Next generation Cloud Computing Architecture

Enabling real-time dynamism for shared distributed physical infrastructure

Vijay Sarathy, Purnendu Narayan and Rao Mikkilineni, Ph D
Kawa Objects, Inc.
Los Altos, CA
vijay@kawaobjects.com, purnendu@kawaobjects.com, rao@kawaobjects.com

***Abstract*— Cloud computing is fundamentally altering the expectations for how and when computing, storage and networking resources should be allocated, managed and consumed. End-users are increasingly sensitive to the latency of services they consume. Service Developers want the Service Providers to ensure or provide the capability to dynamically allocate and manage resources in response to changing demand patterns in real-time. Ultimately, Service Providers are under pressure to architect their infrastructure to enable real-time end-to-end visibility and dynamic resource management with fine-grained control to reduce total cost of ownership while also improving agility.**

The current approaches to enabling real-time, dynamic infrastructure are inadequate, expensive and not scalable to support consumer mass-market requirements. Over time, the server-centric infrastructure management systems have evolved to become a complex tangle of layered systems designed to automate systems administration functions that are knowledge and labor intensive. This expensive and non-real time paradigm is ill suited for a world where customers are demanding communication, collaboration and commerce at the speed of light. Thanks to hardware assisted virtualization, and the resulting decoupling of infrastructure and application management, it is now possible to provide dynamic visibility and control of services management to meet the rapidly growing demand for cloud-based services.

What is needed is a rethinking of the underlying operating system and management infrastructure to accommodate the ongoing transformation of the data center from the traditional server-centric architecture model to a cloud or network-centric model. This paper proposes and describes a reference model for a network-centric datacenter infrastructure management stack that borrows and applies key concepts that have enabled dynamism, scalability, reliability and security in the telecom industry, to the computing industry. Finally, the paper will describe a proof-of-concept system that was implemented to demonstrate how dynamic resource management can be implemented to enable real-time service assurance for network-centric datacenter architecture.

Keywords-Cloud Computing, Distributed Computing, Virtualization, Data Center

I. INTRODUCTION

The unpredictable demands of the Web 2.0 era in combination with the desire to better utilize IT resources are driving the need for a more dynamic IT infrastructure that can respond to rapidly changing requirements in real-time. This need for real-time dynamism is about to fundamentally alter the datacenter landscape and transform the IT infrastructure as we know it [1].

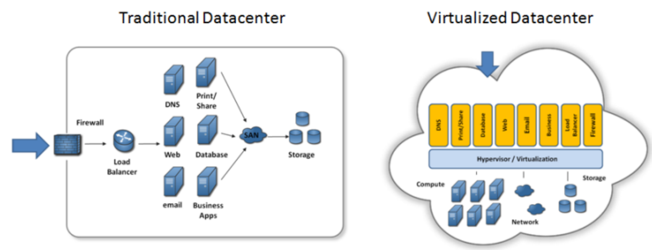

Figure 1: Transformation of the Traditional Datacenter

In the cloud computing era, the computer can no longer be thought of in terms of the physical enclosure – i.e. the server or box, which houses the processor, memory, storage and associated components that constitute the computer. Instead the "computer" in the cloud ideally comprises a pool of physical compute resources – i.e. processors, memory, network bandwidth and storage, potentially distributed physically across server and geographical boundaries which can be organized on demand into a dynamic logical entity i.e. a "cloud computer", that can grow or shrink in real-time in order to assure the desired levels of latency sensitivity, performance, scalability, reliability and security to any application that runs in it. What is truly enabling this transformation today is virtualization technology – more specifically hardware assisted server virtualization.

At a fundamental level, virtualization technology enables the abstraction or decoupling of the application payload from the underlying physical resource [2]. What this typically means is that the physical resource can then be carved up into logical or virtual resources as needed. This is known as provisioning. By introducing a suitable management infrastructure on top of this virtualization functionality, the provisioning of these logical resources could be made dynamic i.e. the logical resource could be made bigger or smaller in accordance with demand. This is known as dynamic provisioning. To enable a true "cloud" computer, every single computing element or resource should be capable of being dynamically provisioned and managed in real-time. Presently, there are many holes and areas for improvement in today's datacenter infrastructure before we can achieve the above vision of a cloud computer. Below we discuss these for each of the key datacenter infrastructure components.

978-0-7695-4063-4/10 $26.00 © 2010 IEEE
DOI 10.1109/WETICE.2010.14

A. Server Operating Systems and Virtualization

Whereas networks and storage resources - thanks to advances in network services management and SANs, have already been capable of being virtualized for a while, only now with the wider adoption of server virtualization do we have the complete basic foundation for cloud computing i.e. all computing resources can now be virtualized. Consequently, server virtualization is the spark that is now driving the transformation of the IT infrastructure from the traditional server-centric computing architecture to a network-centric, cloud computing architecture. With server virtualization, we now have the ability to create complete logical (virtual) servers that are independent of the underlying physical infrastructure or their physical location. We can specify the computing, network and storage resources for each logical server (virtual machine) and even move workloads from one virtual machine to another in real-time (live migration). All of this has helped to radically transform the cost structure and efficiency of the datacenter. Capacity utilization of servers can be increased and overall power consumption can be dramatically reduced by consolidating workloads. Additionally, thanks to server virtualization and live migration, High Availability (HA) and Disaster Recovery (DR) can be implemented much more efficiently [3]. Despite the numerous benefits that virtualization has enabled we are yet to realize the full potential of virtualization in terms of cloud computing. This is because:

- *Traditional server-centric operating systems were not designed to manage shared distributed resources:* The Cloud computing paradigm is all about optimally sharing a set of distributed computing resources whereas the server-centric computing paradigm is about dedicating resources to a particular application. The server-centric paradigm of computing inherently ties the application to the server. The job of the server operating system is to dedicate and ensure availability of all available computing resources on the server to the application. If another application is installed on the same server, the operating system will once again manage all of the server resources, to ensure that each application continues to be serviced as if it has access to all available resources on that server. This model was not designed to allow for the "dial-up" or "dial-down" of resource allocated to an application in response to changing workload demands or business priorities. This is why load-balancing and clustering was introduced. However, that does not alter the association of an application to a server. It just uses more instances of the application – each running in their own server, to try and share any increased burden.

  What is required for cloud computing - where distributed resources are shared amongst applications, is for a way to "mediate" between the applications and the resources by prioritizing the applications' needs based on relative business priorities. Our key observation here is that any sharing of resources will at some point inevitably result in contention for those resources which can only be resolved through a system that performs mediation globally across all the distributed shared resources. Today's operating systems do not natively provide this type of capability. This is often relegated to management systems that are layered on top or orthogonal to operating systems. However, the management system were also designed for a server-centric, configuration-based paradigm and have similar issues which make them ill suited as mediators that can enable real-time dynamism. The issues related to management systems are detailed in a separate section below. Finally, implementing simple server level QoS within the local server operating systems is not the answer as it does not help resolve contention amongst shared or distributed resources globally.

- *Current hypervisors do not provide adequate separation between application management and physical resource management:* Today's hypervisors have just interposed themselves one level down below the operating system to enable multiple "virtual" servers to be hosted on one physical server [4, 5]. While this is great for consolidation, once again there is no way for applications to manage how, what and when resources are allocated to themselves without having to worry about the management of physical resources. It is our observation that the current generation of hypervisors which were also born from the era of server-centric computing does not delineate hardware management from application management much like the server operating systems themselves. It is our contention that management and allocation of a shared infrastructure require a different approach. In an environment where resources are being shared the allocation of resources to specific applications must take into account the application's resource usage profile and business priority relative to other applications also using resources. In the ideal situation the profiles of resources are mediated to match the needs of applications based on their usage profiles and business priorities at run time.

- *Server virtualization does not yet enable sharing of distributed resources:* Server virtualization presently allows a single physical server to be organized into multiple logical servers. However, there is no way for example to create a logical or virtual server from resources that may be physically located in separate servers. It is true that by virtue of the live migration capabilities that server virtualization technology enables, we are able to move application workloads from one physical server to another potentially even geographically distant physical server. However, moving is not the same as sharing. It is our contention that to enable a truly distributed cloud computer, we must be able to efficiently share resources no matter where they reside purely based on the latency constraints of applications or services that consume the resources . Present day hypervisors do not exploit the potential of sharing distributed resources across physical and geographic boundaries and provide latency based composition of logical servers to fully utilize the power of distributed resources.

B. Storage Networks & Virtualization

Before the proliferation of server virtualization, storage networking and storage virtualization enabled many improvements in the datacenter. The key driver was the introduction of the Fibre Channel (FC) protocol and Fibre Channel-based Storage Area Networks (SAN) which provided high speed storage connectivity and specialized storage solutions to enable such benefits as server-less backup, point to point replication, HA/DR and performance optimization outside of the servers that run applications. However, these benefits have come with increased management complexity and costs. In fact SAN administrator costs are often cited as the single most critical factor affecting the successful deployment and management of virtual server infrastructure [6].

C. Network Virtualization

The virtual networks now implemented inside the physical server to switch between all the virtual servers provide an alternative to the multiplexed, multi-pathed network channels by trunking them directly to WAN transport thereby simplifying the physical network infrastructure. With the proliferation of multi-core multi-CPU commodity servers, it has almost become necessary to eliminate the mess of cables otherwise needed to interface multiple HBAs and NICs for each application with a single high speed Ethernet connection and a virtual switch. It is our contention that resultant architectural simplicity will significantly reduce associated management burden and costs.

D. Systems Management Infrastructure

Present day management systems are not cut out to enable the real-time dynamic infrastructure needed for cloud computing [7]. Here are the reasons why:

- *Human system administrators do not lend themselves to enabling real-time dynamism:* Once again, evolution of present day management systems can be traced back to their origin as management tools designed to help human system administrators who managed servers and other IT infrastructure. Even today, the systems have to be configured manually by a human with expert-level knowledge of the various infrastructure pieces and provide minimal automation. The systems often require human intervention to respond to alerts or events resulting in a significant amount of "human latency" which just does not lend itself to enabling the real-time dynamism required by cloud computing.

- *Policy-based management is not really automation:* There are a lot of management systems today that provide policy-based management capabilities. The trouble is the policies have to be programmed by expert system administrators who make judgments based on their experience. This is neither an optimal nor a scalable solution for cloud computing environments where the workload demands are unprecedented and vary wildly.

- *Virtualization compounds management complexity:* Every practitioner of server virtualization is aware of how virtualization can result in "Virtual Machine (VM) Sprawl" and the associated management burden it creates. VM Sprawl is a result of the ease with which new VMs can be created and proliferated on virtualized servers. This is however not the only factor affecting management complexity. Since VMs are cheaper to setup and run than a physical server, load balancers, routers and other applications that required physical servers are all now being run as in VMs within a physical server. Consequently, we now have to manage and route network traffic resulting from all these VMs within a server as well as the network traffic being routed across server. Adding to this confusion are the various OS vendor, each offering the other vendor's OS as a "guest" without providing the same level integration services. This makes the real life implementations, very management intensive, cumbersome and error-prone to really operate in a heterogeneous environment. It is our contention that the cost of this additional management may yet offset any cost benefits that virtualization has enabled through consolidation. A traditional management system with human-dependency is just an untenable solution for cloud computing.

E. Application Creation and Packaging

The current method of using Virtual Machine images that include the application, OS and storage disk images is once again born of a server-centric computing paradigm and does not lend itself to enable distribution across shared resources. In a cloud computing paradigm, applications should ideally be constructed as a collection of services which can be composed, decomposed and distributed on the fly. Each of the services could be considered to be individual processes of a larger workflow that constitutes the application. In this way, individual services can be orchestrated and provisioned to optimize the overall performance and latency requirements for the application.

II. Proposed Reference Architecture Model

If we were to distill the above observations from the previous section, we can see a couple of key themes emerging. That is:

- The next generation architecture for cloud computing must completely *decouple* physical resources management from virtual resource management; and

- Provide the capability to *mediate* between applications and resources in real-time.

As we highlighted in the previous section, we are yet to achieve perfect decoupling of physical resources management from virtual resource management but the introduction and increased adoption of hardware assisted virtualization (HAV) as an important and necessary step towards this goal. Thanks to HAV, a next generation hypervisor will be able to manage and truly ensure the same level of access to the underlying physical resources. Additionally, this hypervisor should be capable of managing both the resources located locally within a server as well as any resources in other servers that may be located elsewhere physically and connected by a network.

Once the management of physical resources is decoupled from the virtual resource management the need for a mediation layer that arbitrates the allocation of resources between multiple applications and the shared distributed physical resources becomes apparent.

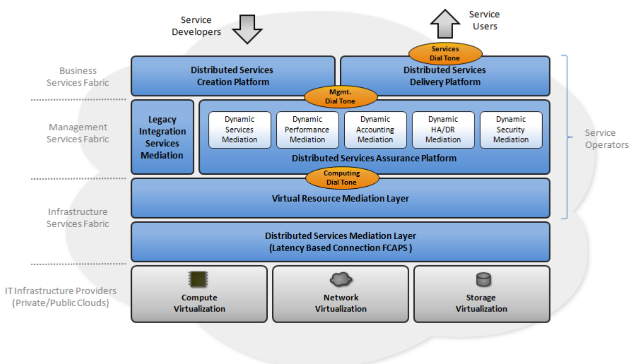

Figure 2: Reference Architecture Model for Next Generation Cloud Computing Infrastructure

- *Infrastructure Service Fabric:* This layer comprises two pieces. Together the two components enable a computing resource "dial-tone" that provides the basis for provisioning resource equitably to all applications in the cloud:

  1. *Distributed Services Mediation:* This is a FCAPS-based (Fault, Configuration, Accounting, Performance and Security) abstraction layer that enables autonomous self-management of every individual resource in a network of resources that may be distributed geographically, and a

  2. *Virtual Resource Mediation Layer:* This provides the ability to compose logical virtual servers with a level of service assurance that guarantees resources such as number of CPUs, memory, bandwidth, latency, IOPS (I/O operations per second), storage throughput and capacity.

- *Distributed Services Assurance Platform:* This layer will allow for creation of FCAPS-managed virtual servers that load and host the desired choice of OS to allow the loading and execution of applications. Since the virtual servers implement FCAPS-management, they can provide automated mediation services to natively ensure fault management and reliability (HA/DR), performance optimization, accounting and security. This defines the management dial-tone in our reference architecture model. We envision that service providers will offer these virtual servers with appropriate management API (management dial-tone) to the service developers to create self-configuring, self-healing, self optimizing services that can be composed to create self-managed business workflows that are independent of the physical infrastructure.

- *Distributed Services Delivery Platform:* This is essentially a workflow engine that executes the application which - as we described in the previous section, is ideally composed as business workflow that orchestrates a number of distributable workflow elements. This defines the services dial tone in our reference architecture model.

- *Distributed Services Creation Platform:* This layer provides the tools that developers will use to create applications defined as collection of services which can be composed, decomposed and distributed on the fly to virtual servers that are automatically created and managed by the distributed services assurance platform.

- *Legacy Integration Services Mediation:* This is a layer that provides integration and support for existing or legacy application in our reference architecture model.

III. DEPLOYMENT OF THE REFERENCE MODEL

Any generic cloud service platform requirements must address the needs of four categories of stake holders (1) Infrastructure Providers, (2) Service Providers. (3) Service Developers, and (4) End Users. Below we describe how the reference model we described will affect, benefit and be deployed by each of the above stakeholders.

Infrastructure providers: These are vendors who provide the underlying computing, network and storage resources that can be carved up into logical cloud computers which will be dynamically controlled to deliver massively scalable and globally interoperable service network infrastructure. The infrastructure will be used by both service creators who develop the services and also the end users who utilize these services. This is very similar to switching, transmission and access equipment vendors in the telecom world who incorporate service enabling features and management interfaces in their equipment. Current storage and computing server infrastructure has neither the ability to dynamic dial-up and dial-down resources nor the capability for dynamic management which will help eliminate the numerous layers of present day management systems and the human latency they contribute. The new reference architecture provides an opportunity for the infrastructure vendors to eliminate current systems administration oriented management paradigm and enable next generation real-time, on-demand, FCAPS-based management so that applications can dynamically request the dial-up and dial-down of allocated resources.

Service providers: With the deployment of our new reference architecture, service providers will be able to assure both service developers and service users that resources will be available on demand. They will be able to effectively measure and meter resource utilization end-to-end usage to enable a dial-tone for computing service while managing Service Levels to meet the availability, performance and security requirements for each service. The service provider will now manage the application's connection to computing, network and storage resource with appropriate SLAs. This is different from most current cloud computing solutions that are nothing more than hosted infrastructure or applications accessed over the Internet.

This will also enable a new distributed virtual services operating system that provides distributed FCAPS-based resource management on demand.

Service Developers: They will be able to develop cloud-based services using the management services API to configure, monitor and manage service resource allocation, availability, utilization, performance and security of their applications in real-time. Service management and service delivery will now be integrated into application development to allow application developers to be able to specify run time SLAs.

End Users: Their demand for choice, mobility and interactivity with intuitive user interfaces will continue to grow. The managed resources in our reference architecture will now not only allow the service developers to create and deliver services using logical servers that end users can dynamically provision in real-time to respond to changing demands, but also provide service providers the capability to charge the end-user by metering exact resource usage for the desired SLA.

IV. Proof of Concept

In order to demonstrate some of the key concepts introduced in the reference architecture model described above, we implemented a proof-of-concept prototype purely in software. The key objectives for the proof of concept were to demonstrate (1) Resource provisioning based on an application profile, and (2) FCAPS-based dynamic service mediation – specifically the ability to configure resources and the ability to dial-up or dial-down resources, based on business priorities and changing workload demand

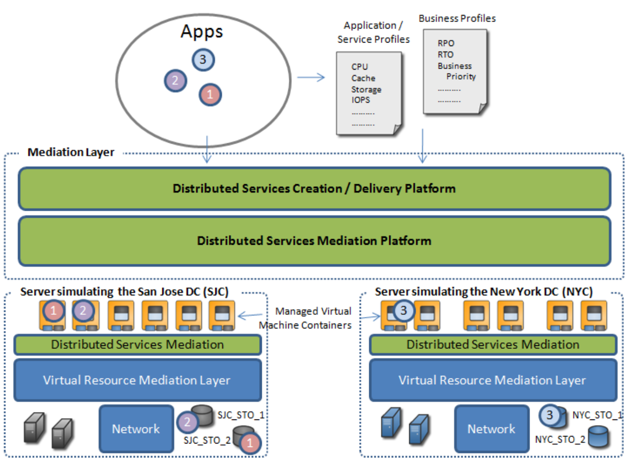

Figure 3 – Diagram showing a logical view of the infrastructure simulated in the proof-of-concept

To facilitate the demonstration of the objectives above, the entire system was implemented in software without deploying or controlling actual hardware resources. For example, all the computing resources i.e. server, network and storage resources, used in the prototype were not real but simulated in software as objects with the appropriate attributes. The proof-of-concept design allowed the resource objects to be distributed across multiple computers that were part of the same Local Area Network (LAN). This simulated a private cloud of physical resources distributed across servers. Each simulated physical resource object was managed independently by a local management agent that wrapped around it to provide autonomous FCAPS-based management of the resource. Another computer was used to run our prototype implementation of the global FCAPS-based mediation layer.

This mediation layer would discover all the physical resources available on the network and become aware of the global pool of resources it had available to allocate to applications. Next, the mediation layer took as input the application profiles i.e. CPU, memory, storage IOPS, capacity and throughput as well as relative business priority and latency tolerance of the various applications that were run in the cloud. The logical representation of the entire proof-of-concept is displayed in Figure 3.

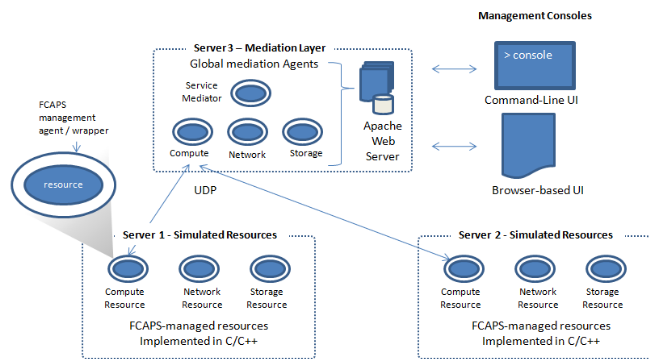

Figure 4 – Diagram showing the physical representation the proof-of-concept

All the simulated software resource objects and mediation layer objects were written using Microsoft Visual C++. The mediation layer provided two interface options i.e. a telnet accessible shell prompt for command line control as well as a browser based console application developed using Adobe Flex UI components that was served up using an Apache Web Server. The physical architecture of the Proof-of-concept is shown in figure 4.

With the above setup in place, we were able to simulate the running of applications and observe dynamic mediation take place in real-time. The first step was to demonstrate the provisioning of virtual server by allocating logical CPU, memory, network bandwidth and storage resources. Figure 5 is one of the many screens from the browser-based management UI console that shows the allocation and utilization of storage resources for one of the running applications.

Next we wanted to demonstrate FCAPS-based service mediation i.e. how logical resources could be dynamically provisioned in response to changing workload patterns. Figure 6 shows a screen from the browser-based management UI console which allowed us to simulate changes in application workload demand patterns by moving the sliders on the left of the screen that represent parameters such as storage network bandwidth, IOPS, throughput and capacity. The charts on the right of figure 6 show how the mediation layer responded by dialing-up or dialing down logical in response to those changes.

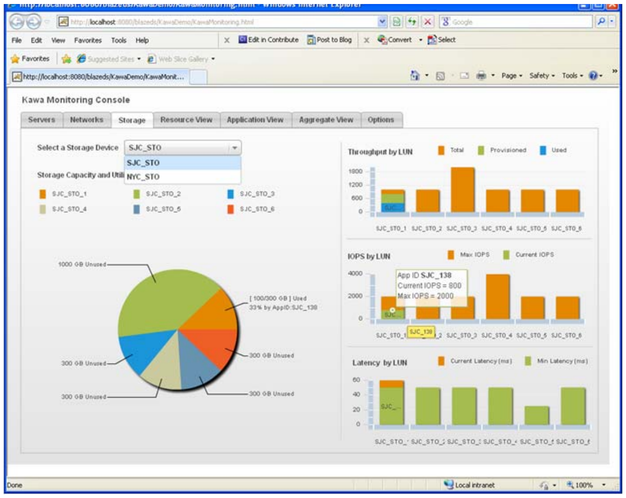

Figure 5 - UI screen from proof-of-concept showing resource allocation to an application in the storage dashboard

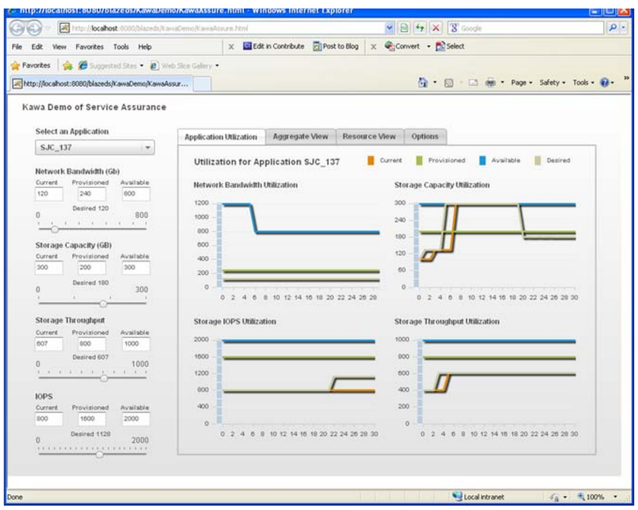

Figure 6- UI screen from proof-of-concept showing dynamic mediation of resources in response to a change in demand for an application.

V. Conclusion

In this paper, we have described the requirements for implementing a truly dynamic cloud computing infrastructure. Such an infrastructure comprises a pool of physical computing resources – i.e. processors, memory, network bandwidth and storage, potentially distributed physically across server and geographical boundaries which can be organized on demand into a dynamic logical entity i.e. "cloud computer", that can grow or shrink in real-time in order to assure the desired levels of latency sensitivity, performance, scalability, reliability and security to any application that runs in it.

We identified some key areas of deficiency with current virtualization and management technologies. In particular we detailed the importance of separating physical resource management from virtual resource management and why current operating systems and hypervisors – which were born of the server-computing era, are not designed and hence ill-suited to provide this capability for the distributed shared resources typical of cloud deployment. We also highlighted the need for FCAPS-based (Fault, Configuration, Accounting, Performance and Security) service "mediation" to provide global management functionality for all networked physical resources that comprise a cloud – irrespective of their distribution across many physical servers in different geographical locations.

We then proposed a reference architecture model for a distributed cloud computing mediation (management) platform which will form the basis for enabling next generation cloud computing infrastructure. We showed how this infrastructure will affect as well as benefit key stakeholders such as the Infrastructure providers, service providers, service developers and end-users.

Finally, we detailed a proof-of-concept which we implemented in software to demonstrate some of the key concepts such as application resource provisioning based on application profile and business priorities as well as dynamic service mediation. Next step is to implement this architecture using hardware assisted virtualization.

We believe that what this paper has described is significantly different from most current cloud computing solutions that are nothing more than hosted infrastructure or applications accessed over the Internet. The proposed architecture described in this paper will dramatically change the current landscape by enabling cloud computing service providers to provide a next generation infrastructure platform which will offer service developers and end-users unprecedented control and dynamism in real-time to help assure SLAs for service latency, availability, performance and security.

References

[1] Rao Mikkilineni, Vijay Sarathy "Cloud Computing and Lessons from the Past", Proceedings of IEEE WETICE 2009, First International Workshop on Collaboration & Cloud Computing, June 2009

[2] Rajkumar Buyyaa, Chee Shin Yeoa, , Srikumar Venugopala, James Broberga, and Ivona Brandicc, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility", Future Generation Computer Systems, Volume 25, Issue 6, June 2009, Pages 599-616

[3] Adriana Barnoschi, " Backup and Disaster Recovery For Modern Enterprise", 5th International Scientific Conference, Business and Management' 2008, Vilnius, Lithuvania. - http://www.vgtu.lt/leidiniai/leidykla/BUS_AND_MANA_2008/inf-communication/630-635-G-Art-Barnoschi00.pdf

[4] Jason A. Kappel, Anthony T. Velte, Toby J. Welte, "Microsoft Virtualization with Hyper-V", McGraw Hill, New York, 2009

[5] David Chisnall, "guide to the xen hypervisor", First edition, Prentice Hall, Press, NJ, 2009

[6] Gartner's 2008 Data Center Conference Instant Polling Results: Virtualization Summary – March 2, 2009

[7] Graham Chen, Qinzheng Kong, Jaon Etheridge and Paul Foster, "Integrated TMN Service Management", Journal of Network and Systems Management, Springer New York, Volume 7, 1999, p469-493