

TCSS 562:
SOFTWARE ENGINEERING
FOR CLOUD COMPUTING

Cloud Enabling
Technology II

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



1

OFFICE HOURS – FALL 2022

■ **THIS WEEK**

■ **Tuesday:**
■ 4:30 to 5:30 pm - CP 229 and Zoom

■ **Thursday***
■ 4:30 to 5:30 pm - CP 229 and Zoom

■ **Or email for appointment**
* Rescheduled due to Veteran's Day holiday - Nov 11th
> Office Hours set based on Student Demographics survey feedback

November 8, 2022

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

L12.2

2

OBJECTIVES – 11/8

■ **Questions from 11/3**

■ Tutorials Questions

■ Class Presentations:
Cloud Technology or Research Paper Review

■ Ch. 5: Cloud Enabling Technology

November 8, 2022

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

L12.3

3

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 13 at 11:59pm | Due Oct 7 at 7:59pm | -10 pts

Tutorial 1 - Linux
Available until Oct 19 at 11:59pm | Due Oct 13 at 11:59pm | -20 pts

Past Assignments

TCSS 562 - Online Daily Feedback Survey - 10/5
Available until Oct 18 at 11:59pm | Due Oct 6 at 8:59pm | -15 pts

TCSS 562 - Online Daily Feedback Survey - 9/30
Available until Oct 18 at 11:59pm | Due Oct 4 at 8:59pm | -15 pts

November 8, 2022

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

L12.4

4

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 8, 2022

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

L12.5

5

MATERIAL / PACE

■ Please classify your perspective on material covered in today's class (**46** respondents):

■ 1-mostly review, 5-equal new/review, 10-mostly new

■ **Average – 6.52 (↑ - previous 6.44)**

■ Please rate the pace of today's class:

■ 1-slow, 5-just right, 10-fast

■ **Average – 5.41 (↓ - previous 5.63)**

■ **Response rates:**

■ TCSS 462: 22/33 – 66.67%

■ TCSS 562: 24/26 – 92.31%

November 8, 2022

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

L12.6

6

FEEDBACK FROM 11/3

..

November 8, 2022

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

L12.7

7

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

October 11, 2022

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

L4.8

8

OBJECTIVES – 11/8

- Questions from 11/3
- **Tutorials Questions**
- Class Presentations:
Cloud Technology or Research Paper Review
- Ch. 5: Cloud Enabling Technology

November 8, 2022

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

L12.9

9

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

November 8, 2022

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

L12.10

10

TUTORIAL 4 – NOV 6

- 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

November 8, 2022

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

L12.11

11

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

November 8, 2022

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

L12.12

12

TUTORIAL 5 – NOV 13

- 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

November 8, 2022

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

L12.13

13

TUTORIAL 6 – NOV 21

- 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

November 8, 2022

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

L12.14

14

OBJECTIVES – 11/8

- Questions from 11/3
- Tutorials Questions
- Class Presentations:
Cloud Technology or Research Paper Review
- Ch. 5: Cloud Enabling Technology

November 8, 2022

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

L12.15

15

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:
TODAY: Wednesday November 16th @ 11:59pm
- Presentation dates:
 - Tuesday November 22, Tuesday November 29
 - Tuesday December 6, Thursday December 8

November 8, 2022

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

L12.16

16

OBJECTIVES – 11/8

- Questions from 11/3
- Tutorials Questions
- Class Presentations:
Cloud Technology or Research Paper Review
- Ch. 5: Cloud Enabling Technology

November 8, 2022

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

L12.17

17

CLOUD ENABLING TECHNOLOGY



18

CLOUD ENABLING TECHNOLOGY

- Broadband networks and internet architecture
- Data center technology
- Virtualization technology
- Multitenant technology
- Web/web services technology

November 8, 2022

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

L12.19

19

3. VIRTUALIZATION TECHNOLOGY

- Convert a physical IT resource into a virtual IT resource
- Servers, storage, network, power (virtual UPSs)
- Virtualization supports:
 - Hardware independence
 - Server consolidation
 - Resource replication
 - Resource pooling
 - Elastic scalability
- Virtual servers
 - Operating-system based virtualization
 - Hardware-based virtualization

November 8, 2022

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

L12.20

20

VIRTUAL MACHINES

- Emulation/simulation of a computer in software
- Provides a substitute for a real computer or server
- Virtualization platforms provide functionality to run an entire operating system
- Allows running multiple different operating systems, or operating systems with different versions simultaneously on the same computer

November 8, 2022

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

L12.21


21

KEY VIRTUALIZATION TRADEOFF

- Tradeoff space:

What is the “right” level of abstraction in the cloud for sharing resources with users?

Degree of Hardware Abstraction



Abstraction Concerns:

- Overhead
- Performance
- Isolation
- Security

November 8, 2022

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

L12.22

22

ABSTRACTION CONCERNS

- Overhead with too many instances w/ heavy abstractions
 - Too many instances using a heavy abstraction can lead to hidden resource utilization and waste
 - Example: Dedicated server with 48 VMs each with separate instance of Ubuntu Linux
 - Idle VMs can reduce performance of co-resident jobs/tasks
- “Virtualization” Overhead
 - Cost of virtualization an OS instance
 - Overhead has dropped from ~100% to ~1% over last decade
- Performance
 - Impacted by weight of abstraction and virtualization overhead

November 8, 2022

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

L12.23

23

ABSTRACTION CONCERNS - 2

- Isolation
 - From others:
What user A does should not impact user B in any noticeable way
- Security
 - User A and user B’s data should be always separate
 - User A’s actions are not perceivable by User B

November 8, 2022

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

L12.24

24

TYPES OF ABSTRACTION IN THE CLOUD

- **Virtual Machines** – original IaaS cloud abstraction
- **OS and Application Containers** – seen with CaaS
 - **OS Container** – replacement for VM, mimics full OS instance, heavier
 - OS containers run 100s of processes just like a VM
 - **App Container** – Docker: packages dependencies to easily transport and run an application anywhere
 - Application containers run only a few processes
- **Micro VMs** – FaaS / CaaS
 - Lighter weight alternative to full VM (KVM, XEN, VirtualBox)
 - Firecracker
- **Unikernel Operating Systems** – research mostly
 - Single process, multi-thread operating system
 - Designed for cloud, objective to reduce overhead of running too many OS instances

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.25

25

VIRTUAL MACHINES

- **Type 1 hypervisor**
 - Typically involves a special virtualization kernel that runs directly on the system to share the underlying machine with many guest VMs
 - Paravirtualization introduced to directly share system resources with guests bypassing full emulation
 - VM becomes equal participant in sharing the network card for example
- **Type 2 hypervisor**
 - Typically involves the **Full Virtualization** of the guest, where everything is simulated/emulated
- Hardware level support (i.e. features introduced on CPUs) have made virtualization faster in all respects shrinking virtualization overhead

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.26

26

TYPE 1 HYPERVISOR

```
graph TD
    VM1[VM guest operating system and application software]
    VM2[VM guest operating system and application software]
    VM3[VM guest operating system and application software]
    VMMH[Virtual Machine Management Hypervisor]
    HW[Hardware virtualization host]
    VM1 --- VMMH
    VM2 --- VMMH
    VM3 --- VMMH
    VMMH --- HW
```

- Host OS and VMs run atop the hypervisor
- The boot OS is the hypervisor kernel
- Xen dom0

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.27

27

TYPE 1 HYPERVISOR

- Acts as a control program
- Miniature OS kernel that manages VMs
- Boots and runs on bare metal
- Also known as Virtual Machine Monitor (VMM)
- **Paravirtualization**: Kernel includes I/O drivers
- VM guest OSes must use special kernel to interoperate
- Paravirtualization provides hooks to the guest VMs
- Kernel traps instructions (i.e. device I/O) to implement sharing & multiplexing
- User mode instructions run directly on the CPU
- **Objective: minimize virtualization overhead**
- Classic example is XEN (dom0 kernel)

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.28

28

COMMON VMMS: PARAVIRTUALIZATION

- **TYPE 1 Hypervisor**
- XEN
- Citrix Xen-server (a commercial version of XEN)
- VMWare ESXi
- KVM (virtualization support in kernel)
- Paravirtual I/O drivers introduced
 - XEN
 - KVM
 - Virtualbox

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.29

29

XEN

- Developed at Cambridge in ~ 2003

```
graph TD
    subgraph Guest_VMs
        US1[User Software]
        US2[User Software]
        US3[User Software]
        US4[User Software]
    end
    subgraph Host_OS
        GOS1[GuestOS XenoLinux]
        GOS2[GuestOS XenoLinux]
        GOS3[GuestOS XenoBSD]
        GOS4[GuestOS XenoXP]
        XAD1[Xeno-Aware Device Drivers]
        XAD2[Xeno-Aware Device Drivers]
        XAD3[Xeno-Aware Device Drivers]
        XAD4[Xeno-Aware Device Drivers]
    end
    subgraph XEN_kernel
        DCI[Domain0 control interface]
        VCPU[virtual x86 CPU]
        VPM[virtual phy mem]
        VNET[virtual network]
        VBDEV[virtual blockdev]
    end
    subgraph Physical_Machine
        HWP[H/W SMP x86, phy mem, enet, SCSI/IDE]
    end
    Guest_VMs --> Host_OS
    Host_OS --> XEN_kernel
    XEN_kernel --> Physical_Machine
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.30

30

XEN - 2

- VMs managed as “domains”
- Domain 0 is the hypervisor domain
 - Host OS is installed to run on bare-metal, but doesn't directly facilitate virtualization (unlike KVM)
- Domains 1..n are guests (VMs) – not bare-metal

```
xeninfo - 17:53:48 Xen 3.1.2-390.el5
3 domains: 1 running, 2 blocked, 0 paused, 0 crashed, 0 dying, 0 shutdown
Mem: 8379564k total, 8377876k used, 1688k free  CPUs: 4 @ 2400MHz

NAME  STATE  CPU(sec)  CPU(%)  MEM(k)  MEM(%)  MAXMEM(k)  MAXMEM(%)  VCPUS
-----
centos-1  --b---  885      1       6313    0.0     1064960     12.7        1
centos-2  --b---  17       0       1056640  12.6    2113536     25.2        1
Domain-0  -----r 2979     19.3    6568960  78.4    no limit    n/a         4
4 1057374 290072  0       0       0       0       0
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.31

31

XEN - 3

- Physical machine boots special XEN kernel
- Kernel provides paravirtual API to manage CPU & device multiplexing
- Guests require modified XEN-aware kernels
- Xen supports full-virtualization for unmodified OS guests in hvm mode
- Amazon EC2 largely based on modified version of XEN hypervisor (EC2 gens 1-4)
- XEN provides its own CPU schedulers, I/O scheduling

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.32

32

TYPE 2 HYPERVISOR

- Adds additional layer

```
graph TD
    VM1[VM  
(guest operating  
system and  
application  
software)]
    VM2[VM  
(guest operating  
system and  
application  
software)]
    VM3[VM  
(guest operating  
system and  
application  
software)]
    VMM[Virtual Machine Management]
    OS[Operating System  
(host OS)]
    HW[Hardware  
(virtualization host)]

    VM1 --- VMM
    VM2 --- VMM
    VM3 --- VMM
    VMM --- OS
    OS --- HW
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.33

33

TYPE 2 HYPERVISOR

- Problem: Original x86 CPUs could not trap special instructions
- Instructions not specially marked
- Solution: Use Full Virtualization
- Trap ALL instructions
- “Fully” simulate entire computer
- Tradeoff: Higher Overhead
- Benefit: Can virtualize any operating system without modification

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.34

34

CHECK FOR VIRTUALIZATION SUPPORT

- See:
<https://cyberciti.biz/faq/linux-xen-vmware-kvm-intel-vt-amd-v-support>
- # check for Intel VT CPU virtualization extensions on Linux
`grep -color vmx /proc/cpuinfo`
- # check for AMD V CPU virtualization extensions on Linux
`grep -color svm /proc/cpuinfo`
- Also see `lscpu` → “Virtualization:”
- Other Intel CPU features that help virtualization:
`ept vpid tpr_shadow flexpriority vnmi`

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.35

35

KERNEL BASED VIRTUAL MACHINES (KVM)

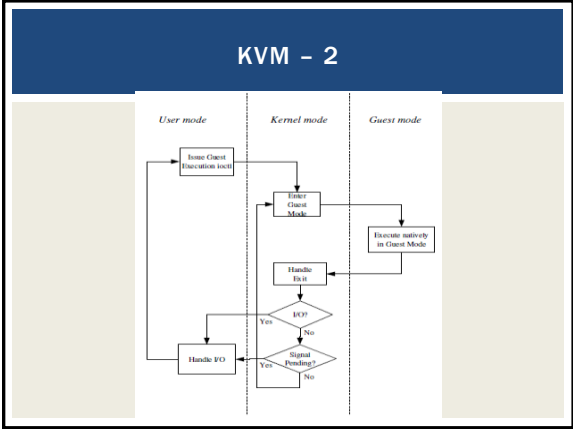
- x86 HW notoriously difficult to virtualize
- Extensions added to 64-bit Intel/AMD CPUs
 - Provides hardware assisted virtualization
 - New “guest” operating mode
 - Hardware state switch
 - Exit reason reporting
 - Intel/AMD implementations different
 - Linux uses vendor specific kernel modules

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.36

36



37

KVM - 3

- KVM has `/dev/kvm` device file node
 - Linux character device, with operations:
 - Create new VM
 - Allocate memory to VM
 - Read/write virtual CPU registers
 - Inject interrupts into vCPUs
 - Running vCPUs
- VMs run as Linux processes
 - Scheduled by host Linux OS
 - Can be pinned to specific cores with “taskset”

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.38

38

KVM PARAVIRTUALIZED I/O

- KVM - Virtio
 - Custom Linux based paravirtual device drivers
 - Supersedes QEMU hardware emulation (full virt.)
 - Based on XEN paravirtualized I/O
 - Custom block device driver provides paravirtual device emulation
 - Virtual bus (memory ring buffer)
 - Requires hypercall facility
 - Direct access to memory

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.39

39

KVM DIFFERENCES FROM XEN

- KVM requires CPU VMX support
 - Virtualization management extensions
- KVM can virtualize any OS without special kernels
 - Less invasive
- KVM was originally separate from the Linux kernel, but then integrated
- KVM is type 1 hypervisor because the machine boots Linux which has integrated support for virtualization
- Different than XEN because XEN kernel alone is not a full-fledged OS

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.40

40

KVM ENHANCEMENTS

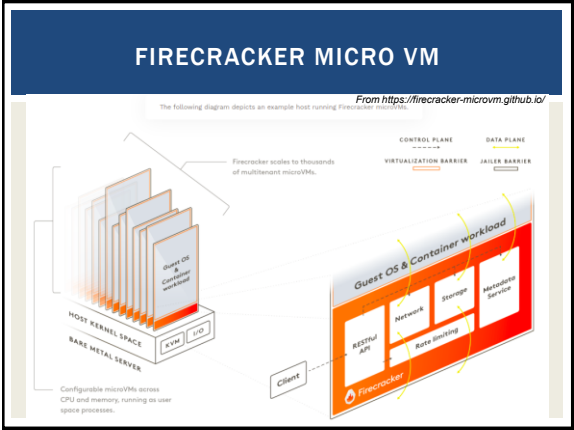
- Paravirtualized device drivers
 - Virtio
- Guest Symmetric Multiprocessor (SMP) support
 - Leverages multiple on-board CPUs
 - Supported as of Linux 2.6.23
- VM Live Migration
- Linux scheduler integration
 - Optimize scheduler with knowledge that KVM processes are virtual machines

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.41

41



42

FIRECRACKER MICRO VM

- Provides a virtual machine monitor (VMM) (i.e. hypervisor) using KVM to create and manage microVMs
- Has a minimalist design with goals to improve security, decreases the startup time, and increases hardware utilization
- Excludes unnecessary devices and guest functionality to reduce memory footprint and attack surface area of each microVM
- Supports boot time of <125ms, <5 MiB memory footprint
- Can run 100s of microVMs on a host, launching up to 150/sec
- Is available on 64-bit Intel, AMD, and Arm CPUs
- Used to host AWS Lambda and AWS Fargate
- Has been open sourced under the Apache 2.0 license

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.43

43

FIRECRACKER - 2

- **Minimalistic**
- MicroVMs run as separate processes on the host
- Only 5 emulated devices are available: virtio-net, virtio-block, virtio-sock, serial console, and a minimal keyboard controller used only to stop the microVM
- Rate limiters can be created and configured to provision resources to support bursts or specific bandwidth/operation limitations
- **Configuration**
- A RESTful API enables common actions such as configuring the number of vCPUs or launching microVMs
- A metadata service between the host and guest provides configuration information

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.44

44

FIRECRACKER - 2

- **Security**
- Runs in user space (**not the root user**) on top of the Linux Kernel-based Virtual Machine (KVM) hypervisor to create microVMs
- Lambda functions, Fargate containers, or container groups can be encapsulated using Firecracker through KVM, enabling workloads from different customers to run on the same machine, without sacrificing security or efficiency
- MicroVMs are further isolated with common Linux user-space security barriers using a companion program called "jailer" which provides a second line of defense if KVM is compromised

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.45

45

UNIKERNELS

- Lightweight alternative to containers and VMs
- Custom Cloud Operating System
- Single process, multiple threads, runs one program
- Launch separately atop of hypervisor (XEN/KVM)
- Reduce overhead, duplication of heavy weight OS
- OSv is most well known unikernel
- Several others exist has research projects
- More information at: <http://unikernel.org/>
- Google Trends OSv →



November 8, 2022

TCSS462/562
School of Eng

46

WE WILL RETURN AT
~7:00 PM



47

VIRTUALIZATION MANAGEMENT

- Virtual infrastructure management (VIM) tools
- Tools that manage pools of virtual machines, resources, etc.
- Private cloud software systems can be considered as a VIM
- Considerations:
- Performance overhead
 - Paravirtualization: custom OS kernels, I/O passed directly to HW w/ special drivers
- Hardware compatibility for virtualization
- Portability: virtual resources tend to be difficult to migrate cross-clouds

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.48

48

VIRTUAL INFRASTRUCTURE MANAGEMENT (VIM)

- Middleware to manage virtual machines and infrastructure of IaaS “clouds”
- Examples
 - OpenNebula
 - Nimbus
 - Eucalyptus
 - OpenStack

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.49

49

VIM FEATURES

- Create/destroy VM instances
- Image repository
 - Create/Destroy/Update images
 - Image persistence
- Contextualization of VMs
 - Networking address assignment
 - DHCP / Static IPs
 - Manage SSH keys

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.50

50

VIM FEATURES - 2

- Virtual network configuration/management
 - Public/Private IP address assignment
 - Virtual firewall management
 - Configure/support isolated VLANs (private clusters)
- Support common virtual machine managers (VMMs)
 - XEN, KVM, VMware
 - Support via libvirt library

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.51

51

VIM FEATURES - 3

- Shared “Elastic” block storage
 - Facility to create/update/delete VM disk volumes
 - Amazon EBS
 - Eucalyptus SC
 - OpenStack Volume Controller

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.52

52

CONTAINER ORCHESTRATION FRAMEWORKS

- Middleware to manage Docker application container deployments across virtual clusters of Docker hosts (VMs)
- Considered Infrastructure-as-a-Service
- Open source
 - Kubernetes framework
 - Docker swarm
 - Apache Mesos/Marathon
- Proprietary
 - Amazon Elastic Container Service

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.53

53

CONTAINER SERVICES

- Public cloud container cluster services
 - Azure Kubernetes Service (AKS)
 - Amazon Elastic Container Service for Kubernetes (EKS)
 - Google Kubernetes Engine (GKE)
- Container-as-a-Service
 - Azure Container Instances (ACI – April 2018)
 - AWS Fargate (November 2017)
 - Google Kubernetes Engine Serverless Add-on (alpha-July 2018)

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.54

54

CLOUD ENABLING TECHNOLOGY

- Adapted from Ch. 5 from Cloud Computing Concepts, Technology & Architecture
- Broadband networks and internet architecture
- Data center technology
- Virtualization technology
- Multitenant technology
- Web/web services technology

November 8, 2022


TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.55

55

4. MULTITENANT APPLICATIONS

- Each tenant (like in an apartment) has their own view of the application
- Tenants are unaware of their neighbors
- Tenants can only access their data, no access to data and configuration that is not their own
- Customizable features
 - UI, business process, data model, access control
- Application architecture
 - User isolation, data security, recovery/backup by tenant, scalability for a tenant, for tenants, metered usage, data tier isolation



November 8, 2022

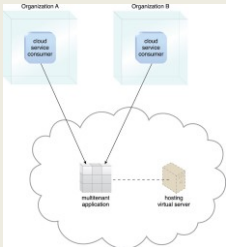
TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.56

56

MULTITENANT APPS - 2

- Forms the basis for SaaS (applications)



November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.57

57

CLOUD ENABLING TECHNOLOGY

- Adapted from Ch. 5 from Cloud Computing Concepts, Technology & Architecture
- Broadband networks and internet architecture
- Data center technology
- Virtualization technology
- Multitenant technology
- Web/web services technology

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.58

58

5. WEB SERVICES/WEB

- Web services technology is a key foundation of cloud computing's "as-a-service" cloud delivery model
- SOAP – "Simple" object access protocol
 - First generation web services
 - WSDL – web services description language
 - UDDI – universal description discovery and integration
 - SOAP services have their own unique interfaces
- REST – instead of defining a custom technical interface REST services are built on the use of HTTP protocol
- HTTP GET, PUT, POST, DELETE

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.59

59

HYPertext TRANSPORT PROTOCOL (HTTP)

- An ASCII-based request/reply protocol for transferring information on the web
- HTTP request includes:
 - request method (GET, POST, etc.)
 - Uniform Resource Identifier (URI)
 - HTTP protocol version understood by the client
 - headers—extra info regarding transfer request
- HTTP response from server
 - Protocol version & status code →
 - Response headers
 - Response body

HTTP status codes:
2xx — all is well
3xx — resource moved
4xx — access problem
5xx — server error

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.60

60

REST: REPRESENTATIONAL STATE TRANSFER

- Web services protocol
- *Supersedes SOAP* – Simple Object Access Protocol
- Access and manipulate web resources with a predefined set of stateless operations (known as web services)
- Requests are made to a URI
- Responses are most often in JSON, but can also be HTML, ASCII text, XML, no real limits as long as text-based
- HTTP verbs: GET, POST, PUT, DELETE, ...

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.61

61

```
// SOAP REQUEST

POST /InStock HTTP/1.1
Host: www.bookshop.org
Content-Type: application/soap+xml; charset=utf-8
Content-Length: nnn

<?xml version="1.0"?>
<soap:Envelope
  xmlns:soap="http://www.w3.org/2001/12/soap-envelope"
  soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">
  <soap:Body xmlns:m="http://www.bookshop.org/prices">
    <m:GetBookPrice>
      <m:BookName>The Fleamarket</m:BookName>
    </m:GetBookPrice>
  </soap:Body>
</soap:Envelope>
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.62

62

```
// SOAP RESPONSE
POST /InStock HTTP/1.1
Host: www.bookshop.org
Content-Type: application/soap+xml; charset=utf-8
Content-Length: nnn

<?xml version="1.0"?>
<soap:Envelope
  xmlns:soap="http://www.w3.org/2001/12/soap-envelope"
  soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">
  <soap:Body xmlns:m="http://www.bookshop.org/prices">
    <m:GetBookPriceResponse>
      <m:Price>10.95</m:Price>
    </m:GetBookPriceResponse>
  </soap:Body>
</soap:Envelope>
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.63

63

```
// WSDL Service Definition
<?xml version="1.0" encoding="UTF-8"?>
<definitions name="DayOfWeek"
  targetNamespace="http://www.copswave.com/soapworks/examples/DayOfWeek.wsdl"
  xmlns:tns="http://www.copswave.com/soapworks/examples/DayOfWeek.wsdl"
  xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:tns1="http://schemas.xmlsoap.org/wsdl/">
  <message name="DayOfWeekInput">
    <part name="date" type="xsd:date"/>
  </message>
  <message name="DayOfWeekResponse">
    <part name="DayOfWeek" type="xsd:string"/>
  </message>
  <portType name="DayOfWeekPortType">
    <operation name="GetDayOfWeek">
      <input message="tns:DayOfWeekInput"/>
      <output message="tns:DayOfWeekResponse"/>
    </operation>
  </portType>
  <binding name="DayOfWeekBinding" type="tns:DayOfWeekPortType">
    <soap:binding style="document"
      transport="http://schemas.xmlsoap.org/soap/http"/>
    <operation name="GetDayOfWeek">
      <soap:operation soapAction="getDayOfWeek"/>
    </operation>
    <input>
      <soap:body use="encoded"
        namespace="http://www.copswave.com/soapworks/examples"
        encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" />
    </input>
    <output>
      <soap:body use="encoded"
        namespace="http://www.copswave.com/soapworks/examples"
        encodingStyle="http://schemas.xmlsoap.org/soap/encoding/" />
    </output>
  </binding>
  <service name="DayOfWeekService">
    <documentation>
      Returns the day-of-week name for a given date
    </documentation>
    <port name="DayOfWeekPort" binding="tns:DayOfWeekBinding">
      <soap:address location="http://localhost:8990/dayOfWeek/DayOfWeek" />
    </port>
  </service>
</definitions>
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]
School of Engineering and Technology, University of Washington - Tacoma

L12.64

64

REST CLIMATE SERVICES EXAMPLE

- **USDA**
Lat/Long
Climate
Service
Demo
 - Just provide a Lat/Long
- ```
// REST/JSON
// Request climate data for Washington
{
 "parameter": [
 {
 "name": "latitude",
 "value": 47.2529
 },
 {
 "name": "longitude",
 "value": -122.4443
 }
]
}
```

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]  
School of Engineering and Technology, University of Washington - Tacoma

L12.65

65

REST - 2

- App manipulates one or more types of resources.
- Everything the app does can be characterized as some kind of operation on one or more resources.
- Frequently services are CRUD operations (create/read/update/delete)
  - Create a new resource
  - Read resource(s) matching criterion
  - Update data associated with some resource
  - Destroy a particular a resource
- Resources are often implemented as objects in OO languages

November 8, 2022

TCSS462/562: Software Engineering for Cloud Computing [Fall 2022]  
School of Engineering and Technology, University of Washington - Tacoma

L12.66

66

REST ARCHITECTURAL ADVANTAGES

- **Performance:** component interactions can be the dominant factor in user-perceived performance and network efficiency
- **Scalability:** to support large numbers of services and interactions among them
- **Simplicity:** of the Uniform Interface
- **Modifiability:** of services to meet changing needs (even while the application is running)
- **Visibility:** of communication between services
- **Portability:** of services by redeployment
- **Reliability:** resists failure at the system level as redundancy of infrastructure is easy to ensure


November 8, 2022

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

L12.67

67

QUESTIONS



November 8, 2022

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

L12.68

68