



managedhosting.de

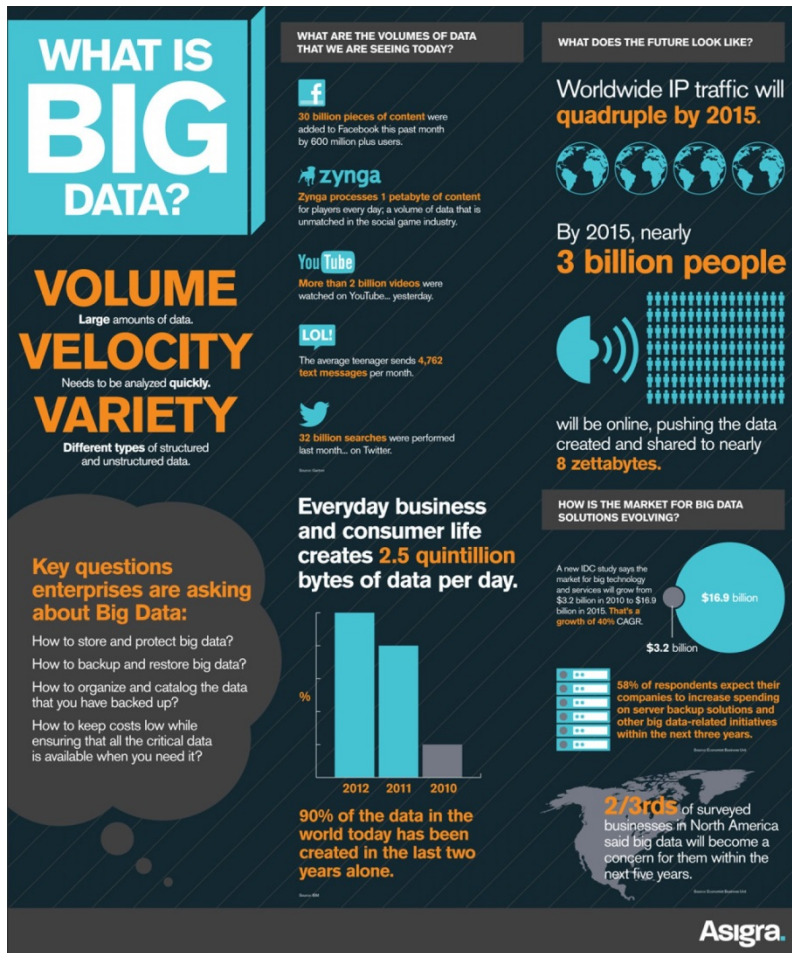
the uptime* company

15. Informatiktag an der HS Mittweida

„Einsatzmöglichkeiten und
Anwendungsbereiche verschiedener
Methoden der Datenspeicherung im Big
Data Zeitalter.“

Andreas Wolske

Big Data



- Server im Jahr **1999**
 - 1-4 GByte RAM
 - Pentium II - 450MHz
- Server im Jahr **2014**
 - 1 TByte RAM
 - 2x 12 Core - 50 GHz
- Festplatten im Jahr **1999**
 - 180 GByte
- Festplatten im Jahr **2014**
 - 10 TByte 3,5"
 - 2 TByte 2,5"
 - 1 TByte SSD

Was ist Big Data?

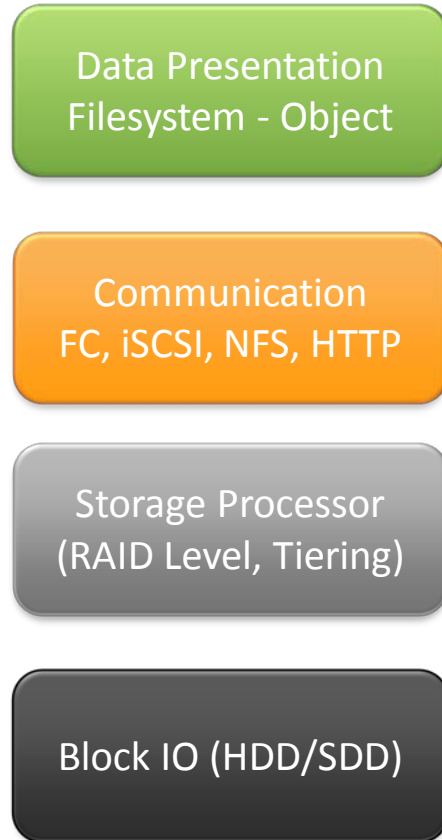
Definition

- Große Mengen an strukturierten und unstrukturierten Daten potentiellen Wissenszuwachs durch Ihre Analyse bergen.
- Rohdaten werden durch Algorithmen nach Mustern durchsucht
- Echtzeitanalyse

Merkmale

- Änderungshäufigkeit, Entstehungsgeschwindigkeit
 - Echtzeit
 - Periodisch
 - Einmalig
- Datenvolumen
 - MB, GB, TB, PB
- Datenart
 - Video, Audio
 - M2M, Logs
 - Soziale Medien
 - Datenbanken

Storage I/O Stack vs. Anforderung



- Größe der einzelnen Objekte: 2x 62TB .vmdk vs. 500 Mio. x 250 kByte .jpg
- Verfügbarkeit (RTO/RPO)
- Latenz (IOPS)
- Geografische Verteilung
- Applikationskonsistenz (Metadaten / Datenbanken)

Ansiedlung der Funktionen

Technologie

- RAID
- Storage Replication
- Filesystem Replication (ZFS)
- Filesystem Snapshots
- Distributed Filesystems
- Clustered Filesystems
- Application Clusters

Anforderung

- Bereitstellungszeit
- Datensicherheit
- Datenverfügbarkeit
- Platzbedarf vs. Dateigröße
- Performance
- Geografische Verteilung
- Metadaten (Versionierung)

Software Defined Storage

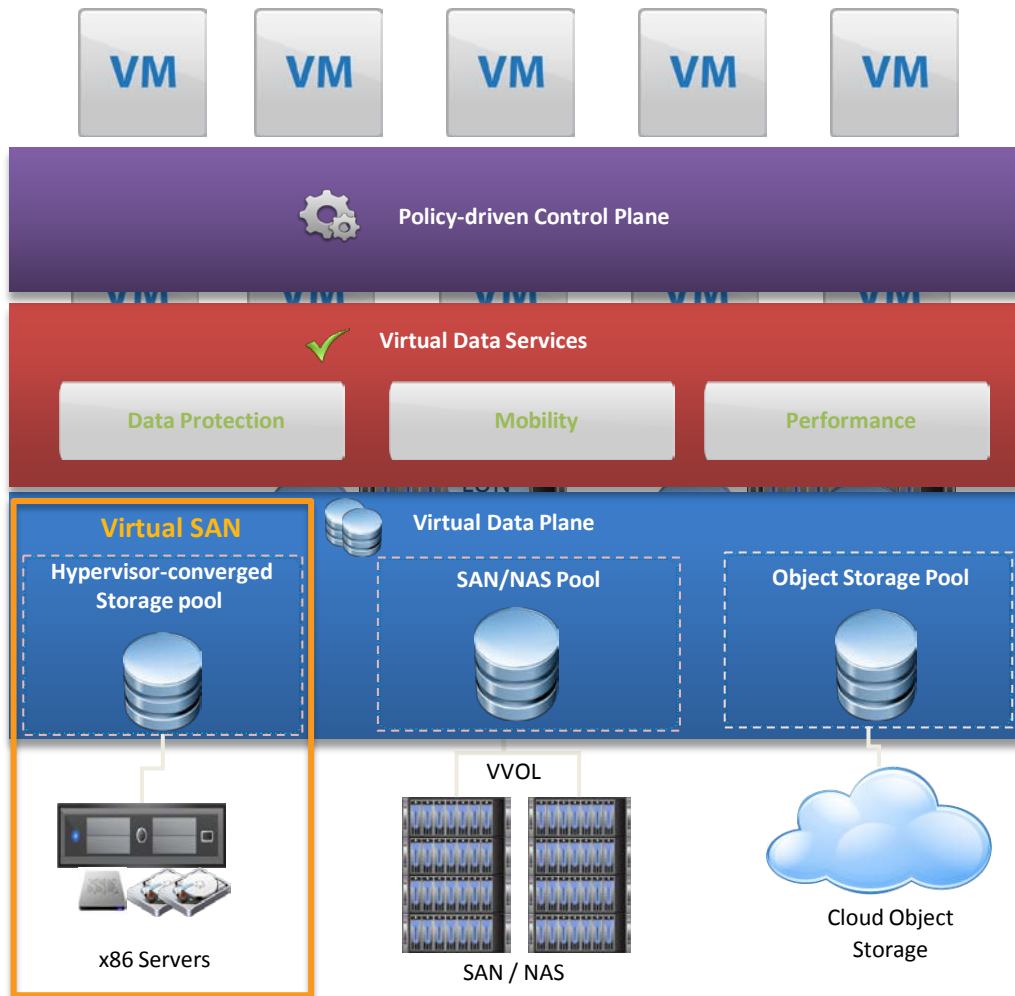
Definition

- Programmierbarkeit
- APIs
- Selbstbedienung
- Abstraktion
- Bereitstellung aller Typen
- Metadaten am Objekt
- Mobilität
- Funktionskonvergenz

Lösungsansätze

- VMware vSAN
- EMC ViPR, XtremIO, ScaleIO
- Hyperconvergent:
 - Nutanix
 - Simplivity
 - Pivot3
- Microsoft Storage Spaces
- Amazon S3
- RedHat / GlusterFS
- PernixData
- Openstack Swift / Cinder

Evolution der Architektur



- Gemeinsame regelbasierte Verwaltung und Automation
- VM – zentrische Datenbereitstellung
- Integration von anderen Technologien
- Abstraktion und Ressourcenpools
- Integration des Hardwarelayers
- Neue Speicherklassen und Qualitäten

Anbieter und Konzepte

Hyper-Converged

NUTANIX
The Virtual Computing Platform

simplivity

SCALE
COMPUTING

Traditionelles Storage

EMC²

hp

IBM

NetApp

DELL

Hybrid Arrays

TINTRI

nimble
storage

coraid

hp HP StoreVirtual 4000
Storage

Windows Server
Storage Space

redhat.
GLUSTER

EMC² ScaleIO
ScaleIO: Distributed File System

nexenta
Enterprise class storage for anyone

pernixdata

hp HP StoreVirtual VSA

atlantis
COMPUTING

Softwarebasiert

Converged

VCE

CISCO

FlexPod™

NetApp

NIMBUS DATA

violin
MEMORY

EMC² XtremIO

WHIP TAIL

PURE STORAGE

All Flash Arrays

Fazit

Data Presentation
Filesystem - Object

Communication
FC, iSCSI, NFS, HTTP

Storage Processor
(RAID Level, Tiering)

Block IO (HDD/SDD)

- Verschiedene Konzepte und Lösungsansätze für verschiedene Anforderungen
- Es gibt kein Allheilmittel
- Genaue Analyse der jeweiligen Anforderungen
- Mehrere Lösungen je nach Anforderung
- Verlagerung der Intelligenz in verschiedene Ebenen

VIELEN DANK!