Distributed Computing in a Hybrid Cloud Environment

Hitarth Himanshu Shah

Abstract— Hybrid clouds have been at the forefront of cloud computing providing better flexibility and interoperability, allowing workloads to move between on-premise servers, private and public clouds. They provide a large number of data deployment options for organisations, helping them make an informed decision on the basis of their requirements ranging from latency and security to control and costs. The general trend for most companies using a hybrid strategy is to use the public cloud for their test and development applications, email, CRM and their private cloud for workloads that tend to comprise of mission critical or sensitive items, plus heavy data analysis applications. However, as is true with any distributed environment, the hybrid cloud environment has its own fair share of challenges that affect availability, security and durability. This paper is an attempt to discuss the migration considerations to the cloud, the challenges faced in a hybrid environment and the solutions that have been made available by the emerging cloud computing services like Amazon Web Services (AWS) and other major providers.

Index Terms— Cloud Computing, Distributed Computing, Amazon Web Services (AWS), Microsoft Azure, VPC, VPN, Scalability.

1 Introduction

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[1]. The use of distributed models to enable lower latency and provide better performance for cloud services is common to various cloud providers. Hybrid clouds are created by combining both, public and private infrastructures, allowing companies and institutions to manage their computing infrastructures in flexible ways and to dynamically take advantage of externally provided resources. The real advantage of hybrid cloud lies in the ability to gain a comfortable and flexible middle ground to ensure all your workloads are treated appropriately and cost-effectively. Considering the growing needs for large computation power and the availability of a growing number of clouds distributed over the Internet and across the globe, this paper addresses the issues encountered in distributed computing in a hybrid cloud environment and how we can leverage the leading container and cloud computing services to effectively solve these issues with minimal effort.

2 The shift from Public to Hybrid Cloud

In late 2010, organisations expressed concerns over the security of public clouds leading to a rise in the popularity of private clouds. Since then, there have often been debates as to how to strike a balance between the public and the private paradigms, eventually leading to the concept of hybrid clouds. Organisations must take into consideration the following factors to determine the viability of hybrid clouds for their operation.

Fit-for-purpose: Employing hybrid clouds facilitates organisations to store their most sensitive data on dedicated hardware while simultaneously benefiting from the public cloud's cost efficiency and on-demand scalability.

Increased Speed to Market: For many organisations with an IT infrastructure that is working near or close to capacity, launching new environments can become a challenge leading to multiple delays that potentially hinder business. Hybrid clouds have allowed resources to be deployed and commissioned in an automated process, yielding results at hugely improved speeds. Thus, companies are no longer limited by their IT footprint, enjoying a lesser time-to-market for their products.

Improved Security Capabilities: In public clouds, data is not stored entirely in a private location across private systems but rather in shared servers accessed via the internet. Hence, there certainly is an argument to be made that network security is compromised in public cloud environments. Hybrid clouds facilitate placing your sensitive customer data on a dedicated server while running your front-end applications in the public cloud — creating a seamless, agile and secure environment.

Cost Benefits: Hybrid clouds reduce an organisation's total Cost of Ownership, match their cost-patterns to their revenue and demand patterns and transition from a capital-intensive cost model to an OpEx-based model.

Business Continuity: A business continuity solution isn't merely backing up and/or replicating content to the cloud, nor is it simply a Disaster Recovery plan. Business continuity refers to the ability of an organisation to continue to do

business during a failure or disaster. In other words, whenever a failure occurs, the operations of the organisation should resume with a minimum downtime. Such a solution therefore needs to be planned to consider key elements such as resilience, recovery and contingency. Where critical data is replicated to a cloud solution in a different location to the primary systems, hybrid cloud solutions are often considered by organisations as a key component of a business continuity solution.

3 Challenges in Distributed Cloud Computing

The primary challenge in hybrid cloud computing is providing a seamless operation across platforms, using cloud application programming interfaces (APIs) and hypervisors thereby making the difference between the public and the private cloud undecipherable to the user. Data center tools are usually preferred to manage hybrid cloud environments. Ideally, there should be provisions to create applications, or move existing applications between the clouds in a hybrid cloud environment, without having to change anything fundamental like networking, security policies, operational processes or management/monitoring tools. This is a problem because, due to issues around interoperability, mobility and differing APIs, tools, policies and processes, hybrid clouds generally increase complexity.

Integration complexity*:* A successful integration of private and public clouds is one of the biggest issues faced by enterprises today. Such issues need to be solved effectively to aid hybrid cloud infrastructure and its effective performance. The technical skills required to integrate multiple cloud services is quite high. While transferring regular applications between clouds is quite easy, moving configurations and Meta data across environments is more difficult, involving higher degrees of complexity.
Another major issue affecting integration is comparability. The feasibility of the hybrid cloud will be determined by the nature of the workload, its size and complexity. The most challenging task, where even the most seasoned technicians fail is to understand the patterns and tools required to move processes. Further, creating and managing infrastructure to handle legacy applications and their migration to the cloud is extremely complex and has hundreds of applications involved.

Networking: Any hybrid cloud computing model that seeks to have long-term success must have an impeccable network design. Implementing hybrid cloud infrastructure resembles opening a new data center. However, instead of opting for multiple internet service providers and network equipment and hardware of your choice, hybrid cloud requires relying on a cloud provider. Because of this restriction, make sure that your selected cloud provider covers the following three aspects. First is the ability to design and implement virtual private cloud networks (VPC) with subnets with a defined range of IP addresses. The second is the ability to set up virtual private network (VPN) connection options compatible with available hardware or software. All in all, network design must take into consideration various factors such as network bandwidth, management between public and private clouds, the impact of location on your network, the network requirements for each individual application, the security requirements of different types of data, and numerous other factors.
The final aspect in case low network latency is needed or to avoid public internet traffic, consider the ability to establish a direct connection between your private data center and VPC (as is provided by cloud computing leaders AWS, Azure and Google). Also, establishing connectivity between your data center on premises and a cloud in your network infrastructure would be completely your responsibility, so plan it in advance.
A network topology must be planned, factoring in its security and the potential latency between the different layers, both between internal and external resources. Mature cloud providers, such as Amazon and Microsoft, already recognize the importance of connectivity and have developed their own hybrid cloud network solutions: DirectConnect and ExpressRoute, respectively.

Migration*:* The process of migration usually takes place in stages. The two typical stages are server migration and data migration. The first stage, server migration refers to moving a fleet of servers from a private data center to the cloud. The main challenge encountered here is the choice of an implementation strategy that availability of resources, expertise, services modernization and time to market. The second stage, data migration, refers to the process of moving data stored in your file system on the database servers to the cloud. Two typical challenges involved are the size of the data volume and the ability to move it without service interruption.

Compliance: The issue of compliance can be viewed from two perspectives, the legal and the technical. To ensure legal compliance, the cloud provider must have all the required certifications to comply with policies and governmental regulations in a target state or country. To ensure technical compliance however, your cloud provider is in charge of hardware, while you are responsible for software configuration and applications security. Technical compliance differs from one cloud provider to the next. Make sure you have enough tools to implement a secure and compliant solution according to your requirements and needs.

4 Solutions For implementing Distributed Hybrid Cloud

Even though implementation of a hybrid cloud might seem rather challenging, numerous cloud providers offer solutions that are capable of overcoming the mentioned issues and ensure successful hybrid cloud implementation. Determine your priorities and needs, build a clear vision of what to expect from a cloud provider and get ready for the benefits of

the hybrid cloud. The names approaches mentioned can differ from one cloud provider to the next, however, the underlying service offered still remains the same. Below, we study the approaches listed by Amazon Web Services, which might be known by different names under Microsoft Azure and the Google Cloud platform.

Amazon VPN

A virtual private network enables the ability to extend a subnet from one geographic location to another geographic location on two seperate networks. This is essentially extending the on-premise network to the cloud, or the cloud to the on-premise network. For AWS, this allows us to communicate with all resources (example EC2 instances) internally without the need for public IP addresses and an internet gateway. The VPN connection has two parallel routes or IPSec tunnels, which is for redundancy. The whole system of the VPN can be set up with three elements, namely the customer gateway, the virtual private gateway and the VPN connection itself.

Customer Gateway: The customer gateway could be a physical device or a software application that forms the part of the VPN connection on the on-premise side.

Virtual Private Gateway: The VPG forms the part of the VPN connection on the Virtual Private Cloud side that has been hosted on the AWS account.

The actual VPN connection that is set up and managed by AWS offers a perfect solution that effectively connects the VPC to the on premise servers.

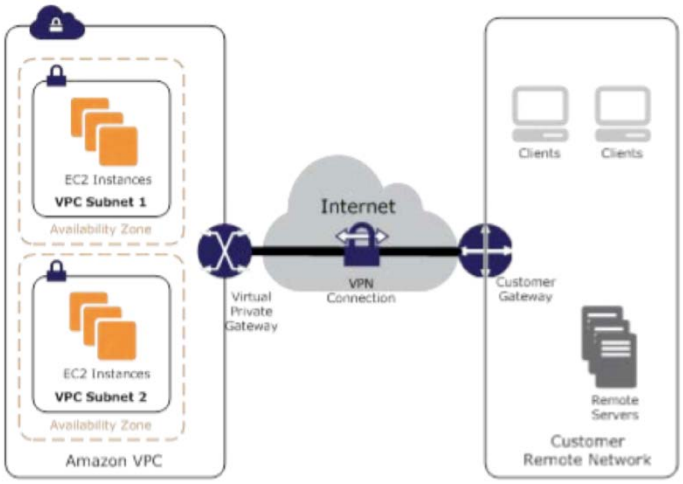
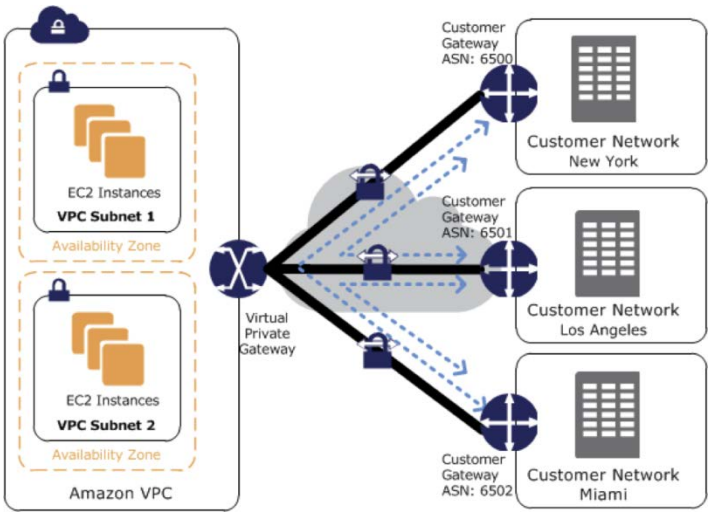


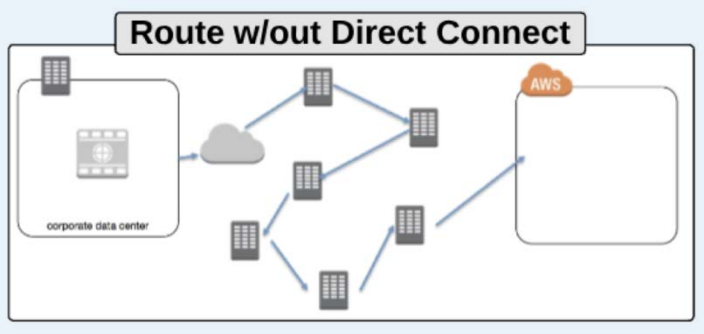
Figure 1: Hardware VPN

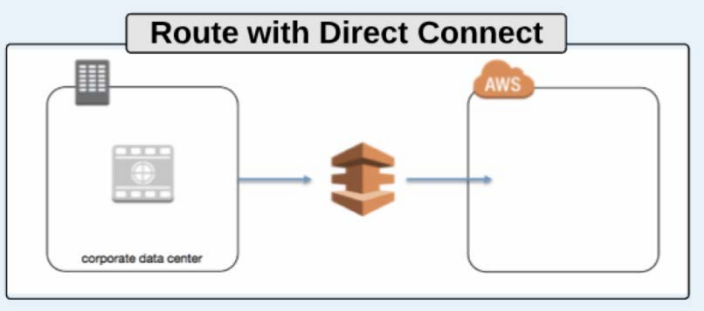


AWS Direct Connect

The Direct Connect service by AWS provides a dedicated network connection between your own network and one of AWS's own Direct Connect locations. This is done via a Direct Connect Provider which is usually Verizon or other ISPs. Direct Connect has the following advantages over standard VPN connections.
Reduce network costs: Reduce bandwidth commitment to corporate ISP over public internet. • Data transferred over direct connect is billed at a lower rate by Amazon (data in/out).
Increase network consistency: Dedicated private connections reduce latency (over sending the traffic via public routing).
Dedicated private network connection to on-premise: Connect the direct connect connection to a VGW in your VPC for a dedicated private connection from on-premise to VPC and use Multiple VIF (Virtual Interfaces) to connect to multiple VPCs.

The following comparison gives us a better idea of the advantages of Direct Connect approach over the VPN approach.





6 Conclusion

To conclude, a hybrid cloud model must be implemented with careful planning, estimation, and understanding of all the associated parameters that have been discussed in this paper. By bridging the gap between the current infrastructure solutions and future-proofed infrastructure technology, organizations can be set up for a more stable as well as innovative future. Owing to its nature, there is no one-size-fits-all solution for hybrid cloud. Due diligence must be exercised by cloud architects and customers alike before deploying a suitable hybrid cloud infrastructure for their business. Proper planning is required to develop a hybrid cloud deployment model. Identifying and selecting the right resource model at the beginning itself is important as course correction at later stage could be disruptive for systems with such complexities.

7.2 Acknowledgments

The work on this paper was carried out under the guidance of Ms. Vijaya Padmadas, Associate Professor at the Computer Engineering Department of Thadomal Shahani Engineering College.

References

[1] Pierre Riteau,Building Dynamic Computing Infrastructures over Distributed Clouds.
[2] L. Wang and G. von Laszewski, Scientific Cloud Computing: Early Definition and Experience," in 10th IEEE International Conference on High Performance Computing and Communications, HPCC '08. , pp.825-830,2008
[3] M. Vouk, Cloud computing- Issues, Research and Implementations. Journal of Computing and Information Technology (CIT) Volume 16, Number4, pp. 235–246, 2008
[4] Q. Zhang, L. Cheng and R. Boutaba, Cloud computing : state of the art and research challenges. Journal of Internet Services and Applications, vol. 1, pp. 7–18, 2010.
[5] Foster, Y. Zhao, I. Raicu and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared," Grid Computing Environments Workshop (GCE '08), 2008.
[6] https://docs.aws.amazon.com/vpc/latest/userguide/vpn-connections.html
[7] https://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html
[8] R.Buyya,R.RanjanandR.N.Calheiros,Modelingandsimulationofscalable Cloudcomputing environments and the CloudSim toolkit: Challenges and opportunities. Proceedings of the Conference on High Performance Computing and Simulation (HPCS 2009), Leipzig, Germany. IEEE Press: New York, U.S.A., 21–24 June 2009; 1–11.
[9] Taha Chaabouni and Maher Khemakhem, CLOUD COMPUTING: A NEW VISION OF THE DISTRIBUTED SYSTEM
[10] T. Dillon, C. Wu and E. Chang, Cloud Computing: Issues and Challenges, in 24th IEEE International Conference on Advanced Information Networking and Applications, pp. 27-33, 2010
[11] M. Munetomo, "An investigation on large-scale distributed design opti- mization framework in the inter-cloud environment (in Japanese)," *IPSJ SIG Technical Report*, vol. 2014-MPS-98, no. 28, pp. 1–2, 2014.
[12] M. Tsugawa, P. Riteau, A. Matsunaga, and J. Fortes. User- level Virtual Networking Mechanisms to Support Virtual Machine Migration Over Multiple Clouds. In *The 2nd IEEE International Workshop on Management of Emerging Networks and Services (IEEE MENS 2010)*, 12 2010.