Optimal Solution for a Disaster Recovery (DR) site Across Multiple Cloud Service

Varun Gupta[1], Anil Kumar Mahto[2]
{vgbharat@gmail.com[1], anil.nitk01@gmail.com[2] }

Student M. Tech. Comp. Sc. & Engg, Dept of CSE, SEST, JAMIA HAMDARD[1],
Asst. Professor, Dept of CSE, SEST, JAMIA HAMDARD [2]

Abstract. Disaster Recovery (DR) is an essential trait for all enterprises to maintain Business Continuity and to survive if disaster happens.

During the past decades, Cloud computing gained popularity for a DR site due to advantages like specialized management dealing with data, specialized & reliable devices with the advantage of economy of scale, optimum use of resources, and it can handle Big Data, Data Mining , Mobile Computing, Block chain, high performance computing tasks efficiently.

Further enhancement in technology and benefits of cloud computing give birth to a new generation of cloud computing technology called Multi cloud computing.

Both models Single-Cloud and Multi-Cloud environment provides DR sites and different geographical zones to deal with disaster, with a difference that Multi-cloud environment operates by different vendors and provides greater flexibility to users.

For a DR site, the Single-Cloud paradigm poses risks for a DR site due to failure of single vendor services, monopoly, damage caused by human errors, fraudulent practices adopted by the vendor etc.

Whereas Multi-Cloud environment, overcome this deficiency with better flexibility , control, cost advantage and decide which workloads are going to be run and where they should be run efficiently.

Due to this Multi-cloud environment is preferred choice for Disaster Recovery (DR) site.

The objective of this paper is to find out the optimal solution for the DR site.

Keywords: Multiple cloud, Polynimbus, DR site, Data Reliability, Sky Computing, Data Replication, Optimal solution for Disaster Recover.

1 Introduction

1.1

Disaster can happen through various reasons – man-made errors, natural disasters (Earth quake, Floods, Fire accident, etc), electrical and power failure problems, temperature and humidity failure, cyber attacks, pandemics and diseases, strikes , warfare, terrorism , sabotage etc.

Performing a careful and complete Business Impact Analysis (BIA) is important to develop an effective Disaster Recovery Plan for an organization or a business.

Business Impact Analysis (BIA) determines two important parameters Recovery Point Objective (RPO) and Recovery Time Objective (RTO) for making backup strategy, whereas the RPO defines the amount of data loss that a business can bear, i.e. the maximum allowable limit or "tolerance capacity" limit to maintain business continuity during an outage.

RTO is the network downtime that a business can tolerate and maintain business continuity, and tells how long it will take to recover after an outage.

Based on RPO frequency and RTO downtime, Businesses can make Backup and DR plans.

Based on RPO frequency, Backup plans can be prepared. For e.g. RPO of 24 hours, Backups can be created on tape libraries at every midnight, shipping of database logs for 1 hour RPO to remote site after every hour, if RPO is on the order of minutes then mirroring of data can be asynchronous, and if the RPO is near zero then the data mirroring can be done synchronously at the remote site.

Based on RTO, Businesses can make plans for DR. For e.g. if RTO is 1 hour, then faster backup services (such as disk-based backup) are required, So that operations can be restored faster than tape based backups.

For a 1-week RTO, the tape may meet backup requirements.

Organizations can make a DR plan on the basis of organizations endurance for data loss (RPO) and downtime (RTO) objectives.

For a stock exchange operation, a momentary downtime can cause a big loss to traders so for such an organization synchronous mirroring of data transfer from a remote site should be performed without fail to deal with any kind of disasters.

1.2

With the advancement of technology, DR services in these days are available at affordable prices in the industry and rapidly adopted by the business community to maintain their Business Continuity.

Cloud computing paradigm comes out as one of the preferred solution to make a DR site and to improve service quality with low latency as services can be efficiently delivered with their geographically dispersed data centers to nearby customer zones.

1.3

However in the past decade it is observed that Single cloud paradigm poses risks for making a DR site due to various reasons - CSP monopoly, Single Point failure if the vendor fails, Lack of customer's choice, Everything moved to one cloud with risk of security and downtime, vendor specific service issues, Flexibility issue that Production cannot be moved to the new environment, if available, Limitations on scalability of current solutions, No multi choices for the Business Organization, Cloud usage with varying levels of security concerns, Geo-location for latency and/or policy compliance etc

In 2009, Microsoft sidekick data outage happens, resulting in loss of personal data (emails, address books, photos etc) of approx. 8 lakh smartphone users from their mobile phone handsets. [8]

Amazon AWS, a large and reputed cloud service provider suffered outages [8] as followed:
-    Suffered an outage on 20 April 2011, resulting in an interruption in the elastic block store (EBS) service and could not execute read / write requests,
-    On June 29, 2012, a severe storm in Northern Virginia impacted AWS's largest data center and disrupted the service of many websites servicing AWS.
-    A major outage occurred on 22 October 2012 and AWS disrupted the services of Reddit, Foursquare, Pinterest and others.
-    On 24 December 2012, due to an outage in AWS, services of many websites including Netflix were disrupted.
-    On February 8, 2017, AWS suffered a massive outage in its Northern Virginia region of S3 services, and that outage was caused by a human error during debugging and replacement of servers.

In 2016, Google's Gmail outage hits US and Europe and services were down for several hours. [8]

Multi-cloud [9][10][11][12][13][14][15][16] computing comes as a solution to the above problems with a choice to choose required services from multiple CSPs with cost advantage, policy, and oligopoly and effectively deal with disaster, if a single CSP fails due to war, cyber attack, natural disasters etc.
Customer can make integration between multiple CSPs to serve customers across different geographical zones with low latency by taking services from nearest data centers of CSPs resulting in improved business performance & efficient delivery of services.

Multi-cloud (Cloud of clouds) environments provide better solutions for data mining, big data analytics, block chain and high performance computing tasks,
Because it has large scalable computing and storage capabilities of multi-cloud with the use of the economy of scale, multi-cloud capability is fully utilized due to virtualization and resource sharing principle, and integrated solutions are made available to users.

1.4

According to a Gartner (A global research and advisory firm) [13] survey that was conducted in the year 2019, Most organizations prefer to work with multiple cloud service providers (CSPs), whereas 81% of respondents stated that they are working with two or more than two CSPs.

According to RightScale survey that was conducted in 2019 [14] -
(RightScale is a software company that offers cloud management and analytics tools for managing public, private and hybrid clouds.)
- 84% respondents avail multi-cloud services.
- 94% respondents use cloud services.
- Priority of enterprises is to maintain the balance between public and private clouds.
- Enterprises avail about 5 clouds on an average.
- Enterprise cloud spending is growing significantly and rapidly.
- Companies run most of the workload in the cloud.
- Management of Enterprises focuses on managing and optimizing IT cloud costs.
- Managing cloud spending and cloud governance are the top challenges for enterprises.
- The topmost priority of cloud service users in 2019 is cloud cost optimization.

As many cloud service providers are providing data storage and computing services through their data centers distributed across the world with varying costs, parameters and capabilities,
Therefore, for cloud customers, it is challenging how to choose CSPs that meet their SLO objectives with optimal costs and allocate data among them.

The purpose of this paper is to find a solution of a disaster recovery site with a low cost and efficient solution i.e. an optimal cost.

2   Latest Literature Review

Cumulus system [17] is proposed to take backup on the single cloud environment and compared various trade-offs between backup on local file system vs cloud platform.
During that time (in 2009) single cloud system was popular and multi cloud environment was at nascent stage. They didn't study about backup on multi-cloud environment.

Wood et. al. [18] in their conference paper (Oct.2011) proposed to backup on a single cloud using a pipeline approach and compare the trade-offs between synchronous versus asynchronous replication on the cloud.
Their disaster recovery cloud model focuses on a single cloud environment and has not addressed single cloud failure issues.
For example, Amazon Web Services (AWS) suffered outage in April 2011 and again in June 2012, Microsoft-owned Sidekick suffered an outage in 2009 for a week and resulting in personal data (e-mail, address book and own mobile handset etc.) of an approximately 800,000 smartphone users in the United States were temporarily lost.
Salesforce.com suffered an outage in January 2011.

Therefore, there is a risk of relying on a single CSP.

Bajpai et al.[19] Proposed remote mirroring of data in such a way to keep the primary mirror system on the local system and secondary mirror system on the single cloud system and therefore Single cloud failure problem not address here also.

Zhang and Zhang [20] proposed inter private cloud storage, and shared this storage with private clouds.
In which each private cloud has local and remote backup servers named LBS and RBS.
After the initial backup on RBS, incremental backup is used to keep, so that network bandwidth capacity can be better utilized and data backup speed is faster.

Y. Tan et al. [21] proposed a causality-based deduplication CABdedupe technology in which redundant data is removed during backup and recovery from the cloud during transmission, and only modified data to be transmitted during backup and recovery using incremental backup approach.
This reduced data for transmission and receiving from the cloud, and led to efficient use of cloud resources, resulting in improved cloud service quality.

Li et al. [22] proposed DATA STORAGE RELIABILITY MODEL to check the reliability of equipment as each device has fixed life and after that failure occurs.
And proposed data replication strategy on such reliable devices with an incremental replication method for optimum use of network bandwidth and storage cost.

R-ADMAD [23] uses error correcting code to convert variable sized objects in fixed sized chunks to remove fragmentation problems; it then distributes this data to storage nodes within the data center of the cloud, And a dynamic restore process is done to perform data recovery.
R-ADMAD based configuration provides dynamic load balancing and the average recovery time decreases by 2 to 6 times.

Nguyen et al. [24] proposed a solution for data replication in cloud computing environment known as Differentiated Replication (DiR), which gives users the option to choose different replication strategies, considering both their needs and system capacity, so that data availability, reliability, durability, and system performance can be increased.
Some researchers focused leveraging advantages of Multi-cloud providers and conducted their research as follows –

Cachin et al.[25] leveraging multi cloud architecture for addressing dependability issues and limitations of isolated single-domain clouds.
He envisioned an inter cloud interface among multi cloud environment, so that client can take more refined services from multi cloud with improved confidentiality, integrity, reliability and consistency.
With the help of fault-tolerant and secure access control protocols, Data integrity and confidentiality across multi cloud platforms is ensured.
Data access priority queue feature reduces communication delays and small size files improves consistency.
However, it lacks cost comparison and recovery parameters of different clouds.

Yu Gu et. al.[26] proposed a multi-cloud based disaster recovery service, which they named DR-Cloud.

DR-Cloud provides an integrated service interface using the resources of many cloud service providers, In which customers can perform data backup and data recovery in multi-cloud storage through the DR-Cloud..

They proposed the cost of data backup storage and retrieval via DR-Cloud.

However the proposed cost does not give the user for option of storing their data backup in a particular cloud with optimal cost.

Zhu Wu et al. [27] proposed SPANStore models to reduce data storage cost in cloud with satisfactory service level objectives (SLO) of latency and failure across multiple CSPs.

They proposed to keep replicated data in geographically distributed data centers across multiple CSPs with minimizing cost and meeting SLO. However for a DR site latency is not a primary goal and replication of data should be in such a way to maintain consistency & appropriate latency of data when disaster happens i.e. a balanced approach for replication can be used. Therefore this solution only helps to know various issues taken for minimizing cost across multiple CSPs.

Guoxin Liu et al. [28] proposed an improved solution for (i) how to allocate data to worldwide datacenters to meet SLO (service level objective) requirements including both data retrieval latency and availability, and ii) how to allocate data and reserve resources in datacenters to reduce payment costs of various CSPs.

Solution proposed by Guoxin Liu et al [28] considers minimizing cost across multiple CSPs but minimizing cost should not be the sole criteria for selection of a DR site.

Selection of a DR site across multiple CSPs should be based on optimal cost, i.e. reduced cost with efficient solution.

3 Proposed Solution

According to the above study, it has been observed that first people used to keep the data in local backup, then in the cloud environment, and now keeping in a multi-cloud environment.

Guoxin Liu et al.[28] Proposed an enhanced solution to minimize cost across multiple CSPs with reduced latency and high availability.

They considered two important factors to minimize the cost of storage across cloud service providers -

(i) Resource reservation (ii) Data center capacity limits

They proposed to minimize the total payment cost Ct to keep data in cloud environment as per following formula -

$$Ct = Cs + Cc + Cg + Cp$$

Where as -

Ct - Total payment cost for a customer (with an objective to minimize it)

Cs - Storage Cost

Cc - Transfer Cost (i.e. transfer cost of data from one data center to other data center running by the same or different CSP)

Cg - Get Operation/Data Recovery Cost

Cp – Put/Data Store Operation Cost

The Service Level Objective (SLO) considers the following constraints to meet the objectives:

(i) Deadline

(ii) Availability

If above SLO satisfied then total minimum cost Ct for a customer can be calculated for cloud services of CSPs.

However, in the solution given by Guoxin Liu et al.[28], the following points were not taken into consideration - Ratings and responses given by CSP users, reliability of CSP devices, customized solutions offered by CSPs, Selection of new optimal CSP solutions when available at a better cost over time, financial capability to run cloud environments, etc.

In terms of device reliability, some cloud vendors may have the latest devices, while some may be providing services through older devices with shorter life. [22]

Therefore, minimum-cost cloud storage service across multiple CSPs with satisfactory SLO cannot be considered as sole criteria to be chosen by a customer.

We have to investigate other parameters like customer's review & rating given to CSP, experience in the field of cloud computing, Capex & Opex for running services, number of customers, number of outages, Geographical spread of data centers in different countries that are in different seismic zones, thereby maintaining the availability of critical data during disasters, etc.

Let say customers give a rating from 1 to 5 star to a cloud service provider through Google Map (or any other mode by rating agencies). –   (A)

Start of operation by CSPs in years: 0 to n (in numbers)   – (B)

Reviews/ Feedbacks containing words like : Ok, Good, Very Good ,Bad, Very Bad, Worst, Cheater, Inefficient, Helpful Management, Superb, Excellent, Best, Problems, not satisfied, Awesome, Affordable Cost, Satisfied, satisfactory solution, Nice service etc. by different customers.   – (C)

Cost of Data Center:

Typically cost involved in a data center can be identified through Capex and Opex parameters.

A capital expenditure (CAPEX) sometimes known as PP&E, short for property, plant, and equipment, is the upfront investment by an organization to acquire, maintain, renovate or to improve fixed assets such as property, building, factory, equipments, technologies etc.

The capex can be found through the cash flow statement of a business organization or obtained through the organization's income statement and balance sheet.

The operating expenditure (OPEX) is a continuous expenditure for running day to day business operations. Opex includes rental fee, electricity charges, equipments, stock costs, advertising/business promotion cost, employee salary etc.

OPEX can be found through line items in the expenditure category in a cash flow statement, CAPEX are usually mentioned in column of, "Investment in property, plant, or equipment."

CHARLES FITZGERALD [31] published a comparative Capex by the top three cloud service providers - Amazon, Google and Microsoft - that these three service providers spent $ 68 billion on Capex in 2018, stating that Amazon spent the most on its cloud infrastructure.

Let say, For a CSP -
Capex - X (in million dollars),   – (D1)
Opex -   Y (in million dollars) per year.   – (D2)

Outage Happened in Data Center: m (in numbers) times   – (E)

Geographical Span in n countries out of which k data centers in
sensitive seismic zone, natural disaster area.   – (F)

No. of customers of a CSP – r (in numbers)   – (G)

To generate optimal solution from above factors:

(A) and (C) can be refined and evaluated in following way:-

Online reviews usually have star ratings and feedback text.
Star rating is simple to understand customer's overall feedback. [29]
But researcher found considerable skewness to make correlation between review text and star rating.
The reason for one person giving '4' star rating may be different from the other. Some people think '5' star rating is an ideal case and nothing is perfect in the world or 100% correct. Some people in anger may give lowest 1 rating while the product is not so bad. People having different mindset and for a same thing their opinion differs.

In general, textual comments are unclear because of their subjective content, and are more ambiguous when compared to numerical ratings. However, rich text content contains more and deeper information.

This gives a better explanation than the star rating. If richer text content aligned with the star rating then it will give clarity & value to online review system.

Sentimental analysis is a good approach to align text content with star rating.
Sentiment analysis defines the intended sentiments or expressions found in a text using traditional natural language processing methods and algorithms. [29]

In this method, a predictive model is generated by the software that learns from the training data to understand the sentiments of the text, then this model is used to determine the sentiments of an additional set.
Training data is created through different sets of user review data.

Our goal is to correctly classify reviews by aligning the review text with the corresponding star ratings.
Better alignment of text with its corresponding star rating will give more accuracy and resulting in better categorization.

Table [1] shows comparison of star ratings across sites [29]:

Table 1. Comparison of star rating guidelines across sites

	Amazon	CNET	Yelp
1 star	I Hate It	Terrible	Eek! Methinks not.
2 star	I Don't Like It	Mediocre	Meh. I've experienced better
3 stars	It's Okay	Good	A-Ok.
4 stars	I Like It	Excellent	Yay! I'm a fan.
5 stars	I Love It	Spectacular	Woohoo! As good as it gets.

As per above,
One can consider the outstanding performance of a product as equal to a 5 star rating,,
Similarly feedback can be related to star rating as below -
 Very Good ≅ 4 Star rating
Good ≅ 3 Star rating
Average ≅ 2 Star rating
Bad ≅ 1 Star rating

Reviews with star rating can be easily collected with the help of Google Map.
There are lot of Web crawler tools available like Octoparse, Cyotek WebCopy, HTTrack,Getleft,Scraper, OutWit Hub, ParseHub etc that automatically extract relevant information from websites.

For better analysis of text reviews, We can convert text into vectors with the help of TegHelper on Weka platform, such as collection of unigrams (single words), bigrams (two word phrases), part-of-speech bigrams (linguistic categories of words), punctuation, and line length. [29]

We compare, review text words with equivalent synonyms of 'Outstanding', 'Very Good', 'Good', 'Average', and 'Bad' performance words and drive rating from 5 to 1 accordingly.

Cohen's kappa coefficient k is a widely used method for assessing inter-rater reliability or agreement between two raters. The higher the level of agreement suggests higher the alignment between ratings.
In above feedback analysis, we can check agreement between text review and corresponding star rating through Cohen's kappa coefficient. The algorithm's classification based on text review analysis can be taken as the first rater, and corresponding star rating given by the reviewer as second rater.
The higher the level of agreement between first rater and second rater means algorithm works efficiently and able to guess product rating on the basis of review text only. [30]

A confusion matrix on the basis of actual and predicted information can be formed as below:

PERFORMANCE MATRIX			
		Predicted Label	
		Positive	Negative
Actual Label	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

The ability of the classifier can be tested through the following metrics:

1) Precision: Precision is the ratio of correct predictions to positive predictions i.e. analyse how correctly classifier predicts -

Precision = TP/TP+FP

2) Recall: Recall is the fraction of correct predictions among actual positive samples, It is also known as sensitivity.
Recall, we can say the knowledge of model that has gained during learning (training) process.

Recall = TP / TP + FN

3) Accuracy: Accuracy is the measurement of correct classification of items and can be calculated from following formula -

Accuracy = TP+ TN / TP+FP+TN+FN

4) F-Measure: F-Measure is a harmonic mean of precision and recall measurement. Our goal is to increase both precision and recall, and due to this harmonic mean is applied to get greater value of F-Measure.

F-Measure = 2 * Precision * Recall / Precision + Recall

5) Kappa:
Kappa score is used to measure agreement between raters, Here in above confusion matrix our goal is to check how closely classifier predict results to actual values.

Kappa score between 0 and 1 or percentage that how correctly classifier (model) predicted correct results.

$K = P(A) - P(E) / 1 - P(E)$

P(A) is relative observed agreement among raters (equivalent to accuracy).
P(E) is the hypothetical probability of possible agreement.
P(A) = (TP + TN) / (TP+TN+FP+FN)
P(E) = Total random agreement probability = P-Yes + P-No
P-Yes = ((TP + FN)/(TP+TN+FP+FN)) * ((TP+FP)/(TP+TN+FP+FN))
P-No = ((FP+TN)/(TP+TN+FP+FN))*((FN+TN)/(TP+TN+FP+FN))

The table[2] below shows the interpretation of kappa values with classification. Higher the agreement means review text is better aligned with corresponding star rating.

TABLE 2 KAPPA SCORE INTERPRETATION

Kappa Score	Description
<0	Less than Chance Agreement
0.01-0.20	Slight Agreement
0.21-040	Fair Agreement
0.41-0.60	Moderate Agreement
0.61-0.80	Substantial Agreement
0.81-0.99	Almost Perfect Agreement

To create a classifier model, a "training set" is constructed from the reviews and their star rating, which is fed to the taghelper.
This classifier model is then tested by putting additional "test sets" to measure how accurately it predicts star ratings. .
Classification model classifies review with corresponding star rating and gives output i.e. better aligned reviews with corresponding star rating.

From above (A), (B), (C), (D1),(D2), (E), (F) and (G), A typical customer can apply conditions on his own to choose the optimum solution for a DR site as per following:
(i) At least 20 customer rated 3 stars or above rating.
(ii) Minimum 5 year operational service (n ≥ 5) by a CSP.
(iii) Minimum 20 Positive Feedback calculated through Data Mining from (C) using sentimental analysis.

(iv)Minimum Capex 15 Million USD by the CSP.
(v) Minimum Opex 1 Million USD per year by the CSP.
(vi)Maximum 1 outage ($m \leq 1$) in a year.
(vii) Geographical span of Data Centers in at least 3 countries ($n \geq 3$) and these are not in sensitive seismic zone ($k=0$).
(viii) At least three replicas of data stored in different data centers of pt (vii) above.
(ix) Serving at least 10 customers ($r \geq 10$).

After applying above conditions/filter , Customer can calculate minimum-cost cloud storage service across multiple CSPs with satisfactory SLO.

The above solution for the DR site selects CSP with experienced and appropriate financial capacity that has received better reviews, having reliable services, proper replication of data in data centers, whereas these data centres exists in different geographical zones to effectively deal with disasters.

However, as described above, the risk of failure to place a DR site on a single cloud [1.3] is high and creating a DR site on more than one cloud can improve business continuity, especially in critical tasks such as stock exchange operations, Keeping data on more than one cloud is necessary to maintain business continuity, as transient downtime causes terrible conditions.
Therefore, depending on the severity of operations and business requirements, cloud service providers can be selected in ascending order costs of optimal solutions to maintain business continuity.

4 Conclusion

The above optimal solution for the DR site can be further enhanced through Artificial Intelligence and data mining techniques, in which the best solution can be selected by properly and comparatively evaluating the various customized features offered by various CSPs.
Cyber attacks are the biggest challenge for privacy and security of the DR site and to deal with these attacks, Consistent research with updated techniques and technologies should be adopted.

References

[1] RAJKUMAR BUYYA, SATISH NARAYANA SRIRAMA et. al. A Manifesto for Future Generation Cloud Computing: Research Directions for the Next Decade, ACM Computing Surveys, Vol. 51, No. 5, Article 105. Publication date: November 2018
http://www.buyya.com/papers/CloudManifesto.pdf (Cloud Computing)
[2] Dinh C. Nguyen, N. Pathirana , Ming Ding and Aruna Seneviratne, Integration of Blockchain and Cloud of Things: Architecture, Applications and Challenges.
IEEE COMMUNICATIONS SURVEYS & TUTORIALS (Block Chain)
[3] Jessica Lyons Hardcastle | Managing Editor

https://www.sdxcentral.com/articles/news/blockchain-for-multi-cloud-management-wandiscos-down-to-boogie/2019/01/ (Block Chain)

[4] Jing Ding and Shanlin Yang, Classification Rules Mining Model with Genetic Algorithm in Cloud Computing , International Journal of Computer Applications (0975 – 888) Volume 48– No.18, June 2012 (Data Mining)

[5] E. Goldin GSTAT, Israel ; D. Feldman ; G. Georgoulas ; M. Castano ; G. Nikolakopoulos, Cloud computing for big data analytics in the Process Control Industry, Published in: 2017 25th Mediterranean Conference on Control and Automation (MED) (Big Data)

[6] Sanae Esseradi ; Hassan Badir ; Abderrahmane Sbihi ; Amjad Rattrout, Mobile cloud computing: Current development and research challenges Published in: 2013 ACS International Conference on Computer Systems and Applications (AICCSA)

[7] JR Raphael, The 10 worst cloud outages , https://www.infoworld.com/article/2622201/the-10-worst-cloud-outages--and-what-we-can-learn-from-them-.html

[8] Mikhail Gloukhovtsev, Sr. Sol. Architect Orange Business Services, WHY YOU NEED MULTI-CLOUD AND HOW TO MANAGE IT , https://education.dellemc.com/content/dam/dell-emc/documents/en-us/2019KS_Gloukhovtsev-why_You_Need_Multicloud_and__How_to_Manage_It.pdf

[9] Multicloud , https://en.wikipedia.org/wiki/Multicloud

[10] Chris Pietschmann, Architecture, Infographic, The Polynimbus Cloud Enterprise, Dec 23, 2016 https://build5nines.com/the-polynimbus-cloud-enterprise/

[11] Chris Pietschmann, Architecture, Infographic, Polynimbus is the future of the cloud, June 15, 2018 https://build5nines.com/the-polynimbus-cloud-enterprise/

[12] Laurence Goasduff in Gartner Publication May 7, 2019, Why Organizations Choose a Multicloud Strategy, https://www.gartner.com/smarterwithgartner/why-organizations-choose-a-multicloud-strategy/

[13] RightScale® , STATE OF THE CLOUDREPORT from Flexera 2019, As Cloud Use Grows, Organizations Focus on Cloud Costs and Governance , http://googliers.net/static/media/uploads/download_files/2019_state_of_the_cloud_report.pdf

[14] RightScale 2018 State of the Cloud Report™ DATA TO NAVIGATE YOUR MULTI-CLOUD STRATEGY https://www.suse.com/media/report/rightscale_2018_state_of_the_cloud_report.pdf

[15] Multi-cloud strategy leads, but hybrid cloud is heating up , https://searchstorage.techtarget.com/infographic/Multi-cloud-strategy-leads-but-hybrid-cloud-is-heating-up

[16] M. Vrable, S. Savage, and G. M. Voelker, Cumulus: Filesystem backup to the cloud, ACM Transactions on Storage (TOS), vol. 5, no. 4, 2009

[17] T. Wood, H. A. Lagar-Cavilla, K. K. Ramakrishnan, P. Shenoy, and J. Van der Merwe, PipeCloud: Using causality to overcome speed-of-light delays in cloudbased disaster recovery, in Proceedings of the 2nd ACM Symposium on Cloud Computing, 2011

[18] A. Bajpai, P. Rana, and S. Maitrey, Remote mirroring: A disaster recovery technique in cloud computing, International Journal of Advance Research in Science and Engineering, vol. 2, no. 8, 2013.

[19] J. Zhang and N. Zhang, cloud computing-based data storage and disaster recovery, in IEEE International Conference on Future Computer Science and Education (ICFCSE), 2011.

[20] Y. Tan, H. Jiang, D. Feng, L. Tian, and Z. Yan, CABdedupe: A causality-based deduplication performance booster for cloud backup services, in Parallel & Distributed Processing Symposium (IPDPS), 2011

[21] W. Li, Y. Yang, and D. Yuan, A novel cost-effective dynamic data replication strategy for reliability in cloud data centers, in IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing (DASC), 2011

[22] C. Liu, Y. Gu, L. Sun, B. Yan, and D. Wang, RADMAD:High reliability provision for large-scale deduplication archival storage systems, in Proceedings of the 23rd International Conference on Supercomputing, 2009.

[23] T. Nguyen, A. Cutway, and W. Shi, Differentiated replication strategy in data centers, in Proc. the IFIP International Conference on Network and Parallel Computing, Guangzhou, China, 2010, pp. 277-288.

[24] C. Cachin, R. Haas, and M. Vukolic, Dependable storage in the Intercloud, IBM Research, vol. 3783, 2010.

[25] Yu Gu, Dongsheng Wang , Chuanyi Liu, DR-Cloud: Multi-Cloud Based Disaster Recovery Service, 2014. TSINGHUA SCIENCE AND TECHNOLOGY I SSNl l1007-0214l l02/ 10ll pp13-23 Volume 19, Number 1, February 2014

[26] Zhe Wu, Butkiewicz M, Perkins D, Katz-Bassett E, Madhyastha HV. Spanstore: Cost-effective geo-replicated storage spanning multiple cloud services. In: Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles. SOSP '13. New York: ACM: 2013. p. 292–308

[27] Guoxin Liu and Haiying Shen , Minimum-cost Cloud Storage Service Across Multiple Cloud Providers, 2016 IEEE 36th International Conference on Distributed Computing Systems.

[28] Mudambi, S.M., Schuff, D., Zhang, Z.: Why aren't the stars aligned? An Analysis of Online Review Content and Star Ratings, In Proceedings of the 47th Hawaii International Conference on System Science (HICSS), IEEE(2014)

[29] A. Alamsyah, Dessy Monica Ginting   Analyzing Employee Voice Using Real-Time Feedback, 4th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia, 2018

[30] CHARLES FITZGERALD, FEBRUARY 12, 2019,
https://www.platformonomics.com/2019/02/follow-the-capex-cloud-table-stakes-2018-edition/