

Soluzioni NAS di seconda generazione.

Per soddisfare la domanda crescente di storage

Sommario

- 1** *Introduzione*
- 2** *Il ruolo tradizionale del NAS*
- 3** *Sfruttamento dei vantaggi del NAS*
- 4** *I limiti del NAS nelle strutture aziendali*
- 5** *Il NAS ad architettura scale-out*
- 7** *Applicazioni idonee ai sistemi NAS ad architettura scale-out*
- 8** *Una progressione logica—NAS ad architettura scale-out come servizio*
- 10** *Soddisfare le esigenze strategiche di storage*

Introduzione

La gestione efficiente dei dati implica una lotta costante fra accesso e scalabilità. L'offerta di accesso ai dati a livello di file – file associati a singoli documenti, file multimediali, database e altre applicazioni – diventa sempre più difficile, a causa dell'accesso di un numero crescente di utenti e dell'aumento dei dati memorizzati. Il raggiungimento della scalabilità necessaria a soddisfare la crescita dei dati comporta al tempo stesso anche un aumento dei costi di hardware e software, oltre a nuove sfide di gestione.

Le soluzioni di archiviazione in rete (NAS, Network Attached Storage) offrono semplicità di gestibilità e accesso.

Finora non possedevano una funzionalità specifica, che ha impedito loro di effettuare la scalabilità.

I nuovi sistemi NAS e i servizi di storage gestiti attraverso queste tecnologie offrono un innegabile valore aziendale alle aziende, grazie al loro ampio spettro di applicabilità, alla competitività in termini di prezzo/performance e alla scalabilità al fine di soddisfare la domanda crescente e mutevole di dati.

Il ruolo tradizionale del NAS

I due modi storage di condividere i dati: l'infrastruttura SAN e l'accesso tramite rete IP (NAS) presentano molti aspetti in comune, ma operano su due livelli di protocollo distinti. I sistemi SAN comunicano con le applicazioni utilizzando protocolli di livello inferiore, che offrono rapidità nelle prestazioni ma una gestione più complessa. Una SAN è un sistema di memorizzazione che identifica i dati per numero di blocco e utilizza il protocollo Fibre Channel o la tecnologia InfiniBand per accedere ai blocchi di dati di tipo raw.

Un NAS è un sistema di classe di dati di tipo multistrato, che identifica le informazioni per tipo di file (per esempio, testo, audio o video). In quanto soluzione completa e multistrato (una SAN può infatti costituire uno dei suoi livelli), un NAS gestisce la protezione dati, l'autenticazione utente e il blocco file. Utilizza protocolli standard come il Network File System (NFS). Grazie al trasferimento dei dati di tipo file-level, il NAS può essere utilizzato per una memorizzazione e un accesso ai dati di tipo file più intelligente. Queste caratteristiche comportano tuttavia un rallentamento nelle prestazioni.

Su piccola scala, il NAS offre inoltre una migliore gestibilità, ma anche in questo caso, l'aumento del volume dei file può gravare sia sulle prestazioni e rendere più complessa la gestione.

In genere, i sistemi SAN sono stati utilizzati per applicazioni aziendali caratterizzate da elevati volumi di dati, mentre i sistemi NAS sono stati impiegati per l'accesso ai dati tipicamente dipartimentali.

I sistemi NAS sono stati solitamente utilizzati per l'accesso ai dati dipartimentali

Sfruttare i vantaggi del NAS

I sistemi NAS offrono vantaggi a livello gestionale e un rapido accesso ai dati grazie alla loro posizione centralizzata.

Il NAS offre numerosi vantaggi in grado di soddisfare le esigenze sia degli utenti finali di dati che dei responsabili IT, chiamati a fornire e gestire le risorse dati. La semplicità che caratterizza il salvataggio, l'accesso, lo spostamento e l'archiviazione file tra il dispositivo e i suoi clienti rende il NAS lo strumento ideale per le attività di condivisione file a supporto del flusso di lavoro dipartimentale. In pratica, gli utenti possono salvare e reperire i file in una posizione centralizzata quando e dove ne hanno bisogno. Il consolidamento su un dispositivo NAS di tutti i file associati ad una particolare applicazione o a un dato reparto o gruppo di utenti aiuta inoltre a semplificare la gestione, perché il file system è ubicato in un unico punto. Le attività di gestione ordinaria, per esempio l'esecuzione di backup, lo spostamento di file all'interno del file system per l'ottimizzazione delle prestazioni e l'invio di file al supporto di archiviazione, sono più rapide rispetto a quanto avviene nell'ambiente distribuito di una SAN.

Le applicazioni che utilizzano lo storage NAS possono in alcuni casi offrire prestazioni migliori rispetto ai sistemi SAN, grazie al controllo di tutte le attività di gestione dei file attraverso il sistema operativo NAS anziché mediante quello del server applicativo. Non dovendo occuparsi delle attività di memorizzazione e recupero dei file, il server applicativo è in grado di dedicare maggiori risorse all'elaborazione delle richieste dell'applicazione, potendo così ottenere prestazioni migliori.

L'instradamento intelligente dei dati al livello del sistema operativo può produrre prestazioni migliori.

I limiti del NAS nelle strutture aziendali

Il problema dei sistemi NAS tradizionali è la mancanza di scalabilità.

I reparti IT chiamati a soddisfare una domanda immediata di ulteriore storage si sono limitati ad aggiungere dispositivi NAS alle proprie reti: si installa un box, si assegna un indirizzo IP e si iniziano a salvare i file. Un simile approccio è ottimale per file affini e relativamente piccoli. Tuttavia, la crescita esponenziale di dati che le aziende di ogni dimensione stanno registrando comporta il rischio di perdere i vantaggi del NAS tradizionale in termini di efficienza e gestione a causa dell'assenza di scalabilità. Un file server che supporta un gruppo di 50 utenti di un reparto contabile, per esempio, diventa inadeguato quando una fusione apporta 50 nuovi dipendenti al suo interno. È possibile aggiungere file server di maggiori dimensioni, ma per le aziende di decine di migliaia di utenti i NAS tradizionali diventano una soluzione inefficace.

Ogni dispositivo NAS offre un livello predeterminato di prestazioni, una capacità di memorizzazione limitata e un numero chiuso di file system disponibili, da ripartire fra le applicazioni e gli utenti desiderati.

La tipica soluzione NAS è limitata a due nodi, pertanto, al raggiungimento della capacità massima, la soluzione logica consiste nell'aggiungere altri dispositivi. Ciò può tuttavia determinare un'interruzione delle operazioni nonché gravare ulteriormente sul personale IT, chiamato a configurare e utilizzare il dispositivo, ridistribuire i file system e creare autorizzazioni di accesso. Gli utenti devono inoltre ristabilire i collegamenti dai loro clienti, con un impatto sulla produttività.

L'aggiunta di dispositivi NAS può effettivamente aumentare la complessità e i costi e impedire un accesso diffuso ai dati. Non è certo la prospettiva ideale per un'azienda in crescita.

Un'altra alternativa consiste nell'architettura verticale, attraverso l'aumento dei volumi di input/output e delle capacità di archiviazione del dispositivo NAS all'aumentare del fabbisogno di dati. Anche questa soluzione presenta tuttavia diversi svantaggi. Uno, è il costo elevato per la sostituzione di un dispositivo con un incremento contenuto della capacità (che può essere successivamente superato). Un altro è che ancora presenta un unico punto di errore con però un aumento dei dati aziendali a rischio in caso di fermo.

L'approccio verticale aumenta la complessità e i costi e comporta un "single point of failure"

Il NAS scale-out

Il NAS di seconda generazione utilizzano architetture scalabili comuni nei servizi Web – pool virtualizzati di risorse a costi limitati.

La soluzione ideale consiste in un sistema che unisce il meglio in termini di accesso e di gestione del file system e che sia scalabile per adattarsi alle esigenze aziendali. Fortunatamente, oggi sono disponibili soluzioni NAS che offrono architetture scalabili e capacità ottimizzate di gestione e che consentono di offrire la scalabilità richiesta per la condivisione dei file a livello di dipartimento, a livello dell'intera azienda e per le applicazioni che necessitano di calcolo ad alte prestazioni (HPC).

Questi sistemi NAS scale-out utilizzano architetture scalabili e tecnologie mutuare dagli ambienti server ed applicate all'ambiente storage. L'obiettivo è offrire un accesso semplificato ai dati e la scalabilità per crescere come imposto dalle esigenze dei singoli utenti, delle applicazioni, del reparto IT e dell'azienda nel suo complesso.

L'architettura scale-out NAS è simile a quella degli ambienti Web. Quando i visitatori si collegano a un sito Web, non necessariamente accedono ogni volta ai dati tramite il medesimo server fisico. I server a costi limitati gestiscono il traffico in modo virtuale. Quando il carico aumenta, non si utilizzano server di maggiori dimensioni, bensì un maggior numero di server gestiti virtualmente come un server unico.

L'applicazione di questa struttura scale-out al NAS, che ha sempre avuto uno sviluppo verticale, rappresenta un'evoluzione emblematica. Il modello verticale è intrinsecamente limitato in termini di dimensioni/capacità (e di costi/affidabilità) da uno o da un paio di failover dei server storage. Una struttura scale-out è invece limitata soltanto dal software, che unisce i molti server economici in un unico server virtuale di grandi dimensioni.

Questi nuovi e più gestibili sistemi NAS devono la loro maggiore scalabilità a tecnologie e pratiche mutuare dai server applicativi, fra cui:

- *Clustering*
- *File system scalabili*
- *Storage Pooling e provisioning.*

Il NAS di seconda generazione rappresenta un'evoluzione perché virtualizza i dispositivi di archiviazione consentendo una maggiore scalabilità.

Clustering

Per clustering si intende l'utilizzo di più nodi di dispositivi NAS che creano un sistema logico storage di maggiori dimensioni, gestibile come un'unica entità. Tra gli altri vantaggi, il clustering offre migliori prestazioni del sistema storage. Le richieste dati e la potenza di elaborazione vengono distribuite su un sistema più ampio, riducendo il bisogno di conservare in memoria cache i salvataggi e le richieste di dati. In un sistema NAS scale-out, i dati vengono inviati direttamente ai dischi, migliorando così la velocità delle richieste e soddisfacendo volumi elevati di richieste.

Poiché il sistema scale-out NAS è gestito come una singola entità; l'aggiunta di nuovi nodi è facilmente riconosciuta dall'interfaccia di gestione e può essere impiegata secondo le necessità. Questa capacità di incrementare rapidamente la capacità di archiviazione o di migliorare le prestazioni e l'affidabilità favorisce un ambiente storage più efficiente. La possibilità di aggiungere nodi con facilità e solo quando necessario rappresenta un notevole miglioramento rispetto alle tradizionali strategie storage, che favoriscono invece acquisti eccessivi con conseguente sottoutilizzo e maggior costo totale di proprietà (TCO).

Le capacità di clustering dei sistemi NAS consentono un rapido incremento delle prestazioni e della capacità totale di archiviazione.

File system scalabili

Un approccio scale-out è intrinsecamente più disponibile perché elimina “il single point of failure” del NAS tradizionale.

Le dimensioni relativamente ridotte dei file system hanno limitato l'utilizzo di NAS nelle strutture aziendali.

In una struttura scale-out, una data istanza di NAS è dinamicamente espandibile (o ridimensionabile) in qualsiasi momento tramite il semplice apporto di server front-end a costi limitati, al fine di aumentare le prestazioni di input/output o, mediante l'aggiunta di un disco di back-end (capacità storage) trasparente per gli utenti finali, di offrire in tal modo una scalabilità pressoché illimitata. È chiaro che questa architettura dinamica è anche facilmente disponibile. In caso di malfunzionamento o rimozione di un qualsiasi server del sistema, i restanti server si assumono il carico di lavoro senza interruzioni.

Questa importante evoluzione è il frutto di due tecnologie fondamentali:

- *File system condivisi altamente scalabili, che prevedono un unico spazio dei nomi per tutti i file*
- *Un metodo per la virtualizzazione del servizio NFS (o per il collegamento del protocollo prescelto), affinché un utente finale possa visualizzare un unico dispositivo anziché i singoli server che compongono il dispositivo NAS virtuale*

IBM General Parallel File System è il software che consente di offrire le nuove funzionalità NAS.

IBM General Parallel File System™ (GPFS) è uno dei file system commerciali maggiormente scalabili. È installato in migliaia di nodi e memorizza petabyte di dati in installazioni distinte. I servizi “scale out” consentono la virtualizzazione del servizio NFS, distribuendo in modo trasparente le richieste di servizio su più server e rendendo così possibile la creazione di un'installazione farm NAS scale-out.

Storage Pooling e provisioning

A causa del numero limitato di nodi di un sistema NAS tradizionale, non si è trovata una soluzione pratica per utilizzare diversi media sullo stesso sistema, al fine di impiegare in modo efficace le strategie di gestione del ciclo di vita delle informazioni. Il NAS scale-out, supportato da GPFS, consente di creare pool di NAS in grado di gestire dischi ad alta e bassa velocità nonché unità a nastro. Permette a chi amministra lo storage di associare un dato attributo di file - per esempio la data del file, l'ID utente, l'ID server, il tipo di file, la directory o la sottodirectory – per poi spostare dinamicamente i dati nel media adatto, in base alle regole di associazione.

La creazione di pool composti da tipologie eterogenee di storage ha incrementato il controllo e la fruibilità dei dati a supporto delle iniziative di gestione del ciclo di vita delle informazioni. Gli amministratori possono offrire risorse in modo più dinamico in risposta ad alcune attività, come i back-up giornalieri o i processi batch eseguiti giornalmente, trimestralmente o annualmente.

IBM General Parallel File System consente l'utilizzo di diversi tipi di media in un unico sistema NAS a supporto delle iniziative di gestione del ciclo di vita delle informazioni.

Applicazioni adatte ai sistemi NAS ad architettura scale-out

I NAS scale-out sono applicabili a quasi ogni tipo di dati, inclusi file di grandi dimensioni, con il vantaggio di un aumento della scala e di un miglioramento delle prestazioni.

Quali tipologie di dati si possono archiviare su un sistema scale-out NAS? Quasi tutte. L'applicazione iniziale del NAS per un'archiviazione centralizzata dei file, dove gli utenti possono trovare i dati con facilità, resta valida. Le differenze attengono nuovamente alla scala e alle prestazioni. Il NAS scale-out è in grado di supportare più utenti, un numero e una dimensione maggiori di file con facilità – oltre a farlo con prestazioni migliori. Un sistema NAS scale-out offre le prestazioni e l'accessibilità, attesa dagli utenti finali, per documenti di produttività personale, fogli di calcolo, immagini e file audio/video.

Il NAS scale-out supporta iniziative IT avanzate come il consolidamento attraverso la virtualizzazione e le SOA.

Il NAS scale-out può soddisfare le richieste di file provenienti da applicazioni in modo più efficiente rispetto alle risorse di storage distribuite. È particolarmente adatto per soddisfare le richieste di volumi elevati di file di piccole dimensioni, per esempio per i servizi online per il mercato consumer. Molti portali online di alto profilo hanno iniziato a utilizzare il NAS scale-out per supportare i loro servizi di posta elettronica, archiviazione multimediale e contenuti Web.

Con l'obiettivo di conformare ulteriormente i servizi IT agli obiettivi aziendali, molte aziende stanno adottando iniziative avanzate come, per esempio, l'architettura orientata ai servizi (SOA) e la virtualizzazione. Spesso, il primo passo in queste iniziative è costituito dal consolidamento delle risorse distribuite in sistemi più concentrati e, tuttavia, più scalabili. Il NAS scale-out supporta il consolidamento attraverso la centralizzazione dei dati in un unico file system e la semplificazione delle attività di gestione associate al salvataggio, allo spostamento e all'accesso dei file.

I settori caratterizzati da elevati volumi di dati e considerazioni sulle prestazioni sono adatti al NAS scale-out.

Le applicazioni HPC richiedono un'elevata potenza di elaborazione del server, ma normalmente presentano anche elevate esigenze in termini di volumi e di prestazioni. Le applicazioni generalmente in uso per l'analisi dei dati o dei risultati dei modelli richiedono da sole molto spazio storage e possono generare rapidamente ingenti volumi di dati. Alcune applicazioni industriali adatte al NAS scale-out sono le seguenti:

- *Petrolio e gas*
- *Geoscienze*
- *Animation rendering*
- *Ricerca*
- *Ricerca in biotecnologie*

Una progressione logica— NAS scale-out come servizio

L'implementazione di un sistema NAS scale-out richiede competenze.

Per usufruire al meglio dei vantaggi strategici e delle risorse tecniche del NAS scale-out sono necessarie competenze notevoli. La corretta progettazione ed esecuzione di un sistema NAS scale-out richiede la conoscenza e l'esperienza concreta in ambienti server high-end e nelle piattaforme di vari sistemi operativi, oltreché competenze di progettazione e gestione dello storage alle quali poche aziende, anche di livello mondiale, hanno attualmente accesso. E malgrado un sistema NAS scale-out, una volta implementato, renda più dinamica la gestione dell'ambiente storage, esso richiederà una gestione costante per mantenere alte le prestazioni. Anche l'eventuale necessità di aggiungere nodi a un'implementazione NAS scale-out e la scelta del momento giusto per farlo richiederanno un'attenta pianificazione ed esecuzione al fine di ridurre il rischio di impatto sul contesto produttivo.

approccio di servizio può aiutare le aziende a diventare rapidamente operative.

Per ragioni sia di gestione IT che di natura strategica, alcune aziende possono decidere di affidare l'implementazione di un sistema NAS scale-out a un fornitore di servizi esperto. IBM sta sfruttando molte caratteristiche dei suoi prodotti e servizi per offrire una soluzione di gestione file scalabile che i clienti possano utilizzare per implementare rapidamente un sistema di spazio file globale per l'archiviazione e l'accesso ai dati.

L' IBM Storage and Data Managed Services – scale out file services è una soluzione NAS per la rapida implementazione di un sistema di spazio file globale e scalabile.

Un componente essenziale della soluzione è il GPFS, un file system a disco condiviso dalle prestazioni elevate in grado di offrire un accesso ai dati rapido e affidabile da tutti i server in un cluster. GPFS consente alle applicazioni parallele l'accesso simultaneo a una serie di file (o anche a un unico file) da qualsiasi nodo integrato nel file system, offrendo al tempo stesso un elevato livello di controllo su tutte le operazioni del file system.

Un'interfaccia utente grafica inclusa nel servizio offre una rapida e chiara visualizzazione delle risorse e delle prestazioni NAS.

La soluzione IBM che utilizza risorse di tipo scale-out è composta da dispositivi hardware storage IBM preconfigurati e dal GPFS, oltre a software aggiuntivi per la gestione delle allocazioni nel file system. Per semplificare l'attività di gestione degli amministratori storage, la soluzione presenta inoltre un'interfaccia utente grafica (GUI) che offre una visualizzazione rapida e chiara delle risorse e delle prestazioni NAS per consentire il rapido adeguamento al fine di soddisfare le dinamiche della domanda degli utenti.

La soluzione IBM offre risorse scale-out per più di 500 nodi e capacità fino a un exabyte.

Benché il NAS scale-out costituisca un approccio in evoluzione, la maggior parte delle soluzioni presenti sul mercato sono ancora limitate, in termini di scalabilità, ad un massimo di otto nodi. I scale out file services di IBM sono stati testati per supportare fino a 500 nodi e offrono il potenziale per un'ulteriore crescita. Sono quindi adatti a grandi aziende con ingenti volumi di dati che intendono impiegare una soluzione unica a livello aziendale. Le soluzioni scale out file services di IBM si basano su una tecnologia che offre prestazioni elevate e che ha dimostrato, in installazioni reali, di riuscire a fornire oltre 130 megabyte al secondo per un solo file tramite connessioni IP. In termini di volume totale di dati memorizzabili, la soluzione scale out file services supporta attualmente fino a un exabyte, pari a un miliardo di gigabyte.

Disponibili in loco presso il cliente o tramite i centri dati IBM, la soluzione scale out file services offre inoltre risorse funzionali che ottimizzano il valore di base del sistema NAS scale-out. Alcuni esempi sono i seguenti:

- *Supporto per la gestione del ciclo di vita delle informazioni affinché i dati siano archiviati in relazione al loro valore*
- *Servizi offerti sulla base della politica di qualità del servizio (QoS) che definisce priorità nella memorizzazione e nelle prestazioni a specifici gruppi o applicazioni*
- *Back-up, ripristino e archiviazione ai fini di protezione dati*
- *Funzioni di snapshot e mirroring per i sistemi ad elevata disponibilità che richiedono finestre di back-up più piccole*
- *Offsite Disaster-Recovery che utilizza il file system globale e consente l'invio dei dati ripristinati ad uno qualsiasi dei numerosi centri dati IBM*
- *Monitoraggio e segnalazione completi a livello di utente e progetto*
- *Funzione di assistenza hardware "Call home" per la manutenzione proattiva dei sistemi.*

Soddisfare la domanda strategica di storage

I sistemi NAS orizzontali sono in grado di affrontare le necessità di scalabilità e prestazioni dal punto di vista tecnico ma, in definitiva, devono sostenere obiettivi aziendali più estesi.

L'utilizzo di architetture avanzate e il consolidamento in un unico file system possono contribuire a ridurre le esigenze in termini di gestione storage, migliorare l'utilizzo e potenzialmente ridurre i costi. I sistemi NAS scale-out possono aiutare a migliorare le prestazioni delle applicazioni e contribuire a inviare file agli utenti in modo più semplice e veloce, sostenendo lo sforzo di miglioramento della produttività individuale e di reattività organizzativa. Questi sistemi storage possono inoltre contribuire a migliorare l'affidabilità e la sicurezza aziendale, perché la loro struttura multi-nodale consente il failover, proprio perché operano come servizio gestito all'interno di un file system globale.

L'assegnazione della quantità esatta di spazio per la memorizzazione dei dati a ogni utente e al momento giusto continuerà a rappresentare una sfida costante per tutte le aziende. Le soluzioni NAS scale-out possono contribuire ad alleggerire alcune di queste sfide, riducendo il compromesso in termini di prestazioni e garantendo un maggior vantaggio strategico.

Per maggiori informazioni

Per maggiori informazioni sui servizi dati e di archiviazione IBM, visita:
ibm.com/services/storage

© Copyright IBM Corporation 2008.

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Prodotto negli Stati Uniti d'America
Tutti i diritti riservati
09-07

IBM, il logo IBM e General Parallel File System sono marchi o marchi registrati di International Business Machines Corporation negli Stati Uniti e/o in altri paesi.

I nomi di altre società, prodotti e servizi possono essere marchi registrati appartenenti a terzi.

I riferimenti a prodotti e servizi IBM non implicano che IBM intenda metterli a disposizione in tutti i paesi nei quali opera.