CVMFS: Stratum0 in Kubernetes

David Schultz, Heath Skarlupka, Vladimir Brik, Gonzalo Merino

WIPAC, UW-Madison
The IceCube Neutrino Observatory

Design and construction

Detector Design
- 1 gigaton of instrumented ice
- 5,160 light sensors, or digital optical modules (DOMs), digitize and time-stamp signals
- 1 square kilometer surface array, IceTop, with 324 DOMs
- 2 nanosecond time resolution

IceCube Lab (ICL) houses data processing and storage and sends 100 GB of data north by satellite daily

Detector Construction
- 7 seasons of construction, 2004-2011
- 28,000 person-days to complete construction, or 77 years of continuous work
- 4.7 million pounds of cargo shipped, 1.2 million of which was the drill
- 48 hours to drill and 11 hours to deploy sensors per hole
- 4.7 megawatts of drill thermal power with 200 gallons of water per minute delivered at 88 °C and 1,000 psi

IceCube

IceTop

86 strings of DOMs, set 125 meters apart

60 DOMs on each string

DOMs are 17 meters apart

firm drill

Enhanced Hot Water Drill

Antarctic bedrock

IceCube Lab

Amundsen-Scott South Pole Station
IceCube CVMFS

Common software base

- 1 major version a year - 6 different platforms
  -- Includes Python, boost, ROOT, …
  -- OS platforms: CentOS 6 & 7, Ubuntu 14 - 18

- Common files:
  -- Test data, lookup tables
IceCube CVMFS

Sizes:

- 32GB shared data
- ~15GB per version per platform
  = 90GB total per version

Total size (raw): 500GB

Total size (/srv/cvmfs): 150GB
Motivations for Kubernetes

- Previous setup used VMs + NFS share
  - Slow disk, long build times
  - VM provisioning difficult - kickstart + puppet
- New setup uses Docker containers
  - Can test on a laptop
  - Easy to upgrade / rollback container version
  - New OS platform can be deployed in ~1 hour
Build Workflow

- Git push or manual trigger
- Software build on multiple platforms in parallel
- CVMFS transaction
- Rsync from shared FS to CVMFS overlay
- CVMFS publish

Sync wait
Build Workflow

- **Build Master**
  - Git push or manual trigger

- **Individual Containers**
  - Software build on multiple platforms in parallel

- **Stratum 0**
  - CVMFS transaction
  - Rsync from shared FS to CVMFS overlay
  - CVMFS publish
Build system - CI

Buildbot - [http://buildbot.net](http://buildbot.net)

- Others would probably work (Jenkins, etc)
- Buildbot chosen for Python config
  -- Easily extensible to perform any action
  -- Handles dependency steps, locking
Buildbot webpage
Reactions after 6 months

- Build automation is great
  -- Nightly builds of the latest internal software
  -- Push-button rebuild/publish across all platforms

- When Kubernetes is down, the Stratum0 is down
  -- It comes up without issue
Backup
Original Setup

- CentOS 6 VM
- CentOS 7 VM
- Ubuntu 14 VM
- Ubuntu 16 VM

SSH

NFS

NFS Server

Stratum0

Local Disk

rsync
# Shared Storage

apiVersion: v1
description: PersistentVolumeClaim
metadata:
  name: cvmfs-buildbot-worker-pv-claim
spec:
  accessModes:
    - ReadWriteMany
  persistentVolumeReclaimPolicy: Retain
#StorageClass: cephfs
resources:
  requests:
    storage: 1Ti

# Stratum0 Worker Deployment
https://raw.githubusercontent.com/WIPACrepo/cvmfs_buildbot/master/kubernetes/worker-cvmfs-centos7-stratum0.json
factory = util.BuildFactory()
factory.addStep(steps.Git(  repourl='git://github.com/WIPACrepo/cvmfs.git',  mode='full',  method='clobber',  workdir='build',))
factory.addStep(steps.ShellCommand(name='build cmvfs',  command=[  'python', 'builders/build.py',  '--src', 'icecube.opensciencegrid.org',  '--dest', '/cvmfs/icecube.opensciencegrid.org',  '--variant', util.Property('variant'),  ],  env={  'CPUS': util.Property('CPUS', default='1'),  'MEMORY': util.Property('MEMORY', default='1'),  },  workdir='build',  haltOnFailure=True,  locks=[  cfg.locks['cvmfs_shared'].access('exclusive')  ],))

factory = util.BuildFactory()
factory.addStep(steps.ShellCommand(  name='open transaction',  command=['cvmfs_server','transaction','icecube.opensciencegrid.org'],  haltOnFailure=True,))
factory.addStep(steps.ShellCommand(  name='rsync',  command=['cvmfs_rsync',  util.Interpolate('/cvmfs-source/icecube.opensciencegrid.org/$(prop:variant)s'),  util.Interpolate('/cvmfs/icecube.opensciencegrid.org/$(prop:variant)s'),],  haltOnFailure=True,  doStepIf=BuildPassed,))
factory.addStep(steps.ShellCommand(  name='publish transaction',  command=['cvmfs_server','publish','icecube.opensciencegrid.org'],  haltOnFailure=True,  doStepIf=BuildPassed,  hideStepIf=lambda results, s: results==SKIPPED,))
factory.addStep(steps.ShellCommand(  name='abort transaction',  command=['cvmfs_server','abort','-f','icecube.opensciencegrid.org'],  haltOnFailure=True,  doStepIf=BuildFailed,  hideStepIf=lambda results, s: results==SKIPPED,))
Improvements

- Git Push
- Software build on multiple platforms
- Test Software
- Snapshot Ceph Volume
- CVMFS Transaction
- Rsync From Shared FS to Block Volume
- CVMFS Publish