International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Analysis of the Technique for Disaster Recovery in Cloud Computing Environment

B. Bamleshwar Rao[1], Dr. Akhilesh A. Waoo[2]

[1]Research Scholar, Department of Computer Science, AKSU, Satna (M.P.), India

[2]Associate Professor, Department of Computer Science, AKSU, Satna (M.P.), India

Abstract: ***Disaster recovery is almost and always a persevering problem in IT platforms. This problem is more crucial and challenging in cloud computing, because at any cost Cloud Service Providers (CSPs) have to and should provide the services to their customers whether or not the knowledge Centre is down, due to a disaster. From the last few years, researchers have shown a profound interest and deepest inclination towards disaster recovery using cloud computing, and an elaborate, vast and ample amount of literature has been published and discussed related to Disaster Recovery. However, there's always been a shortage of relevant topics and points of the analysis of cloud-based disaster recovery. to unravel these queries, this paper provides a complete survey of disaster recovery concepts and research within the cloud computing environments. The main challenges proposed solutions.***

Keywords: Cloud Computing, Disaster Recovery, Replication, Backup, Survey

1. Introduction

Cloud Computing is mostly played on big computers every day because of its ability to share distributed resources around the world. Users can access cloud-based services over the Internet. Large IT companies are expanding their data centers across five continents to support different cloud services. The total market value of computing market capitalization is expected to reach approximately $ 241 million by the end of 2020 (Reid et al., 2011). Rapid advances in cloud computing are motivating many industries to use various cloud services (Arean, 2013), for example, close to 61% of UK businesses that rely on certain types of cloud services (White Paper, 2013).
However, many security challenges have been raised, such as risk management, trusted methods and restructuring to be considered to provide business continuity and better user satisfaction. Disasters, whether man-made or natural, can lead to disruption of expensive resources. Two types of disaster recovery (DR) can be used to prevent network failures or CSPs: Custom and cloud-based service types. The traditional model can be used as a dedicated infrastructure or shared system. Depending on the speed and cost, customers can choose the right model. In a dedicated way, infrastructure is provided to one customer, so both costs and speed are high. On the other hand, in a distributed model (we might also call it a distributed method) infrastructure is provided to many users. This method reduces the cost and speed of recovery. As shown in Figure 1, cloud computing is a way to realize the benefits of a dedicated and shared model. DR can work at low cost and high speeds.

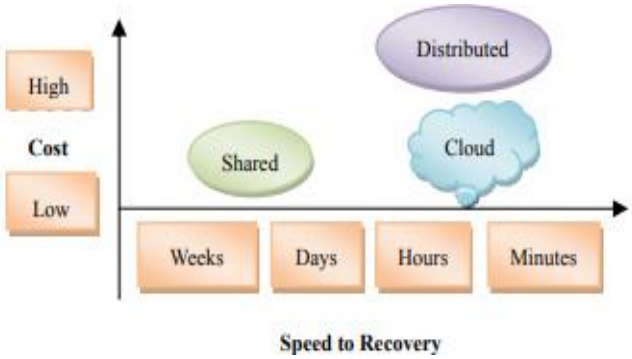

Figure 1: Comparison between traditional and cloud DR models (IBM White paper, 2012)

Table 1 shows the comparisons between these three categories of DR based on different factors. Cloud computing reduces data synchronization between a primary and backup site, reducing a variety of costs while increasing the independence between users' infrastructure and their DR systems.

Table 1: Disaster recovery models (Alhazmi and Malaiya, 2013)

DR Model	Data Synchronization	Independency	Initial Cost	Ongoing Cost	Cost of Potential Disaster
Dedicated	High	Low	High	Depends	High
Distributed	Medium	High	Medium	Depends	High
Cloud	Low	High	Low	Depends	Low

According to a study by IBM (IBM white paper, 2012), only 50% of IBM's disasters are caused by weather and some are due to other causes. For example, cutting down electrical cords, metal server failures, and security breaches. Therefore, DR is not only a process of natural events but also of all major disruptions to cloud systems.

Cloud service providers bring in use DR Services for many Organizations and in different kinds of businesses. And by Using these services, data protection and continuity of

service are guaranteed to customers at different levels. Table 2 shows the different DR services offered by IBM. Also, one critical issue in DR processes is how cloud providers can tolerate disaster to protect lost data and disruption of their service, infrastructure and services. In this paper we discuss both the challenges and solutions of DR machines at the discretion of the cloud provider. For businesses, the main purpose of DR is business continuity which means reclaiming services back online after the disruption. The Recovery Time Objective (RTO) and the Recovery Purpose (RPO) are the two main components that every return method tries to improve. By reducing the RTO and RPO business continuity can be achieved. RTO is the time between interruptions until the service is redeployed, and the RPO refers to the amount of data lost after a disaster. Failover delays include 5 steps depending on the backup level (Alhazmi and Malaiya, 2013):
S1: Hardware setup
S2: OS boot time
S3: Time to process the application
S4: Data recovery time/process
S5: Time to change IP
Therefore, RPO and RTO can be defined as:

$$RPO \propto \frac{1}{Fb}$$

When Fb is in the Backup folder

$$To = fraction\ of\ RPO + \sum_{j\ min}^{ss} Tj$$

Table 2: IBM different DR service level

IBM SmartCloud Recovery Service Level	Recovery Time	Discription
Gold	1 Minute	For mission-critical application
Silver	30 Minutes	For rapid recovery
Bronze	6 to 24 Hours	Assisted failover and failback

The rest of the paper is organized as follows: In part 2 of the computer, the cloud is presented briefly. In section 3 we discuss in detail the DR cloud. In section 4 and section 5 we investigate the major challenges in the DR programs and other proposed solutions, respectively. Following is section 6 which discusses some of the specific cloud-based applications for DR to be presented. In Section 7, open issues are investigated. Finally, the paper ends with the proposed procedure and conclusion of the DR.

2. Disaster Recovery

A disaster is an unexpected event in the life of the system. It can be caused by nature (such as tsunamis and earthquakes), hardware/software failures (e.g. the VM failure of Heroku hosted on Amazon EC2 in 2011) or even human (human error or burglary). It can lead to significant financial losses or even endanger human lives (Kashiwazaki., 2012). Thus, between 2% and 4% of the IT budget for large companies is spent on DR annually (Prakash et al., 2012). Solid-based DR solution is a growing trend due to its ability to withstand disasters and achieve reliability and availability. It can be very useful for small and medium enterprises (SMEs), as they do not have as many resources as large companies do. As shown in Table 4, Data Level, System Level, and Application Level are the three DR levels defined according to program requirements.

Table 4: DR levels

DR Level	Description
Data Level	Security of Application Data
System Level	Reducing Recovering Time as short as possible
Application Level	Application Continuity

DR strategies should have five requirements for optimal performance (Wood et al., 2010):
- You should reduce RPO and RTO
- It has little effect on the performance of the standard system
- It should be geographically separated
- The request must be restored
- You must ensure privacy and privacy

3. Disaster Recovery Plan

To develop a recovery plan in the cloud system there are different DR approaches to go about it which are based on the nature of the system. They are based on the nature of the system. However, in the literature, all these approaches are based on redundancy and backup strategies. The redundancy strategy uses separated parallel sites that can start up the applications after a disaster; whereas backup strategy uses replication technology (Lwin and Thein, 2009). The speed and protection degree of these approaches depends on the level of DR service that is shown in Table 5(Guster and Lee, 2011). Also, three different types of replication technology are available: 1. Host and VM replication, 2. Database replication, 3. Storage replication.

Table 5: Cloud-based DR models

Model	Synchronize Time	Recovery Time	Backup Characteristics	Tolerance Support
Hot	Seconds	Minutes	Physical Mirroring	Very High
Modified Hot	Minutes	1 Hour	Virtual Mirroring	High
Warm	Hours	1-24 Hours	Limited Physical Mirroring	Moderate
Cold	Days	More then 24 Hours	Off site backup	Limited

There are various ways for DR to build a recovery plan to reduce RTO, RPO, cost, and latency by looking at system constraints such as CPU, network requirements and storage. So, we can say that the reorganization of the DR can be considered a problem of efficiency. According to (Nayak et al., 2010), DR strategies include two required phases:

1) Phase alignment: At this stage, all DR solutions should be aligned to the requirements of any data container (data container means data based on the same DR requirements)

2) Planning phase: Selecting the appropriate DR solution that can reduce costs about the required QoS for each data bar.

ENDEAVOR (Nayak et al., 2010) is an outline of the DR planning process. As shown in Figure 1, it consists of three modules:

1) Installation modules: Includes DR requirements (such as type of protection, RTO, RPO, and usage latency), Discovery engine (Access to primary and secondary site configuration information) and database (replication technology, commands, and formatting).
2) Planning modules: Includes Product Design (Analyzing DR Requirements and Comparing Models), Downloading (Classification of DR strategies based on specific features such as cost, risk, and latency (Azagury et al., 2002)) and Global optimization (choosing the right DR Plan (Jaiswal et al., 2011)).
3) Outputs: The ENDEAVOR release is the correct DR plan for each application with details such as target resources and devices, protocol configuration.

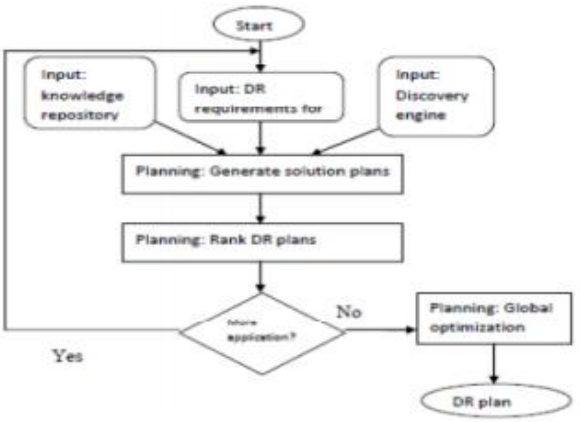

Figure 1: Endeavor flowchart

4. Disaster Recovery Challenges

In this section we investigate some of the common challenges of DR in cloud environments.

4.1 Dependency

One of the disadvantages of cloud services is that customers have no control over the system and their details. Data backup is on the premises of service providers. This issue makes reliance on CSPs for customers (such as organizations) and data loss due to a disaster being a customer concern. Reliance (Javaraiah, 2011) also creates another challenge which is the choice of a trusted service provider.

4.2 Cost

One of the important things to choose the cloud as a DR service is its low cost. Therefore, cloud service providers are always looking for the most cost-effective ways of providing returns by reducing various types of costs. The annual costs of DR programs can be divided into three categories (Alhazmi and Malaiya, 2012):

- Implementation of costs: annual charges
- Ongoing costs: maintenance costs, data transfer costs, and processing costs
- Cost of potential disaster: The cost of the disaster received and the cost of the non-recurring disaster.

4.3 Failure

The early detection failure time significantly affects the recovery time of the system, so it is important to detect and report rapid and accurate DR failure. On the other hand, in many backup sites there is a big question: How to separate network failures and service interruptions.

4.4 Security

As mentioned earlier, DR can be created by nature or can be man-made. The cyber-terrorist attack is one of the most man-made disasters that can occur for a variety of reasons. In this case, protecting and restoring important data will be the main focus of the DR programs other than system restoration.

4.5 Replication Latency

DR techniques depend on the replication process to create backups. Current replication strategies are divided into two categories: synchronous and asynchronous (Ji et al., 2003). However, they both have certain advantages and disadvantages. Synchronized replication, guarantees excellent RPO and RTO, but is dear and might affect system performance because of over-optimization. This problem is incredibly bad for multi-tier web applications because it can greatly increase the trip Time (RRR) between the first and backup sites. On the choice hand, the backup model adopted with async replication is cheaper and also the system has fewer problems, but the standard of the DR service is reduced. Thus, the transaction between costs, the performance of the system and also replication latency is an undeniable challenge in cloud disaster solutions.

4.6 Data Storage

Business database storage is one of the problems of enterprises that can be solved by cloud services. By increasing cloud usage in business and market, enterprises need to store huge amounts of data on cloud-based storage. Instead of conventional data storage devices, cloud storage service can save money and is also more flexible. The architecture of a cloud storage system includes four layers: physical storage, infrastructure management, application interface and access layer. To satisfy applications and also to guarantee the security of data, computing has to be distributed but storage has to be centralized. Therefore, storage a single point of failure and data loss are critical challenges to store data in cloud service providers (Pokharel et al., 2010).

5. Literature Review

Omar H. Alhazmi (2013) This paper discusses the trade-offs involved and provides guidelines for choosing between disaster recovery options. The optimal disaster risk planning should be considered with key parameters including initial costs, data transfer costs, and data storage costs. Organizational data also requires that disaster recovery purposes need to be considered. To assess risk, types of disasters (natural or man-made) need to be identified. The probability of a disaster occurring needs to be evaluated along with the cost of associated failure [1].

Mohammad Matar (2018) Cloud database services are used to reduce storage costs in the fields of information technology and provide other benefits such as online data acquisition. A single cloud is defined as a group of servers whether it is one data center or one hosted by one provider. However, moving from a single cloud to a multi-cloud makes sense and is important for many reasons. Dead cloud services are still subject to data-related exits. In the case of a disaster event, one cloud is subject to partial or total lost data [2].

Abdelfatah A Tamimi (2019) Electronic data created today by large numbers requiring the work of a data service organization can find a different kind of disaster whether it is natural or man-made, which can result in significant data loss. The purpose of the recovery technology is to recover information from a backup server when a large data server is lost in the event of a disaster. There are time and cost difficulties that make it difficult to implement those processes. When using disaster scenarios like a service, these disasters can be remedied and the speed of data recovery at a low cost [3].

Zhang Jian-hua and Zhang Nan (2011) To solve business services service problems such as storage capacity, performance, robustness, security, load, and many other issues, cloud storage has been used to provide a cloud-based data platform. Data services were placed in the cloud, and a robust operating system and platform were used to process the data. This document describes the construction of cloud storage and outlines the deployment of disaster preparedness and other applications that are kept secretly within the cloud, which can access real cloud computing [4].

A. Arul Mary (2011) In the business continuity of business in As many, organizations use the clouds when drawing their drawings and may experience an increase in nature or human-made results in the loss of facts. With the help of using disaster recovery strategies such as backing up information during a disaster. There are a few problems such as time being complex; a poor financial performance that makes the user difficult to manage failures. By using disaster recovery as a provider, the user can manage those failures and can recover facts quickly at a lower cost. Disaster healing is continuing suffering on IT platforms. This suffering is especially important for cloud computing because cloud-based companies (CSPs) need to provide their customers even though the facts are at the bottom because of the disaster. and recovery time from the DR cloud offering completely [5].

Long Wang (2016) Defines disaster risk and renewable business applications both at the cloud infrastructure level and at the application level. We investigate situations that favor one option over another, as well as situations in which your combination of both is required for effective and end-to-end protection. With the research done based on using IBM's Cloud Managed Services (CMS) risk management platform, we highlight the difficulties of protecting enterprise systems in the cloud. For reorganization and optimization, we introduce an algorithm that detects cloud-based machines taking into account performance and business dependencies.

6. Cloud computing

Cloud computing is becoming more common in day-to-day computing due to its ability to share globally distributed resources. Cloud computing may be a set of policies and procedures that are typically supported by a physical or technical infrastructure that allows the corporate to recover quickly from disaster and ensure business continuity if the system crashed or any kind of natural or human-made disaster occurred then there's a chance of knowledge loss and it should also cause the loss. Cloud computing processes are interconnected systems with shared resources many users share identical storage and other computing resources. Therefore, we'd like a strong mechanism to forestall other users from accessing your important and useful data. Cloud-based storage and recovery solutions allow you to copy important business files and restore them if they're compromised. Because of its high flexibility, Disaster Recovery (DR) allows the organization to take care of or resume critical task functions quickly after a disaster. The goal with DR is to stay the corporate working as near normal as possible. Data is stored during a secure cloud environment designed to produce high availability. The service is out there on-demand, enabling organizations of various sizes to style DR solutions in keeping with their needs.

7. Causes of Data Loss

In this section, there are various types of data losses which have been described in the below-following points-

7.1 Natural Disasters

Natural disasters are the foremost uncontrollable cause. Such as, fires, floods, earthquakes, even brownouts, all of them are out of our control. Fortunately, per the survey, the only book of users lost data due to natural disasters.

7.2 Mission-critical - application failure Sudden damage to the applying may occur when left unused for days, leading to loss of knowledge which will be important in some organizations.

7.3 Network failure- When the network crashes, cloud-related systems are disrupted, and cloud-based data and applications are lost because the cloud and clients are connected via the web. If the network fails also will suffer IP-based telephony and telecommunications.

7.4 Network intrusion- When viruses are invaded by applications, a disaster is formed. To avoid a disaster, anti-virus applications are used and programs are placed on the disaster monitoring list.

7.5 Hacking or malicious code- You know that computer viruses can block you and steal Mastercard information. Computer viruses or other malware can also spread like wildfire causing partial or complete damage to your important data. Therefore, it's essential that you just should install good antivirus software and keep it updated.

7.6 System failure: Operating systems are affected if the organization's infrastructure fails, leading to the failure of the organization-wide systems.

7.7 Human Errors: Belief or not, human error is additionally one of the foremost common causes of information loss. Normally, the most reason for the occurrence of a disaster is human, 60% of the info centers are failed. There are two forms of human errors causing data loss, one is clicking Delete or Format button to erase something we do not mean to, and another one is causing physical damages due to dropping or failing our device inadvertently.

8. Disaster Recovery Techniques

In our literature search, we found many ways that have their unique ways to build backup and recovery. Broadly speaking, all those skills focus on three distinct aspects, such as cost control, duplication of data and security issues. Each of these processes is entirely focused on the purpose of support and recovery.

a) Parity Cloud Service

Privacy protection is an important issue for providing personal data access services, the basic data recovery service is not sufficient for public performance. Users are not expected to upload their sensitive data to an online backup server until they have completely relied on the service provider in terms of privacy protection. a framework for the protection of the privacy of personal information has been developed, Parity Cloud Service (PCS). There are four considerations for designing a personal data recovery service. 1-Reliability .2- Economic efficiency.3-Simplicity. 4th privacy protection. PCS is very simple, can completely relieve users of their concerns about privacy protection, easy to use, requires moderate server costs, and can restore user data at a high enough level. Figure 2: shows the conceptual architecture for PCS. [1, 9, 14]

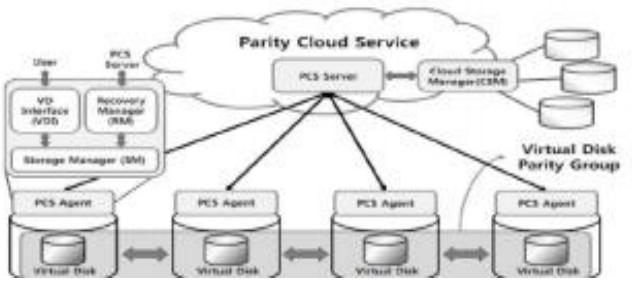

Figure 2: PCS architecture

b) Seed Block Algorithm (SBA):

Is the algorithm used in the proposed system to ensure the secure backing of data to the cloud and the remote server, this method is based on the concept of a computer-based exclusive OR (XOR) process? It consists of three main components 1. Cloud Server Master 2. Cloud clients and 3. Remote Server. The way for this task is to connect each unique client with a new customer and when a new customer registers, a customer ID receives a unique XOR number for recovering lost data in the event of a disaster in a large cloud using XORing data from a particular Seed database for that customer. The advantage of this technology is that it can retrieve data files with high accuracy and maintain its integrity. However, it does not work because the storage space crashes due to the same storage space used in the cloud and remote server [2,7,9] as shown in Fig. 3 clients can access these files from a remote repository if data is not available in a central location.

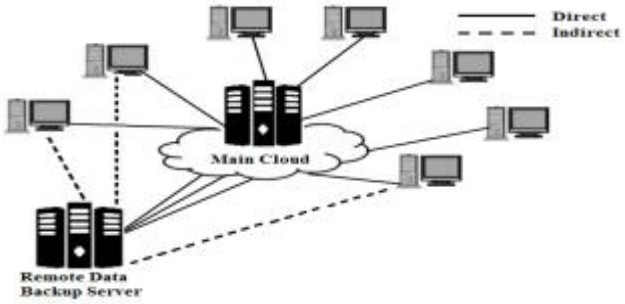

Figure 3: Seed Block Algorithm (SBA)

c) Multi-tier Web Application

Many web applications have web cache interface linked to the cache and are intended to analyze the cost of DR by calculating the cost of replication and reporting failures using the ROBIS standard on the web. We calculate the cost using a subscription service from RUBiS (an e-commerce web application that can be used using multiple Tomcat servers and MySQL data) with 300 clients. As shown in Figure 4, in the normal operation of a primary data center, the cloud can restore post-disaster recovery capabilities using two types of resources: backup mode backup sources for backup before the active disaster; failure mode that will only work after a disaster has occurred.

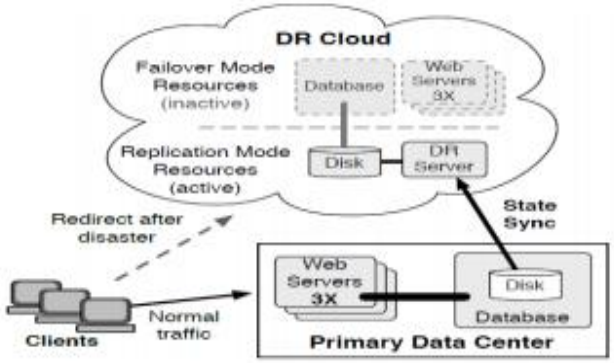

Figure 4: Multi-tier Web Application

d) Full Data

Data storage contains data used by apps and added to the repository from time to time. Reports are created based on an incoming and current data set. The cost of an asset is measured by the size of the data.

e) Carrier Cloud Brokerage

Multi-Cloud Broker Orchestration and Cloud Carriers Architecture can play a growing role in bringing Backup as a Service. To ensure that consumers can not only access their cloud and future guidance in the field of cloud support systems. Finally, it is proposed a DR procedure that can be successfully applied by any DR method.

Figure 5: Multi-Cloud Solution

f) High -Security Distribution and Rake Technology (HS-DRT)

The HS-DRT is a backup concept that uses a very widely distributed data transfer mechanism and high-speed encryption technology. Its advantages include: 1) does not require the use of expensive leased lines. 2) Spatial scrambling and random distribution techniques are used to encrypt important data files. 3) With the increase in the number of users, this technology is safe, and the encryption speed is high. This model can be used for transferees such as laptops and smartphones. This technology cannot be considered ideal for backup and recovery in cloud computing because of the increased cost of data recovery and increased redundancy. When the number of duplicate copies of file data increases from the corresponding processor, the performance will be reduced.

9. Conclusion

In this paper, we have provided an in-depth analysis of the state of the art for DR in cloud computing. First, Then, we discussed the details of cloud-based disaster recovery and compared it with traditional approaches. Also, we derived the main challenges of DR mechanisms and proposed solutions to overcome them. Furthermore, the main DR platforms are discussed, followed by open issues and future direction in the field of cloud-based DR mechanisms. Finally, a DR procedure is proposed which can effectively be utilized by any DR mechanism.

References

[1] Omar H. Alhazmi (2013) Evaluating Disaster Recovery Plans Using the Cloud 978-1-4673-4711-2/13/$31.00 ©2013 IEEE.
[2] Mohammad Matar (2018) Disaster Recovery and Business Continuity for Database Services in Multi-Cloud 978-1-5386-4427-0/18/$31.00 ©2018 IEEE.
[3] Abdelfatah A Tamimi (2019) Disaster Recovery Techniques in Cloud Computing 978-1-5386-7942-5/19/$31.00 ©2019 IEEE.
[4] Zhang Jian-hua and Zhang Nan (2011) Cloud Computing-based Data Storage and Disaster Recovery 978-0-7695-4533-2/11 $26.00 © 2011 IEEE
[5] A. Arul Mary (2017) STUDY ON DISASTER RECOVERY IN CLOUD ENVIRONMENT 978-1-5090-5573-9/16 $31.00 © 2016 IEEE.
[6] Long Wang (2016) Disaster Recovery for Cloud-Hosted Enterprise Applications2159-6190/16 $31.00 © 2016 IEEE.
[7] C.-W. Song, S. Park, D.-W. Kim, and S. Kang, "Parity Cloud Service: A Privacy-Protected Personal Data Recovery Service," 2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, 2011.
[8] N. Vedashree, P. Kumar, G. Anilkumar," Data Recovery in Cloud Environment Using Seed Block Algorithm", (IJCSIT) International Journal of Computer Science and Information Technologies, 2015.
[9] S. Shahzadi, G. Ubakanmay, M. Iqbalz, T, Dagiuklasx "Autonomous, Seamless and Resilience Carrier Cloud Brokerage Solution for Business Contingencies during Disaster Recovery". Long Wang, Harigovind V. Ramasamy, Richard E. Harper, Mahesh Viswanathan, and Edmond Plattier. "Experiences with Building
[10] Disaster Recovery for Enterprise-Class Clouds." In Dependable Systems and Networks (DSN), 2015 45th Annual IEEE/IFIP International Conference on, pp. 231-238. IEEE, 2015.
[11] F. Färber, N. May, W. Lehner, P. Große, I. Müller, H. Rauhe, and J. Dees, "The SAP HANA Database--An Architecture Overview," in IEEE Data Eng. Bull. vol. 35, no. 1, pp. 28-33, 2012.
[12] L. Hossain, J. D. Patrick, and M. A. Rashid, Enterprise Resource Planning: Global Opportunities and Challenges. Hershey Park, PA: Idea Group Publishing. 2015.
[13] Masahiko Jinno, etc, "Spectrum-Efficient and Scalable Elastic Optical Path Network: Architecture, Benefits, and Enabling Technologies," IEEE Communications Magazine, Volume: 47, Page(s): 66-73, November 2009.

Author Profile

Mr. B. Bamleshwar Rao is Assistant Professor and Head, in Department of Computer science in College of Computers and Communication Jabalpur, having 13 years of academic experience. His qualification M.Sc(CS), MCA. He is awarded by the Best Faculty Award.

Dr. Akhilesh A.Waoo is Associate Professor and Head, in Department of Computer science in AKS University, having 20 years of academic and research experience. His qualification includes Doctorate, UGC-NET, M. Tech. (CSE)along with IIT-Bombay and SWAYAM coordinator. Academic Experience is flourished with organization and coordination of national and international events/workshops/seminars. He had published around70 research papers in international journals. He is awarded by the Best Faculty Award. He had published a book on C# platform. He is a member of Easy chair. His research contribution includes supervision of 5 Ph.D. students along with more than 100 dissertations at UG and PG level of students. Also, he is recognized as a reviewer in many international journals.