



WHITE PAPER

INFINIDAT Storage Architektur

ZUSAMMENFASSUNG

Die INFINIDAT® Enterprise-Storage-Lösungen basieren auf der einzigartigen und patentierten INFINIDAT Speicherarchitektur - Ein vollständig abstrahierter Satz von Software-Defined Storage (SDS)-Funktionen wird mit der besten Standard-Hardware integriert. Mit dem Versand der Software liefert INFINIDAT eine ausgiebig getestete Hardware-Referenzplattform mit. Damit stellt INFINIDAT die ersten echten SDS-Lösungen der Enterprise-Klasse bereit.

Der vorliegende Beitrag beschreibt die Technologie, die es INFINIDAT ermöglicht, der einzige Speicheranbieter zu sein, der mit einer beispiellosen Verfügbarkeit von sieben Neunen (99,9999999%), einer „Faster-Than-Flash“-Leistung mit über 1 Mio. IOPS bei einer Latenz von weniger als einer Millisekunde und einer Multi-Petabyte-Kapazität in einem einzigen 42HE-Rack bei geringeren Gesamtbetriebskosten (Total Cost of Ownership - TCO) eine überlegene Zuverlässigkeit erreicht.

KONSTRUKTIONSPRINZIPIEN

Bei der Entwicklung einer Speicherarchitektur, die den Anforderungen moderner Rechenzentren entspricht, müssen mehrere Anforderungen erfüllt sein:

KATEGORIE	ANFORDERUNG
Zuverlässigkeit	Unternehmen arbeiten rund um die Uhr (24x7), Ausfallzeiten werden nicht toleriert
Kapazität	Exponentiell steigende Datenmengen, beschleunigt durch digitale Transformation, disaggregierte Big-Data-Architekturen, Künstliche Intelligenz (KI) und Maschinelles Lernen (ML)
Performance	Da Daten skalieren, muss die Performance Schritt halten, um die gleichen (oder besseren) Ergebnissen in gleichen (oder kürzeren) Zeiträumen bereitzustellen
Einfachheit	Administratoren erwarten einfache Bedienung, umfassende Ökosystem-Integration und integrierte Tools für den Übergang zu DevOps-Modellen, um weniger Zeit mit der Storage-Verwaltung zu verbringen und mehr Zeit für Anwendungen und Businessprozesse zu haben
Konsolidierung	Point-Technologien gehören der Vergangenheit an. Moderne Speicherung sollte in der Lage sein, bei allen Anwendungsfällen zur maximalen Effizienz, Einfachheit und Kostenersparnis beizutragen
Kosten	Budgets sind nicht kapazitäts- und leistungsgerecht skalierbar; disruptive Veränderungen in der Architektur sind notwendig

Gleichzeitig behaupten Amazon, Google und Azure Clouds, die Kosten für den gesamten IT-Stack zu senken, und das tun sie oft für kleinere Kunden, die sich kein großes IT-Team leisten können und sich häufig auf nur ein oder zwei „Alleskönner“ verlassen, um ihren gesamten IT-Betrieb aufrechtzuerhalten. Für große Unternehmen und regionale Cloud-Provider kann die Einführung eines effizienteren IT-Stacks, der ihre geschäftlichen, technologischen und finanziellen Anforderungen erfüllt, bedeuten, dass sie mit allen Vorteilen der Cloud innerhalb ihrer eigenen Infrastruktur ausgestattet sind und dabei gleichzeitig ihre Kosten senken und die Souveränität über ihre Daten wahren.

INFINIBOX ARCHITEKTUR

INFINIDATs Flaggschiff InfiniBox® wurde unter Berücksichtigung der wichtigsten Prinzipien entwickelt, um all diesen Anforderungen gerecht zu werden:

AUFTRAG	ARGUMENTATION	HERAUFORDERUNG
Innovatives Software-Design	Software wird im Gegensatz zu Hardware im Laufe der Zeit optimiert und verbessert die Leistung, anstatt sie zu verschlechtern. InfiniBox basiert auf über 80 erteilten Software-Patenten - die wahre Bedeutung von Software-Defined Storage (SDS).	Performance Einfachheit Zuverlässigkeit Kosten
Design für Stabilität	Wird etwas für die Skalierung entwickelt, ist die Stabilität von entscheidender Bedeutung. Die InfiniBox ist für eine Betriebszeit von sieben Neunen (99,99999%) ausgelegt und verwendet eine dreifach redundante Architektur, bei der alle kritischen Komponenten (Software und Hardware) mindestens zwei (2) Redundanzen (N+2) aufweisen, die vor Ausfallzeiten und Datenverlust schützen.	Stabilität Kosten Einfachheit Konsolidierung
Architektur für Skalierung	Um Kapazität und Leistung zu disruptiven Kosten zu erreichen, bedarf es einer Skalierung. Die InfiniBox wurde für Großkunden mit einer effektiven Kapazität von mindestens 230 TB und einer Skalierung auf bis zu 8,3 PB (mit einer konservativen 2:1 Datenreduktion) in einem einzigen 42HE-Rack entwickelt.	Konsolidierung Kosten Einfachheit Kapazität
Enge Integration von Hard- und Software	Eine der vielen Möglichkeiten, wie InfiniBox diese hohe Ausfallsicherheit erreicht, besteht darin, mehrere Hardware-Komponenten zu testen, aber nur die zuverlässigste auf den Markt zu bringen (Referenzarchitektur). Dieser Ansatz bedeutet, dass die Kunden eine voll integrierte Lösung erhalten, ohne die Kosten und Komplexität der Hardware-Integration vor Ort und ohne den administrativen Aufwand.	Zuverlässigkeit Einfachheit Konsolidierung
Standard- (Commodity off the shelf - COTS) Hardware	Der Einsatz von Standard-Hardware und die Vermeidung langer Entwicklungszyklen ermöglichen eine schnellere Einführung neuer Technologien. Dazu gehören CPUs, Speichertypen und neue Arten von Speichermedien. Die Verwendung von Standard-Hardware sowie der dazugehörigen Software bringt auch mehr Stabilität, da die gleiche Hardware in Tausenden von Systemen weltweit eingesetzt wird.	Kosten Zuverlässigkeit Kapazität Einfachheit Performance

LEISTUNGSBESCHLEUNIGUNG

InfiniBox® ist ein Flash-optimiertes Array, welches eine Kombination aus DRAM, Flash-Medien (SSD) und NL-SAS-Festplatten mit hoher Kapazität zum Schreiben, Lesen und Speichern von Daten verwendet. Im Folgenden wird erklärt, wie Lese- und Schreibvorgänge beschleunigt werden, um maximale Leistung bei minimaler Latenz zu erreichen. Der selbstlernende Algorithmus zur Optimierung der Datenplatzierung heißt Neural Cache. Dieser Abschnitt erklärt, wie Neural Cache Kunden durch die Nutzung intelligenter Software-Algorithmen mit niedrigeren Latenzen als All-Flash-Arrays bereitstellen zu können.

Es ist wichtig zu bedenken, dass die meisten transaktionalen Anwendungen mindestens zwei I/Os benötigen (einen zum Schreiben einer Transaktion in die Protokolle und einen zum Schreiben von Daten in die Datenbank). Damit wird die Latenz zur Schlüsselkomponente bei der Beurteilung sowohl der Anwendererfahrung als auch der maximalen Leistung der Applikation.

Metadaten-Schicht

Die Antwortzeiten der Metadaten-Schicht wirken sich sofort auf die I/O-Latenzzeit aus. InfiniBox beschleunigt Metadaten-Operationen folgendermaßen:

- ▶ **Alle Metadaten befinden sich im DRAM** - Metadaten werden im DRAM gespeichert, wodurch sich sowohl das Lesen als auch das Schreiben beschleunigt
- ▶ **Effiziente Metadaten-Struktur** – alle Einfügungen, Änderungen und Löschungen aus der Metadatenstruktur (genannt 'Trie') arbeiten mit der gleichen Latenzzeit und bieten eine konsistente Performance

Schreibbeschleunigung

Die InfiniBox akzeptiert alle Schreibvorgänge ohne Vorverarbeitung (wie z. B. Pattern-Entfernung, Komprimierung, Verschlüsselung, etc.) in ihrem DRAM und erstellt, bevor die Bestätigung an den Host gesendet wird, eine zweite Kopie des Schreibvorgangs im DRAM eines anderen Knotens über InfiniBand mit niedriger Latenzzeit. Die Übernahme des Schreibvorgangs vom DRAM (direkt an die CPU angeschlossen) anstelle eines externen Flash-Geräts ermöglicht es der InfiniBox, die Schreibvorgänge mit der geringstmöglichen Latenzzeit abzuschließen.

Im Gegensatz zu vielen Architekturen, bei denen der Schreib-Cache in kleine Bereiche aufgeteilt ist (wie bei Matrix- und Dual-Controller-Architekturen), verwendet InfiniBox einen einzigen, großen Speicherpool, um Schreibvorgänge zu akzeptieren. Dies ermöglicht die Aufrechterhaltung größerer Schreib-Bursts, das Überschreiben von Daten, die sich häufig ändern, mit der DRAM-Latenz und der Neural-Cache-Zeit, intelligente Entscheidungen zu treffen und zu priorisieren, welche Datenblöcke von DRAM-Geschwindigkeiten profitieren und welche auf SSDs und HDDs verteilt werden sollen. Indem die Daten länger im Schreib-Cache gehalten werden, vermeidet Neural Cache unnötigen Arbeitsaufwand auf der CPU sowie auf den Backend-Persistenz-Schichten.

Lesebeschleunigung

Im Gegensatz zu herkömmlichen Speicher-Arrays, die darauf abzielen, die aktivsten Daten (auch „Hot Data“ genannt) im Flash-Cache zu platzieren, um eine Performance-Parität mit allen Flash-Arrays (AFAs) zu erreichen, verwendet InfiniBox seinen innovativen Neural-Cache. Dieser zielt darauf ab, alle wichtigen Daten im DRAM zu platzieren. Der InfiniBox Neural Cache ermöglicht die meisten Lesevorgänge mit DRAM-Geschwindigkeit, die 1000-mal schneller als Flash ist.

Seit 2017 umfasst INFINIDATs globale Datenstruktur mehrere Exabytes an Daten, und Neural Cache bietet nachweislich fast alle Lesevorgänge aus DRAM, so dass unsere Kunden ein „All-DRAM-Array“-vergleichbares Erlebnis zu einem Preis erleben können, der unter dem eines AFAs liegt.

Da es sich bei Neural Cache um einen Lernalgorithmus handelt, wird die Leistung im Laufe der Zeit optimiert. Die InfiniBox nutzt eine dicke SSD-Flash-Schicht, die als „Polster“ für DRAM-Fehler dient. Da der neuronale Cache die I/O-Muster lernt und die Platzierung der DRAM-Daten optimiert, ändert die Flash-Schicht ihre Funktion von der Behandlung von DRAM-Fehlern auf die Behandlung von I/O-Muster-Veränderungen, die der Algorithmus möglicherweise nicht vorhersagen kann (z. B. ein periodisches Audit, bei dem Daten nicht im DRAM benötigt werden).

SOFTWARE ARCHITEKTUR

INFINIDAT setzt bei der Entwicklung der InfiniBox sieben Neuen an Zuverlässigkeit ein, um die Unvorhersehbarkeit von Hardware-Ausfällen zu überwinden. Die InfiniBox Software-Architektur nutzt 3 Controller, die Active/Active/Active betrieben werden und das N+2-Design, welches eine ständige Überwachung und Selbstheilung sowie eine problemlose Wiederherstellung nach Hardware-Ausfällen auf allen Ebenen ermöglicht.

Alle Komponenten sind in die Software implementiert, vom RAID bis zu den geclusterten Diensten, um eine ständige Optimierung mit jedem neuen Release zu ermöglichen. In den ersten vier Jahren nach der erstmaligen allgemein verfügbaren Version ist die maximale Leistung der InfiniBox um das Vierfache gestiegen, nur durch ein unterbrechungsfreies Upgrade der Software. Das ist die Stärke einer echten Software-Defined Lösung.

Gebündelte Dienste

Alle Datendienste laufen, in Übereinstimmung mit der N+2-Architektur, auf allen Knoten und sind auf allen Knoten aktiv (keine passiven Knoten im Cluster). Die Datendienste sind so konzipiert, dass sie im Benutzerraum laufen, einschließlich von Low-Level-Komponenten wie Fiber Channel (FC)-Treiber. Da kein Datendienst im Kernel läuft, kann kein Softwarefehler andere Dienste im System oder die Knotenverfügbarkeit beeinträchtigen. Diese Designprinzipien gelten sowohl für Frontend-Dienste wie für Datenprotokolle (NFS, iSCSI, FC, FICON) als auch für Backend-Datendienste wie Neural Cache, InfiniRAID® und InfiniSnap®.

Die Datendienste werden vom Cluster Manager (CLM) gestartet und überwacht, welche alle Serviceprobleme erkennt und bei Bedarf die Dienste neu starten kann. Ein Dienst, bei dem ein Fehler auftritt, startet neu und testet sich selbst, bevor er dem Cluster wieder beiträgt. Jeder Dienst, der nicht korrekt startet, wird dem Cluster nicht beitreten, um zu vermeiden, dass er im Cluster ausfällt („byzantinisches Versagen“). Wenn der Cluster-Manager einen Dienst identifiziert, der mehrmals erfolglos auf einem bestimmten Knoten neu gestartet wurde, stoppt er den Neustart und benachrichtigt den INFINIDAT Support. Jeder Dienstausfall – ob automatisch wiederhergestellt oder nicht – wird an die Datenanalyse-Plattform von INFINIDAT zurückgemeldet, um Softwareprobleme zu erkennen und die Code-Qualität kontinuierlich zu verbessern.

Disk Layout

Das Festplatten-Layout der InfiniBox wird durch die patentierte Software-Innovation InfiniRAID verwaltet. InfiniRAID ist ein Software-definiertes redundantes Array unabhängiger Festplatten (Redundant Array of Independent Disks - RAID), das die gesamte Datenplatzierung, Datensicherung und Wiederherstellung nach Ausfallszenarien steuert.

InfiniRAID ist ein declusterisiertes RAID, welches das Datenlayout von der physikalischen Schicht trennt und Tausende von virtuellen RAID-Gruppen verwendet, die Daten über alle Laufwerke verteilt und Hot Spots verhindert. InfiniRAID erstellt die RAID-Gruppen, so dass alle zwei Laufwerke im System nur bis zu 2,5% ihrer RAID-Gruppen teilen.

Dieser geringe Prozentsatz überlappender RAID-Gruppen hat mehrere Vorteile:

- ▶ **Gleichmäßige Verteilung** – Alle Datensätze, egal wie klein sie sind, werden über alle Laufwerke im System verteilt, wodurch ein maximaler Durchsatz für jede Anwendung erzielt wird
- ▶ **Selbstheilung** – mögliche Hot Spots werden automatisch durch die Optimierung des Datenlayouts gelöst
- ▶ **Virtuelle Ersatzteile** – Die Speicherkapazität wird gleichmäßig über alle Laufwerke im System verteilt. Es gibt keine physischen ‚Hot Spares‘, so dass der Rebuild-Vorgang die Daten neu verteilen kann und unnötige Kosten minimiert. Das System bietet genügend freie Kapazität, um bis zu 12 Laufwerke in einem F6000 ausfallen zu lassen.
- ▶ **Performance-Schutz** – Der Ausfall eines einzelnen Laufwerks (die Daten sind immer noch geschützt) erzeugt nur einen RAID-Neuaufbau mit niedriger Priorität („Rebuild-1“), der die Anwendungsleistung priorisiert und nur die Ressourcen des Systems nutzt.
- ▶ **Schnelle Wiederherstellung** – Wenn ein zweites Laufwerk ausfällt, beschleunigt das System den Neuaufbau für die gemeinsamen 2,5% der RAID-Gruppen, die zwischen den beiden ausgefallenen Laufwerken geteilt werden („Rebuild-2“), bevor es zur niedrigeren Priorität Rebuild-1 zurückkehrt, da es keine ungeschützten RAID-Gruppen mehr gibt.
- ▶ **InfiniSpares** – Über die garantierte Kapazität von 12 Ersatzteilen hinaus kann die InfiniBox bei Bedarf auch freie Kapazität als Reservekapazität nutzen. Diese Innovation ermöglicht den Ausfall von bis zu 100 Festplatten, ohne den Schutz zu verlieren.

Datenschutz-Services

InfiniBox bietet eine Vielzahl von Datenschutz-Services an, um Kunden beim Schutz ihrer Assets zu unterstützen:

- ▶ **Snapshots** – Der InfiniBox Snapshot-Mechanismus heißt InfiniSnap und basiert auf einem nicht-verriegelnden Redirect-on-Write-Mechanismus, der eine konsistente Performance mit oder ohne Snapshots liefert. Jeder Datensatz kann bis zu 1.000 Snapshots haben, jeder kann entweder nur lesend (zur Datensicherung) oder beschreibbar (für Test- und Entwicklungsumgebungen) sein. InfiniSnap führt Snapshots im DRAM durch, ohne dass auf die persistente Schicht geschrieben werden muss.
- ▶ **Niedrige RPO Asynchrone Replikation** – Die asynchrone Replikations-Engine kann Recovery Point Objective (RPO) von 4 Sekunden beibehalten und gleichzeitig die IP-Infrastruktur nutzen, um Kosten und Komplexität zu reduzieren.
- ▶ **Synchrone Replikation** – Die synchrone Replikations-Engine bietet eine synchrone Datensicherung ohne RPO bei einer Latenzzeit von weniger als 400µs (Mikrosekunden). Bei Problemen mit dem WAN (hohe Latenz, Verlust der Konnektivität) fällt die InfiniBox Synchron-Replikations-Engine automatisch in den asynchronen Modus zurück. Wenn das WAN wiederhergestellt wird, repliziert die Replikation automatisch alle fehlenden Daten und setzt die Synchronisationsreplikation ohne Unterbrechung der I/O fort.

Datenreduktion

InfiniBox verwendet mehrere Methoden zur Datenreduktion, um die Kosten für Speicherung weiter zu reduzieren, darunter:

- ▶ **Standardmäßiges Thin Provisioning** – Alle Volumes sind standardmäßig Thin Provisioning Volumes. Da die InfiniBox auch intelligente Kapazitätspools bietet, kann das Risiko einer Über-Allokation / Über-Bereitstellung durch das Setzen von Alarmschwellen und Notpuffern auf dem Pool leicht gemindert werden, wodurch die Verfügbarkeit der Anwendung geschützt wird.
- ▶ **Null-Rückruf** – Als Hosts (physischer oder virtueller) freier Speicherplätze auf einer Festplatte (LUN) schreiben sie Nullen in diesen Speicherplatz entweder durch die gleiche Schreiboperation (effizienter) oder einfach durch Schreiben einzelner Nullen in diesen Speicherplatz. InfiniBox identifiziert beide Fälle und entfernt diesen Speicherplatz, als ob er nie beschrieben worden wäre, was die Thin Provisionierung weiter verbessert.
- ▶ **Komprimierung** – Die InfiniBox komprimiert Daten erst, wenn sie aus dem Schreib-Cache (DRAM) auf die Festplatte transferiert werden. Dies beschleunigt das Schreiben (keine zusätzliche Latenz durch Datenreduktion) und vermeidet das Komprimieren von transienten Daten, die nach einigen Sekunden überschrieben werden (Einsparung von CPU-Ressourcen). Die InfiniBox Komprimierung unterstützt LZ4 mit einer 64KiB Chunk-Größe und erzeugt dabei ein höheres Kompressionsverhältnis als die herkömmliche Small-Block-Komprimierung (üblicherweise in allen Flash-Arrays verwendet).
- ▶ **Snapshots** – InfiniBox Snapshots sind dünn ausgelegt und helfen den Kunden so, die Kapazitätseinbußen einer vollständigen Kopie zu vermeiden.

Netzwerkarchitektur

Bei allen netzwerkbasierenden Diensten ist die Netz-Zugänglichkeit entscheidend für die Verfügbarkeit. Insbesondere bei Internet-Protocol (IP)-basierten Diensten (iSCSI, NFS, asynchrone Replikation, synchrone Replikation) erwarten IT-Administratoren in der Regel, dass das Speicher-System Failover handhabt und Konfigurationsprobleme schnell löst. InfiniBox hat in diesem Bereich Innovationen eingeführt, indem sie im Falle eines Verbindungsproblems ein sofortiges IP-Failover einsetzt und IP-Adressen auf Netzwerkschnittstellen verschiebt, die die entsprechenden Dienste bereitstellen können.

Diese „Instant IP Failover“ gilt für alle Ausfallszenarien, einschließlich Hardware (Knotenausfall, Ethernet-Port/Netzwerkkarten-Ausfall) oder Software (Dienstausfall auf einem bestimmten Knoten). Um die Auswirkungen auf andere Dienste zu minimieren, verschiebt die InfiniBox die minimale Anzahl von IP-Adressen, so dass IPs eines anderen Dienstes auf diesem Knoten oder IPs auf anderen Knoten nicht verschoben werden.

Die InfiniBox nutzt auch virtuelle MAC-Adressen (VMAC) und ordnet jede IP-Adresse einer VMAC zu. Wenn sich IP-Adressen verschieben, werden auch VMAC-Adressen mit verschoben. Dadurch entfällt die Failover-Zeit, so dass die Konfigurationsänderung auf dem Switch erfolgen kann, ohne dass die Änderung an jeden Host weitergegeben wird. Es hilft auch, das unnötige ARP-Problem zu vermeiden und die Verfügbarkeit zu erhöhen.

Die InfiniBox verwendet eine intelligente Netzwerküberwachung (mittels IPv6-Ping), um mögliche Fehlkonfigurationen zu erkennen, wie z. B. das versehentliche Blockieren einer Speicher-Netzwerkschnittstelle für den Zugriff auf ein VLAN, das für Datendienste verwendet wird. Jedes in der InfiniBox konfigurierte Netzwerk wird ständig überwacht und gibt Storage-Administratoren auf die Frage „Warum hat diese Anwendung den Zugriff auf den Speicher verloren?“ die Antwort oft schon lange bevor sie sich diese selbst stellen.

HARDWARE-ARCHITEKTUR

Die InfiniBox ist ein Software-defined Speichersystem, das COTS-Hardware nutzt. INFINIDAT hat im Rahmen des Designs in Software investiert, um COTS-Hardware zuverlässiger, kostengünstiger und einfacher zu verwalten und zu unterstützen. Das kritischste Konstruktionsprinzip heißt N+2: Alle Komponenten haben mindestens eine Dreifachredundanz - zwei "Reserveräder", die für sieben Neunen der Zuverlässigkeit ausgelegt ist. Das InfiniBox System wird vormontiert in einem Rack geliefert, wie im folgenden Bild gezeigt:

Knoten

Die Knoten sind die Speicher-Controller in der InfiniBox. Die drei vollständig redundanten Knoten arbeiten in einem Active-Active-Active-Cluster, so dass I/Os nahtlos durch alle drei Knoten fließen können. Die Knoten sind direkt mit dem schnellen InfiniBand für den direkten Zugriff auf den Speicher mittels RDMA verbunden, was eine schnelle Replikation zwischen den Knoten bei geringstmöglicher Latenz ermöglicht.

Ein Knotenausfall wird von den verbleibenden zwei Knoten behandelt, die ihre Verantwortung übernehmen. Dafür synchronisieren sie erneut einen Teil des Schreib-Caches, der nicht mehr repliziert wird, um die vollständige Datensicherung wiederaufzunehmen und den Betrieb unterbrechungsfrei aufrechtzuerhalten. Die N+2-Knotenarchitektur vereinfacht auch die Wartung eines bestimmten Knotens (z. B. das Ersetzen einer Komponente), da das System noch zwei Active-Active-Knoten hat, die die Daten ausführen und schützen.

3 Automatische Test Switches
 3 Batterie-Speichersicherungseinheiten (Battery Backup Units - BBUs)

3 Knoten

8 Plattenracks



BILD 1

InfiniBox Frontend- und Backend-Konnektivität

Physische Konnektivität

Frontend-Konnektivität von den Knoten zur Kundenstruktur:

- ▶ **Fiber Channel (FC)** – Acht Ports pro Knoten, insgesamt 24 Ports. Alle Ports sind aktiv, so dass jeder Host mehrere Pfade sieht (mindestens einer pro Knoten; zwei pro Knoten werden empfohlen). Durch Multi-Pathing kann ein Port oder HBA-Fehler nur den einzelnen Pfad und nicht die Anwendungen beeinflussen.
- ▶ **Ethernet (Eth) ports** – Vier Ports pro Knoten, insgesamt 12 Ports, die entweder Kupferleitungen oder optische Verbindungen bieten und die Protokolle iSCSI, NFS, synchrone Replikation und asynchrone Replikation sowie die Integration mit InfiniSync (INFINIDATs einzigartige Zero-RPO-Lösung über unbegrenzte Entfernungen) unterstützen. Diese Ports unterstützen ein intelligentes IP-Failover, um zu verhindern, dass ein physikalischer Fehler die Zugänglichkeit des Systems beeinträchtigt.

Intern bieten die Knoten auch die redundante Backend-Konnektivität:

- ▶ **InfiniBand (IB) ports** – Sie werden für die Cluster-Verbindung verwendet. Jeder InfiniBand Fehler, der eine Trennung von einem anderen Knoten verursacht, führt dazu, dass diese beiden Knoten über den dritten Knoten kommunizieren. Wenn ein Knoten von beiden verbleibenden Knoten getrennt wird, wird er aus dem Cluster entfernt, bis die Trennung aufgelöst ist.
- ▶ **SAS ports** – Verbinden der Knoten mit allen Festplattengehäusen. Jeder SAS-Fehler, der zum Verlust des Zugriffs eines bestimmten Knotens auf einige der Platten führt, verwendet das InfiniBand, um über einen anderen Knoten auf diese Platten zuzugreifen.

Die Knoten verfügen über redundante Stromversorgungen und werden von verschiedenen Battery Backup Units (BBUs) gespeist, die wiederum über mehrere Stromeingänge gespeist werden, um eine unterbrechungsfreie Stromversorgung zu ermöglichen.

Automatische Test Switches

Die automatischen Testschalter (Automatic Test Switches - ATS) steuern die Stromzufuhr zu den Battery Backup Units (BBUs) und stellen sicher, dass die Batterie auch bei einem Stromausfall in einer der Stromquellen immer mit Strom versorgt wird. Das ATS kann sofort zwischen zwei Stromquellen umschalten, wenn eine von ihnen ausfällt, so dass die Stromzufuhr zur BBU nicht unterbrochen wird.

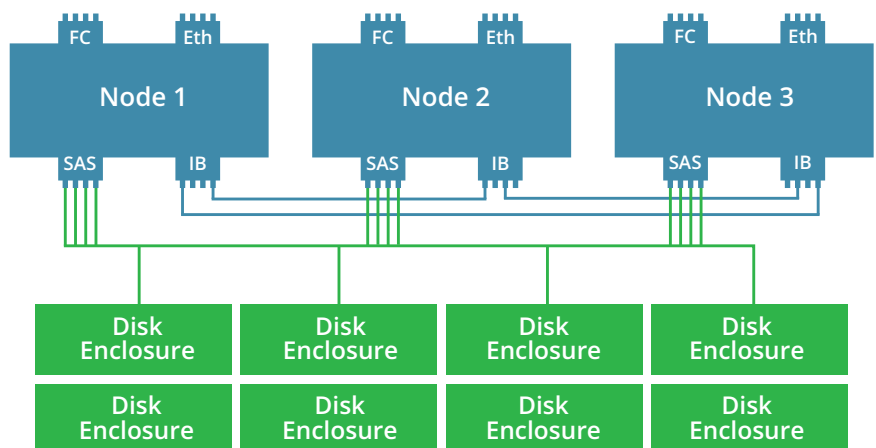


BILD 2 InfiNode Frontend- und Backend-Konnektivität

Battery Backup Units

Die BBUs versorgen die InfiNode Knoten bei kurzen Stromausfällen (z. B. bis die Generatoren voll aktiv sind), ohne dass das System heruntergefahren werden muss. Sie bieten auch die Möglichkeit, Daten im Falle eines längeren Stromausfalls ordnungsgemäß aus dem DRAM-Cache zu entfernen („destage“), so dass die InfiNode immer die richtigen Abschaltvorgänge durchführen kann. Die BBUs werden überwacht und jede von ihnen wird automatisch einmal pro Woche getestet, um sicherzustellen, dass ihre Batterien in Ordnung sind und sie bereit stehen, das System im Falle eines echten Stromausfalls zu schützen.

FAZIT

Die einzigartige InfiniBox Architektur bricht mit den traditionellen Kompromissen zwischen Leistung, Ausfallsicherheit, Kapazität und Kosten. Zum ersten Mal kann die IT/Infrastruktur es Unternehmen ermöglichen, alle wichtigen Daten zu speichern, ohne die IT-Budgets aufzublähen oder ihr Betriebsergebnis zu beeinträchtigen. Mit der InfiniBox sollen sich Unternehmen sicher fühlen, wenn sie ihre digitale Transformation und ihre Big-Data-Initiativen in Angriff nehmen.