



# Dell EMC PowerEdge サーバと スケールアウト NAS “Dell EMC Isilon” による データレイクソリューション 推奨される構成

Kris Applegate  
ソリューションアーキテクト  
Dell EMC Customer Solution Center  
[Kris.Applegate@dell.com](mailto:Kris.Applegate@dell.com)

Boni Bruno  
主任ソリューションアーキテクト  
Dell EMC 新テクノロジーチーム  
[Boni.Bruno@Dell.com](mailto:Boni.Bruno@Dell.com)

Armando Acosta  
プロダクトマネージャ  
Dell EMC コンバージドプラットフォーム部門  
[Armando.Acosta@Dell.com](mailto:Armando.Acosta@Dell.com)

Sai Devulapalli  
データ分析プラクティスリーダー  
Dell EMC 新テクノロジーチーム  
[Saikumar.Devulapalli@Dell.com](mailto:Saikumar.Devulapalli@Dell.com)

## 要約

本ホワイトペーパーでは、Dell EMC Isilon を Dell EMC PowerEdge サーバに接続するための有効な構成について詳述します。推奨される構成を詳述すると共に、お客様の導入事例に合わせたオプション変更についてのガイダンスも提供します。

2016年12月

## 目次

概要 .....	3
対象者 .....	3
企業における Hadoop .....	4
共有ストレージ Hadoop と分散ストレージ Hadoop .....	4
Dell EMC Isilon .....	5
Dell EMC Isilon X シリーズノード .....	5
Dell EMC PowerEdge .....	5
Dell EMC PowerEdge FX2、PowerEdge FC630、および PowerEdge FD332 .....	6
Dell EMC PowerEdge R630 .....	6
Hadoop の役割 .....	6
計算処理ノード .....	6
インフラストラクチャノード .....	6
マネージャノード .....	7
エッジノード .....	7
推奨される構成 .....	8
モジュール型インフラストラクチャ .....	8
ネットワーク図 .....	8
構成 .....	9
ラックサーバインフラストラクチャ .....	11
ネットワーク図 .....	11
構成 .....	12
テスト済みの構成 .....	13
ネットワーク図 .....	13
構成 .....	14
留意事項 .....	15
計算処理ノードおよび Isilon ノードのサイジング .....	15
Isilon プラットフォーム .....	15
サーバプラットフォーム .....	15
サーバ CPU .....	15
サーバメモリ .....	16
サーバローカルストレージ .....	16
ネットワーク .....	16
Dell EMC Customer Solution Center (CSC) .....	16
リンク .....	16

## 概要

現代の企業において分析は重要な役割を果たしています。規模、スピード、およびプライスポイントにおける洞察を引き出すために必要なデータ処理を可能とテクノロジーの1つが Hadoop です。Dell EMC は、純粋な DIY のリファレンスアーキテクチャからほぼ全てのプロジェクト予算に対応可能なターンキーアプライアンスに至るまで、さまざまなソリューションを提供しています。

本書では、® Dell EMC Isilon アレイおよび ® Dell EMC PowerEdge サーバ上において、Cloudera および Hortonworks Hadoop ディストリビューションを実行するための有効な構成について詳述します。さらに、さまざまな導入事例に対応するために使用可能な潜在的な差異も含め、推奨される構成の概要を説明します。お客様の要望に最も柔軟に適応できるように、モジュールおよびラックサーバ両方の構成について説明します。

## 対象者

本書は、独自の Hadoop クラスタをカスタマイズする際に有効な構成を活用しようとするお客様だけではなく、お客様にソリューションを提案しようとしている Dell EMC セールスマーカ―およびパートナーも対象としています。

## 企業における Hadoop

Hadoop とは、分散コンピューティング環境において大規模なデータセットを保存および処理するように設計されたオープンソースプラットフォームです。これには2つの主なサブプロジェクトがあり、データストレージ用の Hadoop 分散ファイルシステム (HDFS) およびデータ処理用の MapReduce です。Hadoop は、大規模なデータセットをサーバ間または共有ストレージ間で分類し、データを並行して処理します。

企業はビジネスおよびテクノロジー両方の利点を得るべく Hadoop に取り組んでいます。ビジネスレベルでは、Hadoop は TCO の見地から魅力的な価値を提供します。Hadoop は業界標準のサーバおよびストレージを使用する事により、従来のビジネス・インテリジェンス (BI) およびアナリティクス・ソリューションに対し、膨大なデータセットを保存し処理にかかるコストを削減します。さらに Dell EMC Isilon のようなスケールアウト NAS を使用する事により、データレイク経由でコスト効率を向上させる事が可能です。

Hadoop を導入するための有効なリファレンスアーキテクチャ、および拡張性のあるソリューションのポートフォリオを提供しています。Dell EMC は多様な専門的コンサルティングおよびサポートサービスにより、これらの提供サービスに対する支援を行っています。

データ分析を成功させるための“正しい答え”はありません。それは絶え間ない発展への旅です。あらゆる組織のデータは固有のものであり、またそのように扱う必要があります。ある企業にとって完璧なソリューションであっても、別の企業のニーズには対処出来ない場合もあります。

この考えを念頭に置いて、Dell EMC はスターターバンドルおよび有効なリファレンスアーキテクチャから統合アプライアンス、およびエンジニアリングソリューションまたは特定の環境向けに完全にカスタマイズされたソリューションに至るまで、多様なビッグデータおよび分析課題に対処するための幅広い製品とソリューションを提供しています。

## 共有ストレージ Hadoop と分散ストレージ Hadoop

予算やパフォーマンス、キャパシティ、および密度といった、さまざまな要件を考慮した複数の導入モデルを許容出来るという事が、Hadoop の持つ柔軟性の証です。Dell EMC Isilon ソリューションは共有ストレージモデルであり、Hadoop 用の永続的なファイルシステムデータが Isilon NAS クラスタに保存されています、これに対して Hadoop ノード自体のローカルストレージにデータが散在しているのが分散モデルになります。

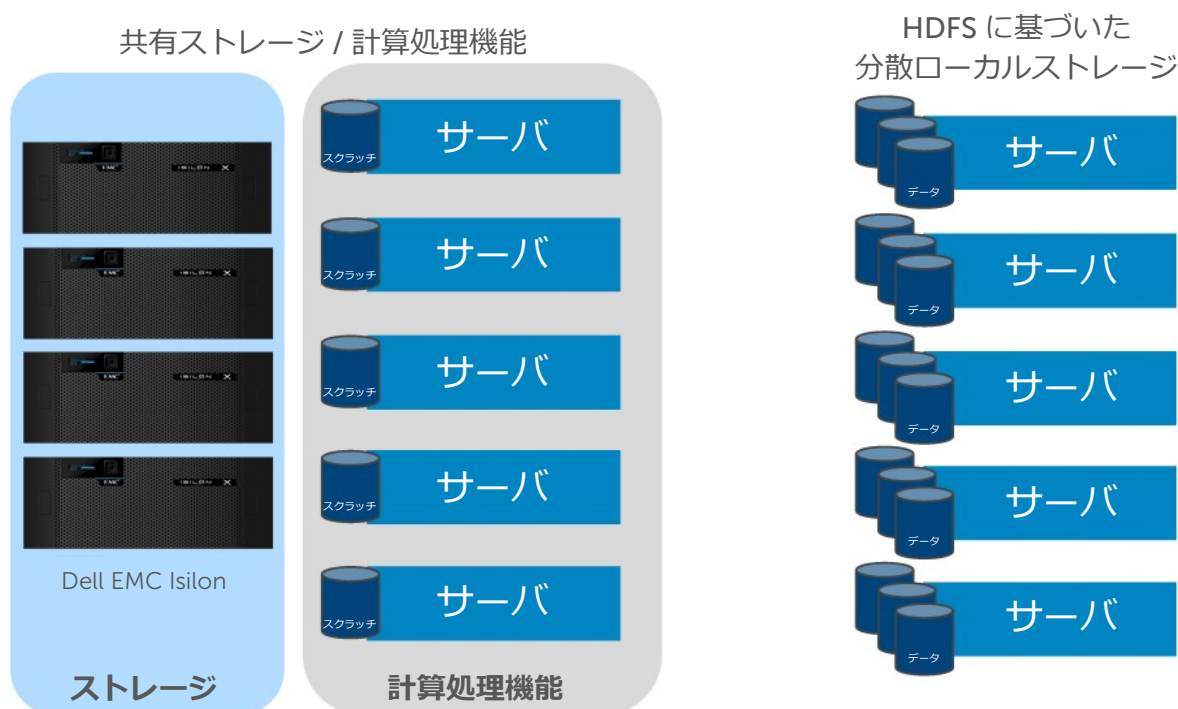


図 1. 共有トポロジと分散トポロジの比較

これら2つのアプローチは、さまざまな利点を提供：

共有ストレージ Hadoop	分散ストレージ Hadoop
単一のデータソース	大規模なスケール (PB+ の数百倍)
データセンターの省スペース化 (ストレージ密度)	一般的なプラットフォーム
ストレージプラットフォーム機能の活用 (パフォーマンス層、代替 RAID、マルチプロトコル)	リニアスケール
ストレージ/計算処理機能の独立したスケーリング	柔軟な複製モデル

表 1. 共有ストレージと分散ストレージの比較

本書では Dell EMC Isilon 構成に重点を置いていますが、Dell EMC はお客様の導入事例に幅広く対応するために、共有モデルおよび分散モデル両方で構築されたソリューションを提供しています。分散ソリューションを詳細に検討されたい場合は、Dell EMC セールsteamまたは Dell EMC Customer Solution Center のソリューションアーキテクトにお問い合わせ下さい。

## Dell EMC Isilon

DELL EMC® Isilon® スケールアウトストレージソリューションは、ストレージではなくデータを管理したい企業向けに設計されています。Isilon のストレージシステムは、実質的にあらゆる規模のインストール、管理、および拡張を強力でありながらも容易に実行可能です。また従来のエンタープライズストレージとは異なりストレージ容量の追加、パフォーマンスの必要性、または将来のビジネスニーズのいかなる変化にもかかわらず、Isilon のソリューションはシンプルのままです。我々は、企業が自身のストレージについて違った観点を持つ事に挑戦しています、なぜなら Isilon という素晴らしく、シンプルなサービスが存在するからです。

### Dell EMC Isilon X シリーズノード

最も柔軟性があり包括的なストレージ製品ラインである Isilon X シリーズは、大容量ストレージと高性能ストレージの適切なバランスを兼ね備えています。非常に汎用的な X シリーズは、ハイスループットおよび高度な並行処理アプリケーション用に理想的なソリューションです。Isilon X シリーズは、ファイルシステムのメタデータおよびファイルベースのストレージワークフロー向けに SSD テクノロジーを導入しているため、ネームスペースの集中的な運用を大幅に高速化します。Isilon は厳しいデータセキュリティおよびコンプライアンス要件を満たすために、X シリーズプラットフォームにおいて自己暗号化ドライブ (SED) オプションを備えた保存データ暗号化 (DARE) も提供しています。



図 2. Dell EMC Isilon X シリーズ

## Dell EMC PowerEdge

Dell EMC のサーバポートフォリオは非常に広範囲をカバーしており、Hadoop ソリューションの計算処理機能に関してさまざまなオプションを備えています。価格、密度、および管理機能を中心に、数え切れないほどさまざまな要件に対応する事が可能なモデルであり、すべての可能なオプションをリストアップするには時間が掛かり過ぎます。そこで、モジュラーインフラストラクチャポートフォリオおよび従来のラックインフラストラクチャポートフォリオという2つの推奨される構成を説明していきます。最初の出発点としてこれらを導入し、お客様の要望に合う仕様にカスタマイズするために、Dell EMC サーバのスペシャリストと連携して下さい。

## Dell EMC PowerEdge FX2、PowerEdge FC630、および PowerEdge FD332

PowerEdge FX2 ファミリーは、あらゆるワークロードの要求を満たせるように調整可能な設定オプションを備えた、完全なモジュラー型エコシステムです。Isilon データレイクの設計ではサーバから Isilon までと同様に、サーバからクライアントまでの堅牢なネットワーク機能だけでなく、柔軟な内部ストレージオプションを兼ね備えた機能を必要とします。PowerEdge FD332 ディスクシェルフを備えた Dell EMC PowerEdge FC630 の計算処理ノードは、容易に管理可能な優れたノードです。

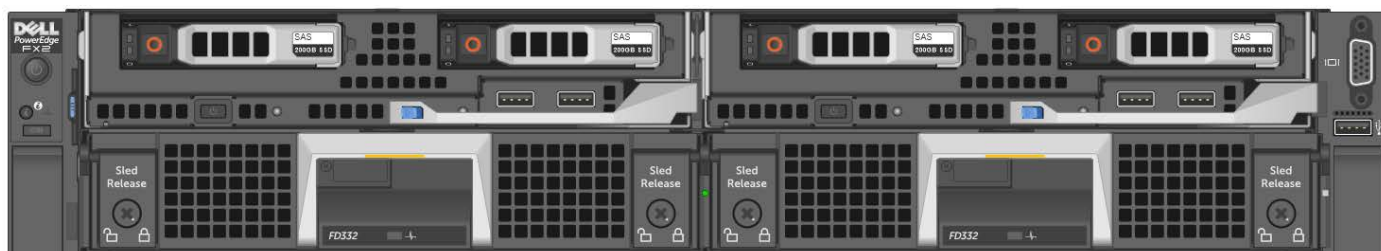


図 3. Dell EMC PowerEdge FX2、PowerEdge FC630、PowerEdge FD332

## Dell EMC PowerEdge R630

最も普及しているサーバプラットフォームとして、R630 はあらゆる導入事例における試練に可能な限り耐えています。この構成では、回転型のメディアまたはソリッド・ステート・ドライブ向けの十分なドライブスロットと、十分なネットワーク帯域幅（データおよびクライアントと直接対応）を活用しています。



図 4. Dell EMC PowerEdge R630

## Hadoop の役割

### 計算処理ノード

全ての共有ファイルシステムの運用は Isilon によって管理されているため、ノードの主な役割はあらゆるデータを綿密にチェックするための計算処理能力を提供することになります。ただし、これらの動作をキャッシュまたは高速化するためには、ローカルストレージを必要とします。ここ数年間にわたるフラッシュの大幅なコスト削減により、一部のお客様はソリッド・ステート・ドライブ (SSD) で構成されるローカルスペースを選択します。SSD の導入は困難な要件ではなく、SSD の価格が下がるにつれ一般的な要件になってきています。

機能	ディスク	タイプ
オペレーティングシステム	2	RAID 1 (ミラー)
Spark Scratch / Map Reduce Spill	2-10 (オプションで SSD)	Non-RAID または RAID 0

表 2. データノードのディスクレイアウト

### インフラストラクチャノード

インフラストラクチャサーバの数は、顧客によって異なります。推奨される構成では4つのノードを割り当てますが、サービスの高可用性の要件が異なるので減らすことも可能です。

## マネージャノード

クラスタ内のマネージャノードは、Cloudera Manager (Cloudera Hadoop)、Ambari (Hortonworks Hadoop)、および Hive、Oozie、Zookeeper のようなサービス向けの主な役割に対する実行の責任を負っています。ノードに障害が発生した場合、高可用性のためにクォーラムを提供するには、上記の3つが必要となります。これらのボックスでは高性能な構成は必要とされず、コスト最適化が優先されます。推奨される構成のために、一般的なプラットフォームを維持するよう計算処理ノードと同じシャーシやサーバタイプを使用しますが、必ずしも必要というわけではありません。また要件によっては、これらの役割を計算処理ノードまたはエッジノードの同一場所に配置することも可能です。

## エッジノード

エッジノードの役割は、データをクラスタ内に送り込んだり、データ結果をクラスタから送り出したりする主要なインターフェースとなります。Isilon ネットワークとデータセンターネットワークは、ほとんどの場合マルチホームになっています。これらのノード構成は、お客様の導入事例により大きく異なる場合があります。例えば、バッチジョブをクラスタ内に段階分けする場合、そのデータを HDFS にコピーする前に、それを保存するための膨大なローカルストレージを必要とします。クラスタ内にデータをストリーミングする場合、膨大なスペースは必要ありませんが、SSD のような高速ストレージを使用すればデータをすばやく移動させる事が可能です。マネージャノードとよく似ていますが、これは導入事例に応じた最適化を優先していきます。我々が推奨する構成は、プラットフォームの共通性を保つためにマネージャノードと同じ構成を維持する事です。最後に、マネージャノードと同様に導入事例に余裕があれば、この役割を計算処理ノードまたはマネージャノードに配置する事も可能です。



## 推奨される構成

### モジュール型インフラストラクチャ

#### ネットワーク図

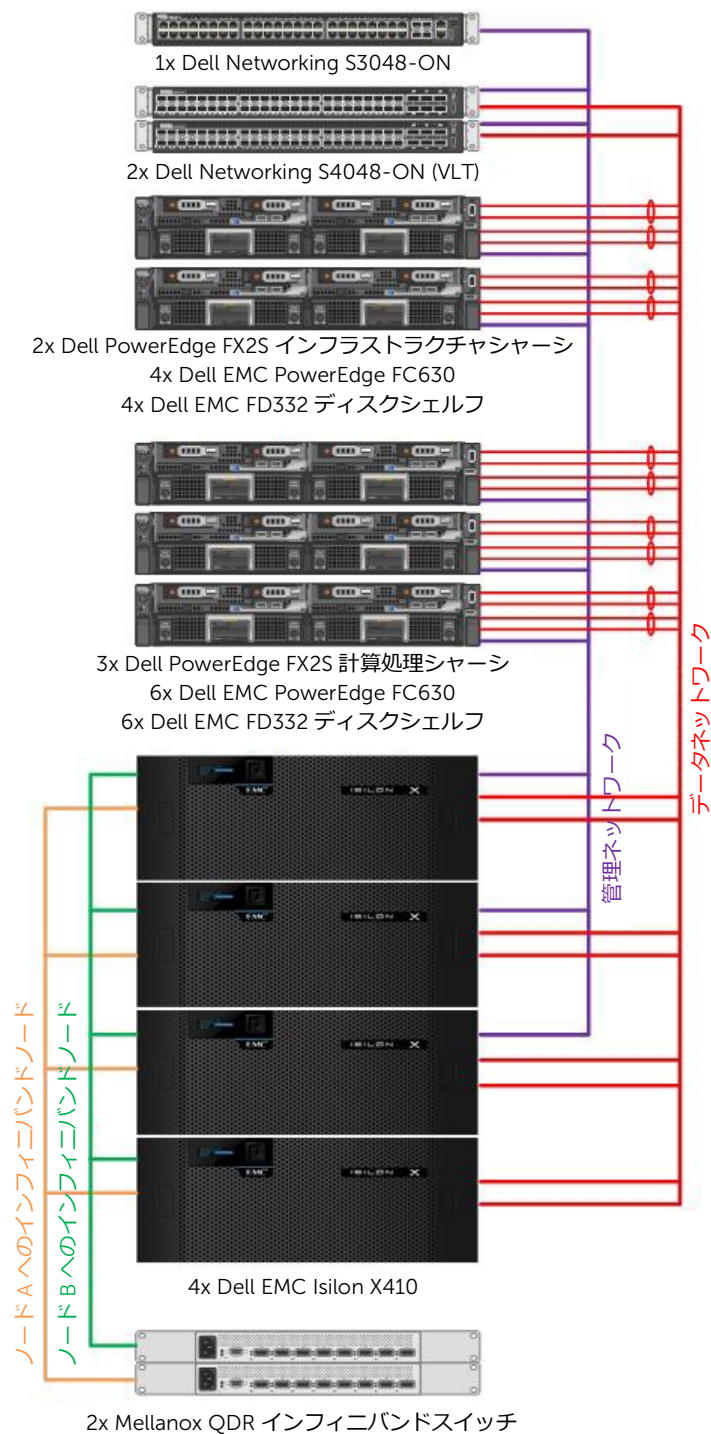


図 5. モジュール型インフラストラクチャ - ネットワーク図



## 構成

Isilon データレイクアレイ	
Isilon ノード	4x Dell EMC Isilon X410 102TB HDD / 3.2TB SSD 256GB 2x10GE および 2x1GE
Isilon スイッチ	2x QDR IB スイッチ - 8ポート、1U、1PS

表3. モジュール型インフラストラクチャ構成 - Isilon データレイクアレイ

ネットワーク	
データネットワーク スイッチ	2x Dell EMC Networking S4048-ON 10GbE スイッチ
管理ネットワーク スイッチ	1x Dell EMC Networking S3048-ON 1GbE スイッチ

表4. モジュール型インフラストラクチャ構成 - Networking

計算処理シャーシ	
計算処理シャーシ	3x Dell EMC PowerEdge FX2s
シャーシ I/O モジュール	2x (シャーシ毎) Dell EMC FX2 10 GbE パススルーモジュール
計算処理 プラットフォーム	2x (シャーシ毎) Dell EMC FC630、2x 2.5インチディスクスロット付き
計算処理 ストレージ	2x (シャーシ毎) Dell EMC FD332、16x 2.5インチディスクスロット付き

表5. モジュール型インフラストラクチャ構成 - 計算処理シャーシ

計算処理サーバ	
計算処理 プラットフォーム プロセッサ	2x (スレッド毎) Intel Xeon E5-2698v4 (20C)
計算処理 プラットフォーム メモリ	256 GB (スレッド毎) - 16x 16GB 2400MHz RDIMM
計算処理 プラットフォーム ディスク	(OS) - 2x (スレッド毎) 200GB ブート MLC 2.5インチ Intel S3610 ソリッド・ステート・ドライブ
計算処理 プラットフォーム ネットワークカード	1x (スレッド毎) Intel X710 デュアルポート 10GbE ネットワーク付属カード

表6. モジュール型インフラストラクチャ構成 - 計算処理シャーシ

計算処理ストレージシェルフ	
計算処理 ストレージシェルフ	8x (スレッド毎) 1.2TB 10K RPM 2.5インチ HDD

表7. モジュール型インフラストラクチャ構成 - 計算処理ストレージシェルフ

インフラストラクチャノードシャーシ	
インフラストラクチャシャーシ	2x Dell EMC PowerEdge FX2s
インフラストラクチャシャーシ I/O モジュール	2x Dell EMC FX2 10 GbE パススルーモジュール

表 8. モジュール型インフラストラクチャ構成 – インフラストラクチャノードシャーシ

インフラストラクチャノードサーバ	
インフラストラクチャノードプラットフォーム	2x (シャーシ毎) Dell EMC FC630、2x 2.5インチディスクスロット付き
インフラストラクチャノードプロセッサ	2x (スレッド毎) Intel Xeon E5-2640v4 (10C)
インフラストラクチャノードメモリ	128 GB (スレッド毎) - 8x 16GB 2400MHz RDIMM
インフラストラクチャノードディスク	(OS) – 2x (スレッド毎) 200GB ブート MLC 2.5インチ Intel S3610 ソリッド・ステート・ドライブ
インフラストラクチャノードネットワークカード	1x (スレッド毎) Intel X710 デュアルポート 10GbE ネットワーク付属カード

表 9. モジュール型インフラストラクチャ構成 – インフラストラクチャノードサーバ

インフラストラクチャノードストレージシェルフ	
インフラストラクチャノードストレージシェルフ	3x (スレッド毎) 1.2TB 10K RPM 2.5インチ HDD

表 10. モジュール型インフラストラクチャ構成 – インフラストラクチャストレージシェルフ

## ラックサーバインフラストラクチャ

### ネットワーク図

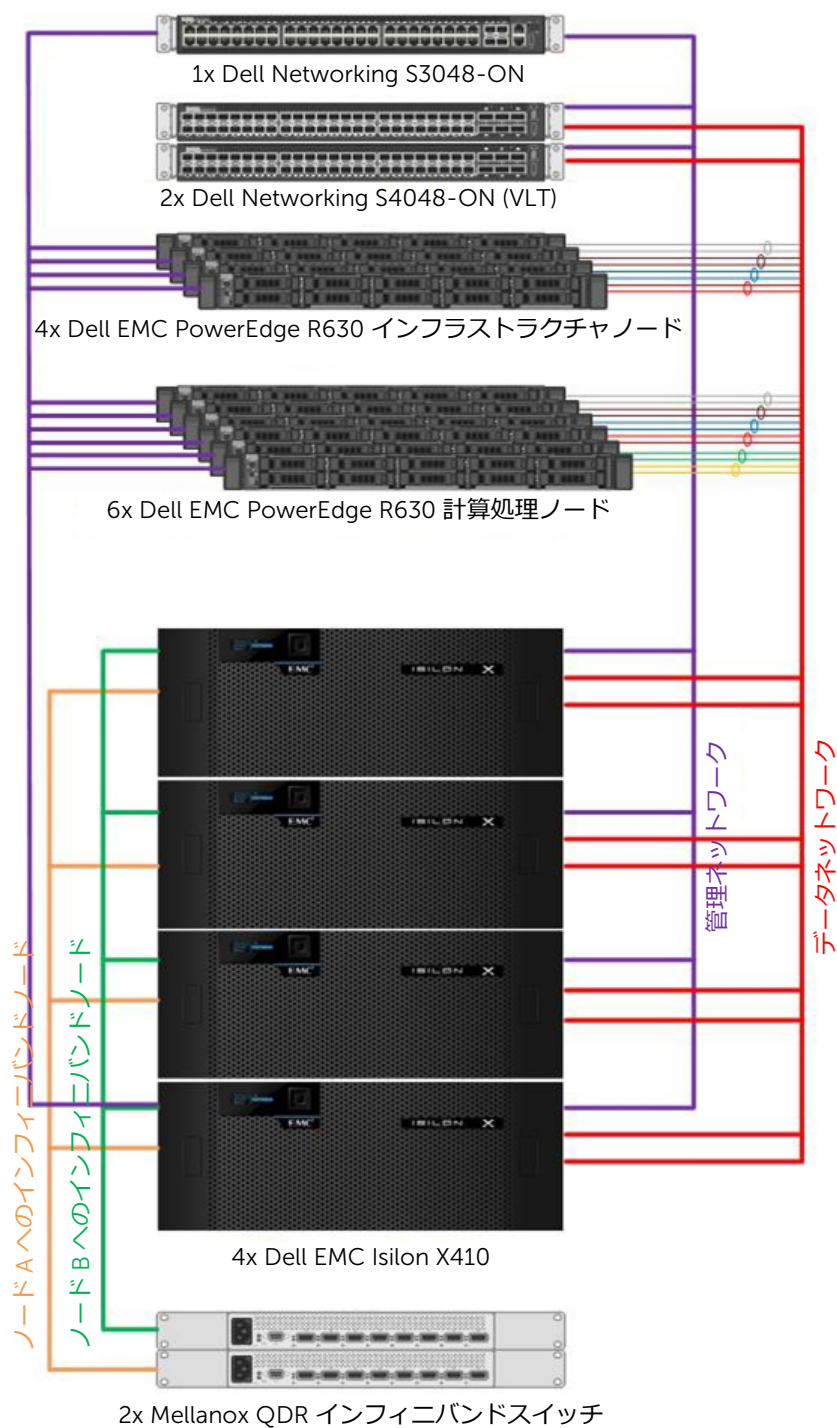


図 6. ラックサーバインフラストラクチャ - ネットワーク図

## 構成

Isilon データレイクアレイ	
Isilon ノード	4x Dell EMC Isilon X410 102TB HDD / 3.2TB SSD 256GB 2x10GE および 2x1GE
Isilon スイッチ	2x QDR IB スイッチ – 8ポート、1U、1PS

表 1 1. ラックサーバインフラストラクチャ構成 - Isilon データレイクアレイ

ネットワーク	
データネットワーク スイッチ	2x Dell EMC Networking S4048-ON 10GbE スイッチ
管理ネットワーク スイッチ	1x Dell EMC Networking S3048-ON 1GbE スイッチ

表 1 2. ラックサーバインフラストラクチャ構成 – Networking

計算処理サーバ	
計算処理 プラットフォーム	6x Dell EMC PowerEdge R630 10ドライブシャーシ
計算処理 プラットフォーム プロセッサ	2x Intel Xeon E5-2698v4 (20C)
計算処理 プラットフォームメモリ	256 GB - 16x 16GB 2400MHz RDIMM
計算処理 プラットフォーム ディスク	(OS) – 2x 200GB ブート MLC 2.5インチ Intel S3610 ソリッド・ステート・ドライブ (データ) – 8x 1.2TB 10K RPM 2.5インチ HDD
計算処理 プラットフォーム ネットワーク付属カード	Intel X710 デュアルポート 10GbE ネットワーク付属カード

表13. ラックサーバインフラストラクチャ構成 – 計算処理サーバ

インフラストラクチャノードサーバ	
インフラストラクチャ ノードプラットフォーム	4x Dell EMC PowerEdge R630 10ドライブシャーシ
インフラストラクチャ ノードプロセッサ	2x Intel Xeon E5-2640v4 (10C)
インフラストラクチャ ノードメモリ	128 GB - 8x 16GB 2400MHz RDIMM
インフラストラクチャ ノードディスク	(OS) – 2x 200GB ブート MLC 2.5インチ Intel S3610 ソリッド・ステート・ドライブ (データ) – 3x 1.2TB 10K RPM 2.5インチ HDD
インフラストラクチャ ノードネットワーク付属 カード	Intel X710 デュアルポート 10GbE ネットワーク付属カード

表 1 4. ラックサーバインフラストラクチャ構成 – インフラストラクチャノードサーバ

## テスト済みの構成

以下の構成は、基本的な機能を検証するために Dell EMC でテストされた内容を文書化したものです。Customer Solution Center が提供するものと同じ機能を利用する事により、お客様自身が概念実証を無償で実行出来るようにしています。

計算処理およびインフラストラクチャの役割は、同じノード間で共有されていました。これは運用環境では推奨されませんが、概念実証のような小規模環境では容認されます。

## ネットワーク図

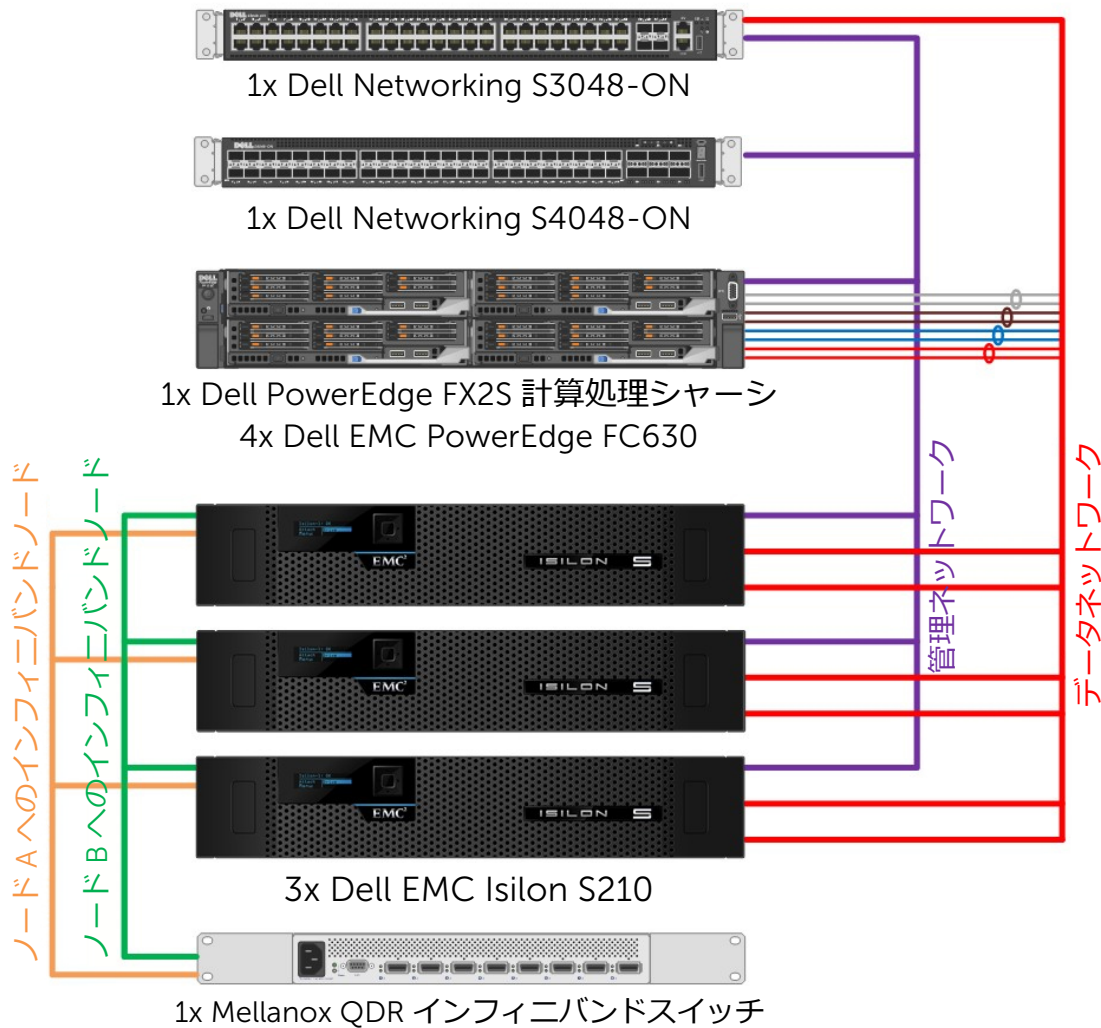


図7. テスト済みの構成 – ネットワーク図

## 構成

Isilon データレイクアレイ	
Isilon ノード	3x Dell EMC Isilon S210 19.8TB HDD / 1.6TB SSD 256GB 2x10GE および 2x1GE (OneFS v8.0.0.2)
Isilon スイッチ	1x QDR IB スイッチ – 8ポート、1U、1PS

表 15. テスト済みの構成 - Isilon データレイクアレイ

ネットワーク	
データネットワーク スイッチ	1x Dell EMC Networking S4048-ON 10GbE スイッチ
管理ネットワーク スイッチ	1x Dell Force10 S60 1GbE スイッチ

表 16. テスト済みの構成 – ネットワーク

計算処理シャーシ	
計算処理シャーシ	1x Dell EMC PowerEdge FX2
シャーシ I/O モジュール	2x (シャーシ毎) Dell EMC FX2 10 GbE パススルーモジュール
計算処理 プラットフォーム	4x (シャーシ毎) Dell EMC FC630、8x 1.8インチ SSD スロット付き

表 17. テスト済みの構成 – 計算処理シャーシ

計算処理サーバ	
計算処理 プラットフォーム プロセッサ	2x (スレッド毎) Intel Xeon E5-2680v3 (12C)
計算処理 プラットフォーム メモリ	256GB (スレッド毎) - 16x 16GB 2400MHz RDIMM
計算処理 プラットフォーム ディスク	(OS) – 2x (スレッド毎) 480GB Intel S3610 MLC 1.8インチソリッド・ステート・ドライブ (データ) – 6x (スレッド毎) 480GB Intel S3610 MLC 1.8インチソリッド・ステート・ドライブ
計算処理 プラットフォーム ネットワークカード	1x (スレッド毎) Intel X520k デュアルポート 10GbE ネットワーク付属カード
計算処理 プラットフォーム オペレーティング システム	RedHat Enterprise Linux 7.2.1511

表 18. テスト済みの構成 – 計算処理シャーシ

## 留意事項

### 計算処理ノードおよび Isilon ノードのサイジング

多くの異なる要因がクラスタのサイジングになります。多くの変数を適切に考慮している事を確認するためには、Dell EMC アカウントチームおよび Dell EMC Customer Solution Center のソリューションアーキテクトと連携する事が重要です。考慮する必要がある変数は以下になります。

- 初期データ量
- 複製の数
- 摂取率
- 保持期間
- スクラッチスペース
- 圧縮
- リード/ライト I/O ミックス

Isilon のノード数に対する計算処理ノード数の初期ガイダンスは、2 : 1の比率になっています。ただし、これは初期のガイダンスに過ぎず、容量、パフォーマンス、および Isilon が提供する可能性がある追加機能への要件を考慮して、より具体的に調整された推奨事項を提案可能な Customer Solution Center と検討する事を強くお勧めします。

### Isilon プラットフォーム

Isilon のクラスタはファイルシステム、ボリュームマネージャ、およびデータ保護を EMC Isilon OneFS® オペレーティングシステムに組み合わせる事によりストレージを簡素化します。EMC Isilon の高性能 X シリーズノード、大容量 NL シリーズ、および高密度 HD シリーズノードによるクラスタ化された使用によって、単一の Isilon クラスタには最高の経済性、スループット、または1秒あたりの I/O 数をペタバイトの範囲で提供する複数の層が混在している場合があります。Isilon クラスタによるストレージの利用率が 80% を超えるため、大抵のストレージシステムと比較して raw 容量が少なくて済みます。従来のダイレクトアタッチトストレージ (DAS) Hadoop と比較すると、Isilon はより安全な一方で、ストレージ容量は3分の1で済みます。非構造化データを Isilon に統合すると、効率性の向上、管理の簡素化、およびコストの削減が可能になります。

### サーバプラットフォーム

Dell EMC PowerEdge のポートフォリオ内部における計算処理ノードおよびインフラストラクチャノードは、多くのオプションを備えています。上記で2つの推奨される構成を詳述しましたが、Dell EMC Customer Solution Center のソリューションアーキテクトと検討可能なオプションも数多く用意しています。下記のオプションをご覧ください。

**PowerEdge ラック/タワーサーバ** – R および T シリーズサーバは、従来の 1U および 2U オプションを必要としているお客様にとって最も一般的なオプションの1つです。高密度向けには1Uの PowerEdge R630、ドライブオプション向けには PowerEdge R730/XD のどちらかを選択するのが最適です。

**モジュラーサーバ** – 堅牢な管理機能および統合ネットワークを必要とするお客様には、Dell EMC モジュール型インフラストラクチャのポートフォリオをご用意出来ます。Dell EMC PowerEdge M1000 ブレードシャーシおよび Dell EMC PowerEdge FX ファミリーは優れた選択肢です。必要なローカルストレージ/スクラッチスペースに対応可能なドライブスロット、またはディスク容量が十分にある事を確認してください。これらは高度なデータセンター密度が必要なインシデント (コロケーション/ホスティング) にも適しています。

### サーバ CPU

お客様によって、サーバコアおよび周波数要件は大きく異なる可能性があります。お客様固有のワークロードに適切なプロセッサを識別するためには、Dell EMC Customer Solution Center のソリューションアーキテクトと緊密に連携することをお勧めします。期待されるパフォーマンスの正確な特性を把握するためには、Customer Solution Center における概念実証を実行できる機能が無償で利用することも可能です。



## サーバメモリ

サーバ CPU と同様に、各お客様や導入事例によっても異なります。一般的には 256GB から開始し、インメモリテクノロジー (Spark、Impala、Alluxio など) の利用が増加するにつれて、メモリを増やす事をお勧めします。

## サーバローカルストレージ

計算処理ノードにはホスト側のキャッシュ/スクラッチスペースが必要になります。大体 5-8TB が、回転型のメモリまたはフラッシュメモリに共通しています。使用可能な Hadoop 容量の約 25% に相当する十分なスクラッチスペースを計算処理ノードに用意する必要があります。フラッシュメモリの価格が急速に下落するにつれて、増加し続ける容量を処理する高速なローカルストレージを得るために、これらのテクノロジーを利用する事は当然の事です。SSD を選択した場合、このローカルスクラッチスペースはドライブベイまたは PCI-E フォームファクタいずれかの SSD になる可能性があります。

## ネットワーク

少なくとも、各ホストから Isilon データノードまでデュアル 10GbE が必要となります。帯域幅の必要性が増すにつれて、フロントサイド (クライアントから計算処理ノードまで) を独自のネットワークカードに分割するか、各ノードへのリンク数および/または速度を増加するかを検討する必要があります。25GbE カードと 40GbE カードの価格は非常に手頃になってきており、複雑さを軽減するために (複雑なボンディングの必要無し) 早期に投資する事を検討したり、絶えず増加する新たなワークロードに対する帯域幅のニーズへの準備を検討してもいいかもしれません。Dell EMC Networking S6100 スイッチは、ホストレベルまたは複数のラックを1つにまとめたアグリゲーション層で高い帯域幅のニーズに対応する優れたスイッチです。

Dell EMC Isilon 製品が進化するにつれて、40GbE Networking への投資は計算処理ノード接続とデータノード間の接続に非常に適している事にも注目してください。

## Dell EMC Customer Solution Center (CSC)

Dell EMC Customer Solution Center (CSC) は、お客様によるソリューションの設計、検証、および構築を支援出来るよう接続可能な世界的なネットワークです。各地域に複数配置されているため、単純なハードウェアプラットフォームから複雑なソリューションに至るまであらゆるものに対応可能です。これらの契約は、非公式な30-60分間のブリーフィングからより長い半日のワークショップ、および顧客が全面的に同意する前にソリューションの検証を行う概念実証にまで及びます。お客様はアカウントチームと連携する事により、無料でこれらのサービスを利用出来るリクエストを提出すること可能です。

## リンク

Dell EMC Customer Solution Center - <http://www.dell.com/customersolutioncenter>

Dell EMC FX PowerEdge Server FX アーキテクチャ - <http://www.dell.com/en-us/work/learn/fx-server-solutions>

Dell EMC Isilon Hadoop 向けハブ情報 - <https://community.emc.com/docs/DOC-39529>

Isilon Hadoop ツール - [https://github.com/Isilon/isilon\\_hadoop\\_tools](https://github.com/Isilon/isilon_hadoop_tools)

Cloudera - <http://cloudera.com>

Hortonworks - <http://hortonworks.com>