Neel Shah **CSCI-4342**
Writing 1 **09/11/2016**

Overview, Key Words, and Subtopic Name:

This product improves scalability, reliability, and security in the cloud by providing new features to a pre-existing add-on, Remus, for the Xen hypervisor. Particularly, this enhancement will provide the ability to manage and maintain state between a cluster of virtual machines (VM) upon a system or security fault and it will enable system checkpointing upon specified events. This product will be beneficial to any entity who utilizes and trusts the cloud to handle their operations.

Keywords: cloud, scalability, reliability, security, virtualization

Subtopic: Cloud reliability and security

Intellectual Merit:

The Small Business Innovation Research Phase I project will provide new methodologies to assure uptime and integrity in cloud systems. Remus already checkpoints and reverts VMs upon faults. Extending this ability such that a cluster of interdependent VMs can be managed, backed up, and restored without disrupting network and filesystem state will be a challenge. With the need to revert many VMs, the time it takes to complete one network or filesystem transaction can increase. Since we do not know where or when in the cluster a fault will occur, it will be a challenge to account for and mitigate time lost due to maintaining cluster state upon a system restore. Also, more efficient checkpointing methods need to be implemented upon the addition of system event based checkpointing.

Upon project completion, the Remus extension for the Xen hypervisor should now be able to protect a cluster of interdependent VMs. If there is a fault in one or many VMs, Remus will revert failed VMs to a known good checkpoint, switch execution to their backup VMs, and it will then assure the network and filesystem state of the remaining VMs is not corrupted after handling a change in cluster topology. Also, Remus will be able to checkpoint VMs if certain predefined events occur on a system. This will enable a more fine grained fault mitigation policy in the individual VM or the cluster of VMs.

To achieve these goals, Remus must first be modified to protect a cluster of VMs in a naïve way, where the entire cluster is rolled back upon a fault in an individual VM. Then, Remus will be extended to efficiently manage cluster state with minimal restores upon a single system fault. Furthermore, after research on the importance and meaning of many system events, Remus will be responsible for monitoring each VM in the cluster for the selected events.

Broader/Commercial Impact:

This innovation will make “the cloud” more reliable. Any entity that leverages the cloud to handle their operations trusts that their resources are always performing without faults to assure a high quality of service. In the perspective of a service provider, the fabric in which they distribute their service is critical in defining the service’s quality. These additions to Remus enables the cloud to scale more efficiently, handle faults better to become more reliable, and have greater system visibility to determine integrity.

This page defines the separation from my old senior design project idea to my new one.

Neel Shah **CSCI-4342**
Writing 1 - Redo **09/25/2016**

Overview, Key Words, and Subtopic Name:

This product accounts for the advent of Internet of Things (IoT) and cloud computing by bringing a new feature to the Xen hypervisor. This new feature emulates Intel's Rack Scale Architecture (RSA) through a software front instead of a hardware one. RSA is a model where one operating system, a virtualized one, can share the hardware resources of other machines to provide one "super" operating system. Mainly, individuals who manage data centers and cluster would benefit from this product since it enables them to design flexible computing fabrics that enable quick and dynamic server configurations to account for incoming workload to scale efficiently.

Keywords: rack scale architecture, virtualization, memory management, internet of things, cloud computing

Subtopic: Rack Scale VMs through a software fabric

Intellectual Merit:

The Small Business Innovation Research Phase I project will provide facilities for data center administrators to create flexible server architectures to account for the rapid growth of the IoT paradigm. This growth coupled with the advent of cloud computing pushes data center administrators to constantly rethink their server architecture to account for dynamic workloads. Currently, resources in a data center are locked at the individual server level. Computing workloads will change with the aforementioned trends, thus requiring resources to adapt for them. Some data centers are unable to upgrade hardware as frequently as their workloads change. This increases strain hindering the data center incapable of efficiently adapting to the new workloads. Furthermore, RSA is typically achieved by hardware solutions. Products such as remote direct memory access (RDMA) enabled network cards and switches are required to enable sharing of computing resources across hosts.

This software solution to rack space architecture through the Xen hypervisor enables data centers to bring flexibility to existing hardware. Administrators can specify one master virtual operating system to share the resources within a specified computing pool. The master operating system will use intelligent algorithms to share and balance memory throughout the pool.

Broader/Commercial Impact:

This innovation will provide facilities for commodity data centers or data centers who are unable to account for modern workloads to reliably scale. IoT and cloud computing are paradigms that are taking over the industry. A multitude of businesses are choosing the two as their main service provider. With this shift, there needs a way for the service provider backend, a data center, to reliably adapt to new workloads. RSA is the solution for this, but it is difficult to attain solely through hardware. This innovation provides the solution through a software fabric which is easy to adapt in data centers that are unable to reorganize resources or upgrade hardware.

Neel Shah **CSCI-4342**
Writing 2 **09/25/2016**

The computing industry is showing trends where developers pursue “Internet of Things” (IoT) and cloud computing. With this movement, service providers need to adapt their hardware to account for the new trend. When many businesses choose IoT devices as their main interface with customers and the cloud as their main backend service provider, they expect data centers to reliably handle all traffic related to their service.

Currently, Intel estimates that the number of IoT devices will reach upwards of 30 billion by 2020. These devices will be communicating with “the cloud” in a data center to provide a service. If data centers are unable to economically adapt to this new workload, then the business will suffer due to a poor quality of service.

Data centers need to constantly adapt to handle this new workload. In most data centers, hardware resources are pinned to individual servers. Data centers are also unable to upgrade resources within individual servers. Virtualization and trends of cloud computing adapt for newer workloads, but suggest that many virtual machines work together to solve simple problems. These problems are simple enough for one operating system, or one virtual machine, to solve, but the singular virtual machine might not enough hardware resources to process the workload. Having many virtual machines over complicates the service model and increase upkeep costs. The simplicity of one virtual machine that is powerful enough to handle the workload would decrease management costs.

Rack scale architecture (RSA) intends on providing a solution to this by giving one operating system access to other servers’ hardware. Intel is a leader in developing this architecture, but their solution requires administrators to change server organization and install specialized hardware. This hardware becomes really expensive, thus making the optimized data center layout unreachable for most consumers.

This innovation brings the novelty of Intel’s RSA solution, but through a solely software solution. Leveraging the open source Xen hypervisor, this product provides a virtual machine to share resources of other physical servers. With it, the data center administrator will be able to select specific physical machines to be places within a conceptual resource pool which the master virtual machine can share resources from.

Specifically, this product focuses on providing intelligent memory management algorithms across physical servers. Traditionally, memory is shared across servers using remote direct memory access (RDMA) enabled network interfaces. These can be expensive. Using Bitdefender’s virtual machine introspection library, libbdvmi, the Xen hypervisor will be able to monitor when the master virtual machine tries to access data from a “remote” page of memory. Then, the hypervisor will be able to copy memory from the remote physical machine into the virtual machine. Furthermore, Xen will be responsible for intelligently mapping memory least used into remote physical machines while keeping the most used memory local to the virtual machines physical host. With this, the virtual machine will not suffer from the costs of retrieving and copying memory from a remote host as much as it could.

This innovation provides mechanisms to handle new workloads that IoT and cloud computing bring.

Neel Shah **CSCI-4342**
Writing 3 **10/31/2016**

Commercial Opportunity:

This innovation, Software Defined Rack Scale Architecture (RSA), will cater to the data center market. The data center market is comprised of systems administrators, systems developers, and Internet of Things (IoT) enthusiasts who will have a need to provide and use a platform that transparently scales to serve the growing number of IoT devices being used. Intel estimates that the number of IoT devices in use will be upwards of 30 billion by year 2020. Since data center and systems administrators will have to create a platform that scales well to account for trends in technology, they will find this innovation ideal to provide the highest quality of service. Systems developers and IoT enthusiasts will always be following the trends in technology to create new products. Since the trends in technology show an increasing demand for IoT devices, developers and enthusiasts will provide. Vendors and retailers in technology are hiring these two types of developers to create more IoT devices. As these products are being released, their developers and their vendors/retailers will expect the end user to enjoy a high quality of service while using the IoT products. If the data centers hosting the backend for IoT devices are not adapting to the new trends in technology, then IoT developers will encourage the data centers to upgrade their platforms.

Data centers must adapt to new trends in technology to provide the highest quality of service to developers and their customers. If data centers fail to do so, the end users of an IoT device will not be happy and will cease to use the device. When end users stop using products, their creators and vendors will be negatively be affected since their profits will decrease. This decrease of profit is caused by the data center that they use to host their IoT device backend. The developer or vendor of the IoT device will pursue new data centers to host their backend. Ultimately, if a data center does not adapt to trends in technology, then they will lose contracts and not make a profit gain.

This innovation provides a flexible, software defined solution to address the IoT trends in technology. With this, data centers do not need to purchase new hardware or remodel existing platforms. Rather, they can provision existing hardware and server configurations into a RSA based setup. This flexible model is beneficial to data centers--now an administrator does not need to spend time and money reconfiguring and rebuilding existing server configurations, rather they simply install and configure software. Installing and configuring software is exponentially trivial compared to installing and configuring hardware. This triviality increases the profit gains that a data center will experience since they are minimizing the work to adapt to the IoT trends in technology.

Currently, there are no other competitors with this innovation. All advances in RSA are hardware defined. This innovation is one of the first software defined approaches to RSA which makes it very attractive. A hardware approach requires a massive redesign of data centers. New hardware needs to be ordered, installed, and configured to enable RSA within a data center. This software defined RSA solution does not need any of that. Instead, data center administrators are able to dismiss the overheads faced with new hardware and add support for RSA with ease. The competitive landscape might change as this innovation nears completion.

It could change such that other companies such as Intel or VMware release software solutions to RSA. Intel is the creator of one of the first hardware solutions to RSA. VMware is a virtualization and cloud giant who is constantly innovating the industry. These two companies have the potential to release a software solution, but as of now, no evidence of this exists on their webpages.

Since there are not any competitors in the market as of now, a release of Software Defined RSA would be widely consumed by data center administrators. This innovation is planned to be an open source product. An open source product enables the release of this toolset to be made prematurely. Open sourcing this product enables the systems community to contribute to its end goal. The goals of this innovation are to provide the foundations to emulate RSA via a software fabric and to provide algorithms to manage memory across distributed physical servers. This is not a full solution to software defined RSA, but it is enough to encourage involvement by the community. Data center administrators and systems developers would readily consume the premature product working together to develop it further.

Releasing this innovation prematurely also enables a quicker adoption of it. Since this product has the ability to adapt data centers to account for new trends in technology, administrators are able to justify investing time into it. If many data center administrators and systems administrators need a flexible and cheap approach to RSA, then the first product that they see will be the foundation to their solution. This means that this innovation will be adopted to be the standard software defined RSA solution.

The open source nature of this innovation means that the revenue gained directly from the use of this product is $0. However, revenue will be gained from a more support based model. Data centers who adopt this innovation will need support to install, configure, and maintain their RSA based servers. Data centers will be charged a service fee that scales according to the number of servers and the number of racks they are using in their RSA. Furthermore, data centers will be charged on the availability of service. For example, a data center who needs 24/7 support of their RSA will be charged more compared to the data center who wishes to have incidental support. This service oriented revenue model will be similar to RedHat's revenue model. Our revenue model will be comprised of various service tiers: tier 0 users will pay $0.00/year for no support; tier 1 users will pay $3,000.00/year for incidental support on each rack or a 10 server cluster; tier 2 users will pay $7,000.00/year for 24/7 remote support on each rack or a 10 server cluster; tier 3 users will pay $15,000.00/year for 24/7 onsite support on each rack or a 10 server cluster.

Providing these service tiers, tier 0 and 1 provide the greatest opportunity for security risks. Tier 0 or tier 1 users may configure their racks in insecure fashions. RSA introduced an assumed trust factor. Since an operating system shares the hardware resources of different physical servers, an assumption regarding trust has to be made. Users of RSA must assume that all servers within a rack are trusted. This means that the contents of memory are not malicious or corrupt. If the RSA based cluster is misconfigured and security holes are introduced, then this trust factor is violated and there is a potential data leak. This risk can affect all users of this innovation since its relative to the data that they are serving on top of this platform.

Societal Impact:

Software defined Rack Scale Architecture (RSA) addresses scalability problems that data centers and service providers face when adapting to the Internet of Things (IoT) trends in technology. The groups of people that would be affected by this innovation are data center administrators or systems administrators, systems developers, IoT developers, and IoT users. These groups of people are a smaller niche within the information technology and the computer science industries. The affected groups of people compose the full stack of an IoT device or application. From IoT users, to IoT developers, to IoT hosts/supporters, and to systems developers, these groups of people will face change due to a software solution to RSA.

A software solution to RSA means enables data center administrator to provide a more scalable and reliable platform to host IoT backends that IoT developers create. Having a more reliable backend implies that IoT users will experience a greater quality of service with their devices. This impact creates a higher demand for IoT devices by end users since they are happier with what developers create. Developers trust their data centers moreso and can create more complex IoT devices and backends. Data centers will be able to scale more efficiently to handle the complex and new IoT devices. Using the methods that this innovation brings and the foundation that it creates for software based RSA, systems developers will be able to provide more features for it and drive the technology to grow even more.

Hardware is ever changing as well, and is being more optimized to handle more data and process it faster. This means that RSA needs to be updated to handle new hardware in technology. Having a software defined RSA implies that updating it is flexible and trivial compared to hardware defined RSA. This means that systems developers will be able to grow this innovation after its open source release to constantly keep it up-to-date with standards in technology.

There are not any environmental issues, health issues, or regulatory issues behind this software. However, there are data centers and cases where this provides some security holes. For example, if the RSA rack is misconfigured and the assumed trust is violated, then there is a critical security breach within the cluster. Data centers that are used by the federal government, health care providers, and other sensitive industries would need to take special care in assuring that trust is not violated. These users of the innovation need to either assure proper configurations or purchase the higher two service tiers to guarantee a secure configuration that does not leak data. In those industries, software defined RSA would need regulatory policies in place. There will be an online wiki outlining proper configuration to assure no security breaches exist in the rack.

The assumed trust constant introduces unethical use cases. For example, if a company purchases the greatest support tier, tier 3, then they receive onsite support. If the support technician has malicious intent, they are able to violate the trust constant and create a security breach. This breach in security can then be used by the technician to steal data or compromise sensitive operations. To account for this, our service technicians will be thoroughly vetted and be subjected to detailed background investigations to assure their compliance to our security policies. Furthermore, we will provide periodic system integrity evaluations to tier 2 and above users to assure that their RSA cluster is free of security violations.