
WAN Optimization – Changes to more than just a box
Edwin Groothuis
AusNOG04 – September 2010
riverbed
Think fast™

Order of appearance

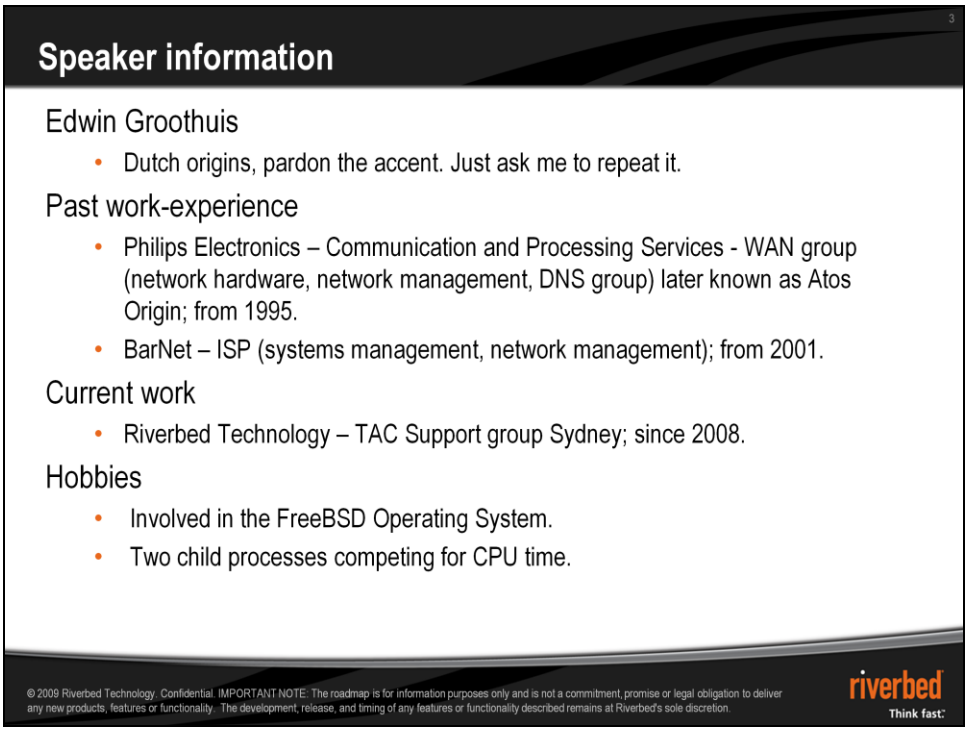
- Speaker introduction
- Background on WAN optimization
- Impacts on the network and monitoring
- Integration with software and services
- Failures and troubleshooting
- Know your networks!
- Questions and answers*

* Or when they come up.

© 2009 Riverbed Technology. Confidential. IMPORTANT NOTE: The roadmap is for information purposes only and is not a commitment, promise or legal obligation to deliver any new products, features or functionality. The development, release, and timing of any features or functionality described remains at Riverbed's sole discretion.

riverbed
Think fast.

-

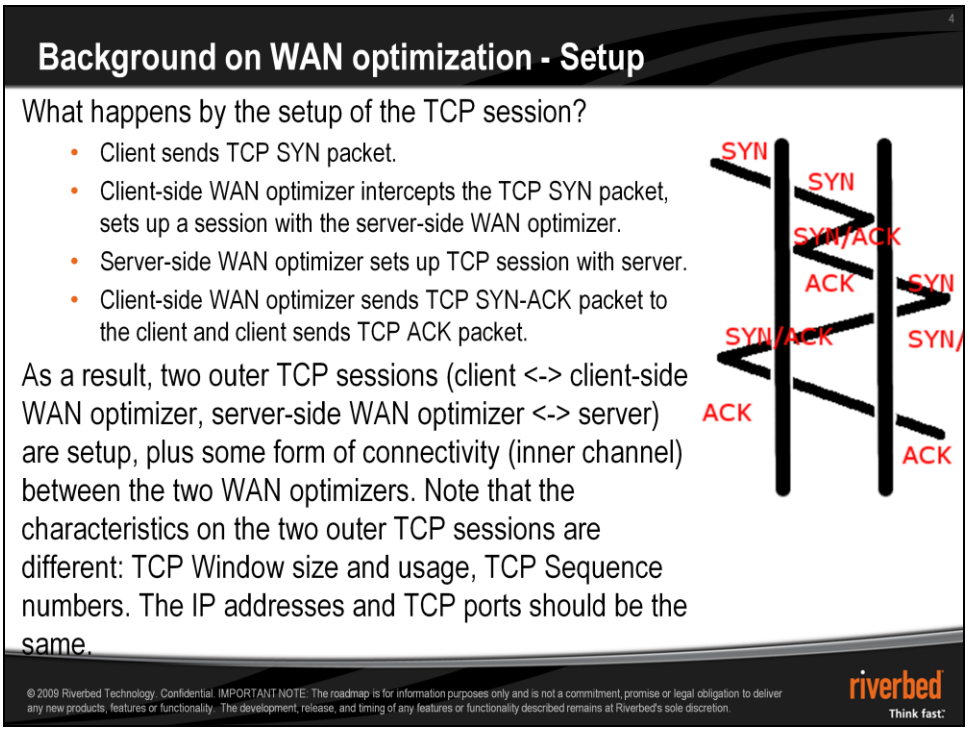


Websites:

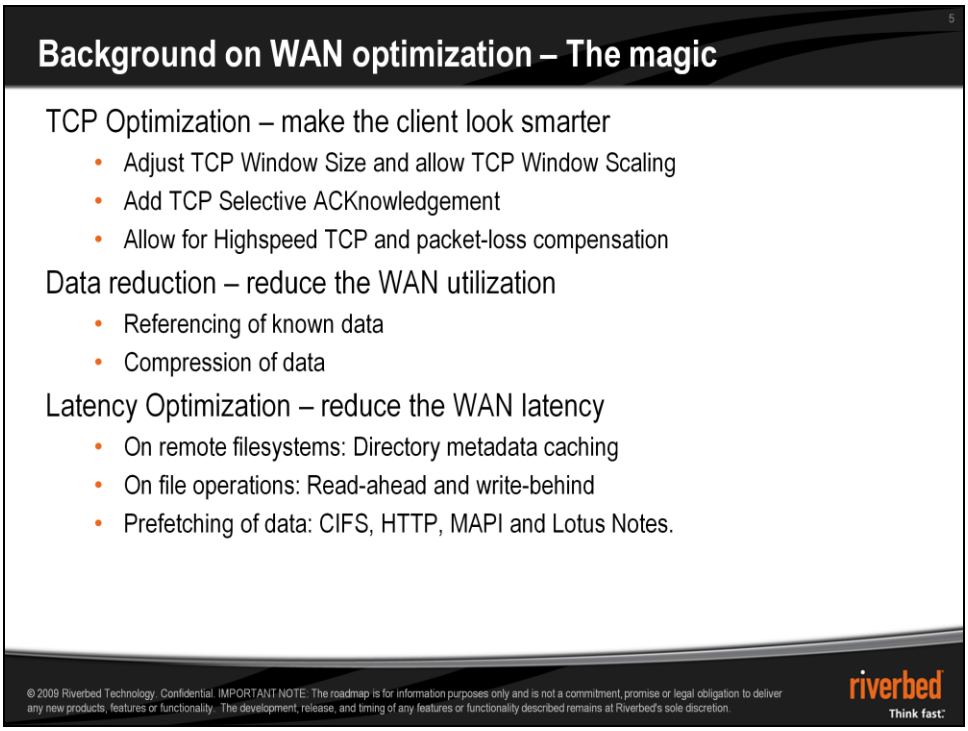
- Personal website: http://www.mavetju.org/
- Weblog: http://www.mavetju.org/weblog/
- Philips Electronics: http://www.philips.com/
- Atos Origin: http://www.atos-origin.com/
- BarNet: http://www.barnet.com.au/
- Riverbed Technology: http://www.riverbed.com/
- The FreeBSD Project: http://www.FreeBSD.org/

Contact me:

- Personal: edwin@mavetju.org
- Work: edwin.groothuis@riverbed.com



- Interception can happen inline (between LAN switch and WAN router) or out-of-path (via WCCP, PBR or proprietary protocols). Failure, asymetric routing, multiple WAN optimizers in the path and high-availability are ignored in this presentation.
- The session between the WAN optimizers can be auto-discovered or configured on the client-side WAN optimizer. Auto-discovery might confuse IPS/IDS/firewalls in the path, configuration is a lot of work.
- The transport layer of the session between the WAN optimizers is different by vendors: TCP sessions, tunnels.
- Once the TCP session is setup and no latency optimization is done, the data-in should be the data-out, although the shape of the delivery might be different: different packet sizes for example.
- For the client and the server, the WAN optimization should be transparent. IP addresses and TCP port numbers on both sides are under normal circumstances the same.



TCP Optimization:

- TCP Window size adjustment and TCP Window Scaling allow the server to send much more data to the “client” (represented by the Server-side WAN Optimizer) before the TCP Window on the server-side if full and the server will stop sending data.
- With TCP Selective Acknowledgement the receiver can request missing packets from the sender before the standard TCP timeout kicks in a the sender resends everything from the last acknowledged packet.
- Highspeed TCP for long fat pipes, with adjustment of buffer-sizes on the WAN Optimizer and the WAN router.
- Packet-loss compensation will not follow the standard TCP Window reduction in case of packet-loss but keeps the TCP Window the same size.

Data reduction:

- Referencing of known data: The first time a new piece of data (string of octets, varying in size) is seen, the data is tagged with a reference. The next time the same piece of data is seen, it only the reference is send. This is protocol independent, so if a file is first retrieved (and thus learned) via HTTP and then retrieved via FTP, the FTP transfer should only be

references.

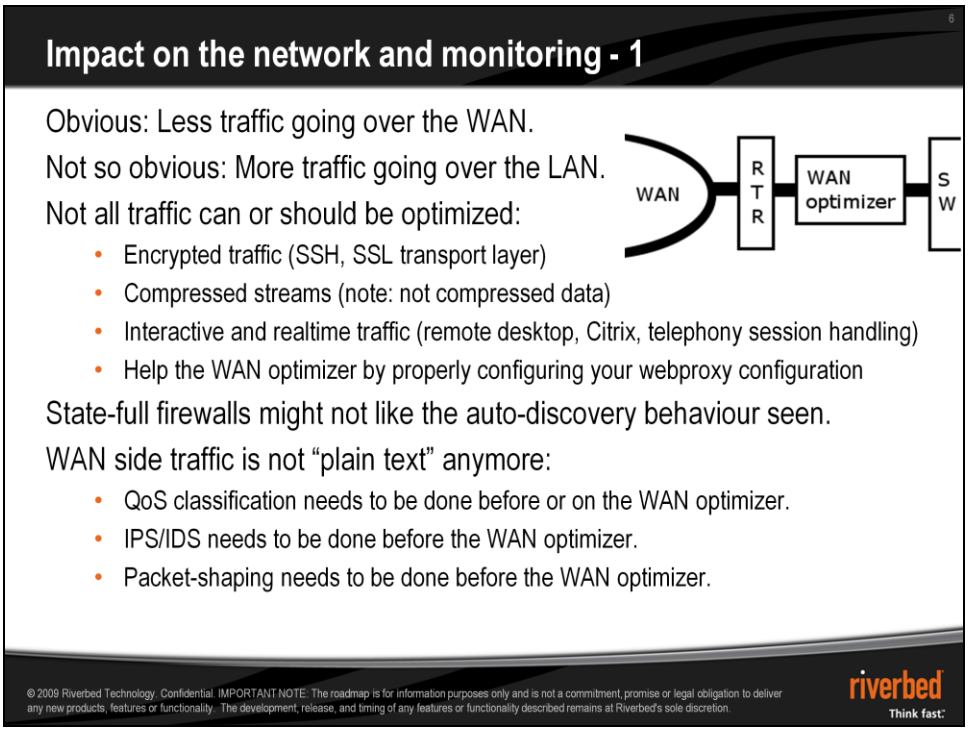
- Compression of new data during the first retrieval, reduces also.

Latency Optimization:

- WAN Optimizer needs to know the protocol and needs to know the behaviour of the clients to be able to predict the next step.
- Behaviour can be different between software releases (Exchange 2000, 2003, 2008), operating systems (Windows clients vs Samba clients) and protocols (SMBv1 vs SMBv2).
- Directory metadata caching:
    - Unoptimized: Client issues a find-first, waits for the answer, issues a find-next, waits for the answer, etc.
    - Optimized: Client issues a find-first, client-side WAN optimizer will ask the server-side WAN optimizer for the directory-list, client-side WAN optimizer will answer the find-first and the following find-nexts for the lifetime of the meta-data.
- Read-ahead and write-behind
    - Unoptimized: Client issues a blockread, waits for the answer, issues another blockread, waits for the answer.
    - Optimized: Client issues a blockread, client-side WAN optimizer will ask the server-side WAN optimizer for a blockread and the next three blocks, client-side WAN optimizer will answer the blockread, client will issue a blockread, client-side WAN optimizer will answer the blockread.
    - Unoptimized: Client issues a blockwrite, waits for an acknowledgement, issues another blockwrite, waits for an acknowledgement.
    - Optimized: Client issues a blockwrite, client-side WAN optimizer will forward it to the server and immediately send an acknowledgement to the client, client issues a blockwrite, client-side WAN optimizer will forward it to the server and immediately send an acknowledge the client.
    - Note that the closing of the file are not locally answered but send all the way to the server.
- Prefetching of data:
    - For CIFS, a central file-share is copied every night through the WAN optimizers to make sure that the data is known and can be

referenced the next day.

- For HTTP, the client-side WAN optimizer will parse the returned HTTP answer and prefetch the static data in advance.
- For MAPI and Lotus Notes: When the client disconnects from the server, the client-side WAN optimizer keeps the TCP session towards the server open and will download the new emails and their attachments received so that the next day when the client connects again the data is known and referenced and thus will be downloaded much faster.



Obvious and not so obvious:

- Not only does WAN optimization reduce the amount of traffic going over the WAN, it might also increase the amount of traffic going over the LAN. If the interface to the LAN router is configured for 100 Mbps and the WAN pipe is 10 Mbps and the WAN optimization is 10 fold, then the LAN interface is suddenly the bottleneck instead of the WAN pipe. Check the reports on the WAN optimizer to see a more granular WAN utilization report than your network management system.

Which traffic?

- Encrypted traffic is by design unique, therefore not repeating and thus will have a very low, it any at all, reduction rate.
- SSL traffic might be optimized if the WAN optimizer supports it. More later.
- Compressed streams are most likely also unique. Disable it in the application if possible.
- Interactive traffic is time-sensitive and shouldn't be delayed by the network.
- Webproxies often do HTTP and HTTPS optimization on the same port, as a result the overall reduction is not optimal. If you split the ports for the HTTP and for the HTTPS traffic and do not optimize the encrypted HTTPS traffic,

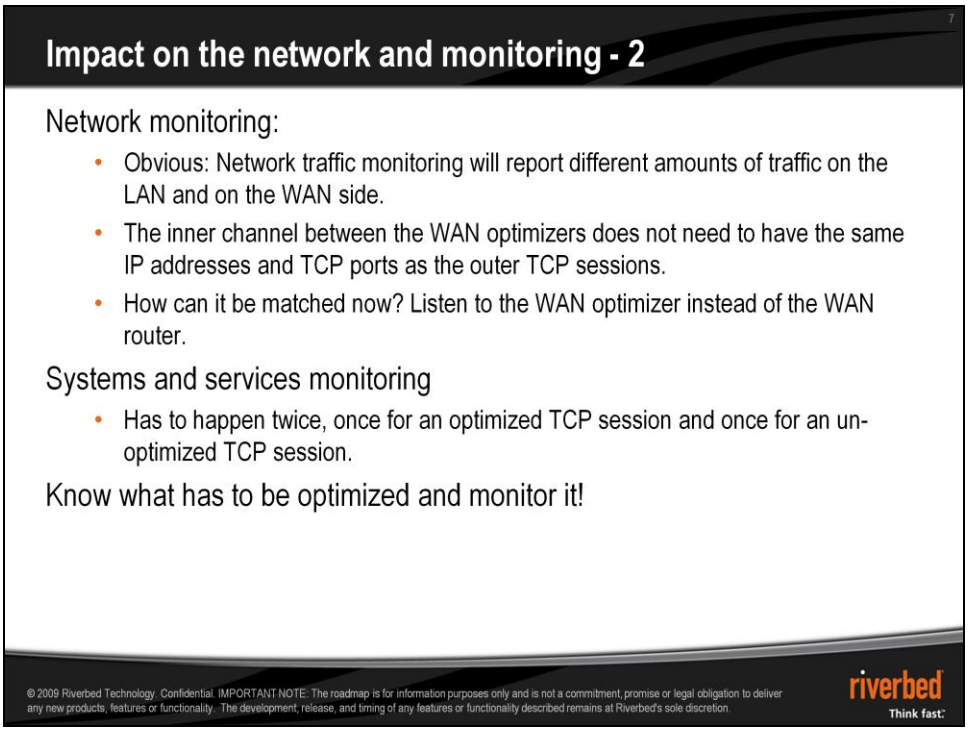
then the HTTP traffic will have a better reduction. Note that certain applications do use the HTTPS proxy for connecting to the outside world although the traffic itself isn't encrypted (Entourage on MacOSX for example). Smart proxy.pac files will help the WAN optimizer to do its stuff.

Firewalls and auto-discovery

- Some firewalls think that auto-discovery produces half-open TCP sessions.
- Some firewalls think that auto-discovery produces double TCP sessions with different sequence numbers.
- Some firewalls don't like the flags or tags used for auto-discovery and zero them out.

Plain text:

- If DSCP QoS Marking is used on the LAN switch, the WAN optimizer will (should) reflect it.
- QoS Classification on the WAN side won't always work because of the inner channel with different IP addresses and TCP ports.
- Packet-shaping on the WAN side will count the size of the reduced packets, not of the real data. Dropping a little packet on the WAN side might result in the loss of multiple packets on the LAN side.
- Packet-shaping on the LAN of the server-side might cause slowness due to the extra traffic caused by latency optimization.
- IPS/IDS – the plain-text part of the TCP session between client and server is not between the WAN optimizers.



Network Monitoring

- Ping, SNMP etc should all be the same as before.
- Old-style networking: The amount of data going into a device in the network is more or less, apart from management of the device, the amount of data coming out of the device.
- New-style networking: It is not one-on-one anymore. And with latency optimization can even cause that the amount of data send by the server is much more than the amount of data received by the client.

Flow exporting:

- The netflow data exported from the WAN optimizer on the WAN side might not match the IP addresses and TCP ports in the netflow data exported from the WAN routers. The two streams should match on size, not on end-points.

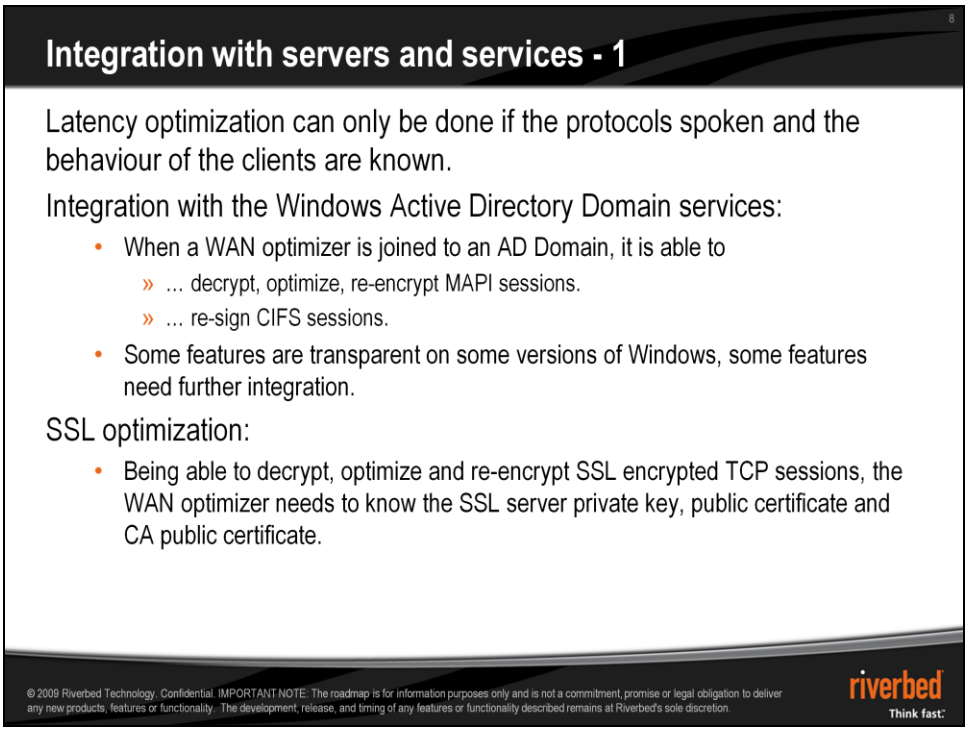
Systems and services monitoring:

- When a normal TCP session is setup, it might be optimized. So you are monitoring the reachability of the service through an optimized TCP session, which might be different from the reachability of an unoptimized TCP session. Monitoring from two different IP addresses, one optimized and one

unoptimized, might give more insight in case of problems.

Know what has to be optimized:

- If you have data-center to data-center replication for SAN, NAS, Database, File Systems, etc, then you want to make sure that the TCP sessions which you expect to find optimized are really being optimized after for example a restart of the WAN optimizers.

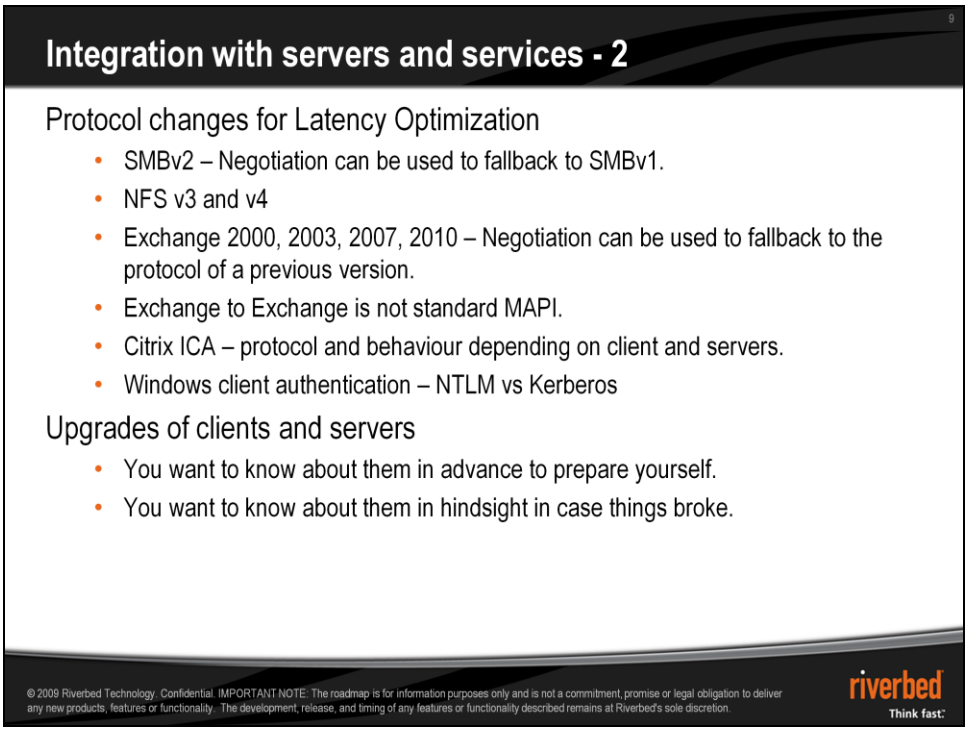


Windows Domain:

- The joining of the WAN optimizer to the domain should only have to happen once, but has the same restrictions as any other device to be joined to a domain:
    - Time on the WAN optimizer needs to be more or less the same as on the Windows machine. Use NTP against the DC controllers in case you don't have access to a public NTP server.
    - IP address of the management port of the WAN optimizer needs to be resolvable into a hostname.
    - The WAN optimizer needs to be able to resolve DNS queries to find out what the DC controllers are for that AD domain.
    - And obvious, the account used must have joining-rights.
- The AD domain of the clients should be the same AD domain(s) the WAN optimizer is joined to, or ones trusted by it.
- Kerberos authentication on the DCs is more work than NTLM authentication, it needs an AD delegate user defined on the WAN optimizer.
- New releases of the Windows OS might interfere with the integration, be careful with integrating them into a production AD domain (Windows 2008R2, Windows 7).

SSL optimization:

- Yes, it’s a MITM approach.
- If the WAN optimizers use some kind of auto-discovery, will you trust the client-side WAN optimizer by default?
- Administrative problem: Expiring certificates after a year or so. Instead of only having to be changed on the SSL server, it now needs to the updated on the WAN optimizers too.
- SSL optimization isn’t only limited to HTTPS, also LDAP over SSL, POP3 over SSL, IMAP over SSL, SMTP over SSL.
- Note that SMTP with TLS and HTTPS going via a webproxy is not fully SSL, it starts with a plain-text session and then switches to an encrypted session. Does the WAN optimizer support it?

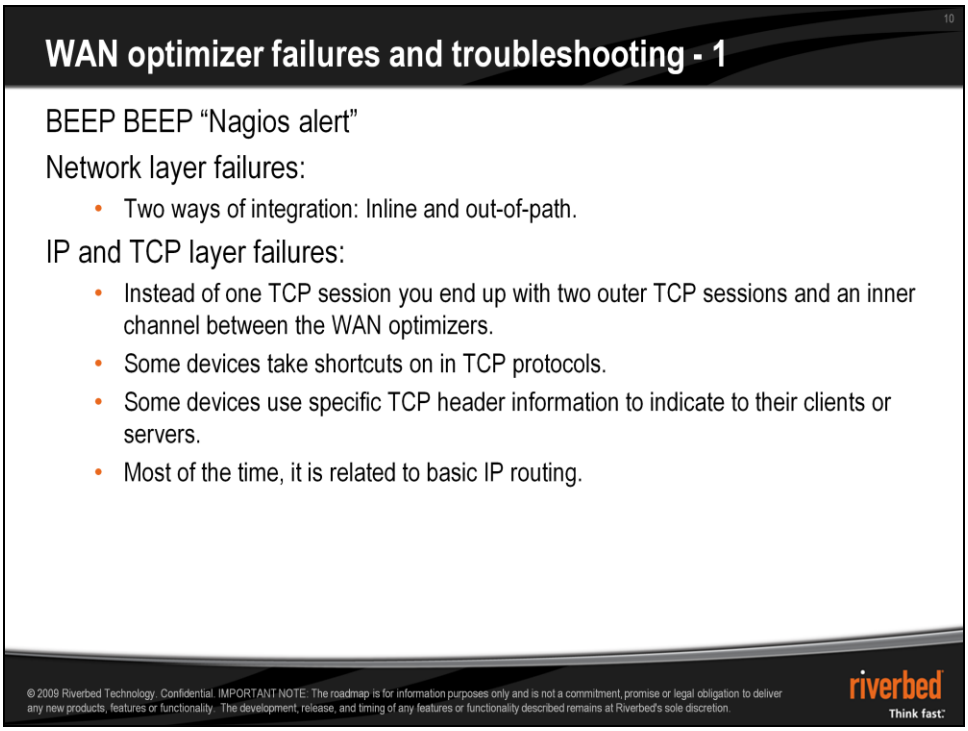


Protocol changes

- Although the SMB version can be negotiated, sometimes it is required to be a certain version.
- NFS has three different versions, v2 is UDP only, v4 is using Kerberos and not supported by all WAN optimizers.
- The MAPI protocol between Outlook clients and an Exchange server and between Exchange servers is different. So is towards Blackberry servers.
- Outlook 2003 had a different network behaviour than Outlook 2007 with regarding to the amount of TCP sessions setup to the Exchange servers.
- The behaviour of Citrix clients, software and hardware ones, is very different between versions. Make sure your WAN optimizer supports the one you want to use.
- Windows 7 clients to Windows 2008R2 servers.

Upgrades

- You really want to know about them, in case you get a complaint from users.



Network Layer failure:

- Inline integration failure can be fail-to-wire or fail-to-block.
- Inline integration failure and recovery require a renegotiation of the ethernet links.
- Inline integration can support Link State Propagation so the WAN router knows when the LAN switch is down and vice-versa.
- Out-of-path integration can be done via PBR, WCCP or vendor specific protocols.
- Out-of-path integration failure takes time to know that the node redirected to is down.
- Out-of-path integration failure can also affect non-optimized traffic if the redirection rules are too broad.

IP layer failure:

- The end-points of the data-path, the WAN optimizers, should be able to talk to each other.
- This is specially fun when
    - the client-side WAN optimizer is integrated on the client machine and

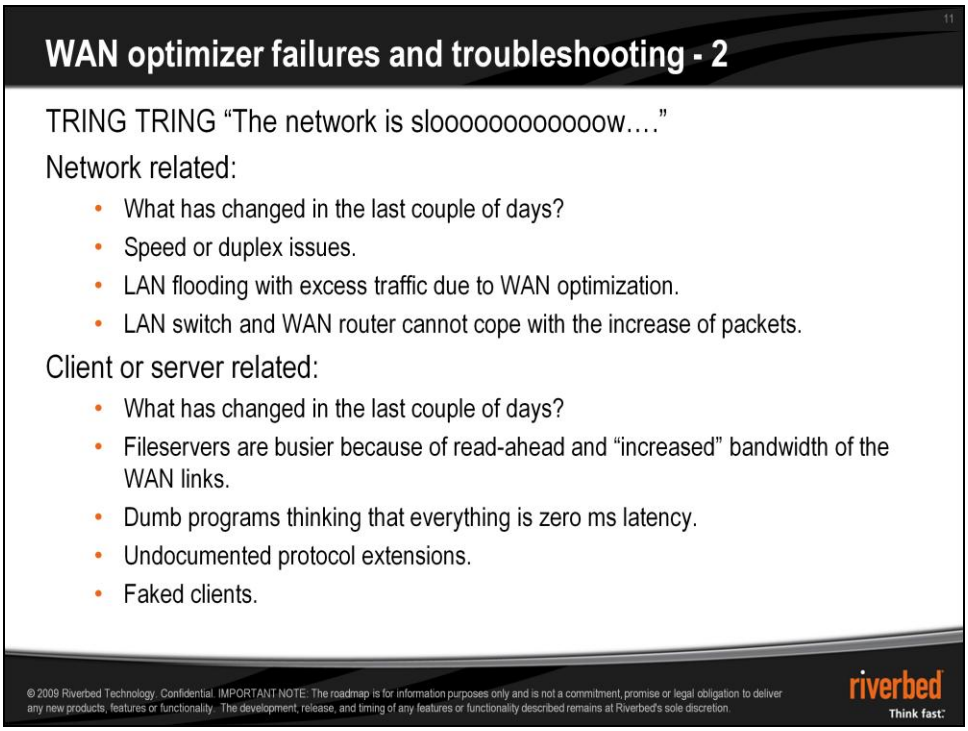
the client machine is behind a NAT device and auto-discovery is used.

    - one or more NAT gateways (client-side NAT, server-side NAT) are involved.

- Don’t use auto-discovery when going through NAT gateways. More preferably: Have a single and unique IP space in your network.
- Basic troubleshooting: Follow the TCP SYN and you can find out where it goes wrong.

TCP layer failures

- Some devices have their own quirks with their TCP stack:
    - Printers spoolers who close the TCP session when they have finished sending the data, but finishing sending doesn’t mean that the printer has received everything yet.
    - Brocade devices use and require flags in the TCP header to indicate certain states. WAN optimizers can lose this information.



What has changed in the last couple of days?

- Upgrades to clients/servers, changes in the network.

Network related:

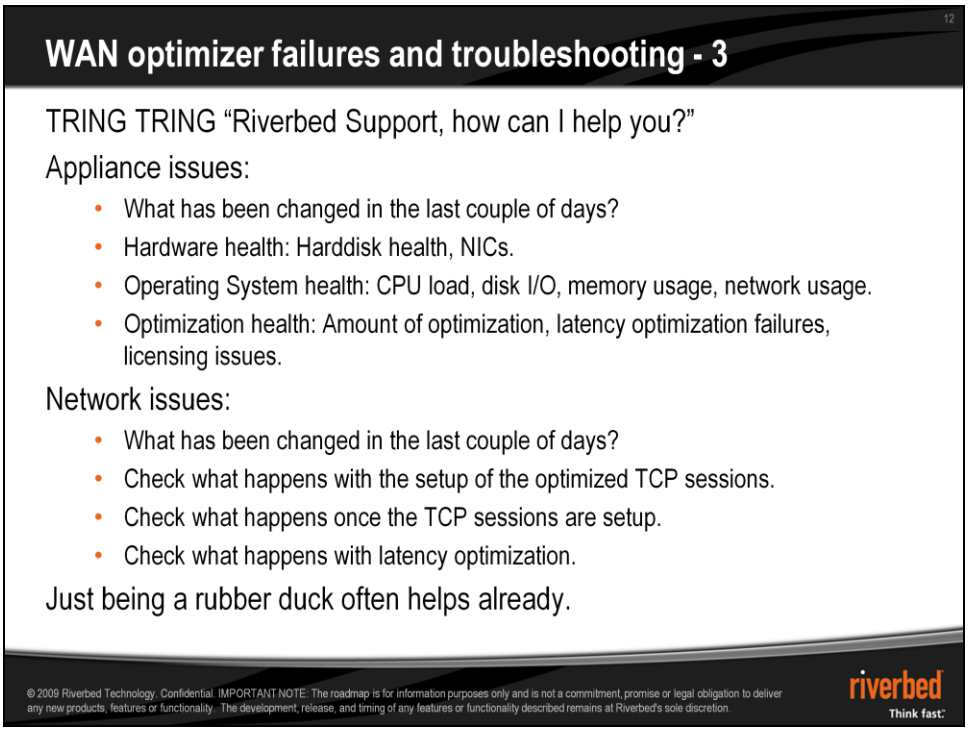
- Just check the speed and duplex issues, they keep popping up and are the easiest way out of a problem. To check the path between client and server, run iperf between two hosts (unoptimized) and the performance should be more or less the same during a 10 second window with 1 second display.
- LAN flooding: check the (one second granular) output on the WAN optimizer instead of the (1, 2, 3, 4, 5 minute average of the) network monitoring system.
- LAN flooding: realtime VoIP traffic will be affected by this first. If the WAN router only supports 100 Mbps, set everything to auto-negotiation so that the LAN side of the WAN optimizer is doing gigabit speeds.
- Not all router-hardware can do BGP and some kind of IGRP and firewalling and also have time to forward packets….

Client and server related:

- The servers will be busier, if not in general due to latency optimization, then

at least in peaks because of a “higher” bandwidth available to the clients.

- Dumb programs ignore latency issues and assume that everything is connection on the local LAN. For example a database client which performs the query and asks for each record sequentially instead of in one batch. On the LAN this doesn’t matter, because the latency is close to zero. On an unoptimized WAN, you only have the delay on the WAN. On an optimized WAN link, you have the naggle-delay of the WAN optimizer, the delay on the WAN, the naggle-delay of the WAN optimizer and the WAN delay. And because of the often small questions and answers the amount of optimization is very small. Consider this interactive traffic and let the developer fix the software!
- Undocumented protocol extensions by third party software. The WAN optimizer doesn’t know about them, therefore it will just pass them to the server. Another example is the McAfee server-side virus-scanner: The client-PC wants to have a file from the CIFS share, copies it and asks the server is the file is scanned there via a “Does this filename exist” where filename is a mix of the real filename and a shared secret). The client-side WAN optimizer knows the directory contents and will answer it locally with a “No, this file does not exist.”. The client-PC will then send the file to the virus scanner on the server and get the answer “Yes, it is clean” and starts it. As a result, a copy of the file from a remote CIFS-share will cause extra LAN traffic.
- CIFS shares on NetApp servers identify themselves as Windows 200x servers, but have a couple of options not the same as Windows 200x servers, Applock for example.



How does your network look like, and what has been changed?

- We would like to know everything about the network. Everything.
- We will find out about it anyway, so let us know in advance.

Hardware

- Harddisks can be dieing but haven’t been kicked out the RAID array yet. They still cause delays when reading from the RAIDed disks.
- NICs: frame errors etc. Yes it is already mentioned on the previous slide but people still don’t check it.

Operating system:

- CPU load: Is the load nicely spread over the CPUs, is the usage pattern different than normal?
- Disk I/O: No I/O wait please.
- Memory Usage: No swapping please.
- Network Usage: Is the usage pattern different than normal?

Optimization:

- Is there a lot of traffic which is optimized but have a bad optimization percentage? That could indicate encrypted or compressed streams.
- Does the latency optimization complain about anything? Are the protocols supported?
- Licensing: The WAN optimizer hardware is capable to do a certain amount of optimization, are we still fine?

Network issues:

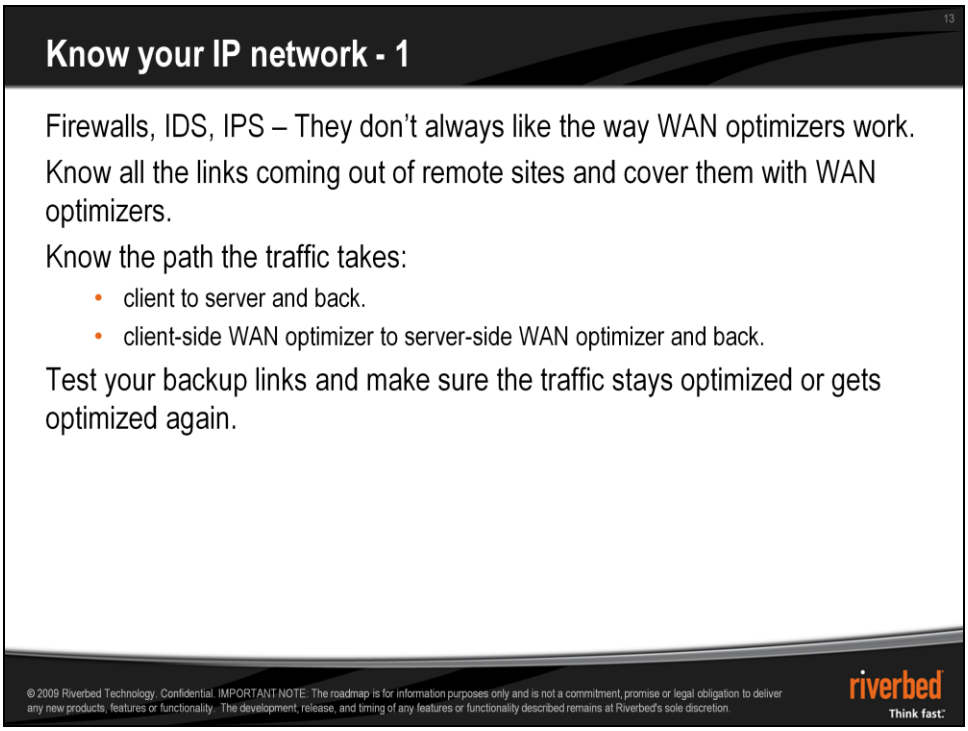
- Does the optimized TCP session get setup? (TCP header pruning ruining auto-discovery, asymetric paths not all covered by WAN optimizers, different network implemented than designed)
- Does it stay up? (TCP RSTs from stateful firewalls because of missing packets in the TCP setup or because of TCP sequence number differences, MTU issues etc)
- Have your network diagram ready and don’t have to extend it while troubleshooting is happening!

Rubber ducking: “Place a rubber duck on your monitor and describe your problems to it. There's something magical about stating your problems aloud that makes the solution more clear.”

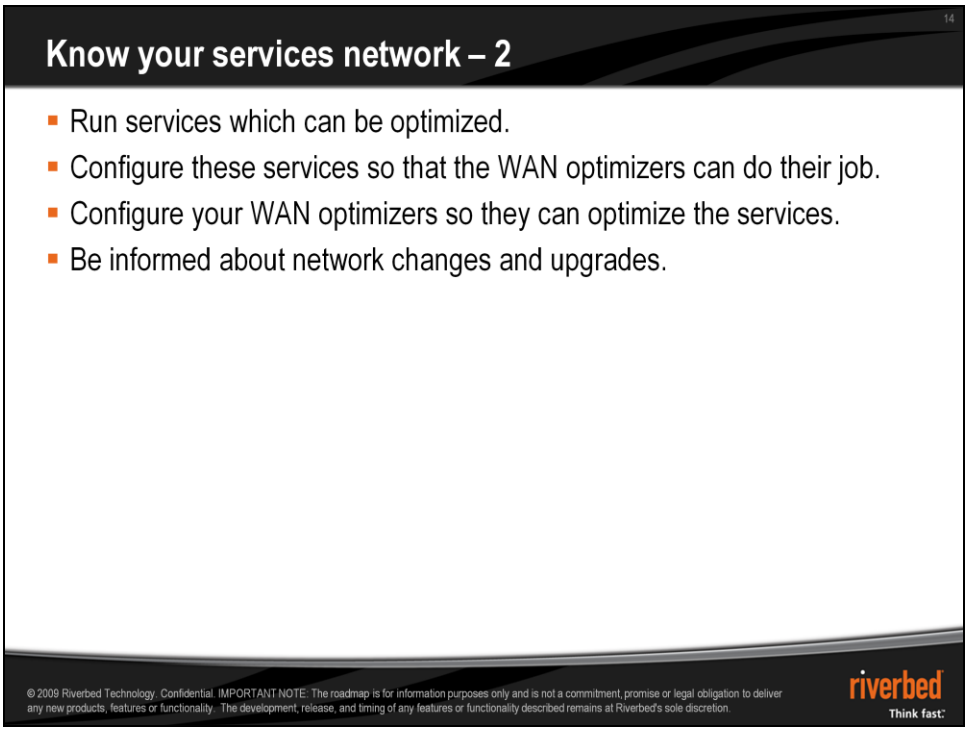
- Person 1: "My code doesn't work! I've got all these objects in an ordered list, and I assume they can all... oh. Ah, yeah. I see the problem. Thanks for your help."
- Person 2: "..."

Sometimes the answer is simple…

… the amount of traffic has grown and the WAN optimizer has become underspec’d.



- Place the security devices at the right side of the WAN optimizer.
- WAN optimizers need to see the full TCP session, use their technologies to catch it! (software or hardware)
- Make sure that the traffic between the client-side and the server-side WAN optimizers takes the path you expect it to take as if it was from the client to the server.
- Testing your backup links, specially if the WAN router is owned or WAN routing is influenced by a third party who can change their configurations so that it looks like standard TCP works fine, but optimized TCP suddenly doesn't work anymore.
- If you have multiple interfaces in the WAN optimizer, then you can have multiple paths for the inner channels between the WAN optimizers.
- If the backup link doesn't have the same operational cost as the normal link (i.e. charged for traffic, capped on a certain speed or amount of traffic), make sure that the optimized TCP sessions fall back to the main link too!

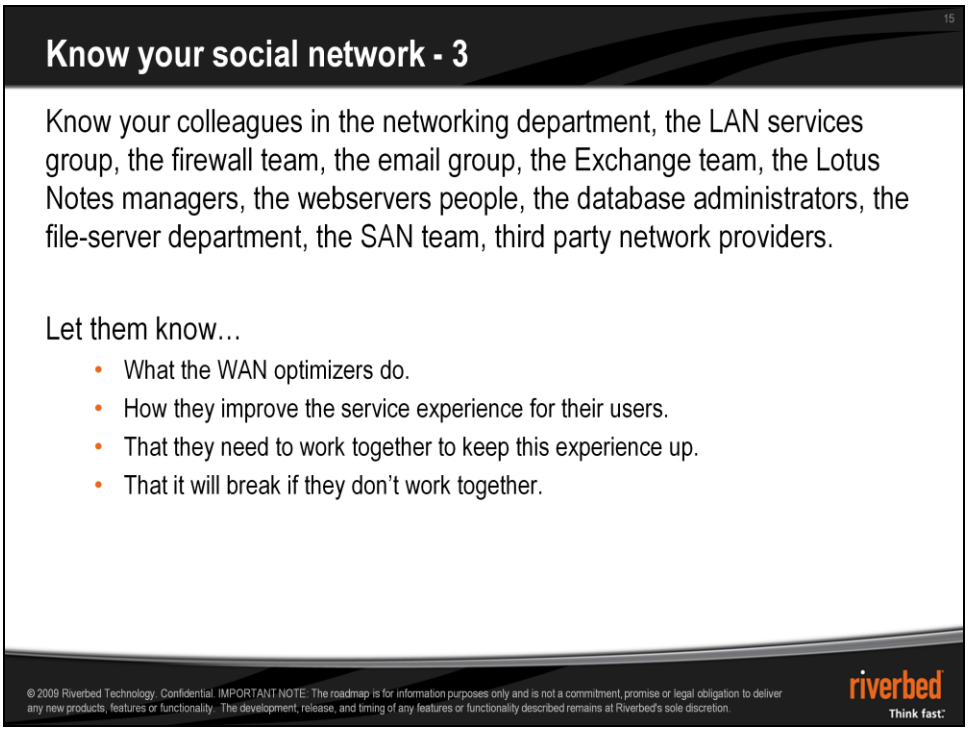


Network changes

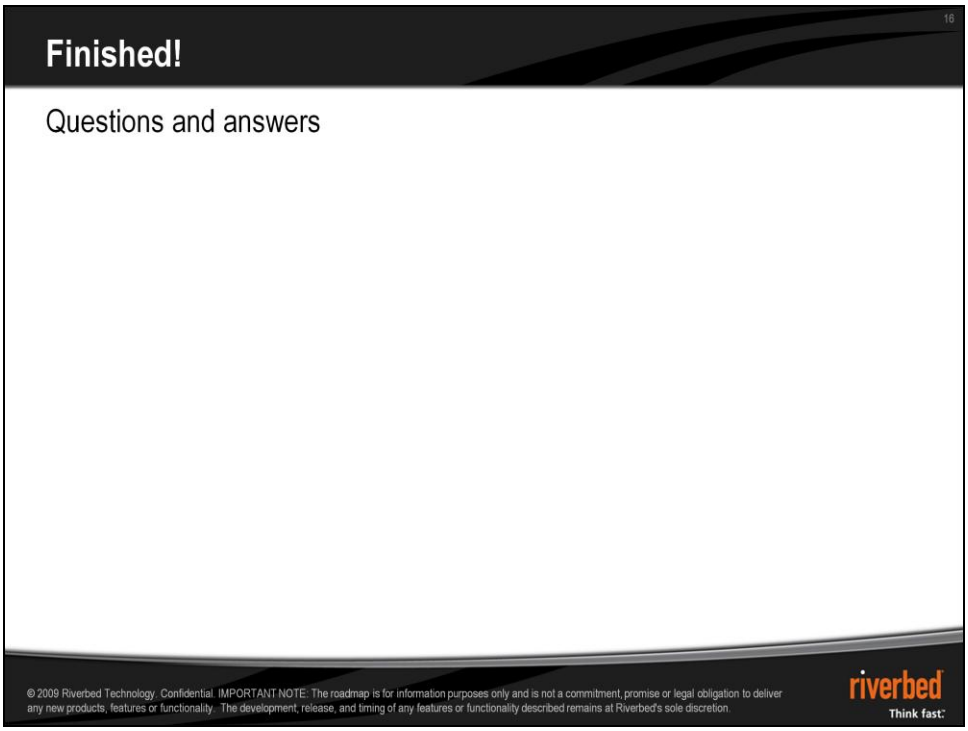
- Additional servers.
- IP address change.

Upgrades

- Software upgrade version.
- Security changes.



When you tell them what the WAN optimizers do, don't only tell them about bandwidth reduction but also about Latency Optimization. That is the part they are mostly interested in, because that is the part which will complement / enhance / interfere with their services.



You can always mail me on my work email address with (technical/sales/info/demos) questions regarding the Riverbed WAN optimizers.