Parallel and Distributed Computing (WQD7008)

Quiz 2

1. The concept of creating a "grid" from the unused resources in a network of computers is known as _____. (4 Scores)
   A. CPU scavenging B. Grid Computing C. Cloud Computing D. Virtual Organization
2. Grid vs the conventional supercomputers, in the context of distributed computing: (4 Scores)
   A. Supercomputers like MPPs in the Top-500 list are more homogeneously structured with tightly coupled operations
   B. Grids are built with heterogeneous nodes running non-interactive workloads. These grid workloads may involve a large number of files and individual users.
   C. The geographically dispersed grids are more scalable and fault-tolerant with significantly lower operational costs than the supercomputers.
   D. All of the above statements.
3. Explain the differences between public private and hybrid clouds. (15 Scores)
4. Explains cloud computing service models in detail. Provide supportive examples. (18 Scores)
5. What network topology is used as the Interconnection of Modular Data Centers. Draw the topology with k=4. (12 Scores)
6. What are the six architectural challenges in cloud architecture development. Explain minimum two of them in detail. (12 Scores)
7. Explain four principles of REST. (8 Scores)
8. Explain features, advantage, disadvantage and limitations of full virtualization and Para visualization. (10 Scores).
9. Describe memory virtualization in Virtual Machines and Virtualization. (5 Scores).
10. Assume a relatively large-scale scenario in which a data processing application is about to deploy on a cloud environment. Many users may access the app at the same time. Can this scenario be deployed on the Microsoft Azure? If yes, provide your justification. If no, discuss all the service models being contributed for such a scenario and justify your answer. (12 Scores).

Total Marks 100.

3. Week 8- Cloud Computing Platforms – power point slide -page 4-7

Public Clouds

- Built over the Internet and can be accessed by any user who has paid for the service.
- Owned by service providers and are accessible through a subscription.
- Many public clouds are available, including Google App Engine (GAE), Amazon Web Services (AWS), Microsoft Azure, IBM Blue Cloud, and Salesforce.com's Force.com.
- The providers of the aforementioned clouds are commercial providers that offer a publicly accessible remote interface for creating and managing VM instances within their proprietary infrastructure.
- A public cloud delivers a selected set of business processes.
- The application and infrastructure services are offered on a flexible price-per-use basis.

Private Clouds

- Built within the domain of an intranet owned by a single organization.
- Therefore, it is client owned and managed, and its access is limited to the owning clients and their partners.
- Its deployment was not meant to sell capacity over the Internet through publicly accessible interfaces.
- Private clouds give local users a flexible and agile private infrastructure to run service workloads within their administrative domains.
- A private cloud is supposed to deliver more efficient and convenient cloud services.
- It may impact the cloud standardization, while retaining greater customization and organizational control**.**

4. Week 8- Cloud Computing Platforms – power point slide -page 16

Infrastructure-as-a-Service (IaaS)

- This model allows users to use virtualized IT resources for computing, storage, and networking. In short, the service is performed by rented cloud infrastructure. The user can deploy and run his applications over his chosen OS environment.
- The user does not manage or control the underlying cloud infrastructure, but has control over the OS, storage, deployed applications, and possibly select networking components. This IaaS model encompasses storage as a service, compute instances as a service, and communication as a service.

Platform as a Service (PasS)

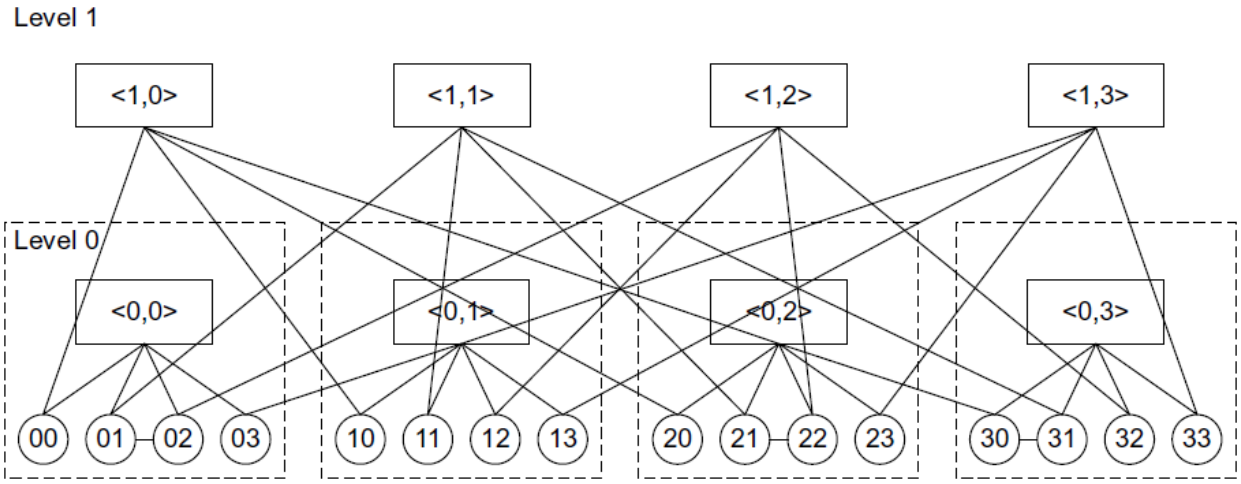
- To be able to develop, deploy, and manage the execution of applications using provisioned resources demands a cloud platform with the proper software environment. Such a platform includes operating system and runtime library support. This has triggered the creation of the PaaS model to enable users to develop and deploy their user applications.

Software as a Service (SaaS)

- This refers to browser-initiated application software over thousands of cloud customers. Services and tools offered by PaaS are utilized in construction of applications and management of their deployment on resources offered by IaaS providers.
- The SaaS model provides software applications as a service. As a result, on the customer side, there is no upfront investment in servers or software licensing.
- On the provider side, costs are kept rather low, compared with conventional hosting of user applications.
- Customer data is stored in the cloud that is either vendor proprietary or publicly hosted to support PaaS and IaaS.
- The best examples of SaaS services include Google Gmail and docs, Microsoft SharePoint, and the CRM software from Salesforce.com.

5. Week - Virtual Machines and Virtualization of Datacenters – power point slide -page 11

A BCube is used as a server-centric network design for modular Data Centers



6. Week 7- Cloud Computing Platforms – power point slide -page 22

Challenge 1—Service Availability and Data Lock-in Problem

The management of a cloud service by a single company is often the source of single points of failure. To achieve HA, one can consider using multiple cloud providers.

Even if a company has multiple data centers located in different geographic regions, it may have common software infrastructure and accounting systems.

Therefore, using multiple cloud providers may provide more protection from failures.

Another availability obstacle is distributed denial of service (DDoS) attacks.

Criminals threaten to cut off the incomes of SaaS providers by making their services unavailable.

Some utility computing services offer SaaS providers the opportunity to defend against DDoS attacks by using quick scale-ups

Challenge 2—Data Privacy and Security Concerns

Current cloud offerings are essentially public (rather than private) networks, exposing the system to more attacks. Many obstacles can be overcome immediately with well-understood technologies such as encrypted storage, virtual LANs, and network middleboxes (e.g., firewalls, packet filters). For example, you could encrypt your data before placing it in a cloud. Many nations have laws requiring SaaS providers to keep customer data and copyrighted material within national boundaries.

Traditional network attacks include buffer overflows, DoS attacks, spyware, malware, rootkits, Trojan horses, and worms. In a cloud environment, newer attacks may result from hypervisor malware, guest hopping and hijacking, or VM rootkits. Another type of attack is the man-in-the-middle attack for VM migrations. In general, passive attacks steal sensitive data or passwords. Active attacks may manipulate kernel data structures which will cause major damage to cloud servers.

Challenge 3—Unpredictable Performance and Bottlenecks

Multiple VMs can share CPUs and main memory in cloud computing, but I/O sharing is problematic. For example, to run 75 EC2 instances with the STREAM benchmark requires a mean bandwidth of 1,355 MB/second. However, for each of the 75 EC2 instances to write 1 GB files to the local disk requires a mean disk write bandwidth of only 55 MB/second. This demonstrates the problem of I/O interference between VMs. One solution is to improve I/O architectures and operating systems to efficiently virtualize interrupts and I/O channels.

Challenge 4—Distributed Storage and Widespread Software Bugs

The database is always growing in cloud applications. The opportunity is to create a storage system that will not only meet this growth, but also combine it with the cloud advantage of scaling arbitrarily up and down on

demand. This demands the design of efficient distributed SANs. Data centers must meet programmers' expectations in terms of scalability, data durability, and HA. Data consistence checking in SAN-connected data centers is a major challenge in cloud computing.

Large-scale distributed bugs cannot be reproduced, so the debugging must occur at a scale in the production data centers. No data center will provide such a convenience. One solution may be a reliance on using VMs in cloud computing. The level of virtualization may make it possible to capture valuable information in ways that are impossible without using VMs. Debugging over simulators is another approach to attacking the problem, if the simulator is well designed.

Challenge 5—Cloud Scalability, Interoperability, and Standardization

The pay-as-you-go model applies to storage and network bandwidth; both are counted in terms of the number of bytes used. Computation is different depending on virtualization level. Google App Engine (GAE) automatically scales in response to load increases and decreases; users are charged by the cycles used. AWS charges by the hour for the number of VM instances used, even if the machine is idle. The opportunity here is to scale quickly up and down in response to load variation, in order to save money, but without violating SLAs. Open Virtualization Format (OVF) describes an open, secure, portable, efficient, and extensible format for the packaging and distribution of VMs. It also defines a format for distributing software to be deployed in VMs. This VM format does not rely on the use of a specific host platform, virtualization platform, or guest operating system. The approach is to address virtual platform-agnostic packaging with certification and integrity of packaged software.

OVF also defines a transport mechanism for VM templates, and can apply to different virtualization platforms with different levels of virtualization. In terms of cloud standardization, we suggest the ability for virtual appliances to run on any virtual platform. We also need to enable VMs to run on heterogeneous hardware platform hypervisors. This requires hypervisor-agnostic VMs. We also need to realize cross-platform live migration between x86 Intel and AMD technologies and support legacy hardware for load balancing. All these issue are wide open for further research.

Challenge 6—Software Licensing and Reputation Sharing

Many cloud computing providers originally relied on open source software because the licensing model for commercial software is not ideal for utility computing. The primary opportunity is either for open source to remain popular or simply for commercial software companies to change their licensing structure to better fit cloud computing. One can consider using both pay-for-use and bulk-use licensing schemes to widen the business coverage.

One customer's bad behavior can affect the reputation of the entire cloud. For instance, blacklisting of EC2 IP addresses by spam-prevention services may limit smooth VM installation. An opportunity would be to create reputation-guarding services similar to the "trusted e-mail" services currently offered (for a fee) to services hosted on smaller ISPs. Another legal issue concerns the transfer of legal liability. Cloud providers want legal liability to remain with the customer, and vice versa. This problem must be solved at the SLA level.

7. Week 5- Service Oriented Architecture– power point slide -page 19

Resource Identification through URIs: Any information that can be named can be a resource, such as a document or image or a temporal service. A resource is a conceptual mapping to a set of entities. Each particular resource is identified by a unique name, or more precisely, a Uniform Resource Identifier (URI) which is of type URL, providing a global addressing space for resources involved in an interaction between components as well as facilitating service discovery. The URIs can be bookmarked or exchanged via hyperlinks, providing more readability and the potential for advertisement.

Uniform, Constrained Interface: Interaction with RESTful web services is done via the HTTP standard, client/server cacheable protocol. Resources are manipulated using a fixed set of four CRUD (create, read, update,

delete) verbs or operations: PUT, GET, POST, and DELETE. PUT creates a new resource, which can then be destroyed by using DELETE. GET retrieves the current state of a resource. POST transfers a new state onto a resource.

Self-Descriptive Message: A REST message includes enough information to describe how to process the message. This enables intermediaries to do more with the message without parsing the message contents. In REST, resources are decoupled from their representation so that their content can be accessed in a variety of standard formats (e.g., HTML, XML, MIME, plain text, PDF, JPEG, JSON, etc.). REST provides multiple/alternate representations of each resource. Metadata about the resource is available and can be used for various purposes, such as cache control, transmission error detection, authentication or authorization, and access control.

Stateless Interactions: The REST interactions are "stateless" in the sense that the meaning of a message does not depend on the state of the conversation. Stateless communications improve visibility, since a monitoring system does not have to look beyond a single request data field in order to determine the full nature of the request reliability as it facilitates the task of recovering from partial failures, and increases scalability as discarding state between requests allows the server component to quickly free resources. However, stateless interactions may decrease network performance by increasing the repetitive data (per-interaction overhead). Stateful interactions are based on the concept of explicit state transfer.

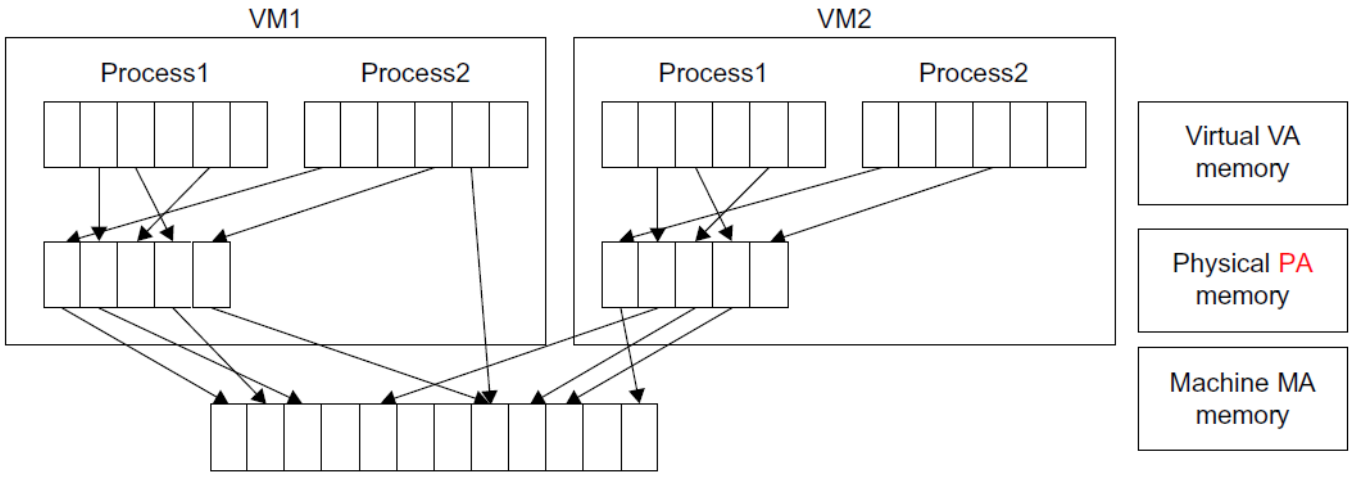
8. Week 7- Virtual Machines and Virtualization of Datacenters – power point slide -page 30

Full virtualization	**Para virtualization**
Guest operating systems are unaware of each other	unlike full virtualization, guest servers are aware of one another
Provide support for unmodified guest operating system.	Hypervisor does not need large amounts of processing power to manage guest OS
Hypervisor directly interact with the hardware such as CPU, disks.	The entire system work as a cohesive unit
Hypervisor allow to run multiple OS simultaneously on host computer.	
Each guest server run on its own operating system	
Few implementations: Oracle's VirtualBox, VMware server, Microsoft Virtual PC	
Advantages	Advantages
This type of virtualization provides best isolation and security for Virtual machine.	As a guest OS can directly communicate with hypervisor
Truly isolated multiple guest OS can run simultaneously on same hardware.	This is efficient virtualization
It's only option that requires no hardware assist or OS assist to virtualize sensitive and privileged instructions.	Allow users to make use of new or modified device drivers
Limitations	Limitations

full virtualization is usually bit slower, because of all emulation. hypervisor contain the device driver and it might be difficult for new device drivers to be installer by users.	Para virtualization requires the guest OS to be modified in order to interact with para virtualization interfaces. It requires significant support and maintainability issues in production environment.

9. Week 7- Virtual Machines and Virtualization of Datacenters – power point slide -page 41

- Virtual memory virtualization is similar to the virtual memory support provided by modern operating systems. In a traditional execution environment, the operating system maintains mappings of virtual memory to machine memory using page tables, which is a one-stage mapping from virtual memory to machine memory. All modern x86 CPUs include a memory management unit (MMU) and a translation lookaside buffer (TLB) to optimize virtual memory performance. However, in a virtual execution environment, virtual memory virtualization involves sharing the physical system memory in RAM and dynamically allocating it to the physical memory of the VMs.
- That means a two-stage mapping process should be maintained by the guest OS and the VMM, respectively: virtual memory to physical memory and physical memory to machine memory. Furthermore, MMU virtualization should be supported, which is transparent to the guest OS. The guest OS continues to control the mapping of virtual addresses to the physical memory addresses of VMs. But the guest OS cannot directly access the actual machine memory. The VMM is responsible for mapping the guest physical memory to the actual machine memory.



10. Week 8- Cloud Computing Platforms – power point slide -page 17 – table 4.2

Yes, the scenario can be deployed. Microsoft Azure is a cloud platform provider that offer a variety of tools and software to develop and deploy any scenario. Its also provide an un restricted model of programming as well as enterprise web application for users.