Containerization: Streamlining Operations and Reducing Downtime

IAM Online
Wednesday, October 9, 2019

Paul Caskey, Internet2
Paul Riddle, University of Maryland Baltimore County
Chris Sutherin, University of Maryland Baltimore County

Welcome!

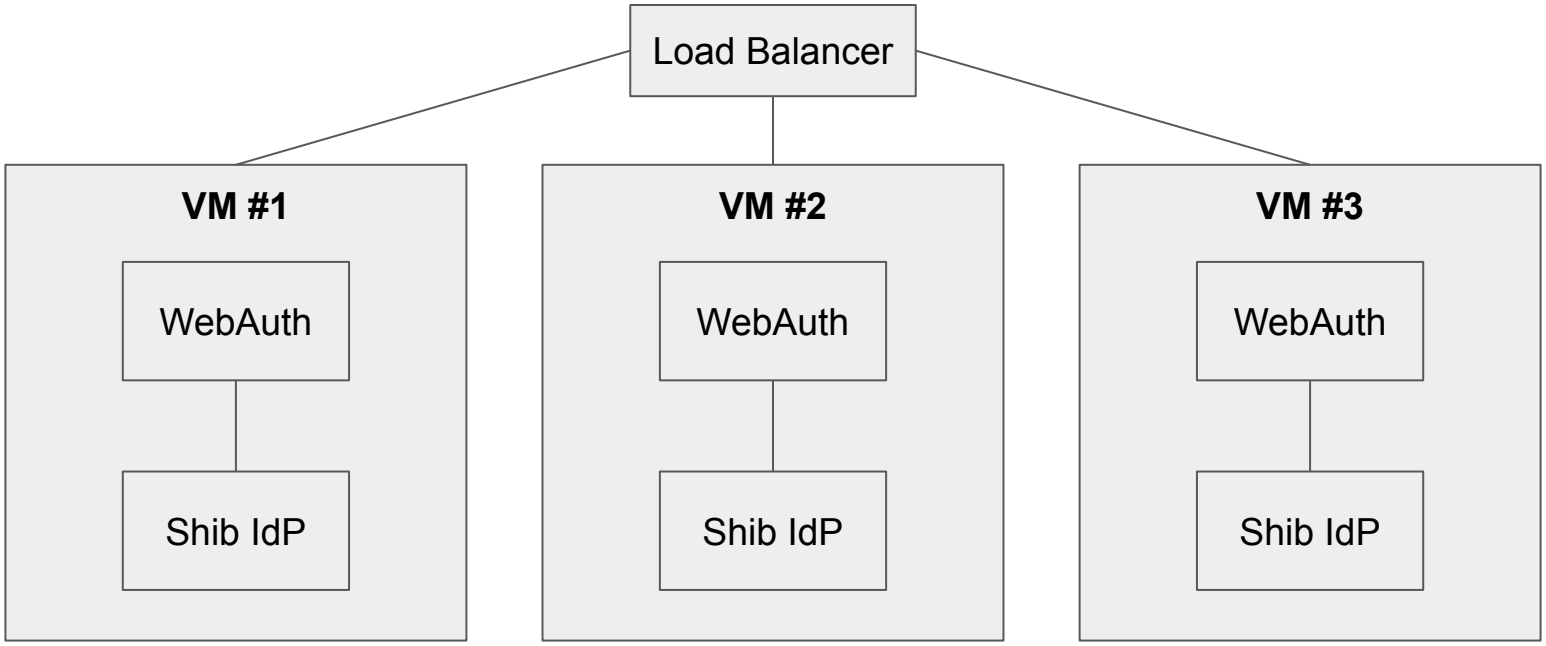
- What is the InCommon Trusted Access Platform (InCTAP)?
- Why are containers useful/important?
- Welcome to today's speakers!

Jump On In - The Water's Fine

"Containerizing" the Shibboleth IdP: Why and How UMBC Took the Plunge

UMBC's IdP Architecture, c.2007-2018

- 3 identical VMs behind a load balancer
- Home-grown Web SSO platform (WebAuth) handles AuthN
- Configuration maintained in a Subversion repository
- Changes propagated manually to each node



It worked OK, but...

- Simple configuration changes were tedious and time-consuming to implement
- Manual change propagation led to configuration inconsistencies across nodes
  - What happens if a node is down when you push changes out?
- Upgrading was tedious
  - Supporting packages (Java, Jetty, etc) tended to get stale over time
- No auto-scaling capability
- Containers can solve all of our problems! What's a container?

A container:

- Can be thought of as a lightweight VM
  - No hypervisor, hardware emulation, or kernel
  - Can run a standard OS, or not
- Typically packages a single application or service
- Can be networked together with other containers to build complex services
- Are built from *images*

An image:

- Provides the “blueprint” for running one or more containers
- Is built using a process similar to a Makefile:
  - Start with a base image (often an OS like CentOS or Ubuntu)
  - Install all of the service’s prerequisite packages
  - Install the service itself
- Can be uploaded to a registry and shared

How do I run this stuff?

- Docker: an ecosystem for building, packaging, and deploying applications using containers
  - Builds images
  - Pulls/pushes images to & from registries
  - Runs containers
  - And much more
- Available for most OS platforms
- Community and Enterprise editions
- There are other container management platforms, but Docker is the most popular and established

Packaged Shibboleth IdP

- Provided by the InCommon Trusted Access Platform
- Base image includes the OS, IdP and all prerequisites
  - No longer necessary to manually install Apache, Java, Tomcat/Jetty, etc.
- IdP works out of the box - just add your configuration
- Upgrading is easy:
  - Pull the latest base image and rebuild
  - Underlying OS and prerequisite packages are automatically kept up-to-date
- No more building servers by hand!!

Sounds great.. but where to start?

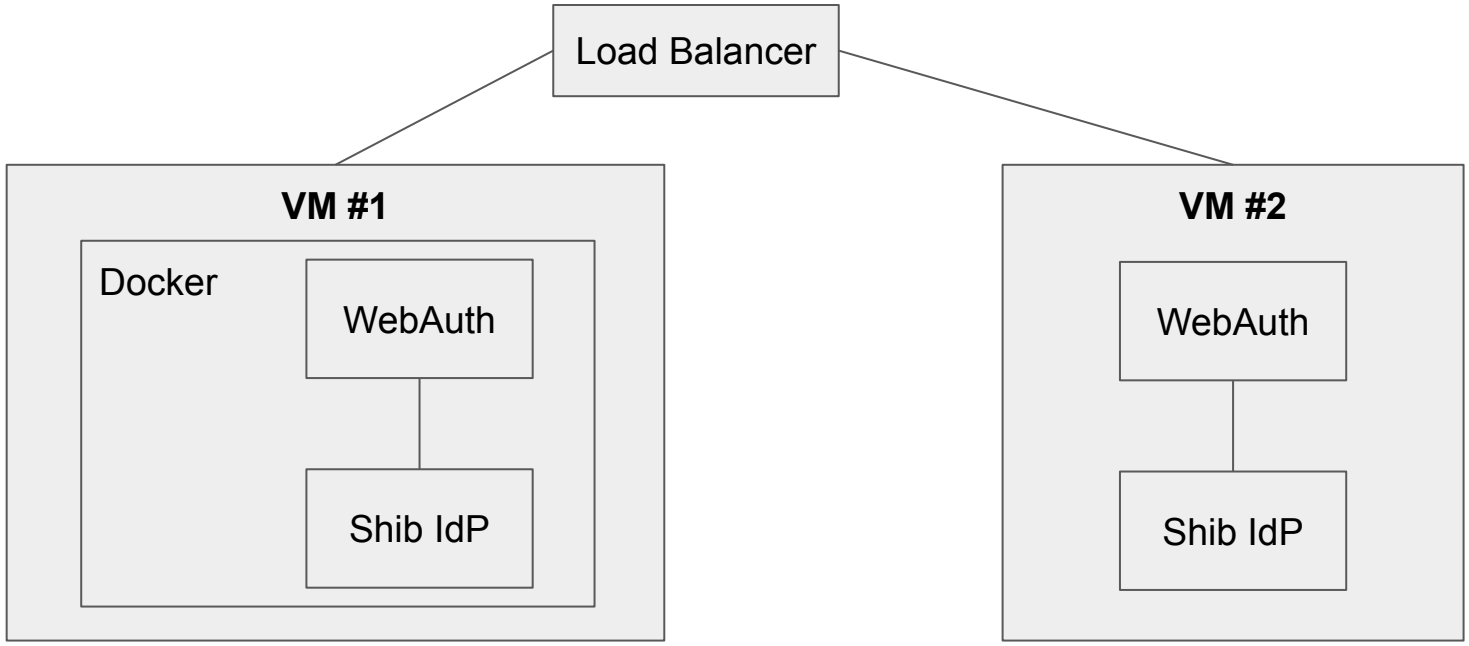
- We had no in-house experience with containers
  - Learned by doing -- Google, Stack Overflow
  - Learned by teaching -- not necessarily recommended :)
- We had (and still have) no organized “DevOps” infrastructure (but making progress)
  - Git server
  - Docker registry
  - etc..
- We decided to “get our feet wet” prior to jumping in
  - Run containerized IdP alongside our existing environment; see how it goes

First Steps (testing the water..)

- Tend to our existing Shibboleth environment
  - Upgrade Java, Jetty, IdP, etc to latest & greatest versions
  - Switch LB from round-robin to single node with failover
- “Containerize” legacy Web SSO software to run alongside IdP in same container
- Bootstrap IdP container with identical configuration to current production
- Run and observe in production for several months while we work through infrastructure challenges

“Transitional” IdP Architecture, c.2018

- Down from 3 VMs to 2
  - 1 VM running Docker and containerized IdP
  - 1 “traditional” VM
- Configuration still maintained and propagated manually

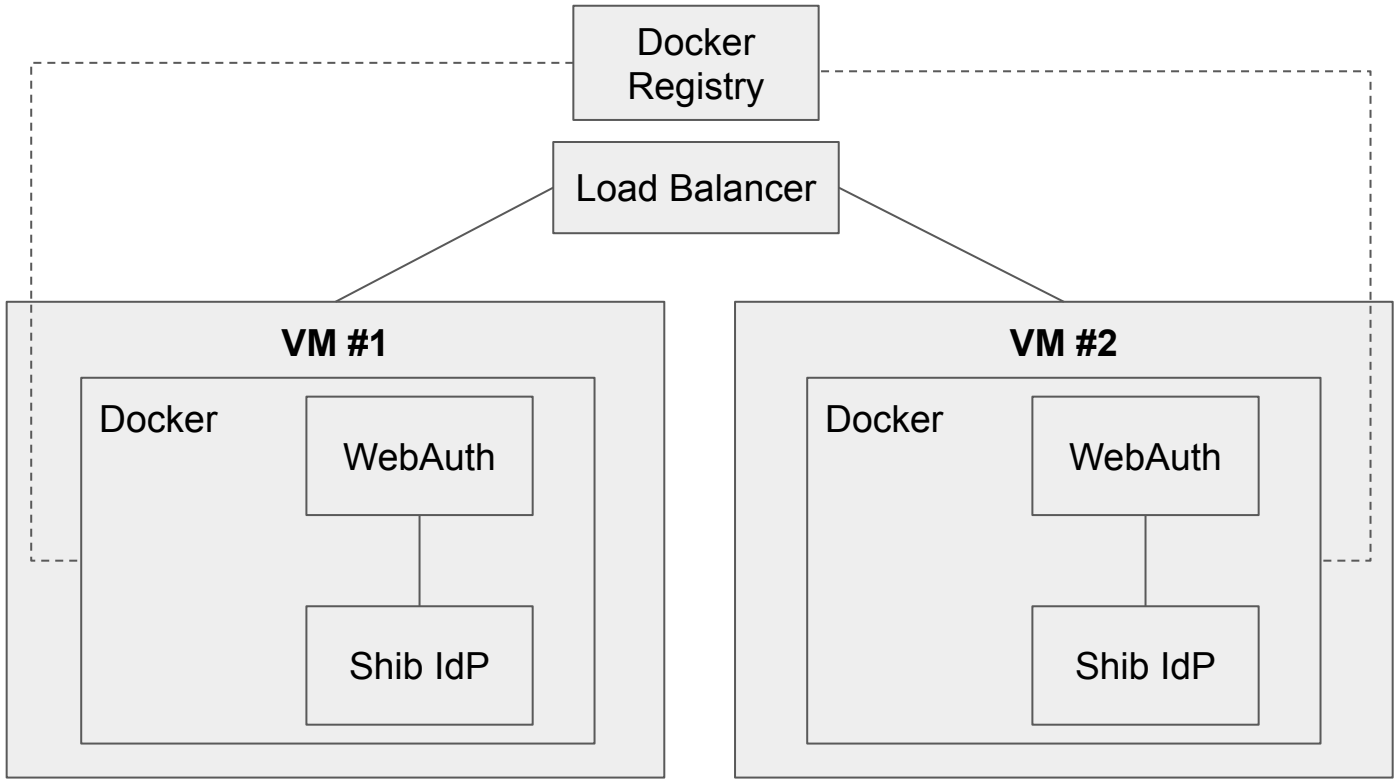


The next step: a container-only deployment

- Once we were confident the container worked well, we phased out the legacy VM
- We now have 2 load-balanced VMs both running the IdP in a container (+ a third for development/staging)
- Same infrastructure (on-premise VMs and LB)
- Docker images hosted in a local registry
  - Both Docker instances run identical containers launched from the same image
  - Solves one of our use cases (synchronization)
- Things are still not as automated as we would like

Present IdP Architecture (2019)

- 2 VMs running Docker and containerized IdP, behind LB
- Containers pushed to a local Docker registry
- Configurations maintained in a local Git repository



Next Steps

- Phase out our legacy Web SSO product
- Get enterprise-level support infrastructure in place
  - Gitlab/Github
  - Docker registry (AWS ECR?)
- CI/CD pipeline
  - Addresses another use case (automation)
- Move IdPs out to the cloud (ECS/Fargate)
  - Some prerequisites must be moved first (LDAP)
  - Want to keep an IdP node on-premise - DNS becomes a challenge
- Container Orchestration?

Topics Discussed

Moving from a manually installed Grouper environment to using delivered containers

Advantages of containerized Grouper

Where we plan on using containers outside of Grouper

- PeopleSoft pain points
- Our desired direction

Transitioning to Grouper Containers

- Prior to using containerized Grouper, we struggled with the initial setup from software installation to correctly configuring the system. We initially spent a lot of time in this phase.
- Once set up, maintenance was neglected. Code bases on different servers would get out of sync.
- Patching was frequently skipped.
- In 2017 we joined the Internet2 Campus Success Program (CSP) for the TIER containers (now InCommon Trusted Access Platform).
- We went live with Grouper containers in the summer of 2018 and never looked back.

Advantages of Containerized Grouper

- I am not a Docker expert and I don't need to be!
- Setup is very simple. Create your Grouper configuration files and copy them into the image and you're off and running.
- Patching is greatly simplified. Just pull the Grouper Docker image you want and it's upgraded. Downtime is minimized.
- There are numerous support options, Slack and email.
- Very scalable. Simply bring up new containers as needed.
- Easily customizable. Run each service together, individually or and combination.

How We Use Grouper

- We pull student, HR and financial data from Oracle (PeopleSoft), campus portal data from MySQL, warehouse data from SQL Server and affiliations from LDAP.
- We provision group membership to the eduPersonEntitlement attribute in LDAP.
- We provision groups to LDAP, AD and Google.
- Examples are email lists for staff, access rights for advisors, students allowed to access campus television and Box.

Where Do We Go From Here

- Today our image and container management is manual.
- To simplify managing images we will create an image repository.
- Ultimately moving to AWS and implementing a container management system.

Using Containers Outside of Trusted Access Platform

- We have recently started the process of creating images to run our PeopleSoft student, financials and HR systems.
- The PeopleSoft maintenance cycle is endless and labor intensive.
- Patching is prone to errors.

PeopleSoft Patching Cycle

Each quarter we must perform the following to patch our PeopleSoft systems

- Apply OS patch to all servers.
- Apply a Java patch to all web servers.
- Apply a WebLogic patch to all web servers.
- Apply a Tuxedo patch to all application servers.
- Apply the PeopleTools patch to all database instances, application and web servers.
- Reconfigure every application server domain and process scheduler.
- Redeploy every web server.
- Copy the SSO files to every application server.
- Copy the SSO files to web server for ever instance

PeopleSoft Patching Cycle Continued

To put this into perspective.

Hardware systems running PeopleSoft

- Web server machines: 19
- Application server machines: 16
- Database servers: 6

PeopleSoft Patching Cycle Continued

Campus Solutions

Human Resources

Finance

Development

Production

Patches Applied:

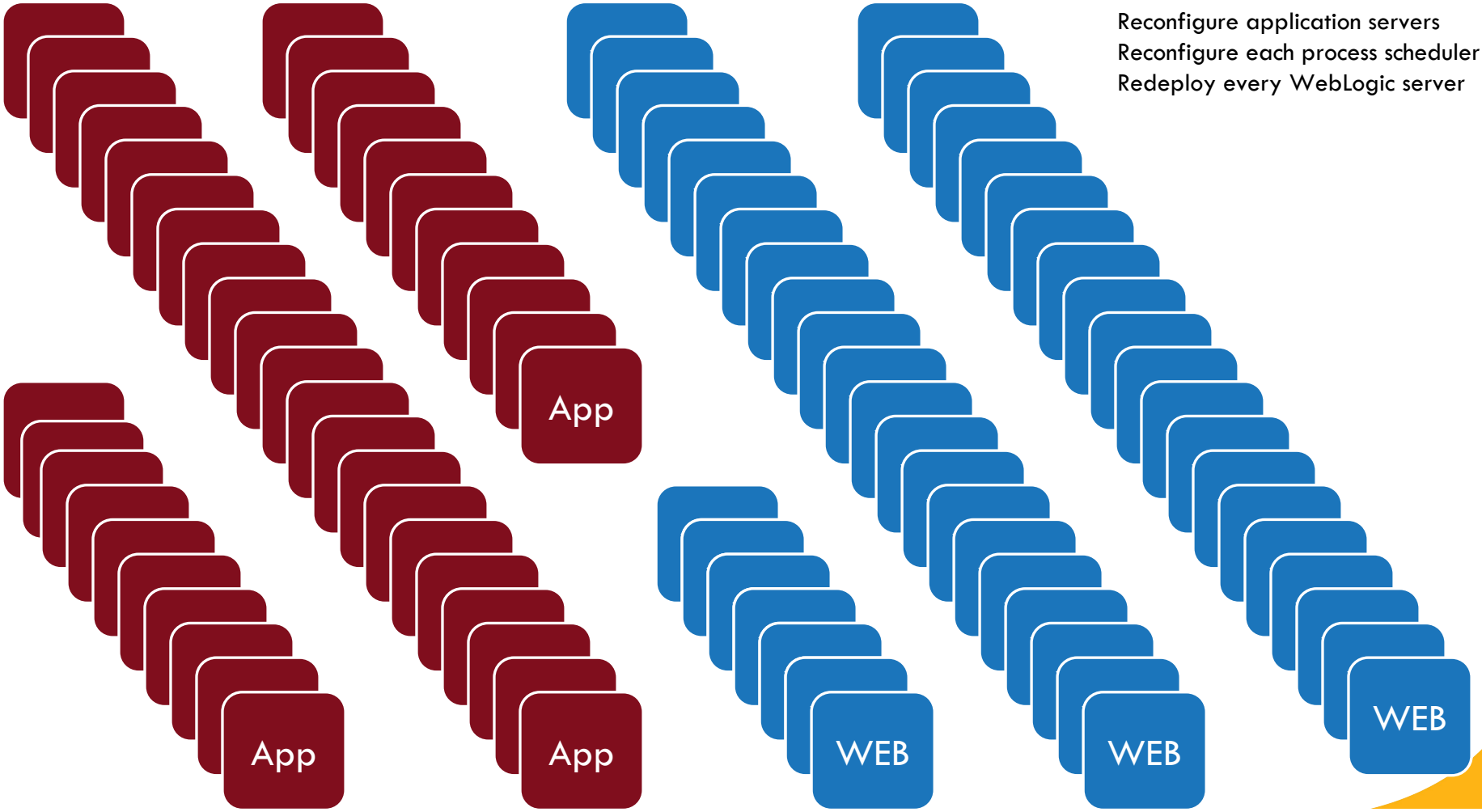
OS Patch
Java Patch
WebLogic Patch
Tuxedo Patch
PeopleTools Patch

PeopleSoft Patching Cycle Continued

PeopleSoft services running

- Total PeopleSoft application: 41
- Application servers: 81
- Process scheduler: 78
- WebLogic web servers: 86
- Databases: 41

PeopleSoft Patching Cycle Continued



Reconfigure application servers
Reconfigure each process scheduler
Redeploy every WebLogic server

PeopleSoft Patching with containers

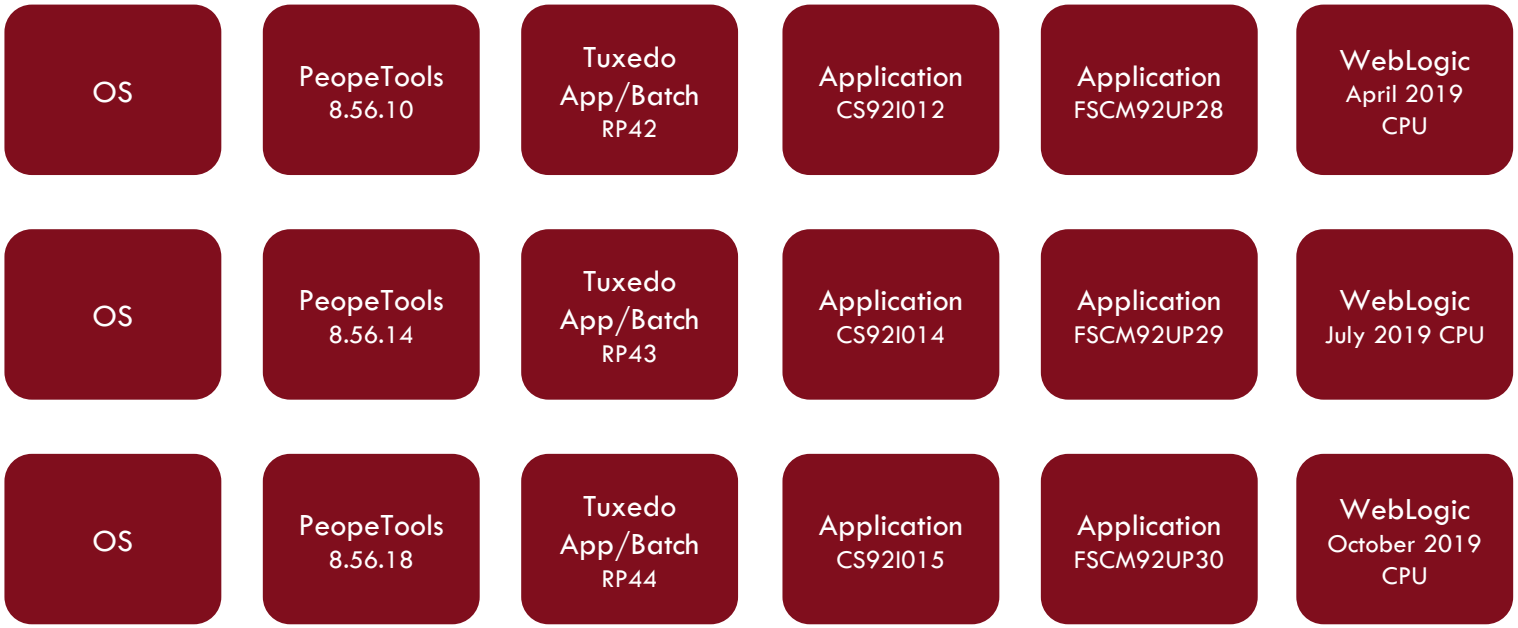
Hardware systems running PeopleSoft

- Build OS image with the packages required by PeopleSoft.
- Build the PeopleTools image that contains all of the PeopleSoft software.
- Build an application image for each system with the current application patch.
- Build WebLogic image
- Build Tuxedo image for the application server and process scheduler
- Use the appropriate image to start WebLogic containers, application server containers and process scheduler containers for each PeopleSoft system

This model will greatly reduce patching efforts with the patch only being applied once per image.

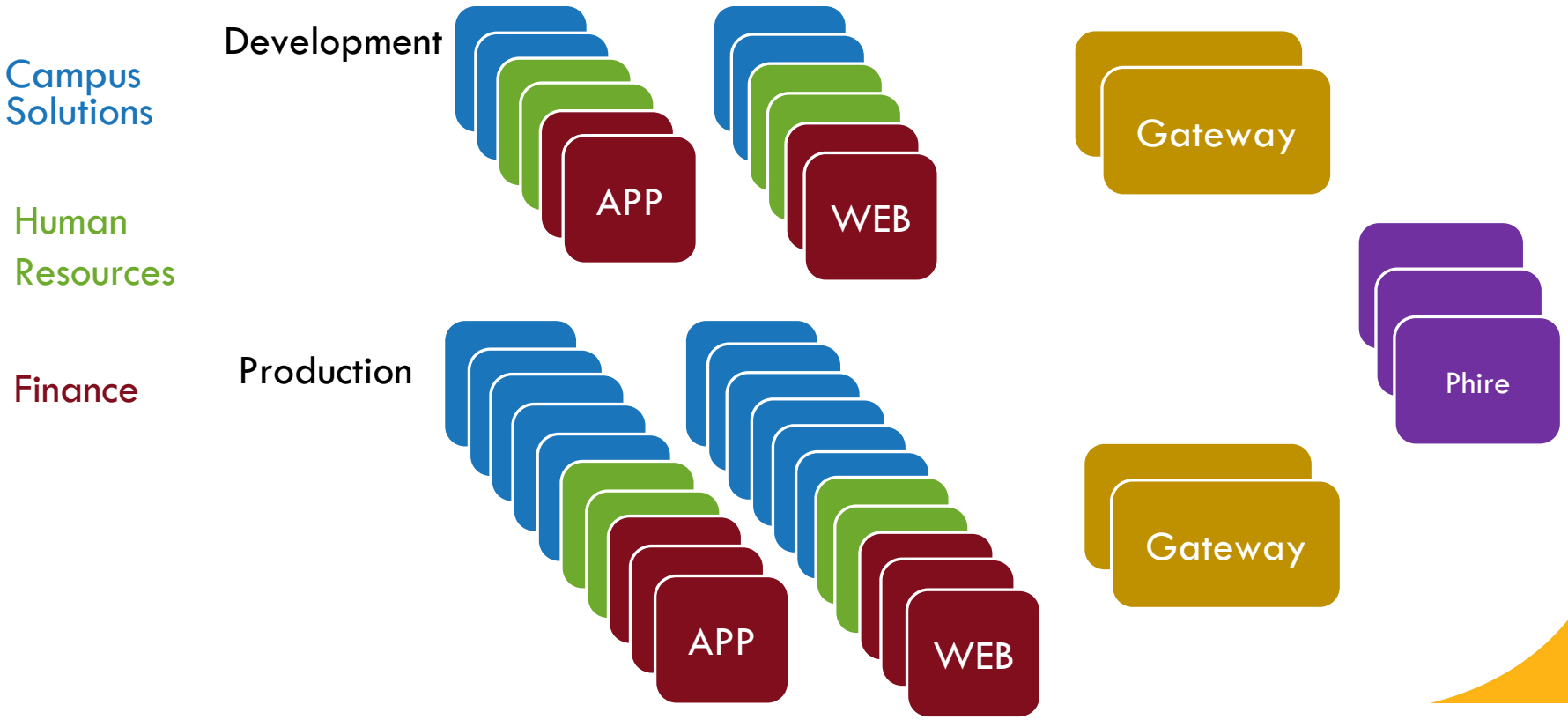
PeopleSoft Patching With Containers Continued

Create a new image for each patch system. Patching only once per image.



PeopleSoft Patching With Containers Continued

Run the new image for each patched system.



Problems

Problems that arise beside the amount of manpower spent quarterly

- Missed redeployment of a web server.
- Missed reconfiguration of an application server.
- Missed redeployment of SSO files.

All of these negatively affect the end user experience.

The Solution

As we move forward with containerizing Peoplesoft our hope is it will make the maintenance process easier.

By building images for PeopleTools, PeopleSoft application, Tuxedo and WebLogic, we will be able to quickly deploy a new container for each system. Resolving all of the issues created from missed items.

2019 Technology Exchange

https://meetings.internet2.edu/2019-technology-exchange/

December 9-13, 2019
New Orleans, Louisiana

CAMP – Two days of campus case studies and key identity management issues

AdvanceCAMP – The premier forward-looking meeting with international IdM thought leaders

Join Us!

Upcoming Training

https://www.incommon.org/academy/library/

Shibboleth Installation Workshop

Denver, Colorado - October 22-23, 2019

Grouper School

November 12-13, 2019 – Philadelphia, Pennsylvania

COmanage Class

November 12-13, 2019 – Philadelphia, Pennsylvania

midPoint Basics

December 3-4, 2019 – Online

Please evaluate today's session

https://www.surveymonkey.com/r/IAMOnline-Oct2019