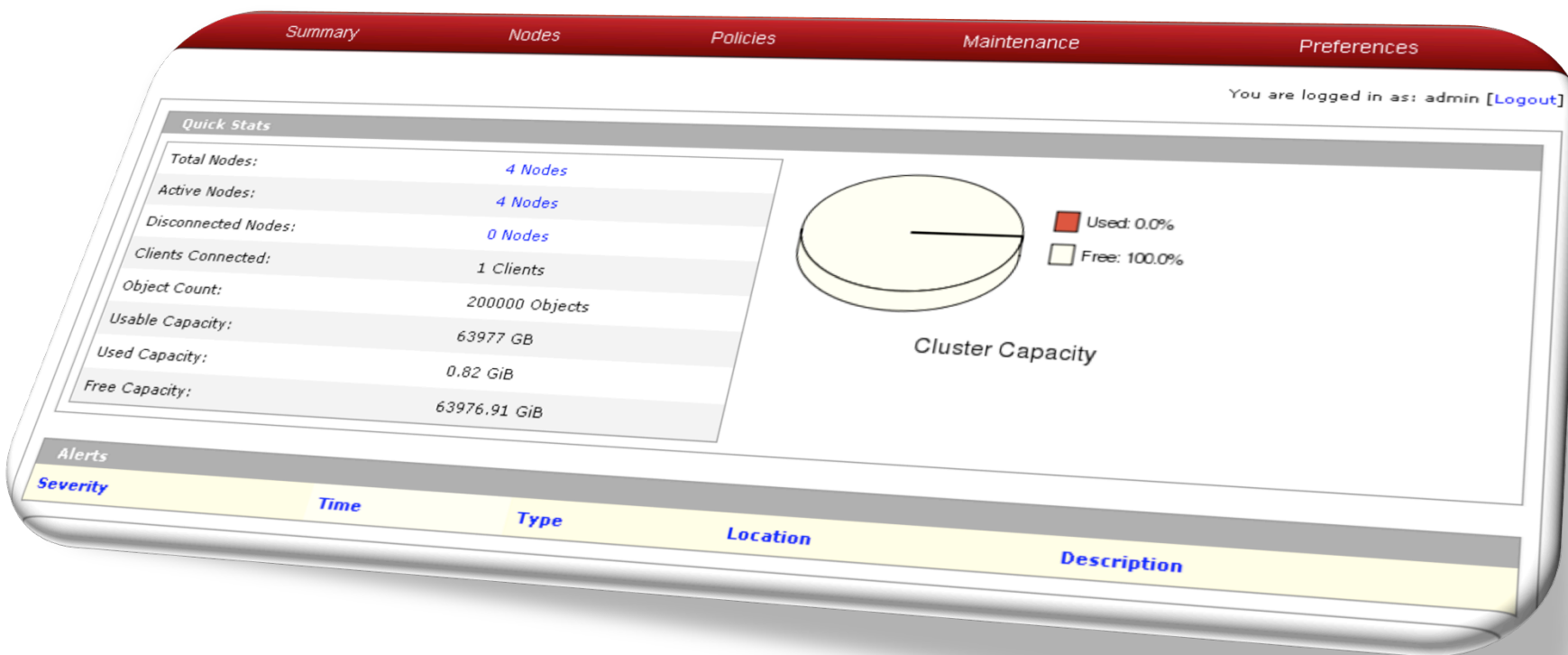
ドライブ間WOSデータリバランス

- ▶ 全てのドライブを利用可能 – どのドライブの空き領域もスペアプール的一部分となり得る
- ▶ リバランスに必要な時間はRAID再構築の50%以下 – 実データのみコピー
- ▶ 短時間でリカバリが行われるために、全体の性能向上とデータロスリスクの減少を実現する
- ▶ ドライブ障害は故障ドライブの容量分だけ全体容量が減少
- ▶ 全体容量は定期保守時に故障ドライブを交換する事によって復旧

WOSの利点

容易な管理

- ▶ 容易なGUI管理
- ▶ “まるでApple製品のような”
 - *Customer quote*



WOS サマリ

スケーラビリティとTCO削減

WOSにより拡張性と性能に対する制限を排除

- クラスタ拡張によりリニアにスケール(性能、オブジェクト数、容量)
ペタバイト環境への容易な拡張

WOSにより緊急サービスコールを排除

- 人手を介さずに保守 (セルフヒーリング)
- 適切な縮退 – 自動的にデータをクラスタ内に再配置
- 従来のRAIDに比べて数倍高速なりカバリ

WOSによる数PBストレージの管理に必要なFTEは1未満

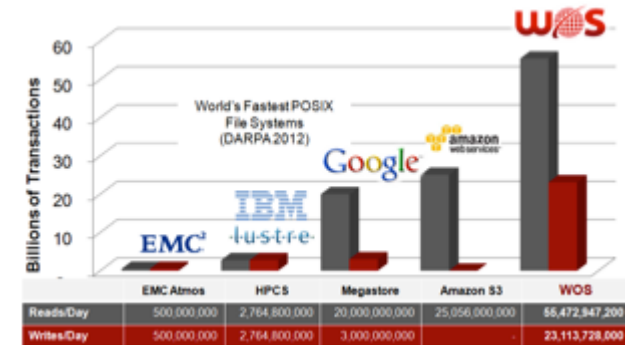
- 単一、グローバルネームスペースによる容易なデプロイ、構成、管理
- ダウンタイム無し、ユーザに影響無しに瞬時に拡張可能

データ保護とDRは追加コスト無しに利用可能

- QoSベースの柔軟なデータ保護方針 – レプリケーションとErasure-Codeを混在可能
- DIFを内蔵、サイレントデータコラプション無し
- 全てのレプリカを利用可能

WOSのよりスペース、電力、冷却コストを半減

- ストレージの効率性は空き容量ではない - 99% WOSドライブ利用率 vs 65% NAS
- 高密度実装 - 180TB/240TBを4Uラックスペースで実現



DataDirect[™]
N E T W O R K S
I N F O R M A T I O N I N M O T I O N[™]

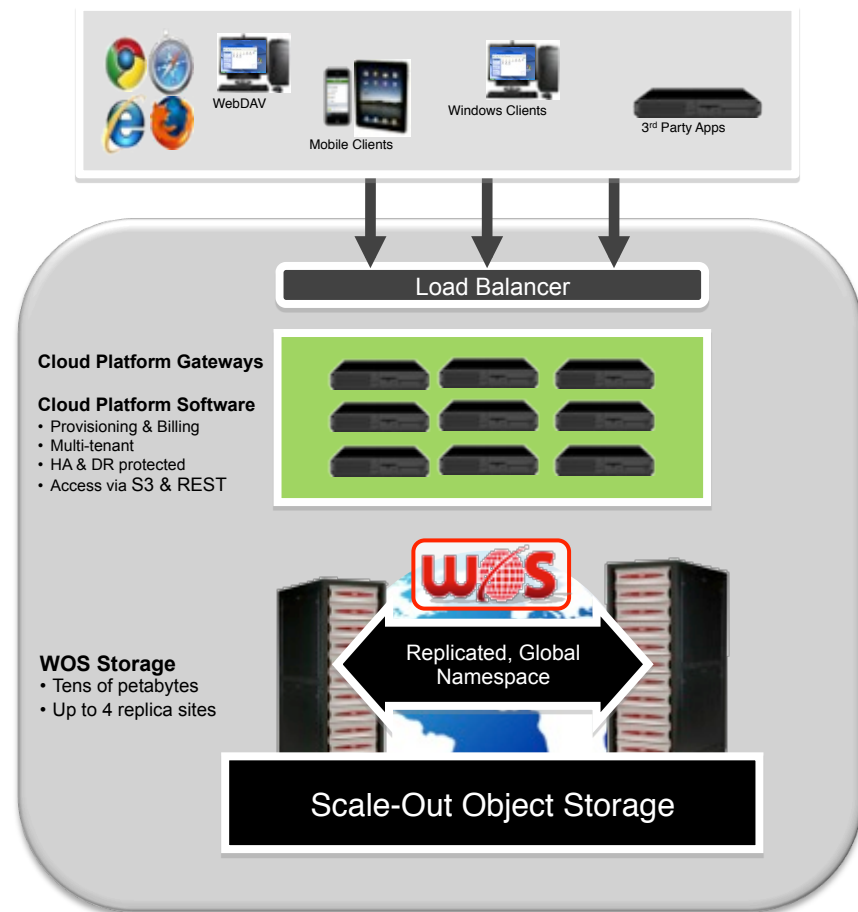
WOS Cloud

Scalable Content Storage
Platform for the Cloud

WOS Cloud概要

サービスプロバイダ、プライベートクラウド、エンタープライズ向けクラウドストレージプラットフォーム

- ▶ クラウドプラットフォームソフトウェア
 - パーソナルクラウドストレージ
 - マルチテナンシーサポート
 - 既存のプロビジョニング、課金システムとの統合
 - S3互換API、WebDAV APIを提供
 - 地理的ロケーションコントロール
 - 複数サイト協調・共有環境
- ▶ Web Object Scaler (WOS)ストレージ
 - インターネットワイドファイルストレージのためのコンテンツ基盤環境
 - グローバルコンテンツ配信
 - 容易なグローバルストレージ管理
 - 業界最高レベルの低TCO



パーソナルクラウドストレージ 柔軟で安全なパーソナルアクセス



パーソナルクラウドストレージ

利点:

- 安全なグローバルファイル共有・協調環境を提供
- ウェブ、スマートフォンからのanytime/anywhereアクセス環境を提供
- ユーザはいつでも再近接のファイルにアクセス
- シンプロビジョニングによるストレージコストの削減
- 大容量ファイルを容易にグローバル共有・公開可能
- 個別ブランディング可能
- 容易なデプロイ・プロビジョニング



White label agents for
PC, laptop, smart
phones, iPad, etc



- ### 共有・協調のための優れたファイルサービス
- ポリシベースのサイト間自動ファイルレプリケーション
 - セキュリティポリシーに則ったユーザファイル共有
 - 再近接コピーへアクセス
 - 複数サイト間協調が可能

WOS Cloud 互換API

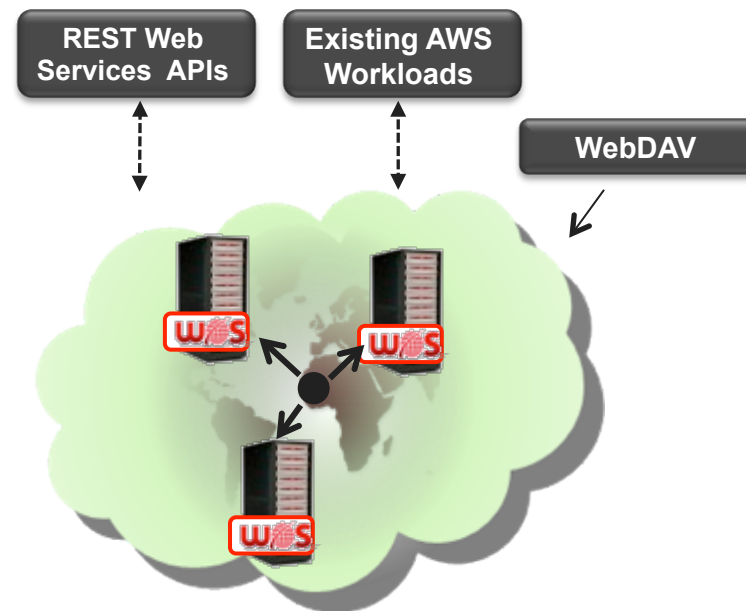
オン・デマンドな優れたサービスとスケーラビリティ



互換APIによるクラウドストレージプログラミング

利点:

- 既存のAmazon Web Service(AWS)を内部へ
 - 他のストレージクラウドを利用するアプリケーションとの相互運用が可能
 - 様々なパートナーソリューションを選択可能
- ベンダ・ロックインを排除
 - 複数APIをサポート
- よりよいサービスの提供
 - データの地理的配置要求に対処
 - URL配信、コンテンツタグ、階層ストレージ等の追加機能が可能
- 複雑さの排除とコスト削減



既存のAWSを内部へ / 3年間のTCOコストがWOSを利用したプライベートクラウドで50%削減

パーソナルクライアント

パーソナルで安全な “Dropbox-like”機能

- ▶ iPhone, iPad, Andriod(将来対応予定), Webクライアント
- ▶ 安全なグローバルファイル共有・協調環境を提供
- ▶ 大容量ファイルを容易にグローバルに共有・公開可能



Webクライアント



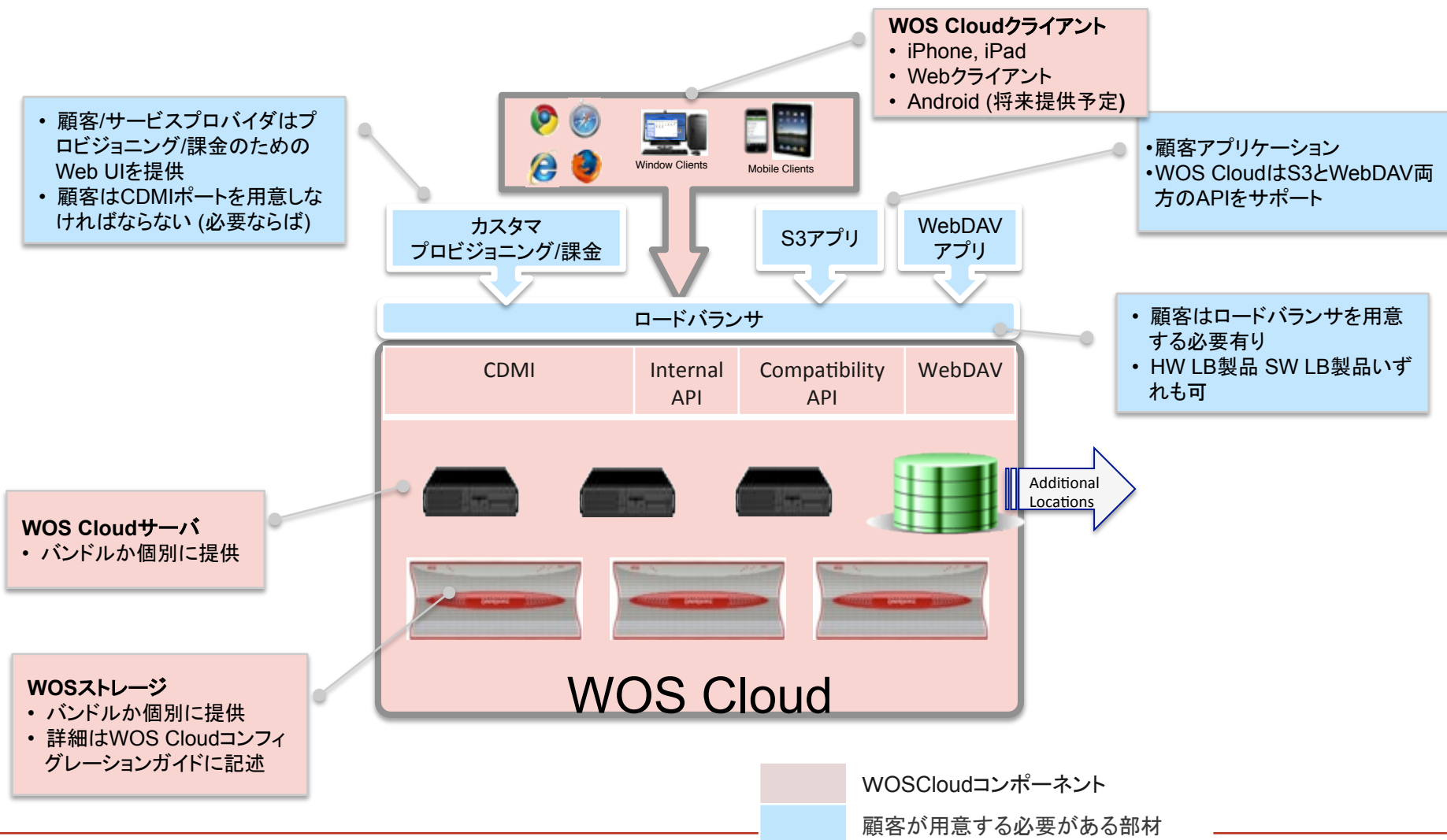
Mobileクライアント

WOS Cloudパーソナルストレージクライアント

- ▶ Apple AppストアよりiPhone, iPadクライアントを入手可能
- ▶ WebブラウザからWOS Cloudパーソナルストレージにアクセス可能
 - WOS Cloudをインストールするだけで利用可能
- ▶ 顧客は個別にWOS Cloudクライアント上で各組織のロゴを利用可能



WOS Cloudアーキテクチャ



DataDirectTM
N E T W O R K S
I N F O R M A T I O N I N M O T I O NTM



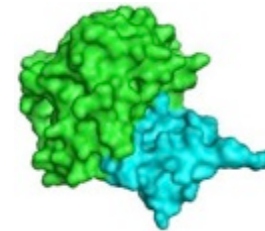
東工大様
事例

WOS Cloud for TokyoTech ACLS

WOS Cloud for TokyoTech ACLS

The Tokyo Institute of Technology Education Academy of Computational Life Sciences

- The Tokyo Institute of Technology (TokyoTech) is the largest institution for higher education in Japan dedicated to science and technology
- TokyoTech's Education Academy of Computational Life Sciences (ACLS) trains students and researchers in both Life Science and Computer Science
- Access to the world famous, award winning, TSUBAME 2.0 supercomputer (Future plan)



WOS Cloud for TokyoTech ACLS

Challenges

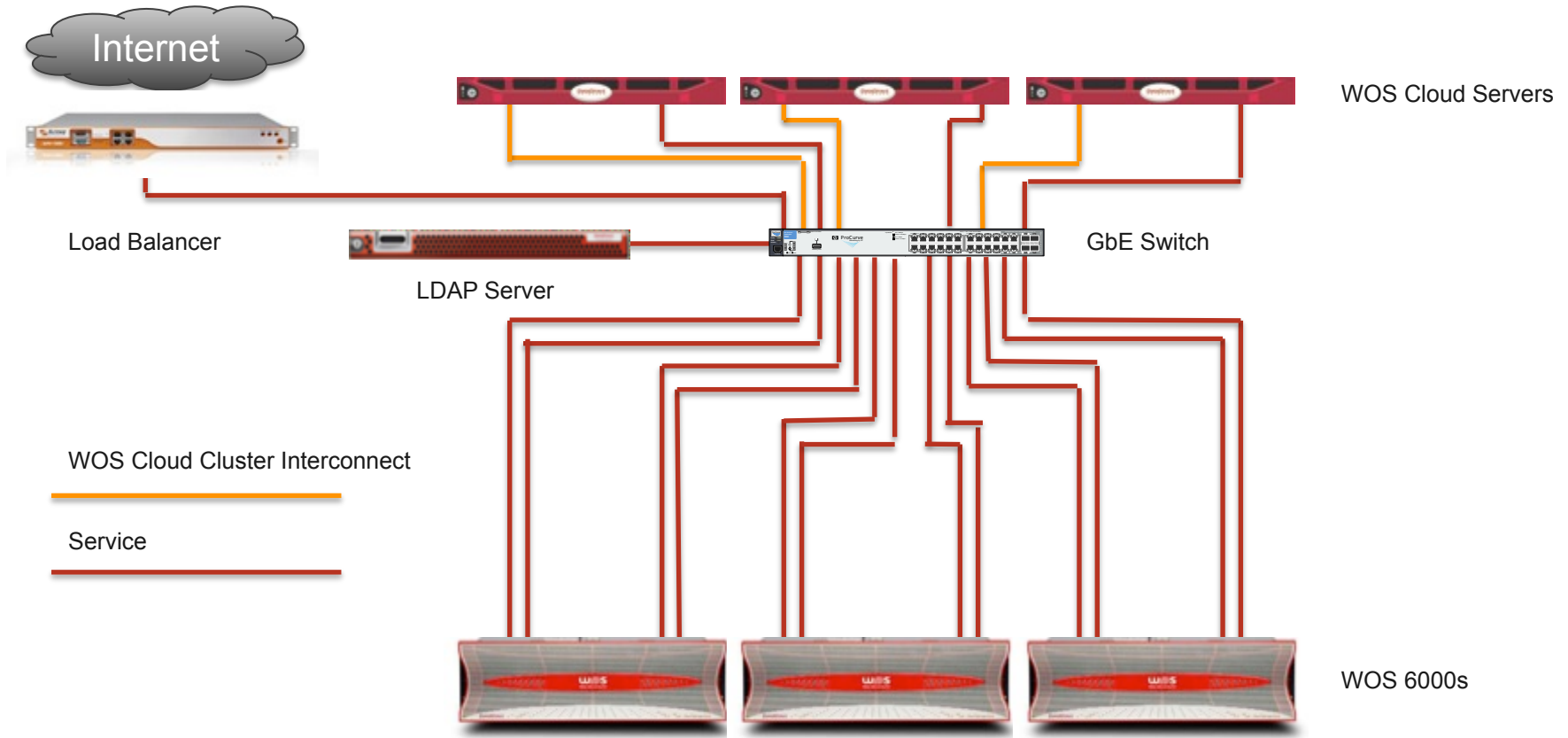
- TokyoTech ACLS wanted a cost effective, highly reliable storage system to cope with the data growth of their data-driven education and research initiatives
- Professors and researchers needed to store high resolution microscopy, genomic data, molecular and macroscopic biologic imagery, and other system biologic sensor output along with meeting and presentation materials
- The ACLS wanted easy file storage and sharing between the PC's, iPhones and iPads used throughout the organization
- The ACLS wanted to have their own service, completely under their control - and in their data center - in order to provide secure file sharing and storage for its researchers and students

WOS Cloud for TokyoTech ACLS

Solution

- The DDN solution was the 360TB WOS Cloud system, consisting of 3 x WOS Cloud servers and 3 x WOS6000 (120TB each)
- Files are uploaded via their PC through the WOS Cloud Client
- Data from a variety of equipment throughout the Academy, including images and other biologic sensor data, voice data, movies and text data are uploaded by S3 applications developed by TokyoTech
- TokyoTech selected WOS Cloud because it offered behind the firewall security, high density, modular scalability and S3 compatibility
- The ACLS has about 600 users that access WOS Cloud
- The ACLS especially liked WOS's object storage replication mechanism as it met their plans to provide future storage and access at an additional location

TokyoTech WOS Cloud Network



TokyoTech WOS Cloud Network

TokyoTech WOS Cloud Installation





WOS Object Storage for OCP

WOS for Open Compute Project (OCP)

WOS for OCP

▶ DDN Joined Open Compute Project

▶ About OCP

- The OCP was initiated by Facebook to reduce equipment and operating costs for hyperscale datacenter deployments. The OCP has established all new specifications for everything from how data centers are built & cooled, to where equipment is located, the size/composition of rack-mounted servers & storage, the physical racks themselves and the power that supplies them.

▶ WOS on the OCP Open Vault Storage Unit

- the Open Vault storage unit draft specification is a 2RU component. <21.1" W x 34.6" D> 30 HDD JBOD SAS storage enclosure, consisting of two identical 1RU HDD trays with 15 x 3.5" HDDs.





Thank You & Demo

DataDirect Networks, Information in Motion, Silicon Storage Appliance, S2A, Storage Fusion Architecture, SFA, Storage Fusion Fabric, Web Object Scaler, WOS, EXAScaler, GRIDScaler, xSTREAMScaler, NAS Scaler, ReAct, ObjectAssure, In-Storage Processing and SATAssure are all trademarks of DataDirect Networks. Any unauthorized use is prohibited.