

A Study on Cloud Computing Quality of Service for Big Data: Contribution to Computation

Gibet Tani Hicham

Laboratory of Informatics Systems and
Telecommunications (LIST)
Department of Computer Engineering, Faculty of
Sciences and Technologies, Abdelmalek Essaadi
University
Route Ziaten, B.P. 416 Tangier, Morocco

El Amrani Chaker

Laboratory of Informatics Systems and
Telecommunications (LIST) Department of
Computer Engineering, Faculty of Sciences and
Technologies, Abdelmalek Essaadi University
Route Ziaten, B.P. 416 Tangier, Morocco

ABSTRACT

Big data is one of the major pioneering technologies of our time. Big data is not only used to refer to the recent massive data growth in various sectors but also to describe a new computing model capable of managing efficiently and accurately the new generation of data sets whatever their volumes, types or characteristics. Running big data inside an organization can be a challenging mission as a consequence to the required infrastructure implementation and management higher cost. From this perspective, cloud computing radiates as a strong candidate to host and manage big data workloads efficiently and at a reduced price. On the one hand, cloud computing is a robust Information Technology solution acclaimed with its extensive computing abilities. On the other hand, moving big data to the cloud can be very complex due to Quality of Service criteria defined between the provider and the end-user. On this context, cloud computing platforms need to be re-evaluated and re-examined to offer a resilient foundation for big data storage and processing. Furthermore, cloud computing QoS stipulations have to be defined, weighed and mapped rationally to big data characteristics in order to highlight the most significant parts that require specific consideration or optimization in pursuance of a cloud computing service dedicated for big data or big data as a service.

Keywords

Big Data, Cloud Computing, Quality of Service, Service Level Agreement, Resources Scheduling.

1. INTRODUCTION

Cloud computing is a service oriented computing paradigm that consists on supplying various Information Technology (IT) services over the internet in a new flexible payment mode (pay per use). This innovative computing model has changed the classical computing perception and usage mode. The primary cloud computing focus is to reduce costs related to IT implementation and management with respect to the Service Level Agreement (SLA) that defines the Quality of Service (QoS) criteria to be respected. With recent IT prevalence, data is no longer constrained to structured databases [1] but also encompasses unstructured data like emails, pictures, videos, audios, binary files, text files, code, scripts... Taking into account the advancements in IT world, almost every business is struggling to manage larger data extents with a diversity of forms and types which requires a transition to new computing solutions such as big data. Despite [2] the fact that cloud computing is an efficient computing model for distributed systems and immense computation, big data hosting on the cloud can be problematic due to QoS characteristics. Cloud computing service delivery model can be an obstacle even for

minor services, thus big data 5Vs porting to the cloud might become an arduous operation.

Big data and cloud computing are the IT tendency of these recent times which are expanding every month in terms of customers and services. Many researchers are investigating cloud computing as a platform for big data storage and processing where the most significant ones that can be mentioned are the contribution of Venkatesh et al. [3] in 2015, where the authors discussed the applications, advantages and challenges faced by big data when used on cloud computing environments. Moreover, the authors proposed some security measures to consider for safeguarding big data on the cloud. The authors discussed big data technology extensively; however, they didn't manage to cover how the cloud would be a solution for big data problematic.

Likewise, R. Kune et al. [4] studied big data computing model and detailed its technologies and characteristics. Later, the authors presented how a cloud computing platform would be utilized to host big data analytics. Additionally, the authors discussed an emerging big data computing platform over clouds. This study was very rich in terms of defining the cloud computing model for big data, nevertheless, cloud computing QoS criteria weren't discussed nor considered during the study.

M. S. Al-Hakeem [5] proposed a new service model for big data on the cloud. The author defined a set of technologies required to integrate big data with cloud computing in order to provide a new cloud service founded on the same basic cloud offerings (Infrastructure as a Service, Platform as a Service and Software as a Service). However, the author did not manage to discuss the various QoS aspects that would affect this new service model.

P. C. Neves et al. [6] described both cloud computing and big data systems where they focused on the issues yet to be addressed. In addition to that, the authors enumerated the issues related to big data computing model and determined the cloud computing solutions. Furthermore, the paper presents a cloud computing model that lists the cloud computing solutions advantages and disadvantages in case of a cloud service for big data.

C.X. Mavromoustakis et al. [7] paper exhibits the challenges related to big data and the opportunities that come with cloud computing model as a solution. Additionally, the authors carried out an expense analysis towards estimating the long term profits of implementing big data as a service on the cloud model. This study lacks the definition of the cloud computing QoS criteria and the SLA stipulations given that

the authors considered reviewing the cost benefits of big data on the cloud.

N. Zanoon et al. [8] managed to elaborate a study on both cloud computing, big data paradigms and concluded that the relationship between them is complementary. The authors developed a compatibility chart between big data and cloud computing in term of features in the interest of revealing the comprehensive relationship between big data and cloud computing.

In this paper, an investigation of big data and cloud computing is conducted in the course of determining the QoS criteria that should be tackled and optimized towards big data hosting on the cloud. Moreover, a set of conceptions has been determined to assess these criteria founded on literature and capability to solve the big data problematic on the cloud. Finally, a discussion of some of the difficulties and issues related to big data processing on the cloud is presented. The rest of the paper is organized as follows: Section 2 & 3 illustrate big data concepts and cloud computing QoS characteristics. Section 4 describe the core considerations to assess QoS criteria for big data on the cloud and results. Section 5 & 6 presents the discussions and evaluation results of the previous section. Section 7 concludes this work.

2. BIG DATA DESCRIPTION

Big data is an expression that is widely used to describe the recent massive data growth, which has become difficult to process using traditional data administration applications or database management systems (DBMS) [9]. Big data is also linked to the analytics activities related to creating value out of a diversity of datasets.

2.1 Big Data characteristics

Generally, big data [10] can be distinguished from classical data with the following basic three parameters:

- **Volume:** Nowadays data has reached enormous sizes. We are no longer talking about giga-bytes (10⁹) or terra-bytes (10¹²) of data but zetta-bytes (10²¹) and yotta-bytes (10²⁴). This immense growth is making data management with classical tools deprecated.
- **Variety:** is the second big data complexity constraint, where data can be in a structured form like text files or unstructured such as XML code, social media feeds...
- **Velocity:** This parameter refers to data generation and arrival speed. For example, data coming from social media or banking transactions.

As stated before, big data is not only about data growth and management but also about helping businesses carrying out significant data analytics in the aim of validating the data and producing value that can be resumed by the following parameters:

- **Veracity:** data condition, quality verification and validation.
- **Value:** Creating valuable data for specific business (analytics) by managing efficiently large and diverse datasets capacities.

Therefore, these two parameters can be added to the previous triad to form the big data 5 Vs.

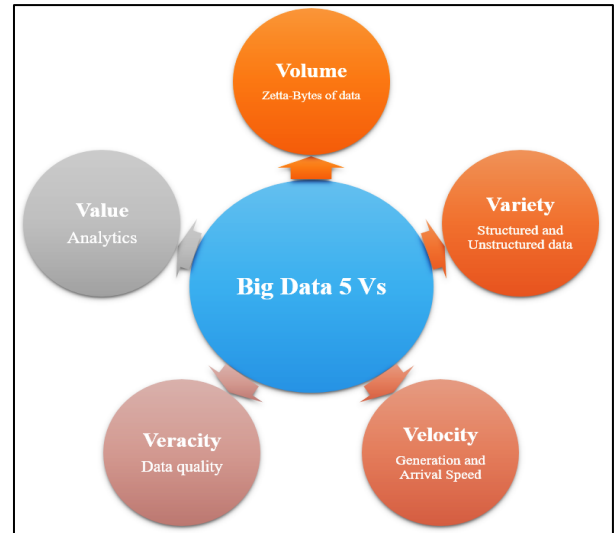


Fig 1: Big data 5 Vs

2.2 Big Data major technologies

In 2003, Google has announced the concept of Google File System known as GFS [11]. Later on, Google also introduced the architecture of MapReduce which was gathered by Doug Cutting in 2005 to create the big data first framework foundation identified as Hadoop.

Hadoop is an Apache project that manages large-scale distributed processing infrastructure for parallel processing of big data on large clustered servers [12]. This Big Data platform incorporates three basic components (Figure 2) [13]:

- **HDFS:** Hadoop Distributed File System (HDFS) was designed to store large files no matter their size or form. HDFS is java-based file system that disperses data as blocks on several server clusters without minimum requirements, where these blocks can be replicated to facilitate their retrieval and parallel processing.
- **MapReduce:** is java based programming model created by Google for processing and generating large sets of data on computer clusters. It is a core component of the Apache Hadoop Software Framework. The Map function allows different points in the distributed cluster to distribute their work. The Reduce function reduces the final form of cluster results into a single result.
- **YARN:** Yet Another Resource Negotiator (YARN) is resource management and task scheduling technology that allocates system resources to different applications running in a Hadoop cluster. It is also used to schedule the execution of tasks on different servers of clusters.

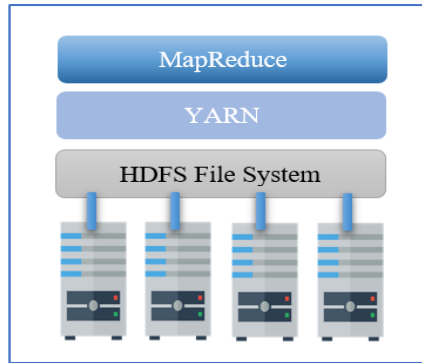


Fig 2: Hadoop general architecture

With big data emergence and development, new database management systems have been revealed which were designated as NoSQL databases or Not Only SQL. NoSQL [14] databases doesn't eliminate SQL completely but compliments it in such a way that the NoSQL database can scale up when it grows and can support a variety of data types in a key-value architecture. NoSQL containers are similar to hash tables where the keys are employed as pointers in a way that make it faster than a classical RDBMS. NoSQL databases can be categorized as:

- Key-Values Databases: stores database as key-value like a hash table.
- Column Databases: Store data in a similar key-value mode, nevertheless the key is a mixture of column, row and timestamp.
- Document Databases: gather documents that incorporate one or more self-contained named fields.
- Graph Databases: stock data in a resilient graph model that scales across multiple machines.

Accordingly, the following are far most used NoSQL databases [14, 15]:

- Cassandra: is a java based database developed by Apache Software Foundation. It involves concepts of both key-value stores and column stores.
- MongoDB: is a C++ based database developed by 10gen. MongoDB is a document database that helps manipulating structured objects without a predetermined schema.
- HBase: is a java based database developed by Apache Software Foundation. It was inspired from Google big table database which is a column oriented database.

3. CLOUD COMPUTING AND QOS

3.1 Cloud Computing Architecture

Cloud computing is an innovative computing model that consists of providing various IT services over the internet in a pay-per-use method. Cloud computing model is generally divided into two sections: frontend and backend, where these two segments are linked with a network which can be the internet in case of public cloud, local area network in case of private cloud or both in a hybrid cloud [16]. The frontend is what the end user sees, commonly a light desktop application to access the cloud service or simply a web browser. The backend is the foundation infrastructure that is composed of

several clustered servers, storages, provisioning and management platforms. The cloud computing offering can be categorized as follow (figure 3):

- Infrastructure as a Service (IaaS): Computing, Storage, Communication.
- Platform as a Service (PaaS): Operating Systems, Database Management Systems, Application Server platforms, Web Server Platforms...
- Software as a Service (SaaS): Email, Text Editing, Enterprise Resource Planning (ERP), Customer Relationship Management (CRM)...

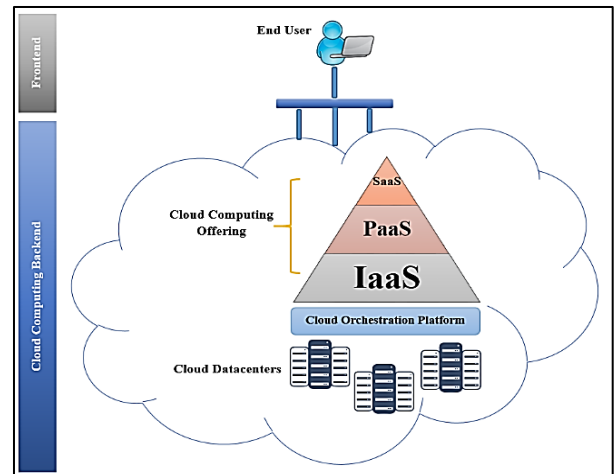


Fig 3: Cloud Computing Architecture

3.2 Cloud Computing QoS Criteria

Cloud computing implements large scale infrastructures to ensure a good Quality of Service (QoS) that is supposed to generate the lowest response time, lowest latency rate and attenuates the packets deprivation. Correspondingly, QoS is defined as a set of criteria capable of measuring the cloud service offering significance.

These criteria can be grouped into the following classes:

Economy class:

The key interest in cloud computing adoption is the economy. Economic criteria are the primary concern for everyone looking forward to utilize this computing model, where the most significant are:

- Cost: The cloud service cost is a decisive metric that helps an organization decide whether to invest in the implementation of a specific IT solution internally or acquire it as a service on the cloud. Furthermore, this metric can be used to evaluate the same service offering of multiple cloud providers [18].
- Payment mode: Actually, cloud providers are offering a diversity of payment modes for each service. The user has the privilege to pay only for what he is using, pay when he finishes an operation or subscribes to a service.
- Return on Investment (ROI): Deploying a solution to the cloud can accelerate an organization time to market of a specific service and consequently help them touch the ROI quickly.
- Green Computing: This property can be vital especially for countries imposing specific

compliance standards regarding datacenters electricity consumption and cooling systems emissions.

Performance class:

For several years, Cloud Computing QoS properties related to performance have been neglected in comparison with the economic characteristics. However, with cloud computing recent evolution users are no longer moving to the cloud to decrease IT expenses, however users are looking forward to exploit the high performance computing offerings:

- **Computation:** Represents the data or task (job) computing and processing in cloud systems [18]. This property can have a massive impact on the cloud service response time and it is generally related to the Core Processing Units (CPU) scheduler residing on the cloud provider infrastructure.
- **Storage:** Designates the storage capacity limits that can be offered to a specific user. This property becomes conclusive for large organizations dealing with massive data growth or big data.
- **Communication:** Denotes the data transfer rate between an organization internal instance and the cloud provider infrastructures [17, 18]. This QoS property is primordial for cloud computing.

Security Class:

Security is crucial when discussing porting an organization data to a third-party infrastructure. Cloud computing security can be assessed through the following criteria:

- **CIA (Confidentiality, Integrity, Availability):** The CIA triad is the first security property to be considered on the cloud. The cloud user data has to be kept confidential inside the cloud provider infrastructure by means of encryption. Moreover, all required communications have to be through secured encrypted protocols to protect the data. Additionally, user data and services have to stay consistent and available at all time.
- **AAA (Authentication, Authorization, Accountability):** This property ensures controlling access to the user cloud resources (data included), enforcing policies, auditing usage, and providing the information necessary for control.

Non-functional class:

Non-functional criteria entitle the properties or characteristics that are used to judge the cloud service, rather than a specific technological requirement [18], which can be resumed on the following:

- **Reliability:** One of the main non-functional cloud properties, which helps define the cloud service ability to keep operating at the required efficiency level.
- **Scalability:** Characterizes the cloud service ability to scale up or down whenever needed by the end-user.
- **Adaptability:** This cloud characteristic represents the cloud service capability to adapt to the user various activities and operational sectors.

4. CLOUD COMPUTING QOS FOR BIG DATA

Cloud computing QoS criteria is an essential aspect that affects massively the continuity and trustworthy of this innovative computing model. According to the latest cloud computing forecasts [19, 20], cloud computing is expanding every year in term of infrastructure investments, services and revenue. This growth is putting the heat more and more on the ability of externalizing big data to the cloud since it is one of the most recognized trends and IT problematics of our time. Without doubt, cloud computing is a durable solution and a fertile ground for big data storage and processing. Nevertheless, cloud computing QoS criteria have to be re-examined to high-light the significant properties that needs more attention or optimization in order to propose an exhaustive cloud computing service dedicated for big data.

As discussed previously, big data characteristics can be resumed on the 5 Vs: Volume, Variety, Velocity, Veracity and Value. Accordingly, these characteristics can be mapped to cloud computing QoS classes and criteria where this mapping can be obtained from the cloud computing QoS properties impact on each one of the big data characteristics (Table 1.).

Table 1. Mapping Big Data Characteristics to the corresponding Cloud Computing QoS Class

BIG DATA CHARACTERISTICS & CLOUD COMPUTING CRITERIA	ECONOMY				PERFORMANCE			SECURITY		NON-FUNCTIONAL		
	Cost	Payment Mode	ROI	Green Computing	Computation	Storage	Communication	CIA	AAA	Reliability	Scalability	Adaptability
	X			X	X	X	X	X	X	X	X	X
	X				X	X		X	X	X	X	X
	X			X	X	X	X			X	X	X
			X		X		X					
	X		X		X		X	X	X			

This mapping represents the implication of each cloud computing QoS class with the big data characteristics. On the same context, a set of initial scores have been associated to the QoS criteria of each class in order to underline its importance (Table 2) in cloud computing model [21, 22]. Each cloud computing QoS class initial score can be summed as:

$$\sum_{n=1}^n sq_n = 1$$

Table 2. QoS criteria initial scores

CLOUD COMPUTING QOS CLASSES	QOS CLASSES CRITERIA	INITIAL SCORE
Economy	Cost	0.5
	Payment Mode	0.1
	ROI	0.3
	Green Computing	0.1
Performance	Computation	0.4
	Storage	0.3
	Communication	0.3
Security	CIA	0.6
	AAA	0.4
Non-functional	Reliability	0.3
	Scalability	0.3
	Adaptability	0.4

Each of the initial scores mentioned on table 2 has been chosen to represent its significance for a big data situation:

- Cost is a characteristic that would affect extremely a decision to move big data to the cloud that is why it weighs more than the other characteristics of the economy class.
- Computation weigh more on the performance class because it is involved on all the big data 5 Vs and also affects storage and communication which relatively attached to the hardware technology.
- CIA features are more demanded on a cloud computing service than the AAA.
- Adaptability is mostly required on cloud computing in order to deliver a big data service, where the cloud implementation needs to consider the client requirements and actual data status, technologies and working procedures and policies

Respectively, a set of theoretic weights have been linked to the big data characteristics according to its standing to this computing model [23, 24]:

Table 3. Big data characteristics weights

BIG DATA CHARACTERISTICS	WEIGHT	WEIGHT DESCRIPTION
Volume	0.25	Management of large data volume is the primary goal of big data.
Variety	0.15	Type of data does not affect big data management at the same level of the other characteristics.
Velocity	0.2	Data generation speed comes slightly after volume and value characteristics.
Veracity	0.15	This is the same as for variety characteristic.
Value	0.25	Producing value out of large amount of data is primary to big data computing model.

The overall weight can be calculated as:

$$\sum_{n=1}^n wb_n = 1$$

Taking into account the big data characteristics mapping to cloud computing criteria (Table 1), the initial scores associated to each cloud computing property (Table 2) and big data characteristics weights (Table 3), the following results can distinguished:

Table 4. Scoring cloud computing criteria and big data characteristics

BIG DATA CHARACTERISTICS & CLOUD COMPUTING CRITERIA	ECONOMY				PERFORMANCE			SECURITY		NON-FUNCTIONAL		
	Cost	Payment Mode	ROI	Green Computing	Computation	Storage	Communication	CIA	AAA	Reliability	Scalability	Adaptability
Volume	0,125	0	0	0,025	0,1	0,075	0,075	0,15	0,1	0,075	0,075	0,1
Variety	0,075	0	0	0	0,06	0,045	0	0,09	0,06	0,045	0,045	0,06
Velocity	0,1	0	0	0,02	0,08	0,06	0,06	0	0	0,06	0,06	0,08
Veracity	0	0	0,045	0	0,06	0	0,045	0	0	0	0	0
Value	0,125	0	0,075	0	0,1	0	0,075	0,15	0,1	0	0	0
Score	0,425	0	0,12	0,045	0,4	0,18	0,255	0,39	0,26	0,18	0,18	0,24
Overall Score	0,59				0,835			0,65		0,6		

The cloud computing QoS score of each property was obtained using the weighted score method in addition to the mapping result which eliminates the QoS property if not important for big data:

$$S = \sum_{n=1}^5 a_n \times sq \times wbn$$

Where:

- S: Score.
- a_n : mapping result (1 if mapping exists, 0 if not).
- sq: Cloud Computing QoS criteria initial score.
- wbn: Big Data Characteristics weights.

The overall score of each cloud computing QoS class was obtained as follow:

$$FS = \sum_{n=1}^n S_n$$

5. BIG DATA ON THE CLOUD EVALUATION

Figure 4 represents the overall impact of cloud computing QoS classes on big data. This result was obtained from the overall score that categorizes the cloud computing QoS aspects which eventually forms the big data essentials. In this figure, the cloud computing QoS performance class can be seen as the major cloud computing feature that has to be tackled when discussing big data porting to the cloud. Additionally, both Economy and Non-functional QoS classes acquires an important share that can affect an organization decision to move big data to the cloud. Finally, security class is one of cloud computing bricks that cannot be ignored especially in case of big data complexity.

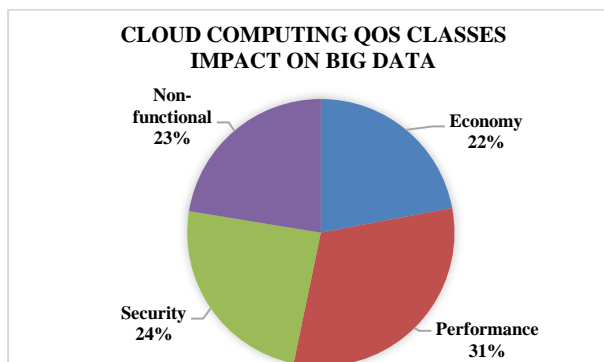


Fig. 4. Cloud computing QoS classes impact on big data

Figure 5 details the cloud computing QoS properties that pose a direct influence on big data. All of these features, characteristics and possibly other criteria need to be contemplated side to side when inspecting cloud computing as a platform for big data.

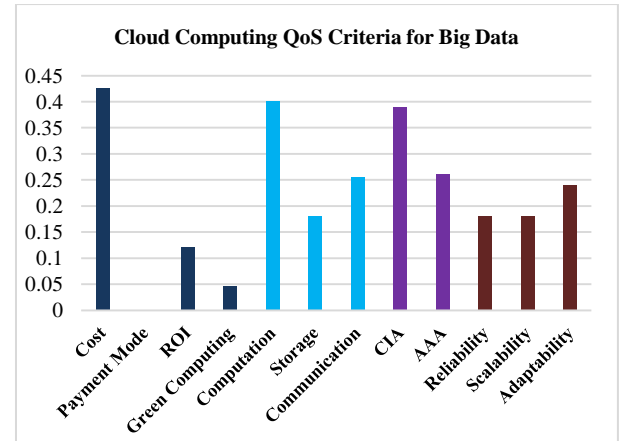


Fig. 5. Cloud Computing QoS Criteria for Big Data

Nevertheless, some of the cloud computing QoS criteria presents more prominence for big data than the others, where the cost and computation properties are both the major concerns for both cloud computing and big data. Computation refers to the processing power (CPUs) that will be allocated for big data implementation and running. This property is assured by the system processors and the compute resources scheduling algorithms present on the operating system that will be used to host big data platforms.

Computation is a cloud computing QoS aspect that has to be engaged in an intelligent matter because of its involvement in each and every one of the other cloud computing QoS criteria that necessary requires a high performance computation to be delivered especially in case of big data. Moreover, storage and communication properties need to be analyzed further to pinpoint the key parts that requires optimization for big data.

Economy and Non-functional criteria comes second on the big data impact scheme (Figure 4), where the cost of big data realization or moving to the cloud constitutes the essential characteristic. The service cost is an essential QoS property that will affect primarily the decision makers to proceed with big data establishment on the cloud or moving the current big data infrastructures to the cloud. This crucial aspect along with ROI will have an enormous effect on all other cloud computing QoS classes and ultimately have to be optimized to bring a competitive cloud offering dedicated for big data. The Non-Functional properties gets their importance from the performance class which defines clearly the response level of each property in this class. Correspondingly, Non-Functional properties definition will eventually fluctuate the cloud computing economy criteria. Finally, the security class properties should not be overlooked because it is a subject that can threaten most decisions to move to the cloud. Consequently, this cloud computing QoS property have to be well defined in the aim of encouraging big data storage and processing on the cloud.

6. COMPUTATION OPTIMIZATION IMPACT: BIG DATA ON THE CLOUD

In this section a simulation of big data on a cloud computing environment using CloudSim simulation toolkit is proposed. The simulation purpose aims to highlight the importance of computation optimization for a big data service on the cloud.

The cloud computing simulation configuration consists on the following [25]:

- 1 datacenter with 100 hosts,
- The host physical configuration is:
 - 16 processors with the capacity of 96900 MIPS (The equivalent of an Intel E7-x870 processor according to Cisco Industry Benchmarks Performance [25]),
 - 65 536 MB of RAM (Random Access Memory),
 - 1 TB of storage.
- 200 testing Virtual Machines were created on the different hosts:
 - 8 processors with the capacity of 96900 MIPS,
 - 16 384 MB of RAM (Random Access Memory),
 - 10 GB of storage.

After the initialization of the cloud computing environment as described above, a set of large processes (big data like processes with massive amount of instructions) has been submitted for execution. In order to bring the computation importance into attention, a comparison of optimized computation algorithm (scheduling algorithm) that has been proposed previously [26] and a simple version of this algorithm which is the Round Robin (RR) computation algorithm and also to the First Come First Served (FCFS) algorithm. The comparison is carried in respect to the average turnaround time which represents the average time amount necessary to carry out the execution of processes.

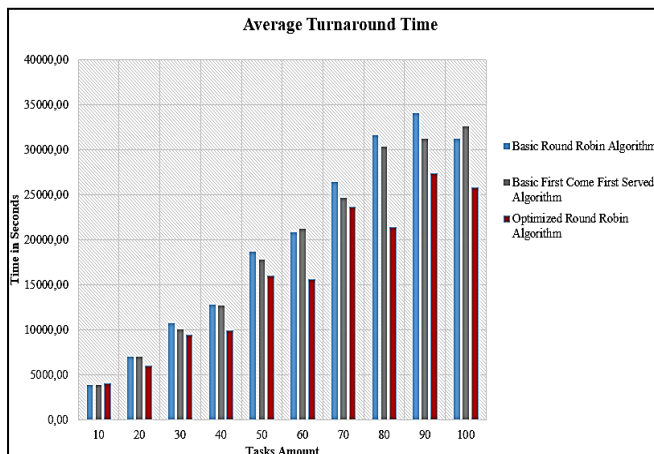


Fig. 6. Comparison of scheduling algorithms

From figure 6 it appears clearly how an optimized computation algorithm would affect cloud computing and big data. On the one hand, an optimized computation algorithm would reduce the cost of new investments in term of infrastructures, energy consumption and eventually provide a service with high computation capabilities. On the other hand, big data could ultimately benefit from optimized computation

algorithm that would help accelerate the processing and response time of its heavy tasks.

7. CONCLUSION

This paper is an investigation of cloud computing and big data technologies with the aim of highlighting the cloud computing QoS criteria that needs to be optimized in order to create a cloud service or offering dedicated for Big Data storage and processing. The study lead to the creation of a mapping scheme between the big data characteristics and the cloud computing QoS criteria extracted from literature and the possibility to solve the cloud computing for big data problematic. Discussing the mapping scheme results in weighing the cloud computing QoS criteria and underlining the substantial ones that requires optimization or special care in order to host a complex paradigm such as big data. From the investigation findings, the performance class and specifically “computation” property is the main QoS characteristic that calls for optimization in the direction of creating a smart cloud computing service for big data. Similarly, economic criteria represent the cloud computing QoS pillar for big data, hence the need for a rational economic offer to promote the cloud computing adoption as a solution for big data storage and processing.

This study will be adopted as a core podium to propose a contribution to the optimization of cloud computing QoS for big data storage and processing taking into consideration the most important features of both computing models and the results obtained on this paper.

8. REFERENCES

- [1] S. A. El-Seoud, H. F. El-Sofany, M. Abdelfattah, R. Mohamed, “Big data and cloud computing: trends and challenges, International Journal of Interactive Mobile Technologies, Vol 11, No 2, 2017.
- [2] J. Changqing, Y. Li, W. Qiu, U. Awada, K. Li. “Big data processing in cloud computing environments”. In proceedings of International Symposium on Pervasive Systems, Algorithms and Networks, I-SPAN 2012.
- [3] H. Venkatesh, D. P. Shrivatsa, J. Nivedita, “A study on use of big data in cloud computing environment”. International Journal of Computer Science and Information Technologies, Vol. 6, 2015.
- [4] R. Kune, P. K. Konugurthi, A. Agarwal, R. R. Chillarige, R. Buyya. “The anatomy of big data computing”. Software Practice and Experience, Wiley Online Library. 9 October 2015.
- [5] M. S. Al-Hakeem. “A proposed big data as a service (BDaaS) model”. International Journal of Computer Sciences and Engineering. Volume-4, Issue-11, 2016.
- [6] P. C. Neves, B. Schmerl, J. Bernardino, J. Camara. “Big data in cloud computing: features and issues”. In Proceedings of the 2016 International Conference on Internet of Things and Big Data. 23-25 April 2016.
- [7] G. Skourletopoulos, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, C. Dobre, S. Panagiotakis, E. Pallis. “Big data and cloud computing: A survey of the state-of-the-art and research challenges”. Advances in Mobile Cloud Computing and Big Data in the 5G Era, Studies in Big Data 22, Springer International Publishing, 2017.

- [8] N. Zanoon, A. Al-Haj, S. M. Khwaldeh. "Cloud computing and big data is there a relation between the two: A Study". *International Journal of Applied Engineering Research*. Volume 12, 2017.
- [9] C. Lakshmi, V. V. N. Kumar. "Survey Paper on Big Data". *International Journal of Advanced Research in Computer Science and Software Engineering*. Vol. 6, Iss. 8, 2016.
- [10] E. Nada, E. Ahmed. "Big data analytics: A literature review paper". Springer International Publishing Switzerland. *Lecture Notes in Computer Science*, 2014.
- [11] A. Anurag, R. Dharmaji Rao. "A review paper on big data: technologies, tools and trends". *International Research Journal of Engineering and Technology*. Vol. 04, Iss. 10, 2017.
- [12] S. Singh, S. Pankaj, G. Rakhi, P. K. Mishra. "Big Data: Technologies, Trends and Applications". *International Journal of Computer Science and Information Technologies*. Vol. 6 (5), 2015.
- [13] Tom White. "Hadoop The Definitive Guide: Storage and Analysis at Internet Scale". O'Reilly Media, 2015.
- [14] N. Ameya, A. Poriya, D. Poojary. "Type of NOSQL databases and its comparison with relational databases". *International Journal of Applied Information Systems*. Vol. 5, No.4, 2013.
- [15] Dadapeer, N. M. Indravan, G. Adarsh. "A Survey on security of NoSQL databases". *International Journal of Innovative Research in Computer and Communication Engineering*. Vol. 4, Iss. 4, 2016.
- [16] P. Singh, E. A. Jain. "Survey paper on cloud computing". *International Journal of Innovations in Engineering and Technology*. Vol. 3, Iss. 4, 2014.
- [17] P. Rajeswari, K. Jayashree. "Survey on QoS metrics and ranking in cloud services". *International Journal of Engineering & Technology*. 7 (1.3), 2018.
- [18] A. KH. Bardsiri, S. M. Hashemi. "QoS metrics for cloud computing services evaluation". *Intelligent Systems and Applications*. 12, 27-33, 2014.
- [19] Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 (White Paper). February, 2018.
- [20] Louis Columbus. "Cloud computing market projected to reach \$411B by 2020". *Forbes*, October, 2017.
- [21] S. G. Kumar, V. Steve, B. Rajkumar. "SMICloud: A Framework for comparing and ranking cloud services". In proceedings of Fourth IEEE International Conference on Utility and Cloud Computing, 2011.
- [22] M. H. Ghahramani, M. Zhou, T. Hon. "Toward cloud computing QoS architecture: Analysis of cloud systems and cloud services". *IEEE/CAA Journal of Automatica Sinica*, VOL. 4, NO. 1, 2017.
- [23] R. Han, Z. Jia, W. Gao, X. Tian, L. Wang. "Benchmarking big data systems: state-of-the-art and future directions". *TECHNICAL REPORT. ICT, ACS*. 2015.
- [24] R. Han, X. Lu. "On big data benchmarking". In proceedings of Workshop on Big Data Benchmarks, Performance Optimization, and Emerging Hardware. 2014.
- [25] N. Sorensen, "Cisco industry benchmarks performance".
- [26] G. T. Hicham, E. A. Chaker, E. Lotfi. "The optimization of compute resources scheduling in cloud computing environments using artificial neural networks". *Journal of Theoretical and Applied Information Technology*, Vol.96. No 13, July 2018.