

Growing Cloud Computing Efficiency

Dr. Mohamed F. AlAjmi, PhD
Vice Dean-Quality and Development
Head of Quality and E-Learning units
Prince Sultan College of EMS
King Saud University
Riyadh, Saudi Arabia

Dr. Arun Sharma Head, Department
of Computer Science, Krishna
Institute of Engineering and
Technology, Ghaziabad-201206,
INDIA

Shakir Khan*
(Corresponding Author)
Researcher at E-Learning and
Distance Learning Deanship, King
Saud University, Riyadh Saudi
Arabia, Nationality Indian

Abstract— Cloud computing is basically altering the expectation for how and when computing, storage and networking assets should be allocated, managed and devoted. End-users are progressively more sensitive in response time of services they ingest. Service Developers wish for the Service Providers to make sure or give the ability for dynamically assigning and managing resources in respond to alter the demand patterns in real-time. Ultimately, Service Providers are under anxiety to build their infrastructure to facilitate real-time end-to-end visibility and energetic resource management with well grained control to decrease total cost of tenure for improving quickness. What is required to rethink of the underlying operating system and management infrastructure to put up the on-going renovation of data centre from the traditional server-centric architecture model to a cloud or network centric model? This paper projects and describes a indication model for a network centric data centre infrastructure management heap that make use of it and validates key ideas that have enabled dynamism, the quality of being scalable, reliability and security in the telecommunication industry to the computing engineering. Finally, the paper will explain a proof of concept classification that was implemented to show how dynamic resource management can be enforced to enable real-time service guarantee for network centric data centre architecture.

Keywords- Cloud Computing; Data centre; Distributed Computing; Virtualization; data centre architecture.

I. INTRODUCTION

The random demands of the Web 2.0 era ingrouping with the desire to better utilize IT resources are driving the need for a more dynamic IT infrastructure that can respond to speedily changing requirements in real-time. This need for real-time enthusiasm is about to basically alter the data centre background and alter the IT infrastructure as we know it. In the cloud computing era, the computer can no longer be assumed in standings of the physical insertion i.e. the server or box, which households the processor, memory, storage and related components that establish the computer. In its place the “computer” in the cloud perfectly includes a group of physical work out resources i.e. work stations, retention, network bandwidth and storage, possibly circulated physically across server and geographical borders which can be planned on demand into a dynamic consistent entity i.e. a “cloud computer”, that can develop or shrink in real-time in order to

promise the desired levels of potential sensitivity, performance, scalability, consistency and safety to any application that runs in it. What is really supporting this alteration today is virtualization technology, more precisely hardware assisted server virtualization. At an ultimate level, virtualization technology allows the abstraction or decoupling of the request payload from the original physical resource. What this typically means is that the physical resources can then be carved up into logical or virtual resources as needed. This is acknowledged 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, each single computing element or resource should be proficient of being enthusiastically provisioned and succeeded in real-time. Currently, there are numerous holes and areas for development in today’s data centre infrastructure before we can attain the above vision of a cloud computer.

A. Server useful systems and virtualization

Whereas networks and storage resources appreciate to advances in network facility management and SANs have already been proficient of being virtualized for a while, only now with the broader acceptance of server virtualization, do we have the complete basic foundation for cloud computing i.e. all computing properties can now be virtualized. Subsequently, server virtualization is the catalyst that is now motivating the transformation of the IT infrastructure from the traditional server-centric computing architecture to a network centric cloud computing architecture. When server virtualization is done, we have the capability to generate whole logical (virtual) servers that are free of the fundamental physical infrastructure or their physical position. We can postulate the computing, network and storage resources for all logical server (virtual instrument) and even transfer workloads from one virtual machine to another in real time (live migration).

All of this has aided deeply to convert the cost structure and competence of the data centre. Despite the many assistances that virtualization has allowed, we are still to realize the complete potential of virtualization with respect to the cloud computing. This is because:

1) *Usual server centric operating systems were not planned to manage collective spread resources:*

The Cloud computing example is all about optimally involving a set of scattered computing resources while the server-centric computing example is about devoting resources to a specific application. The server-centric example of computing fundamentally ties the application to the server. The work of the server operating system is to commit and guarantee to obtain ability of all accessible computing resources on the server to the application. If another application is installed on the same server, the operating system will once again manage the entire server resources to confirm that each application remains to be checked 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 change workload demands or business priorities. This is the reason that load-balancing and clustering was introduced.

2) *Current hypervisors do not supply sufficient division between application management and physical supply 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. While this is great for consolidation, once again there is no way for applications to manage how, what and when resources are assigned to themselves without having to concern 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 define hardware management from application management much similar the server operating systems themselves.

3) *Server virtualization does not yet allow contribution of scattered resources:*

Server virtualization currently permits a single physical server to be structured into multiple logical servers. However, there is no way for example to generate a analytical or computer-generated server from resources that may be physically placed in separate servers. It is true that by virtue of the live migration capabilities that server virtualization technology enables, we are intelligent to move application loads from one physical server to another possibly even geographically distant physical server. However, moving is not the similar as sharing. It is our contention that to enable a truly distributed cloud computer, we must be able efficiently to share resources, no problem where they exist in purely based on the potential constraints of applications or services that consume their sources.

B. Storage set of connections & virtualization

Before the production of server virtualization, storage networking and storage virtualization permitted many improvements have been done in the data centre. The key improvement was the introduction of the FibreChannel (FC)

protocol and Fibre Channel-based Storage Area Networks (SAN) which delivered great speed of storage connectivity and dedicated storage solutions to allow such profits as server-less backup, point to point reproduction, HA/DR and presentation optimization outside of the servers that run applications. However, these pay backs have come with improved management complication and costs.

C. System virtualization

The virtual networks now applied inside the physical server to switch between all the virtual servers to provide a substitute to the multiplexed, multi-patched network channels by trucking them nonstop to WAN transport, thus shortening the physical network infrastructure.

D. Function creation and binding

The existing method of exhausting Virtual Machine images that contain the application, OS and loading disk images is once again born of a server-centric computing model and does not provide itself to enable supply across mutual resources. In a cloud computing pattern, applications should preferably be built as a collection of facilities which can be integrated, disintegrated and distributed on the fly. Each of the services could be measured to be individual procedure of a larger workflow that establishes the application. In this way, individual services can be arranged and provisioned to improve the overall performance and potential requirements for the application.

II. PLANNED SUGGESTION STRUCTURAL DESIGN MODEL

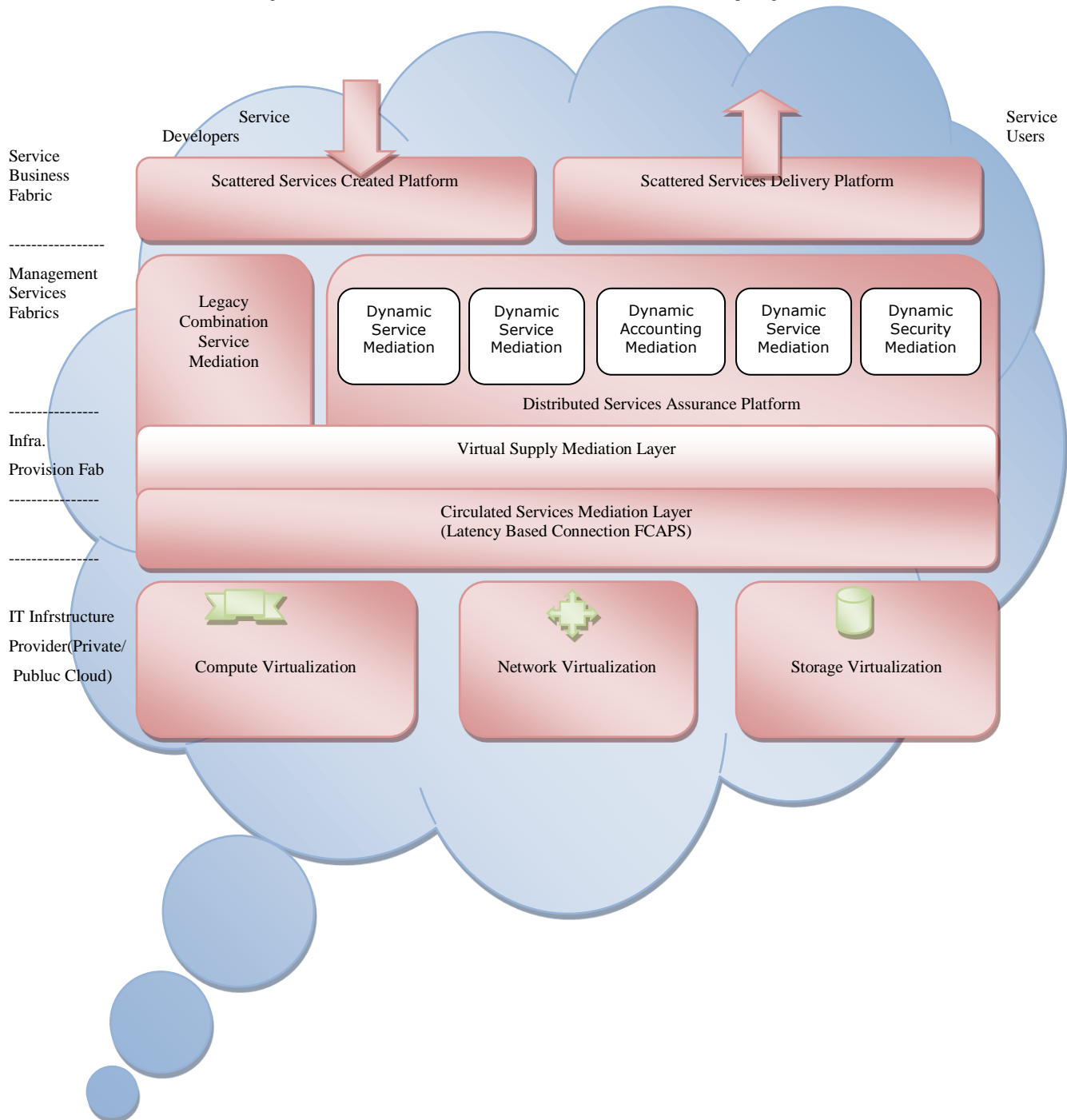
If we were to purify the above interpretations from the previous section, we can realize that a couple of key subjects emerging. That is:

A. *The next generation architecture for cloud computing must entirely decouple physical resources management from virtual resource management.*

B. *Supply the proficiency to intervene between applications and resources in real time.*

As we stressed in the earlier section, we are still to attain perfect decoupling of physical resources management from virtual resource management but the outline and improved acceptance of hardware assisted virtualization (HAV) as a significant and essential step towards this objective. Thanks to HAV, a next generation hypervisor will be capable to achieve and truly guarantee the identical level of access to the fundamental physical resources. Moreover, this hypervisor should be proficient of handling both the resources situated locally inside a server as well as any resources in other servers that may be situated somewhere else physically and linked by a network. Once the controlling of physical resources is decoupled from the virtual resource management. The necessity for a mediation layer that mediates the distribution of resources between various applications and the shared distributed physical resources becomes obvious.

Figure 2 Position Architecture Model for Next Generation Cloud Computing Infrastructure



III. INFRASTRUCTURE PROVISION FABRICS

This layer includes two pieces. Together with the two components allow a computing resource “dial-tone” that delivers the basis for provisioning resource fairly to all applications in the cloud.

A. Scattered 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.

B. Virtual supply mediation layer

This gives the ability to create logical virtual servers with a level of service guarantee those assurances resources such as number of CPUs, memory, bandwidth, latency, IOPS (I/O operations per second), storage through put and capacity.

C. Circulated services Assurance Platform

This layer will allow for creation of FCAPS-managed virtual servers that pack and host the desired choice of OS to allow the loading and execution of applications. Since the virtual servers implement FCAPS-management, they can give automated mediation services natively to guarantee fault management and reliability (HA/DR), performance optimization, accounting and security. This describes the management dial-tone in our orientation architecture model.

D. Scattered Services Delivery Platform

This is basically a workflow engine that implements the application which we described in the previous section, is preferably composed as business workflow that organizes a number of distributable workflow elements. This describes the services dial tone in our reference architecture model.

E. Scattered Services Creation Platform

This layer gives the tools that developers will utilize to generate applications defined as group of services which can be composed, decomposed and scattered on the fly to virtual servers that are automatically shaped and run by the distributed services assurance platform.

F. Legacy Combination Services Mediation

This is a layer that gives addition and support for existing or legacy application in our reference architecture model.

IV. DEPLOYMENT OF THE SUGGESTION MODEL

Any generic cloud service platform requirements must address the needs of four categories of stake holders:

- 1) *Infrastructure suppliers,*
- 2) *Service suppliers.*
- 3) *Facility Developers*
- 4) *End Users.*

Below we explain how the reference model described will affect, benefit and are set up by each of the above stake holders.

A. Infrastructure suppliers

These are vendors who give the underlying computing, network and storage resources that can be fixed up into logical cloud computers which will be dynamically forced to deliver extremely scalable and globally interoperable service network infrastructure. The infrastructure will be utilized by both service creators who develop the services and also the end users who use these services.

B. Service suppliers

With the employment of our innovative reference architecture, service providers will be capable to promise both service developers and service users that resources will be obtainable on demand. They will be capable effectively to determine and meter resource utilization end-to-end usage to allow a dial-tone for computing service while management Service Levels to meet the availability, performance and security needs for each service. The service provider will now handle the application's link to computing, network and storage resource with suitable SLAs.

A. Facility 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.

C. End users

Their demand for selection, mobility and interactivity with sensitive user interfaces will continue to rise. The managed resources in our reference architecture will now not only permit the service developers to generate and distribute services using logical servers that end users can dynamically provision in immediate to respond for changing needs, but also provide service providers the ability to charge the end-user by metering correct resource handling for the required SLA.

V. CONCLUSIONS

In this paper, we have explained the needs for implementing a real dynamic cloud computing infrastructure which contains a group of physical computing resources i.e. processors, memory, network bandwidth and storage, potentially dispersed physically through server and geographical limits which can be controlled on demand into a dynamic reasonable entity i.e. "cloud computer", that can develop or reduce in size immediately in order to give surety about the desired levels of latency, sensitivity, performance, scalability, consistency and security to any application that runs in it. We worked out few key areas of shortage of current virtualization and management technologies. Particularly we explained detail importance of sorting out physical resource management from virtual resource management and why current operating systems are not designed and hence it was suitable to provide this ability for the distributed shared resources especially of cloud deployment. We also painted the need for FCAPS-based (Fault, Configuration, Accounting, Performance and Security) service "mediation" to give global administration functionality for all networked physical resources that include a cloud, irrespective of their allocation across a lot of physical servers in different geographical locations. We then projected an indication architecture model for a distributed cloud computing mediation (management) platform which will outline the foundation for making the possibility of next generation cloud computing infrastructure. We proved how this infrastructure will involve as well as advantage key stake holders such as the Infrastructure providers, service providers, service developers and end-users.

Description in this paper is considerably different from most current cloud computing solutions that are nothing more than hosted infrastructure or applications accessed over the Internet. The proposed architecture in this paper will significantly change the current setting by enabling cloud computing service providers to give a next generation infrastructure platform which will recommend service developers and end users exceptional control and enthusiasm in real-time to assure SLAs for service latency, availability, performance and security.

ACKNOWLEDGEMENTS

The authors acknowledge the continuous support of deanship of research in King Saud University since this research is funded by a grant from the deanship under the number of

RGP – VPP - 150.

REFERENCES

- [1] Adriana Barnoschi, "Backup and Disaster Recovery For Modern Enterprise", 5th International Scientific Conference, Business and Management' 2008, Vilnius, Lithuania. http://www.vgtu.lt/leidiniai/leidykla/BUS_AND_MANA_2008/infcommunication/630-635-G-Art-Barnoschi00.pdf
- [2] David Chisnall, "guide to the xen hypervisor", First edition, Prentice Hall, Press, NJ, 2009.
- [3] Gartner's 2008 Data centre Conference Instant Polling Results: Virtualization Summary – March 2, 2009.
- [4] Graham Chen, Qinzhen Kong, Jaon Etheridge and Paul Foster, "Integrated TMN Service Management", Journal of Network and Systems Management, Springer New York, Volume 7, 1999, p469-493.
- [5] Jason A. Kappel, Anthony T. Velte, Toby J. Welte, "Microsoft Virtualization with Hyper-V", McGraw Hill, New York, 2009.
- [6] RaoMikkilineni, Vijay Sarathy "Cloud Computing and Lessons from the Past", Proceedings of IEEE WETICE 2009, First International Workshop on Collaboration & Cloud Computing, June 2009.
- [7] Rajkumar Buyyaa, Chee Shin Yeo, SrikumarVenugopala, 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.

AUTHORS PROFILE

Dr Mohammed Fahad AlAjmi was born in Kingdom of Saudi Arabia. He received his Ph.D in Pharmacy from King Saud University, Riyadh, Saudi

Arabia in 2007. He chaired many position in the university and currently working as vice dean for quality and development in Prince Sultan College for EMS affiliated to King Saud University. To date he taught many pharmacy students, more than 30 courses. Students' level varies from primary to undergraduate levels.

Dr. Arun Sharma, alumni of IIT Roorkee and Thapar University, received his M.Tech. (Computer Science and Engineering) from Punjabi University, Patiala, INDIA and Ph.D. (Computer Science) Thapar University, Patiala, INDIA. Currently, he is working as head of the department of Computer Science and Engineering Department in KIET school of engineering and Technology at Ghaziabad, India. His areas of interests include Software Engineering, Soft Computing and Database Systems. He has published a number of papers in international Journals and Conferences including IEEE, ACM, Springer, WILEY and others in India and abroad. Dr. Sharma is an active member of IEEE, ACM and Computer Society of India. He is also a member of Board of Studies (BoS) of Mahamaya Technical University (MTU), Noida. He is also on the panel of subject experts and examination for various Universities like IGNOU, BBA University (Central), Lucknow, GGSIP University, Delhi, Thapar University, Patiala, and others. He is also an active member of Editorial Board and Review Committee of several Journals including Journal of Computer Science (USA), International Journal of Computer Science and Security (Malaysia), Research Journal of Information Technology, USA and others.

Shakir Khan was born on 5th Feb, 1978 at Kallanheri in Saharanpur district UP, India. He is working as a Researcher in College of Electronic Learning in King Saud University, Kingdom of Saudi Arabia. He received his Master of Science in Computer Science from Jamia Hamdard (Hamdard University), New Delhi, India in the year 2005, and PhD computer Science, C.S.J.M University Kanpur India. He is member of IEEE. He has actively attended many international conferences and published various research papers in National and International conferences as well as journals. His current areas of interests are in Cloud Computing, Software Engineering, Data Mining and E Learning. Apart from that he worked in the field of Software Development in different MNC companies at Noida India.