International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3, February 2019

Efficient System Performance for Data Replication in Cloud Computing

A Ramamurthy, V.Goutham

Abstract*: Unlike conventional superior registering condition, for example, group and supercomputers, the cloud processing is a gathering of interconnected and virtualized figuring assets that are figured out how to be one brought together high performance processing power. Be that as it may, the Cloud condition establishes a heterogeneous and a profoundly powerful condition. Dissatisfactions at the server ranches centers are standard as an option because of the broad size of real resources and information. Eventually, the cloud condition requires a gainful adaptable insights replication organization saving in musings the quit goal to advance to the natural ordinary for the Cloud circumstance. on this paper, we prompt a records replication way which adaptively picks the measurements data for replication with a specific surrender reason to decorate the general constancy of the system and to satisfy the ideal incredible of organizations. Likewise, the proposed strategy picks effectively the amount of generations and correspondingly the convincing information center points for replication. The generally realized actualities documents are chosen for replication in perspective of using a lightweight time-course of action procedure, which dismembers the proceeded with case of insights certainties requests, and offers estimates to the future realities demands. Exploratory outcomes uncover that the proposed system continues satisfactorily to improve the enduring pleasant of the Cloud structure underneath examination.

Index Choice: System availability, replication, adaptive, cloud computing

I. INTRODUCTION

Distributed computing is a substantial scale parallel and disseminated registering framework. It includes an aggregation of interconnected besides, virtualized preparing resources that are made sense of how to be one bound together enlisting resources. The gave dynamic, virtual resources, for instance, frameworks, servers, accumulating, applications and data, can be passed on as an organization. Organizations are passed on demand to the end-customers over quick Internet as three sorts of handling configuration, to be specific Programming as a Service (SAAS), Platforms as a Service (PAAS) and Infrastructure as an administration (IAAS). The primary objective is to provide customers progressively versatile corporations direct, greater less costly, flexible, substantially open and memorable enlisting resources[1].The software program as a service (SaaS) designing offers programming applications advocated and supervised by means of a master affiliation for the quit-consumer superseding secretly run programs with internet groups programs. in the Infrastructure as a Service (IaaS), Service incorporates administration of equipment and programming for preparing, information stockpiling, systems and any required foundation for sending of working frameworks and applications which would typically be required in an information focus overseen by the client. In the Platform as a Service (PaaS), supplier contains programming tongues and mechanical assemblies and an utility transport level encouraged by method for the authority business venture to help advancement and movement of end-buyer programs[2]. at the point when all is said in accomplished, the Cloud Computing offers the item and contraption premise as organizations using gigantic scale insights centers[3]. at last, Cloud figuring moved away the estimation and records accumulating from the stop buyer and onto extensive wide assortment of server ranches establishment. This mitigates clients of the burdens of framework and readiness provisioning and organization. gadget and writing computer programs are passed on to clients as on-ask for advantages over the web. The Cloud structure can scale out the framework potential to fulfill the ideal non-helpful nature of organizations (QoS). be that as it may, it is endeavoring to give high enduring lovely and skilled get admission to the cloud server cultivates in perspective of the far reaching scale and dynamic nature of the Cloud. Replication is the way toward giving different impersonations of a comparable administration at different nodes[4]. Replication is a connected way inside the advanced restrictive fogs structures, for example, GFS (Google report system) and HDFS (Hadoop disseminated report gadget) [5, 6]. in the cloud, records replication is cultivated through records resource pool and the quantity of measurements duplicates is statically set in gentle of records and delight in[7]. additionally, it isn't urgent to make propagation for all measurements records, uniquely for those non a great deal of the time or as of past due used information insights. In this manner, it's miles vital to adaptively mirror the each now and again used realities data, decide the measure of information impersonations and the actualities center points regions wherein to put the new impersonations as shown by methods for the current cloud circumstances. on this paper, we exhort a flexible replication strategy in a cloud area that adaptively adjusts to the ensuing issues:

• What to duplicate to enhance the non-valuable QoS. The select framework is transcendently depends on assessing the historical backdrop of the measurements needs making utilization of a light-weight time-affiliation gauge estimation. using the normal information ask, we can recognize what information archives expect replication to adorn the structure constancy.

Manuscript published on 28 February 2019.
***** Correspondence Author (s)
Dr A Ramamurthy, Professor CSE, TKR Engineering College, Hyderabad, Telangana, India.
Dr V.Goutham, Professor CSE, SREYAS institute of Engineering College, Hyderabad, Telangana, India.

The quantity of reproductions for each chose information. The situation of the new reproductions on the accessible server farms.

The overhead of replication system on the Cloud framework. This is the most vital factor of the proposed

versatile replication technique where the Cloud has an expansive number of server farms and also an extensive scale information. in this manner, the flexible replication framework ought to be light-weight approach.

The proposed bendy replication method is before everything impelled through the manner that the starting overdue most were given to statistics statistics can be gotten to again within the close to destiny as tested via the accrued estimate bits of expertise of the information get to sample[8, 9]. A replication element is processed in angle of a facts square and the availability of every present multiplication passes a destined area, the replication movement can be actuated. Some other impersonation may be made on any other middle factor which achieves a regular new replication thing. The quantity of latest impersonations may be settled adaptively in light of updating the openness of every file heuristically. This paper relies upon upon the issue formalization delineated in [9]. Regardless, we use a light-weight time-sport plan estimation for predicting the destiny soliciting for of records. The replication choice is on a completely primary stage depending on the gave figures. The heuristic proposed for the dynamic replication framework is computationally decrepit, and can manage huge scale sources and information in a sensible time. Something is left of this paper is dealt with as takes after. Territory 2 demonstrates the associated paintings on statistics collecting and facts replication of circulated registering structures. Phase 3 shows a formalization of a cloud shape seem. Phase 4 depicts the dynamic facts replication framework, inclusive of the replication decision, the amount of duplicates, and the reproduction position. vicinity 5 maintains a watch on the propagation circumstance, parameter setup and execution evaluation of the proposed dynamic information replication strategy. subsequently, finishes and destiny paintings are given in phase 6.

II. RELATED WORK

This area presents two general classes of related work. The main classification examines cloud server farm design, also, the second classification displays the related work to the replication in the Cloud situations.

A. Cloud design

Distributed computing innovation moved computation and records storing frequently from the surrender client onto enormous assortment of backend information facilities. This facilitates clients of the burdens of framework and programming provisioning and the executives. at some point or another, programming and hardwarea re passed on as organizations and wolfed over the net[10-12]. The Cloud server ranches depend on the appropriated record structure development, for instance, Google File System (GFS), the Hadoop conveyed document framework (HDFS). The GFS design comprises of three segments, in particular different customers, a solitary ace server and numerous piece servers[5]. statistics records are disconnected into many settled size portions that are secured within the bump servers. Bumps are secured in undeniable Linux information that are imitated on numerous middle points to offer high-openness and improve execution. The pro server maintains up all the metadata of the file shape, which includes the namespace, the passageway manage records, the mapping from facts to irregularities, and the prevailing zones of pieces. customers ship facts soliciting for to the seasoned server, and the professional server controls the income to the piece servers[13].The HDFS handed on file machine moreover takes after a seasoned/slave designing which entails a unmarried pro server, called the Name node, that deals with the unfold document shape namespace and courses get admission to to information by using clients. In like manner, there are various data nodes, one in every group, which manipulate the plate gathering joined to the middle factors consigned to Hadoop. The Name node chooses the mapping of squares to data nodes [6].

B. Replication in the Cloud

Replication innovation is one of the helpful procedures in disseminated frameworks for enhancing accessibility and reliability[9]. In Cloud processing, replication is utilized for lessening client holding up time, expanding information accessibility and limiting cloud framework data transmission utilization by conferring the supporter severa duplicates of a specific organization on different centers. for instance, at the off danger that one center point comes up brief, a duplicate of the failed organization may be possibly made on a substitute center point keeping up in considerations the stop objective to strategy the requests[12]. severa replication procedures had been proposed in the composition, which can be orchestrated into static and dynamic replication. In a static replication, the amount of duplicates and their zones are at first set early [5, 6, 13]. then again, dynamic replication relentlessly makes and deletes duplicates as shown by method for changing over condition stack circumstances [10, 13]. there was an energizing scope of works for records replication inside the Cloud figuring. For point of reference, in [5], a GFS static cloud records replication figuring is proposed.The GFS gets a replication method wherein a solitary pro places the chose realities piece impersonations on irregularity servers with under common plate space utilization. thus, the measure of replication for every datum throw is controlled by utilizing the customers. So likewise, in [6], a product can choose the scope of impersonations for each document, and the square length and replication thing are configurable per record. In [13], a static included realities replication count units an explicit scope of duplicates in mellow of the base weighted partition. In [10], a dynamic passed on cloud insights replication estimation CDRM depends upon at the HDFS record system where the replication site are picked in perspective of the center points which have low utilization. In [13], six particular powerful data replication figurings, Caching-PP, Cascading-PP, quick spread-PP, Cascading-better, and fast spread-progressively attractive are proposed. In [14], a dynamic united data replication computation is proposed. The computation treats uses a weighting perspective for the replication, and utilizations various figure frameworks.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

III. PROBLEM FORMALIZATION

A cloud data service system typically consists of the scheduling broker, replica selector, replica broker and data centers[8, 11, 12, 14-18], as shown in Figure 1. The scheduling broker manages the replication process, and contains all the information about the number of replicas as well as their location at the different data centers. The formal model about the Cloud data centers architecture is well described at [9] as follows:

Let $U = \{u1, u2, \dots, um\}$ be m users at the Cloud, $TS = \{TS1, TS2, \dots, TSm\}$ be a set of tasks of the user set U, and $TSj = \{tsj1, tsj2, \dots, tsj\ mj\}$ be a subset of tasks of the jth user uj, where mj is the number of subtasks, and tsk is the kth task submitted to the scheduling broker through a Cloud interface. If $u0$ has two tasks, then $TS0 = \{ts01, ts02\}$, and $m0 = 2$. A task tsk is characterized by a 4-tuple $tsk = (tidk, trk, tdk, tfnk)$, where $tidk, trk, tdk\ and\ tfnk$ are the task identification, task generation rate, task deadline time and the number of required files of task tsk, respectively[11, 15, 19]. $DC = \{dnd1, dnd2, \dots, dndn\}$ representsa data center consists of n data nodes on a physical machine PM. Each noderuns a virtual machine, and is characterized by $dndk$ that is a 5-tuple $dndk = (dndk, drk, dstk, dfk, dbwk)$, where dn, drk, $dstk$, dfk and $dbwk$ are the data node identification, request arrival rate, average service time, failure probability and network bandwidth of data node $dndk$, respectively. In order to guarantee the service performance of the data center DC, the task generation rate $trk$of user set U, the request arrival rate $drk$and failure probability$dfk$of DC should meet (1).

$$\sum_{j=0}^{\#subtasks} tr_j \le \sum_{i=0}^{n} dr_i \times (1 - df_i)$$

wheretrj is the task generation rate of task j, dri isthe request arrival rate of task $j$on the node i, dfk is the failure probability of task j.

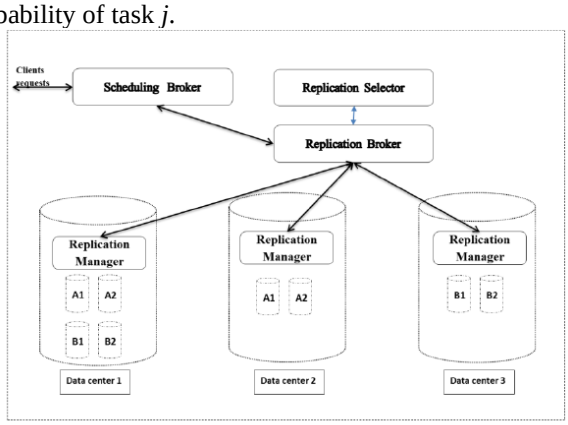

Fig.1. The Cloud data server architecture.

Let $F = \{f1, f2, \dots, fl\}$be a data file set of a datacenter DC. $B = \{B1, B2, \dots, Bl\}$be a set of block sin the data center DC, and $Bi = \{bi1, bi1, \dots, bin1\}$ bethe i-th subset of blocks belonging to the i-th data file fi, which is stripped into ni fixed blocks according to its length .A block bk is characterized by a 5-tuple$bk = (bidk; bpk; bsk; bnk; btk)$, where $bidk$, bpk, bsk, bnk and btk are the block identification, number of requests, block size, the number of replicas and the last access time of block bk, respectively. When user uj requests a block bk from an information node$dndi$with data transfer capacity execution ensure, transmission capacity $bsk/dsti$ ought to be doled out to this session. The total data transmission used to help distinctive solicitations from use set U ought to be less than db, as appeared (2).

$$\sum_{i=0}^{s_i} \frac{bs_k}{dst_i} \le dbw_i$$

wheresi is the maximum number of network sessions of data node $dndi$that can serve concurrently, bsk is the block size of block bk, $dsti$ is the average service time of data node $dndi$, $dbwi$ is the network bandwidth of data node $dndk$. Block availability is the ability of a data block to provide proper service under given constraints. The block availability of a block bk is denoted as B. $(BA)$is the probability of block bk in an available state. $P\ (BA)$is the probability of Block bk in an unavailable state, and $P\ (BAk) = 1 - P(BAk)$.The number of replicas of block bk is bnk. in reality square bk is viewed as inaccessible simply if each one of the reproductions of square bk aren't on hand. So the accessibility and inaccessibility of square bk are determined as 3 and

$$P(BA_k) = 1 - (1 - P(ba_k))^{bn_k}$$

$$\hat{P}(BA_k) = (1 - P(ba_k))^{bn_k}$$

4.

File availability is the ability of a data file to provide proper service under given constraints. The file availability of a data file fi is denoted as FAi. $(FAi)$is the probability of data file fian available state. $(FAi)$is the probability of data file fiin an unavailable state, and $P\ (FAi) = 1 - P(FAi)$. If the data file fi is stripped into ni fixed blocks denoted by $Bi = \{bi1, bi2; \dots; bin1\}$, which are distributed on different data nodes. $Ni = \{bni1, bni2, \dots, bnini\}$is the set of the numbers of replicas of the blocks of Bi. The availability and unavailability of data file fi is given as follows:

$$P(FA_i) = (1 - (1 - P(ba_i))^{bn_i})^{n_i}$$

On the off chance that the information record fi is stripped into niblocks, there are ni reproductions of each square in information document fi,and all squares at a similar site will have indistinguishable accessible likelihood from all squares are put away in information hubs with a similar arrangement in cloud information focuses, the accessible likelihood of every reproduction is $p(bai)$ in information document fi.

IV. DYNAMIC DATA REPLICATION STRATEGY

The proposed versatile information replication has three essential stages: 1) the determination stage: which information documents ought to be chosen for replication; 2) setting the quantity of appropriate reproductions to meet the predefined QoSs.; 3) discovering the best area of generations.. The starter step is to choose which records reproduce and the replication timing. The assurance level separates the authentic background of the information get to requests, and makes utilization of a light-weight time course of action technique to predict the fate get admission to repeat of the records. inside the event that the normal fate sales of information divides outperforms an adaptable part, the certainties irregularities can be decided for replication.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

supply pdk an opportunity to be a notoriety level of a rectangular . pdk is portrayed on the grounds that the predetermination get section to frequency based on the amount of access call for,ank(t)ata time t, the universality recognition $pdk$of a square bk might be found out using Holt's Linear and Exponential Smoothing (HLES).Holt's Linear and Exponential Smoothing (HLES) is a computationally shabby time arrangement forecast system. HLES is chosen for its ability of smoothing and giving here and now expectations to the estimated demands entry rates and administration request rates. Subsequently, HLES engages the proposed framework to screen the entrance statements and organization charges and to give a transient desire to the future touchdown statements and the executives costs with low figuring time. utilizing those anticipations, we can predict the utilization on every server have using condition (2), and expect the future response time of the web advantage making utilization of circumstance (1). HLES smoothes the time course of action and gives a transitory decide in mellow of the example which exists over the time game plan [20]. expect ank t is a period course of action regard at time t. The promptly guess for the m adventures ahead is according to the accompanying:

$$pd_k = an_k(t+m) = L_t + b_t \times m$$

whereL_t and b_t are exponentially smoothed estimates of the level and linear trend of the series at time t:

$$L_t = \alpha an_k(t) + (1-\alpha)(L_{t-1} + b_{t-1})$$

whereα is a smoothing parameter, $0 < \alpha < 1$,

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$

whereβ is a trend coefficient, $0 < \beta < 1$.

A large value of α adds more weight to recent values rather than previous values in order to predict the next value. A large value of β adds more weight to the changes in the level than the previous trend. The replica factor is defined because the ordinary of the share of the reputation degree and the averageavailability of copies at the extraordinary records hubs of all squares l of the records document . it is applied to determine if the information report fimust be imitated, meant as

- Initialize available and unavailable probability of each replica of block b_k,$p(ba_k)$ *and* $\hat{p}(ba_k)$.
- for each data file $f_i$at all data centers DC do
  - Calculate the popularity degree pd_k of a block b_k of data file f_i
  - Calculate replica factor RF_i of data file f_i.
  - If $RF_i$is less than a threshold λ, trigger the replication for the file f_i
- end for
- for each triggered replication for data file $f_i$do
  - for each block $b_k$in the file f_i
    - Calculate the new RF_i by adding a replication on the each data center dc_k.
    - apply the replication which gives the highest new RF_i.
  - end for
- end for
- find the file $f_i$which has the least$\sum_{k=1}^{l} pd_k$.
- for each replica in the file
  - delete the replica which gives the new, without the replica,RF_i bigger than a threshold λ.
- end for

Fig. 2.The proposed adaptive replication strategy.

$$RF_i = \frac{1}{l}\sum_{k=1}^{l} \frac{\sum_{k=1}^{bn_i} pd_k \hat{P}(BA_k)}{bn_i}$$

wherepd, P BA, l and $bni$are the popularity degree, the failure probability of a block bk , number of blocks and number of replicas of data file fi ,respectively. In each time interval T, the replication operation of the data file fi will be triggered if the replication factor RFi is less than a specified threshold. The details of the proposed adaptive strategyare shown in Figure 2.

V. SIMULATION AND PERFORMANCE EVALUATION

This segment assesses the viability of the proposed versatile replication strategy.The CloudSim structure is a Java based recreation stage for the Cloud condition, it bolsters demonstrating and reproduction of vast scale cloud figuring server farms, including clients and resources[21-23].In the CloudSim reproduction, 64 server farms are made with the relating topology appeared in Figure 1. The specialist co-ops are spoken to by 1000 virtual machines, and the preparing components (PEs) assortment of each virtual gadget is in the extent of two to 4. one hundred exceptional data measurements are set in the designated carport circumstance, with every length inside the extent of [0.1, 10] GB. each record is secured in settled size (bs= 0:2 GB) amassing unit called rectangular. Squares of comparable measurements report are dissipated across over various advanced machines. 10000 assignments are submitted to the expert centers using the Poisson dispersal. every task calls for 1 or 2 records reports aimlessly. toward the begin, the measure of generations of each datum record is 1 and set aimlessly. For ease, it's far typical that the thought part of measurements accumulating is rectangular and the part of replication is one blend records record. As appeared in perceive 3, with time passes, the numberof impersonations is expanding inward a snappy length oftime. Then,the scope of duplicates is set aside at arelatively stable stage, that is dealt with the guide of the adaptable parameter □ in the HLES approach. We assume that the additional conspicuous theadjustable parameter □ and the expanding rectangular call for bpk of a positive report, the more prominent proliferations are neededto embellish the archive openness. The reaction time for a realities fileis the interim among the convenience time of the taskand return time of the result. The ordinary responsetime of a system is the suggest estimation of the reaction time for all data request endeavors of the customers, which can be acquired with the guide of the accompanying condition.

$$rt_{avg} = \frac{\sum_{j=1}^{m}\sum_{k=1}^{m_j}(ts_{jk}(rt) - ts_{jk}(st))}{\sum_{j=1}^{m} m_j}$$

where$tsjk$ (st) and $tsjk$ (rt) are the accommodation time and the arrival time of the consequence of assignment k of the client j, separately, and mj is the quantity of the undertakings of client j. As shownin Figure 4, with the quantity of undertakings expanding and □ = 0.7, the reaction time increments drastically. The less the blockavailability, the more drawn out the reaction time will be.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Obviously the proposed versatile replication procedure improves the reaction time and keeps up the responsetime at a steady level inside a brief timeframe.

VI. CONCLUSIONS AND FUTURE WORK

This paper proposes a versatile replication procedure in the cloud circumstance. The method examines the openness what's an increasing number of, a hit get admission to of each report in the server ranch, and thinks the way to enhance the relentless nature of the statistics reviews in mild of choice for the client get admission to to the squares of each archive. The proposed adaptable replication framework redeploys successfully massive scale differing records impersonations on specific statistics middle points with immaterial fee the usage of heuristic output for every replication. The proposed bendy framework is based upon a proper depiction of the problem. The framework perceives the data which are general file for replication in light of looking on the continuous records of the information access to the records the usage of HLES time plan. when a replication issue in angle of the popularity of the files isn't always as a great deal as a particular restriction, the replication banner may be initiated. along those lines, the flexible technique perceives the first-rate replication sector in light of a heuristic look typically favorable replication element of every document. check assessment demonstrates the profitability of the proposed flexible replication technique within the cloud condition. future studies work will base on building a database machine for 3-d fashions. sincerely, fantastic three-D modelsare archived in wonderful facts. these facts are commonly secured in handed on databases which come upon the evil affects of noticing portrayal request and movement over-loading on server ranches. We factor toprovide a framework for quickening records get to, and in addition developing information availability for such databases on a cloud place. further, we are able to consider the usage of Genetic estimations to find the best replication in less time. furthermore, the replication strategywill be despatched and tried on a real cloud enlisting stage. future paintings is also needed to provide the flexible records replicationstrategy as a piece of cloud figuring corporations to meet the characteristics of circulated registering.

REFERENCES

1. Buyya, R., et al., Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Gener. Comput. Syst., 2009. **25**(6): p. 599-616.
2. Armbrust, M., et al., A view of cloud computing. Commun. ACM, 2010. **53**(4): p. 50-58.
3. Mell, P.M. and T. Grance, SP 800-145. The NIST Definition of Cloud Computing. 2011, National Institute of Standards \& Technology.
4. Bj, M., et al., Optimizing service replication in clouds, in Proceedings of the Winter Simulation Conference. 2011, Winter Simulation Conference: Phoenix, Arizona. p. 3312-3322.
5. Ghemawat, S., H. Gobioff, and S.-T. Leung, The Google file system. SIGOPS Oper. Syst. Rev., 2003. **37**(5): p. 29-43.
6. Shvachko, K., et al., The Hadoop Distributed File System, in Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST). 2010, IEEE Computer Society. p. 1-10.
7. Bonvin, N., T.G. Papaioannou, and K. Aberer, Dynamic cost-efficient replication in data clouds, in Proceedings of the 1st workshop on Automated control for datacenters and clouds. 2009, ACM: Barcelona, Spain. p. 49-56.
8. Chang, R.-S. and H.-P. Chang, A dynamic data replication strategy using access-weights in data grids. J.Supercomput., 2008. **45**(3): p. 277-295.
9. Sun, D.-W., et al., Modeling a Dynamic Data Replication Strategy to Increase System Availability in Cloud Computing Environments. Journal of Computer Science and Technology, 2012. **27**(2): p. 256-272.
10. Wei, Q., et al., CDRM: A cost-effective dynamic replication management scheme for cloud storage cluster., in 2010 IEEE International on Cluster Computing. 2010. p. 188 - 196
11. Bonvin, N., T.G. Papaioannou, and K. Aberer, A self-organized, fault-tolerant and scalable replication scheme for cloud storage, in Proceedings of the 1st ACM symposium on Cloud computing. 2010, ACM: Indianapolis, Indiana,USA. p. 205-216.
12. Nguyen, T., A. Cutway, and W. Shi, Differentiated replication strategy in data centers, in Proceedings of the 2010 IFIP international conference on Network and parallel computing. 2010, Springer-Verlag: Zhengzhou, China. p. 277-288.
13. Dogan, A., A study on performance of dynamic file replication algorithms for real-time file access in Data Grids. Future Gener. Comput. Syst., 2009. **25**(8): p. 829-839.
14. Wang, S.-S., K.-Q. Yan, and S.-C. Wang, Achieving efficient agreement within a dual-failure cloud-computing environment. Expert Syst. Appl., 2011. **38**(1): p. 906-915.
15. McKusick, M.K. and S. Quinlan, GFS: Evolution on Fast-forward. Queue, 2009. **7**(7): p. 10-20.
16. Lei, M., S.V. Vrbsky, and X. Hong, An on-line replication strategy to increase availability in Data Grids. Future Gener. Comput. Syst., 2008. **24**(2): p. 85-98.
17. Jung, D., et al., An effective job replication technique based on reliability and performance in mobile grids, in Proceedings of the 5th international conference on Advances in Grid and Pervasive Computing. 2010, Springer- Verlag: Hualien, Taiwan. p. 47-58.
18. Yuan, D., et al., A data placement strategy in scientific cloud workflows. Future Generation Computer Systems, 2010. **26**(8): p. 1200-1214.
19. Litke, A., et al., A Task Replication and Fair Resource Management Scheme for Fault Tolerant Grids Advances in Grid Computing - EGC 2005, P. Sloot, et al., Editors. 2005, Springer Berlin / Heidelberg. p. 482-486.
20. Makridakis, S.G., S.C. Wheelwright, and R.J. Hyndman, eds. Forecasting: Methods and Applications, 3rd Edition.1998.
21. Calheiros, R.N., et al., CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Softw. Pract. Exper., 2011. **41**(1): p. 23-50.
22. Wickremasinghe, B., R.N. Calheiros, and R. Buyya, CloudAnalyst: A CloudSim-Based Visual Modeller for Analysing Cloud Computing Environments and Applications, in Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications. 2010, IEEE Computer Society. p. 446-452.
23. Xu, B., et al., Job scheduling algorithm based on Berger model in cloud environment. Adv. Eng. Softw., 2011. **42**(7): p. 419-425.