

Fünf Möglichkeiten der Wertschöpfung durch eine Data Fabric – wo auch immer Ihre Daten sich befinden

Eine Data Fabric kann das Gewinnen von Erkenntnissen aus Daten unterstützen, die auf lokale, Cloud- und Edge-Bereitstellungen verteilt sind



Ihr Unternehmen konzentriert sich auf Ihre Daten, nicht nur auf die Anwendungen für Ihren Geschäftsbetrieb. In lokalen sowie in Cloud- und Edge-Umgebungen verteilte Daten stellen neue Herausforderungen für Entwickler, Data Engineers und Data Scientists dar. Ohne eine unternehmensweite Ansicht Ihrer Daten wird es zunehmend schwieriger, zu wissen, welche Daten verfügbar und wie diese zugänglich sind, sie zu verschieben sowie ihre Sicherheit und Compliance zu gewährleisten.

Eine Data Fabric ist essenziell für die Handhabung verteilter Datenumgebungen, sagt Joann Starke, Senior Product Marketing Manager von HPE Ezmeral bei Hewlett Packard Enterprise. Eine Data Fabric ist eine Architektur, die Daten mehrerer Standorte in eine einheitliche Umgebung integriert und so Ansicht, Verwaltung, Sicherheit und Government vereinfacht.

„Eine Data Fabric bietet Ihnen die Möglichkeit, Ihre Daten dort zu belassen, wo sie sich befinden, allerdings mit globalem Benutzer- und Anwendungszugriff, unabhängig von ihrem Speicherort“, sagt sie. „Durch die Nutzung einer Data Fabric muss ein Data Engineer oder ein Business Intelligence-Analyst, der die Daten für den Data Scientist aufbereitet, nicht mit zahlreichen Datenverantwortlichen über den Zugriff auf die Quellen verhandeln.“

Dies sind fünf Möglichkeiten zur Wertschöpfung aus Ihren Daten, wo auch immer sich diese befinden.



1 Bessere Datentransparenz

Die Datentransparenz ist ein Maß dafür, wie Ihr Unternehmen Daten von verteilten Quellen überwachen, anzeigen und analysieren kann. Eine gute Datentransparenz bietet viele Vorteile, wie Geschäftsentscheidungen mit besserer Informationsgrundlage. Ist Ihre Datentransparenz jedoch schlecht, agiert Ihr Unternehmen quasi blind und ohne wirklich zu wissen, über welche Daten Sie verfügen, wo sie gespeichert sind und wer sie nutzt. Dies kann zu einer Reihe von Ergebnissen führen, die jedoch alle negativ sind.

Die Datentransparenz wird durch heutige Hybrid Cloud- und Multi-Cloud-Umgebungen kompliziert, in denen Daten über verschiedene geografische Standorte und Architekturen hinweg erzeugt und gespeichert werden. Laut einer IDC-Recherche erzeugt und speichert ein durchschnittliches Unternehmen Daten an 14 bis 20 unterschiedlichen Standorten.¹

„Jeder dieser Standorte verfügt über unterschiedliche Anwendungen mit jeweils eigener Oberfläche, eigenen Formaten und zahlreichen proprietären Tools“, sagt Starke. „Dies bedeutet, dass eine umfassende Ansicht aller Ihrer Daten nahezu unmöglich ist.“

Eine Data Fabric unterstützt die Kombination aus Dateien, Objekten und Streaming-Daten innerhalb einer globalen Datenebene, die sich auf Hybrid Cloud- und Multicloud-Umgebungen erstreckt. Diese Möglichkeit reduziert die Zeit für Discovery, Kopieren, Integrieren und Normalisieren der Daten in lokalen, Cloud- und Edge-Umgebungen.

„Eine Datenebene blendet all diese Infrastruktur aus“, sagt Starke. „Sie führt Dateien, Objekte und Streaming-Daten in einer zentralen Quelle zusammen. Dann ermöglicht der native globale Namensraum autorisierten Benutzern den Zugriff auf die Daten, unabhängig davon, wo sie sich befinden. Der Zugriff ist so einfach wie das Anklicken eines Eintrags in der Verzeichnisstruktur eines Benutzers.“

¹ „[IDC FutureScape Webcast: Worldwide Future of Digital Infrastructure 2023 Predictions](#),” IDC, Januar 2023

2 Unterstützung für Multiformat-Datenzugriff

Mit wachsender Anzahl Ihrer Datenquellen wächst auch die Anzahl der Datenformate und Methoden, die für den Zugriff auf diese Daten erforderlich sind. Damit in einem bestimmten Format vorliegende Daten in einem anderen Format gelesen werden können, müssen sie kopiert und in unterschiedlichen Umgebungen gespeichert werden, wodurch mehr Kapazität für die Handhabung von Datenduplikaten und mehr Zeit für ihre Verarbeitung erforderlich ist.

Die Datenebene überwindet diese Hürden durch die Unterstützung unterschiedlicher Datenformate, so dass Ihre Daten in einem Format gespeichert und in einem anderen gelesen werden können. „Dies bedeutet, Sie können Ihre Daten nativ speichern und anschließend mithilfe von NFS lesen“, erklärt Starke. Dies reduziert den Zeitaufwand für eine Neuformatierung der Daten, bevor sie verarbeitet werden.

3 Integrierte Sicherheit

Die konsistente Anwendung anbieterübergreifender Sicherheitskonzepte stellt ein Risiko dar. Starke nennt intensive Cloud-Anwendungsfälle als Beispiel. Entwickler möchten ihre Anwendungen in der Cloud erstellen, beispielsweise in Amazon Web Services, und anschließend in Azure Testen, um von den besseren Testmöglichkeiten der Plattform zu profitieren. Wenn es jedoch um den Produktionsbereich geht, erfolgt dies lokal. Diese Multicloud-Umgebung ist mit Sicherheits-Komplexitäten und -Risiken verbunden.

Laut eines Berichts von Radware aus dem Jahr 2022 geben 69 % der Unternehmen an, unter Datenschutzverletzungen oder Schwachstellen zu leiden, die durch Variationen von Multicloud-Sicherheitskonfigurationen entstehen.² Eine Data Fabric mit nativ in die Lösung integrierter Sicherheit mindert diese Probleme durch die Möglichkeit einer konsistenten Anwendung von Sicherheits-, Verschlüsselungs-, Authentifizierungs- und Geofencing-Richtlinien in allen Umgebungen.

„Die Sicherheit einer Data Fabric kann in Ihre Authentifizierungs- und Autorisierungssysteme integriert werden“, sagt Starke. „Und sie ermöglicht eine Zugriffssteuerung: Wer wann berechtigt ist, die Daten einzusehen, und ob Geofencing erforderlich ist. Mit anderen Worten: Wenn Sie eine mandantenfähige Lösung in Deutschland benötigen, können Sie diese Daten per Geofencing schützen und vermeiden, dass sie nach außen dringen. Autorisierte Benutzer können jedoch weiterhin darauf zugreifen.“

4 Vereinfachte Datenverwaltung

Das Volumen der von Ihrem Unternehmen erfassten Daten wird ohne Frage weiter steigen, daher ist es entscheidend, die Komplexität der Datenverwaltung zu reduzieren. Mithilfe einer Data Fabric können Sie Dutzende verteilte Toolsets mithilfe einer zentralen Lösung nutzen, die Datenverwaltung und Betrieb mit einer einheitlichen Ansicht, Unterstützung für Industriestandard-APIs sowie einen zentralen Datenzugriffspunkt für Benutzer und Anwendungen bietet.

„Sagen wir, Sie sind ein Fertigungsunternehmen mit mehreren Standorten, von denen jeder einen Edge-Standort darstellt“, sagt Starke. „Mit einer Data Fabric können Sie die Daten verarbeiten, um sofortige Erkenntnisse über diese Produktionsstandorte zu gewinnen, und diese anschließend auf einen anderen Standort abbilden. Sie müssen zum Beispiel nicht 10 bis 20 Petabyte an Daten kopieren, sondern Sie erstellen eine automatisierte Richtlinie, welche die Erkenntnisse für jeden dieser Standorte repliziert, da sie lediglich die gewonnenen Erkenntnisse verschieben, statt der gesamten Daten.“

² „Report: 69 % der Unternehmen berichten, dass Multicloud-Sicherheitskonfigurationen zu Datenschutzverletzungen oder Schwachstellen geführt haben“, VentureBeat, Oktober, 2022

5 Self-Service-Zugriff auf Daten

Zur Optimierung von KI-Modellen benötigen Data Scientists und Data Engineers laut Starke einen mühelosen Zugriff auf die erforderlichen Daten. Wenn die Daten jedoch in einer verteilten Datenumgebung gespeichert sind, können die Benutzer nur schwer bestimmen, welche Datenbestände verfügbar und welche Daten aktuell sind oder wie der Zugriff darauf möglich ist. Eine Data Fabric-Lösung kann einen Self-Service-Zugriff ermöglichen, so dass die jeweiligen Interessengruppen von jedem Ort aus an alle benötigten Informationen gelangen.

„Sie können in der Datenebene nach Metadaten suchen“, erklärt Starke. „Sie können das Alter der Datenbestände und andere Details sehen, um sicher zu sein, dass Sie die aktuellsten und relevantesten Daten nutzen. Dies vereinfacht die Bestimmung der richtigen Datenbestände für jedes Analyseprojekt.“

Wenngleich viel über das stetig steigende Datenvolumen gesprochen wird, ist das größte Problem dabei, eine Wertschöpfung aus den Daten zu erzielen, nicht deren Menge, sondern deren Speicherort. Die Vereinheitlichung Ihrer gesamten Daten mithilfe einer Data Fabric kann den Zeitaufwand für Discovery, Kopieren, Integration und Normalisierung der Daten in Ihren lokalen, Cloud- und Edge-Bereitstellungen reduzieren. Mithilfe einer hochentwickelten Data Fabric können Ihre Benutzer effektiver auf Daten zugreifen und diese nutzen sowie Informationen nutzen, mit denen Sie sich einen Wettbewerbsvorteil sichern.

**Mehr erfahren
unter**

GreenLake.HPE.com/data

Besuchen Sie **HPE GreenLake** 

