



## La Storage Class Memory migliora le prestazioni dei carichi di lavoro chiave

**Di Peter Burris**

6 settembre 2019

Quando le aziende mettono a punto le proprie strategie IT a lungo termine, valutano come modernizzare al meglio le applicazioni tradizionali di elevato valore (HVTA) che rappresentano il cuore pulsante dell'attività e contribuiscono al 40-50% della spesa IT nelle grandi imprese. Le tecnologie hardware Storage Class Memory (SCM) e NVMe-oF renderanno più semplici gli approcci alla modernizzazione delle HVTA.



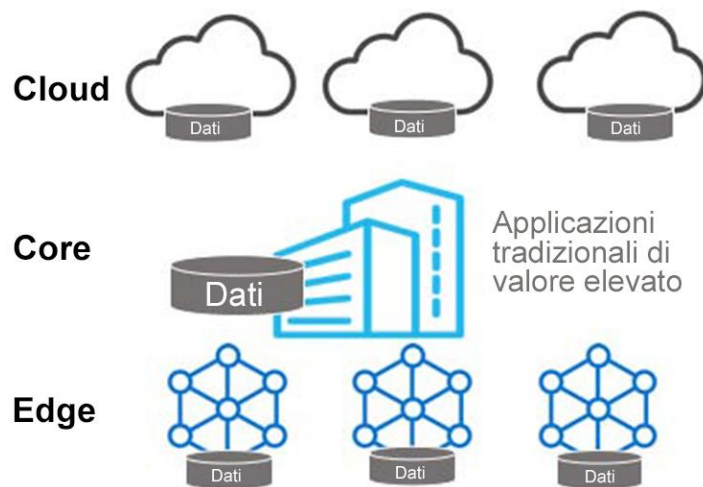
Nuovi tipi di esperienze digitali, carichi di lavoro di analisi, automazione aziendale e integrazione del mercato aumentano i carichi di dati e le transazioni, come pure la complessità del percorso nelle applicazioni tradizionali di elevato valore (HVTA). Queste nuove funzionalità stanno generando nuovi requisiti per modernizzare le applicazioni tradizionali di elevato valore (HVTA) con tecnologie che offrono opzioni superiori in termini di prestazioni e scalabilità.

Le tecnologie destinate a incrementare la produttività dell'infrastruttura HVTA, quali la virtualizzazione, il parallelismo dei carichi di lavoro e i DBMS in-memory, stanno di fatto aumentando l'utilizzo dell'hardware e il throughput aggregato. Tuttavia, possono anche generare complessità indesiderate e problemi legati ai costi all'interno dell'infrastruttura HVTA, in particolare per quanto riguarda le risorse di storage. Le prestazioni e l'affidabilità di ognuna di queste tecnologie sono estremamente sensibili alle prestazioni di I/O all'interno della gerarchia di storage, dai dispositivi esterni alla memoria di sistema. I nuovi tipi di dispositivi di storage, quali le unità SSD basate su memoria flash e NVMe Express over Fabrics (NVMe-oF), contribuiscono ad ottimizzare il bilanciamento del sistema di I/O, ma il divario di latenza tra I/O e memoria si fa sempre più ampio a causa dell'aumento delle richieste di elaborazione distribuita.

Di conseguenza, i maggiori carichi HVTA stanno generando un "effetto farfalla legacy", per cui la minima variazione della latenza dei dati all'interno della HVTA (e di altre applicazioni) può avere ripercussioni sulle prestazioni dei sistemi distribuiti che forniscono attività di automazione aziendale, sistema intelligente ed esperienza del cliente (vedere la Figura 1). Ciò è tanto più vero in quanto le applicazioni vengono sottoposte a refactoring per aumentare in modo selettivo l'agilità aziendale e sfruttare le tecnologie cloud emergenti, aumentando in tal modo la quantità e la distribuzione di macchine software indipendenti impegnate in attività fondamentali. Inoltre, man mano che le aziende passano ai container e alle relative tecnologie, l'effetto farfalla legacy non può che peggiorare.

## LO STORAGE IN SOCCORSO DELLA LATENZA

Gli odierni professionisti IT che gestiscono e utilizzano le applicazioni tradizionali di elevato valore (HVTA) su cui si basano i processi aziendali devono affrontare decisioni estremamente difficili. L'azienda vuole un'esperienza cloud perché promette maggiori opzioni per la scalabilità, i costi operativi e la creazione di valore digitale. Tuttavia, le esigenze di throughput di sistemi distribuiti sempre più complessi stanno spingendo gli architetti di sistema a minare gli ideali del cloud e sostenere gli approcci tradizionali per l'individuazione dei dati nelle risorse del sistema, inclusi i dati oggetto di accesso frequente con caching nella memoria e i dati dei volumi organizzati per il trasferimento in blocco



**Figura 1 - Una maggiore distribuzione dei dati e dell'elaborazione aumenta la sensibilità alle latenze HVTA**

dei dispositivi di storage. Poiché le nuove classi di applicazioni, quali i sistemi di intelligenza artificiale data-first, possono richiedere accesso in lettura frequente ai dati in blocco (il che accentua i problemi di latenza dei dati), gli approcci tradizionali alla mappatura degli schemi di dati fisici alle gerarchie di storage tendono a rivelarsi inefficaci.

Gli odierni professionisti IT che gestiscono e utilizzano le applicazioni tradizionali di elevato valore (HVTA) su cui si basano i processi aziendali devono affrontare decisioni estremamente difficili. L'azienda vuole un'esperienza cloud perché promette maggiori opzioni per la scalabilità, i costi operativi e la creazione di valore digitale. Tuttavia, le esigenze di throughput di sistemi distribuiti sempre più complessi stanno spingendo gli architetti di sistema a minare gli ideali del cloud e sostenere gli approcci tradizionali per l'individuazione dei dati nelle risorse del sistema, inclusi i dati oggetto di accesso frequente con caching nella memoria e i dati dei volumi organizzati per il trasferimento in blocco dei dispositivi di storage. Poiché le nuove classi di applicazioni, quali i sistemi di intelligenza artificiale data-first, possono richiedere accesso in lettura frequente ai dati in blocco (il che accentua i problemi di latenza dei dati), gli approcci tradizionali alla mappatura degli schemi di dati fisici alle gerarchie di storage tendono a rivelarsi inefficaci.

Inoltre, la migrazione delle HVTA agli ambienti di public cloud in genere non è un processo semplice. Ciò dipende in parte dal fatto che le HVTA sono state ottimizzate per ridurre al minimo l'impatto della latenza dei dati, un processo che può coprire anche un decennio di implementazioni tecnologiche, con conseguenti ostacoli pratici alla migrazione. [La nostra ricerca mostra che le imprese stanno valutando un'ampia gamma di approcci alla modernizzazione delle HVTA.](#) Tuttavia, prevediamo che un approccio sarà particolarmente fruttuoso per almeno i prossimi 3-5 anni, ovvero investire in un'infrastruttura in grado di migliorare l'affinità con il cloud.

Esistono numerose tecnologie innovative in grado di supportare i moderni software di applicazioni e sistemi orientati al cloud, consentendo loro di gestire al meglio i requisiti di dati dei carichi di lavoro dell'intelligenza artificiale data-first. Tra queste, spicca in particolare la Storage Class Memory (SCM). SCM aggiunge alla gerarchia di storage un nuovo tier in grado di offrire prestazioni simili a quelle della memoria con costi analoghi a quelli dello storage.

SCM riduce le latenze dei dati fornendo accesso indirizzabile ai byte di tipo memoria ai dati, riducendo così la necessità di leggere ed eseguire operazioni di garbage collection su blocchi di dati di grandi dimensioni. Tuttavia, a differenza delle memorie di sistema (e come i dispositivi di storage), SCM è persistente. Grandi quantità di dati indirizzabili ai byte possono essere mappati nelle risorse sul lato CPU e resi accessibili tramite le istruzioni di caricamento e storage (load-store). Le lunghezze dei percorsi di I/O ai dati cruciali sono drasticamente ridotte quando i dati vengono posizionati in SCM. I test eseguiti da Dell e Intel utilizzando dispositivi SSD Intel Optane a porta doppia mostrano un miglioramento del 1.000% a livello di unità rispetto alle unità flash NAND. Si tratta di un notevole miglioramento che permette di ridurre i tempi di accesso a poche centinaia di nanosecondi, una gamma di prestazioni che contribuisce a migliorare l'effetto farfalla delle HVTA e apre nuove possibilità in termini di aggiunta di esperienze di tipo cloud alle HVTA stesse.

## FASE SUCCESSIVA: SCM NEI SISTEMI DI STORAGE E RELATIVE PRATICHE

L'introduzione di una nuova e potente tecnologia di storage come SCM nelle operazioni esistenti richiede un grado di integrazione che sfrutta la tecnologia senza aumentare esponenzialmente la complessità e compromettere la sicurezza dei dati. Negli ultimi anni, le soluzioni di storage hanno sposato i dispositivi SSD e NVMe, e ciò ha determinato un approccio relativamente fluido al miglioramento delle prestazioni di storage e alla flessibilità. Tuttavia, SCM richiederà un livello di integrazione ancora superiore per fare in modo che applicazioni, sviluppatori di applicazioni e System Engineer possano sfruttarne appieno le prestazioni e la scalabilità.

Crediamo che la famiglia PowerMax sviluppata da Dell EMC di sistemi di storage purpose-built con prestazioni elevate sarà la prima soluzione in cui SCM Intel Optane a porta doppia sarà progettata in modo olistico come storage tier persistente, SSD e NVMe-oF. Quando è stata annunciata, nel settembre 2019, ci aspettavamo che PowerMax avrebbe dato vita a sistemi di storage in grado di fornire 15 milioni di IOPS, banda larga aggregata di 350 GB/sec, tempi di risposta per i dati migliori del 50% per i carichi di lavoro moderni e un significativo passo in avanti in termini di fattori ambientali. Per semplificare l'accesso alla tecnologia SCM, PowerMax offre livelli di servizio definiti dall'utente per prestazioni prevedibili e coerenti.

Di conseguenza, le organizzazioni IT avranno accesso a nuovi approcci per l'adozione congiunta di nuove applicazioni e pratiche. La famiglia PowerMax deve ampliare le opzioni per il consolidamento dei formati di storage di file e blocchi su un sottosistema di storage comune, in particolare migliorando la compatibilità tra lo storage HVTA e altre classi di storage delle applicazioni. Oltre a semplificare l'amministrazione dello storage, questo accorgimento tecnico dovrebbe garantire nuovi livelli di integrazione tra i dati e aprire nuovi vettori di applicazioni per migliorare l'esperienza del cliente, l'automazione e i modelli di business digitale.

## AZIONE

La tecnologia Storage Class Memory (SCM) combina prestazioni simili a quelle della memoria e prezzi analoghi a quello dello storage con una grande facilità d'uso. Questa combinazione sarà essenziale per integrare a livello economico e strategico le applicazioni tradizionali di elevato valore (HVTA) che sono alla base del business e le applicazioni emergenti data-first quali l'intelligenza artificiale, che rappresentano il futuro. I CIO che desiderano offrire un'esperienza cloud per le estensioni HVTA hanno tutto l'interesse a prendere in esame soluzioni quali la famiglia PowerMax di Dell EMC in quanto early enabler di tecnologie SCM quali gli SSD Intel Optane.

Peter gestisce Wikibon ed è responsabile del dipartimento di ricerca presso SiliconANGLE Media. La sua ricerca si concentra sulla trasformazione digitale del business, sul valore dei dati e sull'ottimizzazione del cloud. Vanta oltre 30 anni di esperienza come professionista IT, analista, leader di ricerca ed executive in tecnologia e business. Prima di entrare a far parte di Wikibon, Peter è stato responsabile dei team di ricerca B2B CMO e CIO presso Forrester, Research Director aggiunto presso META Group (ora Gartner) dove ha messo a punto l'importante programma Adaptive Infrastructure Strategies e ha lavorato come leader IT e di marketing presso altre grandi società globali. È un relatore di fama sulle tendenze dell'infrastruttura tecnologica e sulla trasformazione digitale del business. Si è laureato all'Università di Yale e ora vive nella Silicon Valley, in California.



**Peter Burris**

@plburris

peter@siliconangle.co