Scheduling in Cloud Block Storage Systems

Ioannis Papapanagiotou,
Department of Computer and Information Technology
Purdue University
Email:ipapapan@purdue.edu

NSFCloud Workshop 2014 Position Paper

Abstract

The increasing demand for elastic and scalable cloud block storage requires flexible and efficient ways to provision volumes. The scheduling of volume requests in physical storage nodes or virtualized storage pools is usually based on a single criterion, such as the available capacity or the number of volumes per backend. Those properties are exposed to the cloud block storage scheduler through drivers, and may vary based on the workload. Hence, most cloud storage providers refrain from describing Service Level Objectives (SLOs). Our simulation results indicate that our SLA-aware scheduling policies can improve the SLA violation by 20%, require fewer storage nodes (hence lower capital expenses) and can provide higher volume I/O throughput performance compared to the default policies. In this paper, we will present our vision on the design and implementation of new scheduling algorithm for Cloud block storage systems in an actual infrastructure like the CloudLab or Chameleon.

Description

As applications and databases deployed on the cloud grow over time, there is an eventual need to increase the storage capacity. Adding physical storage or upgrading to a storage system with higher capacity requires high capital investment and can be unscalable. Cloud block storage offers a reliable, high-performance, on-demand storage for applications hosted on the cloud. In addition, it can support resource-hungry workloads seamlessly, and can scale very fast by just asking for more space, i.e. a volume extend request. However, the problem of scheduling storage requests in a manner that maximizes the use of resources, minimizes the energy consumption and maintains the Service Level Agreements (SLA) between the tenant and the cloud service provider has not been properly studied. This problem becomes even more relevant, as new open source IaaS solutions, like OpenStack [3], are being deployed in commercial practice by more than 200 companies, or in research cloud infrastructures like CloudLab.

Our work is focused on improving the block storage service scheduler based on the principles of the Cinder project. Cinder virtualizes pools of block storage devices and provides end users with a self service API to request and consume those resources without requiring any knowledge of where their storage is actually deployed or on what

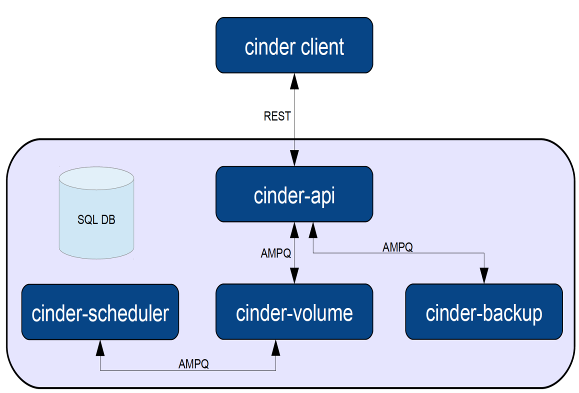

Figure 1: System Components of OpenStack Cinder's Service

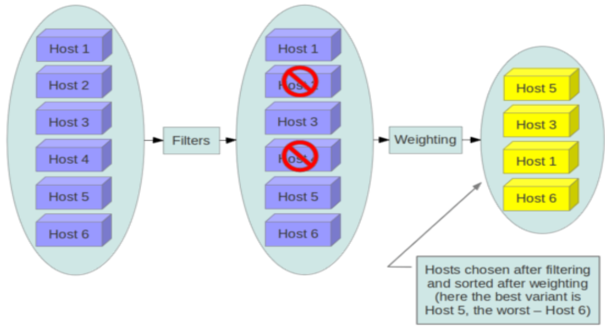

Figure 2: Filtering and Weighting in Cinder's Scheduling

type of device. An overview of the Cinder architecture is shown in Fig. 1. The cinder-scheduler module is responsible for scheduling and routes requests to the appropriate volume service running on storage entity.

Fig. 2 shows the scheduling in OpenStack. It works in two steps: (i) filtering and (ii) weighting. When the Cinder Filter Scheduler receives a request for a logical volume, it first applies filters to determine which hosts are eligible for consideration when dispatching a request. The filter's decision is binary: either a host is accepted by the filter, or it is rejected. It is based on capabilities of the host, namely free space, location and I/O throughput. The default scheduling policy in Cinder is the capacity scheduler. Once the scheduler receives a volume request, it filters out the hosts which do not have sufficient capacity and creates a list of eligible hosts. Then the weight of each host is calculated based on the available capacity. The scheduler selects the host with the largest available capacity as the best candidate to serve the request. However, this is not always the optimal selection as one node may be overloaded with read/write IOPS.

On the other hand, a great amount of work has been devoted in optimizing management of a data center. Most studies have focused on the optimal VM placement in data centers [5], or energy aware VM allocation [1]. More recent studies have focused on adding SLA during the VM placement process [2], or optimizing the job completion time in big data clusters [4]. In our recent work [6], we proposed an SLA-aware resource scheduling for cloud storage, assuming one request per unit time. Hence, the problem could be solved with relatively simple scheduling operations. The main objective of this work is to present our vision, and discuss the research challenges in the volume allocation in cloud storage systems that are comprised of multiple storage entities and multiple requests on a unit time. We will showcase a novel dynamic scheduling algorithm that can provide efficient and rapid allocation of volume requests, and minimize the number of backend systems. We are looking forward to validating our results using the NSF supported Cloud infrastructures.

References

[1] Anton Beloglazov, Jemal Abawajy, and Rajkumar Buyya. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. *Future Generation Computer Systems*, 28(5):755–768, 2012.

[2] Kuan Lu, Ramin Yahyapour, Philipp Wieder, Constantinos Kotsokalis, Edwin Yaqub, and Ali Imran Jehangiri. QoS-Aware VM Placement in Multi-domain Service Level Agreements Scenarios. In *Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on*, pages 661–668. IEEE, 2013.

[3] OpenStack. Openstack. `www.openstack.org`, 2014. [Online; accessed 2-Nov-2014].

[4] Malte Schwarzkopf, Andy Konwinski, Michael Abd-El-Malek, and John Wilkes. Omega: flexible, scalable schedulers for large compute clusters. In *Proceedings of the 8th ACM European Conference on Computer Systems*, pages 351–364. ACM, 2013.

[5] Jing Xu and Jose AB Fortes. Multi-objective virtual machine placement in virtualized data center environments. In *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*, pages 179–188. IEEE, 2010.

[6] Zhihao Yao, Ioannis Papapanagiotou, and Robert D Callaway. SLA-aware Resource Scheduling for Cloud Storage . In *IEEE International Conference on Cloud Networking (CloudNet)*. IEEE, 2014.