International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 02 Issue: 03 | June-2015 www.irjet.net p-ISSN: 2395-0072

A Talk on Cloud Capacity Management

Mr. Narahari Narasimhaiah[1],Dr. R. Praveen Sam[2].

[1]Research scholar, Bharathiar University, Tamilnadu, India.
[2]Professor, Dept. of CSE. Andhra Pradesh, India.

ABSTRACT- ***Optimized resource utilization and cost savings are benefits of the cloud computing, which coincide with the objectives and goals of capacity management[1].Following article briefly describes how various cloud stakeholders perceive capacity[2] and what constitutes the capacity management process to generate infrastructure economies and optimizations. Cloud capacity management explains how the capacity management can be applied to cloud computing.***

Key Words: *Optimized,Computing, Capacityetc…*

1.CAPACITY MANAGEMENT

In a timely manner, all areas of IT, always matches to the current and future agreed upon needs of the business .The goals of the capacity management process is to ensure that cost justifiable IT capacity[2]. As of ITILv3, capacity management is one of the service design processes[4], which is the phase of the IT service lifecycle that converts service strategy into implementable IT services.

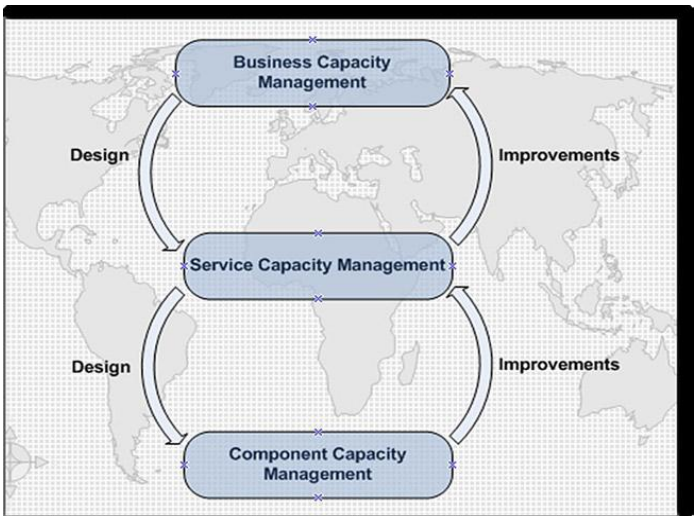

Fig -1 : Layers of Capacity management

In IT service, continuity management, security management, capacity management, andservice availability management are the service design phase[3] principles. Ensuring optimum[5] resource utilization, performance, and cost effectiveness, capacity management plays a vital role in any cloud-based service.

To maintain optimum and cost effective resource capacity is the main purpose of capacity management.

These resources may be facilities[3], hardware, software, or human resources. For budget and planning activities, new IT services provides timely resources and help in resource forecasts[2] ensures the capacity management.

IT service management process areas include demand management, financial management and service portfolio management[6] to ensure that service performance is maintained to avoid running out of resources .

The capacity management process has three interrelated views for a service management,

- Business capacity management
- Service capacity management
- Component capacity management

2. CAPACITY MANAGEMENT IN CLOUD COMPUTING

Asthe complex infrastructure isheterogeneous[3] and the associated toolsets are complex and expensive to implement capacity management,ensuring the highest levels of service performance[2] and continuity and managing their data center capacity, service providers have a plan in a cloud environment. To determine the capacity requirements[7], inputs are taken from tools that monitor resource performance, service level management.The cloud introduces aspects of shared infrastructure[1] and multi tenancy which are rented or leased and changes the way of capacity management is done. To use the pay as you go model and the elasticity model offered in the cloud,redefined process interfaces will be required for enterprises to effectively manage the capacity of the cloud.

2.1 The Capacity-Utilization Curve

The graphic in Figure 2illustrates capacity versus utilization

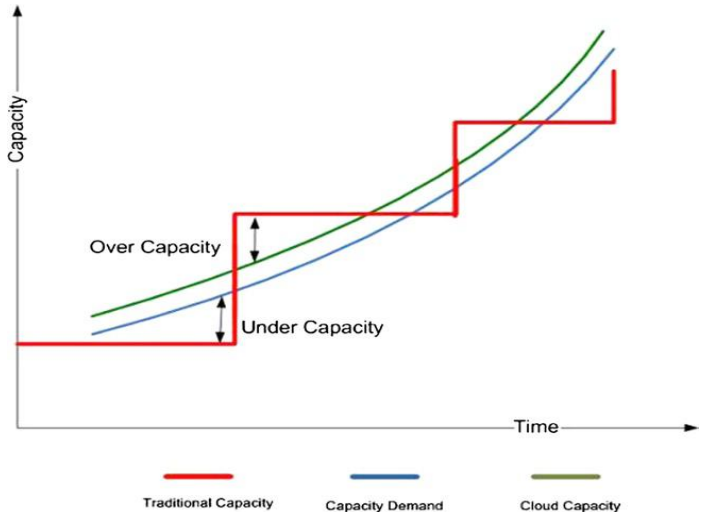

Chart - 1: Capacity versus utilization

This curve shows the core themes around actual service consumption versus cloud services. Cloud economies can be affected[8] by both over and under provisioning of cloud resources which is important to understand.Once the hardware is bought and paid for in a falling demand scenario there will be excess capacity which will go waste in an enterprise.Taking a risk of scenarios of falling demand, the organizations[1] making upfront decision on buying hardware. Getting obsolete or business demand falling is taken by the cloud provider for risk of technology in the cloud model. The organization is making upfront payments for hardware and thus cost of capital needs to be factored for the life of the hardware for a financial management.

3. CONVENTIONAL VS. CLOUD VIEW OF CAPACITY MANAGEMENT

Figure 3showing, depicts the high level private cloud architecture/model which is built upon virtualization, hardware infrastructure, and automation/orchestration layers[9].Private Cloud is the cloud deployment models amongst Large enterprises[3] and SMBs. The basis of application readiness and budget in hand to move applications onto the cloud, typically cloud service providers to these enterprises undertake building their private cloud on. Cloud enablement[2] includes configuring the hardware infrastructure, and building and orchestration/ automation layer and implementing virtualization technology[1].

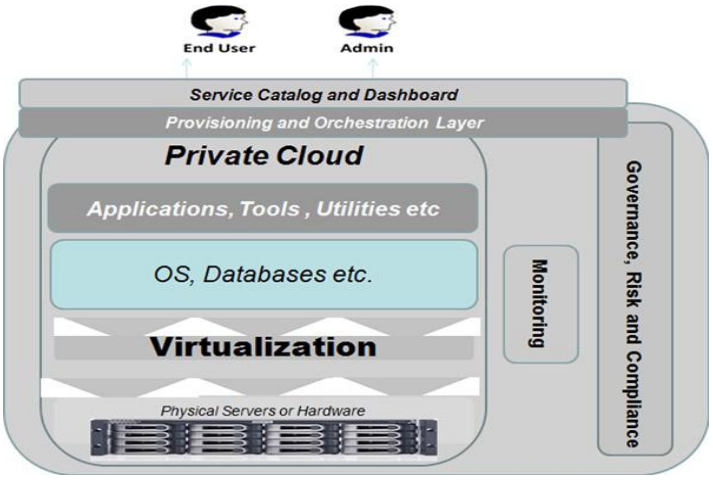

Fig -2: Reference Model for Private Cloud

Capacity management in a cloud environmentmust be able to address the following issues:

- Capacity leading[2] to over-provisioning and issues like VM sprawl.
- Harder to forecast capacity demand.
- Due to multi-tenant hosting in a resource-sharing environment, it hasinefficient and complex chargeback[10] mechanisms.
- Manual allocation inefficiencies is prone to errorsare cause of the failure of dynamic infrastructure.

4. CONCLUSION

All resources must be shared in a virtualized environment, which presents some unique and new challenges tothe data center when determining actual resource requirements.Ensuring that virtualizedbusiness-critical and every virtual server[1] application has the resources, when it requires them presents a complex resourceallocation challenge.To achieve acceptable service levels and consistent, at a known and controlled cost[11], define a capacity managementstrategy that includes:

- In yourvirtualized environment, it can define the processes, workflows, approvals, and schedules associated with the workloads.
- In your environment, to manage the performance[2] and capacity of the resources it can identify and collect the metrics critical.
- Establishing baselines and measuring current performance.
- To establish current and future resourcerequirements, it can collecting and using workload forecasts from consumers.

REFERENCES:

[1]. Cloud Capacity Management, Apress - Navin Sabharwal and Prashant Wali.

[2]. Capacity Management Analytics on System Z - Rajkumar Munsamy.

[3]. Hyperformix - Capacity Management Solution Brief.

[4]. Capacity Management Guidelines For VBLOCK Infrastructure Platforms - White Paper.

[5]. Capacity Management on the Journey to Cloud - CA - Paul Peterson.

[6]. CapacityControl: Why Capacity Management Needs to Evolve - CIRBA.

[7] Cloud Capacity Management - Cloud Computing.

[8]. Capacity Management Ensures Success for Enterprise cloud consumers - Team Quest - by Ron Potter.

[8]. Cloud Computing Architecture and Strategy.

[9]. Implementing Cloud Storage Metrics to Improve IT Efficiency and Capacity Management - Intel White Paper.

[10]. Effective Capacity Management with Modeling and Simulation assisted Performance Testing - Perspective - Infosys.

[11]. Virtualization and Cloud Management Using Capacity Planning - Clabby Analytics.

BIOGRAPHIES

NARAHARINARASIMHAIAH has 17 years of experience in IT industry. At Cognizant Technology Solutions, he is a Principal Architect-Technology with BFS TAO (Technology and Architecture Office). He holds a MBA, MS in Computer Science degree and MSc in Psychology. Currently pursuing PhD in Cloud Computing.

He is a TOGAF 8, PMI, Sun Certified Java Enterprise Architect, SCJP, MCP.NET, Six Sigma AND Certified ScrumMaster as well.

He has experience/knowledge/awareness in Micro Services, Big Data Analytics, Cloud Computing, Mobile Payments, Mobile Development (iOS, Android), Responsive Web/Single Page applications, Solutions Architecture, Service Oriented Architecture, Portfolio Assessments, Technology Rationalization / Consolidation exercises.

Highly astute, goal-oriented and result oriented professional offering a track record as a Technology Partner for a major financial organization. Proven success and experience in managing diverse large teams globally spanning multiple vendors with an on-site and offshore mix. Professional experience in startup as founder/co-founder and CEO organizations. Advising committee member/director in various Non-IT organization(s).

RACHAPUDI PRAVEEN SAM was born in Kurnool City in 1975. He received the B.Tech degree in Computer Science and Engineering with First Class in 1999 from Sri Krishna Devaraya University, Ananthapur, A.P., India; M.Tech degree Computer Science and Engineering with First Class in 2001 from Madras University, Chennai, T.N., India and was awarded Ph.D. degree in Computer Science and Engineering in 2010 from JNTU University, Ananthapur, A.P., India. His Ph.D. specialization is mobile and Ad Hoc Networks(MANETS). His expertise in Computer Networks and Network Security.

He is having 13 years of teaching experience, presently he is working as a professor of Computer Science and Engineering department for G.Pulla Reddy Engineering College (Autonomous), Kurnool City, India. He has a total of 25 publications out of which 13 papers in International and National Journals and 12 papers in National and International Conferences. He is a member of various professional bodies like ISTE, IE, CSI, IAENG, CSTA, and IACSIT.

He received Minor Research Project titled "Developing Disaster Management Applications using Mobile Ad Hoc Network Tested" sanctioned by UGC for a period of 2 years in March 2014.

