The HPC + AI Cloud

Flexible and performant infrastructure
for HPC and AI workloads

Matt Pryor, Senior Technical Lead, StackHPC

zimuth
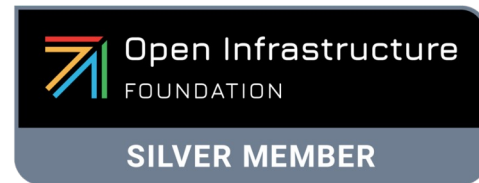
StackHPC

StackHPC Company Overview

- Formed 2016, based in Bristol, UK
  - Based in Bristol with presence in Oxford, Cambridge, France and Poland
  - Currently around 30 people
- Founded on HPC expertise
  - Software Defined Networking
  - Systems Integration
  - Open Research Infrastructure Development and Operations
- Motivation to transfer this expertise into Cloud to address HPC & HPDA (AI)
- "Open" Modus Operandi
  - Upstream development of OpenStack capability
  - Consultancy/Support to end-user organizations in managing HPC service transition
  - Scientific-SIG engagement for the Open Infrastructure Foundation
- Hybrid Cloud Enablement

Getting Maximum Value from your Investment?
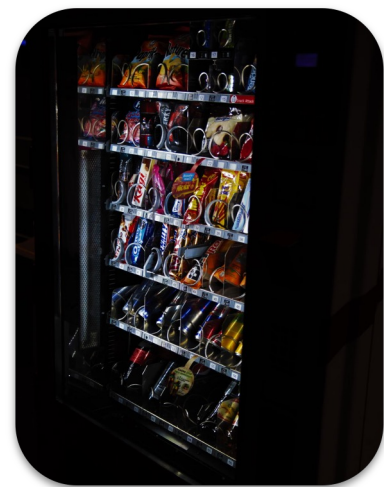
Getting Maximum Value from your Investment

StackHPC Three Pillars

Reconfigurable and isolated infrastructure

Performance to extract maximum value

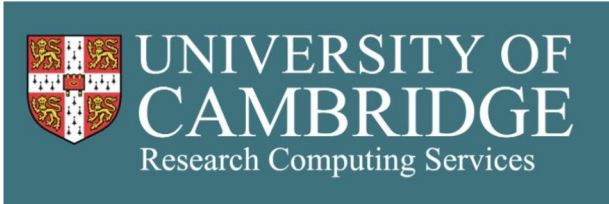Azimuth self-service platforms

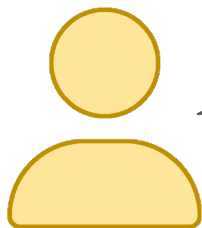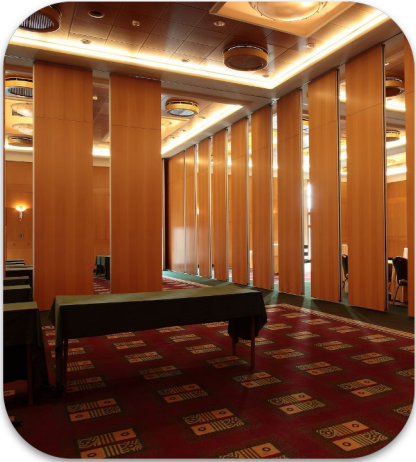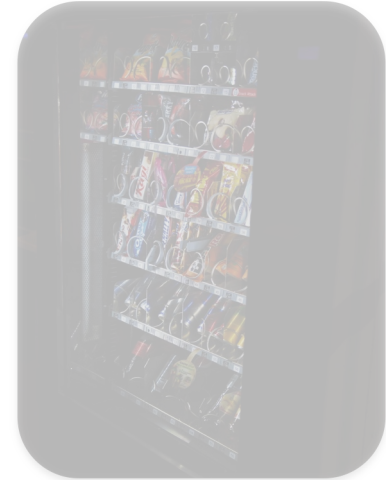# Open Source Co-Development

Diverse Use Cases
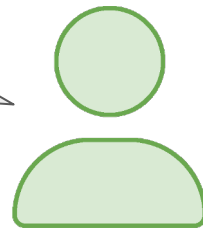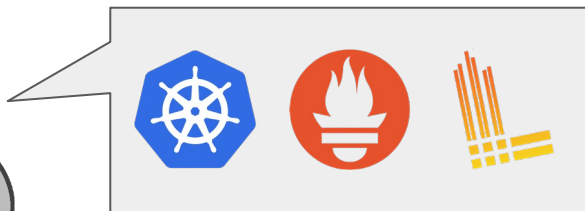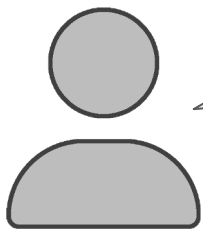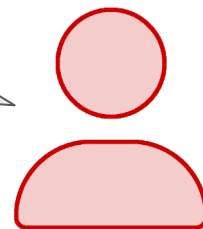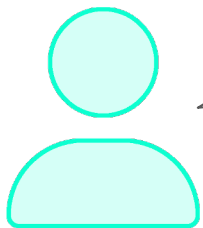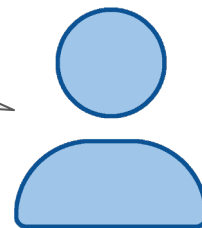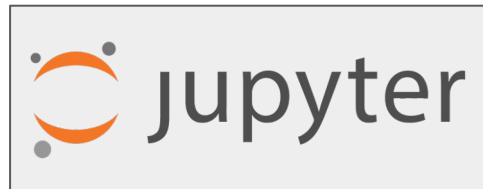
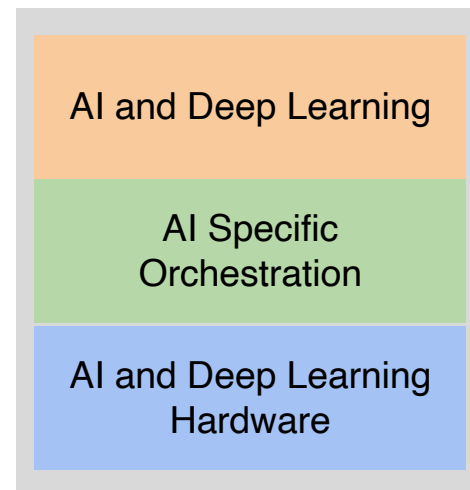# Infrastructure

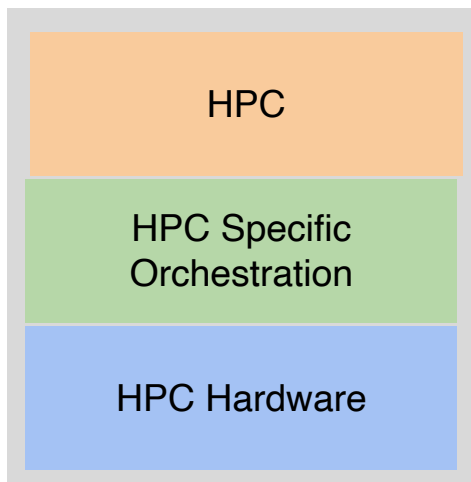**Reconfigurable and isolated infrastructure**

Performance to extract maximum value

Azimuth self-service platforms

Hardware silos

Cloud-native HPC

HPC	Big Data	AI and Deep Learning

APIs: AAAI, Software-defined Networking, Storage Services, Compute Services

Shared Configurable Hardware
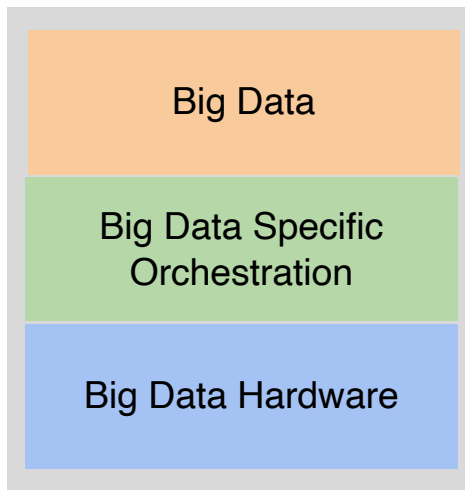
OpenStack to the rescue

- OpenStack is an open-source cloud platform

- Strong multi-tenancy guarantees

- APIs for compute, network and storage

  - Dashboard and command-line interface

- DevOps tools for OpenStack and platforms

  - Code reviewed changes

  - Continuous integration and delivery

Cloud-native HPC site

Cloud-native HPC site

Cloud-native HPC site
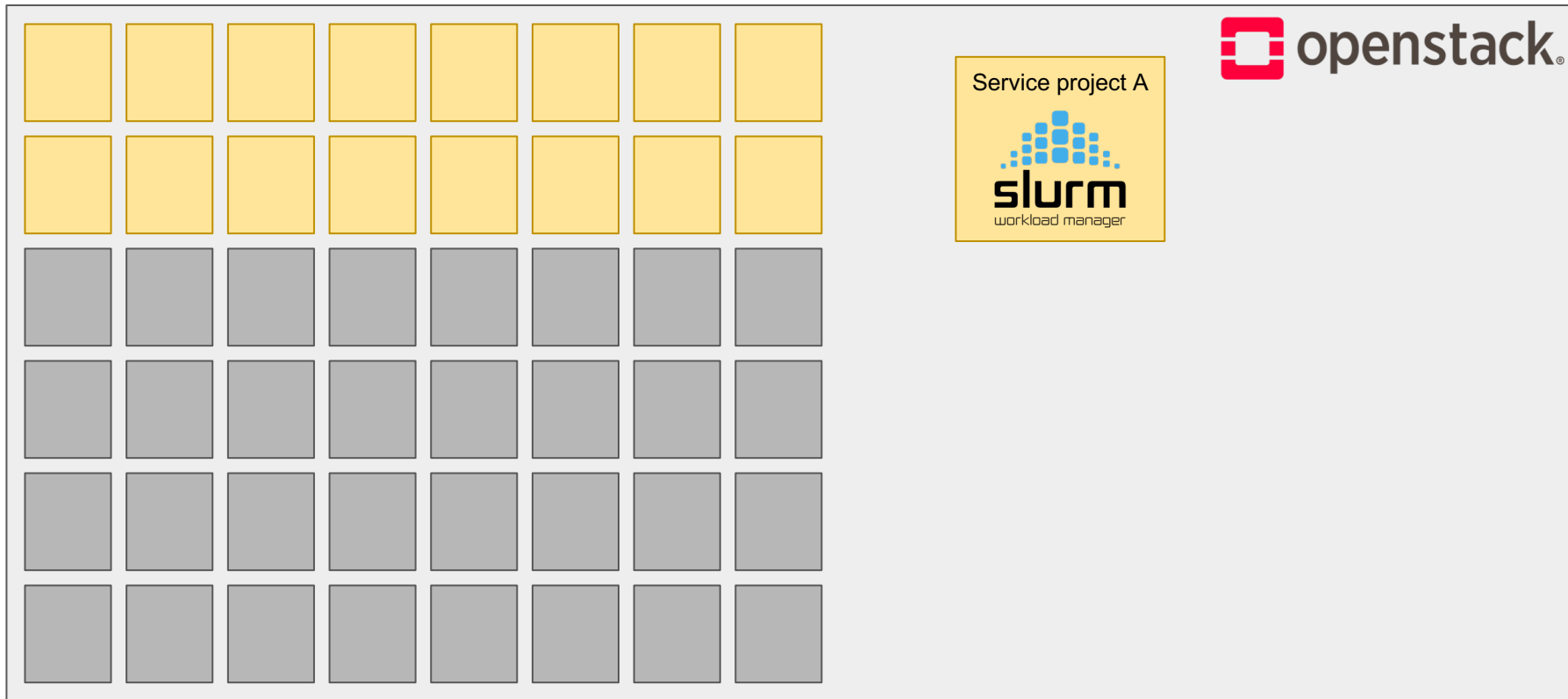
Service project A
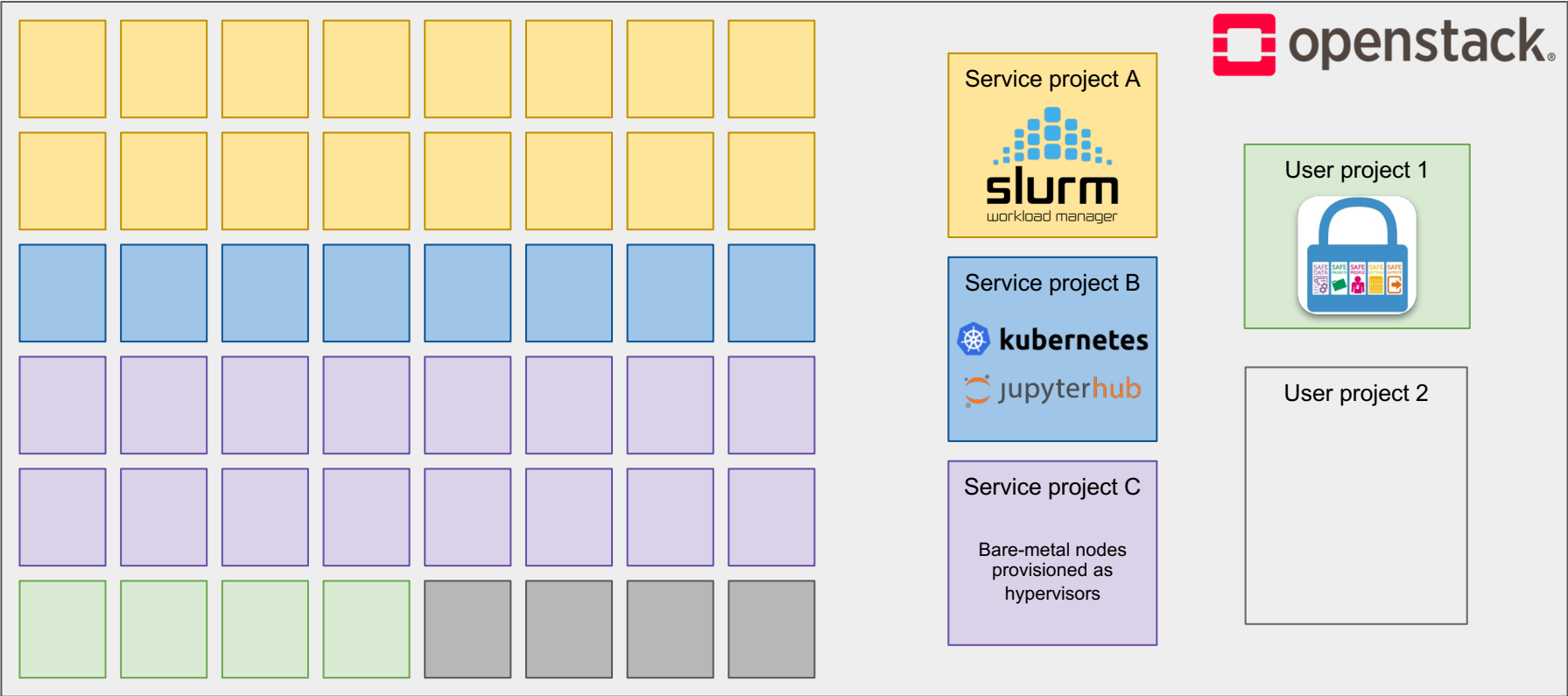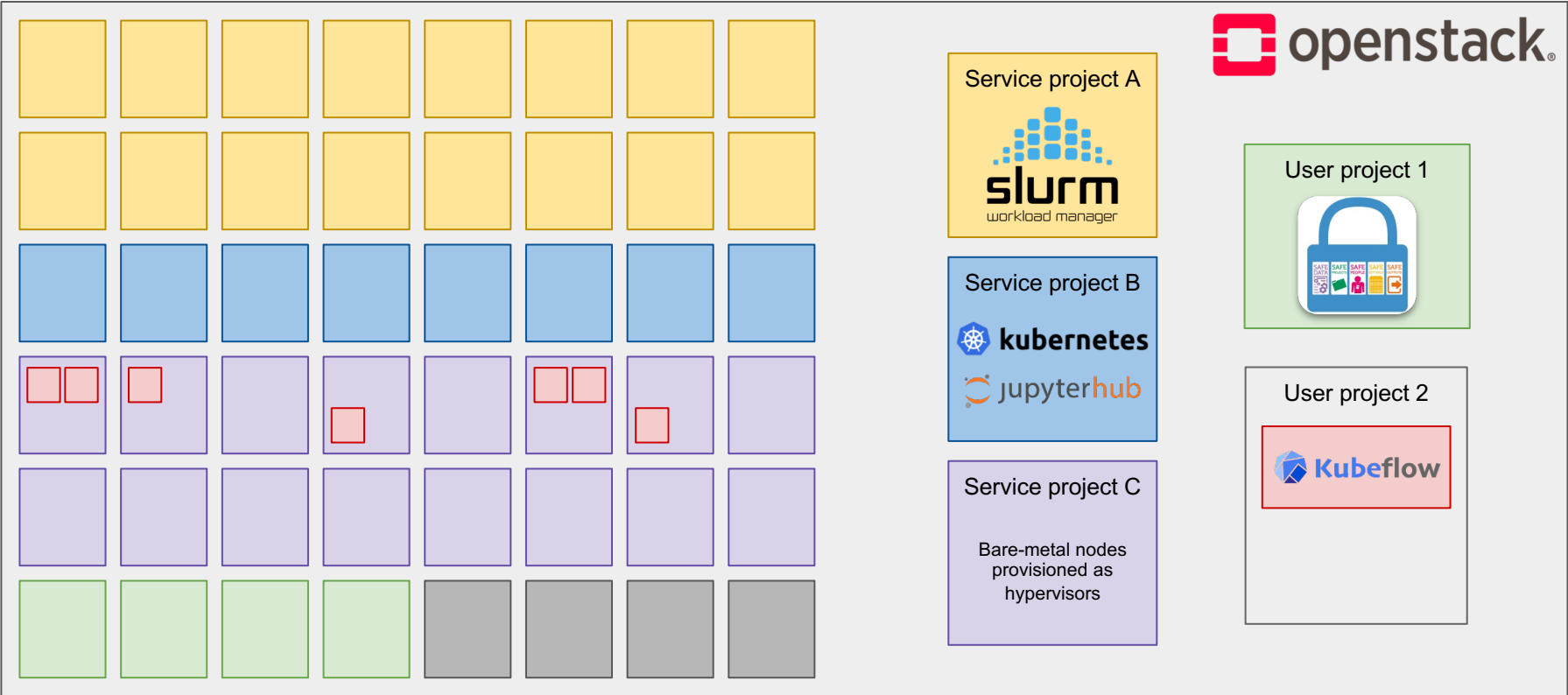
Cloud-native HPC site

Cloud-native HPC site

Cloud-native HPC site

Cloud-native HPC site
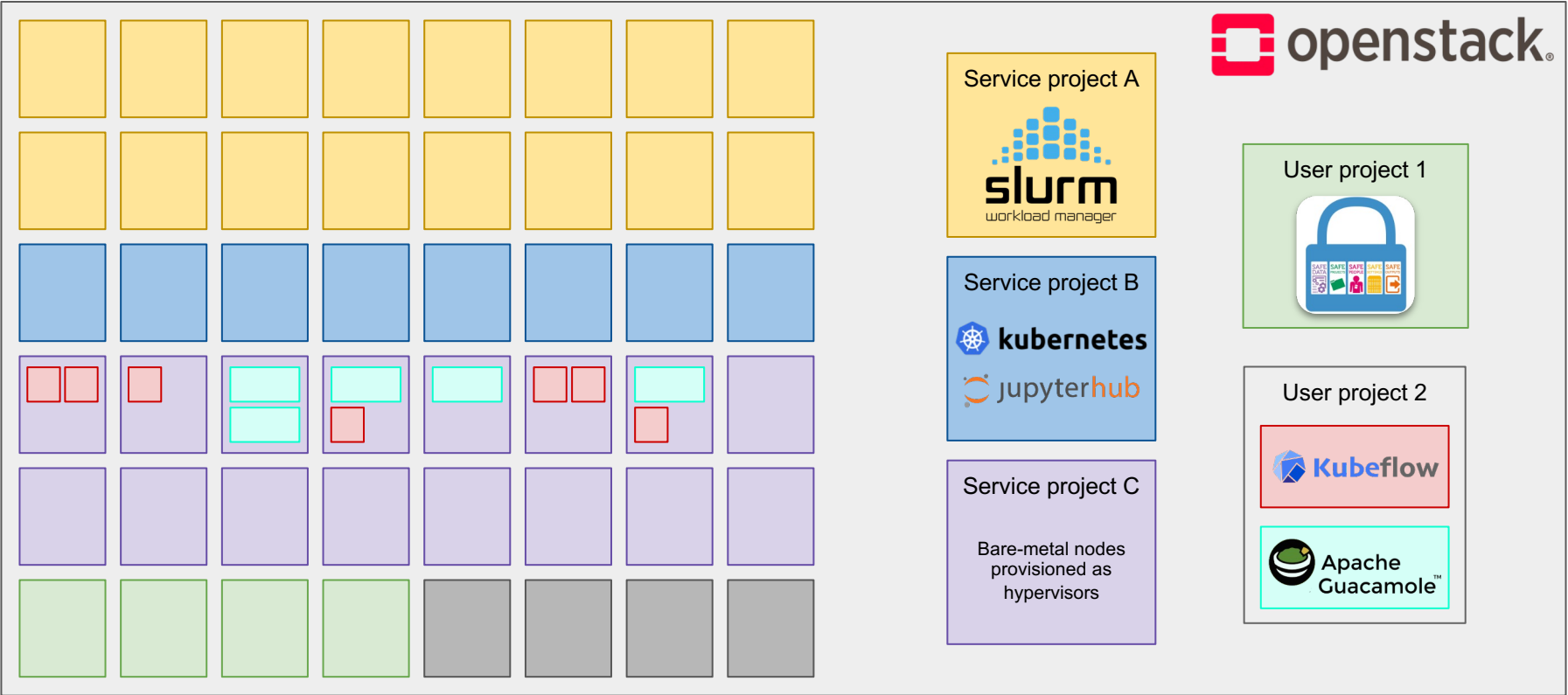
Cloud-native HPC site

Cloud-native HPC site
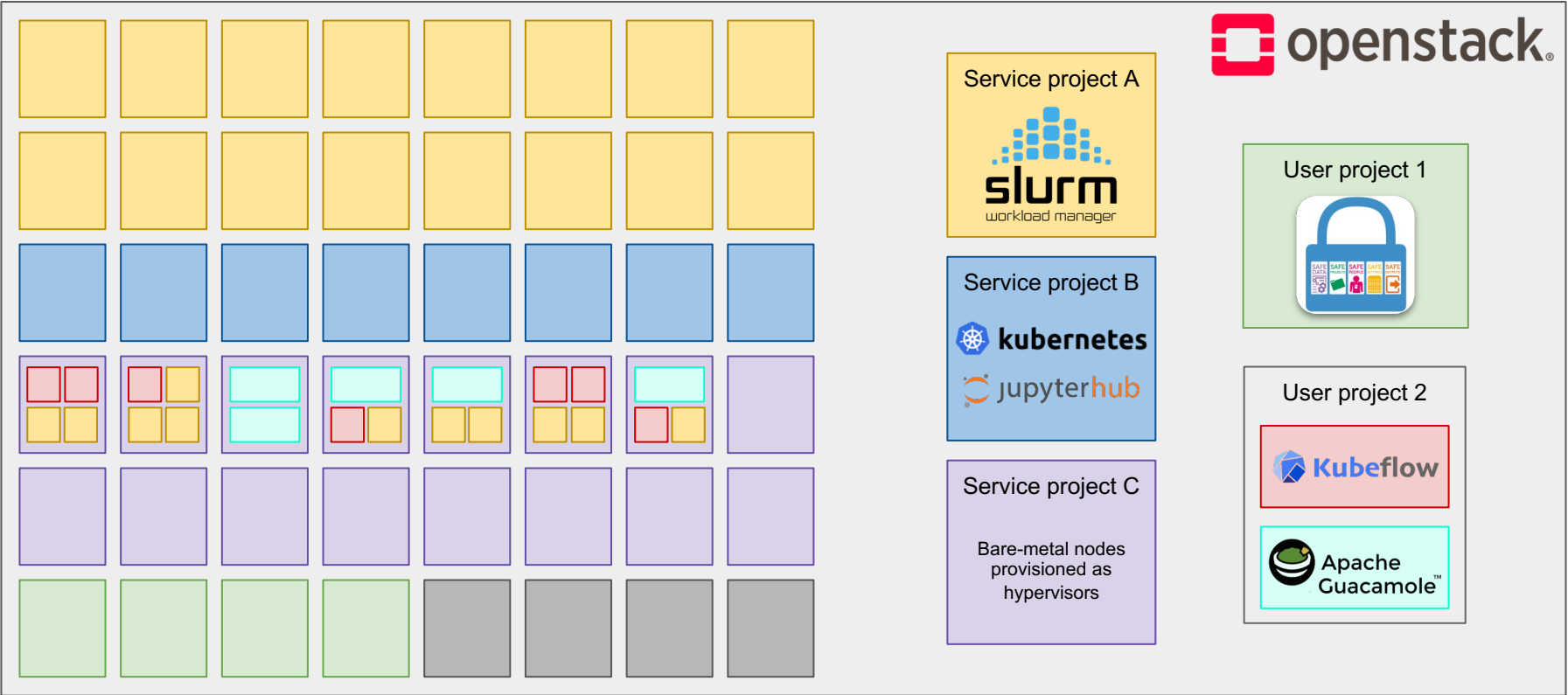
Cloud-native HPC site
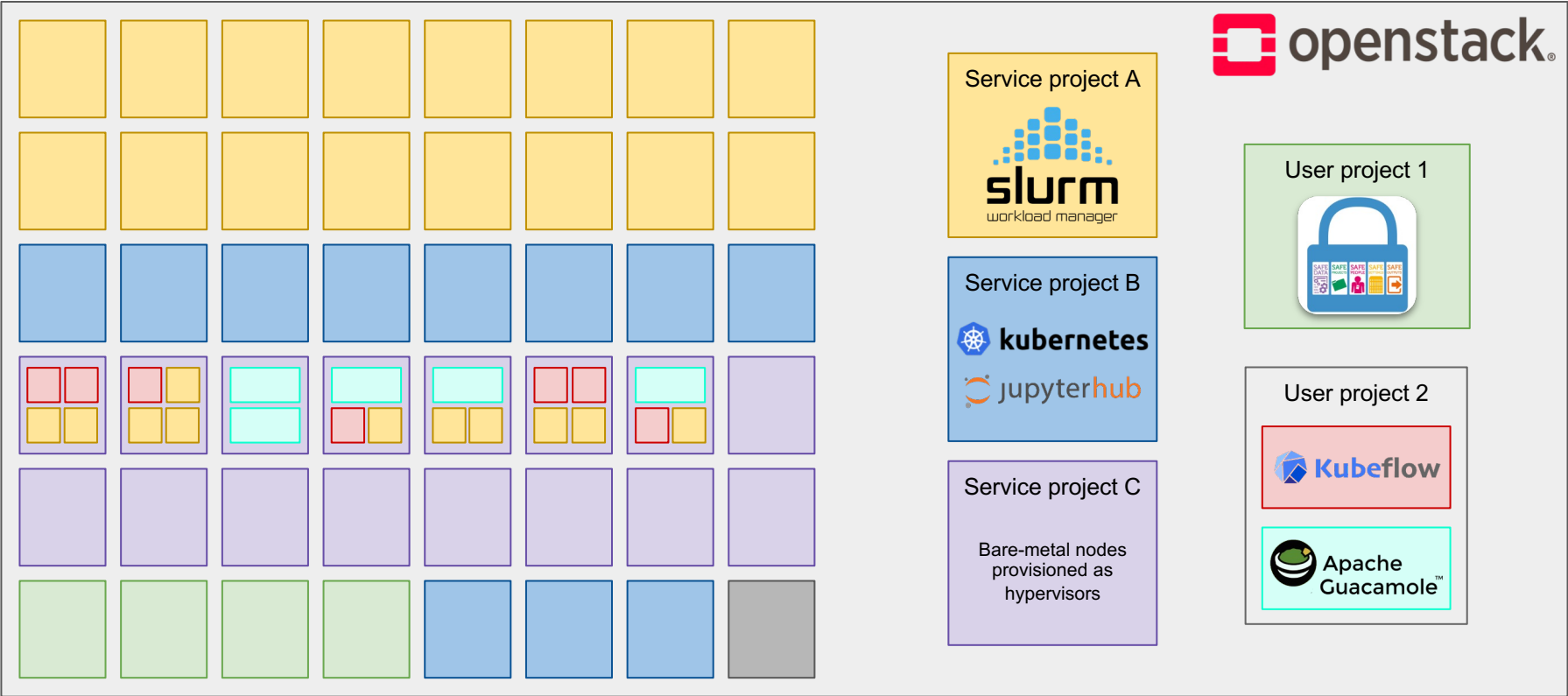
Cloud-native HPC site

Cloud-native HPC site
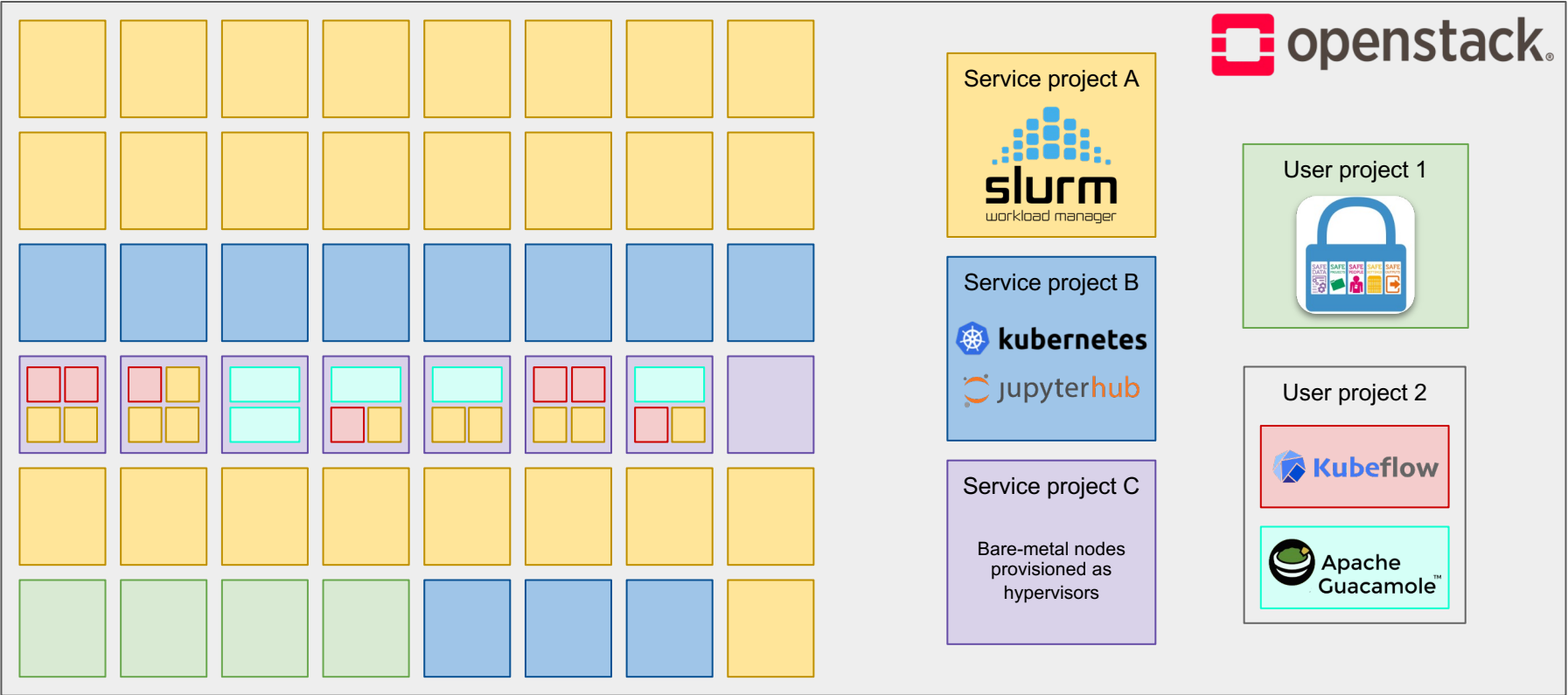
OpenStack Kayobe
Infrastructure as Code
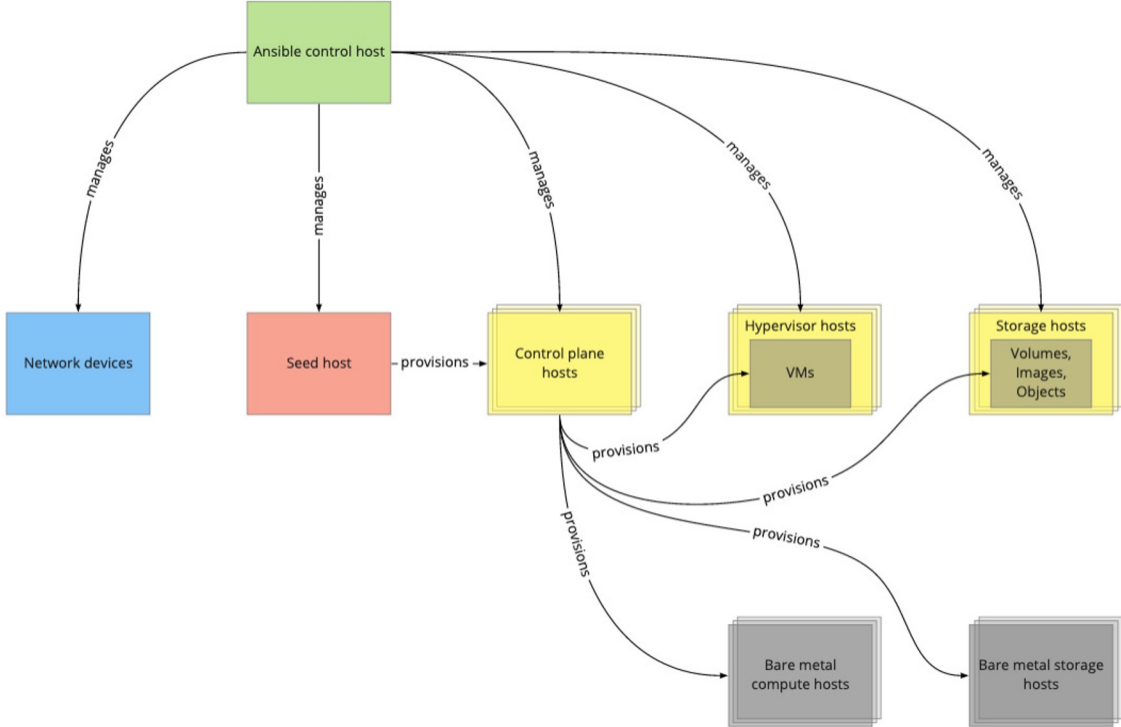
Performance

Reconfigurable and isolated infrastructure

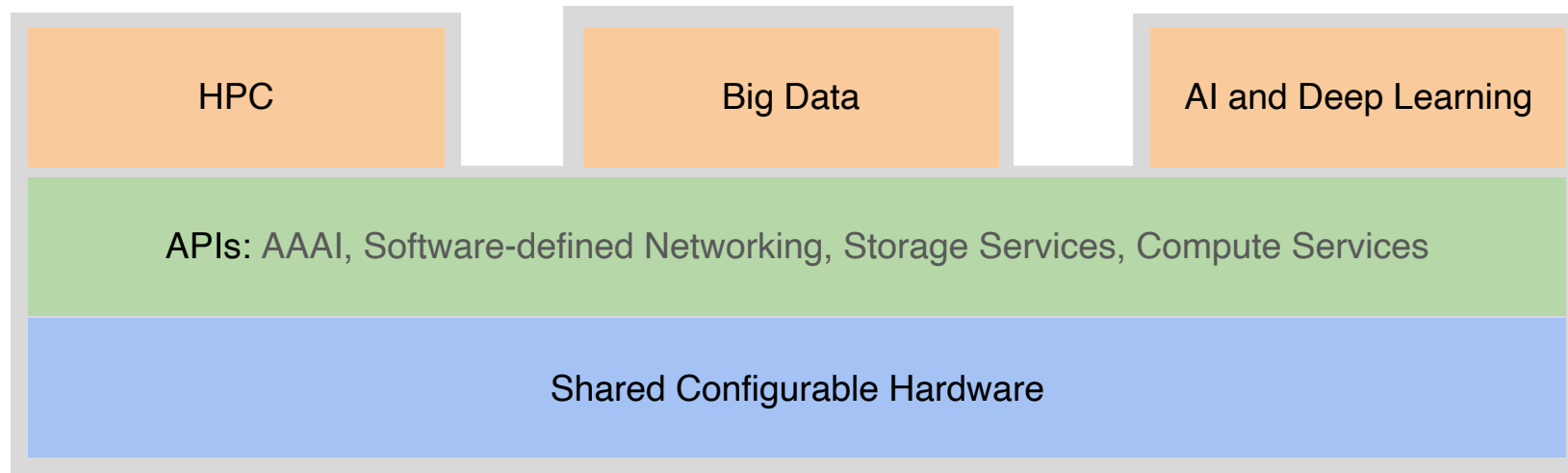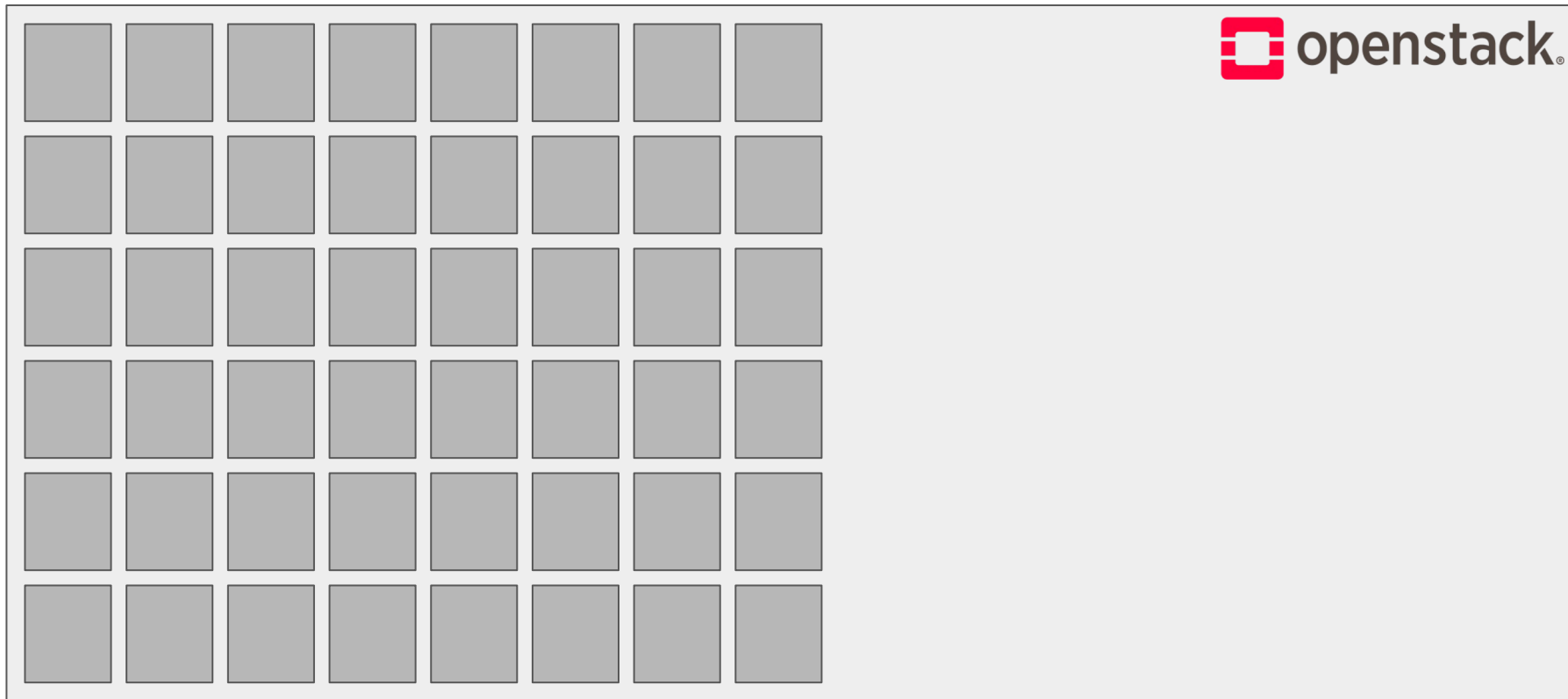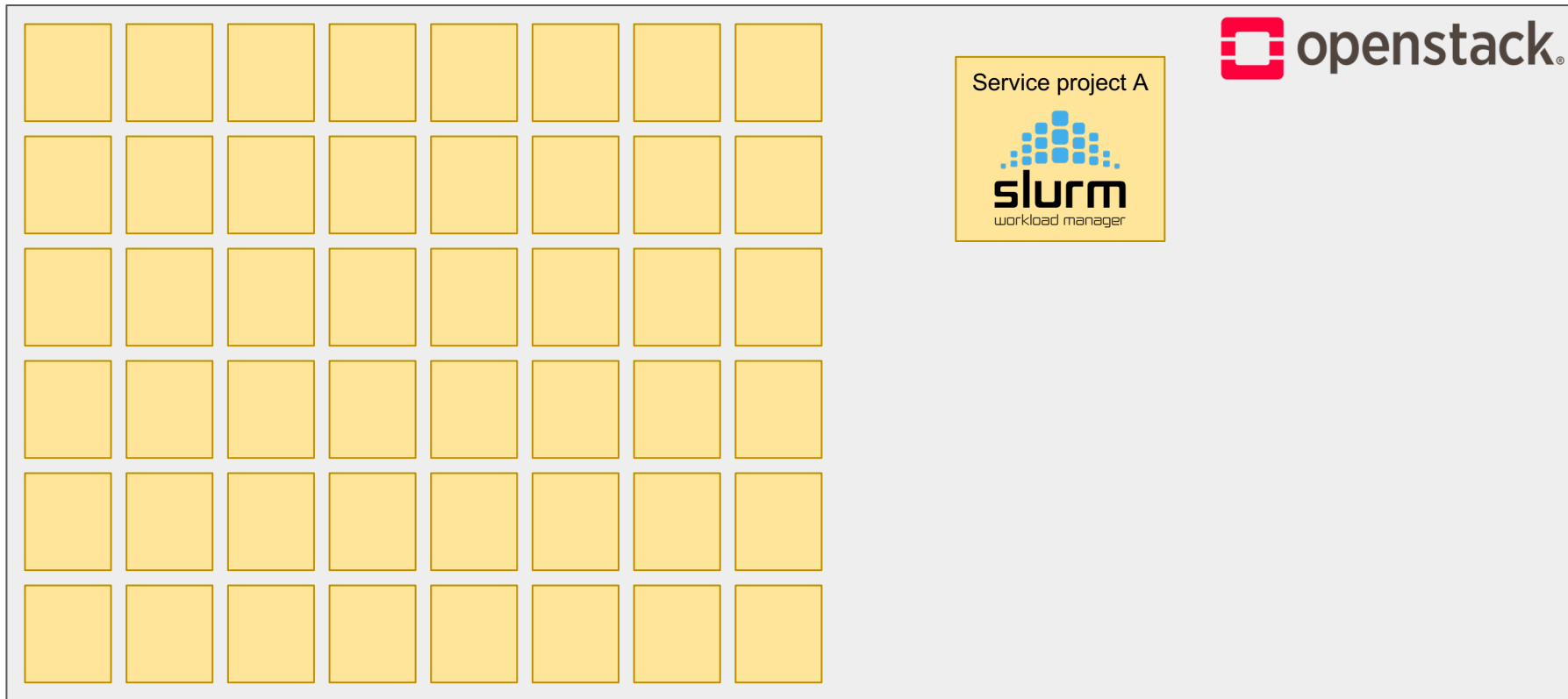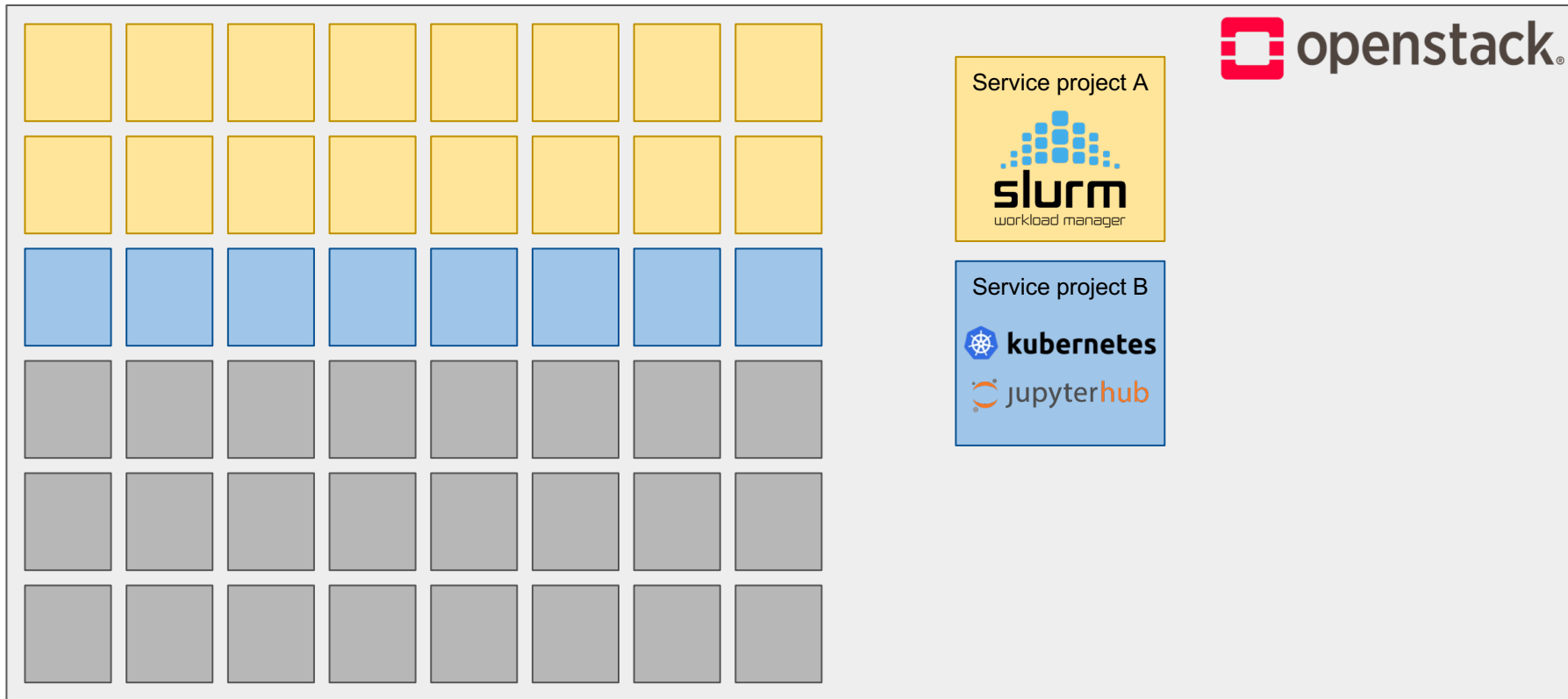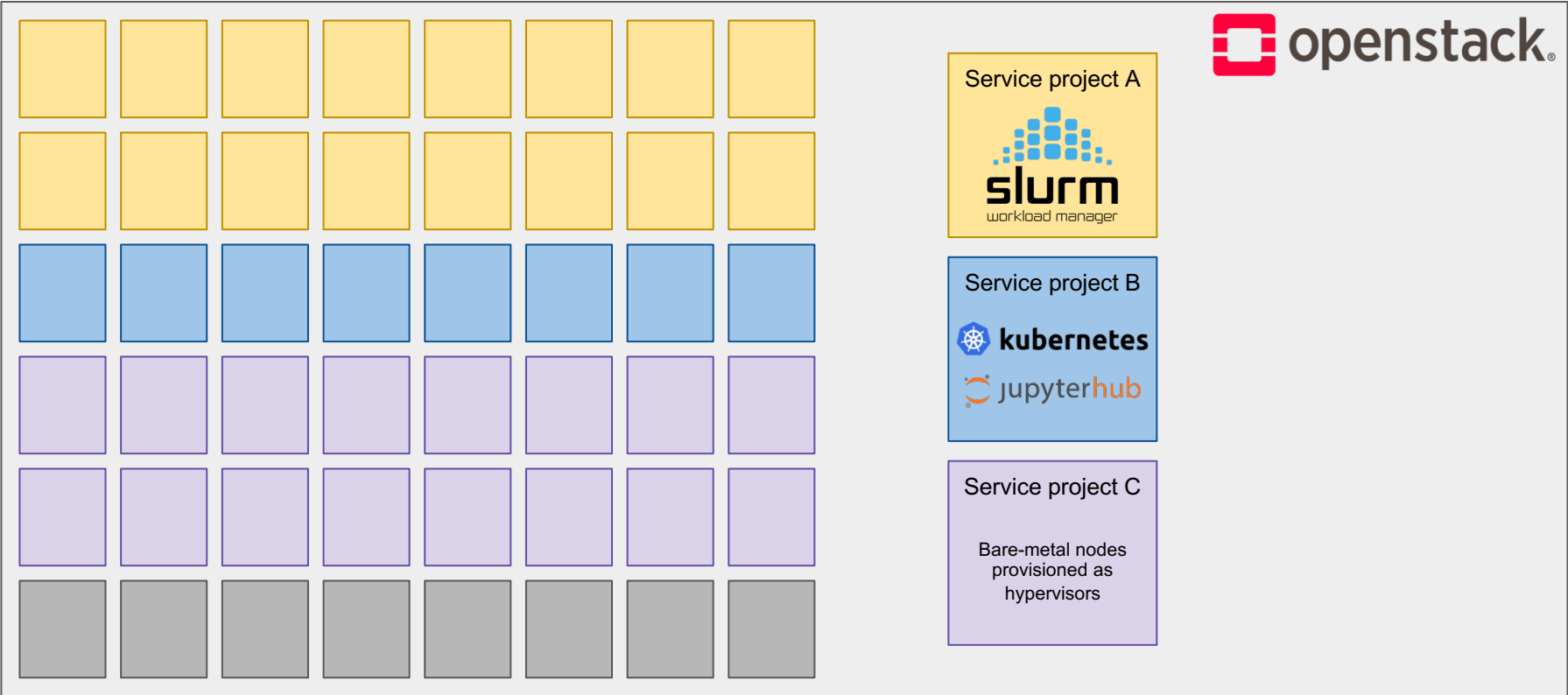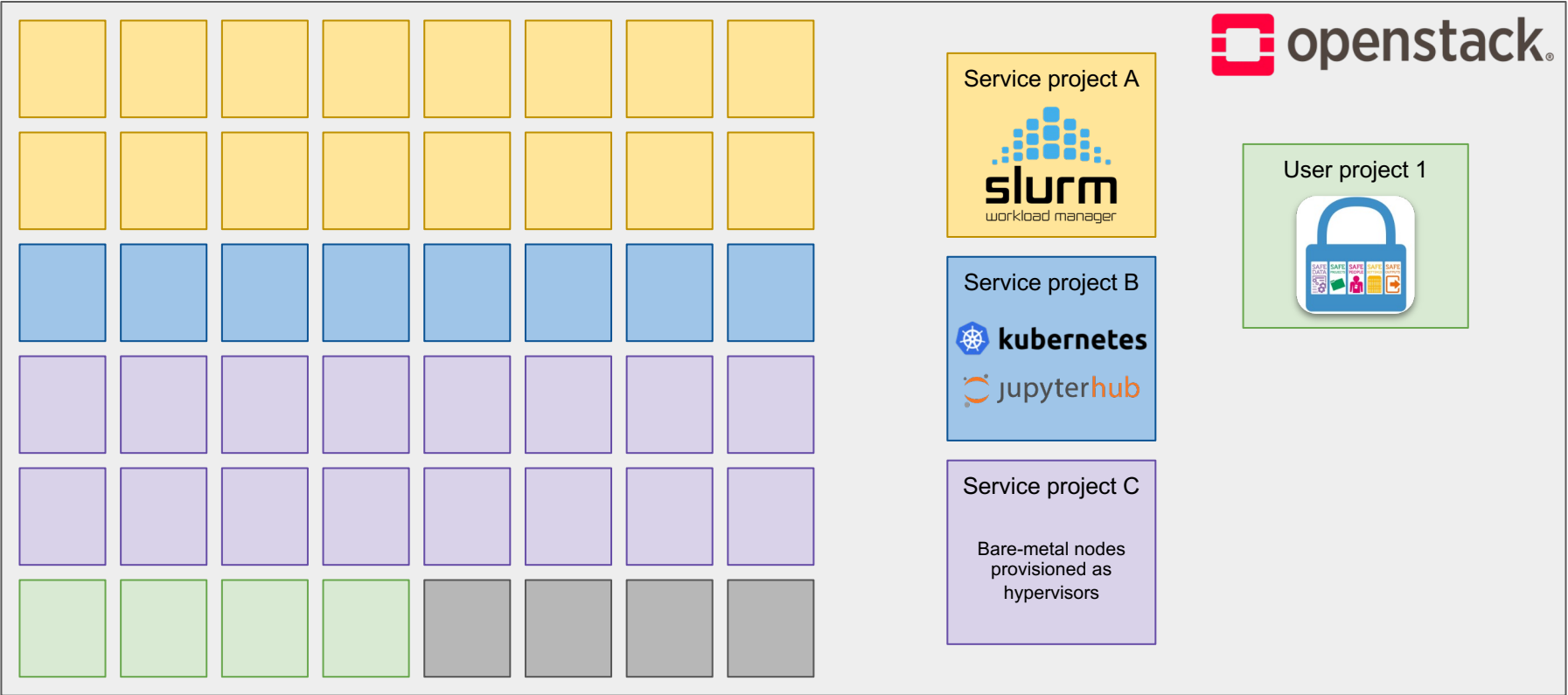**Performance to extract maximum value**
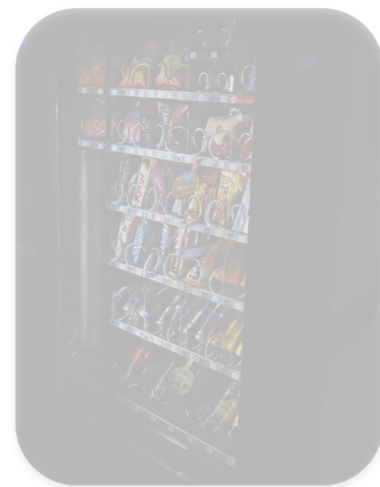
Azimuth self-service platforms

Bare metal vs VMs

- Ironic bare metal adds zero overhead
  - Direct access to hardware
  - Optional dynamic VLAN and cleaning
  - But nodes are big

- Virtual machines
  - Deploy faster
  - GPU passthrough, SR-IOV and pinned CPUs
  - No access to firmware patching

High-speed networking

- Utilise hardware offloads

- Single Root I/O Virtualisation (SR-IOV)
  - VMs can utilise hardware offloads

- Remote Direct Memory Access (RDMA)
  - Low latency, high bandwidth
  - Requires application support
  - MPI and storage transport
  - Typically RoCE (RDMA over Ethernet)
  - Infiniband supported by StackHPC

MPI PingPong

MPI PingPong Latency (small messages) / MPI PingPong Bandwidth (large messages)
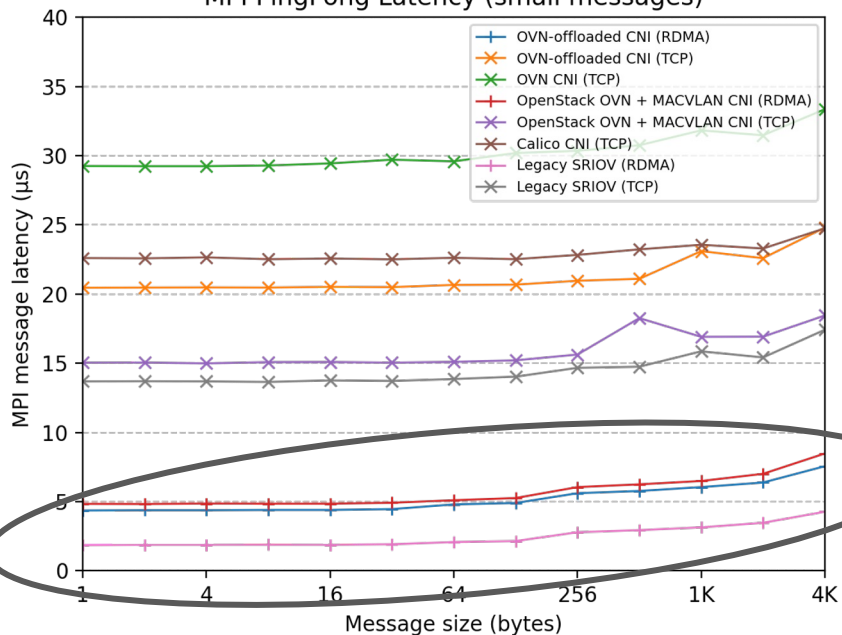
https://github.com/stackhpc/kube-perftest
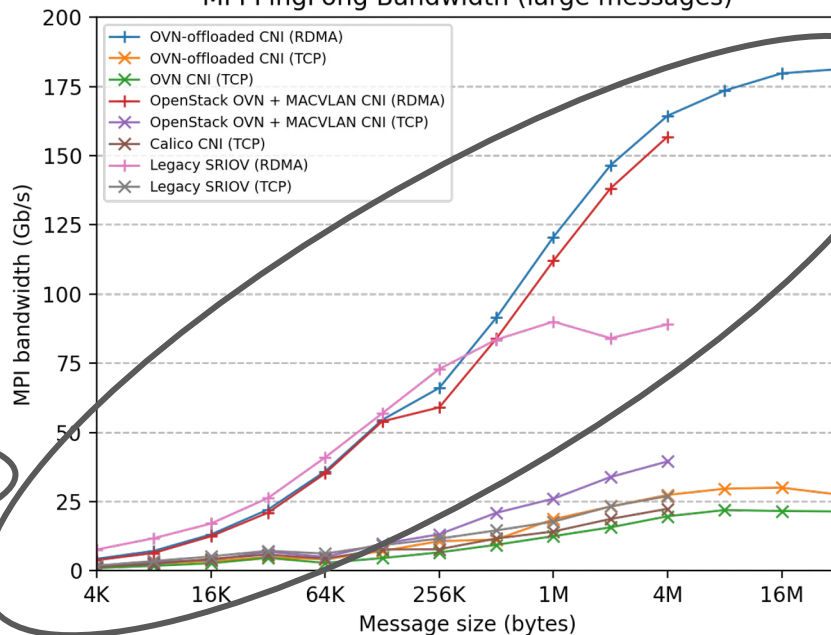
OpenFOAM via kube-perftest

OpenFOAM fluid dynamics simulation (48 parallel processes)

https://github.com/stackhpc/kube-perftest
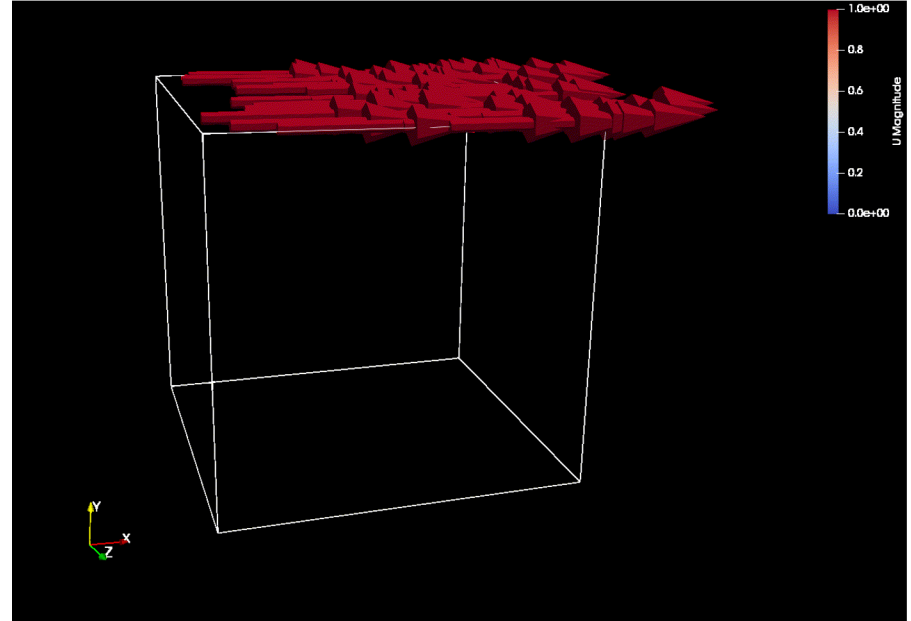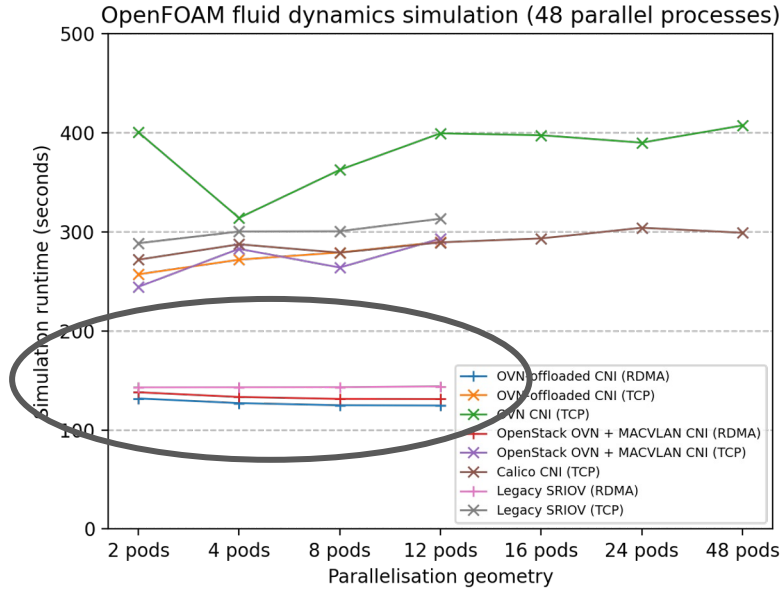
Self-service platforms


Reconfigurable and isolated infrastructure

Performance to extract maximum value

Azimuth self-service platforms

Azimuth

- Web portal for self-service platforms

- Configurable catalogue of curated platforms
  - StackHPC reference platforms
  - Site-optimised platforms
  - Automation using standard tools

- Platform services exposed using Zenith
  - Tunneling application proxy
  - No public IP required
  - SSO and TLS

- Manage platform users with Keycloak
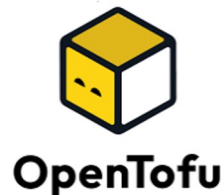
Workstation

- Web-based shell and desktop

- Secure access via Zenith

- Monitoring stack

- Platform lifetime

- User gets sudo

- Apptainer and podman

- Optional SSH with public IP

- Access to project share
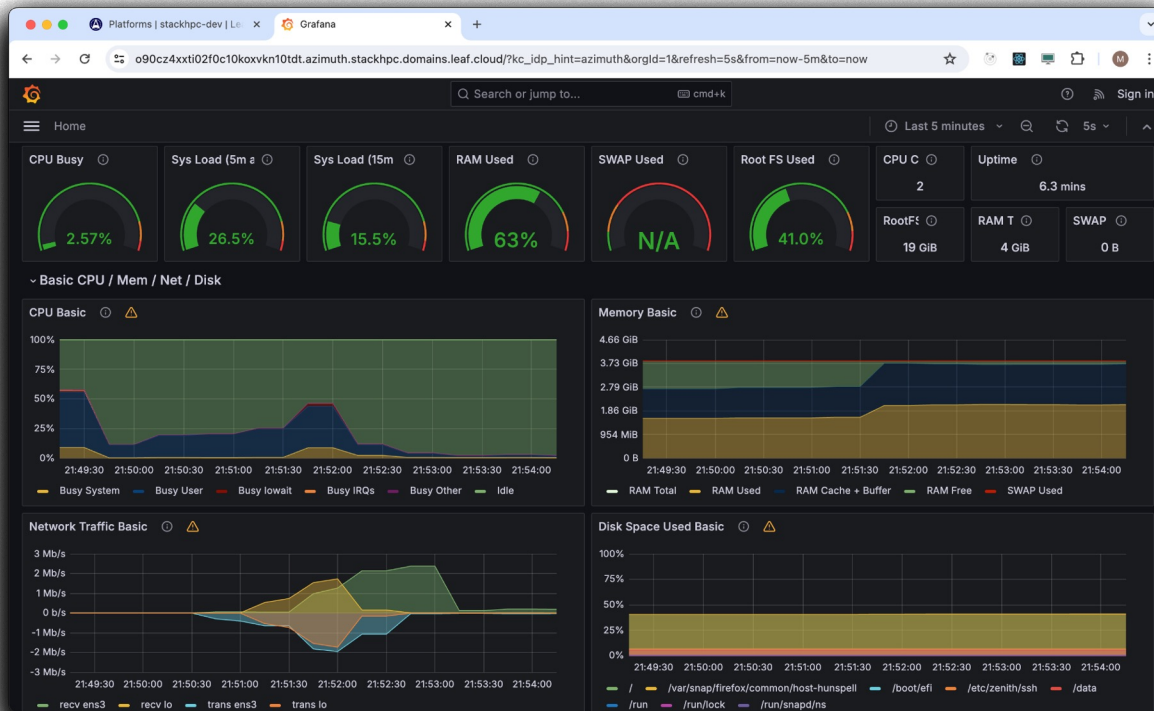
Kubernetes

- Built on Cluster API

- HA control plane

- Multiple node groups

- Download kubeconfig

- Autoscaling, autohealing

- Rolling upgrades

- NVIDIA GPU + NIC support

- Kubernetes dashboard

- Monitoring and logging

- Secure access via Zenith
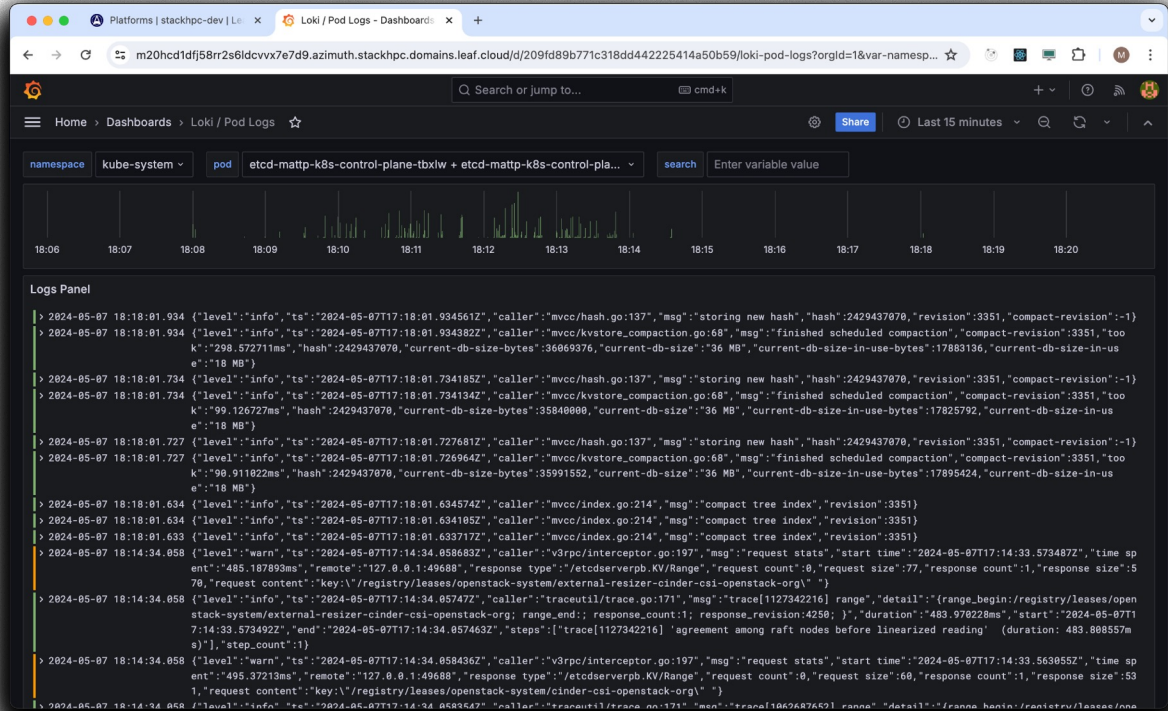
DaskHub

- Runs on Kubernetes cluster

- Each user gets their own notebook server

- Secure access via Zenith

- Grant access to external users using tenancy Keycloak realm

- Dask clusters for parallel computing using Dask Gateway
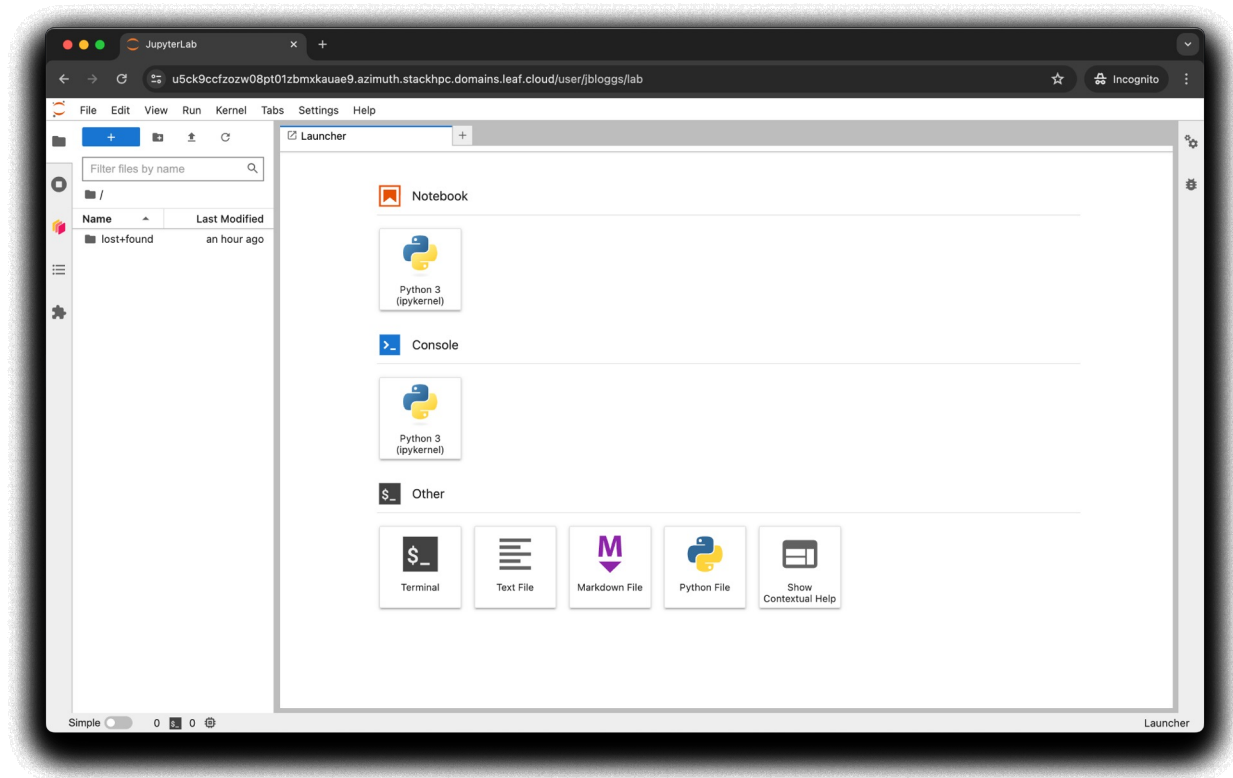
On-demand Slurm

- Single-user Slurm cluster

- No waiting for queues

- Image-based updates

- OpenHPC, Apptainer, EESSI

- Open OnDemand UI

- Job aware monitoring

- Access to project share

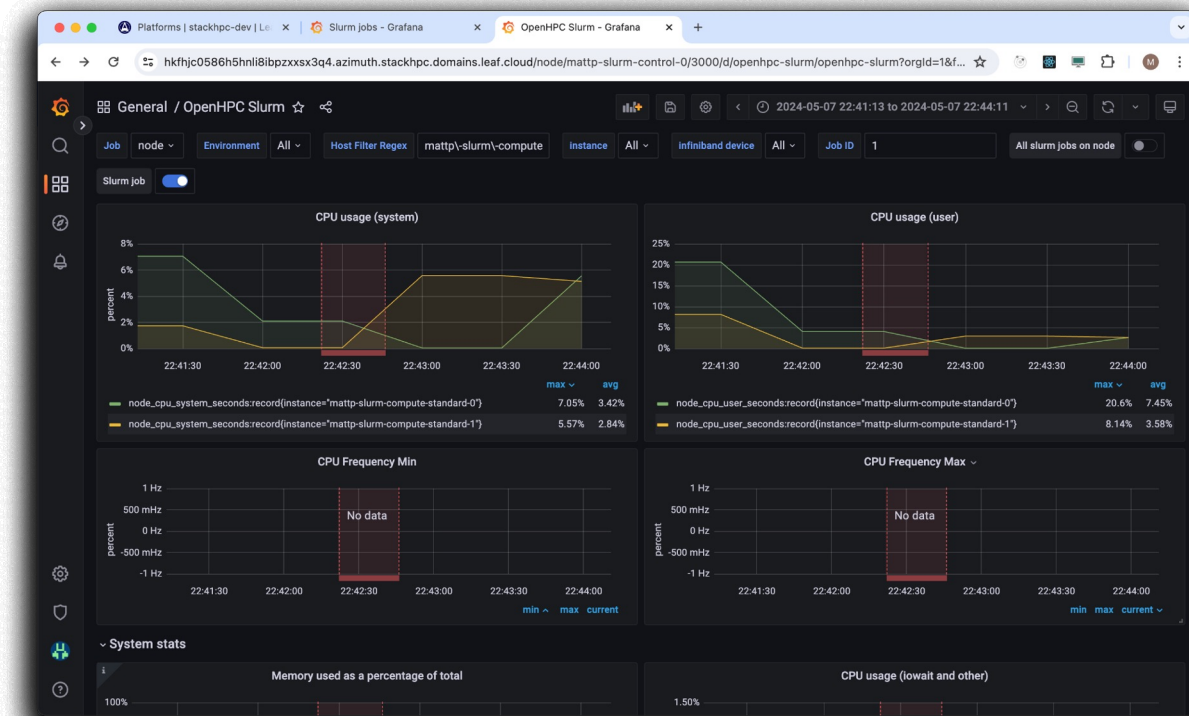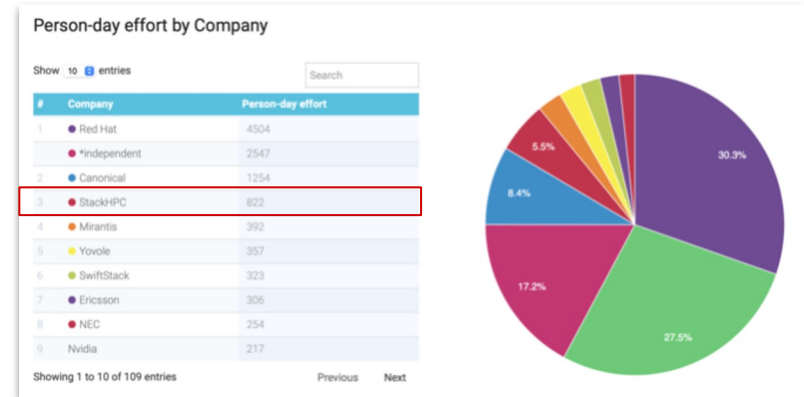# How to get started?

Open Source

- Entire stack is open-source

- OpenStack governed by OpenInfra Foundation
    - StackHPC third-largest contributor

- Azimuth licenced under Apache 2.0
    - Putting together CNCF Sandbox application

- Collaborative development process

Person-day effort by Company

Show 10 entries   Search

#	Company	Person-day effort
1	Red Hat	4504
	*independent	2547
2	Canonical	1254
3	StackHPC	822
4	Mirantis	392
5	Yovole	357
6	SwiftStack	323
7	Ericsson	306
8	NEC	254
9	Nvidia	217

Showing 1 to 10 of 109 entries   Previous Next
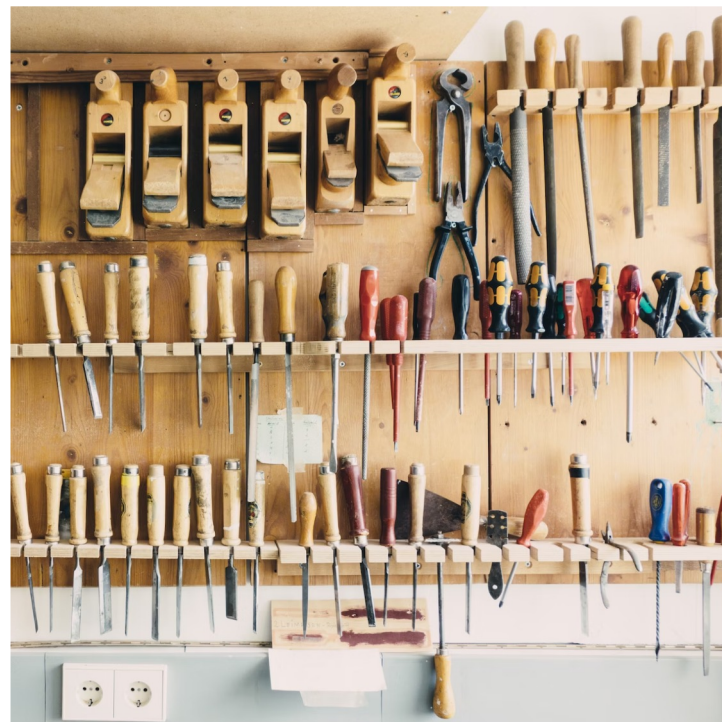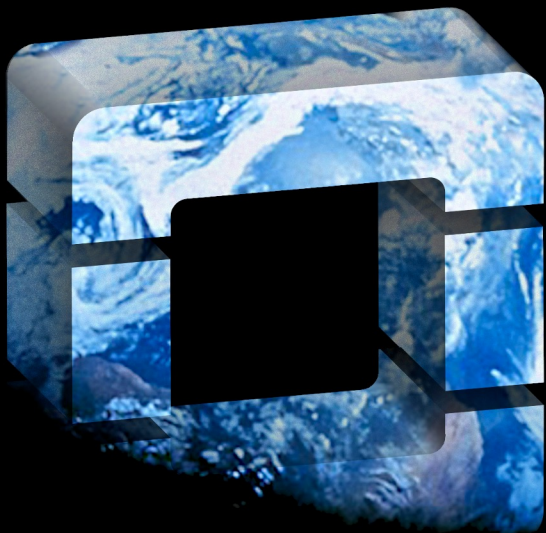
Workshops and design

- Workshops available
  - OpenStack
  - Platforms and Azimuth

- Requirements capture

- Architecture design

- Implementation

- Training and knowledge transfer

- Support and maintenance

Thank You

Come and visit us at
Booth 12!

https://www.stackhpc.com

StackHPC
The Rise of the HPC Cloud