Network Monitoring in the age of the Cloud

Augusto Ciuffoletti

Abstract Network virtualization plays a relevant role in provisioning an Infrastructure as a Service (IaaS), implementing the fabric that interconnects virtual components. We identify the standard protocol IEEE802.1Q [1] , that describes Virtual LAN (VLAN) functionalities, as a cornerstone in this architecture.

We distinguish two aspects of virtual networking: one related to the user of the virtual connectivity, the other related to the provider that implements virtual connectivity. We describe these aspects, and put them into relation with commercial products considered typical: Amazon EC2[2] and VMware[3].

Next we devise network monitoring features that are appropriate in the user and in the provider environments. It turns out that there are significant differences between the two, and we conclude with directions for future research in the field.

1 Introduction

One of the reasons that sustain the peak of popularity reached by the *cloud computing* concept, is that it aggregates a number of extremely effective techniques into a unique abstraction, which is intuitively understandable also without a technical background.

The pieces that come to implement the concept of cloud computing are many, and come from all fields of information technology: from the point of view of the software architecture it is a descendant of Web Services, from that of processing resources it inherits from cluster computing, from the point of view of storage man-

Augusto Ciuffoletti
Dipartimento di Informatica Università di Pisa e-mail: augusto@di.unipi.it

[1] IEEE and 802 are registered trademarks of The Institute of Electrical and Electronics Engineers, Inc.

[2] Amazon Elastic Compute Cloud (Amazon EC2) [2] is a trademark of Amazon.com

[3] VMware is a trademark of VMware, Inc.

agement it takes from distributed storage architectures, to name the more evident, and we find a first crosspoint of all these technologies in the concept of *Grid computing*.

Cloud computing adds another ingredient into the melting pot: resource virtualization. The result is a concept quite effective for the company dedicated to the management of large IT infrastructures, impressive for the manager that doesn't want to invest in the volatile IT technology.

From the point of view of the infrastructure management, the implementation of a Service that offers *on demand* a virtual infrastructure means the maintenance of a unique technology throughout the whole infrastructure, with every available Mips usable to satisfy the next request. There is no resource specialization, since user needs are met when configuring the virtual infrastructure onto generic hardware.

From the point of view of the user, an Infrastructure as a Service (IaaS) provider makes available reliable, low cost resources with unlimited scalability. The know-how needed to exploit an IaaS resource is minimal.

Other aspects are less transparent: from *green* aspects related to energy savings reached optimizing resource utilization (e.g, processing units), to the technological *lock in* deriving from the dependency from a given *IaaS provider* in order to carry on a productive activity.

The same aspects that are now evident from the point of view of the computing activity, are also present from the networking point of view. In this paper we want to give a perspective of how networking issues emerge in a framework that offers IaaS.

We argue that the role played by the hypervisor in computing resources domain is here taken by the VLAN, implemented using networking facilities compliant with the ad-hoc standard IEEE802.1Q [5].

This standard regulates the functionalities offered by VLAN-aware switches, which enable the implementation of VLANs over a switched network: frames originated within the virtual LAN are transparently delivered to every other interface registered in the VLAN. Its history begins in 1998, and a revision has been released in 2005: there is an intense activity around this standard, and several substandards are being developed.

One consequence of the IEEE802.1Q protocol is the introduction of a sort of "two tiers" networking, that hardly fits into the layered ISO/OSI architecture: the Data Link layer is decomposed into two tiers, the *server* implementing an abstraction for use of the *client*.

In figure 1 we see a popular example (see Cisco white paper [3]): a network decomposed into three distinguished physical subnetworks for logistic reasons (the network spans three floors in a building), is rearranged into three VLANs reflecting distinguished offices.

Note that the level 3 router controls routing among the VLANs, and traffic in one VLAN is not visible to interfaces attached to other ones.

The effect of the introduction of VLANs in a complex network is of decoupling the needs of infrastructure management, worried by logistic issues and load balanc-

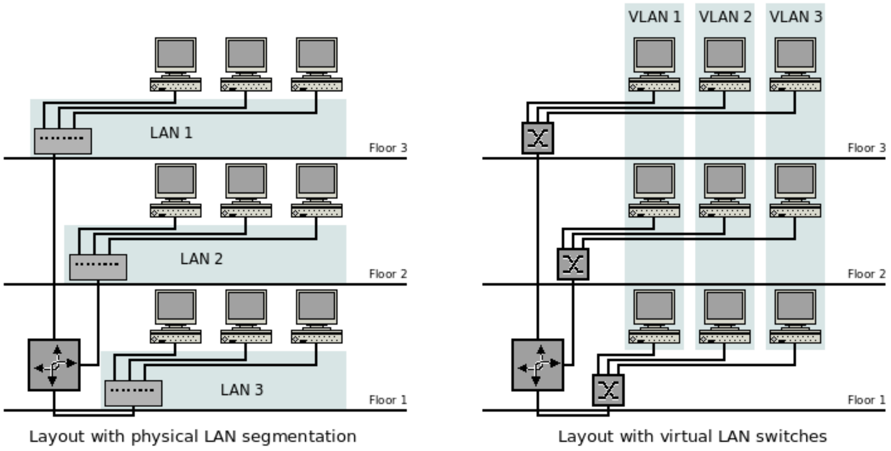
LAN 1
LAN 2
LAN 3
Floor 3
Floor 2
Floor 1
Layout with physical LAN segmentation
VLAN 1
VLAN 2
VLAN 3
Layout with virtual LAN switches

Fig. 1 Two tiers view through using LAN virtualization

ing, and of network-aware applications, that are usually happy with a flat network abstraction.

The role of network monitoring is different in the two tiers: on client's side the user is mainly concerned by the compliance to the agreed quality of service, within the infrastructure its role is to verify the expected performance of the network and help the diagnosis in case of deviation.

The next section is dedicated to an insight of the IEEE802.1Q protocol, in order to introduce some basic concepts for use. Later we proceed analyzing the kind of network abstraction offered by popular cloud providers, and finally we give an outlook of the role of network monitoring in that framework.

2 LAN virtualization

The infrastructure that supports virtual networking is a traditional, geographically extended network, that includes diverse technologies in order to adapt to different demands: distant computer centers, connected by long haul links, segmented into a number of LANs, connected by switches [3].

The technology used to implement virtual networks over a real network is based on a specific protocol, the IEEE802.1Q, and a specific device, the VLAN-aware bridge.

2.1 The IEEE802.1Q protocol

The purpose of the IEEE802.1Q protocol is to allow a network of conformant bridges, the *network infrastructure*, to emulate a number of Virtual LANs. The *network infrastructure* is composed of LAN segments: VLAN-aware bridges route MAC frames so that they are confined within the LAN segments that participate in the implementation of a given VLAN.

In figure 1 the example of a network split into three segments to accommodate logistics, hosting three distinct Virtual LANs. Although some of the advantages of LAN virtualization are not evident in that simple example, we note that:

- performance improves, since physical links that are part of a given VLAN carry only traffic on that VLAN;
- security improves, since it is impossible to interfere (e.g. sniff) traffic on a different VLAN;
- network configuration (e.g., move a host onto a different VLAN) becomes easier since VLAN reconfiguration does not require intervention on cabling.

At this point of our description, each physical link is associated to a single VLAN. To simplify the cabling, IEEE902.1Q defines bridge ports that exchange frames belonging to several VLANs, while ensuring that each VLAN is isolated from the others. This introduces the presence of *trunks* that aggregate the traffic for several VLANs.

The IEEE802.1Q protocol confines communication within one single VLAN: there is no provision for inter-VLAN routing using VLAN-aware bridges. This must be done by a level 3 router: for instance a router with one trunking interface attached to the network infrastructure supporting IEEE802.1Q is split into several virtual interfaces (or sub-interfaces in Cisco jargon) attached to distinct VLANs: packets from one VLAN to the other will cross the router.

The extra functionalities of VLAN-aware bridges are supported by an extra field added to the Ethernet frame header (see figure 2), whose presence is announced by an Ethernet Type specific for the IEEE802.1Q (`0x8100`).

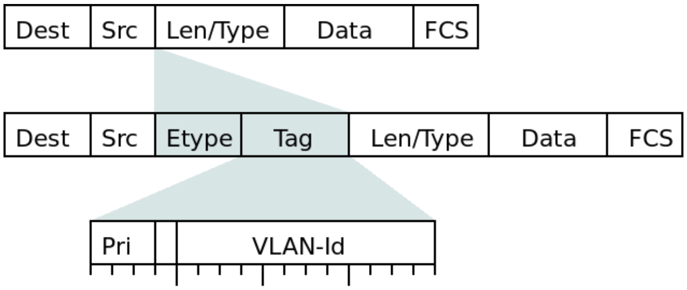

Fig. 2 Frame header with IEEE802.1Q tagging

The tag contains a subfield to indicate a priority (3 bits), and another (12 bits) to indicate the VLAN, using an identifier (or color) unique for a specific VLAN.

The Priority is used by a VLAN-aware bridge in order to select the appropriate outbound queue associated with the output port of the bridge. In its turn, each queue is associated to a given class of service, corresponding to a determined quality of service.

As a consequence, each VLAN is separately manageable, and VLAN-aware bridges are informed about the quality of service associated to a given VLAN. In particular, this can be used to differentiate expedited traffic, like VoIP, from other classes.

A LAN segment that has been selected by network management to receive frames assigned to a given VLAN is said to be a member of the VLAN. Similarly, end stations that are attached to those LAN segments and that can receive frames assigned to the VLAN are said to be attached to that VLAN.

One relevant consequence of the introduction of the IEEE802.1Q protocol is that part of the routing activity is moved from layer 3 (network) to level 2 (data link), using the VLAN identifier recorded in each frame: routing aims at multicasting the frames only within segments that belong to the destination VLAN. Routing activity in IEEE802.1 materializes in a *filtering activity*.

The information for taking the decisions needed to drive the filtering activity are contained in a database, the *Filtering Database*. In principle its content might be compiled by network administrators. In practice this task is delicate and difficult, so that a specific protocol has been designed to gather informations from manager requests, and automatically diffuse Filtering Database updates to the concerned VLAN switches. A standard specification for such protocol has been given (called Multiple VLAN Registration Protocol (MVRP) defined in IEEE802.1ak-2007), and there are also proprietary solution (Cisco[4] VTP [4]).

It is important to note that the presence of these facilities makes practically feasible the *on demand* management of a VLAN aware network.

One of the *caveats* in the configuration of the filtering database contents that implement a VLAN is the presence of loops: the packets originated at an interface attached to a VLAN should be propagated in a tree rooted in the originating interface. An algorithm that conforms to this requirement is embedded in the same standard document. The Multiple Spanning Tree (MSTP) is based on an algorithm found by R. Perlman [7], modified in order to take into account the existence of VLANs.

[4] Cisco is a trademark of Cisco Systems, Inc

3 The VLAN-cloud connection

The VLAN technology is quite powerful, and has a number of potential applications. The point for us is that it converges towards the concept of cloud computing in the IaaS sense.

In fact,

- VLAN trunking allows a single interface to serve several VLANs through virtual interfaces;
- a protocol exists to automatically reconfigure VLANs by managing the filtering databases of the involved switches;
- distinct VLANs run in isolation, without the possibility of leaking, thus ensuring an adequate level of security.

These facts allow to introduce the abstract concept of a virtual host, attached to a virtual LAN through a virtual interface: the building block of a virtual infrastructure. All this can be arranged dynamically, on demand, using the VLAN management protocol.

Here we want to note that the result is that the management of a complex infrastructure is rendered with a very simple metaphor, with the effect of making the infrastructure usable with limited background, without even knowing the name of the IEEE902.1Q protocol, reaching an extremely wide platea of users.

Although appealing, this concept exhibits a potential problem given by the limited scalability. In fact, the number of VLANs in a network is necessarily limited by the length of the VLAN-Id field: 12 bits allow not more that 4096 VLANs to be specified.

In the next section we consider two cases of commercial products that offer virtualization benefits. The two cases are quite different in nature, although they fall in the IaaS category: their study is a way to understand different approaches to virtualization from the point of view of networking, and thus give the basis for our discussion of network monitoring issues.

Amazon EC2 offers a service to the end user wishing to exploit the IaaS technology for its own purpose. VMware is capable of implementing the IaaS service using available technologies, including VLANs. We briefly summarize their characteristics, before proceeding in our discussion on network monitoring.

3.1 Amazon EC2

EC2 gives a minimal control over networking issues, which are almost completely hidden from the user. The user is provided with one pair of IP addresses assigned to the Virtual Machine (VM) when it is created: one accessible only from inside the cloud, the other accessible from outside. The user has no control on how these IPs are generated. In addition, the client may reserve a few additional IP numbers exposed to the Internet.

The user has some possibility to indicate the logistics of a given VM. With this, Amazon wants to meet the practical needs of an enterprise wishing to delocalize computational activities: legal issues concerning the place where processing takes place, and reliability concerns, related to avoiding the loss of data as a consequence of a single failure.

The two aspects are coped with using two distinct abstractions: the *region*, which specifies the geographical region where the computation will take place, and the *availability zone*, that allows to allocate instances so to minimize the possibility that a single failure hits more than one instance.

The *security group* may indirectly serve to implement a sort of VLAN, intended as a set of VMs sharing the same reachability constraints. There are no guarantees of efficiency in the communication, and such tool is primarily intended to simplify the management of security issues.

We note that such abstractions are mostly oriented to a solid and specific market: 3-tiers web servers. Offering a restricted number of functionalities the resulting interface is easy to use, and hides most of the complexities inherent to provisioning an IaaS.

Here we consider EC2 as the representative of a larger class of products with similar characteristics: from our point of view, they share an opaque approach to networking. The service provider may make efforts in order to optimize network utilization, but this is totally out of control from the point of view of the client.

The Open Grid Forum OCCI Working Group [1] is currently pursuing the standardization of an interface for Cloud Computing. The concepts reflected in the interface, for what is concerning networking, are quite similar to those implemented by EC2.

3.2 VMware

VMware offers a quite complete set of tools oriented to exploiting various virtualization technologies; when we focus on networking, we see that the VLAN concepts descending from the standard IEEE802.1Q are easily integrated in VMware infrastructures [10].

Both *virtual adapters* and *virtual switches* [8] are present as abstractions in the toolset. Trunking between ports on virtual switches and adapters is supported as well. However, VMware introduces a notable limitation: within a single host there is no possibility to interconnect two virtual switches.

This option is justified by the improved reliability obtained by forcing the network into a flat structure; on the other hand, complex hierarchical structures are mainly justified by logistics, but a virtualized environment removes such kind of concerns.

In essence, a Virtual Switch can be connected to a number of virtual hosts arranged into distinct VLANs within the server, and to physical adapters or switches outside the server. The typical networking internal to a server is depicted in figure 3.

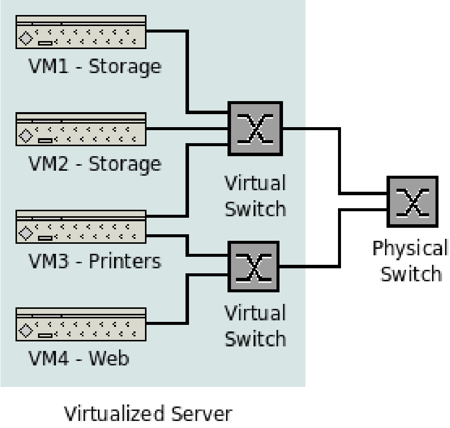

Fig. 3 Typical internal networking of a VMware server

This option makes useless running the cumbersome spanning tree protocol inside the server: the single-tier structure enforced by VMware simply does not allow the introduction of loops, whose avoidance is the main reason for the existence of the spanning tree protocol.

The flexibility inherent to a virtual infrastructure allows the introduction of some limits to the utilization of the traffic trunking technology: these restrictions contribute to a more efficient operation of the server. The recommended organization envisions virtual switches exposing trunking links to the outside (uplinks, in VMware terminology), and untagged communications on ports directed to virtualized hosts within the VMware server.

Overall, the VMware framework allows the exploitation of the VLAN advantage in a virtualized environment: for instance, a virtual server can be moved from one VLAN to another just by reconfiguring switches.

The perspective of a cloud computing facility offering elastic services is envisioned as one of the potential applications, using the *Lab Manager* application. Within this framework the possibility of managing network configuration is retained.

4 Network Monitoring in the age of the Cloud

In order to find the concepts that should guide the design of an effective network monitoring activity in an environment that makes use of virtualization techniques, we need to identify which kind of data do we need to obtain from this activity. We discover that they are different, depending on the layer where the monitored activity takes place: we distinguish, and examine separately, the user layer and the cloud infrastructure layer.

Further, network monitoring activity may be directed to fulfill two distinct purposes: on one side, to detect networking problems and thus improve fault tolerance, on the other to optimize network utilization. Here we mostly focus on the second alternative: the typical scenario is a network intensive application that has alternative ways to carry out its activity. Network availability may bias the decision process.

4.1 Network monitoring on the user side

The user is typically presented with an unstructured set of processing units with no clues about the network infrastructure that binds them together, as seen inspecting the EC2 framework. We may envision two scenarios for the requirements of an end client to a network monitoring infrastructure:

- verify the conformance of the provided service with respect to the Quality of Service (QoS) or for accounting;
- optimize its operation depending on network performance.

We note that the Cloud Service provider may dynamically interfere with both aspects, as seen studying the VMware toolset.

The conformance to QoS is accommodated using passive monitoring tools that inspect traffic across virtual interfaces: being implemented in software, a tempting idea is that such interfaces might be easily instrumented with code used to extract traffic patterns and characteristics. However, the implementation of virtual interfaces turns out to be rather out of control from the point of view of the user, who should rely on features implemented by the IaaS provider.

Network Monitoring tools running in the user space appear to be more appropriate for the task: for instance, inside a virtual server used for load balancing purposes of a number of virtual data servers. In this case, traffic and connections monitoring can be used, for instance, to measure the data transfer rate within the cloud between data servers and the load balancer, or to bill the user accounts according to the quantity of data transferred.

One relevant aspect is that, since one of the major benefits of the cloud computing concept is dynamic adaptation, network monitoring configuration must be dynamic as well. For instance, in the data server example above, when one data server is added or removed to respond to load changes, the network monitoring activity must be adapted accordingly.

We observe that such adaptation should be controlled by an application running in the user space of the load balancer: this enforces the conclusion that network monitoring application should be resident in the user space, and controlled by the user application.

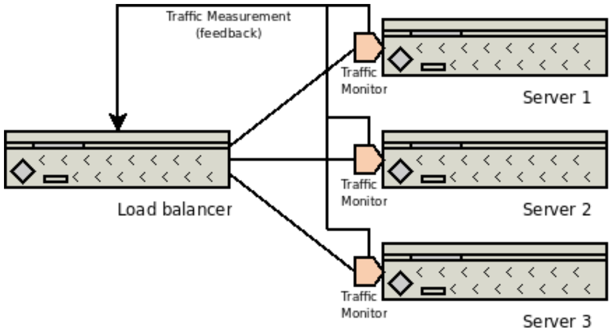

Fig. 4 Closed loop load balancing with user level network monitoring

A consequence of the discussion above is that network monitoring may sit in a critical place concerning resource management of an application that makes use of IaaS, since it implements the feedback control, as seen in figure 4: from control theory, we know that a badly controlled feedback may make the whole system unstable.

One aspect that is not very relevant for the end user is testing liveness and reachability of virtual machines: the probability of host unreachability is drastically reduced by the service provider, which manages the underlying resources and may even relocate the virtual server in case of failure of its physical support.

4.2 Network monitoring on the service provider side

The service provider needs are more composite than those of the end user. We distinguish two aspects: one related to the underlying virtualization infrastructure, the other making reference to the provided service.

As mentioned in the previous section, one of the major tasks of the service provider is to guarantee that VMs are properly working and connected: virtualization infrastructures need to provide tools for such purpose.

For instance, VMware introduces the concept of *beacons* in order to verify the state of health of the virtual infrastructure.

Beacons technique is based on sending packets outside physical servers. In this way the unavailability of uplinks is detected. When coupled with Network Interface Card (NIC) teaming (a materialization of standard IEEE802.1AX [6]) we obtain a substantial increment of reliability of the networking infrastructure: *teams* of links are operated to implement a single virtual link, performing load balancing and excluding a link when diagnosed as failed.

However, pure connectivity does not exhaust the needs of a cloud computing infrastructure. The performance of the networking infrastructure is fundamental in the operation of the cloud both from the point of view of the user of the cloud infrastructure, and from the perspective of the management.

From the point of view of the user, the virtual network performance should be consistent with the agreed QoS. Here the task of network management is to enforce that such constraints are respected: this entails a selective monitoring of the connectivity services offered to a given client.

The task is addressed using end-to-end monitoring between the servers where client instances are running: packet filters are applied in order to select the traffic relevant for the specific client application, and to measure the related network performance.

We note that traffic between physical servers is probably trunked and link aggregation is also used: network monitoring may become difficult in a generic intermediate point. Therefore we conclude that network monitoring should be operated within the virtual environment (e.g., within the server) probably using passive techniques implemented, for instance, within virtual switches. Alternatives are viable, since the virtualization infrastructure may provide a promiscuous mode for switches.

These measurements might be offered as additional services to the user wanting to optimize its computation.

From the point of view of the traffic associated to the management activity, one primary concern is Virtual Machine Image (VMI) displacement: this piece of data amounts to several GBytes, and each VM instantiation entails the displacement of the corresponding VMI from a repository to the server hosting the VM. Commonly used VMIs may be disseminated throughout the Cloud infrastructure, but customized ones must not be too far from the server where the corresponding VM is instantiated. The optimization problem exhibits several degrees of freedom, and its solution is not straightforward. However VMI displacement is the kind of activity that is planned in advance, with little help from network monitoring results.

The event of the displacement of a running VM may be envisioned for fault tolerance, as well as for extreme performance reasons. This event cannot be anticipated, so its execution should be evaluated also considering instantaneous bandwidth availability. Such measurements are extremely difficult to synthesize from a link level view of the traffic: an active end-to-end measurement of residual network availability seems to be a simpler solution for this case. We note that this kind of solution may hinder or make less effective traffic engineering solutions operated on the basis of more predictable traffic.

Looking at the other aspect of network management, the compliance to the QoS agreed with user, we focus on especially demanding applications: those that require audio and video transfer.

The team working around IEEE802.1 is deeply concerned with audio/video streaming over bridged architectures: such interest materializes in a task group specific for this purpose. The activity of the group covers the identification of the components of such a network, the transport of timing information [9], and end-to-end resource reservation.

Network monitoring plays a relevant role in media streaming: as a general rule, the application itself collects statistics in order to optimize buffering or for other reasons related to the real time nature of the stream. The presence of resource reservation protocols should restrict network monitoring to fault tolerance and accounting purposes.

5 Conclusions

Cloud computing is a concept that arises from the aggregation of new and effective techniques. Among these techniques there is a new way for network management, centered around the Virtual Local Area Network technology.

The application of this technology drastically changes the basics of network monitoring design, since tools running at level 3 in the ISO/OSI stack have a picture of the network that does not correspond to the real network, but to a synthetic abstraction.

Network level 2 implements sophisticated traffic engineering techniques:

- trunking allows a single link to support distinct virtual links, managed individually, preventing leaking with high security;
- teaming allows several links to cooperate to the implementation of a single highly reliable link;
- traffic classification allows provisioning several quality of service over the same link.

The consequence is that network phenomena at level 3 are quite different, overall exhibiting a reliable operation, with predictable performance.

The purpose of network monitoring changes accordingly. The user may want to use it primarily for accounting and secondarily to optimize, while fault detection and removal are uninteresting, since the infrastructure provides highly reliable links, and failures are beyond reach when they occur. The service provider needs data that report the dynamic evolution of the link level in order to drive mapping the VM instances requested by users to the servers, and to operate fault tolerance.

We have deliberately disregarded the utilization of network monitoring techniques to enforce security: we considered that this utilization falls outside the scope of this paper, that mostly addresses traffic control. However the potential security of VLAN based architectures must be protected against certain attacks, e.g. VLAN-hopping.

5.1 Looking ahead

Here we envision future directions for research on network monitoring, motivated by the above arguments:

- defining link level network monitoring features that can be embedded into virtual switches and virtual network adapters. The user may take advantage of statistics collected from within the NIC. Using virtual adapters, this feature does not need expensive specialized hardware, but the implementation of such features in the software that implements virtual adapters. Questions arise concerning how to control such features, how to implement them in a efficient way, and how to manage the data produced.

- embedding network monitoring features within developing standards. The IEEE802.1 working groups are currently active in the definition of new standards for VLANs: the inclusion of native network monitoring features within the protocol itself may reduce the overhead introduced to perform measurements;
- studying user level network monitoring infrastructures that dynamically deploy in infrastructures configured on demand. A network monitoring infrastructure may become quite complex, and its management may become a problem for itself. A network monitoring infrastructure that is automatically deployed while new VMs are instantiated may be of interest for those users that want to optimize the utilization of the virtualized infrastructure.

References

1. Open cloud computing interface (OCCI) WG charter, March 2009.
2. Amazon Web Services LLC. *Amazon Elastic Compute Cloud - Getting Started Guide*, 2009.
3. CISCO Systems. *Overview of Routing between Virtual LANs*.
4. CISCO Systems. *Understanding VLAN Trunk Protocol (VTP)*.
5. IEEE Computer Society. *IEEE Standard for Local and metropolitan area networks - Virtual Bridged Local Area Networks*, 2005.
6. IEEE Computer Society. *IEEE Standard for Local and metropolitan area networks - Link Aggregation*, 2008.
7. Radia Perlman. An algorithm for distributed computation of a spanningtree in an extended LAN. *SIGCOMM Comput. Commun. Rev.*, 15(4):44–53, 1985.
8. Jeremy Sugerman, Ayalvadi J. Venkitachalam, Ganesh, and Beng-Hong Lim. Virtualizing I/O devices on VMware workstation's hosted virtual machine monitor. In *USENIX Annual Technical Conference*, page 14, Boston, june 2001.
9. Michael D. Johas Teener and Geoffrey M. Garner. Overview and timing performance of IEEE 802.1AS. In *International IEEE Symposium on Precision Clock Synchronization for Measurement, Control and Communication (ISPCS)*, pages 22–26, Ann Arbor (MI), September 2008.
10. VMware. *VMware Virtual Networking Concepts*.

Acronyms

VM Virtual Machine
VMI Virtual Machine Image
VLAN Virtual LAN
QoS Quality of Service
IaaS Infrastructure as a Service

MVRP Multiple VLAN Registration Protocol
NIC Network Interface Card