

White Paper

Il valore economico di Data Domain e Integrated Data Protection Appliance (IDPA)

Convalida dell'efficienza dei costi delle soluzioni di backup appliance
Dell EMC

Di Vinny Choinski, Senior ESG Lab Analyst;
e Christophe Bertrand, Senior Analyst
Giugno 2018

Il presente white paper di ESG è stato commissionato da Dell EMC
e viene distribuito con licenza concessa da ESG.

Sommario

Sommario.....	2
Introduzione.....	3
Vantaggi delle architetture Data Domain e IDPA	4
Vantaggi economici di Data Domain e IDPA	5
L'efficienza della deduplica è importante.....	6
Velocità e scalabilità.....	8
Considerazioni in merito al risparmio sui costi	9
Una verità più ampia.....	10

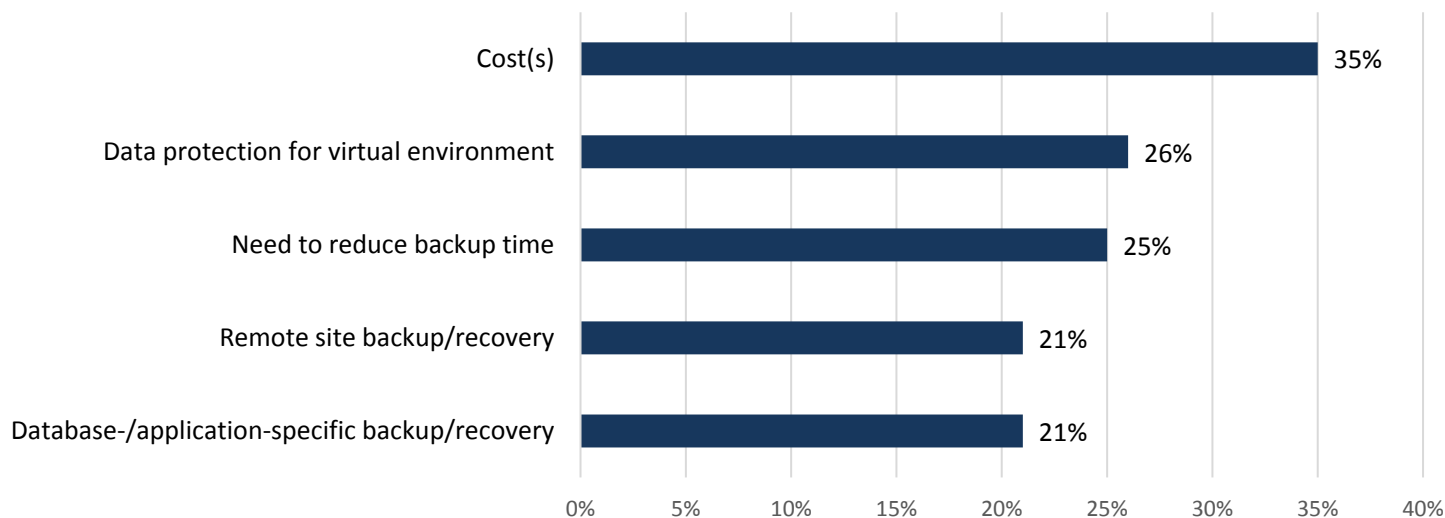
Introduzione

Per innovare di più e superare la concorrenza, le organizzazioni devono intraprendere un percorso coerente e garantire un'infrastruttura moderna. All'IT viene continuamente richiesto di offrire un'infrastruttura ottimizzata per nuove iniziative aziendali e applicazioni di supporto, sempre cercando di contenere o addirittura ridurre i costi. Infatti, gli intervistati della ricerca in corso di ESG definiscono costantemente la riduzione dei costi come uno dei principali fattori trainanti del business con impatto sulla spesa IT. Alla domanda di una recente survey di ricerca su come le organizzazioni intendessero contenere i costi per il 2017, il 27% degli intervistati ha risposto che avrebbe acquistato nuove tecnologie con un ROI migliore.¹

Per esaminare in dettaglio un gruppo specifico, il 35% dei responsabili IT con il compito di implementare processi e tecnologie di protezione dei dati generalmente considera i costi come una delle principali sfide (vedere la figura 1). Sembra che questo gruppo ritenga che anche i problemi incentrati sui workload siano molto complessi, in particolare quelli legati alla protezione di ambienti virtualizzati e uffici remoti. Tuttavia, osservando la classificazione complessiva delle sfide di protezione dei dati, emergono questioni più grandi. I costi e la virtualizzazione sono tra le principali preoccupazioni, ma le tre considerazioni riportate successivamente sono sia *operative* che *funzionali* per natura. Queste rappresentano sfide legate alle prestazioni, al backup e al ripristino dell'architettura distribuita, oltre alla protezione di database e applicazioni.²

Figura 1. Le attuali sfide della tecnologia e dei processi di protezione dei dati

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



Fonte: Enterprise Strategy Group

Queste sfide soddisfano anche i principali requisiti dei leader IT, relativi a un livello superiore di velocità, agilità, affidabilità e controllo dei costi, tutti illustrati in modo coerente nei precedenti report sulla modernizzazione della protezione dei dati di ESG. L'analisi di ESG conferma che le contraddizioni tra i problemi contro cui gli implementatori si scontravano e i requisiti richiesti dai leader causavano spesso un livello effettivo molto ridotto di modernizzazione o trasformazione. Gli implementatori non erano in grado di soddisfare i requisiti dei leader a causa di ostacoli tecnici nei propri ambienti di protezione dei dati.³

Questo documento illustra in che modo i sistemi Dell EMC Data Domain e le soluzioni convergenti realizzate sulla base dell'architettura Data Domain, ad esempio Dell EMC Integrated Data Protection Appliance (IDPA), consentono di offrire agli

¹ Fonte: report ESG Research, [Sondaggio sulle intenzioni di spesa IT per il 2017](#), marzo 2017.

² Fonte: survey ESG Research, [Data Protection Modernization Trends](#), dicembre 2016.

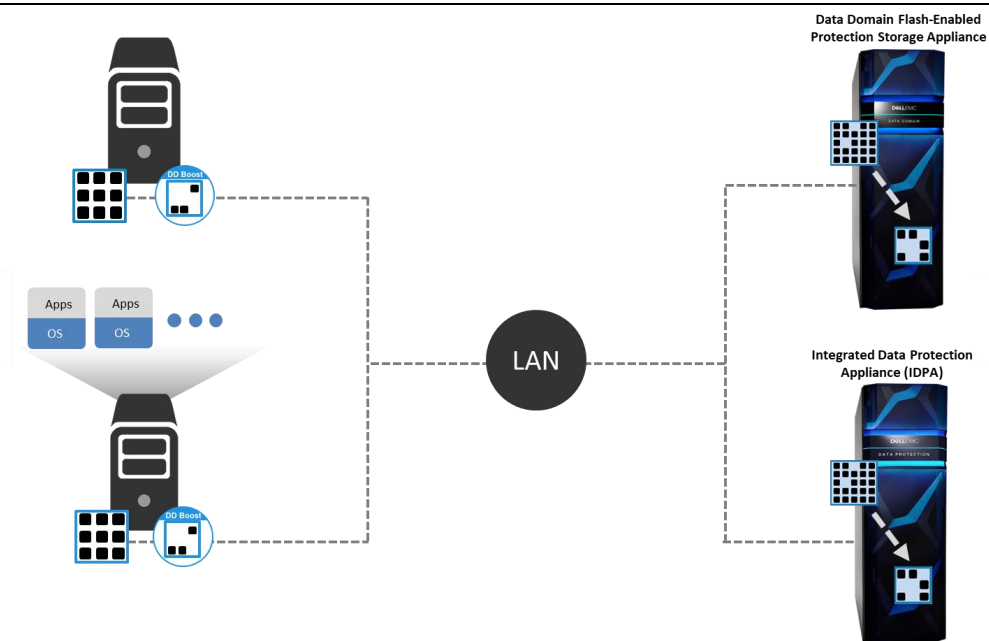
³ Fonte: report ESG Research [2015 Trends in Data Protection Modernization](#), settembre 2015.

implementatori IT l'agilità di cui hanno bisogno per trasformare le infrastrutture e soddisfare i requisiti dei leader IT senior, risolvere le sfide odierne della protezione dei dati e, soprattutto, ridurre i costi di storage per la protezione dei dati. I risultati di ESG si basano su un audit e un'analisi di key performance indicator (KPI): dati reali dei sistemi implementati, tra cui il prezzo di acquisto originale, i costi ambientali, l'utilizzo della capacità e le prestazioni.

Vantaggi delle architetture Data Domain e IDPA

Nel 2006, EMC ha deciso di evitare l'aggiunta di codice di deduplica dei dati alle soluzioni di protezione dei dati esistenti, determinando un cambiamento nella strategia di protezione dei dati. Questo cambiamento ha avuto inizio grazie alla tecnologia di due acquisizioni chiave. Il primo acquisto ha introdotto la tecnologia di deduplica basata sull'origine e il secondo, Data Domain, la tecnologia di deduplica basata sulla destinazione. Le tecnologie di queste due acquisizioni sono state componenti fondamentali nelle soluzioni di protezione dei dati di EMC. Ora appartenenti al marchio Dell EMC, i sistemi Data Domain, tra cui IDPA, sono appliance di protezione dei dati purpose-built, completamente integrati e abilitati per la tecnologia flash, progettati per ridurre la quantità di storage su disco necessaria per conservare e proteggere i dati. Con la deduplica dei dati basata sia sull'origine che sulla destinazione integrata in modo nativo nell'architettura, questi sistemi rendono possibile l'esecuzione di più backup in meno tempo, offrono restore più affidabili e veloci e riducono la quantità di capacità di storage necessaria per la protezione dei dati. Una generazione più recente di sistemi Data Domain e IDPA ha introdotto la tecnologia flash per i metadati, consentendo prestazioni su vasta scala insieme a un accesso e un restore istantanei di virtual machine 20 volte più rapidi direttamente da Data Domain e IDPA rispetto alle generazioni precedenti.

Figura 2. Panoramica delle architetture Data Domain e IDPA



Fonte: Enterprise Strategy Group

Le funzionalità principali dell'architettura Data Domain e IDPA includono:

- **Stream-Informed Segment Layout (SISL):** SISL consente ai sistemi Data Domain e IDPA di eseguire il 99% dell'elaborazione della deduplica all'interno di CPU e RAM. Questo significa che i sistemi non si basano sul numero di dischi per aumentare le prestazioni.
- **Segmentazione a lunghezza variabile:** i sistemi Data Domain e IDPA sfruttano la segmentazione a lunghezza variabile per interrompere i flussi in base alla struttura naturale dei dati per tassi di deduplica ottimali. Questo consente al sistema di stabilire se il segmento è univoco prima di comprimerlo e archivarlo.
- **Deduplica in linea:** i sistemi Data Domain e IDPA eseguono la deduplica all'interno di CPU e memoria quando il flusso di backup viene ricevuto dal sistema. Questo significa che solo i dati univoci vengono inviati al disco e quindi archiviati. Questo elimina anche la necessità di un'area di staging del disco e di risorse di elaborazione per la deduplica post-elaborazione.
- **Data Domain Boost Software:** il software Data Domain Boost distribuisce le parti del processo di deduplica sui client dell'applicazione o sul backup server. Con DD Boost, solo i dati univoci attraversano la connessione tra il backup server o i client e i sistemi Data Domain e IDPA. Questo consente anche di liberare le risorse sulla rete del cliente e i sistemi Data Domain e IDPA per ottenere prestazioni di deduplica lato destinazione migliorate.
- **Architettura per l'invulnerabilità dei dati:** Data Domain e IDPA sono soluzioni realizzate per garantire che i dati vengano ripristinati in modo affidabile. Data Invulnerability Architecture offre la verifica di lettura e scrittura in linea che protegge dai problemi di integrità dei dati durante l'acquisizione e il retrieval e procede al ripristino automatico. Il rilevamento degli errori e il self-healing continui assicurano che i dati siano ripristinabili per l'intero ciclo di vita sui sistemi Data Domain e IDPA.

Vantaggi economici di Data Domain e IDPA

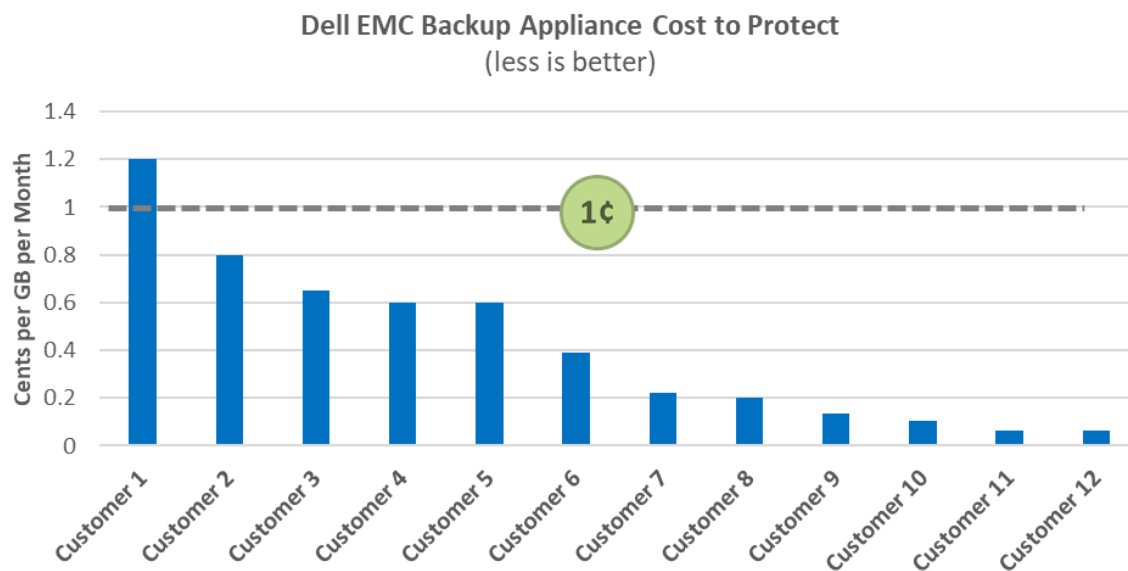
Osserviamo più da vicino il modo in cui le architetture Data Domain e IDPA si traducono in vantaggi economici per gli stakeholder del business. La deduplica dei dati consente ai clienti di archiviare più dati nella stessa quantità di spazio su disco fisico. Questo determina una riduzione dei requisiti di capacità di storage e diminuisce le spese. La deduplica dei dati implementata a livello dell'origine o lato client consente anche un risparmio in termini di capacità, ma con il vantaggio aggiunto di prestazioni migliori del backup. Con la deduplica lato origine, solo i data block univoci vengono inviati dall'origine alla destinazione durante l'operazione di backup e questo determina una riduzione del traffico di rete. La maggiore efficienza della rete consente la crescita dei dati di backup utilizzando l'infrastruttura di rete esistente e può anche eliminare o rinviare la necessità di costosi aggiornamenti della rete. Ovviamente, minore è la quantità di dati da trasferire, maggiore è la velocità delle prestazioni di backup. Durate più brevi di backup consentono anche ai clienti di aumentare la frequenza dei backup, riducendo il rischio di perdite dei dati, che possono essere estremamente costose per un'organizzazione. Con il software DD Boost, i sistemi Data Domain e IDPA supportano la deduplica sia lato origine che lato destinazione, offrendo ai clienti la flessibilità necessaria per implementare la deduplica dove sia più congeniale per gli ambienti.

Come illustrato nella figura 3, l'analisi ESG dei dati reali, tra cui hardware, software, alimentazione, raffreddamento e deduplica, dimostra che i sistemi Data Domain e IDPA sono facilmente in grado di offrire storage agli ambienti di protezione dei dati per pochi centesimi di dollaro per GB al mese.

Si noti come i costi di protezione oscillino tra 0,062 e 1,2 centesimi di dollaro per gigabyte per dodici clienti esaminati da ESG. Questo intervallo relativamente ampio di risparmio è dovuto principalmente alla quantità di risparmio in termini di

capacità nella deduplica raggiunto dai clienti Dell EMC negli ambienti di produzione. La parte restante di questo report osserva più nel dettaglio i costi di protezione con un'attenzione particolare al modo in cui i risparmi della deduplica possono essere incrementati tramite una combinazione end-to-end di software Dell EMC Data Protection Suite e hardware Data Domain e IDPA.

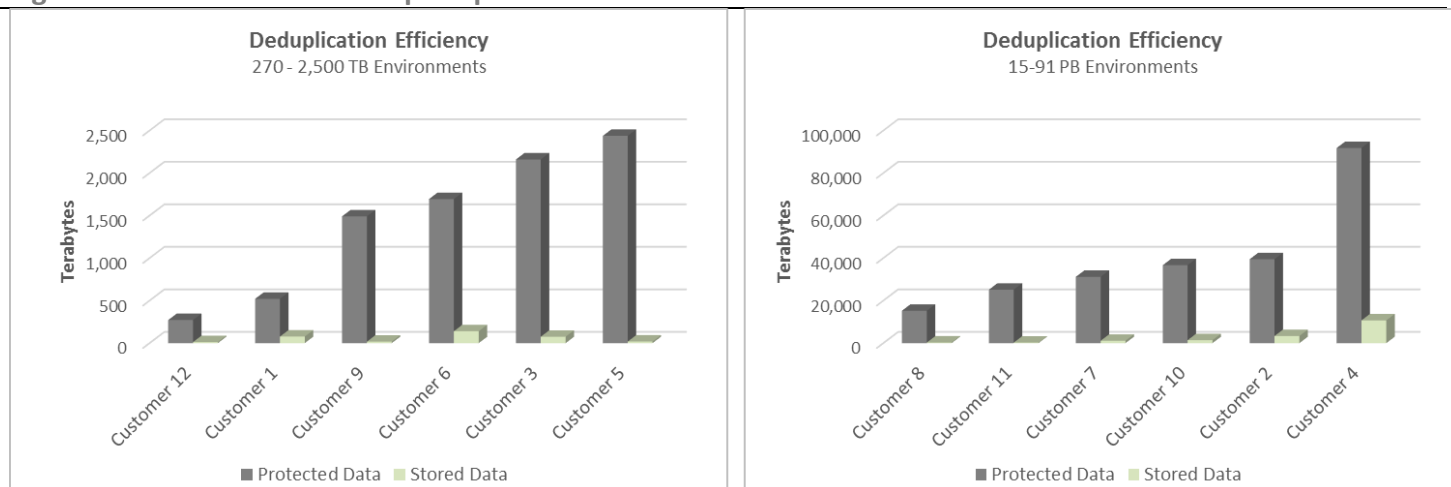
Figura 3. Il valore economico di Data Domain e IDPA



Fonte: Enterprise Strategy Group

L'efficienza della deduplica è importante

ESG ha iniziato a esplorare il valore economico del risparmio in termini di capacità nella deduplica di Data Domain e IDPA eseguendo l'auditing e l'analisi dei dati del supporto call home di 12 clienti Dell EMC attivi. Come illustrato nella figura 4, gli ambienti dei clienti oscillavano tra circa 270 TB e oltre 90 PB di dati da proteggere. I clienti provenivano da più settori, tra cui quello tecnologico, manifatturiero, assicurativo e sanitario. La selezione dei clienti in diversi settori aveva lo scopo di acquisire i risultati della deduplica nei diversi tipi di data set. Le barre grigie nella figura 4 mostrano la quantità di dati protetti per ogni ambiente. Le barre verdi mostrano la quantità di dati archiviati dopo la deduplica.

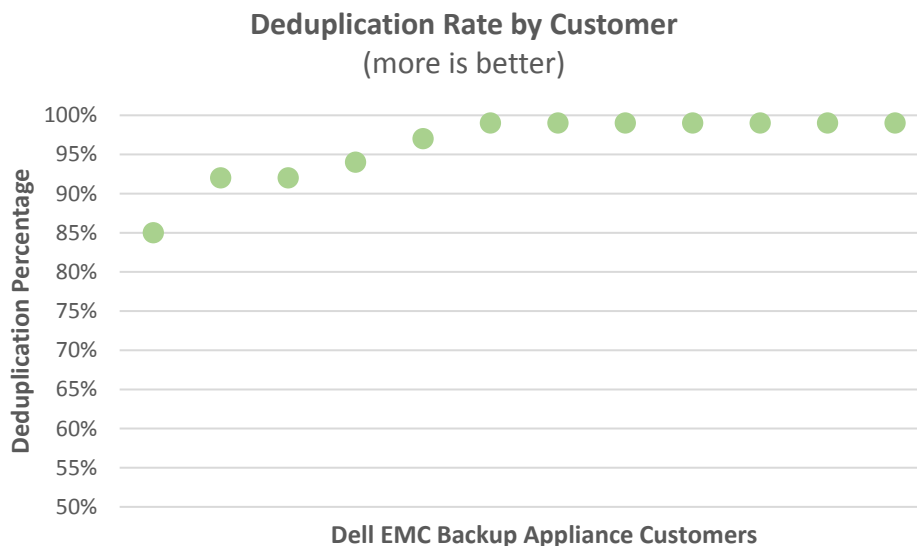
Figura 4. Efficienza della deduplica per cliente

Fonte: Enterprise Strategy Group

- **Deduplica** il tasso di deduplica più elevato osservato nei dati analizzati era di 126:1. Questo risultato deriva dall'ambiente del cliente identificato dal primo data point sul lato sinistro della figura 4. Il tasso di deduplica medio per tutti i clienti analizzati era di 41:1.
- **Protezione**: anche il cliente con il tasso di deduplica più basso osservato approssimativamente di 9:1 poteva proteggere circa 92 PB di dati utilizzando solo 10 PB di capacità. Questo cliente viene identificato dall'ultimo data point sul lato destro della figura 4. Il cliente appartiene al settore manifatturiero e dispone di un ambiente altamente distribuito.

La figura 5 offre un altro punto di vista sull'effetto della deduplica di Data Domain e IDPA. Questa figura mostra la percentuale di deduplica raggiunta dagli stessi ambienti dei 12 clienti. I tassi di deduplica oscillano tra l'85% e il 99%. Generalmente, maggiore è la durata della ricezione dei dati da parte dell'ambiente delle soluzioni Data Domain e IDPA, maggiore sarà il tasso di deduplica. Questo perché la soluzione ha maggiore probabilità di vedere gli stessi schemi di dati nel tempo e deve solo archiviare dati univoci.

La funzionalità di segmentazione a lunghezza variabile dell'architettura Data Domain è un componente chiave per il raggiungimento di questo livello di deduplica. Con la segmentazione a lunghezza variabile, le soluzioni Data Domain e IDPA sono più facilmente in linea con gli schemi naturali, tra cui gli indicatori di timestamp dei database, nelle strutture di dati inviate al dispositivo con finalità di protezione. La segmentazione a lunghezza variabile produce livelli significativamente superiori di deduplica rispetto alle architetture a lunghezza fissa. Consente anche una scalabilità più efficace all'interno di un singolo storage pool. Questo determina anche un numero inferiore di dispositivi da gestire e consente di risparmiare sui costi operativi.

Figura 5. Percentuale di deduplica per cliente

Fonte: Enterprise Strategy Group

Un altro vantaggio della deduplica con segmentazione a lunghezza variabile si verifica quando i dati vengono replicati tra dispositivi Data Domain e IDPA. Una quantità inferiore di dati comporta un numero inferiore di dati da replicare. Tutto ciò offre non solo prestazioni di replica migliori, ma anche meno requisiti di larghezza di banda WAN per il processo di replica.

- **Tassi di deduplica:** i tassi di deduplica sottoposti ad audit e analizzati oscillano tra l'85% e il 99%, con una media del 96%. Questo significa che solo una parte (nel migliore dei casi l'1% o meno) dei dati di produzione va archiviata sui sistemi Data Domain e IDPA con finalità di protezione.
- **Meno del 90%:** dai dati analizzati emerge che solo un cliente aveva un tasso di deduplica inferiore al 90%. Inoltre, questo cliente rappresenta l'ambiente altamente distribuito con pool di protezione distribuiti.
- **90% e oltre:** il resto dei clienti analizzati ha raggiunto tassi di deduplica superiori al 90% e la maggior parte degli ambienti dei clienti un tasso di deduplica del 99%.

Velocità e scalabilità

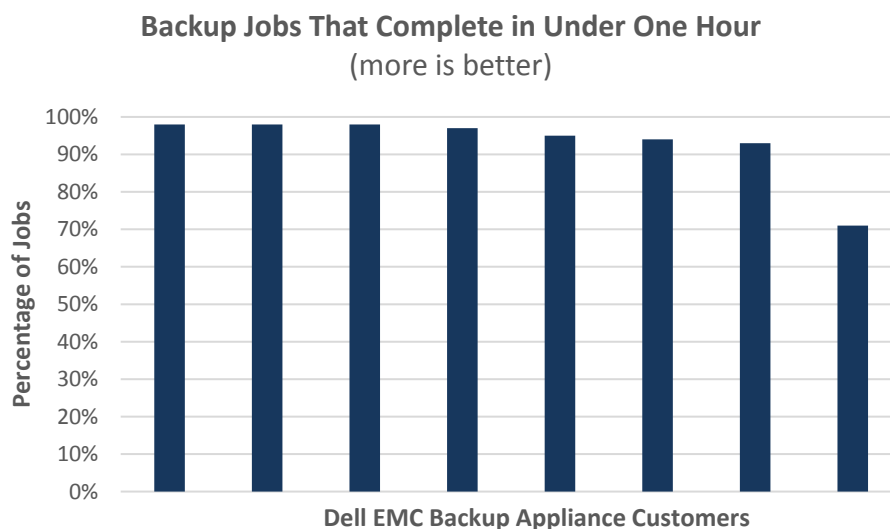
L'aumento della capacità di storage tramite deduplica non va sempre d'accordo con le prestazioni elevate. Data Domain e IDPA affrontano questa sfida sfruttando due elementi chiave delle prestazioni: Stream-Informed Segment Layout (SISL) e il software DD Boost.

SISL è un elemento dell'architettura delle piattaforme Data Domain e IDPA che consente al 99% del processo di deduplica di verificarsi all'interno di CPU e RAM. Questo significa che la soluzione non si basa su un numero elevato di spindle del disco per ottenere prestazioni. Di conseguenza, le soluzioni con ingombro inferiore possono ottenere le stesse prestazioni delle configurazioni con numero di spindle più elevato. Inoltre, poiché i sistemi Data Domain e IDPA utilizzano CPU Intel, sfruttano i vantaggi dati dai miglioramenti apportati alle prestazioni in ogni nuova release dei processori Intel.

Il software DD Boost aiuta anche a offrire prestazioni distribuendo parti del processo di deduplica sui client delle applicazioni o sul backup server. Con DD Boost, solo i dati univoci devono essere spostati dal backup server o dai client ai sistemi Data Domain e IDPA. DD Boost riduce la quantità di dati spostati fino al 99%. Il vantaggio netto del completamento

più rapido dei job di backup è stato oggetto di audit da parte di ESG. Come illustrato nella figura 6, la maggior parte dei clienti solitamente completa i job di backup in un'ora o in meno di un'ora.

Figura 6. Prestazioni del job di backup

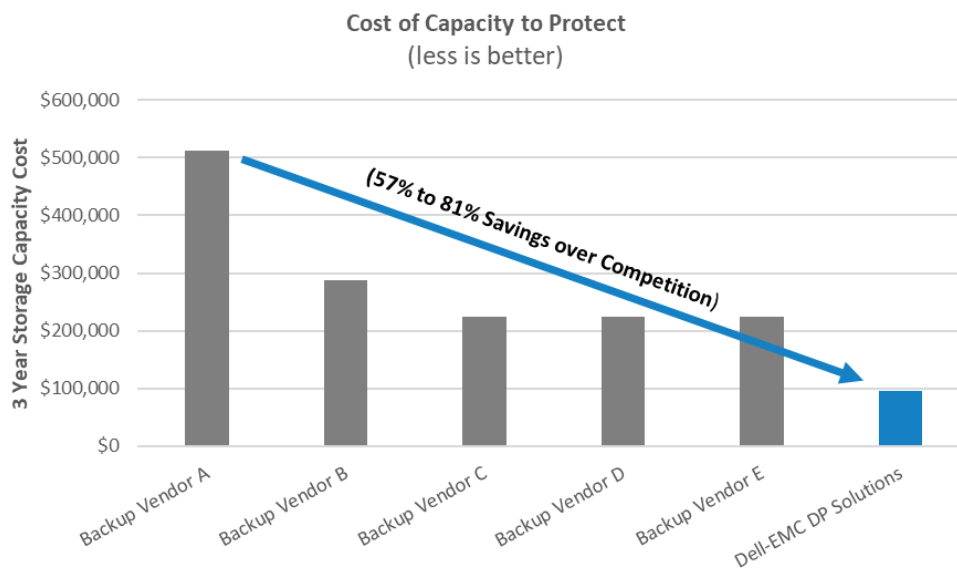


Fonte: Enterprise Strategy Group

- **Meno di 15 minuti:** per un cliente del settore tecnologico, il 76% di tutti i job di backup è stato terminato in meno di 15 minuti.
- **Meno di 1 ora:** per otto dei dodici clienti analizzati, dal 70% al 98% di tutti i job di backup è stato completato in meno di un'ora (come illustrato nella Figura 6).
- **Meno di 4 ore:** per un cliente del settore assicurativo, tutti i job di backup sono stati completati in meno di quattro ore.
- **Più di 4 ore:** per due degli ambienti più grandi (in termini di numero e dimensioni delle applicazioni), sono state impiegate più di quattro ore per completare tutti i job di backup.

Considerazioni in merito al risparmio sui costi

ESG ha completato l'esplorazione del valore economico del risparmio in termini di capacità nella deduplica di Data Domain con l'audit e l'analisi dei dati del supporto call home di oltre 15.000 sistemi Data Domain implementati in tutto il mondo. Questo esempio è stato raccolto dai sistemi che hanno anche registrato il backup software utilizzato e i tassi di deduplica ottenuti. I dati mostrano che l'efficienza di deduplica del backup appliance Dell EMC può essere ulteriormente migliorata con una soluzione di protezione dei dati Dell EMC end-to-end. Come illustrato nella figura 7, l'accoppiamento dell'hardware dell'appliance Dell EMC con il software Data Protection Suite ha amplificato il risparmio di duplicazione e ha ridotto i costi di capacità di storage di tre anni dal 57% all'81% rispetto alle soluzioni di altri fornitori.

Figura 7. Aumento del risparmio Data Domain e IDPA con Data Protection Suite

Fonte: Enterprise Strategy Group

La soluzione Data Domain offre altri vantaggi economici oltre al risparmio in termini di capacità come illustrato nella figura 7. L'analisi di ESG dei dati reali del cliente ha rivelato anche i seguenti vantaggi:

- **Riduzione della larghezza di banda della rete:** la maggior parte delle applicazioni di livello enterprise e delle soluzioni di backup e ripristino si affida alle risorse di rete per funzionare. Questo include sia le risorse LAN per il backup e il restore locali che le risorse WAN per DR e BC. Le soluzioni di protezione dei dati inefficienti possono consumare rapidamente preziose risorse di rete e addirittura influire sulla connettività e la produttività dell'utente. Sfruttando la deduplica di Data Domain e IDPA, un produttore multinazionale ha potuto ridurre l'utilizzo della larghezza di banda per la protezione dei dati in remoto e locale del 98%.
- **Prestazioni:** un'efficiente deduplica dei dati può esercitare un notevole impatto sul miglioramento delle prestazioni complessive della protezione dei dati. Implementando DD Boost, che riduce la quantità di dati da trasferire tra il client e il sistema Data Domain, un produttore di attrezzature pesanti ha potuto incrementare le velocità di backup e restore del 50%. Lo stesso cliente ha migliorato la propria idoneità al DR del 90%.
- **Riduzione del footprint del data center:** l'ambiente fisico, soprattutto la quantità necessaria per supportare i requisiti di spazio, alimentazione e raffreddamento di un Modern Data Center, non è affatto economico. Può essere difficile, talvolta impossibile, espandere fisicamente un data center senza trasferimento. Con Data Domain, ad esempio, una catena di grandi magazzini a livello nazionale ha eliminato il nastro fisico nel proprio ambiente di protezione dei dati e liberato tre file complete di preziosissimo spazio su rack del data center.

Una verità più ampia

I requisiti principali della protezione dei dati dei leader IT sono incentrati sul miglioramento dell'affidabilità e dell'agilità fondamentali della soluzione utilizzata/delle soluzioni utilizzate. Il requisito immediatamente successivo è la riduzione dei costi, vista anche come priorità principale tra gli implementatori della protezione dei dati. Queste sfide non devono essere viste come contraddittorie o che si escludono a vicenda. Infatti, possono essere affrontate tutte con soluzioni di protezione dei dati progettate per garantire efficienza insieme ad affidabilità e capacità.

L'efficienza dipende da molti fattori, spesso frenati dai costi legati al business e all'interoperabilità:

- Capacità di offrire il protection storage in modo economico riducendo i costi della gestione delle applicazioni di protezione dei dati.
- Interoperabilità tra soluzioni di backup software, poiché solo un'organizzazione su quattro presenta un'unica applicazione di backup.⁴
- L'efficienza in termini di scalabilità e deduplica del sistema di protection storage, che può essere utilizzata per ridurre il numero e il costo dei sistemi da implementare e gestire.
- La capacità di sfruttare lo stesso pool di protection storage per backup e dati di archiviazione ottimizzato per la capacità.
- Interoperabilità (e integrazione) tra protection storage e una miriade di piattaforme, in modo che gli owner delle applicazioni (ad esempio Database Administrator e vAdmin) possano utilizzare i propri strumenti per aumentare la protezione o il ripristino, pur restando sempre responsabili dei cittadini IT e sfruttando un protection storage centralizzato ed efficiente.
- La capacità di velocizzare il deployment con soluzioni di protezione dei dati integrate, riducendo il numero di sistemi da implementare, semplificando le licenze con un approccio onnicomprensivo e riducendo i problemi di interoperabilità.
- La capacità di semplificare la gestione complessiva degli ambienti di protezione dei dati con gestione centralizzata e reporting, oltre a un singolo passaggio di aggiornamento/patch per tutti i componenti di protezione dei dati all'interno dell'appliance, con una riduzione del tempo necessario per eseguire la manutenzione dell'ambiente.
- La capacità di integrare abilità più ampie, come cloud (conservazione a lungo termine, disaster recovery per il cloud), ricerca e analisi, che ancora una volta riduce i tempi di implementazione e semplifica la gestione.

Offrendo quello che molti nel settore considerano sinonimo di "protection storage" e/o "deduplica", Dell EMC si è trasformata da semplice fornitrice di componenti per la protezione dei dati a fornitrice di un ecosistema completo che include storage di produzione, vendor di backup, vendor di archiviazione e plug-in accelerator per le applicazioni (ad esempio DD Boost), tra gli altri elementi. Il Dell EMC IDPA è ora parte di questo ecosistema con software di protezione, server e storage, gestione delle singole interfacce utente, funzionalità cloud, di ricerca e di analisi, tutto incluso e pre-configurato. Pertanto, è significativo l'impulso a continuare a innovare in base alle richieste del mercato attuale, con conseguente ulteriore aumento dell'efficienza all'interno di meccanismi di deduplica, miglioramenti con le proprie integrazioni ed espansione delle proprie funzionalità oltre la deduplica solo on-premise, come cloud tiering, replica a/da appliance ospitati nel cloud, Virtual Appliance, disaster recovery nel cloud e così via.

⁴ Fonte: survey ESG Research, *Data Protection Modernization Trends*, dicembre 2016.

Uno dei principali elementi distintivi di una piattaforma di deduplica è il rigore nel garantire l'integrità dei dati all'interno del repository, perché molto dipende dall'integrità di ogni block univoco all'interno del sistema. Pertanto, è cruciale che le organizzazioni indaghino diligentemente il modo zelante e la frequenza con cui il sistema esamina lo storage sottostante per garantire l'integrità. L'approccio di Dell EMC a questo è stato un elemento caratteristico della piattaforma di backup appliance sin dall'inizio tramite ciò che viene definito Data Invulnerability Architecture (DIA).

Sulla base di un audit dei dati raccolti sul campo di oltre 15.000 ambienti di produzione, ESG ha confermato che la combinazione dell'hardware Dell EMC Data Domain e IDPA con il software Data Protection Suite consente di ridurre i costi della capacità di protezione in un intervallo dal 57% all'81% rispetto agli ambienti di backup appliance Dell EMC con un backup software competitivo. Le organizzazioni che esitano a effettuare un investimento in Data Domain o IDPA a causa del "prezzo" dovrebbero riconsiderare i vantaggi economici che questo prodotto può offrire nel tempo, ad esempio il risparmio legato alla riduzione del downtime e della perdita dei dati, il miglioramento delle prestazioni e, soprattutto, la possibilità di proteggere il proprio sistema con pochi centesimi di dollaro per GB al mese.

Tutti i marchi utilizzati appartengono ai rispettivi proprietari. Le informazioni contenute nella presente pubblicazione provengono da fonti che l'Enterprise Strategy Group (ESG) considera affidabili, ma che non garantisce. Questa pubblicazione può contenere opinioni espresse da ESG, soggette a possibili cambiamenti nel tempo. Pubblicazione protetta dal copyright di Enterprise Strategy Group, Inc. La riproduzione o la distribuzione integrale o parziale della pubblicazione, in formato cartaceo, elettronico o altro, rivolta a persone non autorizzate e senza l'espresso consenso di Enterprise Strategy Group, Inc., costituisce una violazione della legge sul copyright vigente negli Stati Uniti ed è passibile di azione legale per il risarcimento dei danni e, se applicabile, di azione penale. Per ulteriori informazioni, rivolgersi a ESG Client Relations al numero 001 508 4820188.



Enterprise Strategy Group è una società di analisi, ricerca e strategia in campo IT conosciuta in tutto il mondo per offrire dati e intelligence fruibili alla community IT globale.

© 2018 by The Enterprise Strategy Group, Inc. Tutti i diritti riservati.

