The Case for Enterprise-Ready Virtual Private Clouds

Timothy Wood[§] Alexandre Gerber[†] K.K. Ramakrishnan[†]
Prashant Shenoy[§] Jacobus Van der Merwe[†]

[§]*University of Massachusetts Amherst*
{*twood,shenoy*}@*cs.umass.edu*

[†]*AT&T Labs - Research*
{*gerber,kkrama,kobus*}@*research.att.com*

Abstract

Cloud computing platforms such as Amazon EC2 provide customers with flexible, on demand resources at low cost. However, while existing offerings are useful for providing basic computation and storage resources, they fail to provide the security and network controls that many customers would like. In this work we argue that cloud computing has a great potential to change how enterprises run and manage their IT systems, but that to achieve this, more comprehensive control over network resources and security need to be provided for users. Towards this goal, we propose CloudNet, a cloud platform architecture which utilizes virtual private networks to securely and seamlessly link cloud and enterprise sites.

1 Introduction

Cloud computing enables enterprises large and small to manage resources better – some because they no longer need to invest in local IT resources and instead can lease cheaper, on-demand resources from providers, and others because they can utilize the flexibility of cloud resources to dynamically meet peak demand without investing in in-house resources. Cloud computing is a natural fit for enterprise customers since it enables outsourcing of another set of non core competencies: IT infrastructure selection, ordering, deployment, and management. Cloud computing allows enterprises to obtain as much computation and storage resources as they require, while only paying for the precise amount that they use. Since cloud platforms rely on virtualization, new resources can be quickly and dynamically added to a customer's resource pool within minutes. From a cloud computing service provider's perspective, server virtualization allows them to flexibly multiplex resources among customers without needing to dedicate physical resources individually. These features have driven the growth of commercial cloud computing services, making them increasingly popular and economical.

However, current cloud computing services need to further evolve to fully meet the needs of businesses. It is highly desirable that cloud resources be seamlessly integrated into an enterprise's current infrastructure without having to deal with substantial configuration, address management, or security concerns. Instead, current commercial solutions present cloud servers as isolated entities with their own IP address space that is outside of the customer's control. The separation of cloud and enterprise resources increases software and configuration complexity when deploying services that must communicate with an enterprise's private network. This can lead to security concerns since enterprise customers must utilize IP addresses on the *public* Internet in order to link application components in the cloud to their own sites. It is left to the customer to manage security on the cloud resources and the enterprise network through firewall rules. Finally, existing cloud services focus on storage and computation resources, and do not allow for control over network resources either within the cloud or the network linking enterprise and cloud sites. The lack of coordination between network and cloud resources leaves the customer again responsible for independently arranging for traffic isolation and bandwidth guarantees with a separate network service provider.

To overcome these deficiencies, we propose the enhancement of the cloud computing framework to seamlessly integrate virtual private networks (VPNs). To this end, we propose CloudNet, which joins VPNs and cloud computing. CloudNet uses VPNs to provide secure communication channels and to allow customer's greater control over network provisioning and configuration.

However, there are challenges to be dealt with when combining VPNs and cloud platforms. First, creating VPN endpoints requires coordination between the network operator and cloud service provider; existing cloud services do not provide sufficient "hooks" to allow cloud resources to be securely attached to a VPN endpoint. Next, provider provided VPNs typically extend only between edge routers within the provider network. Cloud operators must ensure that network isolation extends through any local network infrastrcuture, e.g., switches and routers, within the cloud site itself. Finally, VPNs have traditionally been provisioned at management timescales mainly because the endpoints of a VPN were expected to remain static for long periods of time. In cloud computing, however, flexibility and rapid provisioning are key requirements, and it is essential that the network transparency and secure communication channels provided by VPNs remain effective despite rapid changes in server and network configuration.

In this work we present the CloudNet architecture, which uses the idea of Virtual Private Clouds to create flexible, secure resource pools transparently connected to enterprises via VPNs. CloudNet achieves these goals by automating the creation and management of VPN endpoints and allowing for explicit coordination between the cloud platform and the network service provider.

2 Background & Related Work

Commercial Platforms: Cloud computing has rapidly grown in popularity over the last few years due to the reemergence of virtualization as an efficient method of flexibly sharing resources. There are many different types of cloud computing services ranging from web based word processors and email clients to application development platforms like Google App Engine [6] to virtual infrastructure providers like Amazon EC2 [1] that lease full virtual machines to customers. The authors of [2] provide a good overview of the various types of cloud computing platforms, as well as many of the challenges and benefits of cloud computing. In this work we focus on *Infrastructure as a Service* (IaaS) providers since they provide the greatest flexibility for enterprise users who already have large software systems that they would like to move "to the cloud" with minimal changes.

Existing IaaS platforms such as EC2 already allow customers to lease storage and computation resources on demand. While EC2 allows control over the type of CPU and storage available to each virtual machine instance, it has more limited controls over the network setup, particularly for enterprise customers looking to securely connect cloud resources to their existing infrastructure. EC2 allows for the specification of either public or internal (cloud only) network interfaces for each VM, although the precise placement and IP details are determined by the provider. EC2 also allows for Security Groups to be created which specify firewall rules for each VM. While firewalls provide very fine grain access controls, a higher level of abstraction is useful when trying to cleanly link different enterprise and cloud sites. Managing complex firewall rules as virtual machines are dynamically created and moved between sites can be very difficult. Instead, we propose that virtual private networks be used to provide not only stronger security, but additional features such as network reservation controls and seamless integration of cloud and local resources.

Two additional types of cloud computing environments have been developed to help with these concerns: private and overflow clouds [2]. In a private cloud, the customer is given exclusive access to a portion of a data center which then manages the resources using cloud computing techniques. Overflow clouds are used as backup service pools that are only used when an enterprise's own resources are completely saturated. These techniques allow for enterprises to obtain some of the benefits of cloud computing while reducing the security concerns related to using public clouds. However, the network transparency and resource control issues remain. Our work attempts to make any type of cloud more transparent, secure, and flexible.

Virtual Private Networks: In this work we focus on VPNs provided by a network operator, as opposed to technologies such as IPSec VPNs that create software tunnels between each end host. Network based VPNs are typically realized and enabled by multiprotocol label switching (MPLS) provider networks, following the "hoses model" [4] and are already commonly used by many enterprises. Provider based VPNs can provide either layer-3 VPNs following RFC 2547, or layer-2 virtual private LAN Service (VPLS) VPNs according to RFC 4761. CloudNet relies on network based VPNs since they require no endhost configuration, have lower overheads, and can provide additional services from the network provider such as resource reservation.

Related Work: In this work we focus on providing networking support for enterprise cloud platforms with VPNs. While they did not target enterprise applications, both the Virtuoso and VIOLIN projects address a similar problem of managing the network connecting virtual machines hosted across multiple grid computing sites [10, 9]. Both systems use overlay networks to intercept VM traffic and tunnel it between sites. This requires additional software to be run at each site to create the overlay network. In our work, we leverage existing technology available from network providers such as Layer 2 and 3 VPNs to not only create seamless connections between sites, but to provide greater security, and resource control as well.

3 Challenges & Insights

Cloud computing has seen widespread adoption for public web services and infrequent batch-style applications, but has not yet been accepted as a viable choice for many enterprise uses. To illustrate some of the reasons, consider an enterprise accounting application that consists of a front-end interface, a processing tier, and a back-end database. This is an application that contains private business data, and thus is traditionally run in a secure private network environment within the enterprise. Let us examine the challenges that appear when the enterprise desires to move the processing component of this application to a cloud computing data center in order to achieve greater scalability or to reduce IT costs.

Transparent Cloud Connections: Using current commercial cloud offerings, the enterprise could easily create a set of virtual machines within the cloud to run the application, but would quickly encounter difficulties when trying to link the different application components in and out of the cloud. Both the VM in the cloud and the components still in the enterprise would need to be allocated public IP addresses in order to establish connectivity. Even ignoring the security concerns in this situation, the enterprise will have to make significant changes to its own infrastructure to enable external access to formerly private resources, plus it must modify the application code to handle the new network topology, especially if the existing code assumed all components were within a LAN, for example to utilize broadcasting. This demonstrates a critical limitation of current cloud offerings: the inability to create seamless connections between cloud and enterprise resources.

Security and Isolation: Simply providing the appearance of cloud resources being attached to an enterprise's local network is insufficient (and in fact potentially dangerous) unless the network connections are made via secure channels that safely link only authorized cloud nodes to the enterprise network. To secure the processing component moved to the cloud in our example, an enterprise would need to create firewall rules both within the cloud and at the gateways to its own network in order to securely limit connectivity. While firewall rules can be used to provide fine grained access controls, they can easily expose security holes if misconfigured, and are particularly vulnerable to this when resources are frequently being added or moved as is the case in dynamic cloud computing environments. Managing the set of dynamically changing firewall rules across potentially multiple enterprise sites can be complex. Current cloud systems are ill-suited for dealing with enterprise applications because they cannot provide secure communication channels or strict network isolation in and out of the cloud.

Flexible Cloud & Network Resources: A key motivation for moving applications into the cloud is the ease with which new resources can be allocated or moved. The enterprise may have moved the processing component into the cloud so that new replicas could easily be added on demand. This action may be unreasonably difficult if it requires further reconfiguration of the application code to handle the new VM's IP address, as well as adjustments to both the cloud and enterprise firewalls to support the new network topology. Instead, cloud operators must coordinate with network providers to offer dynamic configuration of server, storage, and network resources to meet enterprise demands.

Additionally, the enterprise may be willing to pay for quality of service guarantees that ensure low latency between the front-end component hosted at the enterprise site and the processing component in the cloud. Alternatively, it may require a large amount of dedicated bandwidth on the link between the processing component and database. Existing cloud services are unable to provide these guarantees because they do not support coordination with network operators to provision resources within the network that links the cloud and enterprise sites.

4 Virtual Private Clouds

To address these challenges, we propose the idea of a Virtual Private Cloud (VPC)[1]. A VPC is a combination of cloud computing resources with a VPN infrastructure to give users the abstraction of a private set of cloud resources that are transparently and securely connected to their own infrastructure. VPCs are created by taking dynamically configurable pools of cloud resources and connecting them to enterprise sites with VPNs. Figure 1 shows a pair of VPCs connected to two different enterprises, each composed of multiple sites. A VCP can span multiple cloud data centers, but presents a unified pool of resources to the enterprise.

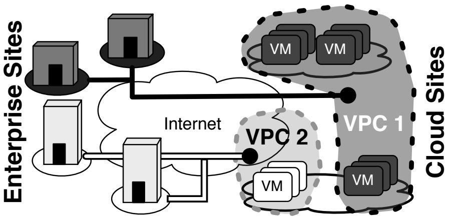

Figure 1: Two VPCs isolate resources within the cloud sites and securely link them to the enterprise networks.

VPNs can be leveraged to provide *seamless network connections* between VPCs and enterprise sites. VPNs create the abstraction of a private network and address space shared by all VPN endpoints. Since addresses are specific to a VPN, the cloud operator can allow customers to use any IP address ranges that they like without worrying about conflicts between cloud customers. The level of abstraction can be made even greater with Virtual Private LAN Services (VPLS) that bridge multiple VPN endpoints onto a single LAN segment. If the cloud provider in the previous section's example used VPCs, a VPLS could be setup so that the processing component could be easily run within the cloud without requiring any modifications since the cloud resources would appear indistinguishable from existing compute infrastructure already on the enterprise's own LAN.

VPNs are already used by many large enterprises to enable *secure any-to-any communication*, and cloud sites can be easily added as new secure endpoints within these existing networks. VPCs use VPNs to provide secure communication channels via the creation of secure, "virtually dedicated" paths within the provider network. This eliminates the need to configure complex firewall rules between the processing component in the cloud and the enterprise, since all sites would be connected via a private network inacessible from the public Internet. The VPC solution must guarantee that the secure VPN links extend to the virtual machines that compose the VPC. In order to ensure that different cloud customers are kept on isolated networks, the cloud provider must segment the LAN for each VPC. These techniques can provide strong security guarantees at a convenient level of abstraction.

VPCs enable *flexible resource control* by utilizing resource reservation mechanisms provided by VPNs. By coordinating with the network provider, the cloud service can offer enterprise users quality of service guarantees along the full path from the enterprise to the cloud site. The virtual network abstraction offered by VPNs also allows for flexibility in response to dynamic VM allocation and placement changes. The seamless LAN environment created by a VPLS service

[1] After using this term, we have since found it also used on a blog post encouraging the use of VPNs and cloud computing [5].

can be exploited to automatically handle routing changes as VMs are moved between sites, allowing them to maintain their identities on the LAN. Once two cloud sites are bridged with VPLS, existing LAN migration techniques can be used to move virtual machines across the WAN. This allows a single VPC to span multiple cloud sites, presenting a pool of geographically distributed servers as a flexible resource seamlessly attached to the enterprise's own network.

5 CloudNet: Towards Enterprise Clouds

We are developing a system called CloudNet which attempts to meet the requirements of an enterprise ready cloud computing environment using VPCs.

5.1 CloudNet Overview

CloudNet leverages existing virtualization technologies at the server, router, and network levels to create dynamic resource pools that can be transparently connected to enterprises. The CloudNet architecture is composed of two intelligent controllers that automate the management of resources in the provider network and in the cloud computing data centers respectively.

The **Cloud Manager** dynamically partitions the cloud computing data centers into Virtual Private Clouds for use by its customers. The Cloud Manager handles the creation of new virtual machines and manages performance within each VPC. The Cloud Manager utilizes several forms of virtualization so that physical resources can be multiplexed across many customers. In our current prototype, Xen is used to virtualize servers, and VLANs are used to partition the local area networks within each cloud data center. The Cloud Manager uses virtual (or "logical") routers to dynamically configure the Customer Edge (CE) routers associated with each VPC. Logical routers are a means to partition physical routers into slices, each with independent control planes. This means that full, physical routers do not need to be dedicated to each VPC, and allows them to be created and reconfigured more rapidly.

The **Network Manager** is run by the network provider and is responsible for the creation and resource provisioning of VPNs. CloudNet utilizes MPLS VPNs that span between the provider edge (PE) routers. The Network Manager dynamically configures the PE routers to create VPN endpoints associated with each VPC. The Network Manager can also be used to specify fine grain access controls that restrict which systems within a single VPN are able to communicate, or to reserve network resources along VPN paths.

Although the Network and Cloud Managers may be controlled by separate entities, communication between them is required. This is necessary to coordinate the link between the network and customer edges, as well as when a virtual machine is migrated between cloud sites.

5.2 CloudNet Usage Scenarios

This section provides further details on our CloudNet prototype implementation and some usage scenarios.

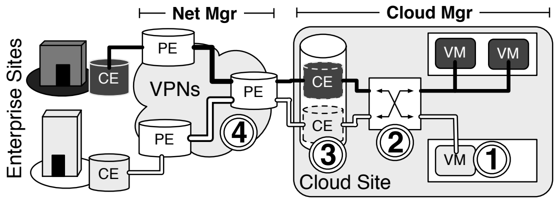

Figure 2: Adding a VM to a new site has four steps done by the Cloud Manager (1. VM creation, 2. VLAN configuration on switch, 3. logical CE creation) and the Network Manager (4. VPN setup on PE routers).

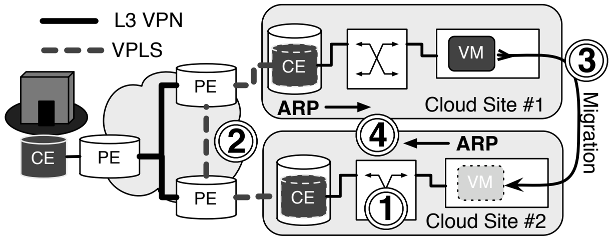

Figure 3: WAN migration using VPLS to move across sites.

5.2.1 Adding Seamless Cloud Resources

Figure 2 illustrates the process of adding a new secure VPC and transparently connecting it to a customer's VPN. First the Cloud Manager uses a placement algorithm to determine a host with sufficient spare capacity to run the VM (Fig. 2 step 1). Once the VM is created, it must be connected to a VLAN to isolate the customer's traffic within the cloud data center (Fig. 2-2). Finally, the Cloud Manager must configure a logical CE router that will be used to connect the new VPC to a VPN endpoint created by the Network Manager (Fig. 2-3). This creates an isolated partition of server and network resources within the cloud site that is dedicated to the VPC.

The Network Manager is responsible for seamlessly connecting the new VPC to the enterprise's network. It does this by defining a new VPN endpoint in the PE router connected to the cloud site. The router advertises the new VPN endpoint via BGP sessions shared with other provider routers connected to enterprise sites. If a Layer 3 VPN is used, then the VPC will be attached to the enterprise network as a new IP routed domain equivalent to any other enterprise site. Alternatively, VPLS can be used to attach the VPC to an existing LAN segment within the enterprise. In either case, the VPC network will be securely attached to the enterprise VPN so that its resources are transparently linked to the enterprise infrastructure, but inaccessible from potentially malicious users on the public Internet.

5.2.2 Cloud-to-Cloud Migration

We envision future clouds as flexible resource pools that seamlessly span multiple data center and enterprise sites. With this kind of architecture, cloud providers and enterprise customers are able to transparently migrate VMs between sites. This can allow a data center operator to perform load balancing between sites without impacting customer applications, or for customers to change placement decisions based on, for example, WAN latency.

CloudNet takes a step in that direction by simplifying the network reconfiguration for migration of virtual servers between cloud sites. Current virtualization software supports transparent VM migration between physical servers on the same LAN [3, 7], but WAN migration remains a challenge due to the need for network reconfiguration and storage migration. CloudNet exploits the benefits of VPLS to tie networks across the WAN into a single LAN, making transitions across the WAN function as if within a LAN, except for greater network delays during the migration.

Figure 3 depicts the steps to prepare for a VM migration across sites. CloudNet first performs the steps necessary to initialize the VLAN endpoint for the destination site (Fig. 3-1). In addition, a VPLS VPN is created to link together the source and destination VLANs (Fig. 3-2). At this point, the virtual machine can be migrated between the two sites (Fig. 3-3). When the VM is transferred to its new host, it will emit an unsolicited ARP message (Fig. 3-4) [3]. The local switch will use this ARP message to establish the mapping between the VM's MAC address and its switch port. The ARP message will also be forwarded through the VPLS to the VM's original site. At that site, the old switch will replace its existing MAC address mapping with the new entry, allowing data to be forwarded through the VPLS to the VM's new site. While we currently do not deal with VM storage, existing storage migration techniques have been considered for data center migration [8].

6 Conclusions

Cloud computing promises to revolutionize computing by providing cheap, flexible, on demand resources. However, current commercial cloud platforms are incapable of meeting the requirements of enterprise customers. In this work we propose the idea of Virtual Private Clouds that combine existing technologies like VPNs with automated controllers to meet three key requirements of enterprise users: 1) transparent connections between cloud and enterprise resources, 2) isolation within the cloud and secure communication channels between sites, and 3) flexible resource allocation schemes capable of responding to dynamic changes of cloud and network resources. We describe these challenges and present how our prototype system, CloudNet, can provide secure and seamless cloud resources to enterprises.

References

[1] Amazon elastic computing cloud. `http://aws.amazon.com/ec2`.

[2] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. Above the clouds: A berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley, Feb 2009.

[3] C. Clark, K. Fraser, S. Hand, J.G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield. Live migration of virtual machines. In *Proceedings of NSDI*, May 2005.

[4] N. G. Duffield, Pawan Goyal, Albert Greenberg, Partho Mishra, K. K. Ramakrishnan, and Jacobus E. Van der Merwe. Resource management with hoses: point-to-cloud services for virtual private networks. *IEEE/ACM Transactions on Networking*, 10(5), 2002.

[5] Elasticvapor blog: Virtual private cloud. `http://www.elasticvapor.com/2008/05/virtual-private-cloud-vpc.html`.

[6] Google app engine. `hthttp://code.google.com/appengine/`.

[7] Michael Nelson, Beng-Hong Lim, and Greg Hutchins. Fast transparent migration for virtual machines. In *ATEC '05: Proceedings of the annual conference on USENIX Annual Technical Conference*, 2005.

[8] K. K. Ramakrishnan, Prashant Shenoy, and Jacobus Van der Merwe. Live data center migration across wans: a robust cooperative context aware approach. In *INM '07: Proceedings of the SIGCOMM workshop on Internet network management*, 2007.

[9] P. Ruth, J. Rhee, D. Xu, R. Kennell, and S. Goasguen. Autonomic live adaptation of virtual computational environments in a multi-domain infrastructure. In *ICAC '06: Proceedings of the 2006 IEEE International Conference on Autonomic Computing*, Washington, DC, USA, 2006.

[10] Ananth I. Sundararaj and Peter A. Dinda. Towards virtual networks for virtual machine grid computing. In *VM'04: Proceedings of the 3rd conference on Virtual Machine Research And Technology Symposium*, 2004.