

Gb Line Setup and Tests

P. de Vicente, L. Barbas, R. Bolaño, F. Colomer

Informe Técnico IT-OAN 2009-05

Revision history

Version	Date	Author	Updates
1.0	01-04-2009	P. de Vicente	First version

Índice

1. Introduction	3
2. Line layout and characteristics	3
2.1. Nearby the OAN observatory	3
2.2. The complete connection layout	7
2.3. TELNET demux equipment	8
2.4. Routers	8
3. Line tests. Breakdowns	9
3.1. First breakdown	10
3.2. Second breakdown	10
3.3. Checks to be done during a breakdown	10
4. Connection setup. Routing	11
5. Tests	12
5.1. First tests	12
5.2. MTU size	14
5.3. Second test	14
5.4. Note on 1024 Mbits/s data flux	17
6. First eVLBI observation	17

1. Introduction

Note: This version is public and some parts have been taken out. Full report is available only privately.

eVLBI observations require to transfer observational data from the radiotelescopes to the correlator in real time. The European Commission has funded a program called Express whose main objective is to connect up to 16 of the world's most sensitive radio telescopes on six continents to the central data processor of the European VLBI Network at the Joint Institute for VLBI in Europe (JIVE). The activities foreseen in this program were basically three:

- Attack the most urgent technical and economical problem at stations: the "last mile" connection.
- Update the correlator at JIVE to process up to 16 data streams at 1 Gbps each in real time
- Investigate possibilities to use distributed computing to replace centralized correlators.

Express started in 2006 and finishes on 2009. More information on Express can be found at <http://www.expres-eu.org/index.html>.

The "last mile" problem usually refers to the connection between the radiotelescopes and the nearest NREN (National Research and Education Node), and contrary to what the name may suggest, the "last mile" connection may be up to 70 km long. NRENs in Europe are grouped into GÉANT2, the European Research Network. GÉANT2 provides the infrastructure connection for 34 countries. Fig 1 shows the connections between all european countries at the time of this report.

Herefter we describe the startup and setup of the OAN last mile connection, that is, the Gb line that connects the OAN 40m radiotelescope to the spanish NREN at Madrid. This connection has been partially funded by Express.

2. Line layout and characteristics

The Gb connection between the OAN and RedIris, the spanish NREN, was awarded to Telefónica. The physical connection is an illuminated optical fiber composed of 8 monomode fibers. Illumination is done by Telefónica's optical network; That means that Telefónica installed two devices at both ends of the monomode fiber which multiplex and demultiplex the optical signal and transform it into a Gigaethernet connection. Transport in the monomode fiber is done using WDM at 2.5 Gb/s.

2.1. Nearby the OAN observatory

The connection layout in the observatory proximity is depicted in figure 2. We have included a short description to help spotting future hardware problems in the cable.

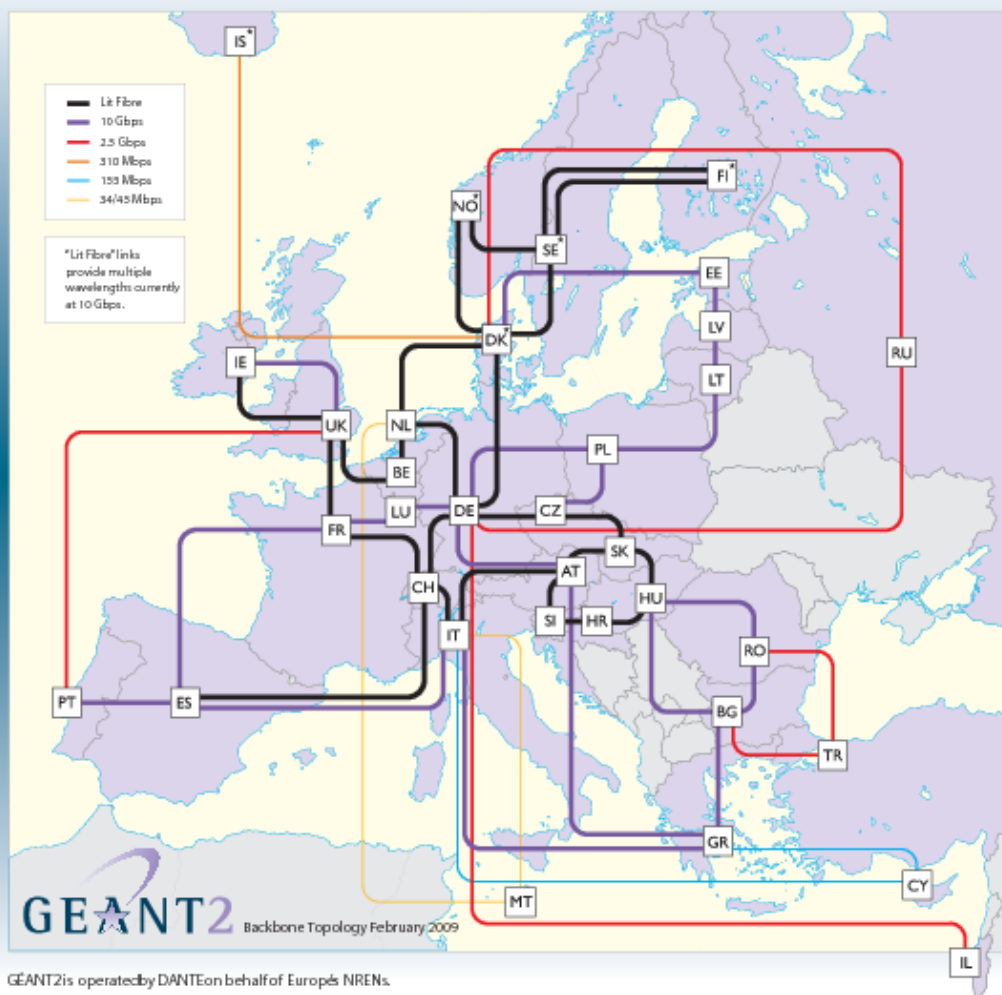
The fiber departs from the control room rack in the 40m radiotelescope. It runs up to the observatory entrance using underground tubes interconnected by coffers on the ground separated 25 m. At the observatory entrance the cable exits one of the coffers and climbs to a wooden



GEANT2

Lighting the pathway for European Research and Education

The first international hybrid research and education network.
Lighting dark fibre for greater network performance.



GEANT2 is operated by DANTE on behalf of Europe's NRENs.

AT Austria	CZ Czech Republic	ES Spain	HR Croatia	IS Iceland*	LV Latvia	PL Poland	SE Sweden*
BE Belgium	DE Germany	FI Finland*	HU Hungary	IT Italy	MT Malta	PT Portugal	SI Slovenia
BG Bulgaria	DK Denmark*	FR France	IE Ireland	LT Lithuania	NL Netherlands	RO Romania	SK Slovakia
CH Switzerland	EE Estonia	GR Greece	IL Israel	LU Luxembourg	NO Norway*	RU Russia	TR Turkey
CY Cyprus							UK United Kingdom

*Connections between these countries are part of NORDUnet (the Nord region network)

GEANT2 is co-funded by the European Commission within its 6th R&D Framework Programme.



This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of DANTE and can under no circumstances be regarded as reflecting the position of the European Union.

Figure 1: GEANT2 infrastructure at the time of this report

mast. From there to Yebes, the cable is aerial, attached to wooden masts separated 50 m approximately and runs parallel to the road. At some point the cable crosses the road. In Yebes the last wooden mast connects to a cement mast at street “El Camino de la Ermita”, and then runs underground up to the small Telefónica prebuilt small house at street Río Tajuña. From there it runs through an underground connection, to the exit of the village, at one side of the road. It crosses a street that drives to a new Yebes district and then rises again to a wooden mast. From there to Horche the cable is aerial again. At Horche the cable goes underground again up to the Telefónica switchboard. From there to Guadalajara the connection is underground and possibly runs together with that one from the satellite station in Armuña de Tajuña.

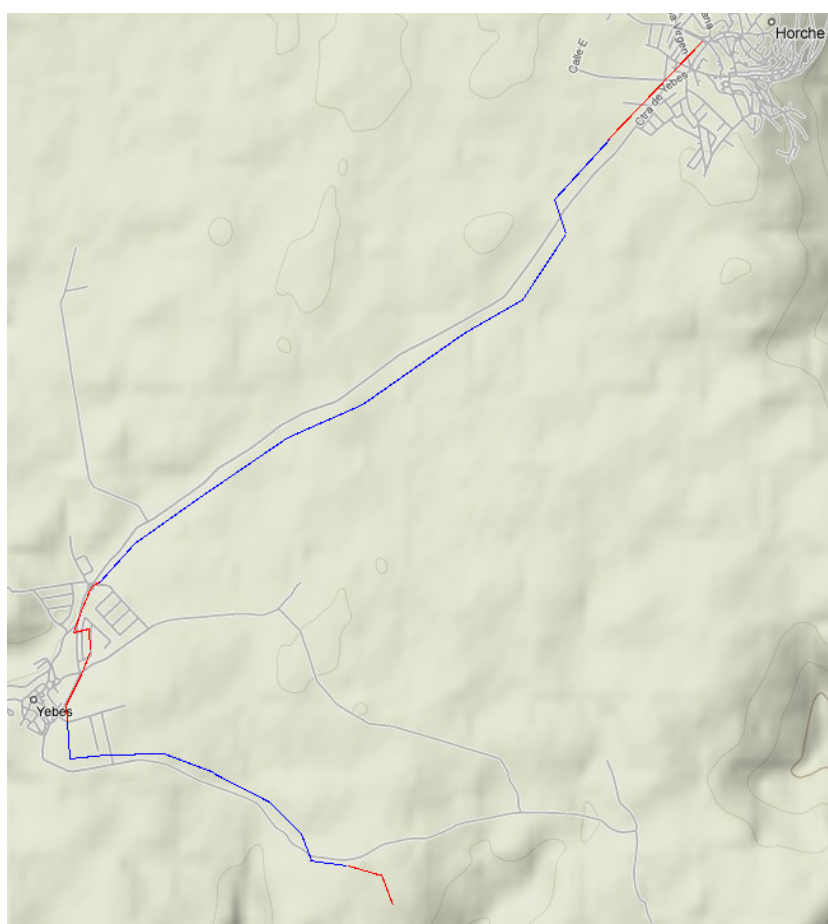


Figura 2: Layout of the Gb line in the nearby of the OAN observatory at Yebes. Underground line is depicted in red, and aerial one is blue.

The optical fiber is soldered at different points along its way. These connections are clearly visible when the trajectory is aerial. Figure 3 shows an example. The fiber makes a circle attached to a wooden mast. Below the circle is a grey small box around 1.5 m above the soil that houses the soldering.



Figura 3: *Wooden mast with a connection box that houses the soldering*

2.2. The complete connection layout

The Gb line runs through several Telefónica facilities from the OAN to the spanish NREN: Horche, Santa María at Guadalajara, Madrid Norte, and then passes through a number of Telefónica small facilities, located in the north of Madrid. In all Telefónica facilities there are patchings which are documented in Figure 4. The last facility is at Telvent in C/Valgrande 7, Alcobendas. This is a building that houses a neutral point, called “meet me room” where the different communication companies can connect among them. The monomode fiber ends at the spanish NREN area. Rediris (the spanish NREN), usually provides a separate space in their restricted area or jail, for communication phone companies. This special area is referred to ETO and is equipped with racks which provide space and electrical power for the end equipments. This special arrangement guarantees that the last patchcord is a short multimode fiber between the last phone company equipment where monomode fibers terminate and RedIris routers.

We ignore the total physical length of the connection, but we estimate it to be around 70 km. Along the way between the Observatory and the ETO area in RedIris jail, there are two devices located at Guadalajara Santa María and Madrid Norte where the signal is amplified. The model of these equipments is unknown to the author but they are referred in the scheme as: "4XANY-1".

Figure 4 shows the original line scheme planned by Telefónica. The end at Telvent does not reflect the reality since the equipment was originally planned to be located at Telefónica's rack in the "meet me" room and later it was moved to RedIris' ETO area for simplicity.

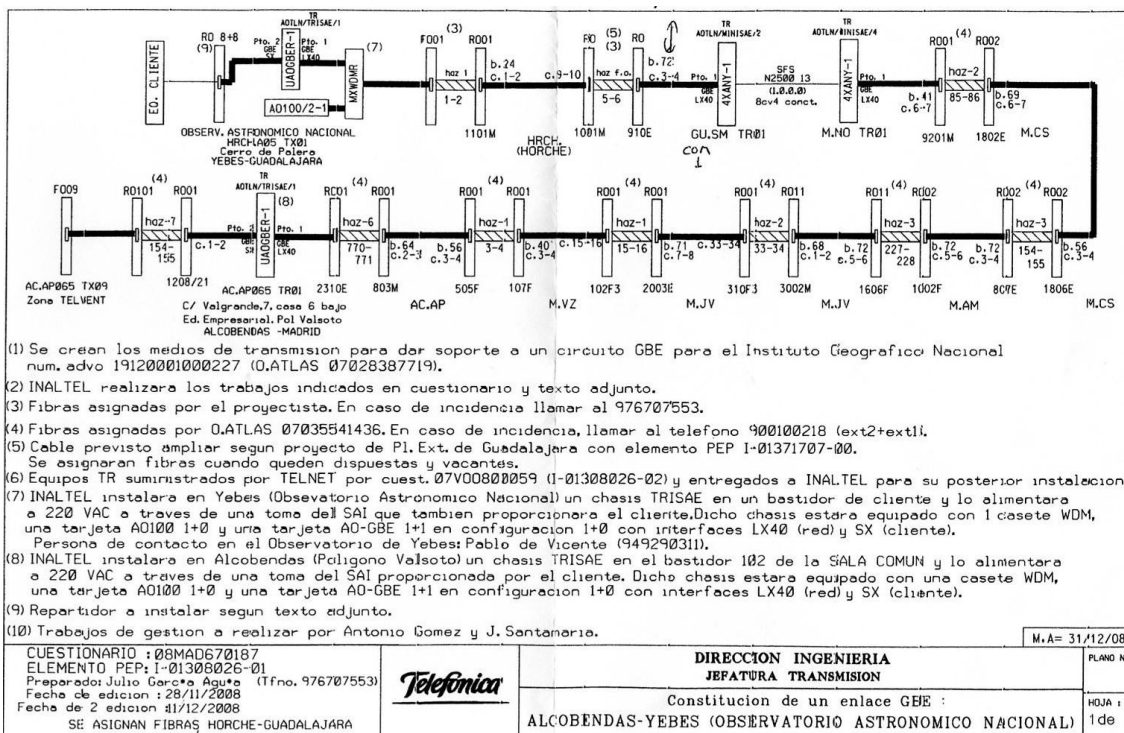


Figura 4: Original line scheme planned by Telefónica

2.3. TELNET demux equipment

Two devices at both ends of the monomode optical fiber emit and receive the optical signals. These equipments belong to Telefónica and are managed remotely by them. Each device is a TRISAE chassis made by spanish company Telnet (<http://www.telnet-ri.es/>) whose main function is to convert from WDM to Gigaethernet. The device contains a mux/demux unit which separates the optical signal in two different wavelengths: 1310 nm and 1550 nm. The first one is used for data transmission and the second one to manage remotely the line using SNMP. There is a second module with two slots for two converters from WDM to Gigaethernet of which one is populated an the second is empty. Fibers out of the Gigaethernet module are multimode, the rest are monomode. The Gigaethernet card uses LC fiber connectors, also known as SFP (Small Form Factor Pluggable) transceivers which work at standard wavelengths.

The chassis has got two RJ45 local ports to manage the equipment. One is for the LAN and the second one is for a serial port connection. The LAN connector is not activated but the serial port is. For security reasons the serial port pinout differs from standard ones.

Optical power at the remote output was measured at 1310 and 1550 nm. Values are summarized in table 1.

Direction	Power at 1310 nm	Power at 1550 nm
Output	-2 dBm	-3 dBm
Input	-17.7 dBm	-18.8 dBm

Cuadro 1: *Detected optical power in mW at two wavelengths: 1310 nm and 1550 nm. For example: -3 dBm is equivalent to 500 μ W.*

2.4. Routers

At both ends of the connection there are two routers which belong to RedIris and the OAN respectively. They are connected to the TELNET devices through a Gigaethernet short multimode patchcord.

The router at the OAN end is a Hewlet Packard Procurve switch model 3500yl. It has got 24 1 Gb ports of which 20 are electrical ones and 4 may be optical (using Gbics) or electrical. Connectors are STP and LC respectively. This switch has 3 level management and hence can work as a router. Connection between the Telefónica equipment and the ProCurve switch/router is achieved by a multimode 50/125 patch cord with LC connectors on both ends. Each switch port has got two small green LEDs with an arrow shape, up and down, that indicate the data flow direction. Up means inbound traffic and down outbound traffic.

The router at RedIris end is a Nortel 8010. We ignore its characteristics but, obviously, it has got at least 1 Gb port to which the multimode fiber is connected.

3. Line tests. Breakdowns

Line was operative since March 11th, 2009. A technician came on March 12th 2009, to measure the optical power on the Observatory end. He coordinated his task with another colleague on the other end. The line was tested creating loops at both ends and measuring the optical power and checking that both TELNET equipments were linked and working.

The line is easily diagnosed to be active by inspecting the arrow shape LEDs on the optical connection port of the OAN router. If there is level 2 connection between RedIris router and the OAN router, the inbound LED at the Procurve switch will light almost constantly. If we are not in front of the equipment two alternative ways will show this same status:

- The Procurve web page shows a view of the switch status. If the link is up the matching port will light in green as seen in figure ??.
- The procurve Command Line allows to monitor the activity of each port. Once we have logged in using ssh, this is easily achieved by typing:

```
show interface 21
```

If the connection is active the port “Link Status” will be Up. If not it will be Down. Below is an example of the information provided by the router. In this case the connection is active:

```
Status and Counters - Port Counters for port 21
```

```
Name      :
Link Status      : Up
Totals (Since boot or last clear) :
  Bytes Rx       : 2,402,444,590      Bytes Tx       : 3,566,931,062
  Unicast Rx     : 222,103,264      Unicast Tx    : 393,950,742
  Bcast/Mcast Rx : 261,943                       Bcast/Mcast Tx : 50,216
Errors (Since boot or last clear) :
  FCS Rx        : 12,166           Drops Tx       : 364
  Alignment Rx  : 0                Collisions Tx  : 0
  Runts Rx     : 0                Late Colln Tx : 0
  Giants Rx    : 100              Excessive Colln : 0
  Total Rx Errors : 12,266         Deferred Tx    : 0
Rates (5 minute weighted average) :
  Total Rx (bps) : 6418360         Total Tx (bps) : 5073368
  Unicast Rx (Pkts/sec) : 136      Unicast Tx (Pkts/sec) : 89
  B/Mcast Rx (Pkts/sec) : 0        B/Mcast Tx (Pkts/sec) : 0
  Utilization Rx : 00.64 %         Utilization Tx : 00.50 %
```

Every time the swith port LEDs were off the circuit had got a broken connection. This shows that the LED status is a very powerful and simple way to diagnose problems in the connection.

Since the startup of the Gb connection the router showed some problems easily discovered inspecting the log. This error was repetitive with a period of 1 to 2 minutes, mostly with data traffic. Below we include an example:

```
W 03/18/09 07:16:44 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:19:21 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:20:45 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:22:08 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:22:29 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:25:59 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:27:43 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:28:56 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:30:20 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:33:28 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
W 03/18/09 07:34:31 00329 FFI: port 21-Excessive CRC/alignment errors. See help.
```

This error refers to Cyclic Redundancy Checks and it means that the checksum algorithm result does not match the result from the detected bits. This is possibly due to a low SNR or, to optical fiber alignment problems.

3.1. First breakdown

First breakdown appeared on March 13th. The link status LED was off. After talking to the technicians they discovered that one SFP connector at the signal amplifier equipment in Madrid Norte was loopy and not properly inserted in the card. The problem was solved after reinserting the SFP module. Problem was solved after a couple of hours.

3.2. Second breakdown

Second breakdown was on March 24th and lasted for 32 hours. The cause is still partially unknown. Optical fibers in Madrid Java were cleaned and reconnected. Telefónica informed the author that after cleaning some optical fiber patchcord ends, the signal increased 10 dB. This caused the disappearance of the CRC errors. However we believe that the breakdown was due to another cause, possibly a disconnected patchcord which was reconnected in a wrong position.

3.3. Checks to be done during a breakdown

If the link LED switches at the OAN Procurve switch off, the most probable cause is a connection breakdown. Before reporting the error to Telefónica two simple tests may be performed to discard that the equipment at the observatory is the cause.

- A loopback can be performed with the multimode optical fiber patchcord. OAN Procurve port 21 output should be connected to port 21 input. If the Procurve is working correctly the LED should light up. If the test is successful the Procurve is working correctly.
- A loopback may be performed at the entrance of the Telefónica equipment in the monomode fiber. In the tray on top of the equipment there is a monomode fiber with a bridge that allows to make a loopback. The basic idea is that the output connector of the line end should be connected to its input. In this way the signal from the Procurve is converted from Gigaethernet to WDM, then it is multiplexed, gets out, and returns through the loopback, is demuxed and converted from WDM to Gigaethernet and sent back to

the OAN Procurve switch. If the Procurve LED lights, the Telefónica equipment is also probably working correctly.

A third test should be performed. It consists on measuring the optical signal power at the entrance of the Telefónica equipment. This is easily done with the OAN optical power meter model OPM5-2D from Noyes. A monomode optical patchcord should be used. Absolute levels should be measured at 1310 and 1550 nm. Correct values are summarized in table 1

If the OAN Procurve switch is working correctly the error should be reported to Telefónica who should spot and repair it. Administrative number and contact telephons are in the internal OAN wiki.

4. Connection setup. Routing

Once the Gb connection was operative we transferred all communications between the Observatory and the rest of the world from the old 2 Mb/s line to the new 1 Gb line. The main usage of the line is for VLBI data transfer but it is possible to reserve a small amount of bandwidth for the daily normal traffic which does not affect VLBI data flux. Original plans were to set a Quality of Service rule which prioritizes the VLBI traffic over normal traffic. For the time being this has proved to be unnecessary but it may be implemented in the future.

The OAN is a C IP class splitted in three subclasses which are used in the three OAN facilities at Spain. Subclass at the Yebes facility, where the 40m radiotelescope is located, corresponds to 193.146.252.0/25. Routing was originally done using a VLAN through the University of Alcalá, and then to our facility in Alcalá that would connect to Internet. So the new line required a new routing directly to RedIris. Routing tables were modified at RedIris, University of Alcalá de Henares and RediMadrid, the scientific research network at the Community of Madrid.

RedIris assigned VLAN 72 for communications from the OAN at Yebes. Traffic between both ends however is not tagged. Figure 5 shows a simple diagram of the connection between RedIris and OAN-Yebes.

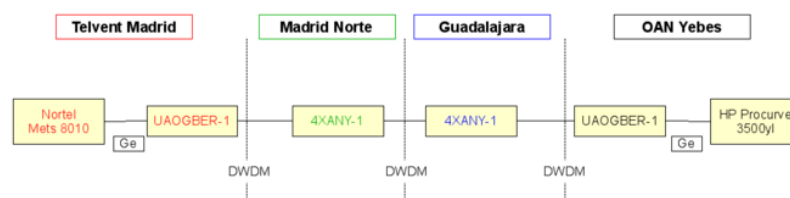


Figura 5: Simplified scheme of the connection between RedIris and OAN-Yebes

To avoid malicious traffic the Procurve Switch uses a restrictive access-list that filters traffic according to its source and destination as well as port used. TCP, UDP and ICMP packets are filtered. In order not to compromise OAN's security the access lists will not be included in this report. However we will mention ICMP filtering here.

ICMP stands for Internet Control Message Protocol. ICMP messages are constructed at the IP layer. IP encapsulates each ICMP message in a single IP datagram. ICMP traffic is very useful to detect problems between routers and hence it is specially important when setting up a new connection. “ping” and “traceroute” use ICMP packages and allow to discover loops or misconfigurations. The ICMP header starts after bit 160 of the IP header traffic and contains type and code. The OAN-Yebes router inspects the packets and allows the following ICMP traffic types and codes, filtering out the rest for security reasons:

- 0 or echo reply
- 8 or echo
- 11 or time exceeded
- 3 or destination unreachable

A more restrictive policy would be to use type 3 4, which means “Fragmentation Needed and Don’t Fragment was Set”. However traceroute may not work properly

5. Tests

Two groups of tests were performed. The first one on March 20th and the second on March 30th and April 1st. The aim of the tests was to measure the quality of the line; that is, the maximum available speed without significant packet losses and improve it if the results were not satisfactory. Between both groups of tests a breakdown (the second) took place and as a consequence the setup was improved. Details are described in this section.

In order to make tests we need used two Linux hosts at OAN and JIVE. These hosts are two Mark5 devices, one recorder at the OAN end and one player at JIVE end. The OAN Mark5B is directly connected to the Procurve 3500yl switch using an STP electrical cable.

Tests were done using tool `iperf`. This tool allows to measure TCP and UDP bandwidth performance, packet loss and delay jitter. On one end the client is started and on the other a server is also started. The latter can be run as a daemon. The server prints periodic intermediate bandwidth, delay jitter and packet loss reports at the specified intervals.

5.1. First tests

These tests were conducted by Mark Kettenis at JIVE who provided instructions to Pablo de Vicente.

Iperf package was installed at the Mark5B host:

```
apt-get install iperf
```

From account oper we first tested UDP packets at 100 Mb/s for 10 seconds with an MTU size of 1500 bytes while at JIVE the server program was started:

```
iperf -c 192.42.120.66 -u -l 1500 -t 10 -b 100M
```

The measured bandwidth was 101 Mbits/s:

Next tests were at

- 500 Mbits/s. Measured bandwidth was 511 Mbits/s.
- 750 Mbits/s. Measured bandwidth was 784 Mbits/s.
- 800 Mbits/s. Packet loss was 0.22 %. Measured bandwidth was 820 Mbits/s.
- 900 Mbits/s. Data at 900 Mbits/s did not achieve the commanded flux rate and kept at 820 Mbits/s. A packet loss of 2 % was measured. This was indicative of some kind of problem in the connection.

The MTU size was changed at the Mark5B interface to 9000 bytes and the test at 900 Mbits/s was repeated with a packet size of 4000 bytes. Bandwidth increased to 913 Mbits/s but no package arrived at JIVE.

Finally two formatter tests were run for 15 minutes: at 256 Mbits/s and at 512 Mbits/s. The formatter test required a clock source and VSI data flux and stopping `dimino` and starting a JIVE program called `jivemark5a`.

The clock and VSI data flux was easily provided by our VSI sampler card, which can be commanded to generate a TVG (Test Vector Generator) pattern from the Field System prompt:

```
> vsi4=tvvg
> vsi4
09:39:20;vsi4
09:39:20/vsi4/tvvg,1,2,0x10
```

`jivemark5a` was started after stopping `dimino` and switching off the key for security reasons. It is important to let pass some time after `dimino` is stopped so that the disks can be properly unmounted:

```
Enddim
```

```
/opt/jive5a-1.0.1 -m0
jive5a Copyright (C) 2007-2008 Harro Verkouter
This program comes with ABSOLUTELY NO WARRANTY.
This is free software, and you are welcome to
redistribute it under certain conditions.
Check gpl-3.0.txt.
```

```
Found 1 StreamStorCard
XLR#1: PCI-816VXF2 Serial: 8182
      0 drives/0.00GB capacity/0 buses/2 ext. ports
```

Results from the tests were successful. Flux was steady at both 256 and 512 Mbits/s and green lights were present at the correlator indicating that data were properly decoded at JIVE.

The next group of tests was delayed until Paul Boven from JIVE was available. The aim of the second group of tests was to investigate if increasing the MTU size reduced the packet loss and improved the quality of the line. In between a breakdown prevented the 40m radiotelescope from taking part in the first eVLBI observation.

5.2. MTU size

MTU stands for Maximum Transmission Unit and it is the maximum unit expressed in bytes used in a given medium. Communication between hosts using ethernet is done segmenting data in 1500 bytes size packets. This unit has not always been so and matches that of ethernet (10 Mb/s). This value is currently used in Internet connections. It is well known that efficiency is limited for this size and speed, see Mathis et al. (1997), for a detailed explanation.

One of the main causes that limit bandwidth when transferring data packets is network congestion. If the network is suffering congestion, the different TCP implementations reduce the data rate at which TCP packets are sent to avoid collisions until a new equilibrium data rate is achieved. This mechanism is inherent to TCP protocol since every packet that is sent over the network requires confirmation from the target.

Many authors have suggested that efficiency in high speed communications (≥ 1 Gb/s) may be improved by increasing the MTU packet size. Collisions are reduced since the number of packets is reduced by 8, and CPU load decreases since the data does not need to be fragmented in so many small parts. A size of 9000 bytes is known as Jumbo frames and seems to be arbitrary. It is somehow limited by the effectivity of the CRC which is known to decrease for 12000 bytes. In any case VLBI data transfer uses UDP which is usually reserved for reliable networks. Packets are sent and no confirmation is required from destination. When TCP and UDP coexist in a network, TCP always adapts to the state of the network, and hence is less competitive than UDP traffic.

5.3. Second test

The second group of tests were done by the author using Yebes Mark5B and Paul Boven at JIVE. The following tests were performed using `iperf`:

- 800 Mb/s with MTU = 1500 bytes. Measured rate was 800 Mb/s. Packet loss was 0.0025 %
- 900 Mb/s with MTU = 1500 bytes. Measured rate was 800 Mb/s. This spotted a problem at Yebes Mark5B. We were using the upper ethernet connector (Intel) which never attains 1 Gb/s. The setup was modified to use the lower ethernet connector (Marvel). The test was repeated and the result was this time succesful: 903 Mb/s i
- 900 Mb/s with MTU = 9000 bytes. This test showed 8 % fragmentation. This required setting all intermediate equipment to be able to use MTUs equal or larger than 9000 bytes.

The Mark5B ethernet interface was modified to accept 9000 bytes packet size:

```
ifconfig eth0 mtu 9000
```

This instruction forces the host to use 9000 bytes for outbound traffic but not for inbound traffic. For outgoing packages equipment in the rest of the way, mainly switches, may split the packets to smaller fragments of 1500 bytes if they disallow big packet sizes.

The 3500yl Procurve was configured to accept large packets on the Mark5B port VLAN and on the Internet port VLAN. Jumbo frames are enabled for all ports associated to a given VLAN since it is not possible to allow them on individual ports:

```

conf term
VLAN 72
jumbo
exit

```

Rediris was contacted and asked for to increase the packet size up to 9000 bytes in all routers used to connect OAN's network with VLBI JIVE recorders.

Some tests using ping with special options were performed by RedIris to check if the whole route between OAN's Yebes router and the JIVE target host was prepared to use an MTU of 9000 bytes after the modified their configuration. For completeness we include here their tests. There are three types:

```

EB-IRIS4.rel> ping 145.145.2.74 do-not-fragment size 8800
PING 145.145.2.74 (145.145.2.74): 8800 data bytes
36 bytes from ae2.500.jnr01.asd001a.surf.net (145.145.80.78): frag
needed and DF set (MTU 1500)
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 227c d480 2 0000 3b 01 fac6 62.40.124.54 145.145.2.74

36 bytes from ae2.500.jnr01.asd001a.surf.net (145.145.80.78): frag
needed and DF set (MTU 1500)
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 227c d4a3 2 0000 3b 01 faa3 62.40.124.54 145.145.2.74

36 bytes from ae2.500.jnr01.asd001a.surf.net (145.145.80.78): frag
needed and DF set (MTU 1500)
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 227c d4d1 2 0000 3b 01 fa75 62.40.124.54 145.145.2.74

36 bytes from ae2.500.jnr01.asd001a.surf.net (145.145.80.78): frag
needed and DF set (MTU 1500)
Vr HL TOS Len ID Flg off TTL Pro cks Src Dst
 4 5 00 227c d4fe 2 0000 3b 01 fa48 62.40.124.54 145.145.2.74

```

```

EB-IRIS4.rel> ping 145.145.80.78 do-not-fragment size 8800
PING 145.145.80.78 (145.145.80.78): 8800 data bytes
8808 bytes from 145.145.80.78: icmp_seq=0 ttl=59 time=47.003 ms
8808 bytes from 145.145.80.78: icmp_seq=1 ttl=59 time=48.234 ms
8808 bytes from 145.145.80.78: icmp_seq=2 ttl=59 time=46.746 ms
8808 bytes from 145.145.80.78: icmp_seq=3 ttl=59 time=47.182 ms

```

```

EB-Madrid0.re0> ping 145.145.80.78 do-not-fragment size 8800 source
130.206.214.1 (el /30 con vosotros)
PING 145.145.80.78 (145.145.80.78): 8800 data bytes
^C
--- 145.145.80.78 ping statistics ---
7 packets transmitted, 0 packets received, 100% packet loss

```

```

EB-Madrid0.re0> ping 145.145.80.78 do-not-fragment size 6000 source
130.206.214.1
PING 145.145.80.78 (145.145.80.78): 6000 data bytes

```



```

6008 bytes from 145.145.80.78: icmp_seq=0 ttl=58 time=43.204 ms
6008 bytes from 145.145.80.78: icmp_seq=1 ttl=58 time=43.159 ms
6008 bytes from 145.145.80.78: icmp_seq=2 ttl=58 time=43.160 ms
6008 bytes from 145.145.80.78: icmp_seq=3 ttl=58 time=43.142 ms
^C
--- 145.145.80.78 ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
round-trip min/avg/max/stddev = 43.142/43.166/43.204/0.023 ms

```

These tests seemed to be unreliable because ICMP filters show incorrect results.

New tests were done from Mark5B using tool `tracpath`. This Linux application shows the route from one host to another while showing the MTU size used along the way. Results were successful:

```

tracpath 192.42.120.66
 1: mark5b.oan.es (193.146.252.31)           0.116ms pmtu 9000
 1: jucar.oan.es (193.146.252.1)             0.518ms
 2: 130.206.214.1 (130.206.214.1)           asymm 3    2.132ms
 3: MAD.XE7-0-0.EB-IRIS4.red.rediris.es (130.206.250.21) asymm 4    3.838ms
 4: rediris.rtl.mad.es.geant2.net (62.40.124.53) asymm 5    2.242ms
 5: so-7-2-0.rtl.gen.ch.geant2.net (62.40.112.25) asymm 6   24.249ms
 6: so-3-3-0.rtl.fra.de.geant2.net (62.40.112.70) asymm 7   32.380ms
 7: so-5-0-0.rtl.ams.nl.geant2.net (62.40.112.58) asymm 8   40.495ms
 8: surfnet-gw.rtl.ams.nl.geant2.net (62.40.124.158) asymm 9   39.882ms
 9: ae2.500.jnr01.asd001a.surf.net (145.145.80.78) asymm 10  39.811ms
10: E8K02.Asd001A.netherlight.net (145.145.19.162) 40.551ms
11: no reply
12: 192.42.120.66 (192.42.120.66)           41.921ms reached
Resume: pmtu 9000 hops 12 back 12

```

From the previous log the route from the OAN mark5B equipment to the correlator is seen. Communication between Spain in The Netherlands is done via Switzerland and Germany. A total of 12 jumps are performed with a total elapsed time of 42 ms. Maximum allowed MTU was 9000 bytes along the whole route.

Fragmentation along the line for different packet sizes was investigated by increasing the packet size in small intervals using "iperf" and a speed of 900 Mbits/s. We tried 8000, 8900, 8960 and 8972 bytes. The latter value was the maximum one allowed to avoid fragmentation.

Final test was 950 Mbits/s during 5 minutes and at the same time normal traffic from the OAN to Internet was flowing. Command at the Yebes' Mark5b was:

```
iperf -c 192.42.120.34 -u -l 8972 -b 950M -t 300
```

which resulted in a total number of 33.4 Gbytes at 957 Mbits/s with a delay of 0.003 ms, and a 1 packet lost over 4000000 which is $2,5 \cdot 10^{-5} \%$. This is an excellent result which proves that the connection between Yebes' Mark5B and JIVE Mark5B is high quality and ready for 950 Mb/s.

Below we summarize hardware and software requirements required to run the connection at 950 Mbits/s with a large MTU:

- Use the lower ethernet connector. Model: Marvel
- MTU size at ethernet device is 9000 bytes

- MTU size at Yebes's router is Jumbo.
- MTU size at GEÁNT2, JIVE and RedIris set to 9000 bytes.
- Cleaning of the optical patchords at Telefónica's optical network.
- Add a public DSA key from JIVE in oper's home directory, subdirectory .ssh.
- Authorize ssh connections (port 22) from JIVE and on port 2620 at OAN's Procurve 3500yl router. The first allows JIVE operators to login into the oper account. This is rarely needed. The second authorization allows to open a dedicated connection to Mark5B JIVE control program. The client runs on JIVE and extracts all data from OAN's Mark5B recorder.

5.4. Note on 1024 Mbits/s data flux

Data flux from Mark5Bs or earlier recorders is limited to powers of 2. Usually stations have got 8 analog BBCs. It is possible to use both, USB (Upper Side Band) and LSB (Lower Side Band), for each BBC with a maximum bandwidth of 16 MHz each. This sums up to 256 MHz bandwidth. Sampling is done at twice the speed and data is coded with 2 bits. This gives a total of 1024 Mbits/s to transfer a total bandwidth of 256 MHz.

In order to transfer 1024 Mb/s it is necessary to use two ethernet cards at each Mark5B or one 10 Gb/s card. If two cards are used, data need to be multiplexed and sent through two Gigaethernet cards. In principle it would be possible to drop one 16 MHz band from the previous setup. That means 64 Mbits/s, which subtracted to 1024 is 960 Mbits/s which fits in 1 Gigaethernet connection. This solution is not operative yet and is being worked at JIVE but loads heavily the CPU at both ends: the correlator and the recorder.

Future recorders will not need to use powers of 2 and this limitation will disappear. This will provide the opportunity to use more efficiently the Gigaethernet lines between telescopes and the correlator.

6. First eVLBI observation

The first real eVLBI observation with the 40m radiotelescope was performed on April 3rd 2009 as part of the 100 hours of Astronomy in the IYA (International Year of Astronomy). It was a demo that lasted 7 hours approximately at C Band (5 GHz). 9 telescopes took part: Mopra, Sheshan, Jodrell Bank, Cambridge, Torun, Onsala, Medicina, Westerbok and Yebes. The first one, being in Australia, stopped early and the second one, in China, after some hours. The target source set below their horizon earlier than in european countries. Two quasars were observed: xxx and 3C120. The first one was used for clock search and the second one was used for the real observation.

Some telescopes transferred data at 512 Mb/s and others at 1024 Mb/s depending on their line capabilities. Figure 6 shows a plot in which the data flux rate summed up for all stations is shown as a function of time. Data flux at 512 Mb/s from Yebes is in fucsia color, and was steady along the whole experiment, which proves that the line worked correctly.

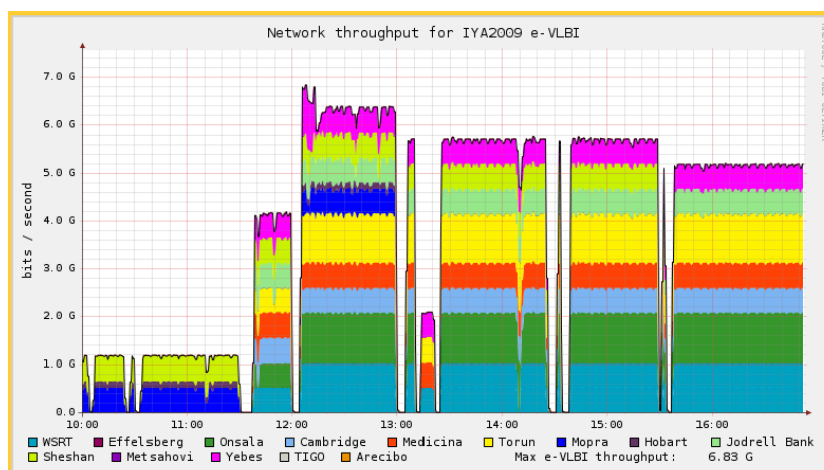


Figura 6: Flux rate from all telescopes as a function of time. Each telescope is assigned a color. Yebes is in the top of the graph and was assigned fuchsia color. The valleys are probably due to the correlator, that failed for short intervals along the experiment.

The demo was setup to observe 4 contiguous frequencies at both LCP and RCP. 8 BBCs were used, each with two 16 MHz bands (Upper Side Band and Lower side Band). That makes 16 subbands of 16 MHz each and hence 256 MHz (128 MHz at each polarization). Since sampling has to be done at double speed that makes 512 Mbits/s and since the signal is coded using 2 bits the total flux required is 1024 Mb/s. Stations which only can transfer 512 Mb/s only used 4 BBCs and dropped the other 4 BBCs.

Fringes with Yebes were found along the experiment for 8 subbands. Figures 7 and 8 show the correlation signal for all subbands at both polarizations as a function of delay step. Each subband is coded with different colors.

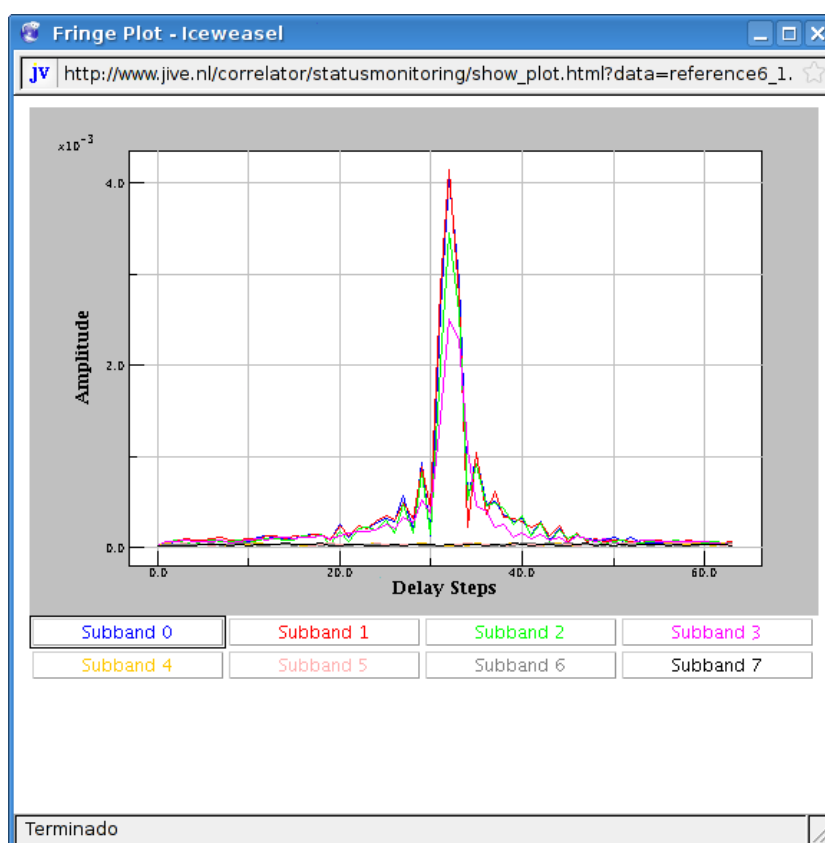


Figura 7: Cross correlation for each subband against Jodrell Bank for RCP. Each subband is assigned a color. Amplitude is in arbitrary units and X axis depicts delay in steps. Each step in time is the inverse of each subband in MHz (62.5 ns).

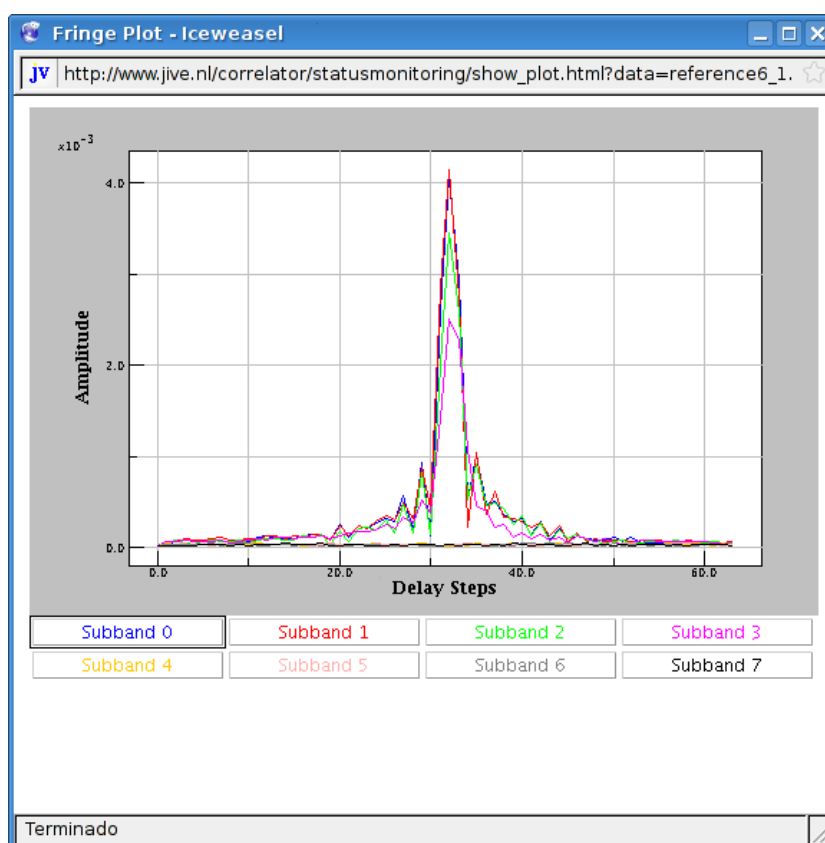


Figura 8: Cross correlation for each subband against Jodrell Bank for LCP. Each subband is assigned a color. Amplitude is in arbitrary units and X axis depicts delay in steps. Each step in time is the inverse of each subband in MHz (62.5 ns).