Availability Considerations for Mission Critical Applications in the Cloud

Valentina Salapura and Ruchi Mahindru
IBM T.J. Watson Research Center, 1101 Kitchawan Rd., Yorktown Heights, NY, U.S.A.

Keywords: Enterprise Class Applications, HA Clusters, ERP Cloud Solutions.

Abstract: Cloud environments offer flexibility, elasticity, and low cost compute infrastructure. Enterprise-level workloads – such as SAP and Oracle workloads - require infrastructure with high availability, clustering, or physical server appliances. These features are often not part of a typical cloud offering, and as a result, businesses are forced to run enterprise workloads in their legacy environments. To enable enterprise customers to use these workloads in a cloud, we enabled a large number of SAP and Oracle workloads in the IBM Cloud Managed Services (CMS) for both virtualized and non-virtualized cloud environments. In this paper, we discuss the challenges in enabling enterprise class applications in the cloud based on our experience on providing a diverse set of platforms implemented in the IBM CMS offering.

1 INTRODUCTION

Cloud computing is becoming the new de facto environment for many system deployments in a quest for more agile on-demand computing with lower total cost of ownership. Cloud computing is being rapidly adopted across the IT (information technology) industry to reduce the total cost of ownership of increasingly more demanding workloads. Various companies and institutions are adopting cloud computing, bringing high expectations of resiliency that have heretofore been associated with dedicated data centers.

Flexibility and elasticity are one of the most important advantages of cloud computing – compute resources are rapidly provisioned on demand. Native cloud applications are designed to tolerate failure, and to minimize state. On the other hand, enterprise-level workloads require High Availability (HA), continuous operation, and long lived virtual machines (VMs).

Enterprises demand usage of enterprise-level workloads - such as Systems Applications and Products (SAP) (Boeder and Groene, 2014) and Oracle (Oracle Corp., 2014), which are becoming a benchmark for running business back-office operations. These applications require an infrastructure with high availability, clustering, shared storage, or physical server appliances. To fulfil such requirements, cloud infrastructure needs to offer features such as high availability clusters, anti-collocation, or shared storage required by enterprise workloads. Anti-collocation requirement could be achieved by using different availability –collocation.

Customers are looking for a common environment to host their virtualized and non-virtualized workloads in an integrated manner. For example, both SAP and some Oracle applications may run on VMs while requiring databases to run on a specialized physical server appliance, or on VMs with larger resources.

IBM Cloud Managed Services (CMS) (IBM Corp., 2014) is a premier cloud offering which provides a unique mix of virtualized and non-virtualized cloud environments for enterprise workloads. IBM CMS enables large installations and service level agreement (SLA) mechanisms.

To satisfy a growing need of enterprise customers for the enterprise-level workloads, we enabled a number of Oracle and SAP workloads in the IBM CMS cloud. IBM CMS cloud offers a fully managed solution for a large number of SAP and Oracle applications in a cloud environment for both virtualized and non-virtualized cloud environments. The solutions cover diverse types of platforms e.g. x86 and Power Systems (Sinharoy et al., 2015). An example of these applications is SAP High-performance Analytic Appliance (HANA) (Färber et al., 2012).

This paper describes how we provided high-availability clustering as a service, and how we integrated physical server appliances on the IBM CMS enterprise cloud.

Salapura, V. and Mahindru, R.
Availability Considerations for Mission Critical Applications in the Cloud.
In *Proceedings of the 6th International Conference on Cloud Computing and Services Science (CLOSER 2016) - Volume 2*, pages 302-307
ISBN: 978-989-758-182-3

2 IBM CMS CLOUD

Enterprise-class customers, such as banks, insurances or airlines typically require IT management services such as monitoring, patching, backup, change control, high availability and disaster recovery to support systems running complex applications with stringent IT process control and quality-of-service requirements. Such features are typically offered by IT service providers in strategic outsourcing (SO) engagements, a business model for which the provider takes over several, or all aspects of management of a customer's data center resources, software assets, and processes. Servers with such support are characterized as being managed.

This should be contrasted with unmanaged servers provisioned using basic Amazon Web Services (AWS) and IBM's SoftLayer offerings, where the cloud provider offers automated server provisioning. In order to make the server managed, these cloud providers have networked with service partners that customers can engage to fill all of the gaps up and down the stack. This enables the user to add services to the provisioned server, but the cloud provider assumes no responsibility for their upkeep or the additional services. Therefore, it puts burden on the customer to obtain a fully managed solution for their enterprise workload rather than the cloud service providing an end-to-end fully managed solution for the customer.

The IBM's CMS is among a small set of industry cloud offerings that support managed virtual and physical servers. It is an enterprise cloud, which provides a large number of managed services that are on par with the ones offered in high end SO contracts. Examples of such services are patching, monitoring, asset management, change and configuration management, quality assurance, compliance, health-checking, anti-virus, load-balancing, security, firewall, resiliency, disaster recovery, and backup. The current product offers a set of managed services preloaded on users' servers in the cloud. The installation, configuration, and run-time management of these services are automated.

3 POSITION: MISSION CRITICAL WORKLOADS REQUIRE ENTERPRISE DATA CENTER RESILIENCE

The main attributes of cloud computing are scalable, shared, on-demand computing resources delivered over the network, and pay-per-use pricing. Typically, one thinks of cloud as on-demand environments which are created and destroyed as needed. This offers flexibility in using as few or as many IT resources as needed at any point in time. Thus, the users do not need to predict resources that they might need in future, which makes cloud infrastructure attractive for businesses.

Cloud native applications take advantage of the cloud's elasticity, and are written in a way to run the application on multiple nodes. The nodes are stateless, and as such tolerate loss of any single node without bringing down the entire application.

On the contrary, enterprise customers require computing infrastructure which is set up infrequently, but is available over a much longer time frame. For example, a database is expected to run continuously, and not to lose any data in the case of infrastructure failure. No response from a database even over a short period of time can result in large business losses for an enterprise.

High availability is an important requirement for running enterprise-level applications. Features like standardized infrastructure, virtualization, and modularity capabilities of cloud computing offer an opportunity to provide highly resilient and highly available systems. Resiliency techniques can be deployed on a well-defined framework for providing recovery measures for replicating unresponsive services, and recovering the failed services.

To achieve application resiliency, high availability clusters are used. Implementing HA clusters requires features such as anti-collocation of VMs – locating VMs on different physical hosts, a requirement which is difficult to guarantee in a cloud environment. For example, VMs are created on physical servers based on hypervisors utilization to achieve balanced and optimally utilized compute environment. Additionally, VMs could migrate between hypervisors for either load balancing or maintenance.

The location of physical servers hosting VMs determines the network latency between the nodes. The latency between the nodes depends on the location of physical servers in a data center – for example, whether the nodes are located in the same row – or on the current network traffic in a data center. For example, ongoing data backup traffic can impact network latency when accessing a DB. Additionally, multiple VMs might need to access the same DB data, and require implementation of a shared storage, a feature which is not typically part of a cloud offering. These cloud properties make implementing resiliency features for enterprise workloads more complicated.

Recently, cloud providers started to support some of these requirements. AWS provides the IT resources so that the customers can launch entire SAP enterprise software stacks on the AWS Cloud. Anti-collocation requirement could be achieved by using different availability zones for VMs (Amazon Corp., 2015).

In addition, there are certain proprietary workloads that are not allowed to run on virtual environment or cannot be supported on the state of the art hypervisors in a cloud environment. Some applications are not certified to run on virtualized servers (e.g. analytic appliances), or would require significant increase in licensing cost if deployed in a cloud environment. Therefore, it is essential to deploy fully managed appliances on physical servers and connect them to the cloud internal network to support applications that cannot be hosted on the cloud, but need to be close to the cloud. A few examples of such applications are SAP Business Warehouse Accelerator (BWA), or Oracle Database Appliances.

Customers owning such applications need an integrated solution which would allow them to use these applications together with the cloud hosted workload. These applications need to run on a physical server which is fully integrated into the management environment of the cloud providing services such as monitoring or backup. It also avoids hybrid solutions where a part of the workload is running in the cloud and the other part running in the non-cloud environment. Such solutions are hard to manage due to different delivery and operation models. Examples of such solutions are HANA appliances and Oracle OVM based systems, which need to be operated in a tight connection to other servers.

4 POSITION: RESILIENCY IN THE CLOUD REQUIRES NEW CAPABILITIES

There are several challenges that have to be considered when providing high availability in the cloud. To implement high availability clustering, high availability software is used. It arranges redundant nodes (two or more OS instances) in clusters to provide continued service in the case of a component failure. OS instances can be accessed by using the same virtual Service Internet Protocol (IP) address. An HA cluster detects hardware or software faults, and performs a failover – it restarts the application automatically on another OS instance. As part of this process, clustering software may configure the nodes to use the appropriate file system, network configuration, and some supporting applications. HA clusters are typically used for critical databases, business applications, and customer services.

IBM CMS cloud provides all infrastructure components needed to create an HA cluster. To support HA clusters for VMs, the virtual infrastructure must provide several important features. First, it must have the capability to anti-collocate the cluster members, that is, to ensure that they are never located on the same physical server during the cluster's entire lifecycle. This, in turn, imposes constraints on the placement algorithms of the virtualization system. Other resiliency scenarios require that the cluster members are in the different building blocks, or even different sites (data centers in different geographical areas).

The environment must also allow shared disks – to allow multiple VMs to concurrently connect to, and share the same physical storage. To avoid a single point of failure, a number of shared disks are arranged in a redundant array of independent disks (RAID). The VMs with access to shared storage should also have one or more private disks for OS image, application, and log files.

One or more virtual Internet Protocol (IP) addresses have to be reserved and assigned to the cluster. The exact usage of vIPs depends on the used HA configuration, if it is arranged as active-passive or active-active cluster. In all configurations, the end user does not see one or the other individual VMs of the cluster, but only the application running in the cluster as available, and accessible via its service IPs. Finally, the HA nodes also have their own IP addresses which are used by an administrator to access individual VMs to set up its configuration.

This HA cluster infrastructure, together with HA clustering software, enables large number of different configurations for high level availability, such as active-passive or active-active configurations. HA clustering software, such as Power High Availability (PowerHA) (Bodily et al., 2009), is installed on top of the HA cluster infrastructure. The HA clustering software provides a heartbeat function, which enables a node to have an awareness of the state of the other nodes in the cluster.

Generally, an HA solution would require a dual-room set up requiring the hardware to be deployed in different buildings and at least 10 km apart. Such HA solution may not be available. Multiple power supplies and multiple networks can be deployed in the same building to provide resiliency. A destruction of an entire room or building is considered a disaster, and a distance above 80 km between the primary and secondary servers in the two datacenters would provide a disaster resilient solution.

5 POSITION: HIGH AVAILABILITY IN THE CLOUD REQUIRES MODIFICATION TO RESOURCE PROVISIONING

In CMS, we implemented cluster support for both Power Systems, and for x86-based virtual systems. To implement HA clusters, we introduced a notion of a two node cluster in the CMS provisioning and management system. The first created VM is denoted 'anchor VM', and a cluster ID is created and assigned to it. The VM is provisioned in the same way as a non-clustered VM. The provisioning system determines the target physical server for any VM at the time of its creation based on the overall utilization and workload distribution of all servers in a data center.

The second VM of a cluster is labelled 'dependent', and it is tagged with the cluster ID of the anchor. Once a dependent VM is provisioned, the system ensures that it is not located on the same physical server as the anchor node. If that is the case, the management system moves the dependent node to a different server, and completes the cluster creation. To fulfil the requirements of HA clustering described above, we extended the provisioning system to enable anti-collocation.

Implementation of a shared storage solution, where the same storage is attached and readable by two VMs, represented a challenge. The storage has to be made available to both nodes via the network, and read/write permissions to the storage need to be defined for both VMs, including conflict resolution, or conflict avoidance.

During provisioning, all requested shared storage is allocated and linked to the anchor VM, together with its private storage. When provisioning a dependent VM, only private storage is allocated and linked to the dependent VM. The shared storage is already created as a part of the anchor VM, and needs to be mounted to the dependent VM. The mounting step is performed automatically after the dependent VM is provisioned. As the last step of provisioning an HA cluster infrastructure, the storage disks to be shared in the cluster are linked to the dependent VM.

The process is expandable to a number of VMs in a cluster larger than two nodes. The provisioning steps are as follows: creating an anchor VM with its all private and to be shared storage, then creating one or more dependent VMs with their private storage. As the last step, all dependent VMs are mapped to the storage shared in the HA cluster. An additional necessary step is to reserve and assign one or more virtual IP addresses to the cluster.

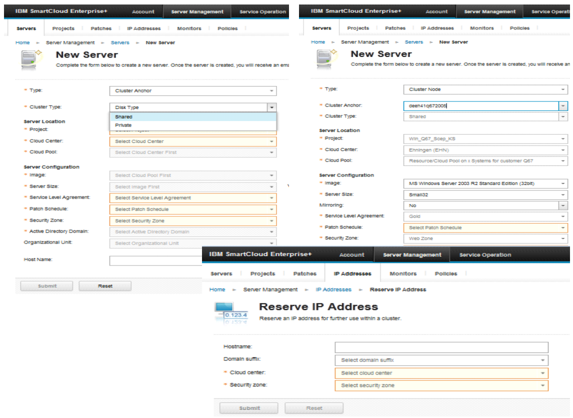

Figure 1: CMS portal for provisioning infrastructure component for an HA cluster.

Figure 1 illustrates the CMS portal when creating the HA infrastructure – requesting two nodes with a number of private and shared disks, and reserving a number of virtual IP addresses. These steps bring additional complexity into the management of the cloud system.

HA cluster nodes can be used in several different configurations. In active-passive configuration, one instance acts as the active instance, while the other one is passive and serves as its back up. Both instances have access to the shared storage. The instances are accessed by a customer via Service IP which points to the active VMs. In the case of a failure of the active VM, failover causes the passive VM to become active. Service IP now points to the second VM in the HA cluster. In this configuration, only active VM has write access to shared storage, whereas the second VM is in the stand-by mode. In the case of failover, the control is transferred to the second VM which then has write control over the shared storage.

In active-active configuration, both VMs are running the application, both are having write access to a part of shared storage (to a resource group), and both act as a backup to each other. All transactions – accesses to the application – are directed to one of the two VMs by a load balancer. In the case of failover, the second VM takes over the write control of the both shared storage resource groups.

Furthermore, as CMS is a fully integrated managed services cloud offering, several enhancements were required in order to enable managed services. For example, a new monitoring solution had to be designed and implemented to monitor HA cluster.

6 POSITION: MULTIPLE LEVELS OF RESILIENCY INCREASE SYSTEM RELIABILITY

SAP application (Boeder and Groene, 2011) typically requires a high level of the workload availability of 99.8% SLA, which defines the maximum allowed down time to less than one and a half hours per month (Schmidt, 2006). To achieve this high SLA objective, a cloud solution for SAP must support HA clusters.

SAP has many configurations, but we describe a two-node active-passive cluster configuration, as supported in IBM CMS. In a typical configuration of a two-node SAP workload, the workload resides on two VMs, both of which contain the complete SAP application stack. In active-passive configuration, at any given time, only one instance of the application is active. The other instance is in a hot standby mode ready to take over the operation if the active instance fails. Both instances have connectivity to a database residing on a number of shared storage devices arranged in RAID, but only the active instance has the read/write access.

HA cluster middleware monitors the internal health of both the applications and the virtual servers hosting them, and performs a failover from the active VM to the passive VM when a failure is detected. For Power AIX systems, CMS makes use of PowerHA (Bodily et al., 2009), and for Windows and Linux systems, CMS employs Veritas Cluster Server (VCS) and Veritas Storage Foundation (VSF) (Symantec Corp., 2009).

Two node clusters offer high availability in CMS because of the additional high availability cloud infrastructure support (Salapura et al., 2013). Failures of a VM or of the hosting physical server are handled by the infrastructure high availability support. For example, during the HA cluster failover in active-passive configuration, the failed VM is automatically restarted at the infrastructure level, either on the original server or on another server. If one VM fails, it is rebooted and the SAP application is restarted. In the case of a failure of a physical server hosting SAP, all VMs from the failed server are restarted on the surviving servers, and the SAP workload is restarted within its VM. In this way, the HA cluster is re-established within a short period of time. Without this feature, HA cluster would be lost.

Multiple levels of resiliency at infrastructure, middleware and application levels increase system reliability. Implementing multiple levels of resiliency delivers a more robust system, while enabling operation of these different levels of resiliency seamlessly.

7 POSITION: ENTERPRISE WORKLOADS REQUIRE MODIFICATION TO CLOUD INFRASTRUCTURE

Enterprise-level customers are looking for a way to operate SAP HANA appliance in the cloud. The SAP HANA appliance (Färber et al., 2012) is in-memory database that allows accelerated processing of a large amount of real-time data. The SAP HANA appliance is operating on non-virtual servers. Integration of the SAP applications running on VMs in the cloud environment with a HANA appliance running the database is business critical for the customers to ensure that they are using the state-of-the-art technology for their transactions and data analysis.

Enterprise-level customers demand a fully managed HANA appliance with managed services like patching, monitoring, health checking, auditing and compliance. HANA has a very strict set of network requirements. A fully managed HANA appliance requires several network interfaces which are used for redundant pairs of customer, management, and backup networks. In addition, a HANA solution requires internal networks for General Parallel File System (GPFS) clustering (Barkes et al., 1998), and HANA clustering and scale-out solutions.

A GPFS cluster had to be established for storage needs. HANA requires guaranteed network latency and bandwidth at any point in time, which is extremely challenging to provide in a shared cloud environment. The large number of network interfaces demands increased network switching. In addition to the switches responsible for providing customer, management and backup networks, HANA requires switches for internal GPFS and SAP HANA networks.

In IBM CMS cloud, to enable integration of SAP HANA databases running on non-virtual servers with SAP workloads running on VMs, the customer's virtual local area networks (vLANs) have to be extended to allow communication from SAP workloads on VMs with HANA databases. Therefore, there are several enablement steps that have to be considered during customer onboarding. For example, all SAP systems of a customer have to be located in the same security zone, and the HANA database appliance server is required to be in the same firewall zone as its corresponding SAP Business Warehouse (BW) Application Server.

There are various deployment modes available for HANA database, from a single node to multi node scale-out deployments. SAP HANA appliance can be a single node server, or a scale-out multi node cluster of multiple servers running one or more SAP HANA systems, depending on the level of resiliency required. The smallest configuration is MCOS (Multiple Components on One OS) and is typically used for development and test systems.

8 CONCLUSIONS

The demand of businesses to take advantage of low cost resources in the cloud, and of the high cost of running their own IT, as well as a tremendous profit opportunity motivates cloud providers to enable enterprise application. Enterprise-level workloads require high availability, clustering, or integration of physical server appliances, features which are not part of a typical cloud offering.

In this paper, we presented how we enabled enterprise-level ERP workloads in the IBM CMS cloud. We implemented several resiliency features, such as HA clustering, shared storage, private network, and physical server appliances, in the CMS cloud. Bringing enterprise level applications into the managed cloud requires enhancing or adapting infrastructure provisioning and management services to fully support it. These features enabled various enterprise applications such as SAP, SAP HANA and Oracle RAC to run in the IBM CMS cloud for both virtualized and non-virtualized environments thus allowing businesses to take advantage of the cloud's flexibility, elasticity, and low cost.

REFERENCES

Boeder, J., Groene, B., 2014. *The Architecture of SAP ERP: Understand how successful software works, 2014*.

Oracle Corp., 2014. Oracle Applications. [Online]. https://www.oracle.com/applications/index.html.

IBM Corp., 2014. Cloud Managed Services. [Online]. http://www-935.ibm.com/services/us/en/it-services/cloud-services/cloud-managed-services/index .html.

Sinharoy, B., Van Norstrand, J. A., Eickemeyer, R. J., Le, H. Q., 2015. IBM POWER8 processor core microarchitecture, *IBM Journal of Research and Development*, vol. 59, no. 1, pp. 2:1-2:21, 2015.

Amazon Corp., 2015. *Amazon Elastic File System – Shared File Storage for Amazon EC2*. [Online]. https://aws.amazon.com/blogs/aws/amazon-elastic-file-system-shared-file-storage-for-amazon-ec2/

Bodily, S., Killeen, R., Rosca, L., 2009. PowerHA for AIX cookbook. *IBM Redbook. 2009*.

Schmidt, K., 2006. High Availability and Disaster Recovery: Concepts, Design, Implementation. *Springer Science and Business Media, 2006*.

Symantec Corp., 2009. *A Veritas Storage Foundation™ and High Availability Solutions Getting Started Guide* [Online]. https://docs.oracle.com/cd/E19186-01/875-4617-10/875-4617-10.pdf.

Salapura, V., Harper, R., Viswanathan, M., 2013. Resilient cloud computing, *IBM Journal of Research and Development, vol. 57* no. 5, 2013.

Färber, F., Cha, S. K., J. Primsch, J., Bornhövd, C., Sigg, S., Lehner, W., 2012. SAP HANA database: data management for modern business applications, in *ACM Sigmod Record, vol. 40*, no. 4, pp 45-51, 2012.

Barkes, J., Barrios, M. R., Cougard F., Crumley. P. G., Marin, D., Reddy, H., Thitayanun, T., 1998. GPFS: *a parallel file system, IBM International Technical Support Organization, IBM Redbook SG24-5165-0, 1998.*