



men, ist identisch. Der Begriff „Big Data“ definiert eine entscheidende neue Technologie in der IT.

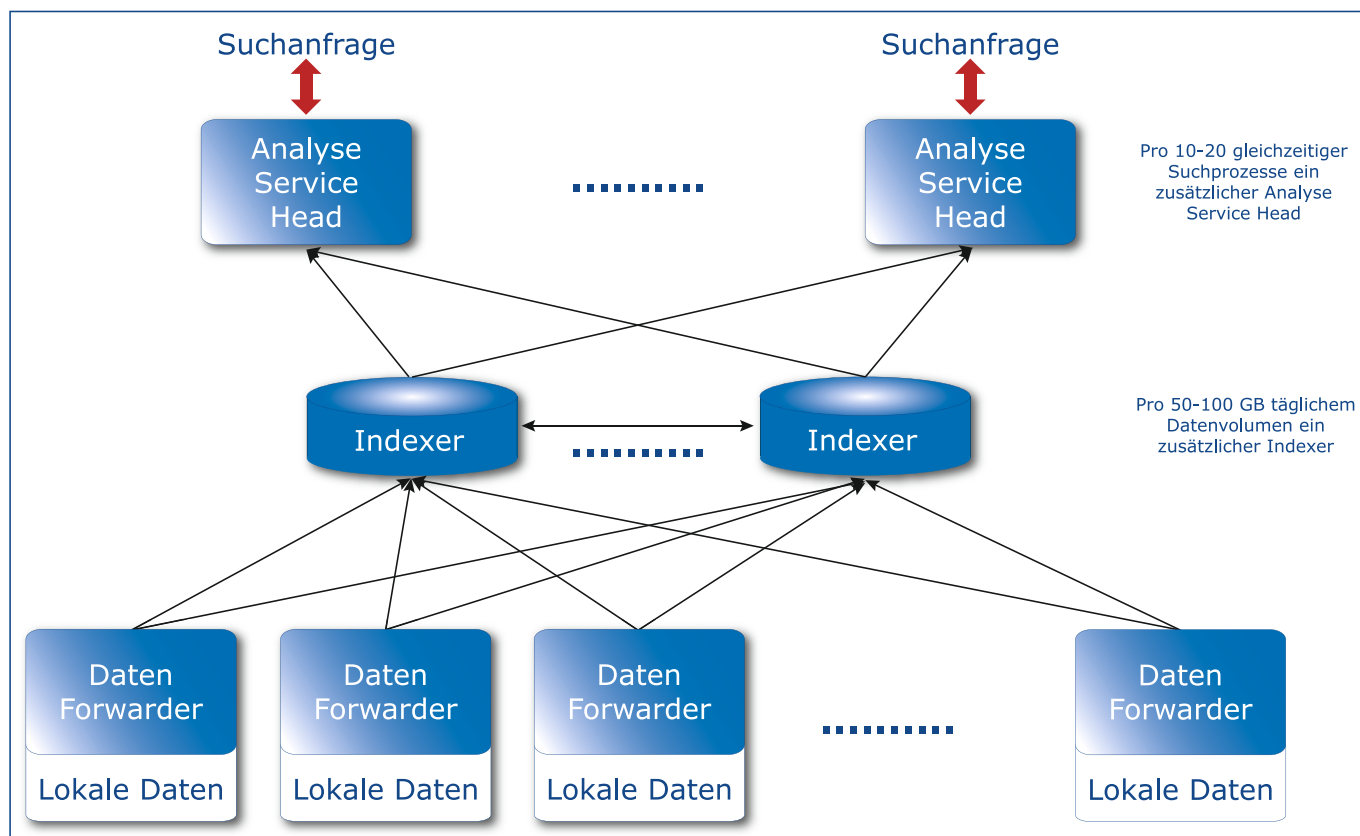
Reduziert man sämtliche Big Data-Definitionen auf einen gemeinsamen Nenner, so versteht man unter Big Data die besonders flexible Analyse von unstrukturierten oder teilstrukturierten Massendaten, wie sie z. B. in Logfiles oder großen Batchfiles auftreten. Internet-Anwender sind z. B. beim „Googlen“ gewohnt, dass schon ohne vollständige Suchbegriffe, ausgehend von den zuerst eingegebenen Buchstaben sofort passende Daten selektiert und angezeigt werden.

Eine weitere Stärke von Big Data liegt in dem komfortablen Handling von besonders großen Datenmengen, die mit dem, was wir heute als Standard-Datenbank verstehen, nicht mehr flexibel analysiert werden können. Die Rede ist von Exa- und Zettabytes – Datenmengen, die jedes relational angelegte Datenbanksystem an die Grenzen bringen.

### Neue Konzepte ermöglichen freie Datenanalyse in Echtzeit

Die uneinheitliche Struktur der Daten verlangt in Kombination mit sehr großen Datenmengen

nach neuen Verarbeitungskonzepten. Ausgelöst von den enormen Datenmengen, die heutige Suchmaschinen wie Yahoo, Google und dergleichen tagtäglich zu verarbeiten haben, ist unter dem Namen Big Data ein neues Datenanalysekonzept entstanden. Wurden die Daten bisher zentralistisch in einem relationalen Datenbankmanagementsystem anwendungsspezifisch indiziert, verwendet Big Data ein anderes, umfassendes und vollautomatisiertes Indizierungsverfahren auf Filesystemebene. Die Indizierung ist skalierbar und kann je nach Datenvolumen auf verschiedene Server verteilt werden. Kombiniert mit moderner in-me-



Grafik: Consist

Dezentrale Datenzugriffe mit Big Data-Technologie

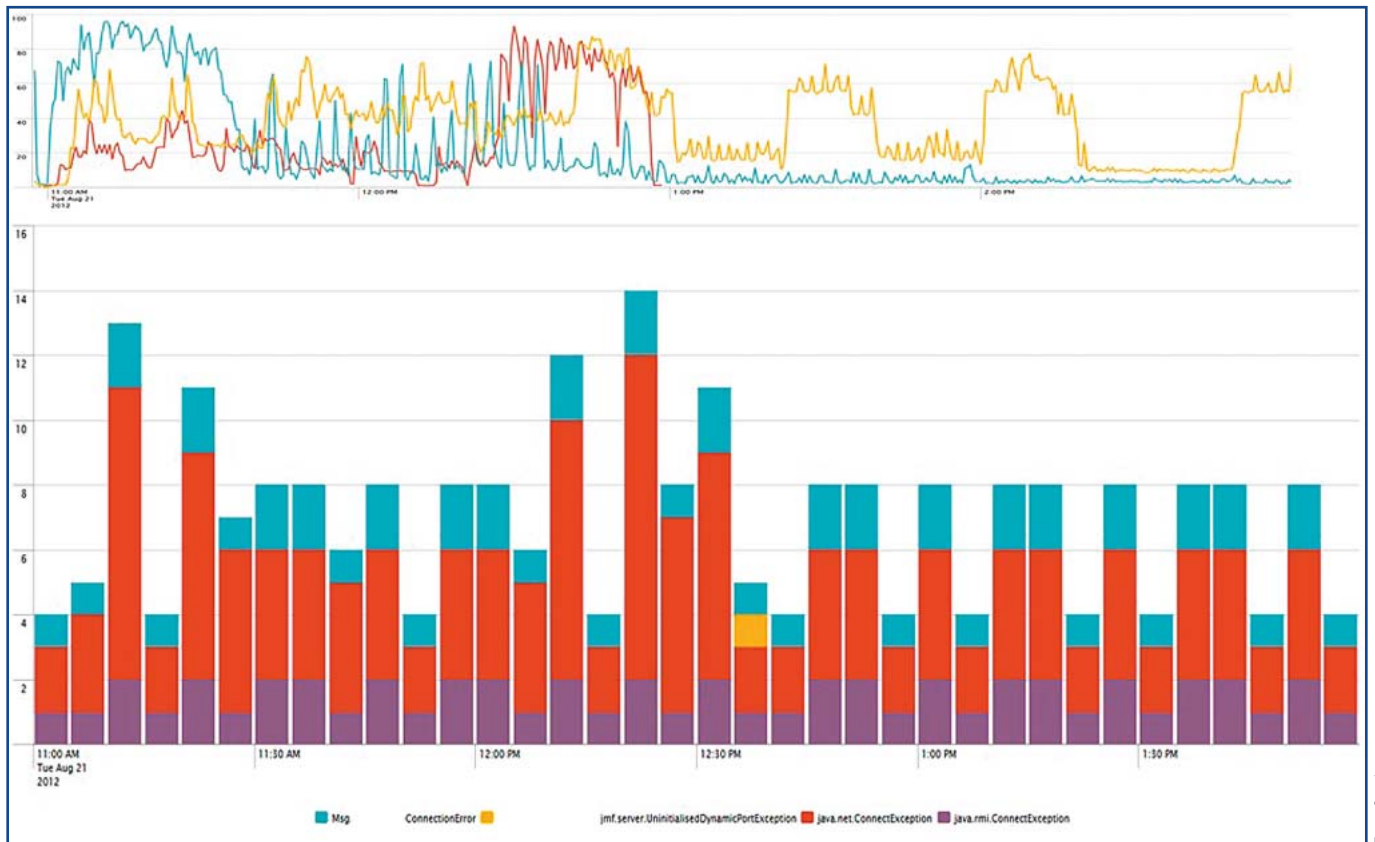


Foto: Consist

Ergebnis einer Datenanalyse mit Splunk

mory-Technologie wird somit eine bislang unerreichte Flexibilität und Performance in der Datenanalyse möglich – quasi in Echtzeit.

## Beschränkungen durch Extraktion und Transformation entfallen

Werden Daten in klassischen DBMS oder BI-Cubes geladen, so werden diese in der Regel entsprechend eines normierten Datenmodells reduziert. Die Datenbank nimmt nur die Daten auf, die in ihre normierte Struktur passen.

Die im Bereich Big Data eingesetzten Technologien arbeiten

anders. Die anfallenden Daten verbleiben unverändert im Raw-Format ohne Extraktions- und Transformationsprozesse. Der maximale Informationsgehalt der Daten bleibt somit erhalten. Damit ist es möglich, die Gesamtdaten jederzeit auf neue, a priori möglicherweise unbekannte Fragestellungen hin zu untersuchen.

## Neue Technologien ermöglichen neue Anwendungsfelder

Mit der neuen Big Data-Technologie lassen sich neue Anwendungsfelder erschließen. Ein wichtiges Thema für Big Data ist die

Analyse von großen Flatfiles im Bereich des Netzwerk-Monitorings wie z. B. Logfiles, Firewall-Protokolle oder Web Requests. Für explorative Fragestellungen, die typischerweise bei der Fehleranalyse oder im Compliance-Umfeld auftreten, hat sich Big Data bereits hervorragend bewährt. Durch die umfassende Datenauswertung in Echtzeit können Ausfallzeiten erheblich reduziert werden und somit zu einer höheren Gesamtverfügbarkeit beitragen.

In der Softwareentwicklung wird Big Data-Technologie mit großem Erfolg im Testbereich genutzt. Auch in diesem Umfeld

## Big Data ermöglicht eine bislang unerreichte Flexibilität und Performance in der Datenanalyse

treten typischerweise zuvor nicht planbare Fragestellungen bei der Analyse von großen Datenmengen auf. Big Data ermöglicht eine erhebliche Zeitersparnis in der Fehlerdiagnose und senkt signifikant die Testkosten.

Auch der Bereich Business Intelligence kann und wird seine Vorteile aus dem Big Data-Bereich ziehen. Big Data eröffnet die Chance, auch außerhalb des starren BI-Cube-Modells explorativ auf den Gesamtdaten zu arbeiten.

### Big Data hat Zukunft

Big Data ist eine junge Technologie, die bereits einen festen Platz in modernen IT-Strategien erobert hat. Sie ermöglicht erstmalig eine

umfassende und besonders flexible Datenanalyse auf Datenmengen beliebiger Größe und ergänzt damit die bewährten Analysemöglichkeiten der traditionellen Datenbankmanagementsysteme.

Consist selbst setzt in Kundenprojekten bereits Splunk ein und hat damit Fehleranalysen komplexer Systeme von mehreren Tagen bis auf Minutenwerte reduzieren können.

### Consist – Ihr Partner für Big Data

Consist hat die Chancen von Big Data erkannt und eine Partnerschaft mit dem führenden Anbieter Splunk Inc. vereinbart. Consist ist damit nicht nur autorisierter Re-

seller von Splunk, sondern steht insbesondere mit zertifizierten Spezialisten für Ihre Big Data-Konzepte und -Projekte zur Verfügung.

Weitere Informationen:

Björn Jensen  
Telefon: 0431/3993-632  
E-Mail: [jensen@consist.de](mailto:jensen@consist.de)



### Über Splunk

Splunk ist der führende Anbieter von Operational Intelligence-Software, mit der Echtzeit-Computerdaten sowie Terabytes von historischen Daten vor Ort oder in der Cloud überwacht, analysiert und in Berichten dargestellt werden. Fast die Hälfte der Fortune 100-Unternehmen und mehr als 4.400 Unternehmen, Dienstleister

und Behörden in 80 Ländern nutzen Splunk, um Service-Levels zu verbessern, IT-Betriebskosten zu reduzieren, Sicherheitsrisiken zu verringern und die Transparenz betrieblicher Prozesse zu erhöhen.

Splunk beschäftigt aktuell mehr als 600 Mitarbeiter weltweit in 14 Büros und im Hauptquartier in San Francisco. Die Aktien wer-

den unter dem Kürzel SPLK an der NASDAQ seit 2012 gehandelt.

[www.splunk.com](http://www.splunk.com)