

ホワイトペーパー

Data Domain と IDPA (Integrated Data Protection Appliance) の経済的価値

Dell EMC バックアップ アプライアンス ソリューションのコスト効率の
検証

Vinny Choinski (上級 ESG ラボ アナリスト)

Christophe Bertrand (上級アナリスト)

2018 年 6 月

この ESG ホワイトペーパーは Dell EMC の委託を受けて作成されたものであり、
ESG から使用許諾を受けて配布されます。

目次

目次.....	2
はじめに	3
Data Domain と IDPA アーキテクチャのメリット	4
Data Domain と IDPA の経済的なメリット	5
重要な重複排除の効率性	6
スピードと拡張性	8
コスト削減の留意事項.....	9
より大きな真実.....	10

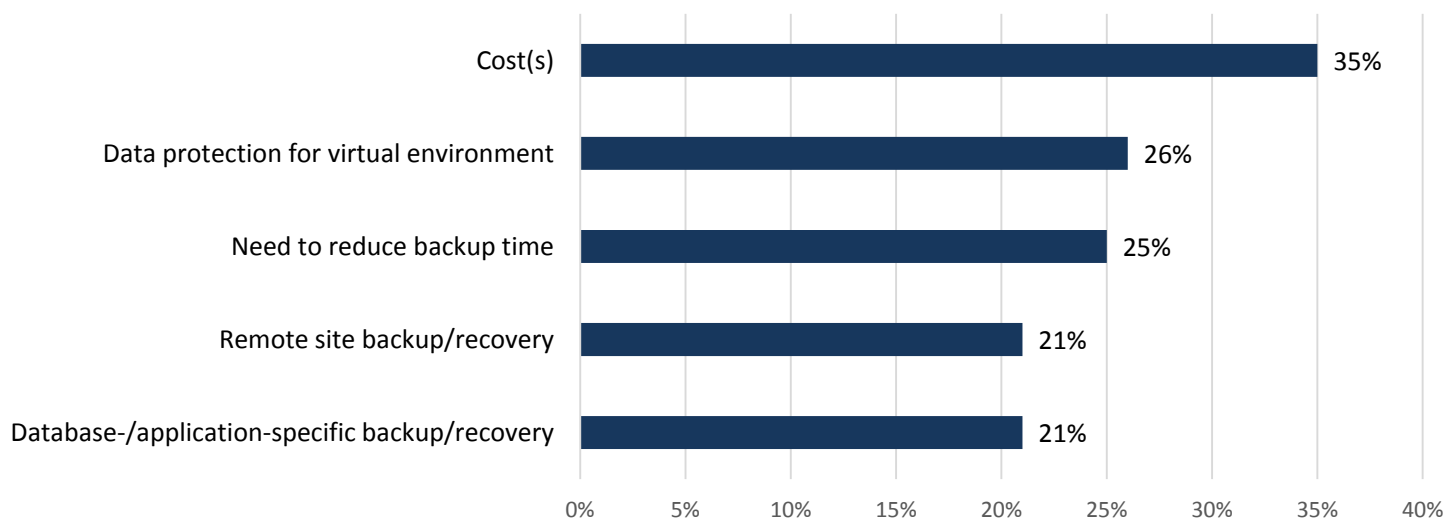
はじめに

競争に先んじてイノベーションを進めるには、一貫した方向性を維持し、インフラストラクチャをモダンに保つ必要があります。新しいビジネス イニシアティブに最適化されたインフラストラクチャを提供し、コストを抑え、削減もしながらアプリケーションをサポートすることが、IT 部門のプレッシャーになっています。事実、ESG が行っている継続的な調査では、IT 支出に影響を及ぼす上位のビジネス促進要素のひとつとして、コスト削減が常に回答者から挙げられています。2017 年の調査では、コストを抑えるために組織がどのような対策をとる予定かという質問に対し、回答者の 27% が ROI を改善するために新しいテクノロジーを購入すると回答しています。¹

ある特定のグループに絞ると、データ保護プロセスとテクノロジーの実装を担当する IT 管理者の 35% が、コストを上位の課題として常に挙げています(図 1 参照)。このグループは、ワークロードを中心とする課題、具体的には仮想環境、さらにはリモート オフィスの保護に関する課題を特に問題と認識しているようです。しかしながら、データ保護に関する課題の全体的なランキングを見ると、さらに重要なストーリーが見えてきます。コストと仮想化が主な懸案事項であるものの、それに続く 3 つの考慮事項は、本質的に運用と機能の両方に関連しています。それらはパフォーマンス、分散アーキテクチャのバックアップ/リカバリ、データベースとアプリケーションの保護に関連した課題です。²

図 1 現在のデータ保護プロセスとテクノロジーの課題

Which of the following would you characterize as challenges with your organization's current data protection processes and technologies? (Percent of respondents, N=387)



出典: Enterprise Strategy Group

これらの課題は、高速化、アジリティー、信頼性、コスト管理の向上に関連する、IT リーダーからの要求事項の上位とも一致します。これらのすべての課題が、過去の ESG データ保護モダナイゼーション レポートで一貫して指摘されてきました。ESG の分析により、実装担当者による取り組みと、リーダーが求めることとの相違が、多くの場合、モダナイゼーションやトランスフォーメーションを阻む非常に大きな原因となったことが分かっています。データ保護環境における技術的な障害が原因で、実装担当者はリーダーの要求に応えることができなかったのです。³

¹出典: ESG 調査レポート、『[2017 IT Spending Intentions Survey](#)』(2017 年 3 月)。

²出典: ESG 調査レポート、『[Data Protection Modernization Trends](#)』(2016 年 12 月)。

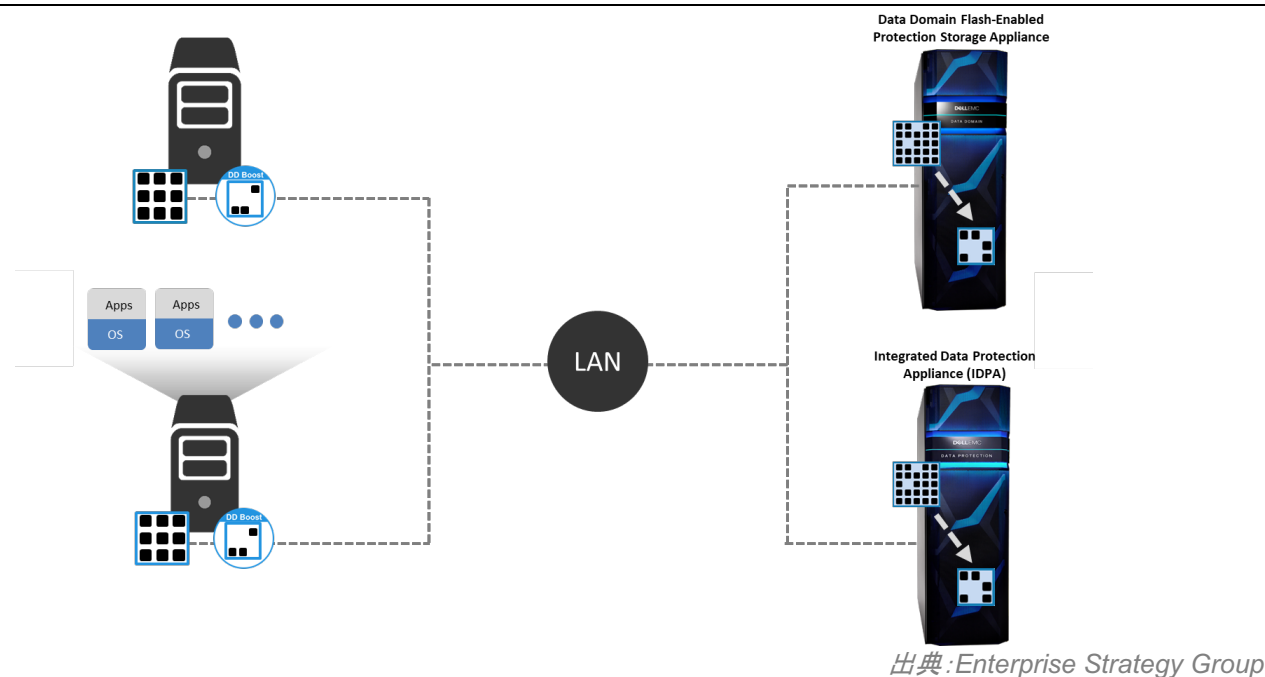
³出典: ESG 調査レポート、『[2015 Trends in Data Protection Modernization](#)』(2015 年 9 月)。

このホワイトペーパーでは、Dell EMC Data Domain システムや Dell EMC IDPA(統合データ保護アプライアンス)などの Data Domain アーキテクチャに構築されたコンバインド ソリューションが、インフラストラクチャを変革する必要がある IT 実装担当者にアジリティーを提供することで、IT リーダーの要求を満たし、現在のデータ保護の課題を解決して、最も重要な課題である保護データを保存するコストの削減をサポートする方法について説明します。ESG の調査結果は、KPI(主要パフォーマンス指標)となる元の購入価格、環境コスト、容量/使用率、パフォーマンスなどの導入されたシステムから得られた実環境のデータの検証と分析に基づいています。

Data Domain と IDPA アーキテクチャのメリット

EMC は 2006 年に既存のデータ保護ソリューションへのデータ重複排除コードの組み込みを回避することを決定しました。これはそのデータ保護戦略を転換する決断でした。この転換は、2 つの主な買収で得たテクノロジーを活用することから始まりました。最初の買収ではソーススペースの重複排除テクノロジーを獲得し、次の買収ではターゲット重複排除テクノロジーを使用した Data Domain を獲得しました。2 つの買収により得たテクノロジーは、EMC のデータ保護ソリューションの基盤となるコンポーネントとなりました。現在は Dell EMC のブランドになっている IDPA を含む Data Domain システムは、フラッシュ対応、完全統合型の専用データ保護アプライアンスです。データを保持および保護するために必要なディスクストレージ量を削減するために設計されています。ソースベースとターゲットベースの両方で、データの重複排除をアーキテクチャにネイティブ統合したこれらのシステムは、より少ない時間でより多くのバックアップを行い、より高速で信頼性の高いリストアを提供して、データ保護に必要なストレージ容量を削減することが可能です。最新世代の Data Domain と IDPA システムは、メタデータ用のフラッシュの導入により、規模に応じたパフォーマンスを実現し、Data Domain と IDPA から仮想マシンへの直接のアクセスおよびリストアが旧世代と比較して 20 倍もの速さで瞬時にできるようになりました。

図 2 Data Domain と IDPA アーキテクチャの概要



Data Domain と IDPA アーキテクチャの主な機能は次のとおりです。

- SISL (Stream-Informed Segment Layout)** : SISL を使うことによって、Data Domain と IDPA システムは重複排除プロセスの 99% を CPU と RAM で実行できるようになります。そのため、システムのパフォーマンスはディスクの数に依存せずに向上します。

- **可変長の区分化**: 重複排除率を最適化するために、Data Domain と IDPA システムは可変長の区分化を利用して、データの構造に基づいてストリームを分割します。これにより、セグメントを圧縮、保存する前に、そのセグメントが一意であるかどうかをシステムが判断できます。
- **インライン重複排除**: Data Domain と IDPA システムでは、バックアップ ストリームをシステムから受信すると、CPU およびメモリで重複排除を実行します。つまり、一意のデータのみがディスクに送信され、保存されることとなります。そのため、ポスト プロセス重複排除のためのディスク ステージング領域とコンピューティング リソースは必要なくなります。
- **Data Domain Boost ソフトウェア**: Data Domain Boost ソフトウェアは、重複排除プロセスの一部をアプリケーション クライアントやバックアップ サーバに分散します。DD Boost により、一意のデータのみがバックアップ サーバまたはクライアント、および Data Domain システム間で転送されます。また、これにより、お客様のネットワーク、Data Domain および IDPA システムのリソースを解放し、ターゲット側の重複排除パフォーマンスを向上させることができます。
- **データ非脆弱性アーキテクチャ**: Data Domain と IDPA は、データを確実にリカバリできるように作られています。データ非脆弱性アーキテクチャは、書き込みと読み取りのインラインの検証を提供し、データの取り込みと回復の際に、データの整合性の問題からシステムを保護し、自動的にリカバリを実行します。継続的な障害検出機能と自動修復機能により、Data Domain と IDPA システムにあるデータはライフサイクル全体を通じていつでもリストアできるようになります。

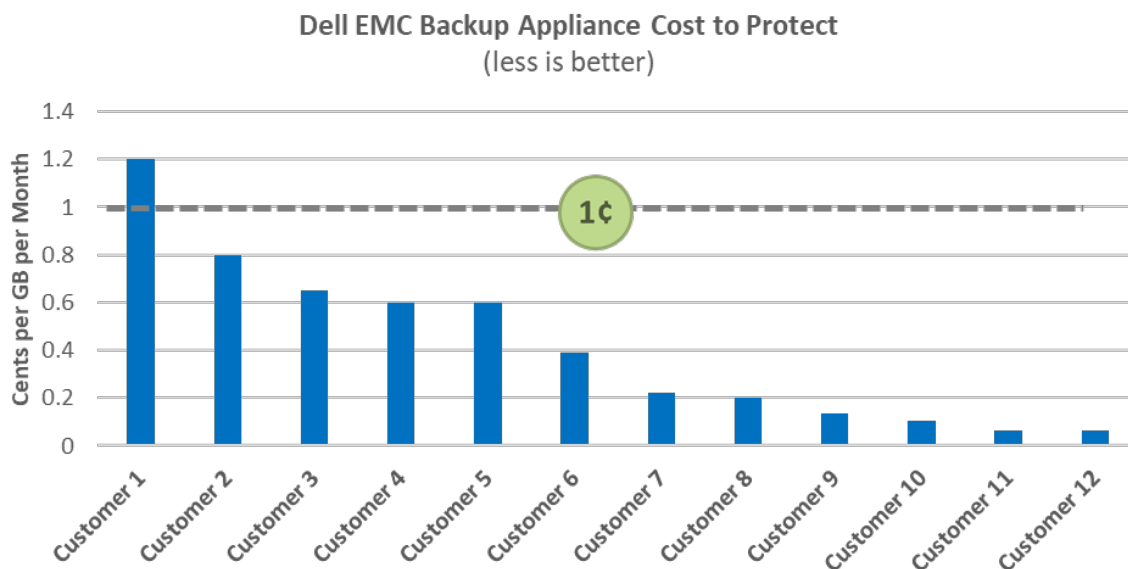
Data Domain と IDPA の経済的なメリット

Data Domain と IDPA アーキテクチャがビジネスの利害関係者の経済的メリットになる理由を詳しく見ていきましょう。データ重複排除機能を使うと、同じ物理ディスク領域により多くのデータを保存できます。これにより、ストレージ容量の要件は少なくなり、コストの削減につながります。ソースまたはクライアント側でデータ重複排除を実装することには、容量の節約だけでなく、バックアップ パフォーマンスが向上するというメリットもあります。ソース サイドの重複排除では、バックアップ中にソースからターゲットに一意のデータブロックのみが送信されるため、ネットワークトラフィックが大幅に削減されます。こうしてネットワーク効率を向上させることにより、既存のネットワーク インフラストラクチャを使用してバックアップ データを増やすことができるようになり、高価なネットワークへのアップグレードの必要性を排除、または延期できます。転送が必要なデータが少ないほど、バックアップのパフォーマンス速度が増すのは明らかです。バックアップ時間が短くなることで、お客様はバックアップの頻度を増やし、組織にとって大きな損失となりえるデータロスリスクを減らすことができます。DD Boost ソフトウェアを搭載した Data Domain と IDPA システムによって、ソースとターゲット両方の重複排除がサポートされ、環境に最適な重複排除機能を柔軟に導入できます。

図 3 に示すように、ハードウェア、ソフトウェア、電源、冷却、重複排除などの実環境データに関する ESG の分析結果によると、Data Domain と IDPA システムでは毎月 1GB あたり 1 セントでデータ保護環境を容易に実現できます。

ESG が調査した 12 社のお客様が保護に費やしたコストは、1 ギガバイトあたり 0.062~1.2 セントの範囲でした。このようにコストの節約量に比較的大きな幅があるのは、Dell EMC の本番環境のお客様により達成される重複排除容量の節約量に相違があることが主な要因です。このレポートの後半では、Dell EMC Data Protection Suite ソフトウェアと Data Domain/IDPA ハードウェアをエンド ツー エンドで組み合わせることで、重複排除により、コストをどの程度まで節約できるかに焦点を当てて詳しく見ていきます。

図 3 Data Domain と IDPA の経済的価値

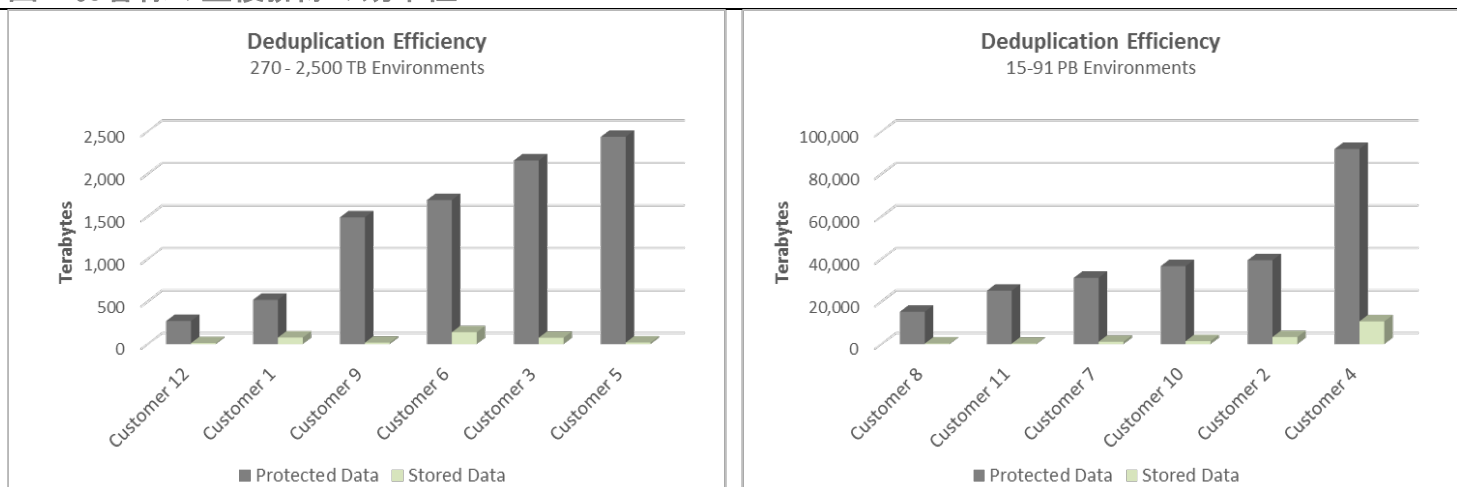


出典: Enterprise Strategy Group

重要な重複排除の効率性

ESG は、Dell EMC の 12 社のお客様から得たオートコール サポート データを検証および分析することにより、Data Domain と IDPA の重複排除容量の節約に関する経済的価値の調査を開始しました。図 4 に示すように、保護対象データのお客様の環境は約 270TB~90PB 以上におよびました。また、お客様の業種はテクノロジー、メーカー、保険、ヘルスケアなど複数にまたがっています。異なる業種のお客様が選ばれたのは、さまざまなタイプの重複排除結果のデータセットを収集するためでした。図 4 の灰色のバーは、各環境で保護されているデータ量を示します。緑色のバーは、重複排除後に保存されるデータ量を示します。

図 4 お客様の重複排除の効率性



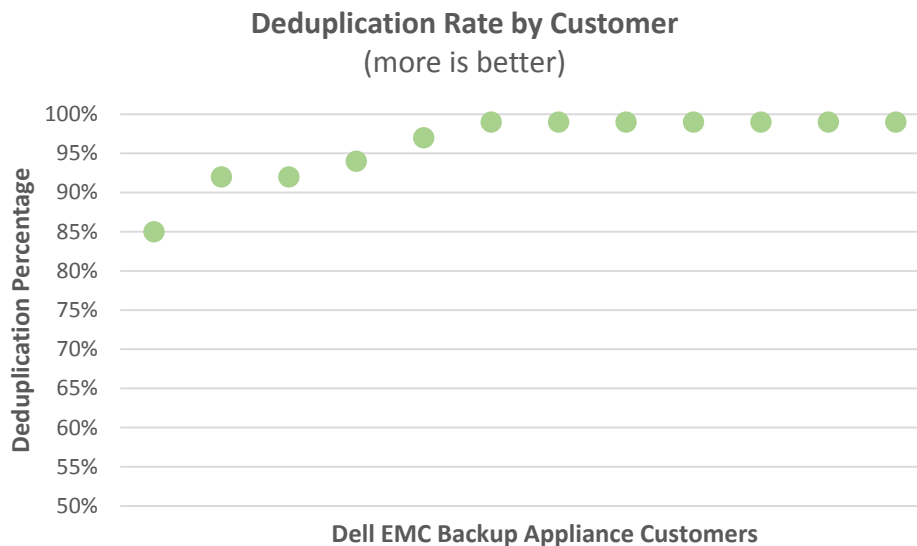
出典: Enterprise Strategy Group

- 重複排除:** 分析したデータで観測された最大重複排除率は 126:1 でした。この結果は、図 4 左側にある最初のデータポイントに該当するお客様の環境から得られたものです。分析が行われたすべてのお客様の平均重複排除率は 41:1 でした。
- 保護:** 約 9:1 という最も低い重複排除率が観測されたお客様でも、わずか 10PB の容量でおよそ 92PB のデータを保護できました。図 4 右側の最後のデータポイントがこのお客様に該当します。このお客様の業種はメーカーに属し、非常に分散した環境を保有しています。

図 5 は、Data Domain と IDPA の重複排除の効率を異なる視点から捉えています。この図は、同じ 12 社のお客様の各環境において実現された重複排除の割合を示します。重複排除率は 85%~99% の範囲です。通常、Data Domain と IDPA ソリューションが環境内でデータを受信する時間が長くなるほど、重複排除率は高くなります。これは、時間の経過とともに同じデータパターンを確認できる傾向が高まり、一意のデータのみを保存するだけで済むようになるためです。

Data Domain と IDPA アーキテクチャの可変長の区分化機能は、このレベルの重複排除を達成する主要なコンポーネントになります。可変長の区分化により、Data Domain と IDPA ソリューションは、保護のためにデバイスに送信されるデータ構造で、データベース タイム スタンプ マーカーを含む自然なパターンをより簡単に一致させることができます。可変長の区分化は、固定長のアーキテクチャよりもはるかに高い重複排除レベルをもたらします。また、単一のストレージプール内でより効果的な拡張性を実現します。これにより、管理するデバイスの数が減り、運用コストの削減につながります。

図 5 お客様の重複排除のパーセンテージ



出典: Enterprise Strategy Group

Data Domain と IDPA のデバイス間でデータをレプリケーションすると、可変長の区分化の重複排除によるメリットがさらに増加します。変更データが少ないほど、レプリケーションするデータは少なくなるためです。これにより、レプリケーションのパフォーマンスが向上するだけでなく、レプリケーション プロセスの WAN 帯域幅の要件を下げることもできます。

- **重複排除率:** 検証および分析された重複排除率は 85%~99%の範囲で、平均は 96%でした。これは、保護のために Data Domain と IDPA システムに保存する必要があるのは、本番データのほんの一部(ベストケースで 1%以下)であることを意味します。
- **90%未満:** 分析されたデータの中で重複排除率が 90%未満だったお客様は 1 社のみでした。繰り返しになりますが、このお客様は分散する保護プールを備えた非常に分散した環境にあることを示しています。
- **90%以上:** 分析が行われた残りのお客様の重複排除率は 90%を超え、ほとんどのお客様の環境で 99%の重複排除率を実現していました。

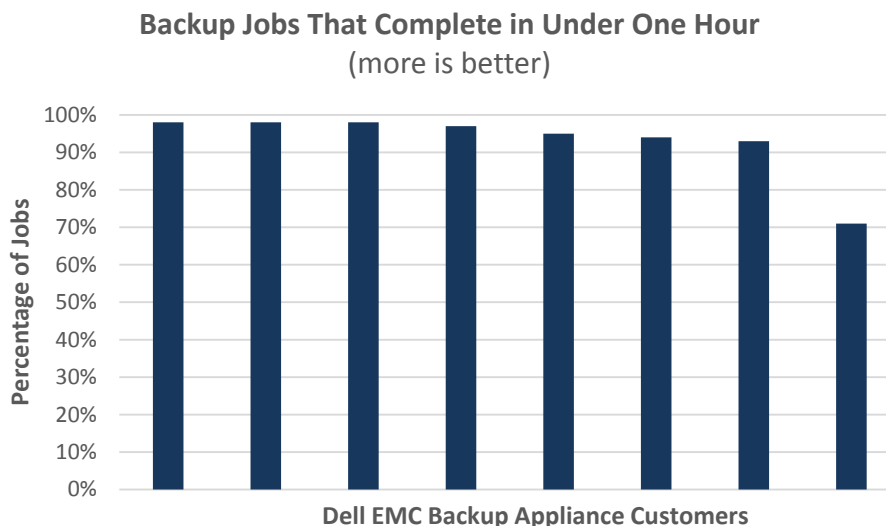
スピードと拡張性

ストレージ容量の拡大が、必ずしも重複排除のパフォーマンスにつながるとは限りません。Data Domain と IDPA では 2 つの主なパフォーマンス要素を活用することにより、この課題に対処しています。SISL (Stream-informed Segment Layout) と DD Boost ソフトウェアです。

SISL は、重複排除プロセスの 99%を CPU と RAM で実行できるようにするための Data Domain と IDPA プラットフォームのアーキテクチャ要素です。つまり、このソリューションでは、パフォーマンスを高めるためにディスク スピンドル数を増やす必要はありません。この結果、より小さいフットプリントのソリューションで、スピンドル数が大きい構成と同じパフォーマンスを達成できます。さらに、Data Domain と IDPA システムは Intel の CPU を使用するため、Intel プロセッサが新しくリリースされるたびにパフォーマンスが向上します。

また、DD Boost ソフトウェアは、重複排除プロセスの一部をアプリケーション クライアントまたはバックアップ サーバに分散させることで、パフォーマンスの実現をサポートします。DD Boost を使用することで、バックアップ サーバまたはクライアントから Data Domain と IDPA システムに移動する必要があるデータは一意のデータのみになります。DD Boost は、移動するデータ量を最大 99%削減します。ESG では、迅速なバックアップ ジョブの完了による総合的なメリットを検証しました。図 6 が示すように、ほとんどのお客様は通常 1 時間以内で定期的にフル バックアップ ジョブを完了しています。

図 6 バックアップ ジョブのパフォーマンス



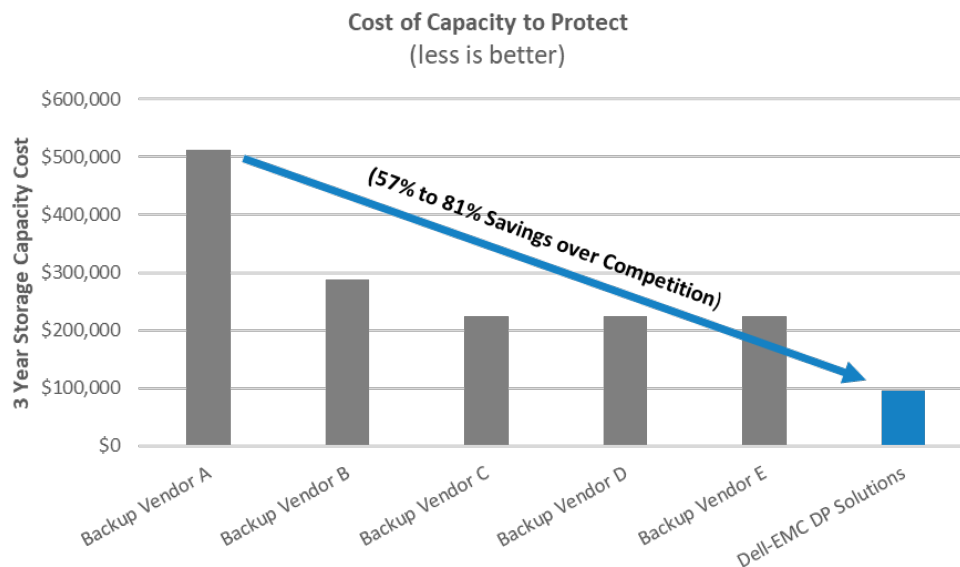
出典:Enterprise Strategy Group

- **15 分未満:**テクノロジー業界のあるお客様では、すべてのバックアップ ジョブの 76%が 15 分未満で完了しました。
- **1 時間未満:**分析対象となった 12 社のお客様のうち 8 社で、すべてのバックアップ ジョブの 70%~98%が 1 時間未満で完了しました(図 6 参照)。
- **4 時間未満:**保険業界のあるお客様では、すべてのバックアップ ジョブが 4 時間未満で完了しました。
- **4 時間超:**2 つの大規模な環境(アプリケーションの数とサイズにおける)では、すべてのバックアップ ジョブを完了するまでにかかった時間は 4 時間以上でした。

コスト削減の留意事項

ESG は、世界中に展開された 15,000 台を超える Dell EMC バックアップ アプライアンスから収集したオートコール サポート データを検証および分析することにより、Data Domain と IDPA の重複排除容量の節約に関する経済的な価値の調査を完了しました。このサンプルは、使用されたバックアップ ソフトウェアと、達成された重複排除率を記録したシステムから収集されました。データは、Dell EMC バックアップ アプライアンスの重複排除の効率が、エンド ツー エンドの Dell EMC データ保護ソリューションにより、さらに向上することを示しています。図 7 に示すように、Dell EMC アプライアンス ハードウェアと Data Protection Suite ソフトウェアを組み合わせることで、他のベンダーのソリューションと比べ、重複排除による削減率が大幅に向上し、3 年間のストレージ容量のコストが 57%~81%削減されました。

図 7 Data Protection Suite により、Dell EMC Data Domain と IDPA のコスト削減率を拡大



出典: Enterprise Strategy Group

Data Domain と IDPA ソリューションには、図 7 に示すストレージ容量のコスト削減を上回るその他の経済的メリットがあります。実際のお客様のデータを ESG が分析した結果、次のメリットが明らかになりました。

- ネットワーク帯域幅の削減:**ほとんどのエンタープライズ アプリケーションと同様に、バックアップおよびリカバリソリューションは、ネットワークリソースに依存して機能します。ネットワークリソースとは、ローカル バックアップとリストア用の LAN リソースおよび DR とビジネス継続性用の WAN リソースの両方を意味します。非効率なデータ保護ソリューションは貴重なネットワークリソースをすぐに消費してしまい、ユーザーの接続性と生産性にも影響を及ぼす場合があります。Data Domain と IDPA の重複排除を活用することで、ある多国籍メーカーはローカルおよびリモートのデータ保護帯域幅の使用率を 98%削減することができました。
- パフォーマンス:**効率的なデータ重複排除は、データ保護パフォーマンス全体の向上に大きな影響を与える場合があります。ある重機メーカーでは、クライアントと Data Domain システム間で転送する必要のあるデータ量を削減できる DD Boost を実装することで、バックアップおよびリストアの速度を 50%向上させることができました。同じお客様の DR への準備状況は 90%改善しました。
- データセンター設置面積の削減:**物理的な場所にかかる費用、特にモダン データセンターのスペース、電源、冷却の要件をサポートするために必要な費用は安くはありません。移転せずにデータセンターを物理的に拡張することは難しく、時には不可能なこともあります。Data Domain により、例えば、全国チェーンのデパートはデータ保護環境の物理テープを排除し、非常に貴重なデータセンターのラックスペースを 3 列フルに空けることができました。

より大きな真実

データ保護に関する IT リーダーからの要求事項で上位を占めるのは、使用中のソリューションの基本的な信頼性とアジリティの向上です。僅差でその次に続くのがコスト削減であり、データ保護実装担当者の間では最優先事項でもあるとみなされています。これらの課題を矛盾する、または相互排他的な要求とみなすべきではありません。実際に、信頼性と機能性ととも、効率性も追求して開発され、改善されたデータ保護ソリューションは、これらのすべての課題に対応できます。

効率性にはさまざまなものがありますが、多くは次のような業務遂行にかかるコストと相互運用性に集約できます。

- データ保護アプリケーションにかかるコストの削減により、経済的に保護ストレージを提供する機能。
- バックアップ ソフトウェア ソリューション間の相互運用性(単一のバックアップ アプリケーションを所有するのは 4 組織のうち 1 組織のみ)。⁴
- 保護ストレージ システムの拡張性と重複排除の効率性。導入、および管理が必要となるシステムの数とコストを削減します。
- 容量が最適化された保護ストレージの同じプールをバックアップおよびアーカイブ データに活用する機能。
- 保護ストレージと多種多様なプラットフォームとの相互運用性(および統合性)。アプリケーション所有者(データベース管理者や vAdmins など)は、独自のツールを活用して補足的な保護/リカバリを実行できるようにしながら、IT 責任者として、一元管理された効率のよい保護ストレージを活用します。
- 統合データ保護ソリューションを使用して導入を迅速化する機能。導入の必要があるシステム数を削減し、オールインワンのアプローチでライセンスをシンプルにし、相互運用性の問題を軽減します。
- 一元化された管理とレポート作成によりデータ保護環境の全体的な管理をシンプルにする機能。アプライアンス内のすべてのデータ保護コンポーネントに対するシングル ステップのアップグレード/パッチを実現して、環境を維持するのに必要な時間を短縮できます。
- クラウド(長期保存、クラウドによるディザスタリカバリ)、検索、分析など、より広範な機能を統合する機能。繰り返しになりますが、導入時間を短縮し、管理をシンプルにできます。

業界の多くの企業が「保護ストレージ」や「重複排除」と同じと捉える機能を提供することにより、Dell EMC はシンプルなデータ保護インフラストラクチャのコンポーネントを提供するベンダーから、本番ストレージ、バックアップ ベンダー、アーカイブ ベンダー、アプリケーション用のプラグ イン アクセラレーター(例: DD Boost)などの要素からなる包括的なエコシステムを擁する企業に拡大しました。Dell EMC IDPA は、保護ソフトウェア、サーバとストレージ、単一の UI 管理、クラウド機能、検索、分析のすべてを含む、事前構成される機能からなる、このエコシステムの一部になりました。このように、最新の市場要求が存在するところには、イノベーションを継続するための大きな推進力があり、重複排除メカニズムの効率向上、それを統合することによる機能強化、オンプレミスの重複排除にとどまらない機能拡張(クラウド階層化、クラウドホスト型アプライアンスとのレプリケーション、仮想アプライアンス、クラウドへのディザスタリカバリなど)がもたらされます。

すべての重複排除プラットフォームで最も重要な差別化要素のひとつに、リポジトリ内のデータ整合性を厳密に確保することがあります。これは多くの機能が、システム内のそれぞれのブロックに固有の整合性に左右されるためです。このため、システムがどの程度忠実に、かつどのくらいの頻度で、基盤となるストレージの整合性の確保を確認しているのか、入念に調査することが重要です。これに対する Dell EMC のアプローチは、DIA(データ非脆弱性アーキテクチャ)と呼ばれる技術により、当初から、そのバックアップ アプライアンスの大きな特徴になっています。

ESG は、15,000 以上の本番環境から収集されたフィールド データの検証に基づき、Dell EMC Data Domain と IDPA ハードウェアを Data Protection Suite ソフトウェアと組み合わせることで、競合のバックアップ ソフトウェアを搭載する Dell EMC バックアップ アプライアンスの環境と比較して保護容量のコストを 57%~81%削減できることを確認しました。

⁴出典: ESG 調査レポート、『Data Protection Modernization Trends』(2016 年 12 月)。

Dell EMC Data Domain または IDPA への投資をその「価格」により躊躇している組織は、ダウンタイムとデータロスにより生じるコストの削減、パフォーマンスの向上、そして最も重要な保護コストが 1 か月あたり 1GB につき 1 セント未満という経済的なメリットについて再検討するとよいでしょう。

すべての商標名は、それぞれの企業が所有権を保有しています。この資料に記載されている情報は、ESG (Enterprise Strategy Group) が信頼できるとみなす出典から取得したものです。ESG が保証するものではありません。この資料には、ESG の見解が記載されていますが、必要に応じて変更される場合があります。この資料は、Enterprise Strategy Group, Inc. が著作権を保有しています。Enterprise Strategy Group, Inc. の明示的な同意を得ずに、ハードコピーの形式、電子的、またはその他の方法を問わず、受領が許可されていない者にこの資料の全部あるいは一部を複製または再配布する行為は、米国著作権法に違反し、損害賠償請求の民事提訴、および妥当な場合は刑事告発の対象となります。ご質問がございましたら、ESG Client Relations (電話: 508-482-0188) までお問い合わせください。



Enterprise Strategy Groupは、IT分野の研究、分析、戦略立案を行う企業として世界的に有名であり、グローバルなITコミュニティに実用的な詳細情報とインテリジェンスを提供します。

© 2018 by The Enterprise Strategy Group, Inc. All rights reserved. (不許複製・禁無断転載)

