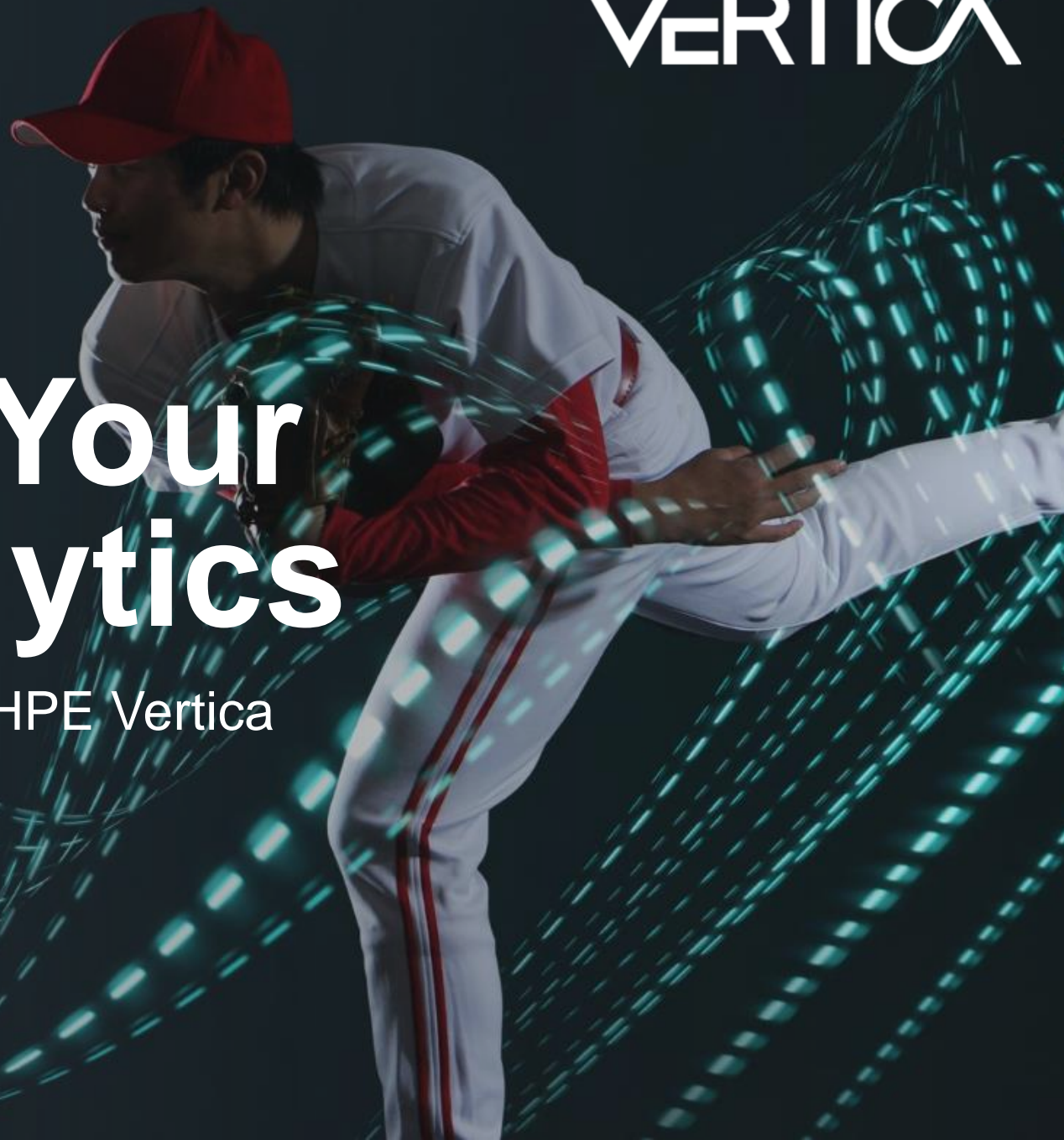


**Hewlett Packard
Enterprise**

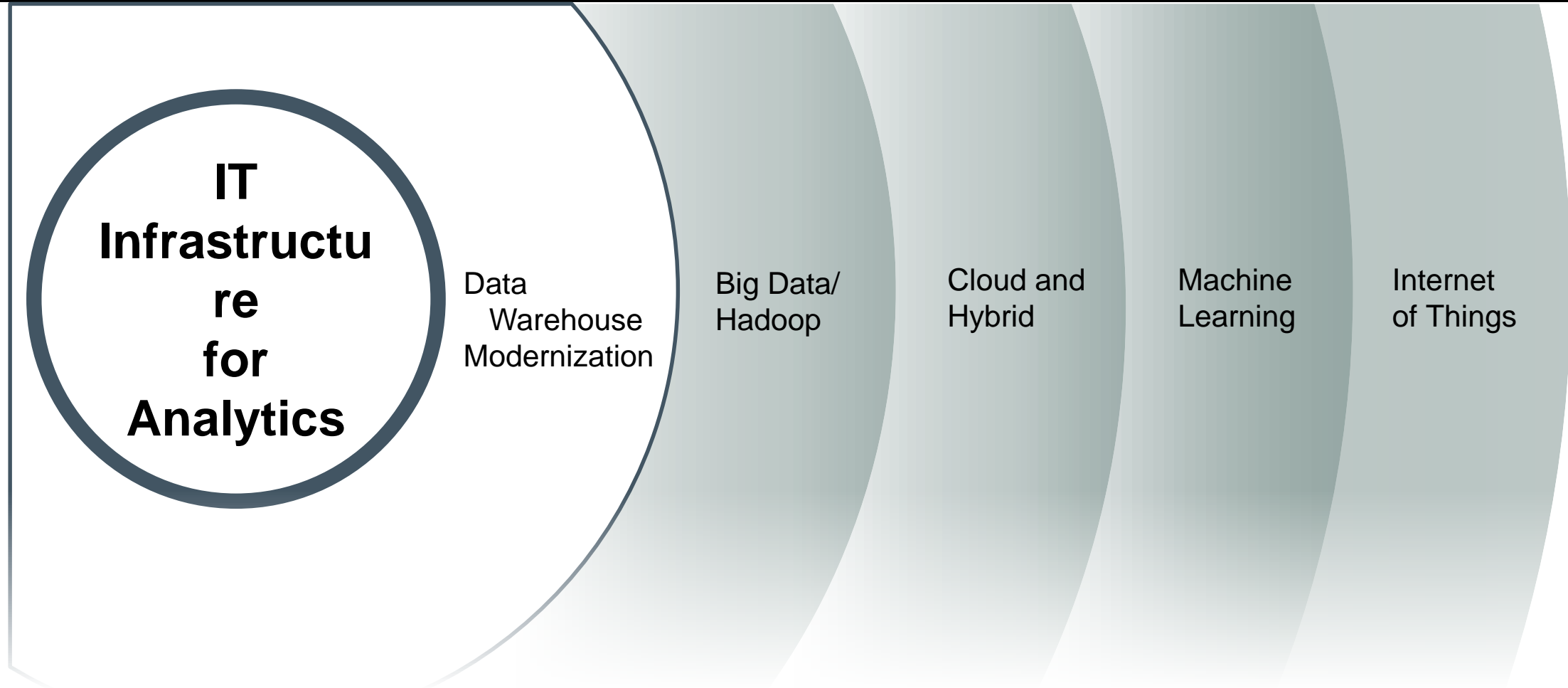
VERTICA

Accelerating Your Big Data Analytics

Jeff Healey, Director Product Marketing, HPE Vertica



Recent Waves of Disruption



Replacing Legacy Solutions that Can Better Scale

VERTICA

**Increasing
data volumes**

**More
concurrent
users**

**High cost of
legacy solutions**

**Analytics
demand
increase**

Legacy Analytical Platform

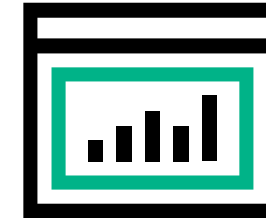


Raw data



Connected Systems

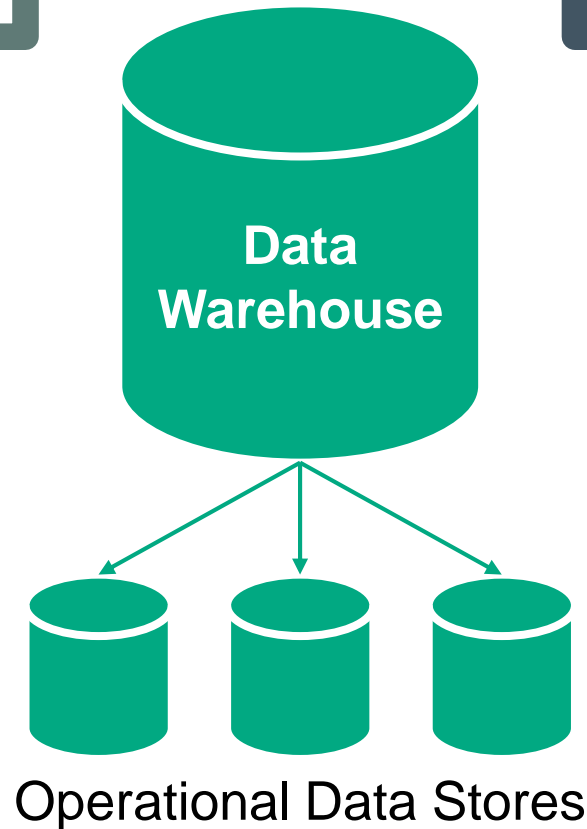
Meaningful Information



Information Consumers

Benefits:

- It's mature and it works, as long as I manage data and users
- I know it well and have staff on hand to run it
- The analytical functions are mature
- It would be costly to replace



Limitations:

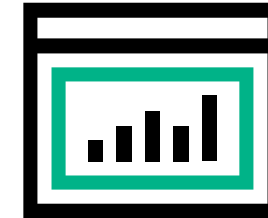
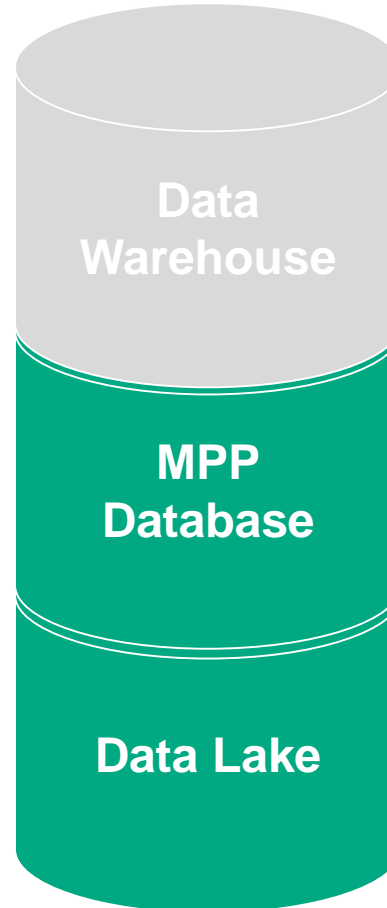
- It's expensive and becomes more so as I increase my data
- It has real limitations on data size, number of users
- My users want more access to analytics, but resources limited
- I spend a lot on consultants who are tuning it

Modernized Analytical Platform

VERTICA



Connected Systems



Information Consumers

Benefits:

- Lowers cost by delivering analytics with matching SLA
- Faster analytics and more users supported with MPP
- Data science supported

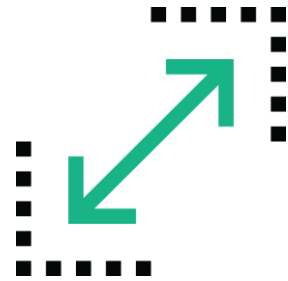
HPE Vertica Analytics Platform

VERTICA



Fast

Boost performance
by 500% or more



Scalable

Handles huge workloads
at high speeds



Standard

No need to learn new
languages or add complexity



Costs

Significantly lower cost
over legacy platforms

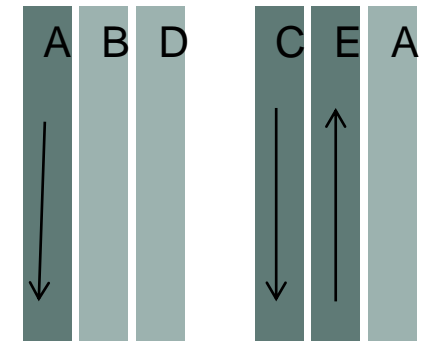
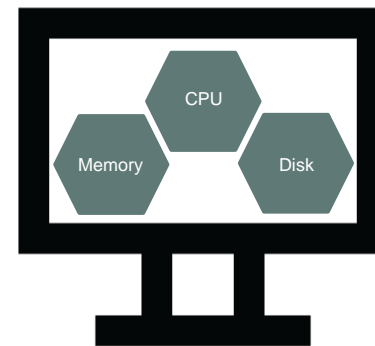
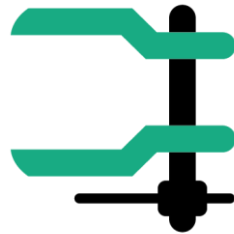
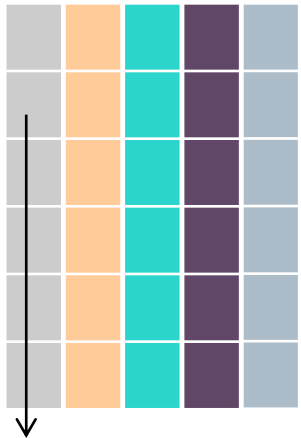
Freedom to Deploy Anywhere – On Commodity Hardware, Across Multiple Cloud Platforms, and Natively on Hadoop

VERTICA



Secrets to Achieving Performance Increases

Columnar Storage	Compression	MPP Scale-Out	Distributed Query	Projections
Speeds Query Time by Reading Only Necessary Data	Lowers costly I/O to boost overall performance	Provides high scalability on clusters with no name node or other single point of failure	Any node can initiate the queries and use other nodes for work. No single point of failure	Combine high availability with special optimizations for query performance

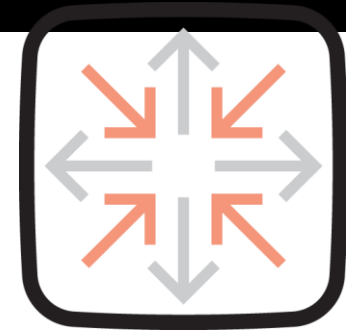
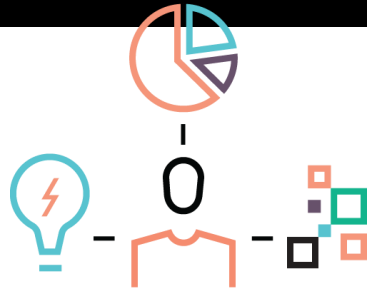


The Appeal of Vertica



Requirement	Proof
Extreme Optimization	<ul style="list-style-type: none">• Columnar design for high performance analytics• Aggressive compression• Scalable to petabyte scale
Total Cost of Ownership	<ul style="list-style-type: none">• Simply and predictable pricing• No penalty for additional hardware or connected users
Ready for your Enterprise	<ul style="list-style-type: none">• SQL compliant to 100% of the TPC-DS benchmark queries• Secure and ACID compliant• No single point of failure
Open and Compatible	<ul style="list-style-type: none">• Open platform – Standards compliant SQL, Python, Java• Working with open source community on Spark, Hadoop, Kafka, etc.

Advanced, In-Database Analytics



SQL '99

- Aggregate
- Analytical
- Window functions
- Graph
- Monte Carlo
- Statistical
- Geospatial

Allows for:

- Standard functionality that performs at scale

SQL Extensions

- Pattern matching
- Event series joins
- Time series
- Event-based windows

Allows for:

- Sessionization
- Conversion analysis
- Fraud detection
- Fast Aggregates (LAP)

SDKs

Analytics

- Java
- C++
- R

Connection

- ODBC/JDBC
- HIVE
- Hadoop
- Flex zone

Allows for:

- Machine learning
- Custom data mining
- Specialized parsers

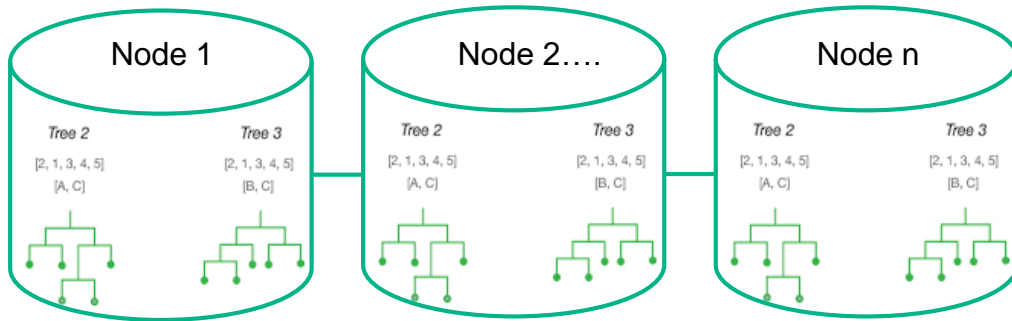
In-database Analytics

- Regression testing
- K-means
- Statistical modeling
- Classification algorithms
- Page rank
- Text mining
- Geospatial

Allows for:

- Statistical modeling
- Cluster analysis
- Predictive analytics

Building Machine Learning into the Core of Vertica



Machine learning functions run in parallel across hundreds of nodes in a Vertica cluster

- Machine Learning algorithms, such as k-means and regression, built into the core of Vertica
- Advanced predictive modeling runs within the database eliminating all data duplication typically required of alternative vendor offerings
- Traditional approaches can't handle many data points forcing data scientists to "down-sample" leading to less accurate predictions
- A single system for SQL analytics and Machine Learning

Why It's an Easy Transition for Analysts

VERTICA

Languages Used

ETL Used

Data Visualization Used

SQL



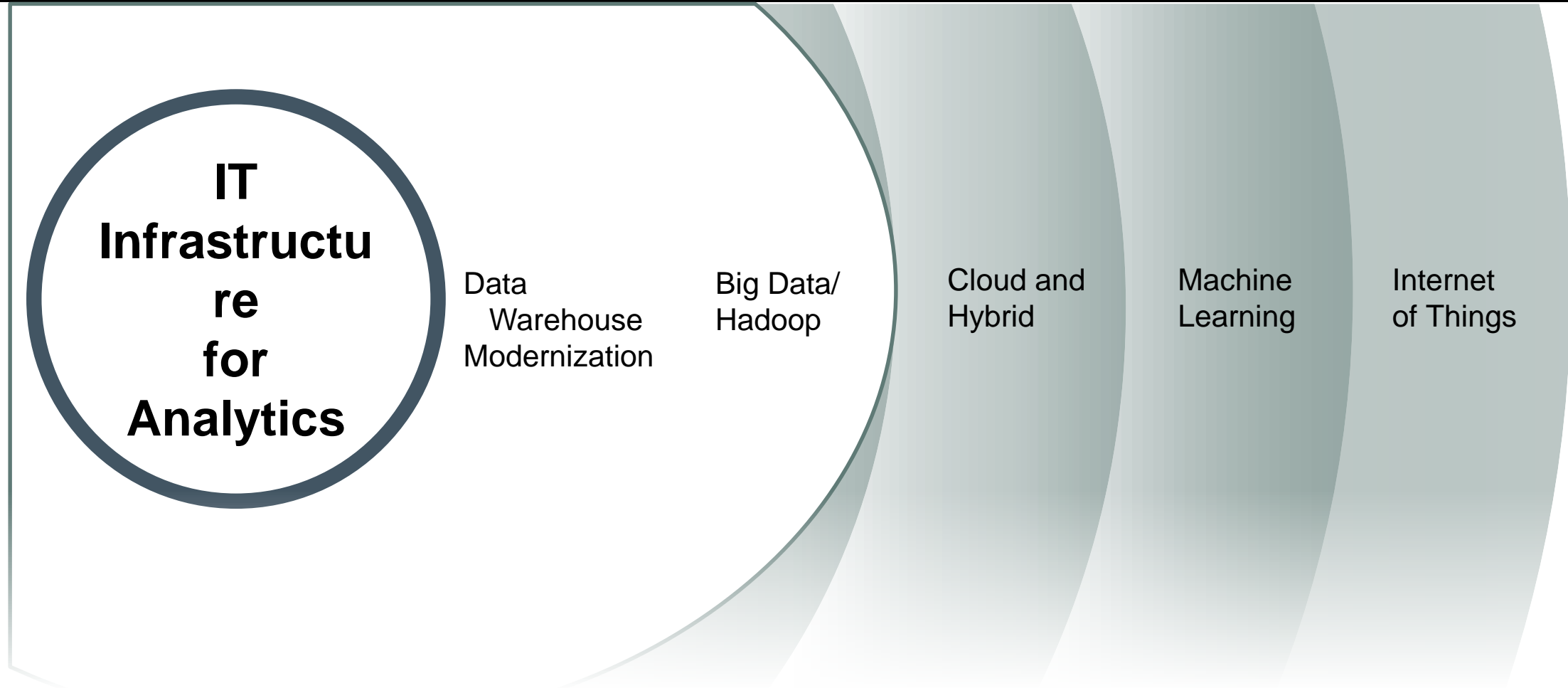
Python



Java



Recent Waves of Disruption



Lowering Costs of Traditional Architectures with Open Source

VERTICA



More types of data

Unstructured data

Storing data of mixed value

The Hadoop bandwagon

Open source economics

Hadoop Hype

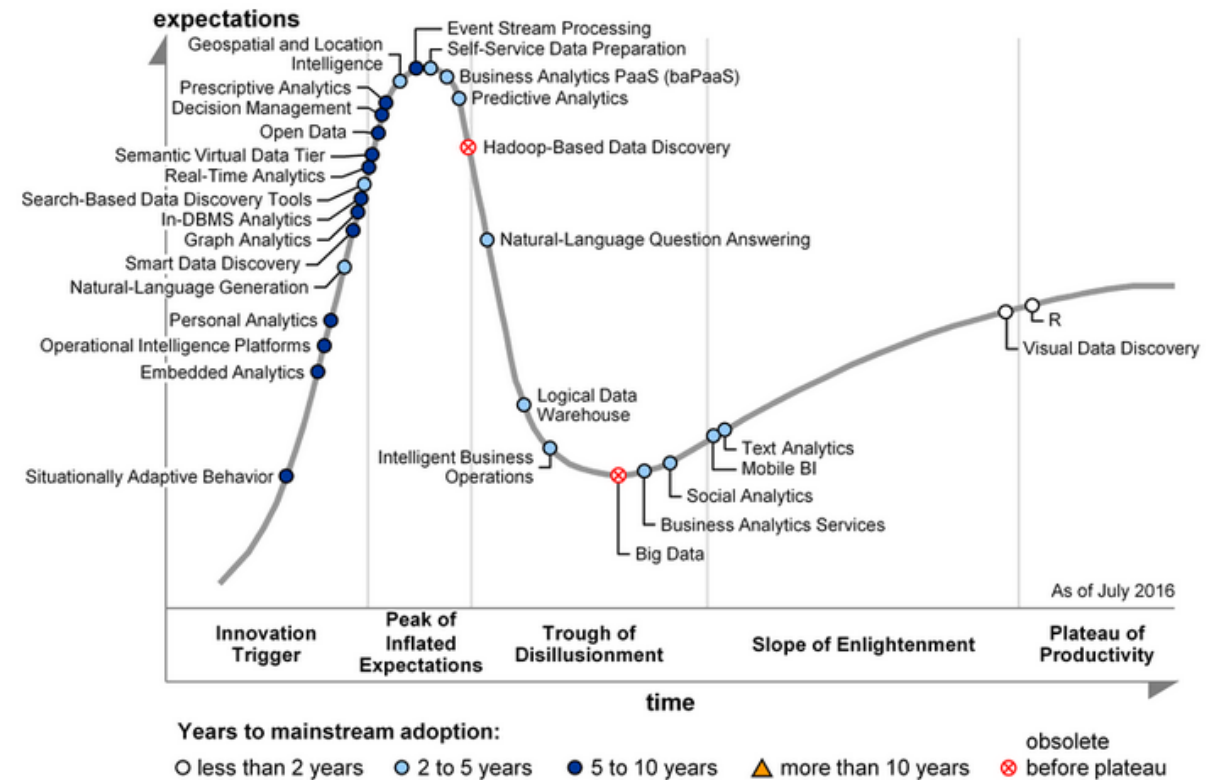
2008 There's this new thing called Hadoop

2011 Can Hadoop replace data warehouse?

2013 Can Hadoop replace ETL?

2017 Has Hadoop failed to deliver the goods?

Figure 1. Hype Cycle for Business Intelligence and Analytics, 2016



Source: Gartner (July 2016)

Hadoop Status

“A smoking heap of cost and complexity”

“...not the technology base the world will be built on going forward.”

datanami
• BIG DATA • BIG ANALYTICS • BIG INSIGHTS •

Home About Whitepapers Events Subscribe

HOME FEATURES ▾ SECTORS ▾ APPLICATIONS ▾ TECHNOLOGIES ▾ VENDORS ▾

March 13, 2017
Hadoop Has Failed Us, Tech Experts Say
Alex Woodie



The Hadoop dream of unifying data and compute in a distributed manner has all but failed in a smoking heap of cost and complexity, according to technology experts and executives who spoke to *Datanami*.

“I can’t find a happy Hadoop customer. It’s sort of as simple as that,” says Bob Muglia, CEO of Snowflake Computing, which develops and runs a cloud-based relational data warehouse offering. “It’s very clear to me, technologically, that it’s not the technology base the world will be built on going forward.”

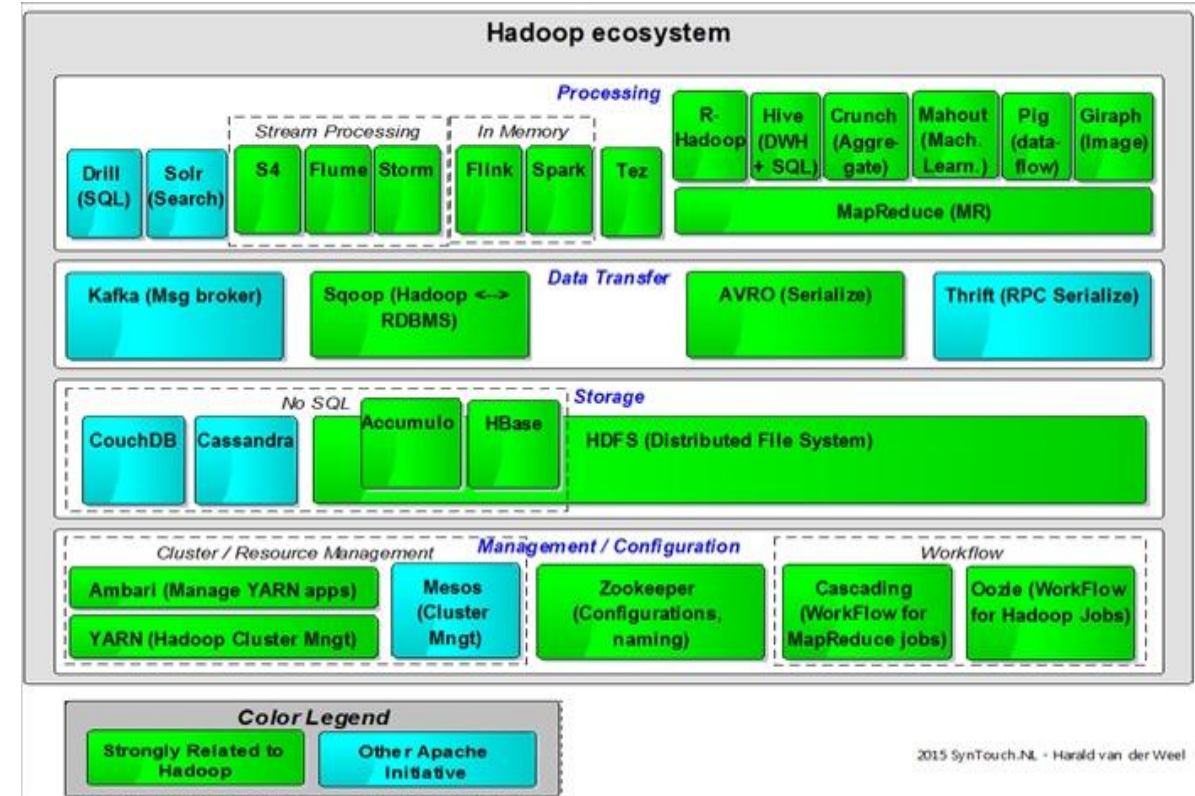
Thousands of organizations store huge amounts of data in Hadoop, and so Hadoop won’t disappear overnight. After all, many companies still run mainframe applications that were originally developed half a century ago. But thanks to better mousetraps like S3 (for storage) and Spark (for processing), Hadoop will be relegated to niche and legacy statuses going forward, Muglia says.

What Went Wrong?

You still need a traditional data warehouse

- Concurrency
- Completeness of analytics
- Standards not followed: ACID, SQL
- Hadoop isn't fast for analytics

Hadoop is difficult to use, requires new staffing



<https://www.syntouch.nl/node/66>

Do You Need Hadoop for Big Data?

VERTICA

LOAD

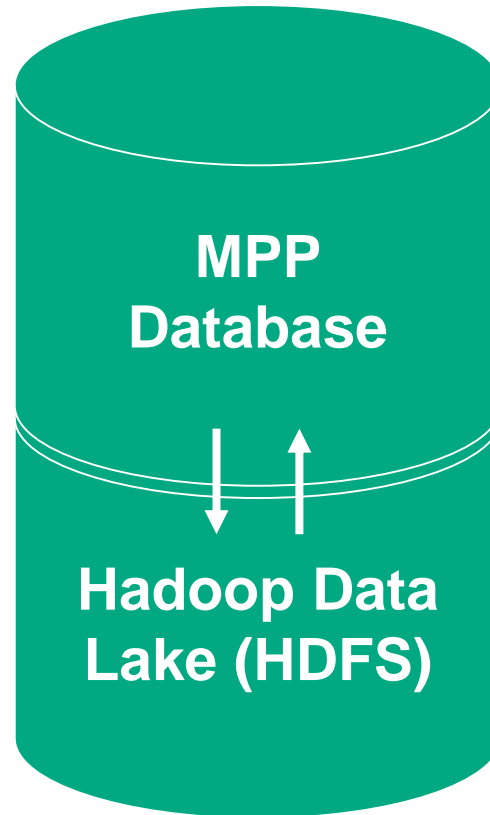
One high-tech company had a service level agreement which called for:

60 TB per hour

QUERY & CONCURRENCY

2,000+ Users

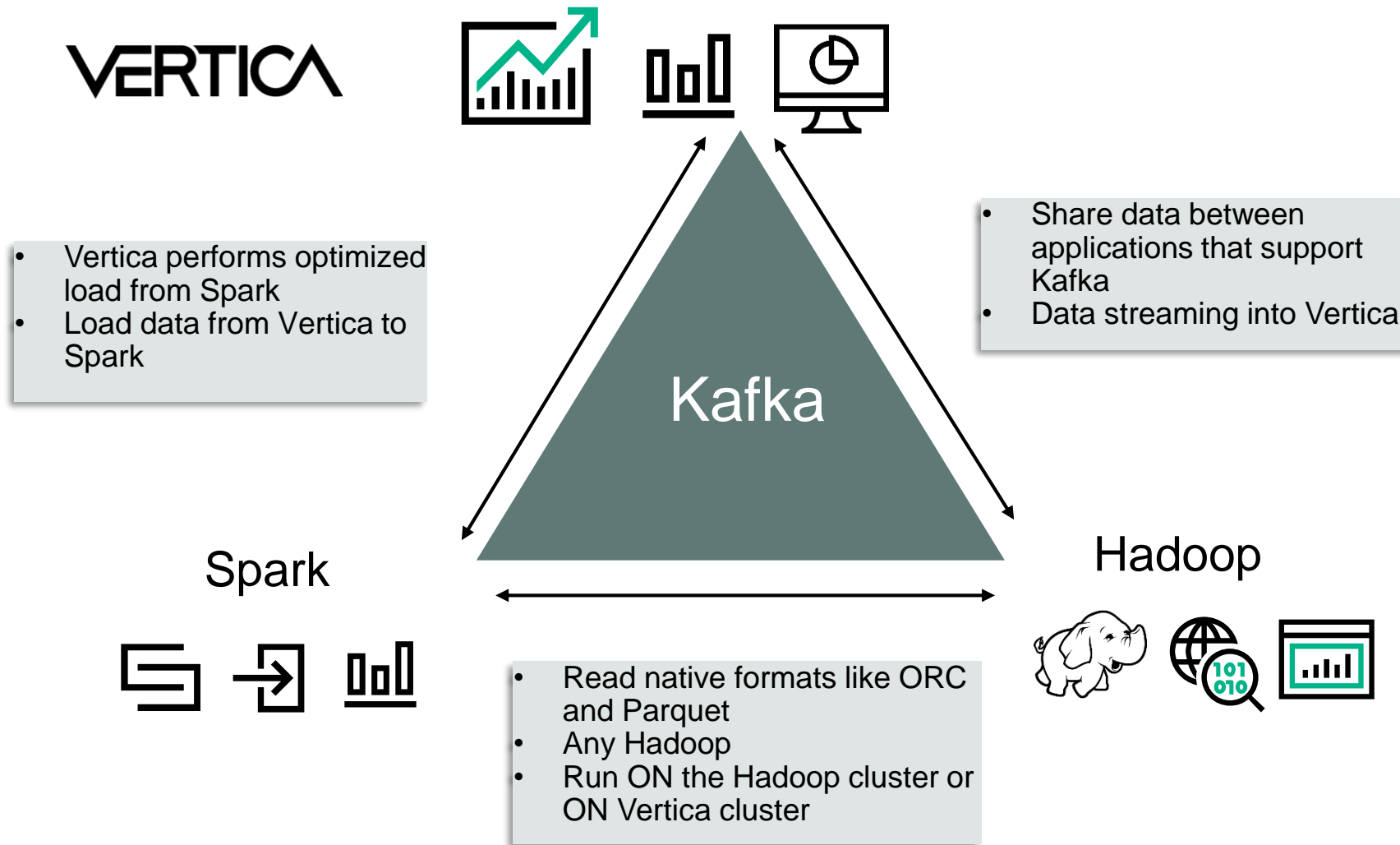
Reaching Across the Data Lake



Important-to-the-business data with tight SLAs

Cooler Data. Unknown Value

Embracing an Open Source Architecture





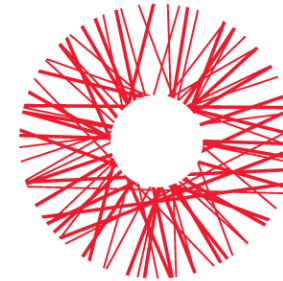
Delivering predictive network analytics for Telecommunications companies

The Challenge

- Data storage requirements increasing exponentially, but customers expecting analytic response times in seconds, not minutes or hours
- With their previous Oracle system, enlarging storage was complicated and time consuming.

The Solution

- Vertica has provided Anritsu with the technology necessary to implement predictive analytics solutions that have only been theoretical until now
- Realized rapid ROI after implementing Vertica in place of a legacy Oracle solution: 351% ROI with a payback of just 4 months



**NUCLEUS
RESEARCH**

ROI: 351%

Payback: 4 months

Annual Benefit: \$3+ million



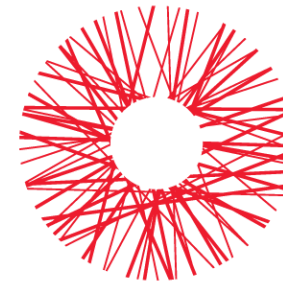
Using customer analytics to eliminate level 1 and level 2 support

The Challenge

- Perform analytics on over 1.5 petabytes of customer product and performance metadata to fine-tune and continue leading-edge product development and evolution

The Solution

- Leverage Vertica for operational analytics to engage in ongoing communications and with customers about storage environments and optimizations
- Use analytics to understand distribution of customers' workloads and how customers access storage, which helps it design storage solutions that match its customers' use patterns



NUCLEUS
RESEARCH

ROI: 287%

Payback: 6 months

Annual Benefit: \$13.6 million

Try Vertica Today



Run Vertica in the Clouds or OnPremise

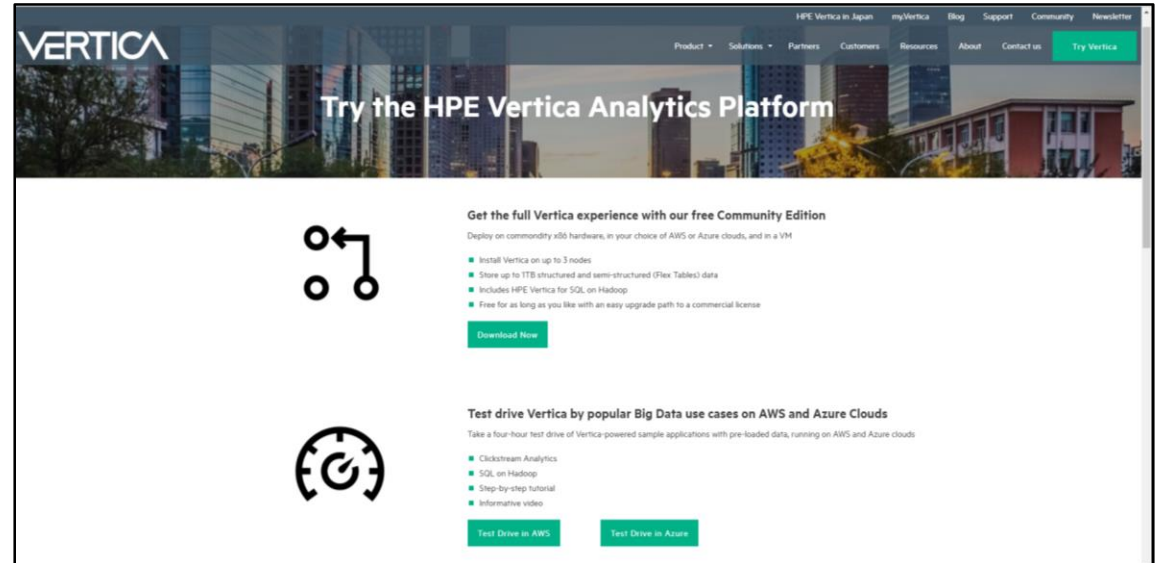
- 3 Node, 1 TB Community Edition
- Bring your own license (BYOL)
- Single or a multi-node cluster
- Choice of many Instance Types
- RHEL 7.0
- CloudFormation template for AWS

Take Test Drives on AWS & Azure

- Clickstream Analytics
- SQL on Hadoop
- Predictive Maintenance

www.vertica.com/try

Hewlett Packard
Enterprise



Thank You!

jeff.a.healey@hpe.c

om