rubrik

Vincere la sfida dei dati non strutturati

eBook

INIZIA →

INDIETRO AVANTI

Introduzione

Usare applicazioni digitali significa avere dati non strutturati.
Petabyte di dati. Ma a conti fatti, sai che dati stai raccogliendo?

Avrai già sentito parlare delle tre V dei dati: **varietà, velocità** e **volume**.

In questo eBook parleremo della quarta V: il **valore**. In particolare, vedremo i 5 step per creare una strategia resiliente per i dati non strutturati. In più, illustreremo come Rubrik può aiutarti a gestire e proteggere i tuoi dati non strutturati.

Non ci resta dunque che esplorare il mondo oscuro dei dati non strutturati.

Cominciamo!

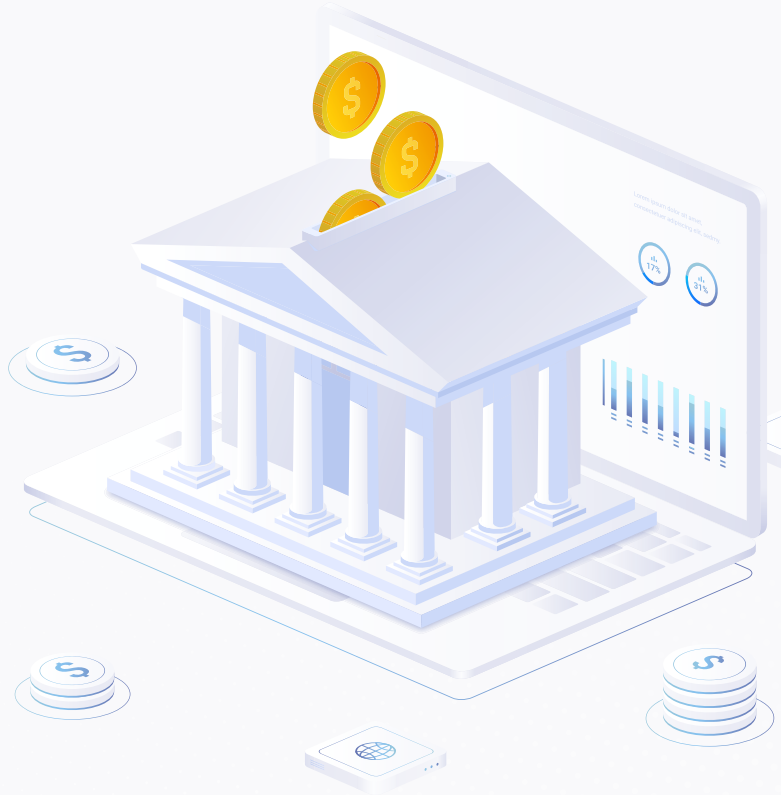
INDIETRO AVANTI

C'era una volta l'istituto MegaBanca

Immagina di essere il CISO di MegaBanca, una banca i cui clienti vantano un patrimonio netto altissimo. Ogni giorno elabori grandi volumi di dati, tra cui molti dati sensibili: numeri di conto corrente, codici fiscali e altre informazioni di identificazione personale dei clienti. Normale amministrazione.

Sai quanto sia importante proteggere questi dati e hai investito molto nella tua strategia di sicurezza, con database di alto livello dotati delle migliori protezioni, un solido piano di Business Continuity e Disaster Recovery, un team molto preparato nel contrastare gli attacchi informatici giornalieri.
La tua banca sembra molto ben difesa.

Ma l'apparenza inganna.



INDIETRO AVANTI

Oggi:

Due dipendenti hanno preso un numero di conto corrente dal database per discutere meglio di un caso.

Qualcuno del team M&A ha mandato via email a un collega la copia di un contratto triennale contenente informazioni proprietarie.

Gli operatori dell'assistenza hanno gestito le chiamate con i clienti, che vengono tutte registrate e spesso contengono informazioni di identificazione personale.

E hai appena saputo che un hacker ha avuto accesso ai tuoi sistemi.

Ora hai un problema. Perché ogni interazione ha generato un file che è andato a finire tra i tuoi dati non strutturati: file non organizzati, difficili da gestire, che molto probabilmente contengono informazioni sensibili e non sono protetti in modo adeguato.

Ti aspetta una nottataccia.

INDIETRO AVANTI

Varietà, velocità, volume e... valore?

È uno scenario da incubo, no? Ma non temere. Con questo eBook ti aiuteremo a sviluppare una strategia per gestire e proteggere i dati non strutturati in modo da non ritrovarti come il CISO del nostro esempio.

Chiunque operi in questo settore da 20 anni ha sentito parlare delle tre V di **varietà, velocità e volume.** Definite da Doug Laney nel 2001[1], le tre V sono le proprietà dei dati raccolti da un'organizzazione.[2]

VARIETÀ

Indica i **diversi tipi** di dati raccolti. Avere informazioni di ogni genere è utile per capire il quadro d'insieme, ma serve un sistema in grado di gestire:

- Dati strutturati e non strutturati
- Formati diversi
- Nomenclature diverse

VELOCITÀ

Indica la **rapidità** con cui vengono generati ed elaborati i dati. In ambito commerciale, la velocità è tutto.

- Un sito di e-commerce B2C che carica i dati in 1 secondo ha un tasso di conversione 2,5 volte superiore a un sito che carica i dati in 5 secondi.[3]
- Negli Stati Uniti, circa 1/3 degli utenti che sceglie una banca verifica se ha una buona protezione dalle frodi,[4] quindi gli istituti devono essere in grado di elaborare i dati in tempo reale per prevenire le frodi.
- Gli hacker sono rapidi e inarrestabili: il 99% dei responsabili IT e della sicurezza è venuto a conoscenza di almeno un attacco informatico nel 2022, con una media di 52 casi affrontati.[5]

VOLUME

Indica la **quantità** di dati raccolti. Indovina un po': sono tantissimi. Secondo il report The State of Data Security 2023 di Rubrik Zero Labs, i numeri sono impressionanti:

- Un'azienda tipica raccoglie in media 239,9 terabyte di dati di backend (BETB).
- Questo numero aumenta notevolmente per chi opera in settori specifici:
  - > Telecomunicazioni: 442,6 BETB
  - > IT e tecnologia: 398,9 BETB
  - > Assicurazioni: 301,5 BETB

INDIETRO AVANTI

Ricapitolando, devi gestire in tempi brevissimi un'enorme mole di dati di diversi formati.

Ma non finisce qui.

Per Rubrik esiste una quarta V da ***prendere in considerazione*** per una corretta gestione e protezione delle informazioni: la V di **valore**.

VALORE

Indica quanto sono utili i dati raccolti per la tua azienda. Devi sapere quali dati hai, perché sono importanti e dove si trovano. Questa enorme quantità di dati business-critical che raccogli è il vero tesoro della tua azienda. Ricordi i 239,9 BETB di dati? Molti di questi sono dati sensibili.

- In generale, un'azienda tipo gestisce in media 24,8 milioni di file sensibili.[6]
- I dati più sensibili raccolti da un'azienda protetta dalle soluzioni Rubrik ammontano a oltre 1,3 miliardi di record.[7]

Più un'azienda cresce, e più cresce la sua impronta di dati sensibili.

24,8 MILIONI

quantità media di file sensibili
di un'azienda tipo

+1,3 MILIARDI

quantità di dati sensibili
di un'azienda protetta dalle soluzioni Rubrik

[1] 3D Data Management: Controlling Volume, Velocity, and Variety
[2] Big Data: The 3 V's Explained; 3 V's (volume, velocity and variety)
[3] Site Speed is (Still) Impacting Your Conversion Rate
[4] New FICO Survey: Americans Value Financial Fraud Prevention More Than Banking Customer Experience
[5] Rubrik Zero Labs, The State of Data Security Report 2023
[6] Ibid.
[7] Ibid.

INDIETRO AVANTI

In conclusione, è il valore dei tuoi dati che stabilisce:

Cosa proteggere

La frequenza di backup

La durata del periodo di conservazione dei dati

La modalità di protezione a riposo per impedire eventi di crittografia, esfiltrazione e ransomware

Ora, dovresti già avere una buona dimestichezza riguardo al valore dei dati strutturati. Hai adottato tutte le misure possibili per gestirli e proteggerli.

E per i dati non strutturati invece? In questo momento sai cosa contengono i file non strutturati della tua azienda e come aumentare la resilienza dei dati a riposo per renderli più difficili da esfiltrare?

Se sei come la maggior parte dei data manager che abbiamo incontrato, la risposta più probabile sarà "non proprio".

La buona notizia è che siamo qui per darti una mano.

INDIETRO AVANTI

5 step per rafforzare la strategia per i dati non strutturati

Gran parte dei dati non strutturati non contiene informazioni sensibili, non ti offre un vantaggio competitivo o non è essenziale per le tue operazioni aziendali. Insomma, molti dei tuoi dati non sono così preziosi.

PUNTO 1 – I tuoi dati non strutturati ***in realtà*** contengono i dati più preziosi per l'azienda.

PUNTO 2 – Gli hacker ***sanno*** che i dati non strutturati possono nascondere informazioni vitali, per questo prendono di mira i backup. Gli autori di attacchi informatici ***accedono*** alla tua rete, non la violano. E ora hanno le autorizzazioni per accedere ai tuoi dati non strutturati.

PUNTO 3 – In più, secondo IDC, entro 5 anni il ***90% dei dati*** sarà di tipo non strutturato.[7]

Se ti affidi al vecchio metodo di copia e replica delle snapshot o al protocollo NDMP, corri rischi enormi, perché non sai dove si trovano i dati e potresti introdurre involontariamente dei malware nei tuoi backup e archivi.

Oggi non puoi più eseguire il backup dei dati non strutturati e sperare che tutto vada bene. Questa è una strategia basata sulla ***continuità***, e non sulla ***resilienza*** dei dati. E nell'era del cybercrime, la ***resilienza è fondamentale***.

I 5 step che descriveremo ti aiuteranno a comprendere, gestire e proteggere i tuoi dati non strutturati.

Vediamoli insieme.

[7] IDC, Meeting the New Unstructured Storage Requirements for Digitally Transforming Enterprises

INDIETRO AVANTI

Step 1: Conosci i tuoi dati

I dati crescono a ritmi incredibili: secondo le stime Rubrik il volume totale dei dati da proteggere di un'azienda tipo aumenterà di **7 volte** nei prossimi 5 anni.[8]

Questo primo step ti permetterà di conoscere il tuo ecosistema di dati non strutturati per assegnarli il giusto valore e protezione.

CAPIRE CHI È ATTUALMENTE RESPONSABILE per la raccolta e la valutazione dei dati. Per molti è compito di chi amministra lo storage e/o il cloud, ma non dovrebbe essere così. Questi amministratori non sono tenuti a conoscere i tipi di dati specifici che gestiscono e la loro provenienza. La raccolta e la valutazione sono a carico dei proprietari dei dati nell'azienda, ma chi se ne occupa davvero?

CAPIRE DA DOVE ARRIVANO I DATI.
Hai il controllo di ogni applicazione che genera dati nella tua azienda o fate un libero uso di shadow IT? Fai un elenco completo. **Poi chiediti:** tutti i punti di origine sono privi di malware? Per tutti i principali bucket di dati, è essenziale che i punti di origine siano noti, valutati e puliti. **Inoltre, controlla:** i proprietari dei dati sono coinvolti e/o prendono decisioni sulla sicurezza dei dati? O queste decisioni vengono prese da chi amministra lo storage e i backup?

[8] Rubrik Zero Labs, The State of Data Security Report 2023

INDIETRO AVANTI

Step 2: Convalida i tuoi dati

A questo punto devi analizzare i dati raccolti e impostare le policy per gestirli al meglio.

Osserva i **dati archiviati** e classificali in ordine di **criticità** per l'azienda. I log delle stampanti? Probabilmente non sono essenziali. Le email invece? Cosa accadrebbe se l'email non funzionasse? O se perdessi i messaggi interni o i dati dei sensori? Nel momento in cui sai quali sono le applicazioni che producono dati importanti, è più facile configurare la strategia di protezione.

Definisci **due ruoli dedicati** nella tua azienda e configura i relativi flussi di lavoro. Entrambi i ruoli saranno responsabili e competenti per il processo di convalida e ti aiuteranno a mettere in pratica la tua strategia di resilienza per i dati non strutturati.

RESPONSABILE ESECUTIVO

Solo il 54% delle organizzazioni esterne ha un responsabile esecutivo senior per la gestione e la sicurezza dei dati, ma il 98% di queste organizzazioni sostiene di avere enormi difficoltà di visibilità dei dati.[9] Se ti riconosci in questa situazione, è il momento di nominare un responsabile esecutivo dei dati. In questo modo potrai, da un lato, ottimizzare e rafforzare la strategia di protezione dei dati e, dall'altro, dimostrare ai team interni, ai clienti e alle parti interessate di gestire con criterio i dati aziendali, ovunque risiedano.

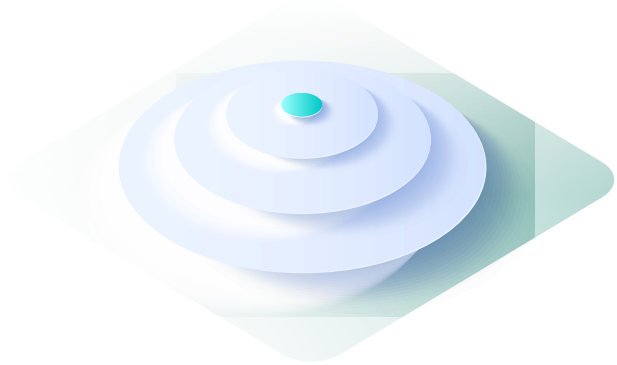
CUSTODI DEI DATI

Se gli amministratori di storage e cloud non sono proprietari dei dati, chi dovrebbe custodirli? Noi consigliamo l'impiego di esperti di dati - i Custodi dei dati - integrati nei team dell'organizzazione e responsabili dell'intero ciclo di vita delle informazioni (ILM). Dal momento in cui i dati vengono generati fino a quando vengono archiviati ed eliminati, questi amministratori sanno perché sono importanti, verificano che siano integri e li proteggono adeguatamente.

[9] Rubrik Zero Labs, The State of Data Security Report 2023

INDIETRO AVANTI

Step 3: Classifica le tue applicazioni



A questo punto devi classificare le applicazioni che generano i dati.

In questo modo puoi standardizzare un flusso di lavoro e gestire dove e come archiviare e proteggere i dati. I livelli delle applicazioni cambiano da un'azienda all'altra, a seconda di quali siano i dati aziendali critici.

LIVELLO ALTO (CRITICI)

I dati prodotti da queste applicazioni sono essenziali. Possono essere file non ricreabili, come le immagini di diagnostica dei pazienti o i dati di navigazione e dei sensori generati in un determinato punto nel tempo, oppure file che devi conservare per motivi di conformità normativa. Anche le email possono rientrare in questa categoria, perché possono contenere informazioni sensibili e perché sono dati essenziali per far funzionare le email. A questo livello devi applicare gli SLA (accordi sul livello del servizio) più rigidi, con la massima protezione e resilienza, per eseguire il recovery al backup più recente in caso di compromissione dei dati.

In conclusione: rientrano in questo livello le applicazioni che, se compromesse, potrebbero causare un'interruzione significativa dei servizi aziendali.

> I dati non strutturati provengono da tutte le aree aziendali: client email, sensori, programmi di social media, stampanti, software di scrittura e per presentazioni, in pratica da ogni applicazione digitale utilizzata.

INDIETRO AVANTI

LIVELLO MEDIO (IMPORTANTI)

I dati prodotti dalle applicazioni di questo livello sono importanti per l'azienda ma non critici. Un esempio sono i dati statistici che consentono di replicare i miglioramenti ottenuti o i report non sensibili utili per i team. Se la tua azienda dovesse subire un attacco informatico e perdere questi dati, potresti comunque ricrearli e riprendere a lavorare senza problemi.

In pratica, le applicazioni in questo livello generano dati utili ma non critici per le attività quotidiane e la stabilità aziendale nel tempo.

LIVELLO BASSO (TUTTI GLI ALTRI DATI)

Le applicazioni del livello basso producono dati di scarso o nessun valore per l'azienda. Nella maggior parte delle aziende questo livello include le applicazioni di stampa e social media. I dati a questo livello non richiedono una protezione avanzata ed è probabile che, in base alla strategia ILM, vengano archiviati e rimossi dall'origine.

In conclusione: rientrano in questo livello le applicazioni con un impatto minimo o insignificante sull'operatività dell'azienda in caso di compromissione da parte di un hacker.

INDIETRO AVANTI

Step 4: Definire e implementare la tua strategia per i dati

Dopo tutte queste fasi preparatorie, è il momento di entrare nel vivo della strategia e dell'implementazione. Utilizzando le informazioni che hai raccolto e organizzato, imposta le tue regole sui dati. Definisci gli standard su ruoli, processi e tecnologie per ogni livello di applicazione determinati nello step 3. Quindi, per le applicazioni di alto, medio e basso livello in ogni team dell'azienda, stabilisci quanto segue:

Chi sono i Custodi dei dati che gestiscono i dati delle applicazioni nei singoli team?

Ogni quanto va eseguito il backup per ogni tipo di applicazione? Quando bisogna trasferire i dati negli archivi? Quando si possono ritirare (eliminare) i dati? Tieni presente che i requisiti legali e normativi possono applicarsi anche a questa fase e influire sulle policy di conservazione.

Qual è la procedura di recovery da adottare per ogni livello in caso di attacco informatico?

In quale punto dell'infrastruttura devono andare i dati provenienti da ogni livello applicativo?

- In che misura sono dati di backup o dati di archivio?
- Puoi permetterti di perdere i dati archiviati (che non essendo sottoposti a backup non sono recuperabili in caso di attacco informatico)?
- Quando stabilisci dove inviare i dati, devi rispettare determinati regolamenti, leggi o criteri applicabili (come HIPAA o CCPA)?

Come verrà gestita la protezione dei dati per ogni tipo di applicazione?

> Ricorda: una volta che avrai definito queste regole, i tuoi Custodi dei dati saranno i **TITOLARI** di gestione, protezione e messa in sicurezza dei dati per l'intero ciclo di vita. Infatti sono loro che monitorano i dati prodotti e devono garantire che arrivino protetti nella destinazione di archiviazione.

INDIETRO AVANTI

Step 5: Mantenere e monitorare la strategia

Il più è fatto. Ora sai da dove vengono i dati non strutturati, chi li gestisce e che valore hanno per la tua azienda, quindi hai sviluppato una procedura per metterli al sicuro, proteggerli e sistemarli nell'infrastruttura aziendale.

Ti resta una sola cosa da fare: mantenere e monitorare la strategia per i dati non strutturati sul lungo periodo. È una fase critica, che non devi sottovalutare, perché i dati continuano ad aumentare in modo esponenziale. Ogni giorno aggiungi nuove applicazioni. C'è un continuo ricambio di personale nei tuoi team. E il cybercrime cresce sempre.

Ecco 4 azioni concrete da adottare per mantenere efficiente e resiliente la tua strategia:

SFRUTTA LA VISIBILITÀ DEI DATI PER ESAMINARE QUELLI SENSIBILI CON REGOLARITÀ

Riduci i dati per semplificarti la vita. Ad esempio, puoi rimuovere i dati sensibili non visualizzati dagli utenti nell'ultimo anno, individuare e cancellare i duplicati, oppure eliminare i dati delle condivisioni di dipendenti, clienti e partner che nell'ultimo anno hanno lasciato l'azienda. Questo vale anche per i dati duplicati in diversi ambienti della stessa azienda.

INDIETRO AVANTI

CONTROLLA LE ANOMALIE PER GARANTIRE LA RESILIENZA DEI DATI

Monitora le attività insolite di aggiunta, eliminazione o crittografia dei dati con il machine learning. Verifica che gli utenti abbiano i giusti diritti di accesso (di sola lettura, modifica, ecc.) in base alle loro esigenze e ruoli. Scegli una soluzione capace di avvisarti subito su possibili attività malevole nei dati di backup durante un attacco ransomware.

CONTROLLA LA CRESCITA DEI DATI

Cerca di impedire un aumento esponenziale dei dati da gestire. Ad esempio, imposta la crescita del cloud a non più del 50% dell'ambiente totale, elimina i dati in base alle policy impostate o archivia i dati sensibili in un unico posto.

RIVEDI LA STRATEGIA REGOLARMENTE

La tua strategia deve essere sempre attuale e in linea con le esigenze aziendali. Esegui una revisione periodica, almeno due volte all'anno, per verificare che i team operino nel rispetto delle policy e che tu possa anticipare le minacce informatiche.

Ottimo lavoro!

Hai completato la transizione da una strategia di **continuità** a una di **resilienza** dei dati non strutturati.

INDIETRO AVANTI

Rubrik: al tuo fianco in questa sfida

Ora che conosci i 5 step e sai come evitare i problemi che invece deve affrontare il CISO del nostro esempio, vediamo le soluzioni offerte da Rubrik.

Rubrik NAS Cloud Direct è una soluzione moderna per gestire e proteggere i tuoi dati non strutturati. Questa VM stateless inclusa nel piano di controllo SaaS di Rubrik, si può distribuire in modalità nativa nel cloud o nel data center e può analizzare miliardi di file, ogni tipo di file, in ogni momento. Basata sui principi Zero Trust, è perfetta per le più complesse sfide sui dati non strutturati:

VOLUME

Hai miliardi di file e ogni giorno ne aggiungi di nuovi? Nessun problema.

- Proteggi i dati dell'ordine di petabyte su tutte le tecnologie NAS con funzionalità efficienti di scansione, indicizzazione e trasferimento dei dati.

VELOCITÀ

Vuoi spostare i file e generare i backup senza interruzioni nell'ambiente di produzione? Con Rubrik puoi.

- Scansiona, indicizza e trasferisci i dati NAS in stream paralleli per massimizzare il throughput di rete.
- Elimina ogni impatto sugli utenti con il throttling o ridimensionamento dinamico.
- Riduci fortemente i tempi di backup e migliora l'efficienza operativa con backup realmente incrementali per sempre.

EFFICIENZA

Vuoi eliminare le complessità e semplificare i processi? Possiamo aiutarti.

- Integra gli strumenti SecOps come SIEM/SOAR per facilitare la collaborazione tra i team ITOps e SecOps per individuare e delimitare rapidamente le minacce.
- Archivia i dati da qualsiasi origine NAS direttamente su archiviazioni on-premise, cloud o private in base a policy da te impostate.

SICUREZZA

Vuoi la protezione migliore? Siamo qui per questo.

- Proteggi i dati NAS con backup immutabili e crittografati a riposo, con l'isolamento delle credenziali per una maggiore resilienza agli attacchi informatici.
- Sfrutta i client NFS e SMB personalizzati, progettati appositamente per una protezione dei dati rapida e scalabile.

GESTIONE

Cerchi un modo per gestire i dati non strutturati invece di lasciarli nel caos? Abbiamo la soluzione che fa per te.

- Cerca e trova file specifici con facilità e recupera le versioni protette in precedenza dei dati NAS.
- Rileva la presenza di dati sensibili con Sensitive Data Monitoring e ottieni informazioni sulla postura di sicurezza degli accessi per limitare il rischio di esposizione.
- Identifica rapidamente i set di dati obsoleti o in crescita con NAS CD Data Discover per decidere se archiviarli, spostarli o eliminarli.

RESILIENZA

La tua azienda ha subito un attacco informatico e deve tornare operativa il prima possibile? Con Rubrik puoi riprendere subito a lavorare.

- Identifica e localizza rapidamente le applicazioni e i file colpiti dal ransomware con Rubrik Anomaly Detection.
- Ripristina con la massima precisione i dati NAS interessati grazie all'analisi dell'impatto di Rubrik Anomaly Detection.
- Automatizza i flussi di lavoro del recovery, come il recupero di massa dei dati NAS, fino alla produzione, comprese le attività post-recovery per un recupero più veloce e una riduzione dei downtime.

Scegli Rubrik NAS Cloud Direct per gestire meglio e proteggere in tranquillità i tuoi dati non strutturati.

Vediamo un caso d'uso concreto.

INDIETRO AVANTI

Rubrik NAS Cloud Direct: vantaggi per un istituto finanziario

Akuna Capital, un istituto che opera nel settore degli hedge fund con sede a New York e Chicago, esegue un sistema di trading ad alta frequenza con 400 TB di dati archiviati, l'equivalente di 2 miliardi di file.

Nel trading ad alta frequenza, le transazioni si concludono in microsecondi, per cui ogni istante è essenziale. Le soluzioni di backup usate da Akuna Capital prima di Rubrik non consentivano di analizzare e proteggere i dati rapidamente e senza interruzioni. I backup erano lenti e inefficienti, con gravi rischi per la sicurezza delle operazioni aziendali.

Grazie all'integrazione di Rubrik NAS Cloud Direct in Pure Flashblade, Akuna Capital ora scansiona 400.000 file al secondo e completa i backup in meno di 2 ore al giorno, con una protezione completa dalle minacce informatiche.

Vuoi scegliere Rubrik?
Era quello che volevamo sentirti dire.

Cambia da oggi la sicurezza dei tuoi dati con Rubrik NAS Cloud Direct.

ULTERIORI INFORMAZIONI →

rubrik