

# A Survey of Various Storage & Virtualization in Cloud Computing

Babu Singh Yadav<sup>1</sup>, Shivank Kumar Soni<sup>2</sup>, Anurag Jain<sup>3</sup>

Department, Computer Science and Engineering

Radharaman Institute of Technology and Science Bhopal Madhya Pradesh INDIA

**Abstract** — Cloud Computing has transformed the information technology by facilitating expandable on-demand prerequisite of computing resources. The creation of cloud computing has consequences in the concern of large-scale data centers around the world holding thousands of compute nodes. Due to the ever-increasing require for storage and computation has determined the growth of large data centers—the massive server areas that run numerous of today's Internet and various E-commerce applications. A data center be capable of consist of thousands of servers and can utilize as much energy as a small metropolitan. The enormous quantity of working out power want to make these methods consequences in several demanding and motivating virtualized data center, capacity management systems and other resource management problems that relate the initial capacity planning need when organizing applications into a virtualized data center. In this paper we are presenting a survey on cloud storage, data center virtualization and capacity management techniques in cloud computing which is valuable for researchers.

**Keywords**— Cloud Computing, Data Centers, Cloud Storage, Virtual Machine (VM), Virtualization.

## I. INTRODUCTION

Cloud computing has grow to be an accepted computing idea that agree to end-users to dynamically scale up or down the resources they utilize cloud computing resource to run their applications easily. Virtualization of resources is a key enabler of such solution mapping of infrastructure calculates capacity. Cloud providers put together big amount data centers to take advantage of cost reward due to economies of level and statistically complex it between applications to maximize effective and communicate the cost profits to consumers. Data centers are being progressively more virtualized on cloud. The proliferation of virtualization technologies and cloud service providers has also made it easy to create or buy virtual machines (VMs) to host applications. As a result, hundreds of virtual appliances with diverse characteristics can be consolidated in one physical server. Consolidation optimizes the utilization of server resources but leaves administrative tasks in a quagmire. While still attractive relative to traditional non-virtualized hosting, VM sprawl and over-sized VMs present problems in terms of capital and operational expenditure at these data

centers. As data centers effort to get better resource utilization all the way through server consolidation, it furthermore develops into essential for data center operators to identify with how the situation of applications impacts show and resource utilization [1].

If we don't mitigated, these problems can potentially forestall the adoption of virtualization methods; in that way regressing to over-conditioned, dedicated systems with higher costs of resources such as CPU, memory, storage, network, and power. Since virtualization technology facilitates several heterogeneous applications to run in a shared environment, careful attention needs to be directed towards the resource consumption characteristics of individual workloads. As the consolidated application VMs are quite diverse, they exhibit varying degrees of resource demands. Without a thorough understanding of the effects of resource allocation on application performance, VM resource provisioning may be sub-optimal. More-over, consolidation creates another challenging problem. An application running inside one VM can interfere with the performance of another application running inside another VM that share physical resources. This performance interference is often significant and applications cannot be modelled correctly without accounting for the interference. Unfortunately, the contemporary server management systems do not explicitly address this contention. To minimize the potential for interference affecting performance, administrators either resort to over-provisioning or require clients to pay more for additional resources. Without faithfully capturing the relationship between resource allocation and application performance and accurately understanding the effect of interference, configuration of cloud VMs and distribution of server resources in a data center are likely to be sub-optimal.

Next, data center workers are required to contract with the consumption and development setbacks speak about to estimating a data center's capacity and initial provisioning for new applications [2]. This may require models of an application's resource conditions and an recognizing of how they are contacted by different hardware configurations [3].

Well-organized source management is an explanation apprehension for data center workers gives the impression of bringing to both get together application SLAs and decrease charges. Distributed hosting proposals attempt to multiplex substantial sources between multiple consumer applications [4]. On the other hand without virtualization, it is complicated to make available strong separation between applications and operating systems are required to be customized to fairly allocate resources. Resource modelling is extremely challenging due to multiple resource types being involved and sometimes with inter-dependence among these (e.g. memory and disk I/O). Moreover, while some types of resources (e.g. CPU time, memory capacity) are easy to partition; other types (e.g. storage and network bandwidth) are not, making it hard to find a parameter that can characterize contention in a shared environment. The contention on these hard-to-partition resources can have a significant impact on a VM's performance and this need to be captured well by the performance model. Virtualization magnifies this impact due to the inherent, underlying sharing and contention [5].

Among the performance and resource management schemes is the relative performance differentiation scheme [6], [7], [8]. It maintains a performance attribute of VMs depending on the differentiation factor or priorities of customer's software application under varying workload conditions by manipulating the resource allocation of each VM. In order to automate the resource management in virtualized and shared resource environments under unpredictable workloads, feedback control methods have been identified as a promising approach because they provide formal methodologies and tools to design the control system and analyze the stability of the management systems [9]. The existing feedback control approaches have used linear modeling and control approaches thus far, disregarding the dominant nonlinear dynamics of shared resource software environments. In particular, the performance properties (e.g., response time) of a VM are nonlinearly related to the shared resource as shown by many existing works (e.g., [10]). In addition, the performance differentiation schemes impose significant nonlinearities on the management system [8]. Consequently, the existing linear feedback control methods typically fail to achieve effective performance differentiation objectives under changing workload conditions and sudden resource demands in virtualized environments.

## II. DATA CENTERS

A data center is defined as "an environmentally controlled centralized facility providing business services by securely

delivering applications and data across a network to remote users". In a traditional data center environment, applications are deployed at different servers to provide necessary security and performance isolation. As more applications are deployed, the number of servers also grows rapidly. This leads to what is referred to as "server sprawl", i.e., a large number of underutilized and heterogeneous servers.

Applications hosted in data center are usually business-critical applications with quality-of-service (QoS) requirements. Such applications typically have time-varying workloads with high peak-to-average ratio, resulting in dynamically changing resource demands. Traditional over-provisioning approaches used for meeting peak demand usually lead to low resource utilization. In addition, the power consumption and cooling costs become great concerns in recent years [12]. According to a report in, the amount of energy used to power the world's data center server's dual in a five-year duration due to primarily to an enhance in require for Internet services, such as music and video downloads. All these difficulties in data center resource management promoted the usage of virtualization technology aiming to produce more cost-efficient data centers, here on referred as virtualized data centers.

## III. DATA CENTER VIRTUALIZATION

Data centers have become increasingly important for hosting business-critical applications. A business relationship typically involves data center owners and application providers. A data center provides resources for hosting applications and application providers pay for what they use. It is often desirable for application providers to be able to lease data-center resources under a "pay-as-you-go" model, and for the data-center providers to be able to multiplex shared resources in a way that guarantees the expected performance of applications. To realize this, the data center must provide flexible and manageable execution environments that are customized for each application without compromising its ability to share resources among applications and delivering to them the necessary performance, security and isolation.

### Benefits of Virtualized Data Centers

There are many benefits to running applications in virtual machines compared to running on physical machines including [13]:

**1. High Utilization:** As sharing resources allows better utilization of overall data center resources.

**2. Performance Isolation:** The applications can be isolated by running them in different virtual machine. Different virtualization technologies differing levels of performance isolation, but most of them offer safety, correctness and fault-tolerance isolation. Though many virtualization technologies can be used for isolation, we choose hypervisor-based virtualization. Though physical machine can offer isolation by running multiple applications in different physical machines, such scenario would waste great amount of resources.

**3. Low Management Costs:** Virtual machines allow easy provisioning of resources in a data center and virtual machine migration allows easy movement of virtual machines to save energy and for other purposes [20].

**4. High Adaptability:** The resources to virtual machines can be dynamically changed allowing highly adaptive data centers. Major portion of this thesis is concerned with creating a dynamic resource control system that can make automatic decisions for high adaptability. However, virtualization is not free, and causes overhead that may not be suitable for some applications.

#### IV. CAPACITY MANAGEMENT TECHNIQUE

Here they discuss about the technology and tools that will help with Capacity Management in your Virtual Data Center and Cloud [11]. They will focus on these three characteristics – Monitoring Capacity, Managing Capacity and Optimizing Capacity.

**Monitoring Capacity:** Monitoring Capacity can be both reactive and proactive. It is reactive when an aware is activated when an upper limit on the exploitation threshold has been violated. The attentive sets off several accomplishments such as recognizing root origin of violate – which VM or Host is/are influenced; which application or service is origin the infringe; is it temporary or stable; etc. It is proactive when you are using past data for investigation to predict future capacity demands and exploitation fashions.

**Managing Capacity:** Managing capacity guarantees that there is always accessible capacity to assemble service demands. To achieve this one has to create use of data from capacity observing, business predict and scheme channels; and input these into a data model to guess how long residual capacity will preceding and when it is essential to get hold of more capacity. Here they can employ a capacity data model using a worksheet to predict capacity or make use of automated capacity planning explanation to help with the model. For an organization promising to cloud services, this

will also assist reduce the possibility and barrier preserved by the organization, thus falling cloud services prices.

Managing capacity also distinguishes and reduces wastage. This occupy detection of VMs in redundant or powered off states for a extended phase; over-assigned or over-sized VMs; and VMs which have expired their rent. In a 3<sup>rd</sup> party cloud, the Service Provider will more often than not have a procedure to get back terminate VMs automatically. Consumer organizations will require monitoring and recognizing redundant, powered-off and enormous VMs which then trigger the suitable process to retrieve the idle capacity to put them back into their resource collections.

**Optimizing Capacity:** Optimizing capacity aspires to maximize the competence and utilization of accessible capacity with no forcing service levels, all the way through the completion of automation and expertise. Example of these automation and expertise would include:

- Compute: over-consign CPU and Memory; and dynamic resource scheduler.
- Storage: thin conditioning, automated storage leveling, storage outlines, storage I/O controller and join together storage resource collections.
- Network: network I/O controller, joined network and WAN optimization.
- In a self-deal with virtual data center and private cloud, optimizing capacity not only assists organizations to enhance the efficiency of their benefits but also decreases their charge of liberation and procedures. For 3<sup>rd</sup> party cloud circumstances, the service provider organizes capacity optimization explanations to improve their liableness as they constrain down their unit charge of calculate, storage and network resources.

In review that capacity management technique, capacity planning and management are an explanation perform in a virtual data center and cloud. Apart from you are administration the virtual data center or cloud infrastructure manually or intense resources from cloud service providers, capacity management will facilitate you accomplish the objective of distributing IT services in the most competent and cost efficient method[11].

Capacity planning for such a situation by using past resource utilization traces to forecast the application resource conditions in the future and to put well-suited positions of applications onto the shared nodes [23]. Such an approach plans to make sure that each node has sufficient facility to meet the comprehensive order of all the applications, while

reducing the number of active nodes. On the other hand earlier period orders are not always precise analysts of expectations demands, in particular for Web-based, interactive applications. But, the CPU and network overheads of VM movement possibly will more degrade application routine on the already-congested node and for this reason VM migration is essentially useful for sustained rather than temporary excess. To overcome the complexity, different tools for organization resources and energy consumption as follow:

- **AutoControl** - resource manager: AutoControl is a feedback-based resource allocation system that deal with dynamic resource sharing inside the virtualized nodes and that match the capacity planning and workload relocation methods others have proposed to accomplish application-level SLOs on shared virtualized infrastructure.
- **LiteGreen** - power manager: LiteGreen is an automated method that combines unused desktops onto a central server to decrease the energy consumption of desktops as a complete. LiteGreen uses virtual machine movement to shift inactive desktops into a central server. The desktop VM is shifted back when the consumer comes back effortlessly and a remote desktop client is used to cover the consequence of movement.

CCM's cloud-scale capacity management solution has three primary allocation phases kesavan et all [14]: (i) Initial Allocation, (ii) Periodic Balancing, and (iii) Reactive Allocation.

(i) **Initial Allocation:** The amount of a resource to be allocated to a cluster or super cluster is captured by the entitlement metric [14]. For an initial allocation that does not yet has runtime resource demand information, the entitlement value is computed using only static allocation constraints.

(ii) **Periodic Balancing:** As mentioned before the periodic balancing algorithm typically runs increasingly infrequently at lower vs. higher levels[14]. Since the granularity of this interval impacts the overhead and accuracy of the CCM capacity balancing solution, administrators are given the ability to configure the resource monitoring and algorithmic intervals to individual deployment scenarios.

(iii) **Reactive Allocation:** In order to deal with unexpected spikes in resource usage CCM uses an additional reactive allocation phase, which is triggered whenever the resource utilization of an entity exceeds some high maximum threshold[14].

## V. RELATED WORK

A cloud computing data centers run various workloads with instance changeable resource demands within virtual machines (VMs). This always shows the ways to regions of

high and low resource use during the cloud infrastructure considerable progress in application resource accessibility, and as a result, taken as a whole data center utilization can be accomplished. But still, they across an entire data center, is a challenging plan. So the author Mukil Kesavan et al was presented a system based on cloud capacity management (CCM) this is on demand CCM enriched with various low-overhead techniques [14]. Self motivated by practical on-field observations and to achieve scalability allocation for thousands of machines. CCM architecture is divided in to three levels that are top-level cloud manager, midlevel super cluster managers, and finally cluster managers at the lowest level. According to the figure 1, hosts are logically grouped into clusters tight with capacity manager (VMware DRS) and these clusters are monitored by super cluster. These super clusters are known as corresponding capacity manager. All these super clusters come under collection of super clusters known as cloud-level capacity manager.

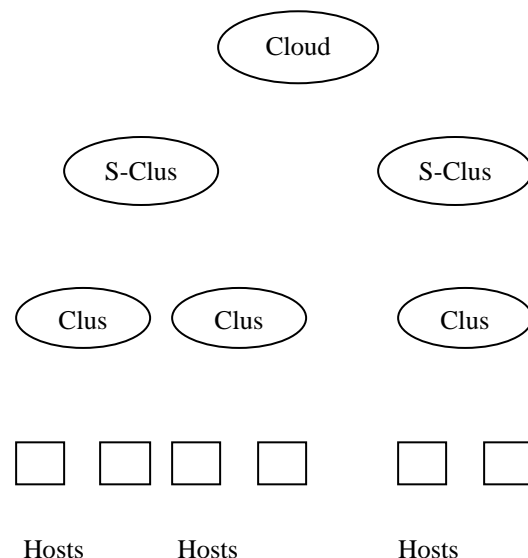


Figure 1: Architecture of Cloud Capacity Management System.

Here author Bryant et al. proposed a prototype of a micro-elastic server called Kaleidoscope to dynamically create small cloned worker VMs to satisfy increased demand on a target VM [15]. They used a novel VM state coloring technique to glean useful semantic information of guest OS page tables and then clone VM states to instantly create replicas of those VMs which can satisfy additional user requests for parent VMs. Although the technique is promising, it did not address how the physical resources should be distributed according to their respective SLAs. Essentially, we view kaleidoscope as a complementary solution that satisfies instantaneous load spikes in user VMs. On the other hand, our revenue driven approach delivers a

more effective resource partitioning when the loads on the VMs are stable and the resource allocation decision is guided by the SLAs.

M. Sedaghat, F. Hernandez-Rodriguez, and E. Elmroth discusses a re-packing approach for coordinating the trade-offs between horizontal and vertical elasticity decisions that must be made to manage the capacity acquired by an elastic application in a cost-effective and on-demand fashion [16]. Adapting application's resource set to changes in load can be quick and cheap in terms of reconfiguration costs following horizontal elasticity decisions; because they only add the extra capacity on the currently deployed VMs; however the resulted resource set could be far from optimal for the aggregated capacity over time. On the other hand, vertical elasticity decisions require frequent, costly and time-consuming reconfiguration of the deployed resource set, such as VMs On/Offs; but they maintain the optimality of the resource set. Therefore there is an inevitable trade-off between costs and benefits of the two methods.

Here author M. Sedaghat, F. Hernandez, and E. Elmroth has addresses the challenge of managing cloud resources in a holistic way to satisfy the business level objectives of an IP [17]. The paper adopts a top-down approach to the development of a unified cloud resource management system, which is outlined. In this approach, the management process is divided across a collection of low level controllers with very distinct responsibilities. Each autonomous controller performs a specific managerial task such as admission, elasticity control, VM Placement, or monitoring and fault management.

Mr. M. Sedaghat, F. Hernandez, and E. Elmroth proposes a Peer to Peer (P2P) resource management framework for cloud data centers [18]. The main objective of the study was to develop a resource management solution that is scalable with respect to both the number of physical servers and incoming Virtual Machine (VM) requests while remaining computationally practical. The framework consists of an agent community that interacts in a goal-oriented P2P fashion and a gossip protocol for information dissemination, discovery and optimization.

Meng et al. pointed out that capacity prediction by sizing individual VMs separately leads to wastage of physical resources [19]. Instead, they proposed a joint VM sizing approach which statistically multiplexes the resource demands of individual clients. They also offered a VM selection mechanism whereby administrators have the ability to group VMs on a physical server based on individual

resource requirements estimated by application-specified SLA models. All of these previous approaches have dealt with VM sizing from the point of view of capacity planning of the data centers. However, none of them have offered any flexibility towards choosing an appropriate VM size based on the clients performance target.

Q. Wang et al explored the problem of providing simultaneous public auditability and data dynamics for remote data integrity check in Cloud Computing [20]. In view of the key role of public auditability and data dynamics for cloud data storage they propose an efficient construction for the seamless integration of these two components in the protocol design. They offered a protocol supporting for fully dynamic data operations principally to support block insertion that is missing in most existing schemes. In the cloud paradigm, by putting the large data files on the remote servers, the clients can be reassured of the burden of storage and calculation. As a client no longer acquires their data locally, it is of vital significance for the clients to ensure that their data are being correctly stored and maintained. Clients should be equipped with certain security means so that they can periodically verify the correctness of the remote data even without the existence of local copies. In this case clients do not unavoidably have time, feasibility or resources to monitor their data; they can delegate the monitoring task to a trusted TPA. They only consider verification schemes with public auditability: any TPA in possession of the public key can act as a verifier [20].

Here Author Armbrust et al presented a survey on cloud computing in [21]. They defined basic terminology of cloud. They also compare cloud with other related technologies. They also try to identifying the top technical and non-technical obstacles and opportunities of cloud computing. Virtualization is primary security mechanism of cloud computing. It is a powerful defense scheme. It capable to defends **alongside** most efforts by customers to show aggression one another or the fundamental cloud infrastructure. One of the common problems is that all recourses are not virtualized. Virtualization software has been acknowledged to restrain listening devices that allow virtualized code to "break loose" to some level. Erroneous network virtualization may perhaps approval to client code right of entry to susceptible sections of the provider's communications or to the assets of other users. Multiple virtual machines (VMs) can share CPUs and main memory surprisingly well in cloud computing. Virtualization is essential to improve architectures and operating systems to efficiently virtualized interrupts and I/O channels [21].

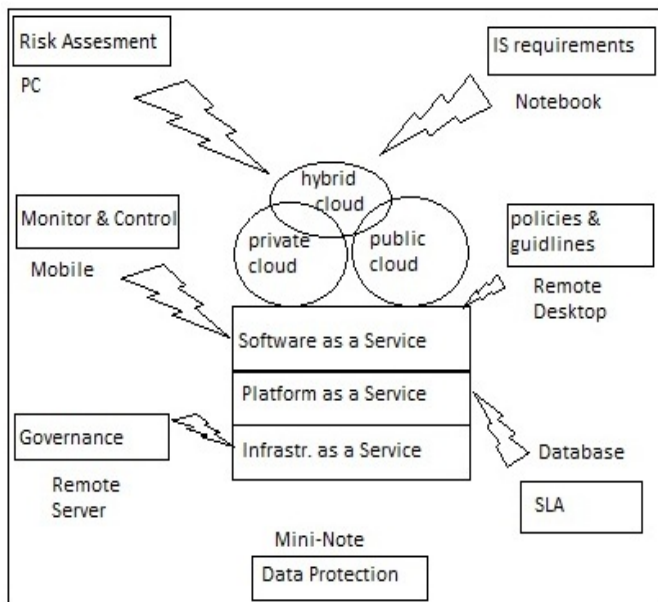


Figure 2: Cloud Data Storage Architecture.

## VI. CLOUD COMPETENCE COMPUTATION

Cloud computing has made the visualization of computing resources as an effectiveness a further step closer to the certainty. As the technology move ahead and network access is converted into earlier and with lower latency, the model of distributing computing power remotely over the Internet will reproduce. Therefore, Cloud data centers are anticipated to produce and collect a superior part of the world's computing resources. In this circumstance, energy-efficient management of data center resources is a fundamental problem in view to both the operating costs Q.Tang et al[12]. The facility to dynamically multiplex datacenter calculate capacity along with workloads with time show a discrepancy requires permits datacenter administrators to oversubscribe resources, evaluated to constituted VM competence or worst-case requires and leads to significant operational efficiency[12]. Transversely a complete datacenter, the monitoring of physical and virtual entities at well granularities creates large network in the clouds. Virtualized datacenter compute capacity multiplexing methods for huge scale situations that focus on accomplishing scalability through easy and uncomplicated techniques. On design of a virtualized management stack that sustains theoretically replicated operations on all virtualized entities like VMs, virtual networks and virtual storage, and the restore of the largest part datacenter automation services like auto-scaling, application deployment and orchestration, rolling upgrades etc.

Cloud computing has its connection to virtualized data centers but with its distinctive characteristics, which makes it to a certain extent demanding. Cloud computing make available different levels of services, which makes observing more difficult. Users are only exposed to high-level servers and do not have the comprehensive information about the resource position[12]. The identical problems probable subsist for Cloud developers and administrators, as the abstract/unified resources more often than not go through virtualization.

- Replication of the global managers would lead to multiple instances of the VM placement algorithm being executed concurrently on multiple controller nodes.
- Scalability and eliminating single points of failure are important benefits of designing a dynamic VM consolidation system in a distributed way.
- Due to usage cloud computing continuous exchange of information between global managers during the process of execution of the VM.

## VII. CONCLUSION

Here in this paper a comprehensive survey of existing techniques that are used for the storage management as well as virtualization of cloud computing is required. Because virtualized data centers have seen explosive growth in recent years, but managing them is still a hard problem. Various factors are analyzed and discussed in this paper so that on the basis of various issues a new and efficient methodology is implemented for the storage management and virtualization of cloud computing to accelerate the progress of autonomous and dynamic resource provisioning of virtual machines in a data center. However, several aspects of our work require further research.

## REFERENCES

- [1] Urgaonkar, B., Rosenberg, A., and Shenoy, P. Application placement on a cluster of servers. In International Journal of Foundations of Computer Science (October 2007), vol. 18, pp. 1023–1041.
- [2] Zhang, Qi, Cherkasova, Ludmila, Mathews, Guy, Greene, Wayne, and Smirni, Evgenia. R-capriccio: a capacity planning and anomaly detection tool for enterprise services with live workloads. In Middleware '07: Proceedings of the ACM/IFIP/USENIX 2007 International Conference on Middleware (New York, NY, USA, 2007), Springer-Verlag New York, Inc., pp. 244–265.
- [3] Stewart, Christopher, Kelly, Terence, Zhang, Alex, and Shen, Kai. A dollar from 15 cents: cross-platform management for internet services. In ATC'08: USENIX 2008 Annual Technical

- Conference on Annual Technical Conference (Boston, Massachusetts, 2008), pp. 199–212.
- [4] Chandra, Abhishek, Gong, Weibo, and Shenoy, Prashant. Dynamic resource allocation for shared data centers using online measurements. In Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and modeling of computer systems (San Diego, CA, USA, 2003), ACM, pp. 300–301
  - [5] Evangelos Kotsovinos. Virtualization: Blessing or Curse? ACM Queue, Jan. 2011.
  - [6] L. Chenyang, T. F. Abdelzaber, J. A. Stankovic, and S. H. Son, "A feedback control approach for guaranteeing relative delays in web servers," in IEEE Real-Time Technology and Applications Symposium, 2001, pp. 51–62.
  - [7] C. Lu, Y. Lu, T. Abdelzaher, J. Stankovic, and S. Son, "Feedback control architecture and design methodology for service delay guarantees in web servers," IEEE Transactions on Parallel and Distributed Systems, vol. 17, no. 9, pp. 1014–1027, 2006.
  - [8] Y. Lu, T. Abdelzaher, C. Lu, L. Sha, and X. Liu, "Feedback control with queueing-theoretic prediction for relative delay guarantees in web servers," in IEEE Real-Time and Embedded Technology and Applications Symposium, may 2003, pp. 208 – 217.
  - [9] T. Patikirikorala, A. Colman, J. Han, and L. Wang, "A systematic survey on the design of self-adaptive software systems using control engineering approaches," in Symposium on Software Engineering for Adaptive and Self-Managing Systems, 2012, pp. 33–42.
  - [10] X. Zhu, Z. Wang, and S. Singhal, "Utility-driven workload management using nested control design," in American Control Conference, june 2006, p. 6 pp.
  - [11] Keng Leong, "Capacity Planning for your Virtual Data Center and Cloud – Part 3 // Cloud, Virtualization" March 12, 2013.
  - [12] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," IEEE Trans. Parallel Distrib. Syst., vol. 19, no. 11, pp. 1458–1472, 2008
  - [13] W. Dally and B. Towles. Principles and Practices of Interconnection Networks. Morgan Kaufmann Publishers Inc., 2004.
  - [14] Kesavan, Mukil, Irfan Ahmad, Orran Krieger, Ravi Soundararajan, Ada Gavrilovska, and Karsten Schwan "Practical Compute Capacity Management for Virtualized Datacenters", IEEE Transactions On Cloud Computing, Vol. 1, No. 1, pp. 88 – 100, 2013.
  - [15] Roy Bryant, Alexey Tumanov, Olga Irzak, Adin Scannell, Kaustubh Joshi, Matti Hiltunen, H. Andres Lagar-Cavilla, and Eyal De Lara. Kaleidoscope: Cloud micro-elasticity via vm state coloring. In Proceedings of the sixth conference on Computer systems (EuroSys), pages 273–286, 2011.
  - [16] M. Sedaghat, F. Hernandez-Rodriguez, and E. Elmroth, "A virtual machine re-packing approach to the horizontal vs. vertical elasticity trade-off for cloud autoscaling," in Proceedings of the 2013 ACM Cloud and Autonomic Computing Conference, p. 6, ACM, 2013.
  - [17] M. Sedaghat, F. Hernández, and E. Elmroth, "Unifying cloud management: Towards overall governance of business level objectives," in 11<sup>th</sup> IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), pp. 591–597, IEEE, 2011.
  - [18] M. Sedaghat, F. Hernández, and E. Elmroth, "Peer to peer resource management for cloud data centers," Submitted to 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), 2014.
  - [19] Xiaojiao Meng, Canturk Isci, Jeffrey Kephart, Li Zhang, Eric Bouillet, and Dimitrios Pendarakis. Efficient resource provisioning in compute clouds via vm multiplexing. In ICAC, 2010.
  - [20] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 5, pp. 847–859, May 2011.
  - [21] Armbrust, Michael, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee et al. "A view of cloud computing." Communications of the ACM, vol. 53, no. 4, pp. 50–58, 2010..
  - [22] Qian Wang, Cong Wang, Jin Li1, Kui Ren, and Wenjing Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing" Proceedings of the 14th European conference on Research in computer security(ESORICS'09), pp. 355–370, 2009.
  - [23] Rolia, J., Cherkasova, L., Arlit, M., and Andrzejak, A. A capacity management service for resource pools. In Proceedings of International Workshop on Software and Performance (July 2005).
  - [24] Clark, C., Fraser, K., Hand, S., Hansen, J. G., Jul, E., Limpach, C., Pratt, I., and Warfield, A. Live migration of virtual machines. In Proceedings of the 2nd Symposium on Networked Systems Design and Implementation (NSDI) (2005).