Slurm + GCP

Brian Christiansen - SchedMD
Keith Binder - Google

SLUG 2019

- GCP Plug/Overview
  - V2 features/func
- Slurm + GCP work
  - HPC Days demo?
    - This is 20-30 minutes
  - V3
- UNC POC
  - Bursting from on-premise
  - scenario/infrastructure
  - results/numbers?
  - nss-slurm

slurm
workload manager

Google Cloud

In 1H 2019

Slurm-GCP scripts spun up over 400K nodes within GCP
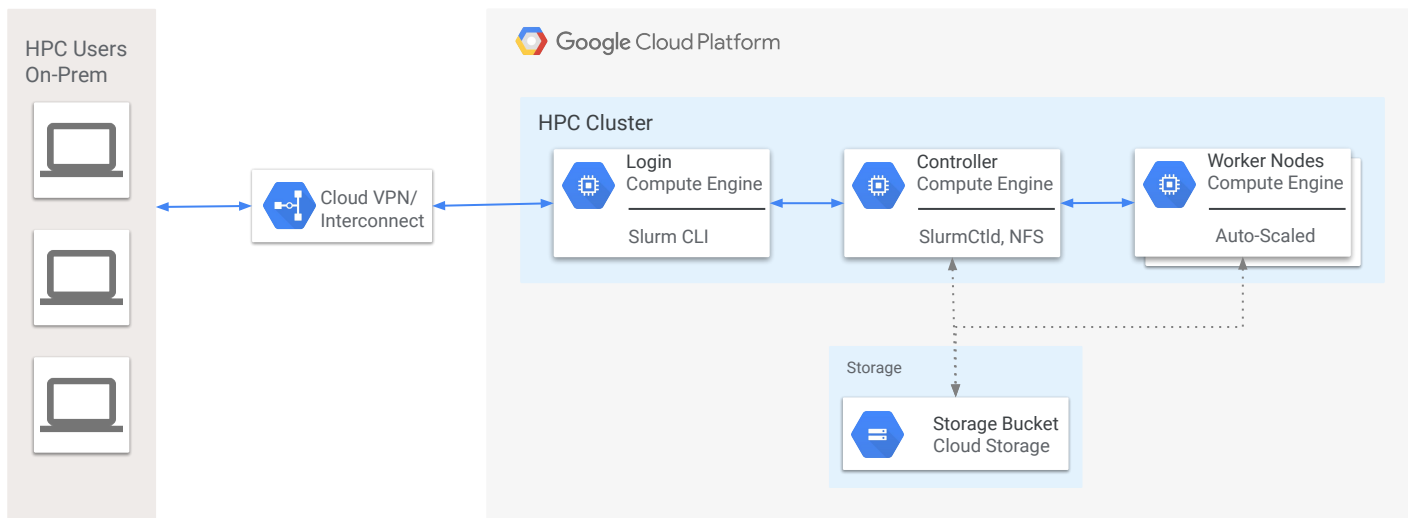
slurm
workload manager

Google Cloud

Slurm GCP Scripts v2 Improvements

- Preemptible VM support
- Image-based scaling
- Attachable GPUs
- Attachable secondary disks
- Scalability
- Burst to 5000 nodes < 10m
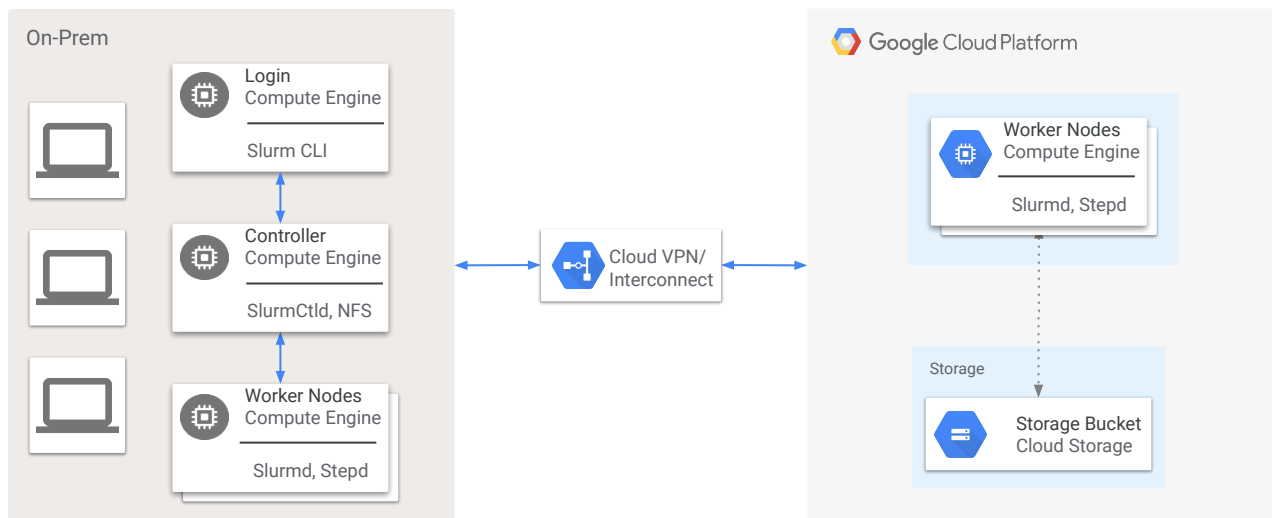- 19.05 cloud capabilities

Open Source on Github: https://github.com/schedmd/slurm-gcp

slurm
workload manager

Google Cloud

Copyright 2019 SchedMD
www.schedmd.com

Cloud Slurm

Architecture: Slurm Auto-Scaling Cluster

HPC Users
On-Prem

Google Cloud Platform

HPC Cluster

Cloud VPN/
Interconnect

Login
Compute Engine

Slurm CLI

Controller
Compute Engine

SlurmCtld, NFS

Worker Nodes
Compute Engine

Auto-Scaled

Storage

Storage Bucket
Cloud Storage

slurm
workload manager

Google Cloud

Hybrid Slurm

Architecture: Slurm Burst to Cloud

Slurm 19.05 Cloud Improvements

- Better responsiveness to creation and deletion of nodes
- Efficient allocation of nodes
- nss_slurm

slurm
workload manager

Google Cloud

Hybrid Slurm: Acme University

Foundational Elements

- Networking
  - VPN / Interconnect for hybrid networking
  - Bi-directional DNS for node resolution
- Storage
  - Mounted on-premise NFS
    - Dedicated Interconnect Recommended (10Gbps)
- Authentication
  - nss_slurm
- Compute
  - Pre-baked image with requisite drivers (i.e CUDA)

slurm
workload manager

Google Cloud

Acme Architecture
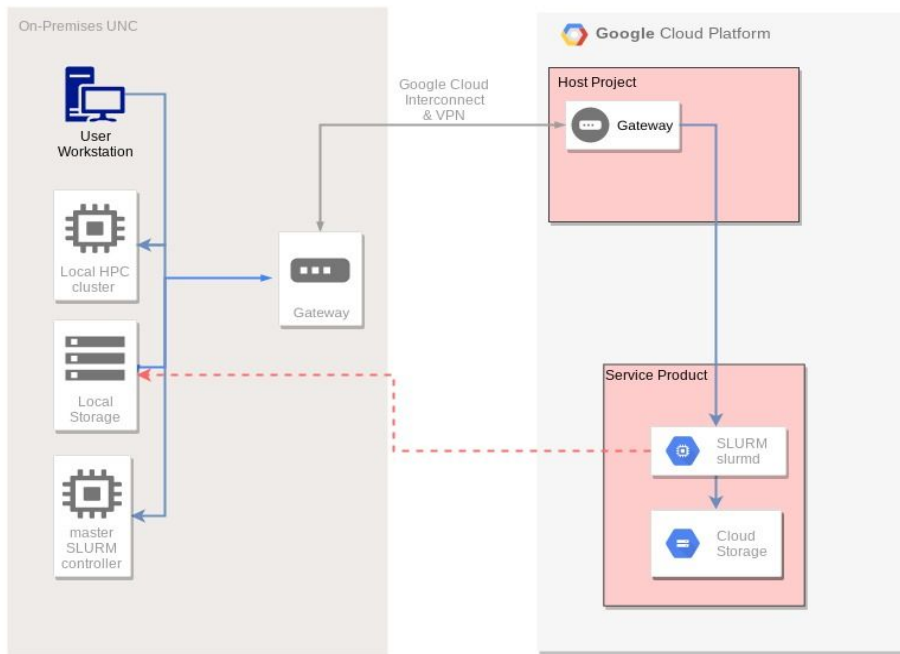
Acme Scale
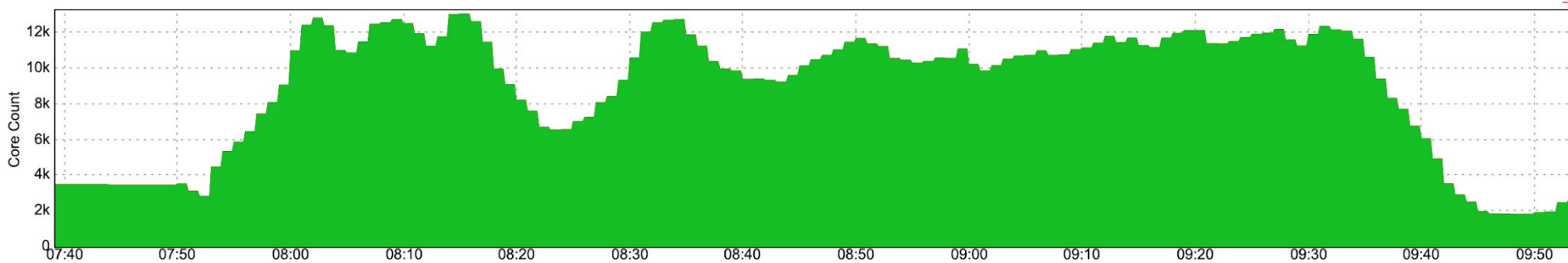
Bursted from 4K to 12K cores in 10 mins

Whats coming in 2020

- Multiple Partitions
- Terraform integration
- Even faster scaling
- Storage options flexibility
- NFS; LUSTRE; GCS
- OMPI Support
- External Database (Cloud SQL)
- Hybrid Configuration

slurm
workload manager

Google Cloud

Demo

Burst from 0 to 1000 nodes in under 10 min

Using Elastifile managed service

slurm
workload manager

Google Cloud

QUESTIONS?

slurm
workload manager

Google Cloud