# Accelerate Innovation in the Enterprise with Distributed ML / DL on GPUs

Thomas Phelan and Nanda Vijaydev – BlueData (recently acquired by HPE)

NVIDIA GTC – March 2019

Hewlett Packard Enterprise

# Agenda

- AI, Machine Learning (ML), and Deep Learning (DL)
- Example Enterprise Use Cases
- Deployment Challenges for Distributed ML / DL
- TensorFlow and Horovod on Containers with GPUs
- Lessons Learned and Key Takeaways

# Game Changing Innovation

## Gartner 2019 CIO Agenda

Q: Which technology areas do you expect will be a game changer for your organization?

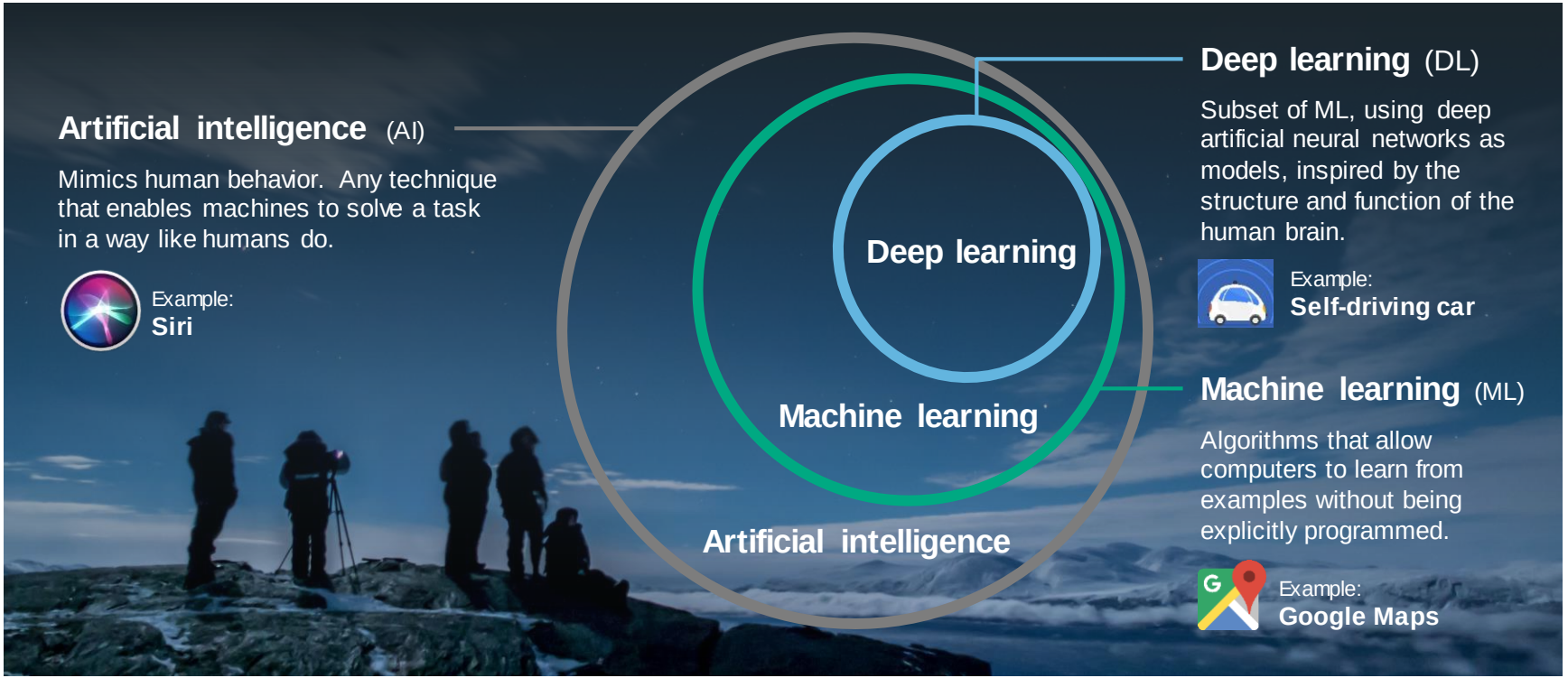
## Answers:

**#1** AI / Machine Learning
**#2** Data Analytics
**#3** Cloud
**#4** Digital Transformation

Source: Gartner, Insights From the 2019 CIO Agenda Report, by Andy Rowsell-Jones, et al.

# AI, Machine Learning, and Deep Learning

# Let's get grounded...what is AI?



# Why should you be interested in AI / ML / DL?

**Everyone wants AI / ML / DL and advanced analytics….**

AI and advanced analytics represent
**2 of the top 3 CIO priorities**

AI and advanced analytics infrastructure could constitute
**15-20%** of the market by 2021[1]

Enterprise AI adoption
**2.7X** growth in last 4 years[2]

**….but face many challenges**

Use cases

New roles, skill gaps

Culture and change

Data preparation

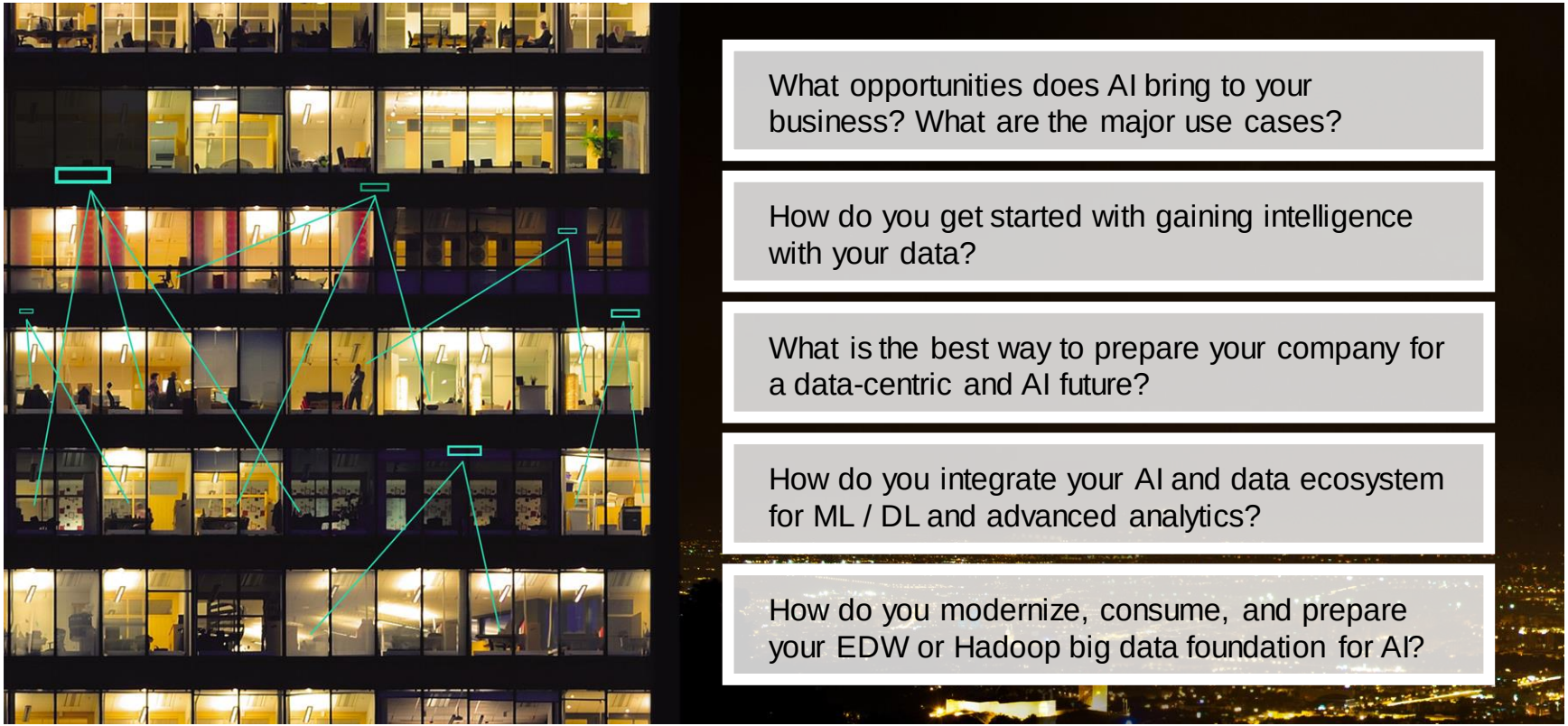
Legacy infrastructure

Hewlett Packard Enterprise

[1] IDC. Goldman Sachs. HPE Corporate Strategy.2018
2 Gartner - "2019 CIO Survey: CIOs Have Awoken to the Importance of AI"

bluedata
recently acquired by Hewlett Packard Enterprise

# Key questions remain



What opportunities does AI bring to your business? What are the major use cases?

How do you get started with gaining intelligence with your data?

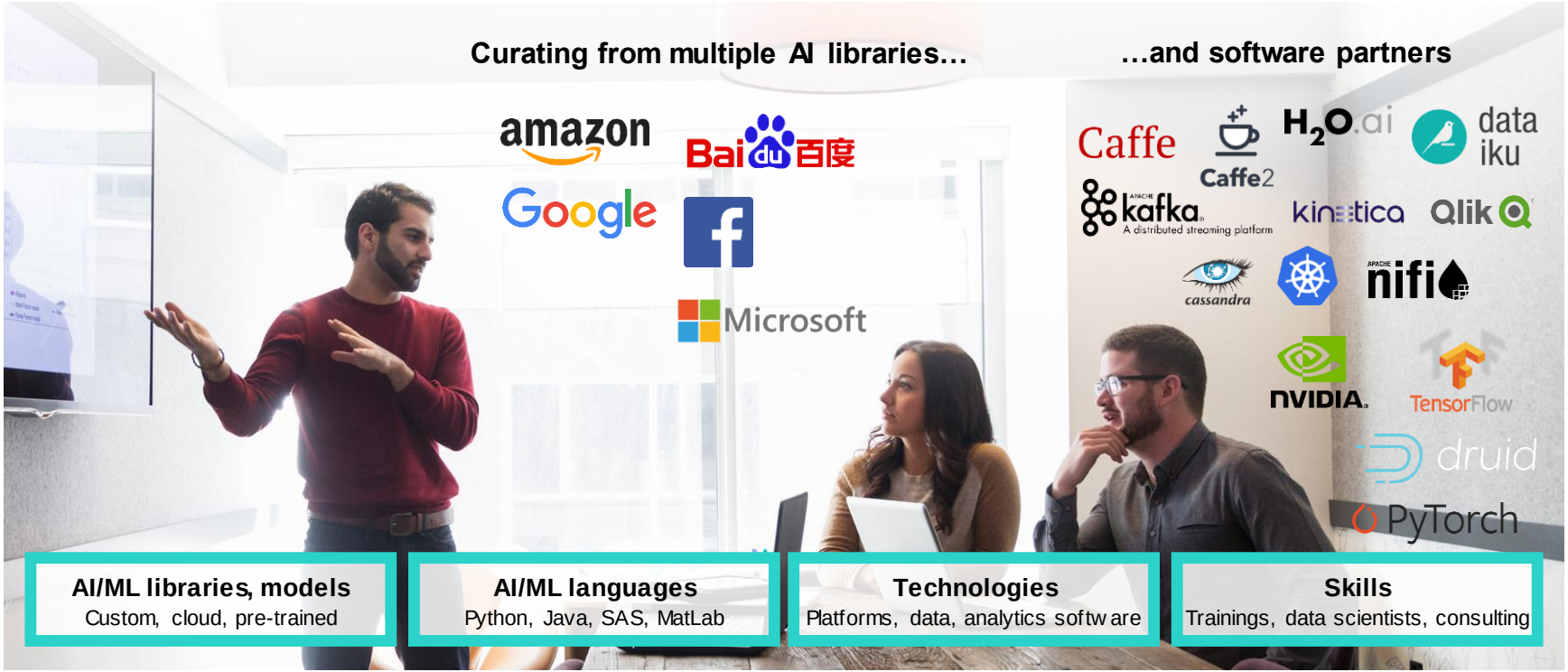
What is the best way to prepare your company for a data-centric and AI future?

How do you integrate your AI and data ecosystem for ML / DL and advanced analytics?

How do you modernize, consume, and prepare your EDW or Hadoop big data foundation for AI?

# HPE can help

Aggregating HPE products and services with our best in class partner and AI ecosystem



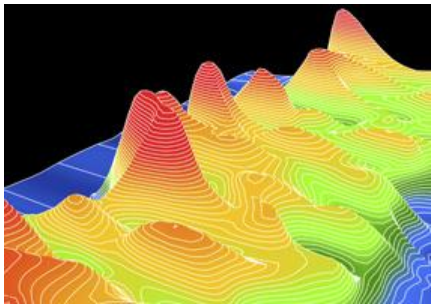

Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# AI / ML / DL Adoption in the Enterprise

**Financial services**
Fraud detection, ID verification

**Government**
Cyber-security, smart cities and utilities



**Energy**
Seismic and reservoir modeling

**Retail**
Video surveillance, shopping patterns

**Health**
Personalized medicine, image analytics

**Consumer tech**
Chatbots

**Service providers**
Media delivery

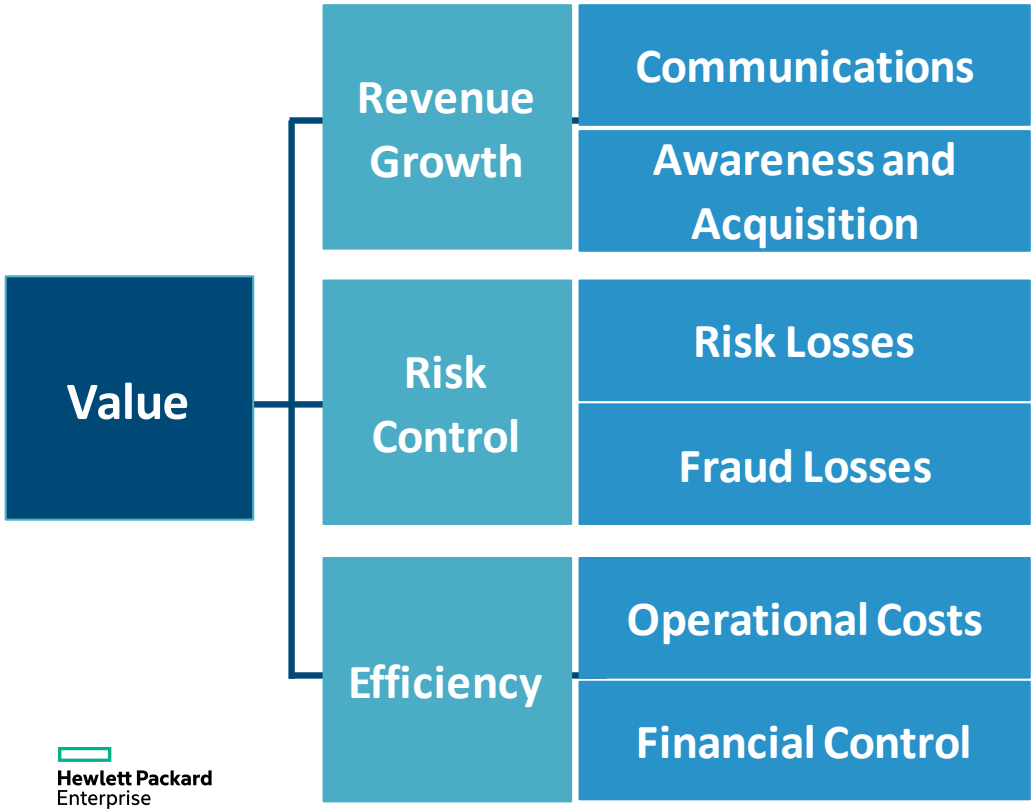
**Manufacturing**
Predictive and prescriptive maintenance

Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# Example Enterprise Use Cases

# ML / DL in Financial Services




## Example Use Cases

- Know Your Customers (KYC)
- Customer Experience
- Customer Value Modeling
- Customer Churn Reduction
- Origination Risk Underwriting
- Credit Risk Assessment
- Fraud Detection / Prevention
- Anti-Money Laundering (AML)
- Capacity Planning
- Automation
- Portfolio Simulation

Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# More Financial Services Use Cases

## Wide Range of ML / DL Use Cases for Wholesale / Commercial Banking, Credit Card / Payments, Retail Banking, etc.

### Fraud Detection

- Real-Time Transactions
- Credit Card
- Merchant
- Collusion
- Impersonation
- Social Engineering Fraud

### Risk Modeling & Credit Worthiness Check

- Loan Defaults
- Delayed Payments
- Liquidity
- Market & Currencies
- Purchases and Payments
- Time Series

### CLV Prediction and Recommendation

- Historical Purchase View
- Pattern Recognition
- Retention Strategy
- Upsell
- Cross-Sell
- Nurturing


### Customer Segmentation

- Behavioral Analysis
- Understanding Customer Quadrant
- Effective Messaging & Improved Engagement
- Targeted Customer Support
- Enhanced Retention

### Other

- Image Recognition
- NLP
- Security
- Video Analysis

CLV: Customer Lifetime Value

# Fraud Detection Use Case

- One of the most common use cases for ML / DL in Financial Services is to detect and prevent fraud
- This requires:
  - Distributed Big Data processing frameworks such as Spark
  - ML / DL tools such as TensorFlow, H2O, and others
  - Continuous model training and deployment
  - Multiple large data sets

# Fraud Detection Use Case (cont'd)

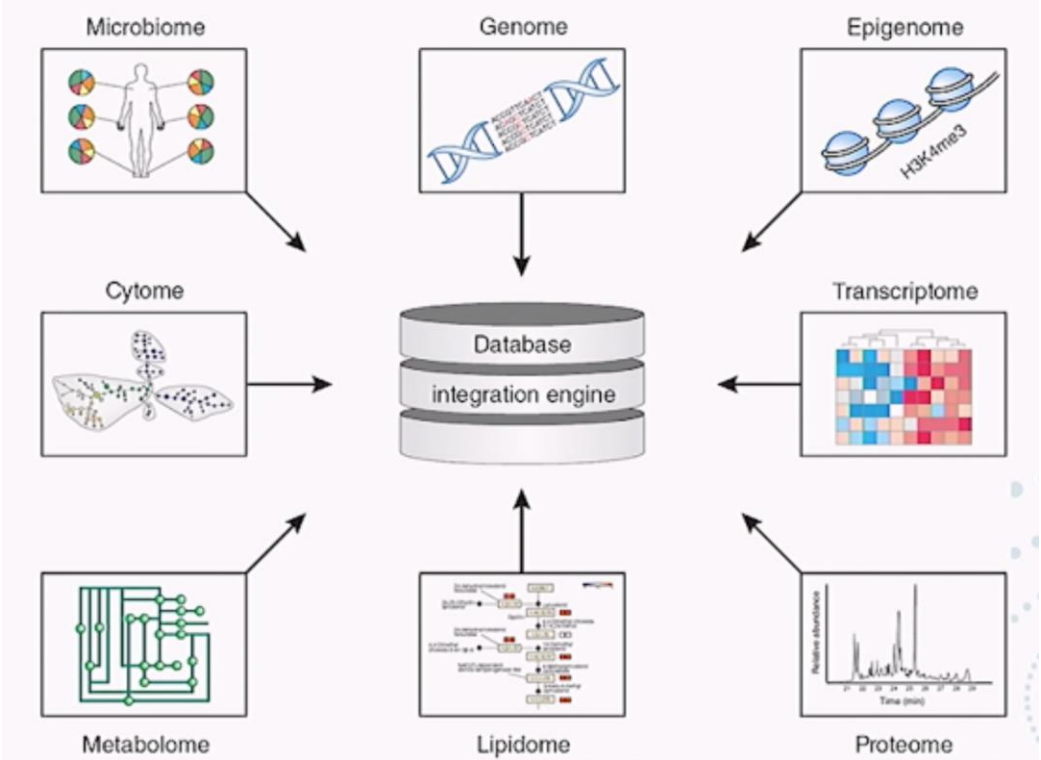
- Data science teams need the ability to create distributed ML / DL environments for sandbox as well as trial and error experimentation
- This requires:
  - Hardware acceleration (e.g. GPUs)
  - Multiple different ML / DL and data science tools
  - Fast and repeatable deployment of clusters

# ML / DL in Healthcare – Use Cases

- Precision Medicine and Personal Sensing
  - Disease prediction, diagnosis, and detection (e.g. genomics research)
  - Using data from local sensors (e.g. mobile phones) to identify human behavior
- Electronic Health Record (EHR) correlation
  - "Smart" health records
- Improved Clinical Workflow
  - Decision support for clinicians
- Claims Management and Fraud Detection
  - Identify fraudulent claims
- Drug Discovery and Development

# Use Case: Precision Medicine

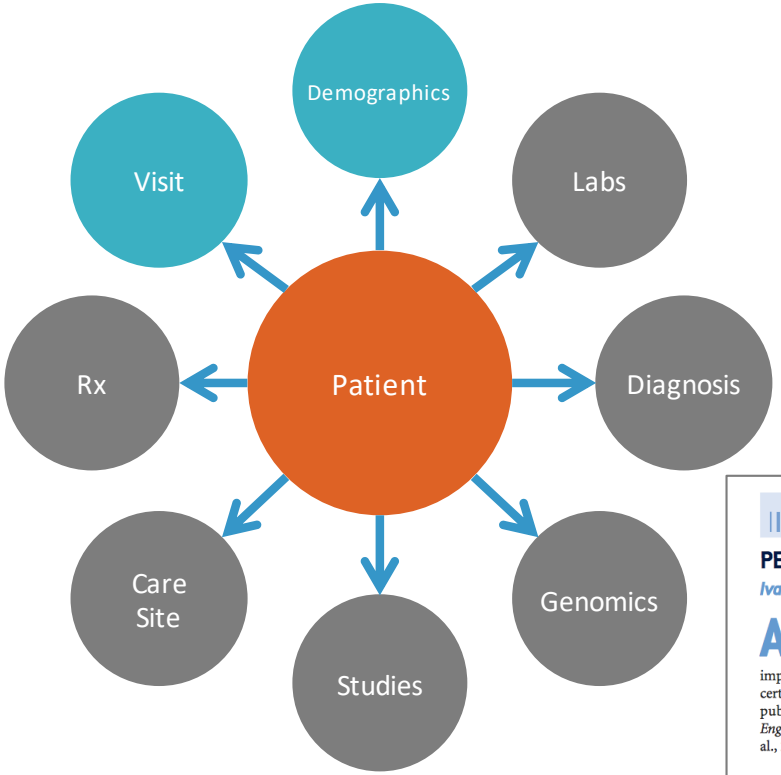
- Many types of data
  - Genomic
  - Microbiome
  - Epigenome
  - Etc.
- Huge volumes of data (petabytes > exabytes)


Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# 360° View of the Patient





**RESEARCH UPDATE**

## PERSONALIZED MEDICINE COMES TO CYSTIC FIBROSIS

*Ivacaftor targets G551D mutation with success*

A phase III study of an experimental, targeted cystic fibrosis (CF) treatment shows sustained lung function improvement for patients with CF with a certain mutation, according to a report published in November in the *New England Journal of Medicine* [Ramsey et al., 2011].

The experimental drug ivacaftor, also known as VX-770, is the first personalized treatment for CF. It targets an underlying genetic cause of CF, the G551D-CFTR mutation, not just its symptoms. About 4-5% of patients with CF have at least one such mutation.

Ivacaftor produced substantial lung function improvement in patients with CF age 12 and older after two weeks. It sustained this improvement through the 48 weeks of the study. Ivacaftor produced a 10.6 percentage point improvement in forced expiratory volume in one second (FEV1) over placebo at 24 weeks, reported by Bonnie Ramsey, MD, Director of the Center

Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

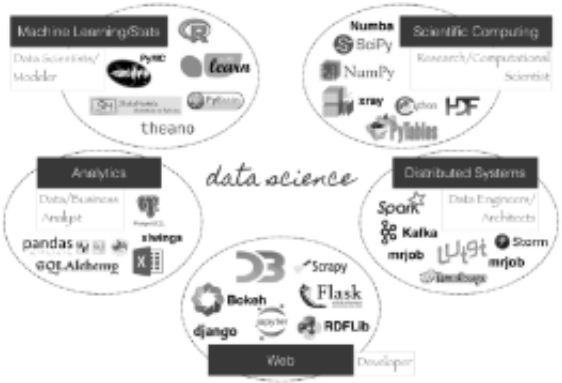
# ML / DL in Healthcare – Requirements

- Data security and data access
  - HIPAA and other regulatory requirements
  - Data is usually in siloes, and data scientists don't want to share their data
- Support for multiple simultaneous clusters with varying QoS
  - Want to offload low priority jobs from production cluster
- Low priority jobs require access to production data
  - Want to avoid repeated copies of production data
- Support for multiple custom tools and analytics applications
  - Need to accelerate the application deployment time

# Deployment Challenges for Distributed ML / DL

# Distributed ML / DL – Challenges

- Complexity, lack of repeatability and reproducibility across environments
- Sharing data, not duplicating data
- Need agility to scale up and down compute resources
- Deploying multiple distributed platforms, libraries, applications, and versions
- One size environment fits none
- Need a flexible and future-proof solution


Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# Example Deployment Challenges

- How to run clusters on heterogeneous host hardware
  - CPUs and GPUs, including multiple GPU versions
- How to maximize use of expensive hardware resources
- How to minimize manual operations
  - Automating the cluster creation and and deployment process
  - Creating reproducible clusters and reproducible results
  - Enabling on-demand provisioning and elasticity

# Example Deployment Challenges

- How to support the latest versions of software
  - Deployment complexity and upgrades
  - Version compatibility
- How to ensure enterprise-class security
  - Network, storage, user authentication, and access

# Docker Containers

**Docker** is a computer program that performs operating-system-level virtualization also known as *containerization*.

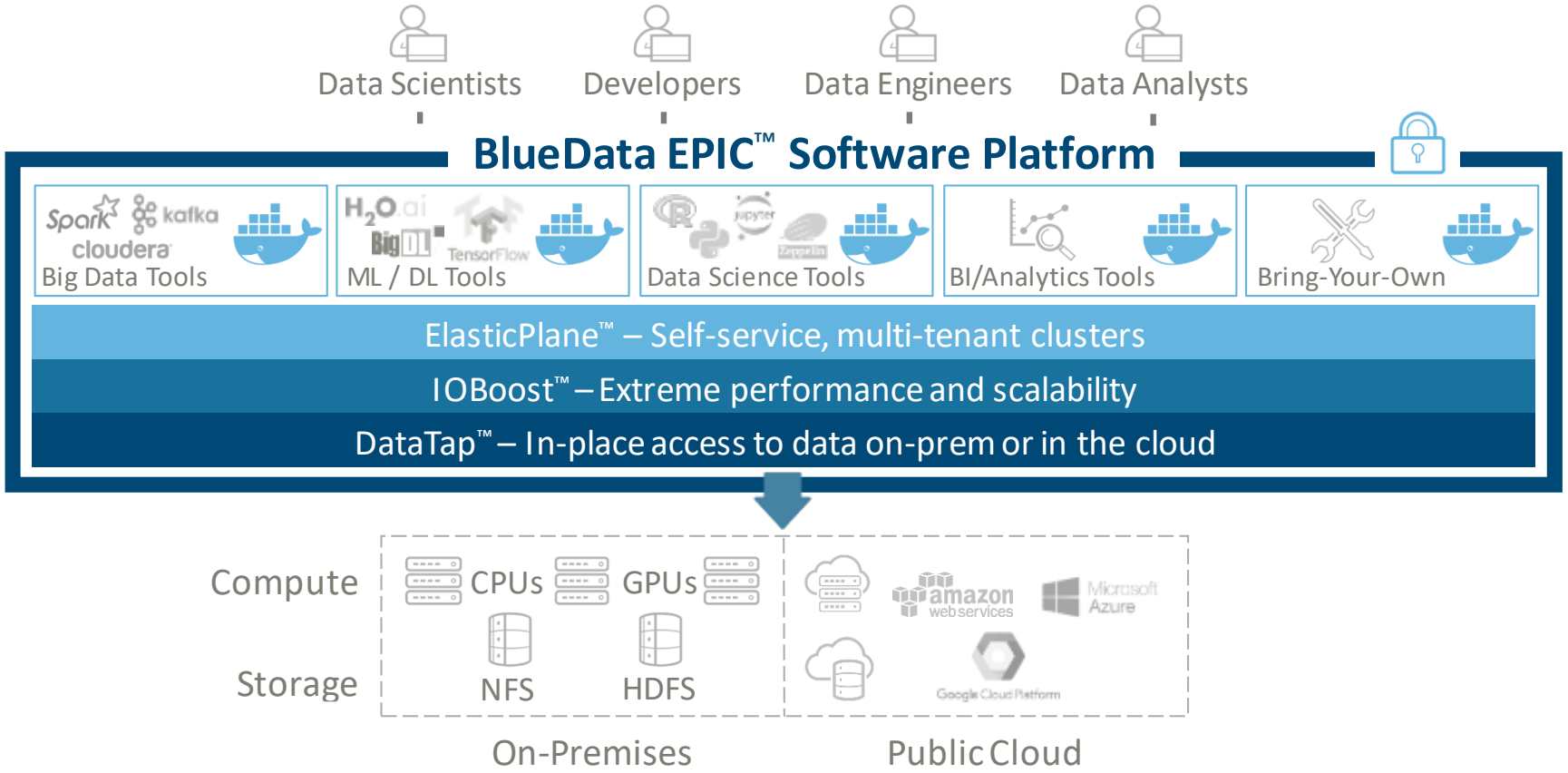
**Containerization** allows the existence of multiple isolated user-space instances.

Source: https://en.wikipedia.org/wiki/docker_(software)

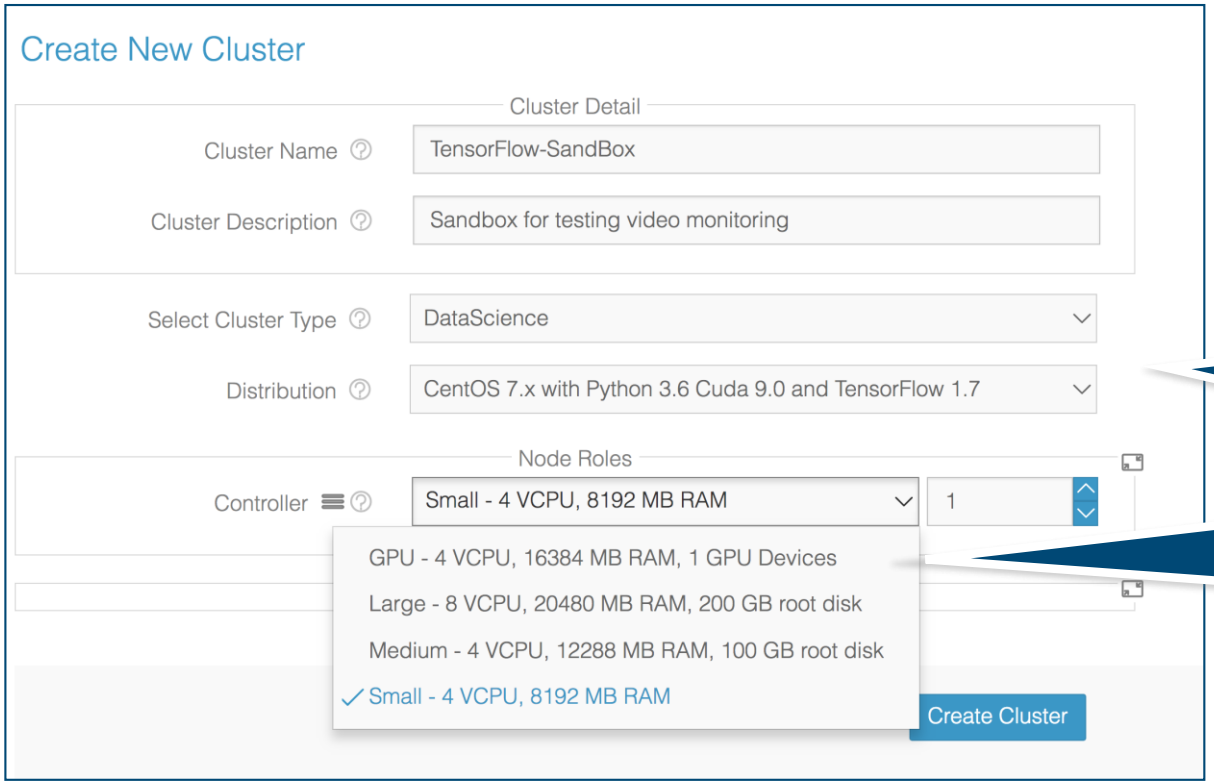
# Distributed ML / DL and Containers

- ML / DL applications are compute hardware intensive
- They **can** benefit from the flexibility, agility, and resource sharing attributes of containerization
- **But** care must be taken in how this is done, especially in a large-scale distributed environment

# Turnkey Container-Based Solution



# One-Click Cluster Deployment



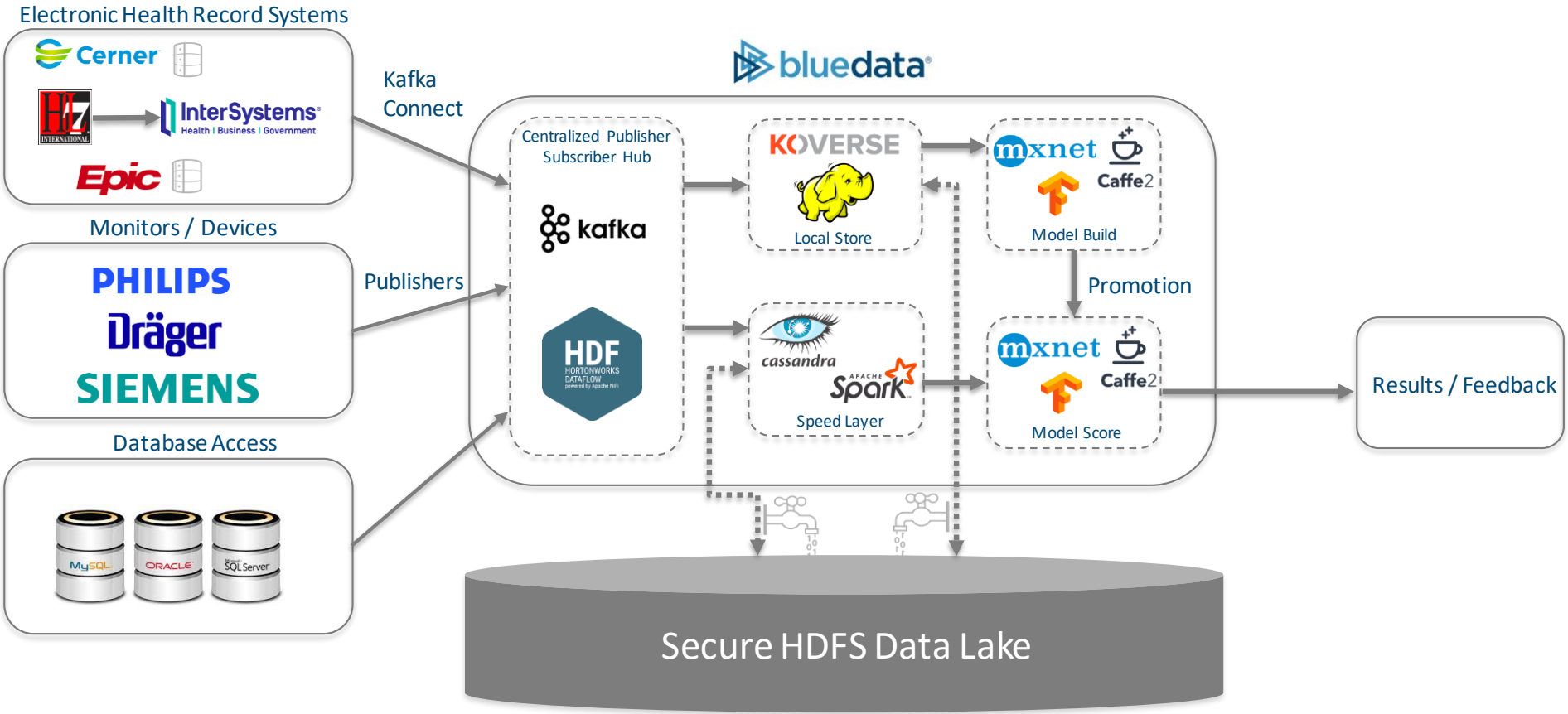

Pick from a list of pre-built and tested Docker-based images

Assign specific resources (GPUs, CPUs) to the cluster, depending on the use case

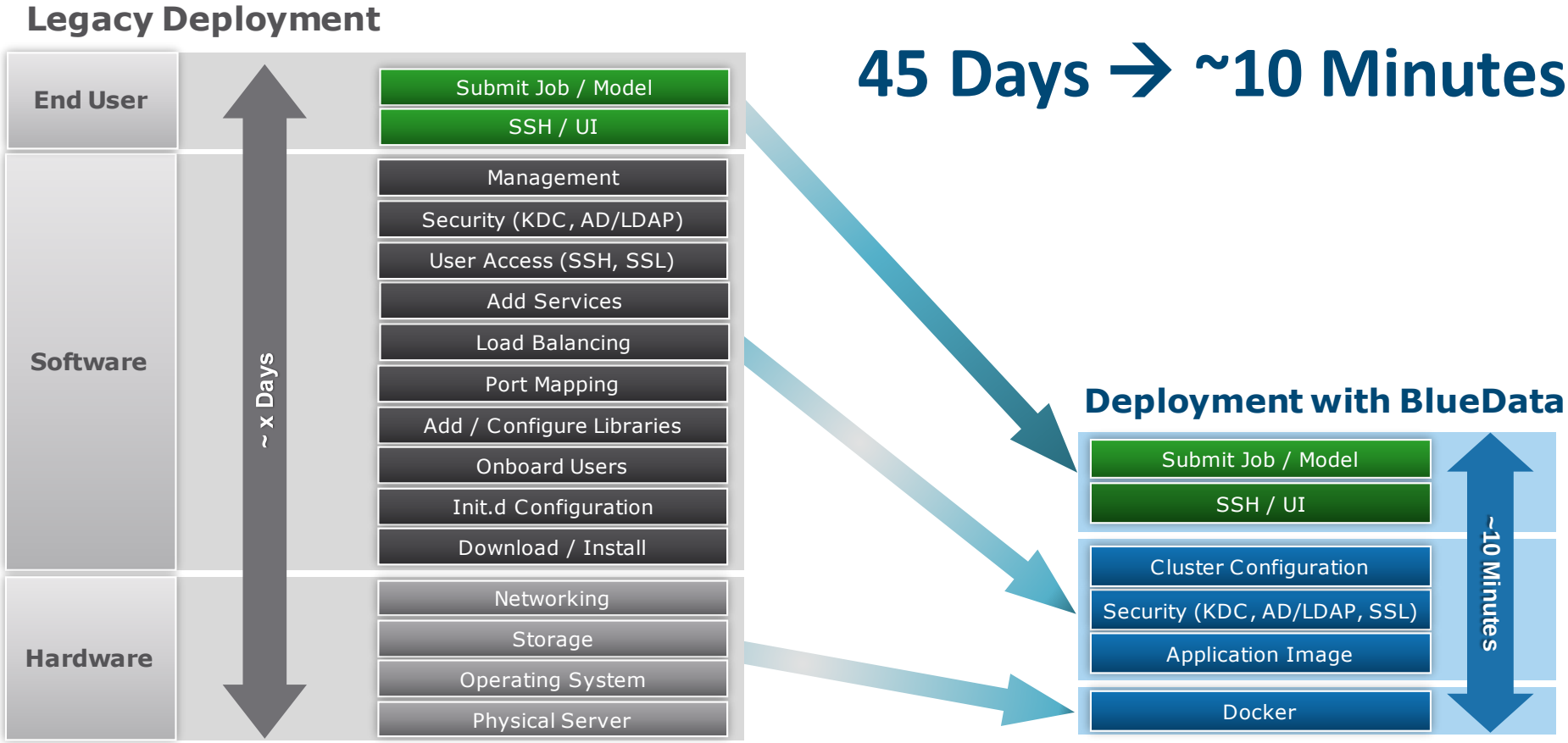
Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# Architecture Example in Healthcare



# Faster ML / DL Deployment Time



# Bringing It All Together

## Building blocks for AI / ML / DL

**Packaging & Deployment**

**Compute**


*Connectors and Accelerators*

**App Stacks**

TensorFlow, Spark, MLlib, H2O.ai, jupyter, Caffe2

**Data**

SQL Server, ORACLE, PostgreSQL, MySQL, SYBASE, IBM DB2, hadoop, kafka, APACHE HBASE, cassandra

Hewlett Packard Enterprise

## Turnkey solution for distributed AI / ML / DL

bluedata®

### Accelerate innovation and time-to-value:

- Speed and agility for data science teams
- Flexibility for architecture teams
- Cost savings for operations
- Enterprise-grade security for IT

bluedata
recently acquired by Hewlett Packard Enterprise

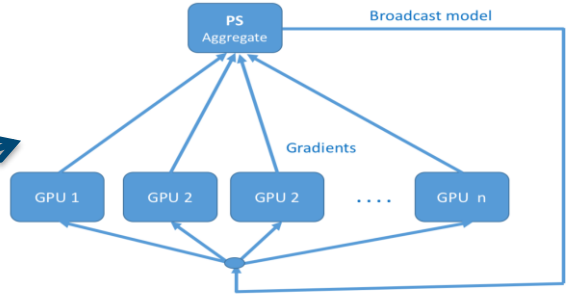
# TensorFlow and Horovod on Containers with GPUs

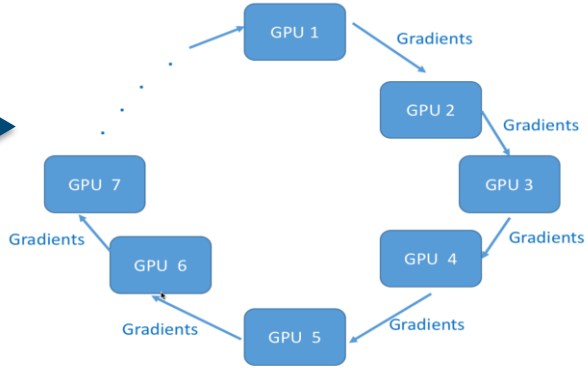
# Distributed Tensorflow – Concepts

- Running TensorFlow training in parallel, on multiple devices, using GPUs
- Goal is to improve accuracy and speed
- Different layers may be trained on different nodes (**model parallelism**)
- Same model can applied on different subset of data, in different nodes (**data parallelism**)

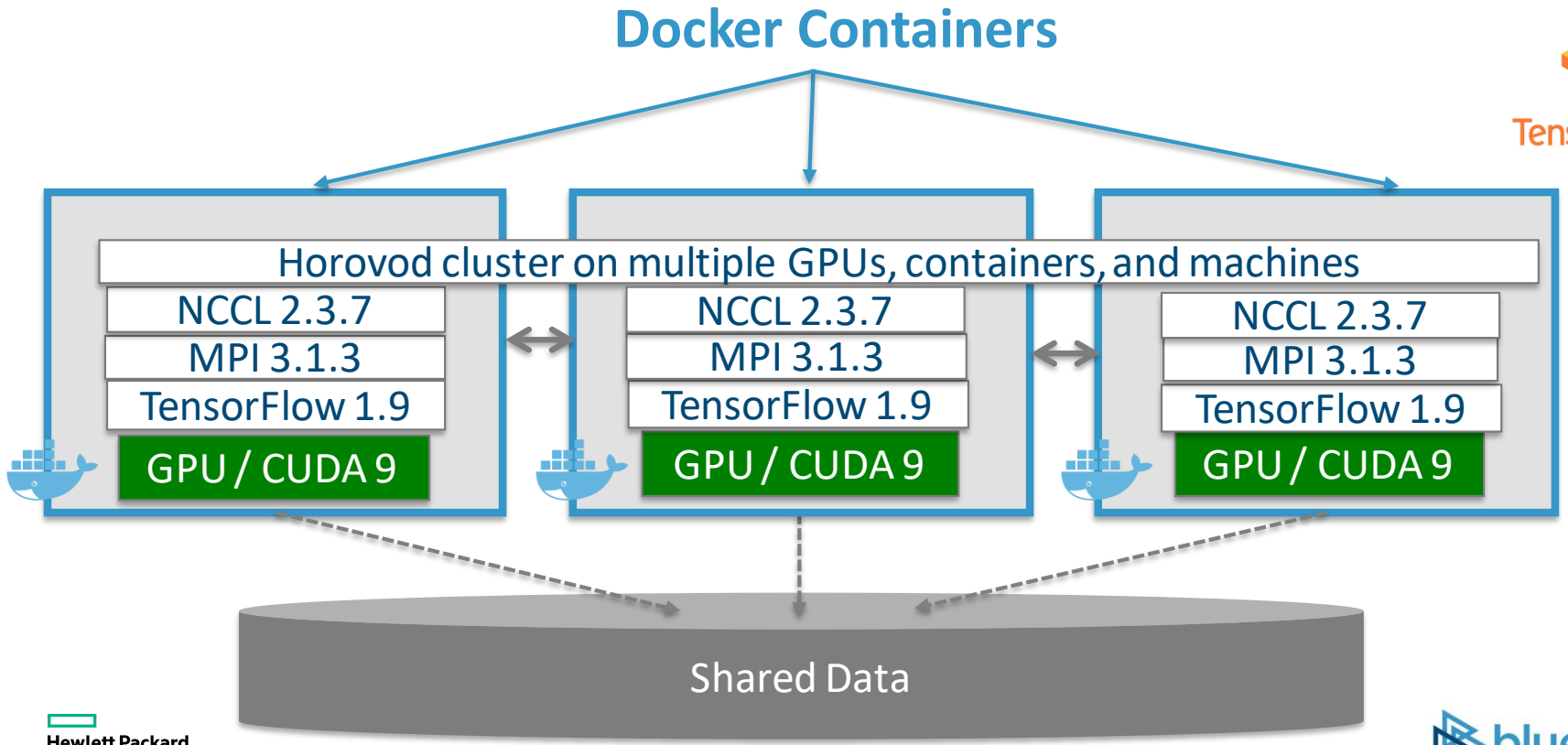
# Distributed Tensorflow – Schemes

- Data parallelism implementation
  - Needs to sync model parameters
  - Uses a centralized or decentralized scheme to communicate parameter update
- Centralized schemes use **Parameter Server** to communicate updates to parameters (gradients) between nodes
- Decentralized schedules use **ring-allreduce** scheme
- Horovod is an open source framework developed by Uber that supports allreduce



# TensorFlow with Horovod on Docker




Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise

# Demo – TensorFlow with Horovod

- tensorflow_wrd2vec.py from git https://github.com/horovod/horovod examples
- Data comes from shared NFS mounts, automatically surfaced by BlueData into containers
- Passwordless ssh setup during cluster creation
- All prerequisites installed all nodes including
  - nccl, cuda driver, cudnn app framework (NVIDIA components)
  - tensorflow, pytorch, scikit-learn, … (compute frameworks)
  - mpi (runtime for distributed jobs)
  - tensorboard for visualization

# Demo – TensorFlow with Horovod

```
mpirun -np 2    /
   --allow-run-as-root /
    -d  -H localhost:1,bluedata-301.bdlocal:1 /
    -bind-to none -map-by slot   /
    -x NCCL_DEBUG=INFO  /
   -x LD_LIBRARY_PATH /
   -x PATH     /
  -mca pml ob1 /
  -mca btl ^openib python tensorflow_word2vec_logs.py
```

# Lessons Learned and Key Takeaways

# Lessons Learned and Takeaways

- Enterprises are using ML / DL today to solve difficult problems (example use cases: fraud detection, disease prediction)
- Distributed ML / DL in the enterprise requires a complex stack, with multiple different tools (TensorFlow is one popular option)
- The only constant is change … be prepared
  - Business needs, use cases, and tools will constantly evolve
- Deployments are challenging, with many potential pitfalls
  - Containerization can deliver agility and cost saving benefits

# Lessons Learned and Takeaways

- Leverage a flexible, scalable, and elastic platform for success
  - BlueData provides a turnkey container-based platform for large-scale distributed AI / ML / DL in the enterprise
  - Enterprise-grade security and performance, proven in production at leading Global 2000 organizations
  - Decouple compute from storage for greater efficiency, and deploy on-premises, in a hybrid model, or multi-cloud
  - Save time, save money, and accelerate innovation

# Thank You


To learn more, visit **BlueData** in the HPE booth (1129)

**www.bluedata.com**

Hewlett Packard Enterprise

bluedata
recently acquired by Hewlett Packard Enterprise