Track 2: Operations: Data Center Architectures and Technologies

SANOG 2006 Tutorials: 1st August 2006

Zeeshan Naseh Asim Khan Bilal Khawaja

© 2006 Cisco Systems, Inc. All rights reserved.

Day Agenda

- Part I Data Center Designs and Services (Zeeshan Naseh)
 Data Center Architectures and Technologies Overview
 Content Switching & Application Optimization
- Part 2 L2 Switching Protocols (Bilal Khawaja)
- Part 3 Fiber Channel and Storage Area Networks (Asim Khan)
- Part 4 Data Center Disaster Recovery (Zeeshan Naseh)

Data Center Designs and Services Agenda

- Data Centers Components
- Data Centers Architectures
- Data Centers Technologies
- Server Load Balancing (Content Switching)
- SSL Offload
- Security (Firewall)
- Integrated Data Center Services Design Options

Data Center Components



© 2006 Cisco Systems, Inc. All rights reserved.

4

Acronyms

- ACE Application Control Engine
- BGP Border Gateway Protocol
- Cat4000 Cisco Catalyst[®] Cat4000
- Cat6500 Cisco Catalyst 6500
- CE Cisco Content Engine
- CSA Cisco Security Agent (Hostbased Intrusion Prevention)
- CSM Cisco Content Switching Service Module on Cat6500
- CSS Cisco Content Services Switch (CSS11000 and CSS11500 family)
- FWSM Cisco Firewall Service Module on Cat6500
- HSRP Hot Standby Routing Protocol
- GSS Global Site Selector
- IDSM Cisco Intrusion Detection Service Module on Cat6500

- LMS Cisco Works LAN Management Solution
- MAC Media Access Control
- MSFC Multilayer Switching Feature Card
- NAM Cisco Network Analysis Service Module on Cat6500
- OSPF Open Shortest Path First
- PBR Policy Based Routing
- SLB Server Load Balancing
- SSL Secure Socket Layer
- SSLM Cisco SSL Offload Service Module on Cat6500
- VMS Cisco Works VPN/Security Management Solution
- VPNSM Cisco Virtual Private Network Service Module on Cat6500

Data Center Residents

Presentation Servers

Web front end servers that provides the interface to the clients, e.g., Apache, IIS, etc.

Business Logic Servers

Also known as middleware custom applications

DB Servers

Oracle, Sybase, etc.

Data

Data Center Elements

Application Solution



Linux/HP, Solaris/SunFire, WebLogic, J2EE Custom App, Etc.

Database Solution

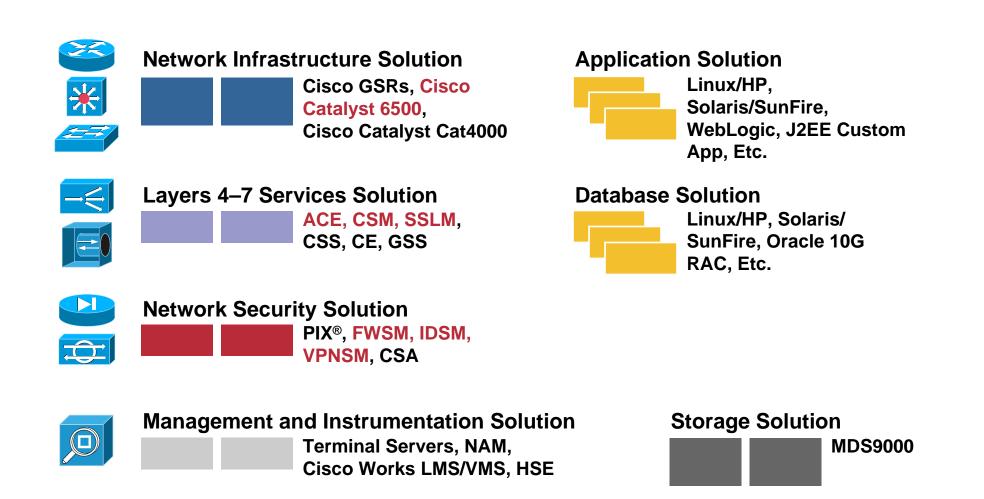


Linux/HP, Solaris/ SunFire, Oracle 10G RAC, Etc.

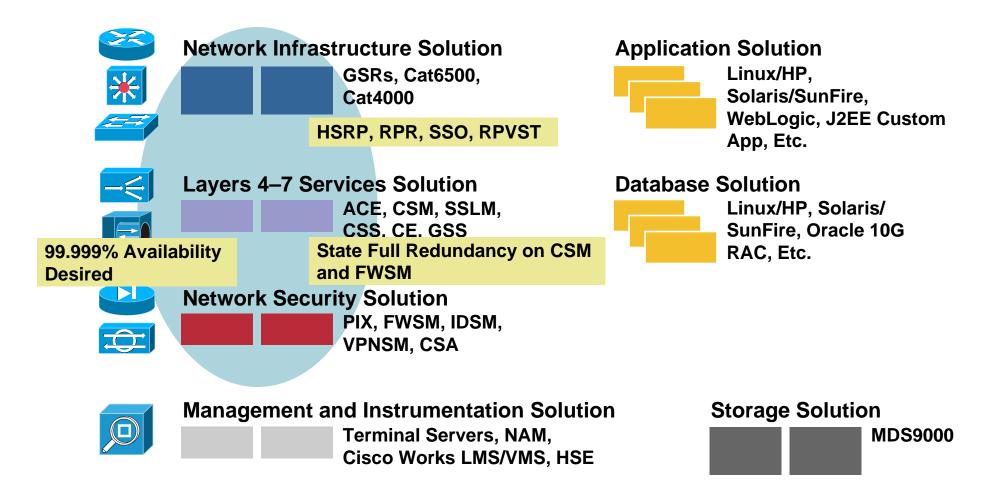


7

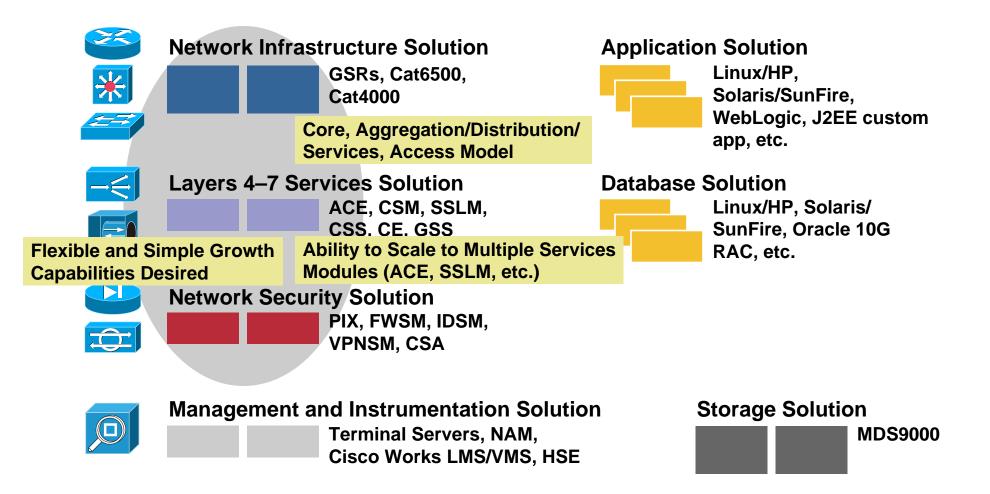
Data Center Elements



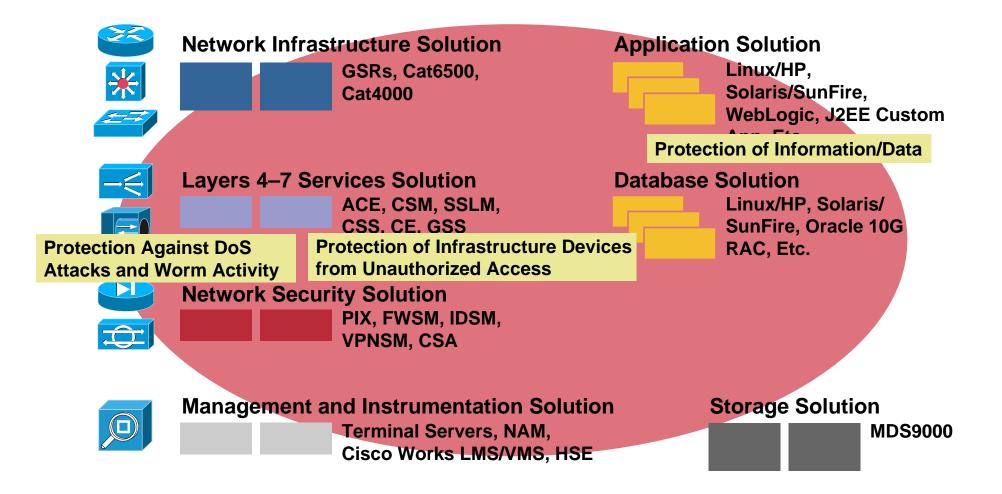
Data Center Elements Redundancy



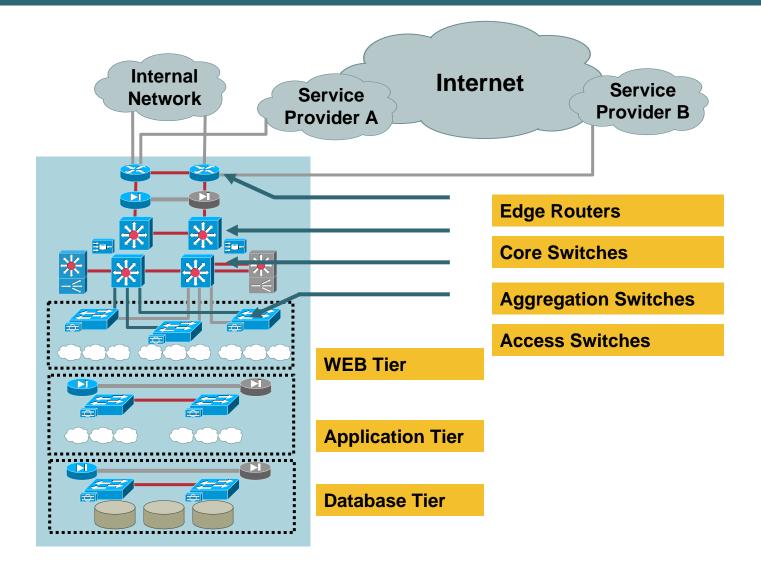
Data Center Elements Scalability



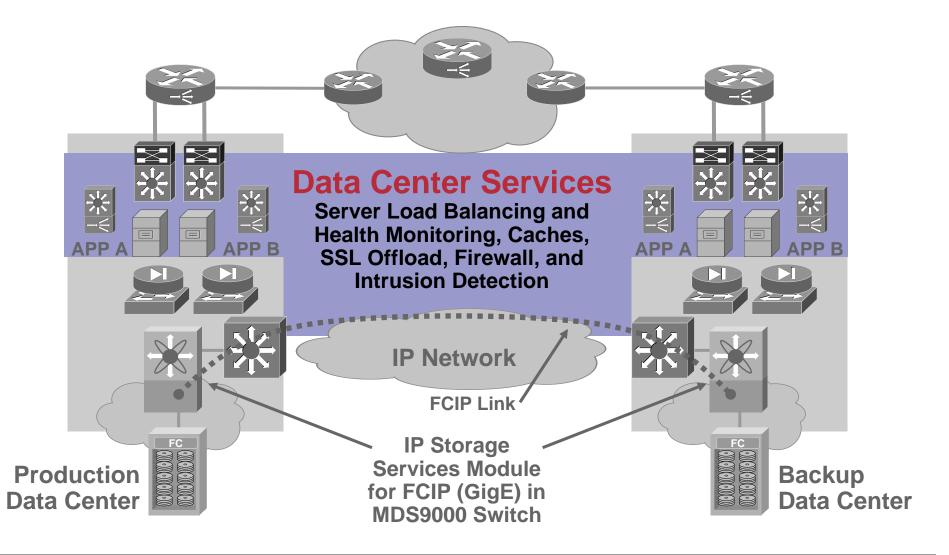
Data Center Elements Security



Typical Data Center Topology



Distributed Data Centers



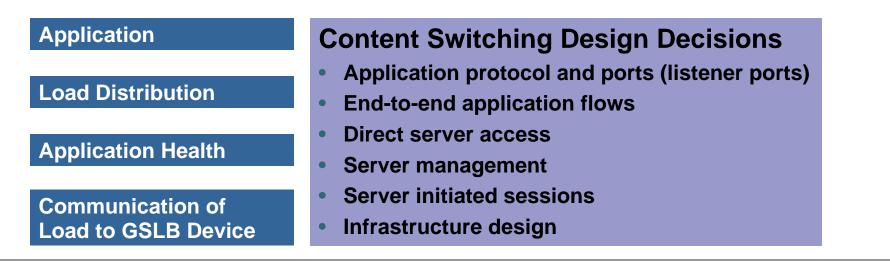
Server Load Balancing



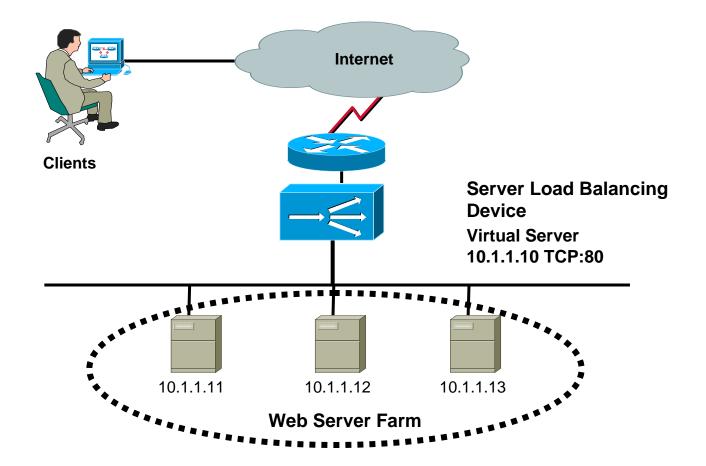
© 2006 Cisco Systems, Inc. All rights reserved.

Server Load Balancing

- Also known as content switching; one of the single most important infrastructure service in the data center
- Key purpose being request load distribution; may that be clients coming from internet, intranet, or extranet
- Layers 3 to 7 content switching capabilities are available with extensive keepalives (server health checks) functionality
- Layer 4 or Layer 7 proxy can be used as a security perimeter



SLB Overview



Content Unaware SLB

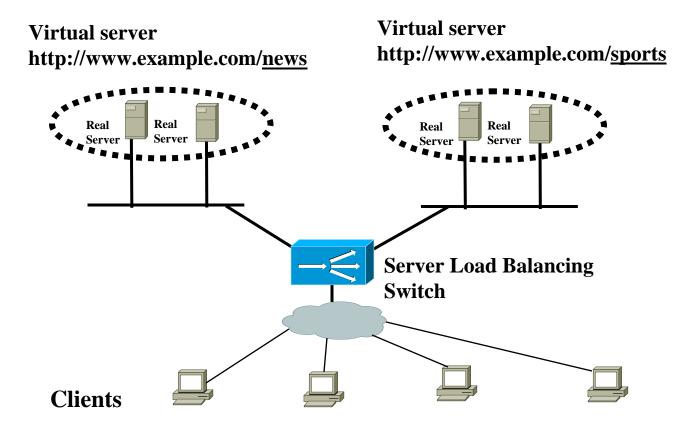
- Allows the balancing of traffic destined to a virtual server across multiple real servers
- Virtual Server / Content Rule = IP address (VIP) & L4 protocol & port
- Virtual server may have 1 to N real servers
- All real servers within a content rule must have the same content
- In the Simplest case, Load balancing decision is made on:
 - initial SYN for TCP (SYN and flow table miss)
 - initial packet for UDP (flow table miss)
- TCP connection state discarded by conn teardown (FINs/RSTs) or idle timer (garbage collection)
- UDP connection state discarded by idle timer (garbage collection)

Content Aware Loadbalancing

- Loadbalancing on anything L5 and above (HTTP cookies, HEADER Fields, HTTP Methods, URLs etc)
- HTTP URL loadbalancing most popular
- Virtual server = IP address & L4 protocol & L4 port & <u>L5-</u> <u>7 info (URL)</u>
- Virtual server is chosen by the longest URL match

Why balance on URLs ?

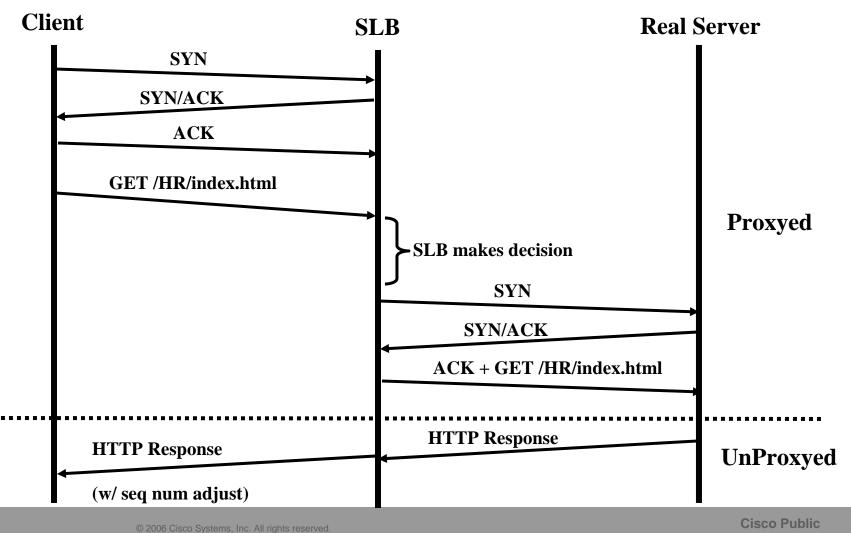
• Distributed content



© 2006 Cisco Systems, Inc. All rights reserved

URL Load balancing Packet Flow (HTTP 1.0)

Client requests http://www.example.com/HR/index.html



20

SLB Modes (Packets from SLB Device to Server)

2 basic Content Unaware SLB modes

Dispatch (VIP not NAT'd)

- rewrites the MAC address of traffic destined for the virtual server to be the real server MAC address

Directed (VIP NAT'd to real server IP)

- rewrites the IP address of traffic destined for the virtual server to be the real server IP address

- Web servers, APP servers

SLB Modes (Packets from SLB Device to Server)

Dispatch Mode (service type transparent or no nat server)

- Requires the real server to have the virtual server IP address
 - loopback interface or secondary IP address
 - a lot of per server configuration, not very popular with web hosting companies
- Requires the real server to be Layer 2 adjacent to the load balancer
- Packets sourced by the real server will contain the virtual IP address as the source (in response to traffic from the load balancer)
- FWs, Caches, SSL Offloaders etc

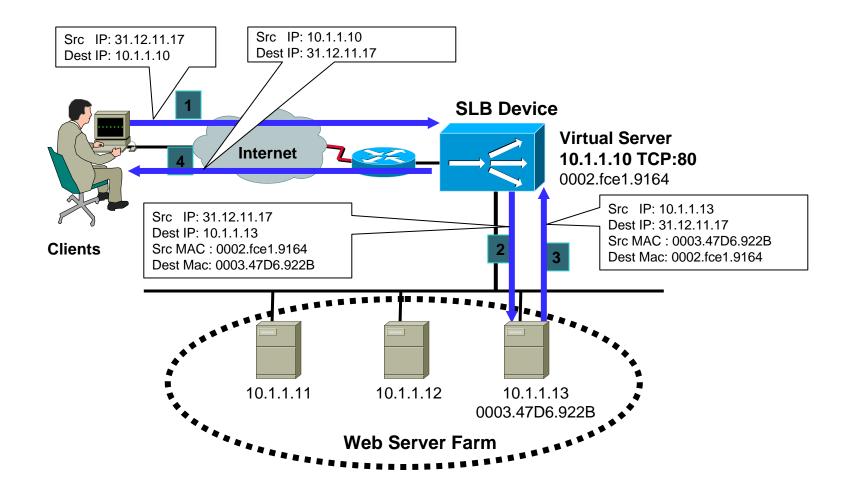
Directed Mode (default in most SLB devices)

- Real server can be anywhere (L3 addressable)
- No additional configuration needed per real server
- Packets sourced by the real server will contain its own IP address as the source
- Optional NAT of the server L4 port (port 80 -> 8080)
- More work for the loadbalancer
 - IP address change, IP hdr checksum, TCP checksum

Source (client) NAT

- Remaps the client's IP address and L4 port to one from the loadbalancer's NAT pool
- Ensures the response packets from the real server traverse the same loadbalancer that handled the request
- Loadbalancer must respond to pings, arps, etc. for addresses within the NAT pool

Typical Load Balanced Session

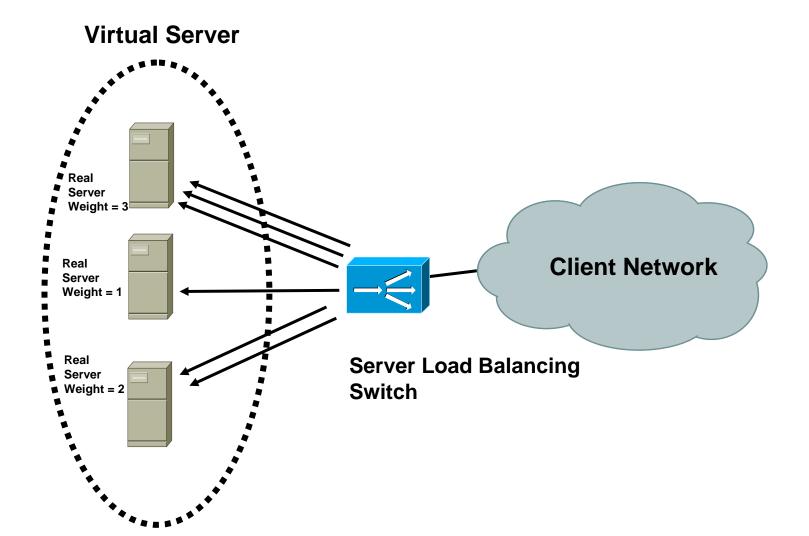


Load Balancing Algorithms

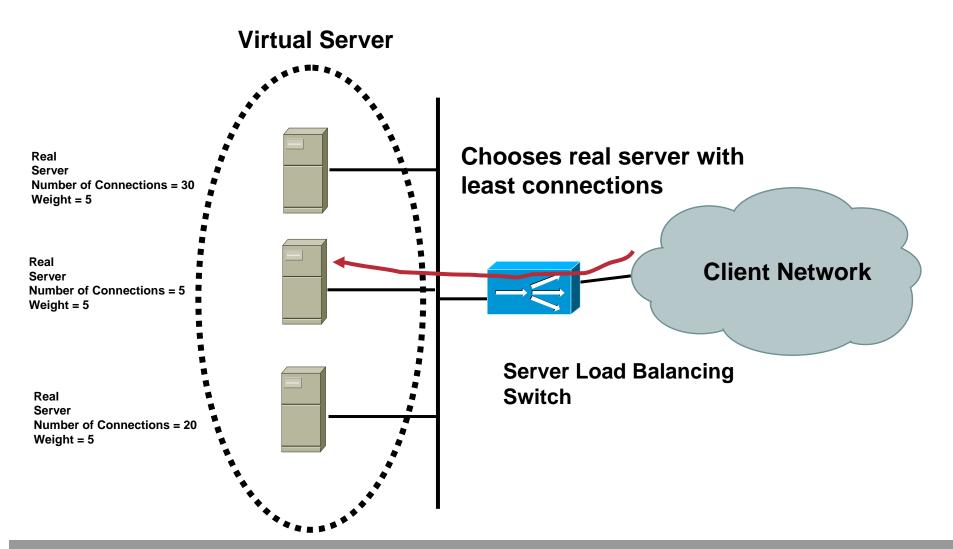
- urlhash
- domainhash
- weightedrr
- leastconn
- url

- domain
- srcip
- destip
- aca
- roundrobin

Weighted Round Robin



Least Connections

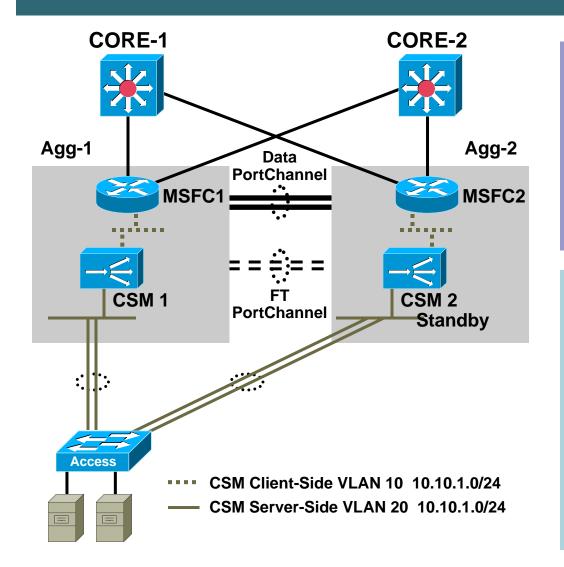


© 2006 Cisco Systems, Inc. All rights reserved.

"Sticky" Connections

- Allows new connections from a client to be sent to the same real server as previous connections from that client
- This binding is aged through the use of a sticky timer
- Configured on a virtual server basis
- Could be
 - Source IP based
 - HTTP Cookie based
 - passive (server inserted cookies)
 - active (SLB device inserted cookies)
 - SSL Session ID based

Content Switching Design Approaches Bridged Mode: Design



Key Content Switching Design Options

- Bridged mode design
- Routed mode design with MSFC on client side
- Routed mode design with MSFC on server side
- One-armed design

(1) Bridged Mode Design Considerations

- Servers default gateway is the HSRP group IP address on the MSFC
- Broadcast/multicast/route update
 traffic bridges through
- No extra configurations for: Direct access to servers Server initiated sessions
- RHI possible
- Load balancer inline of all traffic

Content Switching Design Approaches Bridged Mode: Configuration

CSM

module ContentSwitchingModule 4 ! vlan 10 client ip address 10.10.1.5 255.255.255.0 gateway 10.10.1.1 alias 10.10.1.4 255.255.255.0 !

vlan 20 server ip address 10.10.1.5 255.255.255.0

MSFC

!

interface Vlan10 ip address 10.10.1.2 255.255.255.0 standby 10 ip 10.10.1.1 standby 10 priority 110 standby 10 preempt

!

ACE

interface vlan 10 bridge-group 10 access-group input anyone access-group output anyone no shutdown ! interface vlan 20 bridge-group 10 access-group input anyone access-group output anyone

!

I

interface bvi 10

no shutdown

ip address 10.10.1.5 255.255.255.0 alias 10.10.1.4 255.255.255.0 peer ip address 10.10.1.6 255.255.255.0 no shutdown

ip route 0.0.0.0 0.0.0.0 10.10.1.1

Content Switching Design Approaches Routed Mode: Design

CORE-1

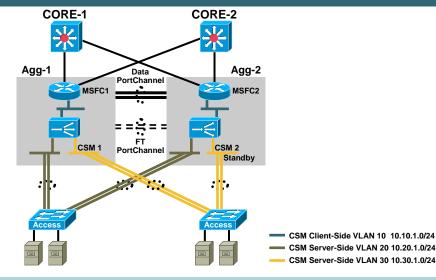
*

Data

PortChannel

PortChannel

Agg-1



(2A) Routed Mode Design with MSFC on Client Side

- Servers default gateway is the alias IP on the CSM/ACE
- Extra configurations needed for:
 - Direct access to servers Non-load balanced server initiated sessions
- CSM/ACE's default gateway is the HSRP group IP address on the MSFC
- RHI possible
- Load balancer inline of all traffic

(2B) Routed Mode Design with MSFC on Server Side

CORE-2

Agg-2

CSM 2 Standby MSFC2

- Servers default gateway is the HSRP group IP address on the MSFC
- Extra configurations needed for (simpler the option 2a):

Direct access to servers Non-load balanced server initiated sessions

- SM's default gateway is the core router
- RHI not possible
- Server to server communication bypasses the load balancer

CSM Client-Side VLAN 5 10.5.1.0/24

CSM Server-Side VLAN 1 10.10.1.0/24

10.20.1.0/24

10.30.1.0/24

Server VLAN 20

Server VI AN 30

Content Switching Design Approaches Routed Mode: Configuration

CSM

module ContentSwitchingModule 4 interface vlan 10 ! vlan 10 client ip address 10.10.1.5 255.255.255.0 ip address 10.10.1.5 255.255.255.0 alias 10.10.1.4 255.255.255.0 gateway 10.10.1.1 peer ip address 10.10.1.6 255.255.255.0 alias 10.10.1.4 255.255.255.0 no shutdown vlan 20 server interface vlan 20 ip address 10.20.1.2 255.255.255.0 ip address 10.20.1.2 255.255.255.0 alias 10.20.1.1 255.255.255.0 alias 10.20.1.1 255.255.255.0 peer ip address 10.20.1.3 255.255.255.0 no shutdown vlan 30 server ip address 10.30.1.2 255.255.255.0 MSFC alias 10.30.1.1 255.255.255.0 interface vlan 30 ip address 10.30.1.2 255.255.255.0 I interface Vlan10 alias 10.30.1.1 255.255.255.0 ip address 10.10.1.2 255.255.255.0 peer ip address 10.30.1.3 255.255.255.0 standby 10 ip 10.10.1.1

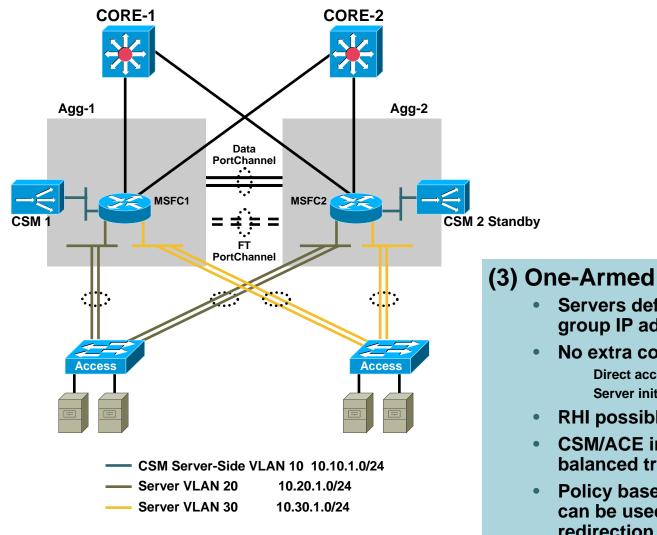
standby 10 priority 110 standby 10 preempt

ip route 0.0.0.0 0.0.0.0 10.10.1.1

no shutdown

ACE

Content Switching Design Approaches One-Armed Mode: Design



(3) One-Armed Design Considerations

- Servers default gateway is the HSRP group IP address on the MSFC
 - No extra configurations for: Direct access to servers Server initiated sessions
- **RHI** possible
- CSM/ACE inline for only server load balanced traffic
- Policy based routing or source NAT can be used for server return traffic redirection to the load balancer

Content Switching Design Approaches One-Armed Mode: PBR Configuration

MSFC

```
!
interface Vlan10
ip address 10.1(
standby 10 ip 1(
standby 10 prio
standby 10 pree
!
```

MSFC interface Vlan20 ip address 10.20.1.2 255.255.255.0 ip policy route-map FromServersToSLB standby 20 ip 10.20.1.1 standby 20 priority 110 standby 20 preempt access-list 121 permit tcp any eq telnet any access-list 121 permit tcp any eq www any access-list 121 permit tcp any eq 443 any access-list 121 deny ip any any route-map FromServersToSLB permit 10 match ip address 121

```
set ip next-hop 10.10.1.4
```

```
CSM - Asymmetric Routing
```

!

module ContentSwitchingModule 4
variable ROUTE_UNKNOWN_FLOW_PKTS 2

ACE - Asymmetric Routing

```
!
```

I

interface vlan 10 ip address 10.10.1.5 255.255.255.0 alias 10.10.1.4 255.255.255.0 peer ip address 10.10.1.6 255.255.255.0 **no normalization** access-group input anyone access-group output anyone no shutdown

Content Switching Design Approaches One-Armed Mode: Source-NAT Configuration

CSM

```
module ContentSwitchingModule 4
natpool SRC_NAT 10.10.1.110 10.10.1.110 netmask
   255.255.255.0
serverfarm SFARM NAT
nat server
nat client SRC NAT
real 10.20.1.11
inservice
real 10.20.1.12
 inservice
probe TCP
```

ACE

policy-map multi-match SLB-TELNET-POLICY class SLB-TELNET loadbalance vip inservice loadbalance policy TELNET-POLICY-TYPE loadbalance vip icmp-reply nat dynamic 1 vlan 10

interface vlan 10 ip address 10.10.1.6 255.255.255.0 alias 10.10.1.4 255.255.255.0 peer ip address 10.10.1.5 255.255.255.0

no normalization

access-group input anyone

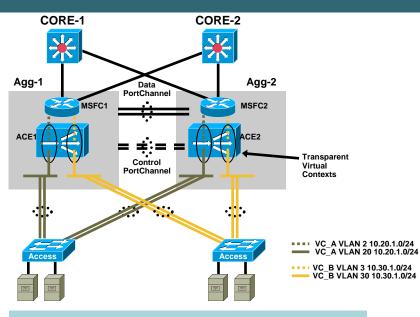
access-group output anyone

nat-pool 1 10.10.1.110 10.10.1.110 netmask 255.255.255.0 pat

no shutdown

2006 Cisco Systems, Inc. All rights reserved.

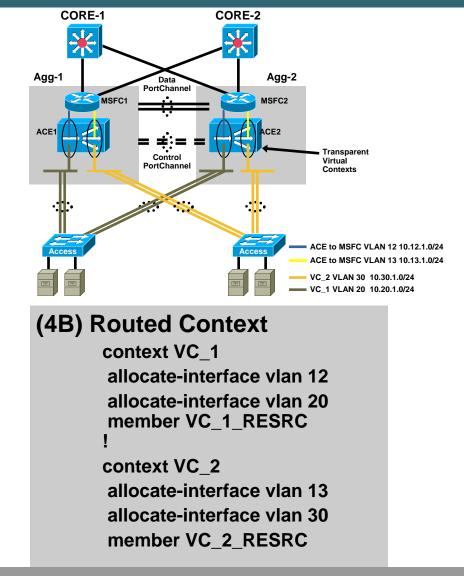
Content Switching Design Approaches Virtual Context in ACE



(4A) Bridged Context

context VC_A allocate-interface vlan 2 allocate-interface vlan 20 member VC_A_RESRC

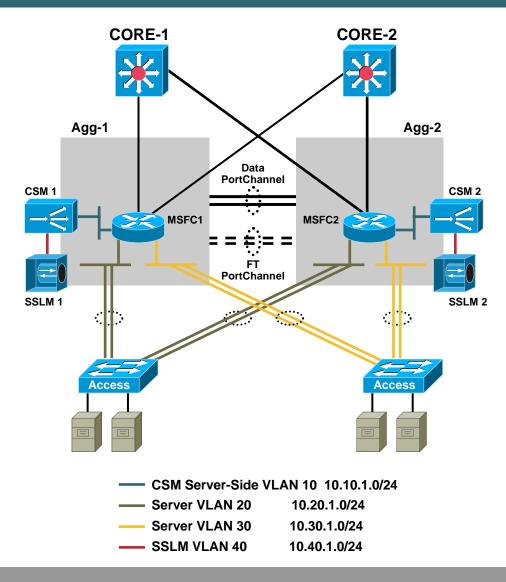
context VC_B allocate-interface vlan 3 allocate-interface vlan 30 member VC_B_RESRC



SSL Offload



Network-Based SSL Offload



Key Motivations

- Offload SSLdecryption/ encryption from servers
- Redundancy
- Scalability
- Unified management of SSL certificates
- Layer 7 based load balancing and sticky possible for HTTPS

SSL Offload Design

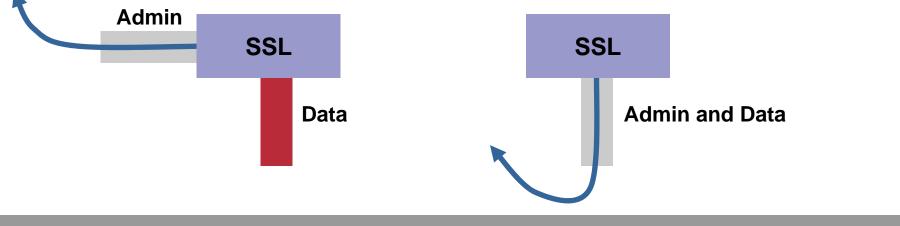
- Simply add the SSLMs on a VLAN connected to the ACE
- SSLMs default gateway would be the alias IP on the ACE
- Backend SSL requires no design change

SSL Services Module Configuration Tips: Admin VLAN and Data VLAN

One VLAN on the SSL Module Has to Be "Admin VLAN"

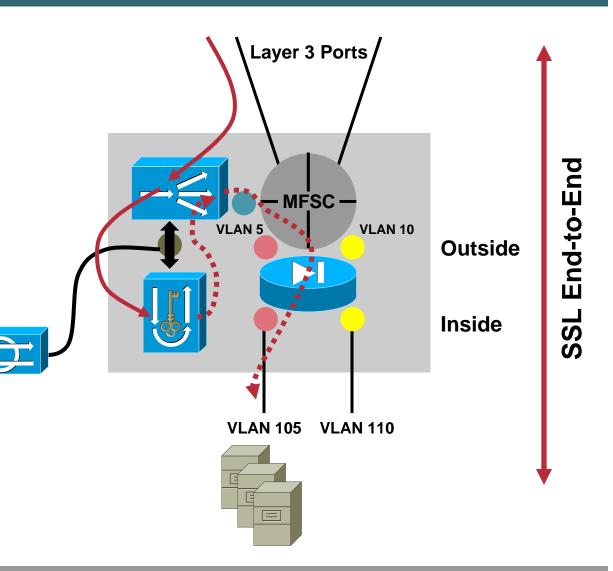
Make Sure That the Admin VLAN Has a Route to the CA, TFTP Server, Management Stations, Etc. The "Admin VLAN" Can Also Carry Data Traffic

The Default Gateway of the Admin VLAN Is the Module Default Gateway



Network Based SSL Offload Intrusion Detection Benefits

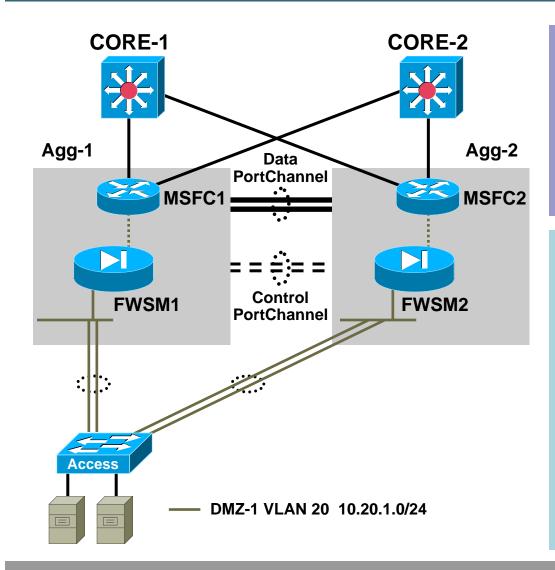
- If SSL offloading uses backend encryption
- More accurate Layer 5 load balancing decisions
- Is there any security advantage?
- Yes—you can monitor the decrypted traffic with an IDS sensor



Data Center Security



Firewall Design Approaches Layer 2



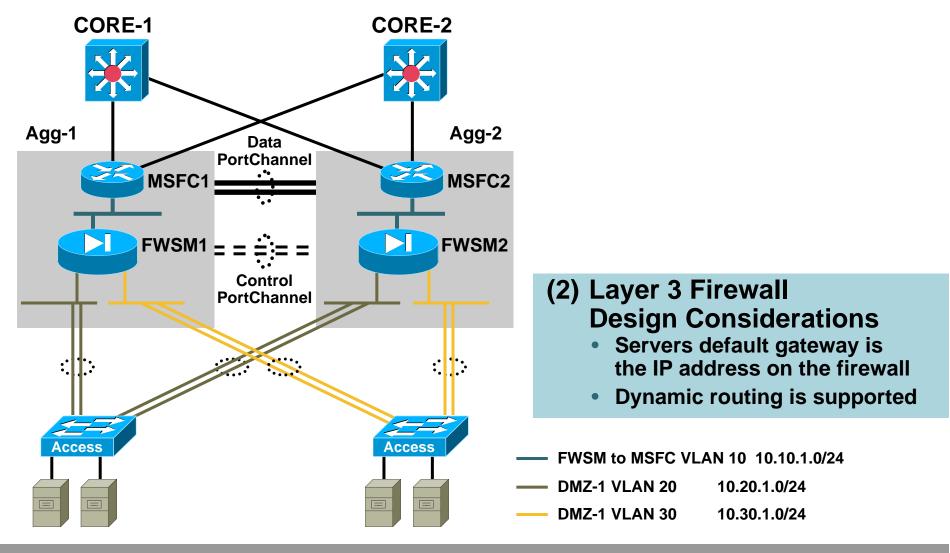
Key Firewall Design Options

- Bridged mode design, also known as transparent or stealth firewall
- Routed mode design, also known as Layer 3 firewall
- Virtual firewall contexts for Layer 2 or Layer 3 mode

(1) Layer 2 (Transparent) Firewall Design Considerations

- Servers default gateway is the HSRP group IP address on the MSFC
- Broadcast/multicast/route update traffic bridges through
- Bump on the wire; easy integration
- Currently two VLANs can be merged

Firewall Design Approaches Layer 3



Firewall Design Approaches Virtual Context

- It's the ability to segment a single physical firewall into multiple virtualized instances
- Multiple interfaces/ VLANs within Layer 3 virtual contexts are supported
- Multiple bridge pairs for Layer 2 virtual contexts are supported

ON MSFC

firewall multiple-vlan-interfaces firewall module 7 vlan-group 100 firewall vlan-group 100 21-25,50-53

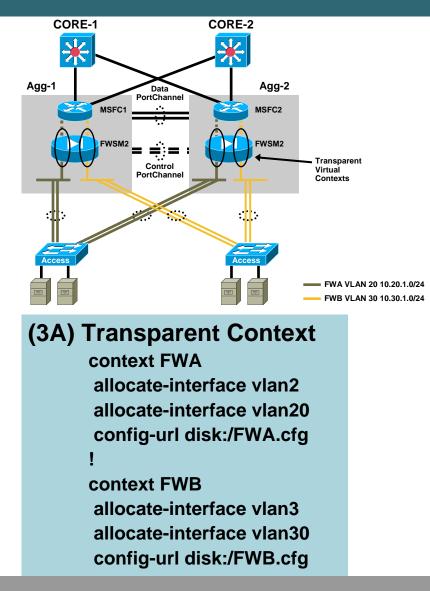
ON FIREWALL

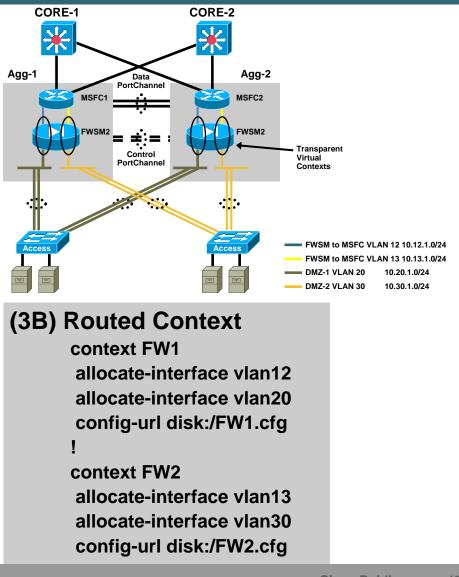
CAT1-FWSM-SYS# conf t CAT1-FWSM-SYS(config)# firewall ?

Usage: [no | clear | show] firewall [transparent] FWSM(config)# FWSM(config)# mode ?

Usage: mode single | multiple FWSM(config)# FWSM#

Firewall Design Approaches Virtual Context



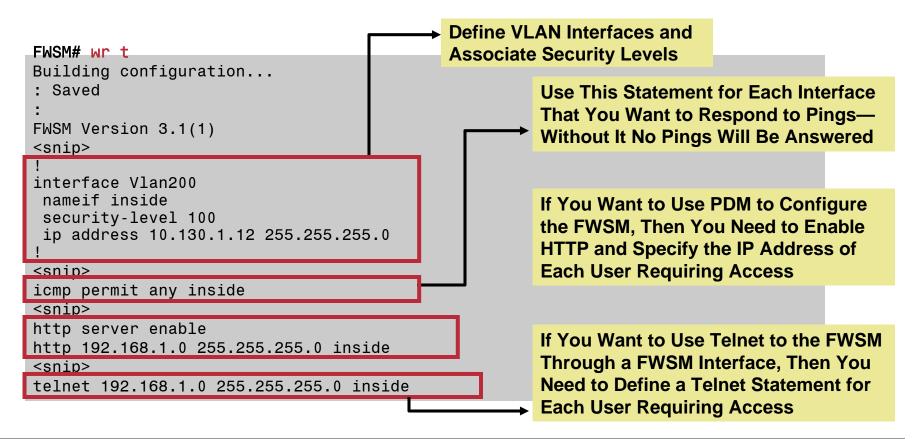


Firewall Services Module (Cont.) Configuration Tips for Getting Started



FWSM

Some Initial Configuration FWSM Configuration Statements



Integrated Data Center Design Options



Data Center Services Design Options

- We understand what products and devices are available in the data center to provide the services of security, server load balancing, SSL offload, IPS, etc.
- We understand design options of individual products
- Let's look at different ways of integrating these products
- Each design consists of three redundant layers—core, aggregation, and access

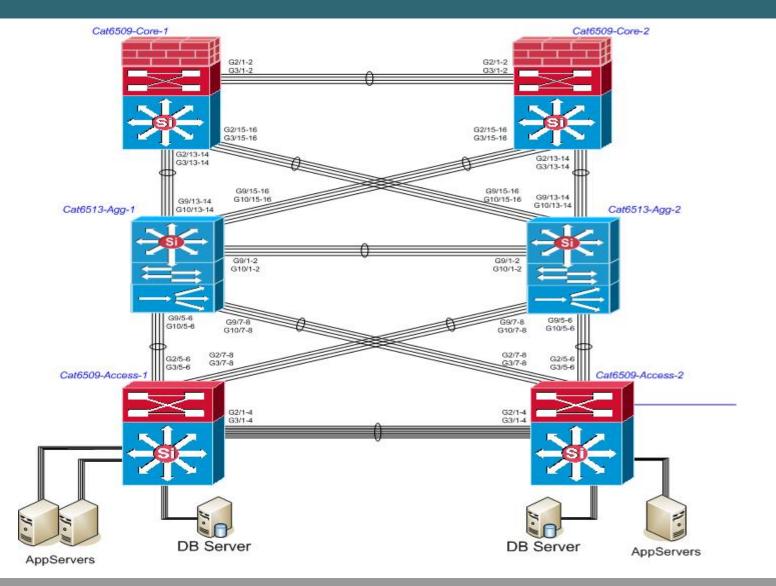
(1) FW on Core With CSM on Aggregation in Layer 3

(2) FW and CSM on Aggregation with CSM in Layer 2 and FW in Layer 3

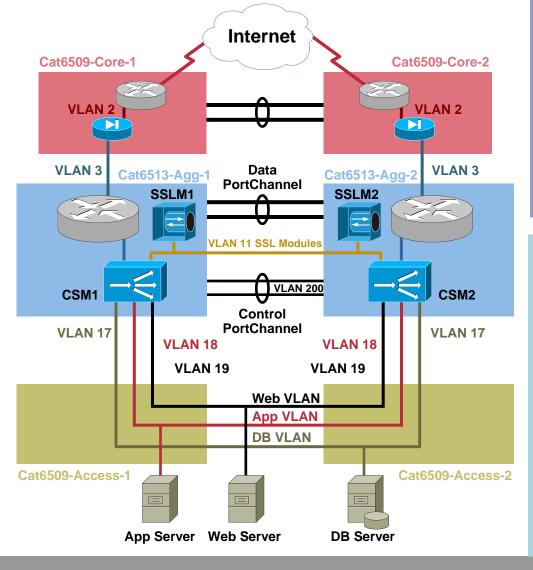
(3) FW and CSM on Aggregation with CSM in One-Armed and FW in Layer 3

(4) FW and CSM on Aggregation with CSM in One-Armed and FW in Layer 2 Secure Internal Segment

Physical Topology



Design (1): Firewall on Core; ACE on Aggregation in Layer 3 Mode



Security Details

- Layer 3 firewall used
- Firewall perimeter at the core
- Aggregation and access are considered trusted zones
- Security permiter not possible between Web/App/DB tiers
- In the aggregation layer, some security using VLAN tags on the CSM is possible

Content Switching Details

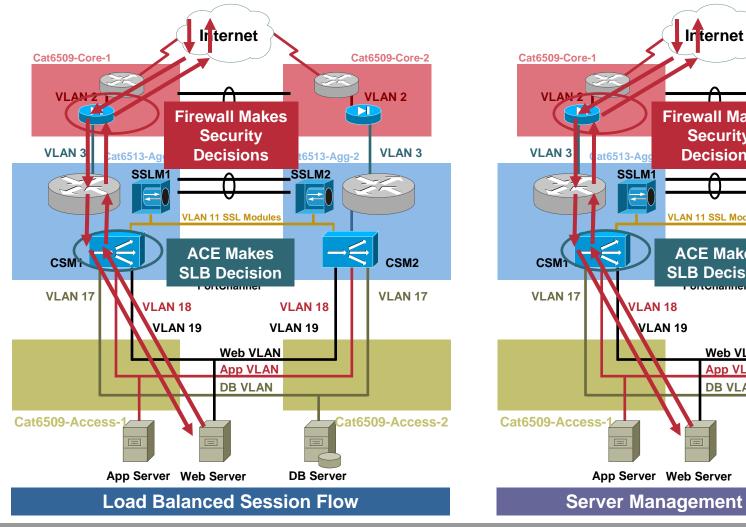
- CSM is used in routed design
- Servers default gateway is the CSM alias IP address
- Extra configurations needed for: Direct access to servers Non-load balanced server initiated sessions
- CSM's default gateway is the HSRP group IP on the MSFC
- Since MSFC is directly connected to the CSM, RHI is possible
- All to/from traffic, load balanced/ non-load balanced servers go through the CSM

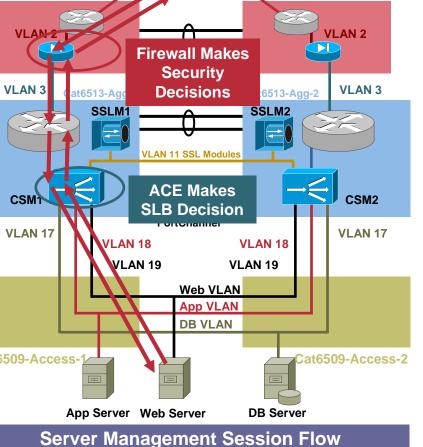
Design (1): Firewall on Core; CSM on Aggregation in Layer 3 Mode

Configuration Snapshots

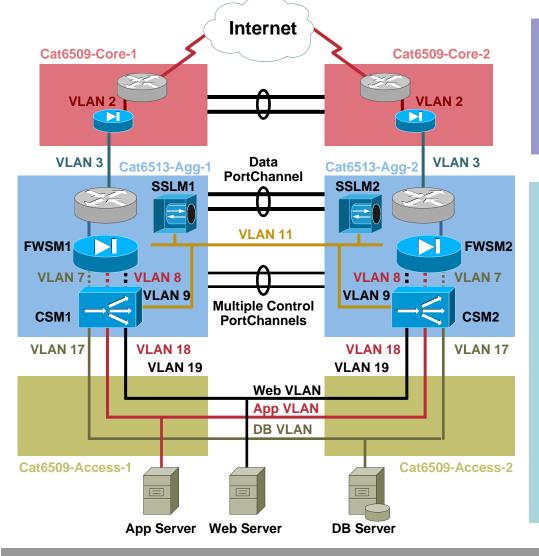
module ContentSwitchingModule 3		MSFC SVI
vlan 16 client		
ip address 10.16.1.12 255.255.255.0		interface Vlan16
gateway 10.16.1.1		ip address 10.16.1.2 255.255.255.0
alias 10.16.1.11 255.255.255.0		standby 16 ip 10.16.1.1
1		standby 16 priority 150
vlan 11 server		
ip address 10.11.1.2 255.255.255.0		
alias 10.11.1.1 255.255.255.0		
1	convorfa	
vlan 17 server	serverfarm ROUTE	
ip address 10.17.1.2 255.255.255.0	no nat server no nat client	
alias 10.17.1.1 255.255.255.0		
1		or forward
vlan 18 server	!	DOUTE
ip address 10.18.1.2 255.255.255.0	vserver ROUTE	
alias 10.18.1.1 255.255.255.0		0.0.0.0 0.0.0.0 any
!		arm ROUTE
vlan 19 server	-	ent rebalance
ip address 10.19.1.2 255.255.255.0	inservice	
alias 10.19.1.1 255.255.255.0		

Design (1): Firewall on Core; CSM on **Aggregation in Layer 3 Mode: Session Flows**





Cat6509-Core-2



Security Details

- Layer 3 firewall used with single contexts
- Firewall perimeter at the core
- Firewall perimeter is used in the aggregation between Web/App/DB tiers

Content Switching Details

- CSM is used in bridged design with multiple bridged VLAN pairs
- Servers default gateway is the firewall primary IP address
- No extra configurations needed for: Direct access to servers Non-load balanced server initiated sessions
- CSM's default gateway is the firewall primary IP address
- Since MSFC is not directly connected to the CSM, RHI is not possible
- All to/from traffic, load balanced/ non-load balanced servers go through the CSM

Configuration Snapshots

```
module ContentSwitchingModule 3
vlan 11 server
ip address 10.11.1.2 255.255.255.0
alias 10.11.1.1 255.255.255.0
vlan 7 client
ip address 10.17.1.11 255.255.255.0
gateway 10.17.1.1
vlan 17 server
ip address 10.17.1.11 255.255.255.0
vlan 8 client
ip address 10.18.1.11 255.255.255.0
gateway 10.18.1.1
vlan 18 server
ip address 10.18.1.11 255.255.255.0
```

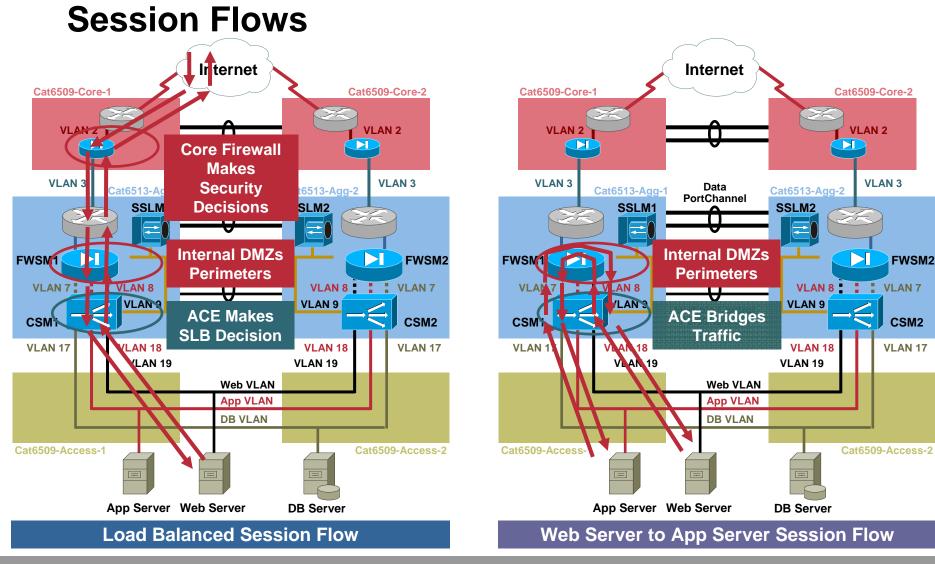
MSFC SVI

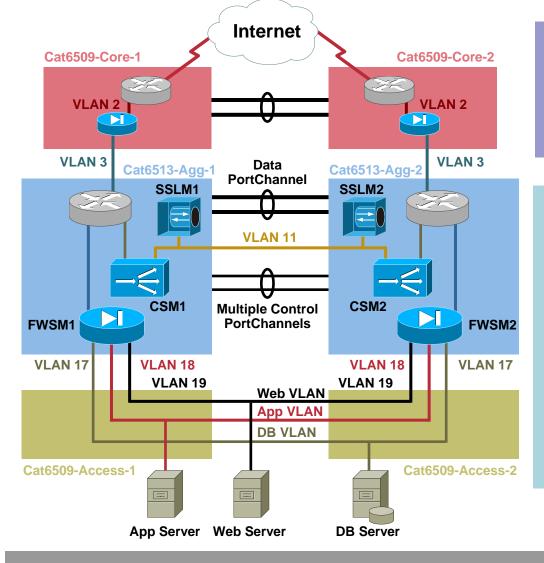
interface Vlan16 ip address 10.16.1.2 255.255.255.0 standby 16 ip 10.16.1.1 standby 16 priority 150

VLANS ON THE FIREWALL

VLAN16 (towards the MSFC)

DMZ VLANs VLAN7 VLAN8 VLAN9





Security Details

- Layer 3 firewall used with single contexts
- Firewall perimeter at the core
- Firewall perimeter is used in the aggregation between Web/App/DB tiers

Content Switching Details

- CSM is used in a one-armed fashion
- Servers default gateway is the firewall primary IP address
- No extra configurations needed for: Direct access to servers Non-load balanced server initiated sessions
- CSM's default gateway is the HSRP group address on the MSFC
- Since MSFC is directly connected to the CSM, RHI is possible
- All non-load balanced traffic to/from servers will bypass the CSM

module ContentSwitchingModule 3
vlan 15 server
ip address 10.15.1.12 255.255.255.0
gateway 10.15.1.1
alias 10.15.1.11 255.255.255.0
!
vlan 11 server
ip address 10.11.1.2 255.255.255.0
alias 10.11.1.1 255.255.255.0
!

MSFC SVI

interface Vlan15 ip address 10.15.1.2 255.255.255.0

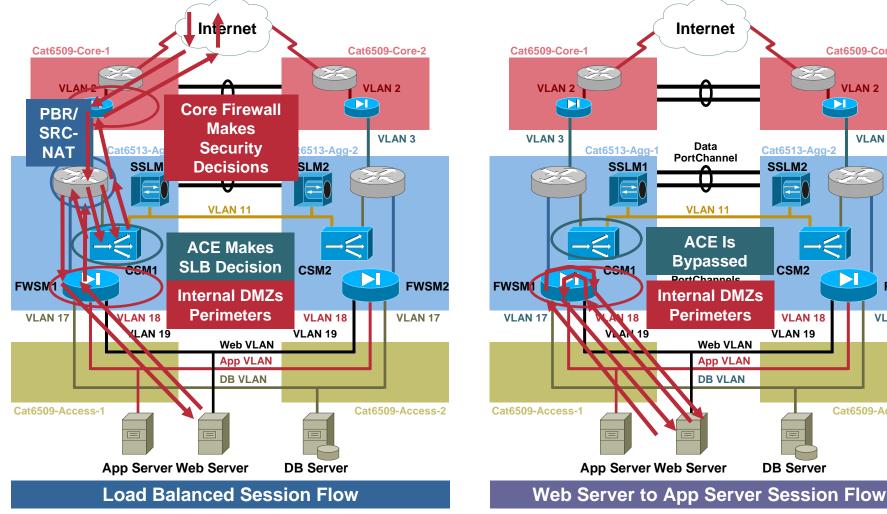
standby 15 ip 10.15.1.1

standby 15 priority 150

interface Vlan16 ip address 10.16.1.2 255.255.255.0 standby 16 ip 10.16.1.1 standby 16 priority 150

VLANS ON THE FIREWALL VLAN16 (towards the MSFC) DMZ VLANs VLAN17 VLAN18 VLAN19

Session Flows (1 of 2)



Cat6509-Access-2

Cat6509-Core-2

VLAN 2

FWSM2

VLAN 17

Cat6513-Agg-2

SSLM2

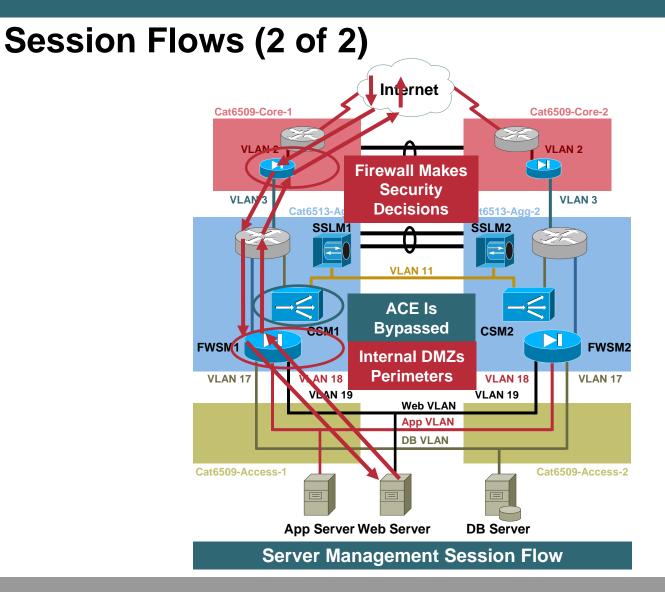
CSM2

VLAN 18

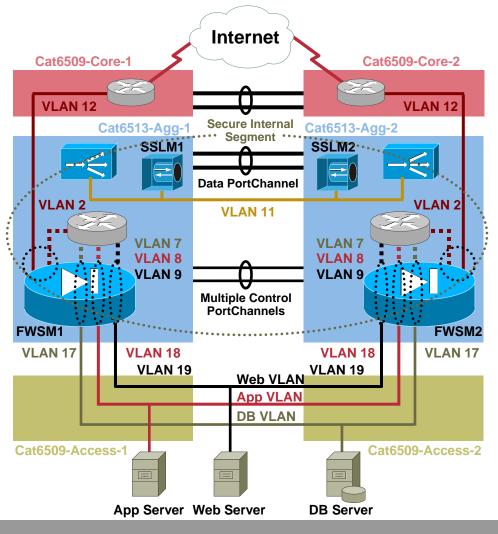
VLAN 19

DB Server

VLAN 3



Design (4): Firewall and CSM on Aggregation; FW in Layer 2 and CSM in One-Armed Mode [Secure Internal Segment]



Security Details

- Layer 2 firewall used with multiple contexts
- Firewall perimeter at outside, internal and each DMZ
- Agg MSFC is a secure internal segment with protection from each connected network
- Secure internal segment is protected from malicious activity from each DC network

Content Switching Details

- CSM is used in a one-armed fashion
- Servers default gateway is the HSRP group IP address
- No extra configurations needed for: Direct access to servers Non-load balanced server initiated sessions
- CSM's default gateway is the HSRP group address on the MSFC
- Since MSFC is directly connected to the CSM, RHI is possible
- All non-load balanced traffic to/from servers will bypass the CSM

Design (4): Firewall and CSM on Aggregation; FW in Layer 2 and CSM in One-Armed Mode [Secure Internal Segment]

module ContentSwitchingModule 3 vlan 15 server ip address 10.15.1.12 255.255.255.0 gateway 10.15.1.1 alias 10.15.1.11 255.255.255.0 ! vlan 11 server ip address 10.11.1.2 255.255.255.0

FIREWALL CONTEXTS

alias 10.11.1.1 255.255.255.0

context DB allocate-interface vlan7 allocate-interface vlan17 config-url disk:/DB.cfg !

context APP allocate-interface vlan8 allocate-interface vlan18 config-url disk:/APP.cfg context WEB allocate-interface vlan9 allocate-interface vlan19 config-url disk:/WEB.cfg

MSFC SVI

interface Vlan15 ip address 10.15.1.2 255.255.255.0 standby 15 ip 10.15.1.1 standby 15 priority 150

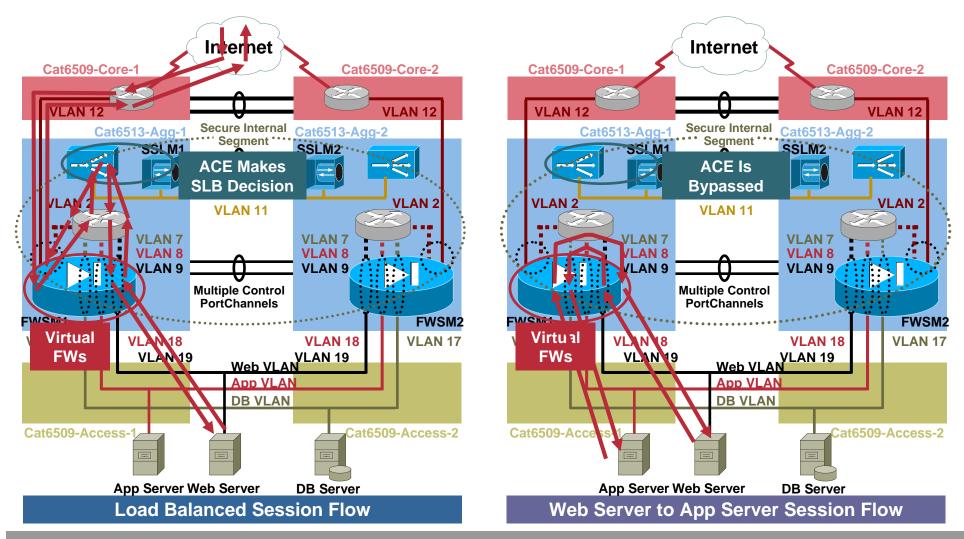
interface Vlan7 ip address 10.17.1.2 255.255.255.0 standby 17 ip 10.17.1.1 standby 17 priority 150

interface Vlan8 ip address 10.18.1.2 255.255.255.0 standby 18 ip 10.18.1.1 standby 18 priority 150

!

interface Vlan9 ip address 10.19.1.2 255.255.255.0 standby 19 ip 10.19.1.1 standby 19 priority 150

Design (4): Firewall and CSM on Aggregation; FW in Layer 2 and CSM in One-Armed Mode [Secure Internal Segment]: Session Flows



Q and A



Recommended Reading

 Designing Content Switching Solutions: ISBN: 158705213X

> By Zeeshan Naseh, Haroon Khan



Designing Content Switching Solutions

A practical guide to the design and deployment of content switching solutions for mission-critical applications in data center environments

ciscopress.com

CISCO SYSTEMS

Zeeshan Naseh, CCIE® No. 6838 Haroon Khan, CCIE No. 4530

CISCO SYSTEMS