

White Paper

Accelerare la modernizzazione AI attraverso l'infrastruttura dati

Sponsorizzato da: IBM Corporation

Ashish Nadkarni Sriram Subramanian Matt Leib
Febbraio 2021

EXECUTIVE SUMMARY

L'intelligenza artificiale (AI), il machine learning (ML) e, spesso, le funzionalità di deep learning (DL) sono componenti essenziali delle iniziative di trasformazione digitale (DX). Le opportunità di business che si possono conseguire con l'AI sono estremamente promettenti. Le aziende sono sempre più consapevoli del fatto che non approfittare dell'AI potrebbe rappresentare una potenziale catastrofe di business, in quanto i competitor riescono a ottenere una quantità di informazioni e possibilità mai avute prima per far crescere e soddisfare i propri clienti. Poche aziende, per non dire nessuna, oggi pensano ancora "l'AI non fa per noi" oppure "l'AI è solo una montatura". Al contrario, iniziative consistenti di AI sono state intraprese in tutto il mondo, in tutti i settori e in aziende di qualsiasi dimensione.

Le aziende stanno cercando una trasformazione e modernizzazione basata sull'AI, il che comporta il passaggio dalla sperimentazione alla generazione di valore per il business a fronte degli investimenti nell'AI. Il successo degli investimenti di business nella trasformazione digitale associata all'AI è direttamente legato al livello di competenza richiesto per sviluppare, implementare e mantenere le soluzioni AI su larga scala. Le linee di business di molte aziende, il personale IT, i data scientist e gli sviluppatori hanno lavorato per studiare l'intelligenza artificiale, individuare i casi d'uso, definire una strategia di AI per il loro business, avviare le prime iniziative e sviluppare e testare le applicazioni AI: queste, utilizzando algoritmi di machine learning e soprattutto di deep learning, forniscono nuovi insight e competenze.

Man mano che le aziende implementano queste iniziative, emergono nuove domande. Esse sono consapevoli - anzi, potrebbero aver sperimentato personalmente - che non possono usare le infrastrutture di calcolo standard generiche e l'infrastruttura di archiviazione esistente o legacy. Inoltre, si rendono conto che sia la formazione AI (la formazione del modello AI) che l'inferenza AI (l'utilizzo del modello per capire o prevedere un evento) richiedono tipi di calcolo ridimensionabili diversi e un'infrastruttura di storage altrettanto ridimensionabile.

IDC ha rilevato che le aziende che cercano di utilizzare le infrastrutture esistenti senza modernizzarle, tutte o in parte, hanno una maggiore probabilità di fallimento. Anche nell'ambito della modernizzazione, l'attenzione varia da azienda ad azienda. Nonostante molte riescano a gestire meglio il proprio sistema di calcolo, spesso sottovalutano il valore dello storage per l'AI. Inoltre, le applicazioni AI e in particolare i sistemi di deep learning, che analizzano

IDC ha rilevato che le aziende che cercano di utilizzare le infrastrutture esistenti senza modernizzarle, tutte o in parte, hanno una maggiore probabilità di fallimento. Anche nell'ambito della modernizzazione, l'attenzione varia da azienda ad azienda. Nonostante molte riescano a gestire meglio i sistemi di calcolo, spesso sottovalutano il valore dello storage per l'AI.

quantità di dati sempre maggiori, sono estremamente sofisticati e richiedono potenti capacità di elaborazione parallela basate su un gran numero di core computazionali. I sistemi di storage standard non riescono a garantire che tali attività vengano eseguite in modo adeguato. Infine, è importante che le iniziative legate all'AI siano parte integrante degli interventi di modernizzazione delle applicazioni che includono Kubernetes e/o container e l'integrazione con uno o più servizi cloud pubblici attraverso un'architettura cloud ibrida.

La ricerca IDC mostra che, in termini di infrastruttura di storage, una scarsa o inadeguata attenzione ai dettagli può far fallire rapidamente le iniziative di trasformazione AI. Per superare questo divario, le aziende che hanno testato l'infrastruttura esistente e sono ora pronte a portarla in produzione devono rinnovare la propria infrastruttura per ottenere le prestazioni di elaborazione parallela richieste e devono farlo investendo in soluzioni di storage più moderne che permettano una scalabilità massiccia e siano integrabili con il cloud, i container e i sistemi di calcolo ad alte prestazioni, sia per la distribuzione a livello globale che per l'accesso ai dati. È qui che soluzioni come IBM Spectrum Scale e IBM Elastic Storage System (ESS) forniscono i componenti necessari per un'architettura informatica AI in quanto sono adatti per i carichi di lavoro AI e la distribuzione sia containerizzata che cloud ibrida, specificamente focalizzata su workload AI.

SITUATION OVERVIEW

L'AI e il ML sono qui e ora

Le aziende di tutto il mondo stanno reagendo attivamente alle nuove opportunità offerte dagli investimenti nell'AI per accelerare le proprie iniziative di trasformazione digitale. L'intelligenza artificiale è un insieme di tecnologie che utilizzano l'elaborazione del linguaggio naturale (NLP - natural language processing), l'analisi di immagini e video, il machine learning, i diagrammi di conoscenza e altre tecnologie per rispondere a domande, individuare informazioni e fornire raccomandazioni. Questi sistemi ipotizzano e formulano possibili risposte sulla base delle conoscenze disponibili, possono essere formati grazie alla grande quantità di contenuti e si adattano e imparano dai propri errori e fallimenti attraverso il retraining o la supervisione umana. IDC prevede che entro il 2022 almeno il 60% di questi casi d'uso incentrati sull'intelligenza artificiale saranno implementati in almeno il 65% delle aziende Global 2000, con una crescita del 34% rispetto al 2019.

IDC prevede che entro il 2022 almeno il 60% dei casi d'uso incentrati sull'AI saranno implementati in almeno il 65% delle aziende Global 2000, con una crescita del 34% rispetto al 2019.

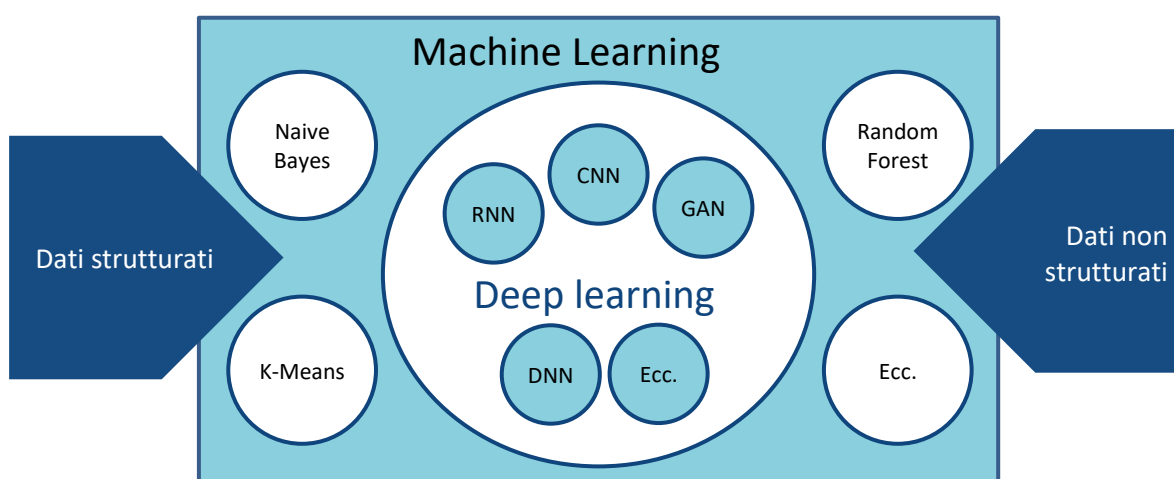
L'AI sta diventando sempre più diffusa in tutte le aziende sia nell'automazione dei processi che dei flussi di lavoro. Nel 2019 IDC ha esaminato 176 casi d'uso di trasformazione digitale in otto aree funzionali delle line-of-business, tra cui l'esperienza del cliente (CX), la strategia legale e aziendale, le facility e il procurement, stimando che circa il 26% di questi casi d'uso erano dipendenti dall'AI ed implementati nella maggior parte delle aziende.

Questo significa che presto la maggior parte delle aziende leader sfrutterà tecnologie AI quali l'elaborazione del linguaggio naturale, il machine learning, il deep learning e lo speech to text, in tutta l'organizzazione al fine di scalare le operation, sfruttare i dati non strutturati e fornire insight di business rilevanti. Intanto le aziende che non hanno ancora stabilito come portare i casi d'uso basati sull'AI dal PoC (Proof of Concept) alla produzione resteranno ancora più indietro, ampliando il divario digitale.

Il machine learning (ML) è un sottoinsieme delle tecniche AI che permette ai sistemi informatici di apprendere e migliorare il proprio comportamento relativamente ad un determinato compito senza la necessità di essere programmati da una persona. I modelli di machine learning sono algoritmi che possono migliorare nel tempo, testando sé stessi ripetutamente, utilizzando grandi quantità di dati sia strutturati che destrutturati, fino a quando si ritiene che abbiano "appreso" un compito, come, ad esempio, riconoscere un volto umano. La figura 1 illustra come il deep learning (DL) sia un sottoinsieme del ML. Le architetture DL tipiche sono deep neural network (DNN), convolutional neural network (CNN), recurrent neural network (RNN), generative adversarial network (GAN) e molte altre.

FIGURA 1

Possibili applicazioni del Machine Learning e del Deep Learning



Fonte: IDC, 2021

Le piattaforme software di AI includono software di AI conversazionale quali, ad esempio: assistenti digitali; predictive analytics per trovare delle relazioni nascoste nei dati e fare previsioni; text analytics e linguaggio naturale per riconoscere, capire ed estrarre informazioni da un testo; voice/speech analytics per riconoscere, identificare ed estrarre informazioni da audio, voci e discorsi; analisi di immagini e video per riconoscere, identificare ed estrarre informazioni, compreso il riconoscimento di modelli, oggetti, colori e altri elementi come persone, volti, emozioni, auto e paesaggi.

Molte aziende sono a buon punto con le iniziative di AI e hanno raggiunto una fase in cui sono pronte a iniziare a implementare l'AI a livello di produzione. Altre realtà stanno ancora sperimentando l'AI, mentre un terzo gruppo di aziende sta attualmente valutando ciò che le applicazioni di AI possono significare per la propria realtà.

Per quanto riguarda il primo gruppo (quello pronto per l'implementazione), IDC sta riscontrando una serie di casi d'uso dell'AI che le imprese, i governi e altre organizzazioni hanno iniziato a implementare. I cinque casi d'uso più comuni oggi, classificati in base all'importo che le aziende spendono per la loro implementazione in termini di hardware, software e servizi, sono:

- **Operatori automatizzati del servizio clienti.** Nel settore bancario, ad esempio, queste applicazioni di AI forniscono un servizio al cliente attraverso un programma di apprendimento che è in grado di comprendere le esigenze e i problemi dei clienti e aiutare una banca a ridurre il tempo e le risorse necessarie per risolvere i problemi dei clienti. Questi agenti automatizzati stanno diventando ampiamente utilizzati in tutte le aziende.
- **Recommendation e automazione dei processi di vendita.** Utilizzate in vari settori, queste applicazioni AI si integrano con i sistemi che gestiscono le relazioni con i clienti (CRM - customer relationship management) per capire il comportamento dei clienti in tempo reale e suggerire azioni adeguate agli agenti di vendita.
- **Sistemi automatizzati di intelligence e prevenzione delle minacce.** Queste applicazioni AI elaborano rapporti di intelligence, estraggono informazioni da essi, stabiliscono relazioni tra diverse informazioni e poi identificano minacce a database, sistemi, siti web e così via, costituendo in tal modo una componente fondamentale nella prevenzione delle minacce per i governi e per le imprese.
- **Analisi e indagini sulle frodi.** Nel settore assicurativo, ma ampiamente utilizzato anche altrove, queste applicazioni di AI utilizzano l'apprendimento basato su regole per identificare le transazioni fraudolente e apprendono automaticamente come identificare vari tipi di frode in ambito assicurativo.
- **Manutenzione preventiva automatizzata.** Nelle industrie manifatturiere queste applicazioni di AI sono basate su algoritmi di apprendimento automatico che costituiscono un modello predittivo accurato di potenziali guasti di impianti e macchinari, riducendo i tempi di fermo e i costi di manutenzione.

Altri casi d'uso dell'AI che hanno guadagnato terreno nelle imprese (classificati in ordine di spesa in hardware, software e servizi) sono:

- Consulenza sui programmi e sistemi di recommendation
- Sistemi di diagnosi e trattamento
- Automazione intelligente dei processi
- Sistemi di investigation e recommendation per la gestione della qualità
- Automazione IT e assistenti digitali per lavoratori aziendali qualificati
- Consulenti esperti di shopping e raccomandazioni di prodotti
- Intelligence normativa, della supply chain e logistica
- Asset/fleet management e gestione automatizzata dei reclami
- Gemello digitale (digital twin) / Simulazione digitale avanzata
- Sicurezza pubblica e risposta alle emergenze
- Apprendimento adattivo
- Rete intelligente
- Gestione di merci, beni e flotte
- Ricerca e scoperta farmaceutica

La trasformazione AI richiede una solida base dati

Le iniziative di ML e DL si basano molto su una combinazione di vari input di dati strutturati e non strutturati (vedere figura 1). I progetti di IA coinvolgono più fasi, tra cui la raccolta dei dati, la configurazione, la depurazione, la verifica, la costruzione del modello, la formazione, il test, l'inferenza e il prelievo dei dati. Dato che ogni fase ha diversi requisiti, dall'accesso più veloce, allo storage a

bassa latenza, allo storage di archiviazione a basso costo, le imprese hanno bisogno di selezionare un'infrastruttura di storage appropriata ed economicamente vantaggiosa. Esse hanno anche bisogno di utilizzare una serie disparata di strumenti per gestire il set di dati AI attraverso il loro ciclo di vita.

Con il rollout del 5G e la proliferazione dei sensori basati sull'IoT, vengono generati e consumati più dati. Telecamere e assistenti intelligenti sono distribuiti sul dispositivo per fornire una risposta più veloce e una maggiore esperienza utente. La necessità di un'AI locale, sviluppata ai margini della rete o sugli endpoint, è in crescita. Le applicazioni AI sensibili alla latenza che girano su dispositivi periferici con connettività limitata richiederanno un'elevata scalabilità. Saranno adottati casi d'uso di edge computing basati sull'AI.

Un'adeguata infrastruttura dati accelera la trasformazione basata sull'AI

Man mano che le aziende investono nell'AI, alcuni fattori, come la progettazione e l'implementazione dell'infrastruttura, acquisiscono maggiore importanza. Ispirandosi al modello dei fornitori di servizi hyperscale e cloud, le organizzazioni stanno avvicinandosi ai requisiti infrastrutturali attraverso la creazione di un'infrastruttura dati unificata. Un'infrastruttura dati è una base comune per le iniziative di IA che include un livello di calcolo e di archiviazione altamente efficiente e scalabile. Piuttosto che approssciare i workload AI come omogenei, un'infrastruttura dati li tratta in modo multiforme, collegando una porzione di quel workload al giusto livello di calcolo, che viene supportato da un livello di storage appropriato a seconda che l'insieme di dati sia strutturato o meno. Questo approccio composito accelera la trasformazione AI in tre fasi:

Difficoltà legate alle iniziative AI

La ricerca IDC mostra che spesso le aziende, nell'ambito delle loro iniziative AI, si trovano ad affrontare le seguenti difficoltà nella gestione dei dati:

- Cicli di inserimento e di preparazione dei dati troppo lunghi
- Silos di infrastrutture per i vari casi d'uso degli analytics
- Copie multiple degli stessi dati senza una fonte di verità univoca
- Necessità di gestire e proteggere in modo sicuro la provenienza dei dati per la ripetibilità
- Bisogno di accessibilità globale (cloud ibrido) e collaborazione
- Integrità dei dati dopo che sono stati raccolti e catalogati

- **Scalabilità.** La dimensione della portata descrive in che misura opera il workload. Le sottodimensioni fondamentali - calcolo, rete e persistenza dei dati (archiviazione) - sono tutte legate all'hardware. Fondamentalmente, le sottodimensioni legate al software, come l'orchestrazione, stanno guadagnando una certa importanza ai fini del mantenimento dell'equilibrio con un aumento delle dimensioni e della complessità dello stack.
- **Portabilità.** Questa è la capacità del workload di essere spostato attraverso le implementazioni core, edge ed endpoint. Oggi molti di questi workload sono di natura statica (cioè, progettati per essere eseguiti in una singola distribuzione). Le aziende stanno cercando sempre più spesso di sviluppare workload in un ambiente (ad esempio, il cloud pubblico) e di installarli (in produzione) in un altro (ad esempio, l'edge). Questo è analogo al modello attuale di sviluppo e distribuzione delle applicazioni mobili.
- **Tempo.** Questa misura si riferisce alla continuità temporale del workload stesso. Molti workload AI prendono spunto dal calcolo ad alte prestazioni o dalla distribuzione di big data e analytics - sono progettati per essere di natura batch. Sempre più spesso - e grazie alla proliferazione di acceleratori ad alte prestazioni - i workload AI possono analizzare i dati in streaming in tempo reale o quasi.

Miti e necessità delle infrastrutture dati AI

Un fondamento essenziale, anche se spesso trascurato, è l'infrastruttura di archiviazione. L'implementazione su larga scala della tecnologia AI spesso richiede maggiori sforzi all'infrastruttura di storage in termini di capacità (crescita) e prestazioni (IOPS e larghezza di banda). Le aziende spesso ritengono che lo storage interno del server o lo storage aziendale utilizzato per altri workload sia sufficiente per l'esecuzione delle applicazioni AI. E una volta che l'infrastruttura è costruita, si rendono conto che lo storage è l'anello più debole della catena. Ogni applicazione basata sull'AI comporta un insieme diverso di requisiti e, quindi, di problematiche per il personale IT. Di conseguenza, sia gli acquirenti che i venditori IT dovrebbero evitare lo stereotipo tradizionale "dato che ho un martello, tutto è un chiodo".

Un atteggiamento più oculato è quello di adottare un approccio olistico all'infrastruttura dei dati. Mentre la persistenza dei dati e i meccanismi d'accesso restano dei punti fermi, le aziende hanno bisogno di espandere la propria visione includendo la rete e l'integrazione tra i livelli di calcolo, il software di archiviazione e i sistemi. Le aziende devono adottare una mentalità coerente con un'infrastruttura dati end-to-end e non solo una che coinvolge "l'inserimento" di un altro sistema di storage. IDC ritiene che i requisiti dell'infrastruttura dati riguardanti in particolare lo storage possano essere sintetizzati nelle aree chiave discusse nelle sezioni che seguono.

Integrazione del calcolo

Si presume spesso - e questo è un presupposto errato - che tutti i workload AI siano containerizzati. Al contrario, molti workload AI sono eseguiti su bare metal o addirittura virtualizzati. In particolare, le applicazioni AI-enabled vengono spesso eseguite su bare metal o su computer virtualizzati. Molti workload AI sono stati ottimizzati per utilizzare gli acceleratori. Questo non significa che tutti i workload AI vengano eseguiti al meglio con un'elaborazione accelerata - in generale l'elaborazione accelerata porta con sé un diverso insieme di difficoltà per i workload.

Persistenza e accesso ai dati

Se cambiano i requisiti di calcolo per i workload AI, cambiano anche i relativi requisiti di persistenza dei dati. Un aspetto sottovalutato e frainteso dello stack dei workload AI è il livello di persistenza dei dati. Spesso si presume - ancora una volta erroneamente - che tutti i carichi di lavoro AI richiedano una grande quantità di storage ad alte prestazioni. Il fatto è che non tutti i workload AI sono "grandi insiemi di dati" - ad esempio potrebbero essere campionati numerosi piccoli insiemi di dati contemporaneamente e per un breve periodo di tempo. Allo stesso modo, i workload bare metal che girano su piattaforme di calcolo a sistemi aperti spesso usano l'accesso a blocchi o file in scale-out. Non è molto raro che i workload virtualizzati vengano eseguiti su un'infrastruttura iperconvergente (HCI - hyperconverged infrastructure).

L'accesso multiprotocollo di fatto è un mix di dati strutturati e destrutturati, che vengono incorporati nell'infrastruttura di storage. Molti dispositivi IoT ed edge comunicano tramite SMB o NFS, e alcuni usano S3. In alcuni casi l'accesso ai dati in streaming è necessario. E in alcuni casi si potrebbe usare anche un client di file system parallelo nativo.

Scaling e tiering

Per supportare le applicazioni AI e ML i sistemi di storage devono fornire prestazioni su scala. Nel caso di archivi di dati non strutturati, si tratta di sistemi di archiviazione che fanno uso di file system paralleli con accesso alla rete. Per i dati strutturati, è l'uso di sistemi di archiviazione basati su flash. Lo scaling è

essenzialmente l'aumento o la diminuzione delle prestazioni e della capacità in maniera indipendente l'una dall'altra per soddisfare i requisiti delle applicazioni AI e ML.

Inoltre, per un'infrastruttura a prova di futuro, il sistema deve anche essere in grado di gestire in modo semplice e conveniente i dati che sono più vecchi o più scadenti su un object storage a basso costo con un'interfaccia di object storage come la S3.

Software-Defined Storage

AI e ML fungono da catalizzatori per il software-defined storage. Essi abilitano il codice dell'infrastruttura e l'automazione attraverso uno strato di controllo software eterogeneo sopra l'hardware. Questo favorisce una migliore integrazione con i flussi di lavoro AI/ML, assicurando così che lo storage si adatti perfettamente alle esigenze dell'applicazione.

Agilità e flessibilità di distribuzione

Le applicazioni che trattano i casi d'uso possono essere sviluppate su misura da un'organizzazione, possono essere basate su software AI commerciale o possono essere fornite come AI SaaS. Le opzioni di implementazione per il software commerciale sviluppato su misura sono on-premise, nel cloud su IaaS, o come cloud ibrido, dove l'ambiente on-premise interagisce con un ambiente cloud pubblico attraverso un livello di automazione e orchestrazione comune.

Data la natura distribuita dell'AI è ragionevole pensare che sia meglio spostare il calcolo più vicino al luogo in cui i dati vengono raccolti o generati piuttosto che il contrario. In tempi recenti, il modello core-edge-endpoint è diventato un modo de facto di descrivere l'AI (dove il core include il cloud e gli endpoint includono l'embedded intelligence). È importante notare che i profili del workload per ogni luogo variano, e di conseguenza cambiano i requisiti dell'infrastruttura sottostante.

A seconda dei vari scenari di implementazione, si devono prendere in considerazione soluzioni per:

- **Elaborare in modo sicuro il volume di dati necessario per la formazione di modelli AI con prestazioni estremamente elevate.** I requisiti di performance per la formazione del deep learning implicano la capacità di eseguire un'elaborazione massiccia e parallela utilizzando GPU combinate con un'ingestione di dati ad alta larghezza di banda.
- **Elaborare in modo sicuro il volume di dati su cui il modello AI eseguirà l'inferenza con prestazioni estremamente elevate.** Le performance di inferenza riguardano la capacità di elaborare i dati in entrata attraverso il modello AI istruito e fornire approfondimenti o decisioni AI quasi in tempo reale.

Per i data scientist e gli sviluppatori, a volte può essere più facile avviare un'iniziativa AI nel cloud, risparmiando loro di dover organizzare il calcolo on-premise che tipicamente, per il deep learning, ha bisogno di essere accelerato. Istanze cloud AI accelerate sono disponibili nella maggior parte dei cloud pubblici, di solito con stack AI open source. Naturalmente, con le istanze accelerate del cloud per l'addestramento AI, il fornitore di servizi cloud (SP - service provider) impone ciò che è disponibile per l'utente finale in termini di processori, coprocessori, interconnessioni, dimensioni della memoria, larghezza di banda I/O e così via. Non tutti i Cloud Service Provider offrono le combinazioni migliori di questi componenti, che alla fine determinano la velocità e la qualità con cui i data scientist possono sviluppare modelli di addestramento. Ne consegue che molte aziende optano per implementazioni on-premise.

Durante i loro esperimenti di AI negli ultimi anni, molte aziende si sono trovate a "sbattere contro il muro" con la loro infrastruttura standard o con le istanze di base del cloud. L'addestramento dei modelli richiedeva troppo tempo e l'inferenza era troppo lenta. La ricerca IDC mostra che il 77,1% degli intervistati dice di essersi imbattuto in una o più limitazioni con la propria infrastruttura AI on premise e il 90,3% degli intervistati dice di essersi imbattuto in limitazioni di calcolo nel cloud.

ARCHITETTURA GLOBALE DELLE INFORMAZIONI AI ESTENDIBILE GRAZIE ALL'UTILIZZO DI IBM STORAGE

Le soluzioni IBM Storage per i dati e l'AI consentono ai clienti di implementare senza soluzione di continuità iniziative AI su scala di produzione in ambienti cloud ibridi. IBM continua a mantenere la sua leadership per i workload scalabili ad alte prestazioni, così come per lo storage ad alta capacità, efficiente, sicuro e scalabile e per soluzioni AI e big data ad alte prestazioni. Il portfolio di storage di IBM fornisce uno storage integrato e la gestione dei dati provenienti dall'edge, dal core datacenter e dal cloud pubblico, accelerando in tal modo la modernizzazione dell'AI. Si presenta pronto con un ampio supporto e integrazione con contenitori Kubernetes e la piattaforma Red Hat OpenShift e può essere distribuito e accessibile nel cloud pubblico o per i workload del datacenter. IBM Storage per i dati e l'AI punta a ridurre la complessità e i costi fornendo una maggiore integrazione con un'architettura di informazioni AI che può essere distribuita su scala per l'intera azienda.

IBM Spectrum Scale

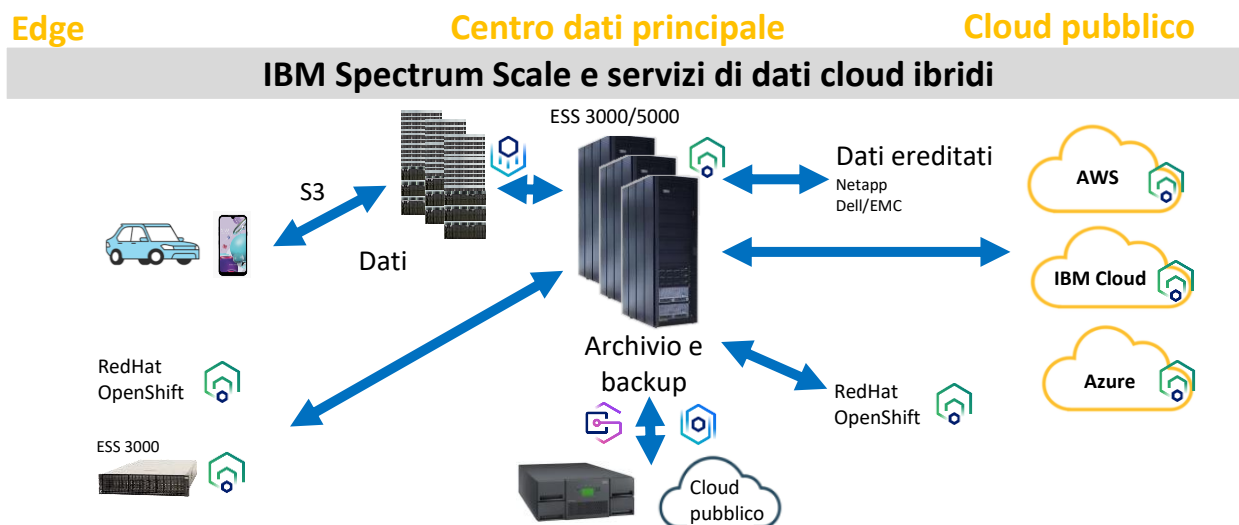
IBM Spectrum Scale è costruito su un'architettura di calcolo distribuita, progettata per qualsiasi workload ad alte prestazioni come AI/ML, modellazione, simulazione e analisi. Si tratta di un file system parallelo e organizzato in cluster che riduce l'opex attraverso una gestione semplice e la scalabilità e il capex attraverso l'ottimizzazione dei dati basata sulle policy e la gestione trasparente del ciclo di vita dei dati. IBM Spectrum Scale fornisce un accesso simultaneo a un singolo file system o a un insieme di file system da più nodi. I nodi possono essere collegati direttamente, collegati in rete, un mix di collegamenti diretti e in rete, o in una configurazione di cluster non condivisa. Questa soluzione scalabile e la piattaforma ad alta disponibilità permettono un accesso condiviso ad alte prestazioni ai set di dati comuni.

IBM Spectrum Scale supporta la replica dei dati, la gestione dello storage basata su policy e le operazioni multi-sito. I team operativi IT possono creare un cluster di nodi container Kubernetes, nodi IBM AIX, nodi IBM Z o LinuxONE, nodi Linux, nodi Microsoft Windows Server, o un mix di tutti e cinque. IBM Spectrum Scale può essere eseguito su istanze virtualizzate o containerizzate, fornendo un accesso comune e condiviso ai dati negli ambienti, sfruttando il partizionamento logico o altri hypervisor. Più cluster IBM Spectrum Scale possono condividere i dati all'interno di un luogo o attraverso connessioni WAN (Wide Area Network) per consentire la collaborazione e l'accesso globale ai dati. IBM Spectrum Scale offre una base robusta per l'infrastruttura AI sviluppata in anni di servizio all'industria informatica ad alte prestazioni (vedi Figura 2).

IBM Spectrum Scale offre una base solida per l'infrastruttura AI costruita in anni di servizio all'industria del calcolo ad alte prestazioni.

FIGURA 2

IBM Spectrum Scale e servizi di dati cloud ibridi



Fonte: IDC, 2021

IBM Spectrum Scale fornisce un namespace globale, accesso condiviso al file system tra i cluster IBM Spectrum Scale, accesso simultaneo ai file da più nodi, alta ripristinabilità e disponibilità dei dati attraverso la replica, possibilità di apportare modifiche mentre un file system è montato e amministrazione semplificata anche in ambienti di grandi dimensioni. I fattori chiave di differenziazione di IBM Spectrum Scale sono:

- Accesso condiviso al file system tra i cluster IBM Spectrum Scale per consentire la condivisione dei dati tra cluster separati all'interno di una sede o attraverso una WAN;
- Prestazioni di sistema migliorate grazie a un file system parallelo brevettato che migliora le prestazioni del sistema;
- Coerenza dei file attraverso l'accesso concorrente e dettagliato ai client in tutto il cluster utilizzando la gestione dei token;
- Maggiore disponibilità e affidabilità dei dati attraverso funzionalità come la tracciabilità del file system e funzionalità configurabili come i supporti intelligenti che possono estendersi su lunghe distanze;
- Flessibilità di sistema migliorata che permette l'aggiunta o l'eliminazione di risorse del disco o del server mentre il file system è installato;
- Gestione semplificata dello storage che aiuta a realizzare una gestione ottimale del ciclo di vita delle informazioni (ILM - information life-cycle management) attraverso una potente e automatizzata gestione dello storage, basata su policy, da flash a HDD, al cloud, al nastro e persino alla riduzione dei dati basata su policy;
- Gestione semplificata tramite numerose interfacce standard di file system che possono essere eseguite direttamente dalla maggior parte delle applicazioni;
- Implementazione del cloud ibrido che garantisce la disponibilità dei dati, l'integrità, la sicurezza e l'archiviazione ottimizzata nativa per container e l'integrazione con Red Hat OpenShift.

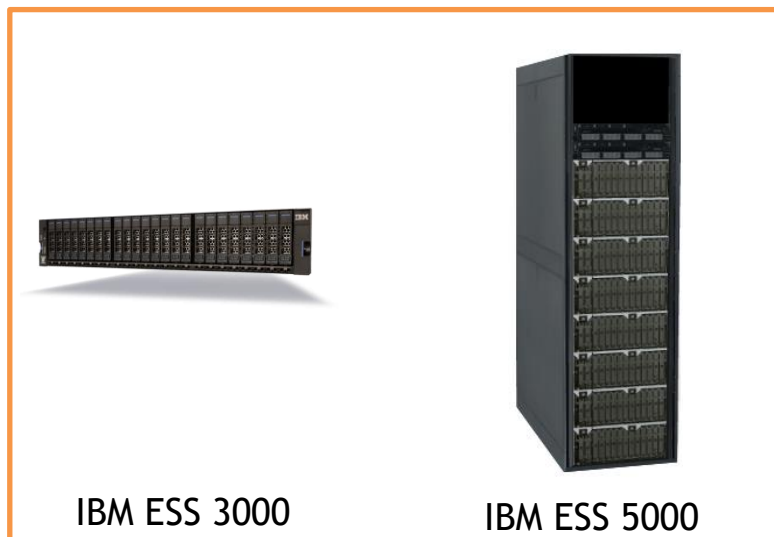
IBM Elastic Storage System

IBM Elastic Storage System (ESS) è una moderna applicazione di storage software-defined progettata per facilitare la configurazione e la gestione dei building block, rendendo più semplice per i dipartimenti IT l'implementazione di uno storage veloce e altamente scalabile per applicazioni di calcolo ad alta intensità di prestazioni, tra cui applicazioni di AI, big data e analytics (vedi Figura 3). IBM ESS:

- È costruito con lo storage flash NVMe per offrire una scalabilità a livello di exabyte e una qualità di servizio costante in tutta l'infrastruttura;
- Può essere integrato con la gestione dei file e le capacità dei servizi dati di IBM Spectrum Scale per fornire un sistema di memoria globale federato;
- Consente alle aziende di consolidare i requisiti di storage dall'edge ai dati core e di integrarsi con il cloud pubblico, riducendo così le inefficienze, abbassando i costi di acquisizione, semplificando la gestione dello storage ed eliminando i silos di dati;
- Offre un accesso coerente ad alte prestazioni per più applicazioni distribuite ed esigenti come bare metal o in ambienti virtualizzati (supporta l'integrazione di Kubernetes e Red Hat OpenShift, rendendo più facile la distribuzione di applicazioni cloud-native).

FIGURA 3

IBM Elastic Storage System



Fonte: IDC, 2021

IBM ESS è disponibile in due formati - 3000 e 5000:

- IBM Elastic Storage System 3000 (ESS 3000) è progettato per affrontare e superare le sfide della gestione dei dati per gli analytics. Confezionato in un compatto alloggiamento 2U, ESS 3000 accelera il time to value per l'intelligenza artificiale, il deep learning e le applicazioni di calcolo ad alte prestazioni grazie al suo storage interno NVMe e alla semplicità e velocità di installazione e aggiornamento del software containerizzato. Il design hardware e software dell'ESS 3000 consente alle aziende l'accesso alle migliori prestazioni del settore, necessarie per garantire che l'elaborazione dei dati sia efficacemente utilizzata.

- IBM ESS 5000 offre nodi con capacità di petabyte scalabili ad alta velocità, combinando lo storage IBM Spectrum Scale definito da software con i server IBM POWER9 basati su processore I/O-intensive. Grazie al raggruppamento dei requisiti di storage in tutta l'organizzazione su IBM ESS 5000 e ESS 3000 basato su NVMe, i team IT possono ridurre l'inefficienza, diminuire i costi di acquisizione e supportare gli esigenti requisiti di AI, HPC, analisi e/o storage ad alta capacità che sono tipici dei settori della sanità, dei media, settore pubblico e servizi finanziari. ESS 5000 può partire da alcuni terabyte e crescere fino a centinaia di petabyte o persino exabyte con un singolo namespace unificato, eliminando così i costosi silos di dati. IBM Spectrum Scale è il file system parallelo che è alla base di IBM ESS 5000 e che incrementa il rendimento man mano che cresce. Esso consente l'integrazione con i precedenti modelli ESS per salvaguardare gli investimenti e fornisce opzioni a basso costo come il cloud storage e IBM Tape. Grazie a IBM Spectrum Scale, l'IT può eliminare i silos di dati e i colli di bottiglia, semplificare la gestione dello storage e ottenere un accesso più rapido ai dati.

IBM Cloud Object Storage

IBM Cloud Object Storage (COS) è una soluzione di archiviazione leader del settore, software-defined, altamente scalabile e conveniente per l'archiviazione di dati non strutturati sull'edge, nel data center centrale e nei cloud privati o pubblici. IBM Cloud Object Storage è ideale per l'implementazione o la modernizzazione di infrastrutture ad alte prestazioni per l'AI, gli analytics, l'IoT, i video e gli archivi di immagini. Fornisce anche un valore senza precedenti che permette alle aziende di abbassare i costi di storage fino al 12%, utilizzando le nuove unità SMR (shingled magnetic resonance) da 18TB e aumentando il throughput fino a 55GBps in un cluster di 12 nodi. Le aziende possono proteggere i loro dati con la soluzione locale o distribuita di data protection di IBM che può essere personalizzata per data lake complessi e requisiti di grande capacità (vedi Figura 4).

FIGURA 4

IBM Cloud Object Storage



Fonte: IDC, 2021

IBM Cloud Object Storage è un sistema fondamentale per l'infrastruttura AI con vantaggi rilevanti come:

- **Scalabilità:** COS supporta la crescita esponenziale dei dati, scalando le prestazioni e la capacità da terabyte a exabyte;
- **Sicurezza:** COS possiede una crittografia incorporata e un'archiviazione WORM (write once, ready many) bloccabile in base alle policy;

- **Semplicità:** COS può accedere ai dati simultaneamente da qualsiasi luogo e fornisce failover automatico, ricostruzione dei dati, espansione automatica e riequilibrio;
- **Efficienze di risparmio:** COS fornisce dati geoprotetti con l'efficienza dell'Information Dispersal Algorithm (IDA) ed è disponibile come soluzione software singola o come soluzione appliance completamente supportata;
- **Capacità di ricerca:** COS offre insight e ricerche personalizzate per risparmiare tempo. Le aziende possono creare metadati personalizzati per migliorare il valore;
- **Accesso ai file migliorato:** Un nuovo gateway software per l'accesso ai file può collegarsi senza problemi a qualsiasi file system Windows o Linux utilizzando l'accesso SMB o NFS, collegando facilmente le applicazioni basate su file all'object storage;
- **Trasferimento ad alta velocità:** L'opzione di trasferimento dati ad alta velocità di IBM Aspera rende facile il trasferimento dei dati e i livelli flessibili di classe di archiviazione aiutano a gestire i costi pur continuando a soddisfare le esigenze di accesso ai dati.

PROSPETTIVE FUTURE

Nel 2022, il 65% del totale del PIL globale sarà digitalizzato, determinando 6,8 trilioni di dollari di spesa IT in tutto il mondo per il 2020-2023 (vedi *IDC FutureScape: Worldwide Digital Transformation 2021 Predictions*, IDC #US46880818, ottobre 2020). L'infrastruttura digitale non si limita ai tradizionali servizi centrali dell'azienda né ai datacenter cloud discreti. Include tutti i mezzi e le risorse che consentono lo spostamento delle applicazioni e del codice per la trasformazione. Sarà la base per migliorare l'esperienza del cliente. Permette anche di incorporare l'intelligenza e/o l'automazione nelle operation aziendali e supporta l'innovazione continua fino ai confini digitali dell'impresa e dell'industria. Una strategia digitale di successo deve trasformare l'infrastruttura per eliminare i silos, abbattere le barriere imposte dalla tecnologia e andare oltre il semplice supporto di strumenti e applicazioni tradizionali.

IDC ritiene che l'AI sarà il fondamento dell'infrastruttura digitale. Il Customer Insights & Analysis Group di IDC ha recentemente condotto una ricerca per analizzare la spesa IT attuale e futura e i piani di acquisizione da parte di organizzazioni di tutte le dimensioni e di vari settori. Lo studio ha evidenziato che quasi il 76% dei 3.600 responsabili IT di diverse aziende di vari settori in tutto il mondo ha dichiarato che l'intelligenza artificiale è già una parte fondamentale della loro strategia DX oppure è previsto che lo diventi nei prossimi uno-due anni. Solo il 22% degli intervistati ha indicato che l'AI sarà un aspetto fondamentale della loro strategia DX nei prossimi tre-cinque anni. Tra quelle aziende che prevedono che l'AI sarà una componente chiave della loro strategia di trasformazione digitale nei prossimi uno-due anni, spiccano in particolar modo le realtà dei settori telecomunicazioni, servizi pubblici, istruzione e servizi professionali.

L'AI è una questione di "time to value" - il valore che le aziende possono ottenere dai dati nel più breve tempo possibile. Il FutureScape di IDC sull'intelligenza artificiale considera "l'intelligenza artificiale come l'innovazione più dirompente della nostra vita". L'AI non è più semplicemente un "nice to have". La pandemia mondiale ha accelerato l'adozione dell'AI, che sta diventando onnipresente in tutti i processi aziendali. Le soluzioni AI alimentate dal machine learning, l'AI conversazionale e la computer vision sono fondamentali per la ripresa delle attività, per accelerare l'innovazione e per trasformare le esperienze di clienti e dipendenti. Circa il 51% degli intervistati di questa ricerca ha dichiarato che sta attualmente valutando l'intelligenza artificiale o la sta già utilizzando; questo dato è in netto aumento rispetto al 34% degli intervistati del 2019. L'impatto maggiore dell'AI è aiutare i dipendenti a migliorare il loro lavoro. L'adozione dell'AI da parte delle imprese continuerà a crescere man mano che i benefici derivanti dalle implementazioni complete diventeranno più tangibili.

IDC stima che gli investimenti in infrastrutture AI continueranno ad essere significativi nei prossimi anni. IDC prevede che il fatturato dell'hardware AI (server e storage combinati) raggiungerà i 13,4 miliardi di dollari nel 2020, con una crescita del 10,3% rispetto all'anno precedente. All'interno del mercato dell'hardware, si prevede che l'AI storage crescerà dell'11,4% nel 2020. Si prevede che il mercato complessivo dell'hardware avrà una forte ripresa nel 2021, con una crescita del 35,5% su base annua, trainata dall'AI storage, che dovrebbe raggiungere il 43,1% su base annua.

Gran parte di questo storage verrà distribuito in ambienti cloud ibridi con un livello di mobilità senza soluzione di continuità che permette di spostare il carico di lavoro dal core al cloud o all'edge e viceversa. Lo storage sarà un pilastro essenziale per il cloud ibrido, in quanto consente lo spostamento del calcolo verso i dati, fornendo al contempo un livello di accesso e di controllo comune.

Lo storage - specialmente quello in ambienti cloud ibridi - continuerà ad essere il fondamento su cui le iniziative AI possono scalare ora e in futuro. Gli investimenti nell'AI e le iniziative di modernizzazione dei dati dovute all'AI spingeranno gli investimenti verso storage scale-out e verso dati non strutturati. IDC ha recentemente intervistato 624 professionisti IT e team operativi a livello globale per determinare le tendenze di adozione delle infrastrutture IT. In questo studio, IDC ha rilevato che oltre il 65% degli intervistati preferisce i file system scale-out a cui si accede localmente o tramite NFS per i workload ad alte prestazioni come l'AI. In effetti, in particolare per i workload AI, che includono carichi di lavoro per la formazione e l'inferenza, le prestazioni erano il requisito principale dello storage. Seguivano la facilità di implementazione in un cloud ibrido e la qualità del servizio.

Si prevede che lo storage AI crescerà dell'11,4% nel 2020. Si prevede che il mercato complessivo dell'hardware avrà una forte ripresa nel 2021, con una crescita del 35,5% su base annua, trainata dall'AI storage, che dovrebbe raggiungere il 43,1% su base annua.

GUIDA ESSENZIALE PER I BUYER IT

Gli acquirenti di soluzioni tecnologiche sono giustamente incerti sul processo di costruzione del proprio stack di infrastrutture AI. Hanno definito i casi d'uso, lanciato iniziative di AI e assunto o addestrato data scientist e sviluppatori di applicazioni, ma si trovano improvvisamente limitati dall'infrastruttura su cui sviluppare modelli di AI. Spesso l'infrastruttura esistente viene sfruttata per poco tempo, seguita da investimenti in infrastrutture accelerate. I data scientist si trovano a mettere insieme stack sull'infrastruttura accelerata e a cercare di farla funzionare, cosa che in definitiva non fa parte della loro mansione. I team delle infrastrutture IT non hanno familiarità con gli stack di cui i data scientist hanno bisogno e non sono in grado di metterli insieme e ottimizzarli. Questo ha portato a un grave divario di competenze che i venditori di server e i Cloud Service Provider hanno tentato di riempire con i loro stack, ognuno a modo suo. Oggi, ci sono tanti stack quanti sono i venditori, spesso sovrapposti se sviluppati da più attori della stessa filiera.

Iniziare dai benefici di business

Le aziende devono iniziare a mettere insieme i requisiti di utilizzo ed i casi d'uso del sistema per identificare quali risultati di business traggono vantaggio dagli investimenti nell'infrastruttura AI. Devono cercare di quantificare e misurare i benefici di tali investimenti. Ad esempio, se l'obiettivo è differenziarsi rispetto alla concorrenza, la domanda dovrebbe essere: di quanto e quando? Questi criteri dovrebbero poi portare alla selezione di un'architettura applicativa. Le aziende devono considerare l'AI quando cercano di migliorare il loro brand tramite una migliore comprensione e una risposta mirata al feeling, ai desideri e ai bisogni dei clienti, che a loro volta portano a un aumento dei ricavi e dei profitti.

Adottare un approccio olistico

È importante vedere l'intero quadro (cioè, avere una visione globale) quando si implementa qualsiasi iniziativa di AI. Considerare un solo problema per volta può portare alla creazione, ancora una volta, di altri silos o, peggio ancora, aumentare la complessità dell'ambiente IT a causa della mancanza di integrazione e interoperabilità tra più architetture e soluzioni. Un'infrastruttura dati deve essere vista come una soluzione globale dall'edge al core al cloud passando per casi d'uso come l'AI e altre applicazioni mission-critical. Deve essere una soluzione globale che supporta casi d'uso come l'AI e altre applicazioni mission-critical collocate nel core, nell'edge e nel cloud.

Sviluppare la giusta architettura di applicazioni e dati

Sviluppare un'applicazione e un'architettura dati AI è un compito complesso che comporta la conversione di esigenze e risultati aziendali in un flusso di lavoro deterministico abilitato dall'AI. Il flusso di lavoro deve descrivere il modo in cui le capacità AI migliorano il comportamento di quell'applicazione, come i dati vengono ingeriti e analizzati e come l'applicazione interagisce con altre applicazioni aziendali e con gli utenti. Ci si deve concentrare su come i dati vengono processati, prodotti e analizzati dalle applicazioni e le relative implicazioni sull'hardware. Nel creare un piano greenfield ci si dovrebbe concentrare sulla combinazione tra componenti software personalizzati (open source o proprietari) e off-the-shelf.

Scegliere lo stack di riferimento giusto

Diversi venditori e fornitori di servizi hanno proposto stack di riferimento per implementare un'infrastruttura AI. Molti di questi sono di natura "aperta", permettendo un'esperienza modulare "plug and play", e possono essere utilizzati come servizio pay-as-you-go per un'implementazione capex-friendly. Questa è una considerazione importante perché gli investimenti in infrastrutture AI possono diventare velocemente molto costosi. IDC prevede di pubblicare la propria opinione sugli stack di riferimento dei venditori più popolari in un prossimo documento.

I benefici IT da tenere a mente quando si esaminano gli stack di riferimento sono la riduzione dei costi, la disponibilità di dati e applicazioni, l'utilizzo efficace dell'infrastruttura e il consolidamento e, dove possibile, una singola piattaforma interoperabile di distribuzione delle applicazioni.

Creare un'architettura delle informazioni per l'AI

Le imprese hanno bisogno di una strategia di gestione dei dati che fornisca un accesso flessibile e organizzato a tutti i dati, di ogni tipo, indipendentemente da dove risiedono e che risponda alle problematiche dell'architettura di riferimento. Uno sviluppo di questo tipo dovrebbe definire e distribuire un'architettura dell'informazione che fornisca una base aperta ed estensibile, con possibilità di scelta e flessibilità, in grado di comunicare con altre piattaforme cloud. La strategia di gestione dei dati ibridi di IBM per velocizzare il viaggio verso l'AI è un approccio prescrittivo definito dalle quattro fasi dell'AI: raccogliere, organizzare, analizzare e infondere.

- **Raccogliere:** Rendere i dati semplici e accessibili nel posto giusto, da qualsiasi database o storage.
- **Organizzare:** Garantire che i dati siano affidabili, completi e coerenti in tutte le fasi del ciclo di vita delle informazioni: definirne il profilo, pulire e catalogare i dati, fornire protezione e conformità e consentire visibilità, identificazione e reporting basati sulle policy.
- **Analizzare:** Creare, distribuire e gestire modelli AI utilizzando strumenti integrati per esplorare e analizzare dati strutturati e non strutturati e distribuirli in modo sicuro.

- **Infondere:** Ottenere fiducia e trasparenza nelle scelte raccomandate dal modello, spiegare le decisioni, rilevare le distorsioni e così via, utilizzando le soluzioni e i servizi forniti.

Il percorso verso l'adozione dell'AI riguarda tutto il percorso dei dati dall'inserimento fino all'ottenimento di insight, con un'architettura informatica che possa essere facilmente integrata in tutta l'azienda. È importante che ogni fase dell'implementazione dell'AI offra un'integrazione per l'intero percorso. Lo storage è stato tipicamente implementato in modo transitorio con soluzioni di storage specifiche che creano silos di dati e soluzioni che non sono integrate tra loro o con un insieme completo di soluzioni di infrastruttura. È possibile che i clienti memorizzino i dati su un file di grandi dimensioni o su un sistema di object storage, ma poi non abbiano dettagli su quei dati o non utilizzino quei dati per ulteriori approfondimenti. I clienti possono ancora iniziare o concentrarsi su una parte del viaggio, ma ogni progetto dovrebbe considerare un'architettura complessiva delle informazioni AI per ottimizzare le risorse e modernizzare l'infrastruttura per l'espansione dei workload AI.

Le aziende che trattano l'IA come un'applicazione composita ad alte prestazioni (fatta di diverse applicazioni interconnesse), prendendo in prestito elementi cruciali come i file system scale-out, il calcolo eterogeneo e le interconnessioni ad alta velocità per l'accesso distribuito al calcolo e allo storage, sono quelle che alla fine possono modulare la propria infrastruttura AI.

Sfrutta le partnership giuste

Per i buyer IT, la collaborazione con un fornitore di soluzioni end-to-end è cruciale per il successo a lungo termine; tuttavia IDC ritiene che al momento nessun fornitore sul mercato sia ancora in grado di fornire quell'ambiente end-to-end, anche se ci stanno lavorando intensamente. Tuttavia, impegnandosi con un partner di fiducia, le aziende possono utilizzare meglio le strategie di AI per far crescere il proprio business. Esse possono diventare agili e adattive e approfittare delle sinergie interne per promuovere la produttività. Infine, possono anticipare le forze destabilizzanti del proprio settore reinventandosi. Un partner ideale dovrebbe fornire:

- Soluzioni collaudate che scalano da piccole distribuzioni di laboratorio fino a grandi distribuzioni globali
- Competenza verticale focalizzata su segmenti che si allineano al focus aziendale
- Accesso integrato a un ecosistema diversificato di ISV e fornitori di infrastrutture
- Un punto di vista data-first per garantire la sicurezza e massimizzare il valore a lungo termine degli investimenti in nuove fonti di dati
- Successo comprovato nel semplificare gli aspetti hardware, software e di sicurezza di grandi progetti

SFIDE/OPPORTUNITÀ

Per le aziende

Questo white paper ha esaminato una serie di problematiche che le organizzazioni devono affrontare quando sono pronte a scalare le loro applicazioni AI per portarle in produzione. Dalla preparazione dei dati allo sviluppo dei modelli, agli ambienti di runtime, all'addestramento, all'implementazione e alla gestione dei modelli di IA, i requisiti per l'infrastruttura sottostante sono in contrasto con i vecchi modelli di hardware general-purpose. Gli investimenti in un'infrastruttura progettata per i carichi di lavoro ad alta intensità di dati, con prestazioni superiori, scalabilità, accesso ai dati e integrazione, e in grado di fondersi in un ambiente cloud ibrido garantiscono valore e qualità del servizio a lungo termine. Le aziende dovranno prendere decisioni in merito alla sostituzione o all'integrazione delle piattaforme di storage

generiche esistenti con sistemi di storage orientati a compiti di elaborazione specifici dell'AI. Questo servirà come base per lo sviluppo e l'esecuzione di applicazioni AI all'avanguardia.

Per IBM

La sfida per IBM è sempre quella della riconoscibilità sul mercato. IBM offre soluzioni di infrastruttura AI eccezionali e complete che sono integrate con stack di software di infrastruttura (come Red Hat OpenShift) e il cloud pubblico, ma che i potenziali clienti - erroneamente - percepiscono come complesse o più costose. La conseguente reazione istintiva di scegliere uno dei grandi fornitori di sistemi di storage per workload estremamente intensi di dati sta privando queste organizzazioni di soluzioni di infrastruttura AI di cui potrebbero effettivamente beneficiare. Le soluzioni IBM Storage for data e AI, per esempio con IBM Spectrum Scale e IBM Elastic Storage System, sono un elemento costitutivo del supercomputing per molti dei più grandi data center. Ora i nuovi workload AI stanno iniziando a imitare seriamente le implementazioni di supercomputing con requisiti HPC e a sfidare l'infrastruttura on-premises e cloud su cui vengono eseguiti; è proprio questo il momento per IBM di entrare in scena e acquisire nuovi clienti.

CONCLUSIONE

Negli ultimi anni IDC ha assistito alla trasformazione dell'AI e a diverse capacità sviluppate da molte aziende su questo ambito. Inizialmente avviate come esperimenti da personale relativamente inesperto ed eseguite su qualsiasi infrastruttura disponibile, queste iniziative hanno ora iniziato a guadagnare importanza. Molte realtà hanno sviluppato una vasta esperienza relativamente all'AI e stanno sperimentando in prima persona la velocità con cui le risorse AI stanno diventando un aspetto fondamentale del proprio business.

Allo stesso tempo anche l'IT ha subito una trasformazione per quanto riguarda l'infrastruttura su cui eseguire l'AI. Oggi c'è molta più chiarezza sui requisiti dell'infrastruttura per la formazione o l'inferenza del deep learning e su come scalare questi ambienti per la produzione. Che l'addestramento del deep learning richieda un'infrastruttura diversa dalle altre applicazioni è abbastanza scontato. L'addestramento del deep learning necessita di nodi in cluster con processori forti, coprocessori potenti, interconnessioni veloci scalabili per lo storage, grande larghezza di banda I/O e molta memoria.

Oggi, la più grande decisione che l'IT deve prendere è come progettare e distribuire al meglio le applicazioni AI dell'infrastruttura di dati con sistemi best-of-breed e come collegarli e ottimizzarli insieme.

IDC ritiene che le soluzioni di archiviazione IBM per i dati e l'AI, tra cui Spectrum Scale, Elastic Storage System e Cloud Object Storage, offrano valore e prestazioni senza precedenti.

Oggi, la più grande decisione che l'IT deve prendere è come progettare e distribuire al meglio le applicazioni AI dell'infrastruttura dati con sistemi best-of-breed e come collegarli e ottimizzarli insieme.

Informazioni su IDC

IDC (International Data Corporation) è il provider globale di altissimo livello di informazioni sul mercato, servizi di consulenza ed eventi per i mercati dell'information technology, delle telecomunicazioni e di prodotti tecnologici di consumo. IDC aiuta i professionisti dell'IT, i dirigenti aziendali e la comunità di investitori a prendere decisioni basate sui fatti per quanto riguarda gli acquisti di tecnologia e la strategia di business. Oltre 1.100 analisti IDC forniscono competenze a livello globale, regionale e locale su opportunità e tendenze della tecnologia e del settore d'industria, in oltre 110 paesi in tutto il mondo. Per 50 anni, IDC ha fornito insight strategici per aiutare i nostri clienti a realizzare i loro obiettivi di business chiave. IDC è una società controllata di IDG, la società leader nel mondo di supporti tecnologici, ricerca ed eventi.

Sede centrale globale

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com

Informazioni relative al copyright

Pubblicazione esterna di informazioni e dati IDC – Qualsiasi informazione IDC che debba essere utilizzata nella pubblicità, in comunicati stampa o in materiali promozionali richiede una previa approvazione scritta dal Vice Presidente IDC o dal Responsabile nazionale competente. A qualsiasi richiesta di questo tipo dovrebbe essere allegata una bozza del documento proposto. IDC si riserva il diritto di negare l'approvazione all'uso esterno per qualsiasi motivo.

Copyright 2021 IDC. La riproduzione senza autorizzazione scritta è assolutamente proibita.

