

A Study & Comparative Analysis of performances of database on Clouds

Gaurav Sharma, R.S. Sharma

Abstract— Cloud computing is an emerging technology that mainly provides a shared pool of configurable computing resources. It is combination of Internet technologies such as SOA, Web 2.0, Web services, System management such as data centers automation, Distributed Computing such as utility & Grid computing and Hardware such as multi-core chip and virtualization. The services are categorized into three distinct environments Software-as-a-service, Platform-as-a-service, and Infrastructure-as-a-service. Now-a-days, a new emerging service called Database-as-a-service which is a part of software-as-a-service. Database-as-a-service is indeed essential for database-driven applications. This paper lists important parameters of database. As a result, Study and analysis of database as Mysql service varies significantly depending on the different cloud infrastructure such as Amazon EC2, Joyent cloud and Rackspace.

Index Terms— Amazon EC2, Joyent Cloud, Mysql, Rackspace Cloud, Xeround, MoNyog, DBaaS.

◆

1 INTRODUCTION

Cloud computing is technology that uses internet and control remote server to maintain data and application. Instead of installing set of software for each computer, you need to load your application and use it without installation every time and access personal files at any computer with the help of internet. The main prominent characteristics of cloud computing are to provide on-demand self-services, resource-pooling and rapid elasticity. Cloud computing is classified into three services: Infrastructure as a Service, Platform as a Service and Software as a Service even database as a Service (DBaaS) is included into Software as a Service. This technology gives us efficient computing by centralized storage, memory, processing and bandwidth.

2 AMAZON ELASTIC CLOUD COMPUTE EC2

Amazon cloud EC2 provides cloud computing solution on the basis of as per usage. Amazon EC2 gives a web service API for manipulating, deprovisioning and provisioning virtual servers inside the Amazon cloud. Amazon EC2 U.S. footprint has several data centres. Out of which, three data centers lie on the East Coast of the U.S. and two lie in Western Europe [3]. We have to sign up separately for separate data centers account. In term of Infrastructure, Amazon itself handles all the hardware and controls the network infrastructure. The servers are operated by Open Source Xen hyper visor that provides facilities of dynamic provisioning, deprovisioning, and isolated computing environment for users [2]. This paper is based on performance evaluation of Database as a service on Amazon infrastructure.

3 JOYENT CLOUD

Joyent also provide cloud computing solution. Joyent provides public, private, or virtual private infrastructures across multiple data centers. Joyent infrastructure as a service incorporates

load balancing and disk caching for improved input-output, enhanced security and reporting capabilities. Joyent SmartOS plays main role in an infrastructure service [4]. It provides both KVM hardware virtualization and Operating system level virtualization on single operating system. Moreover, Joyent Smart OS comprises DTree technology that gives visibility and insight. From security perspective, Joyent SmartOS delivers security and governmental standards including EAL 4+ compliance that isolate network processes storage and memory on virtual server [6].

4 RACKSPACE CLOUD

As we know that rackspace provides cloud computing solution. It provides cloud infrastructure that allows user to deploy and create advanced and high available architecture. Rackspace cloud server runs the latest version of OpenStack Compute - Nova in Essex which is fifth version of OpenStack focuses on quality, usability, extensibility and high performance computing. Rackspace cloud provides security to its user into three ways: Physical security, operational security and system security. Physical security provides security to the locations where the cloud servers and data centers are stored and making people accessible who have the authentication to go there. Operational security provides secure access to important and critical information in the cloud. Last, system security provides security to the client systems and it is maintained by patches and patched operating system [1].

5 XEROUND

Xeround provides Database-as-a-Service (DBaaS) for Mysql database. Xeround delivers configuration and optimization to performance of database and availability on the cloud. Xeround works on tier-II architecture. This architecture is classified into two nodes: Access nodes and Data nodes. Access nodes are used for receiving application requests, communi-

cate with data nodes, perform computations and deliver results while data nodes are used for storing data. The data storage is handled by virtual partitions. Each partition is available to the different data nodes located on separate servers and provide high availability. Xeround first keep our databases in two synchronous in-memory replicas, and then keeps our databases into persistent store such as Amazon EC2, Joyent cloud and Rackspace cloud asynchronously [5], [8].

6 METHODOLOGY

The goal of this performance evaluation is to study about throughput under different cloud infrastructures as Amazon EC2, Joyent Cloud, and Rackspace Cloud. To this end, we ran Mysql databases on different cloud infrastructures such as Amazon and Joyent cloud with the help of Xeround and measures throughputs and analysis which one infrastructure is best suitable for database-driven applications. We have used Webyog for the purpose of monitoring Mysql databases. Moreover, this paper reveals Mysql metrics on the basis of different cloud infrastructure such as Amazon EC2, Joyent cloud, and Rackspace Cloud. We have taken following parameters:

6.1 Threads

Thread is created for each connection to the Mysql server and threads creation take time and resource. Threads cache holds threads that are not being used by any connection. Thread cache used threads that are available and not being used by any connection. We can find out number of threads in the cache by the `thread_cache_size` variable. Each thread normally uses 128KB of memory. The observing point is that always check Threads created and make sure that number of threads created per second should not be more than one. Last but not least, Mysql response should be much faster if it is using threads from cache instead of creating them [7], [8].

6.2 Table Cache

When Mysql access table, it stores table in the cache. It is called table cache. We should always check opened table status variable. If you observe that number is going to be large, you will have to increase the value of table cache by table cache variable. Hence, Mysql queries time will be faster with the help of table cache instead of opening the table file for each query [7], [8].

6.3 Query Cache

The query cache stores the text of select statements together with the corresponding result that was sent to the client. It is very prominent for read-intensive applications. It enables very fast retrievals on cache query hit by storing query cache completely in memory. We can control the size of query cache by `query_cache_size` variables. The value should be in the range of 64MB to 1024MB [7], [8].

6.4 Table Locking

Table locking enables many sessions to read from a table at the same time. If you want to write to a table, you will have to get exclusive access and wait for the other session to finish it. All

other tables have to wait until update is completed [7], [8].

6.5 MyISAM Key Cache

MyISAM has capability of table locking and repair table functionality. It is usually used to minimize disk Input-Output and improve key cache performance. It uses index block where most used index block is stored and data block uses native operating system cache to store most used data block [7].

7 RESULTS

We have observed the following throughput based on Mysql database provided by Xeround on different cloud environments such as Amazon EC2, Joyent cloud and Rackspace cloud with the help of MONyog. It includes

7.1 General Parameters

When Mysql access table, it stores table in the cache. It is called table cache. We should always check opened table status variable. If you observe that number is going to be large, you will have to increase the value of table cache by table cache variable. Hence, Mysql queries time will be faster with the help of table cache instead of opening the table file for each query.

7.2 Data manipulation Language

The total number of data manipulation statements that client have sent to the Mysql server. The data manipulation statements comprise SELECT, INSERT, UPDATE and DELETE commands.

7.3 Thread Cache

Thread cache parameters are classified into three different categories. First, thread cache size which reveals size of threads cached. It should be low as much as possible because thread creation is time consuming activity. Second, thread cache hit ratio which indicates hit rate of thread cached. If the value is low, it is increasing thread cache. Last, thread created which reveals how many threads created to handle connection. The increasing value of thread creation indicates insufficient size for thread cache.

7.4 Query Cache

Query cache points out queries related information. We have taken three parameter related to it. First, query in cache which indicates number of queries currently stored in the query cache. Second, query cache hit which reveals number of queries that served successfully by query cache. Last, query cache hit ratio should be high because query cache is used to improve speed of application.

7.5 Table cache

We have two different sub-parameters of table cache. First, table open reveals number of tables that are currently opened. Second, Number of table cache which points out number of table request that are not handled by table cache.

7.6 Table locking

We used table locks and table lock waited sub-parameters. Table locks indicates number of times table getting

table lock.Latter, it indicates number of times wait was needed before getting table lock.

7.7 Maximum open file

The maximum open files show numbers of files that Mysql open at a time.It would be best to raise value upto 64K.

7.8 Total rows returned

The total rows returned parameter indicates that number of rows returned by all queries.

7.9 Total rows returned via index

The total rows returned via index shows that how many rows

TABLE 1
MYSQL PERFORMANCE PARAMETERS

Parameters	Joyent Cloud	Amazon EC2	Rackspace Cloud
Connection used	60%	20%	60%
Byte Received	6.96K (23.599/sec)	7.83K (7.83/sec)	7.13K (24.182/sec)
Byte Sent	33.81K (114.656/sec)	36.66K (123.88/sec)	34.44K (116.78/sec)
Total number of DML	74 (0.245/sec) 60.163%	22 (0.073/sec) 23.158%	76 (0.252/sec) 60.80%
Thread cache size	0	0	0
Thread Cache hit Rate	0.0%	0.0%	0.0%
Number of Thread created	16.00 (0.053/sec)	20.00 (0.066/sec)	16.00 (0.053/sec)
Table currently opened	29	28	4
Number of table cached	1.76K	1.76K	504

returned by all queries with the help of indexes.

7.10 Average rows per query

The average row per query parameter indicates average number of rows returned for each query.

TABLE 2
MYSQL PERFORMANCE PARAMETERS

Parameters	Joyent Cloud	Amazon EC2	Rackspace Cloud
Number of times table locked acquired immediately	73.00 (0.242/sec)	17.00 (0.056/sec)	75.00 (0.248/sec)
Number of times table locked waited	0.0%	0.0%	0.0%
Total row returned	564	932	506
Total row returned via index	60.00	4.00	0.00
Average row per query	4	9	4
Maximum open file	4.00K	4.00K	1.00K

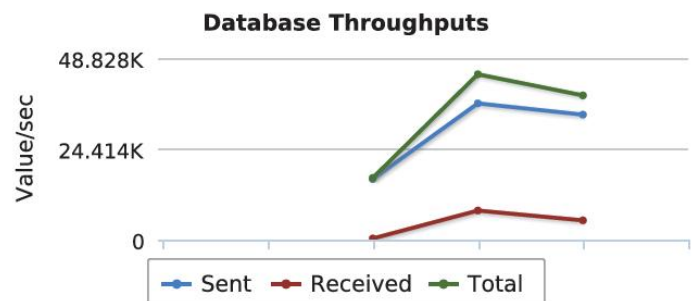


Fig. 1 This graph depicts database throughputs during Mysql performance in Amazon EC2.

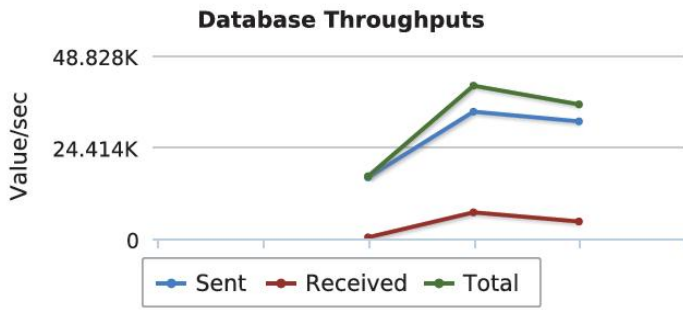


Fig. 2 This graph depicts database throughputs during Mysql performance in Joyent Cloud.

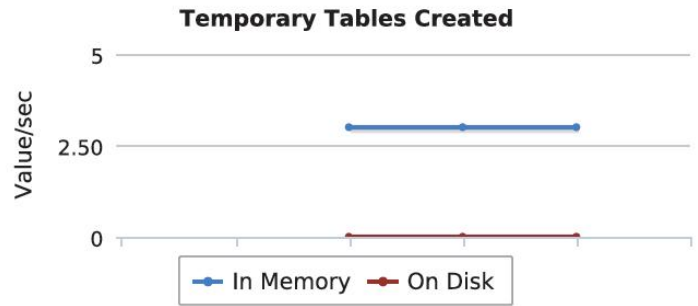


Fig. 6. This graph shows number of temporary table created during Mysql performance in Rackspace Cloud.

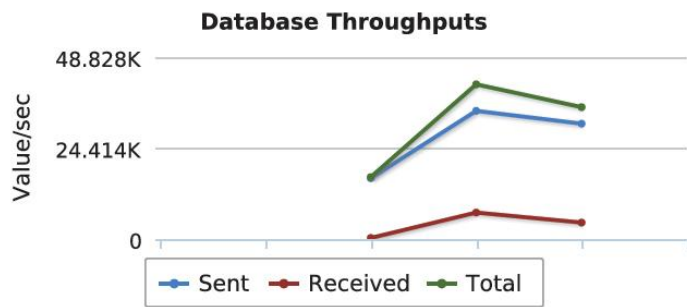


Fig. 3 This graph depicts database throughputs during Mysql performance in Rackspace Cloud.

8 CONCLUSION

This paper represents performances analysis of database such as Mysql on different cloud environments provided by Xeround like Amazon EC2, Joyent Cloud and Rackspace cloud. The main aim of this paper is to assess Mysql database performance because most of the applications are database-driven. For this reason, database-driven applications are directly affected by performance of its databases. Thus, we observed comparison among Amazon EC2, Joyent cloud and Rackspace cloud. I would like to evaluate NoSQL databases comparison like Cassandra, hadoop, Accumulo and Cloudera on different cloud environments in future.

Acknowledgment

This research paper is made possible with the help and support of staff at Computer Department, University college of Engineering, Kota, India. I would like to thank Mr. R.S. Sharma for his most support.

REFERENCES

- [1] Rackspace Cloud, *White paper on Rackspace security-triple strength security*, available at <http://broadcast.rackspace.com/security>
- [2] Guohui Wang, "The impact of virtualization on network performance of Amazon EC2 data center" *Proc. conf. on information communication (INFOCOM '10)* pp.1163-1171, 2010.
- [3] Orna Agmon Ben-Yehuda, Muli Ben-Yehuda, Assaf Schuster and Dan Tsafir, "Deconstructing Amazon EC2 Spot Instance Pricing," *IEEE Trans. Cloud computing technology and science (CLOUDCOM '11)*, pp. 304-311, 2011, doi:10.1109/CloudCom.2011.48.
- [4] Joyent Cloud, *White paper on The Joyent Smart Technologies Architecture for Cloud Computing*, available at: <http://www.joyent.com/>
- [5] Xeround, Technical white paper on cloud database, February 2012, available at: <http://xeround.com/developers/mysql-cloud-db-technical-whitepaper/>
- [6] Joyent Cloud, *Whitepaper on the Joyent Smart data centers*, available at: <http://www.joyent.com/resources/category/white-papers>.
- [7] MONyog, *White paper on MONyog*, available at: <http://webyog.com>
- [8] Gaurav Sharma, "Performance analysis of database on different cloud environments", *Technical-Journal open access*, submitted for publication.

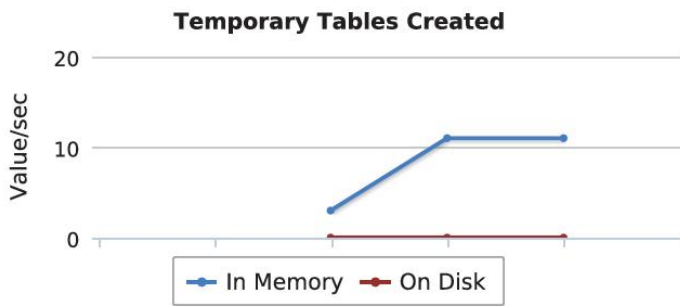


Fig. 4. This graph shows number of temporary table created during Mysql performance in Amazon EC2.

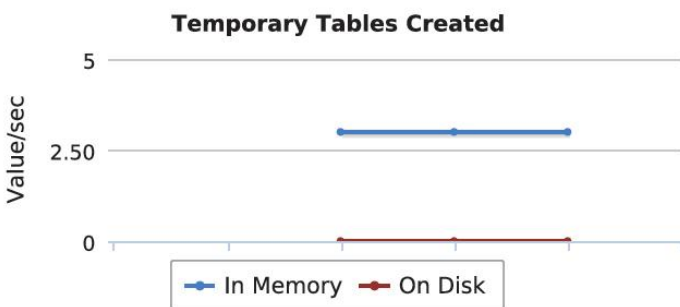


Fig. 5. This graph shows number of temporary table created during Mysql performance in Joyent Cloud.

-
- *Gaurav Sharma, University college of Engineering, Rajasthan Technical University, Kota, India, E-mail: gaurav12mdr@gmail.com*
 - *R.S. Sharma, Senior Member, University college of Engineering, Rajasthan Technical University, Kota, India.*