

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

HAYSTACK OBSERVATORY

WESTFORD, MASSACHUSETTS 01886

Telephone: 978-692-4764

Fax: 781-981-0590

16 September 2003

TO: Distribution
FROM: David Lapsley
SUBJECT: 8 September 2003 e-VLBI telecon summary

Attendees:

Lee Foster, Chuck Foster, Bill Fink, Pat Gary, Kevin Kranacs, Paul Lang – GSFC

Rick Larkin, Peter Schultz – MIT Lincoln Lab

Tom Lehman – ISI-E

Guy Almes, Charles Yun – Internet2

Kevin Dudevoir, Hans Hinteregger, David Lapsley, Arthur Niell, Alan Whitney – Haystack Observatory

This telecon is one of an ongoing series of telecons to prepare for gigabit/sec e-VLBI demonstrations between NASA GSFC and MIT Haystack Observatory using a combination of network facilities including all or part of Glownet, Bossnet, ISI-E, SuperNet, Max and GSFC/HECN.

ACTION ITEMS ARE HIGHLIGHTED IN RED.

Glownet/Bossnet

Tom Lehman reported that he and Peter Schultz had been working on the OC-48 connection. Tom had done some testing over the weekend and had got some good results sending data from ISI to Lincoln lab. Not yet able to test in the other direction, due to a host (SuperGlide) dying.

Peter reported that he had made some adjustments to power levels on amplifiers to get better signal to noise ratio.

Haven't yet tested OC48 at 2.5 Gbps. Currently tested at 1 Gbps. Either use multiple hosts to test this or use VLAN loopback.

Alan Whitney had asked whether it was possible to add an additional wavelength dedicated to e-VLBI. If future of BOSSnet is secure, then we would be happy to explore the idea of maintaining a dedicated e-VLBI wavelength without scheduling (at least 1 x 1 Gbps, possibly 2 x 1 Gbps). Near term e-VLBI goal (for next year) is 2.048 Gbps. OC48 is ideal. In order to achieve this bandwidth from Westford to Lexington possible need to look at purchasing cards for Cisco ONS15454 on GLOWnet. Would also need to purchase equipment for Haystack.

Pat Gary discussed the possibility of a 10 GE lambda on BOSSnet. That could be used on a scheduled basis by e-VLBI. It should not be difficult to add 10 GE lambda from ISI to BOSSnet.

BOSSnet future looks assured. Contracts moving a bit slower than desirable.

NASA/GSFC

Bill Fink reported that NASA/GSFC is purchasing a Force10 GigE switch that will reside outside their firewall router.

Pat discussed changes to the network topology: K1-J5 path would have a 10 GE lambda and Cisco 4912 at J3 has been upgraded to a Cisco 6509 which has the potential to host a 10 GE interface. Pat looking at possibility about having 10 GE capability from GSFC to ISI and to BOSSnet.

NASA/GSFC have ordered 2 x 1p 10GE LAN PHY cards and 1 x 48 port 1 GE card to go with their new Force10 switch. Also looking at 10GE WAN PHYs and would like to test with Nortel 8600 at ISI. Looking at borrowing a WAN PHY card from Force 10 and testing back to back at ISI to find out interoperability. Once interoperability has been established, can then look at interoperability over distance.

David Lapsley suggested that it might be interesting to test MPLS between Force10 and Junipers as well as testing the ability to tunnel 10GE WAN PHY over OC192 SONET network.

Pat mentioned that Force 10 10GE blades would also support IPv6 and multicast. The Force10 switch should be able to route non-blocking from any 10 GE port to any other 10 GE port.

Security

Pat mentioned that they had been testing an Alteon firewall from Nortel. Trying to configure it to host checkpoint firewall software. Physically the Alteon has 8 x 1 GE ports. Hoping to do link aggregation of 2/3/4 Gbps in/out. Theoretical throughput of Alteon is 3.5 Gbps. Alteon can't handle 2 Gigabit link aggregation yet (it breaks). Currently a strong push to ensure that all networks NASA/GSFC participates in are behind stateful firewalls.

Pat mentioned that they still have a Nissan Systems IP Storage Switches on loan. Have a couple locally installed. Testing showed 60 MBps transfers for read and 75-85MBps doing writes over 18 hours between two different SANs. Hoping to test Nissan boxes between GSFC and University of Maryland College Park. Bill has started detailed checkout plans between GSFC and San Diego Supercomputing Center. They have one of these SAN boxes. This testing will help to validate and harden high speed networks.

Guy Almes asked whether GSFC could have some LANs outside firewall. Pat said that currently, e-VLBI is outside firewall. Can keep this for a certain amount of time, but there is pressure to make sure that everything is behind a firewall. Guy mentioned about performance versus security tradeoffs. In particular, Guy mentioned the security workshop held at Lawrence Berkeley labs (DoE lab so high security requirements, where there were examples of people taking both performance and security seriously. One example is using access control lists on routers but not using stateful firewalls.

Bill believes there are ways to make the system more secure without impacting performance by using distributed access control schemes.

10GigE Intel NIC cards

Bill looking at doing some routing tests at the IP layer.

International Connections

Kevin Dudevoir reported on his conversations with people in Germany regarding E3 (34 Mbps) connection to Wetzell telescope site in October.

David reported on his conversations with people from the University of Hawaii and the Australian Academic Research Network regarding possible 10GE links from the East coast of Australia to the West coast of US (via Hawaii). If approved, timescale for installation would likely be 1st quarter of 2004.

David reported on high speed connections to Australia. David looking at running some experiments between the US and Australia. People at Internet2 put him in touch with George McLaughlin who used to be the CEO of AARNET (Australian Academic and Research Network) similar to Abilene/Internet2 here. George was very helpful. George has talked about their plans to install two new 10 GE lambda fibers between East coast of Australia and the West coast of US. Currently the planning is quite advanced but not sure when they find out about whether or not this will happen. These links will be going to Hawaii and landing on the big island there. Last mile problem is an issue in Australia. Cost of connecting at high bandwidth is prohibitive (~ \$1 million for an OC3 per year from mainland to telescope in Tasmania). Telstra, the incumbent carrier has a virtual monopoly on entire market. In Sydney (where some other telescopes located) cost is still prohibitive. They currently only have 512 kbps connection to the Parkes telescope. Planning to have a 1 Gbps connection to Marsfield (correlator site in Sydney). Should be in there now, need to check to see if that is currently in or not. NASA Deep Space Network (Tidbinbilla) is a possibility. The NASA DSN network is used a few times a year for geodetic experiments.

Hans Hinteregger mentioned that connectivity in Australia is a policy problem: most telescopes have Telstra lines running right past them, but don't have permission to use them.

Action Item: David to check on NASA DSN availability for e-VLBI testing.

Performance Monitoring/Testing

Initial discussion regarding Grid/OGSA infrastructure for iperf testing. Charles Yun felt that this was a bit of overkill. Charles discussed a shell script that Internet2 was looking at developing to provide some level of security for iperf test sessions. This script would enable a client and server to securely exchange a random port number to be used for the test and would launch an instance of an iperf server to serve the client request. After the client had finished testing, the server would then shutdown. This is a temporary solution for the short term while the complete End to End performance testing infrastructure is deployed.

Bill mentioned that nuttcp already has server mode and he is considering enhancements that would allow scheduling and queuing as well as authentication. Nuttcp already has limited amount of access control, as it runs as an inetd service- means you can make entries to hosts.allow/.deny. Bill looking at adding to client code to make client wait until port becomes free (if test already in progress with another client). Nuttcp also supports ipv6.

Action Item: Bill to add code to nuttcp client to add queuing mechanism

Action Item: Charles to discuss secure iperf wrapper with David.

David reported that he had checked with Steven Low of Caltech regarding the distribution of FAST within the group. Steven had said it is ok to distribute to people within the group. David has put the kernel on the web. If anybody is interested in obtaining it, send him an email. If you would like your own custom build, then Steven encourages you to go directly to his website (<http://netlab.caltech.edu/FAST/software.html>) and register there.

Performance Test Scheduling

Performance testing with FAST: best result is currently 397 Mbps between here and APAN node in Tokyo. 200 ms delay. Initial results only ~100 Mbps, because default FAST parameters are tuned for LAN. Need to increase rate at which it gets access to bandwidth. Currently, 2 x OC12 between Tokyo and the US soon to be upgraded to an OC48. Best throughput is currently 397 Mbps. Took a couple of debug, recompile cycles to get the correct drivers in the FAST kernel for Turtle. Caltech people were very helpful in doing this and very responsive. Early results had very low throughput: 100 Mbps with a peak of 267 Mbps, while UDP throughput was ~ 500 Mbps. Sent this information back to Caltech people. FAST has a parameter that controls how quickly it can increase the window size, by increasing this parameter, enabled us to increase the throughput to 397 Mbps. May be possible to get higher. Currently to Kashima, currently limited to 100 Mbps, which is bottleneck between Kashima and Tokyo POP, hoping to upgrade, but most likely will not happen for a year. Haven't yet had a chance to compare FAST performance with regular TCP performance yet, on a list of things to do.

Action Item: David to do more detailed performance testing and present the results.

Fastest transfer for Internet2 demonstration would be GGAO to Indianapolis. Bottleneck would be Mark 5 throughput. Main point of demo is not to set a throughput record, but to explain what e-VLBI is about, advantages of e-VLBI, and what is required to realize the potential. Will be emphasizing the last mile connectivity problem to most telescopes. E.g. going to Mauna Kea will be limited 45 Mbps, so this will be the maximum "session" rate.

Kevin Dudevoir reported on his testing to Germany and Hawaii. Wetzell is hopeful to have a 34 Mbps very soon (within 2 weeks latest estimate). Currently seem to have 100 Mbps to a server located on SURFnet. As soon as Wetzell has 34 Mbps connection will test directly to the node. Kevin mentioned that David has secured a test point at the University of Hawaii at Oahu to test to. Still waiting for permission to test to this server. Still measuring end to end 5 Mbps for a single TCP stream. UDP still measuring a 25% packet loss. PMRF has a micro-wave link from PMRF to Kokee. Our contact person Kris Blackstat is leaving. Trying to verify new contact is Patrick Alvarez. Haven't yet been able to contact either Kris or Patrick. Trying to contact either Patrick or his replacement to get a test point deployed at PMRF. Test station is configured with a suite of tools, waiting for confirmation before sending the test station. Test point at University of Hawaii will be useful. Still not sure about whether microwave link is performing properly. PMRF reconfigured microwave link, but not able to see any change in performance. Currently measuring a couple of times per week. Alan suggested it might be good to get people like Patrick involved in our telecons.

Tom had not had a chance to look at web calendar software, but had looked at Globus in detail. Need to look at how to co-ordinate use of iperf to make sure there is no contention when co-coordinating experiments. Need to work out a way to ensure that an experiment can run without being interrupted

by an iperf session. Looks like Internet2 end to end performance infrastructure may solve a lot of our problems.

Charles, Internet2 End-to-end Performance Initiative solving these problems, however, they are looking at a comprehensive solution that will not be deployable for a little while. Probably for another 8-9 months before they reach beta. In the meantime, look at quick hacks to get us through in the short term and get in synch with the work being done in the E2EpI group.

Many of the machines being made available to e-VLBI for testing belong to Internet2 End-to-end Performance Initiative group.

Tom asked if it was sufficient to have a schedule up on the calendar so people can see what to expect in terms of high bandwidth utilization. Do we need to document more clearly?

David said that existing calendar system was good. Would be nice to have an API that would allow applications to download schedule.

Tom suggested that we look at having an overview map that shows e-VLBI test program and then link in Andy's program to show end to end performance, historical data from MRTG. Have a single page to do this. Alan thought that it would be useful to have an e-VLBI map showing all of the e-VLBI test sites.

Action Item: Tom, David, Bill and Andy discuss monitoring.

Internet2 Demo

David reported on preparations for the Internet2 demonstration. He has been working on getting performance servers up and running so that it will be possible to characterize the network performance between all of the sites and the eventual destination in Indianapolis. Signed up already: Mauna Kea, Arecibo, GGAO and also talking to people in Japan for their participation. Demo will involve transferring VLBI data from various remote sites to a Mark 5 in Indianapolis connected to Abilene at 1 Gbps. JIVE interested in principle in participating.

Demo dates: 14th and 15th October. Also been working on software to do RTP framing and transmission protocol for the demonstration. Source code available when it is running.

Alan followed up on the threatened Arecibo OC3 link. University of Puerto Rico with which Arecibo is sharing the link is not able to maintain it. If the OC3 link went away, it would force Arecibo back to a much lower data rate. There has been a couple of months "stay of execution". Alan sent out several emails to Arecibo asking them who their primary link users are without much response. Alan also sent emails to a group of people who use Arecibo. The only response was from Prof. Don Backer from Berkeley, who uses Arecibo quite a bit. He was very supportive. Might not be a lot of users demanding a high bandwidth link to Arecibo. However, they would like to maintain the link for uses that will be coming up including e-VLBI. Contact at Arecibo: Arun Venkataraman.

Charles reported that he has been trying to tackle the disconnect issue for at least a year. Link between UPR and Arecibo is the link at issue; this is the expensive link because it is paid to the Puerto Rican telecom provider.

Rick Larkin said that he would work with Kevin Dudevoir to replace Summit switches with Dell switches. Rick said that the Dell switches are starting to become popular. Need to ensure that they

are configured properly. Throughput performance results will hopefully be out before next telecon.

Action Item: Rick and Kevin to replace Summit switches with Dell switches.

Next telecon

Next telecon is scheduled for Wednesday, 1 October 2003 at 9 am EDT

cc: Steve Bernstein, LL
Jim Calvin, LL
Rick Larkin, LL
Lorraine Prior, LL
Peter Schulz, LL
Leslie Weiner, LL
Herbert Durbeck, GSFC
Bill Fink, GSFC
Lee Foster, GSFC
Pat Gary, GSFC
Andy Germain, GSFC
Chuck Kodak, GSFC
Kevin Kranacs, GSFC
Paul Lang, GSFC
Aruna Muppalla, GSFC
Mary Shugrue, GSFC/ADNET
Bill Wildes, GSFC
Dan Magorian, UMCP
Tom Lehman, ISI-E
Jerry Sobieski, MAX
Guy Almes, Internet2
Charles Yun, Internet2
Richard Crowley, Haystack
Kevin Dudevoir, Haystack
Hans Hinteregger, Haystack
David Lapsley, Haystack
Arthur Niell, Haystack
Joe Salah, Haystack