Computer Networks 93 (2015) 423–434

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Server placement with shared backups for disaster-resilient clouds

Rodrigo S. Couto [a,b,*], Stefano Secci [c], Miguel Elias M. Campista [a], Luís Henrique M.K. Costa [a]

[a] Universidade Federal do Rio de Janeiro, COPPE/PEE/GTA - POLI/DEL P.O. Box 68504 - CEP, 21941-972 Rio de Janeiro, RJ, Brazil
[b] Universidade do Estado do Rio de Janeiro, FEN/DETEL/PEL CEP, 20550-013 Rio de Janeiro, RJ, Brazil
[c] Sorbonne Universités, UPMC Univ Paris 06, UMR 7606, LIP6 F-75005, Paris, France

ARTICLE INFO

Article history:
Received 15 February 2015
Revised 10 August 2015
Accepted 19 September 2015
Available online 17 October 2015

Keywords:
Cloud networking
Resilience
Geo-distributed data centers
Infrastructure as a service

ABSTRACT

A key strategy to build disaster-resilient clouds is to employ backups of virtual machines in a geo-distributed infrastructure. Today, the continuous and acknowledged replication of virtual machines in different servers is a service provided by different hypervisors. This strategy guarantees that the virtual machines will have no loss of disk and memory content if a disaster occurs, at a cost of strict bandwidth and latency requirements. Considering this kind of service, in this work, we propose an optimization problem to place servers in a wide area network. The goal is to guarantee that backup machines do not fail at the same time as their primary counterparts. In addition, by using virtualization, we also aim to reduce the amount of backup servers required. The optimal results, achieved in real topologies, reduce the number of backup servers by at least 40%. Moreover, this work highlights several characteristics of the backup service according to the employed network, such as the fulfillment of latency requirements.

1. Introduction

Many corporations are migrating their IT infrastructure to the cloud by using IaaS (Infrastructure as a Service) services. Using this type of service, a corporation has access to virtual machines (VMs) hosted on a Data Center (DC) infrastructure maintained by the cloud provider. The use of IaaS services helps cloud clients reduce the effort to maintain an IT infrastructure; with IaaS, clients relinquish the control of their physical infrastructures. Therefore, they only rely on IaaS services if providers can guarantee performance and security levels. To encourage IaaS subscriptions, cloud providers usually try to offer high resilience levels of their VMs. To this end, IaaS providers deploy redundancy on their infrastructure to overcome various types of failures, such as hardware (e.g., failure in hard disks, network cables, and cooling systems), software (e.g., programming errors), and technical staff (e.g., execution of wrong maintenance procedures). This strategy, however, does not guarantee service availability under force majeure and disaster events that are out of the provider's control.

Force majeure and disaster events, such as terrorist attacks and natural disasters, are situations outside of the provider's control, which can affect several network links as well as whole buildings hosting data centers. Cloud providers thus generally do not cover this type of event in their SLAs (Service Level Agreements) [1]. Although IaaS providers often do not consider catastrophic events, they can offer recovery services such as VM replication and redundant

* Corresponding author at: Universidade do Estado do Rio de Janeiro - FEN/DETEL/PEL CEP, 20550-013 Rio de Janeiro, RJ, Brazil. Tel.: +55 21 23340027.
E-mail addresses: rodrigo.couto@uerj.br, souza@gta.ufrj.br (R.S. Couto), stefano.secci@upmc.fr (S. Secci), miguel@gta.ufrj.br (M.E.M. Campista), luish@gta.ufrj.br (L.H.M.K. Costa).

http://dx.doi.org/10.1016/j.comnet.2015.09.039
1389-1286/

network components to improve the resilience to clients running critical services. These services can be provided as long as a DC infrastructure resilient to disasters is available, which is generally composed of several sites spread over a region and interconnected through a wide area network (WAN) [2]. Each site has a set of servers interconnected using a local network [3,4]. A resilient IaaS cloud must thus employ a geo-distributed DC to eliminate single points of failure and must employ mechanisms to perform VM backups. Obviously, clients willing to have higher resilience guarantees will pay the cost of maintaining such infrastructure.

In this work, we focus on the design of disaster-resilient DCs with zero VM state loss (e.g., loss of disk and memory content) after a disaster. This means that the provider guarantees zero RPO (Recovery Point Objective) on its VMs. RPO is the time elapsed between the last backup synchronization and the instant when the disaster happens. Hence, it gives an idea of data loss after a disaster [1]. Some critical services demand a low RPO or even zero RPO, such as banking transactions, requiring continuous data replication. Basically, an IaaS with zero RPO consists on VMs that continuously send backups to a server. In this case, an operation demanded by an end user is only accomplished after the VM receives an acknowledgment from the backup site, indicating that the VM state was correctly replicated [5]. As this type of service requires continuous data replication, it requires a high network capacity. Furthermore, as it needs backup acknowledgment, the primary server, i.e., the server hosting the operational VMs, and the backup one must have low latency links between each other.

The literature about resilient physical server placement considers a traditional DC distribution, such as those employed by content delivery networks (CDNs) [6,7]. In these works, the DC services are replicated through a geo-distributed architecture using anycast. Hence, any node that runs the required services are operational and can reply the requests from clients. Consequently, the primary servers and their backups are both running at the same time. Nevertheless, these works do not consider the synchronization of service replicas, disregarding RPO requirements.

This work analyzes the behavior of IaaS services with zero RPO in real WAN topologies. We propose a physical server placement scheme, which designs the DC by choosing where to install the primary servers and their corresponding backups. The placement scheme has to take into account the failure model, in such a way that a disaster does not damage the primary server and its backup at the same time. In addition, the proposed scheme takes advantage of virtualization to reduce the number of backup servers. The basic idea is that a backup server needs to instantiate VMs only after a given disaster occurs [8]. We thus argue that, in a virtualized environment, it is inefficient to provide a dedicated backup server for each primary one. Instead, the proposed scheme aims at sharing backup servers, allowing them to receive replications from different primary servers. To share these resources, the primary and backup servers must not fail at the same time. We apply the proposed scheme in WAN topologies and show that backup sharing can reduce by at least 40% the number of required servers, as compared to the case with dedicated backups. We also quantify the capacity of each WAN topology in terms of number of primary servers supported, which directly affects the number of supported VMs. Using these results, we show that more stringent resilience requirements reduce by at least 50% the number of primary servers supported. Our work differs from the literature by considering the service replication, which incurs in stringent latency and bandwidth requirements. In addition, the current proposals based on anycast do not save backup resources, since all backup servers are also operational. We thus focus on IaaS models, different from traditional CDNs.

This work is organized as follows. Section 2 describes the service model and our design decisions. Based on these decisions, Section 3 introduces the proposed optimization problem. Section 4 shows the results of the optimization problem when applied to real WAN networks. Finally, Section 5 presents related work and Section 6 concludes this work and points out future directions.

2. Modeling and design decisions

The optimization problem proposed in this work distributes primary and backup servers in a given WAN topology. The primary servers are employed to host operational VMs, which are accessed by Cloud users through gateways spread across the WAN; Backup servers receive VM copies from these servers. A VM backup is a complete copy of its primary VM, but it keeps in standby mode in a normal situation. Each primary server replicates VM copies to a single backup server installed in another DC site. This section details the DC design decisions considered in the optimization problem formulation, which is described later in Section 3.

2.1. VM replication

The VM backup scheme considered in this work is based on continuous and acknowledged VM replication, which allows the provider to guarantee zero RPO (Recovery Point Objective) when a disaster occurs. This type of scheme is common in local networks, being natively available in virtualization platforms such as Xen [9]. More recently, VM backup schemes with zero RPO using wide area networks (WANs) started to be addressed in the literature [5,10]. As an example we can cite SecondSite [5], employed as a reference throughout this article. To achieve zero RPO, SecondSite is based on checkpoints. A checkpoint is defined as the VM state (e.g., disk, memory, CPU registers) at a given instant. Such state is continuously sent to a backup server that, in its turn, sends an acknowledgment to the primary server for each received checkpoint. The basic of operation of a VM is to run applications that receive requests from users on the Internet and reply these requests. Before a checkpoint acknowledgment, network packets sent from the VM applications to the users are held in a queue, waiting for the upcoming acknowledgment. When the backup server confirms the checkpoint replication, all packets in the queue are sent to users. Hence, the final user only receives a reply to his requests after the correct replication in the backup server. Note that SecondSite imposes strict bandwidth and latency requirements. The high bandwidth utilization is due to the continuous data replication, which increases with the frequency of changes in the VM state. The strict latency requirements are