



Elastic Analysis Facility

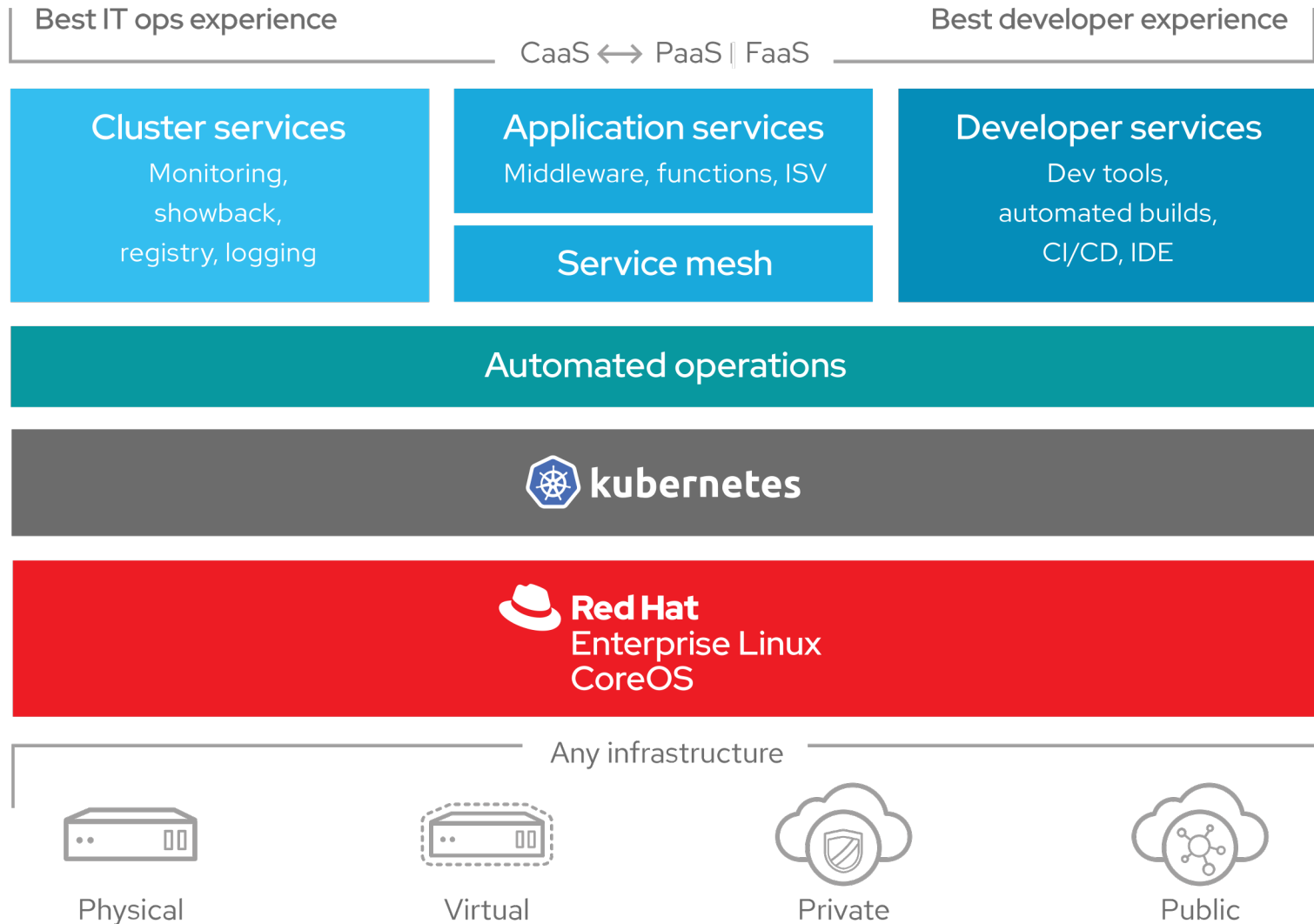
Projects Meeting

Aug 13 2020

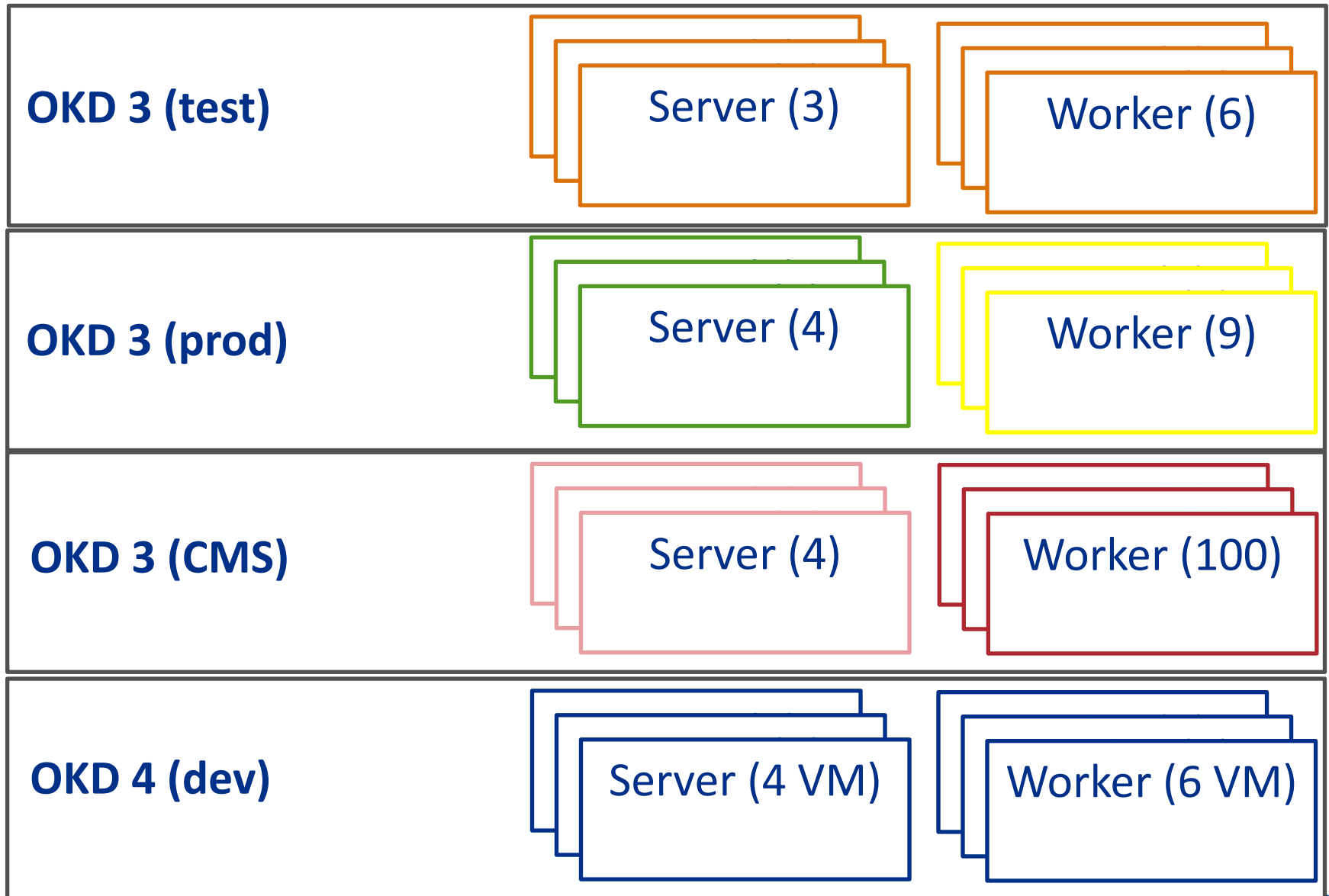
Elastic Analysis Facility goals

- As a foundation layer, use a **containerized infrastructure** (Kubernetes-based). Stable, in-production, operated and supported by routine SCD Facility operations staff. “Routine” corresponds to following a set of documented procedures with responsibilities rotated on a regular basis with staff.
- A **Jupyterhub** deployment that provides a number of different environments with LDAP authentication for Fermilab users. Allow users to customize the user-space environment.
- Support “typical” (**row-wise**) analysis needs
- Support a columnar analysis low latency application (e.g. Apache Spark, **Dask**, ...)
- **Document and package** the Elastic AF so it can be deployed elsewhere (e.g. at a US CMS Tier 2)

Elastic AF: layered architecture



State of the infrastructure



Elastic Analysis Facility: what's new

- Red Hat OKD 4.0 official release
- UNL gave demo of analysis application at PyHEP at a US CMS analysis facility
 - <https://www.youtube.com/watch?v=CDIFd1gDbSc>
 - Condor used as worker scheduler; tokens instead of x509
 - We will test redeployment at FNAL
- Ceph on NVMe storage configured as storage for OKD
 - Minio (object store) successfully installed on top
- Met with Rob Gardner
 - Schedule another “Chicago Kubernetes group” meeting?
 - Show and tell
 - K8s best practices