

Software Defined Environment: la marcia in più dello Storage IBM per l'era dei Big Data

Un software di storage rivoluzionario fornisce scalabilità infinita su tutti i tipi di dati

Milano - 12 mag 2014:

IBM presenta oggi un portafoglio di prodotti Software-Defined Storage, che migliora l'economia del data center e allo stesso tempo consente alle organizzazioni di accedere ed elaborare qualsiasi tipo di dati, su qualsiasi dispositivo di storage, in qualunque parte del mondo.

Una delle tecnologie nel portafoglio è "Elastic Storage", che offre prestazioni senza precedenti, scalabilità infinita e capacità di ridurre i costi di storage fino al 90 per cento, grazie allo spostamento automatico dei dati sul dispositivo di memorizzazione più economico.

Nata nei centri di Ricerca IBM, questa nuova tecnologia rivoluzionaria consente alle imprese di sfruttare - e non solo gestire - la crescita esplosiva dei dati, in svariate forme, generati da innumerevoli dispositivi, sensori, processi di business e social network. Il nuovo software di storage è ideale per le applicazioni a maggiore intensità di dati, che richiedono accesso ad alta velocità a enormi volumi di informazioni: dall'elaborazione dei dati sismici, gestione del rischio e analisi finanziaria, modellazione meteorologica e ricerca scientifica, fino alla scelta dell'azione migliore da intraprendere in situazioni di vendita in tempo reale.

Il Software-Defined Storage è una serie di funzionalità in grado di gestire in automatico i dati a livello locale e globale, fornendo una velocità senza precedenti nell'accesso ai dati, amministrazione semplificata e scalabilità delle infrastrutture tecnologiche rapida e più efficace in termini di costi, di pari passo con l'espansione del volume dei dati. Inoltre, queste novità possono funzionare con i sistemi di qualsiasi azienda, per fornire software automatizzato e virtualizzato.

Una tecnologia rivoluzionaria

L'Elastic Storage è stato utilizzato per la sfida al quiz televisivo Jeopardy! tra Watson di IBM e due ex campioni del gioco. Per la trasmissione, Watson ha avuto accesso a 200 milioni di pagine di dati strutturati e non, incluso il testo completo di Wikipedia. Grazie all'uso dell'Elastic Storage, i cinque terabyte di "conoscenza" di Watson (ovvero 200 milioni di pagine di dati) sono stati caricati sulla memoria del computer nel giro di pochi minuti.

Un motivo chiave per la scelta di tale software per il sistema di Watson è stata la sua scalabilità: i limiti

dell'architettura arrivano infatti alle migliaia di "yottabyte". Uno yottabyte è un miliardo di petabyte, cioè l'equivalente di un data center di dimensioni pari a un milione di isolati di una città, che riempirebbero gli stati del Delaware e del Rhode Island insieme.

I ricercatori IBM hanno dimostrato che l'Elastic Storage è in grado di eseguire con successo la scansione di 10 miliardi di file, su un unico sistema, in appena 43 minuti, questa tecnologia si traduce in prestazioni ineguagliate per le aziende che analizzano enormi repository di dati per estrarre informazioni di business.

In essenza, Elastic Storage si basa su un software "global file system" di IBM per fornire gestione dello storage online, accesso scalabile e strumenti di data governance integrati, in grado di gestire vaste quantità di dati e miliardi di file. Ad esempio, Elastic Storage sfrutta anche la memoria Flash lato server, per un aumento prestazionale fino a sei volte rispetto ai dischi SAS standard. Questa funzionalità riconosce quando un server dispone dello storage Flash e lo utilizza automaticamente come memoria cache per migliorare le prestazioni.

Elastic Storage virtualizza lo storage consentendo a diversi sistemi e applicazioni di condividere pool di storage comuni. Ciò permette un accesso globale trasparente ai dati, senza la necessità di modificare le applicazioni né di ricorrere ad applicazioni supplementari di gestione storage, che spesso creano interferenze. Dal momento che Elastic Storage non si affida alla gestione centralizzata per determinare l'ubicazione e la collocazione dei file, le aziende possono contare su un accesso continuo e ad elevata disponibilità ai dati, anche in caso di guasti del software o dell'hardware.

Per il Computational and Information Services Laboratory (CISL) del National Center for Atmospheric Research la crescita del volume di dati è parte del suo DNA. L'organizzazione, che memorizza e gestisce più di 50 petabyte di informazioni tra i suoi centri in Wyoming e Colorado, si affida a Elastic Storage per offrire ai ricercatori un rapido accesso a grandi quantità di dati eterogenei.

"Forniamo servizi di dati computazionali, didattici e di ricerca per le geoscienze a oltre 1000 utenti, in più di 200 sedi diverse", commenta Pamela Gillman, manager, Data Analysis Services Group, CISL. "Il software di file system globale di IBM ha consentito un accesso scalabile, affidabile e veloce a queste informazioni. Ciò ha migliorato drasticamente le prestazioni delle diverse funzioni, così come dell'organizzazione nel suo complesso".

Un componente chiave dell'Elastic Storage è la sua capacità di trasferire i dati, in modo automatico e intelligente, al sistema di storage più strategico ed economico disponibile. Grazie a funzionalità basate sulle policy e all'analytics in tempo reale, Elastic Storage può ad esempio spostare in automatico i dati raramente utilizzati verso unità a nastro più economiche, memorizzando invece i dati di accesso frequente su sistemi Flash

ad alta velocità per un accesso più rapido. Queste funzionalità basate sulle policy possono fornire un risparmio di costi fino al 90 per cento.

Inoltre, il software è dotato di crittografia nativa e cancellazione sicura, che assicura l'irrecuperabilità dei dati, in conformità a normative quali HIPAA e Sarbanes-Oxley.

Grazie al supporto del software di gestione cloud OpenStack, Elastic Storage permette inoltre ai clienti di memorizzare, gestire e accedere ai dati attraverso cloud privati, pubblici e ibridi, per la condivisione e la collaborazione a livello globale. Oltre a supportare l'accesso a OpenStack Cinder e Swift, Elastic Storage supporta altre API open source, quali POSIX e Hadoop.

Mentre i sistemi di storage tradizionali devono spostare i dati verso sistemi designati separati per il transaction processing e l'analytics, Elastic Storage può bilanciare automaticamente le risorse per supportare entrambi i tipi di carichi di lavoro applicativi, inclusa l'analytics basata su Hadoop. Ciò velocizza enormemente l'analisi ed elimina il processo oneroso, in termini di costi e tempo, di produrre duplicati dei dati.

Il software Elastic Storage sarà disponibile anche come servizio cloud IBM SoftLayer nei prossimi mesi di quest'anno.

<https://it.newsroom.ibm.com/2014-05-12-Software-Defined-Environment-la-marcia-in-piu-dello-Storage-IBM-per-lera-dei-Big-Data>