World Tour 2015

HP Big Data Reference Architecture

Philipp Koik, Strategic Presales
Jochen Mohr, Technology Services

5. Mai, 2015

IT infrastructures trends

How they influence and are influenced by Big Data & Analytics transformation

Server
- Shift from high-end to medium-/low-end servers
- Massive Parallel Processing (MPP), Scale-out computing architecture

Storage
- Shift from high-end centralized storage to local DAS
- Shared-nothing
- Scale-out storage architecture

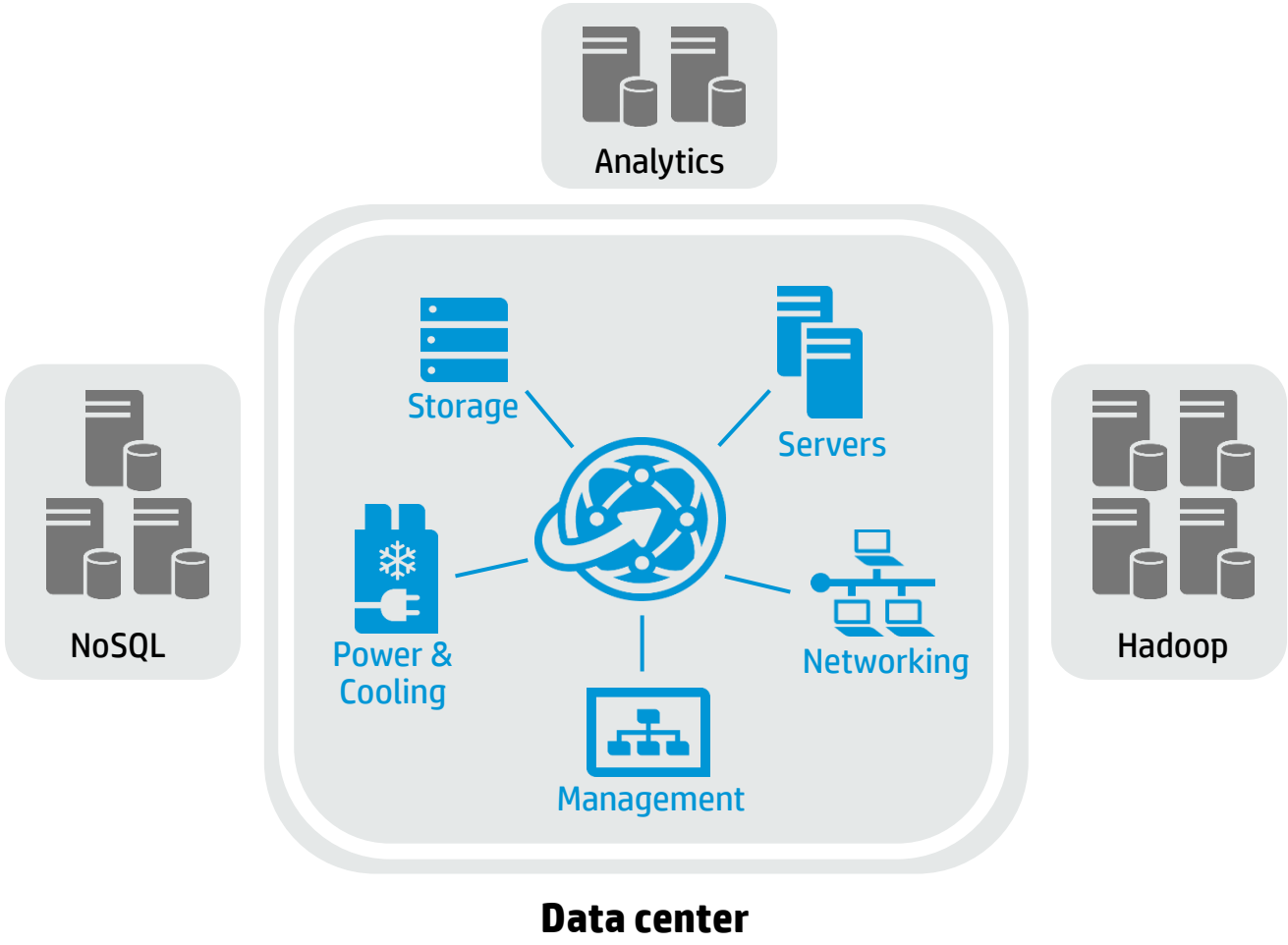
Big Data

Network
- Dramatic increase in fabric speeds and bandwidth demand
- Access from anywhere at anytime
- LAN, WAN, Internet access
- SDN

Software/Application
- Clustering
- NoSQL, NewSQL , Columnar DBs
- BI & BA working with structured & unstructured data
- RT-Analytics
- Data life-cycle Management

Security | **Configuration** | **Monitoring** | **Management** | **BuRA**

IT infrastructures must evolve to handle Big Data demands



Challenges

- Multiple silos with multiple copies of the same data
- Difficult to standardize on a consistent server architecture
- Less elastic than other virtualized or converged infrastructure
- Large scale makes density, cost and power problematic

HP Big Data – Hardware Solution Portfolio

Market-driven offerings and services

ConvergedSystem

HP ConvergedSystem for Big Data

Reference Architecture

cloudera

MAPR TECHNOLOGIES

Hortonworks

DL180

DL380

SL4540

HP Moonshot

HP Apollo

HP ConvergedSystem 300 for Microsoft Analytics Platform

The only appliance with integrated in-memory performance, MPP DW and Hadoop

- Next-generation data warehouse for mission-critical environments
- Factory built, appliance-based on HP Converged Infrastructure
- Pre-loaded with Microsoft software, integrated, tested, and tuned
- Architecture chosen for best data warehouse performance
- Single view of information across the enterprise
- New addition to the Converged Systems “Sharks” family

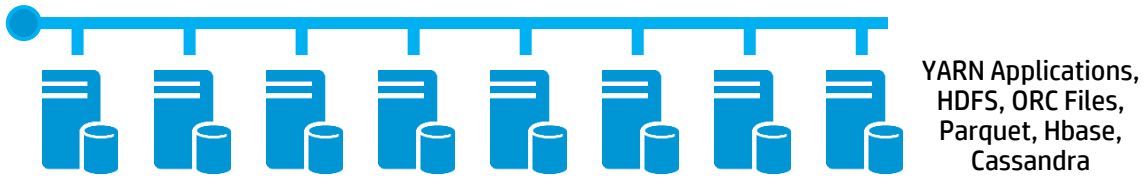
100X	30%	50%
Faster query speed[1]	Better scan rate[1]	Lower cost per TB[2]

[1] Than previous generations
[2] Than competitive offerings

New approach to address Big Data demands

Current traditional Big Data approach

- Compute and storage are always collocated
- All servers are identical
- Data is partitioned across servers on direct-attached storage (DAS)

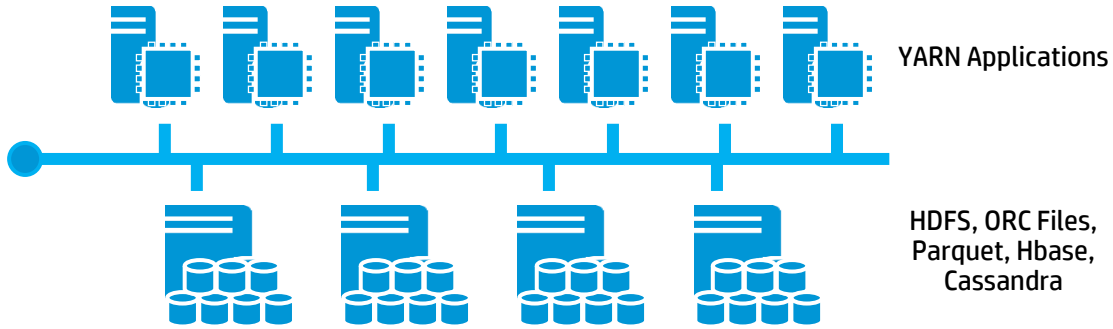


Two Socket, 2U Servers

New HP Big Data approach

- Separate compute and storage tiers connected by Ethernet networking
- Standard Hadoop installed asymmetrically with storage components on the storage servers and yarn applications on the compute servers

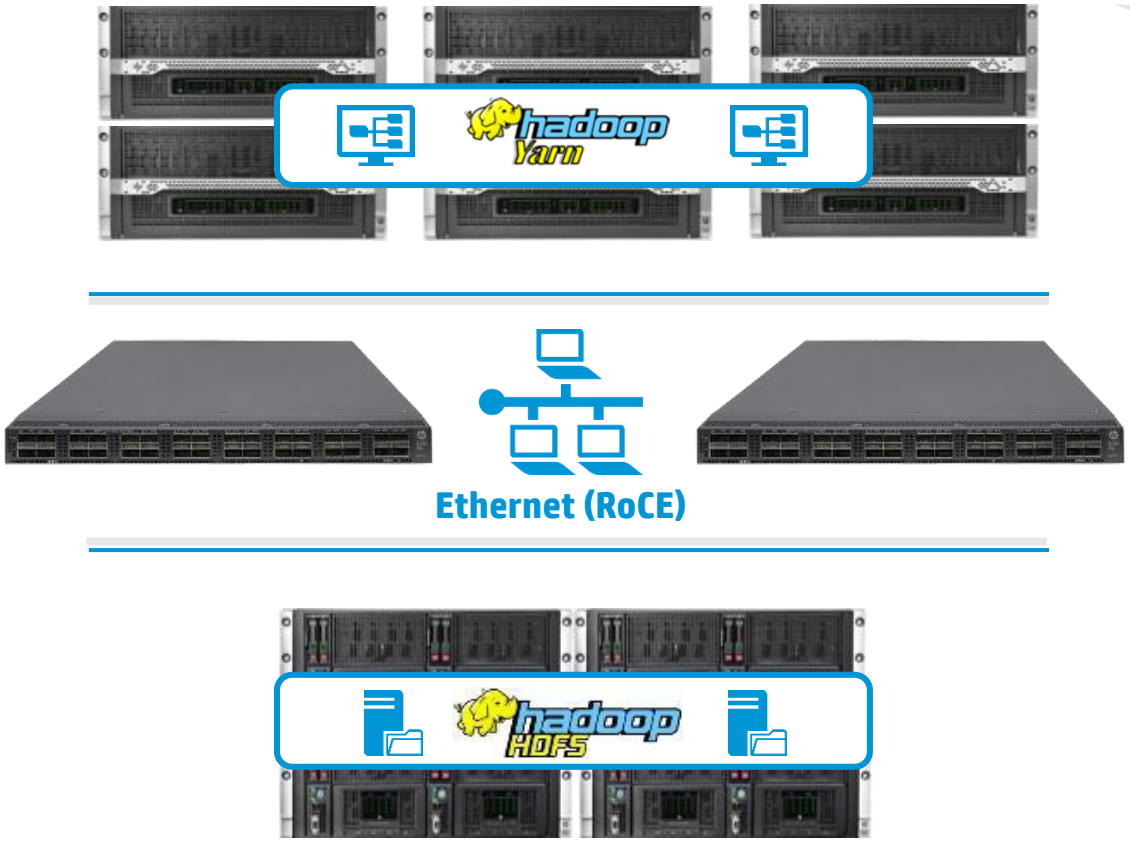
Compute Optimized Servers



Storage Optimized Servers

Benefits of HP Big Data Reference Architecture

HP Moonshot and SL4540 addresses a variety of enterprise big data needs



Cluster consolidation
Multiple big data environments can directly access a shared pool of data

Flexibility to scale
Scale compute and storage independently

Maximum elasticity
Rapidly provision compute without affecting storage

Breakthrough economics
Significantly better density, cost and power through workload optimized components

Building blocks for the HP Big Data Reference Architecture

HP Moonshot System

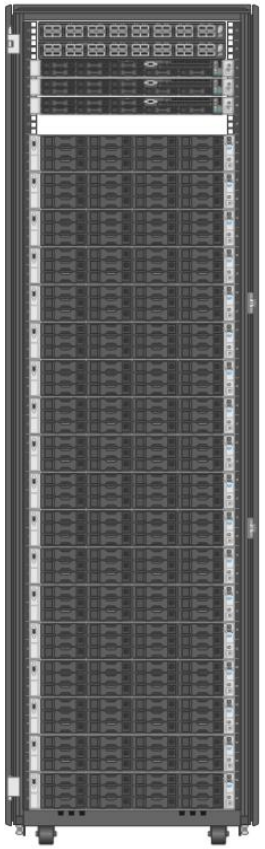
A complete server system engineered for specific workloads and delivered in a dense, energy-efficient package

HP ProLiant SL4540 Scalable System

A cost-effective industry standard storage server purpose built for big data with converged infrastructure that offers high density energy-efficient storage

Independent Scaling of Compute and Storage

Traditional Architecture



HP Big Data Reference Architecture

4x compute
60% of the storage capacity
72% of the Hadoop IO

1.7x compute
1.5x the storage capacity
2.1x the Hadoop IO

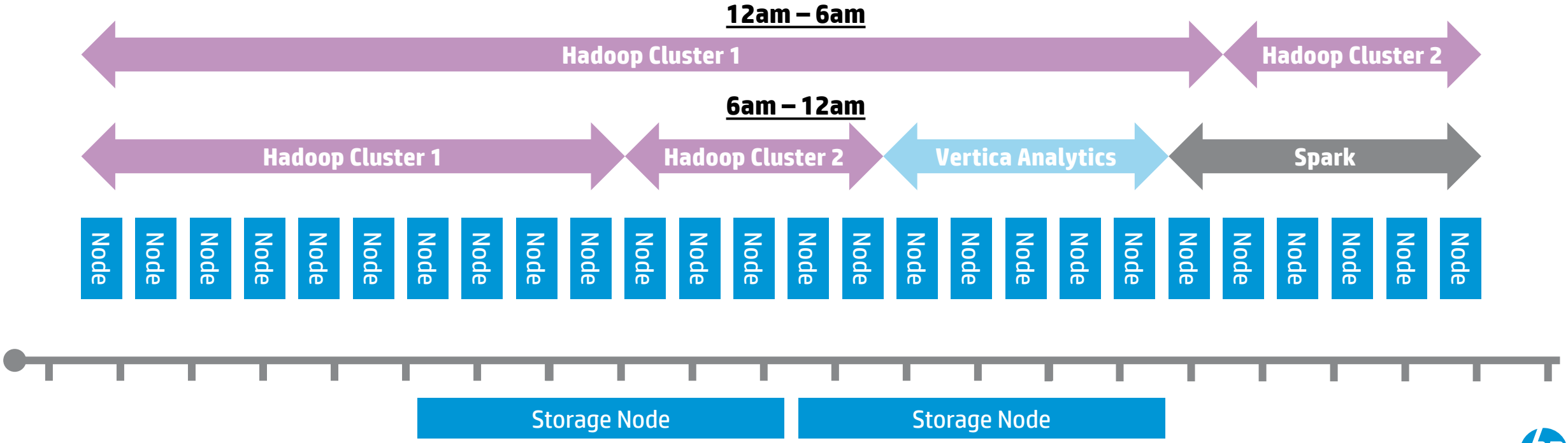
60% of the compute
2x the storage capacity
2.9x the Hadoop IO

Compared with traditional architecture, full rack

Maximum Elasticity for Big Data workloads

Hadoop Labels feature (jira YARN-796)

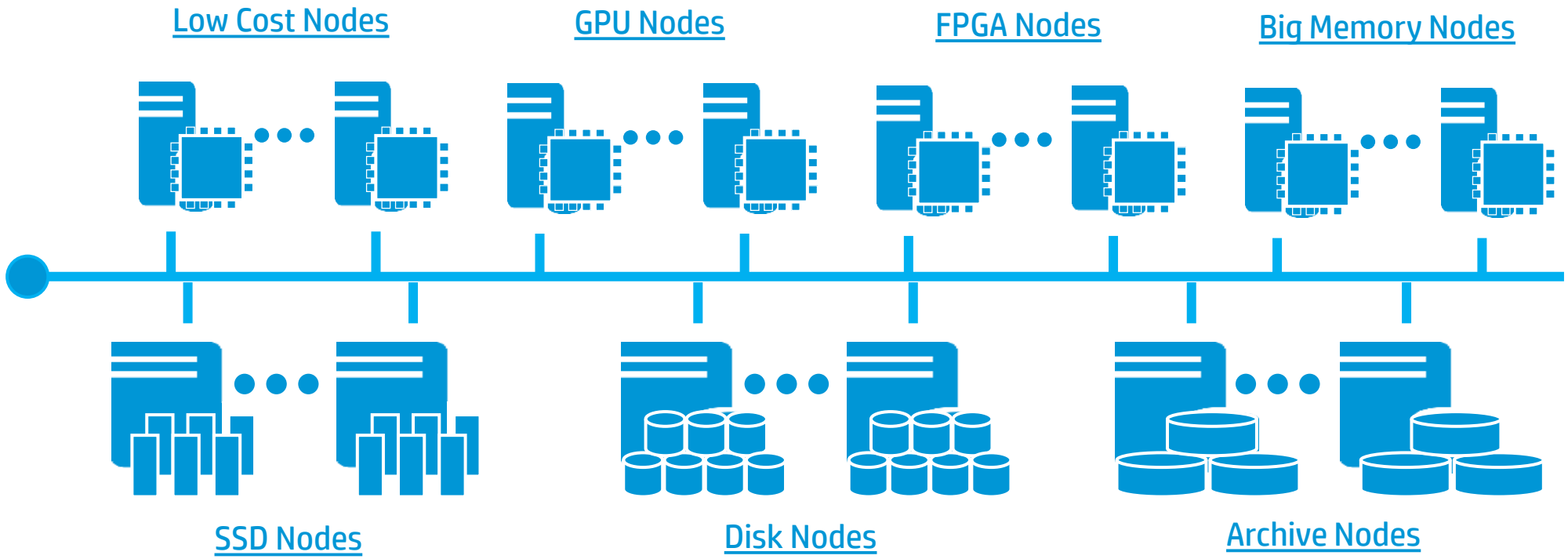
- HP contributed IP into the Hadoop trunk, working with Hortonworks
- Specifying labels on nodes allows for scheduling of YARN containers to specific pools of nodes
  - Admins able to target workloads at optimized platforms
- Combined with the HP Big Data Reference Architecture, compute nodes can be dynamically assigned
  - No data repartitioning



Minotaur CI for Big Data long term view

Evolve to support multiple compute and storage blocks

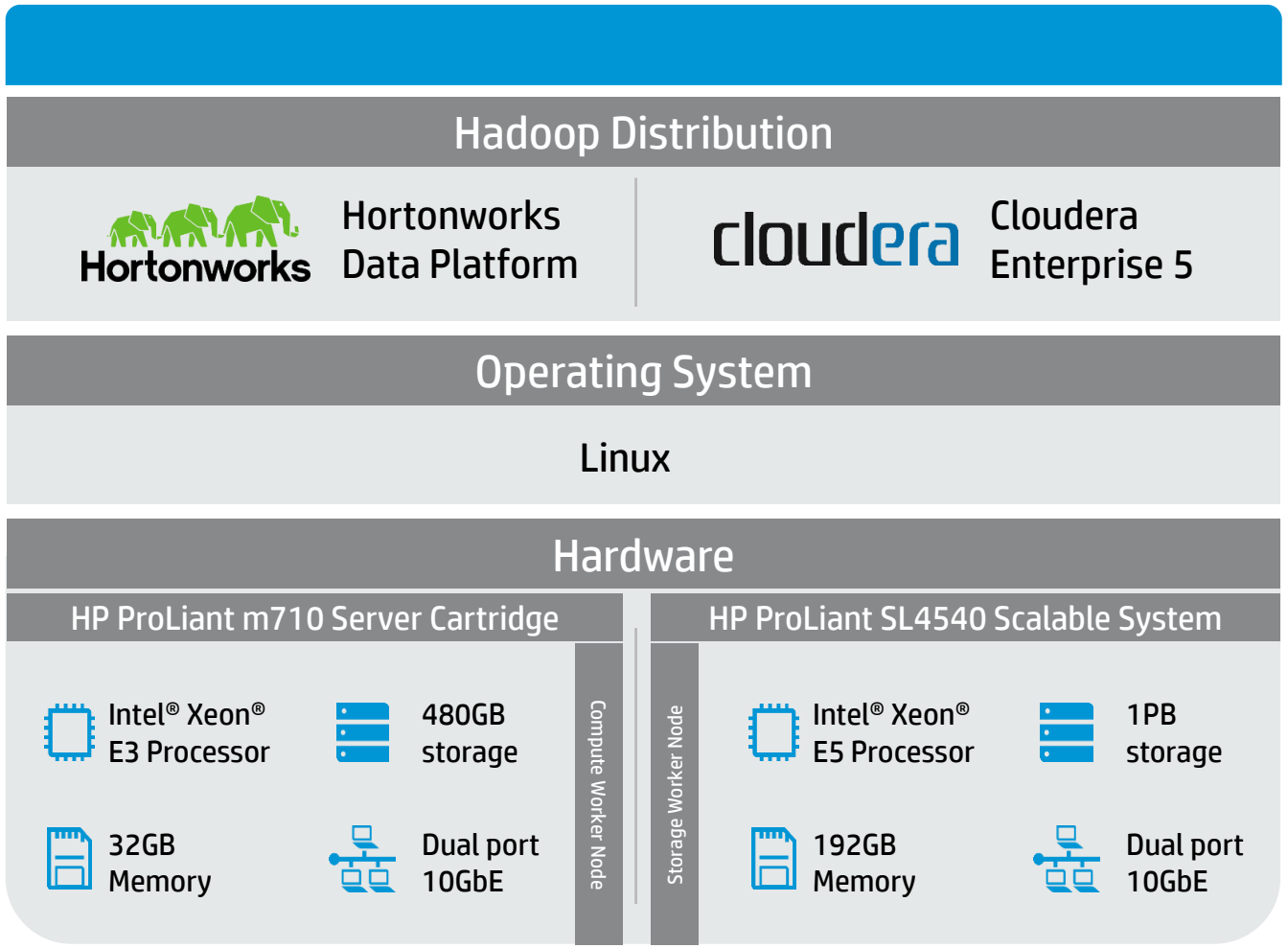
Workload Optimized compute nodes to accelerate various big data software



Multi-temperate Storage using HDFS Tiering, NoSQLs and Objectstores

HP Moonshot System + SL4540 for Big Data

HP Big Data Reference Architecture



Built on Standard Hadoop Distributions

No proprietary software. Leverage the latest versions of Hadoop and consumer plug-ins

Optional HP Consulting Services

Expedite the sizing and configuration of the infrastructure through the Hadoop Reference Architecture implementation service

Optional Factory Build

Hardware racked, wired and tested, delivered to your data center

HP BD Service – An exhaustive Big Data IT portfolio

IT Consulting Services for Big Data Service Delivery Transformation

Advise

Big Data IT strategy and architecture services

- Big Data Strategy Workshop
- Big Data Infrastructure Transformation Experience Workshop
- HP Enterprise Planning for HAVEn
- Big Data Protection and Compliance Analysis
- HP Vertica Deployment Roadmap

Enabling provisioning of Big Data Services to your customers

Transform

Big Data system infrastructure

- HP professional services for HAVEn solutions implementation
- Enterprise Design Service for Hadoop
- Reference Architecture Implementation Service for Hadoop, Microsoft PDW and SAP HANA
- HP Vertica Implementation Accelerator

Big Data protection

- Data Loss Prevention

Helping accelerate adoption and integration of Big Data technologies

Manage

Big Data operation

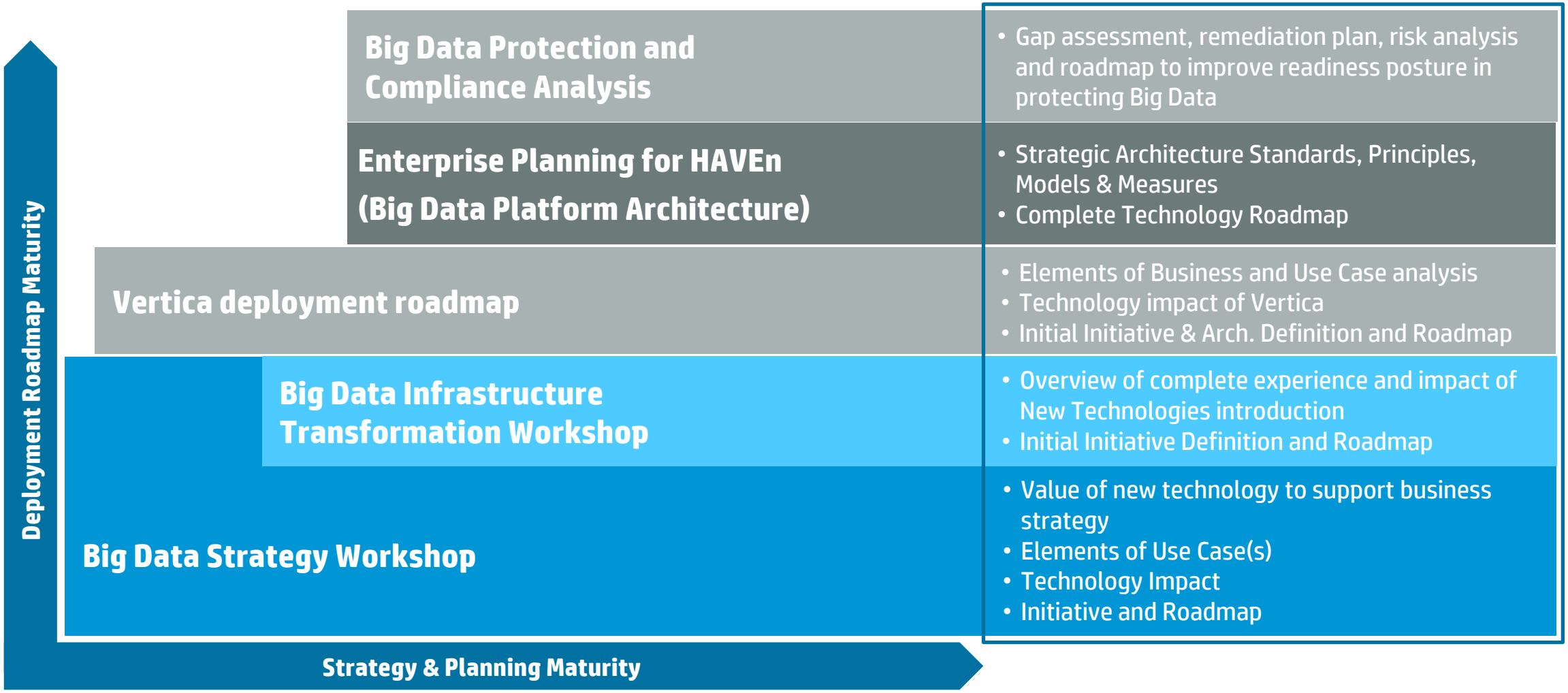
Achieve best-in-class operational efficiency of a client's big data environment leveraging the unique knowledge of HP experts and our global infrastructure

Big Data education

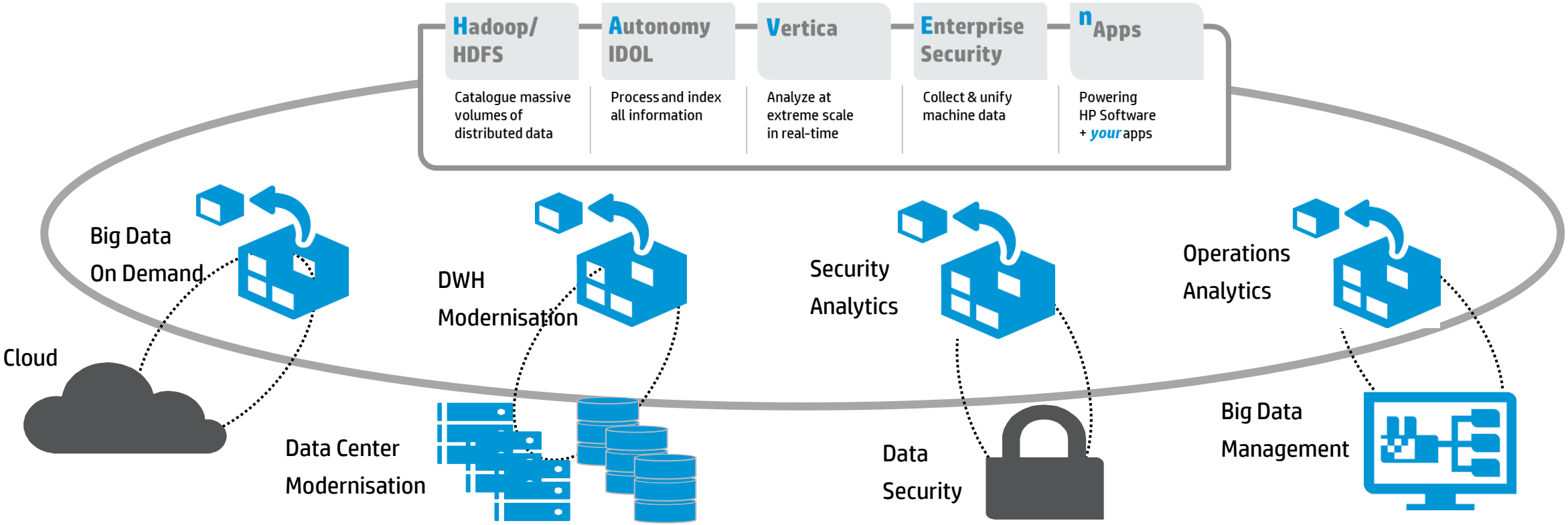
Train and certify a client's IT staff and third-party partners to help them architect, integrate, and administer Big Data solutions. Assist Management of Change

Supporting IT transformation

Big Data Advisory Services Positioning



IT infrastructure modernization & consolidation



Rationales: Standardization, Costs, Control, Monitoring, Elasticity, Data Security, ...

Challenges: In-House-Dev, Performance, Security, Network Integration, Volume, Management Tools, Non-Standard HW, Risks, Interfaces, Backup- and Restore, ...

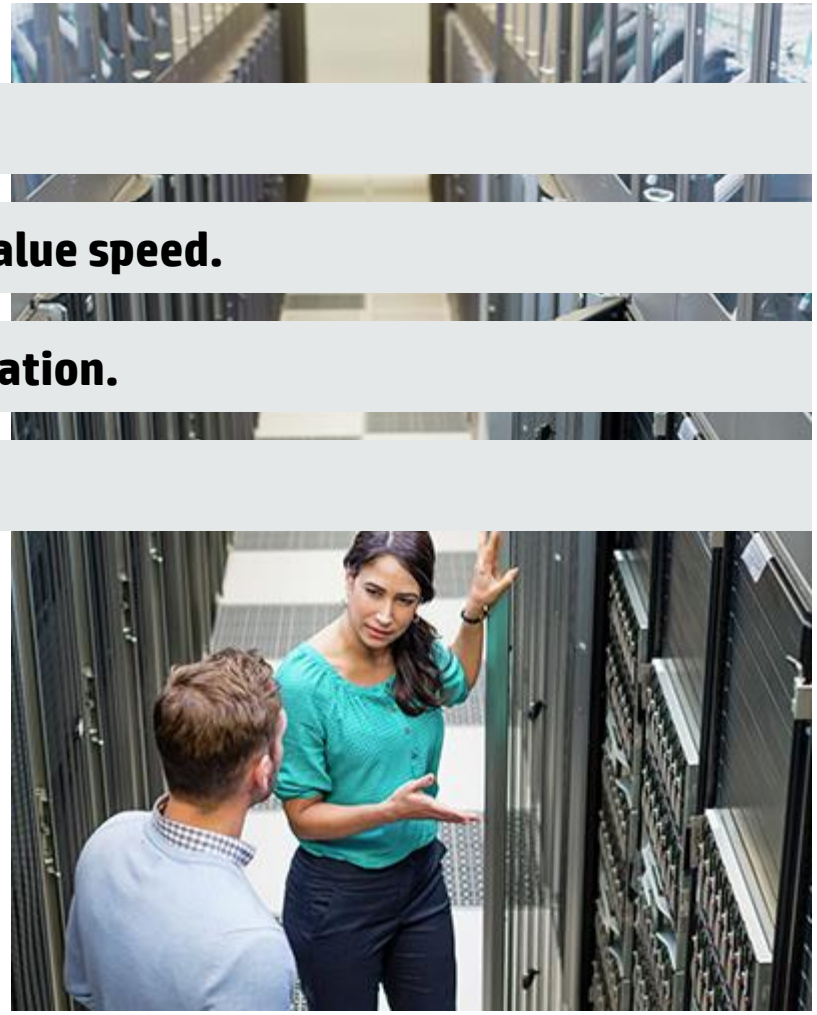
Big Data Technology Consulting solution value

We provide leadership to IT to help achieve business objectives.

We minimize implementation and integration risk, improving time-to-value speed.

We facilitate IT to ramp up on skills to manage transition and transformation.

We will accelerate adoption and integration of Big Data technologies.



Rekordjäger Rainer Zietlow

THANK YOU

BACK-UP

HP Moonshot 1500 Chassis front and rear view

5 hot-plug fan modules

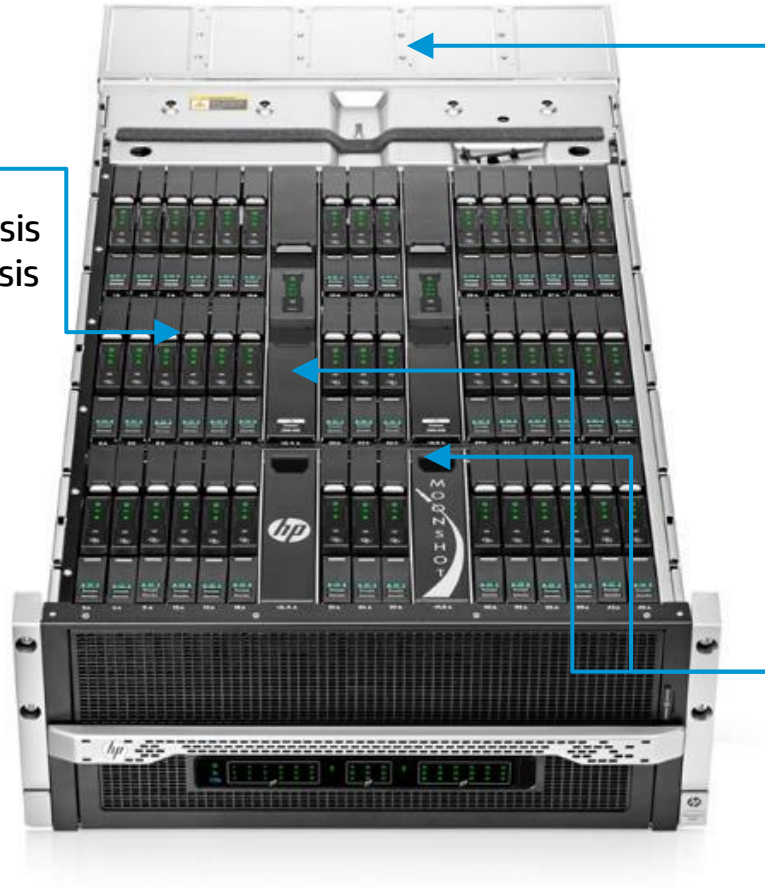
HP Common-Slot Power Supplies

45 hot-plug cartridges

- Single-server = 45 servers per chassis
- Quad-server =180 servers per chassis

Processor

- x86 , ARM, or Accelerator



Dual Network Uplinks

- HP Moonshot-6SFP Uplink Module (6 x10Gb Stackable Uplinks)
- HP Moonshot-4 QSFP Uplink Module (4 x 40Gb Stackable Uplinks)

Dual low-latency switches

- HP Moonshot-45G Switch Module (45 x1Gb downlinks)
- HP Moonshot-180G Switch Module (180 x1Gb downlinks)
- Moonshot-45XGc Switch Module
- (45 x10Gb downlinks)

Big Data Compute Node

HP ProLiant m710 Server Cartridge

CPU	Intel Xeon E3-1284Lv3 with Iris Pro P5200 GPU 4 core / 1.8 GHz (3.4Ghz Turbo) / GPU + 128MB eDDR
MEMORY	Total of 32GB of ECC protected memory, dual-memory channels (4) 8GB LV SO-DIMMs at 1600MHz with (8) embedded DRAM for ECC protection.
NETWORK	Integrated NIC: dual port 10GbE Mellanox CX3 PRO Supported Switch(s): 45 port 10Gb Downlinks, (4) 40GbE QSFP uplinks
STORAGE	Local SSD boot, 480GB m.2 (2280)
POWER	Cartridge: <69W
OS	Ubuntu 14.04 w/KVM, RHEL 6.5,7.0 w/KVM, SLES 11 SP3 w/KISO/KVM, Windows Server 2012 R2, CentOS 6.5, 7.0 Intel Media SDK (media libraries, OpenCL in beta), purchased direct from Intel

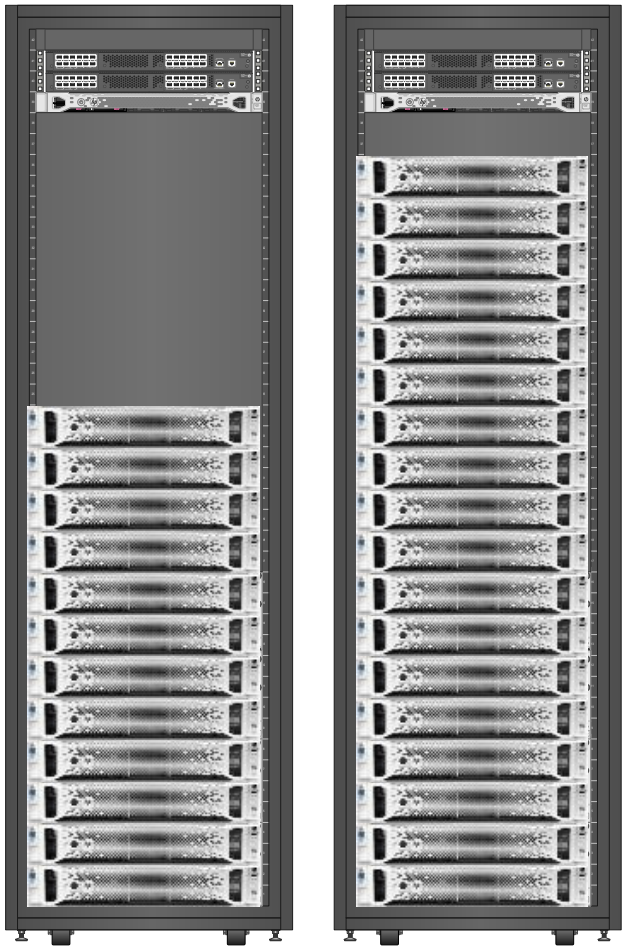
Big Data Storage Node

HP ProLiant SL4500 Scalable System

CPU	Up to two Intel® Xeon® E5-2400 or E5-2400v2 (4, 6, or 8 core) per node	
MEMORY	12DIMMs – Up to 32GB's, 384GB's max 1333/1066 MHz DDR3 RDIMM	
NETWORK	HP Ethernet 1Gb 2-port 361i OR HP Ethernet 1Gb 2-port 361i and HP Ethernet 10Gb 2P 544i Adapter; One 10GbE SFP+ connector; One 10GbE/40GbIB QSFP connector (option to be converted to Infiniband)	
STORAGE	25x 3.5" SAS, SATA, or SATA SSD (hot-plug) 2x 2.5"SATA for boot	
POWER	4x 750W or 1200W hot plug, redundant optional, Platinum Plus supplies	
OS	MS Windows Server 2008 SP2, R2 w/ SP1 (standard, enterprise, datacenter, web server, HPC, embedded), Hyper-V R2 SP1. (64bit only) MS Windows Server 2012 (standard, datacenter, hyper-v, HPC pack) MS Windows Server 2012 R2 (standard, datacenter, hyper-V	RHEL 5.8, 5.9, 5.10 (64bit) RHEL 6.2, 6.3, 6.4 (64bit) SLES11SP2, SP3 (64bit) Ubuntu 12.04.03 LTS VMware ESXi MN 5.0 U3, 5.1 U2, 5.5

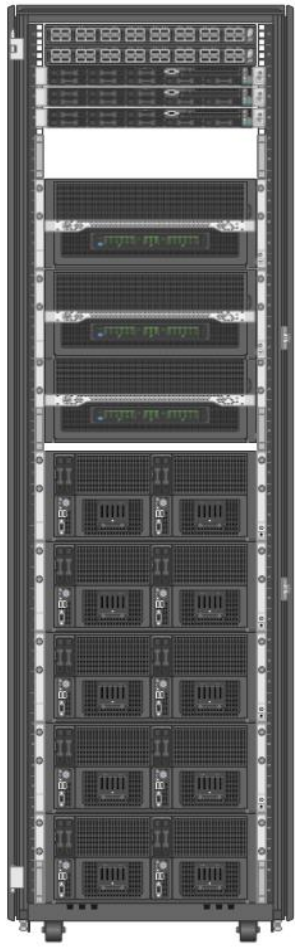
Advantages* of HP Big Data Reference Architecture

A new standard for Big Data delivery at scale



Traditional Architecture

Traditional Hadoop Architecture	HP Big Data Reference Architecture
CPU Performance (SpecINT)	22% better
Data Capacity (TB)	Equivalent
Memory (GB)	20% greater
Power (W)	25% less
Rack Space	60% less
TCO/Performance ($ per MB/s)	**15% better**

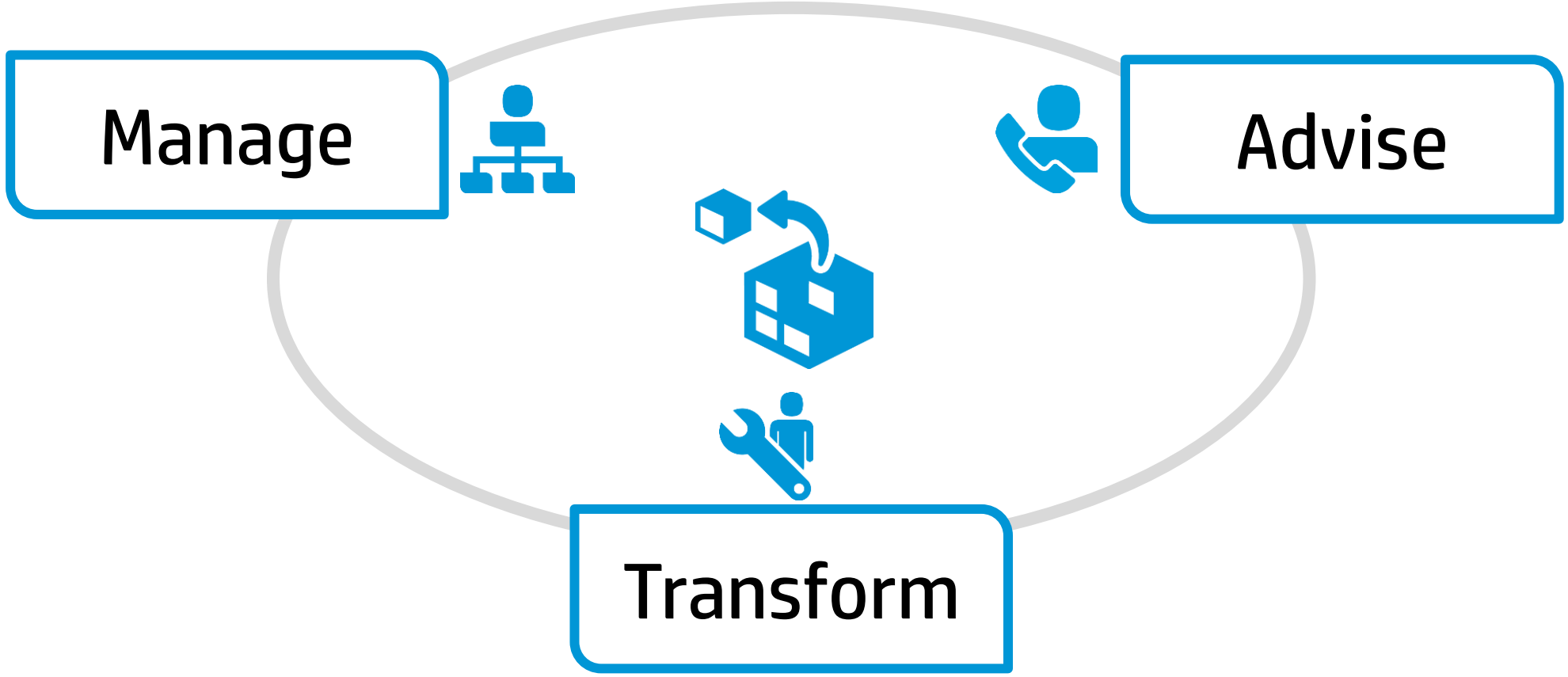


HP Big Data Reference Architecture

* Normalized on capacity

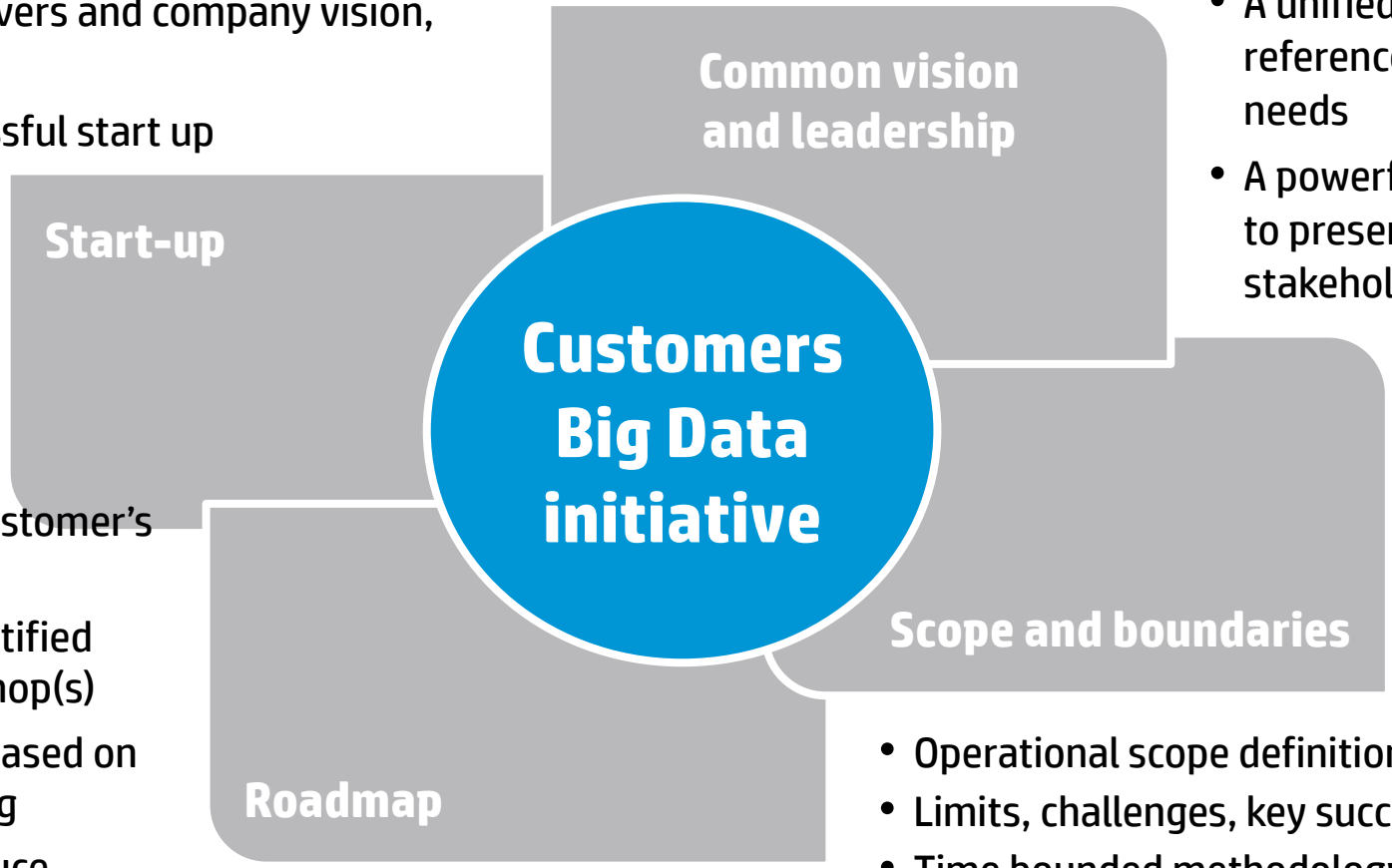
HP TS Consulting Big Data Service Portfolio

An exhaustive Big Data IT portfolio



Benefits & Gains: Value customers will get

- Correlation of business drivers and company vision, use cases and issues
- Quick wins to boost successful start up
- Common, shared strategies



- A unified transformation reference model aligned to business needs
- A powerful and structured tool to present the initiative and gain stakeholders' buy-in

- Initiatives unique to your customer's requirements
- Actions that have been identified and captured during workshop(s)
- Recommended next steps based on HP's experience and offering
- Pragmatic initiatives to reduce risk/cost etc.
- Additional considerations to improve efficiencies

- Operational scope definition (AsIs vs. ToBe)
- Limits, challenges, key success factors
- Time bounded methodology helps focusing on key facts
- Their target vision
- Gaps in their ability to transform

Start your journey with a Transformation vision

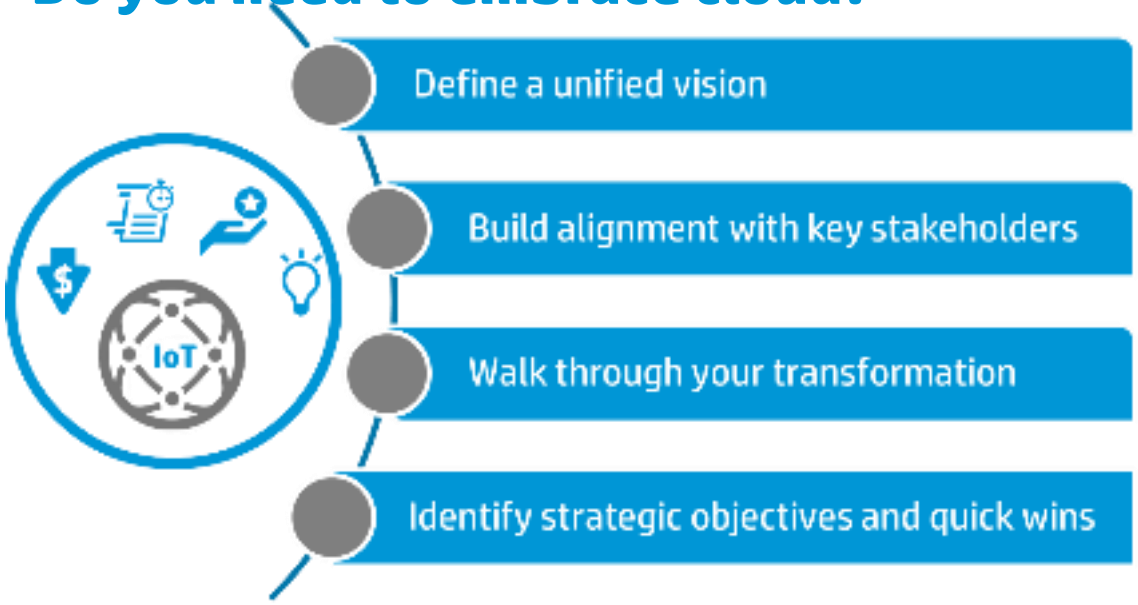
Identify your path, your use case

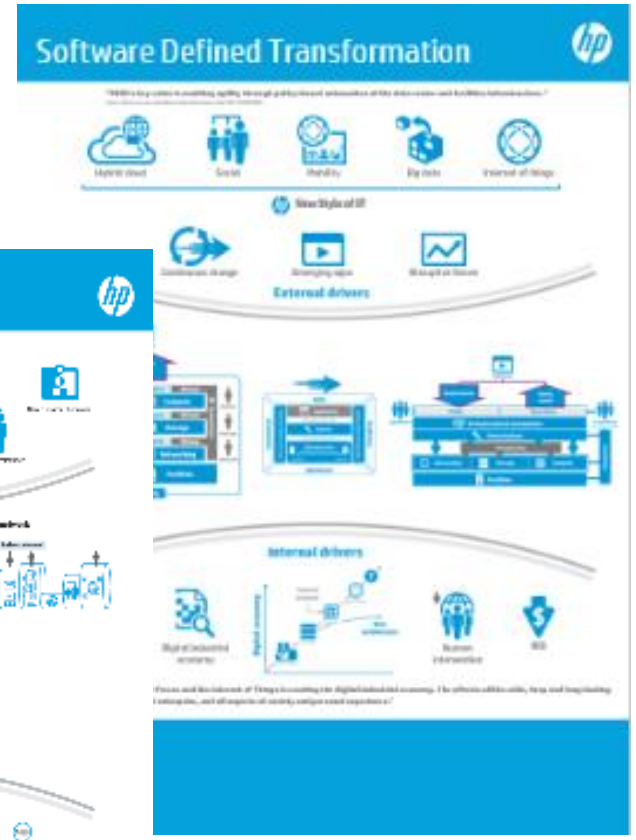
Do you need to stabilize and secure your existing environment?

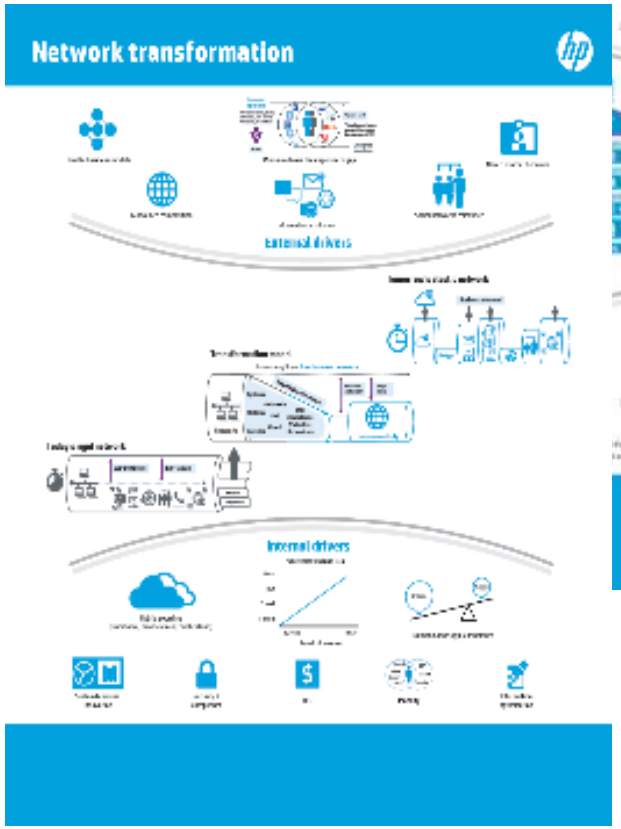
Do you need to transform your network?

Do you need to modernized your data center?

Do you need to embrace cloud?







Thank you