WAN Optimering af Data Center Teknologier

Cisco Expo – March 2006

Mikkel Brodersen
Systems Engineer
Cisco Systems, Danmark

mikkel@cisco.com

Key Application and Infrastructure Issues

APPLICATION PERFORMANCE

- WAN bandwidth constraints
- Network and application latency
- Global collaboration

BRANCH INFRASTRUCTURE COST

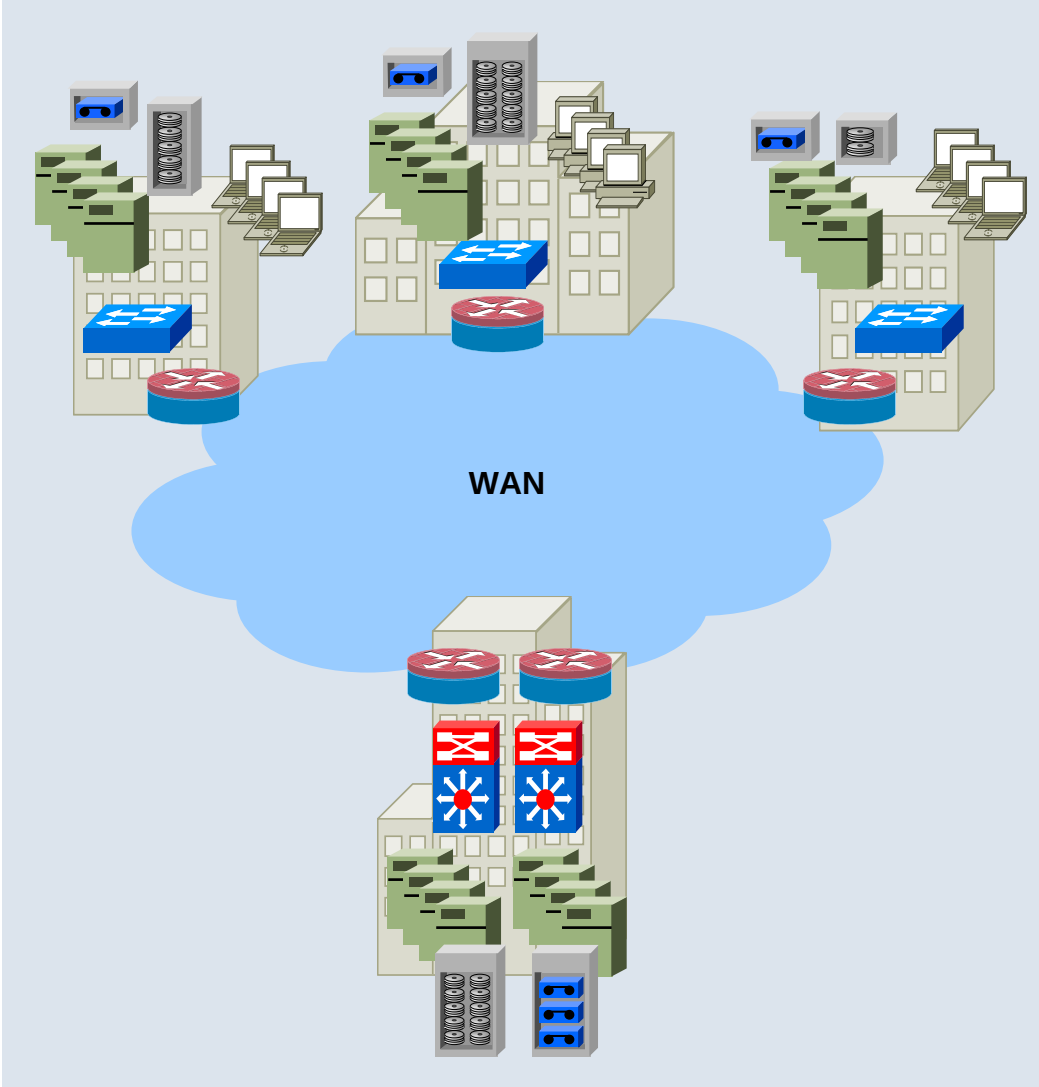
- Server and storage management
- Limited IT resources and budget
- WAN link costs

EXISTING NETWORK DEMANDS

- VoIP, client-server, other apps
- Corporate communications
- eLearning

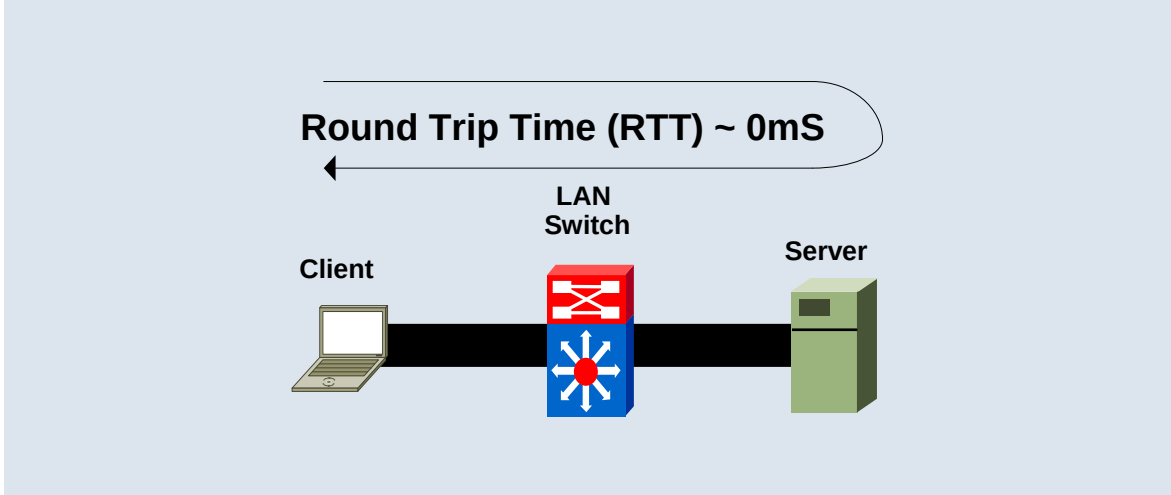
Typical Distributed Enterprise

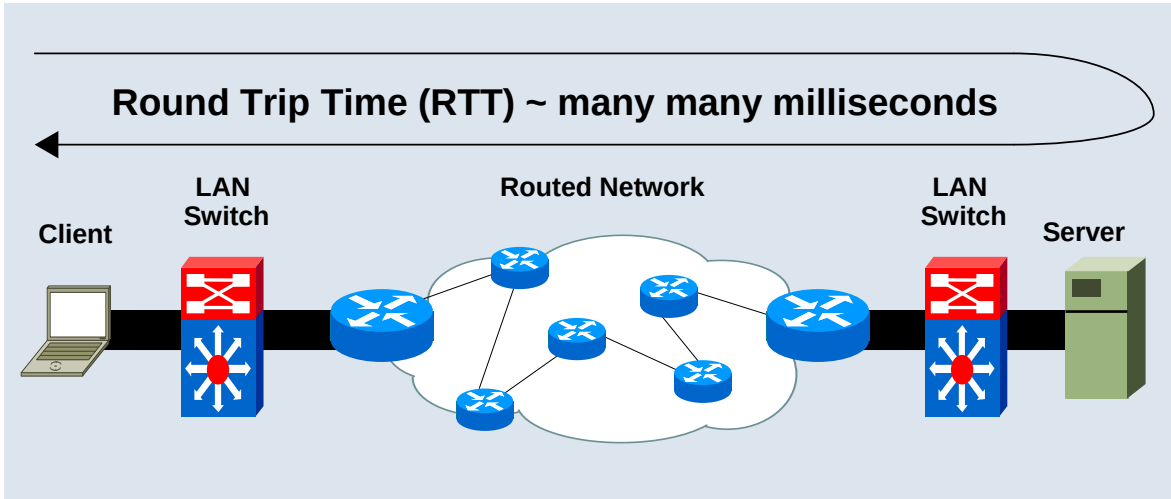
- **Expensive distributed I/T infrastructure**
  - **File and print servers**
  - **Email servers**
  - **Tape backup**
- **Application delivery woes**
  - **Congested WAN**
  - **Bandwidth and latency**
  - **Poor productivity**
- **Data protection risks**
  - **Failing backups**
  - **Costly off-site vaulting**
  - **Compliance**



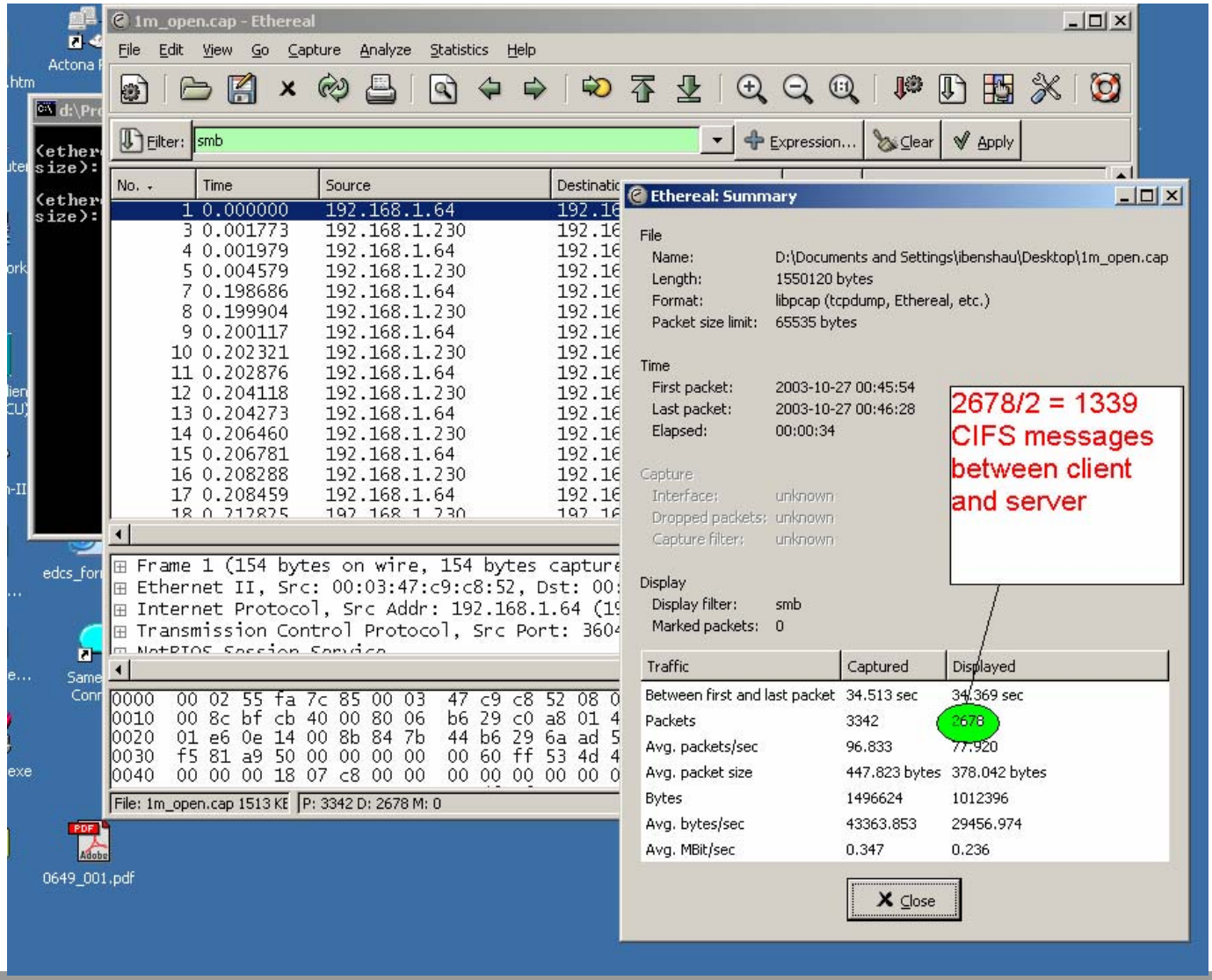
The WAN Is A Barrier To Consolidation

- **Applications are designed for LAN environments**
  - **High bandwidth**
  - **Low latency**
  - **Reliability**
- **WAN characteristics hinder consolidation**
  - **Already congested**
  - **Low bandwidth**
  - **Latency**
  - **Packet Loss**





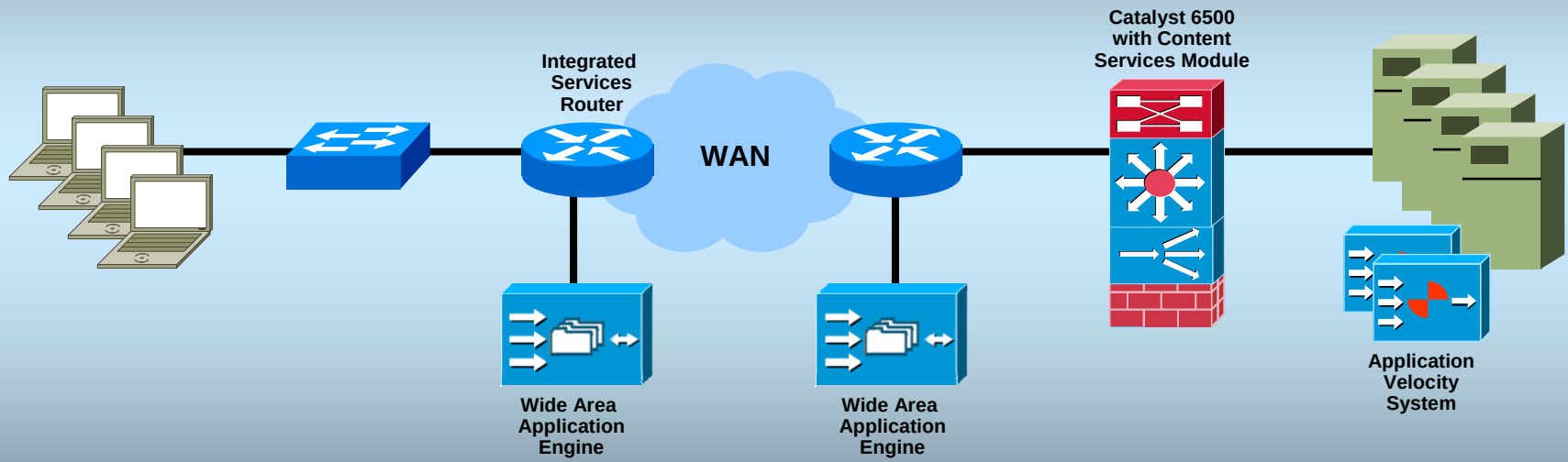
CIFS over WAN is a Challenge
> 1000 messages for single 1 MB Word File



Application Optimization Infrastructure

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
- Service and availability monitoring

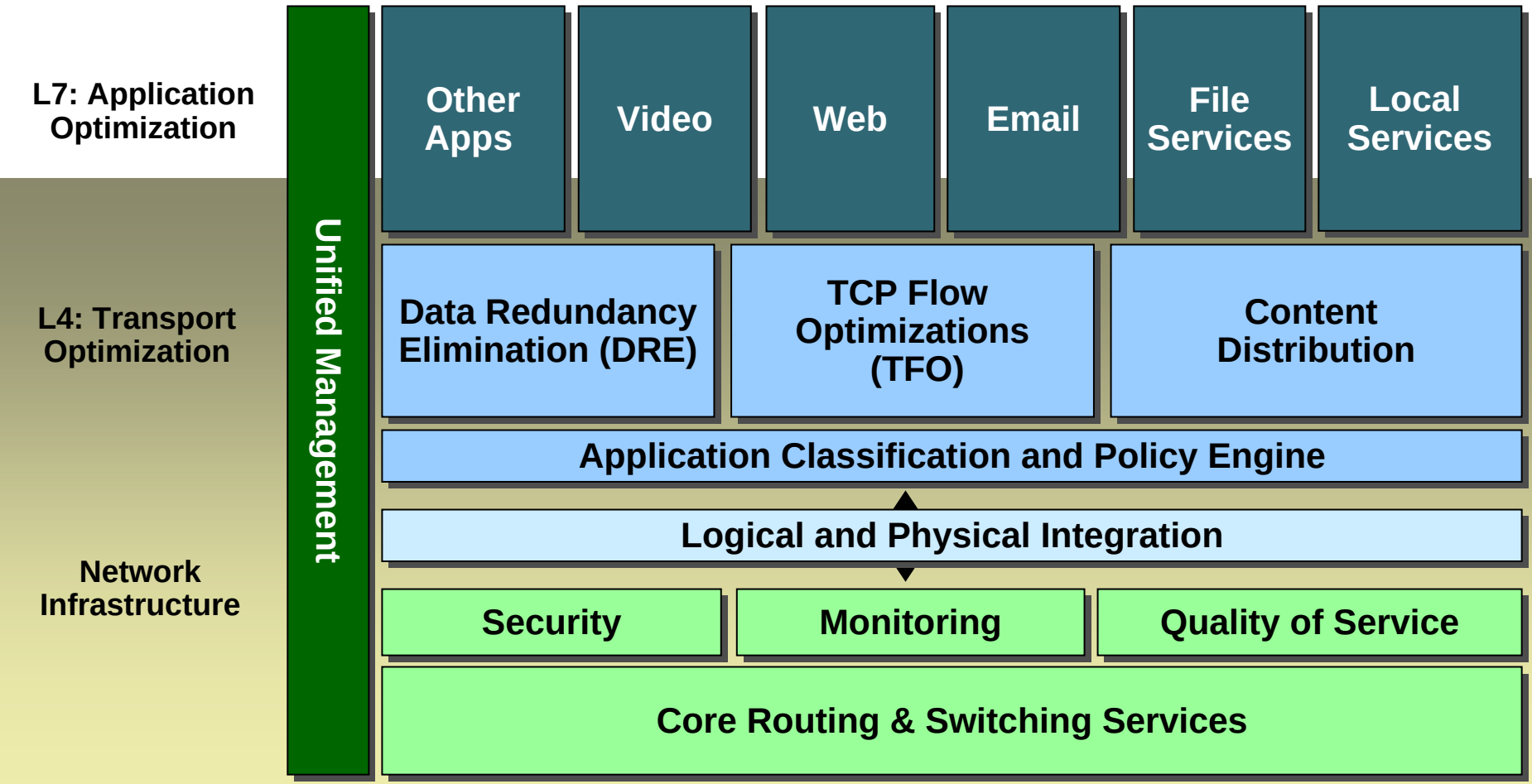


Application Acceleration
- Latency mitigation
- Application data cache
- Meta data cache
- Local services

WAN Acceleration
- Data redundancy elimination
- Window scaling
- LZ compression
- Adaptive congestion mgmt

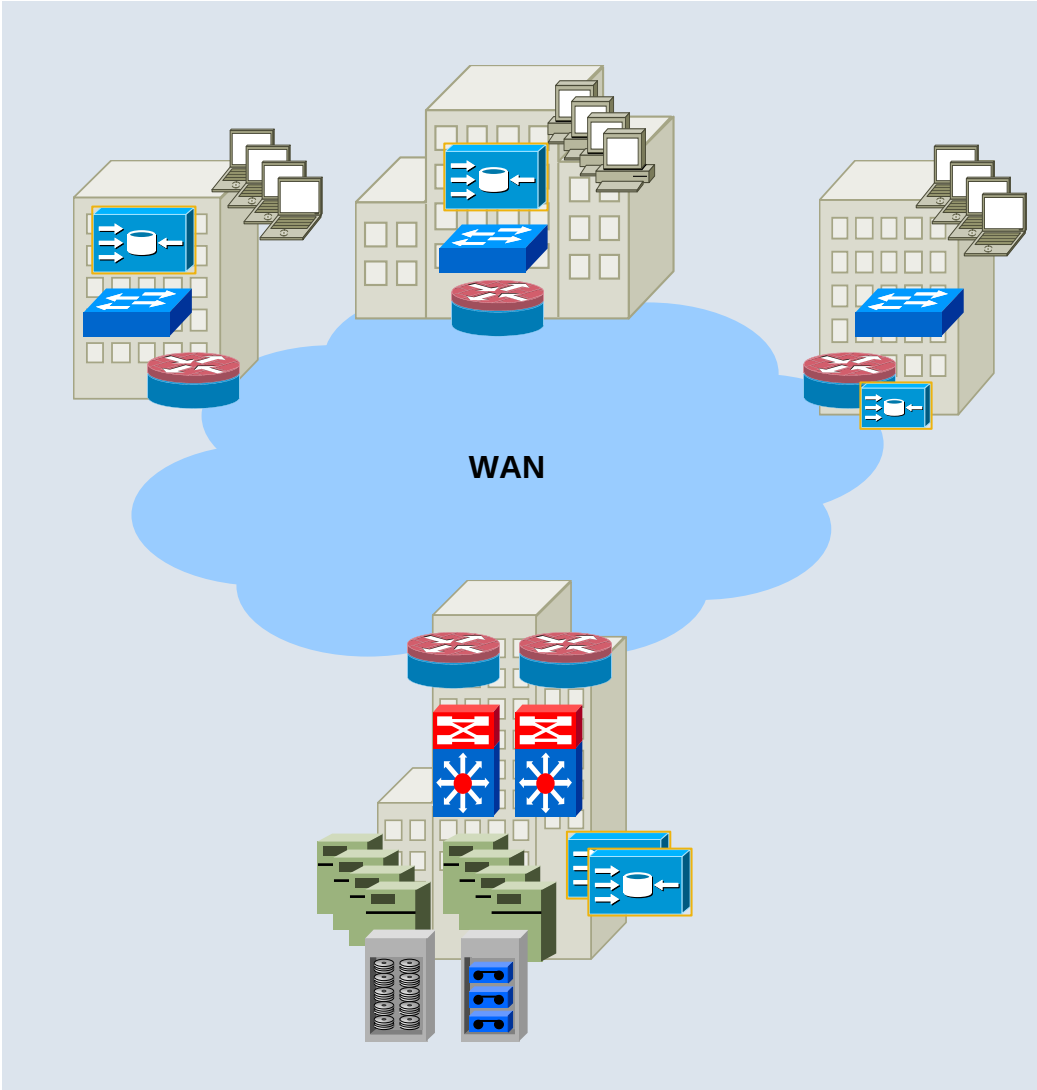
Application Optimization
- Delta encoding
- FlashForward optimization
- Application security
- Server offload

Cisco WAAS Optimization Architecture



Cisco WAAS Enables Consolidation

- **Cisco Wide Area Application Services (WAAS)**
  - **Transparent integration**
  - **Robust optimizations**
  - **Auto discovery**
- **Infrastructure Consolidation**
  - **Remote costly servers**
  - **Centralize data protection**
  - **Save WAN resources**
- **Application Acceleration**
  - **Application adapters**
  - **Advanced compression**
  - **Throughput optimizations**
  - **Policy-based configuration**



WAAS Addresses the WAN Challenge

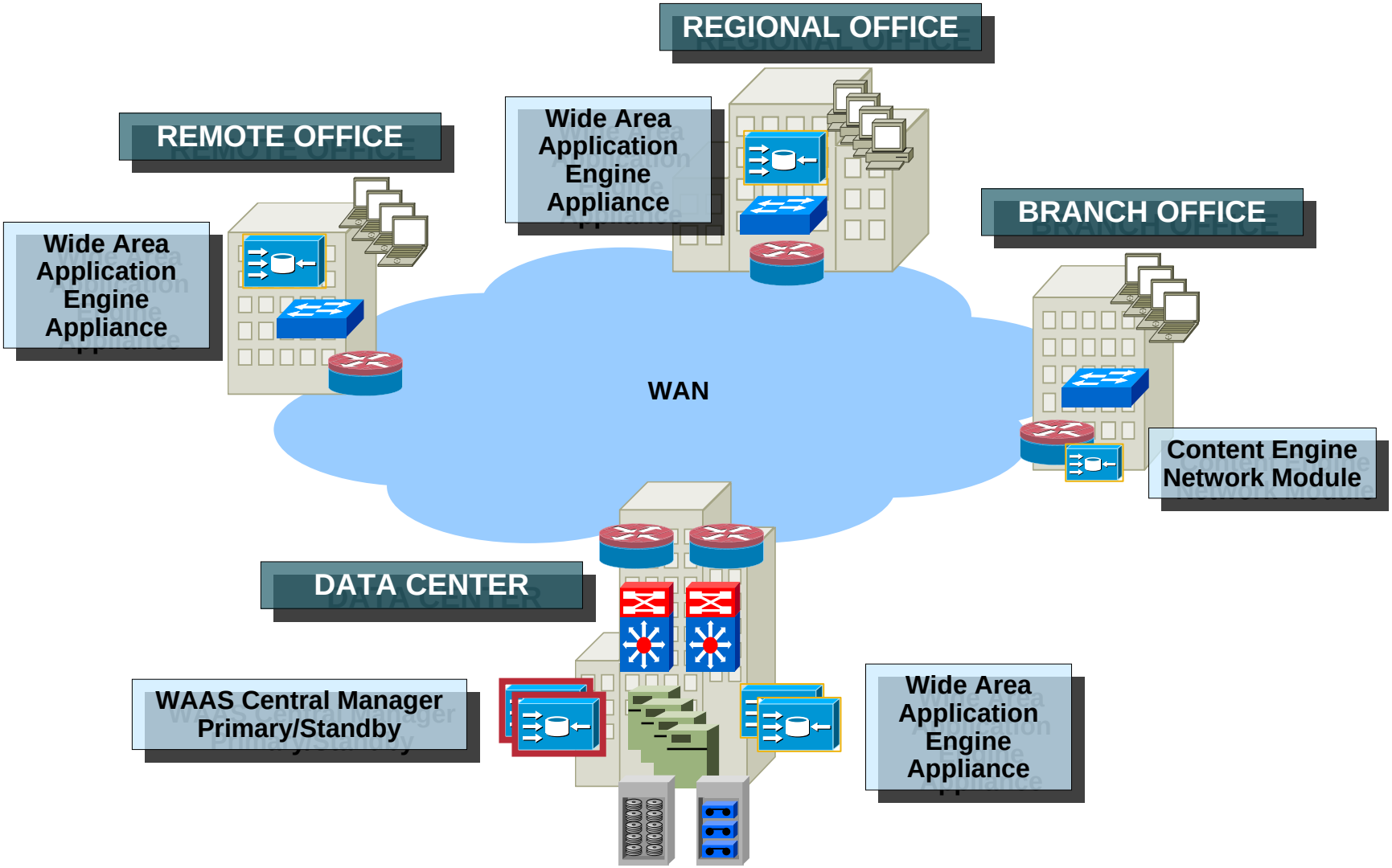
Source	Need	Technology
Latency Mitigation	• Reduced number of network roundtrips from chatty application protocols	• Intelligent Protocol Proxies
Bandwidth Management	• Improve application response time on congested links by reducing the amount of data sent across the WAN	• Application Caching • Data Redundancy Elimination (DRE)
Link Throughput Improvement	• Improve network throughput (total # of data) by reducing TCP-related errors	• Transport Flow Optimizations (TFO)
Network Integration	• Integration into network devices • Compliance with network functions	• Router modules, linecards • QoS, NetFlow, Firewalls
Local Services	• Replacement for services that branch office servers provide	• Centrally managed remote services interface

WAAS Accelerates Broad Range of Applications

Application	Protocol	Typical Improvement
File Sharing	• Windows (CIFS) • UNIX (NFS)	• 2X-100X
Email	• Exchange (MAPI) • SMTP/POP3, IMAP • Notes	• 2X-50X
Internet and Intranet	• HTTP, HTTPS, WebDAV	• 2X-50X
Data Transfer	• FTP	• 2X-50X
Software Distribution	• SMS • Altiris	• 2X-100X
Database Applications	• SQL • Oracle • Notes	• 2X-10X
Data Protection	• Backup Applications • Replication Applications	• 2X-10X
Other	• Any TCP-based Application	• 2X-10X

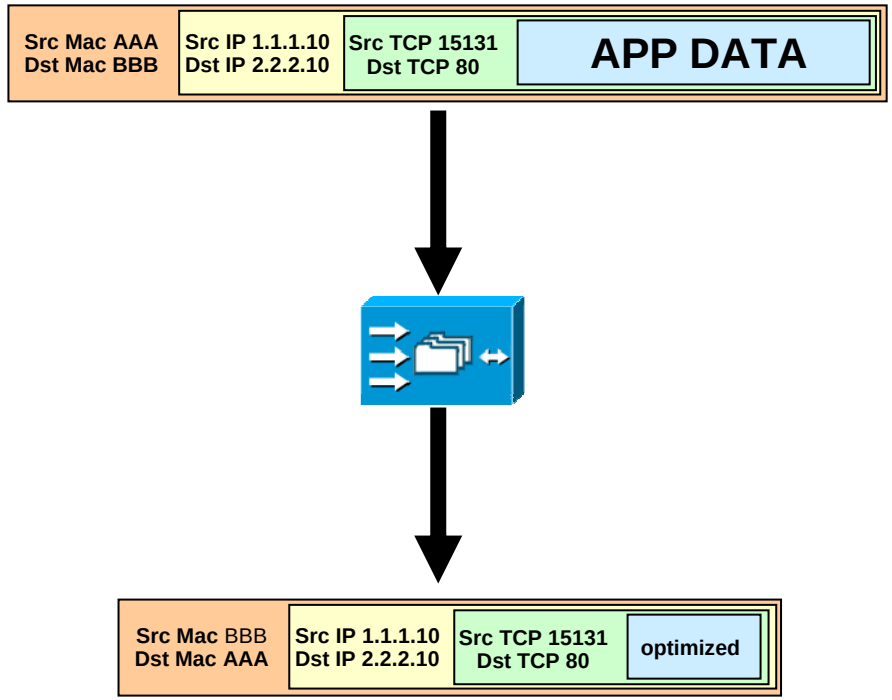
*** Performance improvement varies based on user workload, compressibility of data, and WAN characteristics and utilization. Actual numbers are case-specific and results may vary.**

Cisco WAAS Symmetric Deployment Architecture

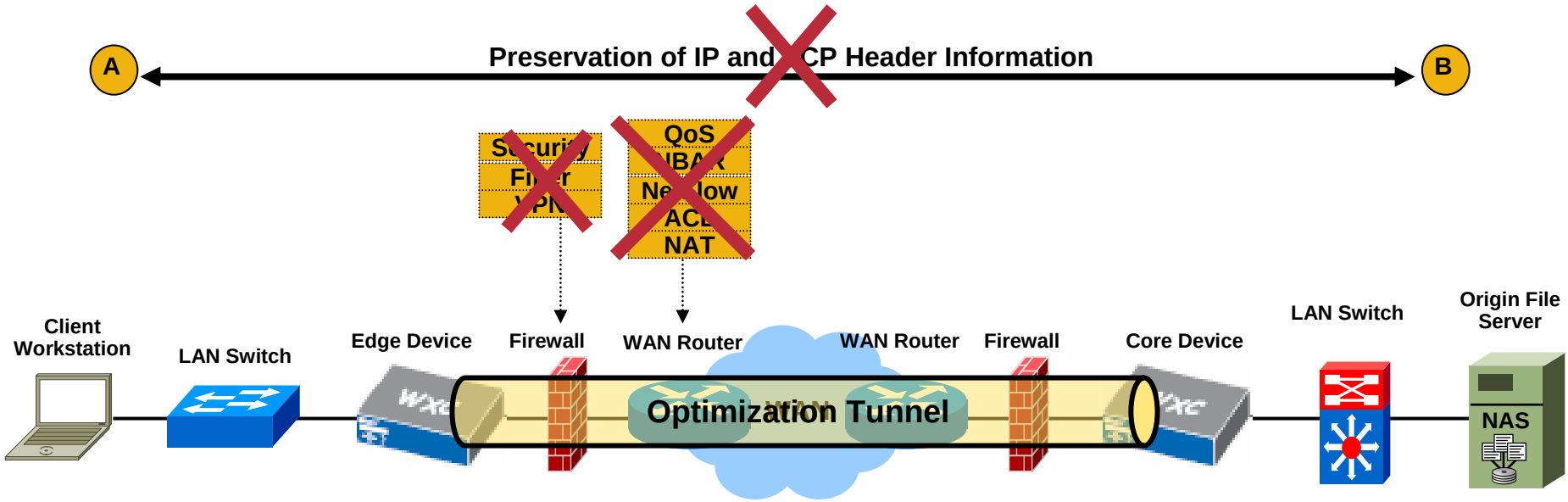


Seamless, Transparent Integration

- **Seamless integration with the packet network with high availability, load-balancing, and failover**
  - **WCCPv2**
  - **Policy-Based Routing**
- **Full preservation of L3/L4 packet header information (IP/TCP)**
- **Compliance with network value-added features**
  - **Classification - QoS, NBAR, Queuing, Policing, Shaping**
  - **Security - Firewall policies, Access Control Lists**
  - **Reporting - NetFlow, monitoring**



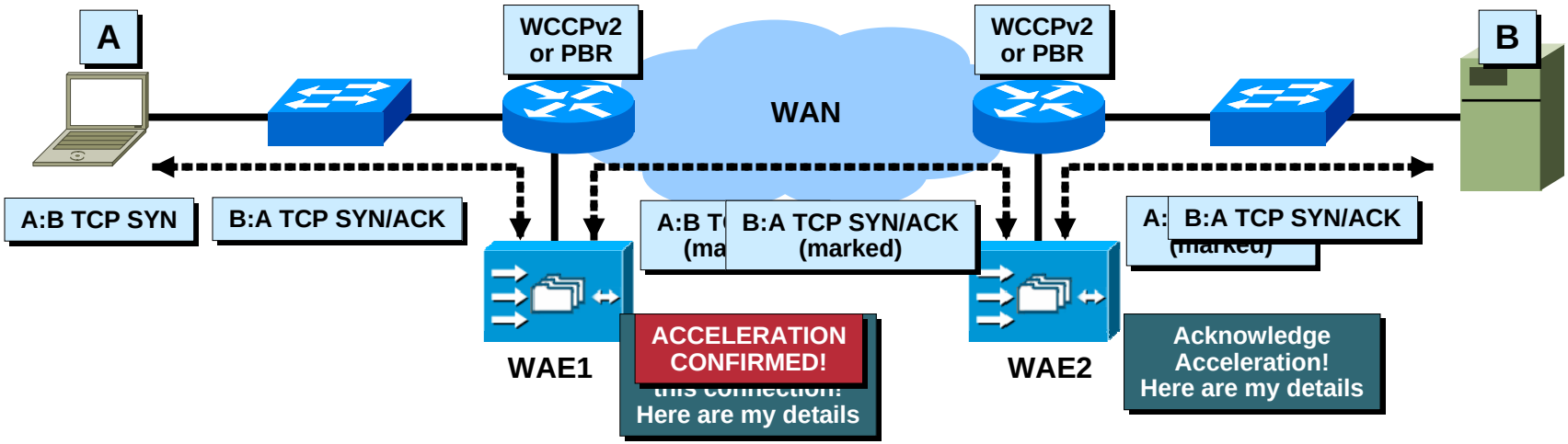
Challenges with Optimizations Using Tunnels



- Networking devices in the path only see traffic amongst optimization appliances
- Difficult to manage and monitor network flows
- Disruptive migration to appliance-based features and security models
- Unnecessarily complex support model

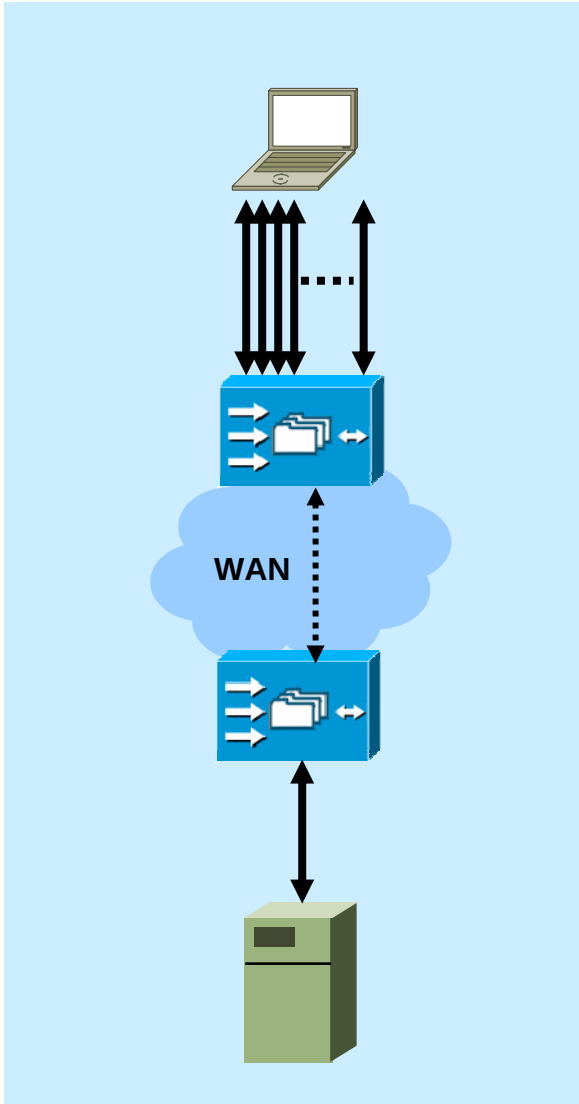
Cisco WAAS Auto-Discovery

- **Cisco WAE devices automatically discover one another and negotiate optimization capabilities**
  - **Performed per TCP connection**
  - **Flexible optimization configuration using ATP**
  - **Exchange of peer capabilities and limitations**



Application Adapters Mitigate Latency

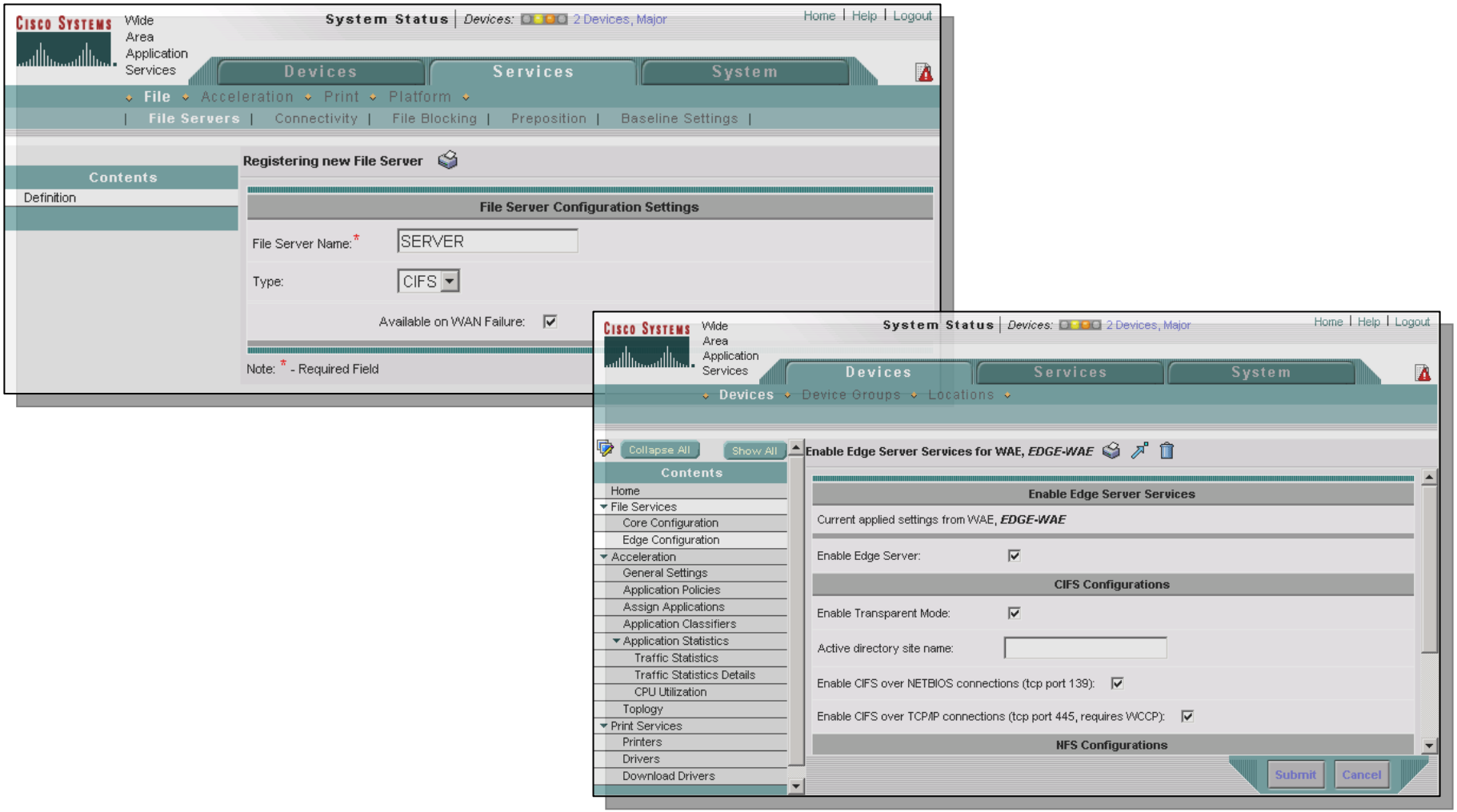
- **Application and protocol awareness**
  - **Eliminate unnecessary chatter and transfer**
  - **Pre-populate edge cache as necessary**
  - **Enable disconnected operations**
- **Intelligent protocol proxy**
  - **Transparent or non-transparent**
  - **Improves application response time**
  - **Provide origin server offload**
- **WAASv4 application adapters**
  - **CIFS (Windows File Services)**
  - **NFS (UNIX File Services)**
  - **Windows printing**



Application Adapters – File Services

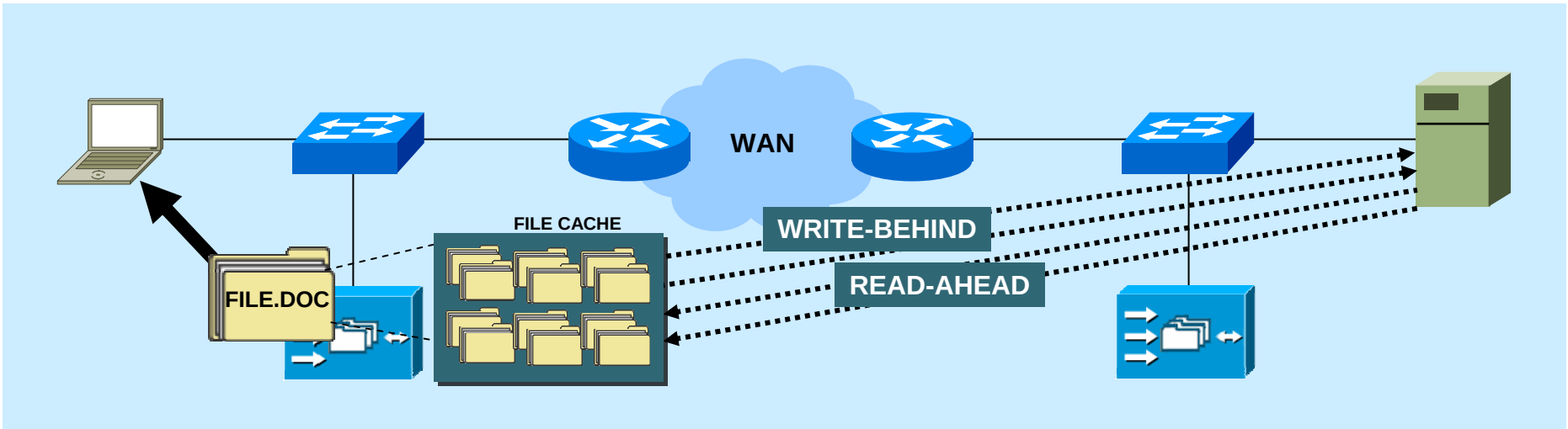
- **Wide Area File Services (WAFS) application adapter provides near-LAN performance to users when accessing remote file server storage**
  - **Latency mitigation – suppress unnecessary file protocol chatter and messaging**
  - **Safe read and write caching – serve validated, unchanged content to requesting users, avoiding WAN transmission**
  - **Scheduled data distribution – prepopulate edge cache to improve performance, i.e. software packages, patches**
  - **Transparent integration – maintains file server security, quotas, and access control**
  - **Disconnected mode of operation – WAFS provides read-only access to cached objects during periods of extended WAN outage**
- **Leverages WAN optimizations such as DRE, TFO, LZ compression provided by Cisco WAAS**

Application Adapters – File Services



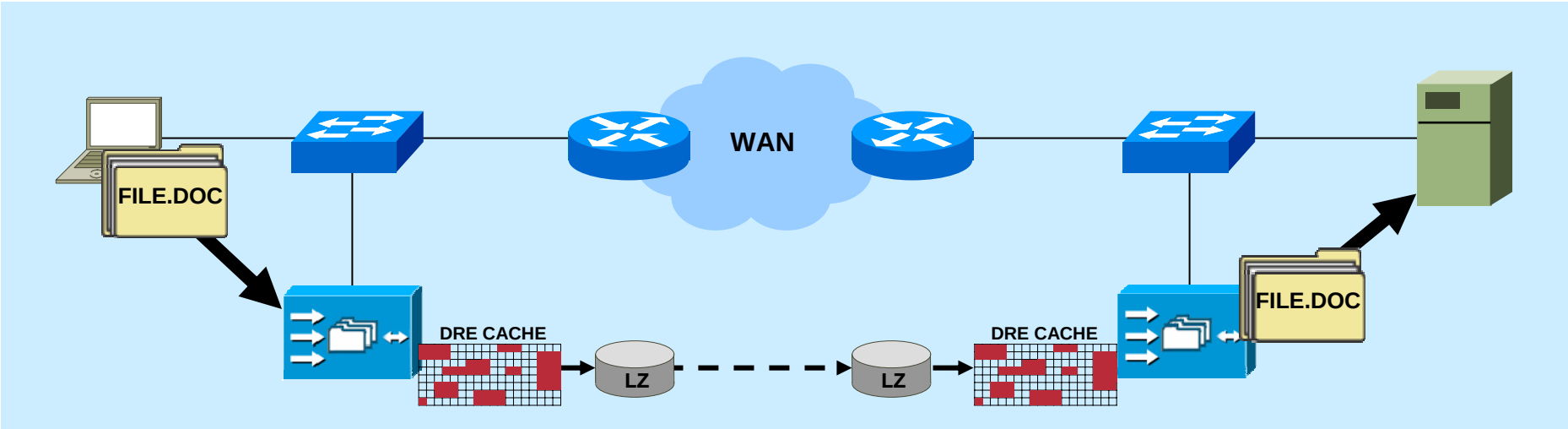
Caching Manages Bandwidth Usage

- **Application-specific caches serve usable, validated content to requestors to minimize bandwidth utilization for application content**
- **Employs segment-level object caching, read-caching (and read-ahead), write-caching (and write-behind) for optimal performance over the WAN**

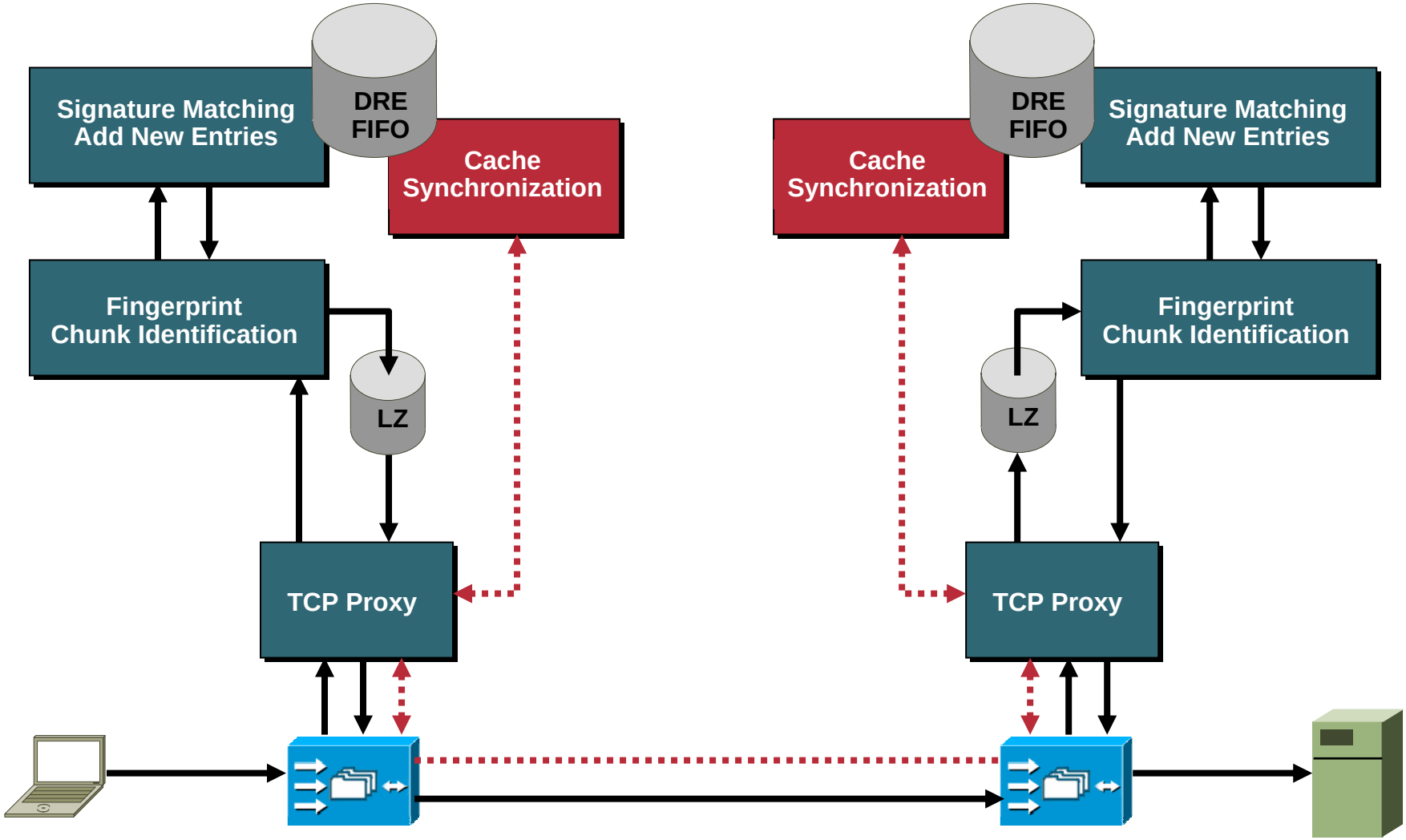


DRE and LZ Manage Bandwidth Utilization

- **Data Redundancy Elimination (DRE) provides advanced compression to eliminate redundancy from network flows regardless of application**
- **LZ compression provides generic compression for all traffic (even traffic with redundancy removed)**

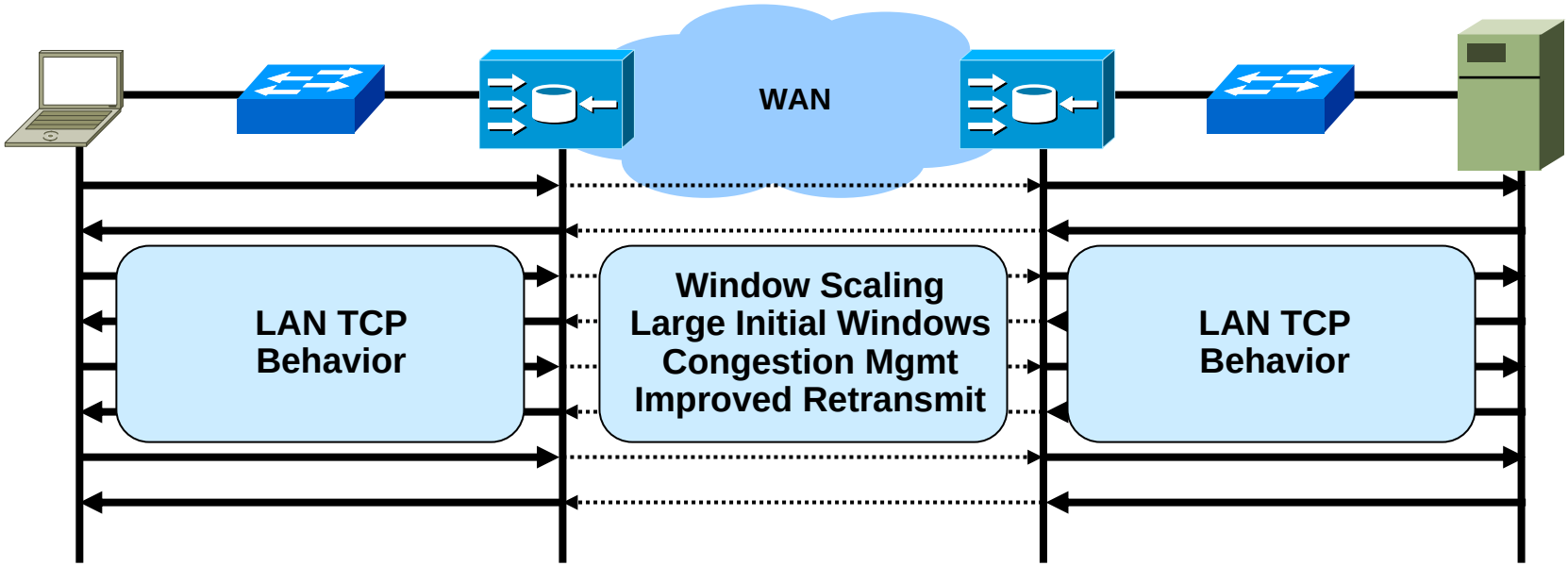


Advanced Compression Block Diagram



TFO Improves Application Performance

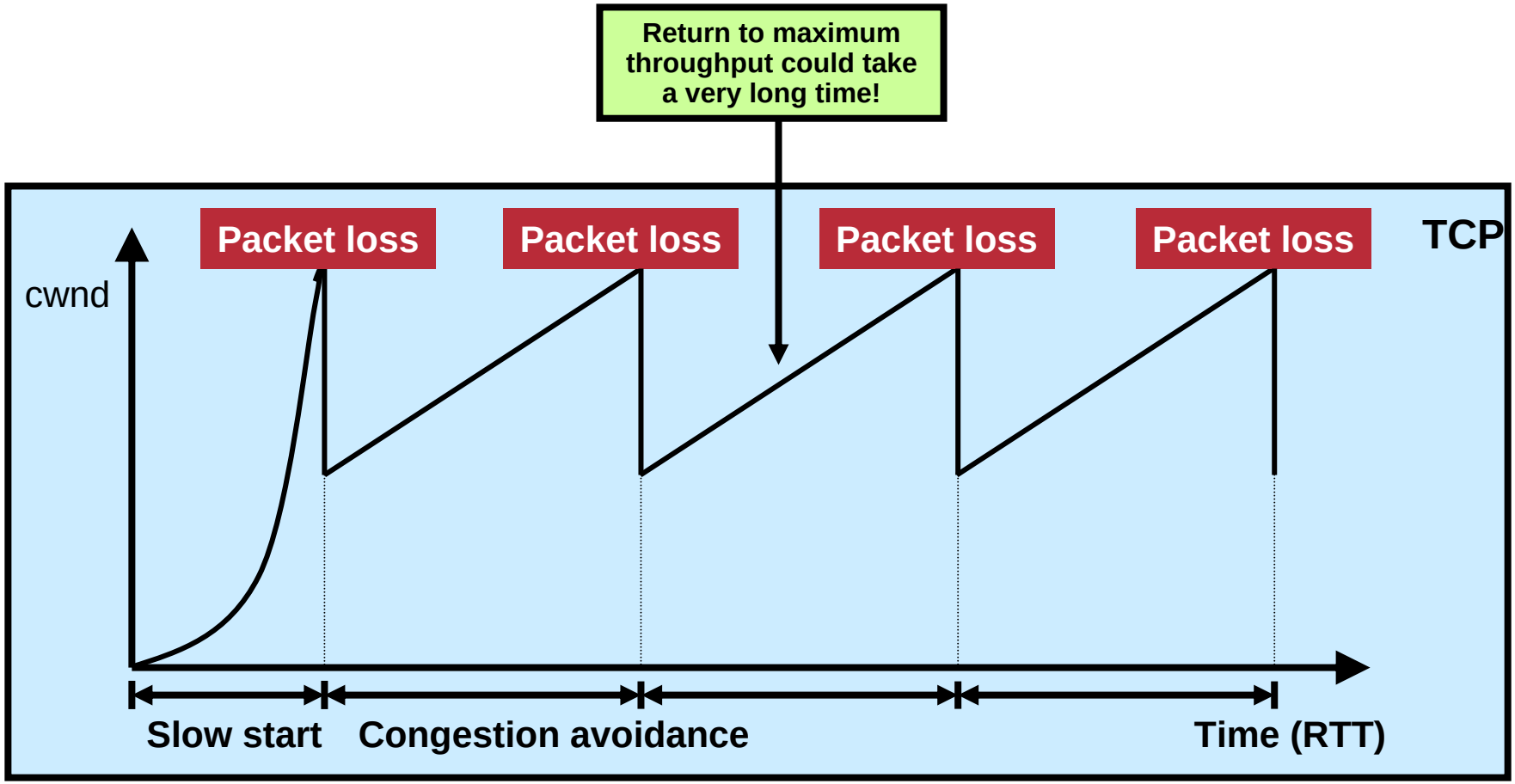
- **TFO overcomes TCP and WAN bottlenecks**
- **Shields nodes connections from WAN conditions**
  - **Clients experience fast acknowledgement**
  - **Minimize perceived packet loss**
  - **Eliminate need to use inefficient congestion handling**



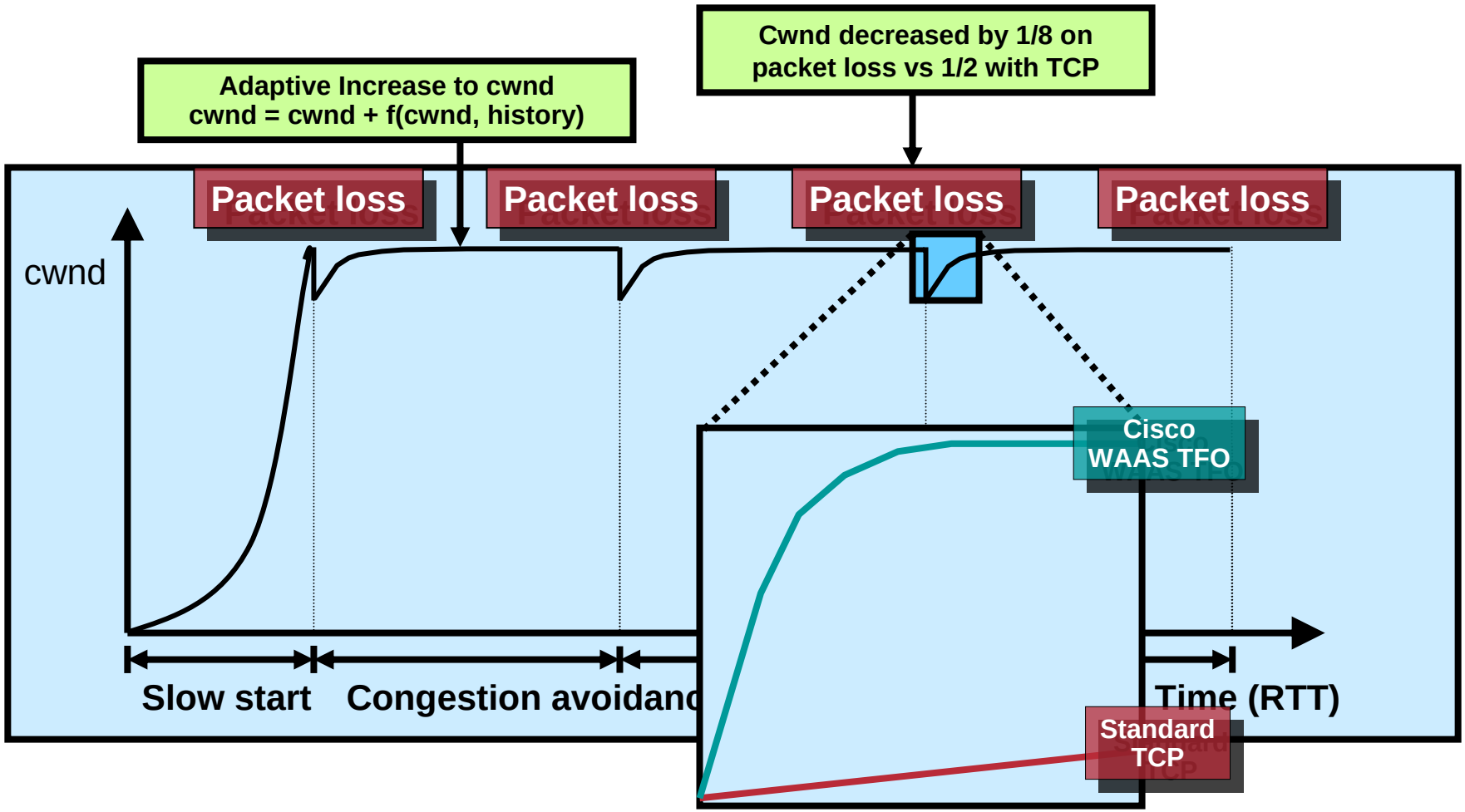
Cisco WAAS Transport Flow Optimizations

- **Cisco WAAS Transport Flow Optimizations (TFO) is designed to overcome common challenges associated with standard TCP implementations**
  - **Window Scaling – capitalize on available bandwidth**
  - **Large Initial Windows – maximize transmission after connection establishment**
  - **Selective Acknowledgement – efficient packet loss recovery and retransmission mechanisms**
  - **Binary Increase Congestion (BIC) – quick return to maximum throughput upon encountering congestion**

Standard TCP Throughput “Saw-tooth”

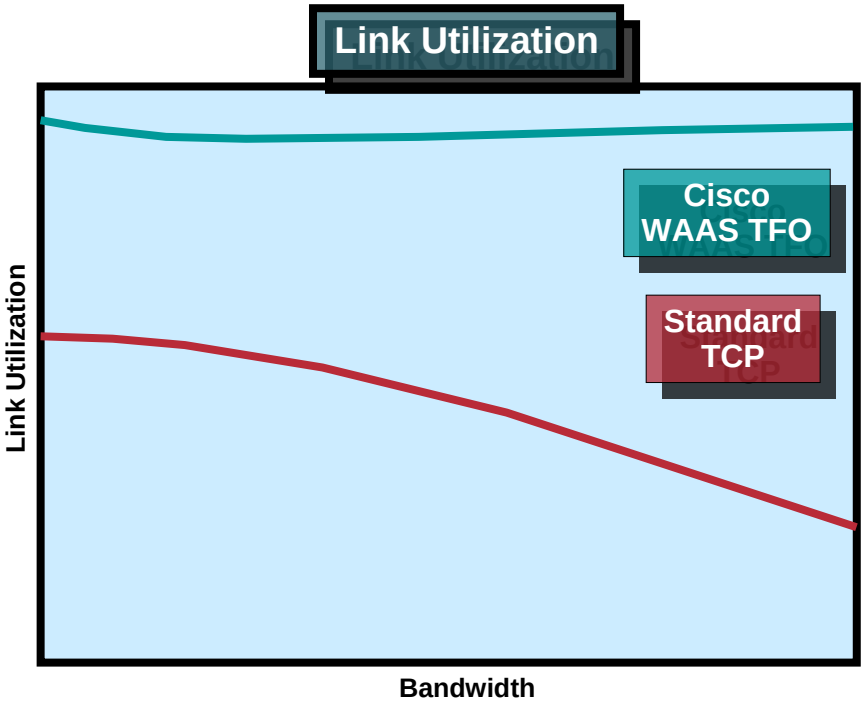


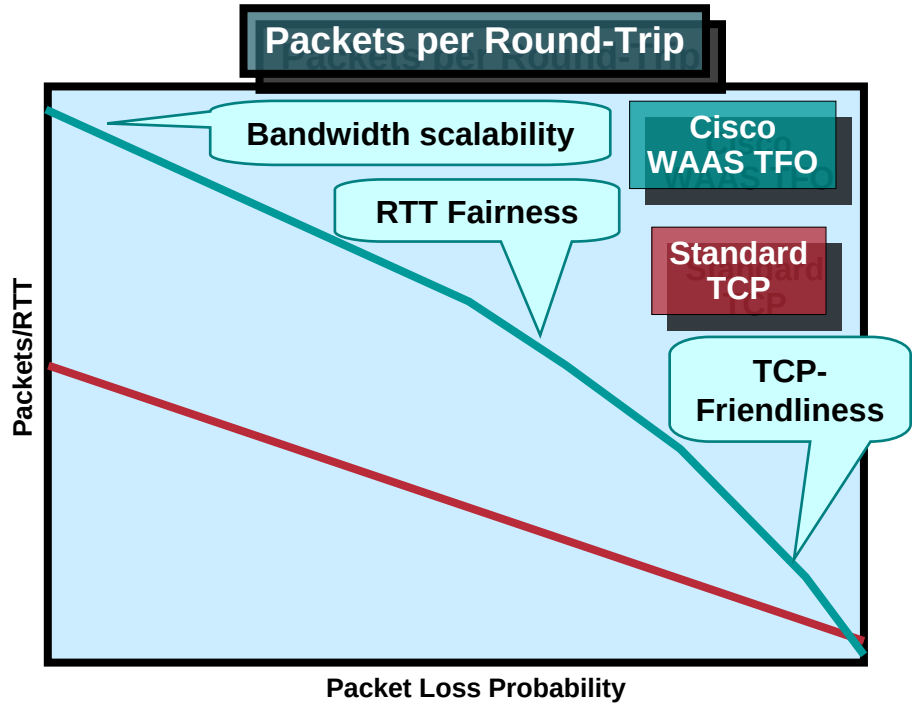
WAAS Throughput and Congestion Avoidance



TCP Throughput and Latency Optimizations

- TCP window scaling improves link utilization and throughput
- Optimized TCP stack improves recovery and congestion handling
- Priority for transactional traffic
- Compatible and friendly to other TCP connections on the network
- Large initial windows improves throughput for short-lived connections





DRE and TFO Application Performance

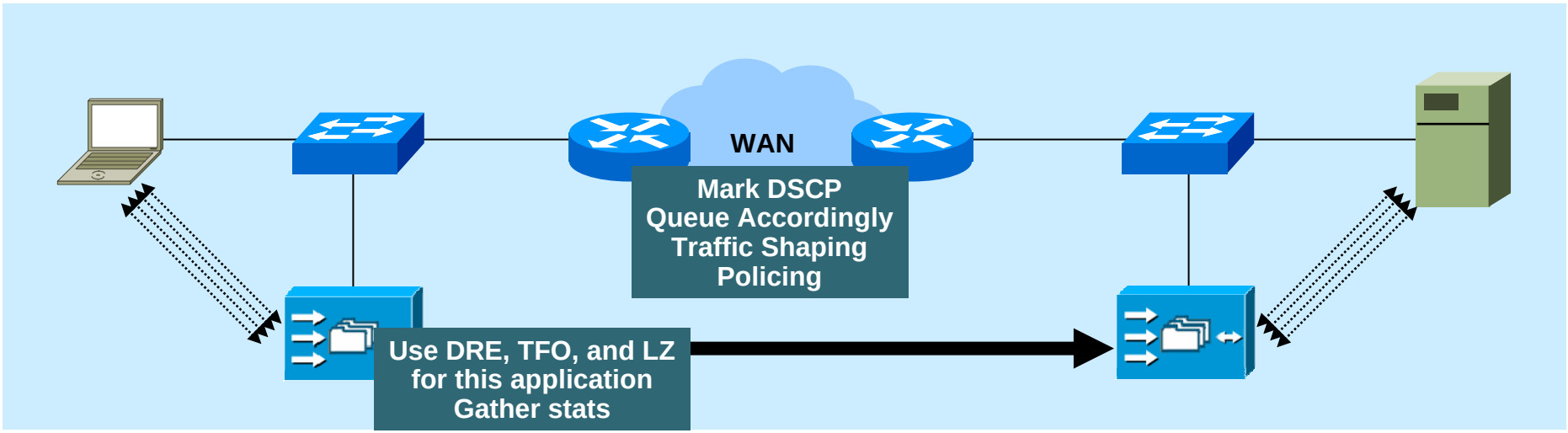
Test environment: 256Kbps circuit, 120mS roundtrip latency, .05% packet loss

	Average Aggregate Data Throughput (Kbps)			Average Response Time (sec)		
Test	Cisco WAAS	Vendor 1	Vendor 2	Cisco WAAS	Vendor 1	Vendor 2
MS Exchange	1218	1204	1193	.19	.2	.2
Internet Mail	111	109	115	.58	.57	.57
Lotus Mail	934	734	384	1.2	1.44	2.82
Oracle	543	533	579	7.82	8.13	7.34
SAP	598	384	474	1.44	2.17	1.81
Lotus Database	1535	1471	1776	2.27	2.64	2.26

*Note: Preliminary DRE and TFO performance test results are from August of 2005

Classification and Prioritization

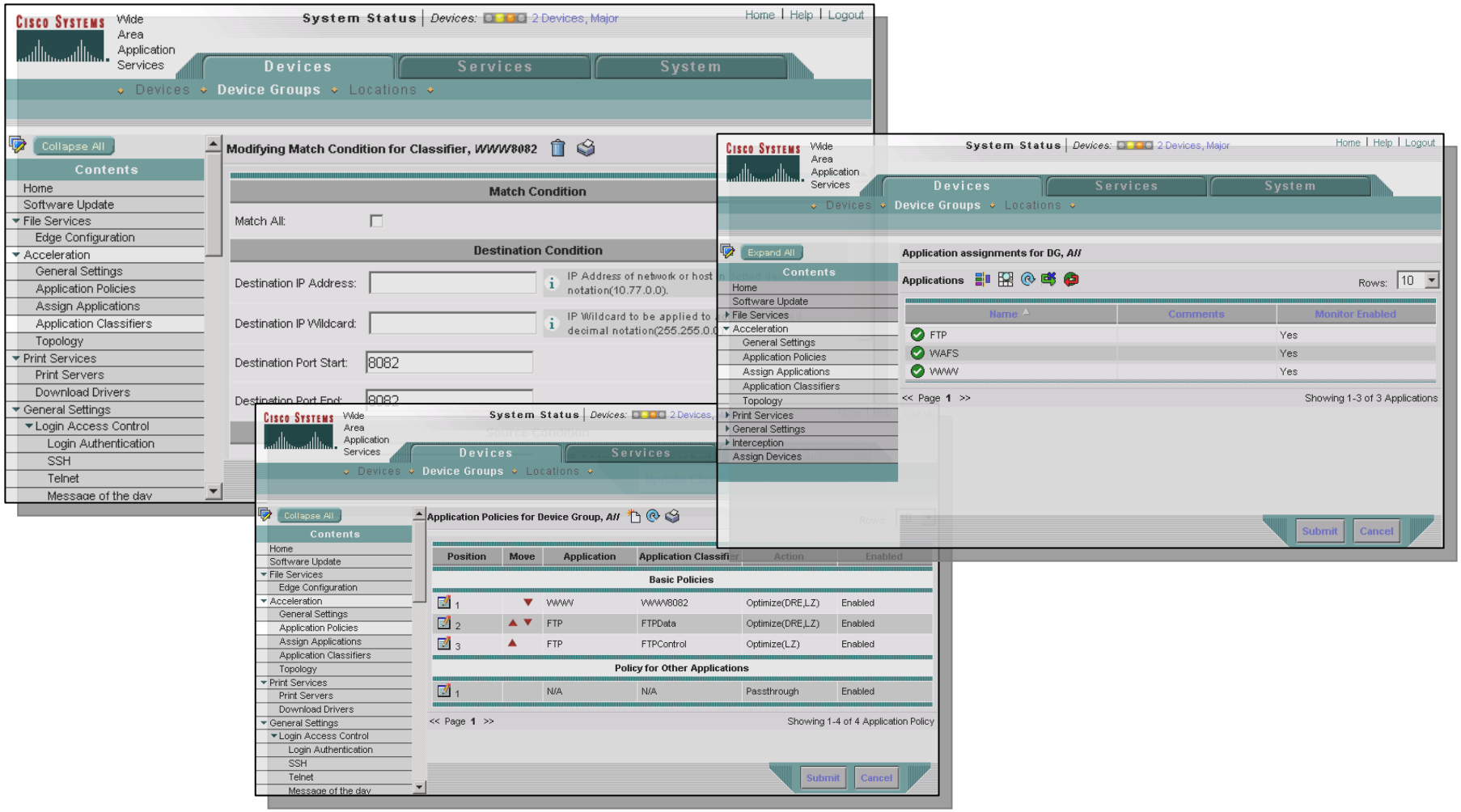
- **Application Traffic Policy (ATP) engine allows for the customizable configuration of optimizations per application protocol, monitoring, statistics**
- **Full support for IOS Quality of Service (QoS) and Network Based Application Recognition (NBAR)**



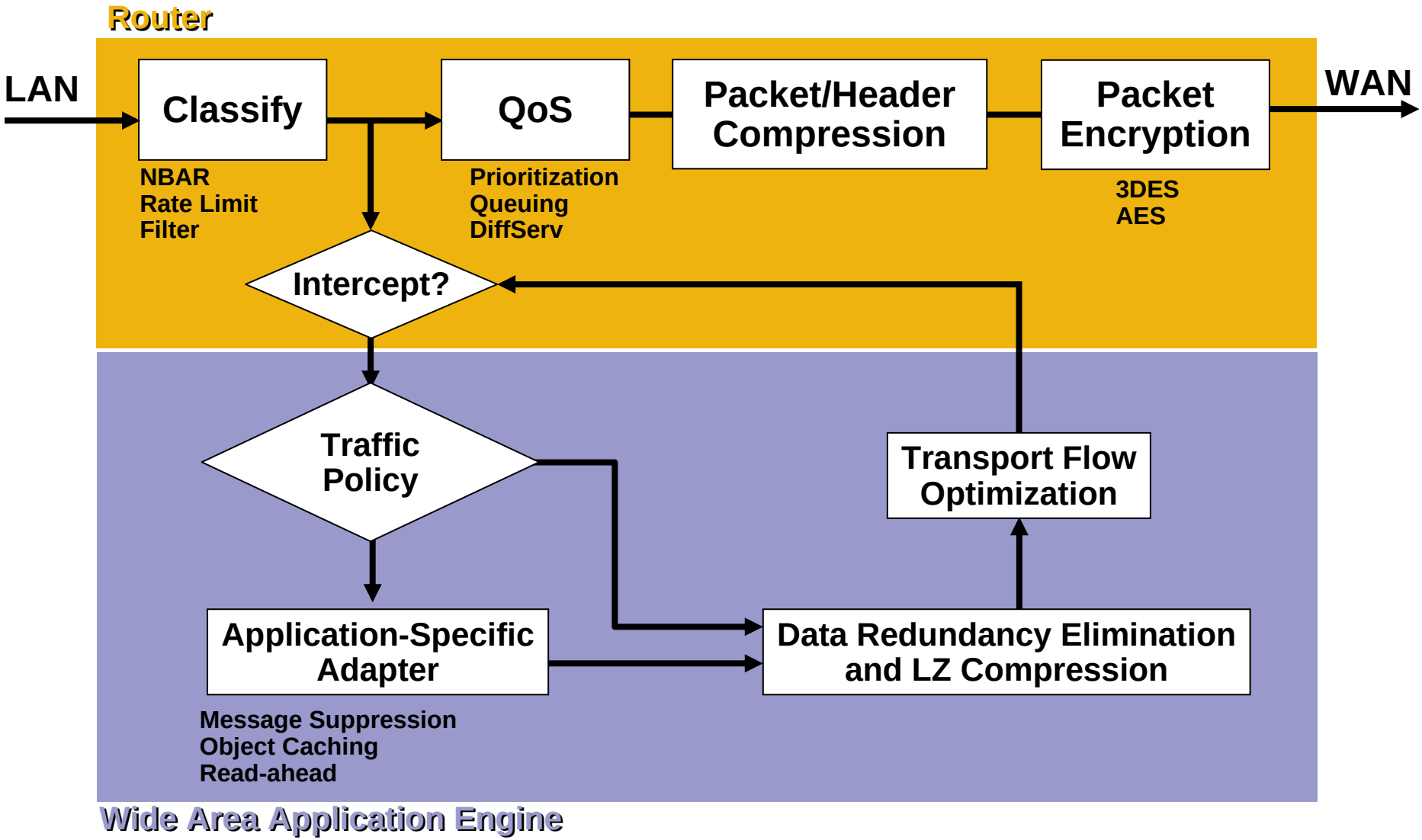
Application Traffic Policy Engine

- **Centrally managed from GUI as well as device CLI**
- **Classify traffic based on L3/L4 attributes**
  - **TCP Protocol Type**
  - **Source/Destination IP Address/Range**
  - **Source/Destination TCP Port**
- **Apply actions to classified traffic**
  - **Actions include: compress, DRE, TFO, protocol specific optimization**
  - **Bypass**
  - **Monitor**
- **Default policies for common applications**
  - **Includes 60+ policies for Citrix, MAPI, SQL, others**
  - **Fully customizable classification and optimizations**

Application Traffic Policy Engine

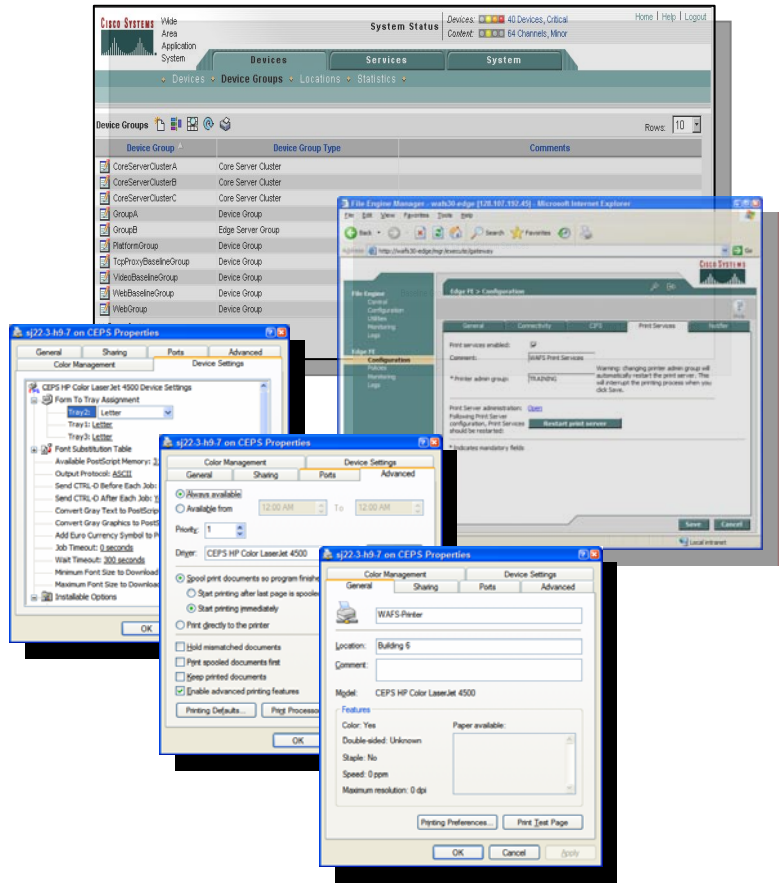


Router-Integrated Packet Flow



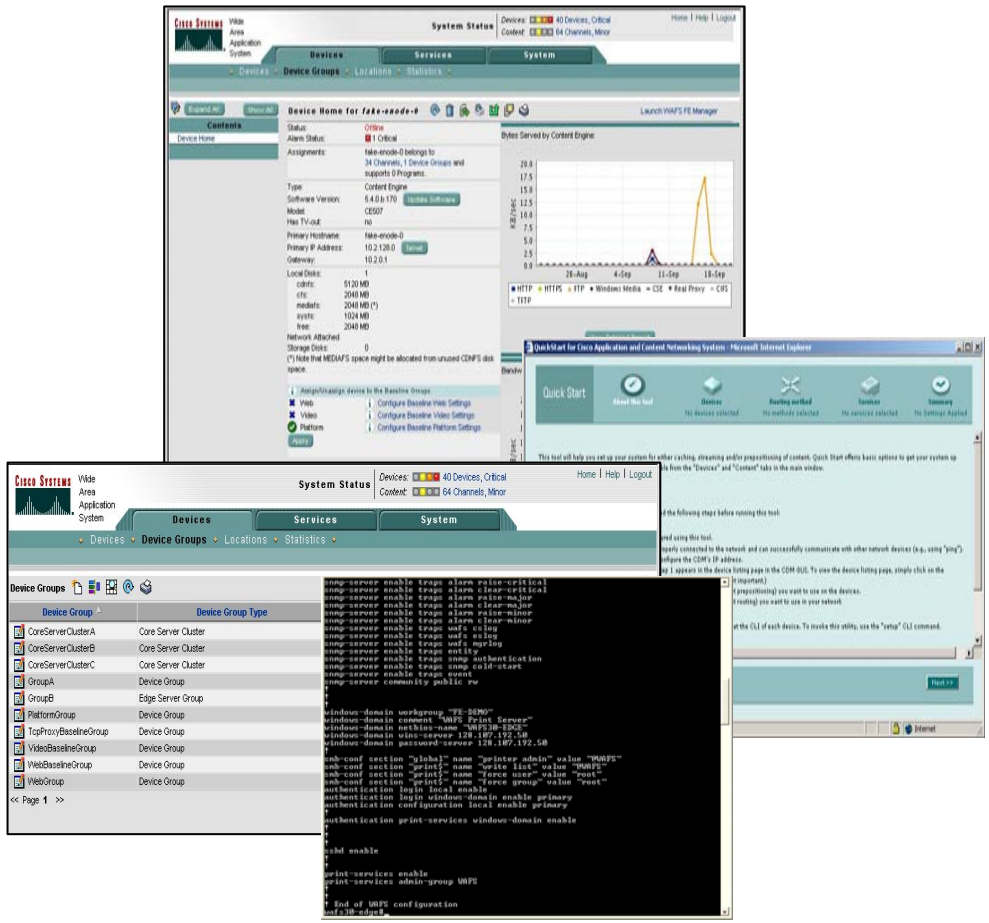
Windows Compatible Print Services

- **Centrally Managed Print Services**
  - Client driver distribution
  - Status and health reporting
- **Supports Any Printer**
  - Full feature compatibility
  - Job control and status monitoring
  - Guest and disconnected printing
- **Print Server Configuration**
  - Network parameters (IP, name, etc)
  - Queue definition and ACLs
  - Standard Windows wizards for settings, accessories



WAAS Intuitive Central Management

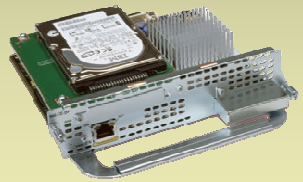
- **Comprehensive Management**
  - **Central configuration**
  - **Device grouping**
  - **Monitoring, statistics**
  - **Alerts, reporting**
- **Easy-to-use Interface**
  - **Graphical U/I, Wizards**
  - **IOS CLI**
  - **Roles-based administration**
- **Proven Scalability**
  - **1000’s of nodes**
  - **Redundancy and recovery**



Cisco WAE Hardware

- **Cisco WAN Application Engine (WAE) hardware family provides unified platform for application optimization and delivery**
- **Appliance models deliver 80GB – 1.8TB of capacity and provide data center scalability, reliability, and delivery for the enterprise edge**
- **Network-integrated module provides optimization and delivery for the enterprise edge**

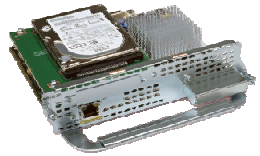
Cisco WAE Appliances



Cisco WAE Network Module

Cisco WAE Hardware Specifications

Platform	Hardware	Positioning
NM-CE	• 500MHz Pentium III 512MB of memory 80GB ATA	• Integrated branch services for small deployment
WAE-511	• 2.8GHz Celeron 1GB memory 80GB to 500GB SATA	• Dedicated appliance for small-medium edge or core deployments
WAE-611	• 3.0GHz Pentium IV 2GB of memory 288GB SCSI	• Dedicated appliance for medium-large edge or core deployments
WAE-7326	• Dual 3.2GHz Xeon 4GB of memory 73GB to 1.8TB SCSI	• Dedicated appliance for enterprise edge or core deployments



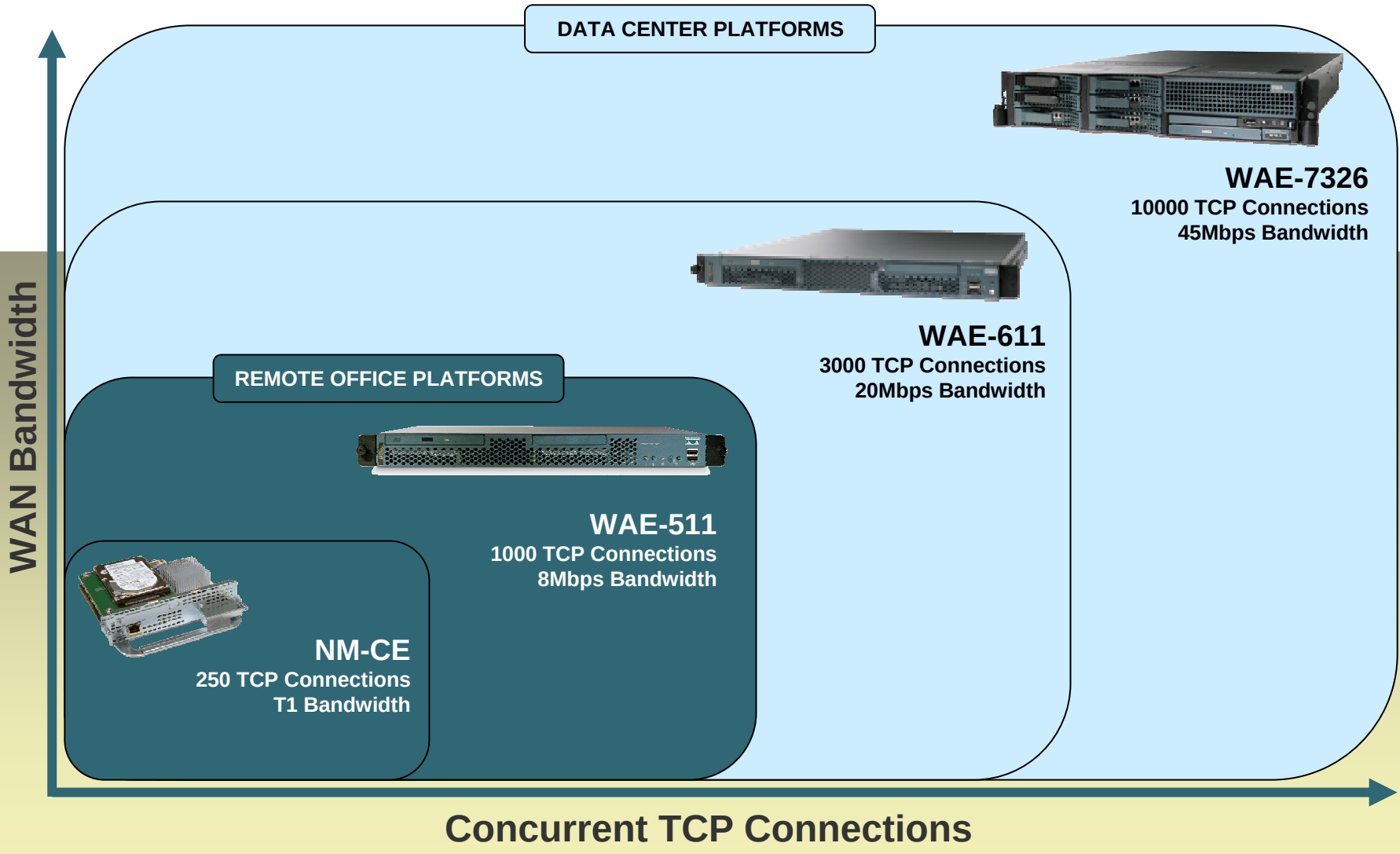
NM-CE

WAE-511

WAE-611

WAE-7326

Cisco WAE Hardware Positioning



Cisco WAE Family Performance and Scalability

Hardware	CIFS Sessions (Edge)	CIFS Sessions (Core)	Optimized TCP Connections	WAN Throughput (Uncompressed)
NM-CE	100	N/A	250	1.5Mbps
WAE-511 512MB Memory	250	1000	500	4Mbps
WAE-511 1GB Memory	500	2000	1000	8Mbps
WAE-611	1250	5000	3000	20Mbps
WAE-7326	2500	10000	10000	45Mbps

*** Recommended capacity numbers, based typical user and branch definitions. Actual numbers are case-specific and results may vary.**

Why Choose Cisco WAAS?

- **Scalability: Proxies offload WAN and scale data center**
  - **Fewer messages reach origin server**
  - **Stateless core lends to simple N+1 core clustering**
- **Availability: Safe and simple failover**
  - **Stateless core implies simple failover with minimal impact**
  - **Allow for offline operation under certain circumstances**
- **Performance: Better and more scaleable performance**
  - **Fewer messages utilize the WAN, decreasing application sensitivity**
  - **Generic network optimizations dramatically improve throughput**
- **Correctness**
  - **Safe read-ahead, write-back, message batching, and pipelining**
  - **Elimination of unnecessary application control traffic**

Why Choose Cisco WAAS?

- **Industry-Leading Network Integration**
  - **Transparent integration – no client or server changes**
  - **Network hardware platform modules and linecards**
  - **Compliance with advanced network functionality and topologies**
- **Scalable Central management**
  - **Proven capability to manage 1000's of devices**
  - **Directory-integrated central management, local appliance GUI**
- **Global Service and Support**
  - **Cisco's award-winning 24x7 Technical Assistance Center**

Trends Impacting Application Delivery

Use of “Chatty” Web Protocols

Geographical Dispersion of Users & Applications

Convergence of Data Centers

Increasing Security Requirements

Result: Performance declines · Security Decreases

Performance Impedes Web-based Business

Applications Increasingly Represent the Business

- Web promoted process automation across the Extended Enterprise
- Pressure to move from staff to software drives productivity

Solutions Should Serve All Users, Everywhere

- Consolidation encouraging one-to-all application support for all
- Fewer and fewer managed endpoints
- Security adds necessary overhead

Applications Give Unprecedented Access to Critical Business Data

- Private Data, Confidential, Intellectual property
- Developers write for functionality
- To expensive to fix all security bugs

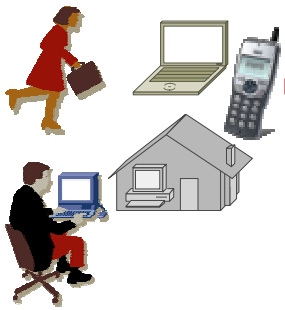
Current IT Approaches Fall Short

Operations

- More servers
- More management tools
- Re-architect infrastructure

Applications

- More testing
- Rewrite applications
- Security patching



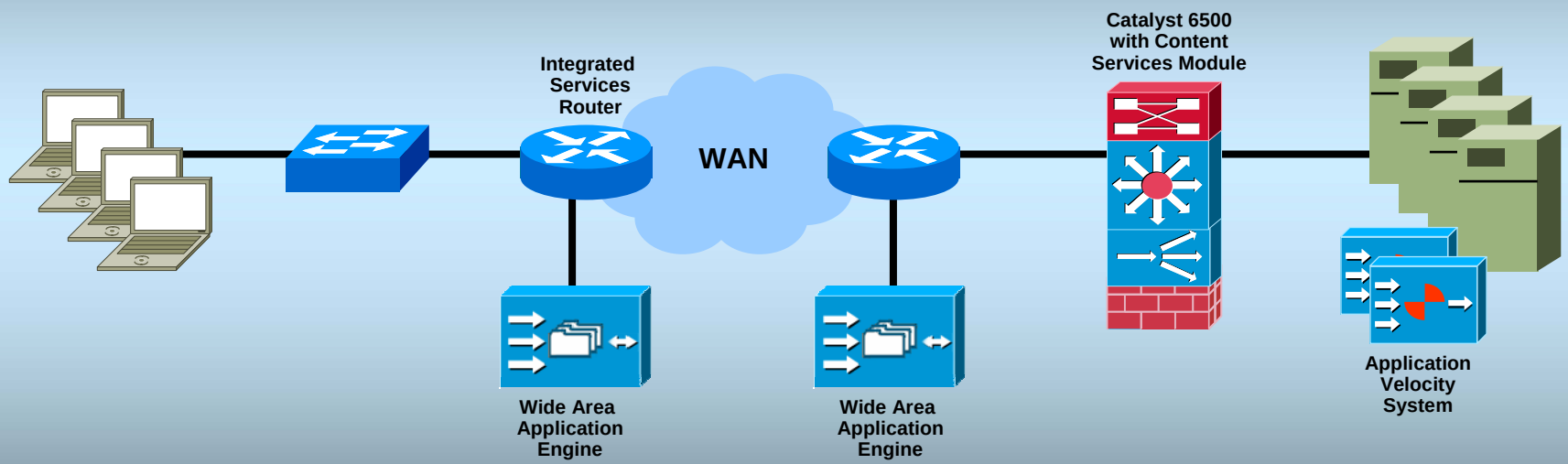
Network

- More bandwidth
- Many point products
- Replicate data centers

Application Optimization Infrastructure

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
- Service and availability monitoring



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

WAN Acceleration
- Data redundancy elimination
- Window scaling
- LZ compression
- Adaptive congestion mgmt

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

Web Application Acceleration & Web Firewall Solution

- **Optimize at Layer-7**
  - **2X–response time improvements**
  - **80% decrease in bandwidth requirements**
  - **80% fewer server cycles**
- **Stop application hacking**
  - **Safely deploy applications**
  - **Secure mission critical data**
  - **Streamline operations**

Secure, Fast & Reliable Applications

Cisco AVS 3120

AVS Delivers Real-World Value

Application	Software	AVS Improvement	Cost of Likely Alternative	Business Impact
Call Center (High tech)	PeopleSoft	↑270%	$4MM (2 New Overseas Data Centers)	• Meet support goals with no additional staffing or costs
Purchasing (Manufacturing)	SAP	↑350%	$5MM (Multiple Overseas Data Centers)	• Increase procurement automation
Mortgage Origination (Financial)	Custom J2EE: WebSphere	↑300%	$2MM (Reengineer Apps and Infrastructure)	• 30% more transactions across same infrastructure
Claims Management (Insurance)	Custom J2EE: WebSphere	↑220%	$3MM (No Reengineering)	• Support “zero-footprint” branch
B2B Operations (Retail)	Plumtree	↑350%	$500K annually (Upgrade 650 Sites)	• Move all costly paper-based processes online
CRM (Financial)	Siebel	↑290%	$2.4MM annually (Upgrade 200 U.S. Locations)	• Immediate jump in CRM usage with Improved account retention

Huge Lead in Acceleration Features

Functional Areas	AVS Acceleration Features
Latency Reduction	▪FlashForwarding* ▪Browser TCP multiplexing* ▪PDF download optimization ▪Response redirection control*
Bandwidth Reduction	▪ GZIP Compression ▪ Delta encoding* ▪ Dynamic browser caching* ▪ Dynamic image optimization ▪ Flexible processing rules
Server Offload	▪ TCP Offload ▪ SSL Offload ▪ RAM Caching ▪ Dynamic caching* ▪ Load-based caching* ▪ Lazy request evaluation* ▪ Single sign-on optimizations ▪ XML merging/transformation

Application Acceleration Examples *FlashForward*

- **Embedded objects referenced in HTML container pages are served with Expires: which sets expiry in the future.**
- **On 2nd visit Browser will not send GET for objects in cache if the current date & time is not greater than the object expiry date.**
- **This reduces the total number of HTTP requests for subsequent visits to the same page.**
- **Benefits:**

  Decreased page download time

  Decreased network congestion

  Decreased number of requests to origin server

Application Acceleration Examples
Delta Encoding

- **HTML pages today are largely dynamically generated making it not cacheable**
- **Browser must download entire page each visit.**
- **Delta works by calculating and sending only the difference between two visits to an dynamic HTML page**
- **Benefits:**

  Reduced bandwidth usage

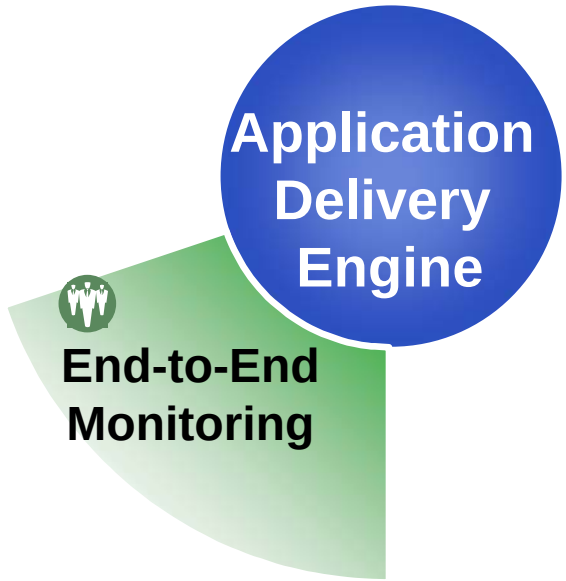
  Reduced page download times

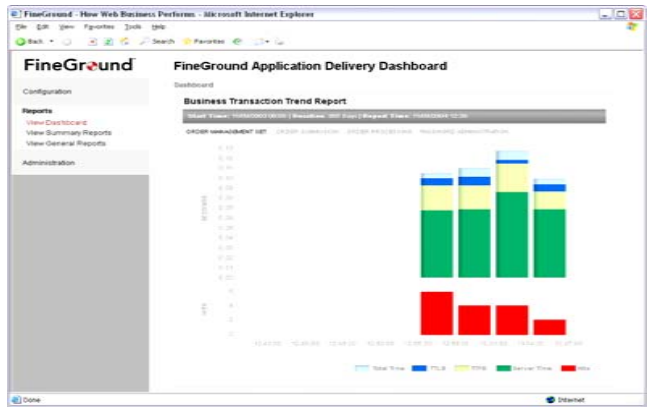
  Works in combination with other optimizations

Application Monitoring

- **End-user response time monitoring**
  - **Actual users and transactions**
  - **Business- and process-level aggregation**
  - **Full drill-down to page and location**
- **“Drop-in” deployment**
  - **No changes to application or desktop**
  - **Data center installation**
- **Delivery Dashboard and flexible reporting**
  - **Wizard-based transaction builder**
  - **Support for Enterprise Consoles (BMC, Tivoli, OpenView…)**
- **Benefits**

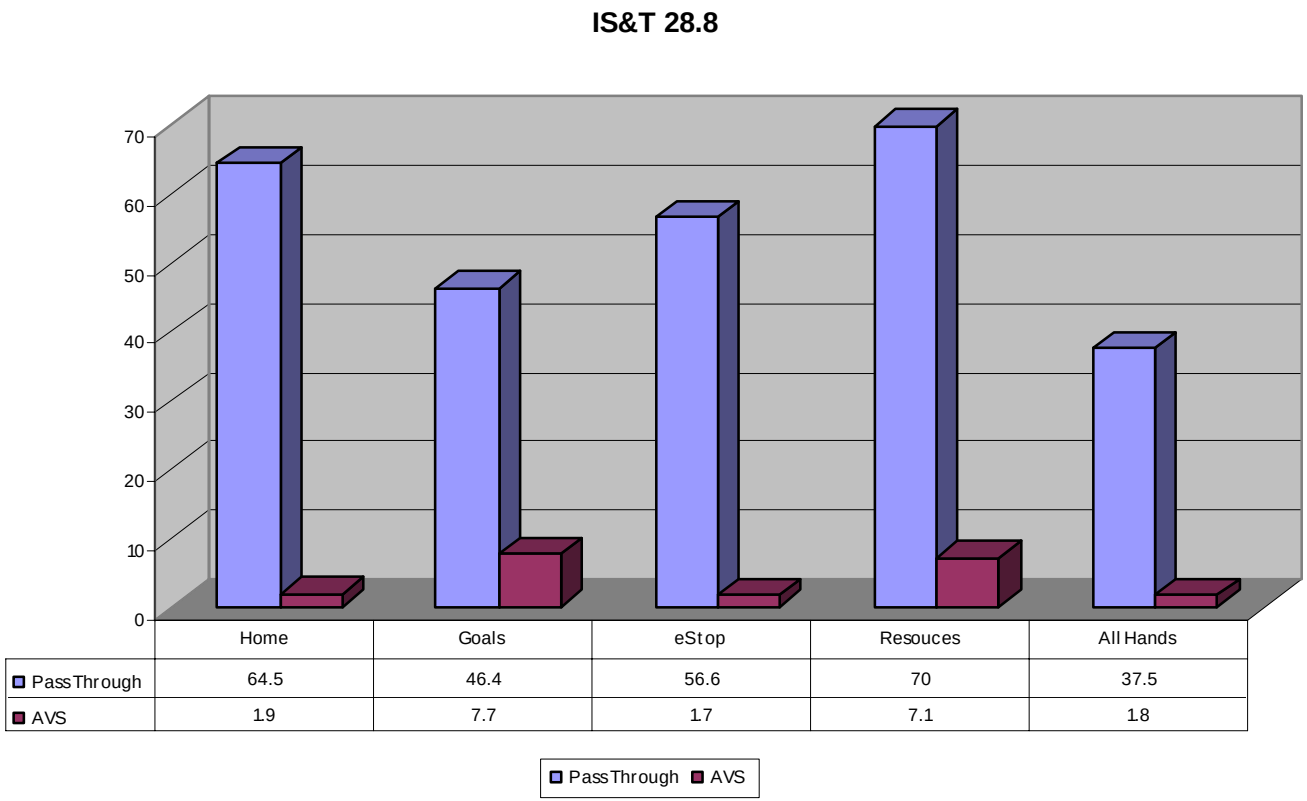
  End-user visibility
  First-line problem triage
  Reduce mean-time-to-repair





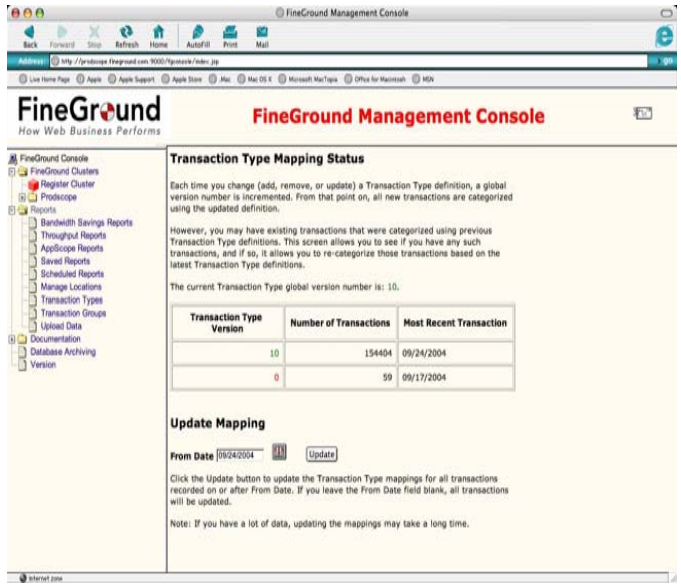
Sample Performance Report

- **Performance Improvement**



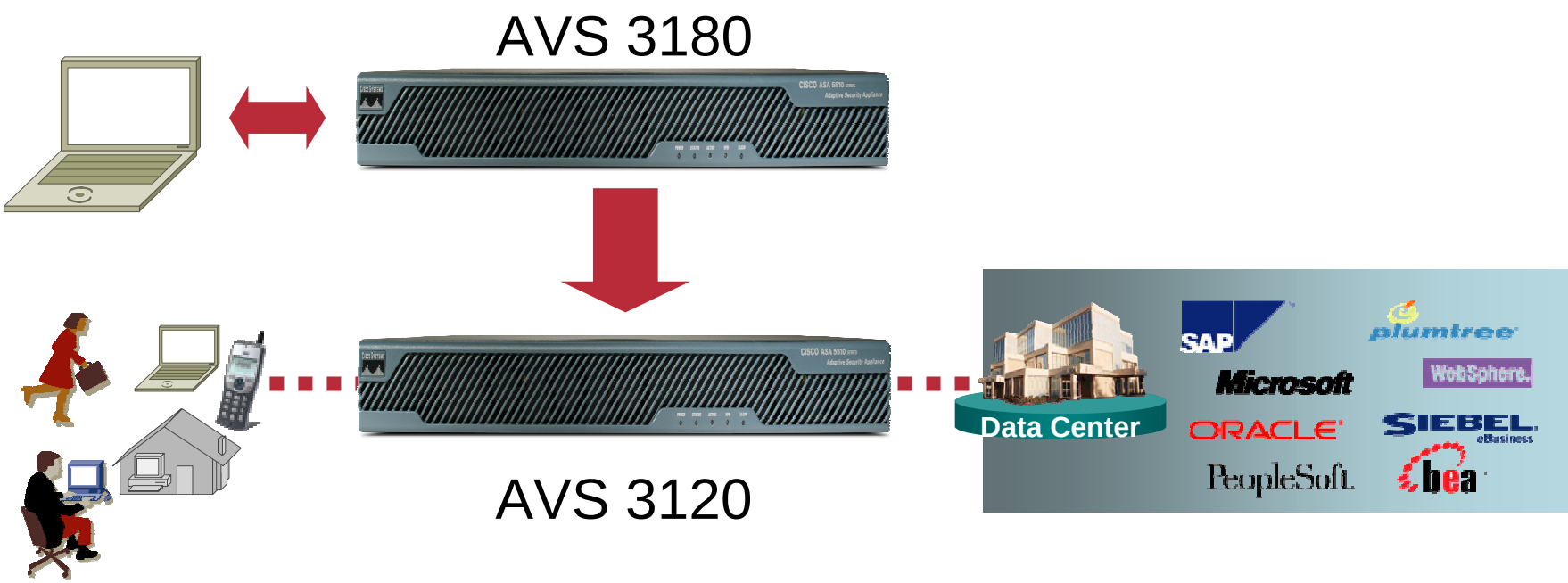
Management Console

- Browser based management
- Scalable and hardened appliance platform
- Powerful management console
  - Local and remote management
  - Integrated reporting, diagnostics, and alert mechanisms
  - Fully SNMP compliant
- Application scalability and failover via multiple certified options
- Integral part of the and products



Monitoring Requires AVS 3120 & 3180

- **AVS 3180 Polls the 3120 for Performance Data**
- **Browser into the 3180**
- **AVS 3180 also allows you to manage one or more AVS 3120**



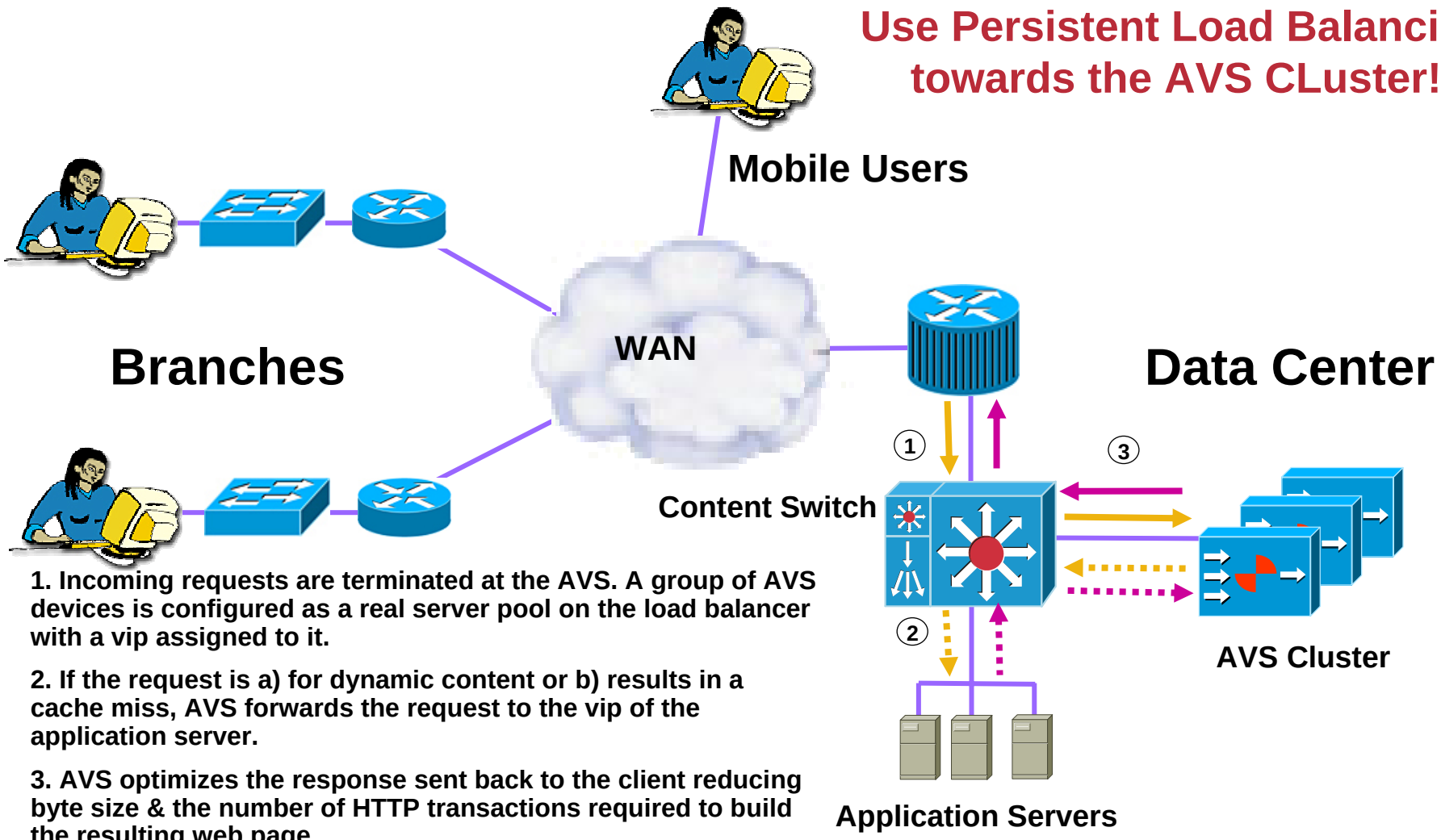
AVS Delivers Applications Securely

INSPECTS FOR:

SQL Injection
Cross-Site Scripting
Command Injection
Cookie/Session Poisoning
Application Reconnaissance
LDAP Injection
Buffer Overflows
Directory Traversals
Attack Obfuscation
Application Platform Exploits
Zero Day Attacks
Cookie Poisoning
Parameter Tampering

- **Bi – Directional Deep Inspection**
- **Positive & Negative Security**
- **Protocol compliance and anomaly detection**
- **Transaction logging and report for application security forensics**

AVS Deployment Scenario

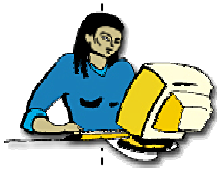


1. Incoming requests are terminated at the AVS. A group of AVS devices is configured as a real server pool on the load balancer with a vip assigned to it.

2. If the request is a) for dynamic content or b) results in a cache miss, AVS forwards the request to the vip of the application server.

3. AVS optimizes the response sent back to the client reducing byte size & the number of HTTP transactions required to build the resulting web page.

Browser behavior w/o FlashForward



GET Index.html

HTTP 200 OK

IMS Foo.gif

304 Not Modified

IMS Foo.js

304 Not Modified

IMS Foo.jpg

304 Not Modified

IMS Foo.css

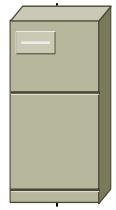
304 Not Modified

IMS bar.gif

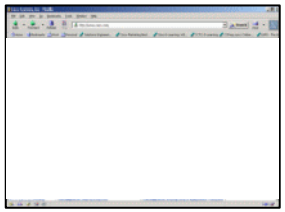
304 Not Modified

IMS bar.js

304 Not Modified

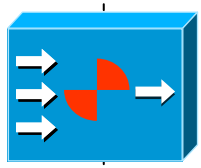


Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired



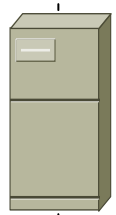
Client Requests HTML page

GET Index.html

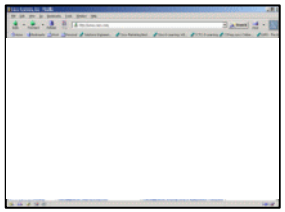


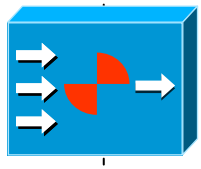
AVS fowards request to Server

GET Index.html



Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired





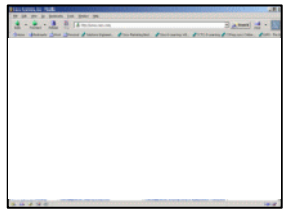
Server returns updated content

GET Index.html

HTTP 200 OK

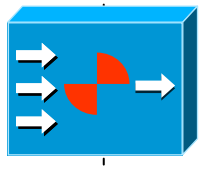
AVS Parses the Container HTML For Embedded Objects

Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired



AVS validates embedded objects

2 Objects have changed



304 Not Modified

304 Not Modified

HTTP 200 OK

HTTP 200 OK

304 Not Modified

304 Not Modified

IMS Foo.gif

IMS Foo.js

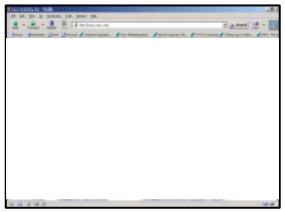
IMS Foo.jpg

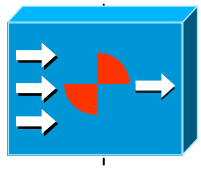
IMS Foo.css

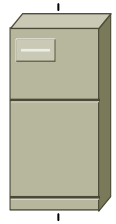
IMS bar.gif

IMS bar.js

Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired







AVS updates the references to embedded objects in HTML

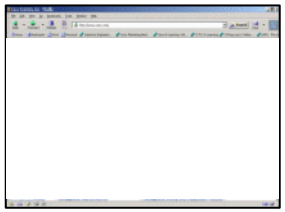
Object Reference:

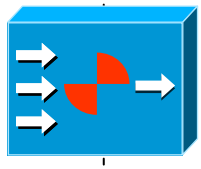
```
<img
src="/images/Foo.jpg">
```

Transformed Obj. Ref:

```
<img
src="/images/Foo_FineGr
ound_vtmmi14xg2fvmlkxsx
uk0ty1xd_FGN_V01.jpg">
```

Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired



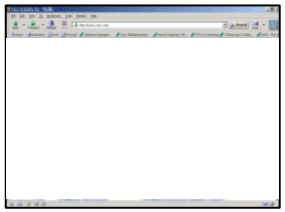


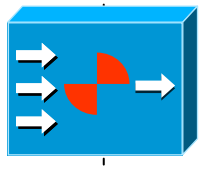
As objects change AVS updates the new object references in HTML

Transformed Obj. Ref:

```
<img
src="/images/Foo_FineGround
_vtmmi14xg2fvmlkxsxuk0ty1xd
_FGN_V01.jpg">
```

Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired



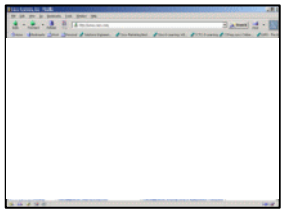


The URL is updated to include a new MD5 hash

New transformed Obj. Ref:

```
<img
src="/images/Foo_FineGround
_erhjk2asfn7pjyyfplfasd12ed
_FGN_V01.jpg">
```

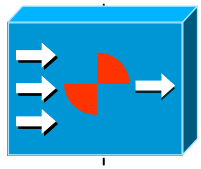
Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired

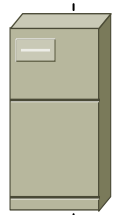



Client downloads HTML page

The requested HTML with updated object references is now sent to the client

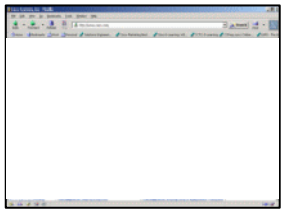
HTTP 200 OK





Client parses HTML and retrieves fresh embedded objects from browser cache

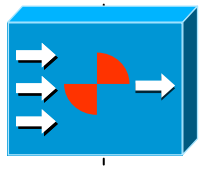
Browser behavior with FlashForward 2nd Visit –AVS cacheTTL is expired





Client downloads new updated objects referenced in HTML Not found in cache.

GET foo_fgn_.jpg



HTTP 200 OK

GET foo_fgn_.css

HTTP 200 OK

Only 3 round trips across the WAN were required to build the updated page.

Round trip latency is avoided with each Flash Forwarded object cached in the browser's cache.