

Optimizacija aplikacij v podatkovnem centru



Silvo Lipovšek

Cisco Application Delivery Networks

Network Classification

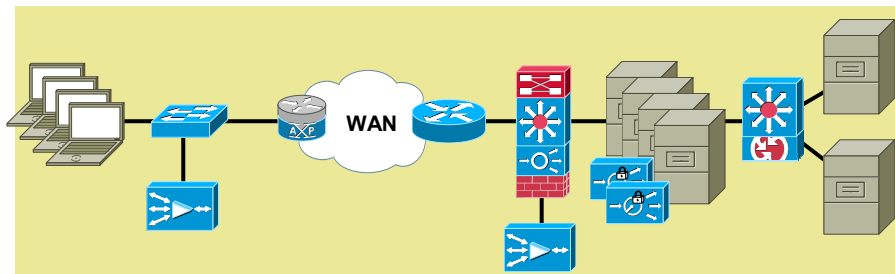
- Quality of service
- Network-based app recognition
- Queuing, policing, shaping
- Visibility, monitoring, control

Application Scalability

- Server load-balancing
- Site selection
- SSL termination and offload
- Video delivery

Application Networking

- Message transformation
- Protocol transformation
- Message-based security
- Application visibility



Application Acceleration

- Latency mitigation
- Application data cache
- Meta data cache
- Local services

WAN Acceleration

- Data redundancy elimination
- Window scaling
- LZ compression
- Adaptive congestion avoidance

Application Optimization

- Delta encoding
- FlashForward optimization
- Application security
- Server offload

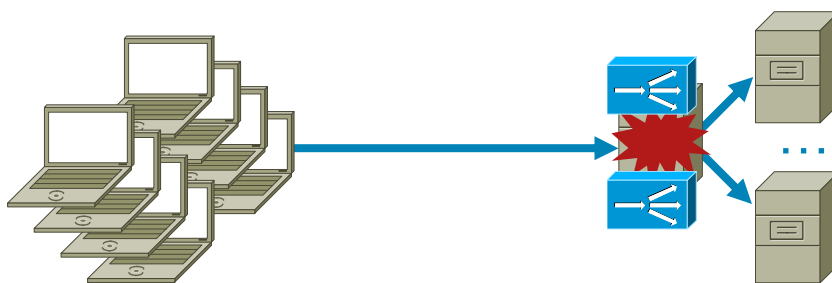
Agenda

1. Introduction
2. Load Balancing and Health Monitoring
3. Flow Management
4. Server Offload
5. Deployments

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

3

Scale and High Availability : The Hardware-Based Load Balancer



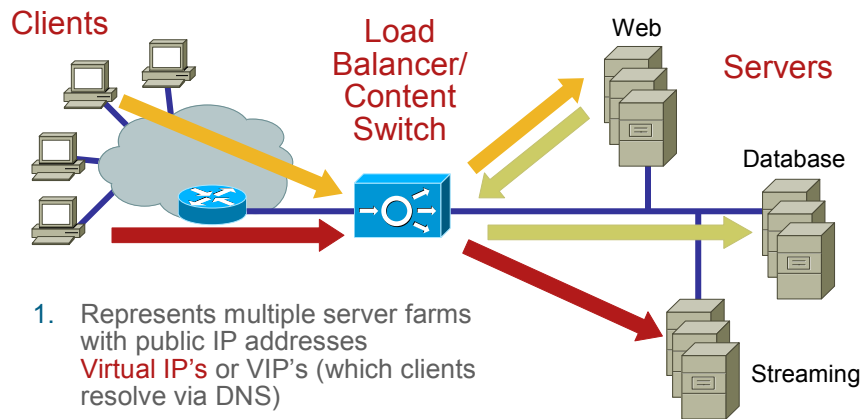
Benefit

1. Addresses fault tolerant, performance and scalability issues
2. Future-proof: architecture includes hardware co-processors to support resource-intensive features (i.e. SSL, compression)

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

4

The Main Functions of a Load Balancer



Devices Being Load balanced

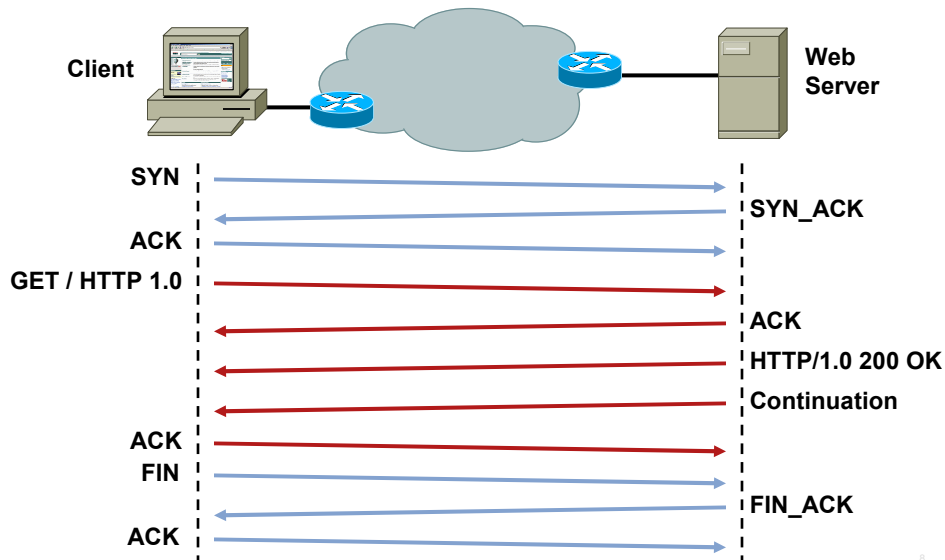
1. Servers
2. Proxies
3. Accelerators (compression engines, SSL offloaders)
4. Caches (reverse and transparent)
5. Firewalls (Layer 3 and Layer 2)
6. VPN concentrators
7. Routers
8. Generic IP device requiring load distribution and/or redundancy

Traffic Being Load balanced

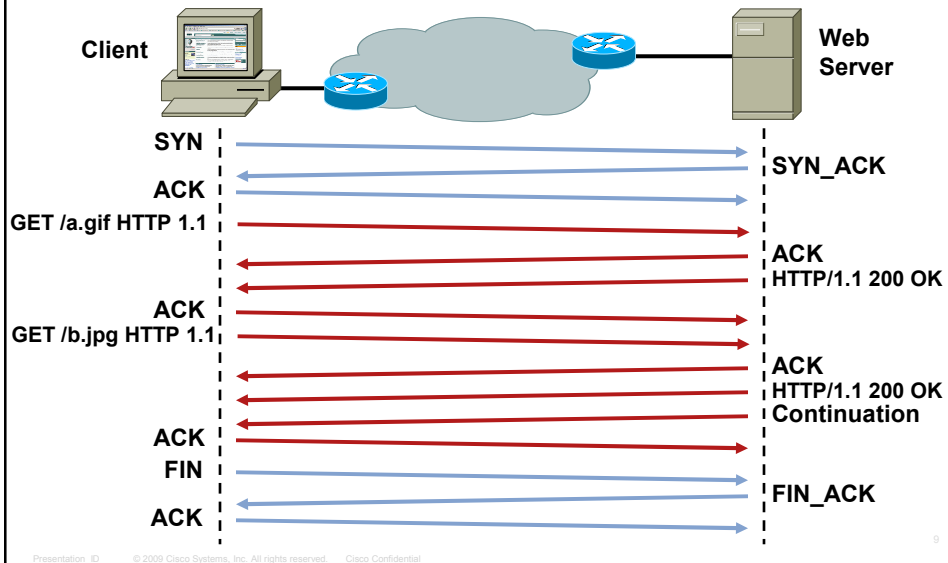
1. Generic IP traffic (i.e. IPsec tunnels)
2. Generic UDP and TCP (i.e. proprietary protocols)
3. Network services (i.e. LDAP, DNS, Radius)
4. HTTP (i.e. Web Presentation Layer, Web Services, SOAP/XML)
5. Voice & Video (i.e. RTSP, SIP, H.323)
6. Remote terminals (i.e. Windows Terminal Services)
7. Multi-connection protocols (i.e. FTP, RTSP)
8. Multi-tier packaged applications (i.e. SAP, Oracle, Microsoft, BEA)
9. Vertical specific applications (i.e. medical, finance, education)



HTTP 1.0—Single Request



HTTP 1.1—Two Requests, No Pipelining



Load Balancing and Content Switching Basic Requirements

1. **High-availability**
Works around server, application and network failures
 2. **No single point of failure**
Failover is transparent to clients
 3. **Disaster recovery**
Fails over across Data-Centers
 4. **High and scalable performance**
Can serve growing number of clients, with more content and transactions
 5. **Intelligent content and load-based decisions**
Selection of the best server
 6. **Transaction assurance**
Entire transaction sent to the same server
 7. **Security**
Protect self, servers and applications
 8. **Flexibility**
Adapt to network topologies and application environments
- Presentation ID: © 2005 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Advanced Requirements: from Load Balancing to Application Delivery

1. **Server offload**
Free up server CPU and resources
2. **Application acceleration**
Better user experience, faster transactions
3. **Bandwidth reduction**
Efficient WAN resources utilization
4. **Application and protocol inspection**
Protection against sophisticated application-specific attacks
5. **Virtualization**
One physical device behaves as many: maximum deployment flexibility and separation of resources
6. **Flexible network management**
Allows multiple users, with different responsibilities, to simultaneously manage the device

11

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Load Balancing algorithms and Health Monitoring

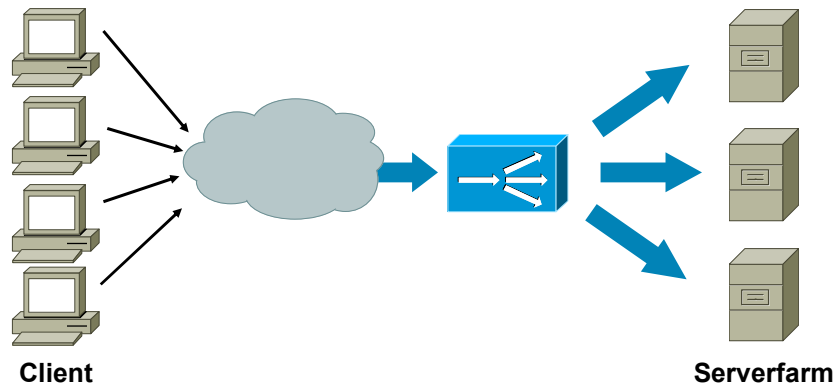


How Connections are Distributed to the Best Available Servers

12

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Load Balancing Algorithms



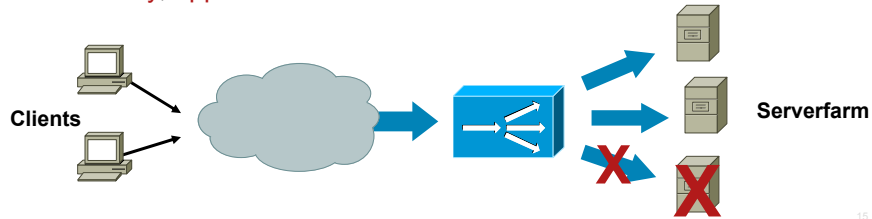
How to Distribute Requests Across Servers ?
Enhanced Predictors Improve Serverfarm Efficiency

Load Balancing Algorithms

1. **Round Robin:** (weighted)
Very simple
2. **Least Connections:** (weighted)
Dynamic, requires slow-start
3. **Hash on IP:** (source/destination, with mask)
No state required for stickiness issues with dynamic changes
4. **Hash on URL:** or portion of URL
5. **Server watermarks:** min and max number of connections per server
6. **Least loaded:** SNMP OIDs based server feedback for obtaining useful information maintained as SNMP Object IDs
7. **Least bandwidth:** connection vs. bandwidth based on the bidirectional traffic flow
8. **Adaptive response predictor:** load-balancing based on server response time
SYN to SYN-ACK
SYN to FIN
Application request to first packet of response

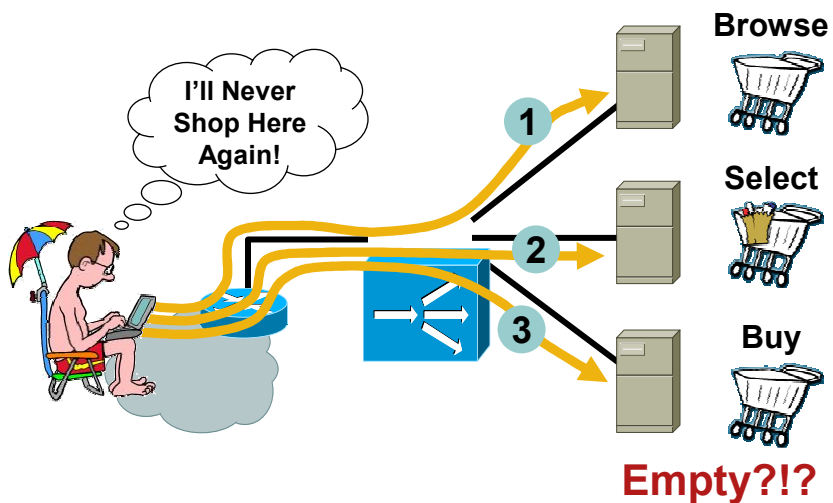
Health Checking

1. The content switch needs to **continuously monitor the back-end servers**
2. Failed servers have to be identified and **removed from rotation**: the load balancing algorithms adapt to the change
3. Server failures should be transparent to clients
4. Servers recovering from failures should be checked and put back in the available pool, **avoiding flapping**
5. Any failure affecting client-server interaction should be detected: **connectivity, application or back-end servers malfunctions**



Session Persistence—Stickiness

The “Shopping Cart” Problem



Session Persistence—Stickiness

1. **Session**: logical aggregation of multiple simultaneous or subsequent connections
2. Sessions are limited in time (timeout)
3. **Servers might keep session state locally**
4. Load distribution across multiple servers introduces the problem

The content switch needs to **identify a session** and send connections belonging to the **same session** (i.e. from the same client) to the **same server**

Methods to identify the session or client:
Source IP address, HTTP session cookie, SIP session ID, SSL ID, generic protocol session data, ...

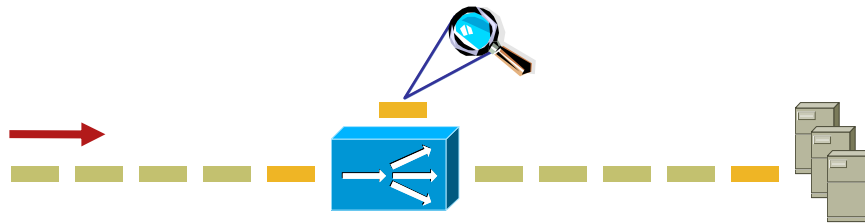
Flow Management



Layer 4 and Layer 7 processing

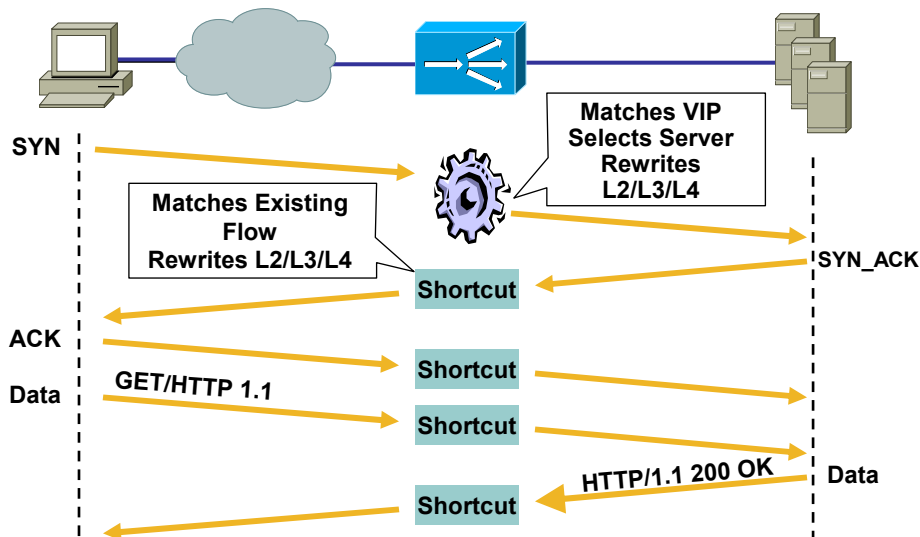
Layer 4 Switching

1. L2–L4 information is always present in the first packet of the flow (unless it is a fragment!)
 - IP protocol
 - Source/destination IP addresses
 - Source/destination L4 ports (for TCP/UDP)
 - Source VLAN, MAC address
2. The load balancing decision can be made on the first packet



19

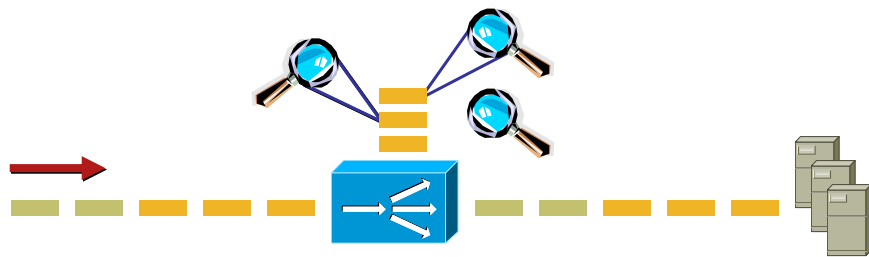
Layer 4 Flow Setup—Basic Load Balancing Decisions Made on First Packet



20

Layer 7 Switching

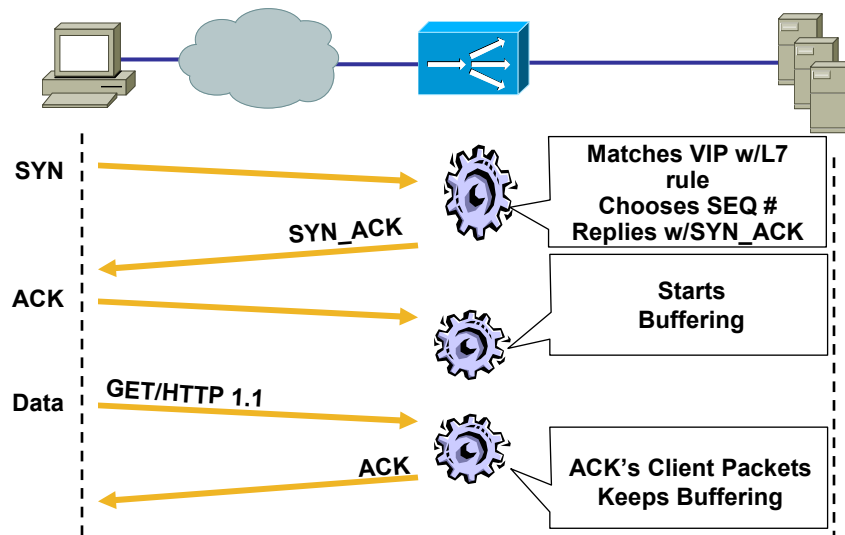
1. L5-L7 information is only received after the TCP setup and might span multiple packets
 - HTTP URLs, cookies, header fields
 - SSL session ID
 - FTP data channel port
 - Generic application data
2. Requires TCP termination and buffering!



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

21

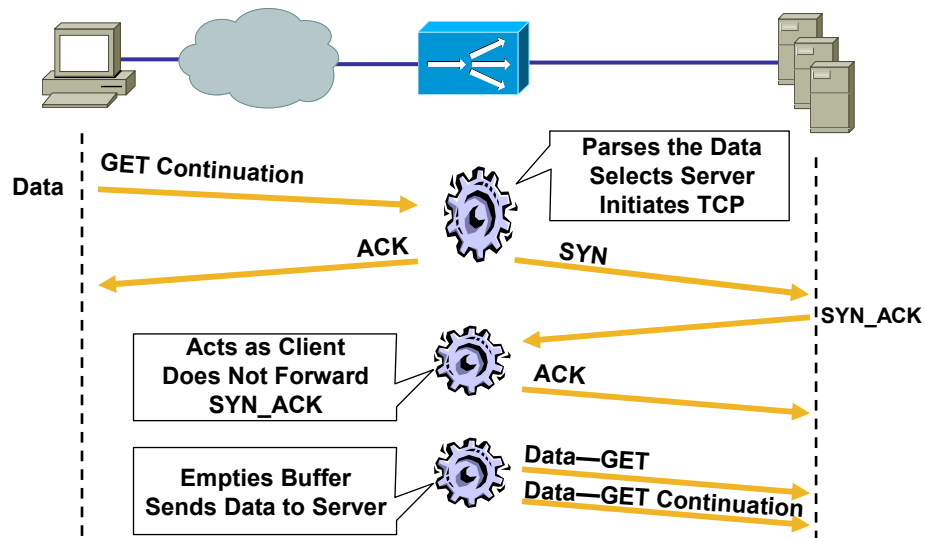
Layer 7 Flow Setup for HTTP (1/3) Load Balancing Decisions Require More Data



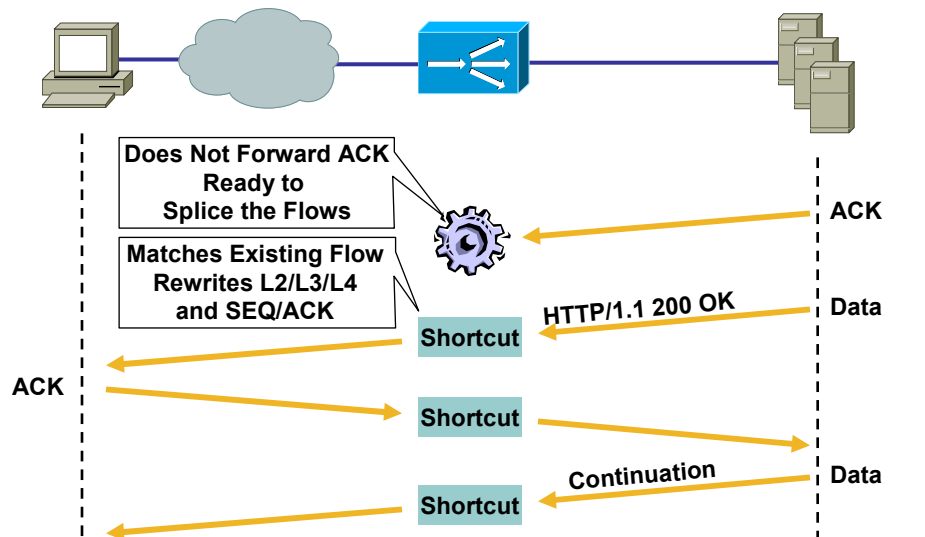
Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

22

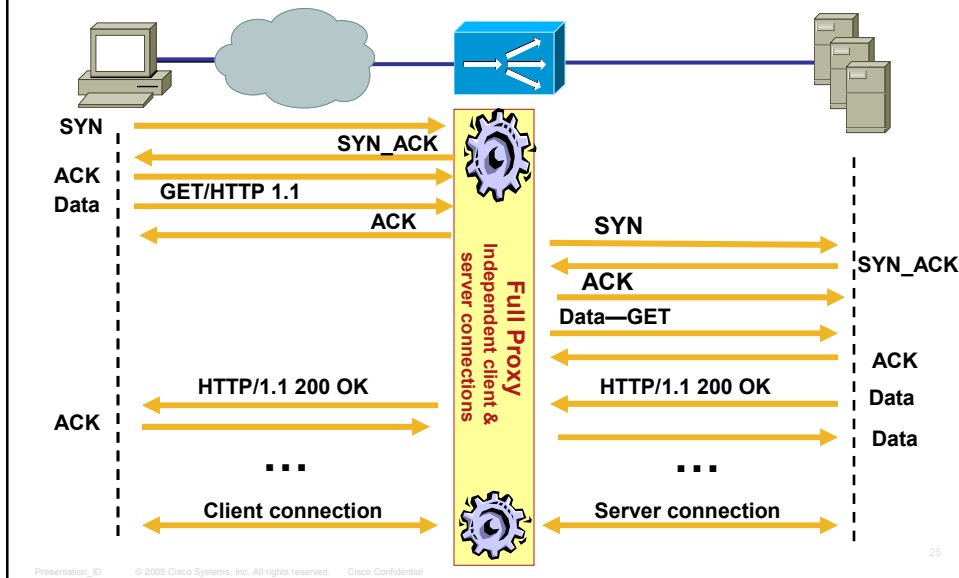
Layer 7 Flow Setup for HTTP (2/3) Load Balancing Decisions Require More Data



Layer 7 Flow Setup for HTTP (3/3) Load Balancing Decisions Require More Data



Layer 7 Flow Setup—Full Proxy The Most Flexible Approach



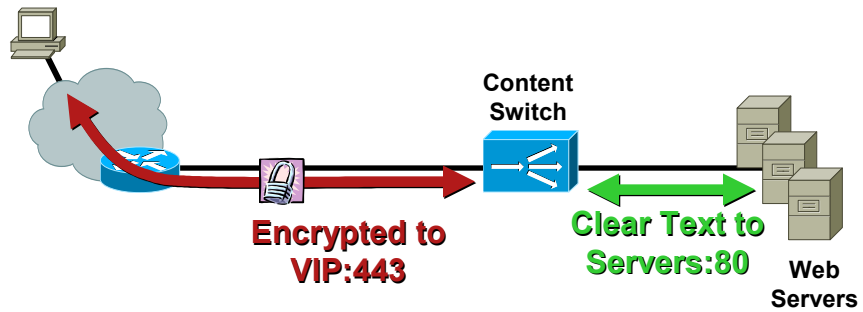
SSL Offload



Freeing Up Server CPU and Resources

Offloading SSL

1. **Offload** CPU-intensive SSL processing
Servers resources are dedicated to serving requests and running applications, rather than encrypting data
2. **Centralized** key/certificate storage/management
3. Allows **advanced content switching** (URL-based, cookie-sticky, payload parsing) and inspection of SSL traffic
4. **Scalability**: easy to add more SSL "performance"

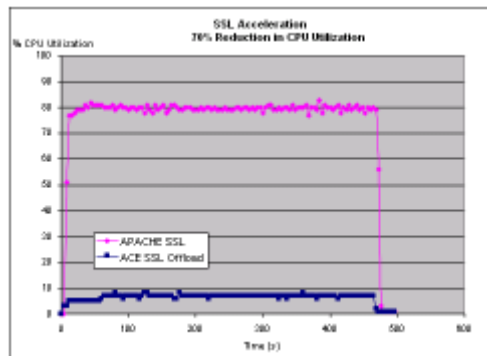


Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Application Acceleration Results

The Impact of Removing the Burden of SSL Transaction Processing from the Apache Server Results in a 70% Reduction in CPU Utilization

Tests performed simulating 1,000 simultaneous users, each generating a single HTTP 1.0 request per second.



Reduction in CPU utilization

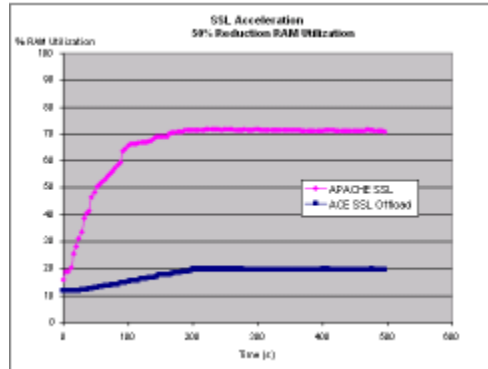
Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

28

Application Acceleration Results

In Addition to the Reduction in CPU Utilization There Is Also a Significant Reduction in Memory Utilization. Figure Below Shows That with ACE Offloading SSL Processing the Apache Server Used 50% Less Memory to Process the Same Number Of HTTP Transactions per Second

Tests performed simulating 1,000 simultaneous users, each generating a single HTTP 1.0 request per second.



Reduced RAM Utilization

29

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

TCP reuse

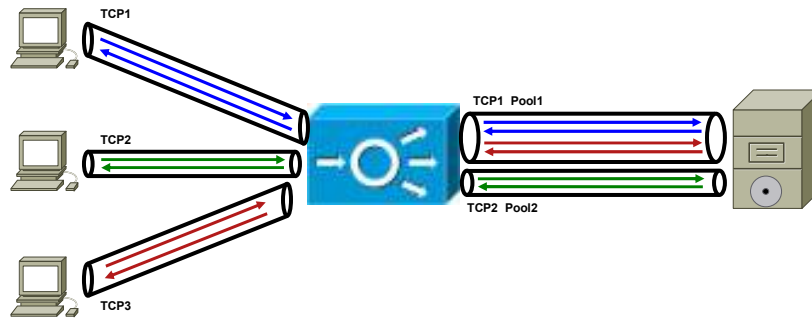


30

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Offloading TCP TCP Reuse (Multiplex)

1. Offload TCP (HTTP) setup processing from servers
Servers resources are dedicated to serving requests and running applications, rather than opening and closing TCP connections
2. TCP connections to the server are kept open (HTTP 1.1 Connection Keepalive)
3. Client requests multiplexed to existing server connections
4. Effective for servers dedicating high percentage of CPU cycles to TCP processing



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

31

TCP Server Offload

1. When the feature is enabled, a server TCP connection may be reused to service a different client TCP connection after the response to the previous HTTP request has been transmitted
2. “**Connection: keep-alive**” is inserted and “**Connection: close**” is removed from the client HTTP request, to avoid closing the server connection early

```
switch/Admin(config)# parameter-map type http PARAM-MAP
switch/Admin(config-parammap-http)# server-conn reuse
```

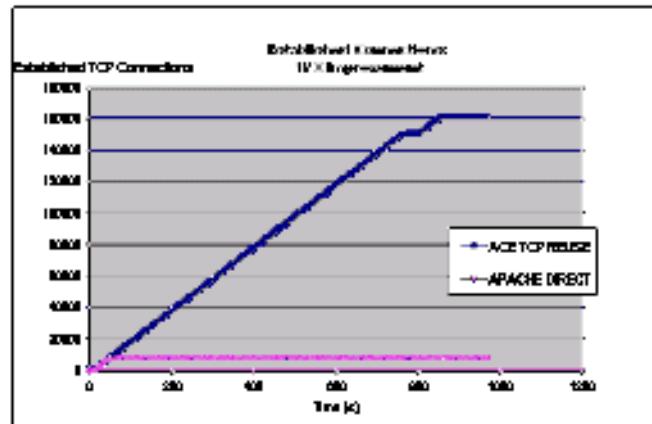
```
switch/Admin# show stats http | include Reuse
Reuse msgs sent      : 1          , HTTP requests      : 4
switch/Admin# show stats http | include Headers
Reproxied requests  : 0          , Headers removed    : 1
Headers inserted    : 1          , HTTP redirects     : 0
```

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

32

TCP-Reuse Acceleration Results

Improvements in Server Capacity, Tested with the Cisco ACE Module Front-ending the Apache Server. The TCP Connection Reuse Feature Was Enabled on the ACE



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

33

HTTP Compression

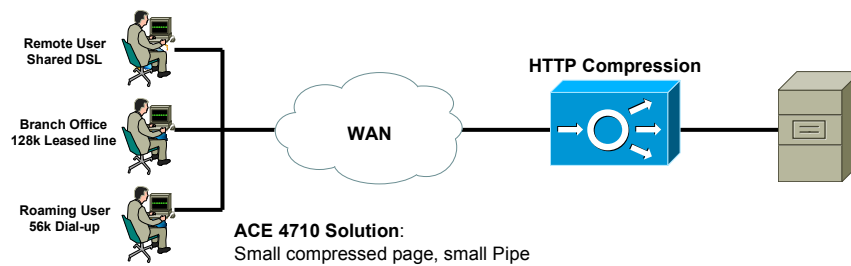


Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

34

Cisco ACE 4710 HTTP Compression

1. Reduces HTTP traffic using GZIP or Deflate compression algorithms which are supported in today's Web browsers
2. Compression is completely transparent to the end user, requiring no downloads or agents
3. Up to 90% reduction in size of web objects such as static and dynamic HTML, Flash, PDFs, text files, XML
4. Optimizes delivery of content for last-mile bandwidth bottlenecks
5. Accelerates end-user experience

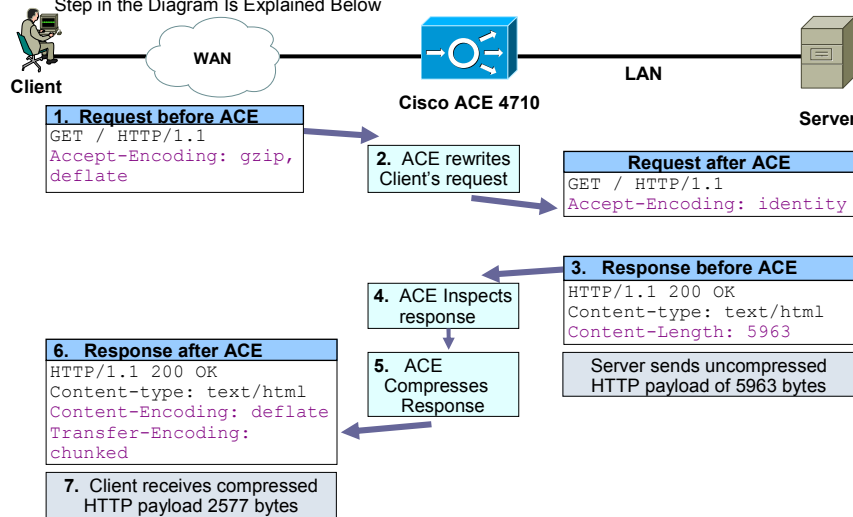


Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

35

Traffic Flow with ACE 4710 Compression

The Following Figure Outlines the Process of How the Server's Response to a Request Made by a Web Browser Is Compressed by the ACE 4710 Appliance. Each Step in the Diagram Is Explained Below

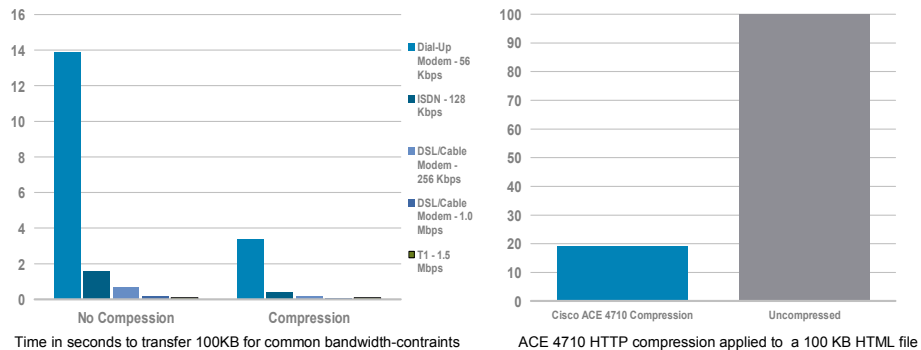


Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

36

Benefits of Compression

1. Compressing web content cuts bandwidth costs by reducing the size of the data delivered to the client by as much as 70 to 90%
2. Cisco ACE 4710 HTTP compression reduces the size of the HTTP payload delivered to the web browser and as a result the overall amount of data transferred across WAN links drops by a factor of 3 to 4
3. This decrease in data sent downstream increases the effective available bandwidth allowing more transactions to be carried across the same links



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

37

What Type of Files Should Be Compressed?

The Following Chart Summarizes How the ACE 4710 Handles the Various Common Web Applications When Compression Is Enabled by the Device Manager

Object Type	Example File Extensions	Gzip/Deflate Compressible	Content-Type	Device Manager enabled default configuration
Static HTML	htm,html,shtml	Yes	text/html	Compressed by default
Dynamic HTML	dhtml,aspx,jsp,cfm,php,asp	Yes	text/html	Compressed by default
XML	xml	Yes	text/xml	Not compressed by default
Images	Gif,jpg,png	No	image/gif image/jpg image/png	Do not apply gzip/deflate compression
Scripts	js	Yes	application/x-javascript text/javascript	Not compressed by default to avoid known browser bugs
Stylesheet	css	Yes	text/css	Not compressed by default to avoid known browser bugs
Downloaded Files	pdf,doc,xls,ppt	Yes	Application/pdf Application/word Application/excel Application/powerpoint	Not compressed by default

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

38

Rate Limiting



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

39

ACE Protects Server Resources

1. Server resources can be overwhelmed with traffic floods – malicious or otherwise
2. ACE can rate limit bandwidth and connections to protect server resources
3. Guaranteed rates for bandwidth and connections per ACE virtual device
4. ACE rate limiting can be applied on virtual contexts, virtual IP addresses and individual servers



Most Granular Protection Of Server Resources

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

40

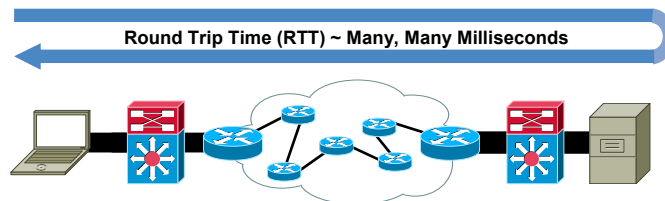
ACE 4710 Application Acceleration



Latency

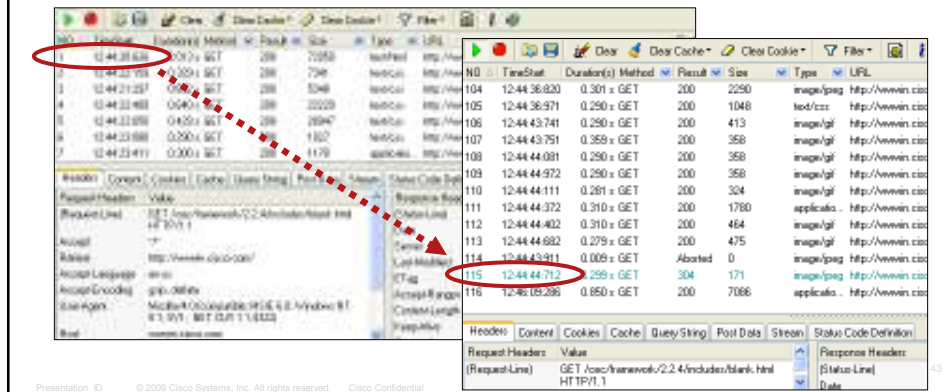
Latency Impairs Application Performance in Three Ways

1. Network latency: the amount of time necessary for a message to traverse the network
2. Transport latency: the amount of time necessary for the transport mechanism (TCP) to acknowledge and retransmit data
3. Application latency: “chattiness” of an application protocol causing messages to be exchanged across the network



Latency Impact on Web Application Front-End (HTTP)

1. HTTP is a chatty protocol
2. Browser parallelism and pipelining is limited
3. Many common web pages include tens or even hundreds of objects
Example: internal Cisco web portal: **115 roundtrips**
24 seconds to render the page with **~250ms** (over VPN, Europe to California)



ACE 4710 Application Acceleration Features

Optimization Feature	Description
Compression	Reduces traffic to web-clients by compressing HTTP response using GZIP and Deflate compression algorithms
Flash Forward	Flash Forward enables acceleration of embedded objects in a web page by caching them locally, resulting in improved application response time
Delta Optimization	Delta feature optimizes the delivery of dynamic web content by only serving differences between visits to a web page.
ETag	Dynamic ETag enables acceleration of non-cacheable embedded objects, resulting in improved application response time.
Dynamic Cache	ACE 4710 optimizes the delivery of dynamic web content by serving data from ACE in memory cache for specified time.

**Benefits from these acceleration functions will vary
by application type and customer use-case**

FlashForward—Key Concept

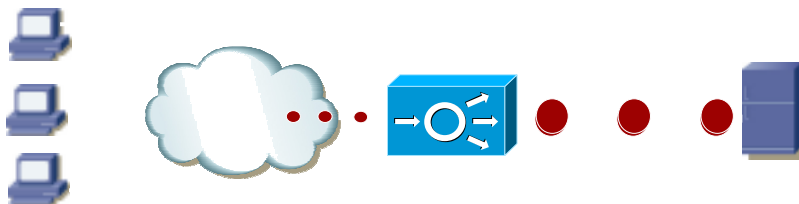
1. Embedded objects referenced in HTML container pages are served with **Expires:** which sets expiry in the future
2. On 2nd visit Browser will not send **GET** for objects in cache if the current date and time is not greater than the object expiry date
3. This reduces the total number of HTTP requests for subsequent visits to the same page
4. Benefits:
 - Significantly decreased page download time
 - and number of requests to origin servers
5. Functions:
 - Cache object on Application Switch
 - Transform object name to include MD5 hash
 - Rewrite reference in page to the transformed object
 - Full benefit after 2nd visit

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

45

Cisco ACE 4710 Delta Optimization

1. ACE delta optimization applied to dynamic web applications such as .Net J2EE SAP Oracle Siebel Lotus
2. Enables dynamic update of client browser caches with content differences or deltas
3. Observes and modifies HTML content that flows through it to achieve bandwidth savings and user download performance.
4. Results in bandwidth savings and improved end-user experience



Problem: Entire 150K page served on each visit

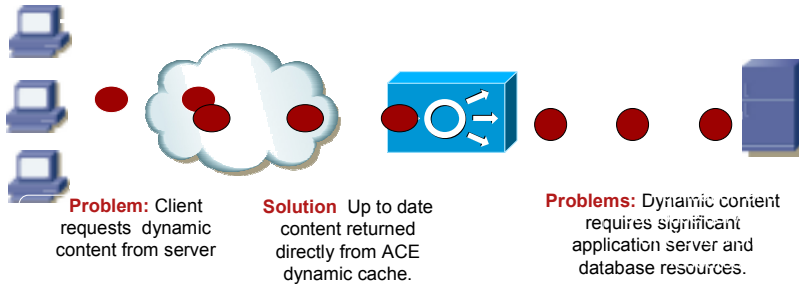
Solution: Only differences sent across the WAN

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

46

Cisco ACE 4710 Server Offload - Dynamic Caching

1. Enables the Cisco ACE 4710 to fulfill requests for dynamic or personalized information
2. Offloads application servers and databases
3. Significantly improves application response time, reduces the server load, and enables more concurrent users to be served.
4. Improved scalability and lower ongoing server upgrade costs.



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

47

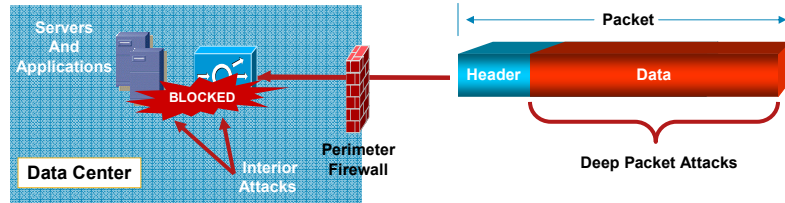
Generic Protocol Parsing



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

48

Generic Protocol Parsing Why Deep Packet Parsing



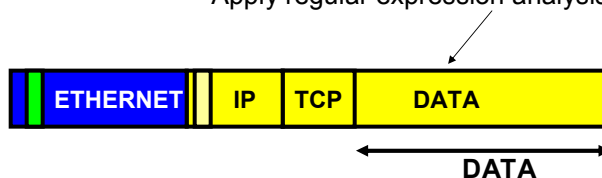
- Basic load balancers support many well-known protocols such as HTTP, FTP, SNMP, etc. However, there are many applications which use proprietary protocols on TCP and UDP.
- Problem is when you want to load balance proprietary protocols, most basic load balancers cannot look at the payload. Users will need to rely on Layer 4 load balancing.
- ACE introduces **Generic Protocol Parsing** which supports load balancing and parsing of the TCP or UDP payload.
- Using generic protocol parsing, ACE can load balance proprietary protocols or proprietary applications and perform deep packet inspection and block application attacks.

Generic Protocol Parsing Use Deep Packet Parsing to Block AIM Traffic

Define a match criteria for Layer 4 payloads by using the **match layer4-payload command** in class-map generic. In this use case, you will be matching a string value called "AIM". Note the regex expression specifies the Layer 4 payload expression that is contained within the TCP or UDP entity body.

```
Ace#show run class-map
class-map type generic match-any L7-generic
  2 match layer4-payload regex ".*AIM*"
```

Apply regular expression analysis to match the data



Generic Protocol Parsing Use Deep Packet Parsing to Block AIM Traffic

1. Once the request matches the following regular expression you want to enforce a policy. This is accomplished using a policy map
2. Due to your policy no "AIM" AOL Instant Messenger type traffic is allowed through ACE. At this point ACE can discard this AIM request using a drop policy-map

```
policy-map type loadbalance generic first-match IM-Inspection  
class L7-generic  
  drop ← Drop the "AIM" request
```

1. Using Generic Protocol Parsing you can match any TCP payload string and execute an action. That action could be load balancing using sticky, permit, deny etc

51

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Inspect HTTP POST body



52

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Parsing of POST arguments in ACE

Sample use case

1. Consider this very simple web form:



2. It contains 3 parameters (param1, 2, 3)

```
<!DOCTYPE html>
<html>
<head>
</head>
<body>
  <div style="border: 1px solid blue; padding: 5px;">
    Enter some text here, I'll echo it back:</div>
    <input type="text" name="param1" value="" />
    <input type="text" name="param2" value="" />
    <input type="text" name="param3" value="" />
    <input type="submit" value="Submit" />
  </div>
</body>
</html>
```

3. We can build an ACE configuration that only accepts 5 digits for param3

Parsing of POST arguments in ACE

Configuration for the sample use case

```
class-map match-any foo
  2 match port tcp eq www
class-map type http inspect match-any foo_inspect
  2 match cookie secondary name param3 value "[0-9]{5}"

policy-map type inspect http all-match foo
  class foo_inspect
  permit log
  class class-default
  reset log

policy-map multi-match foo_mm
  class foo
  inspect http policy foo url-logging

interface vlan2
  ip address 10.10.10.1 255.255.255.0
  service-policy input foo_mm
```

Review following HTTP POST method

Apply Inspection to policy-map multi-match

Some other Security Options



Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

55

Security Features

IP/UDP/ICMP Exploits Blocked by ACE

1. IP checks performed by ACE:

Automatic Anti-spoofing (source IP = dest IP); unicast RPF check

Header length check (min and max lengths, L3 < L2)

IP options control

Drop illicit IP addresses (source IP = class D or broadcast or loopback)

Overlapping fragments dropped, control over max number of fragments

ARP Inspection in transparent mode

2. ICMP checks performed by default:

Requests and responses matching

Prevents injection of unsolicited ICMP errors

Countermeasures specified in draft-gont-tcpm-icmp-attacks.txt

Blocked Attacks: Timestamp/Route Record/Source Routing/Fragment DoS Attacks, IP Spoofing, Ping of Death, ICMP Flood, Smurf, ARP Attacks

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

56

FTP Inspection No PUTs for You...

```

serverfarm host STRICT-FTP-APP-SF
  rserver SERVER2
  inservice
class-map match-all FTP-CM
  2 match virtual-address 172.16.1.52 tcp eq ftp
class-map type ftp inspect match-any NO-PUTS-CM
  2 match request-method put
!
policy-map type loadbalance first-match STRICT-PM
  class class-default
    serverfarm STRICT-FTP-APP-SF
  policy-map type inspect ftp first-match FTP-Inspect-PM
  class NO-PUTS-CM
    deny
  policy-map multi-match CLIENT-VIPS
  class FTP-CM
    loadbalance vip inservice
  loadbalance policy STRICT-PM
  inspect ftp strict policy FTP-Inspect-PM
!
interface vlan 2
  ip address 172.16.1.1 255.255.255.0
  access-group input EVERYONE
  service-policy input REMOTE-MGN
  service-policy input CLIENT-VIPS

```

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

57

“Level 1” HTTP Attacks Prevented by ACE

- 1. Encrypted channel attacks—HTTPS decryption**
ACE is equipped with a powerful SSL offload/termination chip, giving it full visibility into attacks hoping to get around security devices by riding on top of an encrypted channel
- 2. Worms and day-zero attacks**
ACE's HTTP inspection engine contains a powerful fully-customizable regular expression engine. Using regular expressions, users can develop signatures that can block worms and attacks for which no known remedy is published yet! Regexes can be applied on the headers, the URL or even the payload of HTTP traffic
- 3. RFC compliance**
ACE's HTTP inspection engine automatically enforces RFC2616 compliance and can drop any methods, mime-types or transfer encoding as configured by the user
- 4. Buffer overflows**
Maximum HTTP header length can be enforced, avoiding attempts at buffer overflow exploits
- 5. Directory traversals**
An attempt at working one's way up an HTTP server's directory structure by using ../ in GET requests; easily blocked by ACE's regular expression filters
- 6. Malicious URLs**
ACE always canonicalizes URLs, defeating any attacks relying on encoded URL
- 7. Peer-to-peer, instant messaging, HTTP-tunnels**
Traffic tunneled over HTTP can be blocked by ACE's HTTP inspection engine

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

58

HTTP Inspection Components

1. RFC 2616 Compliance and filtering

Protocol Conformance: the 1st line of a REQUEST is "Method SP" and that of RESPONSE is "HTTP-Version SP". etc

De-obfuscation: override attempts to avoid regex searches by encoding the URL

Methods: OPTIONS, GET, POST, HEAD, PUT, DELETE, TRACE, CONNECT

Extensions: INDEX, MOVE, MKDIR, COPY, EDIT, UNEDIT, SAVE, LOCK, NLOCK, REVLABEL, REVLOG, REVNUM, SETATTRIBUTE, GETATTRIBUTE, GETATTRIBUTENAMES, GETPROPERTIES, STARTREV, STOPREV

2. Length and encoding checks

Length: Configurable range for URL and URL Header requests and responses

Encoding: chunked | compress | deflate | gzip | identity

3. Detect HTTP misuse

Peer-to-peer (p2p) applications: KAZAA, GNUTELLA

Tunneling applications: HTTPPort/HTTHost, FireThru

Instant Messaging: (IMI - YAHOO Messenger)

4. MIME type validation and filtering

audio: /*, /midi, /basic, /mpeg, /x-adpcm, /x-aiff, /x-ogg, x-wav (8)

image: /*, /cgf, /gif, /jpeg, /png, /tiff, /x-3ds, /x-bitmap, /x-niff, /x-portable, /x-xpm (11)

text: /*, /css, /html, /plain, /richtext, /sgml, /xmcd, /xml (8)

video: /*, /flc, /mpeg, /quicktime, /sgi, /x-avi, /x-flt, /x-mng, /x-msvideo (9)

application: /msword, /octet-stream, /pdf, /postscript, /vnd.ms-excel, /vnd.ms-powerpoint, /x-gzip, /x-java-archive, /x-java-vm, /zip (10)

1. Regex filtering on HTTP messages

Detect protocol running on top of HTTP – i.e. to detect YAHOO MESSENGER, look for YMSG in the first 4 bytes

59

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Deployments



Network Integration Options and Examples

60

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

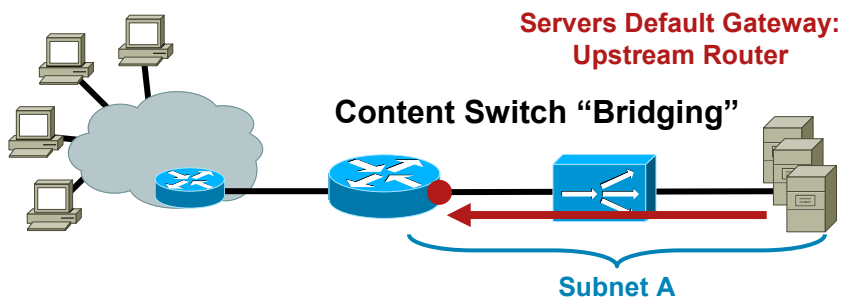
ACE-FWSM Example Design Options

Bridge Mode	Routed Mode	One Arm Mode
Simplified L3 Topology	Removes possibility of loops in Aggregation Layer	Offload SLB from non-load balanced flows
Services separate from Networking	NAT support	Source NAT or PBR required for return traffic
<ul style="list-style-type: none"> •MSFC is default GW, HSRP router, routing peer 	Service Module is Default GW	
Dynamic routing across service infrastructure	Static routing	
Limited NAT support	More complex L3 topology	
Potential for loops – BPDU forwarding		

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

61

Bridge Mode



1. Servers in routable IP subnet
2. VIP's can be in the same or different subnet
3. Requires one IP subnets for each farm
4. Easy deploy for firewall or cache load balancing

© 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

62

Bridge Mode Pros and Cons

Pros

1. Simpler implementation for existing servers
2. Less complex to design and troubleshoot

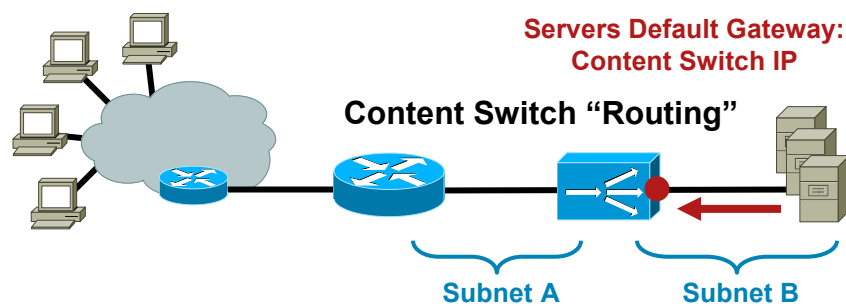
Cons

1. Increases broadcast domain
2. Hard to scale
3. All traffic flows through ACE

63

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Router Mode



1. Servers in private IP subnet
2. VIPs usually in different, routable subnet from servers
3. Requires two IP subnets
4. Easy to deploy with many server IP subnets

64

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

Routed Mode Pros and Cons

Pros

1. Most common approach so easiest to support
2. Scales well since new VLANs are easily added
3. Easy to troubleshoot since client-side and server-side traffic is easily identified

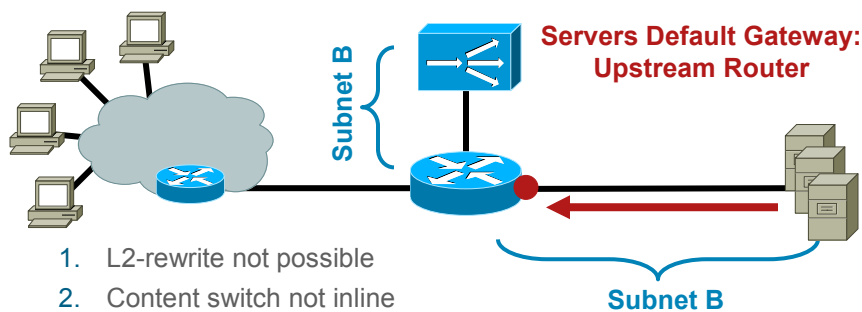
Cons

1. Migrating existing servers requires re-addressing
2. All traffic flows through ACE
3. ACE appliance does not support dynamic routing

65

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

L3 One-Arm Mode



1. L2-rewrite not possible
2. Content switch not inline
Does not see unnecessary traffic
3. Requires PBR, server default gateway pointing to load balancer or client source NAT
The return traffic is needed!
4. Not as common as bridge or routed mode due to problems with forcing traffic back to ACE in return direction

PBR—Policy Based Routing, NAT—Network Address Translation

66

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

One Arm Mode Pros and Cons

Pros

1. Direct access to servers bypasses ACE
2. Server initiated connections bypass ACE
3. Non load-balanced traffic does not utilize licensed resources on ACE

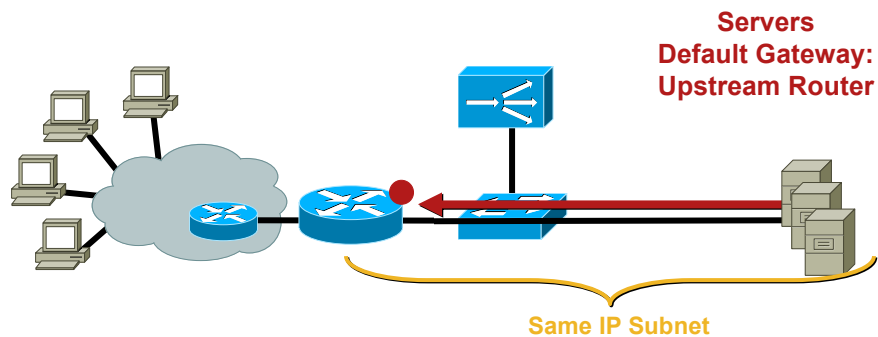
Cons

1. Complex configuration requiring source NAT or PBR
2. Must have greater understanding of application flow
3. May be more difficult to troubleshoot

Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

67

L2 One-Arm Mode Return Traffic Bypassing Load Balancer



1. Bypass for return traffic: high throughput!
2. Requires MAC rewrite, L2 adjacency
3. Servers need identical loopback addresses (one per VIP)
4. TCP termination not possible: **no L7 features!**
5. Load balancer blind to return traffic (inband, accounting)

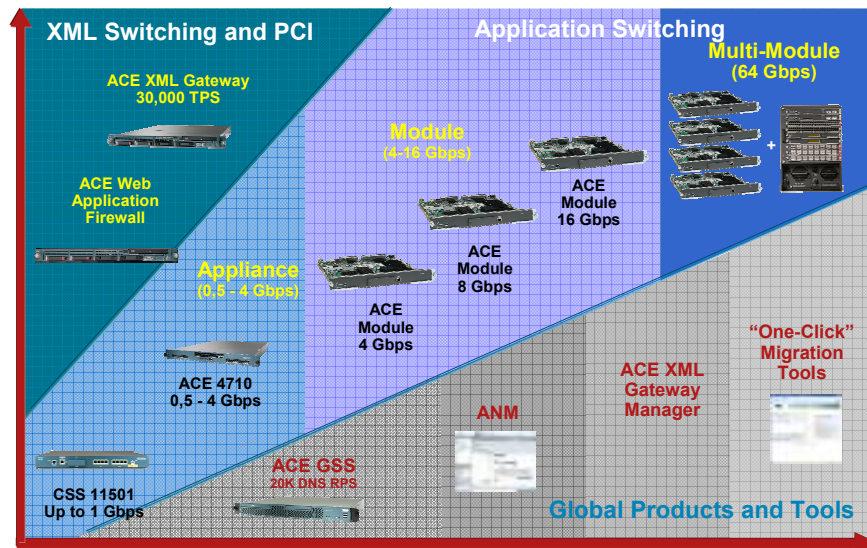
Presentation ID: © 2009 Cisco Systems, Inc. All rights reserved. Cisco Confidential

68

Solution



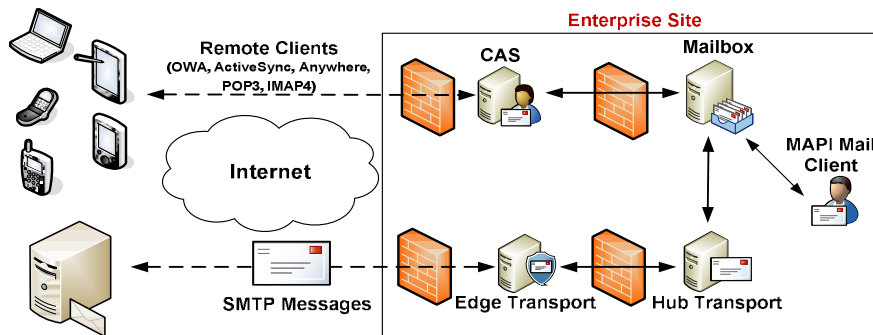
Cisco Application Control Engine Family



ACE and OWA



Microsoft Exchange 2007 Logical Layout

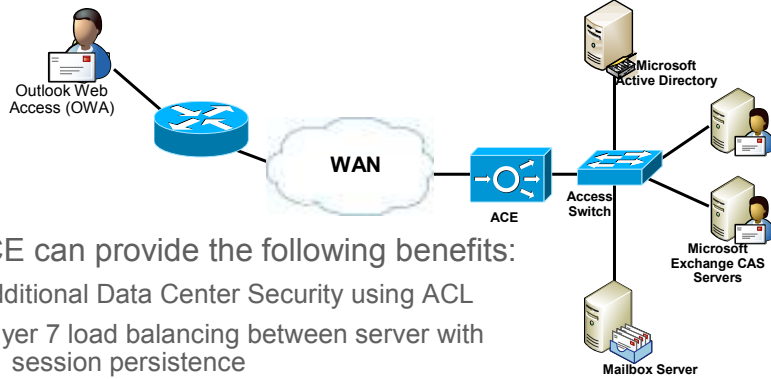


External communications
Internal Exchange Communications



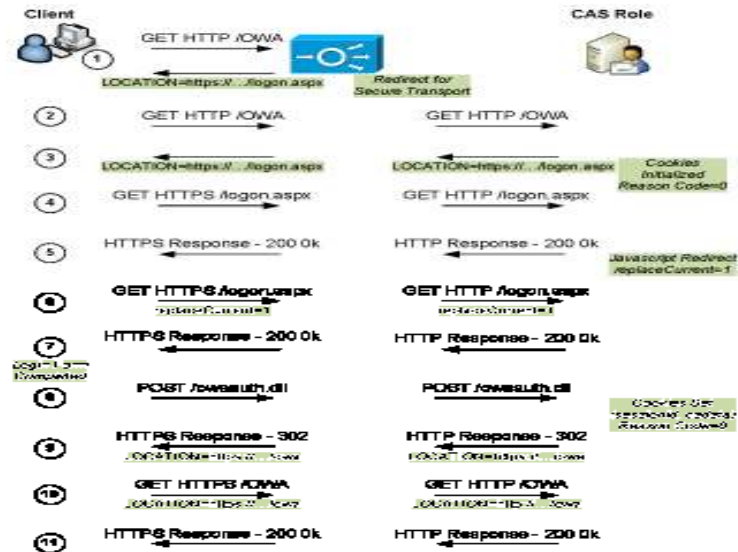
How can ACE provides a highly available and scalable solution from which the Microsoft Exchange 2007 application environment can benefit?

ACE Optimizing Outlook Web Access Traffic

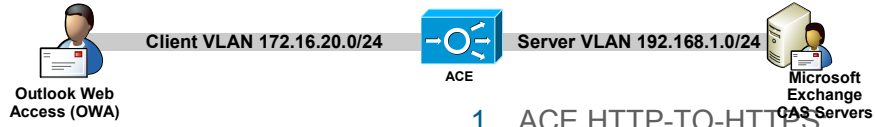


- ACE can provide the following benefits:
 - Additional Data Center Security using ACL
 - Layer 7 load balancing between server with session persistence
 - SSL termination
 - Health monitoring check Client Access Server status
 - HTTP to HTTPS Server Redirection
 - TCP multiplexing and Application Offloading

OWA Login Flow Process



OWA Initial Redirect Clients to HTTPS

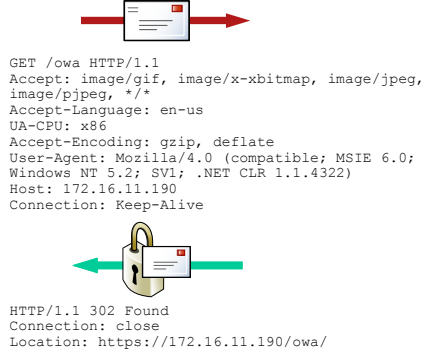


1. ACE HTTP-TO-HTTPS

ACE redirects clients initial request from http-to-https

ACE enforces http-to-https redirects for only /owa. Any other url or get slash to the VIP will be reset and logged. This is applied using HTTP inspection

Match http url /owa will be case sensitive unless you change that in a parameter map



Presentation ID: © 2005 Cisco Systems, Inc. All rights reserved. Cisco Confidential

75

Preventing OWA from Sending HTTP Hrefs



```

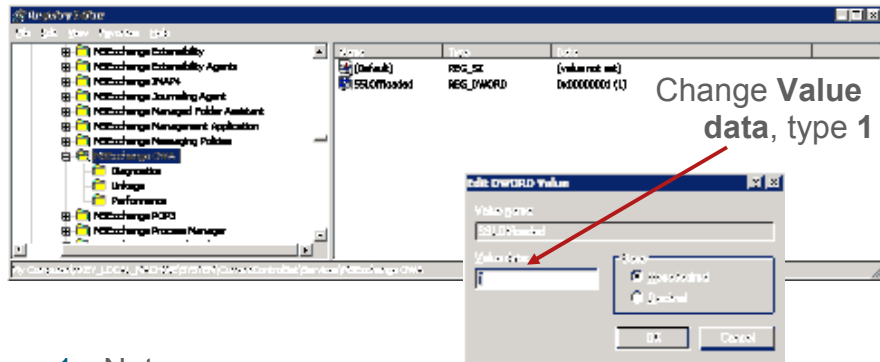
HTTP/1.1 200 OK
Date: Tue, 12 Apr 2005 13:59:37 GMT
Server: Microsoft-IIS/6.0
X-Powered-By: ASP.NET
Content-Type: text/html
Content-Length: 1164
<!--Copyright (c) 2000-2003 Microsoft Corporation. All rights reserved.-->
<!--CURRENT FILE== "IE5" "WIN32" frameset -->
<HTML>
<HEAD>
<META HTTP-EQUIV="Content-Type" CONTENT="text/html; CHARSET=utf-8">
<TITLE>Microsoft Outlook Web Access</TITLE>
<BASE href="http://exampl
</HEAD>
    
```

Incorrectly (Insecure) Formatted Protocol

Presentation ID: © 2005 Cisco Systems, Inc. All rights reserved. Cisco Confidential

76

Configuring the CAS Server for SSL-Offload



1. Note

The CAS role is aware of the SSL-offload functionality of the ACE. To configure support for SSL-offloading on a CAS role, refer to:

<http://technet.microsoft.com/en-us/library/bb885060.aspx>

