Using Virtualization to Ensure Uninterrupted Access to Software Applications for Financial Services Firms

Minhua Yang
Coastal Carolina University
myang@coastal.edu

Yu “Andy” Wu
University of North Texas
andy.wu@unt.edu

Abstract

With inexpensive and stable commodity hardware, virtualization has come of age and is seeing wide implementation. While benefits of virtualization abound, the focus of this paper is on a specific area – maintaining uninterrupted services. We present a set of functionalities of virtualization software (VMware ESX and Microsoft Hyper-V) that can be used to: (a) provide high availability during planned and unplanned downtime and (b) facilitate easier and more cost-effective disaster recovery. We illustrate their relevance to financial services firms, to which high availability and disaster recovery are particularly important given the nature of their business and regulatory pressure.

1. Introduction

As hardware capabilities grow by leaps and bounds at ever-decreasing prices, over the past decade virtualization technologies have come of age and made significant inroads into data centers. When implementing mission critical software applications, virtualization helps to create a highly resilient infrastructure. In a virtualized environment, innovative methods can be used to maintain highly available software services and to achieve expedient recovery if natural or human disasters devastate data centers.

For financial services firms, these features are very attractive because the nature of their business often dictates uninterrupted client access to their software applications. If disastrous events cause service outage, financial services firms cannot rely on a lengthy disaster recovery processes due to the high stakes involved. In fact, if a firm handles a sizable percentage of transactions in critical financial markets, regulatory bodies mandate that the firm establishes redundancy in its data centers, which can be an expensive undertaking even for large firms.

Therefore, in this paper, we organize our discussion of virtualization around two specific areas that are of particular interest to financial services firms: (a) high availability in software services and (b) speedy disaster recovery at reduced costs. After a brief description of virtualization, we discuss how a number of functionalities of current virtualization technologies can work to achieve success in those two areas. Since virtualization often touches on a multitude of technological elements simultaneously, we focus on the mechanisms that are specifically related to providing uninterrupted services. We conclude the paper with a brief discussion of a few relevant considerations, as a caveat that these mechanisms do have some downsides or specific requirements, which could counterbalance some of the benefits.

2. Virtualization

Modern virtualization arguably is not an entirely new technology because it has its roots in IBM VM/370, developed in the early 1970s and now z/VM. The recent resurrection and boom of virtualization took place after commodity x86 processors and other hardware, as well as Windows operating systems, became robust and stable enough [3]. Currently, the three leading players in the virtualization market are VMware, Microsoft, and Citrix. Their flagship products are VMware ESX/ESXi, Hyper-V, and XenServer, respectively. Our discussion in this paper focuses on offerings from VMware and Microsoft.

There are different types of virtualization architectures. The consumer market is more familiar with VMware Workstation, VMware Server, Microsoft Virtual PC, and Microsoft Virtual Server. These products implement what is called full, or hosted, virtualization architecture. They are installed as an application on the operating system (OS) and emulate an entire hardware environment, with devices that are supported by many OSes. This approach provides a...
great deal of flexibility to users, but at the expense of performance. Since access to hardware is emulated in software, what would normally be handled by hardware is now carried out with execution of a large number of software instructions [3].

Targeting the corporate market, VMware ESX/ESXi, Microsoft Hyper-V, and XenServer implement the Type 1 virtualization architecture and are termed hypervisors. The fundamental difference between Type 1 and full virtualization architectures is that hypervisors run directly on the hardware without the intervention of OS [3], [8]. Because of this, they are said to be installed on “bare metal,” as a raw server system and not relying on an existing OS to function [8]. For example, VMware ESX can be installed from installation media. Although it resembles Red Hat Enterprise Linux, it is actually proprietary and should not be considered a flavor of Linux [7].

The case of Hyper-V is more complicated, since it is shipped as part of Windows 2008. Hyper-V can be installed as a role of a Windows 2008 server. Then, during the boot process, Hyper-V inserts the hypervisor binary code between the Windows OS and the physical hardware. Alternatively, the Microsoft hypervisor may be installed as Hyper-V Server. It is a Windows Core installation with only the Hyper-V role installed. As such, it has no graphic user interface and has to perform management through remote management tools. Nevertheless, it has the Windows kernel and all the device drivers supported by the full Windows platform [3].

In virtualization lexicon, the physical server on which virtualization software is installed is a host. Each host can run one or more VMs (also known as guests). Within each VM, an operating system is installed; applications, in turn, are installed on the OS platform.

It is not uncommon to host many VMs on a single host, e.g., on average 20+ VMs on large ESX hosts [7]. A major advantage of virtualization is consolidation of a large number of physical servers onto a much smaller number of hosts, with each server running as a VM. For most applications, execution in a VM does not differ from when they are run on a physical server. A client that requests services from applications usually is not concerned about whether the response originates from a virtualized or physical instance of the application. Virtual service offerings are feasible for a large number of applications in various categories [11]. For financial services firms, two categories are highly relevant – identity management and application servers. Regardless of the application-hosting environment, financial services firms’ business applications need to run continuously and unattended [19]. In the next section, we present virtualization functionalities that work to maintain high availability for clients.

3. High availability functionalities

Many financial services firms operate across a number of time zones globally, leaving razor-thin room for any downtime. In addition, large clients that pay handsomely for financial services frown upon outage of service [14]. Not less importantly, small, individual customers can sustain adverse impact due to software outage. In April 2004, U.S. Central Credit Union of Lexana, KY, caused much customer outcry when its software was overwhelmed by the volume of Automated Clearing House (ACH) transactions it processed [22].

Therefore, financial services firms need to ensure that their applications operate with the least amount of downtime. A corollary is that spikes in the workload should be handled gracefully. The traditional solution to this is to provision additional servers, which typically form a cluster of servers.

Virtualization brings to the table some innovative enhancements to the cluster solution. First, with the assumption that the application runs in a VM, high availability (HA) functionality is a feature of hypervisors that significantly reduces downtime by moving the VM from a failed host to a working host. All major players in the hypervisor market offer HA features, although not all of them are equally equipped for unplanned downtime.

3.1. Reducing unplanned downtime

The first step in preventing unplanned downtime is to balance the request load across VMs so that not a single VM will fail because excessive amount of requests exhaust its resources. The foundation of load balancing is a cluster of servers.

Virtualization software relies on the clustering functionality that is offered as a feature of operating systems. For example, VMware load balancing may be set up on a cluster of servers running Microsoft Windows. Microsoft has been offering the capability of clustering in the advanced versions (Enterprise and Datacenter) of its server operating systems since Windows NT. With Hyper-V and Windows 2008 R2, Microsoft greatly enhances cluster capability and manageability [3].

3.1.1. Microsoft load balancing. Microsoft Network Load Balancing (NLB) aggregates resources of the servers in the load-balancing cluster into a pool with a common network name. In a load balancing cluster,
many, if not all, nodes operates simultaneously. An algorithm determines how service requests are “parceled out” to those servers so that a certain percentage of requests is processed by each server [15]. In a non-virtualization environment, a cluster-aware application runs on servers in the cluster and the best capable server(s) will be called upon to answer new service requests as they occur. In a Hyper-V environment, the “load” that is balanced is in fact VMs, all running the same focal application. Once an additional VM is started on the server with the optimal resources to run the VM, the instance of the application installed in that VM in turn handles new requests for the services provided by the application.

3.1.2. VMware load balancing. In VMware, multiple ESX hosts share the same datastore and form a cluster, which is represented as a cluster object in VMware vCenter. VMware Dynamic Resource Load Balancing (DRLB) automatically distributes the usage of resources over multiple ESX hosts. DRLB closely monitors VMs’ and ESX hosts’ utilization of CPU, memory, disk, and network. These resources are divided into units that are termed “resource pools.” The administrator allocates a finite number of resource pools to each ESX host, based on his or her estimates of the resource usage of the VMs running on that host. Thus, resource pool prevents a single VM or group of VMs from exhausting all available resources. However, flexibility is not lost because the administrator can configure borrowing of resource pools from other ESX hosts or clusters [7], [12]. The workhorse in the DRLB is the VMware Distributed Resource Scheduler (DRS). It determines which host in a cluster should run a VM. DRS also constantly evaluates the load on the hosts and decides whether migration is necessary to balance the load [12]. Three settings of DRS are available. The “manual” setting only warns the administrator that load balancing should take place but does not initiate balancing. If DRS is “partially automated,” it moves a VM to a recommended host in the cluster when the VM is powered on. “Fully automated” DRS monitors the resource utilization and if an ESX host is over-utilized, DRS automatically moves VMs from that host to another ESX host where more resources are available. Automated balancing can be done in either “aggressive” mode, in which VM moving occurs the moment resource utilization goes above the threshold, or “conservative” mode, in which DRS will wait until over-utilization has exceeded threshold substantially [7]. An administrator also can create and apply DRS rules [12].

3.1.3. Microsoft fault tolerance. While load balancing prevents requests from overwhelming an application hence causing an outage in service, HA solutions still need to account for the possibility, however slim, that a VM may fail abruptly. This commonly is achieved using fault tolerance/failover mechanisms.

In a Hyper-V environment, the Failover Cluster Services in Windows 2008 accommodate clusters of between 2 and 16 servers (nodes). When the VMs on one of the nodes become unavailable, Microsoft HA will move those VMs to another node, with an extremely brief interruption in services. The backbone of a Microsoft failover cluster is the share storage, usually a Storage Area Network (SAN), which is divided into logical volumes. Each logical volume is termed a Logical Unit Number (LUN), which, when attached to a server, appears as a local drive.

3.1.4. VMware fault tolerance. VMware also strengthens its ability to handle unplanned downtime with VMware Fault Tolerance (VMware FT). It is based on the same principle behind hard disk mirroring. When VMware FT is turned on, VMware vCenter Server creates a “shadow” VM for the primary VM. Only the primary VM actually responds to service requests, while the shadow VM is kept in sync with the primary VM silently and continuously [7], [12]. If the primary VM fails, the secondary VM takes over immediately with no interruption in service. Since the secondary VM now becomes the primary, a new shadow VM will be created immediately and kept in sync with the new primary VM [12].

3.2. Reducing planned downtime

Compared with unplanned downtime, planned downtime has the welcome characteristic of being initiated by the administrator. From time to time, VMs may need to be moved between hosts. A common reason is that numerous VMs vie for the host’s limited resources. The administrator should diligently find a time to take down the VMs so that interruption of service is minimized. To this end, virtualization vendors have created a number of migration tools, which, when used properly, can migrate VMs in so brief a time that service does not appears to be interrupted at all to the client.

3.2.1. Microsoft migration tools. Quick Migration was the only VM-moving option in the original release of Windows 2008. Because it requires orderly migration of VMs and cannot react to abrupt outages, Quick Migration is more proper as a HA solution in situations of planned downtime. In a Quick Migration, Hyper-V saves the current state of the VM, as a file with the .vsv file extension, to the shared storage of
the cluster. The ownership of the LUN on which the VM resides is then handed over to another node in the cluster, which then reads the folder for the VM and restarts the VM [3], [16].

Windows 2008 R2 added a new feature called Live Migration, which put Microsoft on an equal footing with VMware in terms of minimized planned downtime. When Live Migration starts, it makes iterative copies of a running VM’s memory from its current node to another node. The purpose is to preserve the changes in the memory of the VM, each iteration copying only changes made since the last copy (the concept is similar to the incremental backup). Thus each copy is smaller than the previous one. When the size of the “dirty” memory gets small enough, Live Migration pauses the VM and copies the last portion of changed memory. It then resumes the VM on the node that is the migration target and applies all changes in memory. The result is that the VM now runs on another node, with its memory in the state it was in when it was paused. Since the VM’s folder is stored on the SAN, it does not have to be copied to the second node. The “outage” only occurs after the VM was paused on the original node and before it was resumed on the second node. This takes place in so short a time (less than a second, well within the timeout of a typical TCP request) that, to applications running via TCP, no interruption in service is noticed [3].

Another important facilitation for minimizing of downtime is Clustered Shared Volumes (CSV). Before Windows 2008 R2, when the second node took ownership of the VM’s folder, the takeover had to occur for all the other VMs whose folders were on the same LUN. Thus, all VMs stored on a LUN must be migrated all at once or none will be migrated at all. R2 addresses this inflexibility by introducing CSV to decouple VM migration from LUN ownership. This means that Live Migration may occur on only one VM even when 20 other VMs reside on the same LUN [3].

3.2.2. VMware migration tools. Starting with the letter “v” to be consistent with the rest of the VM nomenclature, vMotion is the VMware counterpart of Microsoft Live Migration. Just like the latter, vMotion enables the administrator to make a copy of a running VM and move it to another host, without perceptible interruption in service. The inner workings of vMotion also are similar to those of Live Migration. Again, vMotion often becomes necessary due to contention for host resources [7], [12].

4. Disaster recovery functionalities

HA is a site-level redundancy solution. In other words, it only protects against interruptions in service to the extent that another server or VM will take over the role of a failed server or VM in the same data center. However, it will be powerless in the case that the entire data center is wiped out or fatally damaged [12], [16]. In fact, the hardware and software used for HA will be part of what is destroyed. To cope with such a disastrous situation, disaster recovery (DR) is in order. DR is about restoring essential operations after the devastation of an extraordinary event, e.g., September 11 or Katrina. To date, the fundamental approach to DR is creation and maintenance of redundancy. A wide variety of measures exist to provide redundancy. Fundamentally, they boil down to two types – redundancy of infrastructure and redundancy of data.

4.1. Infrastructure redundancy

Due to the nature of their industry, financial services firms tend to be better prepared than other firms to stand up against natural or human disasters [5]. For instance, a number of financial services firms managed to resume their operations shortly after September 11, largely owing to the mirror sites [6]. Nevertheless, the satisfactory DR report card turned in by those firms perhaps could be attributed largely to their critical geographic locations and prominent standing in the industry. September 11 exposed the risks of data loss caused by disastrous events. Consequently, on April 7, 2003, the Securities and Exchange Commission (SEC), Comptroller of the Treasury, and the Federal Reserve issued the Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System [4]. Now, financial services firms that account for at least 5% of the transactions in critical financial markets are required to implement sound business continuity practices. One required practice is to maintain sufficient geographically dispersed resources, most importantly, data centers. The targeted recovery time is two hours. Although business continuity practices are not mandatory for other financial services firms, many of those firms are adopting the practices because it is prudent to do so [18]. However, a recent Gartner inquiry [18] found that cost is a concern to both large and smaller financial services firms when they implement and maintain disaster recovery solutions.

Financial services firms can address this concern by using virtualization technologies. Large firms can use virtualization to reduce the cost in maintaining the three data centers. Similarly, medium-sized firms can reduce cost for their two data centers. With enough planning and cost cutting, the three data center strategy, which typically would be out of their reach using only physical servers, can become a possibility.

5626
for medium-size firms. Of course, there are other non-cost factors that determine whether these firms will opt for the three-center strategy; but at least virtualization has the potential of making such a strategy affordable for those firms.

The cost saving mainly comes from the number of servers that a firm is able to maintain with a given amount of hardware. In the traditional configuration, each server occupies a physical machine, with low utilization rates of resources. With virtualization, multiple servers are hosted as VMs on a single physical server, making the fullest possible use of the capabilities of the physical hardware. For example, for each VM host, 80% utilization rate often is used as the threshold, beyond which an additional host becomes necessary [7].

For example, ICICI Bank, the second-largest bank in India, was an early adopter of virtualization. By using VMware ESX, the bank reduced its standalone servers from 350 to 50, running a total of 700 VMs [13]. When Citicorp’s Georgetown, TX data center was completed in 2009, virtualization generated a 30% to 35% reduction in its server count, as compared to the firm’s other, legacy data centers. All told, Citicorp runs more than 4,000 server virtualization instances, which were used to replace many times the number of physical servers; 12,000 to 20,000 servers were either not installed or converted to VMs, thanks to virtualization. The year 2009 marked the first time that Citicorp did not experience any growth in its physical server category [17]. Similarly, Nationwide, the financial services and insurance firm, enjoyed an 80% reduction in floor space, thanks to virtualization with VMware [2].

With fewer physical servers, data centers operate with less energy consumption, which also saves cost. For example, Citicorp’s 100,000-sq. ft. floor Georgetown, TX data center uses, for the same footprint, 800 KW less power than do its other data centers, as the result of being able to maximize the use of virtualization in this newest data center [17]. At Nationwide, virtualization translated into an annual $40,000 reduction in energy consumption [2]. In addition, smaller space is needed for the existing hardware; or the existing space can be used to host more servers, translating into more cost saving [13].

### 4.2. Data redundancy

Above, we described the first advantage virtualization offers in DR, which concerns the firms’ ability to build and maintain multiple infrastructures. The second advantage is about, once the infrastructures are in place, how they can be further utilized to realize the second type of redundancy – data redundancy.

Obviously, multiple data centers in different geographic locations allow the replication of information from one site to another, providing fault tolerance. As the result, financial transaction data stored by servers in the primary location can still be assessed from a secondary location when the primary location is down. With today’s SAN technologies and network bandwidth, enormous volumes of data can be replicated at so fast a speed that not only financial data can be replicated, but VM folders can be replicated between sites to achieve a speedy DR process [12].

**VMware Data Recovery (VDR)** is integrated in the vCenter Server. It allows an administrator to select the VM to protect, schedule the backup, and choose the destination for the backup to go to. Another component of VDR, the VDR virtual appliance, then starts the process by taking and mounting a snapshot of the VM. The snapshot is then streamed to the destination storage. The snapshot is dismounted after it has been completely replicated to the destination [12].

VMware replication also can be done using third-party tools. VMware’s vStorage APIs allow third-party vendors to develop programs that run backups of VMs without involving VMware ESX servers [12].

Microsoft’s DR solution is in fact an enhanced version of its HA offering. Microsoft failover cluster can be set up between two geographically dispersed sites. Key to this is to replicate VMs to the remote site continuously. Currently, this solution relies on independent software vendors (ISVs). When a disaster strikes, the ISV routines running on Failover Cluster service remap the SAN LUNs or storage volumes to be fully read/write LUNs/volumes available to the backup cluster in the remote site. This setup is called a “stretch cluster” and Microsoft has worked with a number of ISVs to provide this capability [3].

The foundation for DR of VMs is the data replication process that replicates the VM folders from one site to another. If disastrous event occur, some replication engines can even automatically redirect service requests to the remote site [16].

### 4.3. Benefits of VM-based DR

A common disaster recovery and business continuity solution is the use of secondary/DR sites. The traditional approach maintains redundant physical servers at the DR site and uses them to replace production servers if the latter become inaccessible due to a disaster. Compared with the traditional approach, virtualization can save costs and reduce some potential problems in maintaining a DR site.

The traditional approach basically duplicates the number of servers at the DR site (a pattern sometimes referred to as “2N,” where N is the number of servers...
at the primary site) [16], [20]. Reduction in this duplication is possible with virtualization. The physical servers at the production site can be virtualized using tools such as VMware P2V (physical-to-virtual) and consolidated on a much smaller number of physical platforms at the DR site [9]. The total number of physical servers at the two sites can go down from 2N to as far as N+1 [20].

If the primary site runs a substantial number of VMs, DR essentially entails firing up backup copies of these VMs at the DR site. This process is potentially speedier and easier, when compared with getting physical backup servers into production mode. A major drawback often exists in using physical backup servers. They are typically retired production servers thus one notch down from those at the primary site. This often causes driver and compatibility issues when trying to restore applications and data on backup servers. Also, since the process often necessitates installation/update of operating systems, installation of applications, restoration of data from backup media, etc., the multitude of hardware and software elements involved adds to the complexity and length of DR [9], [12], [20]. In contrast, with VMs the restoration of a server is simplified as copying the server’s VM folders onto a DR server and starting the VMs [9]. Even though a VM runs just like a physical machine, to the host, it is implemented as a folder in the host’s file system, not any different from other ordinary folders. In each VM, a number of files can be seen, each serving a specific purpose in the operation of the VM. For example, Table 1 is a small sample of VMware file types [7].

Table 1. VM Files

<table>
<thead>
<tr>
<th>Extension</th>
<th>Purpose/Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>.vmdk</td>
<td>VM disk file</td>
</tr>
<tr>
<td>.vmsn</td>
<td>Snapshot</td>
</tr>
<tr>
<td>.nvram</td>
<td>Nonvolatile RAM file</td>
</tr>
<tr>
<td>.vswp</td>
<td>Swap file</td>
</tr>
<tr>
<td>.vmtx</td>
<td>Configuration template</td>
</tr>
<tr>
<td>.vmx</td>
<td>Configuration</td>
</tr>
</tbody>
</table>

Therefore, the folders and files that make up VMs can be replicated from site to site using the same technologies that replicate financial transaction data. Since the primary and secondary sites both have access to the VM folders, the backup site can be brought into action with minimal delay if the primary site is lost to a disaster.

VMs are platform- and hardware- agnostic. Their OS and applications run on a standard set of virtualized hardware emulated by the hypervisor, not on the physical hardware of the host. This in essence abstracts away the difference between primary site and DR site hardware and allows VMs from the production site to be run on DR servers with markedly different hardware and software [9], [16]. The main requirement of the DR server is that it has the proper hypervisor software installed [9]. Provided the hypervisor at the DR site is identical to the primary site installation in terms of version and patch level, one can expect to experience faster DR time and fewer hiccups when restoring virtualized servers.

Faster recovery time and higher success rate were reported in a case in which 78 physical Dell servers at the primary site were supported by a DR site of 38 HP servers. VMs are maintained on EMC Symmetrix SANs and synchronized with EMC remote data replication software. The two sites were connected via a dedicated OC3 connection at 155Mb/s and IP over Fiber. In a recovery test, it took only 16.5 minutes for the VMs at the DR site to become operational successfully. In comparison, when the company outsourced DR and the traditional tape restore method was used, the estimated time to recover was 6.5 days and recovery tests never succeeded [20].

4.4. Cost factors

Critical to the VM-based DR solution is the SAN replication tool. Major virtualization vendors all have design features aimed to minimize the potential interruption caused by DR. However, they still rely heavily on third-party tools for data replication [12], [16]. A good replication engine should be able to work with any storage infrastructure, support automatic failover of resources and redirection of users, and support isolated recovery testing without taking the production network down. When the integration of replication tools and the hypervisor software is ideal, DR could take place without perceptible effects to users [16]. However, better tools usually are more expensive.

Another major cost consideration is the concentration ratio of VMs, i.e., the number of VMs on each host. A smaller number of VMs is less likely to strain the resources on the host. The downside is the related cost to provision the larger number of physical servers required.

Decision on these should be based on a cost-benefit analysis. The analysis can be informed by two major performance metrics – Recovery Point Objective (RPO) and Recovery Time Objective (RTO). RPO describes the acceptable amount of data loss measured in time. RTO is the duration of time and a service level within which a business process must be restored after a disaster [21]. Service level agreement is an important consideration when balancing the amount of physical...
resources at the DR site with the need to restore critical services [16]. A firm may also set targets based on some commonly used security metrics, such as host uptime in percentage, meantime to recovery, etc. [10].

5. Other considerations

As can be seen from the above sections, virtualization vendors offer a number of HA and DR functionalities to ensure uninterrupted access to applications hosted in a virtualized environment. Besides these, financial services firms should consider a few other factors because they can affect the implementation of their HA/DR plans.

First, while virtualization brings about consolidation of servers and saves data center operation costs, a downside to hosting multiple VMs on each host is that it creates single point of failure [3], [11], [16]. This is ironic to the goal of using virtualization features to ensure uninterrupted service. While the cluster solutions mentioned above provide protection against server failures, it is helpful to prevent hardware failure in the first place. A major concern is server overheating. Since virtualization can push the envelope of server resource utilization, overheating becomes more common at higher utilization rates. Solutions to this concern is more environmental than technical – management of data center’s HVAC systems, sufficient spacing between server racks, etc.

Second, HA mechanisms based on server clusters can raise the bar for hardware homogeneity. Clustering requires, at the very least, servers installed with processors of the same architecture. It will not run on mixed architectures, e.g., Intel and AMD. It is also recommended that servers have identical configuration. This ensures that when a VM is migrated to a new host, minimal interruption occurs [11]. Homogeneity requirements can increase hardware acquisition cost, reduce flexibility in hardware provision, or both. These impacts should be weighed against the benefits reaped from clustering before a decision is made.

Third, almost all the solutions we covered rely on a shared SAN or multiple SAN [9], [11], [12]. In the HA (site-level) solutions, all nodes in a cluster share the LUN, which typically stores the VM folders. In the DR (multiple-site) solutions, data from the primary site are replicated to the backup site. Typically, replication occurs synchronously for critical data, including VM folders, while less time-sensitive data are replicated asynchronously. Access speed of the SAN technology used, e.g., FCoE, SATA, iSCSI, etc., can affect the migration and replication. Slower speed may cause more perceptible service interruptions. Also, tools provided by third-party vendors often play an important role, sometimes even as an integrated part of the solutions. Therefore, selection of SAN vendors can be a key success factor.

Fourth, both internal and external factors may affect service availability. Above, we only discussed one type of internal factor – application of virtualization technology – that helps to maintain uninterrupted service. A number of external factors, most notably Distributed Denial of Service (DDoS) attacks and malware, are capable of interrupting service [1]. Virtualization does not have native tools to address these external factors. Prevention comes from information security measures, e.g., firewalls, intrusion detection systems, virus scanners, etc.

6. Conclusion

Modern virtualization technologies afford us innovative methods to maintain uninterrupted software service. This paper conducts an overview of various functionalities in VMware and Microsoft hypervisors that can be applied to ensure high availability of service during downtime and to achieve data and infrastructure redundancy. Proper utilization of these functionalities can minimize interruption due to planned or unplanned downtime. Disaster recovery also can be facilitated due to reduced cost and ease of backup and restoration. Financial services firms can particularly benefit from these functionalities. A few related considerations are discussed.

7. References


