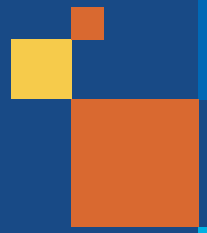
Intel® Network Builders Insights Series

Run Your Containers Anywhere With Amazon Elastic Kubernetes (EKS) on Intel

Mohan Potheri, Cloud Solutions Architect, Intel
Sathish Venkat, Field Applications Engineer, Intel



intel®

Notices and Disclaimers

Intel technologies may require enabled hardware, software or service activation.

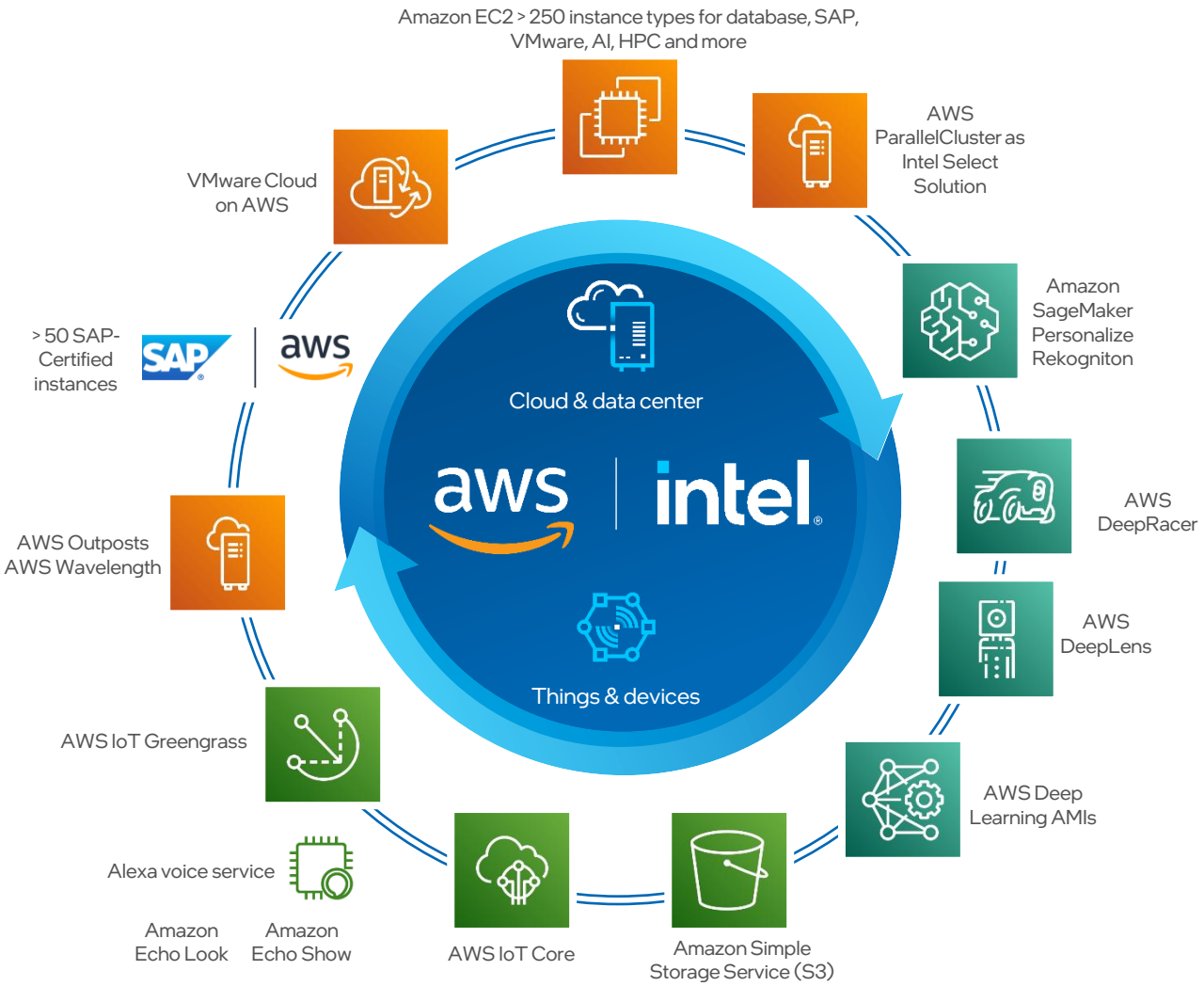
No product or component can be absolutely secure.

Your costs and results may vary.

Agenda

- Intel on AWS
- Kubernetes & Containers
- Amazon Elastic Kubernetes Service (EKS)
- Hybrid Cloud & EKS Anywhere
- Use Cases with EKS Anywhere
- EKS Anywhere Solution for AI/ML
- Key Takeaways

AWS Services Powered by Intel



- 15+ year engineering partnership
- Collaboration with AWS and its partners on Digital Transformation
- Shared customer passion
- High performance + low costs
- World class supply chain

3rd Gen Intel® Xeon® Scalable Processors

Performance made flexible

Only x86 data center processor with

Built-in AI & security solutions

Targeted for 1S-2S systems

Advanced security solutions

- Intel® Software Guard Extensions
- Intel® Crypto Acceleration
- Intel ®Total Memory Encryption
- Intel ® Platform Firmware Resilience

Scalable, flexible, customizable

- Intel® Deep Learning Boost
- Intel® Speed Select Technology
- Intel® AVX-512
- Optimized Software

Breakthrough Data Performance

- OPTANE — Intel® Optane™ persistent memory 200 series
- OPTANE — Intel® Optane™ SSD PS800K series
- Intel® SSD D series

AWS EC2 Instance Offerings - Optimized by Use Case

General Purpose

Balanced workloads

M6i *(NEW)*
Up to 128 vCPUs & 512 GB of memory - ICX

M5
Non-burstable CPU usage
SKX/CLX – 24C

M5d
M5 with local host attached NVMe SSDs
SKX/CLX – 24C

M5zn
M5 with local host attached NVMe SSDs
SKX/CLX – 24C

T3
Burstable CPU usage
SKX/CLX – 24C

Compute Optimized

Compute-intensive, HPC, data lakes, network appliances

C6i *(NEW)*
Up to 128 vCPUs & 256 GiB of memory – ICX

C5
High performance $/performance optimized
SKX – 18C
CLX – 24C

C5d
C5 with local host attached NVMe SSDs
SKX – 18C
CLX – 24C

C5n
C5 with up to 100 Gbps network bandwidth
SKX – 18C

Memory Optimized

High performance databases, in-memory databases

X2i *(NEW)*
Memory-optimized & up to 4,096 GiB of memory- ICX&CLX

R6i *(NEW)*
Up to 128 vCPUs & 1,024 GiB of memory – ICX

R5, R5b
Up to 768GiB of Memory
SKX - 24C

X1
One of the Lowest Price/GiB of RAM
HSX - 16C (4 socket)

X1e
X1 with Extended Memory Footprint
HSX - 16C (4 socket)

Z1d
Single threaded compute optimized with high memory
SKX - 12C

Bare Metal
8 Socket Xeon with 6 TiB Memory up to 24 TiB; SKX/CLX - 28C

Storage Optimized

High IOPS at low cost

I4i *(NEW)*
NVMe SSD Storage
New size with up to 128 vCPUs and 1,024 GiB of memory-ICX

I3
NVMe SSD Storage and Bare Metal Instances
BDX - 16C

I3en
NVMe SSD Storage and Bare Metal Instances
SKX – 24C

H1
Compute and Memory Balanced, Up to 16TB HDD Storage
BDX - 16C

D3, D3en
Up to 366 TB HDD Storage, Lowest Price/Disk Throughput Perf
CLX – 24C

Accelerated Computing

Accelerated WLs machine learning, 3D rendering

P3dn
P3 with Local Host Attached NVMe SSDs and up to 100Gbps Network Bandwidth
SKX – 24C

G4
2 NVIDIA Tesla M60 GPUs per CPU
CLX – 24C

F1
4 FPGAs per CPU
BDX - 16C

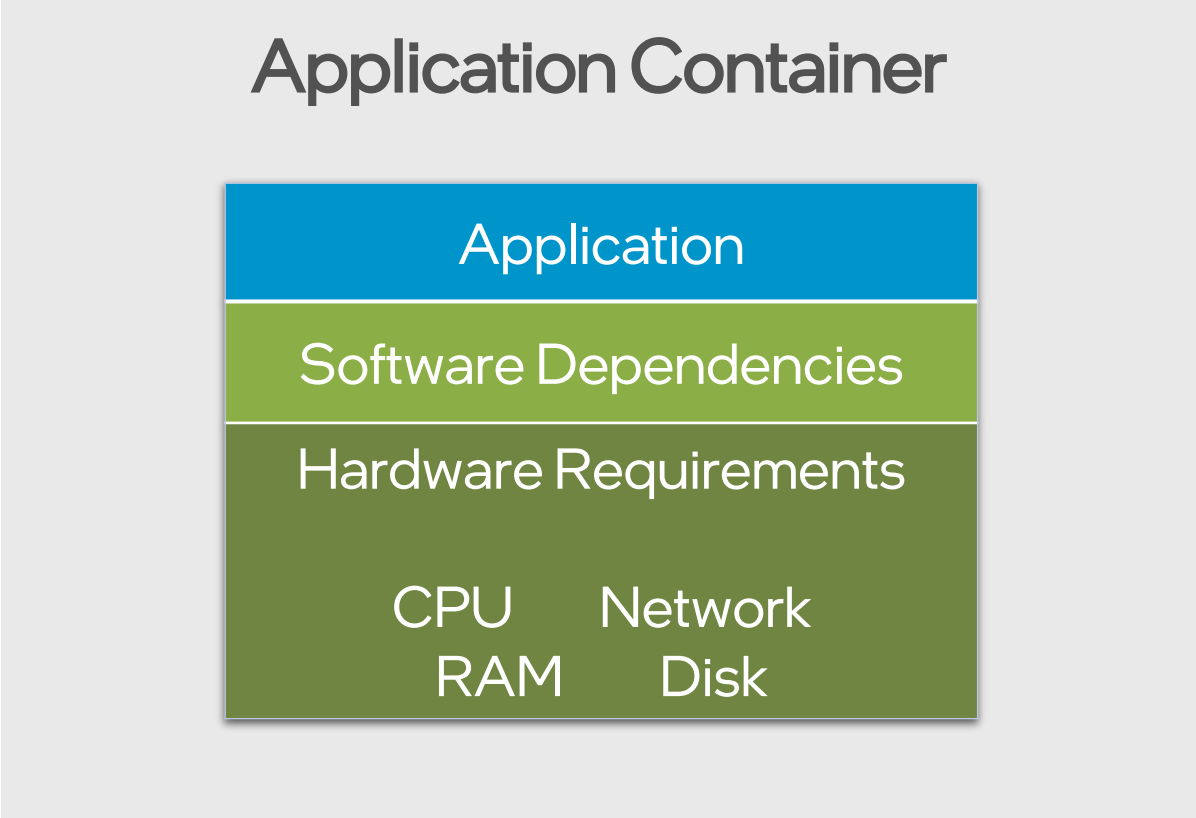
DL1 *(NEW)*
Habana Gaudi AI/ML
Up to 8 accelerators

See https://aws.amazon.com/ec2/instance-types/

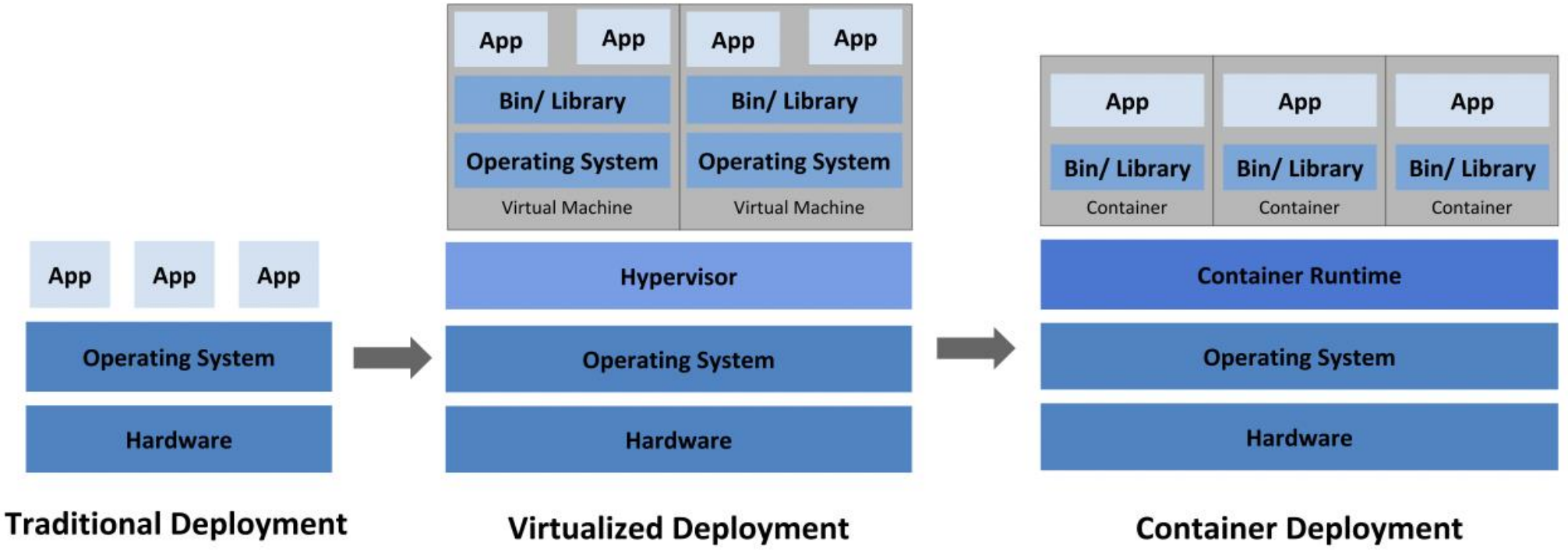
Containers are Revolutionizing Modern Workloads

Containers are

- The predominant mechanism for developing and delivering modern applications
- Highly efficient and can run multiple applications on a single server
- Self-contained as they include everything needed to run an application and hence Extremely portable
- Flexible and easy to use for developers

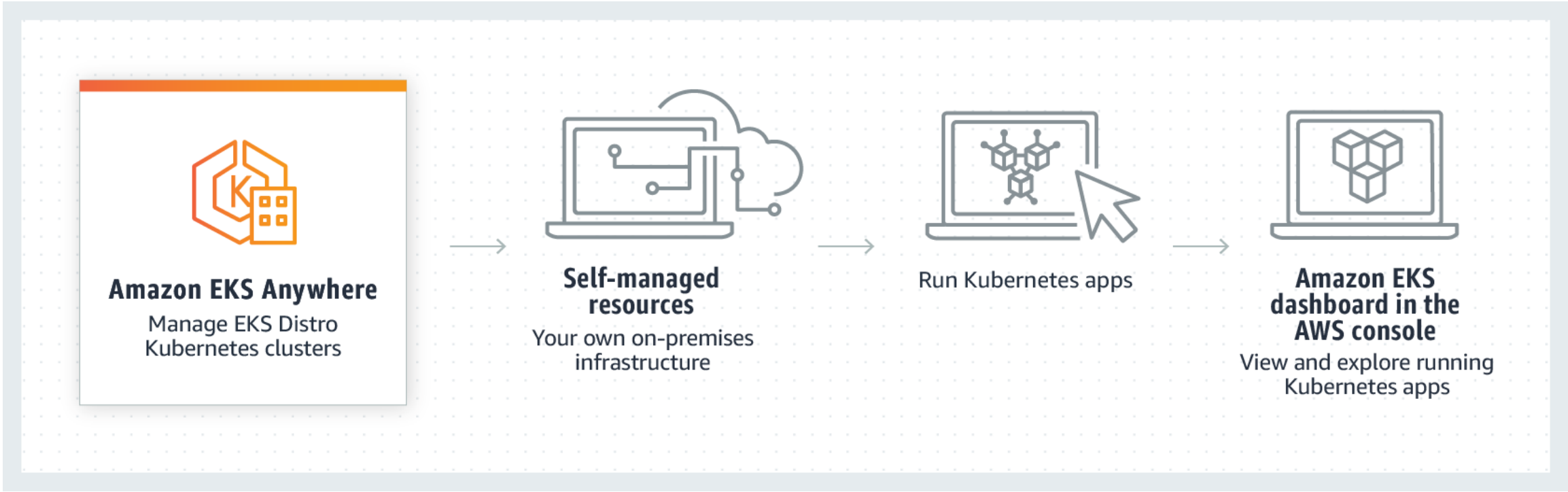


Emergence of Kubernetes



Source: https://kubernetes.io/docs/concepts/overview/

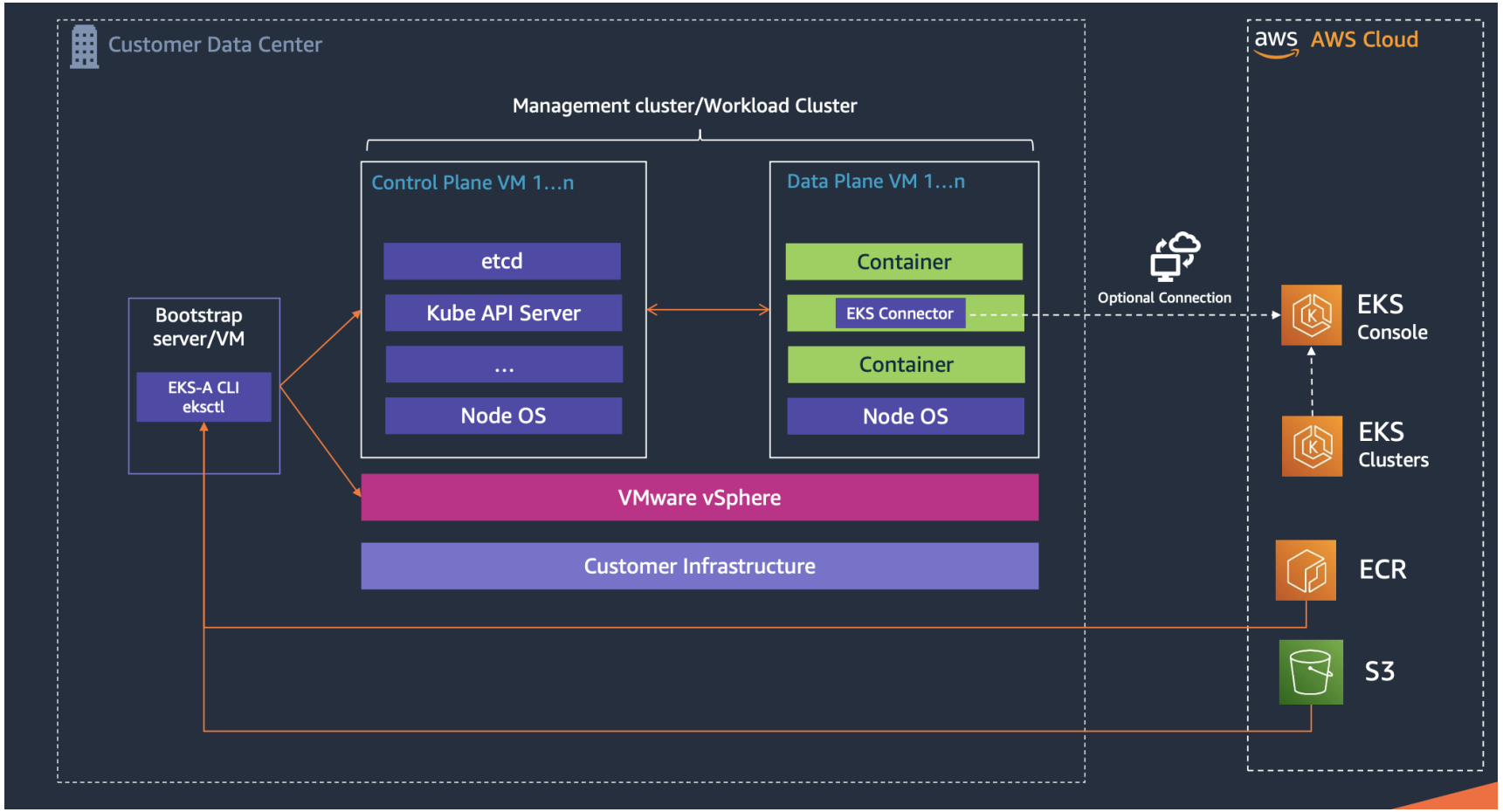
Amazon EKS Anywhere



Hybrid Cloud with AWS

- A hybrid deployment is a way to connect infrastructure and applications between cloud-based resources and existing resources that are not located in the cloud.
- The most common method of hybrid deployment is between the cloud and existing on-premises infrastructure to extend, and grow, an organization's infrastructure into the cloud while connecting cloud resources to internal system.

Hybrid Cloud with AWS



Key Use Cases for EKS Anywhere

Hybrid: Consistently run workloads on cloud and on-premises

Modernization: Containerize existing on-premises apps

IOT: Data Processing on Edge locations

IOT

- Containers are an ideal way to install and update applications that run on IoT devices
- Lightweight nature of containers makes them ideally suitable for resource constrained IOT devices
- Enables easier development of software with easier updates for the endpoint devices
- Containers provide the ability for easily scaling IoT environments

Telecom

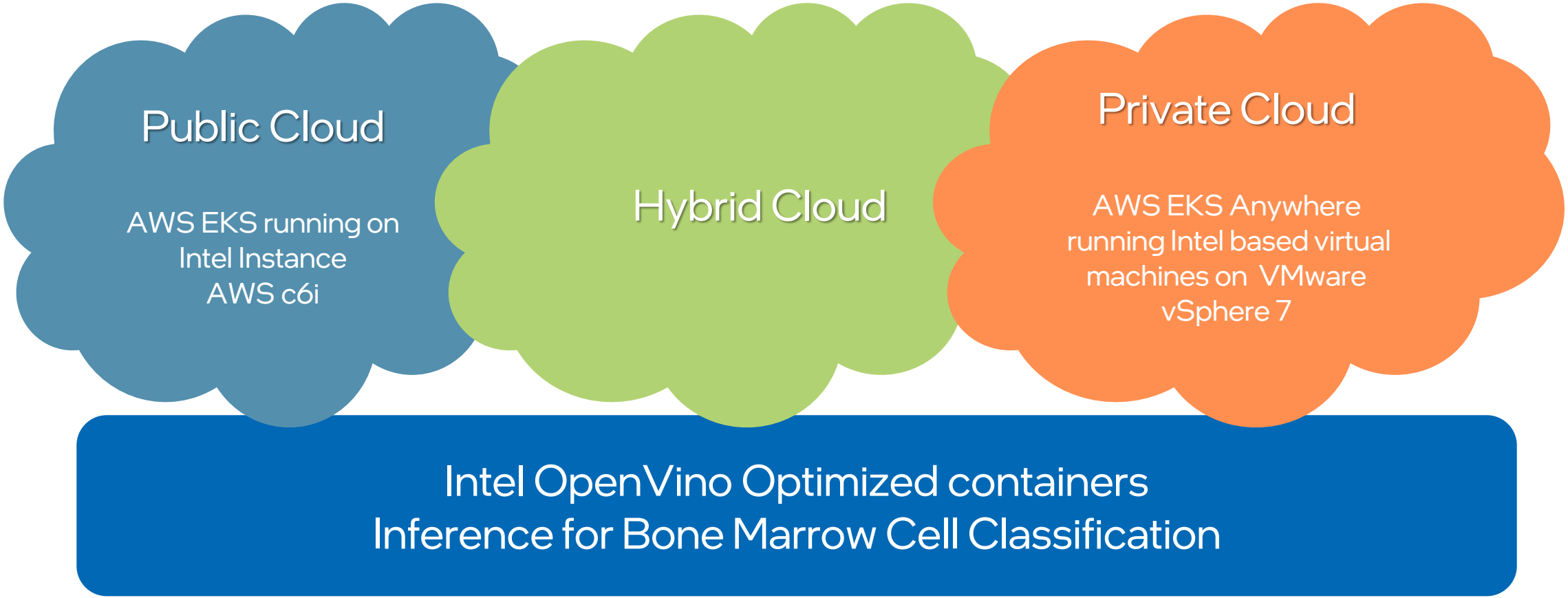
- Open RAN is an industry initiative led by network equipment providers and mobile network operators (MNOs) to design and deploy open, non-proprietary interfaces and protocols-based RAN elements on x86
- With Amazon EKS Anywhere, MNOs can now run, deploy and manage Open RAN solutions, including the distributed unit and centralized unit workloads
- Amazon EKS Anywhere can provide the same unified APIs across the cloud and on-premises environments.
  - RAN,
  - Core,
  - IMS,
  - OSS/BSS

Application Modernization

- Modern Applications leverage Microservices, which focuses on building single-function modules with well-defined interfaces and operations
- Containers are a great vehicle to deploy Microservices
- EKS Anywhere can be leveraged for orchestration of containerized Microservices

EKS Anywhere Hybrid Cloud Solution for Machine Learning

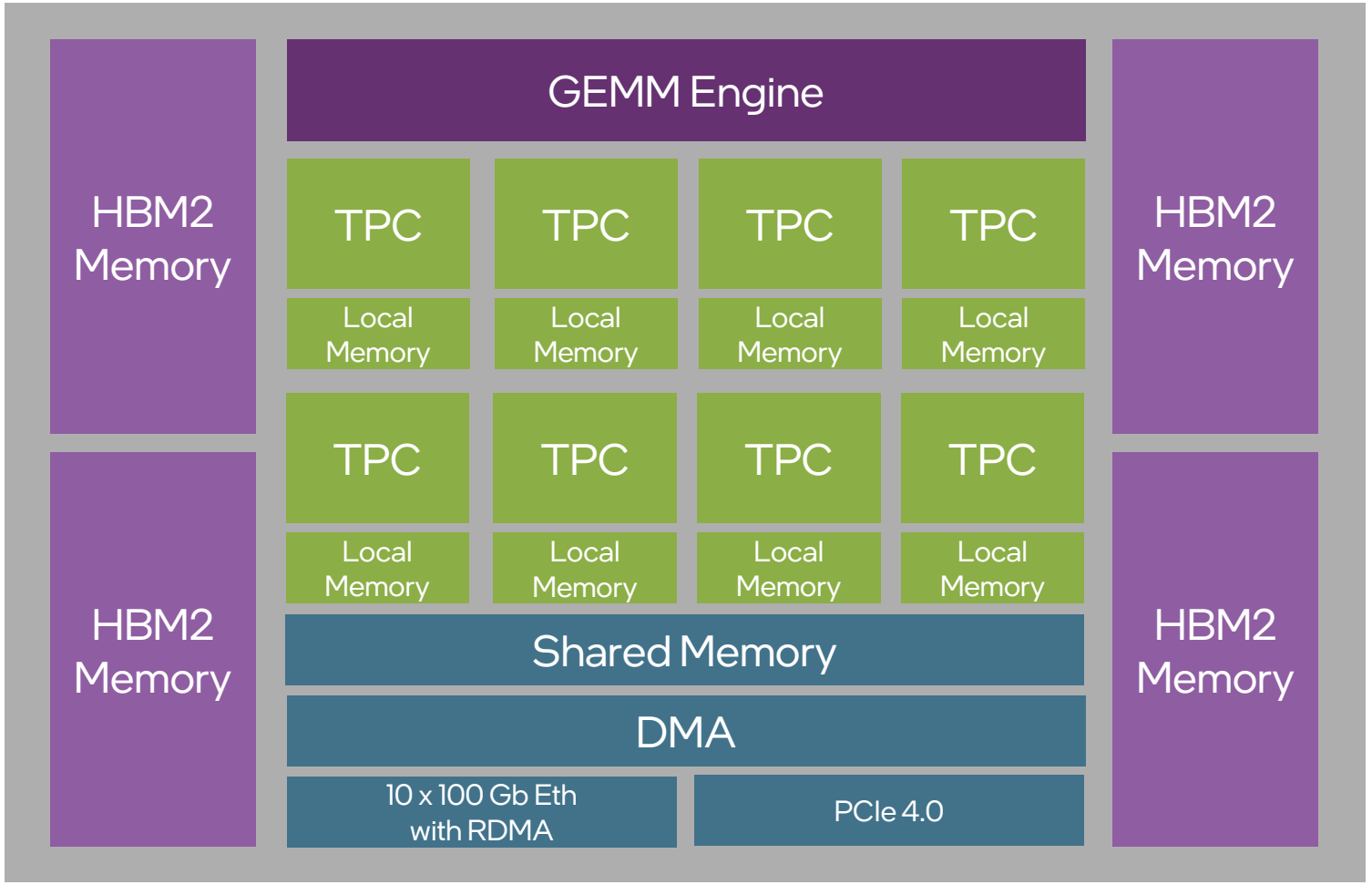
Solution Schematic with 3rd Gen Intel® Xeon® Scalable Processors



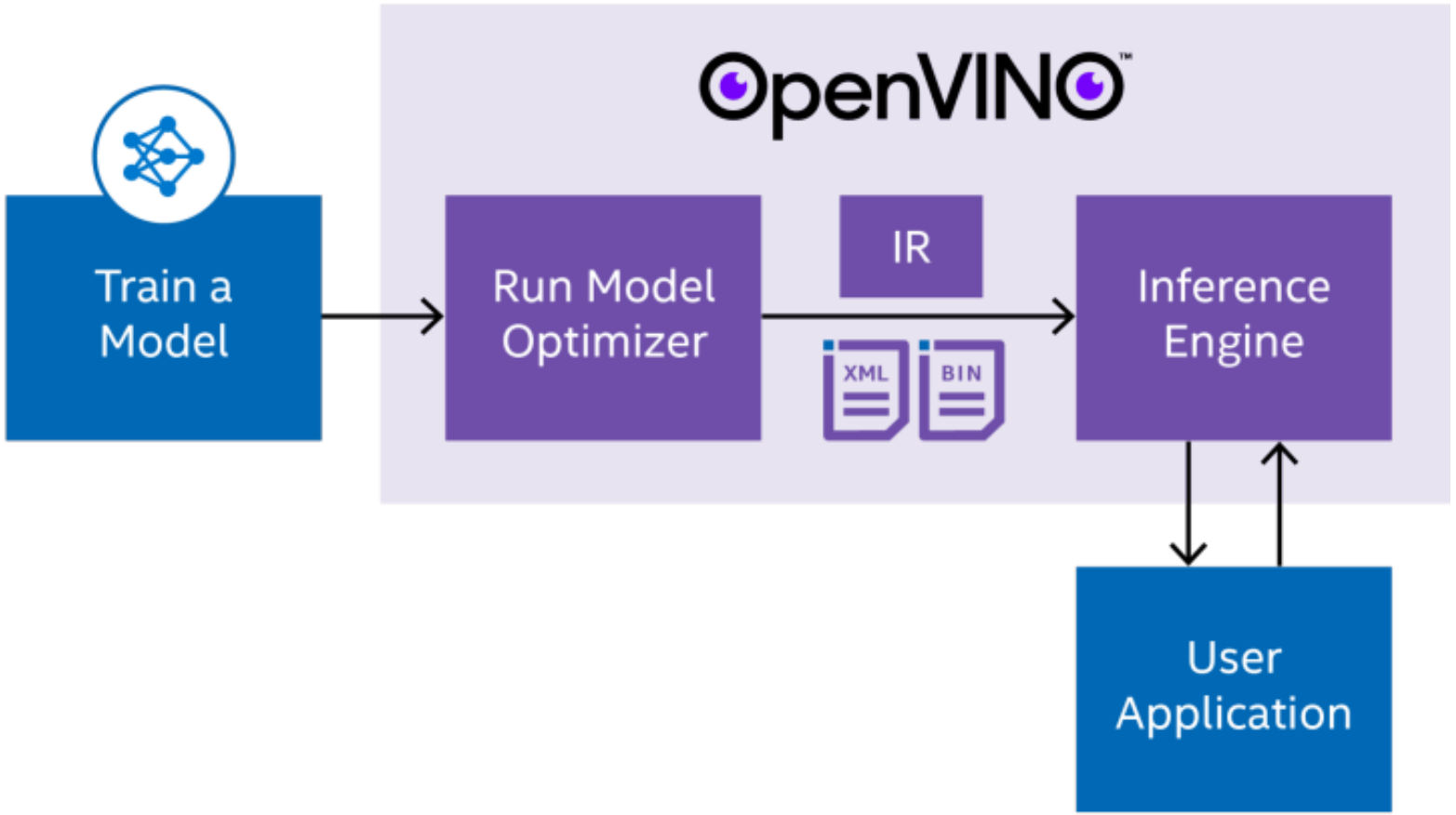
The Solution

- GPUs are needed for training complex machine learning models
- In this solution we take public data
- The use case is to train bone marrow cell classification
- We fit the data leveraging deep learning neural networks until we obtain a desired level of accuracy
- The model will be used for inference across different hybrid cloud endpoints with EKS Anywhere

AWS DL1 Instance with Habana® Gaudi® Processor Training and Inference using OpenVINO™ Toolkit



Machine Learning Training and Inference



Inference in the Cloud and Edge

- The trained models were optimized for inference with OpenVINO
- The models were then deployed on Intel based instances on AWS
- The same models were deployed on-premises
- Performance improvements were seen in inference between 2nd and 3rd Gen Intel® Xeon® Scalable processor instances
- The same model worked flawlessly across the hybrid cloud environment

Key Takeaways

- Intel® Xeon® technology is pervasive across Amazon AWS and on-premises
- Containers are the future of workloads
- AWS EKS and EKS Anywhere provide the ability to consistently deploy and manage across public and private clouds
- EKS Anywhere provides compelling use cases for
  - IOT
  - Telco
  - Application Modernization
  - Machine Learning

Questions?

Xiaojun (Shawn) Li, Sales Director, Next Wave OEM & eODM
xiaojun.li@intel.com

Mohan Potheri, Cloud Solutions Architect
potheri.mohan@intel.com

Sathish Venkat, Field Applications Engineer
venkat.sathish@intel.com

Intel® Network Builders Insights Series

Revisit our previous sessions

https://networkbuilders.intel.com/social-hub/webcast/insights-series


intel