Indian Institute of Information Technology Allahabad

End Sem Question Paper

Course Name: Cloud and Edge Computing **Course Instructor/ Co-ordinator :** Bibhas Ghoshal

Course Code: CEC **Program Name:** B.Tech(IT and ECE) **Exam Date:** 02.05.24 . **MM:** 25

Q1. (a) The datacenter hardware and software is what we call ---------. When a Cloud is made available in a pay-as-you-go manner to the general public, we call it a ------ Cloud; the service being sold is -------------. The -------- refer to internal data centers of a business or other organization,

Soln : The datacenter hardware and software is what we will call a Cloud. When a Cloud is made available in a pay-as-you-go manner to the general public, we call it a Public Cloud; the service being sold is
Utility Computing. We use the term Private Cloud to refer to internal datacenters of a business or other organization

(b) An organization debating whether to install a private cloud or to use a public cloud, e.g., the AWS, for its computational and storage needs, asks your advice. What information will you require to base your recommendation on, and how will you use each one of the following items: (a) the description of the algorithms and the type of the applications the organization will run; (b) the system software used by these applications; (c) the resources needed by each application; (d) the size of the user population; (e) the relative experience of the user population; (d) the costs involved.

Solution. Public clouds have distinct cost advantages over private clouds; there is no initial investment in the infrastructure, no recurring costs for administration, maintenance, energy consumption, and for the user support personnel. The main concern is security and privacy. An organization with very strict security and privacy concerns is very unlikely to use a public cloud. The type of applications play a critical role, scientic and engineering computations which require a low latency interconnection network and enjoy only ne-grained parallelism are unlikely to fare well on either a public or a private cloud. A large user population is more likely to use identical or similar software and to cooperate by sharing the raw data and the results; thus, a private cloud seems more advantageous in this case. Some of the services oered by private clouds target experienced users, e.g., AWS services such as ElasticBeanstalk, while others are accessible to lay persons.

(c) Compare the three cloud computing delivery models, SaaS, PaaS, and IaaS, **from the point of view of the application developers and users.**

Solution. Software-as-a-Service (SaaS) gives the capability to use applications supplied by the service provider in a cloud infrastructure. The applications are accessible from various client devices through a thin-client interface such as a Web browser (e.g., Web-based email). The user does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

SaaS is not suitable for applications that require real-time response or those for which data is not allowed to be hosted externally. The most likely candidates for SaaS are applications for which many competitors use the same product, such as email. Periodically there is a significant peak in demand, such as billing and payroll. There is a need for Web or mobile access, such as mobile sales management software. There is only a short-term need, such as collaborative software for a project.

Platform-as-a-Service (PaaS) gives the capability to deploy consumer-created or acquired applications using programming languages and tools supported by the provider. The user does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, or storage. The user has control over the deployed applications and, possibly, over the application hosting environment configurations. Services offered by this delivery model include session management, device integration, sandboxes, instrumentation and testing, contents management, knowledge management, and Universal Description, Discovery, and Integration (UDDI), a platform-independent Extensible Markup Language (XML)-based registry providing a mechanism to register and locate Web service applications PaaS is not particularly useful when the application must be portable, when proprietary programming languages are used, or when the underlaying hardware and software must be customized to improve the performance of the application. The major PaaS application areas are in software development where multiple developers and users collaborate and the deployment and testing services should be automated.
Infrastructure-as-a-Service (IaaS) offers the capability to provision processing, storage, networks, and other fundamental computing resources; the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of some networking components, such as host firewalls. Services offered by this delivery model include: server hosting, Web servers, storage, computing hardware, operating systems, virtual instances, load balancing, Internet access, and bandwidth provisioning.

The **IaaS** cloud computing delivery model has a number of characteristics, such as the fact that the resources are distributed and support dynamic scaling, it is based on a utility pricing model and variable cost, and the hardware is shared among multiple users. This cloud computing model is particularly useful when the demand is volatile and a new business needs computing resources and does not want to invest in a computing infrastructure or when an organization is expanding rapidly.

Q2. (a) Mention the type of virtualization techniques usually employed.
Soln : System ISA Virtualization

- Instruction Interpretation
- Trap and Emulate
- Binary Translation
- Hybrid Models

(b) Differentiate between full virtualization and para virtualization
Solution :
Full virtualization : Guest OS can run unchanged under the VMM as if running directly on the hardware platform.; Requires a virtualizable architecture, Examples: VMware.

Para virtualization :Guest operating system is modified to use only instructions that can be virtualized. Reasons for
para virtualization:

- Some aspects of the hardware cannot be virtualized.
- Improved performance.
- Present a simpler interface.
- Examples: Xen, Denaly

(c) Consider a Linux-like OS that is built to run on non-virtualized systems, and is compiled into a binary for the x86 architecture. A VMM wishes to run this OS within a guest VM using the **trap-and-emulate method. The underlying hardware has no virtualization support.** Can the VMM run this OS as a guest without having to modify its source code and achieve **correct virtualization?** Answer yes/no and justify. If you answer yes, give an example of aVMM that runs unmodified OS source in this manner. If you answer no, explain why this is not possible to do.

Solution : Yes, VMware workstation hypervisor / Full Virtualization / binary translation

(d) Popek Goldgerg theorem states that a hypervisor for a processor with a given ISA can be constructed if the set of sensitive instructions is a subset of the privileged instructions of that processor. Identify two sensitive instructions for the x86 architecture and discuss the problem each one of these instruction poses.

Soln : Ring deprivileging - a hypervisor forces the guest software, the operating system and the applications to run at a privilege level greater than 0. Recall that the x86 architecture provides four protection rings, 0-3. Two solutions are then possible: (a) the (0=1=3) mode when the hypervisor, the OS, and the application run at privilege levels 0; 1 and 3, respectively; (b) the (0; 3; 3) mode when the hypervisor, a guest OS, and applications run at privilege levels 0; 3 and 3, respectively.
Problems are created when a guest OS is forced to run at a privilege level other than that it was originally designed for. For example, when the CS register2 is PUSHed, the current privilege level in the CR is also stored on the stack.

Address space compression - a hypervisor uses parts of the guest address space to store several system data structures such as the interrupt-descriptor table and the globaldescriptor table. Such data structures must be protected, but the guest software must **have access to them.**
Guest system calls; two instructions SYSENTER and SYSEXIT support low-latency system calls. The rst causes a transition to privilege level 0, while the second causes a transition from privilege level 0 and fails if executed at a level higher than 0. The **hypervisor must then emulate every guest execution of either of these instructions and** this has a negative impact on performance.

Q3.
(a) Mention the steps to migrate a VM from Host A to Host B

Solution : 1. Setup target host B, reserve resources for the VM
2. Push phase: push some memory of VM from A to B
3. Stop-and-copy : stop the VM at A, copy CPU context, and some memory
4. Pull phase: Start VM at host B, pull further memory required from A
5. Clean up state from host A, migration complete

(b) How are dirty pages managed during live migration in Xen Hypervisor?
Soln : Shadow page table constructed on demand for every round
• Dirty bitmap maintained for every round
• Any page access by guest □ page fault to Xen, shadow page table updated
• PTE marked as read-only by default in shadow
• If valid write access, shadow PTE marked writeable, page marked dirty in bitmap
• At end of round, dirty pages are marked for transfer in control software
• Shadow page table and dirty bitmap reinitialized after every round
• Last set of dirty pages copied in stop-and-copy
Guest page table in target host changed based on new physical addresses

(c) Explain with examples the difference between PID namespace and Mount namespace?

Soln : Mount namespace: isolates the filesystem mount points seen by a group of processes. The mount() and umount() system calls only affect the processes in that namespace. Root filesystem seen by a process is constructed from a set of mount points (mount() and umount() syscalls)
• New mount namespace can have new set of mount points
• New view of root filesystem

PID namespace: isolates the PID numberspace seen by processes. E.g., first process in a new PID namespace gets a PID of 1.

Q4. (a) List some of the challenges in Cloud based storage systems
Soln : High performance and reliable; Maintaining consistency among multiple copies of data records ; Sophisticated strategies to reduce the access time and to support multimedia access are necessary to satisfy the timing requirements of data streaming and content delivery; Data replication allows concurrent access to data from multiple processors and decreases the chances of data loss.

(b) Draw the architecture of Google File System (GFS) cluster

(c) Discuss the use of locking in the Bigtable.
Solution. BigTable is a distributed storage system developed by Google to store massive amounts of data and to scale up to thousands of storage servers [94]]. The system uses the Google File System to store user data as well as system information. To guarantee atomic read and write operations, it uses the Chubby distributed lock service; the directories and the files in the namespace of Chubby are used as locks. The system is based on a simple and flexible data model. It allows an application developer to exercise control over the data format and layout and reveals data locality information to the application clients. Any read or write row operation is atomic, even when it affects more than one column. The column keys identify column families, which are units of access control. The data in a column family is of the same type. Client applications written in C++ can add or delete values, search for a subset of data, and look up data in **a row.**

(d) Consider the idea of vector clocks related to the Dynamo DB. Three nodes A, B, C in a distributed key-value store use vector clocks to track the versions of the key-value pairs stored in the system. A client stores various versions of the value corresponding to a key at various nodes in the following manner. Suppose the client stores a value V1 at node A with an initial vector clock of [(A,1), (B,0), (C,0)]. The client then reads the value V1 from A and writes

another value V2 to B. Then, two other clients read V2 from B, and write updated versions V3 and V4 at A and C respectively. Another client then reads the values V3 and V4, reconciles any conflicts, and writes an updated merged value V5 at node A. What is the vector clock of this value V5 stored at A?
Solution : (A,3), (B,1), (C,1)

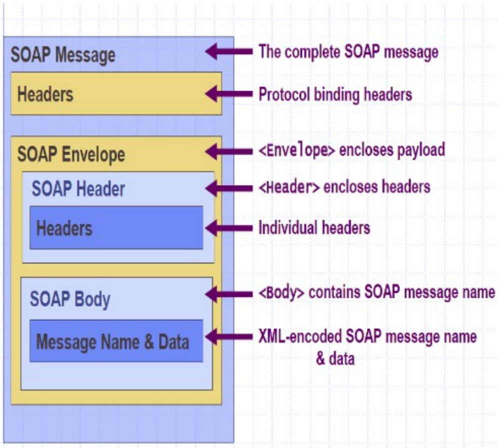
Q5. (a) Mention how the six attack surfaces apply to SaaS/PaaS or Iaas

Solution. The three actors involved in the model considered are: the user,the service, and the cloud infrastructure, and there are six types of attacks possible.
The user can be attacked from two directions, the service and the cloud. SSL certificate spoofing, attacks on browser caches, or phishing attacks are examples of attacks that originate at the service. The user can also be a victim of attacks that either originate at the cloud or spoofs to originate from the cloud infrastructure. In the case of SaaS attacks from the cloud infrastructure and from the service are less likely.

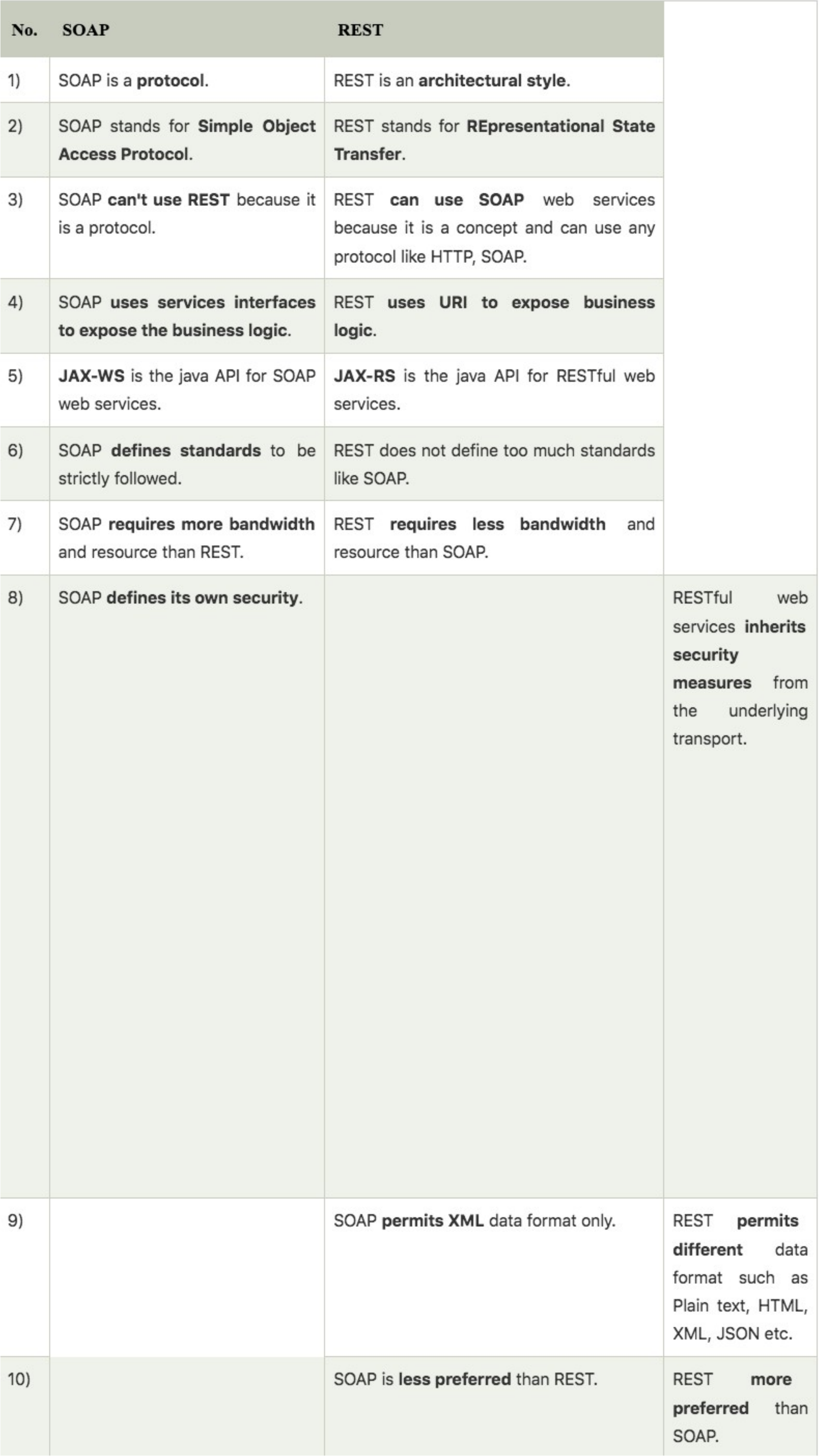
SQL injection, and privilege escalation are the common types of attacks for the service. The service can also be subject to attacks by the cloud infrastructure and this is probably the most serious line of attack; this is less likely in the case of SaaS. Limiting access to resources, privilege-related attacks, data distortion, injecting additional operations, are only a few of the many possible lines of attacks originated at the cloud.
The cloud infrastructure can be attacked by a user which targets the cloud control system; this type of attack is common for all cloud delivery models. The types of attacks are the same a user directs toward any other cloud service. The cloud infrastructure may also be targeted by a service requesting an excessive amount of resources and causing the exhaustion of the resources; this is not realistic for the SaaS cloud delivery model.

(b) Mention the four protocols of web service software stack
Solution :
Web services are based on four protocols:
Extensible Markup Language (XML) : defines complex data structures
Web Service Description Language (WSDL) : Specifies the interface of the web service
Discovery Protocol (DISCO): Pointer to all web services on a particular web site
Universal Description, Discovery, and Integration (UDDI): Central repository of web service descriptions

(c) mention the structure of a SOAP message


SOAP Message
The complete SOAP message
Headers
Protocol binding headers
SOAP Envelope
<Envelope> encloses payload
SOAP Header
<Header> encloses headers
Headers
Individual headers
SOAP Body
<Body> contains SOAP message name
Message Name & Data
XML-encoded SOAP message name & data

(d) List a few differences between SOAP based web services and REST based web services.



No.	SOAP	REST	
1)	SOAP is a **protocol**.	REST is an **architectural style**.	
2)	SOAP stands for **Simple Object Access Protocol**.	REST stands for **REpresentational State Transfer**.	
3)	SOAP **can't use REST** because it is a protocol.	REST **can use SOAP** web services because it is a concept and can use any protocol like HTTP, SOAP.	
4)	SOAP **uses services interfaces to expose the business logic.**	REST **uses URI to expose business logic.**	
5)	**JAX-WS** is the java API for SOAP web services.	**JAX-RS** is the java API for RESTful web services.	
6)	SOAP **defines standards** to be strictly followed.	REST does not define too much standards like SOAP.	
7)	SOAP **requires more bandwidth** and resource than REST.	REST **requires less bandwidth** and resource than SOAP.	
8)	SOAP **defines its own security**.		RESTful web services **inherits security measures** from the underlying transport.
9)		SOAP **permits XML** data format only.	REST **permits different** data format such as Plain text, HTML, XML, JSON etc.
10)		SOAP is **less preferred** than REST.	REST **more preferred** than SOAP.

Q6. (a) Recently Batman has become aware that the level of crime in Gotham City has reached an all-time high. The problem is he is having trouble tracking which citizens of Gotham are committing these heinous crimes. Batman forms a clever plan:

- Perform a thorough scan of each person in Gotham using the bat computer and give them a badness rating.
- Track the citizens with the highest badness rating.

Unfortunately for Batman, the bat computer, while powerful, can only process one person per second and Gotham has a population of one million! It means, that's over a week time! Batman can't sit idle, he needs answers now!

Fortunately for Batman, Alfred has installed a network of one thousand consumer-grade machines (the "bat cluster", naturally). Batman wonders if there isn't a way to leverage this network of machines. Can you help Batman design a solution to the problem using Alfred's network of machines?

Solution : On the bat computer split the citizens into chunks of 1000.

- Send each chunk to a separate node in the bat cluster.
- Each node will scan all the citizens in that chunk and create a badness rating.
- The badness ratings will be returned to the bat computer
- The bat computer will sort the returned pairs by badness rating.
- Batman will track citizens with the highest badness rating.

The **map** algorithm is analyzing a chunk of 1000 citizens and assigns a badness rating. The **reducer** algorithm is sorting the resulting badness ratings and produces a list of villains.

(b) **What is shuffling and sorting in MapReduce?**
Solution : The map tasks generate intermediate key-value pairs and store them on partitions on the local discs. Shuffling is the process of transferring this intermediate data generated on multiple machines to reducers. This data fetched by the reducers are sorted and grouped by the intermediate key, before the data is send as input to the reduce task.

Q7. (a) Consider a company which operates a number of wind farms and needs to constantly monitor and adjust the direction and angles of the wind turbines. The monitoring and decision making is done on the company's cloud or on-premise servers. However, there is a problem with this approach. **The further away from the wind-farm the decision is made, the higher the likelihood of wind conditions changing before the turbine adjustments are made**.

Can you suggest a solution to the company to get rid of the problem?

Solution : Each wind-farm should have its own ability to make these determinations, away from the company's cloud or on-premise servers.

Near-instant decisions such as **sudden wind gusts** can be made using **edge-computing (**computing which takes place using hardware and software embedded in a given wind-turbine) allowing wind-turbines to make quick decisions to prevent damage or to make micro-adjustments increasing the efficiency of the turbine.

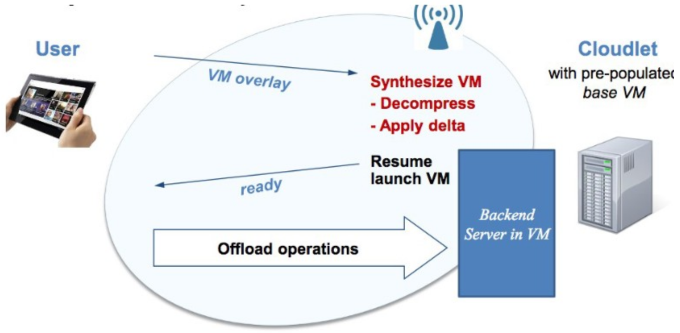
Finer-grained decisions such as **account for differences in weather at different wind farms** can be made using fog-computing

Broad and less time sensitive decisions can be made in the cloud or on-premise

(b) What are cloudlets?
Solution : a nearby offloading site dispersed at the edges of the Internet

(c) What is VM overlay? How is it used in VM synthesis used in Cloudlets (explain steps of VM synthesis)?

Solution :
VM overlay: A binary patch that contains customized parts.
Base VM - Customized VM = Binary Delta which when compressed = VM Overlay


User
VM overlay
Synthesize VM
- Decompress
- Apply delta
Resume
launch VM
ready
Offload operations
Cloudlet
with pre-populatec
base VM
Backend
Server in VM