




# Solve Data Challenges with Edge AI Processing


Presented by Red Hat, HPE, NVIDIA & Carahsoft



Public Sector Webcast  
**Solve Data Challenges with Edge AI Processing**

Thursday, March 25, 2021  
2:00pm ET; 11:00am PT  
*CPE Eligible*

*Thank you for joining us!*



Carahsoft Technology Corp. Proprietary and Confidential



Public Sector Webcast

# Solve Data Challenges with Edge AI Processing

Thursday, March 25, 2021  
2:00pm ET; 11:00am PT

*CPE Eligible*

*Thank you for joining us!*



**Red Hat**



**nVIDIA.**



**Hewlett Packard  
Enterprise**

**carahsoft.**



# About Carahsoft

Carahsoft Technology Corp., **The Trusted Government IT Provider®**, is dedicated to helping public sector IT organizations find the best possible technology solution at the possible value.



# KubeFrame for AI-Edge

*Make real-time, AI-driven edge processing a reality*

Red Hat, Hewlett Packard Enterprise and NVIDIA have collaborated to create a solution that combines easily portable hardware, powerful open source software and industry-leading security to meet the challenges of data processing in the field.



## Featured Experts



**Dan Domkowski**  
AI/ML/DL  
Delivery Specialist  
**Red Hat**



**Trent Boyer**  
North America OEM  
Account Manager  
**NVIDIA**



**Joe Vidal**  
Master Technologist,  
Office of the NA CTO  
**HPE**



# Delivering Edge-based AI/ML with Kubeframe

*Any workload, any footprint, any environment*

## Focus on Meaningful Outcomes



“Data itself are worthless. Absolutely nothing counts in this world except action.”

-Dr. Edmund Jackson, Chief Data Officer, HCA Healthcare

# Getting to Action



Red Hat, NVIDIA, and HPE are uniquely positioned to deliver actionable results with Artificial Intelligence/Machine Learning (AI/ML) and Accelerated Computing...

...in any environment.

# AI/ML Can Lead to Action



## Healthcare

- Increased clinical efficiency
- Faster/better diagnosis
- Improved outcomes



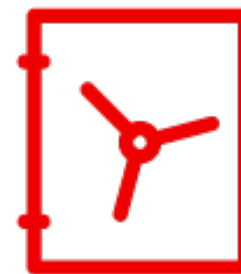
## Financial services

- More personalized services
- Improved risk analysis
- Reduced fraud
- Better predictions



## Telcos

- Better customer insights/experiences
- Optimized network performance
- Improved threat detection



## Insurance

- Automated claims processing and handling
- Usage-based insurance services

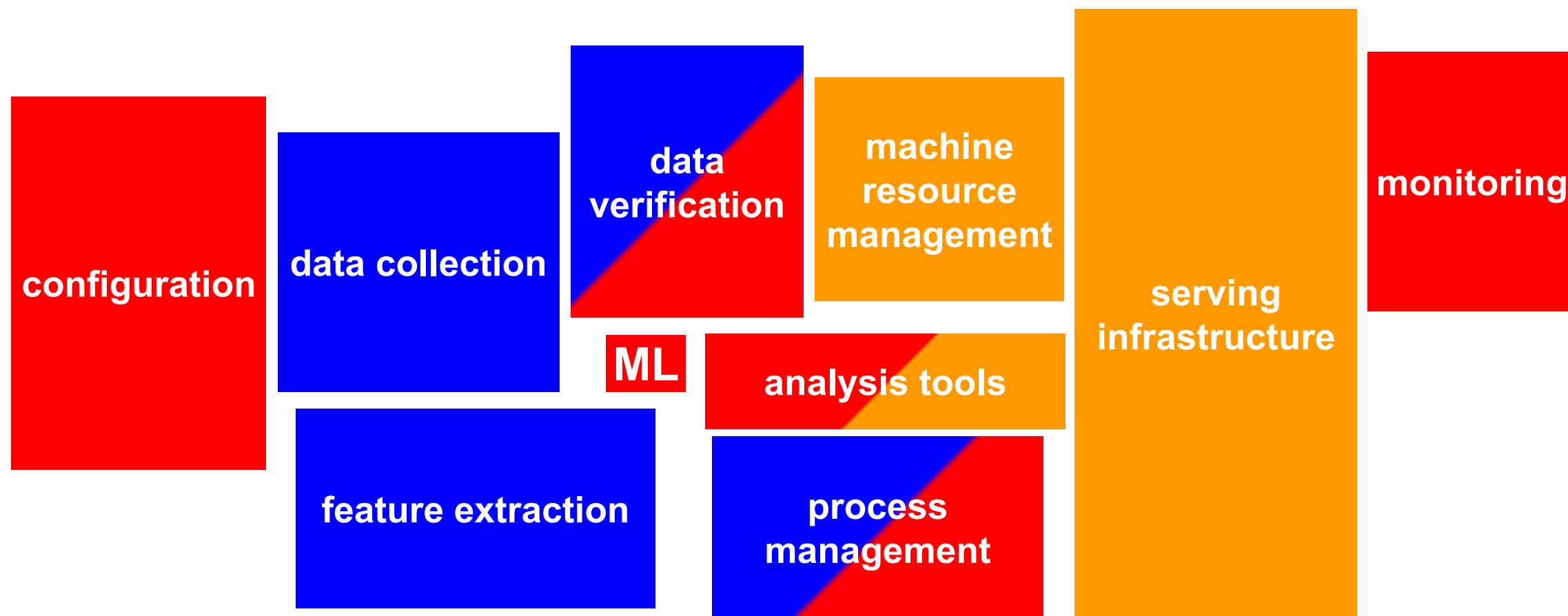


## Automotive

- Autonomous driving
- Predictive maintenance
- Improved supply chains

<sup>9</sup> **Why Now? - Advancements in accelerated computing power, data aggregation, and open source ML tools/frameworks are helping make Meaningful Outcomes real!**

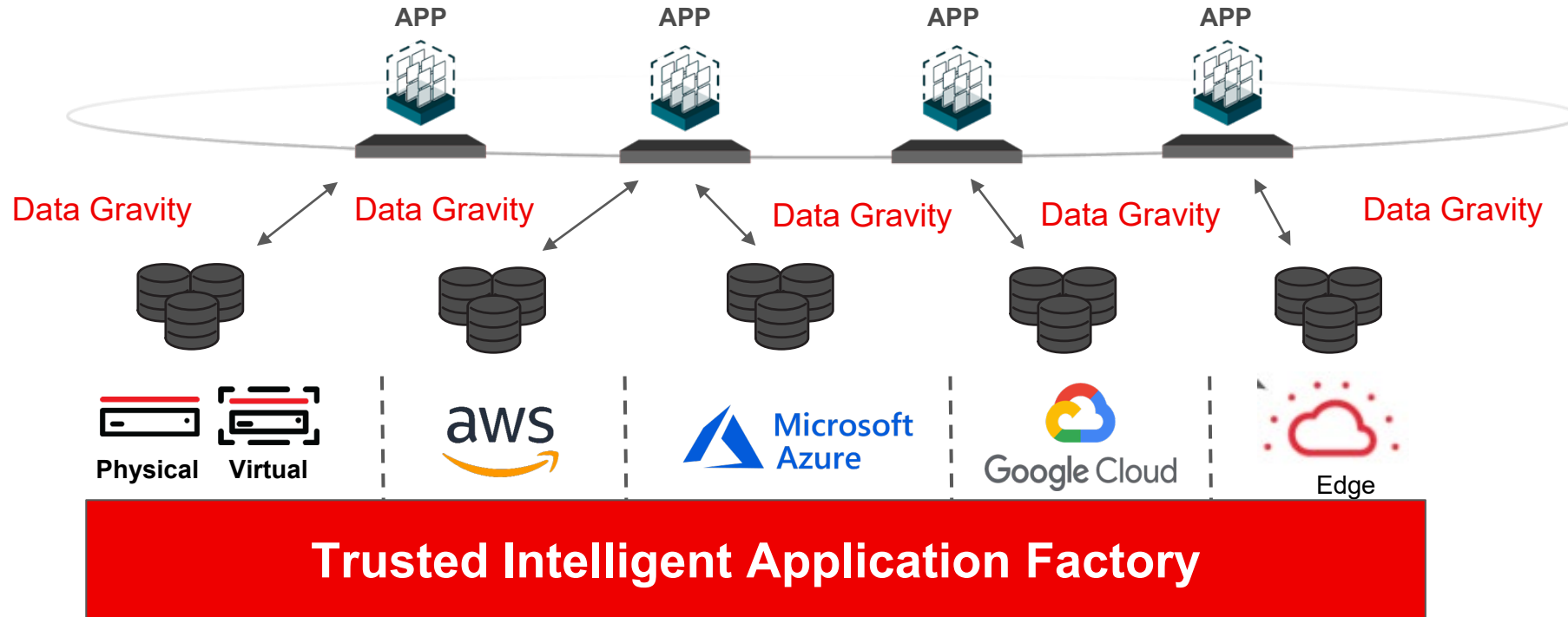
# Delivering AI/ML has its Challenges



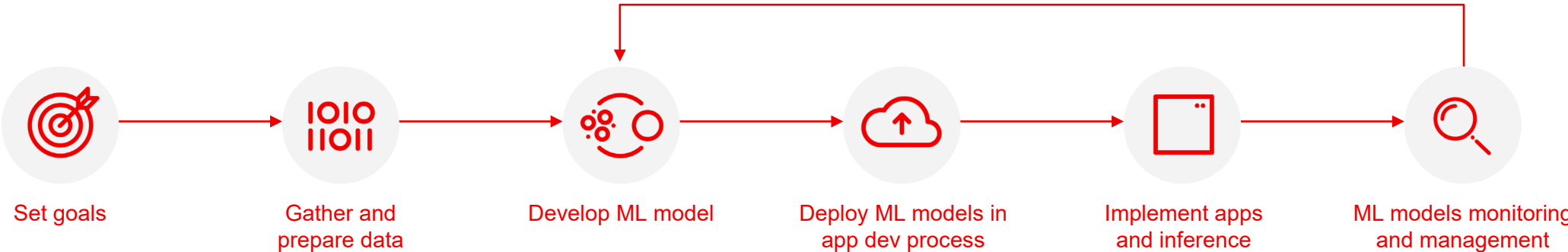
**Infrastructure Operations | Data Engineers | Data Scientists**  
(Adapted from Sculley et al., "Hidden Technical Debt in Machine Learning Systems." NIPS 2015)

# Solution: Deliver Workloads TO the Data

*Flexibly and Consistently in Any Environment*



# DevOps and AI/ML Enforce Good Behaviors



Execution venue	Set goals	Gather and prepare data	Develop ML model	Deploy ML models in app dev process	Implement apps and inference	ML models monitoring and management
Datacenter	X	X	X	X	X	X
Public clouds	X	X	X	X	X	X
Edge computing	X				X	X

Consistent infrastructure platform across data center, edge, and public clouds.

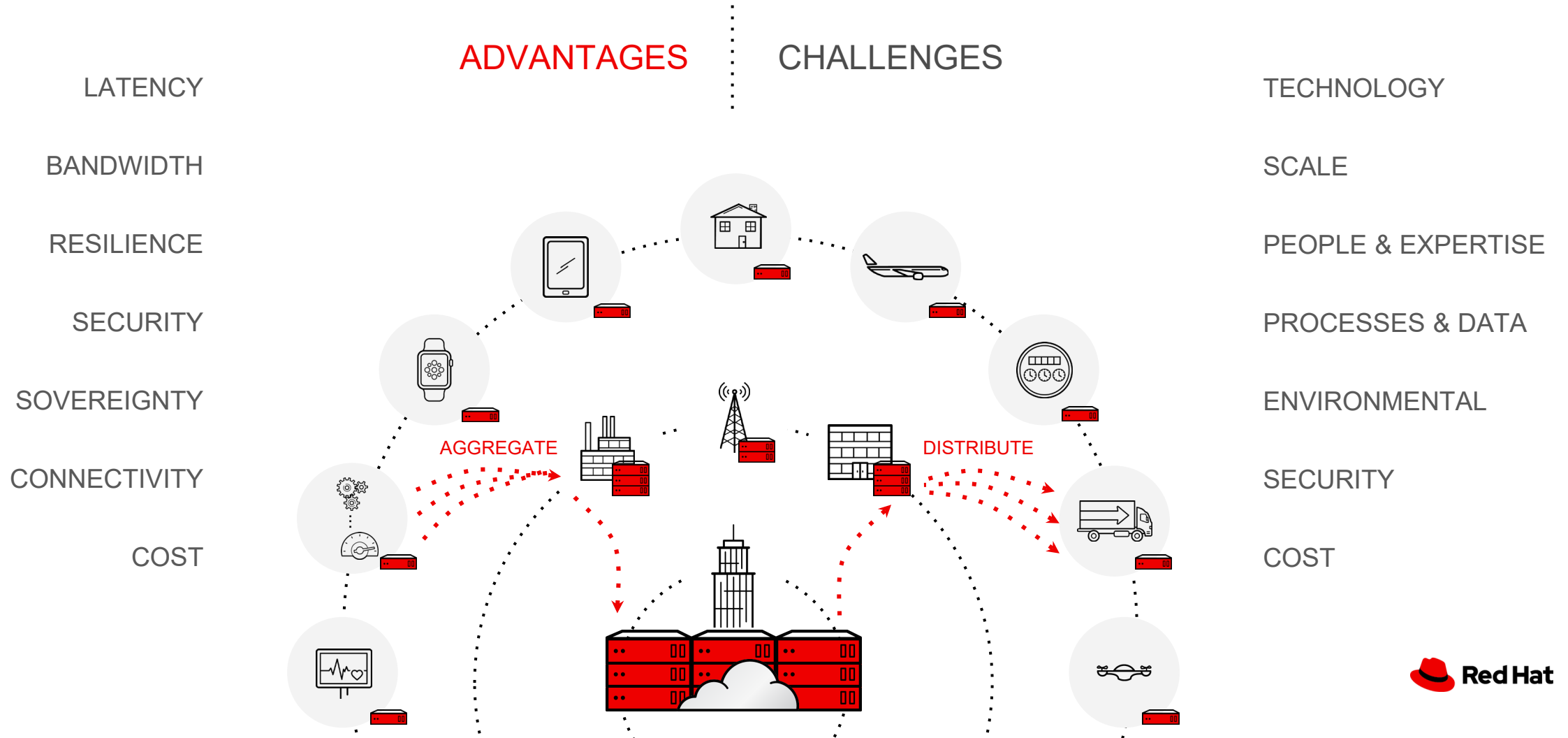
Simplify and accelerate deployment and lifecycle management of:

- ▶ Machine learning and DevOps software tools
- ▶ Machine learning data pipelines and data sources

Integrate compute acceleration with bare metal infrastructure to speed up artificial intelligence inferencing tasks.

# Edge Computing

*Making Decisions Closer to the Data's Source*



# Who is Doing Edge?



**Enterprise**



**Telco 5G Edge**



**Public Sector**



**AI/ML**



**IoT**

(Retail) Remote Office

Radio Access Network  
(vRAN)

Connected Sensors  
and Controllers

Video Surveillance

Trading Sensitive Apps

In-field Operations

Content Delivery Networks

Universal Customer Premise  
Equipment

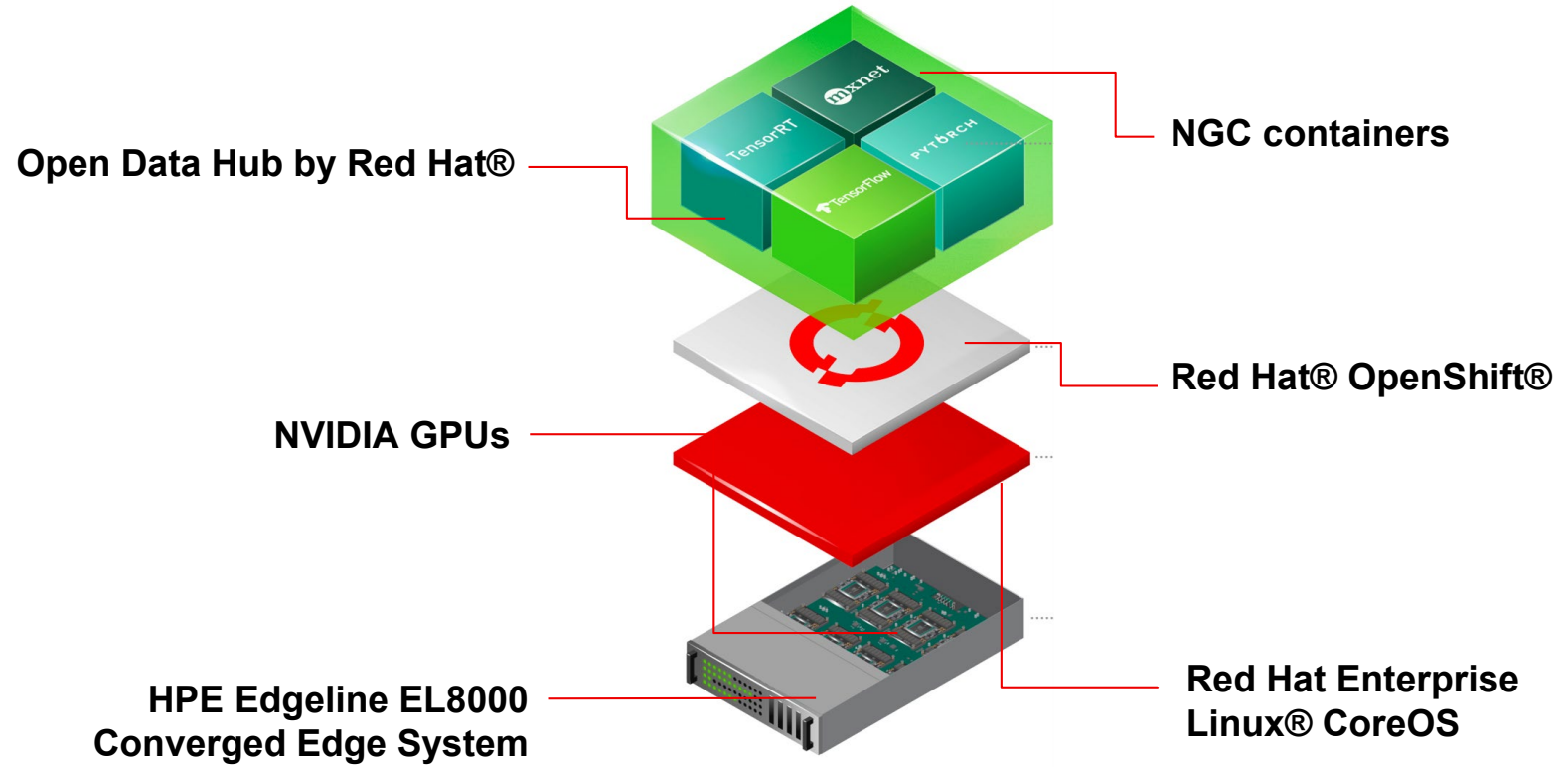
Connected Car, Train, Plane

Healthcare Monitoring and  
Data Processing

and more...

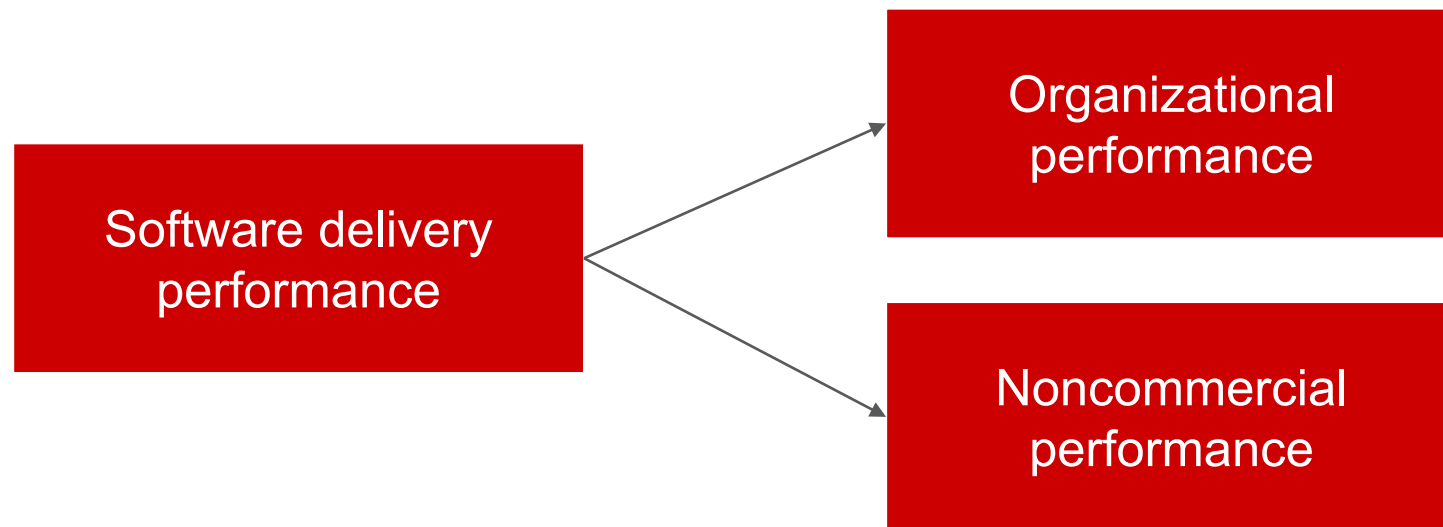
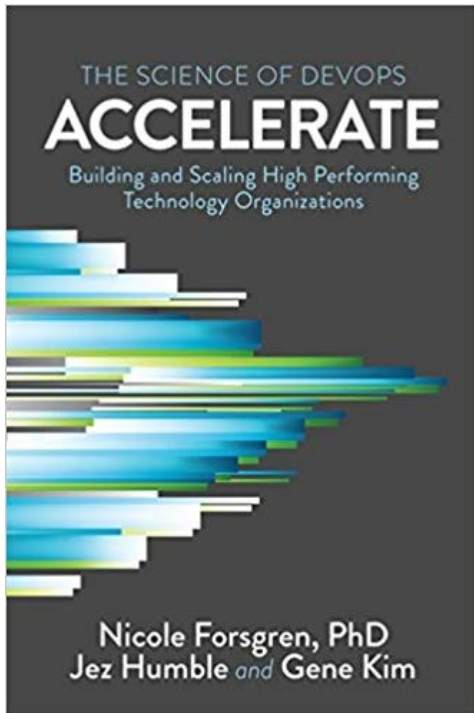
# KubeFrame

*Accelerating Edge-based AI/ML with HPE, NVIDIA, and Red Hat*



AI/ML on Accelerated Computing in Any Environment

# Software delivery performance drives overall performance



# HCA Healthcare Delivering Action!

*Speed up sepsis detection with machine learning software*

Put software in the hands of nurses

**To augment diagnosis and treatment**

Created a portable solution that could be  
**deployed across 100s of sites**

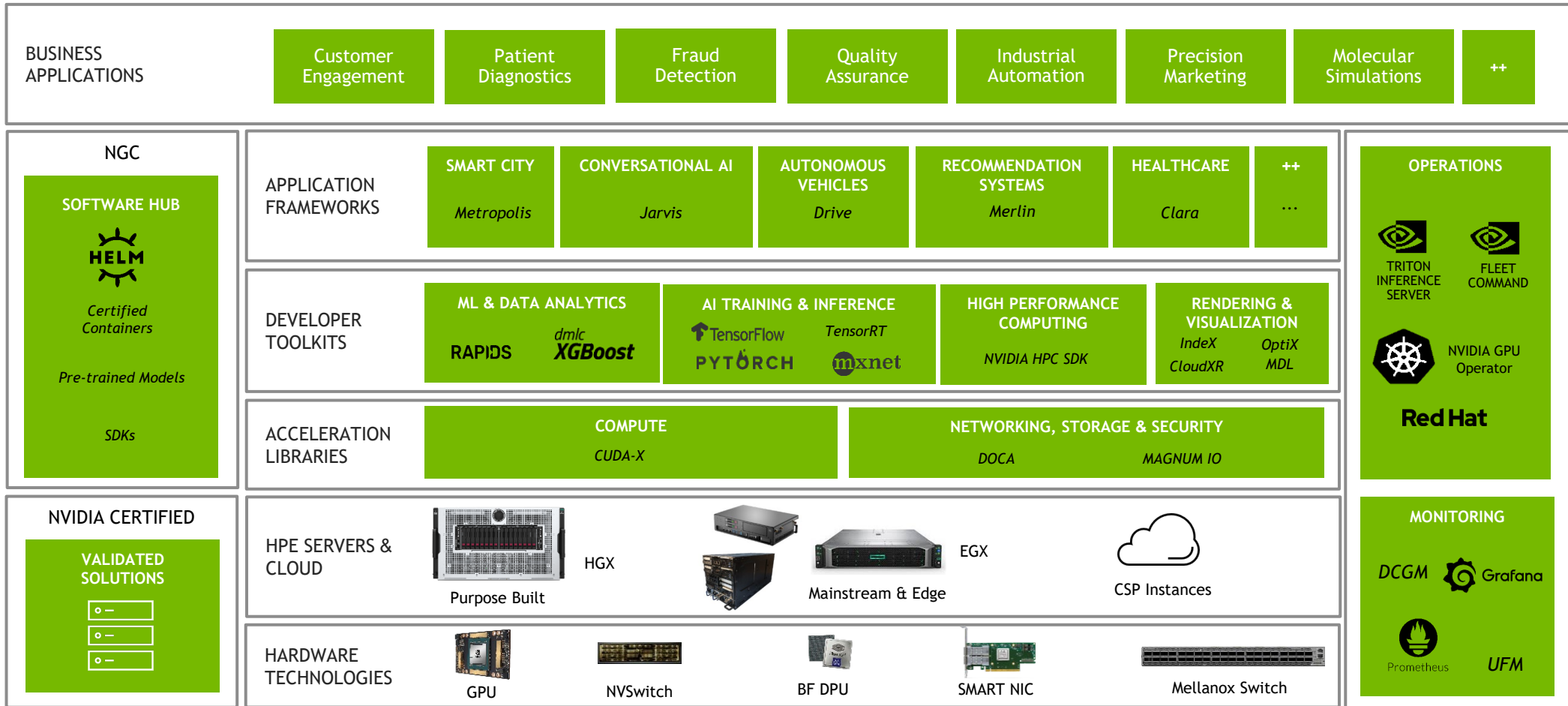
Gave data scientists the ability to rapidly **build, deploy & update ML models to aid Sepsis detection, helping to save lives**





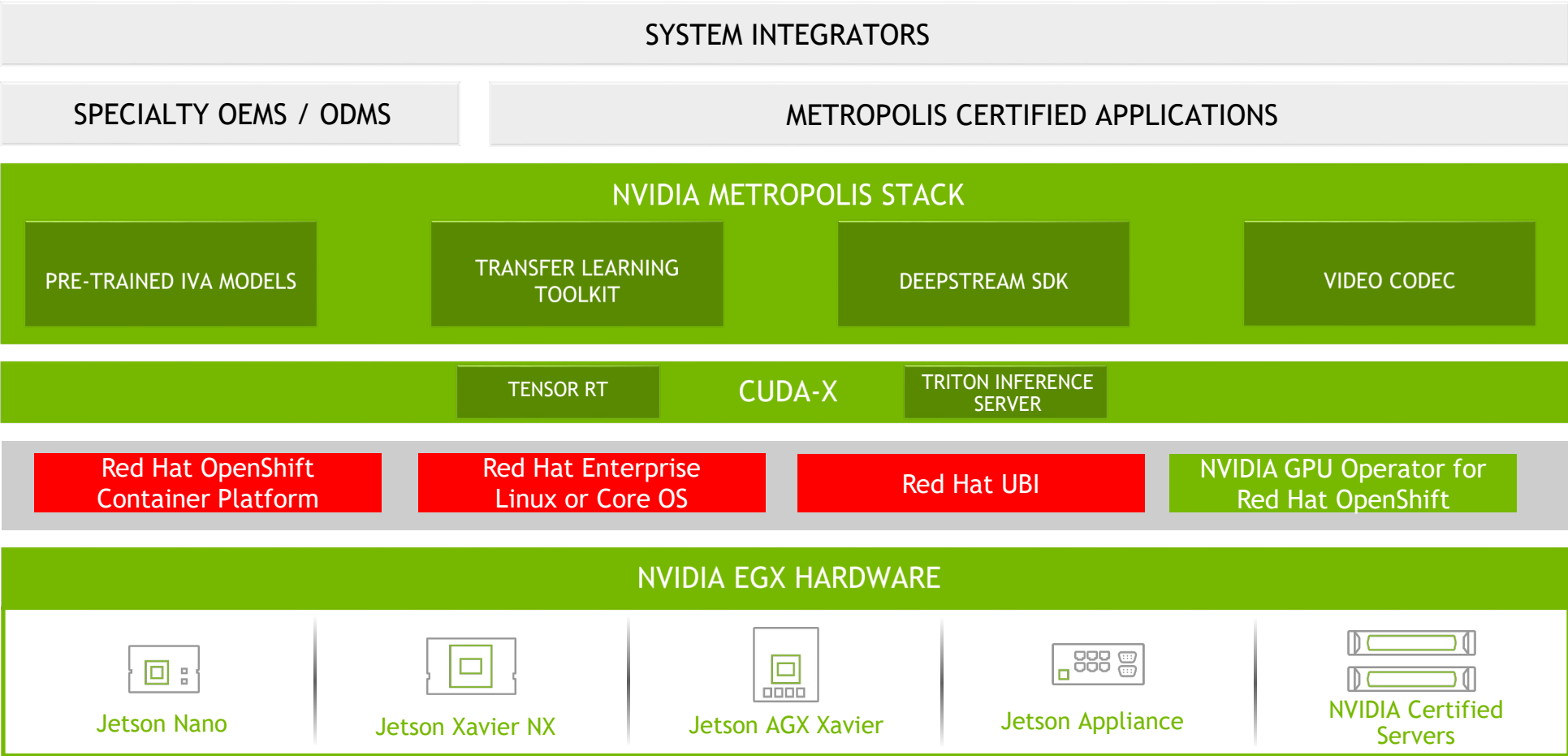
**Trent Boyer**  
North America Sr. OEM  
Account Manager  
**NVIDIA**

# NVIDIA DATACENTER PLATFORM



# NVIDIA METROPOLIS

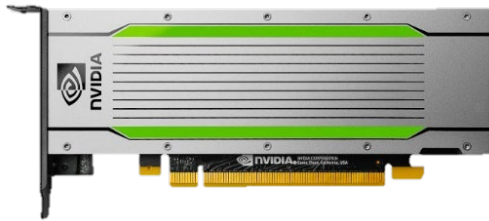
## AI Application Framework for Smart Sensors



# NVIDIA DATA CENTER GPUS

## NVIDIA T4

*Versatile Data Center GPU for Mainstream Computing*

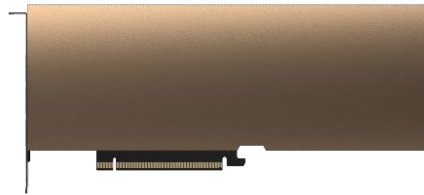


- Multi-Purpose GPU for Enterprise Acceleration, Graphics, Inference
- Entry — Mid Range RTX vWS
- 16GB GPU Memory
- Turing Tensor Cores

## NVIDIA A40

RTX 6000/8000 PASSIVE

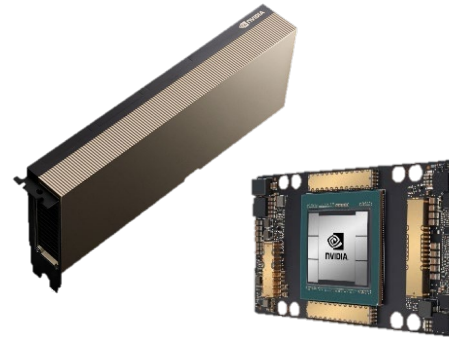
*World's Most Powerful Data Center GPU for Visual Computing*



- Fastest Interactive Graphics with NVIDIA RTX vWS & Ray Tracing
- Largest models, Mid - High-end RTX vWS<sup>1</sup>
- 48GB GPU Memory
- 2-Way NVLink
- Ampere Tensor Cores, RT Cores

## NVIDIA A100 40GB

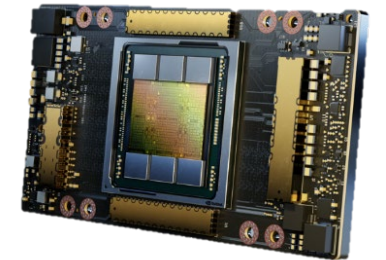
*World's Most Powerful Data Center GPU*



- Fastest Compute
- FP64 Precision
- Virtualization with NVIDIA vCS<sup>2</sup>
- 8-way NVLink and NVSwitch
- 40GB GPU Memory
- MIG for max utilization
- Ampere Tensor Cores

## NVIDIA A100 80GB

*World's Most Powerful Data Center GPU For Largest Workloads and Datasets*



- 80GB GPU Memory
- 1.3X faster memory bandwidth and world's first at >2 TB/s
- 7MIGs at 10GB each

1. NVIDIA A40 support available with vGPU (12.0) Winter 2021 release  
2. NVIDIA A100 support available with vGPU (11.1) Fall 2020 release

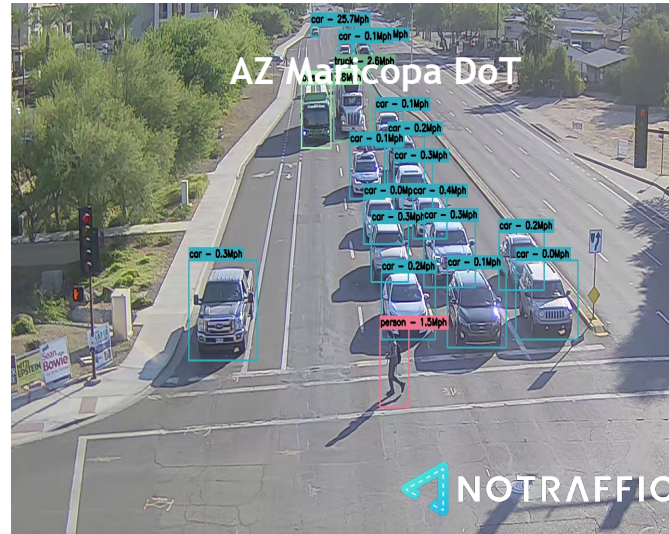
# CUSTOMER CASE STUDIES



**Challenge:** (BART) was losing over \$25M per year in fare evasions across its 48 stations

**Solution:** Irvine Sensor repurposed BART's existing security cameras by applying IVA and NVIDIA GPU's to study consumer behavior and patterns at its most affected stations

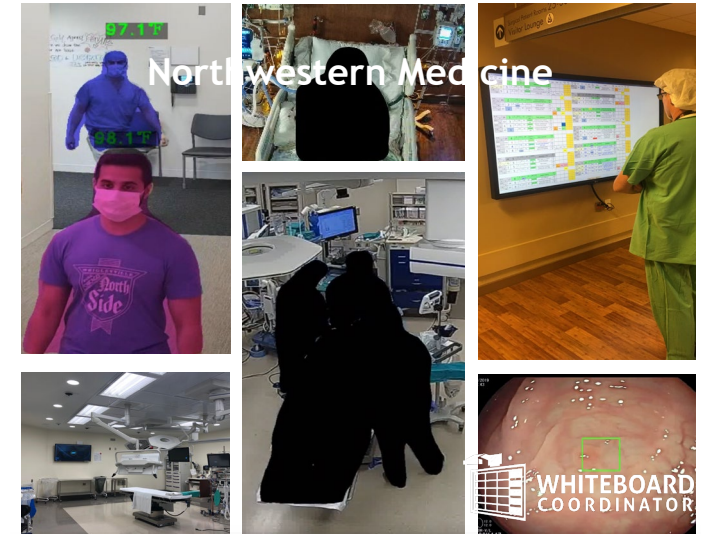
**Result:** Initiative addressed \$25M in losses while improving customer experience with non-intrusive technology



**Challenge:** Serious traffic congestion

**Solution:** NoTraffic, an NVIDIA Metropolis partner, gave them the solution they needed with an AI-based traffic management platform that optimizes flow based on real-time utilization and demand.

**Result:** Estimate's show \$50M/year savings, eliminating 3,957 months of vehicle delay time, and the CO2 effectiveness of removing 4,915 vehicles from their roadways



**Challenge:** Hospitals lack clean data, which negatively impacts patient care, operations, and reimbursement.

**Solution:** Whiteboard Coordinator, a suite of AI solutions for healthcare, uses NVIDIA Clara Guardian SDK on NVIDIA T4 and Quadro RTX to improve efficiency, operations, and patient care in hospitals.

**Result:** 16% operating room productivity gain within 6 months. Reduced nurses' need to enter COVID-19 rooms by as much as 54%

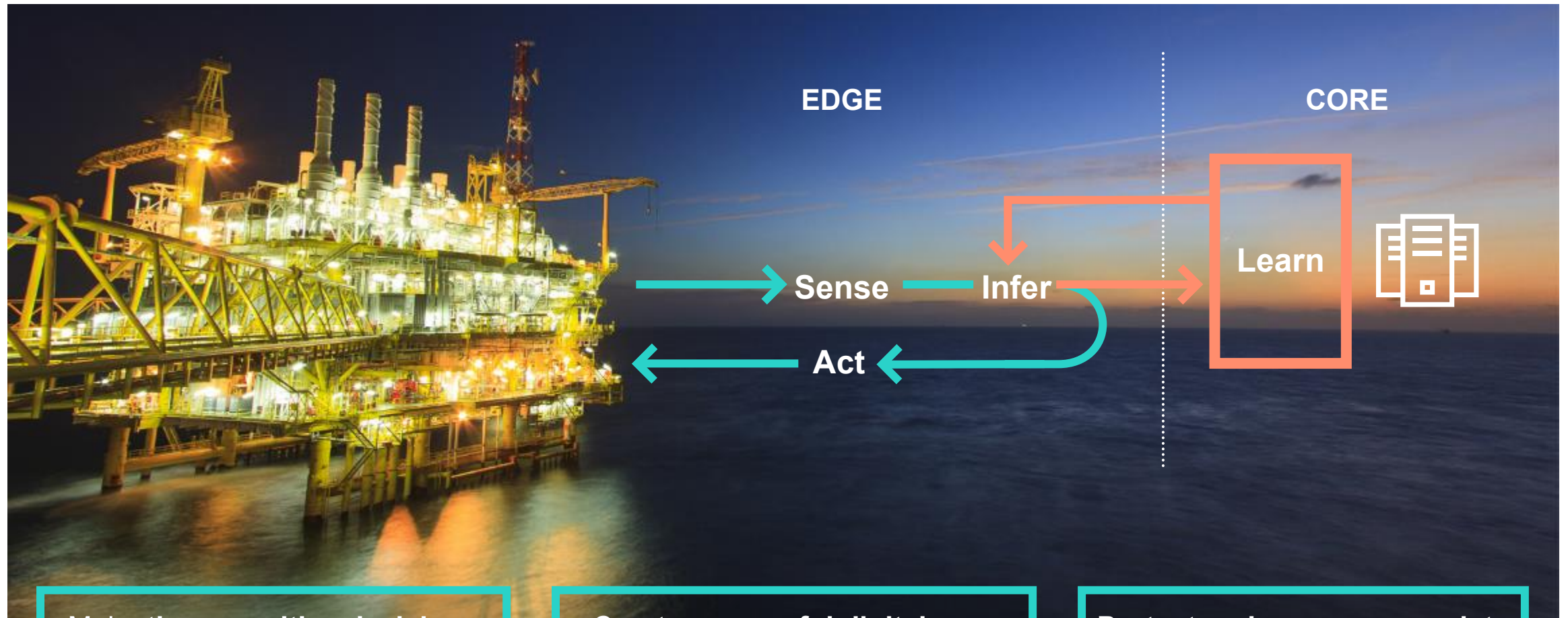


# Hewlett Packard Enterprise



**Joe Vidal**  
Master Technologist,  
Office of the NA CTO  
**HPE**

# Intelligent Edge solutions for Manufacturing



Make time-sensitive decisions  
at the edge of your network  
**HPE Edgeline & NVIDIA**

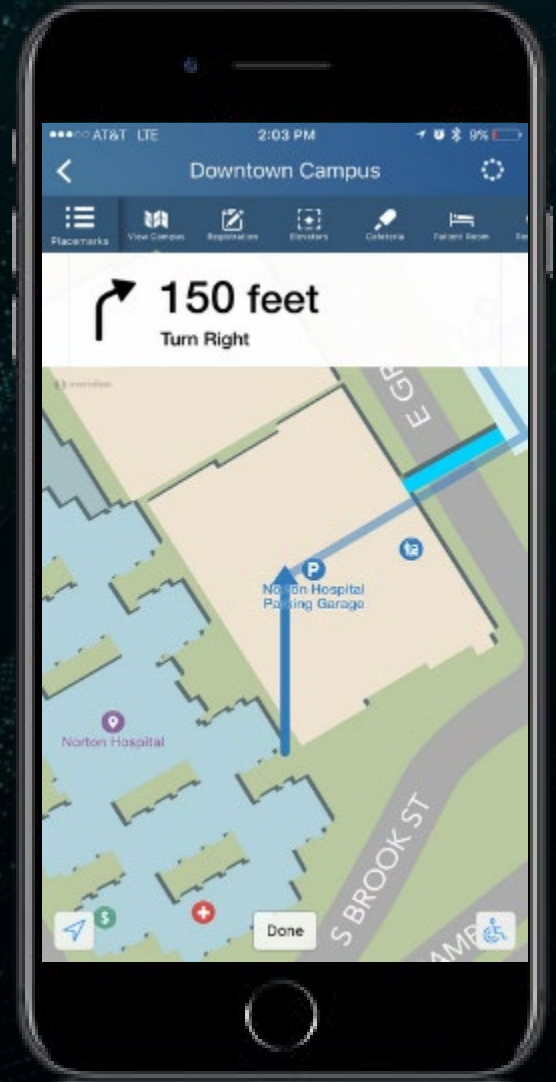
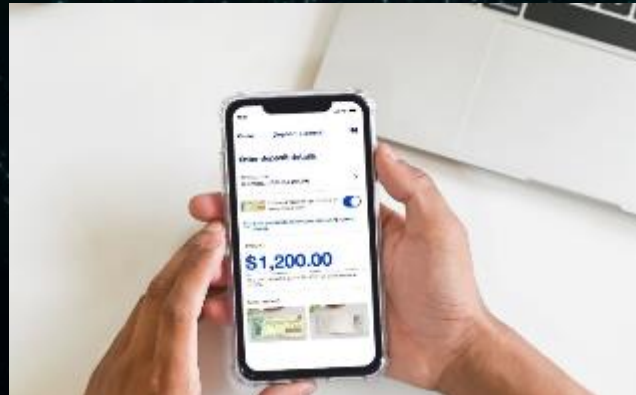
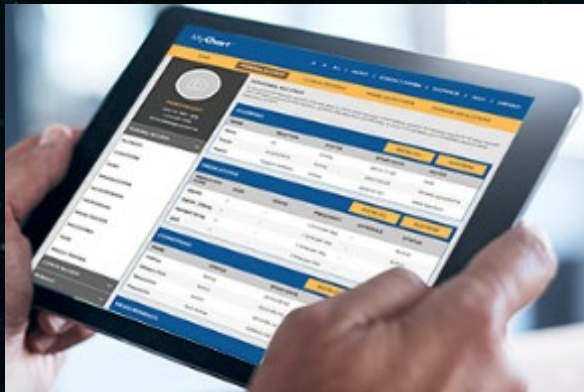
Create a powerful digital core  
**HPE High Performance Computing:  
NVIDIA + Red Hat Containerization**

Protect and manage your data  
**Security Built-In Enabling  
Distance deep learning framework**

# TRANSFORMING FIN-TECH PERSONALIZED BANKING



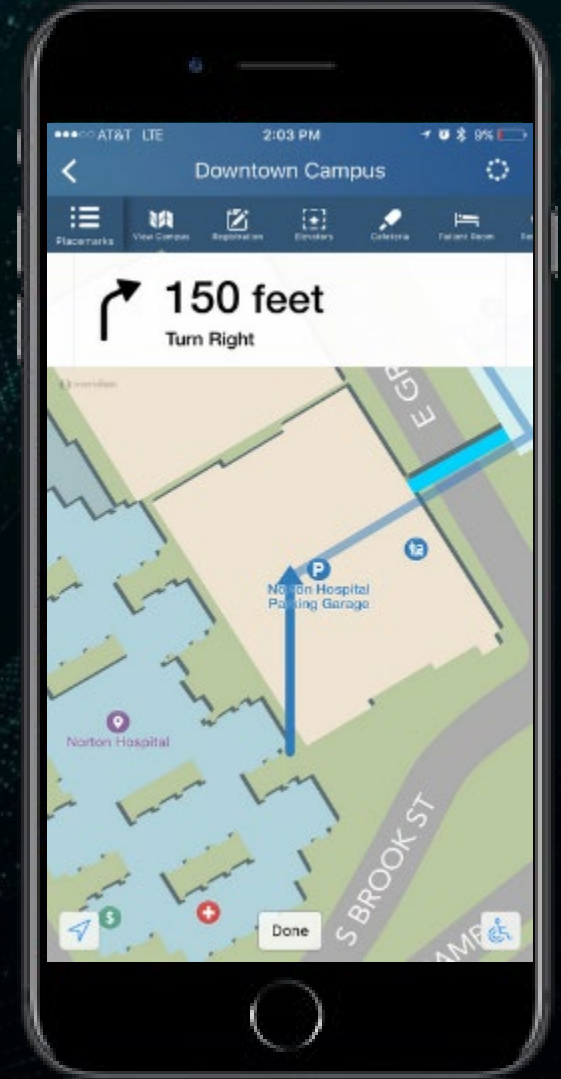
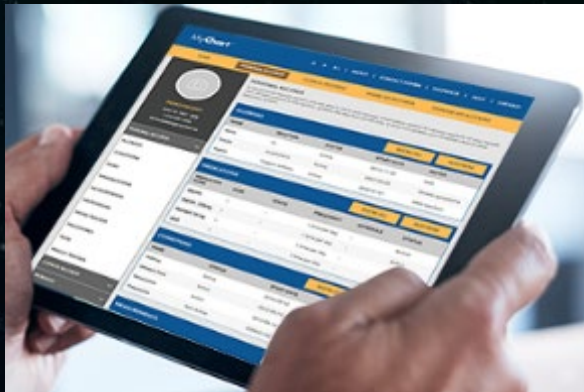
Appointment data, electronic banking records (via Meridian integration) live chat and e-arrival options along with asset tags, indoor navigation and proximity based triggers



# TRANSFORMING HEALTHCARE



Appointment data, electronic health records (via Epic integration) live chat and e-arrival options along with asset tags, indoor navigation and proximity based triggers



# Levi's Smart Stadium

Smartest stadium in the world



## Requirements

- The \$1.2 billion venue has 1.85 million square feet, seats 68,500 and features 165 luxury suites and 8,500 club seats
- Levi's wants to offer a premier experience, capitalizing on its location in Silicon Valley

## Solution Value Prop

- Aruba beacons deliver precise GPS positioning to 70,000 fans
- A new app to connect fans to retail services, facility locations and sports content

## Customer Value

- A heightened stadium experience through:
- GPS accuracy to direct fans to their seats
  - In-seat ordering of food and drinks
  - Information on what food is nearby as a fan walks around
  - Instant streaming replays and sports stats
  - Directions to nearest restroom with the shortest line

- Mobility Uncovering \$2M of New Revenue
- \$1.2M Food, Beverage, Parking
- \$750K In-App Sponsorships and 30% App Adoption Rate

# Super Bowl 50 Aruba – Enhanced Customer Experiences

Shattering Wi-Fi Records at Levi's Stadium and Enabling Mobile Engagement

*Delivering a feature-rich on-site experience that drives higher \$Revenue generation*

The Savings add up weekly

- **\$120K Saved per week** eliminating Ticket fraud
- **30K+ attendees** downloaded feature rich experience
- **HPE Meridian App** is Key
- Delivering **Location-Aware** apps
- **25% Higher Concession Revenue** per game



★ A STADIUM FIRST ★

THE FIRST SUPER BOWL TO DELIVER INDOOR NAVIGATION

Merchandise

97 feet  
Turn Left

2 Min Arrival

Kickoff In  
20 Minutes

★ KEY WI-FI STATS ★  
Crushing all past NFL records

10TB+	TOTAL DATA TRANSFERRED ON WI-FI 60% MORE THAN LAST SUPER BOWL	
3GBPS: FOUR HOURS	CONTINUOUS WI-FI BANDWIDTH	
27,000	UNIQUE WI-FI USERS 38% OF ATTENDEES	
20,300	MAX CONCURRENT WI-FI USERS	

To see how we do it, go to [www.arubanetworks.com/lpv](http://www.arubanetworks.com/lpv)

# LOCATION-AWARE ENHANCED CUSTOMER EXPERIENCE (Consumer Vertical)



Needed to **increase inventory turns & in-store profitability** while better understanding their customer's **buying experience** (both in-store and @home)



## Solution

- Replaced legacy Cisco with HPE Aruba Mobile First infrastructure
- Store wide optimized 802.11AC Wave2 WLAN
- Leveraged all TS portfolio to assess, design, deploy and support every store
- Consulting offered optimized solution for every store

## Value to CTO/CIO

- Modern network for Mobility experience
- Lower cost HPE Aruba products
- Lower Day2 support costs
- Reduction in wired switch ports
- Increased network security

## Value to business

- Reduced costs, improved security, more efficient management of network
- Better employee productivity and customer experience
- New opportunity for targeted marketing on mobile app

## Value to people

- Mobile app to navigate store and find products quickly
- Help from Home Depot associates at the touch of a button
- Special offers targeted to customer requirements
- Enhanced customer experience

Smart Store Solution leverages SAP Appliances from HPE to drive accelerated Inventory Turns and Higher Profitability

# KubeFrame for AI-Edge

**NVIDIA Accelerators, HPE Hardware, Red Hat Software**

*Exclusive to Public Sector through Carahsoft*

## 4x server blades

CPU (per blade)

vCPU (per blade)

RAM (per blade)

GPU (per blade)

SSD (per blade)

NVME (per blade)

## 24 core 2.4 GHz Intel proc

48 vCPUs (with HT enabled)

192 GB

NVIDIA T4

4x 2TB NVME

SSD

## server configuration

1 x bastion node and ingress router

3x bare metal OpenShift cluster


with hyper converged storage

**Built on Project Faros:**

<https://faros.dev>



Half-rack wide  
Half-rack deep  
5U  
~50lbs



**Thank you for attending!**  
*Please contact us with questions.*

**Rich Savage**  
Red Hat Sales Director  
703-871-8629  
877-RHAT-GOV (Team)  
[Rich.Savage@carahsoft.com](mailto:Rich.Savage@carahsoft.com)

Visit <https://carah.io/kubeframe> for more information.

Thank you for viewing this Red Hat, HPE & NVIDIA webcast! Carahsoft is the Master GSA and SLSA Dealer and Distributor for Red Hat Enterprise Open Source solutions available via GSA, SLSA, ITES-SW2, The Quilt and other contract vehicles.

To learn how to take the next step toward acquiring Red Hat's solutions, please check out the following resources and information:



For additional resources:

[carah.io/RedHatResources](https://carah.io/RedHatResources)



For additional Red Hat solutions:

[carah.io/RedHatPortfolio](https://carah.io/RedHatPortfolio)



To purchase, check out the contract vehicles available for procurement:

[carah.io/RedHatContracts](https://carah.io/RedHatContracts)



For upcoming events:

[carah.io/RedHatEvents](https://carah.io/RedHatEvents)



For additional KubeFrame resources:

[carah.io/KubeFrame](https://carah.io/KubeFrame)



To set up a meeting:

[redhat@carahsoft.com](mailto:redhat@carahsoft.com) or 877-RHAT-GOV