

TCSS 562: SOFTWARE ENGINEERING FOR CLOUD COMPUTING

Containerization, Kubernetes

Wes J. Lloyd
School of Engineering and Technology
University of Washington – Tacoma
MW 5:50-7:50 PM



OBJECTIVES – 11/23

- **Questions from 11/18**
- **Quiz 2**– extended until Wed 11/25 @ 11:59p
- **No Office Hours 11/27**
- **Class on 11/25:**
 - Office hours, and finish any remaining lecture from today
- **Introduction to Containerization cont'd**
- **Tutorial 7**
- **2nd hour:**
 - **Introduction to Kubernetes**
 - **Tutorial questions**
 - **Team planning**

November 23, 2020

TCSS562:Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.2

ONLINE DAILY FEEDBACK SURVEY

■ Daily Feedback Quiz in Canvas – Take After Each Class

■ Extra Credit for completing

Announcements

Assignments

Discussions

Zoom

Grades

People

Pages

Files

Quizzes

Collaborations

UW Libraries

UW Resources

▼ Upcoming Assignments

Class Activity 1 – Implicit vs. Explicit Parallelism

Available until Oct 11 at 11:59pm | Due Oct 7 at 7:50pm | ~10 pts

Tutorial 1 - Linux

Available until Oct 19 at 11:59pm | Due Oct 15 at 11:59pm | ~20 pts

▼ Past Assignments

TCSS 562 - Online Daily Feedback Survey - 10/5

Available until Dec 18 at 11:59pm | Due Oct 6 at 8:59pm | ~1 pts

TCSS 562 - Online Daily Feedback Survey - 9/30

Available until Dec 18 at 11:59pm | Due Oct 4 at 8:59pm | ~1 pts

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.3

TCSS 562 - Online Daily Feedback Survey - 10/5

Started: Oct 7 at 1:13am

Quiz Instructions

Question 1

0.5 pts

On a scale of 1 to 10, please classify your perspective on material covered in today's class:

12345678910

Mostly Review To MeEqual New and ReviewMostly New to Me

Question 2

0.5 pts

Please rate the pace of today's class:

12345678910

SlowJust RightFast

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.4

MATERIAL / PACE

- Please classify your perspective on material covered in today's class (22 respondents):
 - 1-mostly review, 5-equal new/review, 10-mostly new
 - **Average – 6.30** (↑ - *previous 6.09*)
- Please rate the pace of today's class:
 - 1-slow, 5-just right, 10-fast
 - **Average – 5.35** (↓ - *previous 5.55*)

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.5

FEEDBACK FROM 11/18

- *Can you please help me to understand how data travels between cloud regions?*
- *Is it the same as how the Internet works (like packets travelling over the network) or any different way?"*
- Yes, travel from one cloud region to another will typically be routed over the Internet as regions span considerable distances (thousands of KMs)
- Please note there are substantial data egress charges
- S3 to off-cloud: 9 ¢/GB
- S3 to distant regions: 2 ¢/GB
- S3 to local regions (Ohio→Virginia): 1 ¢/GB

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.6

FEEDBACK - 2

- Is 3-tier architecture still the dominant architecture in cloud?
- Three tier architecture?
- Do you mean model-view-controller?

- How is that different from service oriented architecture in cloud?
- Model-view-controller is one possible way to compose components of an service oriented application
- Model: data persistence
- View: webservices to render web-based GUI
- Controller: webservices to implement application logic

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.7

TUTORIAL QUESTIONS

- Tutorial 5: Thursday Nov 19th @ 11:59p

- Tutorial 6: Tuesday Nov 24th @ 11:59p

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.8

UPCOMING TUTORIALS

- Tutorial 7 – [POSTED] Introduction to Docker Containerization
- Extra credit tutorials – *submit by Dec 18 @ 11:59p*
- Tutorial 8 – Introduction to FaaS IV: Step Functions and SQS
- Tutorial 9 – Asynchronous Function Profiling with SAAF
- Ungraded tutorials:
- Tutorial 10 – Automating Experiments with SAAF & FaaS Runner
- Tutorial 11 – Scaling beyond a single client – concurrent webservice benchmarking with multiple EC2 instances

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.9

OBJECTIVES – 11/23

- Questions from 11/18
- Quiz 2 – extended until Wed 11/25 @ 11:59p
- No Office Hours 11/27
- Class on 11/25:
Office hours, and finish any remaining lecture from today
- Introduction to Containerization cont'd
- Tutorial 7
- 2nd hour:
- Introduction to Kubernetes
- Tutorial questions
- Team planning

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.10

QUIZ 2

- **Hello Professor – I am not happy with my quiz score, can I retake the quiz?**
- Quiz grade from Canvas is temporary
- Partial credit will be applied for questions having second/third best answers based on complete review of all quiz attempts the class completes the quiz
- Final quiz score will likely be curved
 - This helps to adjust for poorly written/understood questions
- Tutorials 8 & 9 will be available for extra credit

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.11

W Should TCSS 562 Fall 2020 include a Quiz 3?

YES - Include Quiz 3 -
offered week of Dec 7th

No - Skip Quiz 3, Base
the Quiz score off of Quiz
1 & 2 only

Indifference - I am
happy with any outcome
: quiz or no quiz

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

OBJECTIVES - 11/23

- Questions from 11/18
- Quiz 2- extended until Wed 11/25 @ 11:59p
- No Office Hours 11/27
- Class on 11/25:
Office hours, and finish any remaining lecture from today
- Introduction to Containerization cont'd
- Tutorial 7
- 2nd hour:
 - Introduction to Kubernetes
 - Tutorial questions
 - Team planning

November 23, 2020

TCSS562:Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.13

OBJECTIVES - 11/23

- Questions from 11/18
- Quiz 2- extended until Wed 11/25 @ 11:59p
- No Office Hours 11/27
- Class on 11/25:
Office hours, and finish any remaining lecture from today
- Introduction to Containerization cont'd
- Tutorial 7
- 2nd hour:
 - Introduction to Kubernetes
 - Tutorial questions
 - Team planning

November 23, 2020

TCSS562:Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.14

GROUP PRESENTATION

- **TWO OPTIONS:**
- ***Cloud technology presentation***
- ***Cloud research paper presentation***
 - Recent & suggested papers will be posted at:
<http://faculty.washington.edu/wlloyd/courses/tcss562/papers/>
- **Submit presentation type and topics (paper or technology) with desired dates of presentation via Canvas by:
*Monday November 23rd @ 11:59pm***
- **Presentation dates:**
 - Monday November 30, Wednesday December 2
 - Monday December 7, Wednesday December 9

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.15


OBJECTIVES – 11/23

- **Questions from 11/18**
- **Quiz 2– extended until Wed 11/25 @ 11:59p**
- **No Office Hours 11/27**
- **Class on 11/25:**
 - Office hours, and finish any remaining lecture from today
- **Introduction to Containerization cont'd**
- **Tutorial 7**
- **2nd hour:**
 - Introduction to Kubernetes
 - Tutorial questions
 - Team planning

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.16



CONTAINERIZATION

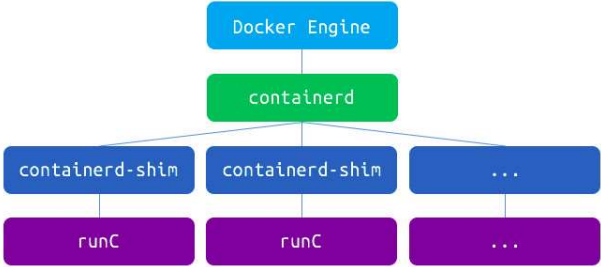
November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.17

OTHER DOCKER TOOLS

- **Docker Machine:**
automatically provision and manage sets of docker hosts to form a cluster
- **Docker Swarm:**
Clusters multiple docker hosts together to manage as a cluster.
- **Docker Compose:** Config file (YAML) for multi-container application; Describes how to deploy and configure multiple containers



```
graph TD;
    DE[Docker Engine] --> C[containerd];
    C --> CS1[containerd-shim];
    C --> CS2[containerd-shim];
    C --> Dots1[...];
    CS1 --> R1[runC];
    CS2 --> R2[runC];
    Dots1 --> Dots2[...];
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.18

CONTAINER ORCHESTRATION FRAMEWORKS

- Framework(s) to deploy multiple containers
- Provide container clusters using cloud VMs
- Similar to “private clusters”
- Reduce VM idle CPU time in public clouds
- Better leverage “sunk cost” resources
- Compact multiple apps onto shared public cloud infrastructure
- Generate to cost savings
- Reduce vendor lock-in

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.19

KEY ORCHESTRATION FEATURES

- Management of container hosts
- Launching set of containers
- Rescheduling failed containers
- Linking containers to support workflows
- Providing connectivity to clients outside the container cluster
- Firewall: control network/port accessibility
- Dynamic scaling of containers: horizontal scaling
 - Scale in/out, add/remove containers
- Load balancing over groups of containers
- Rolling upgrades of containers for application

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.20

CONTAINER ORCHESTRATION FRAMEWORKS - 2

- Docker swarm
- Apache mesos/marathon
- Kubernetes
 - Many public cloud provides moving to offer Kubernetes-as-a-service
- Amazon elastic container service (ECS)
- Apache aurora
- Container-as-a-Service
 - Serverless containers without managing clusters
 - Azure Container Instances, AWS Fargate...

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.21

OBJECTIVES - 11/23

- Questions from 11/18
- Quiz 2- extended until Wed 11/25 @ 11:59p
- No Office Hours 11/27
- Class on 11/25:
Office hours, and finish any remaining lecture from today
- Introduction to Containerization cont'd
- Tutorial 7
- 2nd hour:
 - Introduction to Kubernetes
 - Tutorial questions
 - Team planning

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma


L15.22

TUTORIAL #7

DOCKER, CGROUPS, RESOURCE ISOLATION

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma



L15.23

TUTORIAL COVERAGE

- **Docker CLI → Docker Engine (dockerd) → containerd → runc**
- **Concepts:**
 - Docker installation
 - Working with docker files
 - Docker run – create a container
 - Docker ps – list containers
 - Docker exec –it – run a process in an existing container
 - Docker stop –stop container

November 23, 2020	TCSS562: Software Engineering for Cloud Computing [Fall 2020] School of Engineering and Technology, University of Washington - Tacoma	L15.24
-------------------	--	--------

Commands:

attach

Build an image from a Dockerfile

build

Create a new image from a container's changes

commit

Copy files/folders between a container and the local filesystem

cp

Create a new container

create

Deploy a new stack or update an existing stack

deploy

Inspect changes to files or directories on a container's filesystem

diff

Get real time events from the server

events

Run a command in a running container

exec

Export a container's filesystem as a tar archive

export

Show the history of an image

history

List images

images

Import the contents from a tarball to create a filesystem image

import

Display system-wide information

info

Return low-level information on Docker objects

inspect

Kill one or more running containers

kill

Load an image from a tar archive or STDIN

load

Log in to a Docker registry

login

Log out from a Docker registry

logout

Fetch the logs of a container

logs

Pause all processes within one or more containers

pause

List port mappings or a specific mapping for the container

port

List containers

ps

Pull an image or a repository from a registry

pull

Push an image or a repository to a registry

push

Rename a container

rename

Restart one or more containers

restart

Remove one or more containers

rm

Remove one or more images

rmi

Run a command in a new container

run

Save one or more images to a tar archive (streamed to STDOUT by default)

save

Search the Docker Hub for images

search

Start one or more stopped containers

start

Display a live stream of container(s) resource usage statistics

stats

Stop one or more running containers

stop

Create a tag TARGET_IMAGE that refers to SOURCE_IMAGE

tag

Display the running processes of a container

top

Unpause all processes within one or more containers

unpause

Update configuration of one or more containers

update

Show the Docker version information

version

Block until one or more containers stop, then print their exit codes

wait

Docker CLI

TUTORIAL 7

- Linux performance benchmarks
 - stress-ng
 - 100s of CPU, memory, disk, network stress tests
- Sysbench
- Used in tutorial for memory stress test

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.26

**WE WILL RETURN AT
~6:56PM**



OBJECTIVES – 11/23

- **Questions from 11/18**
- **Quiz 2– extended until Wed 11/25 @ 11:59p**
- **No Office Hours 11/27**
- **Class on 11/25:**
 - Office hours, and finish any remaining lecture from today
- **Introduction to Containerization cont'd**
- **Tutorial 7**
- **2nd hour:**
 - **Introduction to Kubernetes**
 - **Tutorial questions**
 - **Team planning**

November 23, 2020

TCSS562:Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.28




KUBERNETES

from: "The Kubernetes Book", Nigel Poulton and Pushkar Joglekar, Version 7.0, September 2020

L15.29

KUBERNETES

- Name is from the Greek word meaning Helmsman
 - The person who steers a seafaring ship
 - The logo reinforces this theme
- Kubernetes is also sometimes called K8s
- Kubernetes is an application orchestrator
- Most common use case is to containerize cloud-native microservices applications
- What is an orchestrator?
 - System that deploys and manages applications



November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.30

KUBERNETES – 2

Why does Google want
to give Kubernetes away
for free?

- Initially developed by Google
- **Goal:** *make it easier for potential customers to use Google Cloud*
- Kubernetes leverages knowledge gained from two internal container management systems developed at Google
 - Borg and Omega
- Google donated Kubernetes to the Cloud Native Computing Foundation in 2014 as an open-source project
- Kubernetes is written in Go (Golang)
- Kubernetes is available under the Apache 2.0 license
- Releases were previously maintained for only 8 months!
- Starting w/ v 1.19 (released Aug 2020) support is 1 year

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.31

GOALS OF KUBERNETES

1. Deploy your application
 2. Scale it up and down dynamically according to demand
 3. Self-heal it when things break
 4. Perform zero-downtime rolling updates and rollbacks
- These features represent automatic infrastructure management
 - Containerized applications run in container(s)
 - Compared to VMs, containers are thought of as being:
 - Faster
 - More light-weight
 - More suited to rapidly evolving software requirements

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.32

CLOUD NATIVE APPLICATIONS

- Applications designed to meet modern software requirements including:
 - **Auto-scaling:** resources to meet demand
 - **Self-healing:** *required for high availability (HA) and fault tolerance*
 - **Rolling software updates:** with no application downtime for DevOPS
 - **Portability:** can run anywhere there's a Kubernetes cluster

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.33

WHAT IS A MICROSERVICES APP?

- Application consisting of many specialized parts that communicate and form a meaningful application
- Example components of a microservice eCommerce app:

Web front-end	Catalog service
Shopping cart	Authentication service
Logging service	Persistent data store
- **KEY IDEAS:**
 - Each microservice can be coded/maintained by different team
 - Each has its own release cadence
 - Each is deployed/scaled separately
 - Can patch & scale the log service w/o impacting others

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.34

KUBERNETES - 3

- Provides “an operating system for the cloud”
- Offers the de-facto standard platform for deploying and managing cloud-native applications
- OS: abstracts physical server, schedules processes
- Kubernetes: ***abstracts the cloud***, schedules microservices
- Kubernetes abstracts differences between private and public clouds
- Enable cloud-native applications to be cloud agnostic
 - i.e. they don't care *WHAT* cloud they run on
 - Enables fluid application migration between clouds
- Kubernetes provides rich set of tools/APIs to introspect (observe and examine) your apps

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.35

KUBERNETES - 4

- Features:
- A “control plane” – brain of the cluster
 - Implements autoscaling, rolling updates w/o downtime, self-healing
- A “bunch of nodes” – workers (muscle) of the cluster
- Provides orchestration
- The process of organizing everything into a useful application
- And also keeping it running smoothly

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.36

KUBERNETES - CLUSTER MANAGEMENT

- Master node(s) manage the cluster by:
 - Making scheduling decisions
 - Performing monitoring
 - Implementing changes
 - Responding to events
- *Masters implement the control plane of a Kubernetes cluster*
- Recipe for deploying to Kubernetes:
 - Write app as independent microservices in preferred language
 - Package each microservice in a container
 - Create a manifest to encapsulate the definition of a Pod
 - Deploy Pods to the cluster w/ a higher-level controller such as “Deployments” or “DaemonSets”

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.37

DECLARATIVE SERVICE APPROACH

- **Imperative definition**: sets of commands and operations
 - Example: BASH script, Dockerfile
- **Declarative definition**: specification of a service's properties
 - What level of service it should sustain, etc.
 - Example: Kubernetes YAML files
- Kubernetes manages resources **declaratively**
- How apps are deployed and run are defined with YAML files
- YAML files are POSTed to Kubernetes endpoints
- Kubernetes deploys and manages applications based on declarative service requirements
- If something isn't as it should be: *Kubernetes automatically tries to fix it*

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.38

KUBERNETES MASTERS

- Provide system services to host the control plane
- Simplest clusters use only 1 master – no replication
 - Suitable for lab and dev/test environments
- Production environments: masters are replicated ~3-5x
 - Provides fault tolerance and high availability (HA)
 - Cloud-based managed Kubernetes services offer HA deployments

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.39

MASTER SERVICES

- API Server
- Cluster store
- Controller Manager
- Scheduler
- Cloud controller

Kubernetes Cluster

```
graph TD; subgraph Master ["Kubernetes Master Server(s)"]; etcd; API_Server["API Server"]; Scheduler; CM["Controller Manager"]; end; LS1["Linux Server(s)"]; subgraph Node1 ["Kubernetes Node"]; D1["Docker"]; K1["Kubelet"]; KP1["Kubernetes Proxy"]; end; subgraph Node2 ["Kubernetes Node"]; D2["Docker"]; K2["Kubelet"]; KP2["Kubernetes Proxy"]; end; subgraph Node3 ["Kubernetes Node"]; D3["Docker"]; K3["Kubelet"]; KP3["Kubernetes Proxy"]; end; LS1 --> Node1; LS1 --> Node2; LS1 --> Node3;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.40

API SERVER

- Can run on 1-node for lab, test/dev environments
- Default port is 443
- Exposes a RESTful API where YAML configuration files are POST(ed) to
- YAML files (manifests) describe desired state of an application
 - Which container image(s) to use
 - Which ports to expose
 - How many POD replicas to run

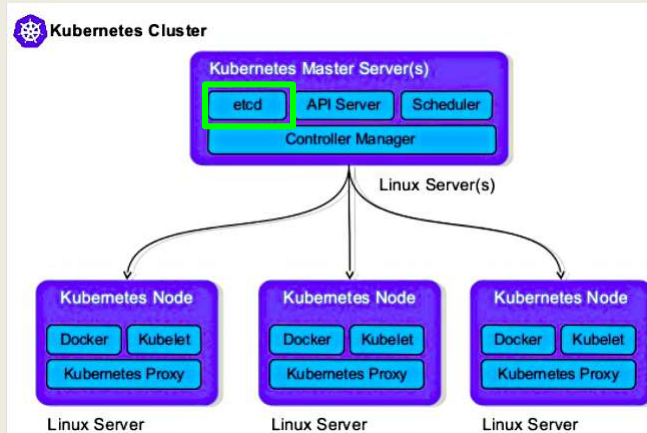
November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.41

MASTER SERVICES

- API Server
- Cluster store
- Controller Manager
- Scheduler
- Cloud controller



November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.42

CLUSTER STORE

- Used to persist Kubernetes cluster state
- Persistently stores entire configuration and state of the cluster
- Currently implemented with **etcd**
 - Popular distributed key/value store (db) supporting replication
 - HA deployments may use ~3-5 replicas
 - Is the authority on true state of the cluster
- etcd prefers consistency over availability
- etcd failure: apps continue to run, nothing can be reconfigured
- Consistency of writes is vital
- Employs RAFT consensus protocol to negotiate which replica has correct view of the system in the event of replica failure

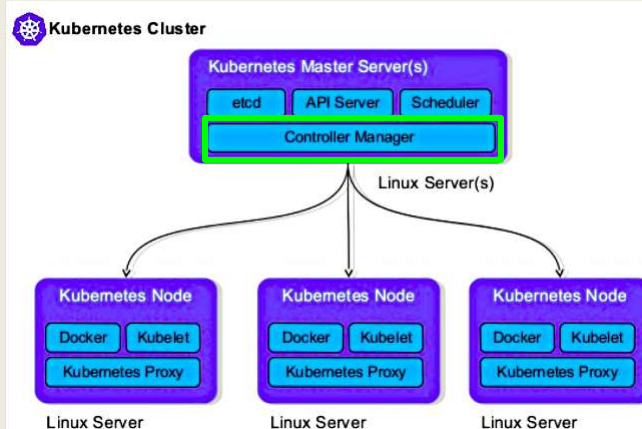
November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.43

MASTER SERVICES

- API Server
- Cluster store
- **Controller Manager**
- Scheduler
- Cloud controller



November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.44

CONTROLLER MANAGER

- Provides a “controller” of the controllers
 - Implements background control loops to monitor cluster and respond to events
 - Control loops include: node controller, endpoints controller, replicaset controller, etc...
- **GOAL: ensure cluster current state matches desired state**
- **Control Loop Logic:**
 1. Obtain desired state (defined in manifest YAMLs)
 2. Observe the current state
 3. Determine differences
 4. Reconcile differences
- Controllers are specialized to manage a specific resource type
 - They are not aware/concerned with of other parts of the system

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.45

MASTER SERVICES

- API Server
- Cluster store
- Controller Manager
- **Scheduler**
- Cloud controller

```
graph TD; subgraph Master ["Kubernetes Master Server(s)"]; etcd; API_Server[API Server]; Scheduler; CM["Controller Manager"]; end; subgraph Nodes ["Kubernetes Nodes"]; direction TB; N1["Kubernetes Node"]; N2["Kubernetes Node"]; N3["Kubernetes Node"]; end; N1 --- LS1[Linux Server]; N2 --- LS2[Linux Server]; N3 --- LS3[Linux Server]; Master --> N1; Master --> N2; Master --> N3;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.46

TASK SCHEDULER

- Scheduler's job is to identify the best node to run a task
 - Scheduler does not actually run tasks itself
- Assigns work tasks to appropriate healthy nodes
- Implements complex logic to filter out nodes incapable of running specified task(s)
- Capable nodes are ranked
- Node with highest ranking is selected to run the task

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.47

ENFORCING SCHEDULING PREDICATES

- Scheduler performs predicate (property) checks to verify how/where to run tasks
 - Is a node tainted?
 - Does task have affinity (deploy together), anti-affinity (separation) requirements?
 - Is a required network port available on the node?
 - Does node have sufficient free resources?
- Nodes incapable of running the task are eliminated as candidate hosts

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.48

RANKING NODES

- Remaining nodes are ranked based on for example:
 1. Does the node have the required images?
 - Cached images will lead to faster deployment time
 2. How much free capacity (CPU, memory) does the node have?
 3. How many tasks is the node already running?
- Each criterion is worth points
- **Node with most points is selected**
- If there is no suitable node, task is not scheduled, but marked as pending
- **PROBLEM:** *There is no one-sized fits all solution to selecting the best node. How weights are assigned to conditions may not reflect what is best for the task*

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.49

MASTER SERVICES

- API Server
- Cluster store
- Controller Manager
- Scheduler
- Cloud controller

```
graph TD; subgraph "Kubernetes Cluster"; subgraph "Kubernetes Master Server(s)"; etcd; API_Server[API Server]; Scheduler; Controller_Manager[Controller Manager]; end; subgraph "Linux Server(s)"; direction TB; Node1["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; Node2["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; Node3["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; end; end; "Kubernetes Master Server(s)" --> Node1; "Kubernetes Master Server(s)" --> Node2; "Kubernetes Master Server(s)" --> Node3;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.50

CLOUD CONTROLLER MANAGER

- Abstracts and manages integration with specific cloud(s)
- Manages vendor specific cloud infrastructure to provide instances (VMs), load balancing, storage, etc.
- Support for AWS, Azure, GCP, Digital Ocean, IBM, etc.

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.51

MASTER SERVICES

- API Server
- Cluster store
- Controller Manager
- Scheduler
- Cloud controller

```
graph TD; subgraph "Kubernetes Cluster"; subgraph "Kubernetes Master Server(s)"; etcd; API_Server[API Server]; Scheduler; Controller_Manager[Controller Manager]; end; subgraph "Linux Server(s)"; direction TB; Node1["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; Node2["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; Node3["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; end; end; "Kubernetes Master Server(s)" -- "Linux Server(s)" --> Node1; "Kubernetes Master Server(s)" -- "Linux Server(s)" --> Node2; "Kubernetes Master Server(s)" -- "Linux Server(s)" --> Node3;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.52

WORKER NODES

- Nodes perform tasks (i.e. host containers & services)
- Three primary functions:
 - Wait for the scheduler to assign work
 - Execute work (host containers, etc.)
 - Report back state information, etc.
- Nodes are considerably simpler than masters

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.53

WORKER NODES

- Kubelet
- Container runtime (*Docker, etc.*)
- Kubernetes Proxy

```
graph TD; subgraph Master ["Kubernetes Master Server(s)"]; etcd; API[API Server]; Scheduler; CM[Controller Manager]; end; subgraph Nodes ["Kubernetes Nodes"]; direction TB; N1["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; N2["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; N3["Kubernetes Node<br/>Docker, Kubelet, Kubernetes Proxy"]; end; Master -- "Linux Server(s)" --> N1; Master -- "Linux Server(s)" --> N2; Master -- "Linux Server(s)" --> N3;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.54

KUBELET

- Main Kubernetes agent
- Runs on every node
- Adding a new node installs the kubelet onto the node
- Kubelet registers the node with the cluster
- Monitors API server for new work assignments
- Maintains reporting back to control plane
- When a node can't run a task, kubelet is NOT responsible for finding an alternate node


November 23, 2020

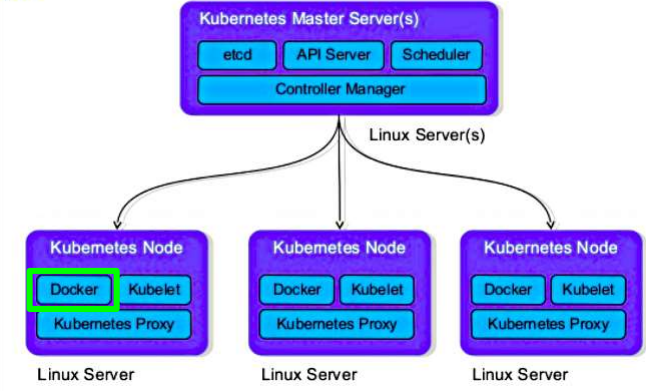
TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.55

WORKER NODES

- Kubelet
- Container runtime (*Docker, etc.*)
- Kubernetes Proxy

 **Kubernetes Cluster**



```
graph TD; Master["Kubernetes Master Server(s)  
etcd, API Server, Scheduler, Controller Manager"] -- "Linux Server(s)" --> Node1["Kubernetes Node  
Docker, Kubelet, Kubernetes Proxy"]; Master -- "Linux Server(s)" --> Node2["Kubernetes Node  
Docker, Kubelet, Kubernetes Proxy"]; Master -- "Linux Server(s)" --> Node3["Kubernetes Node  
Docker, Kubelet, Kubernetes Proxy"]; Node1 --- LS1[Linux Server]; Node2 --- LS2[Linux Server]; Node3 --- LS3[Linux Server];
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.56

CONTAINER RUNTIME(S)

- Each node requires a container runtime to run containers
- Early versions had custom support for a limited number of container types, e.g. Docker
- Kubernetes now provides a standard Container Runtime Interface (CRI)
- CRI exposes a clean interface for 3rd party container runtimes to plug-in to
- Popular container runtimes: Docker, containerd, Kata

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.57

WORKER NODES

- Kubelet
- Container runtime (*Docker, etc.*)
- **Kubernetes Proxy**

```
graph TD; subgraph "Kubernetes Cluster"; MS[Kubernetes Master Server(s)]; MS --- CS[etcd]; MS --- AS[API Server]; MS --- S[Scheduler]; MS --- CM[Controller Manager]; MS -- "Linux Server(s)" --> N1[Kubernetes Node]; MS -- "Linux Server(s)" --> N2[Kubernetes Node]; MS -- "Linux Server(s)" --> N3[Kubernetes Node]; subgraph "Linux Server"; N1 --- D1[Docker]; N1 --- K1[Kubelet]; N1 --- KP1[Kubernetes Proxy]; N2 --- D2[Docker]; N2 --- K2[Kubelet]; N2 --- KP2[Kubernetes Proxy]; N3 --- D3[Docker]; N3 --- K3[Kubelet]; N3 --- KP3[Kubernetes Proxy]; end; end;
```

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.58

KUBE-PROXY

- Runs on every node in the cluster
- Responsible for managing the cluster’s networking
- Ensures each node obtains a unique IP address
- Implemented local IPTABLES and IPVS rules to route and load-balance traffic
- IPTABLES (ipv4) – enables configuration of IP packet filtering rules of the Linux kernel firewall
- IPVS – IP Virtual Server: provides transport-layer (layer 4) load balancing as part of the Linux kernel; Configured using ipvsadm tool in Linux

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.59

CORE KUBERNETES COMPONENTS

- Kubernetes DNS
- Pods
- Services

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.60

KUBERNETES DNS

- Every Kubernetes cluster has an internal DNS service
- Accessed with a static IP
- Hard-coded so that every container can find it
- Every service is registered with the DNS so that all components can find every Service on the cluster by NAME
- Is based on CoreDNS (<https://coredns.io>)

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.61

CORE KUBERNETES COMPONENTS

- Kubernetes DNS
- Pods
- Services

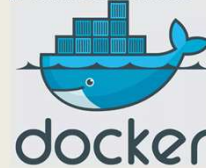
November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.62

PODS

- Pod – atomic unit of deployment & scheduling in Kubernetes
- A Kubernetes Pod is defined to run a containerized application
- Kubernetes manages Pods, not individual containers
- Cannot run a container directly on Kubernetes
- All containers run through Pods
- Pod comes from “pod of whales”
- Docker logo shows a whale with containers stacked on top
- Whale represents the Docker engine that runs on a single host
- Pods encapsulate the definition of a single microservice for hosting purposes
- Pods can have a single container, or multiple containers if the service requires more than one



November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.63

PODS - 2

- Examples of multi-container Pods:
 - Service meshes
 - Web containers with a helper container that pulls latest content
 - Containers with a tightly coupled log scraper or profiler
- YAML manifest files are used to provide a declarative description for how to run and manage a Pod
- To run a pod, POST a YAML to the API Server:
“kubectl run <NAME>” where NAME is the service
- A Pod runs on a single node (host)
- Pods share:
 - Interprocess communication (IPC) namespace
 - Memory, Volumes, Network stack

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.64

PODS - 3

- Pods provide a “fenced” environment to run containers
- Provide a “sandbox”
- Only tightly coupled containers are deployed with a single pod
- Best practice: decouple individual containers to separate pods
 - *What is the best container composition into pods? (1:1, 1:many)*
- Scaling
 - Pods are the unit of scaling
 - Add and remove pods to scale up/down
 - Do not add containers to a pod, add pod instances
 - Pod instances can be scheduled on the same or different host
- Atomic Operation
 - Pods are either fully up and running their service (i.e. port open/exposed), or pods are down / offline

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.65

PODS - 4

- Pod Lifecycle
 - An application should not be tightly bound or dependent on a specific Pod instance
 - Pods are designed to fail and be replaced
 - Use of **service objects** in Kubernetes help decouple pods to offer resiliency upon failure
- Deployments
 - Higher level controllers often used to deploy pods
 - Controllers implement a controller and watch loop:
 - “Deployments” – offer scalability & rolling updates
 - “DaemonSets” – run instance of service on every cluster node
 - “StatefulSets” – used for stateful components
 - “CronJobs” – for short lived tasks that need to run at specified times

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.66

CORE KUBERNETES COMPONENTS

- Kubernetes DNS
- Pods
- Services

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.67

KUBERNETES “SERVICES”

- Pods managed with “Deployments” or “DaemonSets” controllers are automatically replaced when they die
 - This provides resiliency for the application
- KEY IDEA: Pods are unreliable
- Services provide reliability by acting as a “GATEWAY” to pods that implement the services
 - They underlying pods can change over time
 - The services endpoints remain and are always available
- Service objects provide an abstraction layer w/ a reliable name and load balancing of requests to a set of pods

November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.68

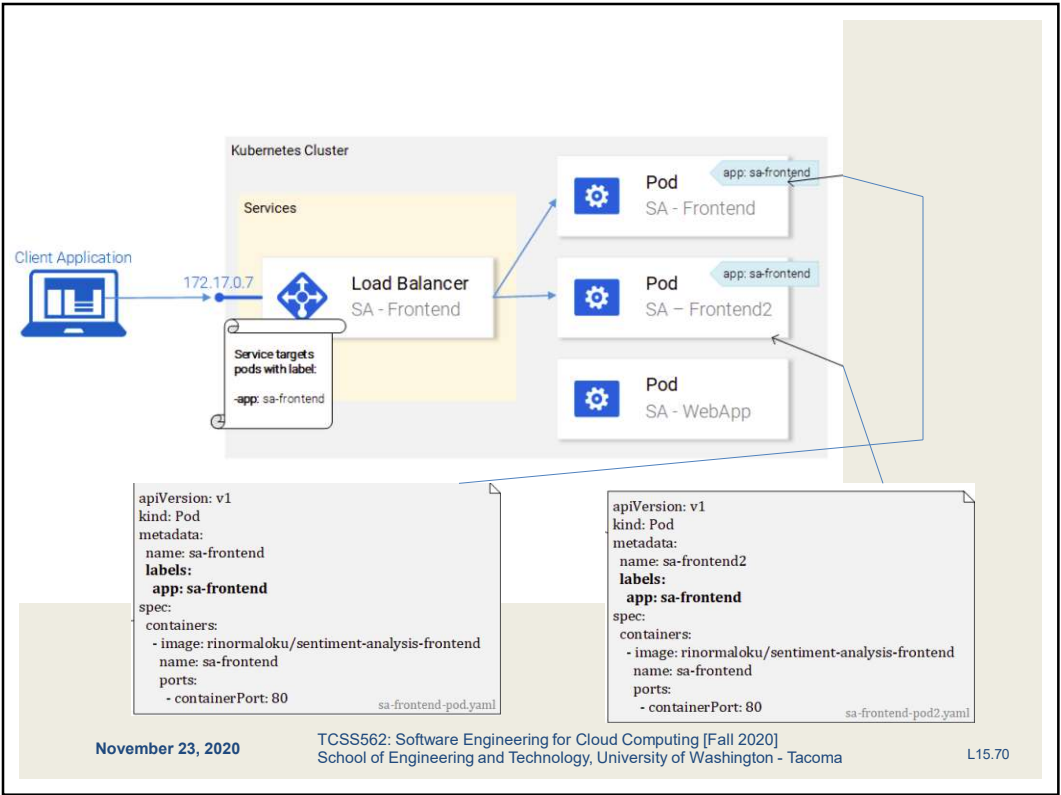
SERVICES

- Provide reliable front-end with:
 - Stable DNS name
 - IP Address
 - Port
- Services do not posses application intelligence
- No support for application-layer host and path routing
- Services have a “label selector” which is a set of lables
- Requests/traffic is only sent to Pods with matching labels
- Services only send traffic to healthy Pods
- **KEY IDEA:** Services bring stable IP addresses and DNS names to unstable Pods


November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.69



QUESTIONS



November 23, 2020

TCSS562: Software Engineering for Cloud Computing [Fall 2020]
School of Engineering and Technology, University of Washington - Tacoma

L15.71