



Broschüre

Hadoop der Enterprise-Klasse

Wertschöpfung und Leistungsfähigkeit von Apache Hadoop optimal einsetzen und von einer Skalierbarkeit ohne Kompromisse profitieren



Hewlett Packard
Enterprise



Moderne Unternehmen wollen den Nutzen großer Datenmengen – der Big Data – effizient ermitteln und davon profitieren. Mittlerweile suchen viele Unternehmen deshalb nach agilen und leistungsfähigen Analysesystemen. Apache Hadoop ist ein Software-Framework, das messbare Einsparungen und einen erheblichen Vorteil bei Speicherprozessen und der Datenverarbeitung bringt. Viele Unternehmen haben im Rahmen von Pilotprojekten Hadoop als Datenrepository für einfache Workloads getestet. Hadoop bietet jedoch weit mehr, wenn man die auf der Plattform gespeicherten Daten besser einsetzt, mit ihnen interagiert und so neue aussagekräftige Informationen offenlegt. Es gibt jedoch auch Herausforderungen, durch die dieser Nutzen eingeschränkt wird und Chancen ungenutzt bleiben. Hewlett Packard Enterprise hat hierfür eine Lösung entwickelt, mit der sich diese Herausforderungen bewältigen lassen. Es handelt sich hierbei um ein leistungsfähiges, flexibles Angebot für den Aufbau einer zukunftssicheren, datengesteuerten Basis, die mit Ihren Geschäftsanforderungen wächst.

Die heutigen geschäftlichen Herausforderungen

Daten sind die treibende Kraft für die Idea Economy von heute. Um alle relevanten Daten aber auch nutzen zu können, brauchen Sie neue Antriebskomponenten, die helfen, herausragende Geschäftsergebnisse zu erzielen. In den vergangenen 10 Jahren ist Hadoop zu einer solchen wichtigen Antriebskomponente geworden, die von zahlreichen Entwicklern, Datenwissenschaftlern und IT-Spezialisten gleichermaßen geschätzt wird. Den Visionären hat Hadoop sehr geholfen, jedoch blieb dieses Konzept für die meisten Unternehmen, die sich sehr früh mit Hadoop befassten, bisher größtenteils immer auf wissenschaftliche Experimente, unkomplizierte Workloads und Pilotprojekte beschränkt. Hewlett Packard Enterprise kennt die Herausforderungen und Chancen durch Hadoop und kann auch Ihrer Organisation helfen, Hadoop zu einem unternehmensweiten Meilenstein Ihrer datengesteuerten Plattform zu machen.

60 %
HDFS-Performance¹

100 %
Ihrer relevanten Daten

Bis zu 8 X
schnellere Analysen²

HERAUSFORDERUNGEN FÜR KUNDEN

Über vorhandene Technologien kann das Datenwachstum nicht verarbeitet werden: Die meisten der heute verfügbaren Daten wurden in den beiden vergangenen Jahren angehäuft. Diese stammten häufig aus neuen Datenquellen und -typen. Die vorhandenen Architekturen, Datenbanken und Datenmanagementkonzepte erreichen langsam aber sicher ihre Grenzen der Belastbarkeit. Eine Umfrage zeigte, dass bei 41 Prozent der befragten Unternehmen die vorhandenen Systeme große Datenmengen aus unterschiedlichen Quellen nicht verarbeiten können.³

Hadoop-Skalierbarkeit ohne Kompromisse: Unternehmen wollen aus ihren Hadoop-Investitionen einen größeren Nutzen ziehen und ihre Datenmengen optimal skalieren können. Hadoop bietet jedoch nicht das Analysepotenzial und die Funktionen zur Performance-Optimierung von komplexen Workloads. Workloads. Zudem fehlt es an stabiler Sicherheitsfunktionalität, da mehr Daten auch mehr Risiken mit sich bringen.

Ermittlung des optimalen Nutzens von Hadoop: Unternehmen haben oft Probleme, die richtigen Projekte und Anwendungsfälle zu erkennen, um die Hadoop-Technologie optimal nutzen zu können. Das Hadoop-Framework unterstützt zahlreiche Workloads und Anwendungsfälle. Trotzdem haben die Unternehmen nach wie vor Probleme, zu erkennen, welche umgesetzt werden sollen, und zu beurteilen, welche geschäftlichen Auswirkungen und welchen Nutzen dies bringt.

KUNDENANFORDERUNGEN

Konsolidiertes, skalierbares und kostengünstiges Repository für Big Data: Unternehmen benötigen einen zentralen Standort über Hadoop, um kosteneffektiv ein breites Datenspektrum speichern und verwalten zu können. Dies reicht von Geschäftsdaten, die losgelöst von den engen Strukturen traditioneller Datenbanken sind, bis zu strukturierten, teilstrukturierten und unstrukturierten Daten. Zum Einsatz kommen dabei Standard-SQL-Tools und auch leistungsfähigere Tools.

Workload-Management, hohe Performance, Verfügbarkeit und Sicherheit im ganzen Unternehmen: Unternehmen wollen flexible und gleichzeitig leistungsfähige Lösungen, die auf bewährten Infrastrukturen und Analysetechnologien aufbauen. Diese Technologien sollen ANSI-basierte SQL-Abfragen ermöglichen und je nach Anwendungsfall immer das richtige Rechen-, Speicher- und Analysepotenzial bieten. Sie wollen in Hadoop integrierte Sicherheitsfunktionen, die ihre aktiven, ruhenden und genutzten Daten ausreichend schützen.

Ein Partner, der die Anforderungen des Unternehmens kennt und Erfahrung mit komplexen, heterogenen IT-Umgebungen hat: Unternehmen brauchen Zugang zu erfolgreichen Hadoop-Anwendungsfällen und müssen diese priorisieren. Dann werden die Anwendungsfälle in bestehende Anwendungen und das gesamte BI- (Business Intelligence) und Analyseumfeld eingebunden. IT-Umgebungen sind schwierige und komplexe Umgebungen. Der richtige Partner muss daher auch hier ein fundiertes Wissen mitbringen, um anstehende Probleme lösen zu können.

Nutzenversprechen

Hewlett Packard Enterprise weiß, was Ihr Unternehmen benötigt, um in der Idea Economy von heute nicht nur zu überleben, sondern auch erfolgreich zu sein. Ideen müssen zuerst in erfolgreiche Experimente und dann in umfassende Geschäftslösungen umgesetzt werden. Dieser zweistufige Prozess ist die Folge einer zunehmenden Nachfrage der Kunden, aus Daten verlässliche Informationen zu machen. Die aussagekräftigen Informationen sollen dann helfen, bessere Geschäftsergebnisse zu erzielen. Hewlett Packard Enterprise steht Ihnen hierfür als Partner zur Verfügung, um zusammen mit Ihnen den wertschöpfenden Nutzen Ihrer Daten zu ermitteln. Wir können Sie beim Aufbau einer datengesteuerten Basis unterstützen, damit aus Daten aussagekräftige Informationen werden. Diese Informationen bilden die Grundlage für Ihren Unternehmenserfolg – nach Ihren individuellen Vorstellungen. Dabei wird Hadoop als wichtige Komponente Ihrer Big Data-Analyselösung dienen.

¹ **HPE Big Data Reference Architecture: A Modern Approach, 4AA5-6141ENW, Oktober 2015, Rev. 2**

² HPE interne Studie.

³ „Capitalizing on the promise of Big Data“, PWC Paper, Januar 2013.

Hadoop-Know-how: Bei Hewlett Packard Enterprise hat man alle Aspekte der Hadoop-Umgebung optimiert, damit die Kunden mehrere Anwendungen gleichzeitig ausführen und mehrere Datenspeicher in einem einzelnen System konsolidieren können. HPE Referenzarchitekturen helfen, den neuen Anforderungen gerecht zu werden. Mit diesen Architekturen können Sie unabhängig voneinander die Rechen- und Speicher-Performance skalieren. So lässt sich die gesamte Rechen- und Speicherkette bei den standardmäßigen lokalen Hadoop-Knoten aufbrechen. Traditionelle Architekturen lassen sich optimieren. Mit diesen Erweiterungen können Sie YARN, Spark, Kafka, Mahout und natürlich die Kernmodule – Hive, HBase, MapReduce und Hadoop Database File System (HDFS) – optimal nutzen. Darüber hinaus hat Hewlett Packard Enterprise leistungsfähige Software im Angebot, die über die Hadoop-Funktionalität hinaus weitere Funktionen bietet: Vertica for SQL on Apache Hadoop, Vertica Enterprise Edition für vorausschauende Analysen sowie verschiedene Produktsuites für Information-Governance, Management und Datensicherheit. So steht ein umfassendes unternehmensweites Funktionspektrum zur Verfügung.

Während sich Hadoop im Rahmen der „Datenrevolution“ im Big Data- und Analyse-Markt sehr schnell als tragfähige Open-Source-Technologie etabliert hat, traten in diesem Zusammenhang auch einige deutlich sichtbare Herausforderungen auf. Aufgrund ihrer Größe und den umfassenden Sicherheits- und Zuverlässigkeitsfunktionen bringen Hadoop-Implementierungen komplexe Herausforderungen bei Planung, Bereitstellung und langfristigem Management mit sich. Aktuell ist auf dem Markt praktisch kaum Hadoop-Know-how verfügbar. Die Herausforderungen, die sich durch dieses fehlende Wissen ergeben, werden durch die stetige Weiterentwicklung der Hadoop-Technologie verstärkt, da damit auch mehr Probleme und Unsicherheiten verbunden sind.

Hewlett Packard Enterprise hat hierfür einen bewährten vierstufigen Ansatz für die Implementierung hoch entwickelter Daten- oder Analysesysteme entwickelt: Erkennung, Entwicklung, Integration und Implementierung. Dieser Ansatz hilft zahlreichen kleinen und großen Unternehmen weltweit, noch mehr von Hadoop zu profitieren und bessere Geschäftsergebnisse zu erzielen.

Zugang zu qualifizierten Analyseexperten auf der ganzen Welt: Hewlett Packard Enterprise implementiert unternehmenskritische Lösungen für Kunden auf der ganzen Welt und setzt dafür über 3.500 Analyseexperten ein, die im Durchschnitt über 18 Jahre Erfahrung im Analysebereich haben und in neun Centers of Excellence auf vier Kontinenten zur Verfügung stehen.

Integrierte Lösungen: Unser leistungsfähiges Portfolio an offenen Standardprodukten, offenen Architekturen und Services sowie unser fundiertes Know-how helfen Ihnen bei der Integration von Analysefunktionalität in Standardprozesse und -workflows. So können Sie die ermittelten aussagekräftigen Informationen in die Tat umsetzen und sind nicht von unflexiblen proprietären Lösungen abhängig, die auf lange Sicht nicht rentabel sind.

Über 1200
Analyseexperten weltweit

Über 18
Jahre an Analyseerfahrung

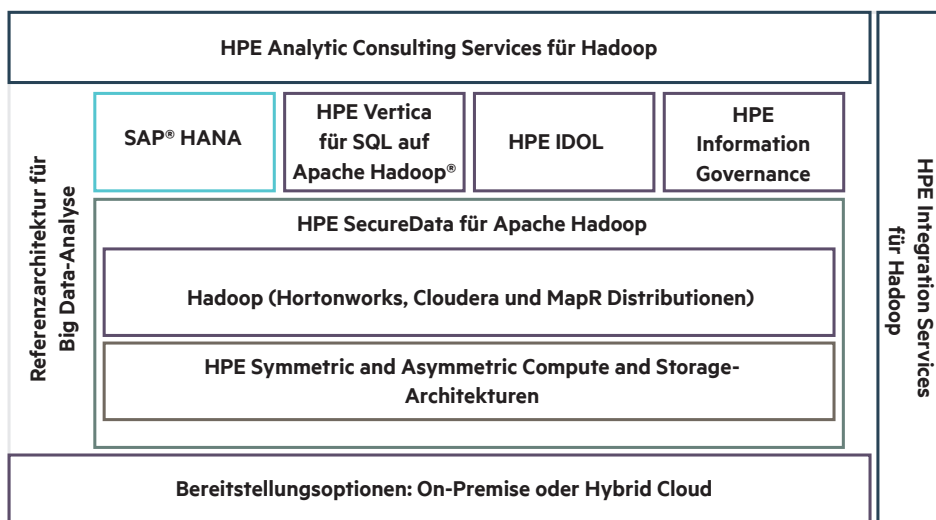


Abbildung 1: Überblick über die Lösung

Hadoop der Enterprise-Klasse

Es herrscht zunehmend die übereinstimmende Meinung, dass für eine datengesteuerte Basis zur Verarbeitung großer Datenmengen im TB-/PB-Bereich mit unterschiedlichen Datentypen (Maschinen-, Human- und traditionelle Geschäftsdaten) unweigerlich Hadoop die zentrale Komponente sein wird. HDFS ist solch ein „Data Lake-Ziel“ für Daten, die aus vorhandenen schemabasierten Datenbanken extrahiert werden. HDFS dient aber auch als Ziel für zahlreiche neue, weniger strukturierte und unterschiedliche Datenquellen wie soziale Medien, Maschinenprotokolle und andere neue Big Data-Quellen. Hewlett Packard Enterprise baut auf das Open-Source-Konzept, um eine unternehmensweite Hadoop-Lösung zu ermöglichen.

KOMPONENTEN DER LÖSUNG	BESCHREIBUNG
------------------------	--------------

HPE As a Service-Lösung für Hadoop	Die HPE As a Service-Lösung für Hadoop ist eine auf große Unternehmen zugeschnittene, auf Open-Source-Technologie basierende Big Data-Plattform. Es geht hier um eine als Service implementierte, cloudbasierte Lösung und ein Verbrauchsmodell, die neben unseren gehosteten und On-Premise-Lösungen angeboten werden. Unsere Lösung stellt den Kunden eine konfigurierte Hadoop-Plattform in der HPE Helion Cloud zur Verfügung – ohne hohe Vorlaufkosten, hohe Implementierungsrisiken, Mehrbelastung der IT-Ressourcen sowie „Shelfware“. So können Sie sich voll und ganz auf die Umsetzung Ihrer Geschäftsziele konzentrieren. Die HPE Lösung ist in verschiedenen geschichteten Workloadoptionen verfügbar und kann so ganz flexibel an Veränderungen angepasst werden. Risiken durch veraltete Technologien sind kein Thema mehr. Die HPE As a Service-Lösung für Hadoop kann durch die Datenbankanalysefunktion von HPE Vertica erweitert werden. So können dann auch größere Datensets nahezu in Echtzeit gehandhabt werden. Mit IDOL können zudem alle Arten von unstrukturierten Daten verarbeitet werden.
HPE Analytic Consulting Services	Mit HPE Analytic Consulting Services kann der Nutzen Ihrer gesamten Daten innerhalb weniger Stunden oder Tage statt bisher Wochen oder Monaten bestimmt werden. Wir helfen Ihnen auch, die geeigneten Komponenten einer datengesteuerten Basislösung zu ermitteln, Analysen auf der Grundlage dieser Basislösung genau dort in Ihrem Geschäftsprozess durchzuführen, wo dies erforderlich ist, und unterstützen Sie natürlich auch bei der Implementierung der Lösung. Auch bei der Ermittlung der optimalen Kombination aus bestehenden BI-/EDW-Lösungen (Enterprise Data Warehouse), der Abwägung von Investitionen im Vergleich zu Batchanalysen in Echtzeit sowie zur Nutzung von Hadoop als Data-Lake- und Analyseplattform stehen wir Ihnen beratend zur Seite.
Hadoop	Hadoop ist ein Apache-basiertes, modulares Open-Source-Software-Framework für das Management großer Datenmengen und für Analyse Zwecke. Mit dieser Software lassen sich große Datenmengen bis in den Petabyte-Bereich auf offener Standardhardware verarbeiten.
Datengesteuerte Hadoop Integrationservices	Mit HPE Discovery and Transformation Services für Hadoop können Sie von folgenden Vorteilen profitieren: Aufbau eines formalen und strukturierten Ansatzes für die Datenverarbeitung und -analyse. Hierzu gehört auch eine strukturierte Methodologie, die Nutzung spezieller Tools für Datenvirtualisierung und gemeinsame Datennutzung im Rahmen einer integrierten Plattform. Hinzu kommen Schulungsmaßnahmen für den weiteren Wissensaufbau bei Ihren Mitarbeitern sowie Methoden für eine effizientere Zusammenarbeit. Und nicht zu vergessen die Best Practices bei Experimentierphasen, die ambitioniert aber auch kosteneffektiv erfolgen müssen. HPE Integration Services für Hadoop helfen Ihnen bei der Optimierung und Modernisierung Ihrer Infrastruktur. Sie unterstützen zudem bei allen weiteren Systemanforderungen, um auch auf der Hadoop-Plattform unternehmensweite Funktionalität bereitstellen zu können.
HPE IDOL 10	Mit HPE Intelligent Data Operating Layer (IDOL) können Sie auf große Mengen an Humandaten im Petabyte-Bereich zugreifen, diese analysieren, besser verstehen und nutzen. Diese Daten stammen in der Regel aus den unterschiedlichsten Quellen wie der Cloud oder aus internen Quellen und sind in Ihrem HDFS Data Lake – kontextbezogen und in mehreren Sprachen – gespeichert. Hinzu kommen 500 Analysefunktionen, die auf Text-, Bild-, Audio-, Video-, Social Media- und strukturierte Daten in Datenbanken angewendet werden können. IDOL enthält direkt einsatzbereite Konnektoren für über 1000 Standarddateiformate und 400 Anwendungs- und Datenbankschnittstellen. IDOL kann zusammen mit Hadoop verwendet werden, um Ihren Hadoop Data Lake zu kategorisieren, zu indexieren und sinnvoll nutzen zu können.
HPE Vertica for SQL on Apache Hadoop	Mit HPE Vertica for SQL on Apache Hadoop können Sie unternehmensweit SQL-Abfragen für Ihre Hadoop-Daten durchführen. Dabei steht Ihnen das gesamte Funktionsspektrum von ANSI SQL-Syntax auf MapR, Hortonworks und Cloudera zur Verfügung. Mit HPE Vertica for SQL on Apache Hadoop bieten sich Ihren Geschäftsanalysten, Datenwissenschaftlern und Entwicklern sofortige Integrations- und Zugriffsmöglichkeiten für Ihre traditionellen BI-Tools und die aktuellen Analyse- und Visualisierungstools nach Branchenstandard. HPE Vertica ermöglicht unternehmensweite, innovative Analysen, die alle Anforderungen von ANSI SQL sowie Atomicity, Consistency, Isolation, Durability (ACID) erfüllen. Zudem werden auch komplexe Datentypen unterstützt und weitere Funktionen geboten, die nur über unsere SQL on Hadoop-Implementierung zur Verfügung stehen. Die Lösung überzeugt durch Optimierungsmöglichkeiten wie Komprimierung, Tabellenspeicher und Projektionen, durch die ein Performance-Niveau erreicht wird, das deutlich über dem von SQL on Hadoop-Standardlösungen liegt.
HPE Referenzarchitekturen für Hadoop	Voraussetzung für eine wirksame Risikominimierung und die Optimierung von Humandaten, Finanzdaten und Rechenzentrumsressourcen ist auf jeden Fall die Wahl der richtigen Architektur, die Ihre individuellen Hadoop-Anwendungsfälle unterstützt: Hive, SQL, Spark In-Memory-Verarbeitung, Kafka-Streamingdaten, Mahout Machine Learning usw. – Hewlett Packard Enterprise bietet eine Vielzahl von Referenzarchitekturen, die die Workloadoptimierung für Hadoop basierend auf Kombinationen aus Moonshot, HPE Apollo 2000 und 4000 Gen9 Serie und ProLiant XL Servern ermöglicht. Dabei werden Rechen- und Speicheranforderungen für Hadoop-Cluster mit allen drei Hadoop-Distributionen aufeinander abgestimmt: MapR, Hortonworks und Cloudera. Die Minotaur-Referenzarchitektur erweitert die Hadoop-spezifischen Referenzarchitekturen auf SAP HANA, HPE Vertica und HPE IDOL.
HPE Financial Services	Die richtige Balance aus dem neuen Hadoop Data Lake und den neuen Analysefunktionen mit der vorhandenen Infrastruktur und Systemumgebung zu finden, erfordert finanzielle Flexibilität. Mit der richtigen Finanzierung können Sie im gesamten Implementierungsprozess schnell auf Veränderungen reagieren und sich darauf einstellen sowie Innovationen besser umsetzen und ein ausgewogenes Kosten-Nutzen-Verhältnis erreichen. Darüber hinaus können Ihnen die Experten bei HPE Financial Services helfen, Risiken in Bezug auf den Unternehmensruf und die Einhaltung gesetzlicher Bestimmungen in den Griff zu bekommen, die bei der Entsorgung alter Systemkomponenten auftreten können. Hierfür stehen spezielle HPE Asset Recovery Services zur Verfügung.
HPE SecureData	HPE SecureData ist ein datengesteuertes Sicherheitskonzept für Hadoop, bei dem die Daten verschlüsselt und mit Tokens versehen werden. Die Daten werden dadurch – eng an der Datenquelle ausgerichtet – unkennlich gemacht. Die vertraulichen Datenelemente werden in nutzbare, jedoch nicht erkennbare Äquivalente transformiert, die ihr Format, ihr Verhalten und ihre Bedeutung beibehalten. Diese Form von Datenschutz kann in Anwendungen, Analyseengines, bei Datenübertragungen und in Datenspeichern verwendet werden. Die Daten können dann sehr schnell und sicher wieder für die Anwendungen und Benutzer, die sie brauchen, kenntlich gemacht werden. HPE SecureData ist in der Lage, eine nahezu unbegrenzte Anzahl an Datentypen auf Datenebene zu schützen.

Optimieren Sie Ihre IT-Investitionsstrategie mit neuen nutzungsbasierten Technologiefinanzierungsmodellen im Einklang mit Ihren Geschäfts- und Transformationszielen.
hpe.com/solutions/hpefinancialservices

Was kommt als nächstes?

Nehmen Sie an einem Transformations-Workshop teil. Wenden Sie sich an Ihren HPE Vertriebsmitarbeiter, um sich für einen Transformations-Workshop anzumelden.

Weitere Informationen unter
hpe.com/us/en/solutions/empower-data-driven.html



Melden Sie sich noch heute an.