

# TCSS 562: SOFTWARE ENGINEERING FOR CLOUD COMPUTING

## AWS Demo, Cloud Enabling Technology

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

MW 5:50-7:50 PM



## OBJECTIVES - 11/4

- **Questions from 11/2**
- Quiz 1
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
- Tutorial #5
- **From: Cloud Computing Concepts, Technology & Architecture:**  
Chapter 5 - Cloud Enabling Technology
- Tutorial questions
- Team planning

November 4, 2020

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

L11.2

# 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 4, 2020

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

L11.3

## 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 4, 2020

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

L11.4

MATERIAL / PACE

- Please classify your perspective on material covered in today's class (19 respondents):
  - 1-mostly review, 5-equal new/review, 10-mostly new
  - Average – 6.74 (↑ - *previous 6.60*)
- Please rate the pace of today's class:
  - 1-slow, 5-just right, 10-fast
  - Average – 5.58 (↑ - *previous 5.52*)

November 4, 2020

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

L11.5

FEEDBACK FROM 11/2

- I'm not very clear about the placement groups when creating instances. Can you please explain.
- EC2 Placement groups allow VMs to be launched into groups that apply a recommendation regarding where VMs are allocated
- From the EC2 launch wizard:

Placement group ⓘ

Placement group name ⓘ

Placement group strategy ⓘ

Capacity Reservation ⓘ

☒ Add instance to placement group

☐ Add to existing placement group.

☒ Add to a new placement group.

new\_group\_1

cluster

cluster

spread

partition

November 4, 2020

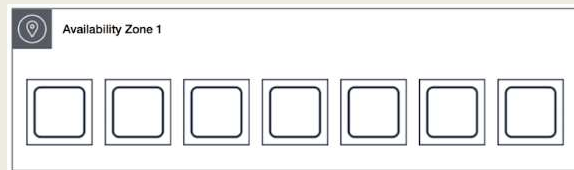
TCSS562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.6

WJL6

## AWS EC2 PLACEMENT GROUPS - SPREAD

- **Spread Placement:** Instances placed on distinct servers located on different server racks
  - Guarantees that each VM in the group DOES NOT co-reside with any other
  - Maximum of 7 VMs allowed per availability zone
  - In the US-EAST-2 Ohio region there are 3 zones, limiting spread placement to a maximum of 21 VMs



November 4, 2020

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

L11.7

WJL6

## AWS EC2 PLACEMENT GROUPS - CLUSTER

- **Cluster Placement:** Instances are packed together as closely as possible inside an Availability Zone
  - Cluster placement groups enjoy a higher per-flow throughput limit for TCP/IP traffic
  - Cluster placement groups are recommended for applications that benefit from low network latency, high network throughput, or both



November 4, 2020

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

L11.8

## Slide 7

---

**WJL6** May consider cutting out partition placement since we don't use it in any experiments

Wes J. Lloyd, 7/22/2020

## Slide 8

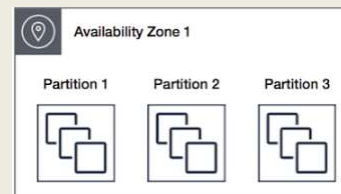
---

**WJL6** May consider cutting out partition placement since we don't use it in any experiments

Wes J. Lloyd, 7/22/2020

## AWS EC2 PLACEMENT GROUPS - PARTITION

- **Partition Placement:** a group with partitions where each partition has its own set of racks with a distinct network and power source
- No two partitions in a group share the same racks, allowing isolation from hardware failure
- Partitions can guarantee that “components” of multi-tier applications DO NOT SHARE hardware by using different partitions, or that individual “components” DO SHARE hardware by sharing a partition
- Users are limited to seven partitions per availability zone



November 4, 2020

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

L11.9

## “df” LINUX COMMAND

- What is the difference between the `df -Th` and `df -ih` commands?
- **df -Th**
  - “T” shows the filesystem types
  - “h” provides human readable output
- **df -ih**
  - “i” shows the available number of inodes on the filesystem
  - An inode is a file record
  - A file record is required to track a file on the file system
  - It is possible to run out inodes to track files before running out of actual disk space

November 4, 2020

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

L11.10

OBJECTIVES – 11/4

- Questions from 11/2
- Quiz 1
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
- Tutorial #5
- From: Cloud Computing Concepts, Technology & Architecture:  
Chapter 5 - Cloud Enabling Technology
- Tutorial questions
- Team planning

November 4, 2020

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

L11.11

QUIZ 1

- Question 8 : partial credit added
- Question 14 : everyone received 1 point
- Curve: added 2.2 points so that class average is 85%

November 4, 2020

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

L11.12

## OBJECTIVES - 11/4

- Questions from 11/2
- Quiz 1
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
- Tutorial #5
- From: Cloud Computing Concepts, Technology & Architecture:  
Chapter 5 - Cloud Enabling Technology
- Tutorial questions
- Team planning

November 4, 2020

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

L11.13

## AWS DEMO





## LIST OF TOPICS

- AWS Management Console
- Elastic Compute Cloud (EC2)
- Instance Storage: Virtual Disks on VMs
- Elastic Block Store: Virtual Disks on VMs
- Elastic File System (EFS)
- Amazon Machine Images (AMIs)
- EC2 Paravirtualization
- EC2 Full Virtualization (hvm)
- EC2 Virtualization Evolution

- (VM) Instance Actions
  - EC2 Networking
  - EC2 Instance Metadata Service
  - Simple Storage Service (S3)
  - AWS Command Line Interface (CLI)
  - Legacy / Service Specific CLIs
  - AMI Tools
  - Signing Certificates
  - Backing up live disks
  - Cost Savings Measures

November 4, 2020

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

L11.1  
5

## EC2 INSTANCE: NETWORK ACCESS

- Public IP address
- Elastic IPs
  - Costs: in-use FREE, not in-use ~12 \$/day
  - Not in-use (e.g. “paused” EBS-backed instances)
- Security groups
  - “Firewall” restricts access to AWS resource TCP/IP ports & protocols
- Identity access management (IAM)
  - AWS accounts, groups
- Virtual Private Cloud (VPC) / Subnet(works) / Internet Gateway / Router
- Network-Address-Translation NAT-Gateway

November 4, 2020

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

L11.16

# SIMPLE VPC

- Recommended when using Amazon EC2
- 1 VPC, 1 Availability Zone, 1 Subnet

Destination	Target
10.0.0.0/16	local
0.0.0.0/0	igw-id

November 4, 2020

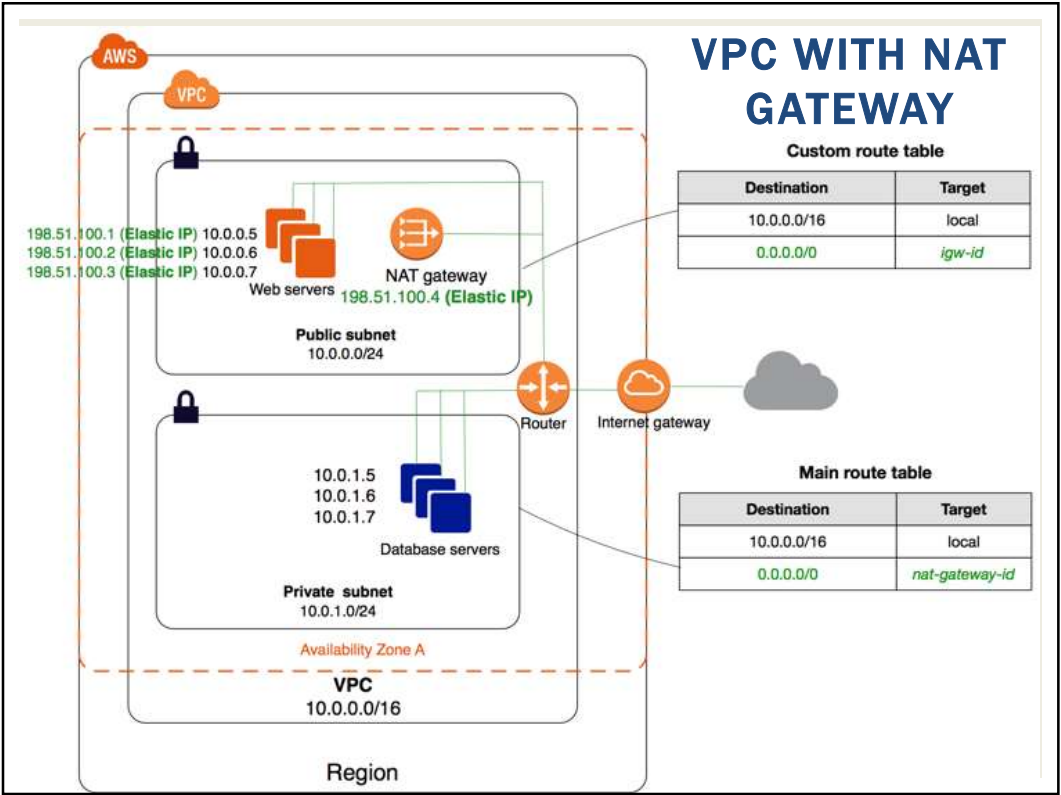
TCSS562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.17

# VPC SPANNING AVAILABILITY ZONES

- 1 VPC
- 2 Availability Zones
- 2 Subnets
- Router interconnects subnets

Destination	Target
10.0.0.0/16	local



INSPECTING INSTANCE INFORMATION

- EC2 VMs run a local metadata service
- Can query instance metadata to self discover cloud configuration attributes
- Find your instance ID:

```
curl http://169.254.169.254/  
curl http://169.254.169.254/latest/  
curl http://169.254.169.254/latest/meta-data/  
curl http://169.254.169.254/latest/meta-data/instance-id  
; echo
```
- ec2-get-info command
- Python API that provides easy/formatted access to metadata

November 4, 2020

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

L11.20

## SIMPLE STORAGE SERVICE (S3)

- Object storage service
- Indexed using key-value pairs
- Considered as blob storage as can store any type of data
- No considered as a full-featured NoSQL DB, but can be coupled with Amazon Athena to support interactive queries (SQL)
- Can mount an S3 bucket as a volume in Linux
  - Supports common file-system operations
- Data replicated for performance w/ eventual consistency
- Frequently used w/ Lambda to persist function state

November 4, 2020

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

L11.21

## AWS CLI

- Launch Ubuntu 20.04 VM
  - Instances | Launch Instance
- Install the general AWS CLI
  - `sudo apt install awscli`
  - `aws configure`
- Creates config files under `~/.aws` hidden directory
- Credentials file:  
[default]  
`aws_access_key_id = <access key id>`  
`aws_secret_access_key = <secret access key>`
- Config file:  
`region = us-east-1`

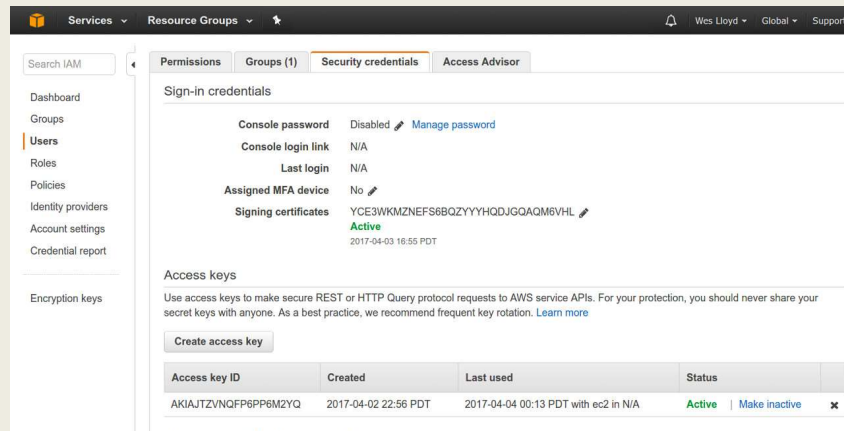
November 4, 2020

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

L11.22

## AWS CLI - 2

- **Creating access keys:** IAM | Users | Security Credentials | Access Keys | Create Access Keys



November 4, 2020

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

L11.23

## AWS CLI - 3

- **Export the config file**
  - Add to `/home/ubuntu/.bashrc`

```
export AWS_CONFIG_FILE=$HOME/.aws/config
```

- **Try some commands:**
  - `aws help`
  - `aws command help`
  - `aws ec2 help`
  - `aws ec2 describes-instances --output text`
  - `aws ec2 describe-instances --output json`
  - `aws s3 ls`
  - `aws s3 ls vmscaleruw`

November 4, 2020

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

L11.24

## LEGACY / SERVICE SPECIFIC CLI(S)

- `sudo apt install ec2-api-tools`
- Provides more concise output
- Additional functionality
- Define variables in `.bashrc` or another sourced script:
  - `export AWS_ACCESS_KEY={your access key}`
  - `export AWS_SECRET_KEY={your secret key}`
- `ec2-describe-instances`
- `ec2-run-instances`
- `ec2-request-spot-instances`
- EC2 management from Java:
  - <http://docs.aws.amazon.com/AWSJavaSDK/latest/javadoc/index.html>
- Some AWS services have separate CLI installable by package

November 4, 2020

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

L11.25

## AMI TOOLS

- Amazon Machine Images tools
- For working with disk volumes
- Can create live copies of any disk volume
  - Your local laptop, ec2 root volume (EBS), ec2 ephemeral disk
- Installation:
  - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ami-tools-commands.html>
- AMI tools reference:
  - <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ami-tools-commands.html>
- Some functions may require private key & certificate files

November 4, 2020

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

L11.26

## PRIVATE KEY AND CERTIFICATE FILE

- Install openssl package on VM

```
# generate private key file
```

```
$openssl genrsa 2048 > mykey.pk
```

```
# generate signing certificate file
```

```
$openssl req -new -x509 -nodes -sha256 -days 36500 -key  
mykey.pk -outform PEM -out signing.cert
```

- Add signing.cert to IAM | Users | Security Credentials |  
- - *new signing certificate* - -

- From: [http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/setup-ami-tools.html?icmpid=docs\\_iam\\_console#ami-tools-create-certificate](http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/setup-ami-tools.html?icmpid=docs_iam_console#ami-tools-create-certificate)

November 4, 2020

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

L11.27

## PRIVATE KEY, CERTIFICATE FILE

- These files, combined with your `AWS_ACCESS_KEY` and `AWS_SECRET_KEY` and `AWS_ACCOUNT_ID` enable you to publish new images from the CLI

- Objective:

1. Configure VM with software stack
2. Burn new image for VM replication (**horizontal scaling**)

- An alternative to bundling volumes and storing in S3 is to use a containerization tool such as Docker. . .

- Create image script . . .

November 4, 2020

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

L11.28

## SCRIPT: CREATE A NEW INSTANCE STORE IMAGE FROM LIVE DISK VOLUME

```
image=$1
echo "Burn image $image"
echo "$image" > image.id
mkdir /mnt/tmp
AWS_KEY_DIR=/home/ubuntu/.aws
export EC2_URL=http://ec2.amazonaws.com
export S3_URL=https://s3.amazonaws.com
export EC2_PRIVATE_KEY=${AWS_KEY_DIR}/mykey.pk
export EC2_CERT=${AWS_KEY_DIR}/signing.cert
export AWS_USER_ID={your account id}
export AWS_ACCESS_KEY={your aws access key}
export AWS_SECRET_KEY={your aws secret key}
ec2-bundle-vol -s 5000 -u ${AWS_USER_ID} -c ${EC2_CERT} -k ${EC2_PRIVATE_KEY}
--ec2cert /etc/ec2/amitools/cert-ec2.pem --no-inherit -r x86_64 -p $image -i
/etc/ec2/amitools/cert-ec2.pem
cd /tmp
ec2-upload-bundle -b tc5562 -m $image.manifest.xml -a ${AWS_ACCESS_KEY} -s
${AWS_SECRET_KEY} --url http://s3.amazonaws.com --location US
ec2-register tc5562/$image.manifest.xml --region us-east-1 --kernel aki-
88aa75e1
```

November 4, 2020

TCS5562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.29

## COST SAVINGS MEASURES

- **From Tutorial 3:**
- **#1: ALWAYS USE SPOT INSTANCES FOR COURSE/RESEARCH RELATED PROJECTS**
- **#2: NEVER LEAVE AN EBS VOLUME IN YOUR ACCOUNT THAT IS NOT ATTACHED TO A RUNNING VM**
- **#3: BE CAREFUL USING PERSISTENT REQUESTS FOR SPOT INSTANCES**
- **#4: TO SAVE/PERSIST DATA, USE EBS SNAPSHOTS AND THEN**
- **#5: DELETE EBS VOLUMES FOR TERMINATED EC2 INSTANCES.**
- **#6: UNUSED SNAPSHOTS AND UNUSED EBS VOLUMES SHOULD BE PROMPTLY DELETED !!**
- **#7: USE PERSISTENT SPOT REQUESTS AND THE "STOP" FEATURE TO PAUSE VMS DURING SHORT BREAKS**

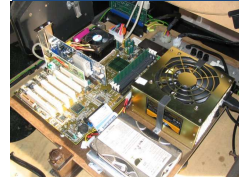
November 4, 2020

TCS5562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.30



WE WILL RETURN AT  
~7:10PM



## OBJECTIVES - 11/4

- Questions from 11/2
- Quiz 1
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
  - **Tutorial #5**
  - From: Cloud Computing Concepts, Technology & Architecture:  
Chapter 5 - Cloud Enabling Technology
  - Tutorial questions
  - Team planning

November 4, 2020

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

L11.32

## OBJECTIVES – 11/4

- Questions from 11/2
- Quiz 1
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
- Tutorial #5
- **From: Cloud Computing Concepts, Technology & Architecture:  
Chapter 5 - Cloud Enabling Technology**
- Tutorial questions
- Team planning

November 4, 2020

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

L11.33

## CLOUD ENABLING TECHNOLOGY

November 4, 2020

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

L11.34



## CLOUD ENABLING TECHNOLOGY

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

November 4, 2020

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

L11.35

## 1. BROADBAND NETWORKS AND INTERNET ARCHITECTURE

- Clouds must be connected to a network
- Inter-networking: Users' network must connect to cloud's network
- Public cloud computing relies heavily on the **internet**

November 4, 2020

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

L11.36

PRIVATE CLOUD NETWORKING

The diagram illustrates a private cloud network architecture. On the left, a cloud icon contains three server racks and three desktop computers, labeled "private cloud network". A blue router connects this network to a vertical firewall. To the right of the firewall is a "corporate Internet connection" represented by a globe and two red routers. On the far right, two human figures are shown: one with a desktop computer and one with a smartphone, both labeled "users accessing cloud services remotely". Blue lines represent network connections between the private cloud, the corporate Internet, and the remote users.

November 4, 2020

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

L11.37

PUBLIC CLOUD NETWORKING

The diagram illustrates a public cloud network architecture. It features two main cloud components. The top component, labeled "cloud consumer network", contains three server racks and three desktop computers labeled "cloud consumers". It is connected via a blue router and a firewall to a "corporate Internet connection" (globe and red routers). Remote users, labeled "users accessing cloud services remotely", are shown on the right. The bottom component, labeled "cloud provider network", is a cloud icon containing multiple server racks. It is connected via a blue router and a firewall to its own "cloud provider Internet connection" (globe and red routers). Blue lines represent the network connections between the consumer network, the corporate Internet, the provider network, and the remote users.

November 4, 2020

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

L11.38

## INTERNETWORKING KEY POINTS

- Cloud consumers and providers typically communicate via the internet
- Decentralized provisioning and management model is not controlled by the cloud consumers or providers
- Inter-networking (internet) relies on connectionless packet switching and route-based interconnectivity
- Routers and switches support communication
- Network bandwidth and latency influence QoS, which is heavily impacted by network congestion

November 4, 2020

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

L11.39

## CLOUD ENABLING TECHNOLOGY

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


November 4, 2020

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

L11.40

## 2. DATA CENTER TECHNOLOGY

- Grouping servers together (clusters):
- Enables power sharing
- Higher efficiency in shared IT resource usage (less duplication of effort)
- Improved accessibility and organization
- Key components:
  - Virtualized and physical server resources
  - Standardized, modular hardware
  - Automation support: ease server provisioning, configuration, patching, monitoring without supervision... *tools are desirable*

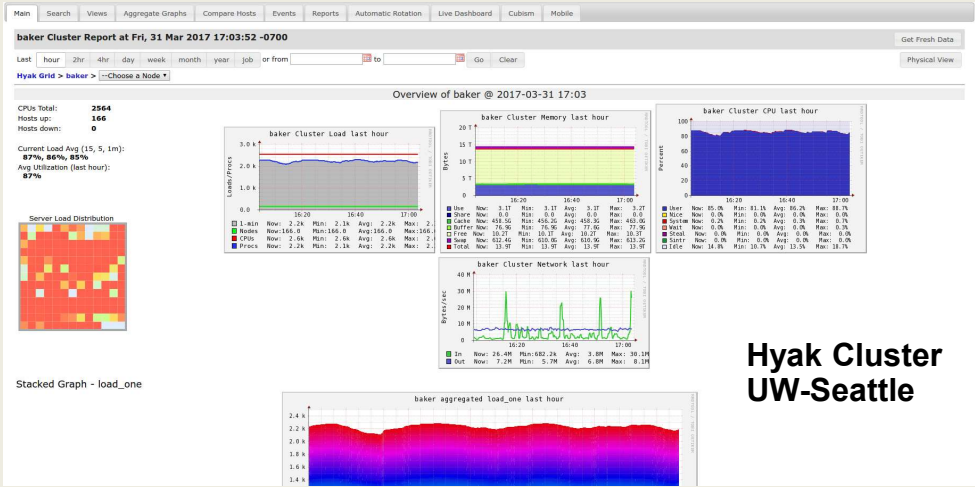


November 4, 2020

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

L11.41

## CLUSTER MANAGEMENT TOOLS



November 4, 2020

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

L11.42

## DATA CENTER TECHNOLOGY – KEY COMPONENTS

- Remote operation / management
- **High availability support:** \*\*redundant everything\*\*  
Includes: power supplies, cabling, environmental control systems, communication links, duplicate warm replica hardware
- **Secure design:** physical and logical access control
- **Servers:** rackmount, etc.
- **Storage:** hard disk arrays (RAID), storage area network (SAN): disk array with dedicated network, network attached storage (NAS): disk array on network for NFS, etc.
- **Network hardware:** backbone routers (WAN to LAN connectivity), firewalls, VPN gateways, managed switches/routers

November 4, 2020

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

L11.43

## CLOUD ENABLING TECHNOLOGY

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

November 4, 2020

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

L11.44

### 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 4, 2020

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

L11.45

### 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 4, 2020

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


L11.46



# KEY VIRTUALIZATION TRADEOFF

■ Tradeoff space:  
What is the “right” level of abstraction?

Degree of  
Hardware  
Abstraction



Abstraction  
Concerns:  
Overhead  
Performance  
Isolation  
Security

November 4, 2020

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

L11.47

# TYPE 1 HYPERVISOR

VM  
(guest operating  
system and  
application  
software)

VM  
(guest operating  
system and  
application  
software)

VM  
(guest operating  
system and  
application  
software)

Virtual Machine Management  
Hypervisor

Hardware  
(virtualization host)

■ Host OS and VMs run atop the hypervisor

■ The boot OS is the hypervisor kernel

■ Xen dom0

November 4, 2020

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

L11.48

Slides by Wes J. Lloyd

L11.24

## 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 4, 2020

TCS5562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.49

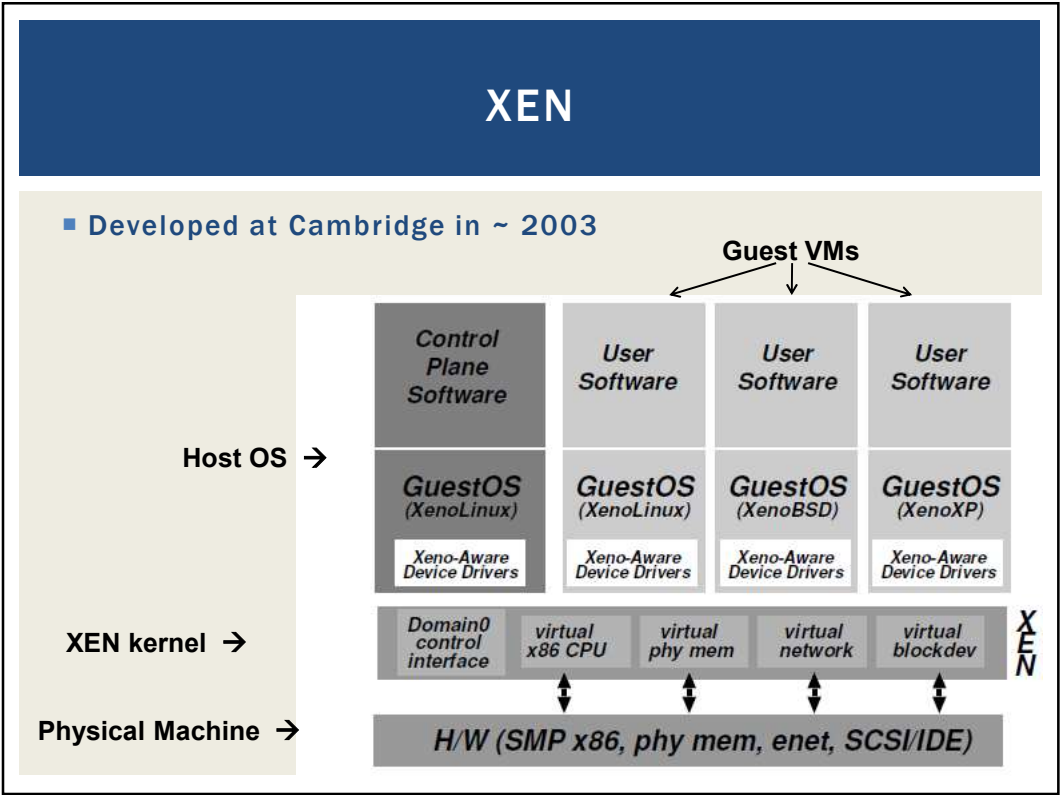
## COMMON VMMS: PARAVIRTUALIZATION

- TYPE 1
- 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 4, 2020

TCS5562: Software Engineering for Cloud Computing [Fall 2020]  
School of Engineering and Technology, University of Washington - Tacoma

L11.50



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

```
xentop - 17:53:48 Xen 3.1.2-398.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
NETS NETTX(k) NETRX(k) VBDS VBD OO VBD RD VBD WR SSID
centos --b--- 46 0.0 532352 6.4 1064960 12.7 1
1 27960 885 1 0 6313 37119 0
centos-2 --b--- 17 0.0 1056640 12.6 2113536 25.2 1
1 50 0 1 0 3981 541 0
Domain-0 -----r 2979 19.3 6568960 78.4 no limit n/a 4
4 1057374 290072 0 0 0 0 0
```

November 4, 2020

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

L11.52

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 4, 2020

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

L11.53

TYPE 2 HYPERVISOR

- Adds additional layer

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

November 4, 2020

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

L11.54

## 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 4, 2020

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

L11.55

## KERNEL BASED VIRTUAL MACHINES (KVM)

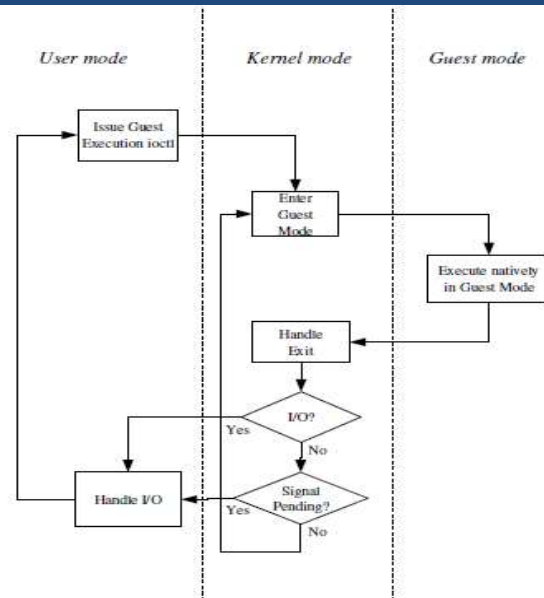
- **x86 HW notoriously difficult to virtualize**
- **Extensions added to 64-bit Intel/AMD CPUs**
  - **Provides hardware assisted virtualization**
  - **New “guest” operating mode**
  - **Hardware state switch**
  - **Exit reason reporting**
  - **Intel/AMD implementations different**
    - **Linux uses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

November 4, 2020

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

L11.56

## KVM - 2



## 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 4, 2020

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

L11.58

## 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 4, 2020

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

L11.59

## 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 4, 2020

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

L11.60

## 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 4, 2020

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

L11.61

## 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 4, 2020

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

L11.62



# VIRTUAL INFRASTRUCTURE MANAGEMENT (VIM)

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

November 4, 2020

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

L11.63

## 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 4, 2020

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

L11.64

## 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 4, 2020

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

L11.65

## VIM FEATURES - 3

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

November 4, 2020

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

L11.66

## CONTAINER ORCHESTRATION FRAMEWORKS

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

November 4, 2020

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

L11.67

## 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 4, 2020

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

L11.68

## CLOUD ENABLING TECHNOLOGY

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

November 4, 2020

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

L11.69

## 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



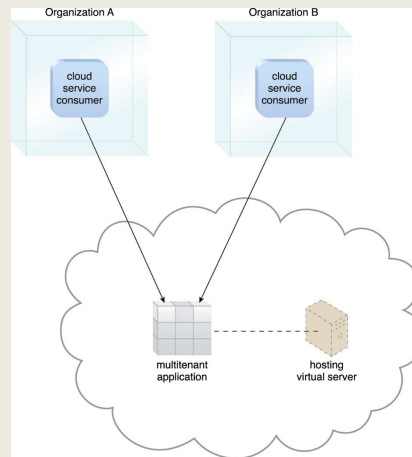
November 4, 2020

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

L11.70

## MULTITENANT APPS - 2

- Forms the basis for SaaS (applications)



November 4, 2020

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

L11.71

## CLOUD ENABLING TECHNOLOGY

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

November 4, 2020

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

L11.72

## 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 4, 2020

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

L11.73

## 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 4, 2020

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

L11.74

## 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 4, 2020

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

L11.75

// 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 4, 2020

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

L11.76

```
// 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 4, 2020

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

L11.77

```
// WSDL Service Definition
<?xml version="1.0" encoding="UTF-8"?>
<definitions name="DayOfWeek"
targetNamespace="http://www.roguewave.com/soapworx/examples/DayOfWeek.wsdl"
xmlns:tns="http://www.roguewave.com/soapworx/examples/DayOfWeek.wsdl"
xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns="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"/>
      <input>
        <soap:body use="encoded"
namespace="http://www.roguewave.com/soapworx/examples"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding"/>
      </input>
      <output>
        <soap:body use="encoded"
namespace="http://www.roguewave.com/soapworx/examples"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding"/>
      </output>
    </operation>
  </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:8090/dayofweek/DayOfWeek"/>
    </port>
  </service>
</definitions>
```

November 4, 2020

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

L11.78



## REST CLIMATE SERVICES EXAMPLE

- **USDA**  
**Lat/Long**  
**Climate**  
**Service**  
**Demo**

```
// REST/JSON
// Request climate data for Washington

{
  "parameter": [
    {
      "name": "latitude",
      "value": 47.2529
    },
    {
      "name": "longitude",
      "value": -122.4443
    }
  ]
}
```
- **Just provide**  
**a Lat/Long**

November 4, 2020

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

L11.79

## 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 4, 2020

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

L11.80

## 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 4, 2020

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

L11.81

## OBJECTIVES – 11/2

- Questions from 10/28
- Term Project Proposals
- Quiz 1 Review
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
  - Tutorial #5
  - AWS overview and demonstration
  - Cloud Enabling Technology
  - Tutorial questions
  - Team planning

November 4, 2020

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

L11.82

OBJECTIVES – 11/2


- Questions from 10/28
- Term Project Proposals
- Quiz 1 Review
- AWS overview and demonstration
- 2<sup>nd</sup> hour:
  - Tutorial #5
  - AWS overview and demonstration
  - Cloud Enabling Technology
  - Tutorial questions
  - Team planning

November 4, 2020

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

L11.83

QUESTIONS



November 4, 2020

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

L11.84