[[1]](#footnote-1)

1. *Spoorthi, K. Priya, B.C. Sai Vyshnavi*

*Department of Computer Science and Engineering,*

*Panimalar Engineering College, Chennai, Tamil Nadu, India*

Disaster recovery mechanism with active storage replication and DNS failover in cloud virtual machine

K. SPOORTHI, K. PRIYA, B.C. SAI VYSHNAVI

***Abstract***--**The Domain Name System (DNS) is a vital service in the Internet. Much more than a simple translation mechanism, it also allows higher profile functionalities such as load balancing and enhanced content distribution. In the scope of cloud computing, DNS is foreseen as an elastic and robust service, supports failover mechanisms, decentralized configuration and multi-tenant isolation. The problem arises where a particular DNS is not able to point an IP address of the server because of heavy traffic, DDoS attack, system crash and much more reasons. During this period there are huge chances for an organization to be a victim of huge, valuable data loss and have a downtime for a certain period. This will cause a billion dollar business impact and risk of losing data. To overcome this, we come up with an approach of DNS Failover with active Replication, which is a solution designed to help keep the services online and prevent from Data losses. When systems and services go down, DNS failover is used to direct the users to another resource with little to no disruption which has active replication of Data with the current production Server. The Replicating server is named as Contingency server (CNR) which has a active storage replication with the Production server preventing Data loss.**

*Index Terms*-- internet server, cloud computing, DNS(domain name server), Ddos attack, data loss, failover, Contingency server (CNR), storage replication, Production server

# Introduction

Back when mainframe computers were the norm, disaster recovery strategy relied heavily on tapes. Later, WAN/LAN technologies came into use and took the advantage of replication. As the technologies are expensive to maintain and time-consuming, the next generation of Disaster Recovery technologies has made significant progress. This backup technology has made it reliable and efficient. Additionally, data virtualization and cloud features acquired a modern backup and recovery alternative.

Disaster Recovery, on the other hand, is a strategy and procedures for rapidly restoring the access to data, Information Technology resources following an outage. This entails switching to a backup set of servers and storage systems until the primary data centre can be restored. In general, Disaster Recovery involves provisioning and setting up the main server

and backup server, synchronization in which the data is replicated actively from primary site to backup site, and failover which refers to recovery of main server at the backup server.

In this work, we describe Disaster Recovery mechanism in cloud virtual machine, which includes active storage replication and DNS failover. Domain Name System (DNS) is essential source of the internet. DNS is an automatic process that translates the requests for domain names into the respective IP addresses, controlling which server an end user will reach when they type a domain name into their web browser. Due to an increase in traffic, such as the number of users sending requests to the specific DNS, the server may crash, making it difficult to connect to the requests server and possibly resulting in data loss. In order to overcome, the process of failover and the replication is done by using the SDR algorithm.

To achieve redundancy in the event of failure or termination of a formerly active system; a standby database, system, server, or other network must always be ready to automatically switch into action as failover is critical to Disaster Recovery. This includes the failover automation, as the domain name is connected to the two or more servers in a network with the live server always up and running. As long as the main server is active, the secondary server merely rests. However, if the secondary server perceive any changes from the primary failover server, some systems are automated with manual approval configuration. This simply alerts the technician or the data center instead, requesting the change to the server takes place manually. Instead of performing manually, Virtualization simulates a computer environment using a Virtual Machine or a Pseudo Machine running the host software. The failover process, thus can be independent of the physical hardware components of computer server systems.

# literature survey

Disaster recovery mechanism in cloud has currently received a lot of attention. A handful of research has been done on data replication algorithm and cloud environment. In [1] the author proposed an algorithm, namely RSPC, that introduced a replication strategy that fulfills the inhabitant targets and the supplier benefit. It demonstrates a popularity based file selection strategy through analyzing data access by users in a determined period of time.

1. Mansouri, et al [2] addressed the approach for secure data replication. They proposed the Secure Data Replication (SDR) algorithm, which identifies popular files as replication candidates based on the number of file-access times, where the higher the number of file-access times, the higher the popularity. Following that, only 20% of the candidates will be chosen for replication. Moreover, to improve the performance in cloud, they also proposed [3] a dynamic replication algorithm namely Hierarchical Data Replication Strategy, that identifies the popularity of a file based on a predicted number of file-access times then it replicates that popular file into the best site using network level locality.

In paper [4], Yanling Shao, et al presented a replication management system that is cost effective and reliable. It is implemented utilising the Dynamic Replica Creation Algorithm (DRCA) and Data Replica Scheduling Algorithm (DRSA), which are based on a meta-heuristic method in which files are selected depending on their popularity. This improves system's performance and allows for the search of a higher-quality replica placement solution while also lowering total data access costs while meeting the deadline.

Xiong Fu, et al [5] used Replica Placement Based on Load Balance strategy (RPBLB) to reduce remote users access time. RPBLB recommended replicating the most frequently used data to new storage while keeping storage load balancing in mind. This also ensures the load balance between data centers.

A.U. Rehman, et al [6] in his work, addressed the Software Defined Networking (SDN) fault-tolerance and discussed the OpenFlow fault-tolerance support for failure recovery. They proposed the make function such as firewall and Domain Name Service (DNS) to their customers. They further highlighted the SDN-specific fault-tolerance issues and provide a comprehensive overview of the state-of-the-art SDN fault tolerance research efforts.

S. Zheng, et al [7] presented Ziggurat, a multi-tiered file system that combines main memory which is non-volatile and slow disks to create a storage system with near-NVMM (Non Volatile Main Memory) performance and large capacity. Depending on application access patterns, write size, and the likelihood that the application will stall until the write completes, Ziggurat routes incoming writes to NVMM, DRAM, or disk. They proposed Emerging fast, byte-addressable Non-Volatile Main Memory (NVMM) that provides huge increases in storage performance compared to traditional disks.

V. Bindhu [8] the author in her paper proposed zero-downtime deployment, would ideally not cause any outage to the end users. The old version continues to run till the new version is ready. Perhaps the most strong innovation in improving the transmission capacity usage of the next generation network is cognitive radio network (CR-N). Anyway, the conventional CR-N is significantly compelled in getting to and the range detecting, because of its restricted, handling power and the stockpiling capacities.

Ernesto Garbarino [9] provided a guide to automating application deployment, scaling and management. This helps to learn and set up Kubernetes and achieve zero-downtime deployments using the service controllers. The combination of micro service architecture, containerization, orchestration and public cloud compute resources have significantly reduced the time-lines. However, the modern businesses are extremely dynamic and fluid in nature. They demand changes in the web applications continuously and rapidly.

Jayashree Mohan, et al [10] presented CheckFreq, which is an automatic, fine grained checkpointing framework that algorithmically determines check pointing frequency at the granularity of iterations using online monitoring. Moreover, it carefully pipelines check-pointing with computation to reduce the checkpoint cost by introducing two-phase check-pointing. They investigated various check pointing approaches such as asynchronous check pointing in High Performance Computing, and among those approaches, synchronous check pointing occurred large checkpoint stalls.

# proposed methodology

In this work, we come up with an approach of DNS failover with active Replication which is a solution designed to help keep the services online and prevent from Data losses. When systems and services go down DNS failover is used to direct the users to another resource with little to no disruption which has active replication of Data with the current production Server. The Replicating server is named as Contingency server which has active storage replication with the Production server preventing data loss. To direct the traffic to another resource with less traffic, we use concept of Forward Domain Name Service (FDNS) controlled with Apache KNOX to perform Automated DNS failover to less traffic backup server. To avoid risk of losing production valuable data we use virtual shared disk storage attached to cloud VM having Network Attached Storage (NAS) which uses SCP and Elastic File system (EFS) to transfer Data back and front to the Cloud storage disk. We believe using the following method servers can be prevented from DDoS attack and can have auto recovery mechanism which eliminates service downtime. For more security a SSH key handshake mechanism will be implemented in the server to prevent Trespassers to override the Production Data.

Furthermore, we use Secure dynamic replication (SDR) algorithm for securely replicating the file systems. SDR algorithm enhances the security level of data without any encryption technique since each data center has only one part of the file, and as a result, even in a successful attack, useful information is not discovered. SDR is a heuristic and a security-based replication method that has three main steps as follows: i) Determination of popular file: The most needed files are selected for replication; ii) Replica placement and optimize fuzzy rules with Competitive Swarm Optimizer (CSO): The popular file is divided into some fragments, and then, the suitable locations based on the T-coloring concept and CSO technique for storing fragments are determined; and iii) Replica replacement Due to the limited storage space, less valuable replicas are determined for deletion.

# system architecture

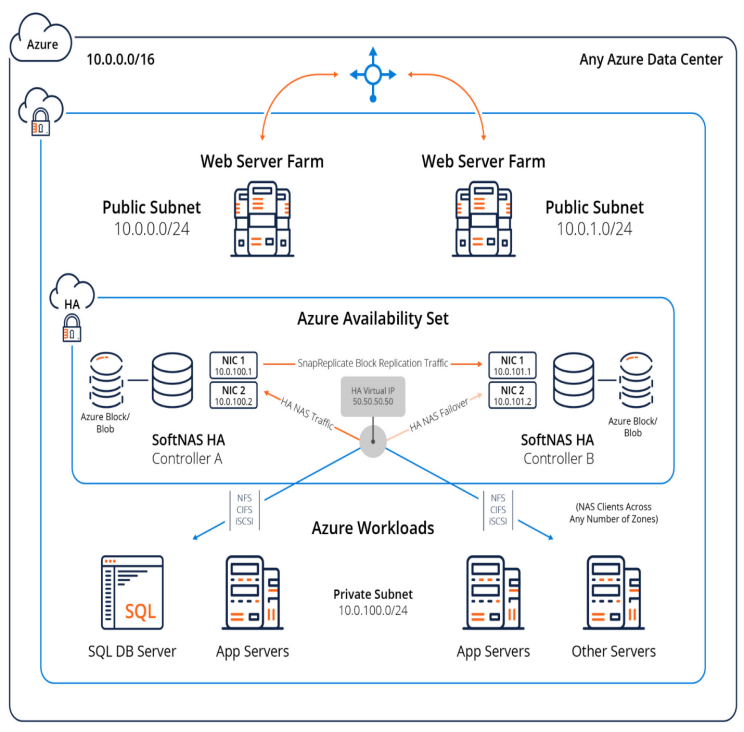


Fig 1. Architecture diagram

The figure illustrates the overview of disaster recovery mechanism with active storage replication. SoftNAS HA integrates seamlessly into today's virtualized data centers. In today's data center, VMware is frequently used, with a network architecture consisting of several VLANs used to separate various types of network traffic. This architecture implements the best practice for SoftNAS HA. The DNS IP address in this model is 10.0.0/16, which is assigned to the Azure Data Centre. A switch is used to connect two web server farms of different servers with unique IP addresses. One server serves as the primary server, while the other, with IP address 10.0.1.0/24, serves as the backup server. Public user subnets are assigned outside the data centre network, and they must pass through one or more routers to get to the data centre. If server A crashes, the DNS name is internally redirected to the backup server.

Azure Availability Set acts as storage unit with High Availability HA lock system, which acts as hardware architecture and are not interdependent. Here, SoftNAS block replication traffic is routed through a dedicated Replication VLAN and subnet, with SnapReplicate configured to flow across it. This prevents data replication traffic between controllers from interfering with storage or other data centre services. This is done using public Virtual IP address, which acts as an interface validating the data. Furthermore, all SNAP HA failover and takeover operations are controlled and witnessed by the HA Controller and also keeps a record of which storage controller is acting as the primary controller. Moreover, workloads are the programmes that run on the server. To virtualize the storage traffic, separate storage VLAN and subnets are assigned. Storage access is handled through NFS, CIFS and iSCSI file system between VMware vmKernels on each VM host.

# results & discussion

An elastic and scalable architecture for DNS as a Service, suitable for cloud-based platforms, was presented and validated. The provided assessment, obtained through experimentation, demonstrated the employed scaling practices for cloud-based services. The obtained results further proved the proposed architecture’s elasticity and flexibility, which takes into account the specificity of the DNS service while not being tied to a particular cloud-computing platform. Regarding the performance evaluation of the presented architecture, it became clear that the service is able to accommodate variable loads of DNS queries per second, always keeping satisfactory levels of performance in terms of DNS queries throughput and low latency DNS answers. This performance level was maintained by DNSaaS resorting to a horizontal scaling approach, instantiating additional resources whenever required. The DNS Service Cluster has a greater number of SQRs and reduces failover time by 94% compared to the Real Servers.

# conclusion

As we know that the DNS service is so vital.If the network traffic increases as the multiple users are connected to the srever the server may crash resulting in the failover delay and the data loss.This might happen to the banking servers , business servers etc.. and this will lead to a huge economic loss to the company or the service providers In this work ,we presented the load balancing and the data replication process this results in the very few to no time delay and due to replication the data is replicated in both the servers for regular intervals ,so there will be no data loss.

# References

1. S. N. John and T. T. Mirnalinee, “**A novel dynamic data replication strategy to improve access efficiency of cloud storage**,” Information Systems and e-Business Management , vol. 18, pp. 405-426, July. 2019.
2. N. Mansouri, M. M. Javidi and B. M. H. Zade, “**A CSO-based approach for secure data replication in cloud computing environment**,” The Journal of Supercomputing, vol 77, pp. 5882-5933, 2021.
3. N. Mansouri, M. M. Javidi and B. M. H. Zade, “**Hierarchical data replication strategy to improve performance in cloud computing**,” Frontiers of Computer Science, vol 15, Article 152501, 2021.
4. Yanling Shao, Chunlin Li, Zhao Fu, Jia Leyue and Luo Youlong, “**Cost-effective replication management and scheduling in edge computing,”** Journal of Network and Computer Applications, vol 129, pp. 46-61, March. 2019.
5. Xiong Fu, Jian Li, Wenjie Liu, Song Deng and Junchang Wang, **“Data replica placement policy based on load balance in cloud storage system,”** IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2019).
6. A. U. Rehman, R. L. Aguiar and J. P. Barraca, **“ Fault-Tolerance in the Scope of Software-Defined Networking (SDN),”** IEEE Access, Vol 7., to be published.
7. S. Zheng, M. Hoseinzadeh, and S. Swanson, **“Ziggurat: A tiered file system for non-volatile main memories and disks,”** in *Proc. FAST’19,* Boston, MA, USA, February 25–28, 2019.
8. V. Bindhu, **“Constraints Mitigation in Cognitive Radio Networks Using Cloud Computing,”** Journal of trends in Computer Science and Smart technology, vol 2,pp. 1-14, 2020.
9. Ernesto Garbarino, **“Beginning Kubernetes on the Google Cloud Platform: A Guide to Automating Application Deployment, Scaling, and Management,”** pp. 129-154, 2019.
10. Jayashree Mohan, Amar Phanishayee, Vijay Chidambaram, **“CheckFreq: Frequent, Fine-Grained DNN Checkpointing,”** in *Proc.* 19th USENIX Conference on File and Storage Technologies, February 23–25, 2021.
11. L. Y. Chen, W.M. Li, and Z. M. Lei**, “Alleviating the Impact of DNS DDoS Attacks,”** Networks Security, Wireless Communications and Trusted Computing, International Conference, Apr. 2010, pp. 240- 243.
12. A. J. Kalafut, C. A. Cole, C. Lei, M. Gupta, and N. E. Myers, **“An empirical study of orphan DNS servers in the internet,”** In IMC '10: Proceedings of the 10th annual conference on Internet measurement, Melbourne, Nov. 2010, pp. 308-314.
13. P. Vixie, **“DNS Complexity,”** ACM Queue vol. 5, no. 3,Apr. 2007.
14. D. Wessels, **"A Recent DNS Survey,"** DNS-OARC,Nov. 2007.
15. N/A, **“Comparing DNS resolvers in the wild,”** In IMC '10: Proceedings of the 10th annual conference on Internet measurement, Melbourne, Nov. 2010, pp. 15-21.

1. [↑](#footnote-ref-1)