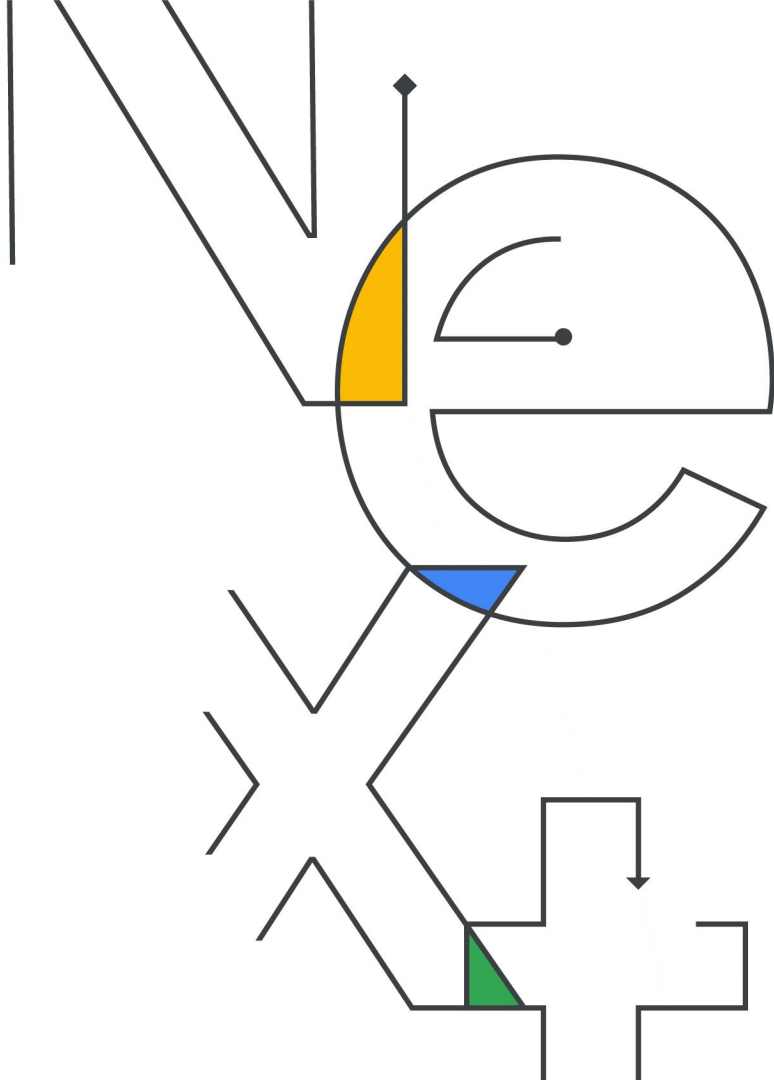


Google Cloud

Next '22

What's new in Kubernetes:
Run batch and high performance
computing on GKE

Oct/
11-13





Louis Bailleul

Chief Enterprise Architect
PGS



Maciek Różacki

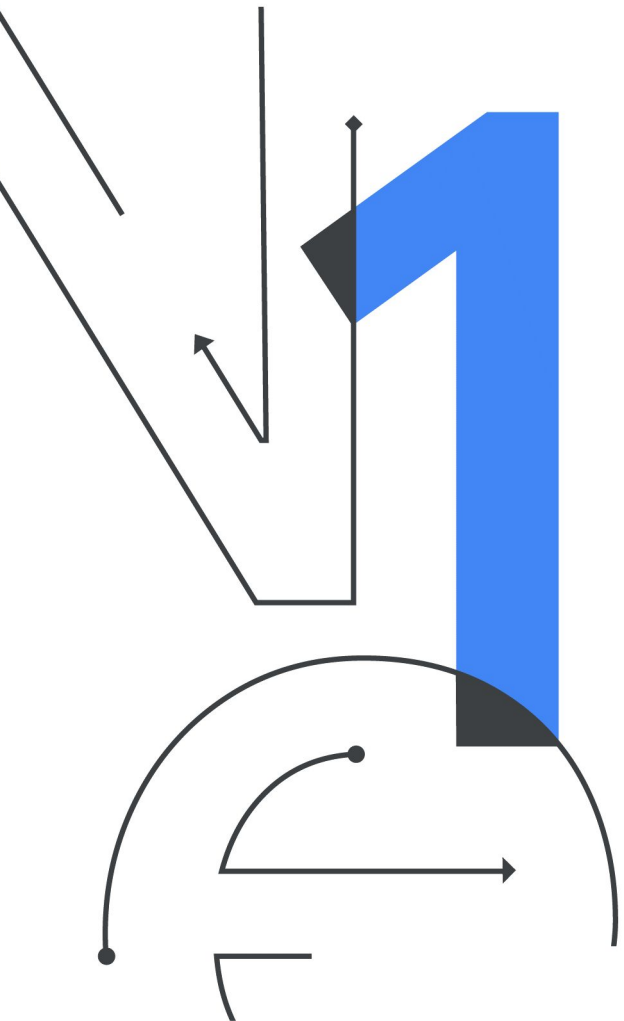
Senior Product Manager,
GKE
Google Cloud

Contents

01 Best practices for running Batch & HPC on GKE

02 PGS high performance computing on GCP





Best practices

Batch&HPC on GKE

What are batch jobs



Completion

Workloads that run to completion. A group of pods that may collaborate with each other or run very independently.

Flexible

Frequently will be flexible on time when they run, location or failure tolerant.

Diverse

Useful in many use cases - high performance computing, machine learning training and inference, extraction transformation and loading and other.

Scalability and performance

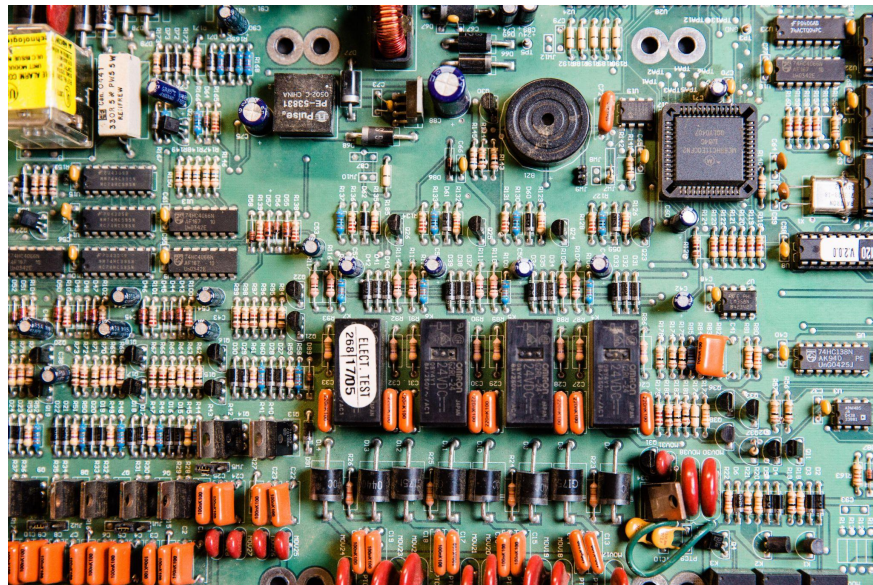
Mega-clusters - 15 000 nodes

Use smaller VMs - 1-2 worker pods

100 pods/s peak throughput

30-50 pods/s as average*

Image Streaming



Use Spot VMs

Up to 91% discount*

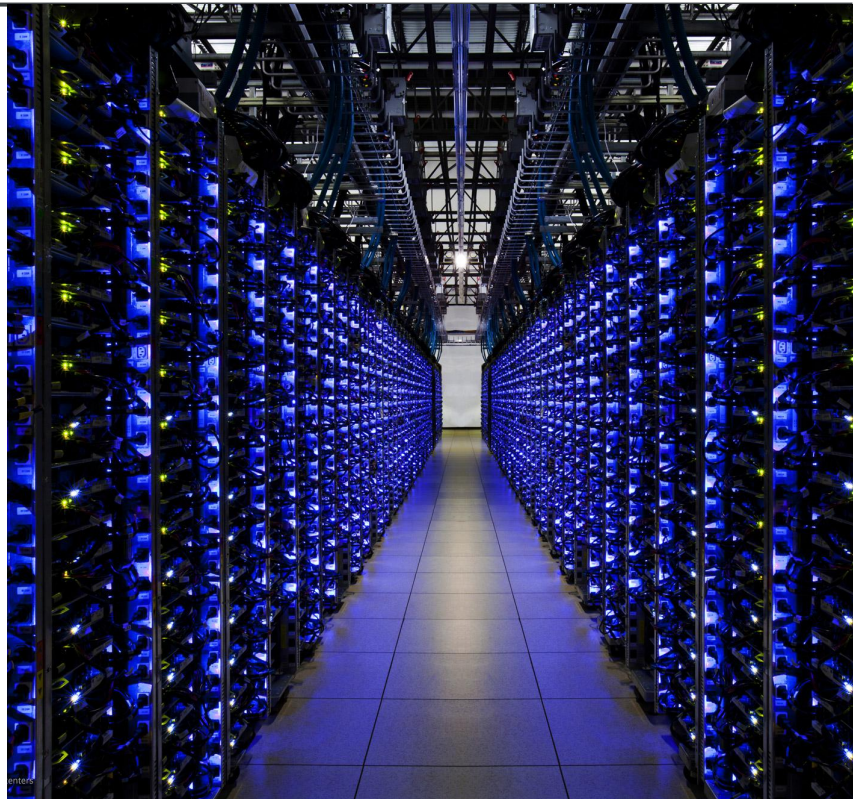
Checkpointing & graceful termination

GKE automatically heals failed VMs

GKE seeks the cheapest VMs



[*https://cloud.google.com/spot-vms](https://cloud.google.com/spot-vms)



Key settings



Node Auto Provisioning

The most powerful binpacking feature



Optimize utilization profile

~20% better binpacking
Adjusts Autoscaling and Scheduling behavior.
Fit for most use cases



ANY spreading

Choice of zone when scaling up driven by capacity and unused reservations.
Especially important for Spot.

Don't want to manage node VMs?

Use Autopilot

GKE Autopilot is a mode of cluster operations where you don't have to worry about nodes
- just run your workloads

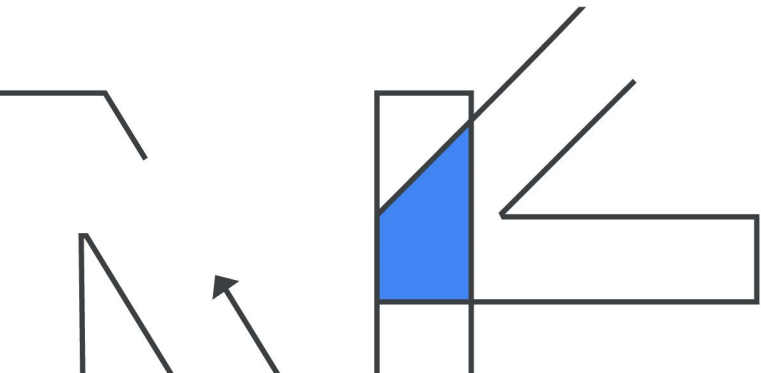
Node VMs are hidden. All provisioning and reliability of nodes is Google's responsibility. Users just need to create workloads.

Less control but significantly less work for platform admins.

Recent launches oriented on batch and HPC workloads:

- Compute classes
- GPU support

45% of GKE Clusters could see 20% or more savings on GKE Autopilot



Stick to Kubernetes core features

Job API

Scheduler

Cluster autoscaler

Cloud integration

Ecosystem compatibility



Job Queueing

Use Kueue.sh

Slim

Reuse of core K8S

Compatibility

Resource fungibility

Open source, reusable anywhere

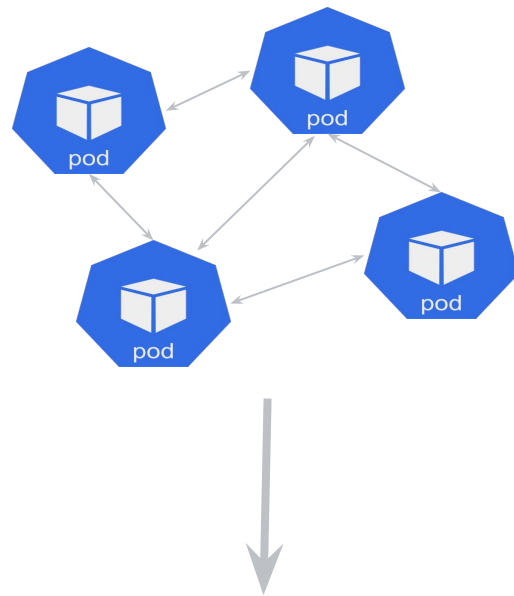


Tightly coupled workloads

Use key settings - optimized network latency

- Compact placement
- Host networking
- Busy polling
- Increase TCP memory
- gVNIC

Kubeflow's MPI Operator V2



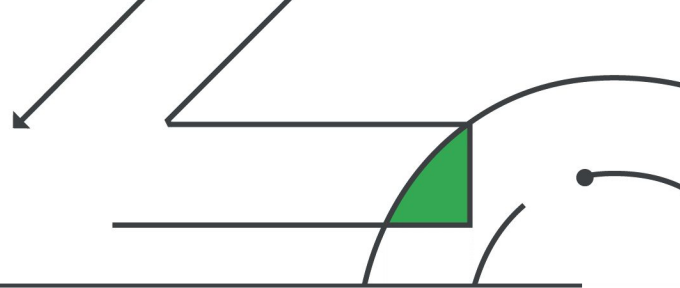
PGS high performance
computing on GCP





What is seismic acquisition ?

What is PGS ?



Integrated marine geophysical company

Acquires seismic data of the ocean subsurface process the data into images of the subsurface layers

- Energy exploration: offshore renewables, carbon capture and storage, as well as oil & gas
- Data volume processed can be multiple PetaBytes
- Requires massive amounts of compute power

Highly specialized hardware

PGS historical compute

Cray supercomputers

XC40 and XC30 supercomputers

- Abel (XC40) with 147,456 cpu cores ranked 12 on top500 Supercomputer in June 2015 (5.37 PFlops)
- Galois (XC30) with 55,296 cpu cores added another 2.1 PFlops
- 12PB of Lustre file system to 257GB/s of throughput

202K
cores

7.47
PFlops

12PB
Storage

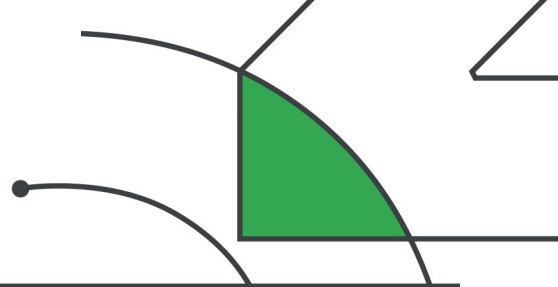
257GB/s





The journey

Why cloud ?



2019

Signed a contract with GCP to move 20% of compute as a “lift and shift”



2021

Economic downturn and early successes transform the ambition to 80% of compute on GCP



May 2022

Last Cray supercomputer decommissioned

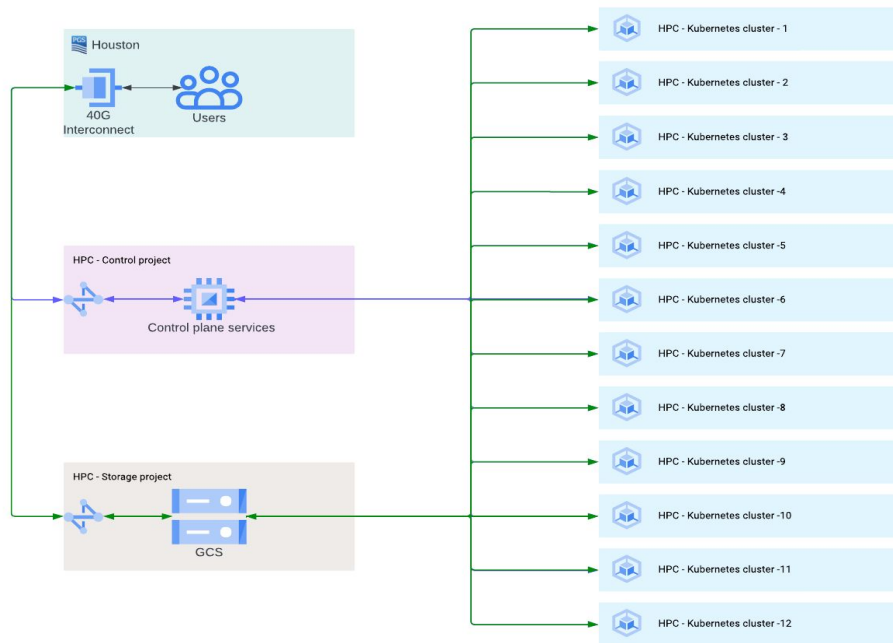
Up to 2.8 Millions vcpu

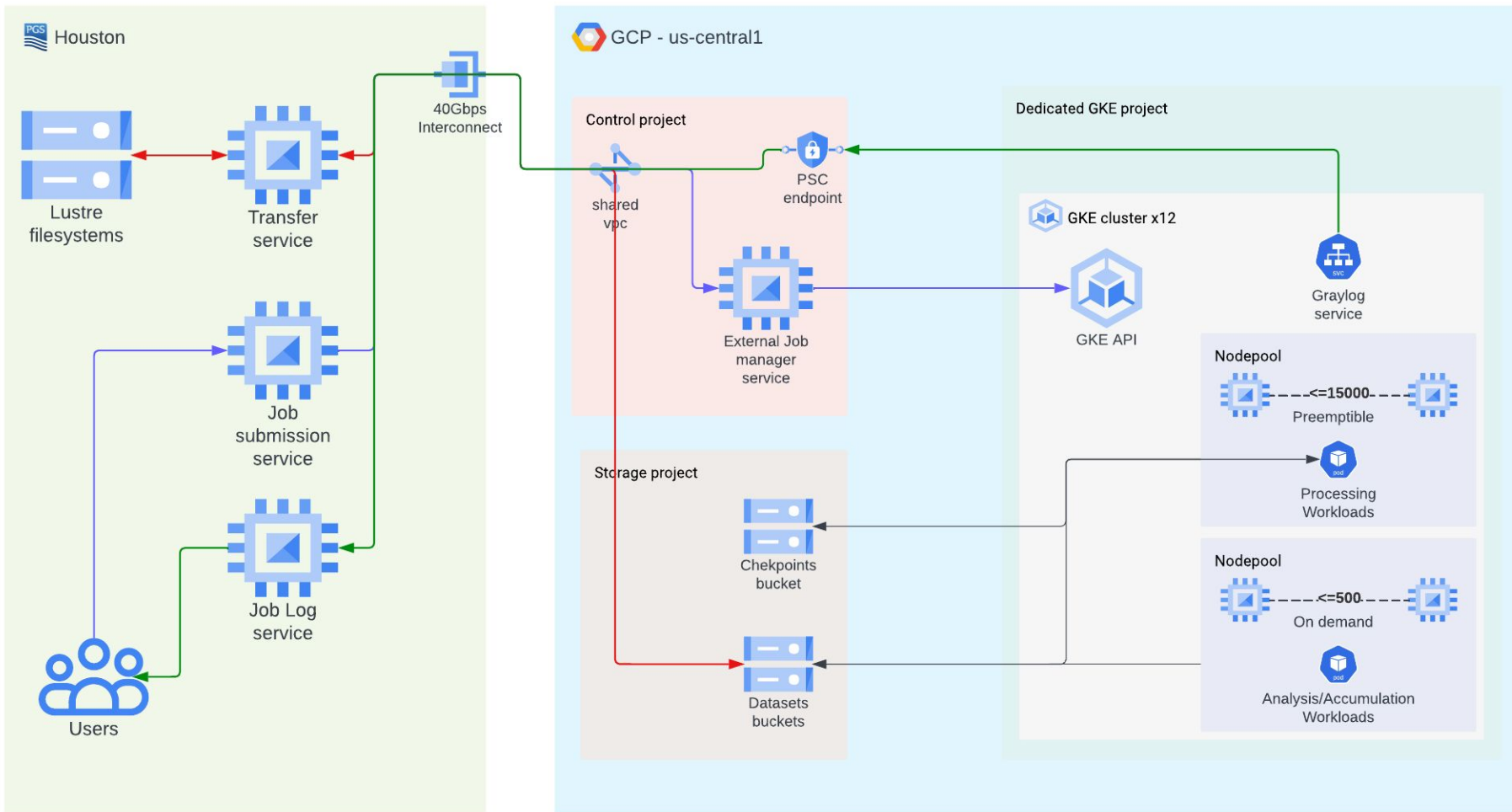
GKE as a supercomputer

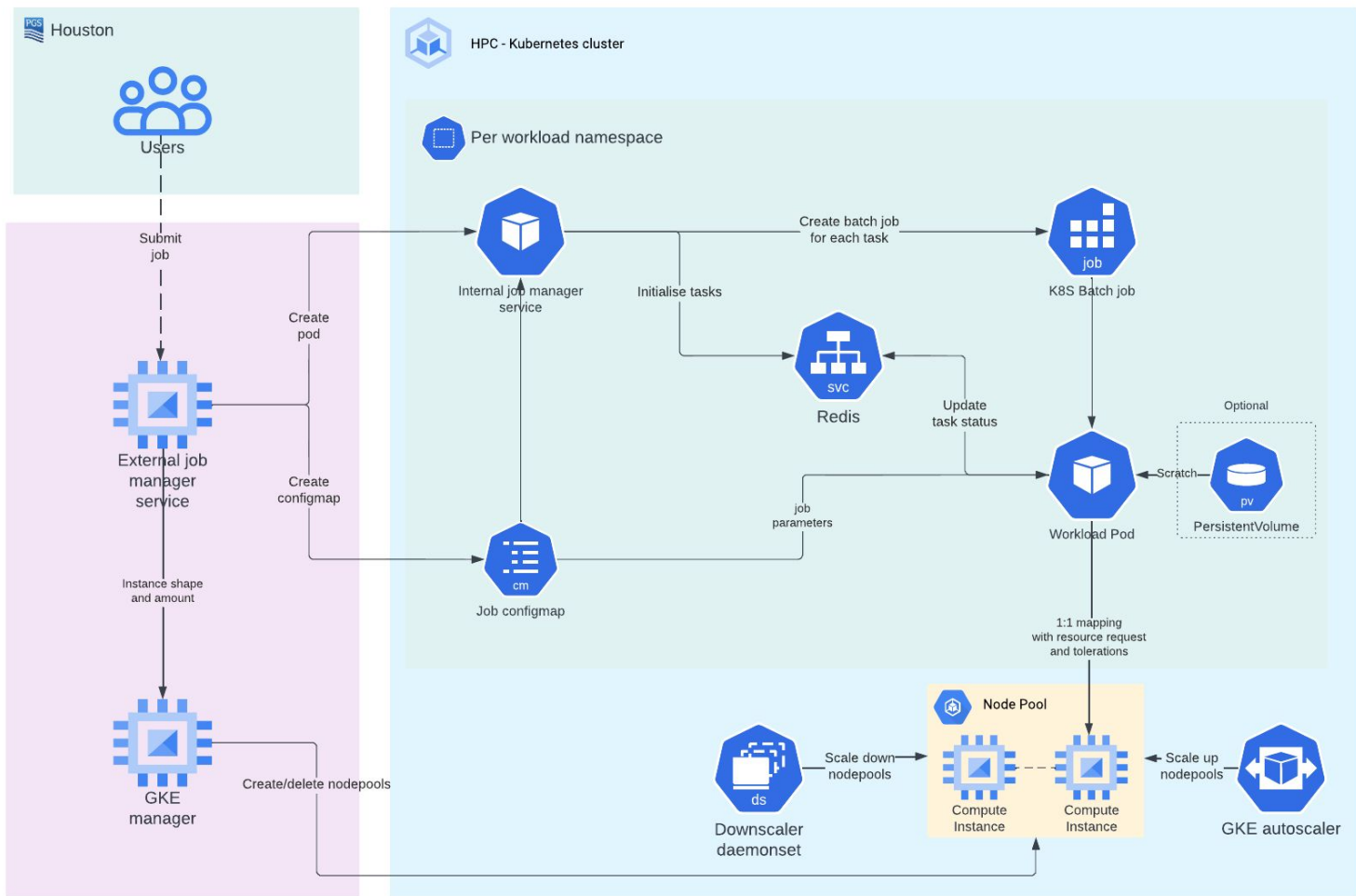
A 7X capacity increase

	Maximum	Peak
GKE nodes	180000	53417
Total vcpu	2.88 Millions*	855536*
Performance	72.02PFlops*	21.14PFlops*
Top500 (June 22)	Top 7	Top 24
GCS bandwidth	1.2 TBps	503 GBps

* using n2d-standard-16







Avoiding CAPEX

The power of the cloud

Limited CAPEX

Hardware flexibility

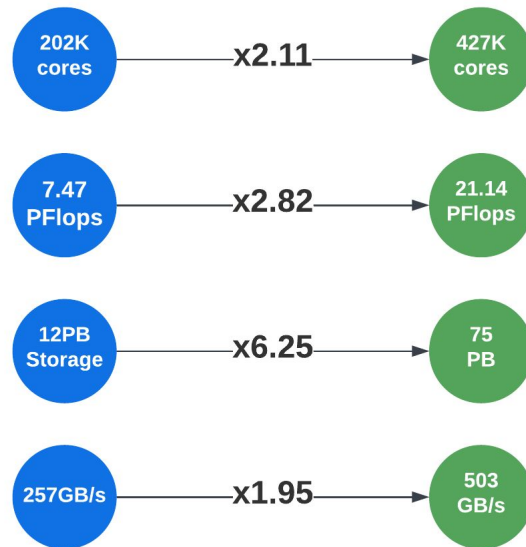
Modernisation

Using low carbon footprint

(us-central1 97% carbon free energy*)

On Premise

GCP

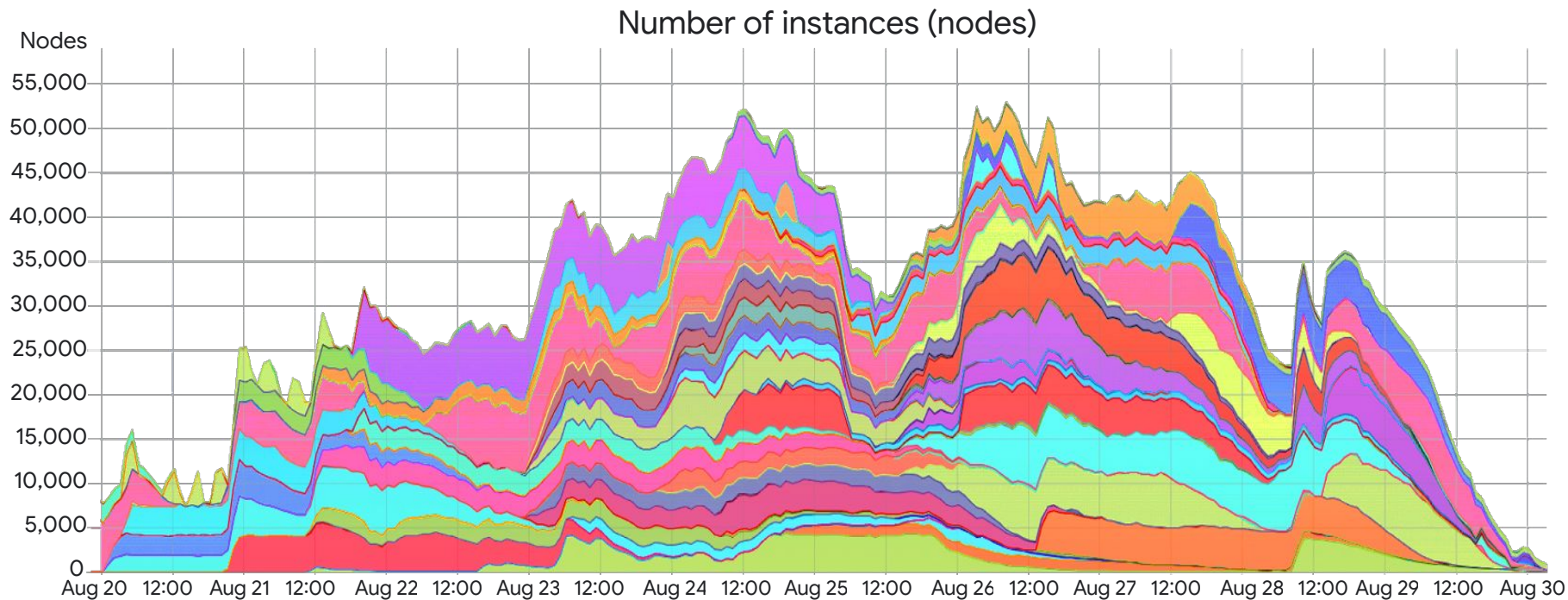


[*https://cloud.google.com/sustainability/region-carbon#data](https://cloud.google.com/sustainability/region-carbon#data)

Google Cloud

855,536 vCPUs

Scale up to 53,417 instances



Thank you

Google Cloud

Next '22

