WORKLOAD CONSOLIDATION FOR CLOUD DATA CENTERSWITH GUARANTEED QOS USING REQUEST RENEGING

AUTHOR

Dr.V.K.Manavalasundaram
Assistant Professor
Department of IT
Velalar College of Engineering and Technology
Erode

STUDENTS

D.Vinodhini
Final B.Tech IT
Velalar College of Engineering and Technology
Erode

G.Sandhiya
Final B.Tech IT
Velalar College of Engineering and Technology
Erode

B.Maneshraja
Final B.Tech IT
Velalar College of Engineering and Technology
Erode

ABSTRACT

Cloud computing is rapidly growing and many more cloud providers are emerging.Cost efficiency and resource cost maximization become two major concerns of cloud providers to remain competitive while making profit.Outline novel economics-inspired Work load consolidation mechanisms to tackle the profit maximization problem from the perspective of a cloud provider acting solely. Admission control mechanisms tailored within a Profit management framework to maximize resource cost has been proposed. Existing abstractions for in-memory storage on clusters, such as distributed shared memory, keyvalue stores, databases, and Piccolo offer an interface based on fine-grained updates to mutable state (e.g., cells in a table). Cost effective Work load consolidation based on following strategies are Cost Efficiency of the Cloud: Cost reductions and profit increases, Pay-as-you-go pricing, Implications of multi tenancy. Scheduling and Work load consolidation as a cost efficient solution: Exploitation of application characteristics, Explicit consideration of user experience/satisfaction.

INTRODUCTION

Cloud computing providing unlimited infrastructure to store and execute customer data and program. Customers do not need to own the infrastructure, they are merely accessing or renting they can forego capital expenditure and consume resources as a service, paying instead for what they use. Benefits of Cloud Computing: Minimized Capital expenditure, Location and Device independence, Utilization and efficiency improvement. Very high Scalability. High Computing power. Using a rich set of operators. The main challenge in designing RDDs is defining a programming interface that can provide fault tolerance efficiently. Existing abstractions for in-memory storage on clusters, such as distributed shared memory, key value stores , databases, and Piccolo , offer an interface based on fine-grained updates to mutable state (e.g., cells in a table). Both approaches are expensive for data-intensive workloads, require copying large amounts of data over the cluster network, whose bandwidth is far lower than that of RAM, and incur substantial storage overhead. RDDs provide an interface based on coarse-grained transformations (e.g., map, filter and join) apply the same operation to many data items. Allows to efficiently provide fault tolerance by logging the transformations used to build a dataset (its lineage) rather than the actual data.

WORKLOAD CONSOLIDATION COST OPTIMIZATION

Cloud computing has emerged as important computing technology and its pay-as-you-go cost structure enabled the providers to offer computing service on demand and pay for the resources just as utility computing. Virtualization is the important process which allows the sharing of computing resources in online. The computing resources are of different types. These includes Infrastructure as a service (Iaas) which provides the capability to the consumer to provision network, storage and processing. It can include the operating system and applications. Eg Amazon EC, Open Nebulla, Eucalyptus. Platform as a service(Paas)provides the capability to the consumer to acquire applications created using programming languages, deploy onto the cloud infrastructure and tools supported by the provider. Eg., Hadoop, Microsoft Windows Azure, Google App Engine.

Software as a service (Saas) provides the capability to the consumer to use the applications of the provider which runs on cloud infrastructure. Eg., Google Apps, SalesForce.com, Eye OS. When there is any requirment for the users in the cloud, the cloud system provides the required resources to the users by creating virtual machines(VM) in the host machine.The tasks of the users are in the form of workflow.

GROUP AND REAL WORKFLOW OPTIMIZATION ON CLOUD

A workflow is a depiction of a sequence of operations, declared the work of a person, work of a simple or complex mechanism, work of a group of persons, work of an organization of staff, or machines. Workflow may be seen as any abstraction of real work, segregated in work share, work split or whatever types of ordering. Workflows are designed to achieve processing intents of some sort, such as physical transformation, service provision, or information processing.

DATA PROCESSING FLOWS ON THE CLOUD

Scheduling data processing workflows (dataflows) on the cloud is a very complex and challenging task. It is essentially an optimization problem, very similar to query optimization, that is characteristically different from traditional problems in two aspects: Its space of alternative schedules is very rich, due to various optimization opportunities that cloud computing offers; its optimization criterion is at least two-dimensional, with monetary cost of using the cloud being at least as important as query completion time.Scheduling of data flows that involve arbitrary data processing operators in the context of three different problems:

1)Minimize completion time given a fixed budget,2)Minimize monetary cost given a deadline, and 3)Find trade-offs between completion time and monetary cost without any a-priori constraints

EVALUATION OFRESOURCE PROVISIONINGALGORITHMS

A recent advancement wherein IT infrastructure and applications are provided as‘services’ to end-users under a usage-based payment model. It can leverage virtualized services even on the fly based on requirements (workload patterns and QoS) varying with time. The application services hoste dunder Cloud computing model have complex provisioning, composition, configuration, and deployment requirements. Evaluating the performance of Cloud provisioning policies, application workload models and resources performance models in a repeatable manner under varying system and user configurations and requirements is difficult to achieve.

To overcome this challenge, CloudSim: The CloudSim toolkit supports both system and behavior modeling of Cloud system components such as data centers, virtual machines (VMs) and resource provisioning policies.It implements generic application provisioning techniques that can be extended with ease and limited effort.

Currently, it supports modeling and simulation of Cloud computing environments consisting of both single and inter-networked clouds (federation of clouds). It exposes custom interfaces for implementing policies and provisioning techniques for allocation of VMs under inter-networked Cloud computing scenarios. Several researchers from organizations, such as HP Labs in U.S.A., are using CloudSim in their investigation on Cloud resource provisioning and energy-efficient management of data center resources. The usefulness of CloudSim is demonstrated by a case study involving dynamic provisioning of application services in the hybrid federated clouds environment. The federated Cloud computing model significantly improves the application QoS requirement sunder fluctuating resource and service demand patterns

PROPOSED SYSTEM

Proposed framework through large-scale simulations, driven by cluster-usage traces that are provided by Google. A Event handler-DAG based DHT scheduling algorithm that generates VM requests based on the user resource usage in these traces. Under pricing conditions that are aligned with those of Amazon EC21, our admission control algorithms substantially increase resource cost for the provider.To maximize the profit, a service provider should understand both service charges and business costs, and how they are determined by the characteristics of the applications and the configuration of a Work load consolidation system. Event handler-DAG is to treat a Work load consolidation system is a queuing model, such that our optimization problem can be formulated and solved analytically.

EXISTING SYSTEM

Scientific applications partially or entirely shifting from traditional computing platforms (e.g., grid) to the cloud. Due to the pay-as-you-go computational behaviour, performance and (monetary) cost optimizations have recently become a hot research topic for workflows in the cloud. To address the limitations of current approaches, propose Profit Maximization, a transformation-based optimization framework for optimizing the performance and cost of workflows in the cloud.

ADVANTAGE

The advantages of request reneging on minimizing the required processing rates and average response times with QoS guarantees. We also compare the performance of the system with and without request reneging using two different service types. In each run, we calibrated the maximum processing rate of each VM by assigning a cap to each VM’s Virtual CPU (VCPU) that limits the maximum amount of processing rates a VM receives from its allocated physical core. We then applied the traditional binary search method to find the minimum required processing rates that can meet the given QoS conditions.

DISADVANTAGE

Having validated the performance improvement of request reneging, we now compare the completion ratios and average response times when serving requests in a multiplexed manner (e.g., GWPC) and a split manner(eg..SPT) The purpose of the research work targeted in the project was not intended to develop a problem-specific algorithm but rather to investigate how a generic optimization tool based on cloud can be used to solve task planning optimization problems without major modifications to the optimization algorithm itself.

CONCLUSION

Building a distributed computing infrastructure using smart phones for enterprises, technical challenges in building such an infrastructure. The viability and efficacy of various components within novel scheme (Min-Min ToF) for virtual Work load consolidation on a SOC, with three key contributions listed below. Optimization of task's Work load consolidation under user's budget.

It proves its optimality using the CWC conditions in the convex-optimization theory. Maximized resource utilization based on ToF: In order to further make use of the idle resources, Design a dynamic algorithm by combining the above algorithm with ToF and the arrival/completion of new tasks.

Min-Min could get an improvement on Mobile throughput by 15 percent 60 percent than the traditional methods used in P2P Grid model, according to the simulation. Experiments confirm the designed Min-Min protocol with lightweight query overhead is able to search qualified resources very effectively.

REFERENCES

1 Herald Kllapi and Eva Sitaridi "Schedule Optimization for Data Processing Flows on the Cloud "/," in Proc. Int., 2011.

2 Maciej Malawski,E.-K. Byun, Y.-S. Kee, J.-S. Kim, and S. Maeng, "Cost optimized provisioning of elastic resources for application workflows, "Future Gen. Comput. Syst., vol. 27, pp. 1011–1026, 2011.

3 H. Wang, Q. Jing, R. Chen, B. He, Z. Qian, and L. Zhou,"Distributed systems meet economics: pricing in the cloud," inProc. HotCloud, 2010, pp. 1–7.2013.

4 Hero dotos Hero dotou and S. Papadimitriou, "Profiling, What if Analysis, and Cost based Optimization of Map Reduce Programs," in Proc. Int. Workshop Data Manage. New Hardware, 2011, pp. 50–55.

5 F. Busching, G. Berriman, S. Schildt, and L. Wolf, "Cost-driven Scheduling of Grid Workflows Using Partial Critical Paths," in Proc. 32nd Int. Conf. Distrib. Comput. Syst. Workshop, Jun. 2012, pp. 114–117

6 Jim Gray, Goetz Graefe,"The Five-Minute Rule Ten Years Later, and Other Computer Storage Rules of Thumb"-0911b.pdf, 1997.

7 Richard T.B. Ma, Dah-mingChiu,"Internet Economics: The use of Shapley value for ISPsettlement_"-0911b.pdf, 2011.

8 M. Y. Arslan, S. Abrishami, Jia Yu, S. Singh, H. V. Madhyastha, K. Sundaresan, and S. V. Krishnamurthy, "Computing while charging: Building a distributed

9 computing infrastructure --using smart phones," in Proc. 8th Int. Conf. Emerging Netw. Experiments Technol., Dec. 2012, pp. 193–204.

JiaYu ,Rajkumar Buyya, and L. Wolf, "A Taxonomy of Workflow Management Systems for Grid Computing," in Proc. IEEE Int. Conf. Green Comput. Common. IEEE Internet Things and IEEE Cyber, Phys. Social Comput., Aug. 2013, pp. 1986–1991.