PARALLEL PRECEDENCE CONSOLIDATION FOR SIMILAR WORKLOAD IN CLOUD

A. Srinivas[1], Stalin Babu J[2], A. Praveen Kumar [3], K. Gurunadha Gupta[4]

[1]*Assistant Professor, Computer Science Engineering,*
Sri Indu College of Engineering and Technology, Telangana, India

[2]*Assistant Professor, Computer Science Engineering,*
Nalla Malla Reddy Engineering college, Telangana, India

[3]*Assistant Professor, Computer Science Engineering,*
Swami Ramananda Tirtha institute of Science &, Telangana, India

[4]*Assistant Professor, Information Technology,*
Sri Indu College of Engineering and Technology, Telangana, India

Abstract:
Cloud computing allows business customers to scale up and down their resource usage based on needs. Many of the routed gains in the cloud model come from resource multiplexing through virtualization technology. In this paper, we present a system that uses virtualization technology to allocate data center resources dynamically based on application demands and support green computing by optimizing the number of servers in use. We introduce the concept of "skewness" to measure the unevenness in the multi-dimensional resource utilization of a server. By minimizing skewness, we can combine different types of workloads nicely and improve the overall utilization of server resources. We develop a set of Green Computing techniques that prevent overload in the system effectively while saving energy used. Trace driven simulation and experiment results demonstrate that our algorithm achieves good performance by Resource Management.
Key words: Cloud Computing, Resource Management, Virtualization, Green Computing.

1. INTRODUCTION:

Security and privacy is one fundamental obstacle to cloud computing success. In this context, we've discussed several critical security challenges that current research thrusts aren't yet addressing. Some researchers report finding serious problems with two cloud systems and say these flaws probably exist in other cloud architectures. The team said the vulnerabilities could let attackers gain administrative rights to host systems. The investigators found flaws with Amazon Web Services (AWS) and informed Amazon, which has since patched the problems. They also discovered vulnerabilities with the open source Eucalyptus private-cloud software framework. The Ruhr University team used XML signature-wrapping attacks to gain administrative access to AWS customer accounts. In these attacks, a hacker changes the content of the signed part of a message sent via the SOAP protocol without invalidating the signature itself. The investigators say they are working on techniques that can be used with existing XML security approaches to eliminate this flaw. They also report that AWS was susceptible to cross-site scripting attacks, in which hackers inject a malicious script into web pages. According to team members, their cloud attacks were general in nature and did not target Amazon's system and thus would probably be effective against multiple cloud architectures. Within the cloud computing world, the virtual environment let users access computing power that exceeds that contained within their own physical worlds. To enter this virtual environment requires them to transfer data throughout the cloud. Consequently, several data storage concerns can arise. Typically, users will know neither the exact location of their data nor the other sources of the data collectively stored with theirs.

For example, is security solely the storage provider's responsibility, or is it also incumbent on the entity that leases the storage for its applications and data? Furthermore, legal issues arise, such as e-discovery, regulatory compliance (including privacy), and auditing. The range of these legal concerns reflects the range of interests that are currently using or could use cloud compute. These issues and their yet-to-be-determined answers provide significant insight into how security plays a vital role in cloud computing continued growth and development. From current cloud computing practices and inherent risks involved, it is clear that at present there is a big scope for unsolved security threats and lack of risk analysis approaches in the cloud computing environments. A proper risk analysis approach should be developed that helps to both the service providers and the customers. Therefore, it is inevitable and should find solutions for threats and risks due to vulnerabilities analyzed. Hence the customers can be guaranteed data security and the service providers can win the trust of their customers. Also the cloud users can perform the risk analysis before placing their critical data in a security sensitive cloud. Further additions to the matrix and inclusion of additional variables for assessment should be considered as cloud computing progresses to advanced levels where new risks could materialize.

2.EXISTING:

Existing parallel scheduling mechanisms normally take responsiveness as the top priority and need nontrivial effort to make them work for data centers in the cloud era. In this paper, we propose a priority-based method to consolidate parallel workloads in the cloud.

4.PROPOSED:

The cloud computing paradigm is attracting an increased number of complex applications to run in remote data centers. Many complex applications require parallel processing capabilities. Parallel applications of certain nature often show a decreasing utilization of CPU resources as parallelism grows, mainly because of the communication and synchronization among parallel processes. It is challenging but important for a data center to achieve a certain level of utilization of its nodes while maintaining the level of responsiveness of parallel jobs. To overcome these and other concerns, we must develop a security model that promotes CIA. This model could enable each cloud to offer a measure of its to date and projected CIA, but the obvious difficulty is that obtaining security data is difficult, if not impossible. This problem has existed since computing advent due to financial, business, and national security concerns. It might be exacerbated in cloud computing because the need to provide data confidentiality can also impact incident reporting

To ensure data confidentiality, integrity, and availability (CIA), the storage provider must offer capabilities that, at a minimum, include a tested

1.encryption schema to ensure that the shared storage environment safeguards all data;
2.stringent access controls to prevent unauthorized access to the data; and
3.Scheduled data backup and safe storage of the backup media. Security is implicit within these capabilities, but further fundamental concerns exist, which requires attention.

5.METHODOLOGY:

Cloud computing is the provision of dynamically scalable and often virtualized resources as a services over the internet Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them. Cloud computing represents a major change in how we store information and run applications. Instead of hosting apps and data on an individual desktop computer, everything is hosted in the "cloud"—an assemblage of computers and servers accessed via the Internet.

6. KEY CHARACTERISTICS:

1. Agility improves with users' ability to re-provision technological infrastructure resources.

2. Cost is claimed to be reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure. This is purported to lower barriers to entry, as infrastructure is typically provided by a third-party and does not need to be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is fine-grained with usage-based options and fewer IT skills are required for implementation. The e-FISCAL project's state of the art repository contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.

3. Virtualization technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.

4. Multi tenancy enables sharing of resources and costs across a large pool of users thus allowing for:

5. Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)

6. Utilization and efficiency improvements for systems that are often only 10–20% utilized.

7. Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.

8. Performance is monitored and consistent and loosely coupled architectures are constructed using web services as the system interface.

9. Security could improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. However, the complexity of security is greatly increased when data is distributed

over a wider area or greater number of devices and in multi-tenant systems that are being shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

10. Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

6.1. Priority Based Scheduling:

1. Each process is assigned a priority. Process with highest priority is to be executed first and so on.

2. Processes with same priority are executed on first come first serve basis.

3. Priority can be decided based on memory requirements, time requirements or any other resource requirement.

6.2 First Come First Serve (FCFS):

1. Jobs are executed on first come, first serve basis.

2. Easy to understand and implement.

3. Poor in performance as average wait time are high.

3.1.3. Parallel job scheduling

1. System performance depends on the workload Analogy: algorithm performance depends on the input
2. Evaluation workload should be representative of real workloads
3. In our case, the workload is a sequence of jobs to run
4. Each job is a rectangle in processors times pace
5. Given many jobs, we must schedule them to run on available processors

6.3 Backfilling:

Quinn Snell et al. [5] proposed a backfilling scheme. Backfilling is a space sharing optimization that tries to balance between the goals of utilization and maintaining FCFS order. It allows small jobs to move ahead and run-on processors that would otherwise remain idle. This is done to avoid situations in which the FCFS order incompletely violated and some jobs are never run. There are two types of backfilling algorithms:

- Conservative Backfilling
- Aggressive Backfilling

1) Conservative Backfilling: The conservative backfilling approach. In this scheme, jobs are scheduled according to the order of arrival time when there is enough number of processors. If not,

another job with later arrival time and smaller jobs are scheduled torn. It provides reservation to all jobs and limits the slowdown.

2) EASY Backfilling: They also proposed an aggressive approach, provides a reservation to only the job at the head of the job queue and only allow job at the head of the queue can be pre-empt other jobs. It does not have a guaranteed response time of the user job at the time of job submission.

6.4 Gang Scheduling

The main alternative to batch scheduling is gang scheduling that schedules related threads or processes to run simultaneously on different processors. It is a time sharing optimization technique. This scheduling is used if two or more threads or processes are communicated with each other. The problems with gang scheduling is that the requirement that all a job's processes always run together causes too much fragmentation and context switching overhead. Gang scheduling is based on a data structure called ouster haut matrix. In this matrix each row represent time slice and each column represent a processor. The threads or processes of a job are packed into a row of the matrix.

6.5 Backfilling Gang Scheduling

The Backfilling gang scheduling (BGS) method. It is an optimization technique which combines gang scheduling and backfilling scheduling. This scheduling can be done by treating each of the virtual machines created by gang-scheduling as a target for backfilling. Which produce better results than individual approaches gang scheduling or backfilling.

6.6 Migration Gang Scheduling:

The Migration Gang Scheduling method. The process of migration embodies moving a job to any row in which there are enough free processors to execute that job. There are two options for migrate a job from a source row to a target row.

6.7 Migration Backfilling Gang Scheduling:

The method that the migration embodies moving a job to any row in which there are enough free processors to execute that job. If we cannot replicate a job in a different row because it's set of processors are busy with another job, attempt to move the blocking job to a different set of processors. A job can appear in multiple rows of the matrix, but it must occupy the same set of processors in all the rows. This rule prevents the Ping-Pong of jobs.

6.8 Job kill based EASY Backfilling (KEASY):

The scheduling scheme Job kill based EASY backfilling (KEASY). It is capable of dispatching a job to run in background VMs while it is not qualified for backfilling according to EASY. There is a chance that the corresponding foregrounds VMs of those Background VMs are idle during the jobs lifetime, which leads to performance improvement.

6.9 Reservation based EASY Backfilling (REASY):

The Reservation based EASY backfilling (REASY) scheme. In this scheme job kill is not allowed in the scheduling; once a job is deployed onto background VMs of a set of pro- cessors, its run is pinned onto this set of processors. Only the entire foreground VMs of this set of processors are available can this job run in the foreground. For the reservation making, if a reservation is being made for a job running in the background tier, the shadow time is the last termination time of the jobs running in its foreground VMs; the extra foreground VMs are the ones now idle and no process of the job is running in their background VMs.

6.10 Conservative Migration Supported:

The Migration Supported Backfilling (CMBF), which is same as Conservative backfilling. Only the difference is, the scheduler is able to suspend a job and resume it on other nodes in a later time. This algorithm avoids starving a preempted job. When the number of jobs in the queue is large, the cost can be high because CMBF requires tracking backfilling jobs for each job in the queue.

6.11 Aggressive Migration:

The Aggressive Migration Supported Backfilling (AMBF). It only tracks backfilling jobs for the job at the head of the queue and allows the head-of queue job to preempt other jobs. The rest of jobs in the queue are not allowed to pre-empt jobs.

6.12 Priority-based Consolidation Method:

In priority-based method to consolidate parallel workloads in the cloud, dividing computing capacity into two tiers: foreground tier and background tier. The VM running in foreground is assigned a high CPU priority and the VM running in background is assigned a low CPU priority. There are two priority-based methods to consolidate parallel workloads in the cloud: Conservative migration and consolidation supported backfilling (CMCBF) and Conservative migration and consolidation supported backfilling (AMCBF).

1) Conservative Migration and Consolidation supported Backfilling that allows jobs to run in background VMs simultaneously with those foreground VMs to improve node utilization. It ensures that a job is dispatched to foreground VMs whenever the foreground VMs are idle or that job satisfies the node requirement. It allows jobs to run in background VMs simultaneously with those foreground VMs to improve node utilization. Compared to CMBF, CMCBF also deals with how to ensure that the background workload does not affect the foreground job. CMCBF only dispatches a job to run in background VMs when the corresponding foregrounds VMs have utilization lower than a given threshold. The foreground VM utilization can be obtained from the profile of foreground jobs, or from the runtime monitoring data.

2) Aggressive Migration and Consolidation supported Backfilling: The priority based consolidation method, CMCBF faces similar problem as CMBF when racking backfilling jobs for each job. To reduce the cost, new modified algorithm is Aggressive Migration and Consolidation supported Backfilling (AMCBF). Only the jobs in head-of queue can preempt other jobs in AMCBF.

7.RESULTS:

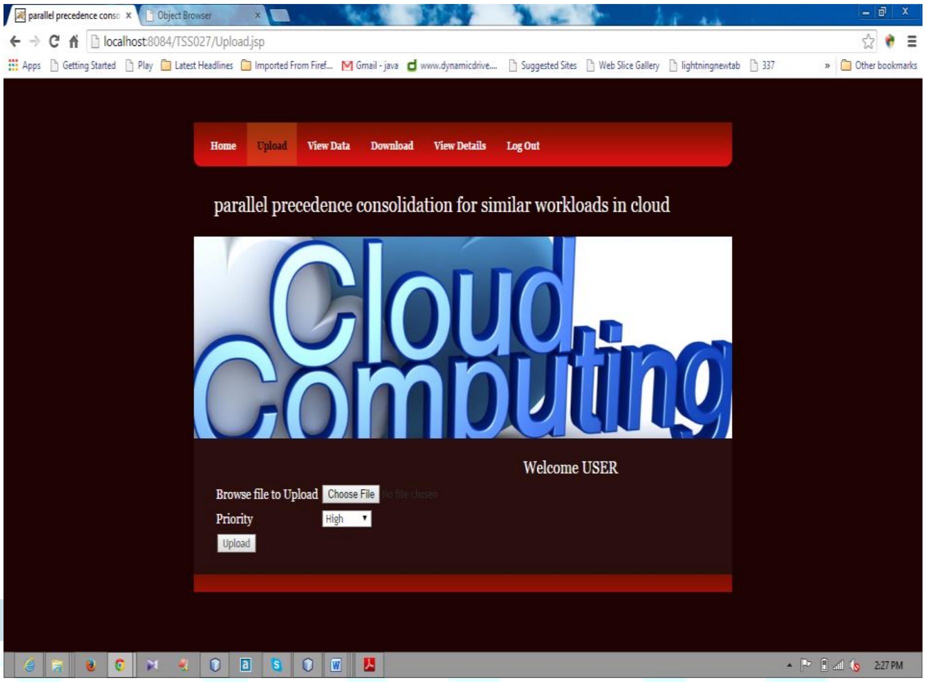

Fig1.Priority based file

Fig 2 Priority based file uploaded

Fig 3. Priority based file extracted

8.CONCLUSION:

Although security and privacy services in the cloud can be fine-tuned and managed by experienced groups that can potentially provide efficient security management and threat assessment services, the issues we've discussed here show that existing security and privacy solutions must be critically reevaluated with regard to their appropriateness for clouds. Many enhancements in existing solutions as well as more mature and newer solutions are urgently needed to ensure that cloud computing benefits are fully realized as its adoption accelerates. Cloud computing is still in its infancy, and how the security and privacy landscape changes will impact its successful, widespread adoption. We have presented the design, implementation, and evaluation of a resource management system for cloud computing services. Our system multiplexes virtual to physical resources adaptively based on the changing demand. We use the skewness metric to combine VMs with different resource characteristics appropriately so that the capacities of servers are well utilized. Our algorithm achieves both overload avoidance and green computing for systems with multi-resource constraints.

REFERENCES:

[1] M. Armbrust *et al.*, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep., Feb 2009.

[2] L. Siegele, "Let it rise: A special report on corporate IT," in *The Economist*, Oct. 2008.

[3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho,R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of
virtualization," in *Proc. of the ACM Symposium on Operating Systems Principles (SOSP'03)*, Oct. 2003.

[4] "Amazon elastic compute cloud (Amazon EC2),http://aws.amazon.com/ec2/."

[5] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proc.of the Symposium on Networked Systems Design and Implementation (NSDI'05)*, May 2005.

[6] M. Nelson, B.-H. Lim, and G. Hutchins, "Fast transparent migration for virtual machines," in *Proc. of the USENIX Annual Technical Conference*,2005.

AUTHORS:

A.Srinivas, Post Graduated in Computer Science & Engineering (M.Tech) From JNT University, Hyderabad in 2009 and Graduated in Computer Science & Information Technology (B.Tech) from JNTU, Hyderabad in 2004.He is currently working as an Assistant Professor, Department of Computer Science & Engineering in Sri Indu College of Engineering & Technology (SICET), (V) Sheriguda, (M) Ibrahimpatnam, R.R.Dist, and Telangana, India. He has 9+ years of Teaching Experience. His research interests include Cloud Computing, Data Mining, Information Security,Software Testing, Wireless Networks and Software Quality.

Stalin babu J, Post Graduated in Computer Science & Engineering (M.Tech) From JNT University, Kakinada in 2011.He is currently working as an Assistant Professor, Department of Computer Science & Engineering in Nalla Malla Reddy Engineering College (NMREC), (V) Kachvanisingaram, (M) Ghatkesar, R.R.Dist, Telangana, India. He has 8 years of Teaching Experience. His research interests include Cloud Computing, Data Mining, and Information Security.

A.Praveen kumar, Post Graduated in Computer Science & Engineering (**M.Tech**) From JNT **University**, Hyderabad in 2010.He is currently working as an Assistant Professor, Department of Computer Science & Engineering in Swami Ramananda Tirtha Institute of Science & Technology, Nalgonda, Telangana, India. He has 8 years of Teaching Experience. His research interests include Cloud Computing, Data Mining.

K Gurnadha Guptha, working as Assistant Professor in Sri Indu College of Engineering and Technology He has done a post-graduate from Swarnandhra College of Engineering and Technology, Narasapur, west Godavari,AP. He has done a graduate from ANNA UNIVERSITY, CHENNAI, TN. His main research interests are Data warehousing and Mining, Distributed Database System.