TCSS 462/562: (SOFTWARE ENGINEERING FOR) CLOUD COMPUTING

Containerization II & Kubernetes

Wes J. Lloyd
School of Engineering and Technology
University of Washington – Tacoma

TR 5:50-7:50 PM

2

OFFICE HOURS – COMING UP

- **Monday 11/28 with Zening Zhao**
  - 12:30 to 1:30 pm - Zoom
- **Tuesday 11/29**
  - 3:30 to 5:30 pm - CP 229 and Zoom
- **Thursday 12/1**
  - 3:30 to 5:30 pm - CP 229 and Zoom
- **Or email for appointment**

**- Extra Office Hours ADDED: moving to 5/hrs/wk for remainder of quarter*
> Office Hours set based on Student Demographics survey feedback

3

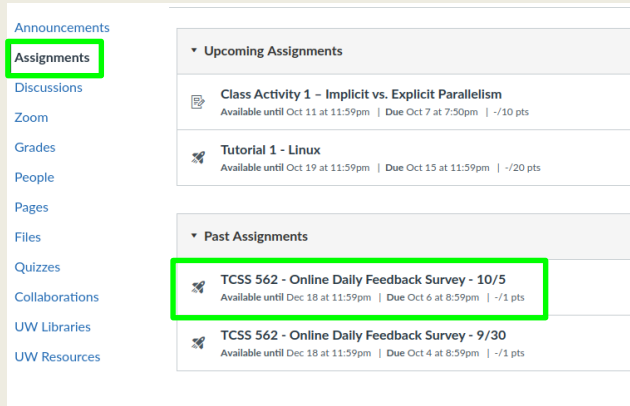
OBJECTIVES – 11/22

- **Questions from 11/17**
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- Kubernetes

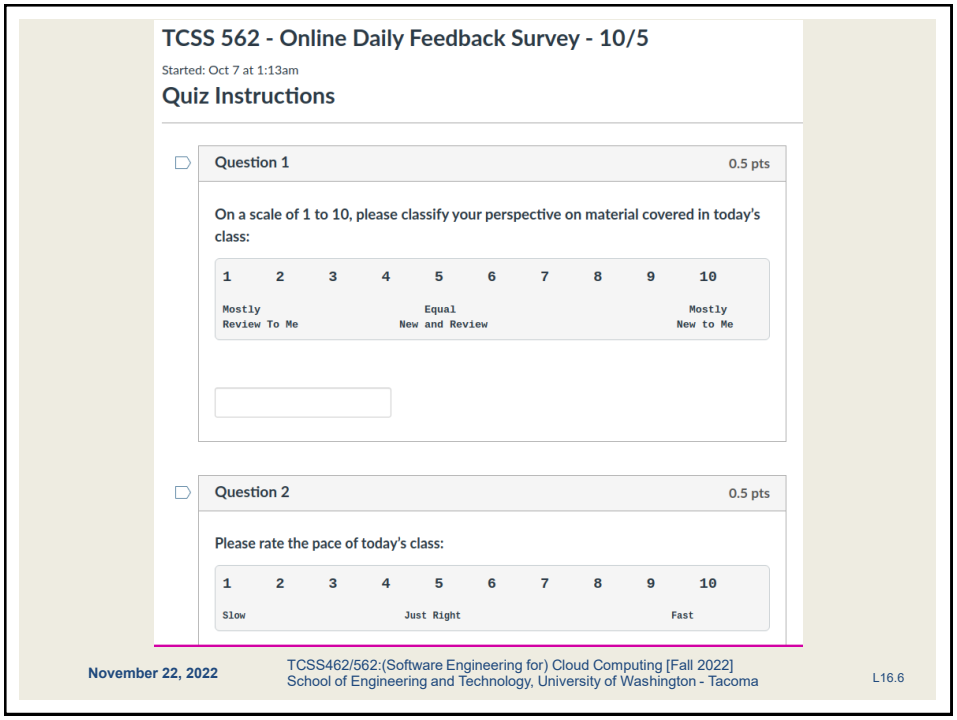
4

ONLINE DAILY FEEDBACK SURVEY

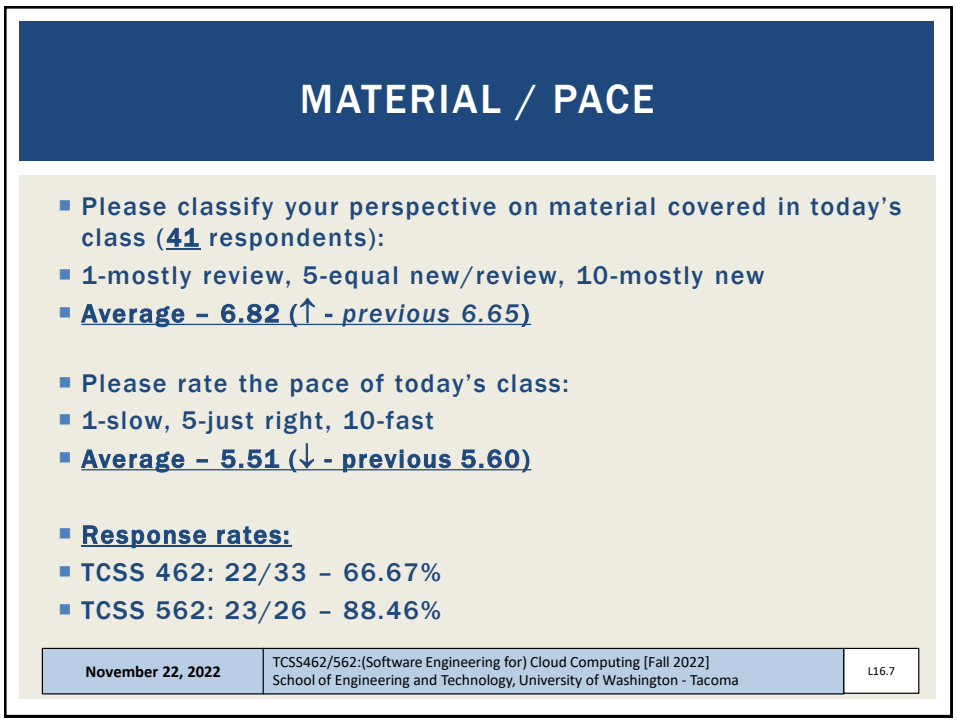
- Daily Feedback Quiz in Canvas – Take After Each Class
- Extra Credit for completing



5


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:
1 2 3 4 5 6 7 8 9 10
Mostly Review To Me
Equal New and Review
Mostly New to Me
Question 2
0.5 pts
Please rate the pace of today's class:
1 2 3 4 5 6 7 8 9 10
Slow
Just Right
Fast
November 22, 2022
TCSS462/562:(Software Engineering for) Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma
L16.6

6


MATERIAL / PACE
Please classify your perspective on material covered in today's class (41 respondents):
1-mostly review, 5-equal new/review, 10-mostly new
Average – 6.82 (↑ - previous 6.65)
Please rate the pace of today's class:
1-slow, 5-just right, 10-fast
Average – 5.51 (↓ - previous 5.60)
Response rates:
TCSS 462: 22/33 – 66.67%
TCSS 562: 23/26 – 88.46%
November 22, 2022
TCSS462/562:(Software Engineering for) Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma
L16.7

7

FEEDBACK FROM 11/17

- ***For tutorial 8, if I decided not participate in the study, what should I do during the class time on Dec. 1st***
- All students will complete tutorial 8 for credit, regardless if they opt-in to the study
- Opting-in means your data will be included in the assessment, and you will work to complete the tutorial during the scheduled session on December 1st
- If opting-out, a best effort attempt to complete tutorial 8 is still required to receive tutorial credit.
- If opting-out, completing the tutorial during the session on Dec 1st would help to facilitate asking questions

8

FEEDBACK - 2

- ***What are the differences between an OS container (other than the fact containers can be used to run many different OS services) and a VM?***
- The goal of operating system containers (i.e. LXC, V-Server) is to provide a lighter-weight alternative to VMs
- OS containers host an entire Linux system (all the standard processes, files, etc.)
- Differences from an actual VM are then →:
  - OS containers share the same Linux kernel (virtualized OSes with VMs each have a separate kernel)
  - OS containers must use the same version of kernel – no ability to run a mixture of versions like you can with VMs (though OS containers can host different Linux flavors because each has a distinct filesystems)
  - Containers sharing the same Linux kernel are all scheduled together – so there is less isolation. Greedy processes could steal resources from others resulting in interference and slower performance
  - OS containers will be Linux only. No Windows.

9

FEEDBACK - 3

- ***When is one better than the other - if we only need one OS service?***

10

CLOUD AND DISTRIBUTED SYSTEMS LAB

WES LLOYD, WLLOYD@UW.EDU,
HTTP://FACULTY.WASHINGTON.EDU/WLLOYD

- **Weekly Research Group Meetings**
- **Wednesdays at 3:30 pm (via Zoom)** no meeting this week

- **Looking for Winter 2023 and beyond:**
- **BSCSS students**
  - Independent Study (TCSS 499)
  - Honors Thesis

- **MSCSS students**
  - MS Thesis (TCSS 700)
  - MS Capstone (TCSS 702)
  - Independent Study (TCSS 600)
- Email wlloyd@uw.edu to follow-up and learn more

11

CLOUD AND DISTRIBUTED SYSTEMS LAB

WES LLOYD, WLLOYD@UW.EDU,
HTTP://FACULTY.WASHINGTON.EDU/WLLOYD

- **Serverless Computing (FaaS):**
- Service composition, performance and cost optimization/modeling /analytics, application migration, mitigation of platform limitations, vendor lock-in, observability/monitoring, influencing infrastructure, FaaS at the edge (IoT), fog, and cloud, resource federation, function/load balancing/scheduling, what are the best abstractions?, side channels, resource contention/heterogeneity, autonomic configuration/deployment, software tools

- **Containerization (Docker):**
- Containers, container orchestration frameworks, observability/ monitoring, resource allocation, checkpointing
- **Infrastructure-as-a-Service (IaaS) Cloud:**
- Application/workload deployment, performance and cost optimization/ modeling/analytics, infrastructure management, resource contention detection/mitigation, HW heterogeneity, observability/ monitoring, side channels to infer characteristics of the host & VM placement, virtualization overhead with increasing vCPU density

12

AWS CLOUD CREDITS

- IAM User Accounts Create – please let me know of any issues with these accounts
- If you did not provide your AWS account number on the AWS CLOUD CREDITS SURVEY to request AWS cloud credits and you would like credits this quarter, please contact the professor

13

TUTORIAL 5: When copying the CreateCSV class to create ProcessCSV it is imperative that the S3 PutObject call is DELETED !

In Tutorial 5 an EventBridge Rule is created to trigger ProcessCSV each time a PutObject event occurs on the bucket / file

If ProcessCSV generates PutObject(s) this results in a circular/endless call and will exhaust cloud credits quickly

AWS CREDITS → → → → → → → →

14

Please verify there are no unusual billing issues in your account every couple of days – Click on your name in the upper right hand corner of the AWS console

**Select 'Billing Dashboard'.
Check charges for services used in tutorials.**

**Tutorial 3: ec2; Tutorial 4: Lambda;
Tutorial 5: Simple Storage Service, Lambda, CloudWatch, CloudTrail; Tutorial 6: RDS, Lambda**

AWS CREDITS → → → → → → → →

15

Don't be sorry. Check your AWS bill and credits early and often

AWS CREDITS → → → → → → → →

16

OBJECTIVES – 11/22

- Questions from 11/17
- **Tutorials Questions**
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- Kubernetes

November 22, 2022	TCSS462/562:(Software Engineering for) Cloud Computing [Fall 2022] School of Engineering and Technology, University of Washington - Tacoma	L16.17

17

TUTORIAL 4 UPDATE

- Tutorial 4 (originally due Nov 6, extended to Nov 23)
- Please drop into office hours, or contact instructor via Canvas/email to resolve issues and complete Tutorial 4 !

18

TUTORIAL 5 & 6 UPDATE

- Tutorial 5 (originally due Nov 13, now extended to Nov 29)
- Tutorial 6 (originally due Nov 20, now extended to Dec 2)
- The final term project is due **Friday December 16 at 11:59pm**
- It is important to **complete** Tutorials 4, 5, and 6 ASAP to give time to segue to working on the term project
- Please note while assignment extensions are possible, a rushed term project can be obvious and fail to deliver many case study insights
- Tutorial 4 & 5 required for Tutorial 8 'hackathon' on Dec 1st

19

TUTORIAL 0

- Getting Started with AWS
- http://faculty.washington.edu/wlloyd/courses/tcss562/tutorials/TCSS462_562_f2022_tutorial_0.pdf
- Create an account
- Create account credentials for working with the CLI
- Install awsconfig package
- Setup awsconfig for working with the AWS CLI

20

TUTORIAL 4 – ~~NOV 6~~ NOV 23

- Introduction to AWS Lambda with the Serverless Application Analytics Framework (SAAF)
- https://faculty.washington.edu/wlloyd/courses/tcss562/tutorials/TCSS462_562_f2022_tutorial_4.pdf
- Obtaining a Java development environment
- Introduction to Maven build files for Java
- Create and Deploy "hello" Java AWS Lambda Function
  - Creation of API Gateway REST endpoint
- Sequential testing of "hello" AWS Lambda Function
  - API Gateway endpoint
  - AWS CLI Function invocation
- Observing SAAF profiling output
- Parallel testing of "hello" AWS Lambda Function with faas_runner
- Performance analysis using faas_runner reports
- Two function pipeline development task

21

IAM USERS – TUTORIAL 4

- Students completing tutorial 4 with an IAM user account may encounter permission issues
- Please contact the instructor if encountering any issues

22

TUTORIAL 4 - RESUBMISSION

- For tutorial 4 submissions, several submission indicate **Thread.sleep(10000)** was added but the results for the question 6 do not confirm this.
- It is possible that:

1. The provided results from the SAAF Report Generator were from a test run before the **Thread.Sleep()** statement was added to the code
   - OR -
2. The **Thread.Sleep()** statement was added in the incorrect location of the code
   - OR -
3. When opening the CSV output from the Report Generator, the file separator characters were set incorrectly.

- The only separator for a CSV file is the comma ",“
  Be sure to correctly open the CSV file in the spreadsheet.
  Columns can be offset resulting in the wrong answers being provided for Question 6.

23

TUTORIAL 4 – RESUBMISSION - 2

- The sleep statement must go between the START FUNCTION and END FUNCTION comments in the handleRequest() method specified as the AWS Lambda function's handler under runtime settings in the AWS Lambda GUI.

```
//****************START FUNCTION IMPLEMENTATION*************************
    try
    {
        Thread.sleep(10000);
    }
    catch (InterruptedException ie)
    {
        System.out.println("Interruption occurred while sleeping.");
    }
//****************END FUNCTION IMPLEMENTATIO N**************************
```


24

TUTORIAL 4 – RESUBMISSION - 3

- **SANITY CHECK:** consider that adding 10 seconds of sleep to your AWS Lambda function will cause the function to run for at least 10 seconds. This will impact the outputs requested for Question 6:

- **avg_runtime** is the server-side (cloud) runtime of the function
- This is the time it takes for the function to run on AWS Lambda (cloud)
- Adding sleep of 10 seconds should increase a function's **avg_runtime**

- **avg_roundTripTime** is the total time for a request from a client (laptop?) to travel to the server (cloud), make the function call, and return.
- If trying to make 50 calls at once on a laptop with a small # of CPU cores this time may be slow
- Adding sleep of 10 seconds should increase a function's **avg_roundTripTime**

25

TUTORIAL 4 – RESUBMISSION - 4

- **avg_cpuIdleDelta** time is the amount of time the Lambda function's Firecracker vCPUs are idle during the function call on the server measured in centiseconds:

  100 centiseconds = 1 second

  100 centiseconds = 1000 milliseconds

- By default, AWS Lambda functions with 512 MB run in a runtime environment with access to two vCPU cores
- This is the total vCPU idle time for both cores (it is doubled)
- Adding sleep of 10 seconds should increase your function's avg_cpuIdleDelta
- **How much should avg_cpuIdleDelta increase ?**

26

TUTORIAL 5 – ~~NOV 13~~ NOV 29

- Introduction to Lambda II: Working with Files in S3 and CloudWatch Events
- https://faculty.washington.edu/wlloyd/courses/tcss562/tutorials/TCSS462_562_f2022_tutorial_5.pdf
- Customize the Request object (add getters/setters)
  - Why do this instead of HashMap ?
- Import dependencies (jar files) into project for AWS S3
- Create an S3 Bucket
- Give your Lambda function(s) permission to work with S3
- Write to the CloudWatch logs
- Use of CloudTrail to generate S3 events
- Creating CloudWatch rule to capture events from CloudTrail
- Have the CloudWatch rule trigger a target Lambda function with a static JSON input object (hard-coded filename)
- **Optional**: for the S3 PutObject event, dynamically extract the name of the file put to the S3 bucket for processing

27

TUTORIAL 6 – ~~NOV 21~~ DEC 2

- Introduction to Lambda III: Serverless Databases
- https://faculty.washington.edu/wlloyd/courses/tcss562/tutorials/TCSS462_562_f2022_tutorial_6.pdf
- Create and use Sqlite databases using sqlite3 tool
- Deploy Lambda function with Sqlite3 database under /tmp
- Compare in-memory vs. file-based Sqlite DBs on Lambda
- Create an Amazon Aurora "Serverless" v2 MySQL database
- Using an ec2 instance in the same VPC (Region + availability zone) connect and interact with the database using the mysql CLI app
- Deploy an AWS Lambda function that uses the MySQL "serverless" database

28

OBJECTIVES – 11/22

- Questions from 11/17
- Tutorials Questions
- **Class Presentations Schedule - Cloud Technology / Research Paper Presentations**
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- Kubernetes

29

GROUP PRESENTATIONS

- TWO OPTIONS:
- ***Cloud technology presentation***
- ***Cloud research paper presentation***
  - Recent & suggested papers will be posted at: http://faculty.washington.edu/wlloyd/courses/tcss562/papers/
- **Presentation dates:**
  - Tuesday November 22, Tuesday November 29
  - Tuesday December 6, Thursday December 8
- **Peer Reviews**
  - Word DOCX form will be provided, fill out, submit PDF on Canvas
  - Feedback shared with groups
  - TCSS 462: 1 review/day required, additional are extra credit
  - TCSS 562: same as 462, but no peer review req'd on day of your talk

30

GROUP PRESENTATIONS

- 11 Presentation Teams
- 3 Cloud Technology Talks
- 8 Cloud Research Paper Presentations

- Thank you for the submissions

- Two students – we are not sure of your team and need clarification:
- Alan Liu (team 15 ???)
- Andrew Moreno-Escareno (team 8 ???)

31

PRESENTATION SCHEDULE

- **Tuesday November 22**

1. Bhagyashree Aras, Dhruvi Kaswala (team 3)

Research paper: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline

- **Tuesday November 29**

1. Divya Jacob, Nehaa Vuppala, Nandhini Dhanasekaran (team 10)

Research paper: Efficient GPU Sharing for Serverless Workflows

2. Jasleen Kaur, Naman Bhaia (team 4)

Research paper: A Serverless Publish/Subscribe System

3. Jeffrey Stockman, Rick Morrow, Mahmoud Ali Elkamhawy (team 1)

Research paper: Migrating from Microservices to Serverless: An IoT Platform Case Study

32

PRESENTATION SCHEDULE - 2

- **Tuesday December 6**

1. Yuan Huang, Yifan Xie (is Alan Liu in this group?) (team 15)

Research paper: A Prediction based Autoscaling in Serverless Computing

2. Angela Mu, Xiaojie Li, Ruigeng Zhang (team 6)

Research paper: Apollo: Modular and Distributed Runtime System for Serverless Function Compositions on Cloud, Edge, and IoT Resources

3. Jui Wang, Jinming Yu (team 7)

Cloud Technology: AWS Rekognition

- **Thursday December 8**

1. Mohammed Alshayeb (team 2)

Research paper (2021 list) Towards Federated Learning using FaaS Fabric

2. Nicole Guobadia (team 8)

Cloud Technology: AzureML (Machine Learning as a Service)

3. RamaSoumya Naraparaju, Sathwika Suddala, Chhavi Gupta (team 12)

Cloud Technology: Amazon Redshift

4. Yafei Li, Sue Yang (team 5)

Research paper: Cypress: Input size –Sensitive Container Provisioning and Request Scheduling for Serverless

33

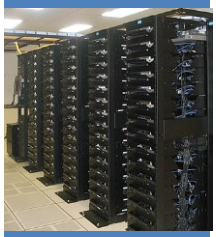
OBJECTIVES – 11/22

- Questions from 11/17
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- Kubernetes

34

TUTORIAL #7

DOCKER, CGROUPS, RESOURCE ISOLATION



35

TUTORIAL 7 – DEC 5

- Introduction to Docker
- https://faculty.washington.edu/wlloyd/courses/tcss562/tutorials/TCSS462_562_f2022_tutorial_7.pdf

- Must complete using Ubuntu 22.04 (for cgroups v2)
- Use docx file for copying and pasting Docker install commands
- Installing Docker
- Creating a container using a Dockerfile
- Using cgroups virtual filesystem to monitor CPU utilization of a container
- Persisting container images to Docker Hub image repository
- Container vertical scaling of CPU/memory resources
- Testing container CPU and memory isolation

36

TUTORIAL COVERAGE

- **Docker CLI → Docker Engine (dockerd) → containerd → runc**

- Working with the docker CLI:

- docker run — create a container
- docker ps -a — list containers, find CONTAINER ID
- docker exec --it — run a process in an existing container
- docker stop — stop a container
- docker kill — kill a container
- docker help — list available commands
- man docker — Docker Linux manual pages

37

```
Commands:
  attach      Attach local standard input, output, and error streams to a running container
  build       Build an image from a Dockerfile
  commit      Create a new image from a container's changes
  cp          Copy files/folders between a container and the local filesystem
  create      Create a new container
  deploy      Deploy a new stack or update an existing stack
  diff        Inspect changes to files or directories on a container's filesystem
  events      Get real time events from the server
  exec        Run a command in a running container
  export      Export a container's filesystem as a tar archive
  history     Show the history of an image
  images      List images
  import      Import the contents from a tarball to create a filesystem image
  info        Display system-wide information
  inspect     Return low-level information on Docker objects
  kill        Kill one or more running containers
  load        Load an image from a tar archive or STDIN
  login       Log in to a Docker registry
  logout      Log out from a Docker registry
  logs        Fetch the logs of a container
  pause       Pause all processes within one or more containers
  port        List port mappings or a specific mapping for the container
  ps          List containers
  pull        Pull an image or a repository from a registry
  push        Push an image or a repository to a registry
  rename      Rename a container
  restart     Restart one or more containers
  rm          Remove one or more containers
  rmi         Remove one or more images
  run         Run a command in a new container
  save        Save one or more images to a tar archive (streamed to STDOUT by default)
  search      Search the Docker Hub for images
  start       Start one or more stopped containers
  stats       Display a live stream of container(s) resource usage statistics
  stop        Stop one or more running containers
  tag         Create a tag TARGET_IMAGE that refers to SOURCE_IMAGE
  top         Display the running processes of a container
  unpause     Unpause all processes within one or more containers
  update      Update configuration of one or more containers
  version     Show the Docker version information
  wait        Block until one or more containers stop, then print their exit codes
```

Docker CLI

38

TUTORIAL 7

- Tutorial introduces use of two common Linux performance benchmark applications
- stress-ng
- 100s of CPU, memory, disk, network stress tests
- Sysbench
- Used in tutorial for memory stress test

39

OBJECTIVES – 11/22

- Questions from 11/17
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- **Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)**
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- Kubernetes

40

WE WILL RETURN AT ~7:00 PM

41

OBJECTIVES – 11/22

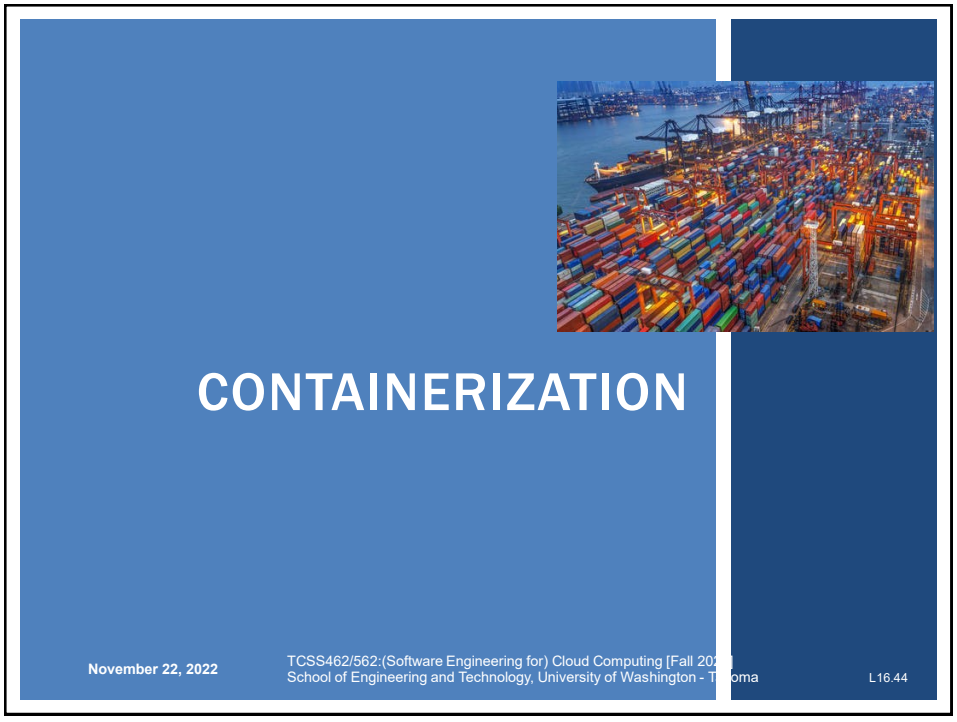
- Questions from 11/17
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- **Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala**
- Containerization
- Kubernetes

42

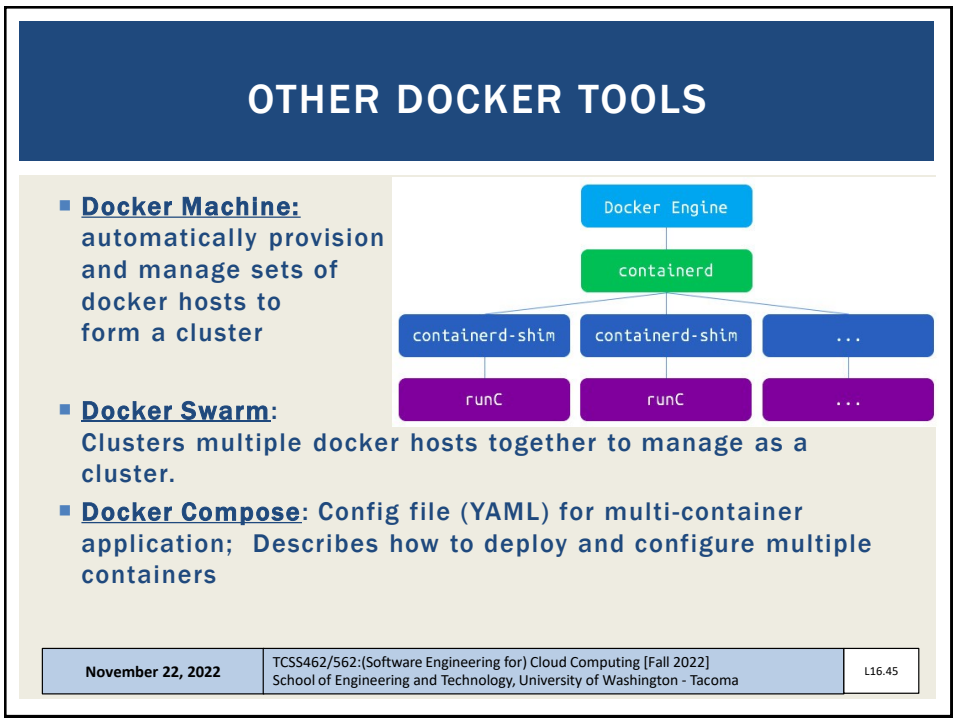
OBJECTIVES – 11/22

- Questions from 11/17
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- **Containerization**
- Kubernetes

43


CONTAINERIZATION
November 22, 2022
TCSS462/562:(Software Engineering for) Cloud Computing [Fall 202
School of Engineering and Technology, University of Washington - T oma
L16.44

44


OTHER DOCKER TOOLS
Docker Machine: automatically provision and manage sets of docker hosts to form a cluster
Docker Engine
containerd
containerd-shim
containerd-shim
...
runC
runC
...
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
November 22, 2022
TCSS462/562:(Software Engineering for) Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma
L16.45

45

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

46

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

47

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
  - Serverles containers without managing clusters
  - Azure Container Instances, AWS Fargate…

48

OBJECTIVES – 11/22

- Questions from 11/17
- Tutorials Questions
- Class Presentations Schedule - Cloud Technology / Research Paper Presentations
- Tutorial 7
- Tutorial 8: Addressing Serverless Computing Vendor Lock-In through Cloud Service Abstraction (UW Research Study)
- Cloud Research Paper Presentation: Towards a Serverless Bioinformatics Cyberinfrastructure Pipeline (Team 3) Bhagyashree Aras, Dhruvi Kaswala
- Containerization
- **Kubernetes**

49

KUBERNETES

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

L16.50

50

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 22, 2022	TCSS462/562:(Software Engineering for) Cloud Computing [Fall 2022] School of Engineering and Technology, University of Washington - Tacoma	L16.51

51

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

52

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

53

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

54

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

55

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

56

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

57

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"

58

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*

59

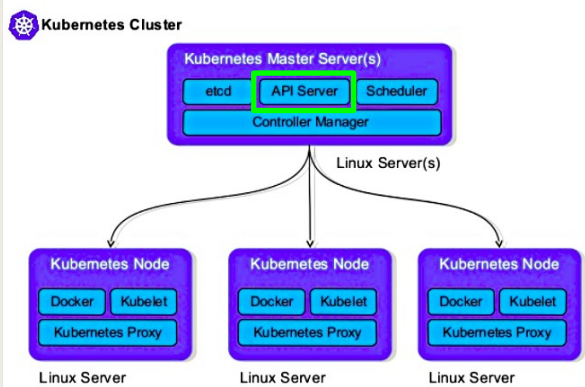
KUBERNETES MASTERS

- 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

60

MASTER SERVICES

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



61

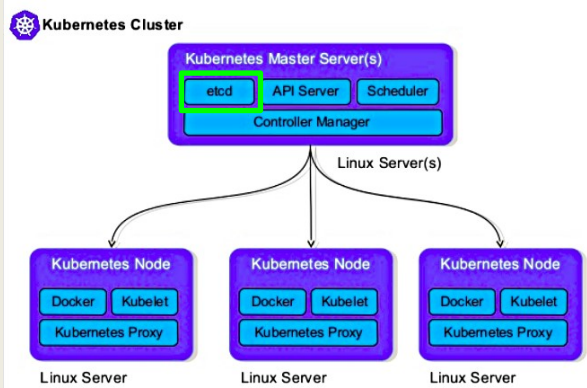
API SERVER

- 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

62

MASTER SERVICES

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



63

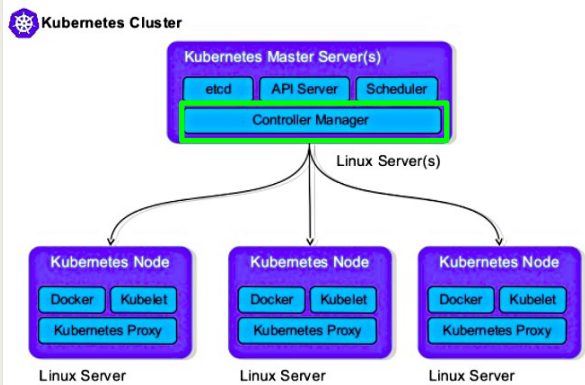
CLUSTER STORE

- 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

64

MASTER SERVICES

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



65

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:

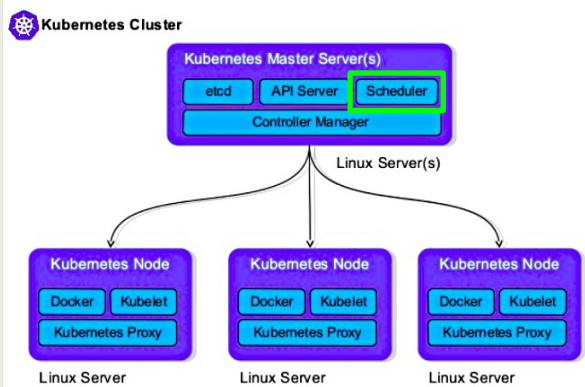
1. Obtain desired state (defined in manifest YAMLs)
2. Observe the current state
3. Determine differences
4. Reconcile differences

- Controllers are specialized to manage a specific resource type
  - They are not aware/concerned with of other parts of the system

66

MASTER SERVICES

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



67

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

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

RANKING NODES

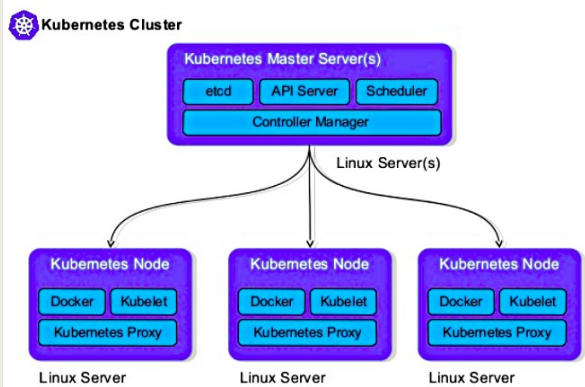
- Remaining nodes are ranked based on for example:
1. Does the node have the required images?
   - Cached images will lead to faster deployment time
2. How much free capacity (CPU, memory) does the node have?
3. How many tasks is the node already running?

- 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*

70

MASTER SERVICES

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



71

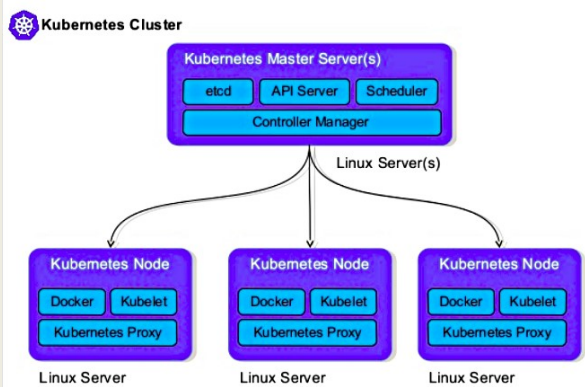
CLOUD CONTROLLER MANAGER

- 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.

72

MASTER SERVICES

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



73

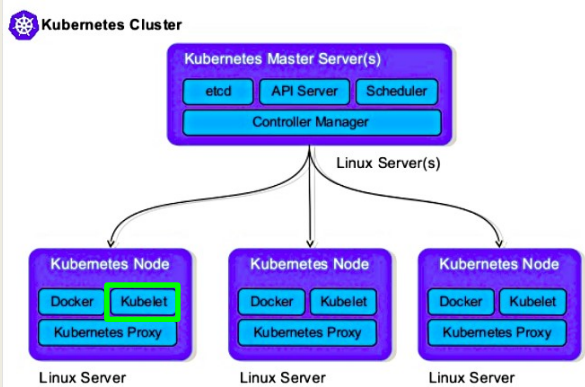
WORKER NODES

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

74

WORKER NODES

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



75

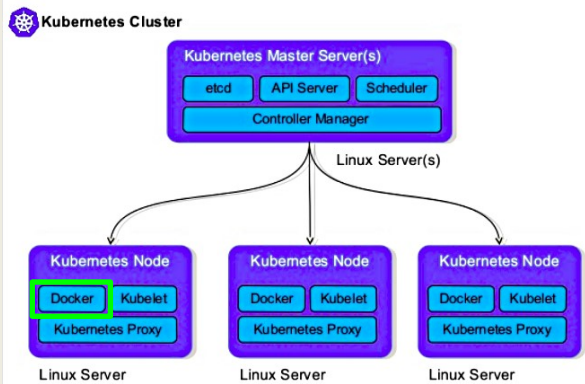
KUBELET

- 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

76

WORKER NODES

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



77

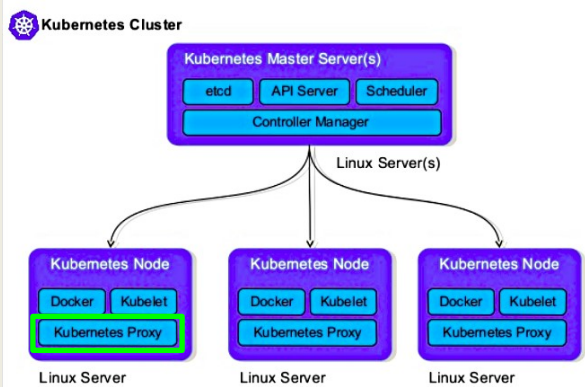
CONTAINER RUNTIME(S)

- 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

78

WORKER NODES

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



79

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

80

CORE KUBERNETES COMPONENTS

- Kubernetes DNS
- Pods
- Services

81

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)

82

CORE KUBERNETES COMPONENTS

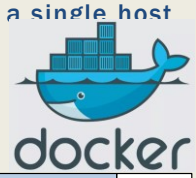
- Kubernetes DNS
- Pods
- Services

83

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



84

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

85

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

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

CORE KUBERNETES COMPONENTS

- Kubernetes DNS
- Pods
- Services

88

KUBERNETES "SERVICES"

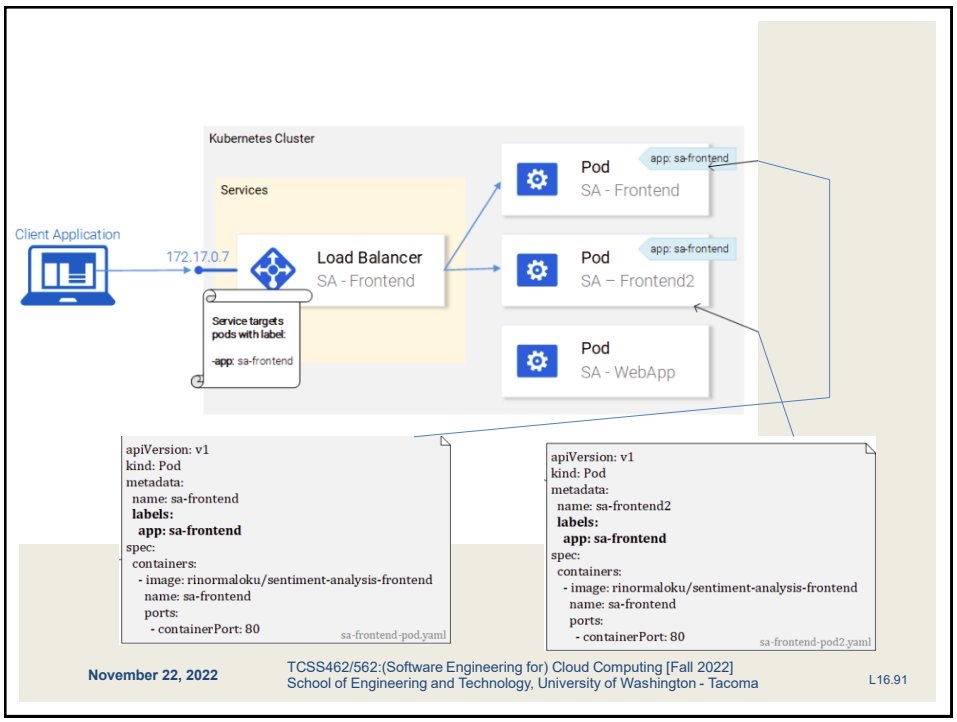
- Pods managed with "Deployments" or "DameonSets" controllers are automatically replaced when they die
  - This provides resiliency for the application
- **KEY IDEA**: Pods are unreliable
- **Services** provide reliability by acting as a "GATEWAY" to pods that implement the services
  - They underlying 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

89

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

90



91

QUESTIONS

November 22, 2022

TCSS462/562:(Software Engineering for) Cloud Computing [Fall 202
School of Engineering and Technology, University of Washington - T oma

L16.92

92