Local Disaster Recovery Using Virtualization Technology

Aye Myat Myat Paing
University of Computer Studies, Yangon
paing.ayemyat@gmail.com

Abstract

Business applications running on IT infrastructure necessitate high levels of availability in order to minimize the amount of downtime experienced during any planned and unplanned outages. As a result, disaster recovery has gained great significance in IT. Exploiting virtualization and ability to automatically reinstall a host, where the action on a virtual machine is performed only when a disaster occurs. Virtualization affords significant cost and performance advantages over more traditional disaster recovery options such as tape backup or imaging. Our approach is to design and implement a continual migration strategy for virtual machines to achieve automatic failure recovery. By continually and transparently propagating virtual machine's state to a backup host via live migration techniques, trivial applications encapsulated in the virtual machine can be recovered from hardware failures with minimal downtime while no modifications are required. Moreover, our framework intends to monitor virtual machines for problems such as CPU utilization, I/O activity, and memory utilization. This raises a difficult problem, since it is quite difficult to discriminate based on these measures between a virtual object that is performing properly, and one that is quite ill. We apply the out-of-band monitoring using virtualization and machine learning can accurately identify faults in the guest OS, while avoiding the many pitfalls associated with in-band monitoring.

Keywords: virtualization, availability, fault-tolerance, machine learning

1. Introduction

Virtualization has been widely adopted by data centers for transparent load balancing, application mobility, server consolidation and secures computing [1]. With one physical machine hosting many virtual machines, a single node failure may result in more severe disruption to hosted services, which brings great challenge for automatic failure recovery in virtualized computing environments. One of the most general solutions for failure recovery is to replicate the states of the protected virtual machine to a backup host, with which the virtual machine may be recovered from host failures. Although virtual machine replication can be done in various ways, its consistency and efficiency are not guaranteed. Most virtual machines maintain memory and external storage states in separate ways, which may result in an inconsistency when replicated to the backup host. In the mean time, fast and transparent failure recovery requires the backup to be synchronized with the primary virtual machine; however, synchronizing on every change brings too much performance tradeoffs [9].

In this paper, we present a continual migration mechanism for virtual machine based fault tolerant systems by continually and transparently propagating virtual machine's states to a backup host via live migration [3] techniques. While other migration based high available systems [9] maintains a secondary disk image file on the backup host, continual migration propagates disk states via network attached storage, which is more common in a data center configuration and provides better agility in both configuration and deployment. Some research work achieves failure recovery by taking and restart from checkpoints [7], [4], [8]. Continual migration aims to provide transparent fault tolerance for commodity applications in data centers or virtual computing environments [5].In addition, our framework make use of out-of-band monitoring system using machine learning and virtualization. Traditional approaches use in-band monitoring agents. However in-band agents suffer from several drawbacks: they need to be written or customized for every workload (operating system and possibly also application), they comprise potential security liabilities, and are themselves affected by adverse conditions in the monitored systems. Therefore, this paper describes one approach to out-of-band monitoring that performs this discrimination based on statistical analysis, as implemented using machine learning.

2. Related Work

Remus provides an extremely high degree of fault tolerance, to the point that a running system can transparently continue execution on an alternate physical host in the face of failure with only seconds of downtime, while completely preserving host state such as active network connections. This system encapsulates protected software in a virtual machine, asynchronously propagates changed state to a backup host at frequencies as high as forty times a second, and uses speculative execution to concurrently run