Apache CloudStack

Orquestrando a sua nuvem IaaS

Marcelo Lima

GTER 40
GTS 26
10 E 11 DE DEZEMBRO

Sobre ...

- Marcelo Lima marcelo.lima@shapeblue.com
  - Arquiteto de Cloud & Data Center da ShapeBlue Brasil
  - Especializado em...
    - Design & Building de Clouds baseadas em Apache CloudStack / Citrix CloudPlatform
    - Infraestruturas de Data Center
    - Automação de Infraestrutura
    - Ambientes de alta disponibilidade e escalabilidade
  - Certificações
    - Apache CloudStack Certified Professional
    - Citrix Certified Professional – Networking (CCP-N)
    - F5 System Engineer – LTM/GTM/ASM
    - Cisco Certified Internetwork Expert – Data Center (CCIE-DC Written)
    - Cisco Certified Network Professional – Data Center (CCNP-DC)
    - Cisco Certified Network Associate – Data Center (CCNA-DC)
    - Cisco Data Center Unified Computing Support Specialist
    - Cisco Data Center Unified Computing Design Specialist
    - Cisco Data Center Unified Fabric Support Specialist
    - Cisco Data Center Unified Fabric Design Specialist

Sobre a ShapeBlue

"Somos um time de arquitetos e engenheiros especialistas em construção de infraestruturas para nuvens IaaS tanto públicas quanto privadas.

Somos líderes globais em consultoria e integração do Apache CloudStack"

Sobre a ShapeBlue

Principais clientes

apachecloudstack
open source cloud computing

Ementa da apresentação

- O que é o Apache CloudStack?
- Arquitetura
- Tipos de redes
  - Básica
  - Avançada com Grupos de Segurança
  - Avançada (Compartilhada, Isolada, VPC e VPC roteado)
- SDN no Apache CloudStack
- Autenticação/Cloud federada

GTER 40
GTS 26
10 E 11 DE DEZEMBRO

O que o Apache CloudStack?

O que é o CloudStack

- Plataforma de orquestração de infraestrutura, multiusuário e segura (multi-tenant)
  - Plataforma confiável para entrega de Nuvem IaaS
  - Agnóstico quanto ao hypervisor, ou seja, independente do software de virtualização utilizado
  - Escalável, flexível e seguro
  - Código aberto, padrões abertos
  - Implementação privada (no próprio Data Center) ou como uma solução hospedada

Plataforma aberta flexível

GTER 40 GTS 26
10 E 11 DE DEZEMBRO

Compute

XenServer | Vmware (vCenter) | KVM | LXC | Hyper-V | Xen Project | Bare metal (IPMI)

Storage

Disco local | iSCSI | FC / FCoE | NFS | CIFS/SMB (Hyper-V) | CEPH GlusterFS (KVM) | NFS | CIFS/SMB (Hyper-V) | Swift | S3

Storage Primário | Storage Secundário

Rede

Tipo de rede | Isolamento | Firewall | SLB/GSLB | SDN | VPN

ShapeBlue The Cloud Specialists

apachecloudstack
open source cloud computing

Provedor de Serviços de Rede

- Um *appliance* físico ou virtual que forneça serviços de rede para o CloudStack, como por exemplo:

- Virtual Router
- VPC Virtual Router
- Internal LBVM
- Citrix NetScaler
- F5 BigIP - LTM
- Juniper SRX Firewall
- Palo Alto
- Cisco VNMC/ASA 1000v
- MidoNet
- BigSwitch Vns
- Nuage VSP
- VMware NSX
- Juniper Contrail
- Ovs
- VXLAN
- OpenDaylight (experimental)

O que é possível fazer com o CloudStack?

- *Criar máquinas virtuais a partir de Templates (modelos) ou imagens ISO*
- *Iniciar e desligar máquinas virtuais*
- *Criar redes isoladas, compartilhadas e multicamadas (multi-tiered)*
- *Gerenciar regras de firewall e de redirecionamento de portas (port forwarding)*
- *Gerenciar serviços de rede como Load Balancing, Static e Source NAT, VPNs, Global Load Balancing e Autoscaling (escalonamento automático)*

GTER 40
GTS 26
10 E 11 DE DEZEMBRO

Arquitetura

Arquitetura do CloudStack

- Sua estrutura hierárquica possibilita escalar massivamente
  - Região
    - Um grupo de zonas de disponibilidade dentro de uma mesma área geográfica
    - Necessário um servidor de gerenciamento do CloudStack para gerenciar cada região
  - Zona de disponibilidade
    - Tipicamente uma ou mais Zonas por Data Center
    - Contém pelo menos um Pod, um Cluster e uma unidade de Storage Secundário

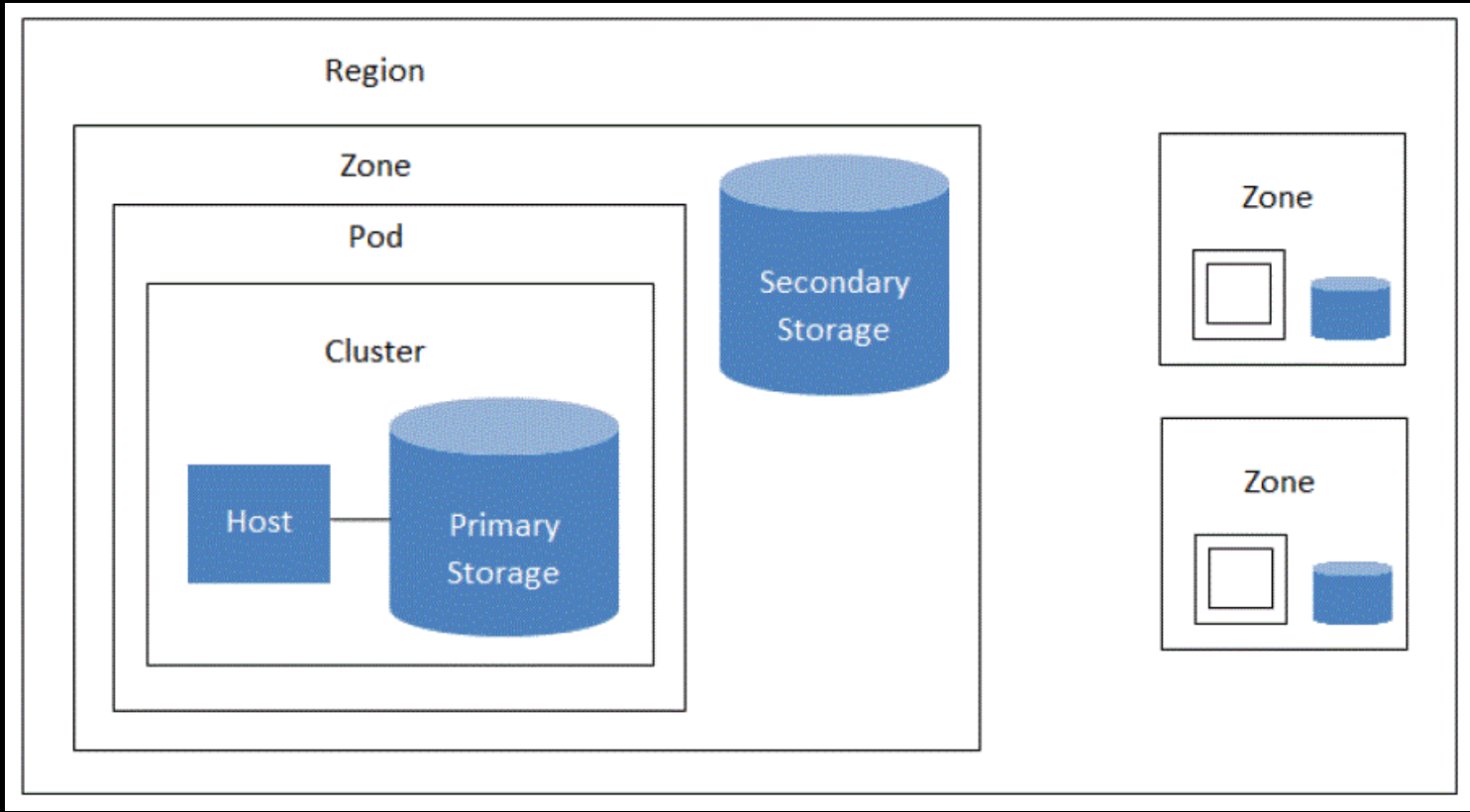
Arquitetura do CloudStack

- Pod
  - Entidade lógica, normalmente, um *rack* contendo um ou mais Clusters e infraestrutura de rede
- Cluster
  - Grupo de servidores (*Hosts*) idênticos executando o mesmo software de virtualização (*Hypervisor*)
  - *Storage* Primário

Arquitetura do CloudStack

- Storage Primário
  - Tradicionalmente único para cada Cluster
  - Tanto KVM quanto VMware suportam *Storage* Primário *Zone-Wide*
  - Hospeda os discos das Instâncias de VMs e *Snapshots das Maquinas Virtuais*
  - Podem ser em qualquer formato que o software de virtualização suportar
- Storage Secundário
  - *Zone-Wide (Region-Wide utilizando S3)*
  - NFS + S3 ou NFS + Swift para replicação dentro de uma mesma região (*Region Wide Replication*)
  - Armazena *Templates*, imagens ISO e *Snapshots* de volumes/discos (*backups*)

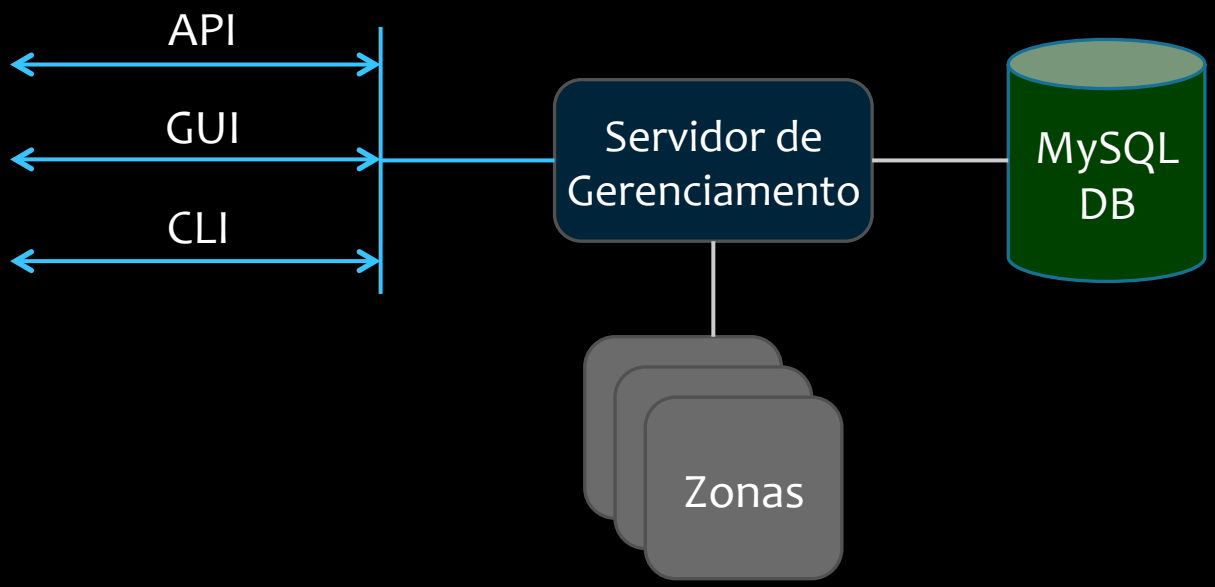
Arquitetura do CloudStack



Arquitetura de implantação dos Servidores de Gerenciamento

GTER 40
GTS 26
10 E 11 DE DEZEMBRO

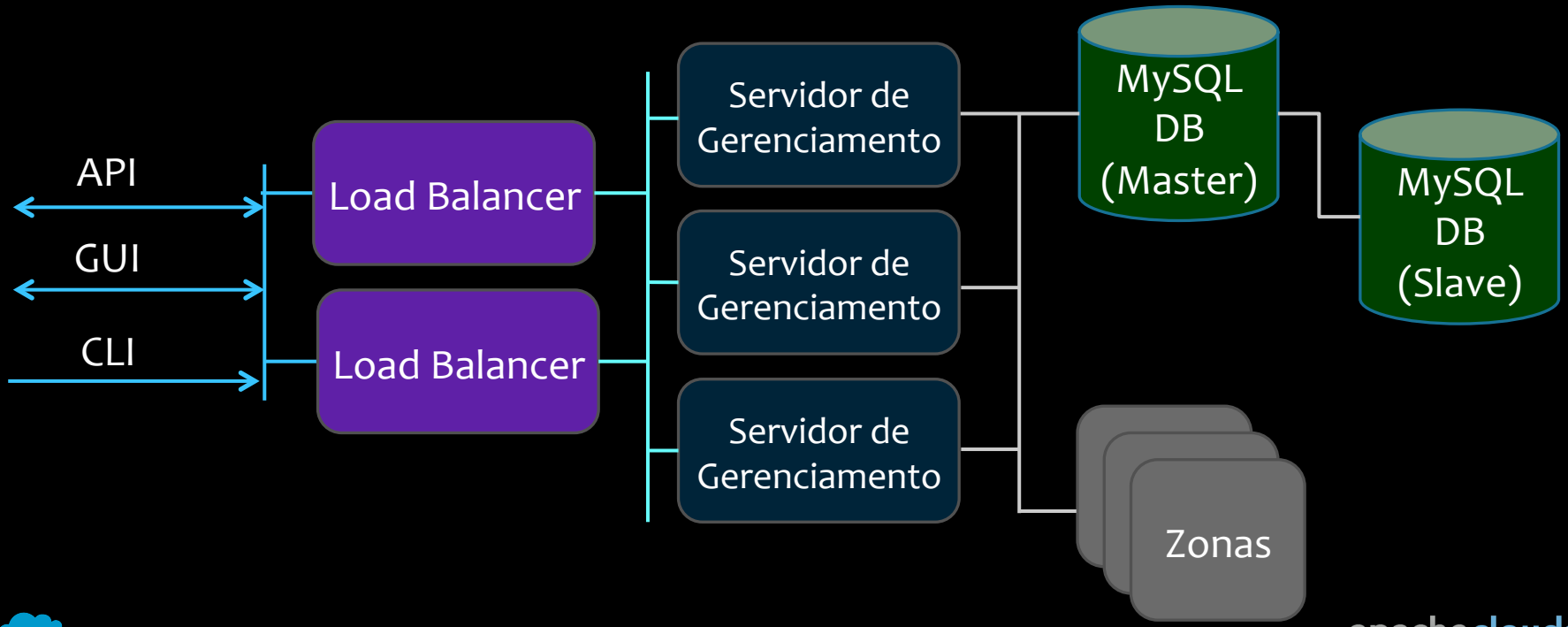
Exemplo de implantação 'Single-Node'



Arquitetura de implantação de Servidores de Gerenciamento

GTER 40 GTS 26
10 E 11 DE DEZEMBRO

Exemplo de implantação 'Multi-Node'

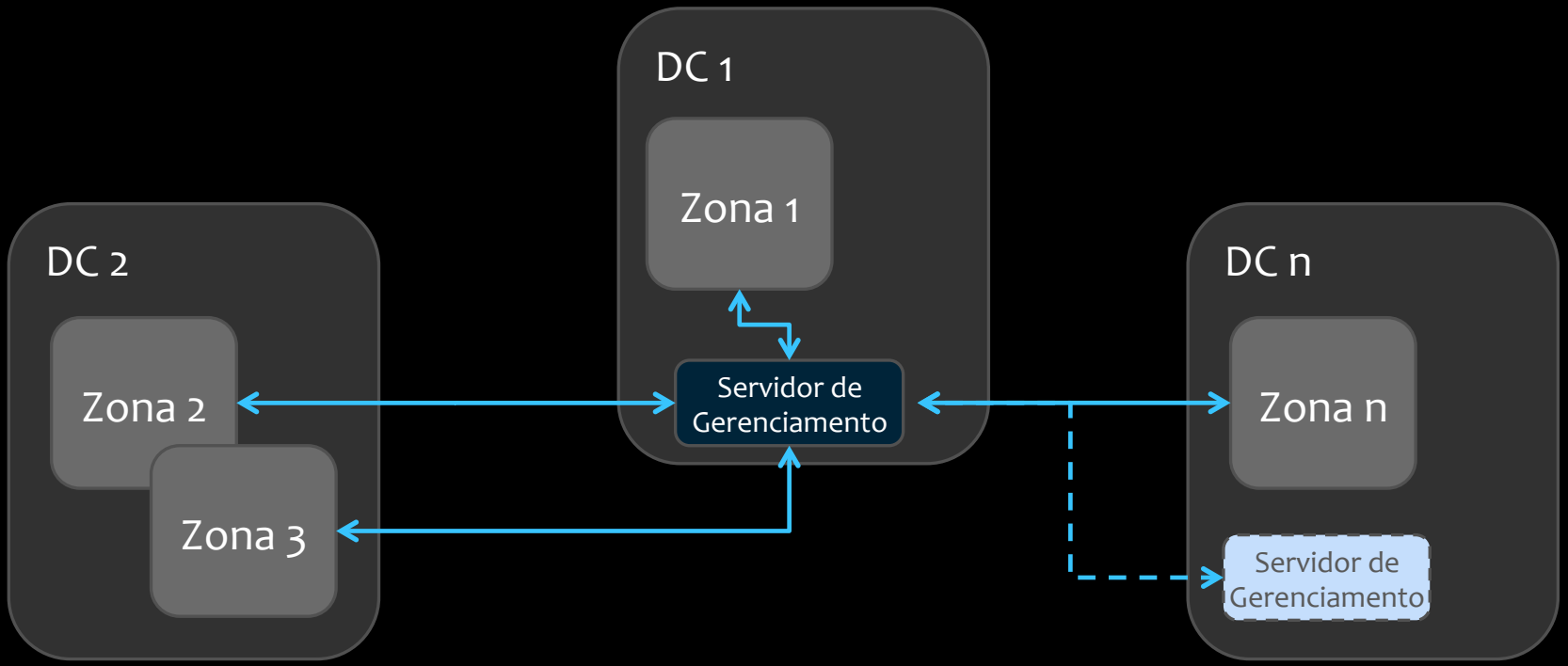


ShapeBlue The Cloud Specialists

apachecloudstack
open source cloud computing

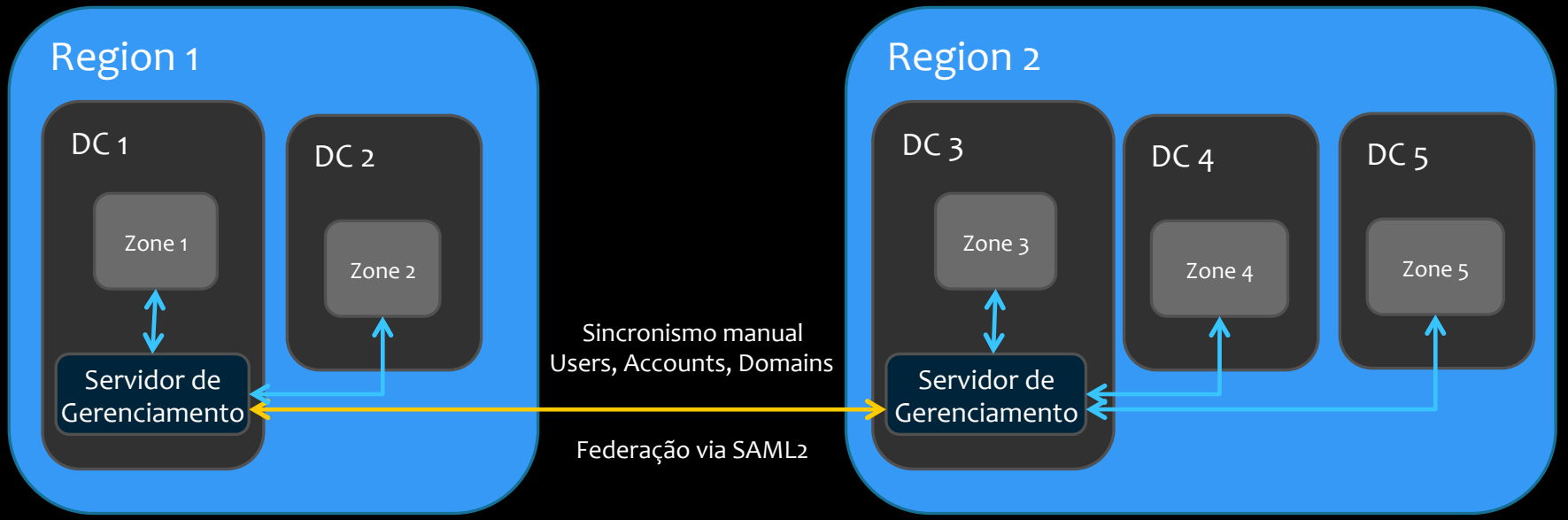
Arquitetura do CloudStack

Exemplo de múltiplas Zonas de Disponibilidade dentro de uma região



Arquitetura do CloudStack

Exemplo de múltiplas Zonas de Disponibilidade dentro de uma região



GTER 40
GTS 26
10 E 11 DE DEZEMBRO

Tipos de rede no CloudStack

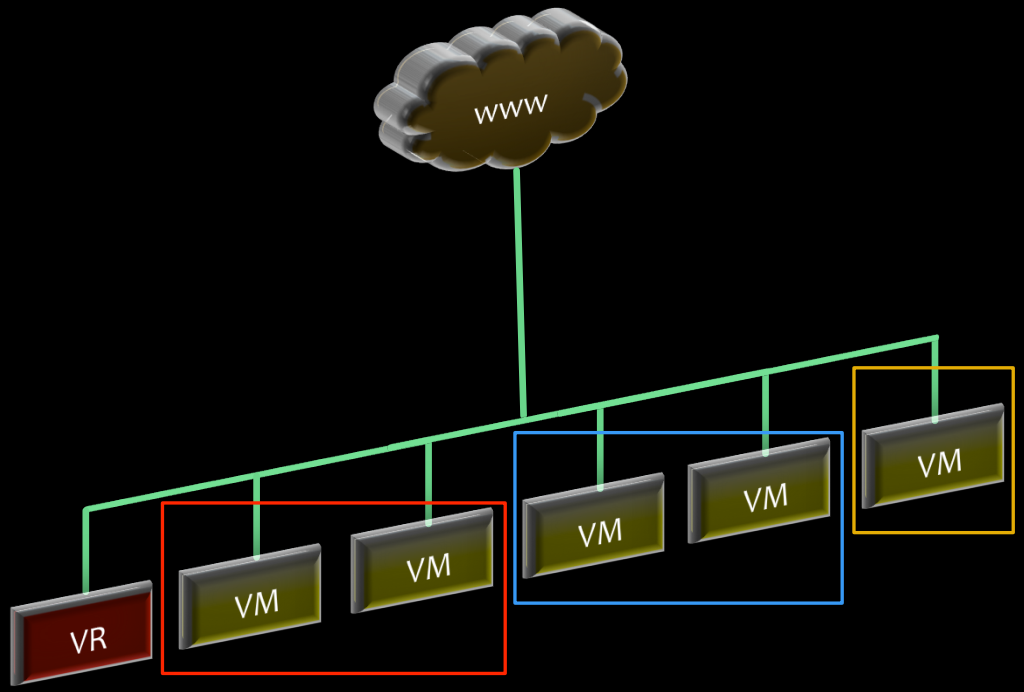
Tipos de rede no CloudStack

- Zonas
  - Zona Básica – Compartilhada com Grupos de Segurança
  - Zona Avançada – Grupos de Segurança
  - Zona Avançada – Compartilhada, Isolada, VPC

Zona Básica

- Zona Básica – Compartilhada com Grupos de Segurança
  - Fornece uma única rede onde o isolamento dos hóspedes pode ser fornecido através da camada 3 (IP) utilizando grupos de segurança (filtragem por IP de origem)
  - Uma única rede Flat

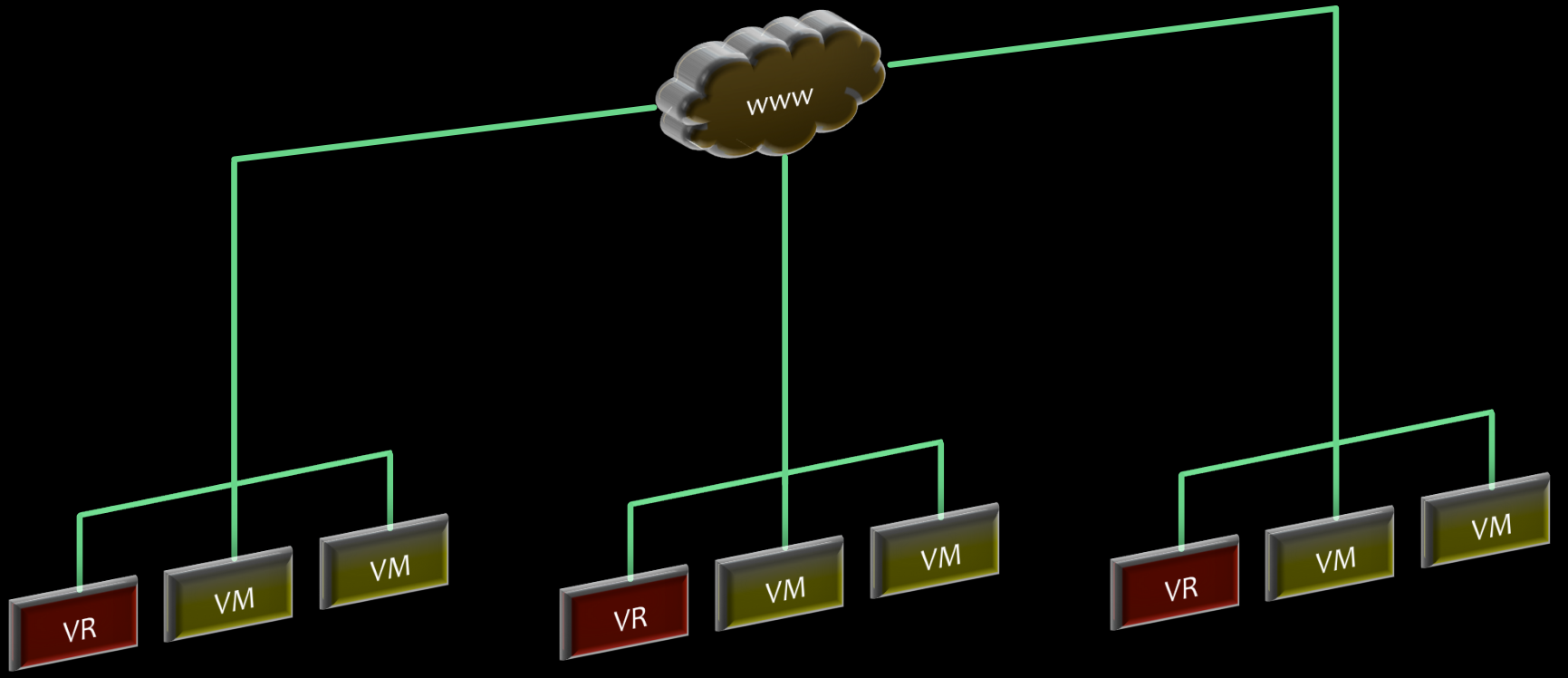
Rede ‘Compartilhada’ – Zona Básica



Zona Avançada com SG

- Zona Avançada – Compartilhada com Grupo de Segurança
  - Semelhante a zona Básica porém fornece mais de rede onde o isolamento dos hóspedes pode ser fornecido através da camada 3 (IP) utilizando grupos de segurança (filtragem por IP de origem)
  - Mais de uma rede Flat

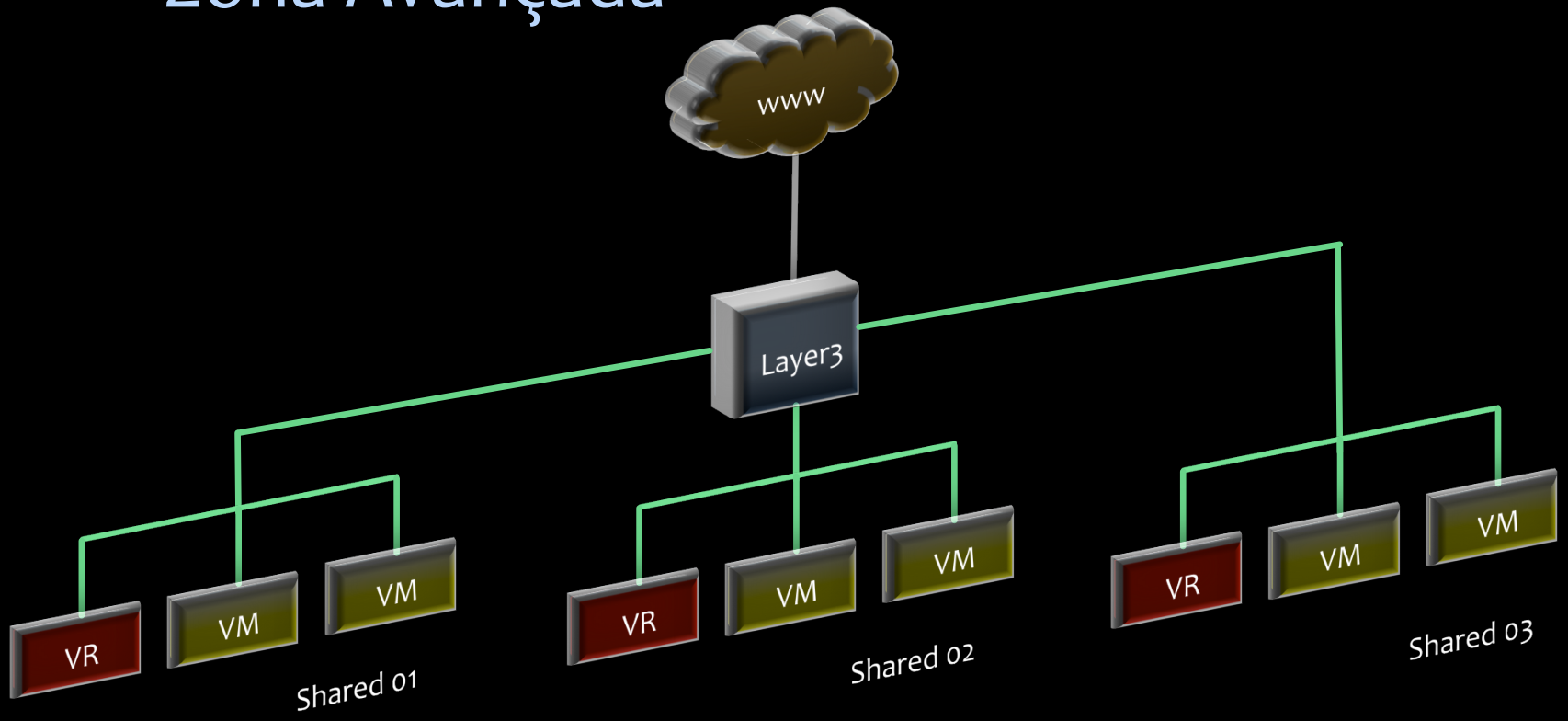
Zona Avançada com SG



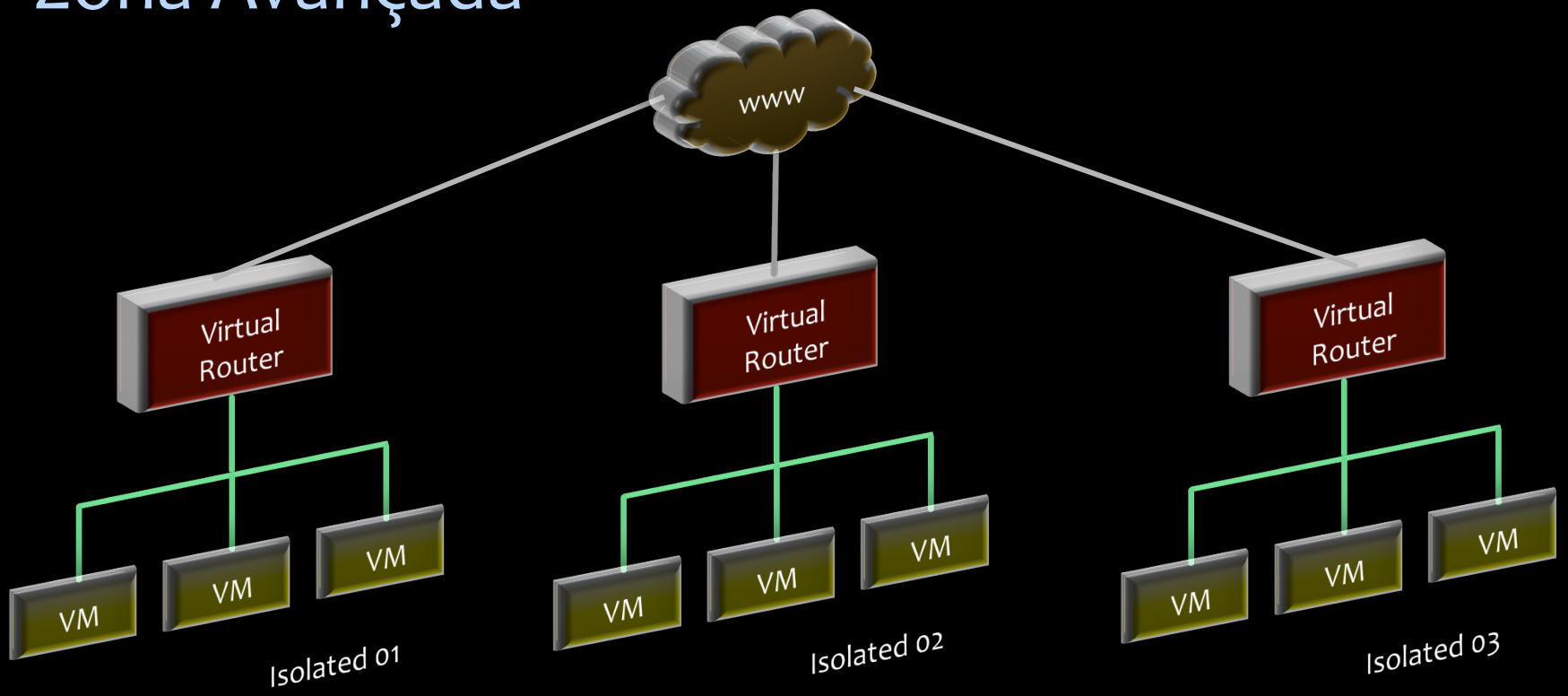
Zona Avançada

- Zona Avançada – Compartilhada, Isolada ou VPC
  - Esse modelo de rede fornece maior flexibilidade na definição de redes Guest e fornece ofertas/serviços de rede personalizadas, tais como firewall, VPN, *Load Balancer* e recursos de VPC. O isolamento dos Guests é realizado através da camada-2 (Ethernet) usando VLANs, por exemplo, VLANs ou SDN.

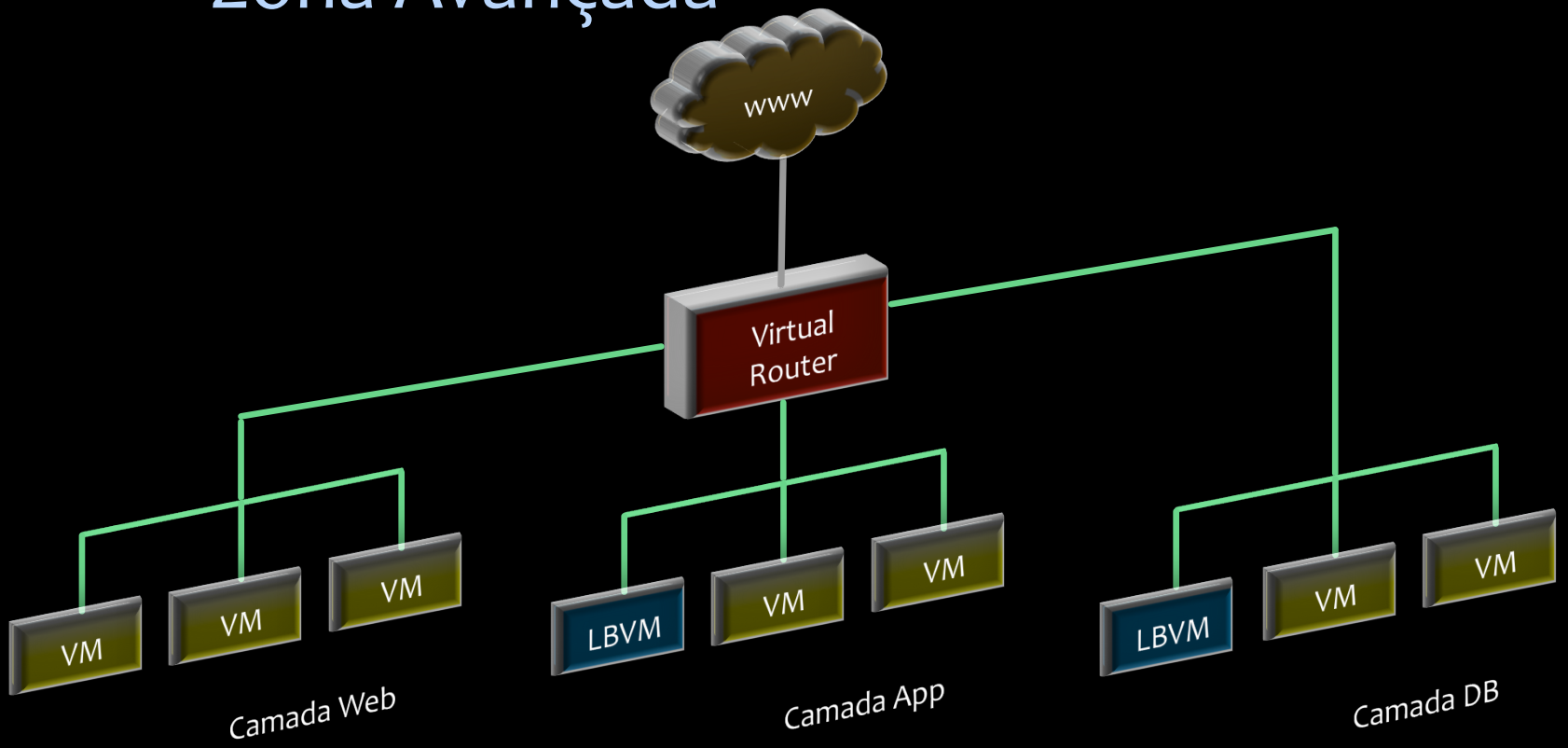
Rede ‘Guest Compartilhada’ Zona Avançada



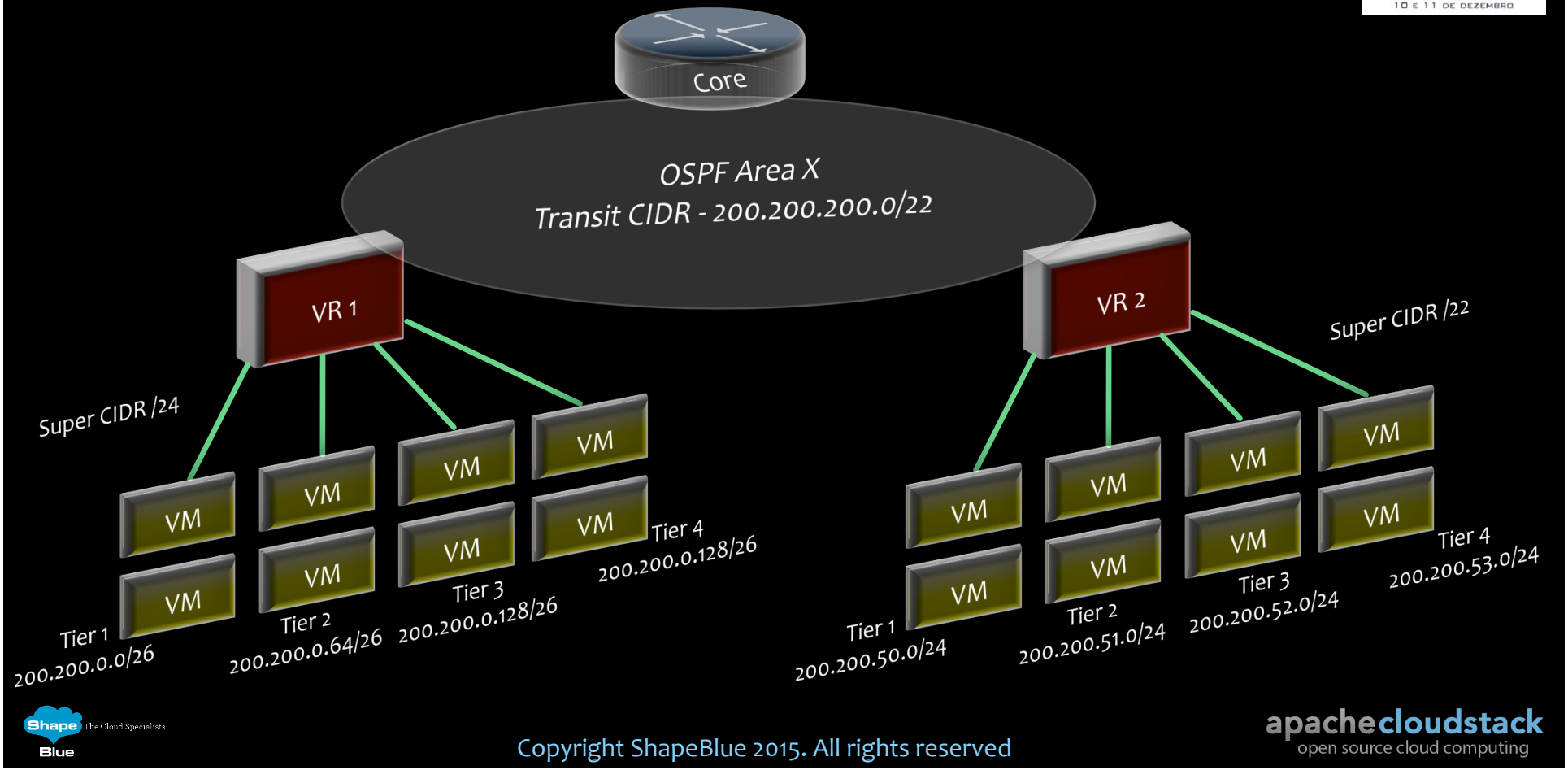
Rede ‘Guest Isolada’ Zona Avançada



Rede 'VPC – Virtual Private Cloud'
Zona Avançada



Rede 'VPC – Roteado/IPv6' Roadmap



SDN no CloudStack

SDN no CloudStack

- O limitador – 802.1q (Vlan)
  - Limita o isolamento entre clientes/usuários ao número de vlans suportado pelos switches L2 utilizados na infraestrutura. No melhor dos casos 4094 redes.
  - Extensão de vlans entre Data Centers pode ser um problema.

SDN no CloudStack

- SDN em Cloud Computing
  - Não podemos limitar SDN apenas em gerenciamento de fluxos. SDN em cloud não se limita apenas em OpenFlow.
  - Serviços gerenciados via software, entre eles Firewall (filtros e nat), Load Balancing (slb e gslb), VPN (c2s, s2s), AutoScaling, OpenFlow, VLAN, VXLAN, NVGRE, STT, GRE, ODL, etc.

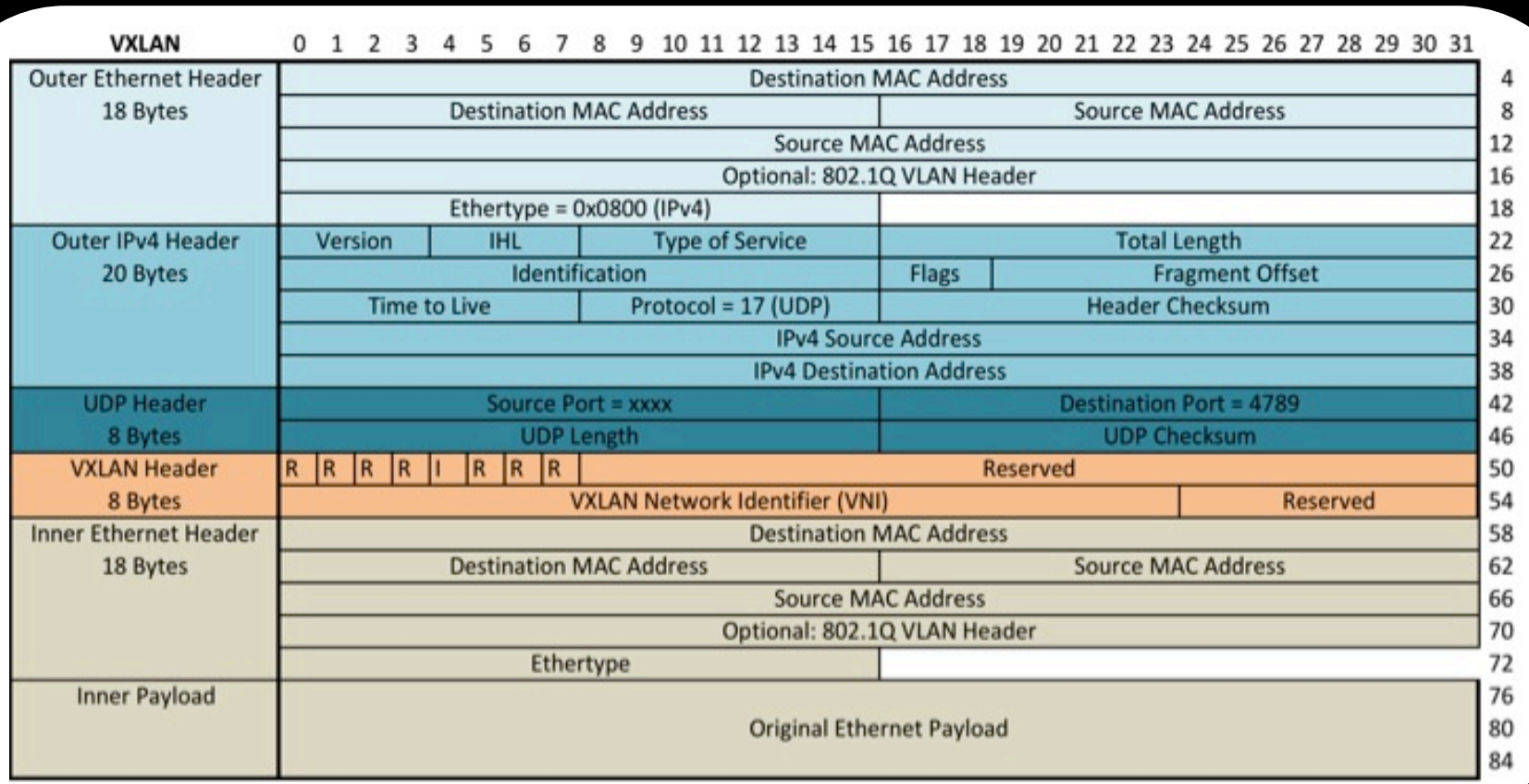
SDN no CloudStack

- As alternativas
  - VXLAN – Virtual eXtensible Local Area Network
  - GRE Ovs – Generic Routing Encapulation over Open vSwitch
  - STT – Stateless Transport Tunnel

SDN no CloudStack

- VXLAN - Virtual eXtensible Local Area Network
  - O frame L2 original é encapsulado dentro de pacote UDP.
  - O switch não precisa aprender o mac de cada VM, consequentemente reduz o tamanho da tabela mac dos switches.
  - Necessita roteamento multicast caso haja equipamentos L3 entre VTEPs.
  - Não necessita Gateway VXLAN pois é somente para tráfego Guest.
  - Altamente escalável, 2^24 (16M).
  - VXLANs podem ser extendidas entre Data Centers via L3.
  - + 54 bytes por pacote.

SDN no CloudStack



VXLAN	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	
Outer Ethernet Header 18 Bytes	Destination MAC Address	4
	Destination MAC Address / Source MAC Address	8
	Source MAC Address	12
	Optional: 802.1Q VLAN Header	16
	Ethertype = 0x0800 (IPv4)	18
Outer IPv4 Header 20 Bytes	Version / IHL / Type of Service / Total Length	22
	Identification / Flags / Fragment Offset	26
	Time to Live / Protocol = 17 (UDP) / Header Checksum	30
	IPv4 Source Address	34
	IPv4 Destination Address	38
UDP Header 8 Bytes	Source Port = xxxx / Destination Port = 4789	42
	UDP Length / UDP Checksum	46
VXLAN Header 8 Bytes	R R R R I R R R / Reserved	50
	VXLAN Network Identifier (VNI) / Reserved	54
Inner Ethernet Header 18 Bytes	Destination MAC Address	58
	Destination MAC Address / Source MAC Address	62
	Source MAC Address	66
	Optional: 802.1Q VLAN Header	70
	Ethertype	72
Inner Payload	Original Ethernet Payload	76 80 84

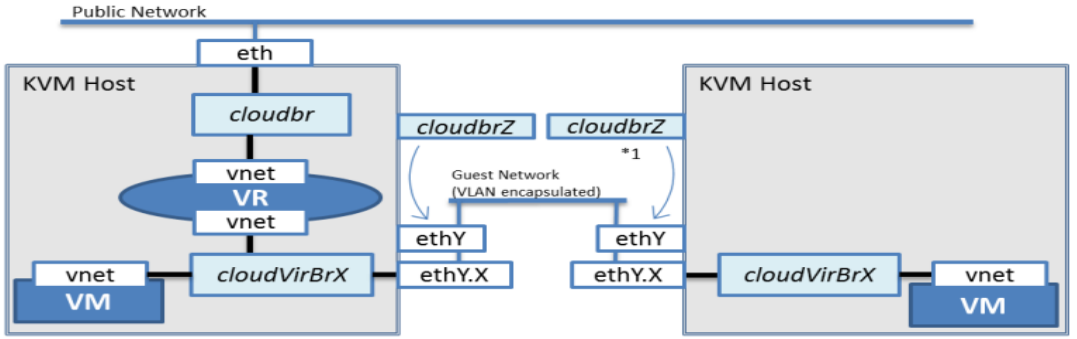
SDN no CloudStack

Compare VLAN/VXLAN bridging in KVM

VLAN Isolation

cloud-agent create vlan interface (ethY.X) on physical interface (ethY) which associated to guest traffic label (cloudbrZ), created vlan interface will be associated to cloudVirBrX.

Frame sent via ethY.X will be encapsulated with vlan header and go out from physical interface (ethY).

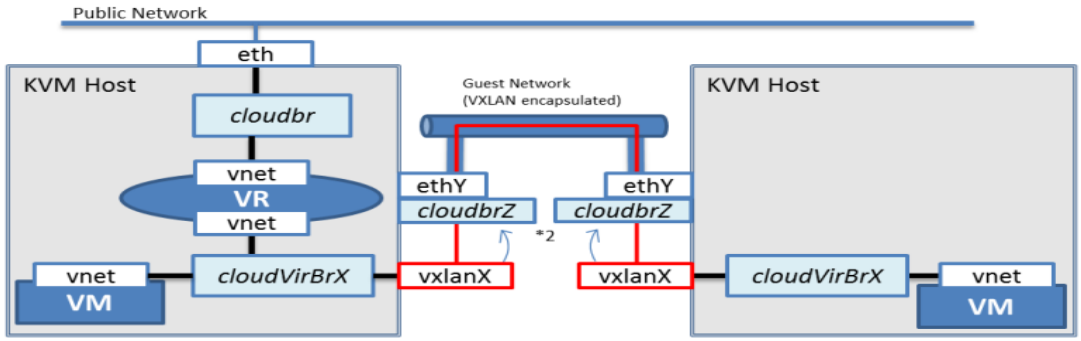


*1: Guest Traffic go through physical interface ethY, member of cloudbrZ.
cloud-agent uses name of bridge to identify physical interface.

VXLAN Isolation

cloud-agent create vxlan interface (vxlanX) on bridge interface (cloudbrZ) specified by guest traffic label, created vxlan interface will be associated to cloudVirBrX.

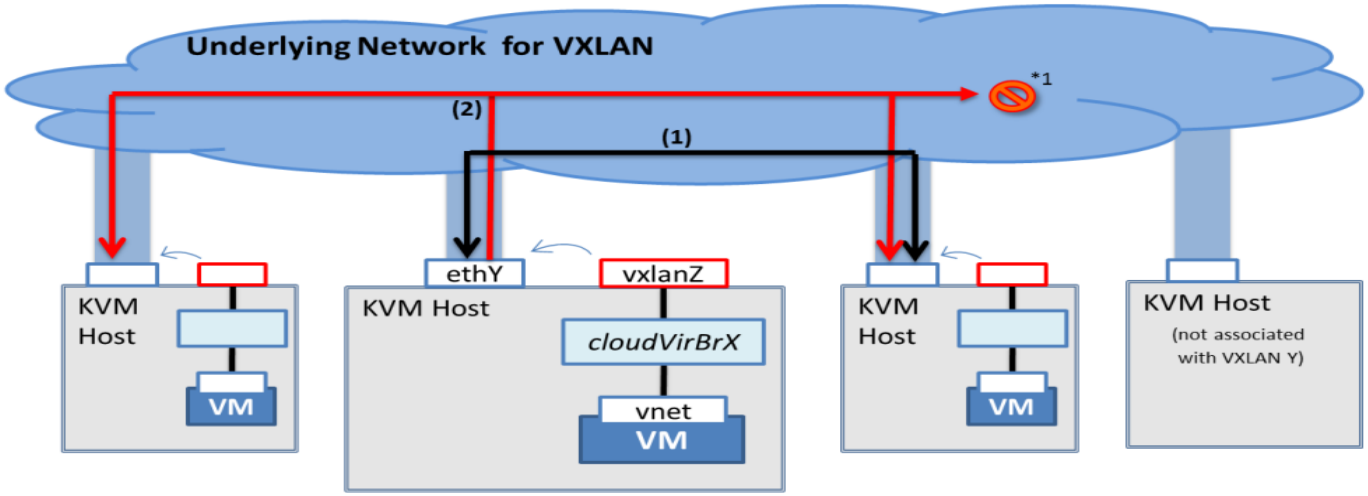
Frame sent via vxlanX will be encapsulated with vxlan header and go out from physical interface (ethY).



*2: Guest Traffic go through bridge interface after encapsulation.
That's because Linux kernel assign IP address to bridge itself, not to member of bridge.

SDN no CloudStack

How traffic flows with VXLAN



① If Unicast and KVM host (Src) learned mapping between VM and KVM host (Dst)
➔ VXLAN uses **Unicast**

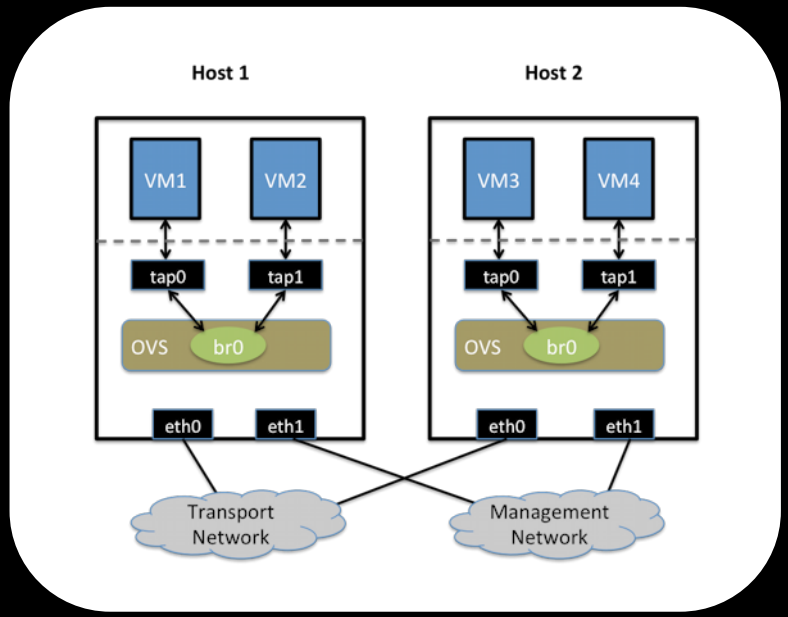
② If multicast or broadcast or Unicast but KVM host (Src) doesn't know mapping
➔ VXLAN uses **Multicast**
KVM host (Dst) learn mapping between VM and KVM (Src)

*1: If underlying Network supports IGMP/MLD snooping and/or Multicast routing.

SDN no CloudStack

- GRE - Ovs
  - CloudStack programa e executa a criação de túneis GRE entre Hypervisores utilizando Ovs. Forma uma estrutura full-mesh.
  - Ovs protege contra loopings e storms.
  - Sem delay para novos fluxos.
  - Altamente escalável, 2^24 (16M).
  - Túneis podem ser extendidos entre Data Centers via L3.
  - + 24 bytes por pacote.

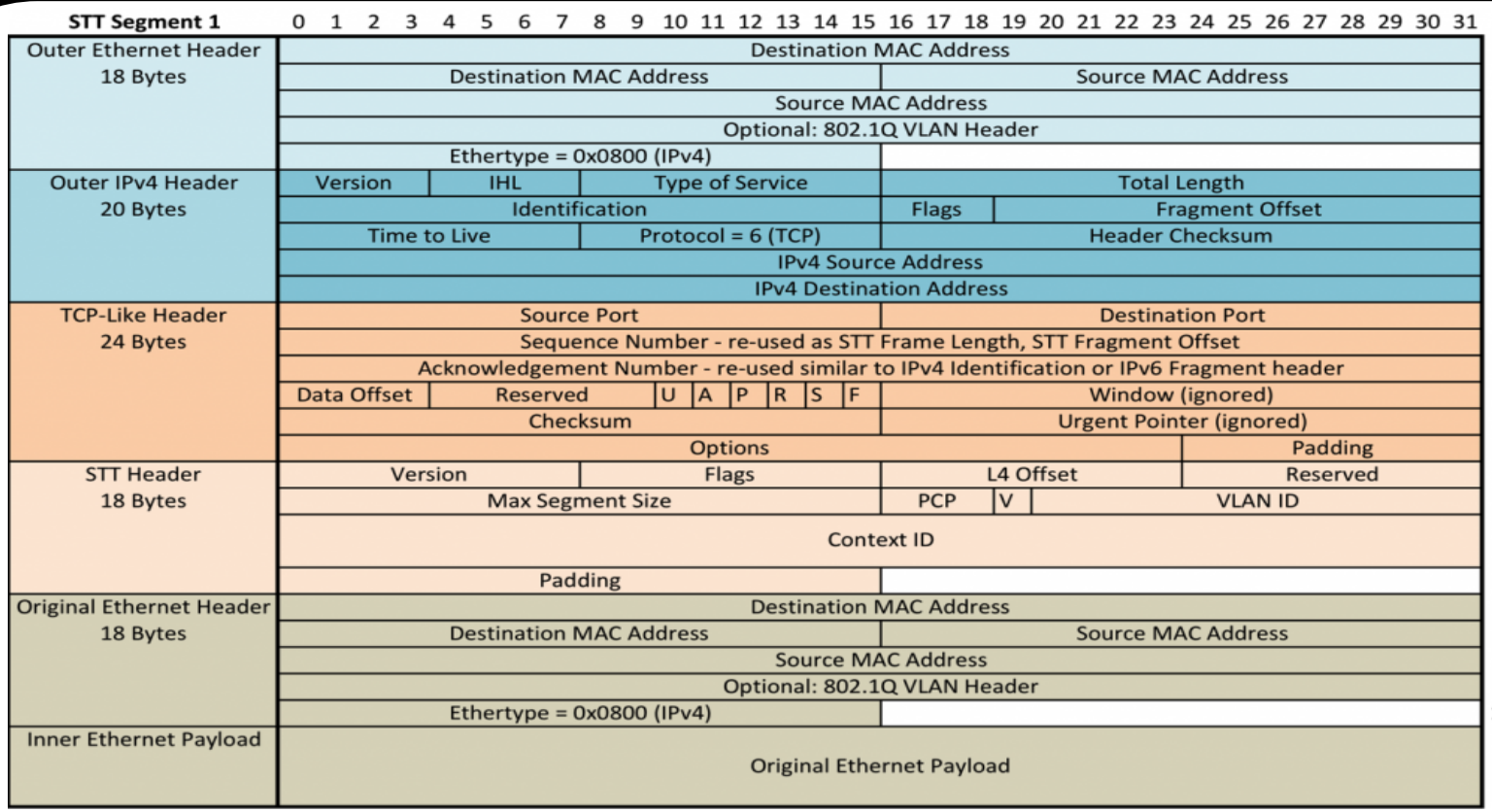
SDN no CloudStack



SDN no CloudStack

- STT
  - O frame L2 original é encapsulado dentro de pacote TCP.
  - Para ter performance necessita de placas de rede aceleradoras (TSO) para montagem dos frames fragmentados.
  - Funciona muito bem dentro do mesmo domínio com MTU grande e NICs TSO.
  - Altamente escalável, 2^64 (18Q).
  - Depende do VMware NSX.
  - + 80 bytes no primeiro pacote e 62 bytes nos seguintes.

SDN no CloudStack



STT Segment 1	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	
Outer Ethernet Header 18 Bytes	Destination MAC Address	4
	Destination MAC Address (16) | Source MAC Address (16)	8
	Source MAC Address	12
	Optional: 802.1Q VLAN Header	16
	Ethertype = 0x0800 (IPv4) |	18
Outer IPv4 Header 20 Bytes	Version (4) | IHL (4) | Type of Service (8) | Total Length (16)	22
	Identification (16) | Flags (3) | Fragment Offset (13)	26
	Time to Live (8) | Protocol = 6 (TCP) (8) | Header Checksum (16)	30
	IPv4 Source Address	34
	IPv4 Destination Address	38
TCP-Like Header 24 Bytes	Source Port (16) | Destination Port (16)	42
	Sequence Number - re-used as STT Frame Length, STT Fragment Offset	46
	Acknowledgement Number - re-used similar to IPv4 Identification or IPv6 Fragment header	50
	Data Offset (4) | Reserved (4) | U | A | P | R | S | F | Window (ignored) (16)	54
	Checksum (16) | Urgent Pointer (ignored) (16)	58
	Options (48) | Padding (16)	62
STT Header 18 Bytes	Version (16) | Flags (16) | L4 Offset (16) | Reserved (16)	66
	Max Segment Size (32) | PCP (3) | V | VLAN ID (12)	70
	Context ID	74
		78
	Padding (32) |	80
Original Ethernet Header 18 Bytes	Destination MAC Address	84
	Destination MAC Address (16) | Source MAC Address (16)	88
	Source MAC Address	92
	Optional: 802.1Q VLAN Header	96
	Ethertype = 0x0800 (IPv4) |	100
Inner Ethernet Payload	Original Ethernet Payload	

SDN no CloudStack

STT Segment 2..N	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	
Outer Ethernet Header 18 Bytes	Destination MAC Address	4
	Destination MAC Address / Source MAC Address	8
	Source MAC Address	12
	Optional: 802.1Q VLAN Header	16
	Ethertype = 0x0800 (IPv4)	18
Outer IPv4 Header 20 Bytes	Version / IHL / Type of Service / Total Length	22
	Identification / Flags / Fragment Offset	26
	Time to Live / Protocol = 6 (TCP) / Header Checksum	30
	IPv4 Source Address	34
	IPv4 Destination Address	38
TCP-Like Header 24 Bytes	Source Port / Destination Port	42
	Sequence Number - re-used as STT Frame Length, STT Fragment Offset	46
	Acknowledgement Number - re-used similar to IPv4 Identification or IPv6 Fragment header	50
	Data Offset / Reserved / U / A / P / R / S / F / Window (ignored)	54
	Checksum / Urgent Pointer (ignored)	58
	Options / Padding	62
Original Ethernet Header 18 Bytes	Destination MAC Address	66
	Destination MAC Address / Source MAC Address	70
	Source MAC Address	74
	Optional: 802.1Q VLAN Header	78
	Ethertype = 0x0800 (IPv4)	80
Inner Ethernet Payload	Original Ethernet Payload	

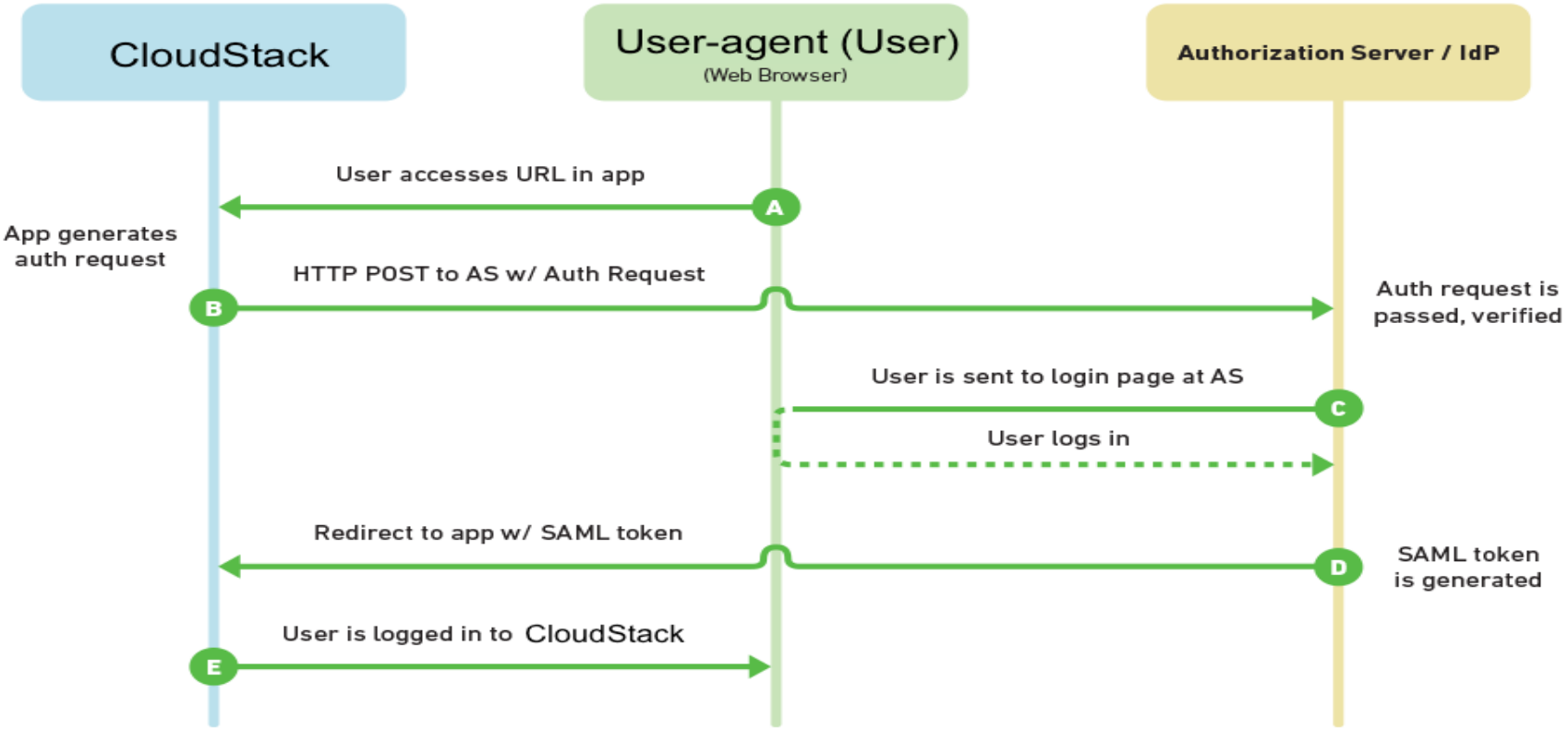
GTER 40
GTS 26
10 E 11 DE DEZEMBRO

Autenticação federada

Autenticação federada

- Instituições diferentes podem utilizar sua base de usuários para autenticação cruzada.
- Utiliza o protocolo SAML2 (ADFS, Shibboleth, OpenAM, etc)
- Compartilhamento de recursos entre instituições
- Single Sign-On
- Integração com a CAFe – Comunidade Academica Federada

SAML 2.0 Flow



GTER 40
GTS 26
10 E 11 DE DEZEMBRO

Demonstração de uma Cloud Federada

apachecloudstack
open source cloud computing

Perguntas?

marcelo.lima@shapeblue.com
Skype: mlimadc

www.cloudstack.org
www.shapeblue.com

Shape Blue The Cloud Specialists

apachecloudstack
open source cloud computing