



SVA WHITEPAPER

BIG DATA²

SCHNELLERE ANALYSEN DANK FLASH-BASIERTEM SPEICHER



**NUR EINE ZEITNAHE
ANALYSE LIEFERT AUCH
ERKENNTNISSE**

VORWORT

Big-Data-Analysen sind für viele Unternehmen eine große Herausforderung: Welche Big-Data-Technologie soll genutzt werden? Apache Hive, Hadoop oder Cassandra? Sollen wir in die Cloud gehen oder On-Premises nutzen? Welches Daten-Management ist das richtige?

Es gibt viele wichtige Fragen, die Unternehmen im Vorfeld beantworten müssen. Eine Herausforderung ist vor allem, wie Sie riesige unstrukturierte Datenmengen in kürzester Zeit analysieren können. Denn nur eine zeitnahe Analyse liefert Ihnen die nötigen Erkenntnisse und Handlungsempfehlungen.

Dauert die Big-Data-Analyse zu lange, kann es sein, dass die Daten bereits veraltet sind und schon neuere vorliegen. Gerade Echtzeitanalysen sind wichtig für Fachbereiche, die mit „predictive Analytics“ oder „predictive Maintenance“ arbeiten, z. B. im Bereich Finanz- und Risiko-Controlling, Wetter, Logistik und Verkehr, IoT u. v. a. Hier gilt buchstäblich das Motto „Zeit ist Geld“.

Aus diesem Grund überlegten wir uns, wie wir Big-Data-Analysen noch schneller und effizienter machen können. Das Ergebnis ist ein gemeinsamer Test von SVA und NetApp, in dem wir zeigen, dass der Betrieb von Hadoop Clustern auf Flash-basiertem Storage von NetApp performanter, einfacher und flexibler ist als herkömmliche Hadoop Installationen auf lokalen Drives.

Der Fokus lag auf der transparenten Nutzung des NFS-Protokolls für Hadoop unter Verwendung des Moduls „NetApp In-Place Analytics Module“ (oder auch „NFS Connector für Hadoop“). Das vorliegende Whitepaper schildert die Testumgebung, die Durchführung der einzelnen Tests, die Ergebnisse der TeraSort-Benchmarks und die daraus abgeleitete optimale Vorgehensweise beim Betrieb von Hadoop auf einer konvergenten Infrastruktur-Plattform von NetApp und Fujitsu.

**DATEN LASSEN SICH MIT
MODERNEN BIG-DATA-
ANWENDUNGEN EINFACH
AUSWERTEN**

**STETS AKTUELLE DATEN
ZU ALLEN WICHTIGEN
BUSINESS-BEREICHEN**

1. VORGEHENSWEISE

Im Zuge der „digitalen Transformation“ beginnen viele Unternehmen, ihre bereits bestehenden Daten aus Kunden- und Lieferantenbeziehungen zu analysieren. Selbst große, über die Jahre gewachsene Datenbestände aus ERP-Systemen, Produktions- und Logistikdatenbanken lassen sich mit Hilfe moderner Big-Data-Anwendungen wie des hier verwendeten Hadoop mit kalkulierbarem Aufwand an Zeit und Geld auswerten.

Hadoop verwendet viele parallel laufende Auswertungsprozesse und liefert damit in Echtzeit Informationen über zeitliche und inhaltliche Zusammenhänge: z. B. Einkaufspreise, Logistikbewegungen, Produktionszahlen, Kaufverhalten und Umschlagzeiten. Business Intelligence mit Hadoop macht auch subtile Abhängigkeiten sichtbar.

Flexibel und leicht selbst anzupassende Visualisierungswerkzeuge im modernen Look and Feel gängiger ERP-Lösungen liefern stündlich statt täglich die jeweils aktuelle Datenbasis für sichere, umsatzrelevante Entscheidungen rund um:

- Bestandsoptimierung
- Preisoptimierung
- Predictive Maintenance
- Prozessoptimierung
- Kundenverhalten
- Betrugserkennung

Die Nutzung von Informationen sowie Schnelligkeit und Nachhaltigkeit sind Kernziele der digitalen Transformation. Anhand der positiven Erfahrungen mit gemeinsamen Kunden in zahlreichen Big-Data-Projekten haben die langjährigen Partner NetApp und SVA eine schnell zu integrierende Big-Data-Lösung für Hadoop entwickelt. Diese verbindet Flexibilität, Performance und Skalierbarkeit bestmöglich. Zentrales Element der Lösung ist neben NetApp All-Flash Storage das „NetApp In-Place Analytics Module“. Dieses kompakte Software-Modul versetzt z. B. die Analytics Software Cloudera Hortonworks HDP bis Version 3.1 in die Lage, transparent und performant beliebige Daten über das NFS-Protokoll zu analysieren und dabei die bekannten Mehrwerte der NetApp ONTAP Data Management Software zu nutzen (u. a. Komprimierung, Snapshots, Cloning, Replikation, integriertes Backup und Recovery).

Die ONTAP Data Management Software ist das Betriebssystem der NetApp FAS- und AFF-Storage-Plattformen. Verfügbar ist es auf nativer NetApp Storage Hardware, als Software-Defined Storage (SDS) auf geeigneter X86-Server-Hardware und als virtuelle Cloud-Storage-Instanz bei Cloud-Anbietern wie AWS, Microsoft Azure und Google GCP. NetApp In-Place Analytics Module setzt also in der einen oder anderen Form NetApp Speicher voraus.

HADOOP INSTALLATIONEN SIND OFT SEHR ZEITAUFWENDIG

DIE IDEE: EINE BASIS- INFRASTRUKTUR FÜR BIG-DATA-Projekte ZU SCHAFFEN

Das vorliegende Dokument beschreibt die Lösungs-Architektur, erklärt die nachvollziehbaren Mehrwerte für den produktiven Betrieb von Hadoop im Unternehmen, schildert die Durchführung sowie die Ergebnisse der bekannten TeraSort Benchmarks für Big Data. Ebenso werden die daraus abgeleiteten besten Vorgehensweisen beim Betrieb von Hadoop auf der konvergenten Infrastrukturplattform von NetApp und Fujitsu dargestellt.

1.1 HINTERGRUND

Nach den Erfahrungen der SVA Spezialisten fällt unabhängig von der Größe und Laufzeit eines Big-Data-Projektes immer ein beträchtlicher zeitlicher Aufwand für die Vorbereitung einer leistungsfähigen Hadoop Infrastruktur an. Je nach Ausstattung des Kunden werden bestehende Komponenten neu konfiguriert, teilweise mehr oder minder aufwendig erweitert – zum Beispiel mit lokalen SSDs geeigneter Größe – und nach erprobten Vorgehensweisen zu einer leistungsfähigen Hadoop Plattform zusammengefügt. Mitunter liefert SVA zusätzliches Equipment für einen Proof of Concept zum Verbleib beim Kunden oder stellt im Rahmen einer Leihstellung eigenes Material zur Verfügung.

Das führt dazu, dass jedes Hadoop Projekt sehr individuell ist. Folglich sind die Möglichkeiten eingeschränkt, die Inbetriebnahme der Umgebung zu automatisieren. Nicht alle optimierten Vorgehensweisen lassen sich in jeder Umgebung realisieren, Abläufe müssen zeitaufwendig mit dem vorhandenen Material neu erlernt werden. Damit steigt auch die Wahrscheinlichkeit, dass Fehler und entsprechende Verzögerungen und Folgekosten entstehen.

Daher entstand schließlich die Idee, gemeinsam eine Basisinfrastruktur für Big-Data-Projekte auf Basis von Hadoop zu verifizieren, die den genannten Einschränkungen nicht unterliegt.

Die Infrastrukturebenen „Compute“ und „Netzwerk“ sind in ihren typischen Leistungsparametern wie CPU Cores, Taktrate, RAM, Bandbreite und Latenz praktisch von allen Herstellern in vergleichbarer Qualität zu bekommen. Einzig im „Storage Layer“ sind eindeutige Differenzierungsmerkmale für NetApp auszumachen, die über reine Leistungsdaten hinaus vor allem Einfluss auf den Betriebsaufwand und damit die laufenden Kosten für eine produktive Hadoop Umgebung haben.

Für den Endkunden ergeben sich folgende Vorteile durch Verwendung einer Standardinfrastruktur:

- Schneller, zuverlässiger und planbarer Projektbeginn durch immer gleiche Konfiguration
- Unmittelbar betriebsbereit als Leihstellung (von SVA) bzw. geringster Aufwand bei Neukauf
- Bekannte Leistungsparameter, erprobte Skalierbarkeit in Performance und Kapazität
- Geringstes Fehlerpotenzial durch (teil-)automatisiertes Befolgen von Best Practices
- Fertige Betriebsmodelle für Monitoring, Backup, Disaster Recovery inkl. Lifecycle Management
- Bekannte und eingespielte Support-Abläufe zwischen SVA und Hersteller – vom Proof of Concept bis zur Überführung in den Produktivbetrieb und darüber hinaus

**HADOOP WURDE FRÜHER
PRIMÄR ZUR ANALYSE
EXTERNER DATEN GENUTZT**

Im vorliegenden Fall profitiert man aufseiten von SVA und ihren Kunden von der langjährigen Partnerschaft zwischen NetApp und Fujitsu einerseits und der intensiven Partnerschaft beider Hersteller mit SVA andererseits.

Mit der konvergenten Infrastrukturplattform der beiden Hersteller hat SVA bereits eine ausgezeichnet geeignete Lösung im eigenen Portfolio, die Compute, Netzwerk und Storage mit exzellenten Ergebnissen in Bezug auf Leistungsfähigkeit, Skalierbarkeit und Flexibilität vereinigt. Einfach zu bestellen, zu implementieren, zu betreiben und mit Rundum-Support abgesichert. Die konvergente Architektur erfüllt alle aktuellen Anforderungen an eine moderne Datacenter-Infrastruktur von Compute über Netzwerk bis zum Speicher.

An erster Stelle steht für SVA jedoch immer die Sicherheit des Kunden. Für die SVA Big-Data-Spezialisten war es daher eine Selbstverständlichkeit, die konvergente Infrastrukturplattform bei einem aufwendigen Proof of Concept für den Betrieb von Hadoop intensiv zu testen. Die nachfolgenden Abschnitte dokumentieren die Vorbereitungen, den Ablauf sowie die Ergebnisse dieser Tests.

1.2 NETAPP IN-PLACE ANALYTICS MODULE

Historisch betrachtet, hat man Hadoop primär zur Analyse von „externen Daten“ genutzt, d. h. von aus anderen Systemen exportierten oder aggregierten Daten. Die neuen Daten wurden dann in ein HDFS-Filesystem kopiert und anschließend ausgewertet. Bewährt, aber nicht immer optimal.

Parallel dazu gibt es immer auch den Bedarf, bereits im Unternehmen vorhandene große Datenbestände auszuwerten, die meist in Network-Attached-Storage-Systemen (NAS) liegen. Nach dem beschriebenen Muster würde man diese Datenbestände in ein vorhandenes oder neu anzulegendes HDFS-Silo hineinkopieren, um anschließend die Auswertung zu starten. Dieser Prozess muss administriert werden. Aufgrund der doppelten Datenhaltung in NAS und HDFS ist dieser Vorgang ineffizient und belastet unnötig die Ressourcen der beteiligten Systeme und des Netzwerkes.

Das „NetApp In-Place Analytics Module“ für Hadoop erlaubt es jedem Kunden mit NetApp NAS Storage, die vorhandenen ebenso wie neu eintreffende Daten durch Big-Data-Analysen auszuwerten. Und zwar ohne diese Daten zu bewegen oder gar zu kopieren! Kostentreibende Faktoren, wie z. B. die Erstellung von Datenkopien, fallen weg, und zusätzlich kommt man ohne schwierige Synchronisierungsoperation zu einem „Single Point of Truth“ für alle Daten im Unternehmen. Der Aufbau eines unternehmensweiten „Data Lake“ wird zum sprichwörtlichen Kinderspiel. Network Attached Storage (NAS) von NetApp verbindet über das NFS-Protokoll problemlos all die vielen Einzeldaten im Unternehmen: z. B. Datenbanken, Fileshares, Export- und Log-Verzeichnisse.

Das NetApp Whitepaper „[Designing and Building a Data Pipeline for Your AI Workflows](https://www.netapp.com/us/media/wp-7264.pdf)“ (<https://www.netapp.com/us/media/wp-7264.pdf>) gibt einen ausgezeichneten Überblick über die Vorteile einer durchgängigen Data Pipeline „from Edge to Core to Cloud“ für datenzentrierte Projekte wie Big Data, AI, Machine Learning und Deep Learning.

**DATEN MÜSSEN WEDER
BEWEGT NOCH KOPIERT
WERDEN**

**MEHR CPU- UND
MEMORY-RESSOURCEN
STEHEN ZUR VERFÜGUNG**

Im Finanzumfeld wie im Gesundheitswesen reduziert dies außerdem den immensen Abstimmungsaufwand zur Einhaltung rechtlicher Vorgaben für die Datenspeicherung. In diesem Szenario analysiert das „In-Place Analytics Module“ die Daten einfach auf dem Original-Speicherort.

Das NetApp In-Place Analytics Module entkoppelt ganz bewusst den reinen Analyse-Vorgang (bspw. den MapReduce-Prozess) von Storage-Management-Aufgaben (bspw. die HDFS-Verwaltung). Für den Analyseprozess stehen so mehr CPU- und Memory-Ressourcen zur Verfügung. Das Modul nutzt die offiziellen File System Interfaces von Hadoop. Die aktuelle Version basiert auf dem NFSv3-Protokoll und unterstützt die AUTH_NONE und AUTH_SYS Security Policies für die UNIX-spezifische Authentifizierung.

Abbildung 1 zeigt die Integration des In-Place Analytics Modules unterhalb des Hadoop File-Systems bzw. Storage Layers. Auf diese Weise lässt sich NFS im einfachsten Fall parallel zu HDFS betreiben, kann HDFS aber sogar komplett ersetzen. Je nach Anwendungsfall und abhängig von den Performance-Anforderungen der Applikation kann die eine oder andere Implementierung von Vorteil sein.

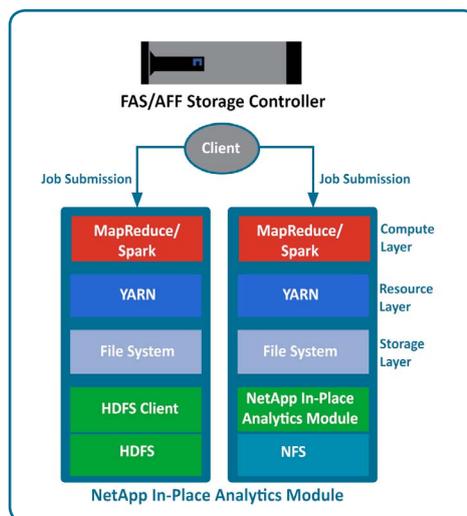


Abbildung 1: NetApp In-Place Analytics Module

Unabhängig davon sind keinerlei Änderungen innerhalb der Applikationen notwendig. Die Verwendung des NFS-Protokolls anstelle von HDFS ist völlig transparent für die darüber liegenden Layer innerhalb von Hadoop. Die Hadoop Betriebsmannschaft, ebenso die Data Engineers und Data Architects, können unverändert die nativen Hadoop Kommandos nutzen. Die etablierten Prozesse für Monitoring, Datenmanagement und Datensicherung müssen nicht geändert werden.

**KEINERLEI ÄNDERUNGEN
INNERHALB DER ANWEN-
DUNG NOTWENDIG**

**DAS NETAPP IN-PLACE
ANALYTICS MODULE
BIETET VIELE VORTEILE**

Auf der Habenseite stehen zudem zahlreiche Speicher-Management-Funktionen der Enterprise-Klasse von NetApp zur Verfügung, die den laufenden Betrieb, Replikation und Disaster Recovery auch größter Datenbestände beherrschbar machen. NetApp Storage und das NFS-Protokoll haben ihre Fähigkeiten und ihr Leistungsvermögen seit langen Jahren in Tausenden von ORACLE Datenbank-Umgebungen nachgewiesen, ebenso bei großen und größten VMware Server-Farmen. Weltweit und auch in Deutschland.

Die Vorteile:

- Keine Änderungen bei vorhandenen Applikationen nötig.
- Keine Änderungen bei vorhandenen Hadoop Clustern nötig – nur zwei Konfigurationsdateien müssen angepasst werden.
- Keine aufwendiges Kopieren, da Daten am Original-Speicherort analysiert werden.
- Datensicherheit der Enterprise-Klasse – mit RAID-Schutz; keine drei HDFS-Repliken notwendig.
- Effizienztechniken wie Deduplizierung, Kompression, Compaction sparen Kapazität und Kosten.
- Entkoppelt Compute und Storage – mehr Ressourcen für Analytics und unabhängige Skalierung.
- Unified Storage mit allen relevanten Speicher-Protokollen – any protocol, any time, anywhere.
- NetApp Speicher-Erweiterungen und Upgrades bleiben transparent für Hadoop.
- Netzwerkverkehr wird automatisch über mehrere Netzwerk-Interfaces verteilt.
- Hortonworks Zertifizierung für die Hortonworks Data Platform (HDP) liegt vor.
- Installation und Konfiguration über ein mitgeliefertes Apache Ambari-Modul.
- Integriertes Management von User, Access und Configuration Management im UI.
- Azure Integration: Azure HDInsight Cluster arbeiten mit NFS-Daten.
- „NetApp Data Fabric“ – hybride Analyse-Lösungen aus On-Premises- und Cloud-Instanzen.

**LATENZEN UNTER EINER
MILLISEKUNDE MIT ALL
FLASH FAS (AFF)**

1.3 NETAPP CLOUD-CONNECTED ALL-FLASH STORAGE

Das intelligente, modulare Design der Speicher-Systeme der NetApp FAS- und AFF-Familien ist auf hohe Performance, Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit ausgelegt.

Bei den Systemen der FAS-Serie handelt es sich um sogenannte hybride Speicher-Systeme, die sowohl hochkapazitive und kostengünstige Festplatten als auch schnelle Flash-SSDs unterstützen. So kann man SSDs als Cache für die langsameren Festplatten nutzen oder bei Bedarf auch separate All-Flash-Speicherbereiche mit SSDs für hohe Performance schaffen.

HOCHVERFÜGBARKEIT UND GRENZENLOSE SKALIERBARKEIT

AFF-Systeme unterstützen hingegen keine Festplatten mehr. Es handelt sich um reinen All-Flash-Speicher. Das eröffnet den Weg zu gezielten Optimierungen hinsichtlich Latenz und Durchsatz, wodurch AFF-Systeme nochmal verbesserte Performance-Werte mit SSDs gegenüber FAS aufweisen. All Flash FAS (AFF) vereint führendes Datenmanagement mit Latenzen unter einer Millisekunde, sodass sich Daten nach Bedarf in die gewünschte Umgebung verschieben lassen, um die optimale Kombination aus Performance, Speicher-Kapazität und Kosteneffizienz zu erzielen.

Als Teil eines Scale-out-Clusters können alle NetApp FAS-Systeme mit All-Flash-basierten AFF-Systemen kombiniert werden. ONTAP Cluster stellen Hochverfügbarkeit bei Wartungsvorgängen und praktisch grenzenlose Skalierbarkeit sicher. Das aktuelle Flaggschiff FAS9000 beispielsweise lässt sich mühelos mit NetApp AFF-Arrays auf bis zu 172 PB skalieren und zudem mit vorhandenen Storage Arrays von Drittanbietern zusammen nutzen.

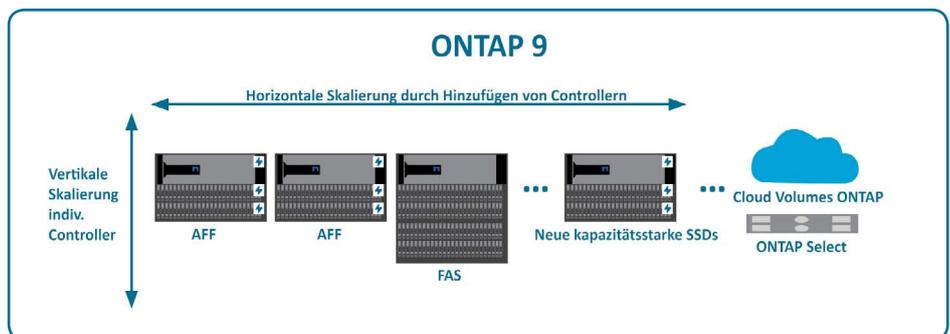


Abbildung 2: Nahtlose Skalierbarkeit

EINFACH HORIZONTAL ODER VERTIKAL SKALIEREN

Unternehmen können horizontal skalieren, indem sie ihre Flash- und Hybrid-Flash-Nodes kombinieren, Upgrades für Hardware und Software durchführen oder auch vertikal skalieren, ohne Benutzer zu stören. Und sie haben die Möglichkeit, Flash der nächsten Generation zu integrieren, auch Software-definiert und in der Cloud. Beide, die FAS- und die AFF-Serie, werden durch die Datenmanagement-Software NetApp ONTAP unterstützt und bilden eine Unified-SAN- und NAS-Infrastruktur in einem Speichersystem. Durch einen umfassenden Ansatz zur Ausfallsicherheit erfüllen alle FAS- und AFF-Modelle höchste Ansprüche an die Verfügbarkeit von mindestens 99,9999 %. Dieser Ansatz kombiniert hochgradig zuverlässige Hardware, innovative Software und ausgereifte Service-Analysefunktionen.

Merkmale und Funktionen von NetApp ONTAP

NETAPP ONTAP	
Universelle Datenplattform	<ul style="list-style-type: none"> ■ Alle Protokolle: CIFS, FCoE, FCP, iSCSI, NFS, NVMe ■ Alle Datenträger: NVMe- und Flash-SSD, SAS, NL-SAS, SATA ■ Alle Workloads: kleine oder große Blöcke, zufällig oder sequenziell ■ Maximale Benutzerfreundlichkeit: ein System für alle Anforderungen
Flexible Skalierbarkeit	<ul style="list-style-type: none"> ■ Erweiterbare Performance: horizontale und vertikale Skalierung bei Controllern und Festplatten ■ Erweiterbare Kapazität: horizontale und vertikale Skalierung ■ Betriebliche Effizienz: Wachstum ohne zusätzliche Managementpunkte
Integrierte Datensicherung	<ul style="list-style-type: none"> ■ Native, platzsparende Snapshot-Kopien ■ Enge Applikations-Integration (applikationskonsistente Snapshots etc.) ■ Native Disaster-Recovery-Replizierung: SnapMirror ■ Natives Disk-to-Disk-Backup: SnapVault
Offene und erweiterbare Plattform	<ul style="list-style-type: none"> ■ Leistungsstarke APIs für Drittanbieter- und individuelle Integrationen ■ Verschiedene Support-Tools und -Umgebungen ■ Skalierbarkeit von kleinen auf größte Umgebungen
Lebenszyklus-Wartung ohne Ausfallzeiten	<ul style="list-style-type: none"> ■ Austausch aller Komponenten ohne Ausfallzeit oder teure Migrationen ■ Hinzufügen von Performance, Anschlüssen und Kapazität ohne Unterbrechung
Sichere Mandantenfähigkeit	<ul style="list-style-type: none"> ■ Shared Storage für maximale Effizienz bei gleichzeitig sicherer Trennung ■ Sichere und skalierbare mandantenfähige Lösung mit voller Funktionalität ■ Leistungsfähige QoS und rollenbasierte Zugriffssteuerung zur Erfüllung von Service-Leveln
Geschäftsfördernde Effizienz	<ul style="list-style-type: none"> ■ Deduplizierung auf 4-KB-Blockebene mit hoher Granularität für NAS- und SAN-Primärdaten ■ Cache-Optimierung dank deduplizierungsfähiger SSD Flash Pools ■ Nahezu sofortige Klone zur schnellen Erstellung von Entwicklungs-/ Test- und VM-Umgebungen
Vollständige Datenportabilität	<ul style="list-style-type: none"> ■ Verschieben von Daten zu jedem Zeitpunkt und auf jeden Festplattentyp ohne Unterbrechung ■ Verschieben von Daten zwischen Private, Hybrid und Public Clouds ■ Hocheffiziente, deduplizierungsfähige Replizierung

**DIE SKALIERBARKEIT IST
INDIVIDUELL ANWEND-
BAR AUF COMPUTING,
ARBEITSSPEICHER UND
STORAGE**

**GEDACHT UND GEMACHT
FÜR ENTERPRISE-
UMGEBUNGEN**

1.4 KONVERGENTE INFRASTRUKTURPLATTFORM VON FUJITSU UND NETAPP

Die gemeinsame, konvergente Infrastrukturlösung (CI) von Fujitsu und NetApp basiert auf einer einfachen, modularen und multidimensionalen Skalierung, die individuell auf Computing, Arbeitsspeicher und Speicher anwendbar ist – ein Alleinstellungsmerkmal im Vergleich zu anderen CI-Lösungen. Das bedarfs- und nutzungsbasierte Preismodell ist für Entscheidungsträger und Käufer von IT-Infrastrukturen zudem ein entscheidender Faktor.

Je nach Kundenanforderung stehen im Rack montierte Systeme und Konfigurationen auch ohne Rack zur Verfügung. Die konvergente Infrastrukturlösung wird auf Ihre individuellen Bedürfnisse konfiguriert: die Anzahl der Server Nodes, Memory- und Speicherkapazitäten sowie die Art des Storage – Hybrid oder All-Flash – lassen sich ganz flexibel konfigurieren. Auch wenn neue Geschäftsanforderungen aufkommen, kann Ihr Unternehmen Rechen- und Storage-Kapazitäten nach Bedarf hinzufügen. Mit einer präzisen Skalierung wird gewährleistet, dass nur die wirklich benötigten Kapazitäten berechnet werden. Gleichzeitig wird die Interoperabilität mit bestehenden IT-Umgebungen garantiert.

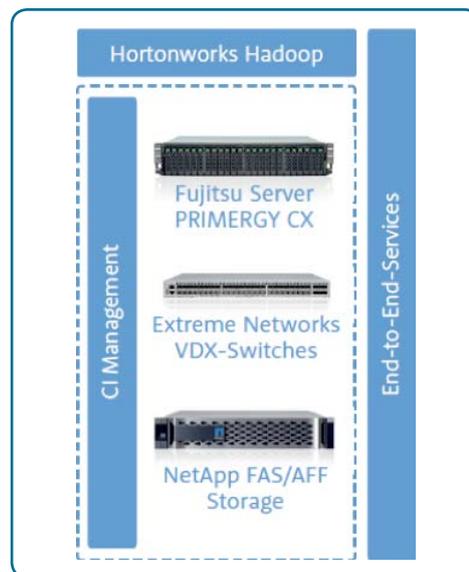


Abbildung 3: Konvergente Infrastrukturplattform

Konvergente Infrastrukturplattform von NetApp und Fujitsu – gedacht und gemacht für Enterprise-Umgebungen:

- Vorkonfiguriert, getestet und abgenommen verlassen die Konfigurationen die Produktion
- Betriebsbereit ab Tag 1
- Best-in-Class-Komponenten von Fujitsu and NetApp:
 - Fujitsu Server PRIMERGY CX
 - NetApp AFF- und FAS-Speicher – einschließlich aller Funktionen
 - Extreme Networks VDX Switches
 - Startklar für VMware vSphere, Hortonworks HDP und viele andere

**BARE-METAL-INSTALLATION
AUF BASIS VON CENT OS
LINUX UND HORTON-
WORKS HDP**

2. TEST-AUFBAU

Die Architektur des Proof of Concept wird folgendermaßen dargestellt:

2.1 COMPUTE-ARCHITEKTUR

Im Proof of Concept wurde eine Bare-Metal-Installation auf Basis von Cent OS Linux und Hortonworks HDP verwendet. Die Installationen verteilen sich auf die vier Server des Typs Fujitsu PRIMERGY CX. Folgend eine Darstellung der in Hortonworks verwendeten Rollen, die auf die Server verteilt wurden.

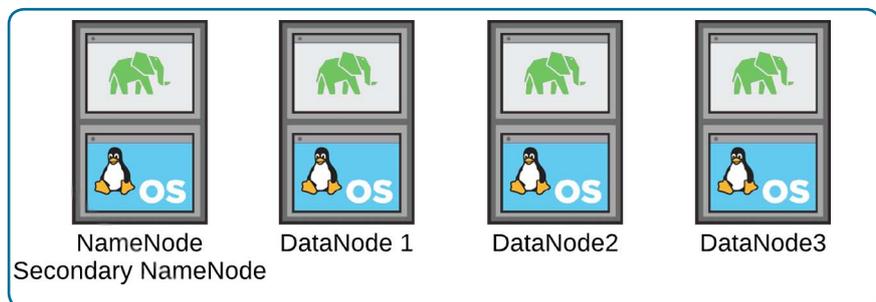


Abbildung 4: Compute-Architektur

Der erste Server erhielt die Rollen des Name Nodes und des Secondary Name Nodes. Die restlichen Hosts dienen als sogenannte Hadoop Worker Nodes, welche die eigentliche Datenverarbeitung während des Tests durchführten.

2.2 NETZWERK-ARCHITEKTUR

Das Netzwerk wurde wie im folgenden Bild dargestellt für die Tests konfiguriert:

**ACHT 10GBE-PORTS
WURDEN VERWENDET**

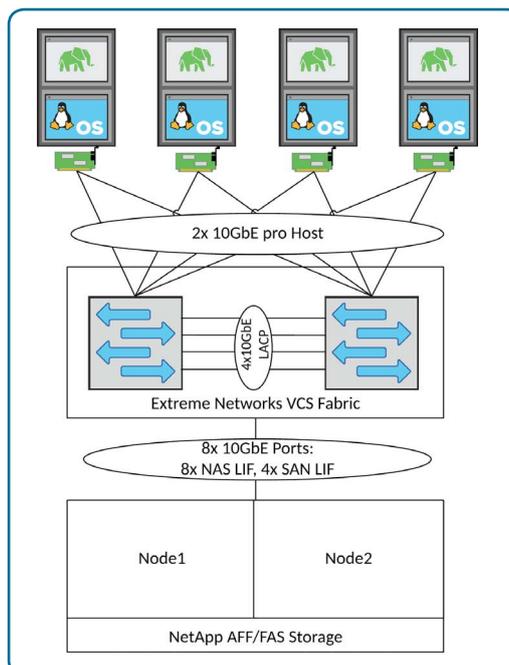


Abbildung 5: Netzwerk-Architektur

**REDUNDANZ UND
PERFORMANCE-VORTEIL
DURCH SWITCH-ÜBER-
GREIFENDE KANAL-
BÜNDELUNG**

Das Netzwerk wurde über zwei Switches des Typs Extreme Networks VDX 6740 bereitgestellt. Beide Geräte wurden als eine VCS Fabric konfiguriert. Mit dieser Konfiguration wird sowohl Redundanz als auch ein Performance-Vorteil durch Switch-übergreifende Kanalbündelung mittels LACP für die angeschlossenen Hosts ermöglicht. Die Hosts wurden jeweils mit zweimal 10GbE im LACP-Verbund angebunden. Aufseiten des Storage-Systems wurden insgesamt acht 10GbE-Ports verwendet. Diese stellten die Grundlage für insgesamt acht NAS LIFs (Logical Interface) und vier SAN LIFs dar.

2.3 STORAGE-ARCHITEKTUR

Die Storage-Architektur wurde für die Tests wie im folgenden Bild dargestellt verwendet:

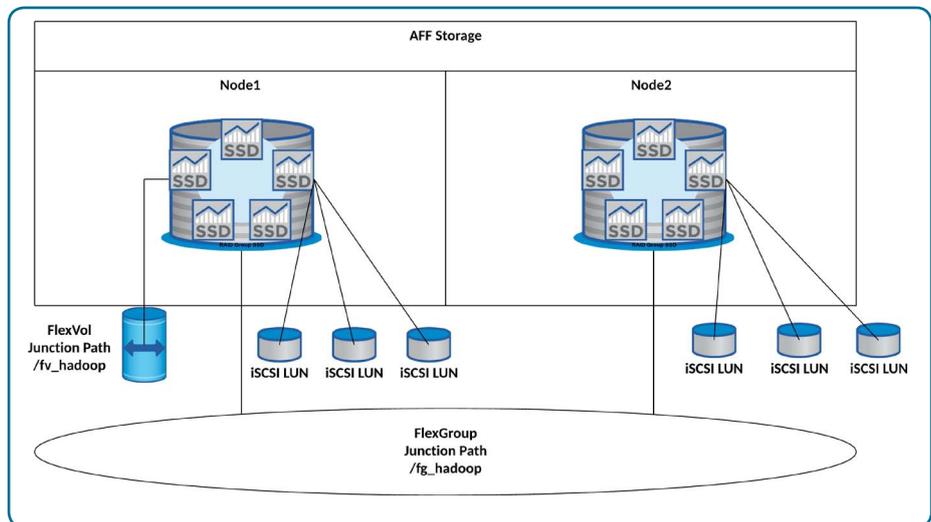


Abbildung 6: Test-Storage-Architektur NetApp AFF A220

Im ersten Test wurden via NFS und iSCSI bereitgestellte Speicherbereiche verwendet.

NFS:

- NetApp FlexVolume – genutzt als Speicherbereich via NetApp In-Place Analytics Module, innerhalb eines NetApp Nodes
- NetApp FlexGroup – genutzt als Speicherbereich via NetApp In-Place Analytics Module, über beide NetApp Nodes verteilt

iSCSI LUN:

- genutzt als Datenträger via HDFS, über beide NetApp Nodes verteilt

Für die weiterführenden Tests wurde aufgrund der ersten Ergebnisse ausschließlich die Verwendung von NFS mit FlexGroups als Datenbereich betrachtet. Der im Hadoop Konstrukt notwendige YARN Cache wurde ebenfalls via NFS auf Basis von FlexVolumes angebunden.

ZUSÄTZLICH MIT EINEM NETAPP STORAGEGRID

Die nächste Abbildung stellt die verwendete Speicher-Architektur im Test der NetApp AFF A700 dar. In dieser Konstellation wurde zusätzlich ein NetApp StorageGrid eingebunden, um den Einfluss des NetApp StorageGrids als FabricPool zu untersuchen.

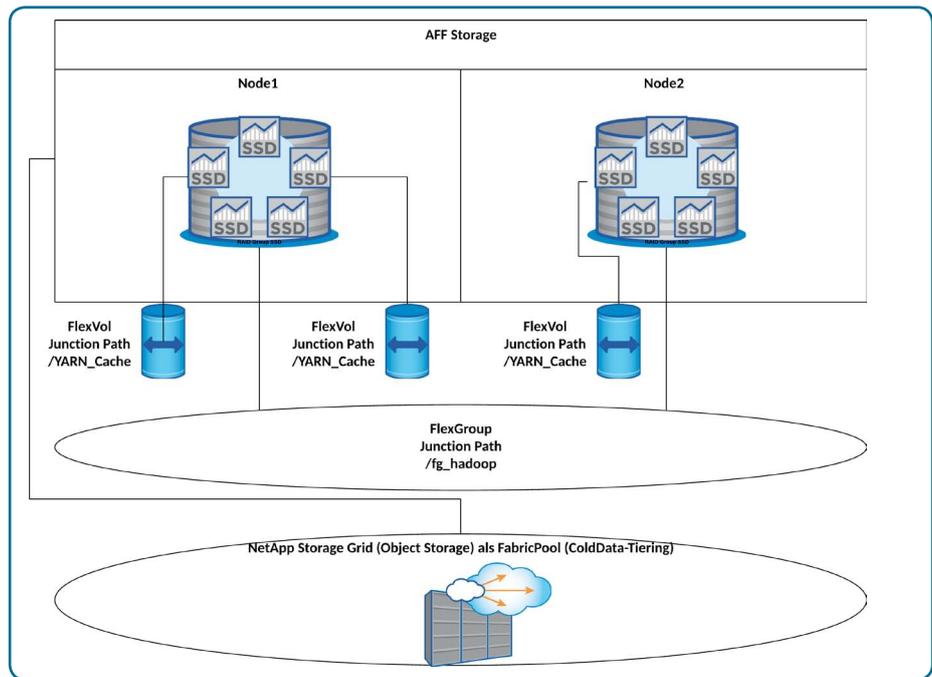


Abbildung 7: Test-Storage-Architektur NetApp AFF A700

Weiterhin wurde auch ein Test mit einer NetApp FAS (basierend auf SAS- und NL-SAS-Festplatten) durchgeführt. Die Darstellung bezüglich der Speicher-Architektur sieht folgendermaßen aus:

EIN WEITERER TEST MIT EINER NETAPP FAS

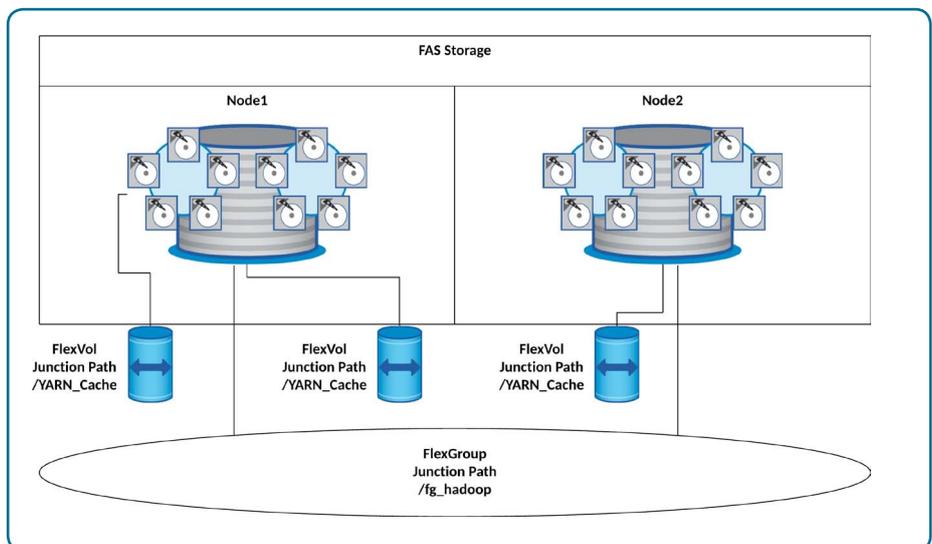


Abbildung 8: FAS-Speicher-Architektur

2.4 HARDWARE- UND SOFTWARE-KOMPONENTEN

Die Hardware- und Software-Produkte, mit denen die Benchmarks ermittelt wurden, werden hier aufgeführt. Im Einzelfall finden sich weitere Informationen zu den installierten Versionen und Einstellungen in den nachfolgenden Unterabschnitten.

Hardware- und Software-Komponenten für den ersten Proof of Concept

Komponente	Produkt oder Lösung	Details
Server	4x Fujitsu PRIMERGY Server CX2560 M4, jeweils 384 GB RAM, je 2 Stück 12 Core Intel CPU 2,1 GHz	Intel „Xeon Silver 4116“
Netzwerk	2x Extreme Networks VDX6740 Switch, 1x Extreme Networks VDX6740T Switch	Extreme VDX6740 für Daten als VCS Fabric Extreme VDX6740T für Management
Speicher	AFF 220A, 12x 960 GB SSD, 2x 10GbE Cluster Interconnect, 8x UTA2 (16 Gbps FC/10GbE) für Host Connectivity	ONTAP 9.4, 2-Node Switchless Cluster, Flash Bundle, iSCSI und NFSv3, UTA2 als 10GbE, weitere Anpassungen gemäß Technical Report TR-4382 Abschnitt 5.1 https://www.netapp.com/us/media/tr-4382.pdf
Server-Betriebssystem	CentOS Linux Release 7.5.1804 (Core)	
Java	OpenJDK Runtime Environment	Build 1.8.0_161-b14
Hadoop Distribution	Hortonworks 2.6.5.0-292 Rollout mittels Apache Ambari 2.6.2.0	2 Name Nodes, 3 Data Nodes
TeraSort Benchmark	TPCx-HS_Kit_v2.0.3	
In-Place Analytics Module	hadoop-nfs-connector-3.0.0, NetApp_Ambari_Plugin_1.0	
Monitoring und Dokumentation	NABOX, NetApp SystemManager, ONTAP CLI, Ganglia	

ALLE KOMPONENTEN DES PROOF OF CONCEPTS AUF EINEN BLICK

**ALLE KOMPONENTEN DES
ZWEITEN PROOF OF
CONCEPTS**

Hardware- und Software-Komponenten für den zweiten Proof of Concept

Komponente	Produkt oder Lösung	Details
Server	4x Fujitsu PRIMERGY Server CX2560 M4, jeweils 512 GB RAM, je 2 Stück 12 Core Intel CPU 2,1 GHz	Intel „Xeon Silver 4116“
Netzwerk	2x Extreme Networks VDX6740 Switch, 1x Extreme Networks VDX6740T Switch	Extreme VDX6740 für Daten als VCS Fabric Extreme VDX6740T für Management
Speicher	AFF A700A, 24x 960 GB SSD, 2x 40 GbE Cluster Interconnect, 8x UTA2 (16 Gbps FC/10GbE) für Host Connectivity FAS8040A, 24x 2TB NL-SAS/SAS, 2x 10GbE Cluster Interconnect, 4x UTA2 Port 10GbE 8x UTA2 (16 Gbps FC/10GbE) für Host Connectivity 4x StorageGrid SG5712, 12x 4TB NL-SAS	ONTAP 9.5, 2-Node Switchless Cluster, Flash Bundle, iSCSI und NFSv3, UTA2 als 10GbE konfiguriert, weitere Anpassungen gemäß Technical Report TR-4382 Abschnitt 5.1 https://www.netapp.com/us/media/tr-4382.pdf
Server-Betriebssystem	CentOS Linux Release 7.5.1804 (Core)	
Java	OpenJDK Runtime Environment	Build 1.8.0_161-b14
Hadoop Distribution	Hortonworks 2.6.5.0-292 Rollout mittels Apache Ambari 2.6.2.0	2 Name Nodes, 3 Data Nodes
TeraSort Benchmark	TPCx-HS_Kit_v2.0.3	
In-Place Analytics Module	hadoop-nfs-connector-3.0.0, NetApp_Ambari_Plugin_1.0.1	
Monitoring und Dokumentation	NABOX, NetApp SystemManager, ONTAP CLI, Splunk	

**IDEALER WERT FÜR MAP-
UND REDUCE-PROZESSE:
15,14 GB**

2.5 KONFIGURATION VON JAVA UND HADOOP

Die Installation und Konfiguration von Hortonworks wurde mittels Ambari bewerkstelligt. Bis auf eine Anpassung blieben die Standardeinstellungen erhalten. In Ambari wurde im Abschnitt MapReduce2 der Speicher für die Map- und Reduce-Prozesse auf 15,14 GB festgelegt. Dieser Wert ist bei der eingesetzten Infrastruktur ideal.

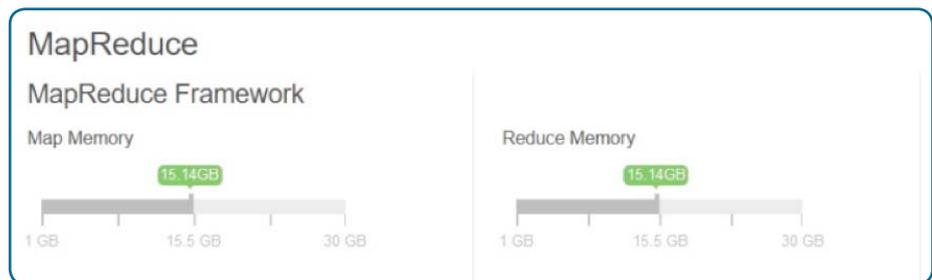


Abbildung 9: Einstellungen MapReduce

2.6 KONFIGURATION DES NETAPP IN-PLACE ANALYTICS MODULES

NetApp stellt das Modul und ein zugehöriges Ambari-Plugin zur grafischen Administration kostenlos auf seiner Support-Webseite support.netapp.com zum Download zur Verfügung. Sie finden beide im Bereich „Tools“ in der „Toolchest“ unter „Administration“. Die Installation des Modules wird im „NetApp In-Place Analytics Module Installation Guide“ beschrieben, der im Download-Archiv enthalten ist. Die Installation des Moduls kann entweder über CLI erfolgen oder über die grafische Ambari-Oberfläche im Browser. Für die Browser-Installation ist im ersten Schritt das Ambari-Plugin zu installieren. Die weitere Vorgehensweise ist im „NetApp Ambari UI Service Plugin Installation Guide“ detailliert dargestellt. Beide Anleitungen erläutern ausführlich die Konfigurationsparameter für das Analytics Module.

**PLUGINS STEHEN ZUM
DOWNLOAD BEREIT**

Download: NetApp In-Place Analytics Module			
Platform: Data ONTAP			
Task	Type	Description	Download
Administration	Client Tool	Ambari UI plugin for In-Place Analytics Module for Hadoop 1.0.1	NetApp_Ambari_Plugin_1.0.1.zip (1.82 MB)
Administration	Client Tool	NFS Connector for Hadoop Software Image	hadoop-nfs-connector-3.0.0.zip (779.09 KB)

Abbildung 10: Download NetApp In-Place Analytics Module

JSON FILE AUF BASIS EINER FLEXGROUP

Alle Konfigurationsinformationen über CLI oder Ambari werden zentral in der JSON-Datei `nfsmapping.json` abgelegt. Die abgebildeten Einstellungen bis auf die Pfadangaben wurden für alle Benchmarks verwendet. Bei der Konfiguration der Pfade wurden für den Test auf FlexVolumes und FlexGroups unterschiedliche Einträge verwendet. Da das Analytics-Modul intelligente Optimierungen für NetApp ONTAP Storage verwendet, sind die Parameter vom Aufbau der Infrastruktur abhängig und vom Einsatzbereich gemäß der Dokumentation einzustellen.

Anbei exemplarisch das JSON File auf Basis einer FlexGroup:

```
[root@dn1 conf]# cat /etc/NetAppNFSCconnector/conf/nfs-mapping.json
{
  „spaces“: [
    {
      „endpoints“: [
        {
          „exportPath“: „/fg_hadoop_both“,
          „hosts“: [
            „nfs://192.168.55.115:2049/“,
            „nfs://192.168.55.116:2049/“,
            „nfs://192.168.55.117:2049/“,
            „nfs://192.168.55.118:2049/“,
            „nfs://192.168.55.119:2049/“,
            „nfs://192.168.55.120:2049/“,
            „nfs://192.168.55.121:2049/“,
            „nfs://192.168.55.122:2049/“,
            „nfs://192.168.55.123:2049/“,
            „nfs://192.168.55.124:2049/“,
            „nfs://192.168.55.125:2049/“,
            „nfs://192.168.55.126:2049/“,
            „nfs://192.168.55.126:2049/“,
            „nfs://192.168.55.127:2049/“,
            „nfs://192.168.55.128:2049/“,
            „nfs://192.168.55.129:2049/“,
            „nfs://192.168.55.130:2049/“
          ],
          „path“: „/fg_hadoop_both_alias“
        }
      ],
      „name“: „space1“,
      „options“: {
        „nfsAuthScheme“: „AUTH_SYS“,
        „nfsExportPath“: „/fg_hadoop_both“,
        „nfsGid“: 0,
        „nfsGroupname“: „root“,
        „nfsMountPort“: -1,
      }
    }
  ]
}
```

```
    „nfsPort“: 2049,  
    „nfsReadMaxPoolThreads“: 128,  
    „nfsReadMinPoolThreads“: 32,  
    „nfsReadSizeBits“: 20,  
    „nfsRpcbindPort“: 111,  
    „nfsSplitSizeBits“: 30,  
    „nfsUid“: 0,  
    „nfsUserCacheTimeout“: 15,  
    „nfsUsername“: „root“,  
    „nfsWriteMaxPoolThreads“: 128,  
    „nfsWriteMinPoolThreads“: 32,  
    „nfsWriteSizeBits“: 20  
  },  
  „uri“: „nfs://192.168.55.115:2049/“  
}  
]  
}
```

Um das Verständnis der abgebildeten Konfiguration zu erleichtern, folgt eine Übersicht der verfügbaren Parameter und ihrer Funktionen.

Details zur Konfiguration sind im „NetApp In-Place Analytics Module Installation Guide“, der dem Module beiliegt, nachzulesen. Die Vorgehensweisen wurden dem NetApp Technical Report TR-4382 entnommen: <https://www.netapp.com/us/media/tr-4382.pdf>

PARAMETER NIAM JSON 1

Parameter NIAM JSON 1 Auszug Technical Report TR-4382

Parameter	Standardwert	Beispielwert	Beschreibung
Globale Optionen			
Name	N/A	Hadoop cluster	Descriptive name for the NFS storage (SVM name) that you are connecting from the Hadoop servers.
Uri	N/A	nfs://192.168.120.1:2049/	Primary IP of the NetApp ONTAP NFS server.
NFS Configuration			
nfsExportPath	/	/hadoop	The default export path of your NFS storage.
nfsReadSizeBits	16	20	The size (bits) of the NFS read request. This value is auto- negotiated with the NFS server.
nfsWriteSizeBits	16	20	The size (bits) of the NFS write requests. This value is auto- negotiated with the NFS server.
nfsSplitSizeBits	20	28	The size of the input split used.
nfsAuthScheme	AUTH_NONE	AUTH_SYS	Authentication scheme to use with the NFS request. Only AUTH_SYS and AUTH_NONE are currently supported.
nfsPort	2049	2049	Port used for NFS connections. This parameter is used if the PortMapper service is not running.
nfsMountPort	-1	12345	Port used for mount in case the PortMapper is not running.
nfsRpcbindPort	111	111	Port used for the PortMapper connections.
nfsUserCache-Time out	1	5	Timeout (in minutes) after which the user and group information cache is refreshed. It's an optional field and the default value is 1 minute.
nfsUsername	root	root	User name used to mount and access data (for AUTH_SYS). It's only required for the 2.0 version.
nfsGroupname	root	root	Group name used to mount and access data.

PARAMETER NIAM JSON 2

Parameter NIAM JSON 2

Parameter	Standardwert	Beispielwert	Beschreibung
			AUTH_SYS). It's only required for the 2.0 version.
nfsUid	0	0	User ID used to mount and access data (for AUTH_SYS). It's only required for the 2.0 version.
nfsGid	0	0	Group ID used to mount and access data (for AUTH_SYS). It's only required for the 2.0 version.
"nfsUserCache-Time out":30, „nfsReadMin-PoolTh reads“:4, „nfsRead-MaxPoolTh reads“:128, „nfsWriteMin-PoolTh reads“:4, „nfsWrite-MaxPoolTh reads“:128			This advanced parameter does not need to be included in the nfs-mapping.json, but you can update it if needed.
Endpoint Configuration (for each item in list)			
hosts	NA	[„nfs://192.168.1.1:2049/“, „nfs://192.168.1.2:2049/“]	An array of data LIFs for a given NFS path specified in exportPath.
exportPath	/	/nfs_volume_name	The absolute export path of the NFS volume. NetApp recommends making this path name either the same as the volume name or the same as the subfolders of the nfsExportPath value. Make sure the exported volume or folder exists before using it in the NetApp In-Place Analytics Module configuration.
path	/	/nfs_volume_name	The volume name in the NFS SVM where the data resides for analytics. This name is an alias for the exportPath. The value of path should not be /.

Hinweise:

- nfsExportPath und exportPath sollten nicht gleich lauten.
- nfsExportPath sollte mindestens eine Verzeichnisebene höher liegen als exportPath. Falls dies nicht möglich ist, bei “/” belassen.

**TERASORT IST IN JEDER
HADOOP INSTALLATION
ENTHALTEN**

3. BENCHMARKING MIT DER TERASORT SUITE

Bei dem verwendeten Benchmark handelt es sich um den TPCx-HS v2. Dieser ist ein standardisierter Benchmark des bekannten TeraSort. TeraSort mit seinen 3 Teilen (TeraGen, TeraSort, TeraValidate) ist standardmäßig in jeder Hadoop Installation enthalten. TPCx-HS konsolidiert diese drei Schritte zu einem Vorgang und wiederholt die Ausführung des Benchmarks nach Nutzeranforderung. Es werden Daten einer bestimmten Größe generiert, im HDFS abgespeichert, sortiert und final validiert.

3.1 AUFBAU UND VORBEREITUNG DER BENCHMARKS

Wie zuvor erwähnt, besteht der TPCx-HS v2 Benchmark aus drei elementaren Teilen. Zuerst werden die Testdaten generiert und in ein definiertes Zielverzeichnis geladen. Die Größe des gesamten Datensatzes ist ebenfalls definierbar. Im zweiten Schritt werden die Daten mithilfe des MapReduce Algorithmus sortiert. Durch die Nutzung von YARN stehen eine definierbare Anzahl von Mappern und Reducern zur Verfügung. Abschließend wird die Validierung der Ergebnisse vollzogen. Hierbei wird die Richtigkeit der sortierten Daten geprüft. Der gesamte beschriebene Vorgang wird einmal wiederholt.

Der Benchmark TPCx-HS v2 hat einige Konfigurationen, die zu beachten sind. In der Konfigurationsdatei Benchmark_Parameters.sh werden Parameter gesetzt, mit denen der Benchmark ausgeführt werden soll. Im Verlauf des Tests stellten sich die folgenden Parameter als für das Testsetup optimal heraus.

Benchmarkparameter

Parameter	Wert
HADOOP_USER	root
HDFS_USER	root
HDFS_BENCHMARK_DIR	nfs://192.168.55.115:2049/<flexgroup/volume>/tpc
SLEEP_BETWEEN_RUNS	60
NUM_MAPS	64
NUM_REDUCERS	64

3.2 DOKUMENTATION DER ERGEBNISSE

Die Ergebnisse werden mit der Metrik HSph@SF des TPCx-HS v2 Benchmark dargestellt. Je höher der Wert, desto besser die Performance. Über folgende Formel erfolgt die Berechnung:

$$HSph@SF = \frac{SF}{\frac{T}{3600}}$$

SF bedeutet der Skalierungsfaktor. (1 entspricht 1 TB, 3 entspricht 3 TB usw.)
T ist die Gesamtzeit in Sekunden.

Diese Metrik wird nach jedem Durchlauf auf der Konsole ausgegeben.

Weiterhin werden in den Ergebnissen die Laufzeit des Benchmarks, die CPU-Auslastung des primären Speicher-Systems sowie dessen maximaler Durchsatz in den Test-Einstellungen dargestellt.

100 GB TESTDATEN WURDEN VERWENDET

3.3 DURCHFÜHRUNG DER BENCHMARKS UND TESTERGEBNISSE

Die Benchmarkdurchläufe wurden mit 100 GB Testdaten (erzeugt von TeraGen) durchgeführt. Im ersten Schritt wurde der Benchmark auf einem NetApp FlexVolume, iSCSI LUNs und NetApp FlexGroup ausgeführt, um die am besten geeignete Speichertechnologie zu definieren. Im Verlauf des PoC wurde das YARN UserCache Verzeichnis, welches das temporäre Arbeitsverzeichnis des Benchmarks darstellt, von im HDFS befindlichen iSCSI LUNs auf via NFS bereitgestellte Flex Volumes ausgelagert. Wie bereits erwähnt wurden die ersten Tests auf folgenden Speicherbereichen durchgeführt:

NFS:

- NetApp FlexVolume – genutzt via NetApp In-Place Analytics Module für Testdaten und später direkt via NFS als YARN UserCache, innerhalb eines NetApp Nodes
- NetApp FlexGroup – genutzt via NetApp In-Place Analytics Module für Testdaten, über beide NetApp Nodes verteilt

iSCSI LUN:

- genutzt via HDFS für Testdaten und YARN UserCache, über beide NetApp Nodes verteilt

Die Ergebnisse des ersten Tests:

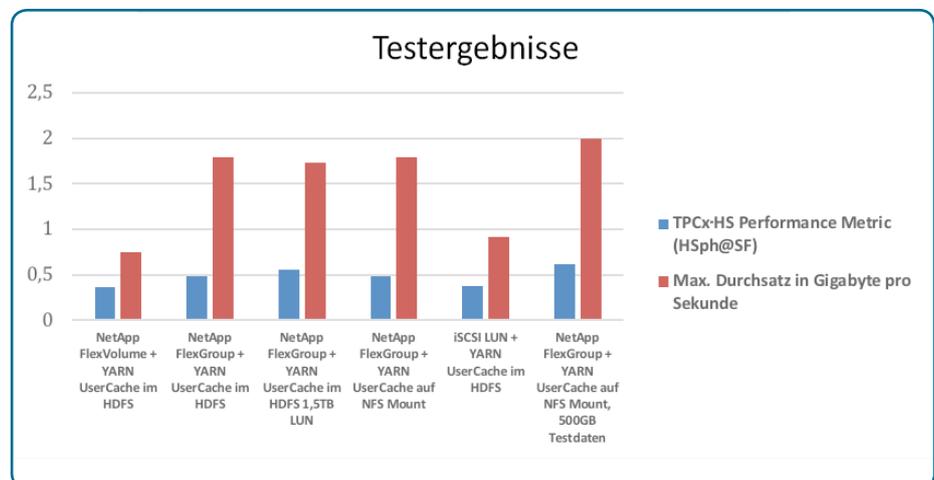


Abbildung 10: Testergebnisse

**FÜR DIE DATEN-
VERARBEITUNG EIGNET
SICH EINE FLEXGROUP
(VIA NFS)**

**IDEAL IST AUCH NFS FÜR
DEN YARN USERCACHE**

Detaillierte Ergebnis-Übersicht des ersten Tests:

Beschreibung	HSph@SF	Max. Durchsatz in GB/s
NetApp FlexVolume + YARN UserCache im HDFS	0,3605	0,75
NetApp FlexGroup + YARN UserCache im HDFS	0,48665	1,79
NetApp FlexGroup + YARN UserCache im HDFS mit 1,5 TB LUN	0,5515	1,73
NetApp FlexGroup + YARN UserCache auf NFS Mount	0,4839	1,79
iSCSI LUN + YARN UserCache im HDFS	0,3731	0,92
NetApp FlexGroup + YARN UserCache auf NFS Mount, 500GB Testdaten	0,6095	1,99

Die Ergebnisse zeigen auf, dass unter Berücksichtigung von Performance und Speicherplatzverbrauch die Wahl einer via NFS angebotenen FlexGroup und die Verwendung von NFS für den YARN UserCache die am besten geeignete Variante für die Daten-Verarbeitung ist.

Im Folgenden wurde der Test auf weitere NetApp Speichersysteme ausgeweitet.

Für die Auswertung des Benchmarks werden folgende Aspekte herangezogen:

- Wert HSph@SF
- Laufzeit
- Maximal erreichter Durchsatz
- CPU-Auslastung NetApp System

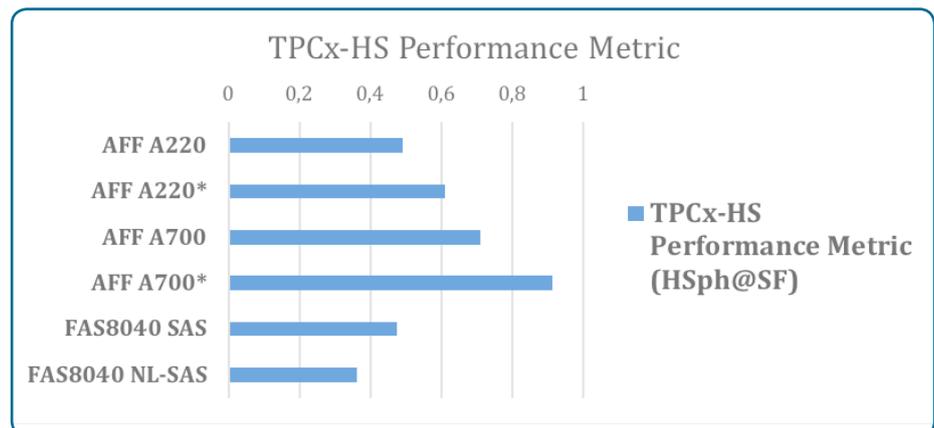


Abbildung 11: Testergebnisse TPCx-HS Performance Metric
*500 GB Testdaten

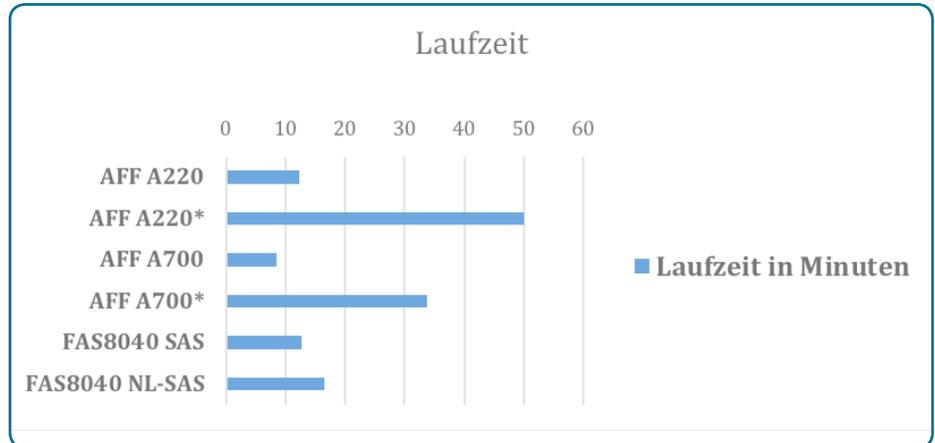


Abbildung 12: Testergebnisse Laufzeit

*500 GB Testdaten

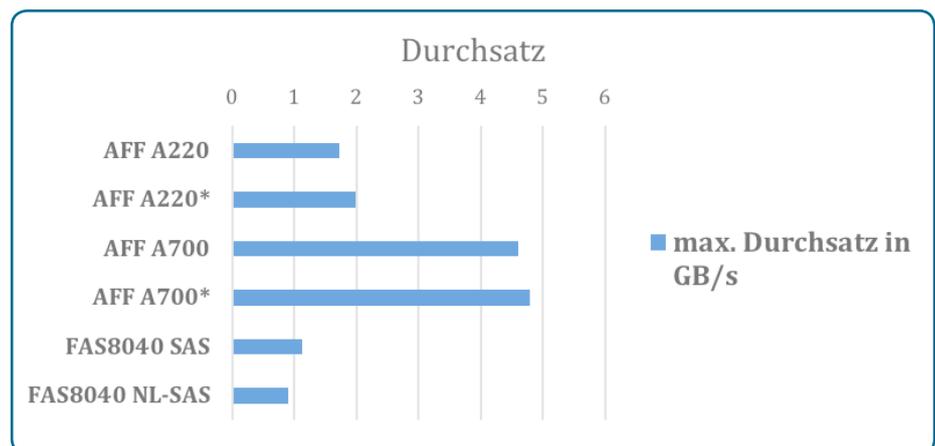


Abbildung 13: Testergebnisse max. Durchsatz

*500 GB Testdaten

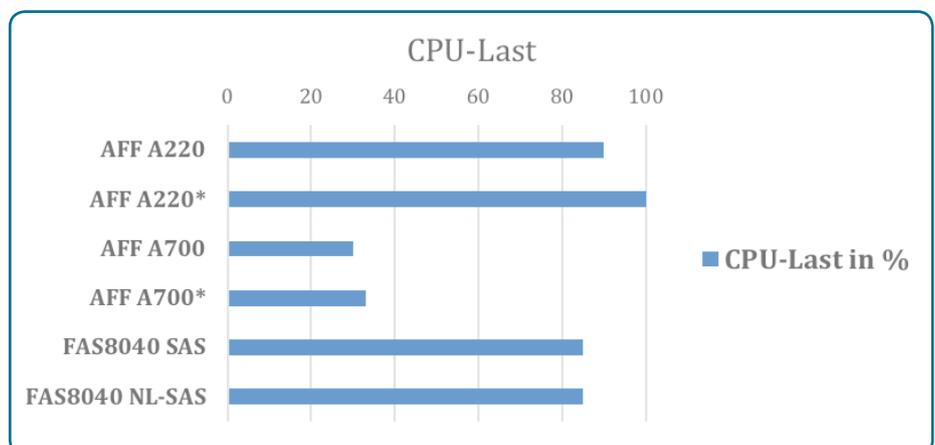


Abbildung 14: Testergebnisse CPU-Last

*500 GB Testdaten

**JE MEHR PARALLEL
GENUTZTE RESSOURCEN,
DESTO BESSER DIE
ERGEBNISSE**

Modell	HSph@SF	Laufzeit in Minuten	max. Durchsatz in GB/s	CPU-Last in %
AFF A220	0,4839	12,4	1,79	90
AFF A220*	0,6095	49,9	1,99	100
AFF A700	0,7089	8,5	4,6	30
AFF A700*	0,9129	33,7	4,8	33
FAS8040 SAS**	0,4761	12,6	1,13	85
FAS8040 NL-SAS**	0,3625	16,6	0,91	85

Abbildung 15: Gesamtübersicht Testergebnisse

*500 GB Testdaten

**Test auf einem NetApp Node

3.4 PERFORMANCE

Die Bewertung aufgrund des Durchsatzes und des Benchmarkwertes ergibt ein klares Bild. Je mehr Ressourcen seitens des Speichersystems parallel genutzt werden konnten, desto besser waren die erreichten Benchmarkergebnisse.

Ein via NFS und das NetApp In-Place Analytics Module bereitgestelltes FlexVolume nutzt die Ressourcen eines NetApp Controllers. Auch wenn die Workload über drei Hadoop Worker Nodes verteilt wird, ist die maximale Performance hierbei durch die Ressourcen des im Test verwendeten NetApp AFF A220 Controllers begrenzt. Bei der Verwendung eines FlexVolumes ist die Parallelisierung über beide Storage Controller also nicht gegeben. Mit einem Wert von 0,75 GB/s stellt die Verwendung eines NetApp FlexVolumes hiermit die langsamste Variante dar.

Als zweite Variante bekam jeder Hadoop Worker Node eine iSCSI LUN, die der Data Node mit HDFS für die Benchmarks verwendete. Dies entspricht der Verwendung lokaler Datenträger mittels HDFS. Die LUNs wurden auf beide NetApp AFF A220 Controller verteilt. Die Testergebnisse sind etwas besser: ca. 0,92 GB/s. Da in einem HDFS-Aufbau mit LUNs der Replikationsfaktor von 3 genutzt wird, liegt die Vermutung nahe, dass hierbei das schwächste Glied in der Infrastruktur die Maximalgeschwindigkeit vorgibt. Die Limitierung liegt hierbei auf der Performance eines einzelnen NetApp AFF A220 Nodes. In einem fortführenden Test mit einer AFF A700 wurde diese Vermutung bestätigt. Die Verwendung von iSCSI LUNs ist dennoch nicht zu empfehlen, da der Performance Overhead durch die CPU basierten iSCSI-Berechnungen der Hadoop Worker Nodes im Vergleich zur Verwendung von NFS das Gesamtergebnis negativ beeinflusst. Ein weiterer Punkt stellt der erhöhte Speicherbedarf dar, da für jeden Hadoop Worker Node eine dedizierte LUN bereitgestellt werden muss.

Der dritte Test in diesem Setup ist die Nutzung einer NetApp FlexGroup via NFS und des NetApp In-Place Analytics Module. Durch dieses Setup werden die Ressourcen eines NetApp Clusters parallel über alle Controller genutzt. Weiterhin wird der Replikationsfaktor RF=3 transparent für Hortonworks auf RF=1 im Storage reduziert, sodass die Daten nur einmal in der FlexGroup abgelegt werden. Hierdurch werden die für eine klassische HDFS-Implementierung genutzten Ressourcen frei. Auch die Konsistenz des HDFS ist aus Sicht von Hortonworks

DER EINSATZ VON NL-SAS-FESTPLATTEN REDUZIERT DIE PERFORMANCE UM CA. 25 %

stets gewährleistet. Die Tests unter Verwendung der NetApp FlexGroup auf einer NetApp AFF A220 erreichten im Maximum ein Ergebnis von knapp 2 GB/s. Da der Testsetup aus zwei NetApp AFF A220 Controllern bestand, stellt dieses Ergebnis auf Basis der im Test ermittelten Begrenzung von 1 GB/s für einen Controller das Maximum des Testaufbaus dar. Eine Anpassung der Konfigurationen auf eine noch höhere Parallelisierung oder Erhöhung der verfügbaren Ports aufseiten des Speichersystems ermöglichte keine besseren Werte in diesem Testszenario. Der limitierende Faktor ist die CPU in den NetApp Nodes.

Im Verlauf weiterer Tests mit einer NetApp AFF A700 und einer FAS8040 auf Basis von SAS- und NL-SAS-Festplatten konnten weitere Erkenntnisse bezüglich Performance gewonnen werden. Das Benchmarking der FAS8040 mit SAS-Festplatten zeigte unter Nichtberücksichtigung von Latenzen die gleichen Ergebnisse in Durchsatz und TPCx Performance Metric wie die AFF A220. Auch hier stellte sich die CPU als limitierender Faktor heraus. Dies bedeutet, dass nur bei der Notwendigkeit geringer Latenzen die Wahl von SSDs im Einstiegsbereich der NetApp Systeme sinnvoll ist und SAS-Festplatten eine valide Möglichkeit darstellen. Der Einsatz von NL-SAS-Festplatten reduziert dagegen die Performance um ca. 25 % und sollte nur bei großen Datenmengen, mit Blick auf günstigen Speicher, in Betracht gezogen werden.

Eine interessante Alternative zum Aufbau einer NL-SAS-basierten Speicherlösung stellt die Verwendung von NetApp StorageGrid dar, dem Objektspeicher von NetApp. Dieser kann sowohl nativ in Hadoop als Massenspeicher eingebunden werden sowie als transparentes Tiering für SSD-Speicherbereiche auf NetApp ONTAP Systemen verwendet werden (NetApp FabricPool). Die Performance wurde unter Verwendung des gleichen MapReduce-Algorithmus der NetApp AFF A700 getestet. Die Laufzeit des Tests war bei der Verwendung von NetApp FabricPool ca. 30 % höher im Vergleich zur Konfiguration der NetApp AFF A700 ohne FabricPool.

MIT NETAPP FABRICPOOL WAR DIE PERFORMANCE CA. 30 % HÖHER

Die Tests auf Grundlage einer NetApp AFF A700 waren unter Verwendung von FlexGroups mit einem maximalen Durchsatz von bis zu 4,8 GB/s und einer Reduzierung der Laufzeiten um ca. ein Drittel die schnellsten. Das Limit der NetApp AFF A700 wurde nicht erreicht. Der Engpass entstand durch die 10-Gb-Netzwerkinfrastruktur und die verwendeten CPUs der Server des Testaufbaus. Sowohl die CPUs der drei Hadoop Worker Nodes als auch die Netzwerkports waren komplett ausgelastet. Auf Serverseite wurde die maximale Bandbreite von 6x 10GbE voll ausgelastet.

Um die Performance noch weiter zu erhöhen, gilt es, die Reduktion des globalen HDFS-Replikationsfaktors auf RF=1 zu untersuchen. Ein weiterer Ansatz stellt die Erhöhung der Server-Bandbreite durch die Verwendung von Ethernet Adapter >10 Gb dar. Auf Basis dieser Erkenntnisse empfiehlt es sich, bei NetApp Systemen ab dem Modell NetApp AFF A300 eine Netzwerkinfrastruktur mit Geschwindigkeiten >10GbE zu nutzen.

***DIE VERWENDETEN
SPEICHERSYSTEME
NUTZEN ALLE EFFIZIENZ-
MECHANISMEN AUS***

***HADOOP SERVICES SIND
AUCH IN DER PUBLIC
CLOUD MÖGLICH***

3.5 STORAGE EFFICIENCY UND REPLIKATION

Die verwendeten Speichersysteme (Basis NetApp ONTAP 9) verwenden standardmäßig abhängig von dem AFF- oder FAS-Modell alle in NetApp ONTAP verfügbaren Effizienz-Mechanismen.

Die Effektivität ist bei allen verwendeten Szenarien im Speicherbereich aus Sicht des Storage-Systems identisch. Der Vorteil bei der NFS-Implementierung ist die direkte Verfügbarkeit des eingesparten Platzes im Host-System, da der Speicherplatz auf FlexVolume- und FlexGroup-Ebene freigegeben wird. Diese Ebene ist gleichzeitig das verfügbare Dateisystem auf dem Linux Host. Weiterhin werden bei der Nutzung von NFS im Vergleich zu iSCSI die Daten nur einmal anstatt dreimal auf das Speichersystem geschrieben.

Die Verwendung der Snapshot- und FlexClone-Technologie von NetApp ermöglicht es, die Datenbestände innerhalb von Sekunden unabhängig von Größe und Platzbedarf erneut bereitzustellen. Dies wurde ebenfalls untersucht, indem die Datenbestände (read-only) durch die Verwendung eines Snapshots (auch writeable) mittels eines FlexClone an Hadoop Worker Nodes angebunden wurden. Auf diese Weise ist sowohl eine einfache wie auch eine mehrfache Bereitstellung der Daten für Dev und Test ohne Veränderung der ursprünglichen Daten möglich.

Die Verwendung der NetApp SnapMirror-Technologie wurde ebenfalls betrachtet. Nachdem die Datenbestände auf ein anderes NetApp System repliziert wurden, war es mittels des NetApp In-Place Analytics Modules möglich, direkt auf diese Daten aus dem Hadoop Cluster zuzugreifen.

Dieses Szenario ist beliebig auf Hadoop Cluster unterschiedlichster Standorte adaptierbar. Aufgrund der Verfügbarkeit von NetApp ONTAP basierten Systemen in der Public Cloud, wie z. B. Amazon WebServices, Microsoft Azure und anderer, ist auch die Einbindung von Hadoop Diensten in der Public Cloud möglich.

ALLE VORTEILE AUF EINEN BLICK

4. DIE VORTEILE VON HADOOP CLUSTERN AUF FLASH-BASIERTEM SPEICHER

Für die getesteten Systeme von NetApp ergeben sich folgende Erkenntnisse:

- Ausgezeichnete Performance-Ergebnisse bereits mit dem NetApp All-Flash-Einstiegsmodell A220.
- Integrierte, leicht zu konfigurierende Datensicherung für Backup und Disaster-schutz (u. a. Performance-neutrale Snapshots, Disk-to-Disk Backup, Replikation).
- Shared Storage mit NetApp über NFS gestattet unabhängige Skalierung von Speicherkapazität und -Performance, ohne gleichzeitig Compute-Ressourcen hinzufügen zu müssen.
- Zusätzliche Hadoop Worker Nodes lassen sich jederzeit problemlos per NFS mit dem Shared Storage verbinden. Der Aufbau eines zentralen „Data Lake“ und darauf aufbauend die Realisierung einer durchgehenden „Data-Pipeline“ im Unternehmen wird erleichtert.
- Einfache Bereitstellung eines Shared Storage auf Basis von NFS für die Index files (files und edit) der Primary und Secondary Name Nodes – die Kapazität kann jederzeit transparent angepasst werden.
- Weniger Ressourcenbedarf an CPU und RAM in den Hadoop Worker Nodes durch Verkleinerung der Replizierung im Hadoop Cluster auf RF=1. Die Datensicherheit wird durch die integrierte RAID Protection von NetApp ONTAP gewährleistet.
- Unter Berücksichtigung des Anwendungsfalls ist die Nutzung von SAS-Festplatten im Einstiegsbereich der NetApp Systeme eine valide Option.
- Bei der Verwendung von Systemen > NetApp AFF A220 ist eine Netzwerkinfrastruktur ab 25 GbE empfehlenswert.
- Bei der Einbindung von NetApp StorageGrid als FabricPool für SSD reduziert sich die Performance um ca. 30 % bei Zugriff auf „kalte“ im StorageGrid befindliche Daten.

Folgende Erkenntnisse wurden für den Betrieb und die Administration eines Hadoop Systems gewonnen:

- Compute-Ressourcen (CPU + RAM) können dank zentralem Shared Storage höher ausgelastet werden, weil keine lokalen Ressourcen für die Speicherverwaltung notwendig sind. Die bekannte Daumenregel „Storageverwaltung kostet 1 Core pro installierte Platte/SSD“ ist damit obsolet.
- Anhand der Messwerte ist abzusehen, dass die Performance der Hadoop Worker Nodes durch mehr RAM deutlich erhöht werden kann. Die RAM-Ausstattung sollte 512 GB oder mehr pro Host betragen.
- Höhere Performance durch Verwendung von NetApp FlexGroups anstelle von iSCSI LUNs pro Worker-Node.
- Direkte und transparente Bereitstellung von Daten im Hadoop Cluster mittels des NetApp In-Place Analytics Modules.

- Einfacher Zugriff auf durch NetApp SnapMirror replizierte Datenbestände an anderen Standorten, bei anderen Service Providern oder in der Public Cloud.
- Sekundenschnelle Bereitstellung auch großer Datenbestände mittels NetApp Snapshots und NetApp FlexClone für Dev- und Test-Szenarien.
- Abbildung einer Mandantentrennung mittels Storage Virtual Machines für unterschiedliche Hadoop Cluster auf einer Shared-Plattform.
- Niedrigere Kosten für lokalen DAS-Speicher in den Hadoop Worker Nodes, da nur Boot Drives nötig sind. Optional können hierzu auch NetApp Boot LUNs genutzt werden (über iSCSI oder FC). Hierdurch ist eine weitere Kostenreduzierung durch Wegfall des DAS-Speichers möglich. Weiterhin kann die notwendige Anlaufzeit bei Server-Ausfällen und Compute-Erweiterungen z. B. mittels NetApp FlexClone reduziert werden. Dieser Aspekt ist in einem weiterführenden Test zu validieren.
- Validierte „Building Blocks“ in verschiedenen Größenabstufungen erleichtern die Dimensionierung im Vorfeld von Projekten, bieten gesicherte Erfahrungswerte für Performance und Optimierungspotenziale und erleichtern den Aufbau von bewährten Methoden im gesamten Team.
- Die Auslagerung des YARN UserCaches auf via NFS bereitgestellte NetApp FlexVolumes reduziert den Speicherbedarf und die Performance-Anforderungen an die Name Nodes. Klassisch wird der YARN UserCache auf dem HDFS-Bereich abgelegt und somit ebenfalls dreifach repliziert.
- Die Größe des YARN UserCaches ist der Quotient aus den aktiven Daten und der Anzahl der Hadoop Worker Nodes.
- Bei der Verwendung von LUNs unter HDFS hat sich die Verwendung von nur einer LUN pro Hadoop Worker Node als ideal erwiesen. Auf diese Weise wird ein unnötiges Rebalancing innerhalb der Worker Nodes vermieden und somit die Performance im Cluster erhöht.
- Die Einbindung von NetApp StorageGrid als FabricPool für SSD-basierte Speicherbereiche kombiniert die Möglichkeiten eines Data Lake und einer Echtzeit-Analyse, die sich für den Hadoop Cluster transparent darstellt. Gleichzeitig kann das NetApp StorageGrid auch nativ als Objektspeicher in Hadoop angebunden werden.

Folgendes Verbesserungspotenzial wurde für NetApp In-Place Analytics Module identifiziert:

- Das vom NetApp In-Place Analytics Module eingebundene NFS-File-System wird aktuell nicht in der Ambari HDFS Summary angezeigt. Aufgrund dessen ist die Überwachung des Speicherplatzes nur außerhalb von Ambari möglich (z. B. Netapp OnCommand Unified Manager, Nagios etc.). Dieser Umstand wurde als „Request for Engineering“ an NetApp weitergeleitet.
- Zukünftige Versionen des NetApp In-Place Analytics Modules könnten möglicherweise gesteigerte Lese-Performance durch Kenntnis der Datenlokalität in der FlexGroup erzielen. Auch dies wird bereits als „Request for Engineering“ innerhalb von NetApp weiter verfolgt.

EINGEREICHTE REQUESTS FOR ENGINEERING

DIE ERWARTUNGEN WURDEN ÜBERTROFFEN

VIELVERSPRECHENDE AUSSICHTEN FÜR DIE ZUKUNFT

SIE HABEN NOCH FRAGEN ZUM THEMA BIG DATA ANALYSEN?

Ich berate Sie gerne:
 Marc Riedel,
 Partnermanager NetApp,
 Tel: +49 6122 / 536-0,
 marc.riedel@sva.de

5. FAZIT

Zusammenfassend kann man sagen, dass die untersuchte konvergente Infrastrukturplattform auf Basis von NetApp und Fujitsu die Erwartungen übertroffen hat. Gleich in mehrerer Hinsicht weist der integrierte und validierte Stack aus Server, Netzwerk und Storage bessere Eigenschaften im Vergleich zu herkömmlichen Implementierungen mit X86-Servern und lokalem Storage (SSD) auf. Und zwar in den Bereichen:

- Performance
- Effizienz
- Flexibilität
- Skalierbarkeit
- Implementierungsaufwand

Anhand der Beobachtungen und Testergebnisse dieses PoCs ergeben sich vier attraktive Anwendungsfälle für die konvergente Infrastrukturplattform in zukünftigen Analytics-Projekten:

- Default-Plattform für Analytics-PoCs: klein, einfach, zuverlässig, flexibel, skalierbar – und schnell
- Ideal für Kunden ohne verfügbare Ressourcen:
Einfache Bestellung, bewährte Vorgehensweisen, Erfahrungen einfach übertragbar.

Für NetApp Bestandskunden gestaltet sich die Nutzung von auf NetApp NFS befindlichen Daten mittels des Einsatzes des NetApp In-Place Analytics Modules sehr einfach, da diese direkt in den Hadoop Cluster eingebunden werden können.

- Aufgrund der Flexibilität und einfachen Skalierbarkeit ideal für Hadoop Projekte mit nicht kalkulierbarem Speicherbedarf, z. B. Predictive Maintenance, Social-Media-Anbindung etc.
- Bereits die NetApp All-Flash-Einstiegsmaschine AFF A220 ist als Dev-Umgebung für DL (Deep Learning) geeignet. Dank vertikaler und horizontaler Skalierbarkeit der NetApp Speichersysteme ist sie auch hohen und höchsten DL-Anforderungen gewachsen.
 Siehe auch: „NetApp Verified Architecture: NetApp ONTAP AI“
 (unter: <https://www.netapp.com/de/media/nva-1121-design.pdf>)
 oder „Cisco FlexPod AI“

6. AUSBLICK

Einen elementaren Bestandteil des NetApp Produktportfolios stellt die NetApp Data Fabric dar. Mit dieser ist die hybride Nutzung der Public Cloud und eines On-Premises-Rechenzentrums, Service Providers etc. möglich.

DIE NETAPP DATA FABRIC

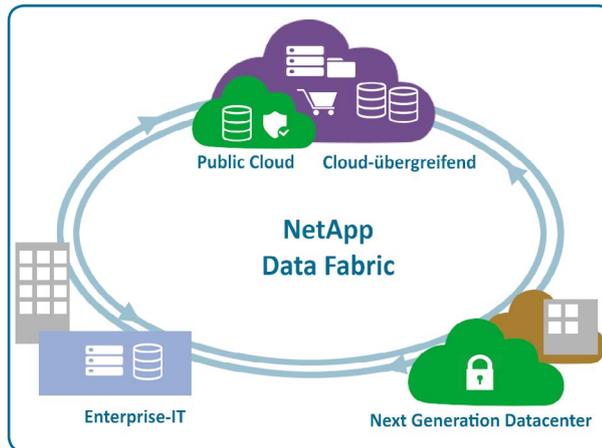


Abbildung 16: NetApp Data Fabric

Adaptiert auf den Anwendungsbereich Hadoop kann bei Bedarf ein im Folgenden dargestelltes Szenario umgesetzt werden:

ANSÄTZE FÜR NEUE INFRASTRUKTUR-MODELLE

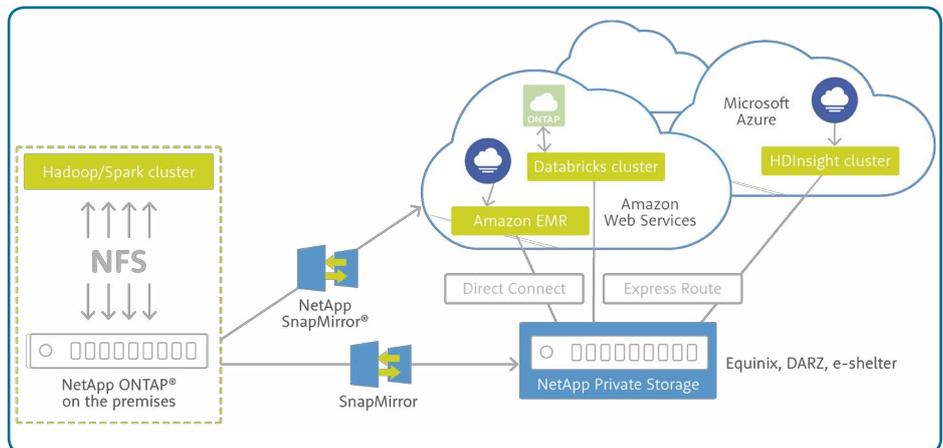


Abbildung 17: Hadoop@ NetApp Data Fabric

Auch die Sammlung von IoT-Daten in der Cloud unter Verwendung von NetApp Cloud Volumes ONTAP oder die Nutzung der hohen Performance der Public Cloud in Verbindung mit NetApp Cloud Volume Services stellen valide mögliche Ansätze dar.

ZUKÜNFTIGE TEST- SZENARIEN UND HERAUS- FORDERUNGEN

Folgende Aspekte sind in einem fortführenden Proof of Concept zu betrachten:

- Einbindung der NetApp Data Fabric auf Basis von Cloud Volume Services und Cloud Volumes ONTAP für Analytics-Aufgaben.
- Test eines leistungsfähigeren NetApp AFF-Systems, um die Eignung für AI-Ansätze, wie Deep Learning und Machine Learning, zu validieren.
- Betrachtung von NetApp HCI als Alternative für Fachabteilungen, um einen einfachen und sicheren Einstieg in bspw. ML zu garantieren. Durch die Verfügbarkeit von QoS im Speicherbereich ist es möglich, eine bestimmte Performance zu garantieren.
- Der Fokus des PoC lag auf der Funktionalität in Zusammenhang mit NetApp ONTAP Speichersystemen und deren infrastrukturellen Möglichkeiten. In der Folge wurden Aspekte bzgl. Security aufgenommen. Diese sind in dem NetApp Technical Report „TR-4744 Secure Hadoop using Apache Ranger with NetApp In-Place Analytics Module“ ausgeführt.

Sie haben Fragen zu Hadoop Clustern auf Flash-basiertem Speicher? Oder möchten von uns zu Big-Data-Analysen beraten werden? Wir unterstützen Sie gerne.

Version	Datum	Details
Version 1.0	November 2018	Erstversion, Auswertung Hortonworks auf Basis konvergente Infrastruktur Fujitsu
Version 1.1	April 2019	Erweiterung um Test mit AFF A700, StorageGrid und FAS8040 auf SAS- und NL-SAS-Basis

8. TEAM UND BETEILIGTE UNTERNEHMEN

Name	Firma
Christian Biermann	SVA System Vertrieb Alexander GmbH
Antonios Genios	SVA System Vertrieb Alexander GmbH
Tobias Nirschl	SVA System Vertrieb Alexander GmbH
Kai Paschke	SVA System Vertrieb Alexander GmbH
Hans Reinhardt	SVA System Vertrieb Alexander GmbH
Nico Seiberth	SVA System Vertrieb Alexander GmbH
Konstantinos Stergakis	SVA System Vertrieb Alexander GmbH
Christopher Voss	SVA System Vertrieb Alexander GmbH
Alexander Wicklein	SVA System Vertrieb Alexander GmbH
Steve Guhr	NetApp Deutschland GmbH
Jonathan Tourtellot	NetApp Deutschland GmbH
Roger Zink	NetApp Deutschland GmbH

SVA (www.sva.de) ist einer der führenden System-Integratoren Deutschlands im Bereich Datacenter-Infrastruktur. Das Unternehmen mit Hauptsitz in Wiesbaden beschäftigt mehr als 980 Mitarbeiter an 18 Standorten in Deutschland. Mehr als 80 Spezialisten erarbeiten auf die Kundensituation zugeschnittene Proofs of Concept, Test- und Migrationsszenarien, Architekturen und Lösungen für Big Data und Analytics. Bei Bedarf unterstützen die IT-Spezialisten von SVA Operational Services auch beim produktiven Betrieb der Big-Data- und Analytics-Infrastruktur. Das Spektrum der Unterstützung reicht von der gelegentlichen telefonischen Hilfe bis zum hochverfügbaren Betrieb ganzer Systemlandschaften.

NetApp (www.netapp.com) ist die Instanz für Datenmanagement in der Hybrid Cloud. Mit einem breiten Portfolio an Hybrid-Cloud-Datenservices und On-Premises-Speicherlösungen, die das Management von Applikationen und Daten über

Cloud- und On-Premises-Umgebungen hinweg vereinfachen, beschleunigt NetApp die digitale Transformation. Gemeinsam mit hochrangigen Partnern wie SVA hilft NetApp Unternehmen weltweit, das volle Potenzial ihrer Daten auszuschöpfen und so Verbindungen zu Kunden zu finden, Innovationen voranzutreiben und Betriebsabläufe zu optimieren.

Fujitsu Central Europe (Deutschland, Österreich und die Schweiz) ist ein führender europäischer Informations- und Telekommunikations-Komplettanbieter (www.fujitsu.de). Das Unternehmen bietet ein umfassendes Portfolio von Technologieprodukten, Lösungen und Dienstleistungen, das von Endgeräten über Rechenzentrumslösungen, Managed und Maintenance Services und Cloud-Lösungen bis hin zum Outsourcing reicht. Die beiden Unternehmen Fujitsu und NetApp verbindet eine mittlerweile 20-jährige Technologie- und Vertriebspartnerschaft. Aus der globalen Zusammenarbeit resultieren wegweisende Produkte, Services und Lösungen.

9. COPYRIGHT UND TRADEMARKS

© SVA GmbH 2019

Alle Marken- und Produktnamen sind Warenzeichen und werden als solche anerkannt.

NetApp, das NetApp Logo, die Gateway-Gestaltung, Data ONTAP, DataFabric, FA-Server, FilerView, FlexClone, FlexVol, MultiStore, NearStore, NetCache, SecureShare, SnapDrive, SnapLock, SnapManager, SnapMirror, SnapMover, SnapRestore, SnapValidator, SnapVault, SyncMirror, VFM und WAFL sind eingetragene Marken und ApplianceWatch, BareMetal, Camera-to-Viewer, Center-to-Edge, ContentDirector, ContentFabric, EdgeFiler, FlexShare, HyperSAN, InfoFabric, NetApp Availability Assurance, NetApp ProTech Expert, NOW, NOW NetApp on the Web, RoboCache, RoboFiler, SecureAdmin, Serving Data by Design, SharedStorage, Smart SAN, SnapCache, SnapCopy, SnapDirector, SnapFilter, SnapMigrator, Snapshot, SnapSuite, SohoCache, SohoFiler, Unbound Cloud, Vfiler, Virtual File Manager und Web Filer sind Marken der NetApp Inc. in den USA und anderen Ländern. Alle anderen Marken oder Produkte sind Marken oder eingetragene Marken der jeweiligen Rechtsinhaber und werden hiermit anerkannt.

10. KONTAKT

Marc Riedel
Partnermanager NetApp

SVA System Vertrieb Alexander GmbH
Borsigstraße 14
65205 Wiesbaden, Germany

Tel: +49 6122 / 536-0
Fax: +49 6122 / 536-399
E-Mail: marc.riedel@sva.de
www.sva.de

Die SVA GmbH ist einer der führenden System-Integratoren Deutschlands im Bereich Datacenter-Infrastruktur und beschäftigt mehr als 1000 Mitarbeiter an 18 Standorten.