

# Application Intelligence in the Network

Issy Ben-Shaul

CTO, Application Delivery

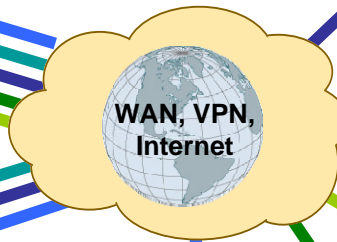
Cisco Systems, Inc.

**INTEROP**<sup>®</sup>

BUSINESS. TECHNOLOGY.  
ONE WEEK. ONE PLACE.

# The Globally Integrated Enterprise

- Hire globally based on skills, cost, and business
- Operations horizontally and globally integrated
- Strategy and management shaped globally



ORACLE SAP plmtree  
PeopleSoft SIEBEL WebSphere

Enterprise Applications

HTTP, HTTPS

MAPI, IMAP, WebDAV

Microsoft Exchange Server 2003 Lotus

E-mail Servers

ICA, TN3270

CITRIX RUMBA  
EXTRA!

Legacy Application Servers

MMS, RTSP/RTP

Windows Media real

Streaming Media Servers

CIFS, NFS, WebDAV

NetApp EMC<sup>2</sup> Windows

Servers

DATA CENTER

- Multiple applications
- Distributed users – partner, supplier
- Complex application environments
- Security and data management concerns

# What is App Intelligence ?

- Passive
  - Classify apps (payload inspection)
- Active
  - Control (police) admission of flows/apps
  - Prioritize resource allocation per app/flow
  - Optimize apps
- Adaptive
  - Learn the characteristics of flows/apps
  - Adjust active policies based on identified characteristics
- **PERFORM ALL OF THE ABOVE at L7/8**

# Why App Intelligence: Passive Classification

- App-level visibility
  - monitoring & troubleshooting
  - Traffic engineering: SLAs, bandwidth allocation & latency prioritization
  - App-based optimization
  - Comment: tunneling hinders visibility!
- L7/8 (payload inspection) essential
  - Multiple apps on same L7 protocol/port
  - Dynamic port apps
  - Same app may run on different ports

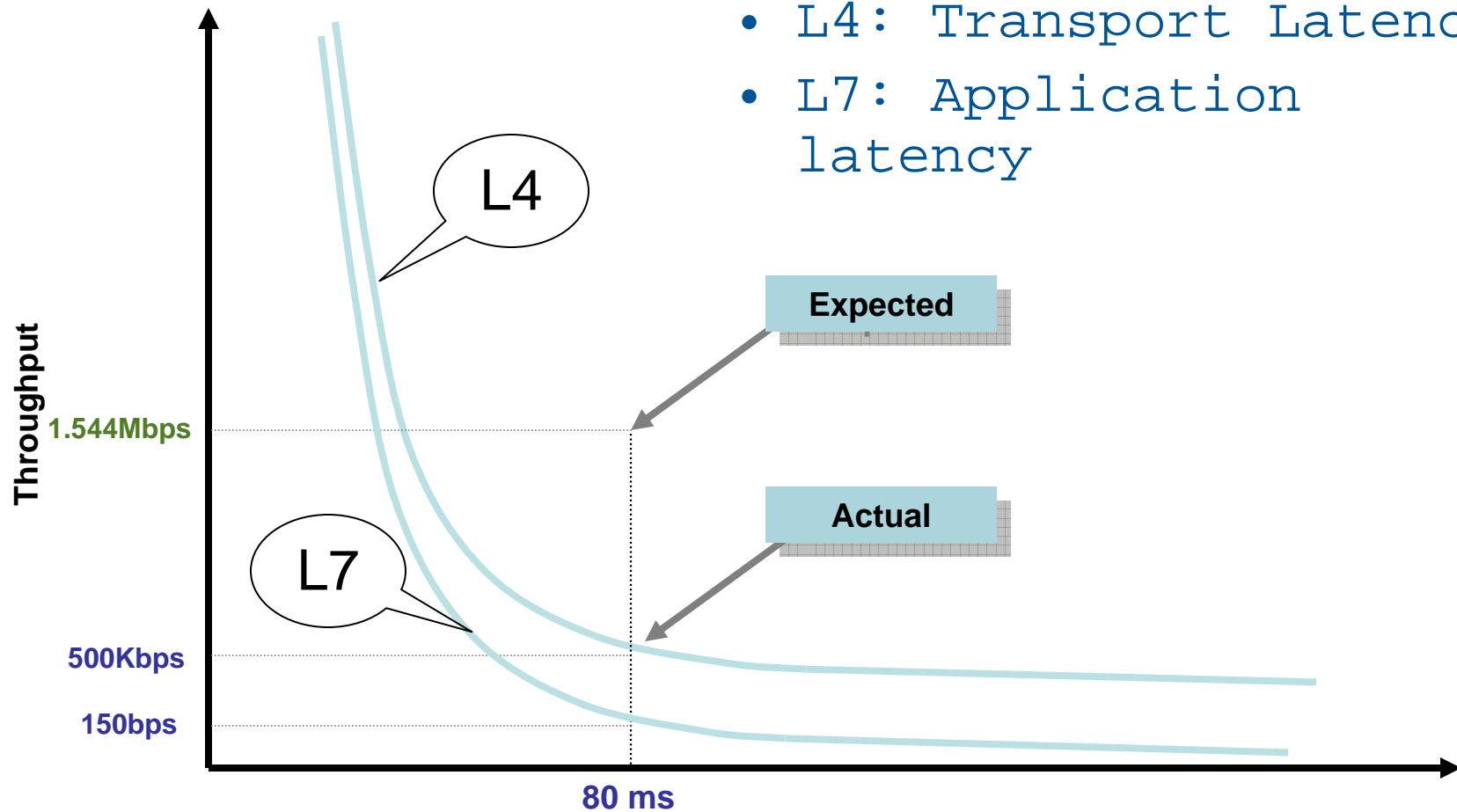
# App Intelligence: L7/8

## WAN Optimization

- Business Drivers
  - Branch and data-center consolidation
  - Application Acceleration
  - Bandwidth Savings
- L3/4 optimization is not sufficient due to app latency

# The Impact of Latency

- L1: Network latency
- L4: Transport Latency
- L7: Application latency

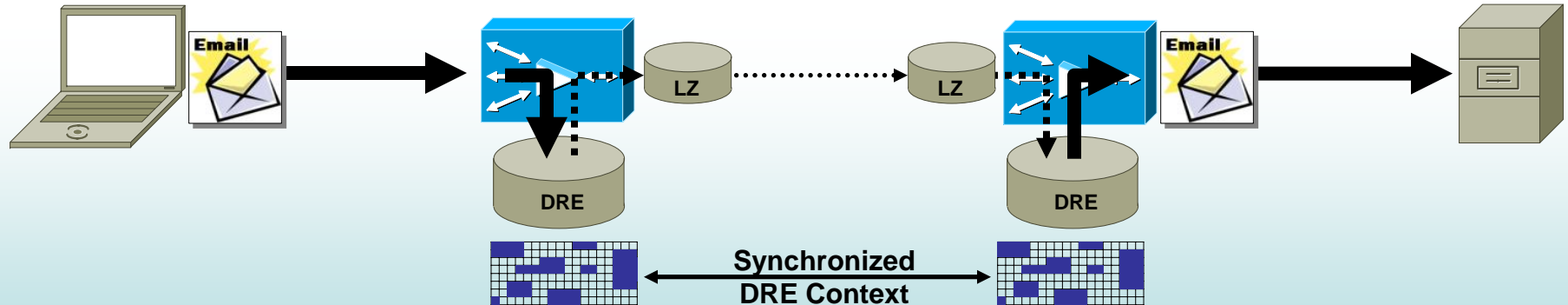


Round Trip Time (RTT)

INTEROP

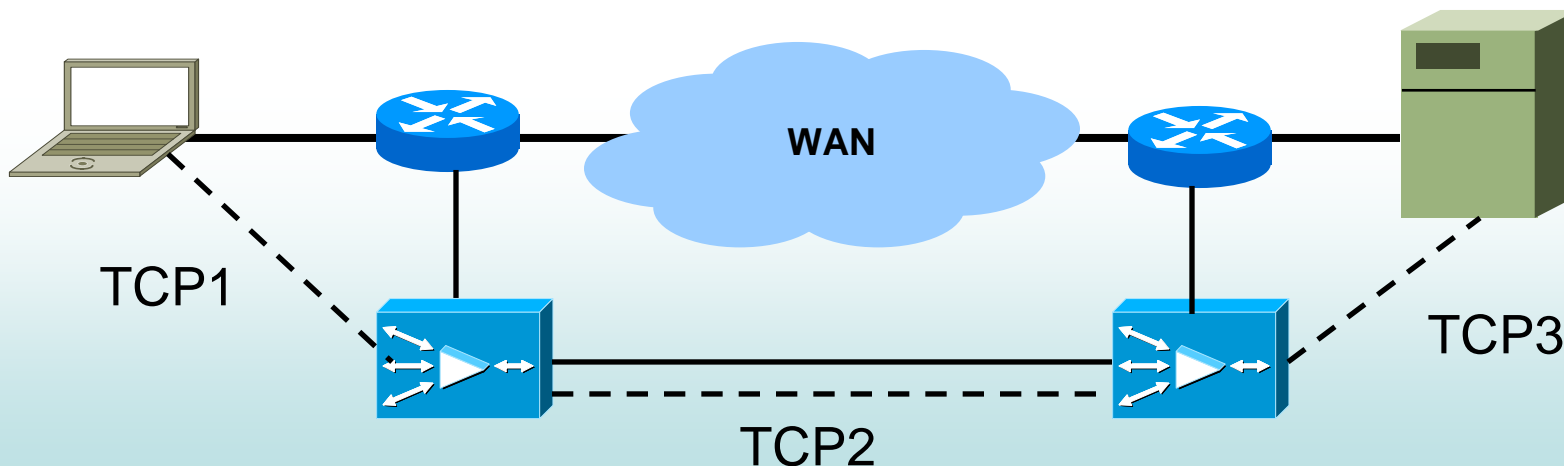
# L4 Optimization: Data Redundancy Elimination

- Divide stream to small chunks
- Store in local DB chunk signature and destination peer
- If first time, store and send to peer
- If seen before, send signature only



# L4: TCP Flow Optimization

- Fill the pipe: Improved TCP window management
  - pipelining (local acknowledgement)
  - Window scaling
- Improved congestion control:
  - Smaller drop, faster climb, fairness
  - Less sensitive to packet loss
- Prioritize processing of latency-sensitive apps
- Flexible policy configuration
- Pacing based on DRE



# Example: CIFS Latency

- Open a 2MB Word File: ~1000 sequential messages
- L4 optimization does not address app latency

The screenshot shows the Wireshark interface with a capture file named '1m\_open.cap'. The main packet list shows a series of sequential packets from 192.168.1.64 to 192.168.1.230. The 'Ethereal: Summary' pane provides the following statistics:

Traffic	Captured	Displayed
Between first and last packet	34.513 sec	34.369 sec
Packets	3342	2678
Avg. packets/sec	96.833	77.920
Avg. packet size	447.823 bytes	378.042 bytes
Bytes	1496624	1012396
Avg. bytes/sec	43363.853	29456.974
Avg. MBit/sec	0.347	0.236

A callout box highlights the displayed packet count with the calculation:  $2678/2 = 1339$  CIFS messages between client and server.

## L7: CIFS Optimizations

- Understand protocol to perform:
  - Safe object and meta-data caching
  - Read ahead
  - Write behind (flush on close)
  - Local handling of operations
  - Scalable: 10x offload of core devices
- Bottom line: > 90% WAN Latency reduction

# Optimized CIFS: only small fraction with WAN delay

The image shows a Wireshark network traffic analysis. The main window displays a list of captured packets. A filter is applied: `smb.time > 0.06`. The packet list shows various SMB transactions between 192.168.1.230 and 192.168.1.1. A summary window titled "Ethereal: Summary" is open, showing details for the selected packet (No. 25). The summary includes file information (D:\Documents and Settings\ibenshou\Desktop\1m\_open.cap), time (2003-10-27 00:45:54 to 00:46:28), and capture details. A red callout box points to the summary window, stating "Only 42 messages with WAN delay". The summary window also shows a table of traffic statistics:

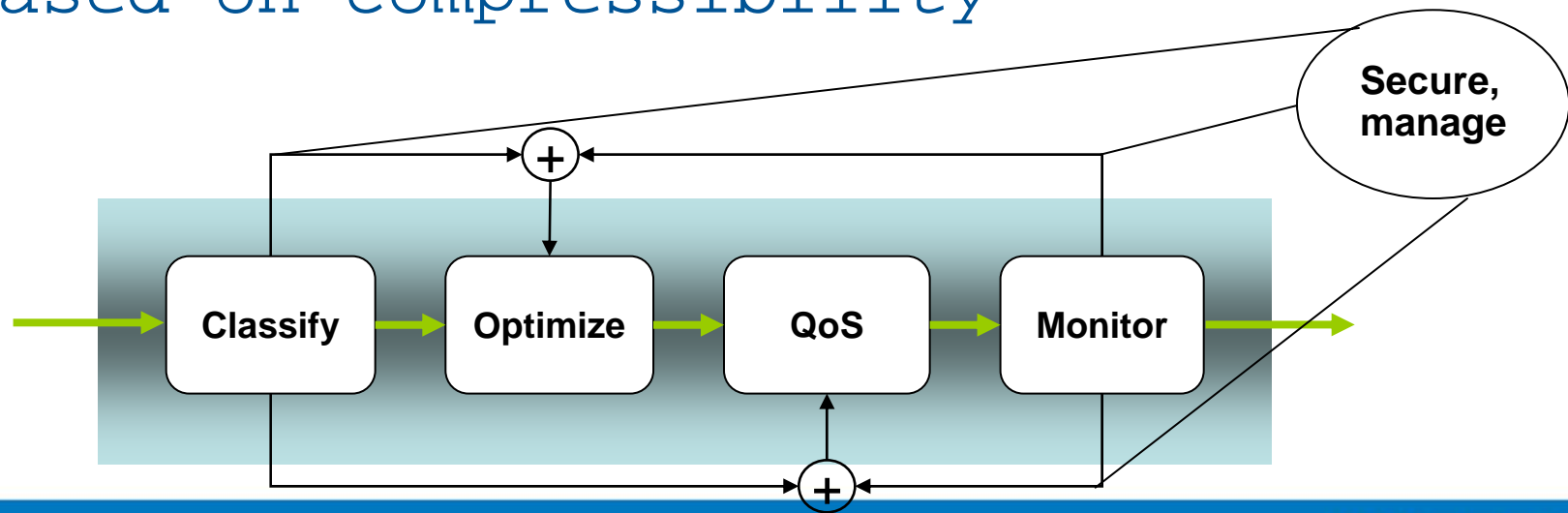
Traffic	Captured	Displayed
Between first and last packet	34.513 sec	21.354 sec
Packets	3342	42
Avg. packets/sec	96.833	3.606
Avg. packet size	447.823 bytes	170.857 bytes
Bytes	1496624	13156
Avg. bytes/sec	43363.853	616.083
Avg. MBit/sec	0.347	0.005

# Example: L7/8 HTTP Optimizations

- Pre-fetch embedded, related URLs
- Reuse Connections
- Cache objects based on app-knowledge

# Towards Adaptive Network Management

- Behavioral Classification: discover application characteristics (chattiness, compressibility)
- Prioritize/optimize based on application type and importance
- Dynamically allocate bandwidth per app based on compressibility



# Summary

- Most enterprise apps are network-based
- The network is the only common platform for all apps
- App-intelligence is essential to
  - Monitor, troubleshoot, operate
  - Optimize, Accelerate
  - Control, Secure
  - Prioritize based on app importance
  - MATCH APP and CUSTOMER NEEDS with available network resources