Cisco Public.

US WAN Rearchitecture

Cisco on Cisco Technology Seminar

Rich Gore: Hello and welcome to this Cisco on Cisco Technology Seminar. Today, we're gonna be talking about how Cisco IT rearchitected the North America WAN to support peer-to-peer applications. I'm Rich Gore and with me is Don Layton.

Don Layton: And Rich and I were co-project managers on this project. Rich was responsible for the development of the architecture and I was responsible for the deployment and implementation of the solution. Today, we're gonna talk about the benefits we derived from the new peer-to-peer network we put in place.

Rich Gore: Sounds good.

Agenda

Don Layton: The agenda for today: We'll first cover some drivers around the rearchitecture itself. We'll talk about the WAN design. We'll talk about the hub site-specific design and we'll talk about the migration steps that got us there and, finally, we'll talk about the results.

Americas WAN Rearchitecture Drivers

Rich Gore: So now let's talk about the America's WAN rearchitecture, the business drivers behind this, because rearchitecting the WAN is a huge, huge job and you don't do it lightly, so we had a client/server network that was working very well and why did we want to move to a peer-to-peer architecture? What does that mean anyway?

Don Layton: Well, the peer-to-peer architecture suggests that you're having more regional connections to each other, so closer together and various regions are talking to each other rather than talking back to the main site or the home site. For us, the biggest push was IP telephony.

We were just introducing the Call Managers at that time and putting them out in the field but we were also looking at video conferencing. We were looking at collaboration tools and other such peer-to-peer-type things that were really changing the face of the network and changing the face of the type of traffic we were putting over the network.

Rich Gore: Okay, yeah. Yeah, also throughout this whole process, we were facing rapid growth, rapid utilization increases and that was true whether we were in client/server mode to peer-to-peer mode and so we were looking at ways of growing the network quickly.

Don Layton: Yeah. We were also looking at pushing some of our applications services out to the regional networks as well, regional sites.

Rich Gore: Okay, yeah. Like the Content Distribution network.

Don Layton: Yes and, again, rapid growth was a big issue here.

Cisco North America WAN: Client/Server

Rich Gore: So this is a map of the North American WAN in its client/server mode in the year 2001 and it looks confusing, but actually it was very, very flexible, scalable, a very useful network that really shows that all of our servers, all of the data center stuff was in San Jose.

All of our clients were at those 120 sites out there throughout North America and they needed the shortest possible path from wherever they were back to the servers in San Jose, back to where all the email stuff and all the files were.

So what we did is we built Frame Relay, for the most part, and some ATM when Frame Relay would sort of max out at T1 speeds and a little bit of leased line but mostly Frame Relay and which is pretty much a straight shot back from each one of those offices back to San Jose and this is a very good client/server network.

So when you're looking at different business's networks, if you look at this kind of network, this sort of client/server everything hubbing into one point network, you'll understand the sorts of applications that are going on it. So let's take a look at what it became.

Cisco North America WAN: Peer-to-Peer

Don Layton: So the new network was based on a service provider style or mentality where we put in a small service provider network with a large backbone with regional hub sites and we would move or hub each of the individual 120 sites into its closes hub site that was out there on the network.

Rich Gore: Yeah, it does really look like a service provider network with a relatively large backbone with everything coming into regional hubs, so like little stars as opposed to that one major star we had back there.

WAN Rearchitecture Design Summary

Rich Gore: So as a quick summary, really what we ended up with was a U.S. WAN where we had regional hubs, hubs in Raleigh and San Jose and Chelmsford and Kanata, which is just outside of –

Don Layton: Ontario.

Rich Gore: Ontario. Thank you. I keep forgetting, and Chicago and Dallas and Denver and I think that covers it – and Atlanta. And we were also – I remember considering some other sites as well.

Don Layton: We had talked about either Seattle or Miami as possibilities, even Irvine

as well.

Rich Gore: That's right, Southern California and we'll talk about that in a little bit more.

Don Layton: The main pieces here, though, were the OC12 backbone running east and west, both a north version and a south version and then using OC3s to connect those backbones inter-regionally in the north-south fashion and then using pairs of T1s or multiple pairs of T1s or pairs of DS3s, depending on the size of the location, to connect the individual sites back into the hub sites.

Rich Gore: And that was the physical infrastructure. On top of that, we had all those peer-to-peer applications which required a really much more rigorous QoS policy to be applied and we'll talk about that, too, in a little bit.

And with all this added back – or bandwidth coming into the network and utilization increasing, we, of course, had to increase the WAN or the WAN bandwidth pretty much everywhere.

Comparing WAN Architecture Advantages

Rich Gore: So in comparing the two architectures, the before and the after, the 2001 and the 2003, I have to say I really do like the client/server network, the one that we had built back in the 1990s and then eventually moved away from in the year 2001. It was optimized for client/server traffic and it did that very well.

Don Layton: It was a good network. It serviced us for many years and many of our customers out there still have that type of network but, again, as the peer-to-peer-type applications come down the line, that type of network shows its problems and we could not only handle the client/servers with the new network; we could also handle the peer-to-peer applications with the new network.

Rich Gore: That's the nice thing. You're not really giving up anything by moving to the peer-to-peer network. You're just increasing your capabilities. The nice thing about the client/server network is that it's very easy to manage and very easy to do capacity planning. Any Frame Relay link needs to be bumped up, you bump it up. It's pretty easy to do.

Don Layton: So very true. It is a little bit easier to do capacity planning and to do the general management of the network with the old Frame Relay style, but it wasn't that big of a difference for our folks when we moved into that. We did – the main thing was to support those regional data centers that we were trying to pull out there and to support the client-to-client communication that was growing out of the peer-to-peer-type applications.

Rich Gore: Yeah. I remember one of the surprises was when we looked at the two architectures we realized that the client/server, that star network, was very cost effective for smaller networks when the links of the lines were relatively small, when you had fewer sites to connect to, but that turned out not to be the case as the network grew and grew.

Don Layton: Yeah, as we did the cost analysis, which I think you did most of that, we found there was a breakeven point somewhere in the 90- to 100-site range within the United States where the regional hub network actually was cheaper than or almost as cheap as the star network, the Frame Relay network.

Rich Gore: Yeah, for much greater bandwidth.

Don Layton: Yeah.

Rich Gore: So client/server network excellent for – very good performance for what we call near real time applications, things like email, things where if it is delayed an extra quarter second you don't really notice it. You don't really care.

Don Layton: Your ERP applications and such things like that. Whereas with our newer applications that are coming online such as IP telephony or frame – or video conferencing, we really had to have real time performance that was available on the network and this new network provided that type of capability.

Rich Gore: Yeah, additional latency on something like voice or video is just very painful to experience.

Don Layton: Yes and, of course, both of them are very flexible and very scalable in their own ways.

WAN Design

Rich Gore: Yeah. So we'll talk about the WAN design. We'll drill down from the very top down to the actual hub sites and Don has pretty much – I remember working on this part, so let me take these two slides.

WAN Route: Hub Locations on Underlying Fiber Path

Rich Gore: When we first started trying to figure out how to design this new peer-to-peer network, we said, "Okay, it needs to have a heavy backbone and it needs to be able to connect to regional hubs, but how can we build this network while still maintaining the lowest possible latency?"

And so and so we said, "Okay, let's look at the underlying fiber because that's where it's all gonna be riding on. It's gonna be riding on the service provider provided fiber."

We looked at a particular or really a generic sort of vendor's map of where the fiber is and we said, "Where does the fiber actually end up crossing? Where are the major cross points? Let's think about locating our hubs at those."

So we see right there a map of – a fiber map of the U.S. and where we ended up putting our regional hubs and you'll see that for almost all of them they really are running or riding on the real major intersection points of the fiber map.

Don Layton: So on this map, Rich, I notice that there's nothing in Seattle. There's nothing in Miami.

Rich Gore: Yes, we have a lot of people in Seattle, a lot of people in the Miami area and also a lot of people in Southern California and we ended up deciding not to put hubs in those locations and we'll see a little bit later sort of where the populations are and it'll kind of make sense.

But really we cast and recast this network several different ways, several different times, each time generating the costs, the leased line costs, for this network 'cause this is all based on leased line. It's all based on leased Sonnet, leased TDM links, and there's ways of doing it more expensively.

WAN Route: Hub Site Populations/Sites

Don Layton: So here you see the distribution of Cisco offices across North America. The first number you see there is the population and the second number is the number of offices associated with that population and this really drove us to how we would break up the network and attach a particular

remote site to the backbone itself.

Rich Gore: Now remember it was looking at this, especially that big mass of population in San Jose, which really means the mass of population connecting into San Jose and a lot of them were in Southern California.

A lot of them were in northern California, Oregon and Washington and we were thinking maybe we should have hub sites in Southern Cal; maybe we should have a hub site up in Seattle.

And then we just ran the numbers and it looked like you're adding a lot of extra backbone length especially up into Seattle area and it doesn't really significantly improve latency for the people in Seattle and Southern California, so we ended up not doing it; same with Florida, same with Miami.

Don Layton: Yes and then also we had – you notice we're in Raleigh-Durham RTP area, and for some of our customers, they may be more interested in the Washington, D.C. area for a hub site, but for us in particular, because of the high concentration of people in the Raleigh-Durham area, it made more sense to do it on our campus.

Rich Gore: Yes. Yeah, that was the one exception to keeping it on the fiber map for us because there just wasn't a big fiber nexus in Raleigh, but we ended up with an access line out of Washington, D.C., I think.

Don Layton: Yes.

Branch Office Site Connections

Rich Gore: So this was the branch office.

Don Layton: So the branch office sites originally were, obviously, attached via Frame Relay. A large majority of them were sub-T1 rate, even sub-256 rate. We were running into issues of constant bandwidth utilization issues on the Frame Relay and we needed to alleviate that problem.

So as part of the design, we were able to move from the Frame Relay to a TDM T1-type solution for each location where each site had a minimum of two T1s to it, diversely routed as much as possible, and we'll talk about that a little later.

Several sites beyond that, got multiple T1s, again diversely routed as much as possible but treated in pairs – two, four or six at a time – and our much larger sites would get either DS3 or dual DS3 or even OC3 capability as the population required.

Rich Gore: I'm gonna be spoiling a later slide, but it still is amazing to me that we went from 256 and 384 and 512K Frame Relay and went to dual T1 and we pretty much didn't spend any more money to do that. That still surprises me.

Don Layton: That was amazing.

Rich Gore: Yeah.

Don Layton: And then the final thing on there is the average length of our individual circuits to each small location; went from an 1,800 mile average down to 260 miles and that helped drive down that latency.

Rich Gore: There's – I think the secret in that cost – the lack of a cost increase was that all of those Frame Relay links were coming from – a lot of them were coming from the East Coast and those were long and Frame Relay prices reflect the underlying length of the circuit frequently.

Branch Office Backup Connections

Don Layton: So on – the backup solutions for a branch office with our Frame Relay network were based on BRI ISDN backup and we had some several – we had several issues around that, not least of which was price and cost, ongoing cost.

One of the things that came out of the analysis for this was that our testing of our BRI circuits every day for about four minutes per circuit averaged out to almost $5,000.00 a month alone in testing cost which were able to apply back into the cost or to alleviate the cost of the new network.

Rich Gore: Yeah, I remember that the little 256K Frame Relay circuits, we had, I think, two BRI circuits to back them up and we went to the engineers and said, “Okay, we’re gonna have T1s to each one of these sites. How many BRI backup channels do you need?”

And they were looking at, oh, my God. Do we need six? Do we need eight? How are we gonna manage a minimum of six to eight BRIs per location times 120 locations? They were screaming.

Don Layton: Yes and for any of those of you out there who have had to manage a BRI network, you know the issues and come around with a unknown dialing, unknown connections that are up for a long time and you get a very high bill at the end of the month and that’s something we were able to avoid with the new network.

Rich Gore: Yeah. We ended up deciding that along with the primary T1 we would have – at first, we thought we were gonna have a backup T1 to take the place of all those BRIs and it was gonna be a little bit more expensive but much, much less of a headache, and then we thought if we got two T1s, why don’t we just have them both be active.

Don Layton: And it worked out well.

QoS on the WAN

Rich Gore: Yeah. So QoS on the WAN. QoS was pretty important and, frankly, because especially at the time when were undergoing congestion in the first network, we realized that we needed to protect the traffic on the WAN, but we also had QoS on the LAN as you’ll see in the next slide.

We ended up with seven different layers of QoS, seven different classes, that starts with network control, routing control, our voice traffic, our video traffic, voice and video signaling.

Then we broke out data into high priority and default just about everything like all of your email and all of your file transfer stuff and then batch which usually that meant PC backup stuff.

Don Layton: Yeah. You have to remember in a peer-to-peer network that QoS is extremely important to keep your links, these sensitive applications and traffic moving through the network.

Rich Gore: Yeah, that’s right ’cause the whole reason for going to peer-to-peer in the first place is its latency sensitive traffic and so to protect that and to protect it in a variety of ways, you use QoS.

One thing I didn’t understand when I looked at QoS – I just have to say this – is I thought that the classes of QoS service meant that the traffic was more important. QoS6 was more important than QoS5 and that’s not really at all the case.

I was surprised to find that out. It’s just a way of marking them, so

probably the most important stuff up there is – actually maybe it's the routing and control traffic but –

Don Layton: For us network folks; it's the routing and control traffic.

Rich Gore: Yeah, yeah. But maybe after that it might be the voice and video signaling traffic to keep the links alive, so – and batch certainly was – should be at the bottom, but it's just an arbitrary way of marking them.

LAN and WAN QoS

Rich Gore: So this is a picture of how we do marking of traffic in the LAN and in the WAN, so in the LAN, we look to a trusted edge switch and we just selected certain edge switches as being our trusted edge switch and that's very critical, to know which ones to trust, and the edge switches were basically looking at the traffic.

The phones, the IP hardware phones, were marking the voice traffic with the voice bearer channel and the voice signaling information for the switch itself, so the switch would see those markings and then use class of service, CoS, to mark those as signaling traffic or voice traffic.

And then it would look on a purport basis to see about other traffic like video traffic. They're also; I found out – some of our early call center traffic was marked by the switch just according to the ports so that they were given additional precedence.

And then that gets sent across to the router and the router uses all of those trusted edge markings plus – it uses ACLs really to look at all the other stuff, to decide based on port, based on protocol what's the high priority data, what's the low priority data, and then looks to the switch to tell it what's the voice and video.

Vendor Strategy

Don Layton: So our vendor strategy was fairly simple and straightforward. We wanted to work with a single vendor as much as possible to help drive down costs and also have a single throat to choke when things happened.

We used the vendor to drive diversity through their own network. Most vendors have a diverse network out there. You just need to work with them to get that diversity working for you as well.

We work with third-party vendors where necessary in certain theaters or certain locations where we couldn't get the diversity we were looking for from our single vendor, but that was kept to a very small and minimum number. We co-located and consolidated the contracts so that we could again drive down cost –

Rich Gore: It makes it simpler.

Don Layton: – and make it easier and then, obviously, this gives a much better business partnership with whoever your vendor is.

Rich Gore: So I have to ask this. With a single vendor you're sort of putting all of your eggs in one basket. How has that worked for us over the last five years?

Don Layton: So for the last five years, it's actually worked very well. We've had very minimal outages across the network. The biggest issue with them is watching out for grooming. Often you're –

Rich Gore: Grooming.

Don Layton: Grooming, so your vendors may do something called grooming where they will go back without your knowledge and choose new paths for your T1s or your E1s and you need to work with your vendor to identify that you do not want to be regroomed. You do not want to be –

Rich Gore: To maintain that diversity of path.

Don Layton: To maintain the diversity of path.

Rich Gore: Okay, interesting. Have they done anything surprising like jack up the price at any point or –?

Don Layton: We haven't had any price changes. We've just had some path changes and we've found those out in different times.

Hub Site Design

Rich Gore: Pretty good. So now let's talk about the hub site design. We've talked about the network and the backbone. We're gonna drill down a little bit more and I'm gonna pass the baton over to Don because he did all the work on the implementation of the hub site design.

Hub Sites-Equipment

Don Layton: So our hub sites were based on the standard three-layer access distribution and core type of routing. We used the Cisco 7600s for the connections to the backbone, the OC3, OC12 backbones. We used the Catalyst 6500 switches in the center there and we used the Cisco 7200 routers to connect our individual sites into each of the hub sites.

Rich Gore: Okay. So every one of those 120 sites was coming into a 7200 and then going through the layer to network and then coming out on the – or rather all the – the WAN itself was being accessed by the 7600s.

Don Layton: Correct.

Rich Gore: Okay.

Hub Sites-Telco Co-location

Don Layton: So for an individual hub site, this compares to what you would typically have where you would host things at your location, at your site, versus doing a co-location or what we were calling hoteling at the time.

So when you do a standard implementation for a site, you would typically bring those OC12s, which are very expensive circuits to bring down to your site, right into your site and you would end up paying the access cost for those.

By co-locating those pieces in the providers' areas, in their data centers, we were able to avoid that extra cost of those OC12 and OC3 access points.

At that time, the cost of an OC12 access was $25,000.00 a month, so for one of our hub sites such as Denver where we have two OC12s and two OC3s coming in, you're talking about some $75,000.00 a month just in access cost that we were able to save and apply that towards the cost of what is really an inexpensive Teleco co-location site.

Rich Gore: I also sort of see from this diagram that all of the traffic that would normally be coming down those access lines and being routed and switched – or routed really inside your office is no longer being routed in the office. It's being routed up in the vendor's data center, which looks a lot nicer.

It reminds me of what happened in the U.K. when we had all of our access lines from about 20 different sites coming into one site and all of our routing equipment coming into that one site which was a leased office, and when we lost the lease we had to move all of that routing and switching equipment, all of those access lines from one location to another location a few miles down the road.

Don Layton: So this definitely provides you a more flexible solution for issues around facilities like that, but it also – on the other side of it, you're much more tightly tied to your vendor at that point in time.

Rich Gore: Yeah, that makes sense, too.

Don Layton: It gets much harder to extract yourself from the vendor if that ever had to happen.

Rich Gore: How did our engineers feel about letting go of the equipment that used to be in our locations and was now off in the vendor's data center?

Don Layton: Like any network engineer, they like to be able to go and hug their router and they were very uncomfortable at first. One of the ways we were able to get around that is we were able to put web cameras in each one of the Teleco hotel sites or co-location sites so that when we actually had an issue where we needed to use the local smart hands to come in and fix something, we could actually see what they were doing and tell them, "No, no, no. Not that router, the one to your left."

Rich Gore: Okay, so good. So, you know, it's like video camera daycare so you could see your routers even if you couldn't touch them.

Don Layton: That's right.

Rich Gore: That's nice.

IDC to Remote Design

Don Layton: So this is a generic design for the diversity that we used to go from one of our co-location sites down to an individual Cisco site. Using the vendor's specific diversity that's out there, we tried to go to different POPs within them and different Teleco sites and we were able to achieve this in about 70 percent of our sites.

Rich Gore: Why only 70 percent?

Don Layton: Well, about 30 percent of our sites are in locations where it just could not get diversity at all.

Rich Gore: Okay, just didn't exist.

Don Layton: Somewhere in the middle of Idaho or places like that where that type of diversity just didn't exist. We couldn't achieve it.

IP Addressing

Rich Gore: Okay. So IP addressing. If you haven't readdressed your network in a long time, and most people don't because it's an incredibly difficult and tedious task, then it's gonna look a lot like your attic or your garage.

There's gonna be stuff piled on top of stuff, address blocks piled upon address blocks, with very little relationship to each other and Cisco was no exception. We had – from acquisition, from growth; we just had lots and lots of different addresses.

We had done kind of a readdressing. Actually, we'd done a very thorough one in Europe. We'd done another very thorough

readdressing plan in our engineering sites and that included Raleigh, North Carolina and also, to a large degree, San Jose, but all of our other Internet or our WAN sites in the United States and Canada still had kind of a mess of IP address blocks.

And to make this peer-to-peer network work really well, we had to make sure you could summarize routes going – you know, going the north route or the south route, going east of me or west of me in a relatively simple way and so we ended up doing a major IP addressing-readdressing scheme.

We gave each hub site a slash 19 or more depending on the number of addresses that was gonna be there, also an RFC 1918, a 10 net block of addresses for their own lab use and remote access use, and we also – and this was, I think, key – we left room in this addressing plan for growth and, in fact, you mentioned Seattle, Miami and Irvine.

When we did the addressing scheme, we addressed it as if we had Seattle and Irvine and Miami as separate hub sites, so we gave them separate blocks and then when we merged Seattle and Irvine into San Jose, we just merged those – well, we kept those blocks there kind of separate in the off chance that some day we might want to put a new hub site in Irvine and Seattle.

And, in fact, it turns out there's a possibility we might want to put a new hub site in Irvine.

Don Layton: That may come.

Rich Gore: So we'll have the addressing scheme ready for it.

Don Layton: Yeah. Route summarization was a large driver for this network as well, and this is the time to do that route summarization through IP readdressing and IP address planning.

Rich Gore: Yes, which is very painful, but at least it sorts everything out.

Don Layton: Well and it made our networks much more simple and much easier to troubleshoot when there were problems.

Rich Gore: Very true. Yeah. And also we got – we found a lot of address blocks that weren't being used, so we were able to reclaim those and, yeah, just do a spring-cleaning.

Migration Steps

Rich Gore: So now let's talk about the migration steps. I'll lead this off because I know the network as it was and Don, who was instrumental in creating it as it is today, will lead us through the different steps we went through in migration.

Americas US/Canada WAN Backbone

Rich Gore: So this is that cluttered map with all the clutter kind of removed, so underneath all that clutter of all those different 120 links from each of the end offices back in San Jose, there was a triangle of OC3.

Actually, it was T1 a long time ago and then it became DS3 and then back in 2001, it had – actually, I guess '99 through 2000, it was OC3 and even then, we were starting to – we were nearly doubling traffic every – almost every two years and we were starting to run low on capacity on our major backbone between the East Coast and West Coast.

And then, of course, everything else all hubbed back into San Jose.

There were a few exceptions. There was a couple of engineering sites that were hubbing back into Raleigh and then all of our Canada sites were hubbing back into the Kanata site which then came into I think Chelmsford or Boston.

Don Layton: Comes from Boxboro.

Rich Gore: Oh, Boxboro, thank you, our Boxboro site. And that's mostly because the cost of coming across the Canadian border is very expensive so we tried to limit the number of circuits that did that, but that was pretty much the network we were starting out with, so, Don, tell us about the next one.

Phase 1 Rearchitecture-Sep '01-Nov '01

Don Layton: So we started by bringing more of the individual sites, especially around the Raleigh-Durham area, around the Boxboro area, and especially around San Jose – we started bringing those onto the TDM network and bringing them onto the network in the new site at the new location.

We also looked at where we wanted to hub the Canadian network, not just out of Boxboro but where else were we gonna put that redundant link.

Phase 2 Rearchitecture-Nov '01-Apr '02

Don Layton: So due to contractual reasons, we actually started with the northern half of the backbone.

The southern half of the backbone had a little bit longer contract on it and so we started our work by building out the co-location facilities there in New York, in Chicago, in Denver.

And, again, once we had those locations built up so that the Cisco gear was in place, we would start taking the local sites around there and move them over from being on the Frame Relay back to San Jose to TDM back into their closest location.

Rich Gore: And I see we finally got diversity from our Kanata site to –

Don Layton: This is where we put in the diversity so that we had two cross border circuits that were no longer riding the same cross border fiber, we found out, and –

Rich Gore: Oh, great.

Don Layton: Yeah, for a couple of years, and we did put that diverse route into the Chicago co-location facility.

Rich Gore: Nice and in the meantime everything that wasn't connected to those hub sites was still coming back into San Jose just like it had been before.

Don Layton: And we did this with a team of about 10 or 11 network engineers who worked as pairs, and went out to each of the individual sites to do the upgrade that was required on the hardware to bring up the new network and to turn up the connection into the new backbone.

Rich Gore: This was a big project. I remember that.

Phase 3 Rearchitecture-Apr '02-Sep '02

Don Layton: It was. So after we'd gotten the backbone ready on the northern side, we were able to start working on the backbone on the southern side and we went in and turned up the co-location facilities in Atlanta and in Dallas and we ran that OC12 through those back to San Jose and to

Raleigh.

And then we turned up the DS3s, the first couple of DS3s that were connecting the north-south routes to provide diversity and to provide direct access between locations in those areas.

And, again, all during this time, we kept sending out our teams of two men at a time to these various sites and various regions to cut them over from the old Frame Relay to the new TDM network.

Rich Gore: So by September '02, the backbone was done. Most of the offices were linked into their respective hubs and we were just chasing – rather you; you were doing the work and your team was doing the work – chasing after those last offices that were difficult for one reason or another.

Don Layton: Absolutely and as with most projects, the last 20 percent is mostly 80 percent of the work.

Results

Rich Gore: So that was pretty much what led up to our new WAN. Let's talk about the results and I was doing the financial analysis, so I'll talk about some of the financials and then we'll move into some of the other results.

Round-trip Performance Improvement in 15

Rich Gore: So the major reason for doing this was to reduce the latency and, obviously, the latency from the earlier network is going to be especially improved for all East_Coast-to-East_Coast traffics.

So we looked at different pairs or sites, especially those going from one East Coast site to another East Coast site, which had traditionally been going all the way to San Jose and all the way back, so like a 3,000 – no, 6,000-mile roundtrip.

And I said this is where we're gonna get the best possible latency for all peer-to-peer connections and, in fact, that's what we found. We found huge improvements in latency for any site connecting in Chelmsford to any site also connecting into Chelmsford because it was just 260 miles in, 260 miles out, on average, and the response time was significantly better.

Also true for RTP or Raleigh-Durham. Also true for anybody calling from the Chelmsford area to Kanata and that's – the list was – of all possible pairs was much longer than this. I just peeled off the top ones to show you where the significant improvements were made.

Bandwidth Increase: 368%

Rich Gore: And as we mentioned, we were increasing the bandwidth. The spectacular increase in bandwidth was that OC12 backbone, but migrating from a bunch of 256K and 384K and 512K Frame Relay circuits to a minimum of dual T1s that was incredible.

And so when I totaled up all the access bandwidth, just, you know, how many 256s times 256, how many 384s times 384, just totaled up all the bandwidth, and then I totaled it up again with the dual T1 and the multiple T1 and occasional DS3 access, I found that our bandwidth had increased by more than a factor of three, which we talked about earlier, and, as we said, well, sounds kind of expensive.

Don Layton: It did.

Rich Gore: Sounded really expensive.

Cost Increase: 4%

Don Layton: We were scared, but we found on the cost increase that it was very little, somewhere between 4 and 5 percent overall, but the biggest change here was that your cost increase went from those access lines to your backbone, so your backbone cost went up dramatically during this time; whereas, our individual access cost went down.

So for those of you who budget those areas differently, this is something you need to look at and prepare your organization for and the changes that'll happen in that budget.

Rich Gore: Yes. For us, we did have some pain to go through because we had one group responsible for paying for the backbone and other groups responsible for paying for access costs to their individual sites.

And so all of a sudden the backbone group found that their budget was now being blown away by a factor of four or five and all the access – the people responsible for paying for the access were finding incredible relief that no longer was their access costs at all _____.

[Crosstalk]

Don Layton: _____ they had extra money, didn't they?

Rich Gore: Yes. Yeah, so we actually had to go around, you know, moving money from one set of budgets to the other. And this is where I have to say your mileage may vary.

Cost Fundamentals

Rich Gore: Now in doing this financial analysis, we've discovered that the Frame Relay links that we were paying for in 1999 and 2000, 2001 were being priced kind of as if they were a thousand-mile-long leased lines.

And so when we reduced or we moved from one Frame Relay link to a leased line link, that access cost, which is – or cost by mile; the price is per mile – dropped significantly, so, of course, our access cost drops significantly even though the bandwidth was multiplied by a factor of three and a half to four.

But, of course, our backbone was – costs ate up a lot of that savings, all of that savings, in fact. And what we discovered in playing with the spreadsheets, playing with the analysis, was that in the year 2000 to 2002, when we do – when we looked at our 120 sites that were averaging 260, 270 miles from hubs that we were breaking even with dual T1 access.

We found that the more sites you had and the shorter the access lines for a leased line, of course, the better off this particular – or the cheaper this particular architecture is gonna be.

So for customers who have a small number of sites who have already a relatively localized network like all in one state or in Europe all in one country, then the costs are gonna change dramatically but for us the costs were pretty significant and justified our paying for an OC12 which in 2000, 2001 we thought was just really cool.

Don Layton: Yeah. We have to remember in that timeframe the cost of TDM circuits was falling quickly, almost daily at that point in time.

Rich Gore: Yes, that's true.

Don Layton: So that had some effect.

Rich Gore: It had a big effect, yeah.

Don Layton: And we have to remember that the Teleco hoteling or the co-location that we did did save us a significant amount of money by pulling those in-circuits out of our individual sites and bringing them back into the providers' locations.

Rich Gore: Yeah. Yeah, that was big.

Cisco Global WAN Backbone-2007

Don Layton: So here's the picture of the global WAN as it exists right now. You can still see that same OC12, OC3 backbone in the middle of the North American area. You can see that we're also using the TDM network in the Asia-Pac area and in EMEA, we're now using MPLS and that's working very well. All of that works as a good peer-to-peer network, MPLS as well.

Rich Gore: Yes. Yeah, so this is our network optimized for peer-to-peer today and it's done by TDM and Sonnet or it's done by MPLS VPN, but it all works very nicely.

Global Peer to Peer Services in place

Rich Gore: So these are the things we've been able to do with our new peer-to-peer network, actually starting back in the year 2000 before we'd had the WAN completely rearchitected. We were already putting IP telephony onto the network.

Within less than a couple of years, we had completely migrated globally all of our 262 PBXs and key systems off the network and substituted them for 13 production Call Manager clusters, just Call Manager servers sitting in data centers pretty much where servers belong.

And we also are now running 70,000 IP telephones, 30,000 software-based IP communicators on this network and saving a lot of money in terms of PBX maintenance costs, PBX lease costs and a variety of other things and then there's a lot more. There's a lot more applications running on this.

We have all of our voice mail running on IP. It's Unity voice mail. We have all of our contact centers, 17 contact centers, running on IP using IPCC. We have video.

We have actually a variety of different video services running on the network but I chose one, our Cisco Unified Video Advantage, partly 'cause I like it and probably 'cause it runs on a Call Manager. It's basically part of our telephone system. You pick up the phone. If the other person has their camera on, it's a video call. That's pretty cool.

Don Layton: And, of course, the big push we're having in Cisco right now is with TelePresence and TelePresence uses up a lot of bandwidth.

Rich Gore: Yes, it is. Yes, it does, although I put it up here because it requires peer-to-peer network and also because it uses the Call Manager to set up and tear down calls, so I thought that was pretty cool.

Slide 33

Rich Gore: Well, that's about it. That's really all we have to talk about.

I wanted to let you know that if you're interested in learning more about the WAN rearchitecture or about how Cisco IT deploys any of Cisco's products within our own internal network and the benefits that we've gained and the lessons we've learned from doing these deployments,

you can go to www.cisco.com/go/ciscoit.

You'll get to the Cisco on Cisco Web site and learn all about what's going on inside Cisco IT.

Don Layton: Great. Thanks, Rich.

Rich Gore: Yeah, Don. Thanks very much.

Don Layton: Been a pleasure.

Rich Gore: Appreciate it.

Don Layton: All right.

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Europe Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: +31 0 800 020 0791
Fax: +31 0 20 357 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C78-331727-01 10/06

