College Dublin
Computing • IT • Business

Cloud Computing Fundamentals & Platforms

Data Centre Research

Continuous Assessment 1

Programme:	Bachelor of Science in Computing and Information Technology
Lecturer:	Mr. Michael Weiss
Student:	Leisly Alitzel Pino Duran
Number Student:	2020303

BLADE SERVER RESEARCH CONDUCTED FOR PRIVATE ON-PREMISE NETWORK

Network Equipment Item:	Description:	Price per item:	Price per 5:
	PowerEdge MX740C - Full Configuration Base: PowerEdge MX740C Server Trusted Platform Module: Trusted Platform Module 2.0 Shipping: Shipping Material, Individual Blade, PowerEdge MX740C, EMEA1 Systems Docs Chassis: 2.5" Chassis with up to 6 SAS/SATA/NVMe Hard Drives MLK Memory: 32GB RDIMM, 3200MT/s, Dual Rank, 16Gb BASE x8	€50,828.74	€254,143.70
	Intel® Xeon® Gold 6148 2.4G,20C/40T,10.4GT/s, 27M Cache,Turbo,HT (150W) DDR4-2666	Included in price	

Hard diks	1. 92TB SSD SATA Read Intensive 6Gbps 512 2.5in Hot-plug AG Drive, 1 DWPD	€1,758.90	
	Windows Server® 2019 Standard,16CORE,FI,No Med,No CAL, Multi Language	Included in price	
	C4, RAID 5 for 3 or more HDDs or SSDs (Matching Type/Speed/Capacity)	Included in price	
Warranty	Prosupport Plus and 4Hr Mission Critical, 36 Month(s)	Included in price	

Purchase specifications and price, in Dell Products:: https://www.dell.com/en-ie/work/shop/servers-storage-and-networking/poweredge-mx740c-full-configuration/spd/poweredge-mx740c/emea_mx740c_vi_vp?view=configurations&configurationid=09703e45-93e2-44f7-95c0-e25f636415f7#features_section

INSIDE A MODERN DATA CENTRE

A data center, we know it by the place where we store and secure our information in the cloud, but this transcends more than a concept. A data center has a physical place where by security protocols, it is only accessible to authorized persons. Physical security begins at the perimeter of the facility, with security access gates and fences. As we get closer to the buildings inside, there are multiple layers of additional physical security. There are even biometric scanning systems so that only authorized employees can access certain areas. Microsoft is one of the companies that has a large data center for Azure Cloud services, which is a network of more than 60 data center regions, spread over 34 countries. That is, more than 200 data centers. The choice of the location of the data centers responds to several criteria. One of them is the proximity to population centers and because there is a market demand, as well as the location having multiple high-capacity network connections and that there is an important qualified workforce. Within these, there are server rooms, scattered around racks or closets. For standard Azure storage, the company employs a high-capacity connection blade that supports 88 three-and-a-half-inch hard drives logically divided into 422 hard drive segments. Meanwhile, for Cloud workloads, it uses other types of hardware, with more memory and storage. They are multipurpose Gen7 servers have 64 cores with up to 768 GB of RAM, 7 TB of SSD storage and 50 GB per second network cards equipped with an FPGA. These equipment are specially designed by providers for them. Microsoft has an environmental commitment to have energy efficiency and in the future to achieve renewable energy. It also has an efficient use of water, especially when it comes to cooling the servers. There are three types of service that Cloud Computing handles, which are:

• IaaS (Infrastructure as a Service): This hardware allows you to choose the capacity of processors, RAM and hard disk storage. As well as virtualization services such as virtual machines, firewalls, backup systems or load balancers. Clients can run their business applications on this infrastructure. Microsoft Azure stands out for the large number of resources that can be contracted, which makes it ideal for companies with very demanding needs in terms of infrastructure.

• PaaS (Platform as a Service): It offers platforms as services and applications such as databases, middleware, development tools, intelligence services can be launched. This type of

service is for developers who just want to focus on implementing and managing their applications.

• SaaS (Software as a Service): This software-as-infrastructure model hosts the company's software, as well as its data, on external servers, and pays a fee for its use. Any employee of a company will be able to access the company's applications from anywhere without having to install them on a local computer. All users access SaaS applications daily since it can be said that, any online application that runs in the cloud and that can be accessed through the internet, is a SaaS application.

Depending on the needs of the companies, this is how an IaaS, PaaS or SaaS model can be adapted. The fundamental difference between these three types of platforms is given by the type of maintenance and support offered by the service provider. In an IaaS platform, the user has access to the installed software and its configuration, something that users of a Saas service do not have. On the other hand, a PaaS service only has access to the software that makes up the development platform to be used.

In terms of safety there are also notable differences. In an IaaS environment, the security of installed applications rests with the user, who must be aware of which version to install and that it remains up-to-date. But in a SaaS or PaaS environment, the security of the installed applications rests with the service provider, who is responsible for their installation and updating.

Cloud Computing services are used daily by millions of users and more and more companies want not only their data in the cloud, but their entire IT environment.

CLOUD MIGRATION STRATEGY

Migrating operating system instances across distinct physical hosts is a useful tool for administrators of data centers and clusters: It allows a clean separation between hardware and software, and facilitates fault management, load balancing, and low-level system maintenance. The migration is used to distribute virtual machines load efficiently across servers in a cloud and system maintenance. Google has this type of service on Google Cloud Platform and has the following advantages:

- Get more performance, scalability and security.
- Facilitates proactive resource management and maintenance.
- Billing per second of use, even as needs increase.
- Custom machines and automatic size recommendations.
- Data is encrypted both in transit and at rest.
- Ability to resize discs on the fly.

There are different methods and types of migration, the live migration of virtual machines is one of them and allows a separation of concerns between the users and operator of a data center or cluster. It is working for various reasons such as evacuating a host machine for maintenance, balancing the workloads on host machines, optimizing physical resource utilization, and meeting the custom demands of the user applications.

Live operating system migration is a extremelely powerful tool for cluster administrators, allowing separation of hardware and software considerations, and consolidating clustered hardware into a single coherent management domain. If a physical machine needs to be removed from service an administrator may migrate operating system instances including the applications that they are running to alternative machine, freeing the original machine for maintenance.

In order to comply with the above benefits, this migration is carried out in phases. In this way, we will have a sequence of steps to comply with, which it will help to optimize time and it will give efficiency and effectiveness in the process:

Pre-Migration: We begin with an active VM on physical host A. To speed any future migration, a target host may be preselected where the resources required to receive migration will be guaranteed.

Reservation: A request is issued to migrate an OS from host A to host B. We initially confirm that the necessary resources are available on B and reserve a VM container of that size. Failure to secure resources here means that the VM simply continues to run on A unaffected.

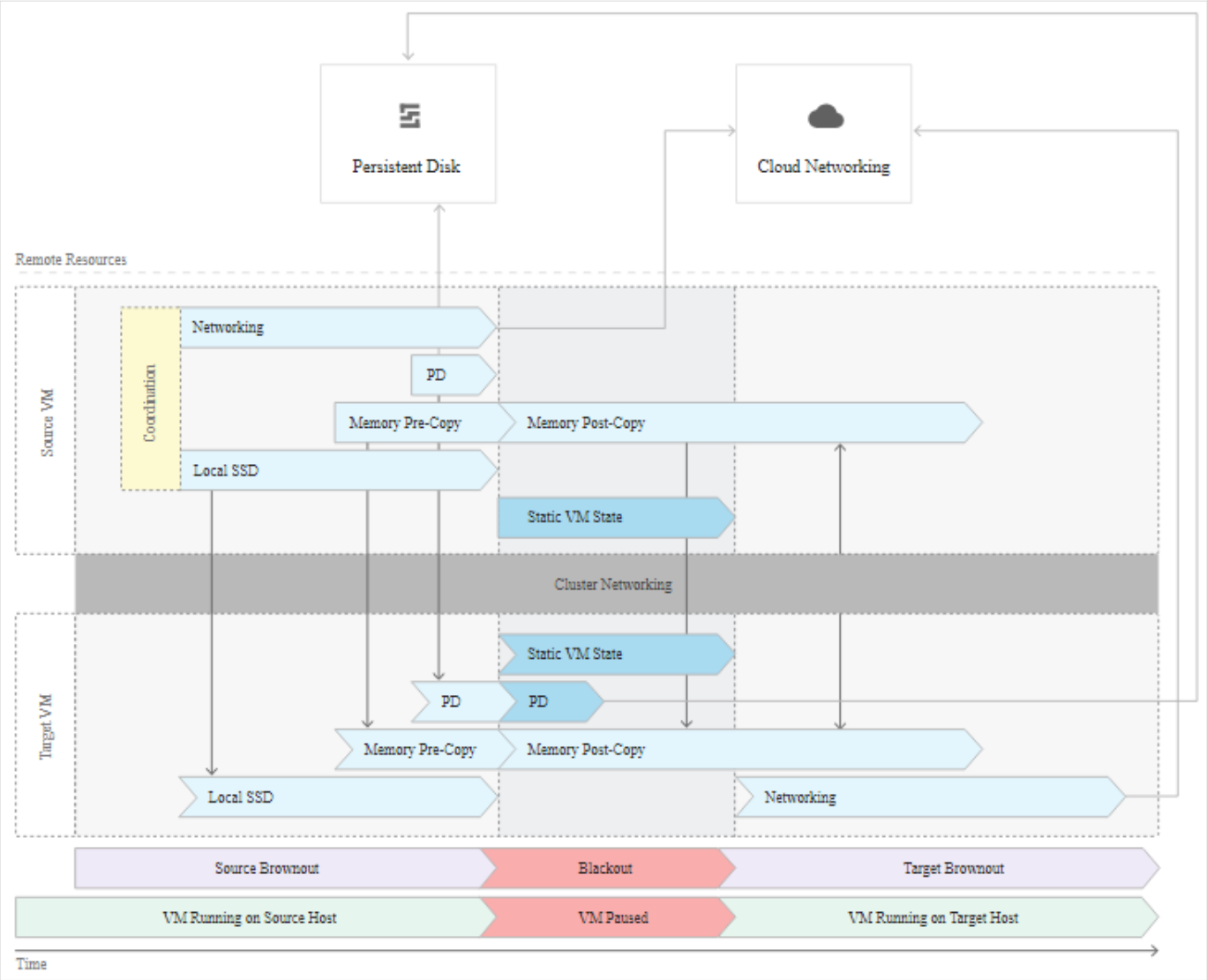
Iterative Pre-Copy: During the first iteration, all pages are transferred from A to B. Subsequent iterations copy only those pages dirtied during the previous transfer phase.

Stop-and-Copy: We suspend the running OS instance at A and redirect its network traffic to B. As described earlier, CPU state and any remaining inconsistent memory pages are then transferred. At the end of this stage there is a consistent suspended copy of the VM at both A and B. The copy at A is still considered to be primary and is resumed in case of failure.

Post-copy:The memory contents are transferred by a background process, it has a short and constant downtime and transfers each memory page only once.

Commitment: Host B indicates to A that it has successfully received a consistent OS image. Host A acknowledges this message as commitment of the migration transaction: host A may now discard the original VM, and host B becomes the primary host.

Activation: The migrated VM on B is now activated. Post-migration code runs to reattach device drivers to the new machine and advertise moved IP addresses.



The duration of migration is a key indicator of the impact and is strongly influenced by the parameters such as the page dirty rate, the size of the writable working set, the available network bandwidth, and size of the virtual machine. In this process there are different metrics, parameters that are important and are based on the function and at the time that the process is executed:

1. Preparation Time: This is the time between initiating migration and transferring the VM's processor state to the target node, during which the VM continues to execute and dirty its memory. For pre-copy, this time includes the entire iterative memory copying phase, whereas it is negligible for post-copy.

2. Downtime: This is time during which the migrating VM's execution is stopped. At the minimum this includes the transfer of processor state. For pre-copy, this transfer also includes any

remaining dirty pages. For post-copy this includes other minimal execution state, if any, needed by the VM to start at the target.

3. Resume Time: This is the time between resuming the VM's execution at the target and the end of migration altogether, at which point all dependencies on the source must be eliminated. For pre-copy, one needs only to re-schedule the target VM and destroy the source copy. On the other hand, majority of our post-copy approach operates in this period.

4. Pages Transferred: This is the total count of memory pages transferred, including duplicates, across all of the above time periods. Pre-copy transfers most of its pages during preparation time, whereas post-copy transfers most during resume time.

5. Total Migration Time: This is the sum of all the above times from start to finish. Total time is important because it affects the release of resources on both participating nodes as well as within the VMs on both nodes. Until the completion of migration, we cannot free the source VM's memory.

6. Application Degradation: This is the extent to which migration slows down the applications executing within the VM. Pre-copy must track dirtied pages by trapping write accesses to each page, which significantly slows down write-intensive workloads. Similarly, post-copy requires the servicing of major network faults generated at the target, which also slows down VM workloads.

REFERENCES

Clark, C., Fraser, K., Hand, S., Hansen, J.G., Jul, E., Limpach, C., Pratt, I., & Warfield, A. (2005). Live migration of virtual machines. NSDI. http://cs.brown.edu/courses/cs227/archives/2012/papers/migration/p20-clark.pdf

Gilesh, M. P., Jain, S., Madhu Kumar, S. D., Jacob, L., & Bellur, U. (2019). Opportunistic live migration of virtual machines. Concurrency and Computation: Practice and Experience. doi:10.1002/cpe.5477

Google Cloud, (no date). Live migration | Compute Engine Documentation | Google Cloud [online]. Google Cloud. [Accessed November 5, 2021]. Available at: https://cloud.google.com/compute/docs/instances/live-migration

Google Cloud, (no date). Migrate VM with Migrate for Compute Engine: Introduction [online]. Google Cloud. [Accessed November 5, 2021]. Available at: https://cloud.google.com/architecture/migrating-vms-migrate-for-compute-engine-getting-started?hl=en

Google Cloud, (no date). Virtual Machine Migration | Google Cloud [online]. Google Cloud. [Accessed November 5, 2021]. Available at: https://cloud.google.com/vm-migrate?hl=en

Google Cloud Tech, (2021). What is a Data Center? [online]. Youtube. [Accessed November 5, 2021]. Available at: https://www.youtube.com/watch?v=Amow8BJm5Go

Google Workspace, (2014). Inside a Google data center [online]. Youtube. [Accessed November 5, 2021]. Available at: https://www.youtube.com/watch?v=XZmGGAbHqa0

Gopalan, K., (2008). Post-Copy Based Live Virtual Machine Migration Using Adaptive Pre-Paging and Dynamic Self-Ballooning [online]. Computer Science, SUNY Binghamton. [Accessed November 5, 2021]. Available at: https://kartikgopalan.github.io/publications/hines09postcopy.pdf

Hacking, S., & Hudzia, B. (2009). Improving the live migration process of large enterprise applications. Proceedings of the 3rd International Workshop on Virtualization Technologies in Distributed Computing - VTDC '09. doi:10.1145/1555336.1555346

Hines, Michael & Gopalan, Kartik. (2009). Post-copy based live virtual machine migration using pre-paging and dynamic self-ballooning. Proceedings of the 2009 ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments, VEE'09. 51-60. 10.1145/1508293.1508301.

Hines, M. R., Deshpande, U., & Gopalan, K. (2009). Post-copy live migration of virtual machines. ACM SIGOPS Operating Systems Review, 43(3), 14. doi:10.1145/1618525.1618528

Huang, D., Ye, D., He, Q., Chen, J., & Ye, K. (2011). Virt-LM. Proceeding of the Second Joint WOSP/SIPEW International Conference on Performance Engineering - ICPE '11. doi:10.1145/1958746.1958790

IBM, (2018). IBM cloud learn hub [online]. IBM - Deutschland | IBM. [Accessed November 5, 2021]. Available at: https://www.ibm.com/pe-es/cloud/learn/iaas-paas-saas

Jo, C., Cho, Y., & Egger, B. (2017). A machine learning approach to live migration modeling. Proceedings of the 2017 Symposium on Cloud Computing - SoCC '17. doi:10.1145/3127479.3129262

Juniper Networks, (2018). Global Leader in AI Networking, Cloud, and Connected Security Solutions. [Accessed November 5, 2021]. Available at: https://www.juniper.net/assets/us/en/local/pdf/datasheets/1000439-en.pdf

Microsoft, (2018). Project Denali to define flexible SSDs for cloud-scale applications [online]. Cloud Computing Services | Microsoft Azure. [Accessed November 5, 2021]. Available at: https://azure.microsoft.com/en-us/blog/project-denali-to-define-flexible-ssds-for-cloud-scale-applications/

Microsoft, (2020). We Live in the Cloud | Microsoft Story Labs [online]. Stories | Microsoft news, features, events, and press materials. [Accessed November 5, 2021]. Available at: https://news.microsoft.com/stories/microsoft-datacenter-tour/

Microsoft Azure, (2021). Difference between IAAS, PAAS and SAAS offerings on Microsoft Azure [online]. Youtube. [Accessed November 5, 2021]. Available at: https://www.youtube.com/watch?v=WAqVsb7QhXw

Mythili Vutukuru, (2021). Virtualization and Cloud Computing Lecture 9: VM Live Migration [online]. Youtube. [Accessed November 5, 2021]. Available at: https://www.youtube.com/watch?v=npT3smBqgAc

Nathan, S., Bellur, U., & Kulkarni, P. (2015). Towards a comprehensive performance model of virtual machine live migration. Proceedings of the Sixth ACM Symposium on Cloud Computing - SoCC '15. doi:10.1145/2806777.2806838

OpenStack Docs: Configure live migrations [online]. (without date). OpenStack Docs: Xena. [Accessed November 5, 2021]. Available at: https://docs.openstack.org/nova/pike/admin/configuring-migrations.html#configuring-migrations-xenserver-shared-storage

Techopedia, (2015). What is Live Migration? - Definition from Techopedia [online]. Techopedia.com. [Accessed November 5, 2021]. Available at: https://www.techopedia.com/definition/16813/live-migration

Tech Vision, (2020). Inside The World's Largest Data Center [online]. Youtube. [Accessed November 5, 2021]. Available at: https://www.youtube.com/watch?v=g7JaN3rTK2A