Content Switching and Application Optimization Technologies and Design Approaches within Data Centers

Content Track

APRICOT 2006

Zeeshan Naseh

Technical Leader, Data Center Networking Practice, Cisco Systems

Agenda

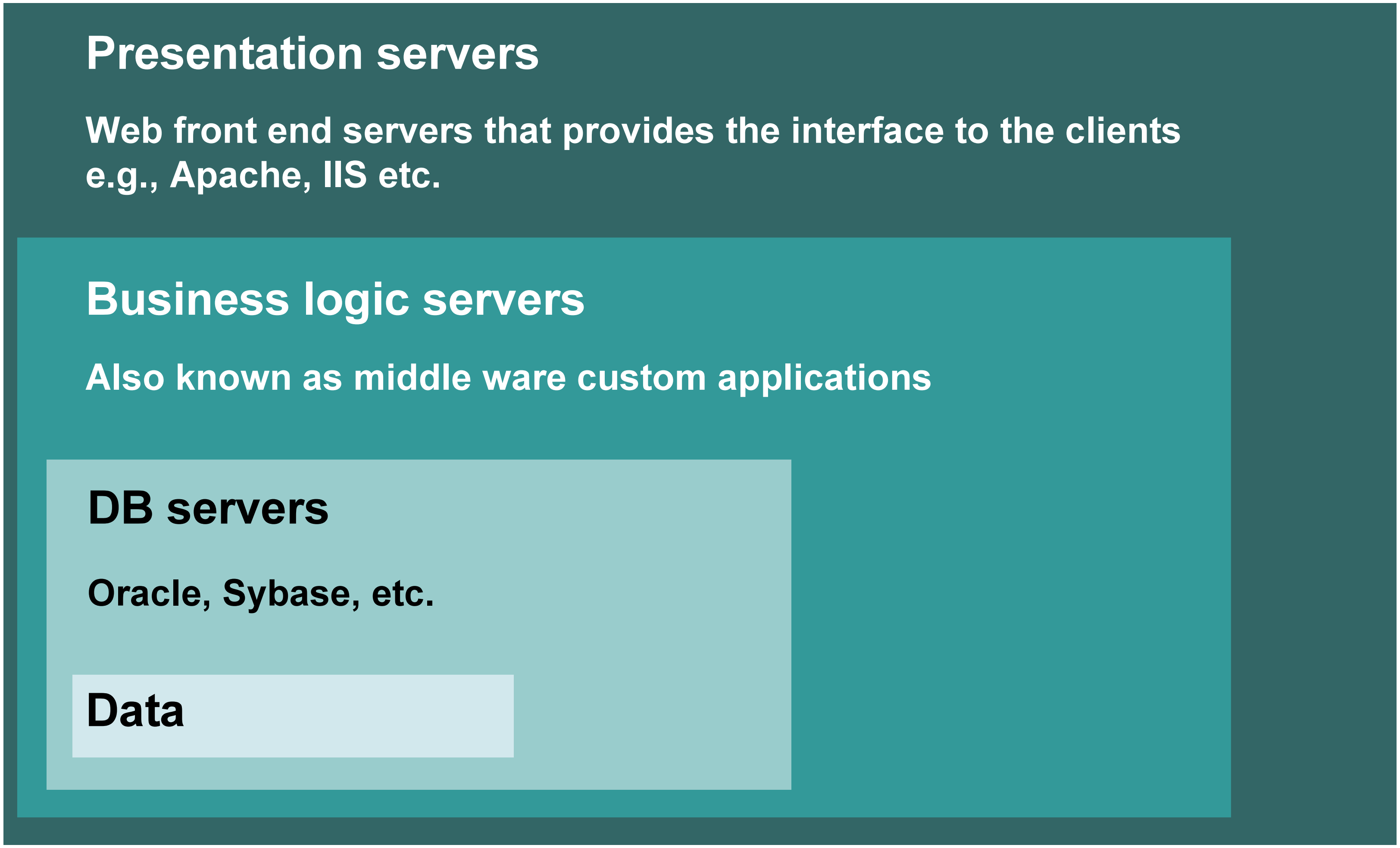
- Data Centers Components
- Server Load Balancing (Content Switching)
- SSL Offload
- Security (Firewall, Intrusion Detection, VPN)
- Integrated Data Center Services Design Options
- Real World Deployments

DATA CENTER COMPONENTS

Acronyms

BGP	Border Gateway Protocol
Cat4000	Cisco Catalyst® Cat4000
Cat6500	Cisco Catalyst 6500
CE	Cisco Content Engine
CSA	Cisco Security Agent (Host-based 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

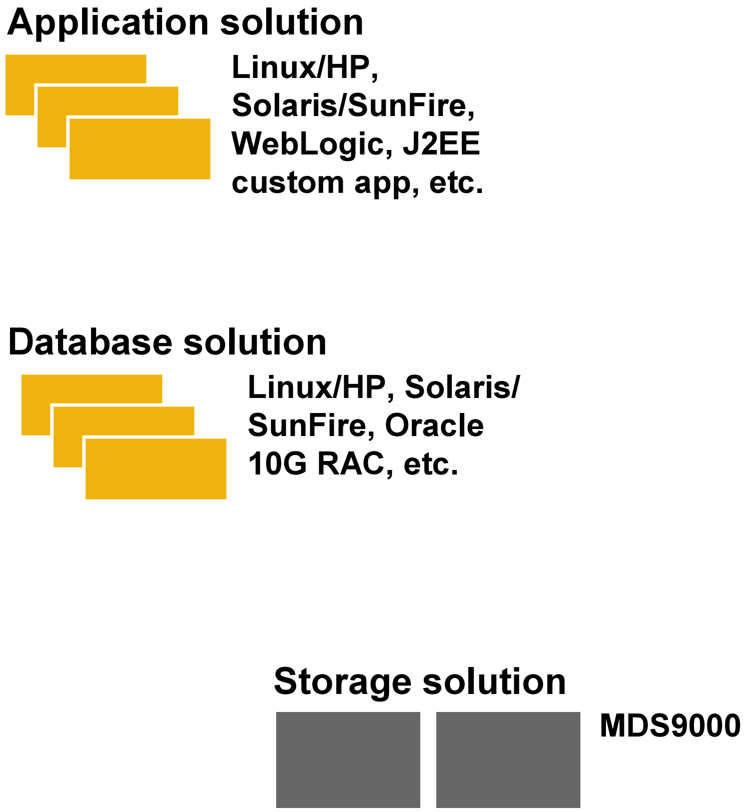
Also known as middle ware custom applications

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.

Storage solution

MDS9000

Data Center Elements

Network infrastructure solution

Cisco GSRs, CISCO CATALYST 6500, Cisco Catalyst Cat4000

Layer 4–7 services solution

CSM,
SSLM,
CSS,
CE, GSS

Network security solution

PIX®,
FWSM,
IDSM,
VPNSM,
CSA

Management and instrumentation solution

Terminal servers, NAM, Cisco Works LMS/VMS, HSE

Application solution

Linux/HP, Solaris/SunFire, WebLogic, J2EE custom app, etc.

Database solution

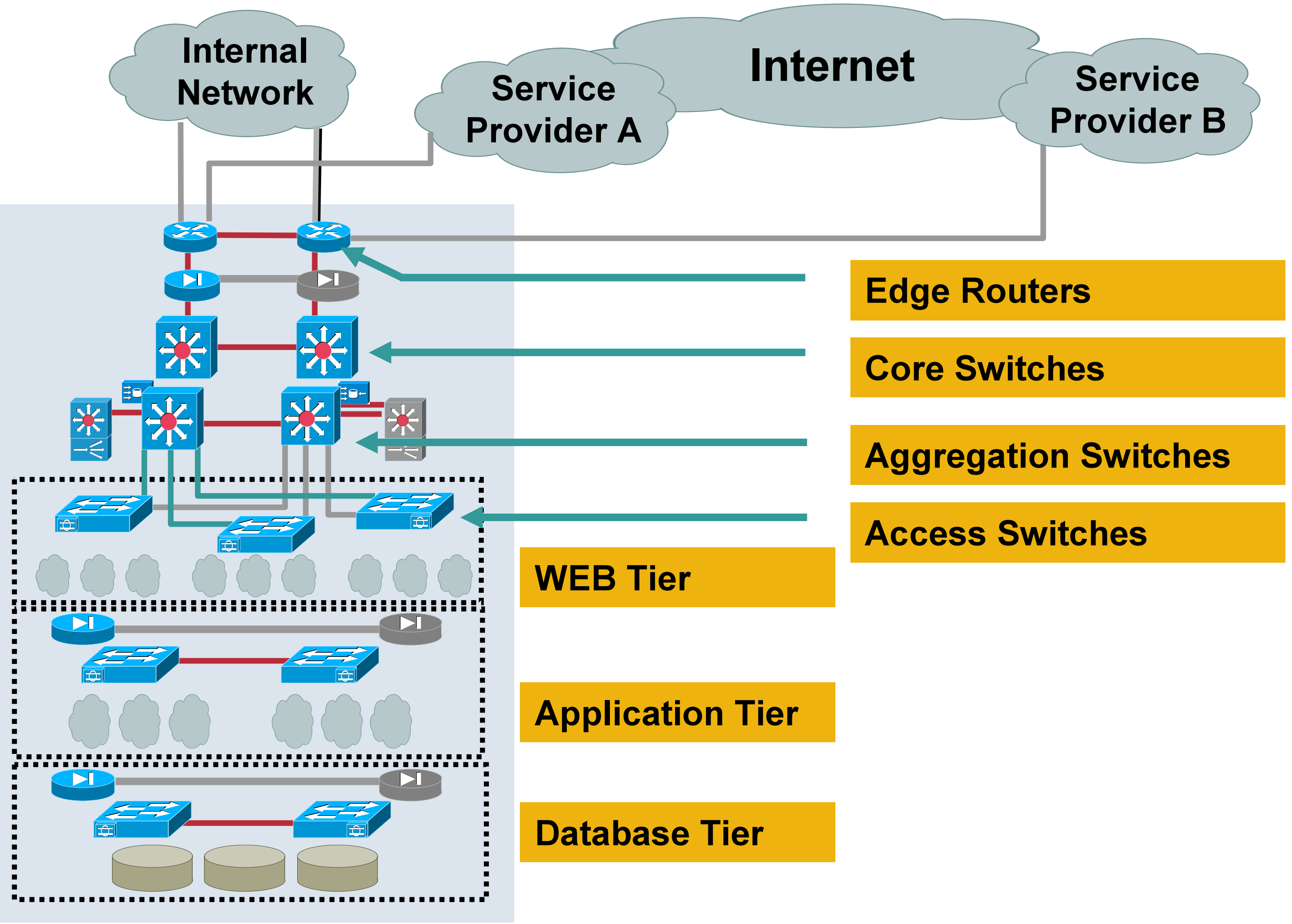
Linux/HP, Solaris/ SunFire, Oracle 10G RAC, etc.

Storage solution

MDS9000

- Redundancy
- Scalability
- Security

Typical Data Center Topology



Internal Network

Internet

Service Provider A

Service Provider B

Edge Routers

Core Switches

Aggregation Switches

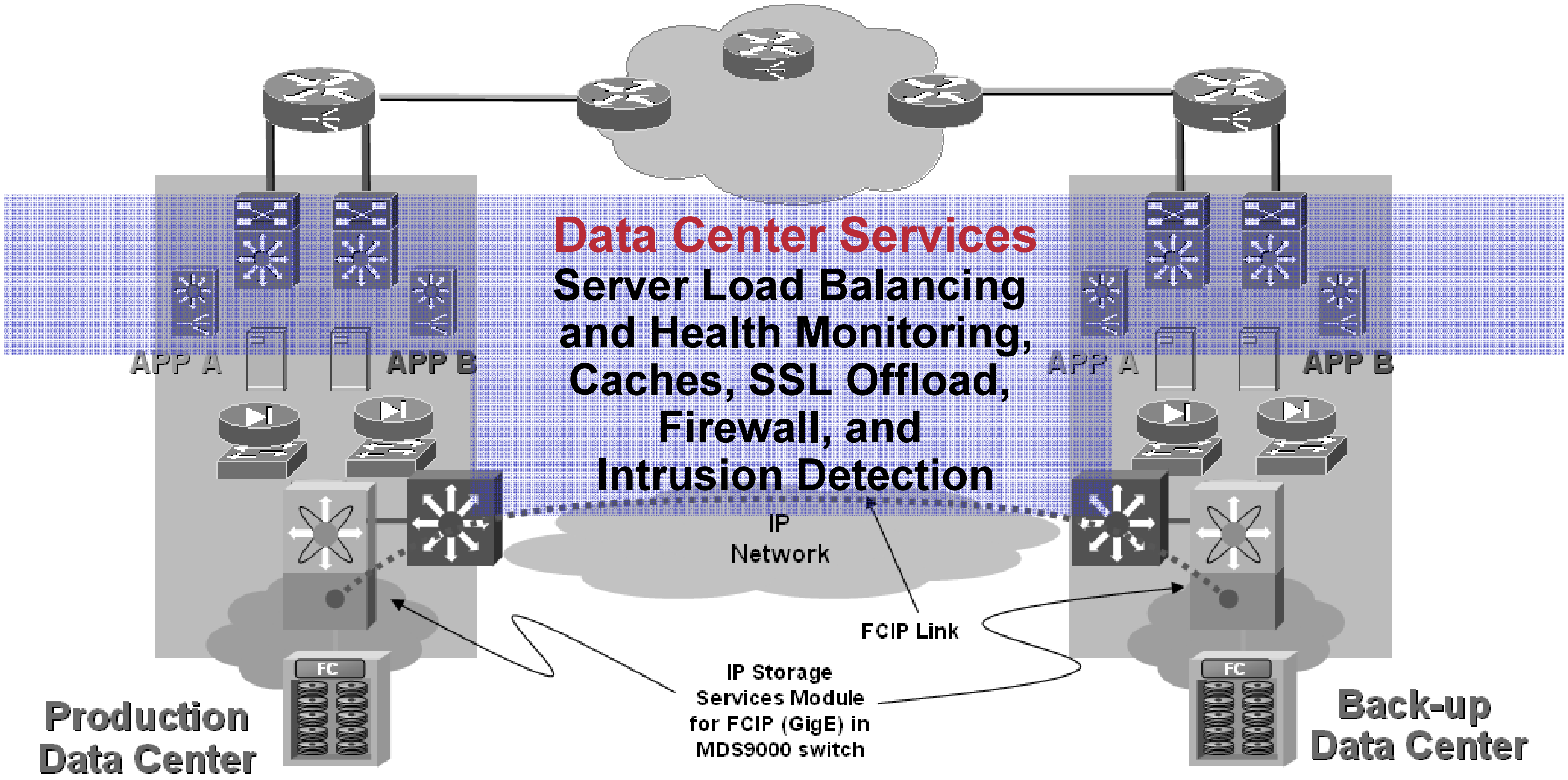
Access Switches

WEB Tier

Application Tier

Database Tier

Distributed Data Centers



Data Center Services

Server Load Balancing and Health Monitoring, Caches, SSL Offload, Firewall, and Intrusion Detection

SERVER LOAD BALANCING

Server Load Balancing

- a.k.a. 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
- Layer3 to layer7 content switching capabilities are available with extensive keepalives (server health checks) functionality
- Layer4 or layer7 proxy can be used as a security perimeter

Application Redundancy

Load Distribution

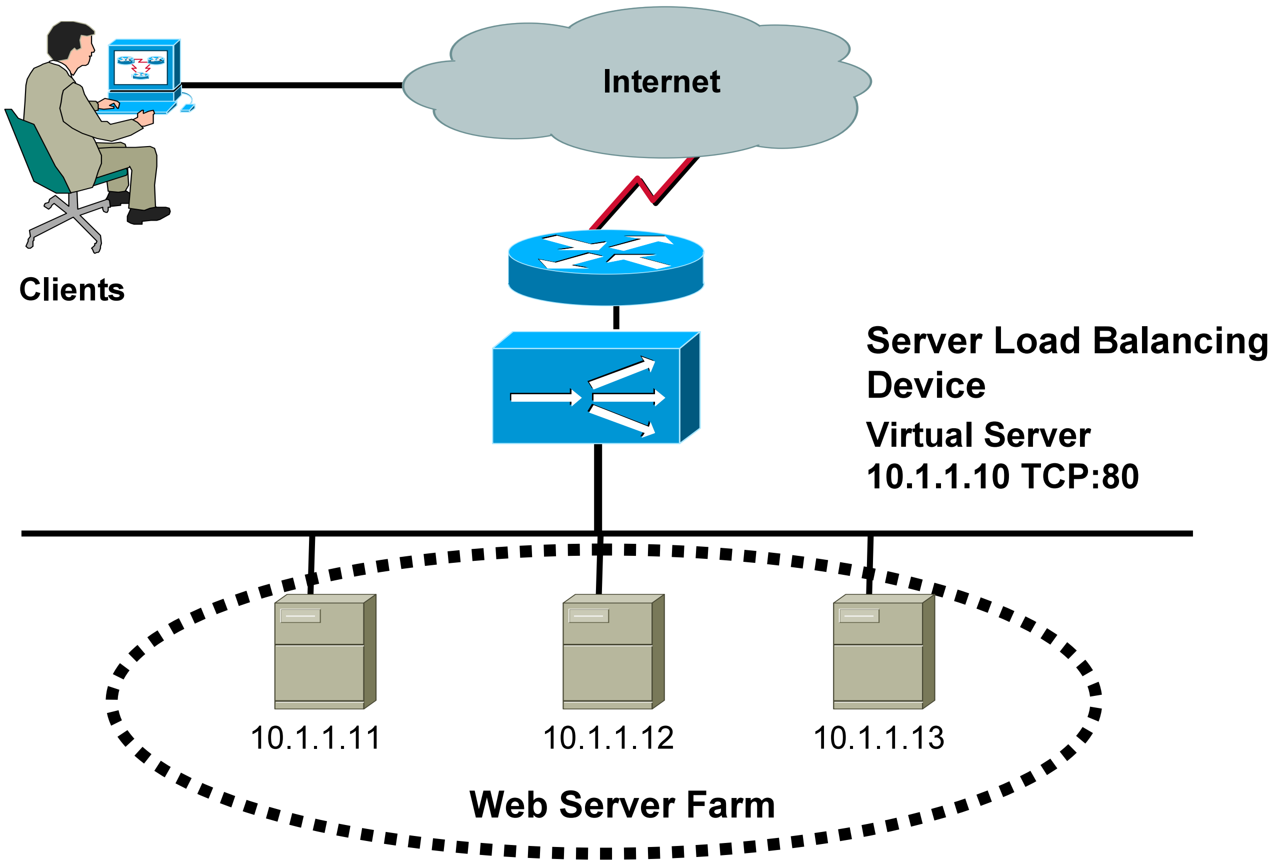
Application Health Checks

Communication of Load to GSLB Device

Content Switching Design Decisions

- Application protocol and ports (listener ports)
- End-to-end application flows
- Direct server access
- Server mgmt
- Server initiated sessions
- Infrastructure design

SLB Overview



Clients

Internet

Server Load Balancing Device

Virtual Server
10.1.1.10 TCP:80

10.1.1.11

10.1.1.12

10.1.1.13

Web Server Farm

Content Unaware SLB

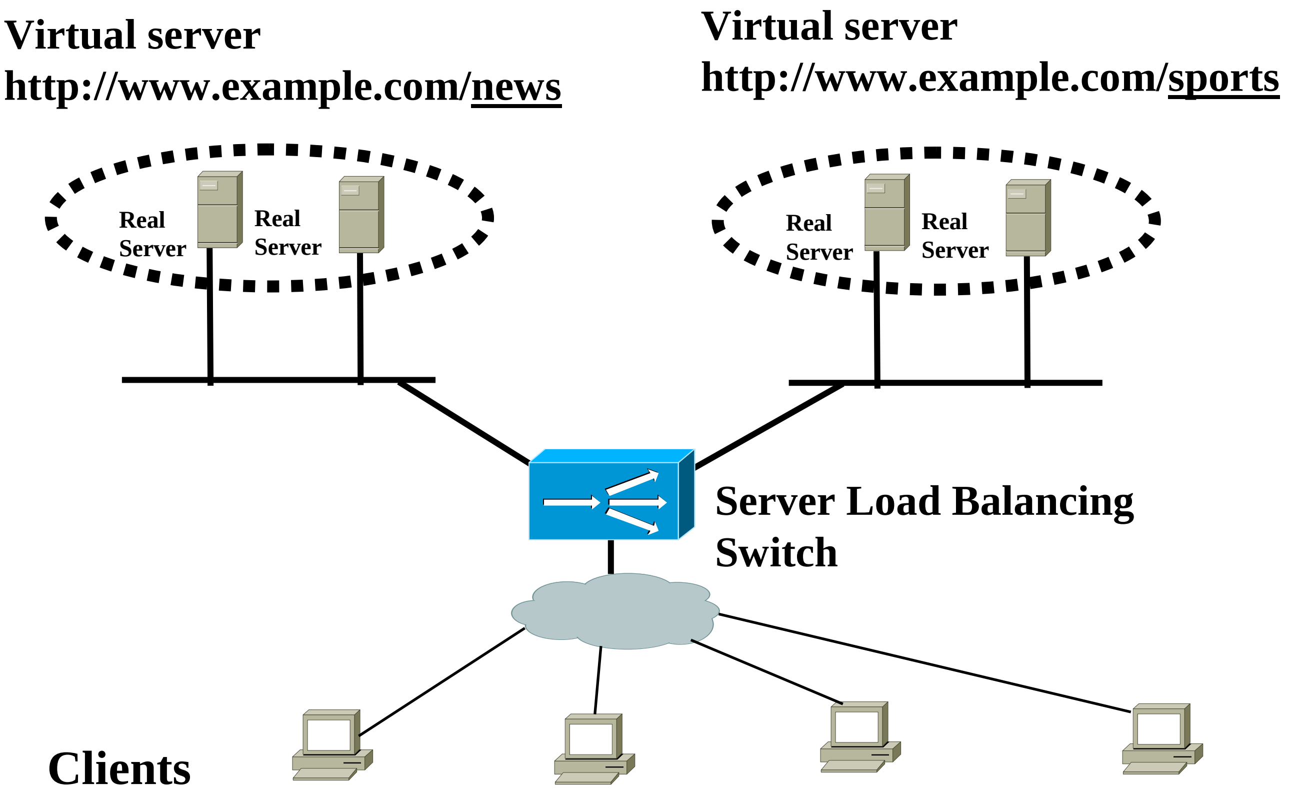
- Allows the balancing of traffic destined to a virtual server across multiple real servers
- Virtual Server / Content Rule = IP address (VIP) & L4 protocol & port
- Virtual server may have 1 to N real servers
- All real servers within a content rule must have the same content
- In the Simplest case, 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 & L5-7 info (URL)
- Virtual server is chosen by the longest URL match

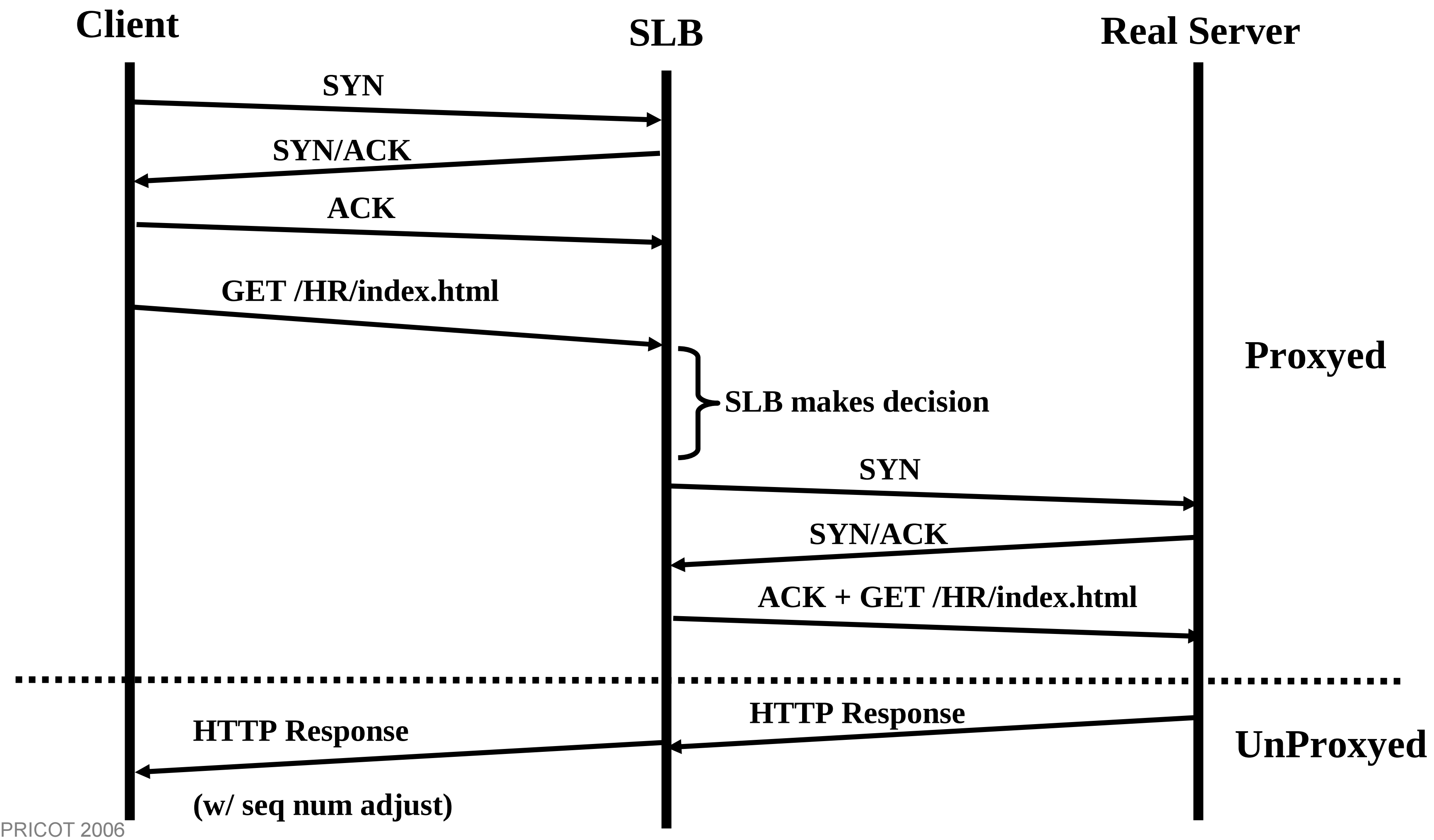
Why balance on URLs ?

- Distributed content


Virtual server
http://www.example.com/news
Virtual server
http://www.example.com/sports
Real Server
Real Server
Real Server
Real Server
Server Load Balancing Switch
Clients

URL Load balancing Packet Flow (HTTP 1.0)

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



Client

SLB

Real Server

SYN

SYN/ACK

ACK

GET /HR/index.html

Proxyed

SLB makes decision

SYN

SYN/ACK

ACK + GET /HR/index.html

HTTP Response

HTTP Response

UnProxyed

(w/ seq num adjust)

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

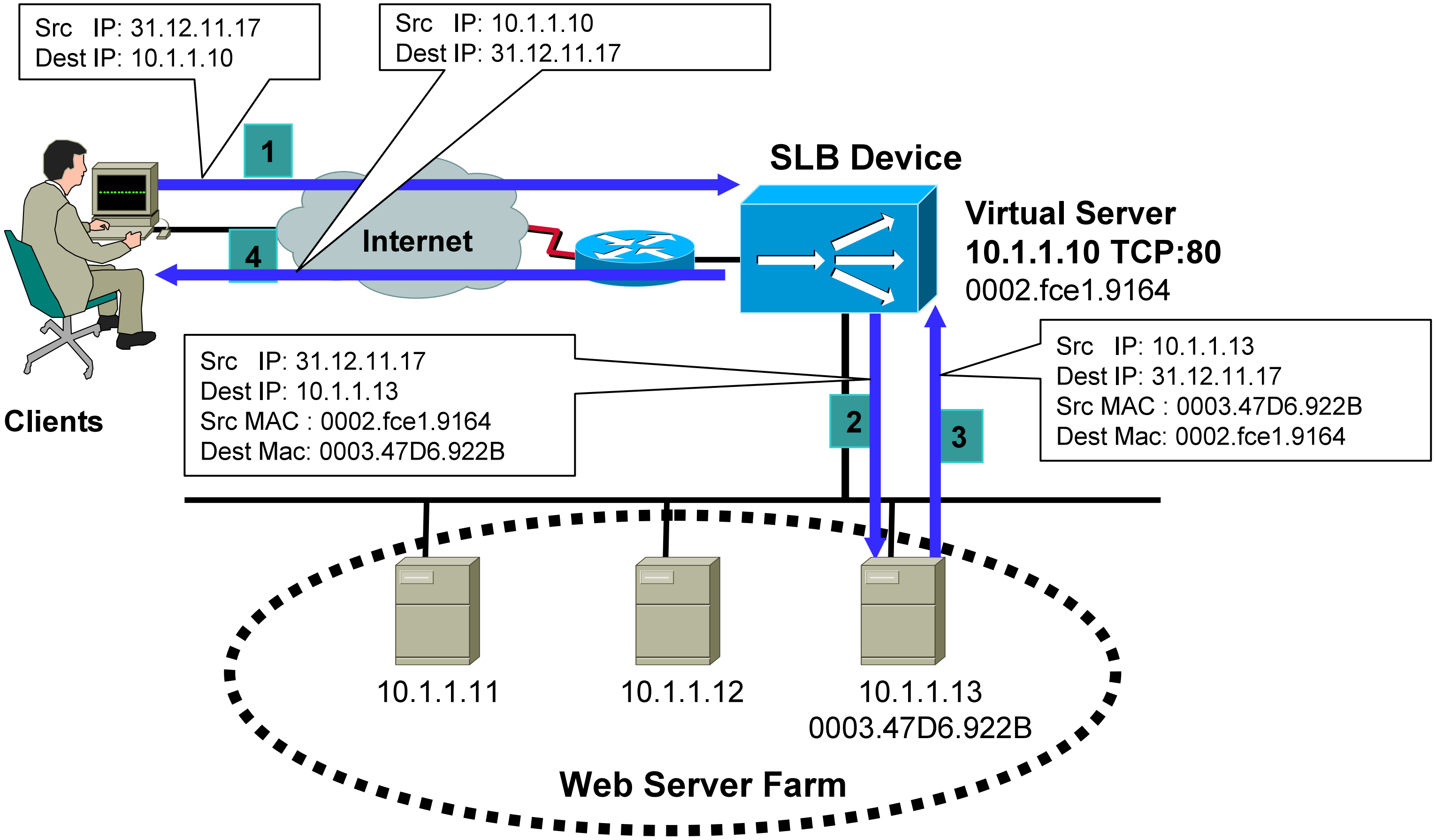
Directed (VIP Nat'd to real server IP)

- rewrites the IP address of traffic destined for the virtual server to be the real server IP address
- Web servers, APP servers

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



Src IP: 31.12.11.17
Dest IP: 10.1.1.10

Src IP: 10.1.1.10
Dest IP: 31.12.11.17

Src IP: 31.12.11.17
Dest IP: 10.1.1.13
Src MAC : 0002.fce1.9164
Dest Mac: 0003.47D6.922B

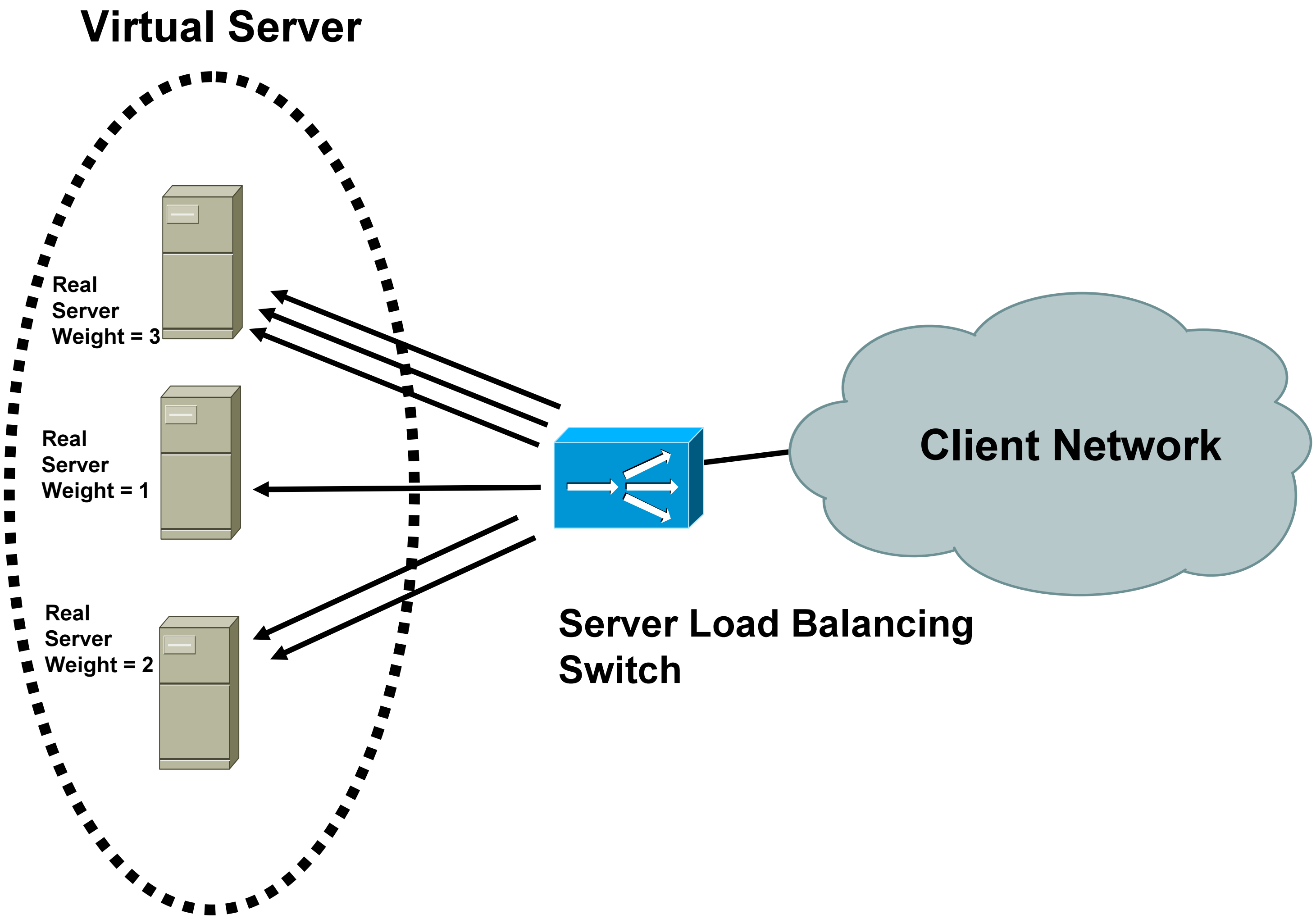
Src IP: 10.1.1.13
Dest IP: 31.12.11.17
Src MAC : 0003.47D6.922B
Dest Mac: 0002.fce1.9164

Load Balancing Algorithms

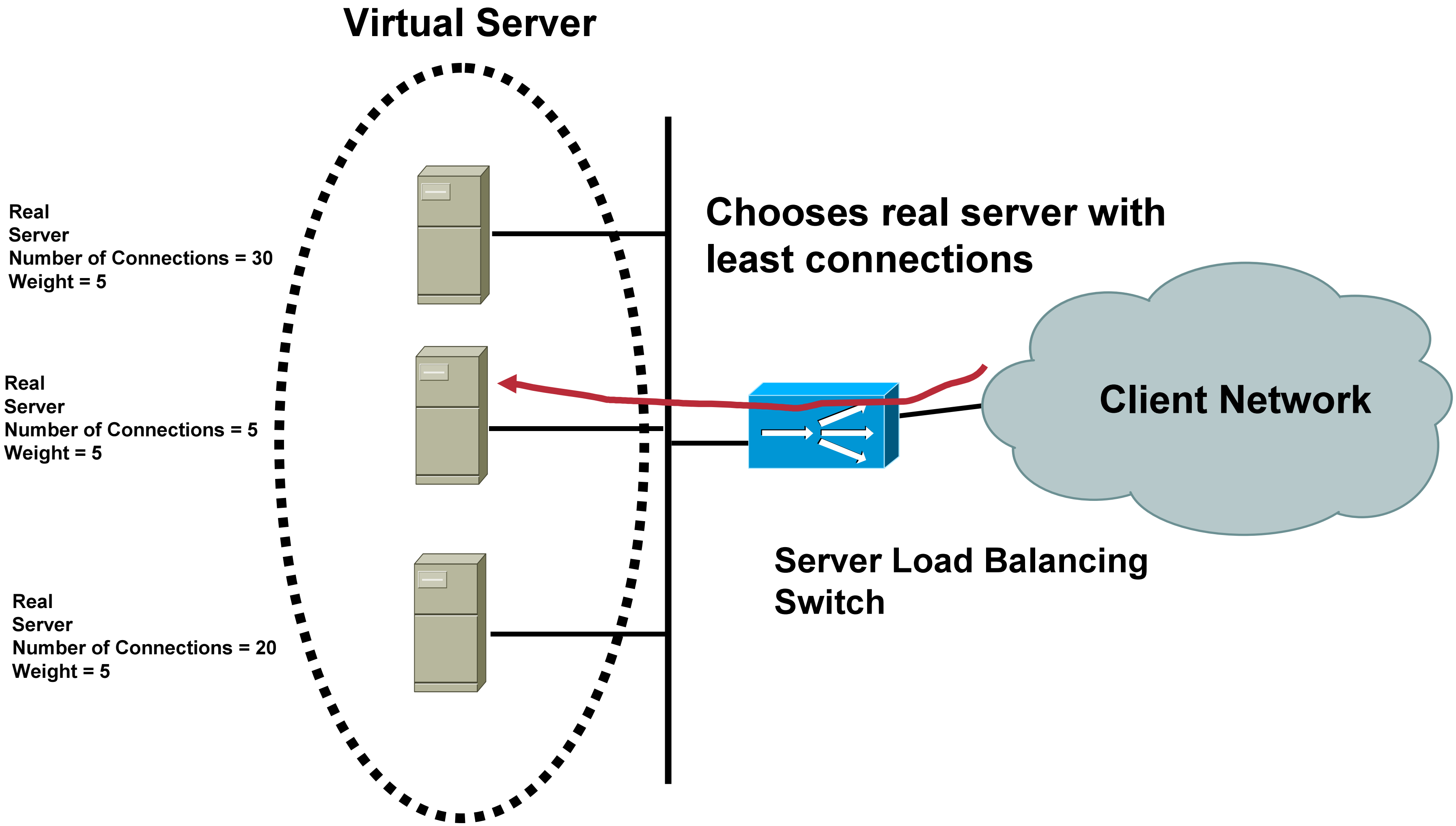
- urlhash
- domainhash
- weightedrr
- leastconn
- url

- domain
- srcip
- destip
- aca
- roundrobin

Weighted Round Robin



Least Connections



Virtual Server

Real
Server
Number of Connections = 30
Weight = 5

Real
Server
Number of Connections = 5
Weight = 5

Real
Server
Number of Connections = 20
Weight = 5

Chooses real server with least connections

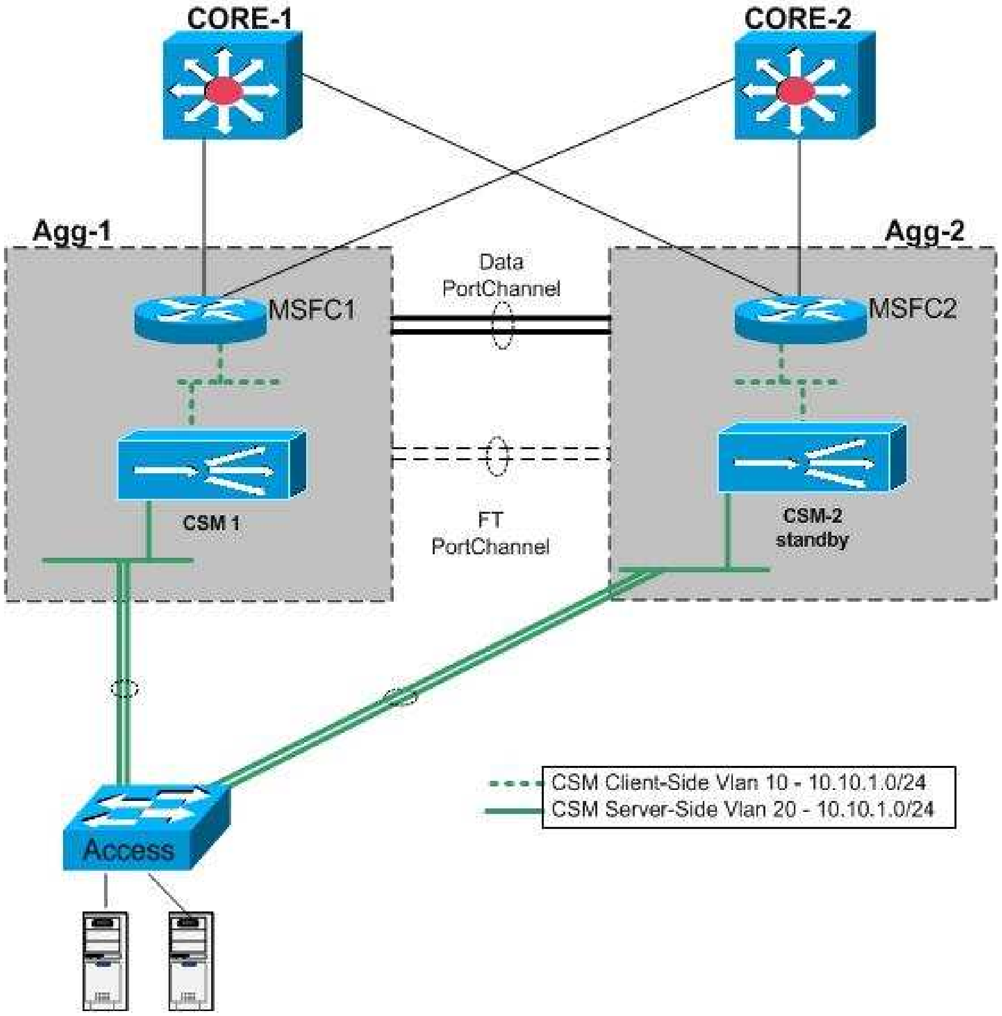
Client Network

Server Load Balancing Switch

“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



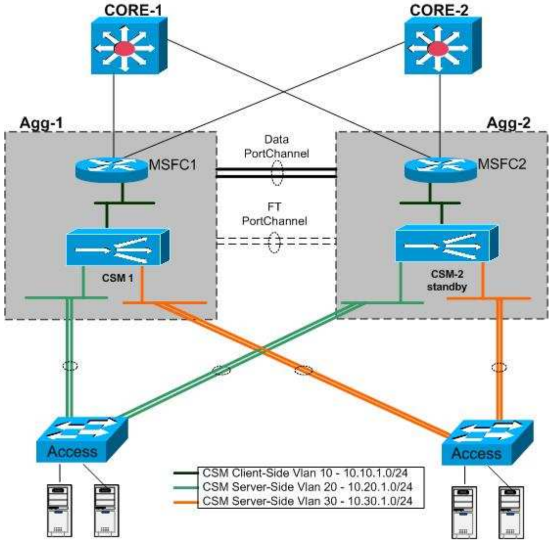
Key Content Switching Design Options

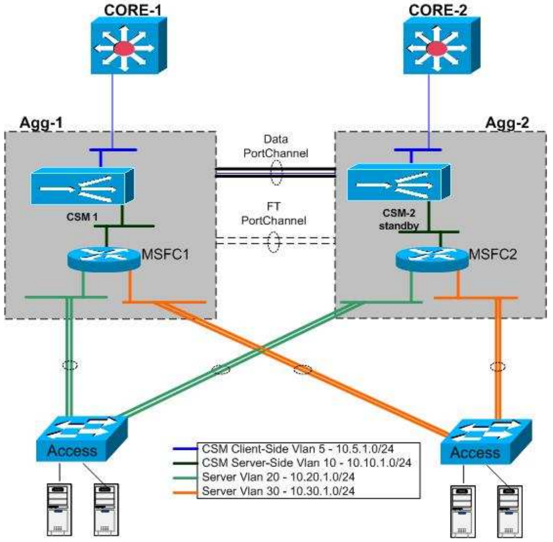
- Bridged Mode Design
- Routed Mode Design with MSFC on client side
- Routed Mode Design with MSFC on server side
- One-Armed Design

(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
- CSM inline of all traffic

Content Switching Design Approaches


CORE-1
CORE-2
Agg-1
Agg-2
Data
PortChannel
FT
PortChannel
MSFC1
MSFC2
CSM 1
CSM-2
standby
Access
Access
CSM Client-Side Vlan 10 - 10.10.1.0/24
CSM Server-Side Vlan 20 - 10.20.1.0/24
CSM Server-Side Vlan 30 - 10.30.1.0/24


CORE-1
CORE-2
Agg-1
Agg-2
Data
PortChannel
FT
PortChannel
CSM 1
CSM-2
standby
MSFC1
MSFC2
Access
Access
CSM Client-Side Vlan 5 - 10.5.1.0/24
CSM Server-Side Vlan 10 - 10.10.1.0/24
Server Vlan 20 - 10.20.1.0/24
Server Vlan 30 - 10.30.1.0/24

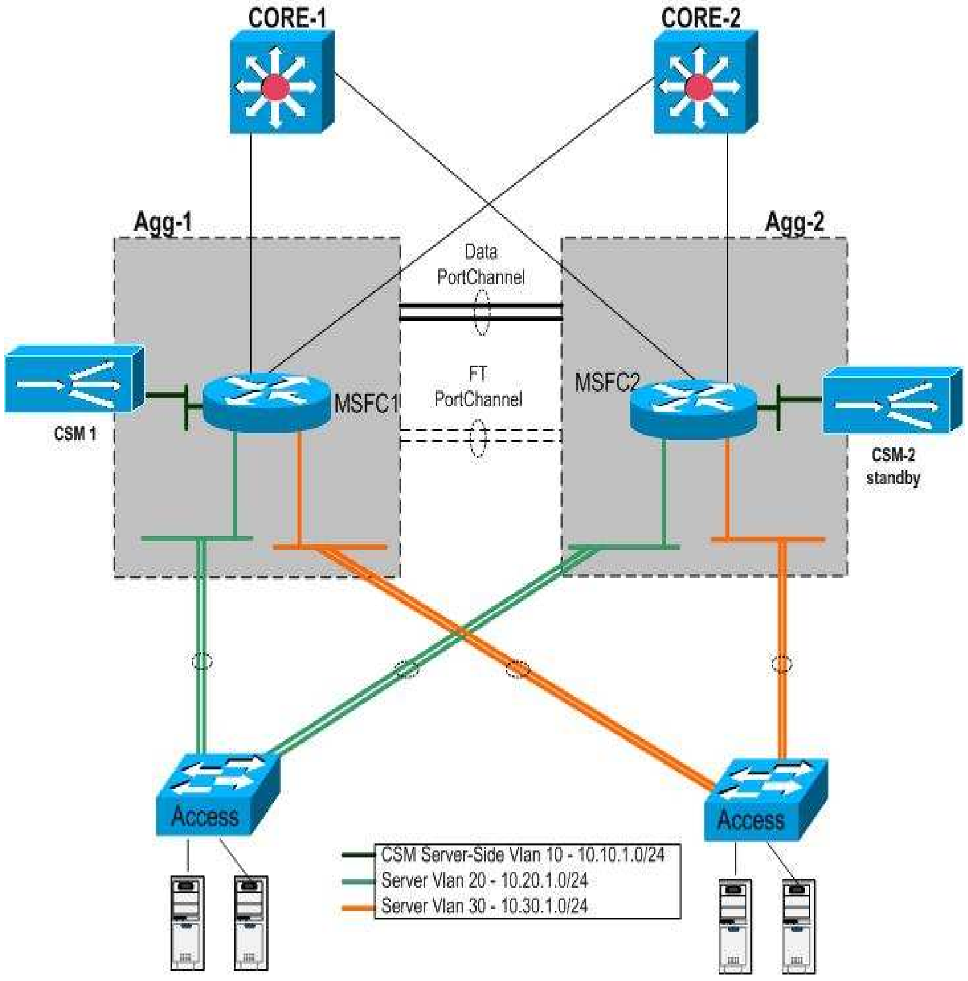
(2A) ROUTED MODE DESIGN WITH MSFC ON CLIENT SIDE

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

(2B) ROUTED MODE DESIGN WITH MSFC ON SERVER SIDE

- 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 CSM

Content Switching Design Approaches


CORE-1
CORE-2
Agg-1
Agg-2
Data PortChannel
FT PortChannel
MSFC1
MSFC2
CSM 1
CSM-2 standby
Access
Access
CSM Server-Side Vlan 10 - 10.10.1.0/24
Server Vlan 20 - 10.20.1.0/24
Server Vlan 30 - 10.30.1.0/24

(3) ONE-ARMED DESIGN CONSIDERATIONS

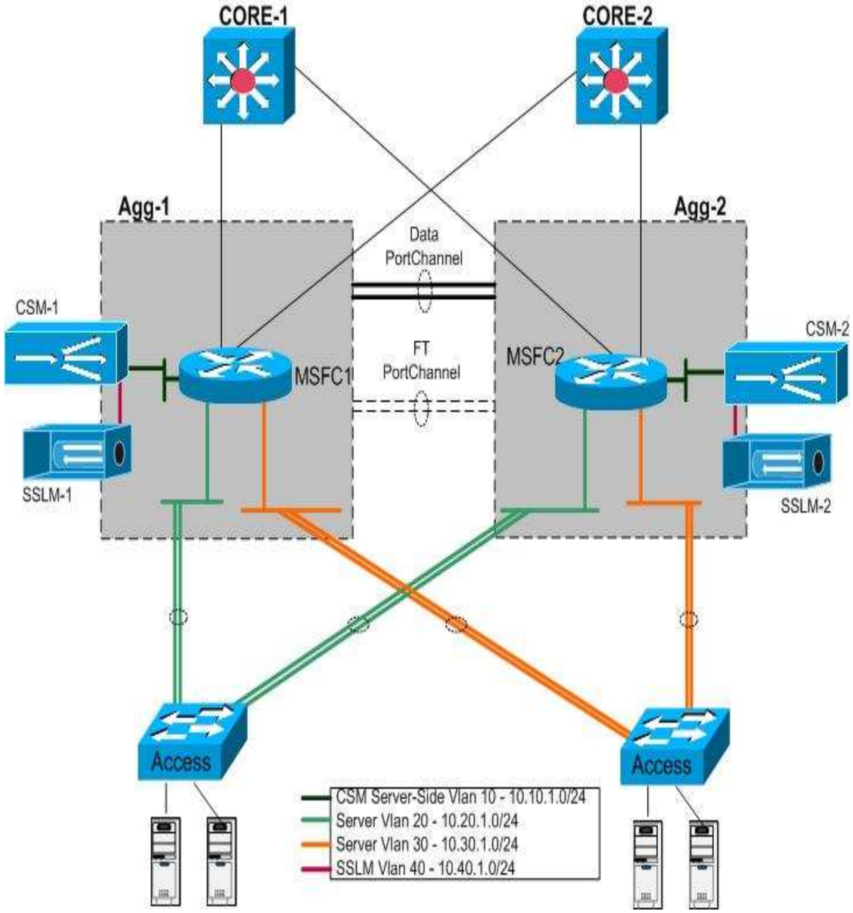
- Servers default gateway is the HSRP group IP address on the MSFC
- No extra configurations for:
  - Direct access to servers
  - Server initiated sessions
- RHI possible
- CSM inline for only server load balanced traffic
- Policy based routing or source NAT can be used for server return traffic redirection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

Content Switching Designs Summary

	(1) Bridge Mode	(2a) Routed Mode MSFC on Client Side	(2b) Routed Mode MSFC on Server Side	(3) One-Armed
Default Gateway of Servers	HSRP IP on MSFC	Alias IP on CSM	HSRP IP on MSFC	HSRP IP on MSFC
Direct Access to Servers	No extra configuration needed	Extra configuration needed	Extra configuration needed, may bypass CSM	CSM is bypassed
Servers Originated Connections	No extra configuration needed	Extra configuration may be needed	Extra configuration may be needed, may bypass CSM	CSM is bypassed
Multicast Support	Supported, bridges through	Not supported	Not supported, server to server works	Supported as CSM is bypassed
Layer 2 Loops	Possible if misconfigured	Not possible	Not possible	Not possible

SSL OFFLOAD

Network-Based SSL Offload


CORE-1
CORE-2
Agg-1
Agg-2
Data
PortChannel
FT
PortChannel
CSM-1
CSM-2
MSFC1
MSFC2
SSLM-1
SSLM-2
Access
Access
CSM Server-Side Vlan 10 - 10.10.1.0/24
Server Vlan 20 - 10.20.1.0/24
Server Vlan 30 - 10.30.1.0/24
SSLM Vlan 40 - 10.40.1.0/24

Key motivations

- Offload SSL decryption/encryption from servers
- Redundancy
- Scalability
- Unified mgmt 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 CSM
- SSLMs default gateway would be the alias IP on the CSM
- Back end SSL requires no design change

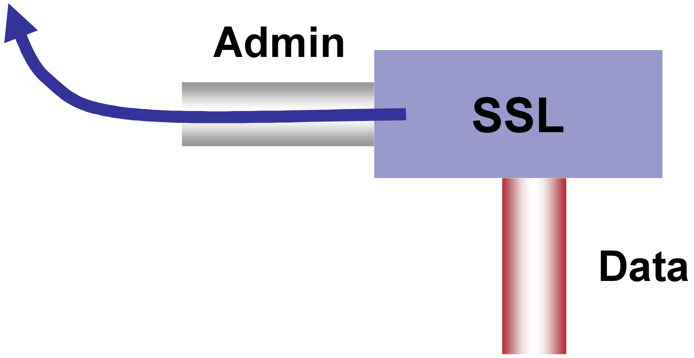
SSL Services Module
Configuration Tips: Admin VLAN and Data VLAN

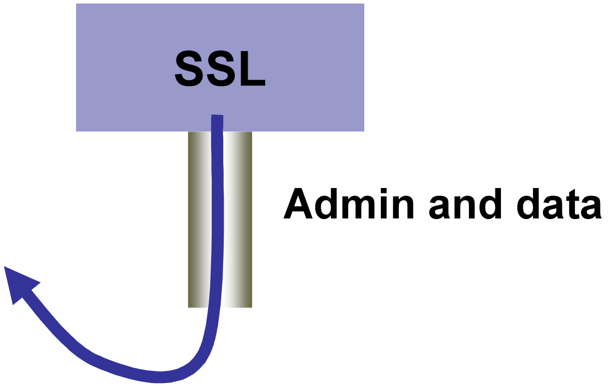
One VLAN on the SSL module has to be “admin VLAN”

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

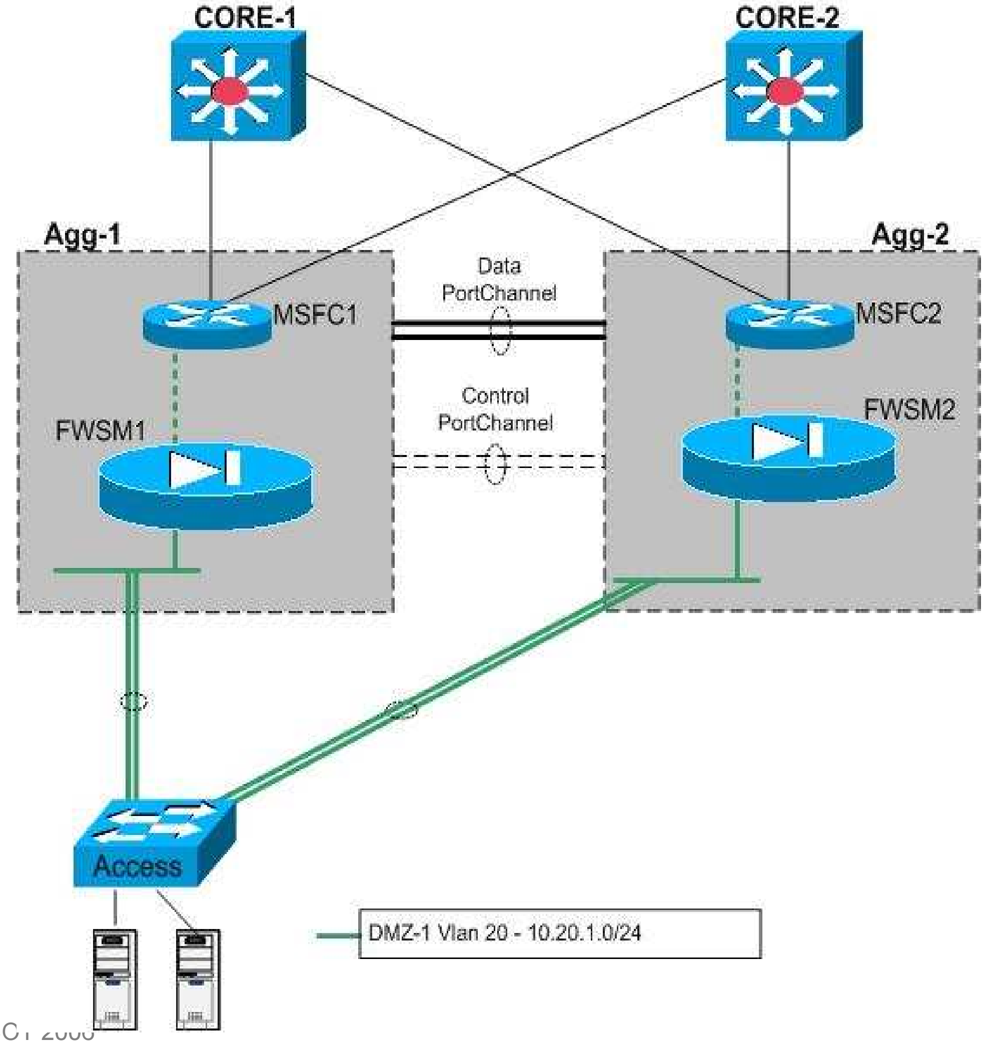
The default gateway of the admin VLAN is the module default gateway





DATA CENTER SECURITY

Firewall Design Approaches: Layer2



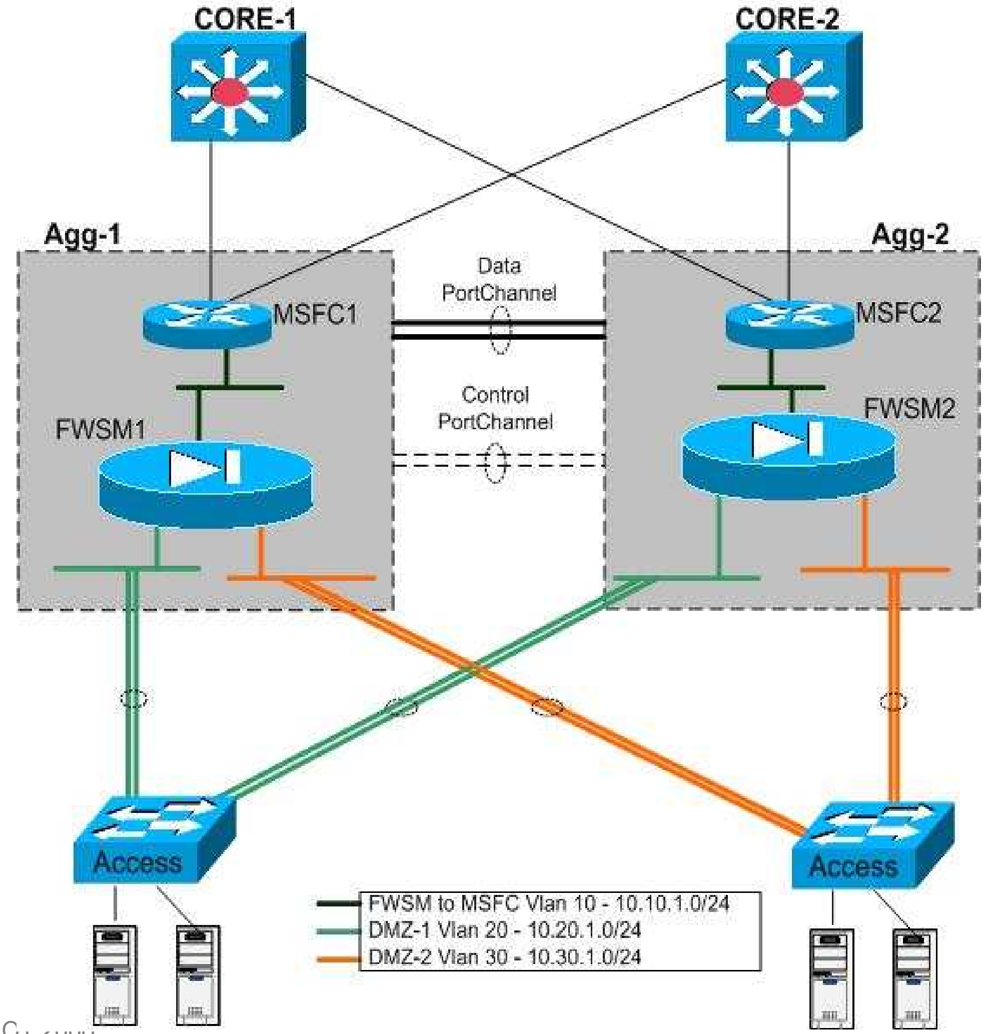
Key Firewall Design Options

- Bridged mode design, a.k.a. transparent or stealth firewall
- Routed mode design, a.k.a. layer3 firewall
- Virtual firewall contexts for L2 or L3 mode

(1) LAYER2 (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

Firewall Design Approaches: Layer3


CORE-1
CORE-2
Agg-1
Agg-2
MSFC1
MSFC2
Data PortChannel
Control PortChannel
FWSM1
FWSM2
Access
Access
FWSM to MSFC Vlan 10 - 10.10.1.0/24
DMZ-1 Vlan 20 - 10.20.1.0/24
DMZ-2 Vlan 30 - 10.30.1.0/24

(2) LAYER3 FIREWALL DESIGN CONSIDERATIONS

- Servers default gateway is the IP address on the firewall
- Dynamic routing is supported

Firewall Design Approaches: Virtual Context

- It’s the ability to segment a single physical firewall into multiple virtualized instances
- Multiple interfaces/VLANs within layer3 virtual contexts are supported

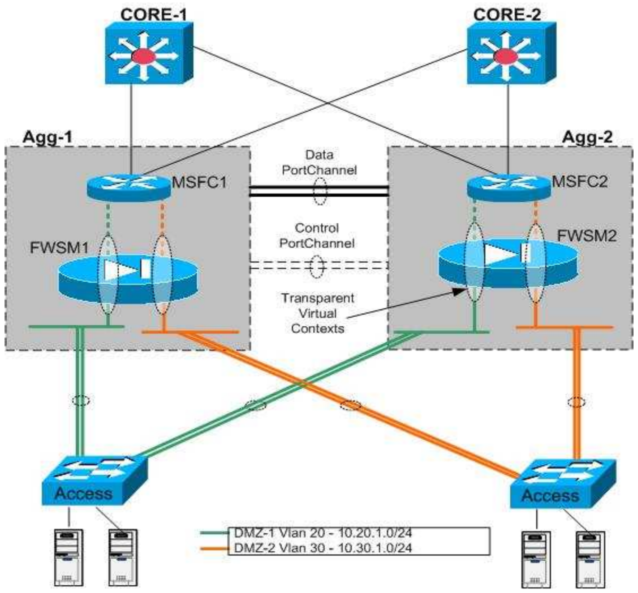
ON MSFC

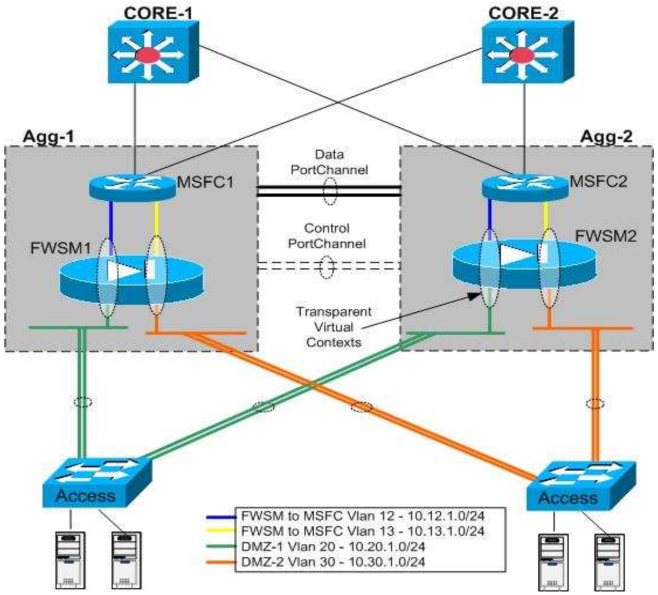
```
firewall multiple-vlan-interfaces
firewall module 7 vlan-group 100
firewall vlan-group 100  21-25,50-53
```

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




CORE-1
CORE-2
Agg-1
Agg-2
MSFC1
MSFC2
FWSM1
FWSM2
Data PortChannel
Control PortChannel
Transparent Virtual Contexts
Access
Access
FWSM to MSFC Vlan 12 - 10.12.1.0/24
FWSM to MSFC Vlan 13 - 10.13.1.0/24
DMZ-1 Vlan 20 - 10.20.1.0/24
DMZ-2 Vlan 30 - 10.30.1.0/24

(3A) TRANSPARENT CONTEXT

```
context FWA
  allocate-interface vlan2
  allocate-interface vlan20
  config-url disk:/FWA.cfg
!
context FWB
  allocate-interface vlan3
  allocate-interface vlan30
  config-url disk:/FWB.cfg
```

(3B) ROUTED CONTEXT

```
context FW1
  allocate-interface vlan12
  allocate-interface vlan20
  config-url disk:/FW1.cfg
!
context FW2
  allocate-interface vlan13
  allocate-interface vlan30
  config-url disk:/FW2.cfg
```

INTEGRATED DATA CENTER DESIGN OPTIONS

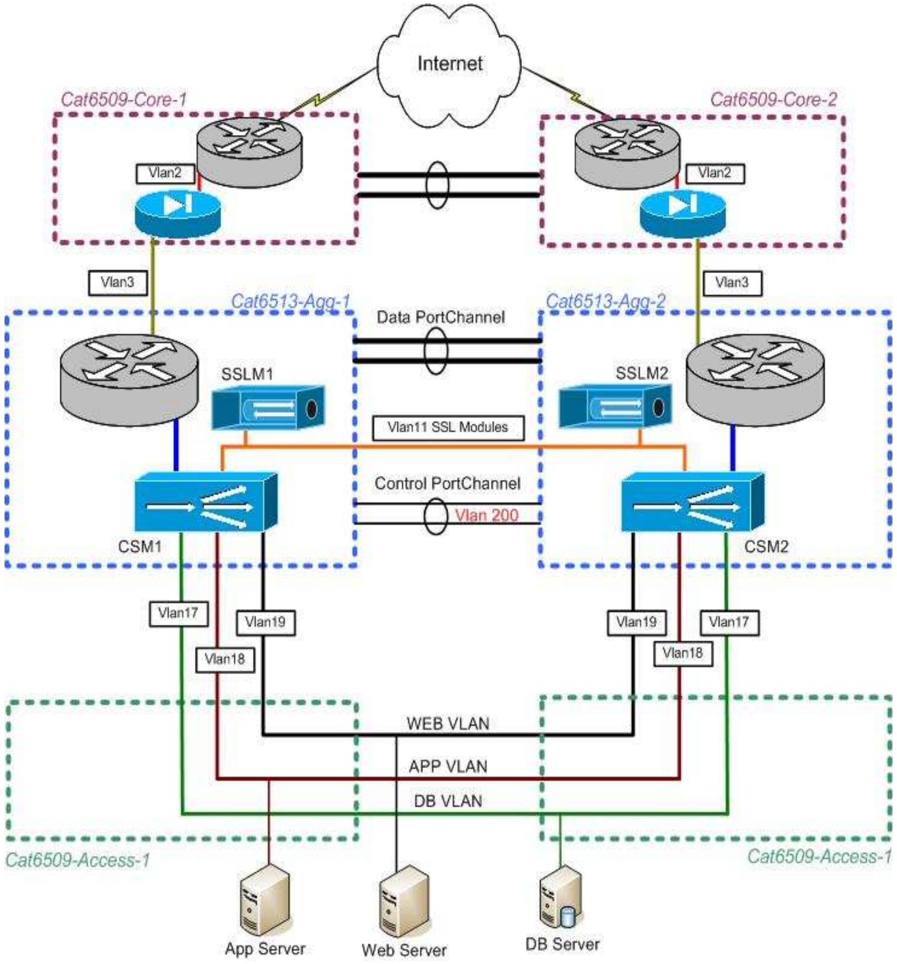
Data Center Services Design Options

- 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 layer3

(2) FW and CSM on aggregation with CSM in layer2 and FW in layer3

Design (1): Firewall on Core; CSM on Aggregation in Layer3 Mode



Security Details

- Layer3 firewall used
- Firewall perimeter at the core
- Aggregation and access are considered trusted zones
- Security perimeter 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-loadbalanced servers go through the CSM

Design (1): Firewall on Core; CSM on Aggregation in Layer3 Mode Configuration Snapshots

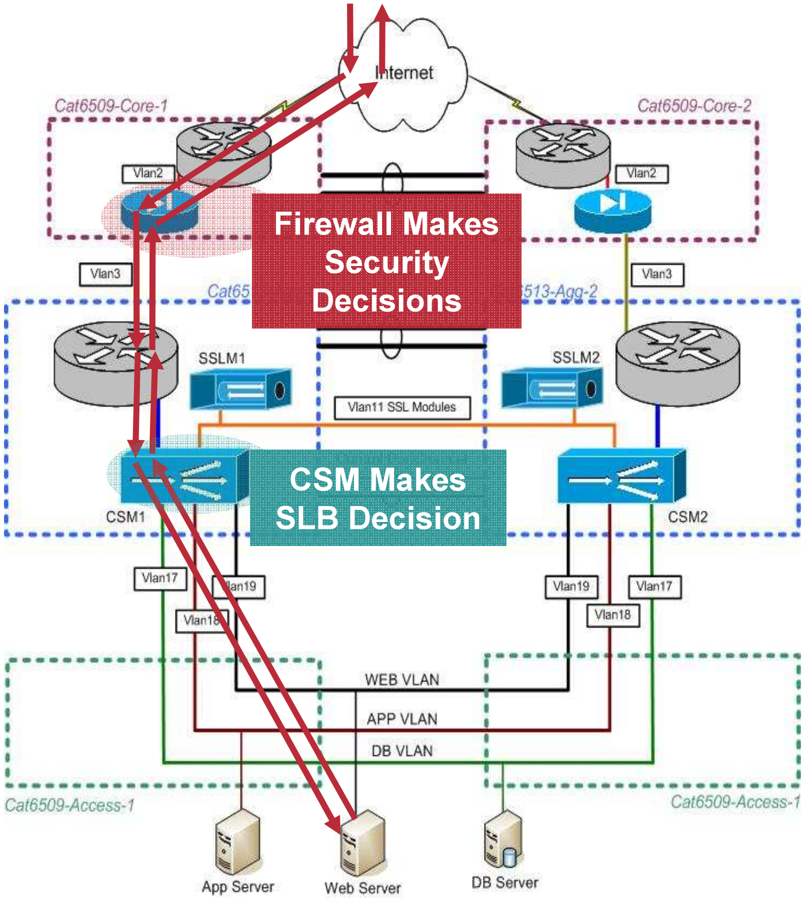
```
module ContentSwitchingModule 3
 vlan 16 client
  ip address 10.16.1.12 255.255.255.0
  gateway 10.16.1.1
  alias 10.16.1.11 255.255.255.0
!
 vlan 11 server
  ip address 10.11.1.2 255.255.255.0
  alias 10.11.1.1 255.255.255.0
!
 vlan 17 server
  ip address 10.17.1.2 255.255.255.0
  alias 10.17.1.1 255.255.255.0
!
 vlan 18 server
  ip address 10.18.1.2 255.255.255.0
  alias 10.18.1.1 255.255.255.0
!
 vlan 19 server
  ip address 10.19.1.2 255.255.255.0
  alias 10.19.1.1 255.255.255.0
```

```
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
```

```
 serverfarm ROUTE
  no nat server
  no nat client
  predictor forward
!
 vserver ROUTE
  virtual 0.0.0.0 0.0.0.0 any
  serverfarm ROUTE
  persistent rebalance
  inservice
```

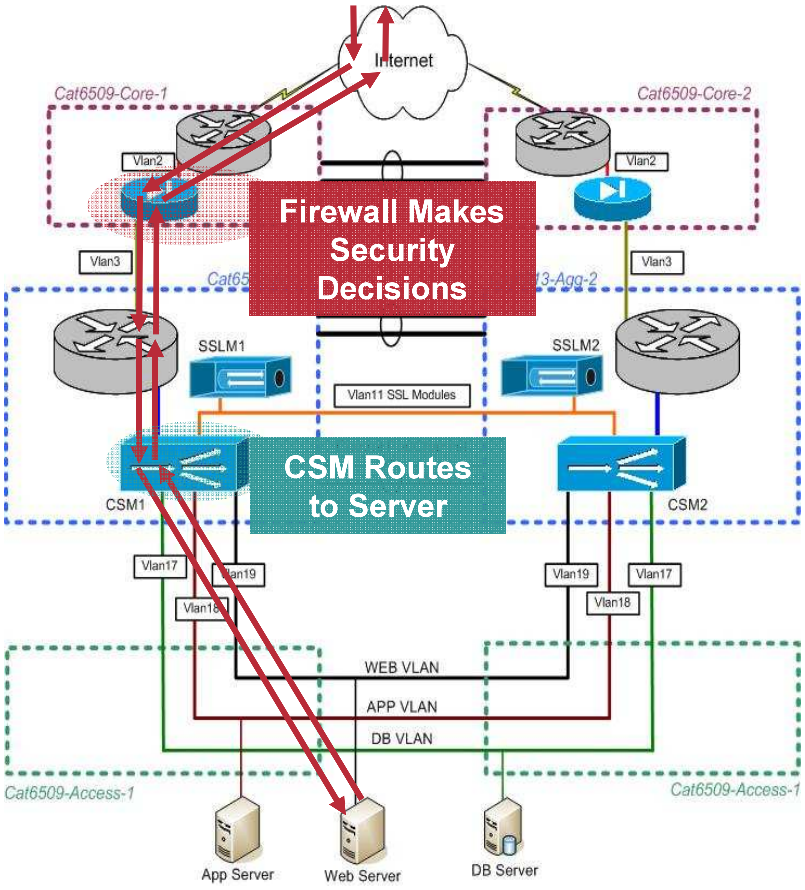
Design (1): Firewall on Core; CSM on Aggregation in Layer3 Mode: Session Flows



Firewall Makes Security Decisions

CSM Makes SLB Decision

Load Balanced Session Flow

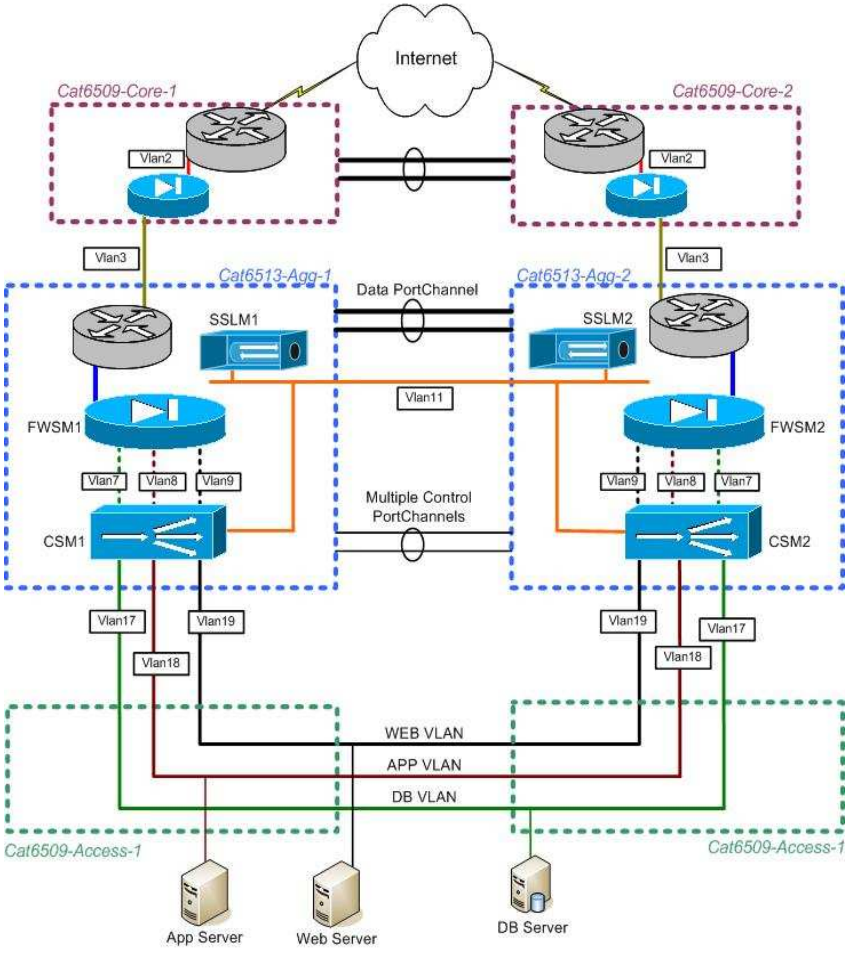


Firewall Makes Security Decisions

CSM Routes to Server

Server Mgmt Session Flow

Design (2): Firewall and CSM on Aggregation; FW in Layer3 and CSM in Layer2 Mode



Security Details

- Layer3 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-loadbalanced servers go through the CSM

Design (2): Firewall and CSM on Aggregation; FW in Layer3 and CSM in Layer2 Mode: 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 server
  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 server
  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)

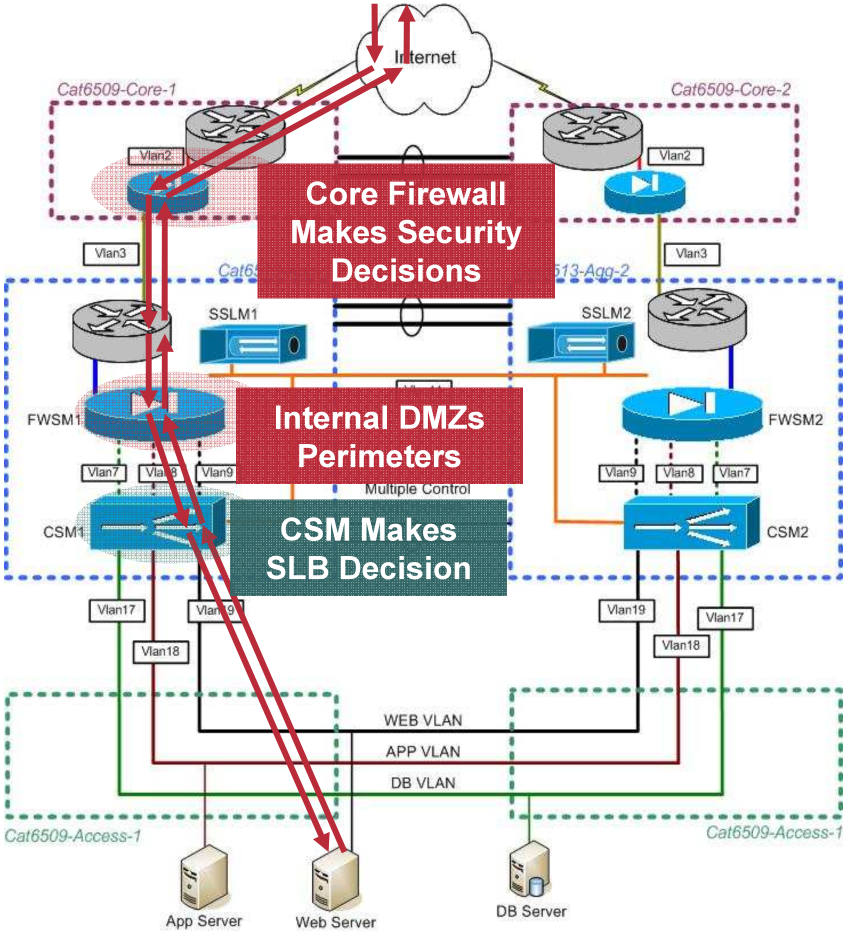
DMZ VLANs

VLAN7

VLAN8

VLAN9

Design (2): Firewall and CSM on Aggregation; FW in Layer3 and CSM in Layer2 Mode: Session Flows

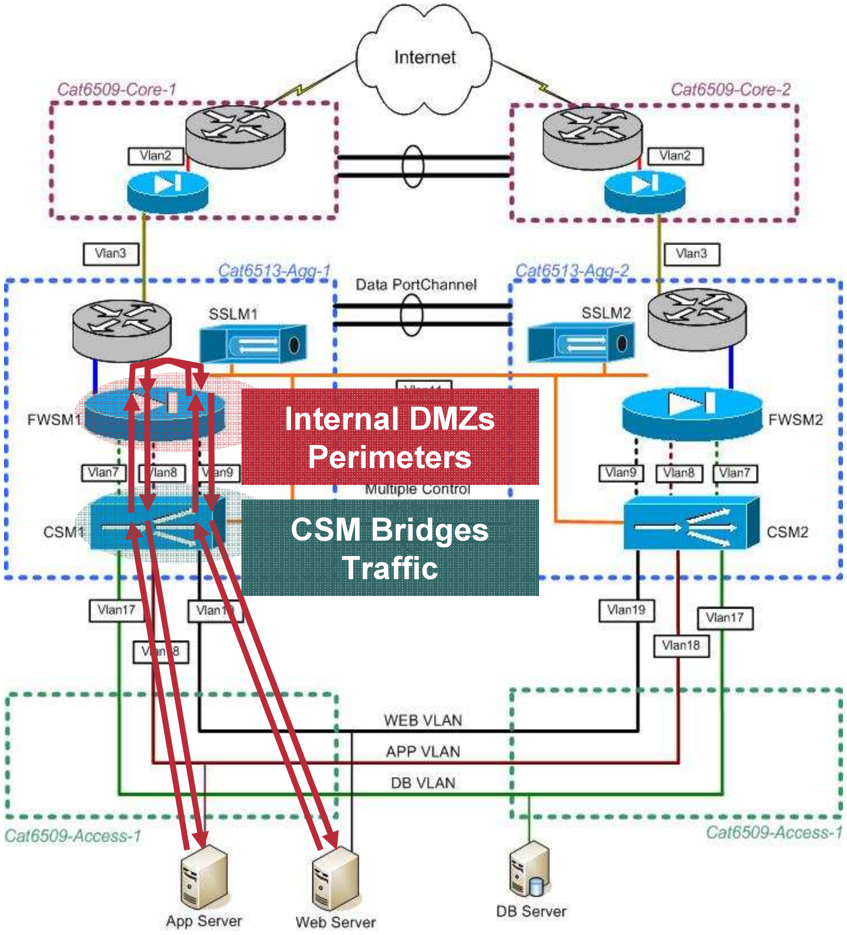


Core Firewall Makes Security Decisions

Internal DMZs Perimeters

CSM Makes SLB Decision

Load Balanced Session Flow



Internal DMZs Perimeters

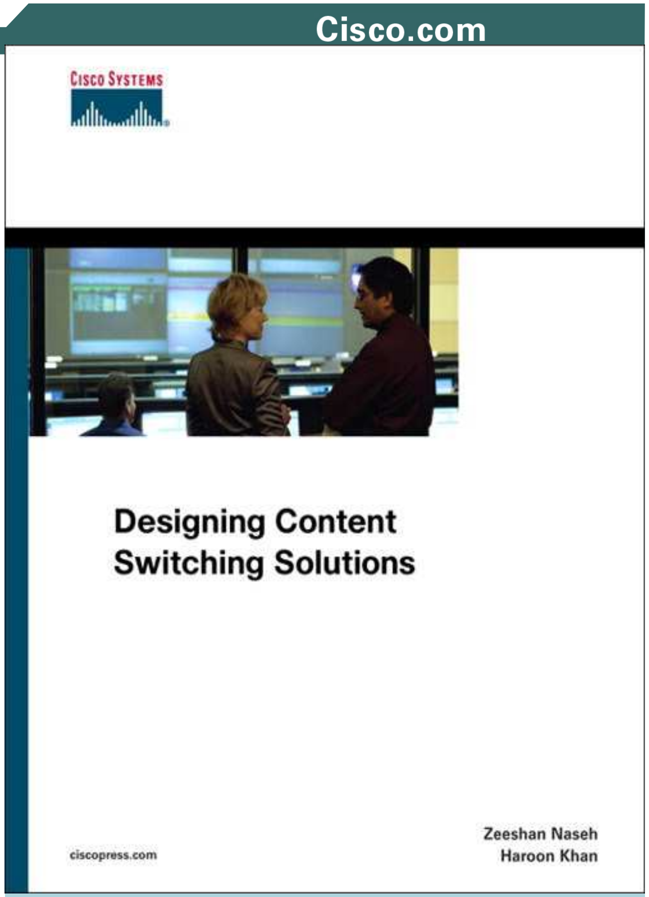
CSM Bridges Traffic

Web Server to App Server Session Flow

Recommended Reading

Designing Content Switching Solutions

ISBN: 158705213X



Q and A

CISCO SYSTEMS