Extending Containerized Infrastructure to the Edge with GPUs

Jered Floyd
Technology Strategist,
Office of the CTO
Edge Computing

MOTIVATION

LATENCY
BANDWIDTH
RESILIENCE
SECURITY
SOVEREIGNTY
INTEROPERABILITY
COST

CHALLENGES

TECHNOLOGY
SCALE
PEOPLE & EXPERTISE
PROCESSES & DATA
ENVIRONMENTAL
SECURITY
COST

AGGREGATE
DISTRIBUTE
Edge Tiers

Centralize where you can, distribute where you must.

GPU-enablement at any level.

Provider/Enterprise Core

Provider Edge

End-User Premises Edge

Device or Sensor

“last mile”
Extending the Open Hybrid Cloud Vision with Edge Computing

Any workload, any footprint, any location.
Edge Infrastructure Topologies

**Standalone cluster(s)**
- 100-1k’s of sites / 3-20 nodes per site

**Remote Nodes**
- 1k-10k’s of sites / 1-5 nodes per site

**Device Edge**
- 100k-10M’s of sites / 1 node per site
Distributed Nodes

- Single cluster deployment
- Primary site has shared control plane (and resource nodes)
- Remote sites have only resource nodes

Benefits
- Smaller footprint at the remote sites
- Faster to scale to new location (resource scale out)
- Easier operational management (single cluster, single config)

Complications
- Control plane is still a single point of failure
- Network drop affects management of workloads
5G Requirements

Real Time requirements
  Latency sensitive => 1ms processing budget at DU node

Timing synchronization
  vRAN applications require host and app time synchronization with <1us accuracy

Accelerator support
  Data plane accelerators for Split RAN architecture

High throughput
  Dedicated CPU and thread pinning and other performance tuning
Platform for 5G – Cloud Native is Key

5G Platform

- Load Balancing, Scaling / Elasticity
- Discovery
- Resilience
- Tracing
- Invocation Messaging / IPC
- Container Mgmt.
- Build, Deployment Pipeline
- Monitoring
- Authentication
- Logging

- PaaS (cloud-based)
- Edge Compute
- Application lifecycle management
- Mobile
SDOs, Industry Alliances and Open Source Communities

Standards, Specifications & Use Cases
- 3GPP
- ITU
- IEEE
- NGFI, TSN, PTP
- IETF
- CPRI
- Common Public Radio Interface
- ETSI
- World Class Standards
- NFV, MEC, ZSM
- SMALL CELL FORUM
- Solving the HotNet puzzle
- GSMA
- ngmn
- the engine of broadband wireless innovation

SW & HW Ref. Implementations, PoCs
- O-RAN
- TELECOM INFRA PROJECT
  vRAN FH, OpenRAN, DCSG
- OPEN AIR INTERFACE
- NW & Telco-WG
- SUPERFLUIDITY
- 5TONIC

Management, Orchestration, Control, Infrastructure Development & Integration
- ONAP
- OPEN DAYLIGHT
- Ansible
- OPNFV
- Edge, CRAN, VCO
- ANKAINO
- EDGE STACK
- CLOUD NATIVE COMPUTING FOUNDATION
- Prometheus
- UCSC/5G, Edge

Node Level Mechanisms
- OPEN
- CONTAINER INITIATIVE
- Real-Time LINUX
- QEMU
- KVM
- libvirt
- DPDK
- DATA PLANE DEVELOPMENT KIT
- Open vSwitch
- vowed.io
- redhat.

Use Cases
- UCSC/5G
- Edge, CRAN, VCO
Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.