



最先端シミュレーションとAIにおける これからのストレージとは

2024年3月13日

スキヤリティジャパン株式会社



Agenda

1. Scality会社紹介
2. 最先端シミュレーションとAI分野での課題
 - 検討すべき項目は何か？
3. まとめ



Scality 会社案内



- **Scality**

- 設立: 2009年
- 本社所在地: フランス パリ
- 拠点: パリ、サンフランシスコ、ロンドン、ワシントンDC、東京
- 製品: Scality RING (オブジェクトストレージソフトウェア)
- Gartner MQ for Distribution Filesystem and Object Storage にて8年連続 Leaderポジションを獲得

- **Scality Japan**

- 設立: 2015年3月
- 所在地: 東京都千代田区大手町
- 国内の主なお客様: メディア系、研究機関、大学、中央政府系機関、医療機関、通信事業者、サービスプロバイダ、金融業、製造業など
 - 日本国内で100PB以上 (実効容量) の出荷実績
 - 2021年の日本のオブジェクトストレージ市場でシェア1位 (スマプラ総合研究所調べ)



最先端シミュレーションとAI分野での課題

- シミュレーションとAIにおけるデータに対する共通の課題
 - より精度の高い結果を得たい=より多くのデータを用いて行う必要がある

この課題を解決するための技術として

- IoTによるリアルタイムでの詳細な情報取得
- 計算機リソースや通信技術の進化による大容量のデータ処理/データ転送



技術の進歩により、シミュレーションやAI分野も加速的に良い方向へ進んでいる

ただし技術が進歩するほど、データ容量は増加する

- データ増加に対する検討すべき項目は何か
 - 巨大化するデータの長期的な保管
 - 柔軟なデータへのアクセスと管理
 - データ主権



巨大化するデータに対するScality

大規模データとの10年以上の歩み

Scality RINGの特徴

- 大規模なLinuxコンピュートクラスタ/GPUファームの計算結果蓄積
- 最新のAI/MLライブラリやツールを使用するHPCの結果も蓄積
- S3とRINGの組み合わせでインターネットアクセスを広げる
- RINGはTieringされた領域としてアーカイブ
- スケーラビリティ（拡張性）
- 長期にわたっての保管

実績のある業種（おもに欧州）

- 大手通信業
- 動画配信基盤
- 大手製造業（衝突解析データ蓄積）
- 国有鉄道の取得データ
- 最近は航空宇宙開発企業
- 監視カメラ

※ 10年以上、拡張を十数度行い、数十PBのデータ蓄積を行っている実績も多数

- 永年保管に近いアーカイブが求められる → HPCやAI領域での永年アーカイブにも応用



事例紹介: 某自動車メーカー (Simulation Data Archive/36PB/S3, SMB, FUSE)

■ 基本情報

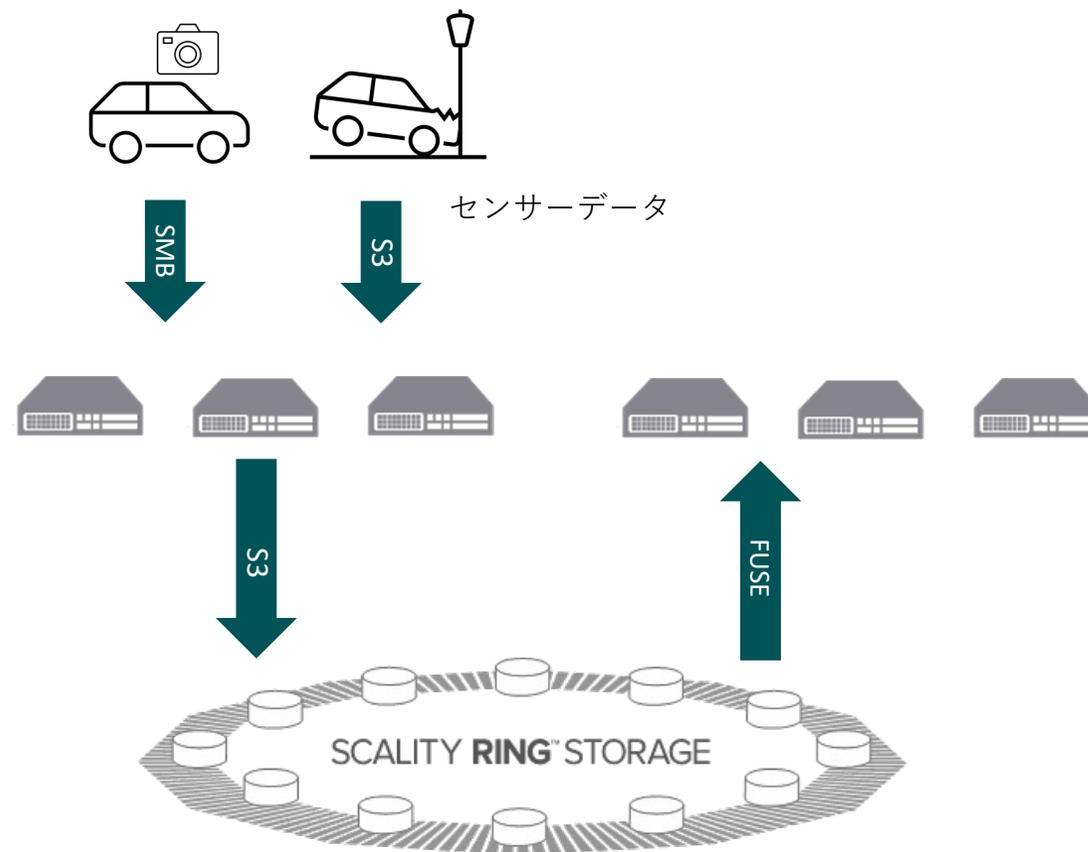
- フランスに本社を置く多国籍メーカー
- 自動運転などの次世代自動車の開発を先進的に取り組む
- 実走行データの格納先としての基盤を模索

■ 悩みと課題

- 高い書き込み性能をもつストレージ基盤
- 優れたコスト効率をもつストレージであること
- 将来的な容量拡張（ペタバイト以上）を実現できること

■ RING導入後の成果

- RINGの採用により利用部門が要求するストレージ性能を実現
- 汎用的なIAサーバーを利用できることにより大幅なGB単価低減を実現
- 大容量に耐えうるストレージ基盤のため、過去11回の増設をへて実行容量36PBのストレージ基盤に成長
- 当初はファイルプロトコルのみを利用していたが、S3 APIも利用することでシステム全体の最適化を実現



事例紹介:海外某医療機関 (HPC Archive /Data:5PB/S3)

■ 基本情報

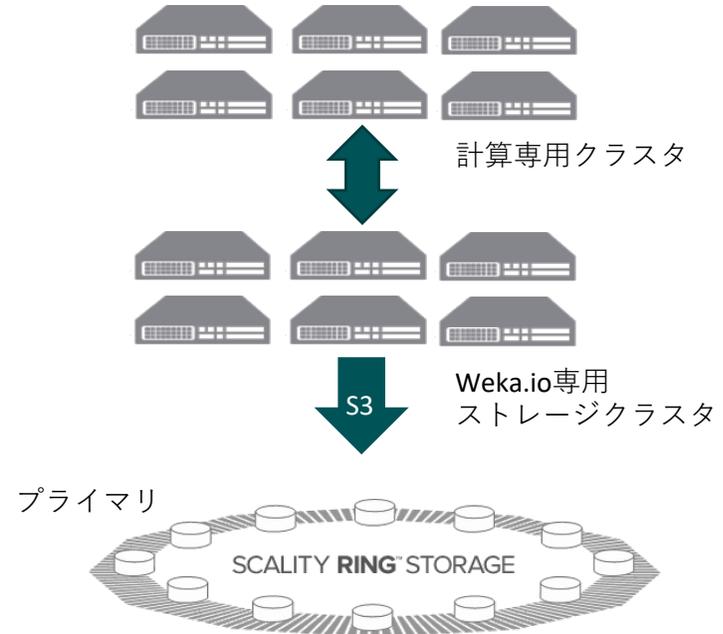
- ボストンに本社を構える米国のバイオ医薬品企業
- 研究開発におけるHPCシステムを保有
- 高速ストレージ領域としてWeka.ioを選定
- 5PBを超える対象データを保有

■ 悩みと課題

- アーカイブ領域として将来の拡張も見据えた大容量なデータ保管ができること
- コスト効率よくデータアーカイブすること
- Weka.ioから透過的にアクセスできること

■ RING導入後の成果

- 大規模な拡張実績を持つRING採用することで将来の拡張計画が可能に
- アーカイブ容量のTB単価の低くすることでコストの最適化を実現
- Weka.ioからの透過的なアクセスの実現により運用コスト低減を実現



事例紹介:海外某民間宇宙開発企業(Telemetry Data Archive/9PB/S3)

■ 基本情報

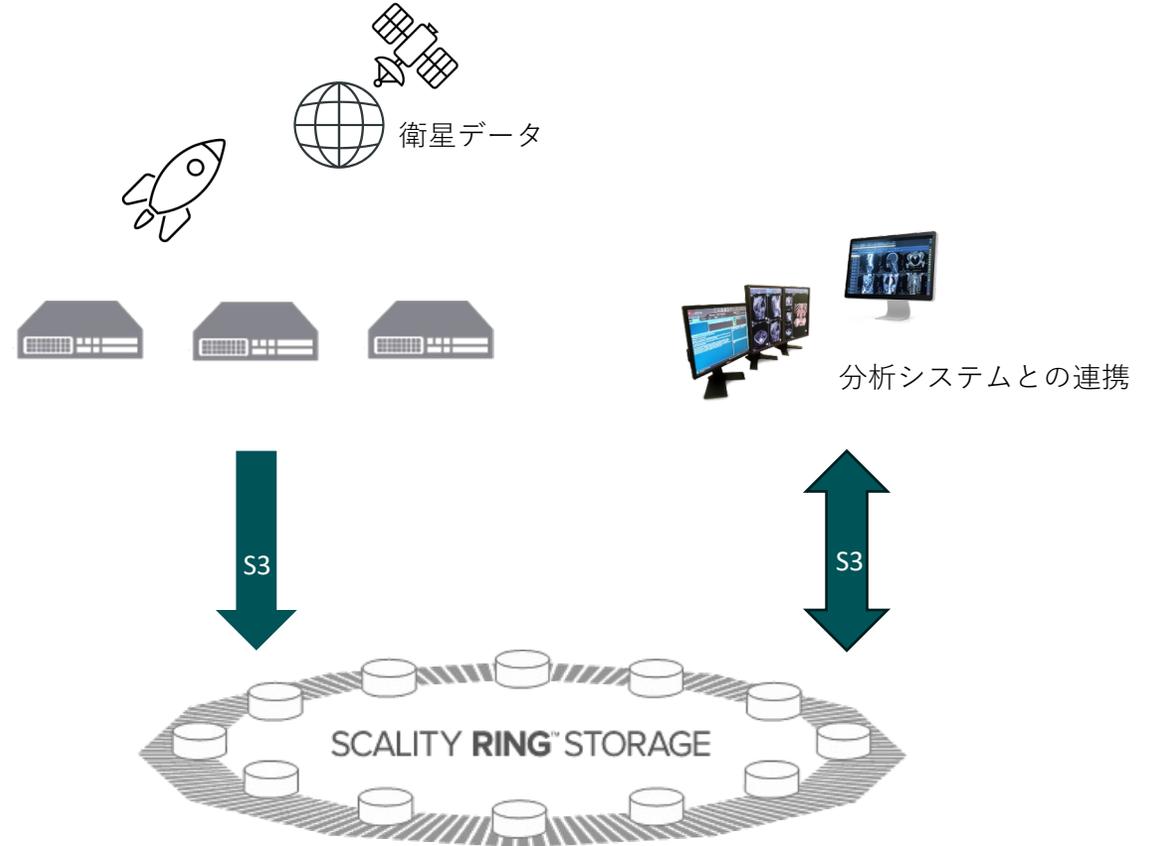
- 米国の某民間宇宙開発企業
- テレメトリーデータ、衛星画像などを保管ならびに解析
- S3 APIを利用したアプリケーションを自社で開発ならびに利用

■ 悩みと課題

- コスト効率のよい（AWS S3より単価の低い）S3準拠のオブジェクトストレージサービスを自社内にて展開したい
- テレメトリーデータ取得時の性能要件
- 他社での豊富な実績（大容量で顧客運用に耐えられるか）
- 将来の容量増加に現実的に対応できる拡張性を持った製品

■ RING導入後の成果

- HW部分に汎用IAサーバーを利用できるため、ストレージコストの低減を実現
- 社内でAWS S3よりも安価なストレージサービスの提供が可能に
- 当初の想定を上回る性能を実現
- 複数回に渡る容量拡張も容易に行なえるため、ストレージを複数セット用意する必要がなくなり運用効率が向上



「データ主権」と言った新たな課題

- 「データ主権」と長期保管
 - 「**データ主権 (data sovereignty)**」は世界中で大きな課題
 - プライベートクラウドにデータを保護
 - 地政学リスク、欧州のGDPR、個人情報の保護
 - データ＝誰が何のために何処に保護するのか
 - シミュレーションやAIにおけるデータセットは知的資産であり社会的源泉



まとめ

最先端シミュレーションとAIにおけるこれからのストレージのポイント

- 1、巨大化するデータの長期的な保管
- 2、柔軟なデータへのアクセスと管理
- 3、データ主権

1. 今後もデータ増加は止まりません、そのため**将来を見据えた拡張性やコスト、災害対策などの考慮が重要**
2. リアルタイムでのデータ収集や広範囲で容易にアクセスできる環境を考え、S3プロトコルのような**インターネットアクセスを主体した設計が必要**
3. シミュレーションやAI分野でもパーソナルデータの取り扱いを考慮すると**データ主権の問題にも適切な対応が重要**



事例集とホワイトペーパー



SCALITY RING
活用事例集

2024



SCALITY RING
Technical White Paper

2024



フォローアップセミナーの実施

下記のとおり、Supercomputing Now! 2024 のフォローアップセミナーを実施します。

本日ご説明したScalityが選ばれる3つのポイントに関する詳細をご紹介します

- 巨大化するデータの長期的な保管
- 柔軟なデータへのアクセスと管理
- データ主権

- 4月10日（水） 12:00 – 12:30
- 4月17日（水） 16:00 – 16:30
- 4月24日（水） 19:00 – 19:30

オンライン形式での実施となります

※いずれも同じ内容にて実施させていただきます。

ご参加登録フォームは[こちら](#)

もしくは右記URLコードからご登録をお願い致します





Thank you