



Exploring SLATE/K8s to deploy CMS Tier 3 services at Notre Dame

Kenyi Hurtado



khurtado@nd.edu

January 20, 2020

K8s-HEP Meetup ('Chicagoland')

USCMS Tier 3s

The point of a CMS Tier 3 is to utilize local computing resources to do any of the following tasks: develop analysis code, access CMS data, submit jobs to remote CMS resources and/or store analysis outputs.

USCMS Tier 3s and Support

Doug Johnson, Carl Lundstedt, Kenyi Hurtado (Operations Team)
James Letts (Coordinator)

- 10-20 Tier 3s in USCMS
- Some can have login nodes with some batch system and cvmfs, and/or a Storage Element or Compute Element, or all of it (Tier-2 like).
- Tier 3 support team composed by 3 people + coordinator
 - Ticket based system to address issues with Tier 3 related software: GridFTP, XRootD, HTCondor-CE, etc

Motivation

Wouldn't it be nice to have a uniform layer for deploying all these services? Enter Kubernetes....



Easier to ask for system administrators to maintain an industry standard product like Kubernetes than installing and maintaining several specialized grid related components (HTCondor-CE, GridFTP, XRootD, etc).

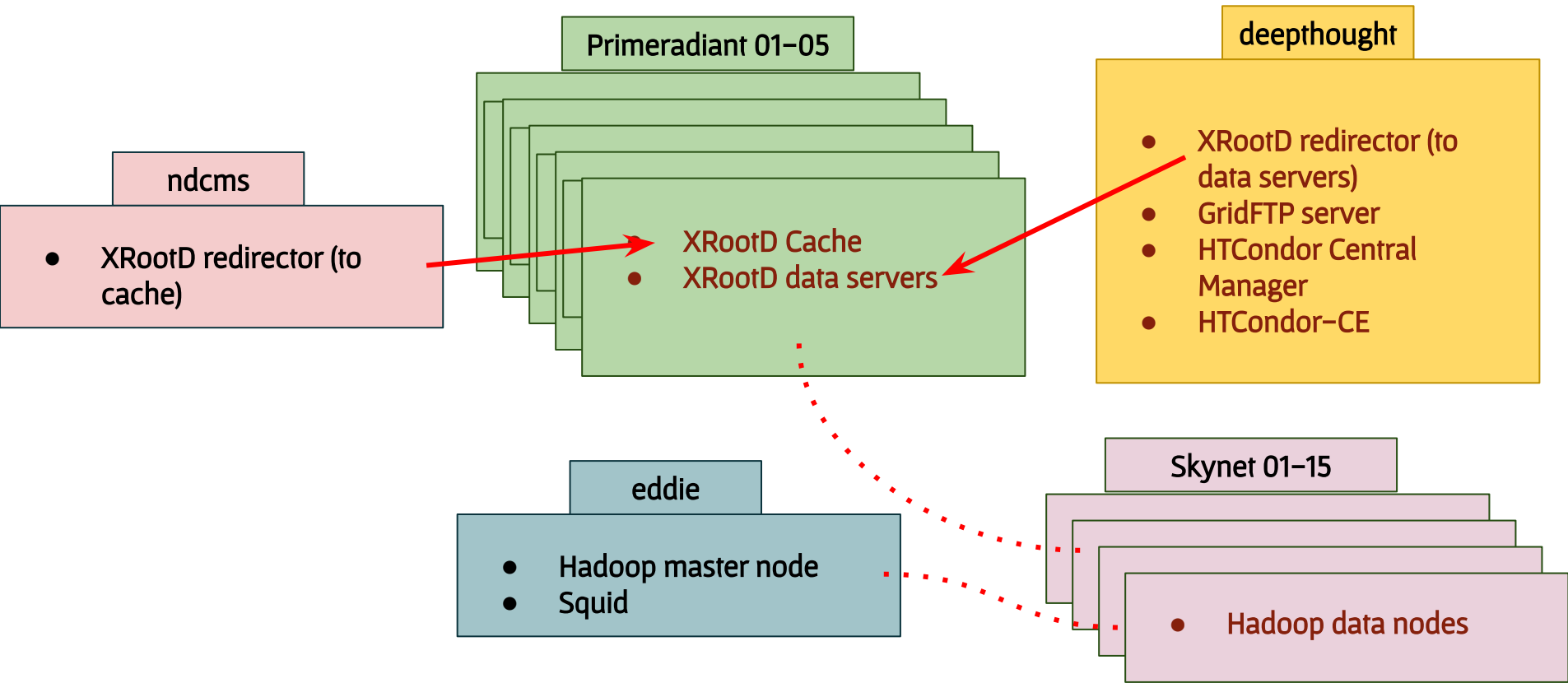
Motivation

IDEA...

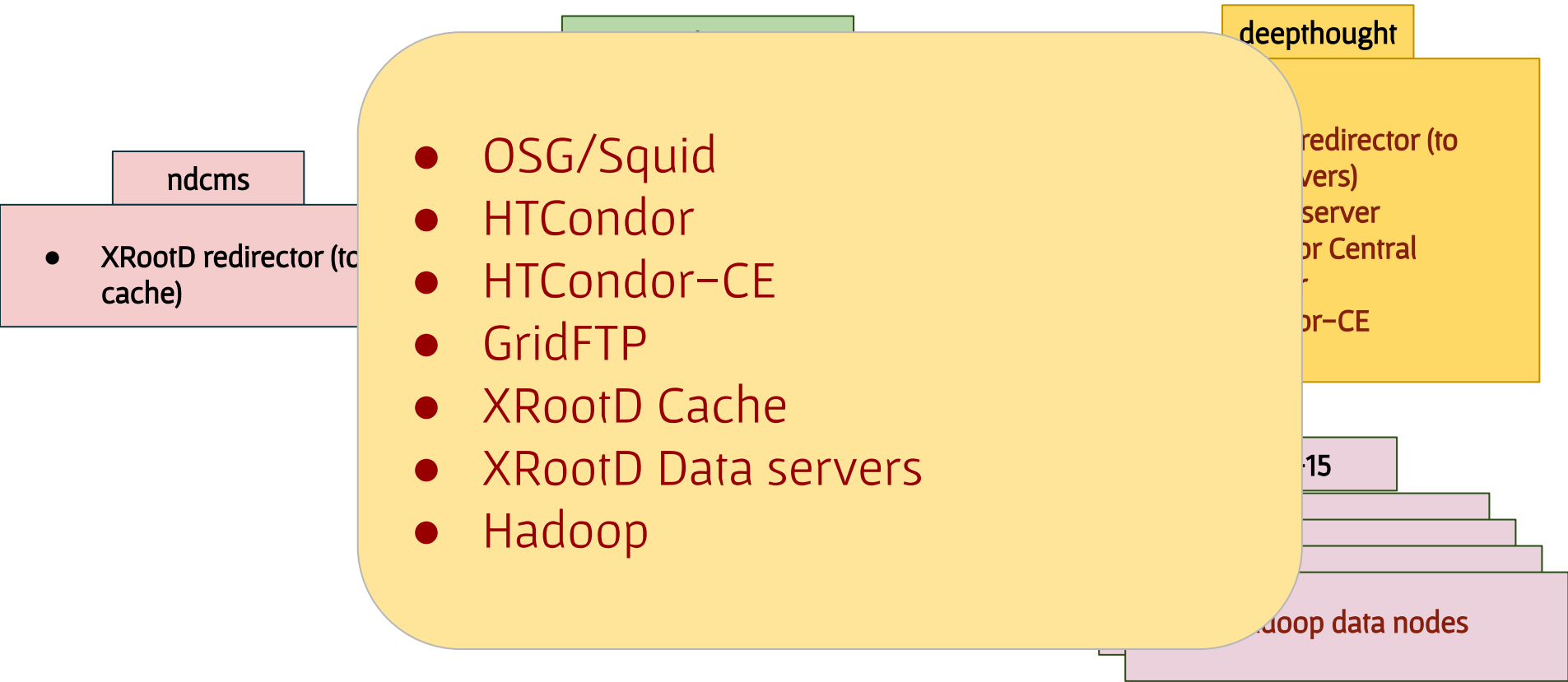
Have a (guinea pig) Tier 3 Site and deploy all components via K8s. Document and use it as model for other Tier 3s.

Chosen guinea pig: Notre Dame

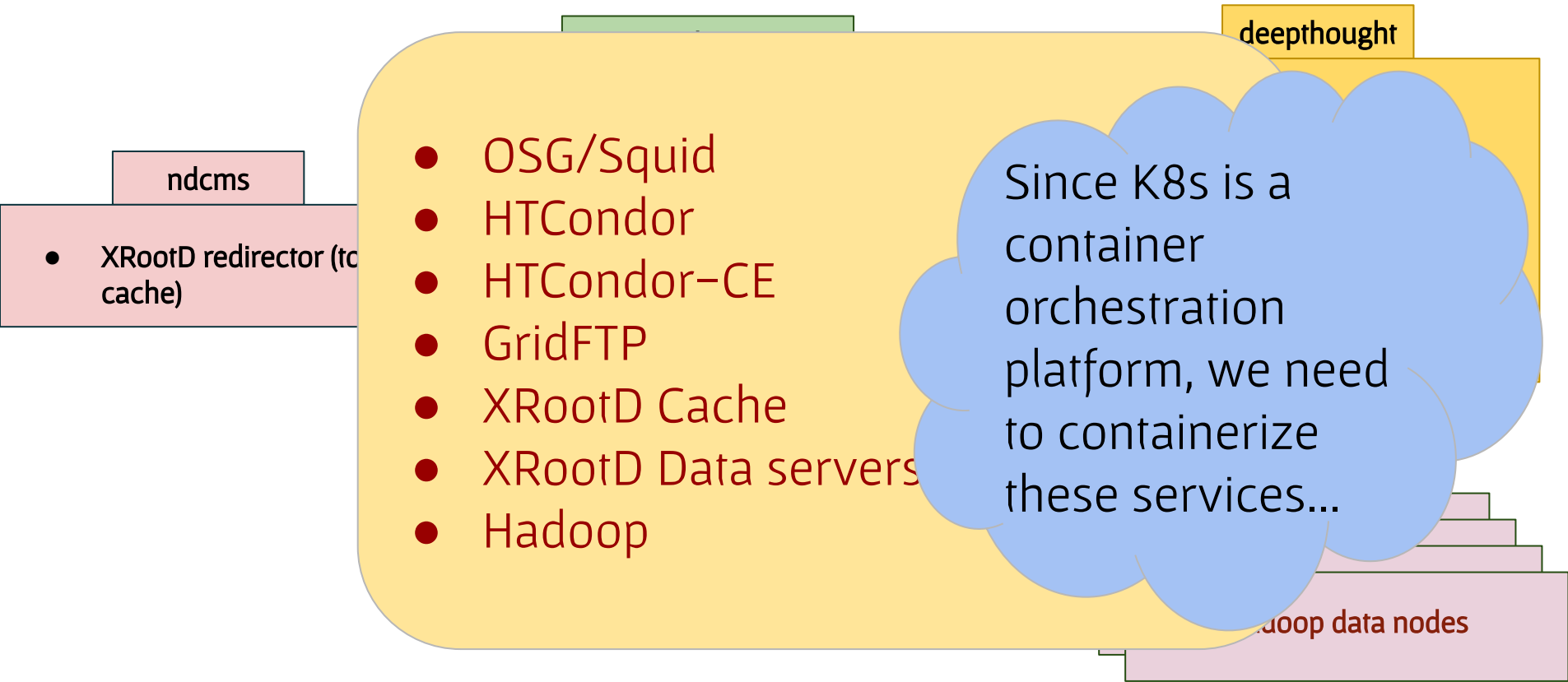
NDCMS Tier 3



NDCMS Tier 3



NDCMS Tier 3





SLATE

Federated "NoOps" operations model
containerizing several services already.

Slate apps



- **GridFTP**
- **XCache**
- **OSG-Frontier Squid**
- **HTCondor**
- HTCondor-CE
- Hadoop
- Spark
- Jupyterhub

SLATE	
gridftp	Globus GridFTP data management system
htcondor	HTCondor distributed high-throughput computing system
minio	MinIO is a high performance distributed object storage server, designed for infrastructure.
osg-hosted-ce	OSG Hosted Compute Element
perfsonar-testpoint	Perfsonar Testpoint Deployment
stashcache	StashCache is an XRootD-based caching service
grafana	The leading tool for querying and visualizing time series and metrics.
xcache	XCache is a xrootd based caching service for k8s

The Plan

- Deploy all T3-related services via Slate/Kubernetes at Notre Dame
 - Maintain K8s cluster
 - Profit from e.g.: upgrades & security updates being pushed to container images (and get them by simply re-deploying their pods)
- If this works out:
 - Document SLATE/k8s-based T3 deployment model for other T3s in USCMS.

What we have so far

- Barely starting looking into this...
 - Got 3 new machines to deploy the K8s cluster at ND a couple of weeks ago
 - 32 cores, 256 GB, 6 TB, 10 Gbps...
 - K8s cluster is working (master and 2 workers)

What we have so far

- Can't use SLATE yet though
 - Requested public IPs / domain names for them.
 - Need at least the K8s master to be public for SLATE API server to reach it.

To do list...

- Once that is done, the plan would be to deploy all T3 services one by one, starting with:
 - Squid
 - XCache
 - GridFTP

To do list...

- Eventually working with SLATE folks for the apps in curated stage, like HTCondor-CE, XRootD servers, Hadoop?....
 - Helm charts sanitisation, etc.

Thanks!
Questions?