OpenShift Commons Briefing:
Kubernetes 1.8: Features and Future

Derek Carr - Lead Engineer, Kubernetes
Clayton Coleman - Architect
What’s new this time around?
BIGGER THAN EVER

● 2000+ pull requests and 2500 commits
● 380+ committers
● 39 features across 29 SIGs and 5 WG
● 4 features to stable, 16 features to beta

And that’s just the technical part
COMMUNITY IS THE BEST FEATURE

- Kubernetes Steering Committee Elections
- New top level SIG - SIG-Architecture
- Formalizing our proposal process (KEP)
- Continuing investment in the community

Everyone is welcome here!
Focus on Stability
STABILITY IS A FEATURE

● Strong focus on fixing bugs
1450 issues resolved since 1.7 (June 30th)

● Mature existing features to beta or stable
20 features moved to beta or stable

● Production matters
Refine, polish, scale, tighten
EVENTS AND SCALE

- Events expose important transient info to users
- Key debugging tool
- In large clusters, excessive events observed
- FIX: Streamline sources
- FIX: Client-side rate limit
- 1.9: Server-side throttle
DENSE CLUSTER SCALING

- On very dense clusters clients retrieve a lot of data
- Alpha: gzip compression
- Alpha: clients can receive results in chunks
- Result: reduce 95th percentile latencies and improve client experience
MONITORING

- Prometheus cluster monitoring tech preview in OpenShift 3.7
- Fixed 10+ bugs around component monitoring
- Used experience at scale to guide new metrics
- Improve observability

Improvements in network and memory use before and after etcd3 migration
Features
METRICS AND AUTOSCALING

- Beta metrics API for pods and nodes, replacing heapster
- Custom metrics API to expose arbitrary application metrics
- HPA v2 autoscaler targeting custom metrics
- Improve visibility into HPA status
- Lay foundations for usage based scheduling
EXTENSIBILITY EVERYWHERE

- FlexVolumes can more easily be deployed
- New opportunities for node extension
- Custom resource definitions support validation
- Initializer and admission webhook bug fixes
- CRI continues to mature (see CRI-O)
RESIZE AND SNAPSHOT VOLUMES

- Grow PVCs by updating request size
- Transparent increase by provider
- Ask for a volume to be snapshotted
- Create a new PVC from an existing snapshot

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: snapshot-pv-provisioning-demo
  namespace: myns
  annotations:
    snapshot.alpha.kubernetes.io/snapshot: pv1-1
spec:
  accessModes: [ "ReadWriteOnce" ]
resources:
  requests:
    storage: 4Gi
  storageClass:
    storageClassName: snapshot-promoter
```
NETWORKING

- Alpha IPVS kube-proxy implementation
  - Promises better failover handling and detection

- Network policy improvements
  - Pod egress support added to NetworkPolicy
  - CIDR rules for pod matching
PERFORMANCE SENSITIVE APPS

● Goal
  ○ Support more workload types
  ○ With better performance
  ○ Without sacrificing reliability

● Focus areas
  ○ CPU management
  ○ Device plugins
  ○ HugePages
CPU PINNING POLICIES

- Static CPU pinning (Alpha)
  - Pods can run on exclusive core for their life
  - Improve latency for Guaranteed QoS
  - Takes physical processor topology into account
- Future CPU management policies proposed
DEVICE PLUGINS

- Plugin model (Alpha)
  - Discovery of external devices
  - Making devices available to pods
  - Health checking of devices
- Work done across ecosystem to validate model against GPUs, NICs, FPGAs, etc.
HUGEPAGES

- Alpha support to enable pods to request and consume huge pages
- Benefits applications that have a large memory working set
  - Database management systems
  - Java applications
  - Packet processing systems
CRI-O is nearing production readiness

Performance, reliability, and overhead improvements. Stay tuned!
Kubernetes 1.9
(SOME) 1.9 GOALS

- Stability and bug fixes (obv!)
- Everything is extensible
- Scaling improvements
- Initializers and webhook admission being used internally
- De-scheduler
- Priority and preemption
- More node level improvements in resource management
- CRI-O to stable on 1.8
- Get volume snapshots and resizing to beta
- Workload APIs to stable?
- Metrics used in the scheduler
- Better Prometheus integration into metrics
- PodSecurityPolicy to stable?
- Block device support
Questions?

1.8 releases today!

Clayton Coleman - @smarterclayton
Derek Carr - @derekwaynecarr