Framework Architecture on High Data Availability Server Virtualization for Disaster Recovery

Murizah Kassim[1], Maznifah Mohd Sahalan[2], Nur Izura Uzir[3]
[1]*Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 UiTM Shah Alam. Selangor, Malaysia.*
[2]*Department of Infostructure, Universiti Teknologi MARA, 40450 UiTM Shah Alam. Selangor, Malaysia*
[3]*Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.*
maznifah@salam.uitm.edu.my

***Abstract*—This paper presents a design and tested framework architecture on High Data Availability Server Virtualization for Disaster Recovery at one campus network centre. Today's data information and knowledge are becoming most valuable commodity in business exchange and transactions where data need to be secured from cyber attackers and information security is a crucial needed. A propose architecture using server virtualization to provide high availability of data, through fast and high data through fast and high data recovery on virtual infrastructure for disaster recovery is done. The architecture uses multi side network RAID to achieve return of time objectives (RTO) and return of point objectives (RPO) of the application in the organization. Method presents servers consolidation and multiple physical server applications are deployed onto the virtual machines (VM), which then run on a single (many real server apps usually are started on top of this exclusive products (VM), which then function using one) or fewer real high-end servers to achieve better performances compared to utilizing several or even hundreds of traditional servers. Security perimeters are used in the proposed architecture to maximize the data protection in the organization. The setup experiments of virtualization technologies using VMWare, Ranger Pro for backup and Trend Micro Deep Security tools. Three architecture framework are tested which presents automated data replication simulations from production site to disaster recovery site that creates an active-active environment. Results presented that Recovery Point Objective (RPO), Recovery Time Objective (RTO), data loss and data availability at 99.91 % of data are recovered during recovery process using multi side network RAID. Thus this technique protects a larger share of disaster recovery workloads in terms of high availability and data protection.**

***Index Terms*—Disaster Recovery; High Data Availability; Network Framework Architecture; Security; Virtual Machine (Wms); Virtualization.**

I. INTRODUCTION

Data protection systems include high availability (HA), backup, disaster recovery (DR), data archive and security systems. Each of these systems is attempted to give a particular business unit access to the data that needs within a time frame acceptable for that particular business unit [1, 2]. Failure systems face downtime that includes catastrophic, hardware failures, software failures, errors by human interruption, and cyber-attacks. Given these risks, most organizations protect their information by using strategies such as remote mirroring [3, 4], including snapshots in addition to recurrent backups to help recorded argument or maybe disk [5], system test [6] or backup to clouds. These types of strategies possess various qualities, benefits as well as expenses. For instance, utilizing synchronous remote control reflecting enables programs to become rapidly unsuccessful more than as well as started again in the remote control area. Snapshots internal to a disk array tend to be space-efficient and invite quick recuperation of the constant current edition from the information. Backups to the tape or even disk permit a mature edition from the information to become recovered. These types of strategies possess restrictions. Remote mirroring reflecting generally offers higher source needs, local snapshots do not protect against failure of the disk array, as well as coping with backups can lead to substantial lack of current improvements. [7]. Ensuring data availability in back up and restoring processes and securing data from attackers are the focus of this research. The restoration processes need to overcome the issue of unavailability of data to ensure successful data restore and recovery process. Mitigate the risk of losing data process involved server administrators on backup copies of data stored on various storage devices. Users sometimes have no mechanism to test the backup data and once the backup doesn't work it turns to disaster when the backup process is not able to be used. Thus every systems must be tested wither on customised software or system [8]. During the gathering information from reference, it is found that certain organisation mentioned and worries of data loss if anything happen to the application system [9].

This research presents a design framework on information availability using server virtualization on disaster recovery that was simulated and tested as an actual environment. This research comprised of 5 phases which involves understanding the information availability during disaster recovery. It also identified the requirements and implementation techniques. Based on the requirements gathered in the first phase, the proposed architecture is designed, and later the proposed and the existing architecture are implemented and simulated as actual server's environment in the third phase. Three types of architecture framework were tested and identified parameters and measured. Lastly, the architecture is evaluated where data are collected and analyzed on the comparison framework. Best results on best architecture are present with improved performance on High Data Availability Server Virtualization for Disaster Recovery.

II. LITERATURE REVIEW

Research presents that information security held few important key concepts identified from information security core principles that are confidentiality, integrity and availability [10]. Data availability is important to ensure no

disruption of data in a long run. One of the elements on DR plan is to ensure availability of data after any disaster happen to make sure no disruption in business operation. DR may be the procedure by which an organization may recuperate the information following any kind of catastrophe occasions occurred [11]. DR designs are utilized in some instances to give extra safety towards lack of information because of problems, not just within the computer program on their own however in the surrounding environment [12] Information security is described to be keeping information systems through unauthorized entry, use, disclosure, interruption, customization, or even damage. This means to secure the information and the systems from those who might look for in order to improper use and even more from unauthorized individual. Accessibility describes the ability to access the information whenever required. Information security involved integrity which provides information that can't be produced, transformed, or even erased without authorization. For instance, losing integrity may appear whenever a database system is not correctlyshut down before maintenance or the database server suddenly loses its electrical energy. There was a promising and positive impact from data protection for organizations in Malaysia and other countries. Without data, system cannot be operated and it will affect the whole operation cycle. Due to unavailability of data, organisation will suffer and may cause the business closure and lead to unemployed workers. Back up and restoration process include of data availability to complete the cycle. Incomplete data cycle will cause the data loss and system cannot be operated as the previous operation. Replication of data also does survive the information [13].

Data protections methods can be classified into three categories which are restore, recover, and overcome. Restore methods can restore the systems' functionality after an incident or event occurred but the organizations require a significant amount of time to do so. Recovery methods allow the organizations or business unit to continue functioning after such an event, but would require a minor disruption in service before doing so. Finally, methods that allow to completely overcoming an event are typically the most expensive, but this method would allow the organizations or business unit to continue functioning uninterrupted through any kind of event [14]. Typical strategies need exact same equipment standards as well as settings, regular administration, higher energy as well as air conditioning expenses. Typical strategies produces bottleneck as well as overall performance issue as well as consider additional time to accomplish. Application and information restoration techniques via image tools as well as tape backups tend to be complicated as well as sluggish. Typical back-up as well as restoration techniques include complicated process operating-system back-up agents, scheduling and performing backups, rebuilding information, testing and verification backups[11]. Data protection is really a crucial facet of just about all computing environments. Through the years, it's transformed along with the aim of keeping a good enterprise's data from device failure to include software program failing, human mistakes, website failure as well as fraud. The term data protection for an IT department is to ensure the data that the organization needs is available when it is needed and is not made available to those entities that should not be given access. Data protection system includes high availability (HA), backup, disaster recovery, and archive and security systems.

Data protection is very important for the commercial sectors where data recovery is crucial in case of disaster to minimize data losses. In fact, many companies like small and medium-sized businesses (SMB) nowadays still rely on the traditional backup technology like tape data storage backup for data protection. Nowadays, there were limited ways of architectures for backup to recover data. Most people think that, once the backup was successful, the data are always available, whenever disaster happens. Therefore, many organizations have turned to corresponding replication and high availability solutions to minimize downtime and ensure critical applications and important data is protected [15, 16]. Due to this, the increasing amount of data imposes challenges for traditional data protection during a disaster. Thus, the information technology (IT) department needs to ensure the availability of data with larger backup set and need longer recovery time. Some global demand has determine the important steps such as to restore the information after disaster occurs, manage more backup with less time and protect the data due to the attackers' activity. This challenges may pave ways for an alternative approaches for data protection. Thus, solutions of architectures of data protection using server virtualization technology in DR were proposed in this research.

III. METHODOLOGY

Research method consists of 5 phases of architecture framework designed as in Figure1.

A. Phase 1: Analyzing and Identifying Research Requirement

Phase 1 presents a comprehensive study on precious methods on server virtualization, disaster recovery, server virtualization and data protection. The problems are identified and analyzed which meet three research objectives. It is identified that Virtual Server Environment (VSE) component will be used in the implementation as a part of DR site. VM is a robust soft partitioning as well as virtualization technologies that gives operating systems isolation, distributed CPU (along with sub-CPU granularity), distributed I/O, and automatic, dynamic resource allocation that is built in. Just one server operating VMs can make several digital machines or even devices, using their personal individual "guest" operating-system situations along with various operating-system variations, applications and users.

The physical resources of the server are shared amongst any of the VMs it hosts, based on demand and entitlement. Each VM hosts its own applications in a fully isolated environment. The VSE for servers is an integrated virtualization solution for the server environment. The VSE technology enables the organization to increase server consumption by giving virtual servers which automatically grow and shrink based on organizations priorities. Via restricted integration along with higher accessibility, partitioning and utility pricing, the VSE allows the organization to preserve support levels in case of unpredicted down time as well as to cover extra capability with an as-needed foundation. By reducing the number of servers and through efficient utilization, the organization may significantly decrease equipment purchase as well as server administration expenses. VSE also should allow the organization to continuously analyze and optimize its Adaptive Infrastructure[17].

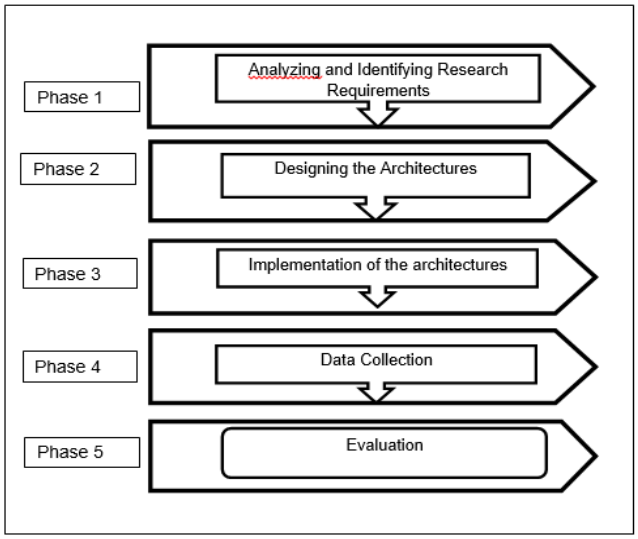

Figure 1: Research Methodology Framework

The organization should be able to address the request for additional server capacity, by using the VSE spare capacity that is available. The VSE functionality must provide intelligent, simplified control of the organizations virtual servers through a fully integrated software family for planning, management and automation. The partitioning, availability capabilities delivered by VSE are tightly integrated with its planning, management and automation capabilities to provide intelligent control for a broad range of virtualized resources adapted from HP website. All but not limited to the factors above will be taken into consideration during research simulation to devise the best solution design and migration and implementation plan, with minimal downtime to business applications.

B. Phase 2: Designing the Architectures

The second phase focuses on designing the simulation of existing and proposed architecture. During this phase, the components are identified and simulation of the implementations environment architectures is created. Each architecture uses the same specification parameters for virtual machine simulation. There are heterogeneous types of virtual machines with different operating system (OS) and system resources (CPU, Memory). This research has created four test virtual machines. Table 1 shows the virtual machines' specification for the simulation architecture.

Table 1
Virtual Machine Specification for the Simulation Architectures

No	CPU	OS	Memory	HDD	Application	Database
VM 1	1	Windows 2003 32 bit	4Gb	20GB	IIS Web Server 6	MsSOL Express 2005
VM 2	2	Windows 2003 32 bit	8Gb	20GB	IIS Web Server 6	MsSOL Express 2005
VM 3	1	Windows 2003 32 bit	4Gb	20GB	IIS Web Server 7	MsSOL Express 2008
VM 4	2	Windows 2003 32 bit	8Gb	20GB	IIS Web Server 7	MsSOL Express 2008

In the simulation, we have simulated a full backup for Architecture 1, Architectures 2 and the proposed Architecture 3. The details of the architectures are listed below:

i. Architecture 1: Hardware – Four Physical Machines (as in Table I), SAN Storage and tape library.
ii. Architecture 2: Hardware – Four Virtual Machine, SAN Storage and SAN Switch while software as stated in Table 3.1 using Continuous Access (CA) as the software for SAN to SAN replication.
iii. Proposed Architecture 2: Hardware - Four Virtual Machine, SAN Storage and SAN Switch while software as stated in Table 1 using VMWare Site Recovery, Trend Micro Deep Security and backup software of Range Pro.

C. Phase 3: Implementation of the Architectures

The thrid phase presents three simulation architectures that has been design in the proposed architectures at Production Site and DR Site. The steps involved in implemented approach such as preparation for the suitable platform that involves setting up the software and hardware component and installing respective software. Next is constructing the simulation architecture according to current IT environment.

a. Production Site

In the simulation of the Production Site, this research uses VMware vSphere 4 as the virtual infrastructure for all architecture. The virtual infrastructure consists of two virtualization hosts (or physical servers) with Storage Area Network (SAN). The SAN is really a storage equipment to servers so the equipment as locally connected to the operating systems. The SAN generally has its own connection of storage equipments which commonly are not obtainable with the normal connection through normal equipments. The SAN on it's own doesn't supply the "file" abstraction, just block-level procedures. Nevertheless, file systems techniques constructed along with SANs perform supply this particular abstraction, and therefore are referred to as SAN file systems or shared disk file systems [18]. Table 2 is the breakdown of the hardware specifications of the virtualization infrastructure at the simulation production site.

Table 2
Hardware specification for virtualization infrastructure

Machine	# of processors	Processor Type	Total Memory	NIC Ports	Internal Hard Disk	Connector SAN
ESX Server Host 1	2 x quad core processors	2.5 GHz	32GB	8	73GB x 2	FC port x 2
ESX Server Host 2	2 x quad core processors	2.5 GHz	32GB	8	73GB x 2	FC port x 2
Management Server	As Virtual Machine					
SAN Switch	2 x Gigabit Ethernet Switch					
SAN Storage	HP EVA 8400					
Tape Drive	HP Storage Works 1/8 Tape Autoloader (Architecture 1&3 only)					

Two servers are installed with VMware ESX4 part of VMware vSphere 4 virtualization suites. The two servers are called Host servers. The host server then hosts a guest machine called virtual machine on the host. All architecture use the above hardware specification except tape drives which is not applicable for Architecture 2. All the data are stored in SAN in the form of virtual machine disk (VMDK). VMware products has developed the virtual appliance using the Virtual Machine Disk (VMDK) file format that is open and documented. The SAN storage and the host servers are interconnected via Internet Small Computer System Interface (iSCSI) Network which is an Internet Protocol (IP)-based storage networking standard for connecting data storage services. Through having SCSI instructions more than IP systems, iSCSI can be used in order to help data exchanges

more than intranets and also to handle storage space more than lengthy miles. iSCSI may be used to transfer data over local area networks (LANs), wide area networks (WANs), or the Internet and can enable location-independent data storage and retrieval.

The actual process enables clients (known as initiators) in order to deliver SCSI instructions (CDBs) in order to SCSI storage devices (targets) upon remote servers. It is a popular SAN protocol, permitting organizations in order to combine storage in to data center storage space arrays whilst supplying hosts (such as database and web serversmachines) using the false impression associated with locally-attached hard disks. In contrast to conventional Fiber Funnel (FC), that demands special-purpose wiring, iSCSI could be go beyond lengthy miles utilizing current system national infrastructure [18]. Different from the Architecture 1 is that there must be link connectivity between both sites for Architectures 2 and the proposed Architecture. In this research methodology used existing bandwidth 1Gbps link in between to cater the needs of the storage replication for both architectures. In this virtualization host, three virtual machines are being created for testing and benchmarking purposes. Each of the virtual machines serves its own unique purposes and is used to differentiate the time taken.

b. Disaster Recovery Site

Setup for simulation of DR Site is identical as the one for simulating the Production Site. Identical hardware and setup are used to provide a seamless and bare-metal recovery for the virtual machine OS. Bare-metal restore is a method in the field of data retrieval and restoration where the copied data is accessible in a form that enables one to recover a computer system from " bare metal", i.e without the specifications concerning earlier installed software or operating system. Generally, the particular copied data are the essential operating system, applications and data elements to restore or recover the backed up system to a totally separate piece of hardware. In a few adjustments, the hardware obtaining the recover needs to have an identical setting for the hardware that was the source of the back up, despite the fact that virtualization strategies and also mindful organizing can easily permit any bare-metal restore to a hardware configuration diffrenet from the original.

D. Phase 4: Data Collection

During this data collection phase, there are three architecture simulation implemented and data are collected based on each architecture.

i. Architecture 1: Backup and Restore using Traditional Method;
ii. Architecture 2: SAN to SAN Replication; and
iii. Proposed Architecture : Multi side network RAID and backup method.

E. Phase 5: Evaluation

Phase 5 consists of evaluation phase calculating the performance metrics consists of Recovery Time Objective (RTO), Recovery Point Objective (RPO), data availability and data loss. There are a few steps to evaluate the simulation of the architectures. This step consists of Capacity Planning and P2V Migration.

a. Capacity Planning

Firstly, capacity planning for virtual infrastructure has been conducted. This research has to make sure capacity planning for virtual infrastructure is being done precisely and effectively. A successful Virtual Infrastructure deployment is preceded by a Basic Consolidation Estimate (BCE), where existing IT infrastructure as well as immediate future IT requirements will be assessed with the focus to better design a Virtual Infrastructure in the direction of specific IT and business goals. The actual VMware BCE is made to supply present environment's virtualization possible by providing an accurate assessment of Windows. The end result of the VMware BCE is really a clear path togetting started with a successful Virtual Infrastructure Deployment. This particular statement is actually produced, utilizing a VMware item known as Capacity Planner. VMware Capacity Planner is to provide quick and accurate virtualization analysis. The report produced is a plan depending on environment and goes beyond the results that an interview process can provide. VMware Capacity Planner Collector collects inventory and performance data that's delivered to theInformation Warehouse, analyzed by the Data Analyzer. The info gathered through simulation environment for analysis includes hardware and software inventory to provide capacity and system purpose, hardware resource utilization as well as software particular usage.

b. P2V Migration

P2V migration is an important step to migrate all application from physical environment to virtual without having to reinstall and reconfigure at DR site. P2V migration involves two main approaches which are hot cloning and local hot cloning.For migration strategy, cloning all physical systems is required at production site before performing P2V migration at DR site. Sufficient disk capacity is required to store physical server images before P2V process. The P2V migration tools used largely depends on the type of operating system being migrated. In this research, the main P2V tools used are VMware vCenter Converter for Windows servers. The value proposition of P2V and Virtual-to-virtual (V2V) is to protect multiple servers using a consolidated recovery environment and reduce RTO compared to imaging or tape backups. There are two types of cloning which consists of hot and local hot-cloning.

i. Hot-cloning: - also called live cloning or online cloning, entails cloning the source machine while it is running its operating system. Hot-cloning is generally preferred for source machines running applications with minimal data changes on disk, or where files or database services downtime is affordable.
ii. Local hot-cloning **:-** For local hot-cloning, the migration is performed with the application running on the source machine. It involves the installation of VMware vCenter Converter on the source server, in which system reboot is required for Windows 2000 and below. This architecture is preferred when the source machine sits in a separate subnet and remote hot-cloning is not feasible due to tight policies on network security. This research used live cloning to do the migration.

The performance metrics consists of Recovery Time Objective (RTO), Recovery Point Objective (RPO), data availability and data loss is calculated to get the result of this research.

Business continuity practitioners currently use two metrics to capture data dependability objectives [19].

i. Recovery time. The Recovery Time Objective (RTO) specifies the maximum allowable delay until application service is restored after a failure event. The RTO can range from seconds to days.
ii. Data loss. Recovery may require reverting to some consistent point prior to the failure, discarding updates issued after that recovery point. The Recovery Point Objective (RPO) gives the maximum allowable time window for which recent updates may be lost. The RPO can range from zero (no loss is tolerable) to days or weeks.

The formula used to get the data restoration for all architecture in hour is as Equation 1.

$$H = (TB+TR+TT)/60 \quad (1)$$

where H is hour, TT is Time to Back up In Seconds, TR is Time to Restore in Seconds and TT is Transportation Time.

IV. RESULT

A framework Architecture on High Data Availability Server Virtualization for DR is presented. Four servers are designed which each server consists of four virtual machines with architectures of 32 bit and 64 bit. The simulation process concentrates with only one physical server. Three architectures in the simulation are created which is the backup and restore using traditional architecture, SAN to SAN replication, and Combination of auto replication and backup method.

A. Architecture 1: Backup and Restore using Traditional Approach

Figure 2 presents architecture 1which is called cold site. The data is being transported as a backup from the production state in the production computer into the media called tape. Tape is then physically transported by hand to the DR site. The tape then be restored in the event of disaster recovery. This architecture is a proven architecture as it has been the main means of DR for many years [19]. However due to the fact that the backup could only be performed once a day the Recovery Point Objective (RPO) would also be as long as 24 hours up to Service Level Agreement (SLA) of the organizations. In addition to that the restoration time would be lengthy as tape has never been classified as something fast. Results present that Architecture 1 achieves BCDR by executing backup jobs to push data into tape media. The tape media is transported to the DR (DR) daily. Data is restored in the event when needed.

B. Architecture 2: SAN to SAN Replication

Figure 3 presents architecture 2 where storage is automatic replicated from Production to DR. However, this only ensure that data availability but not server availability. Manual step would needs to be taken to stop the replication, enable storage access from DR sites and manually powering up each of the VMs. This architecture is Active-Passive between SAN storage. This architecture achieves not only business continuity disaster recovery (BCDR) but creates an Active-Active environment by replicating data that is change by block level from the production site to the DR site by means of Storage Network Raid found in HP EVA 8400 . With this in place data are automatically transferred to the DR site without the need of having tape transportation and doing away with the restoration jobs previously mentioned in Architecture 1. With this in place, the data would then be garbage in garbage out. An additional layer of protection would then be required to restore the data from tape for specified different generation of data, for example 3 weeks ago.

C. Proposed Architecture: Multi side Network RAID and backup method

Figure 4 presents the multi side network RAID and backup method architecture. The enhancements made to this architecture from Architecture 1where two servers are equipped with 10 Gbps link and data is secured with Trend Micro Deep Security as a security tool to ensure the security of VM. The EVA SAN Storage's uniqueness is abled to perform a Network Mirror RAID that compromises different nodes at different location using 10Gbps Internet connectivity. It is identified that no downtime in the SAN storage when failover over via SAN to SAN replication as there is a period of time where by the SAN storage administrator will need to change the replication as well as turning the passive read only SAN to read-write SAN. All the ESX hosts are connected to centralized 1Gbps iSCSI SAN storage as this is a pre-requisite to exploit all key features of VMware. VMware acts to continuously monitors utilization across a resource pool and intelligently allocates available resources among virtual machines, no impact to end users, to provide additional power savings and makes DR faster, reliable and manageable, so that organizations can meet their recovery objectives.

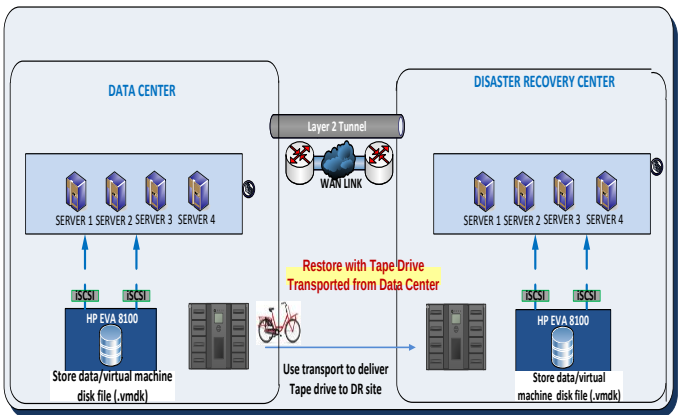

Figure 2: Architecture 1- Backup and Restore using Traditional Architecture

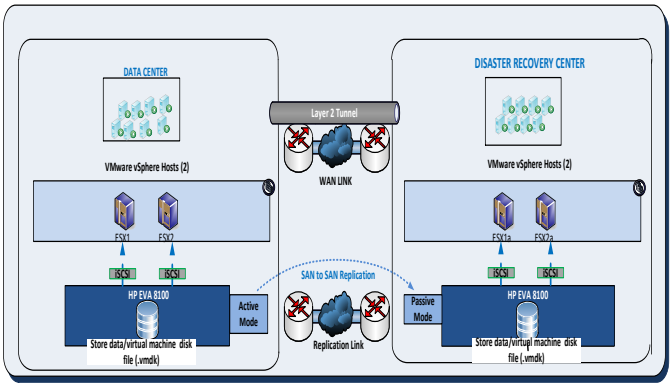

Figure 3: Architecture 2: SAN to SAN Replication

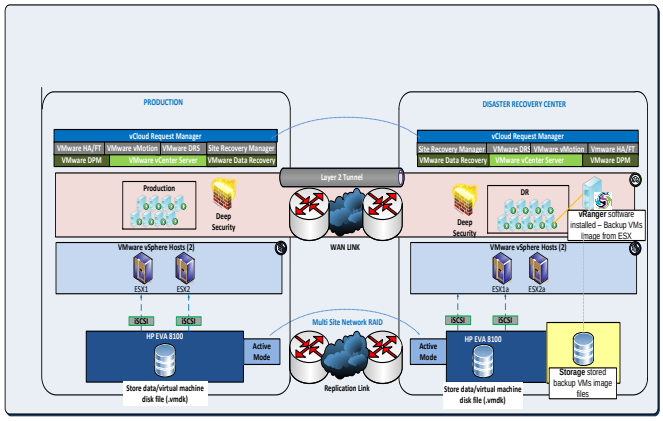

Figure 4: Proposed Architecture 3: Multi Side Network RAID and backup method

D. *Proposed Architecture Framework*

Final framework presents the proposed Architecture where servers consolidation architecture is planned using the capacity planner to get the accurate performance of the server's virtualization. The architecture is then designed accordingly based on the research requirements.

a. *Server Consolidation*

Server consolidations consist of consolidated of multiple operating system (OS) and applications stacks onto a single platform for disaster recovery. It is identfied as one of the most challenging consolidation strategies available to the enterprise IT managers. It can deliver substantial benefits where costs and risks can be high especially when it involves the expense and complexity of a migration different platform. Virtualization as a recovery platform protects a larger share of data site workloads without having to invest in costly duplicate hardware and software. Therefore, this virtualization research simulation mainly focuses on a virtual DR infrastructure in the DR simulation by providing a single site failure solution. Through a number of defined phases, one unit of new ESX servers are configured in the dummy DR site as Dummy DR Cluster to provide sets of the predefined VMs. This VMs are converted from physical servers, and to be replicated and managed by one unit of new ESX servers in the DR site as a DR Cluster. A total of four units of existing physical servers and VM in the organisation Production site which consists of Windows OS platform are virtualised as a set of VMs in the DR site based on the concept P2V (Physical to Virtual) and V2V (Virtual to Virtual) to the said virtual DR infrastructure. This will formulate the defined virtual DR infrastructure.

The results consists of capacity of data transfer for the backup and capacity of the data restore, the duration of the backup and restoration, and the percentage of data availability and data loss.

i. Figure 5 shows the high availability of data recovery after disaster using server virtualization technology. The result computed from the test has shown that the server 4 of Proposed Architecture achieved up to 99.91% data availability during the restoration process, as compared to server 4 of Architecture 1 achieved the 92.80% while server 4 of Architecture 2 yields the lowest result at 95.73%. Data loss for the server 4 of Proposed Architecture is only 0.09% while others are higher than that.

ii. Figure 6 shows the proposed an efficient architecture in time and speed in server virtualization technology. The shortest time taken for each backup process means the efficiency of the architecture. In this research the Proposed Architecture shows the RPO comprises of Server 1: 23 seconds, Server 2: 22 seconds, Server 3: 21 seconds and Server 4: 20 seconds. It shows the Proposed Architecture takes shorter time to backup data for server 1, 2, 3 and 4. In contrast, Architecture 2 shows few different time of backup process Server 1: 31 seconds, Server 2: 30 second, Server 3: 29 seconds and Server 4: 28 seconds. Hence, server 4 leads the good result at 28 second for 2691 byte data for 1 month simulation. While server 4 with Proposed Architecture given much better result at 20 seconds to complete the backup process.

iii. Proposed a secured virtualized environment architecture that can protect data from attacker which suits the security parameter for the virtualization technology for the simulation.

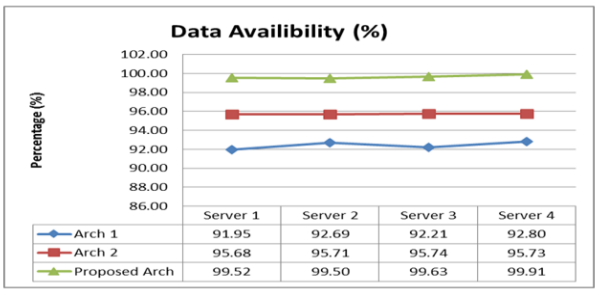

Figure 5 : High Availability of Data Recovery After Disaster Using Server Virtualization Technology

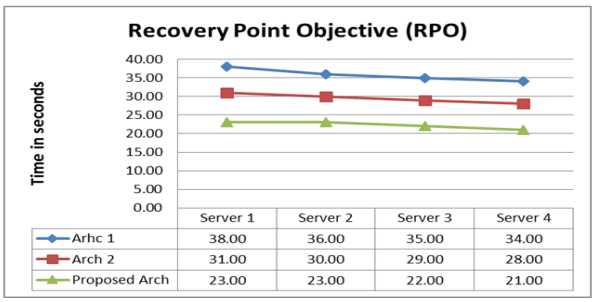

Figure 6: An Efficient Architecture in Time and Speed in Server Virtualization Technology

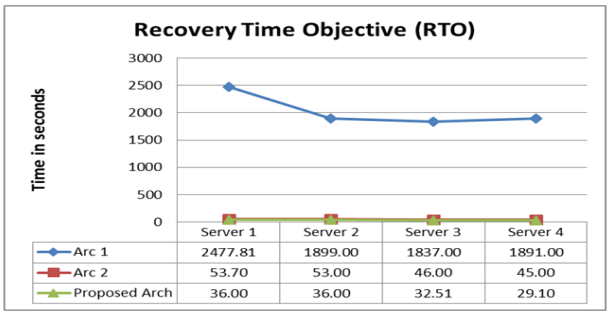

Figure 7: Comparison of Recovery Time Objective (RTO)

V. CONCLUSION

Research presents design framework architecture on High Data Availability Server Virtualization for DR where existing architecture relying on the tape backup which does not provide 100% protection. Architecture 1 presents Tape

backup approaches but find some limitations such as backup is prone to human error, Tape records are sometimes unreadable, especially on old tapes and data is only current to the last backup. Architecture 2 also shows limitation where approaches using SAN to SAN replication in active-passive mode on the data transfer and if interception happened, it does not real time in terms or data transfer. Thus, this research presents architecture framework with a management and recovery system of data losses, increased data protection and data availability of information systems after DR in time and speed in server virtualization technology in the organizations and secure virtualized environment architecture that can protect data from attackers using Trend Micro Deep Security as a tool. It is to achieve a secure virtualization environment and data protection. The Recovery Time Objective (RTO) specifies the maximum allowable delay until application service is restored after a failure event as in Figure 7. The RTO can range from seconds which is 2910 seconds for server 4 in Proposed Architecture. Recovery Point Objective (RPO) gives shortest time for back up process either incremental or full back up of 21 second for server 4 in Proposed Architecture as in Figure 6. It is identified that data availability is 99.91% for server 4 in Proposed Architecture and the loosing of data at 0.09% due to either human intervention or restoration process and services. The efficient architecture also presents using of 64bit arch, SAN to SAN replication with VMware Software and Trend Micro Deep Security. The concept involves automated replicating data from production to disaster recovery. The Proposed Architecture contributed good result in data availability after restoration process. Research significant where two objectives is presented which managed to contribute the suitable architecture for the organization. In future, cloud technology can be explored for data protection for DR good archiving and DR resource. Research towards the server performance of cloud technology, the impact of the effectiveness and efficiency of cloud technology when disaster happened can also be explored. Definitely for cloud technology, the cost depends on the organization's requirements.

ACKNOWLEDGEMENT

Authors would like to thank Universiti Teknologi MARA for support grant ARAS number 600-RMI/DANA 5/3/ARAS927/2015) for this research.

REFERENCES

[1] H. Smith, *Data center storage: cost-effective strategies, implementation, and management*: CRC Press, 2016.

[2] T. Lumpp, J. Schneider, J. Holtz, M. Mueller, N. Lenz, A. Biazetti, and D. Petersen, "From high availability and disaster recovery to business continuity solutions," *IBM Systems Journal*, vol. 47, pp. 605-619, 2008.

[3] A. Singh and K. Chatterjee, "Cloud security issues and challenges: A survey," *Journal of Network and Computer Applications*, vol. 79, pp. 88-115, 2017.

[4] V. Catrinescu and T. Seward, "Implementing High Availability and Disaster Recovery," in *Deploying SharePoint 2016*, ed: Springer, 2016, pp. 349-363.

[5] Z. Gao, H. Min, X. Li, J. Huang, Y. Jin, A. Lei, S. Bourbonnais, M. Zheng, and G. Fuh, "Optimizing Inter-data-center Large-Scale Database Parallel Replication with Workload-Driven Partitioning," in *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXIV*, ed: Springer, 2016, pp. 169-192.

[6] N. A. Sulaiman, M. Kassim, and S. Saaidin, "Systematic test and evaluation process (STEP) approach on shared banking services (SBS) system identification," in *Education Technology and Computer (ICETC), 2010 2nd International Conference on*, 2010, pp. V5-219-V5-223.

[7] S. Gaonkar, K. Keeton, A. Merchant, and W. H. Sanders, "Designing Dependable Storage Solutions for Shared Application Environments," *IEEE Transactions on Dependable and Secure Computing*, vol. 7, pp. 366-380, 2010.

[8] N. A. Sulaiman and M. Kassim, "Developing a customized software engineering testing for Shared Banking Services (SBS) System," in *System Engineering and Technology (ICSET), 2011 IEEE International Conference on*, 2011, pp. 132-137.

[9] K. Padmaja and R. Seshadri, "A Review on Cloud Computing Technologies and Security Issues," *Indian Journal of Science and Technology*, vol. 9, 2016.

[10] M. E. Whitman and H. J. Mattord, *Principles of information security*: Cengage Learning, 2011.

[11] R. Sindoori, V. P. Pallavi, and P. Abinaya, "An overview of disaster recovery in virtualization technology," *Journal of Artificial Intelligence*, vol. 6, p. 60, 2013.

[12] K. Schmidt, *High availability and disaster recovery: concepts, design, implementation* vol. 22: Springer Science & Business Media, 2006.

[13] J. McDermott, "Replication does survive information," *Database Security XI: Status and Prospects*, p. 219, 2016.

[14] M. Poess and R. O. Nambiar, "Large scale data warehouses on grid: Oracle database 10 g and HP proliant servers," in *Proceedings of the 31st international conference on Very large data bases*, 2005, pp. 1055-1066.

[15] P. S. Weygant, *Clusters for High Availability: A Primer of HP Solutions*: Prentice Hall Professional, 2001.

[16] N. C. Brazelton and A. M. Lyons, "Downtime and Disaster Recovery for Health Information Systems," *Health Informatics: An Interprofessional Approach*, p. 337, 2017.

[17] N. I. Uzir, M. Salam, and M. Sahalan, "Analysis for Scalable Performance with Server clustering."

[18] D. Mishchenko, *VMware ESXi: Planning, implementation, and security*: Cengage Learning, 2010.

[19] K. Keeton, C. A. Santos, D. Beyer, J. S. Chase, and J. Wilkes, "Designing for Disasters," in *FAST*, 2004, pp. 59-62.