Unas experiencias con el intercambio de datos por GRID en el marco de la colaboración con el experimento ALICE en el Large Hadron Collider en CERN.-

L.Diaz,L. Nellen, G.Paic

Instituto de ciencias Nucleares

UNAM

CERN

We're in a particular position because we are the only discipline-that will live or die with the Grid (Carminati)

- CERN, as the collection point for data, is the Tier-0 center of the LHC computing grid.
- distributes the data to 11 Tier-1 centers around the world.
- The amount of data is on a much larger scale than most people ever encounter.
  - More than 10 petabytes (PB) of data will be collected each year by the experiments at CERN. For comparison, one gigabyte (GB) of data, a little more than can be stored on a CD, is one thousandth of a terabyte(TB), which in turn is one thousandth of a petabyte.

Nodes features

- ICN-UNAM

citli.nucleares.unam.mx

Scientific Linux 4.5

kernel: 2.6.9-55.ELsmp

Memory: 1GB

Firewall: Outside ICN firewall

- CERN

hufsa.cern.ch

Debian

kernel:2.6.18-4-686 SMP

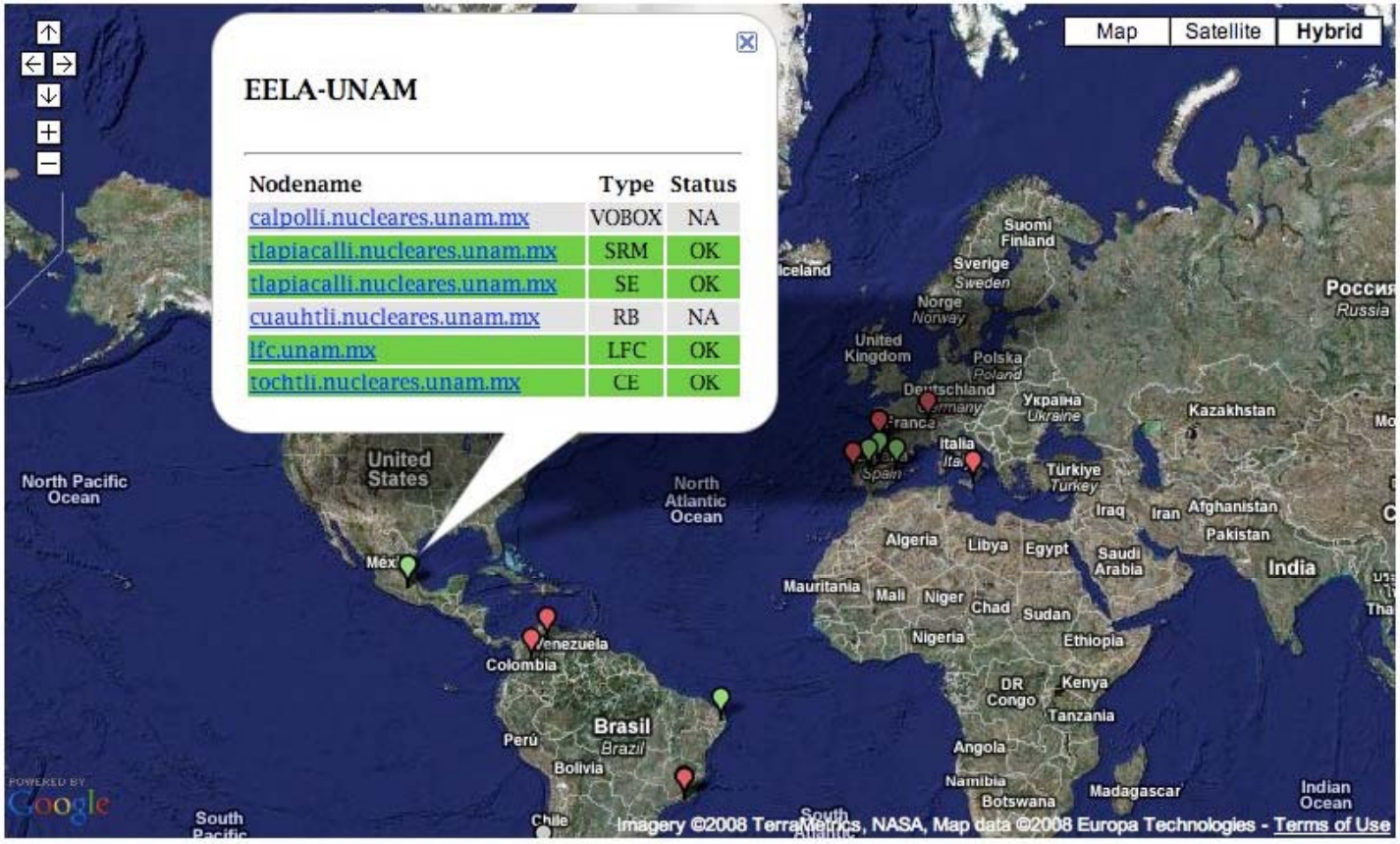
Memory: 1GB

Firewall: Outside CERN firewall

Tochtli is alive and well with all services

Last update: Thu Mar 6 13:01:17 UTC 2008

Updated hourly based on the Service Availability Monitoring (SAM) results.



Site Ok Site with problems Test results not available.

Comments and/or suggestions: Alexandre.Duarte@cern.ch

- Network tests between ICN - CERN
- ICN ALICE grid

TCP/IP settings

- ICN

/proc/sys/net/core/wmem_max
734003200
/proc/sys/net/core/rmem_max
734003200
/proc/sys/net/ipv4/tcp_rmem
4096 7340032 734003200
/proc/sys/net/ipv4/tcp_wmem
4096 7340032 734003200

- CERN

/proc/sys/net/core/wmem_max
734003200
/proc/sys/net/core/rmem_max
734003200
/proc/sys/net/ipv4/tcp_rmem
4096 7340032 734003200
/proc/sys/net/ipv4/tcp_wmem
4096 7340032 734003200

Test configuration

- Tool: Iperf
- TCP Window size: 256, 512, 1024 KB
- Number of threads: 1,4,8,16 simultaneus connections

Threads	TCP Window size (KB)			Time
1	256	512	1024	15min
4	256	512	1024	15min
8	256	512	1024	15min
16	256	512	1024	15min
			Total:	1hr

Number of tests: 24

- Total time : 1 day
- Test performed in both directions at the same time but not synchronized

TEST ID: 5

DATE: Tue Mar 4 12:24:01 CET 2008

Test performed after Opera Oberta transmission, network routing through **Public Internet**

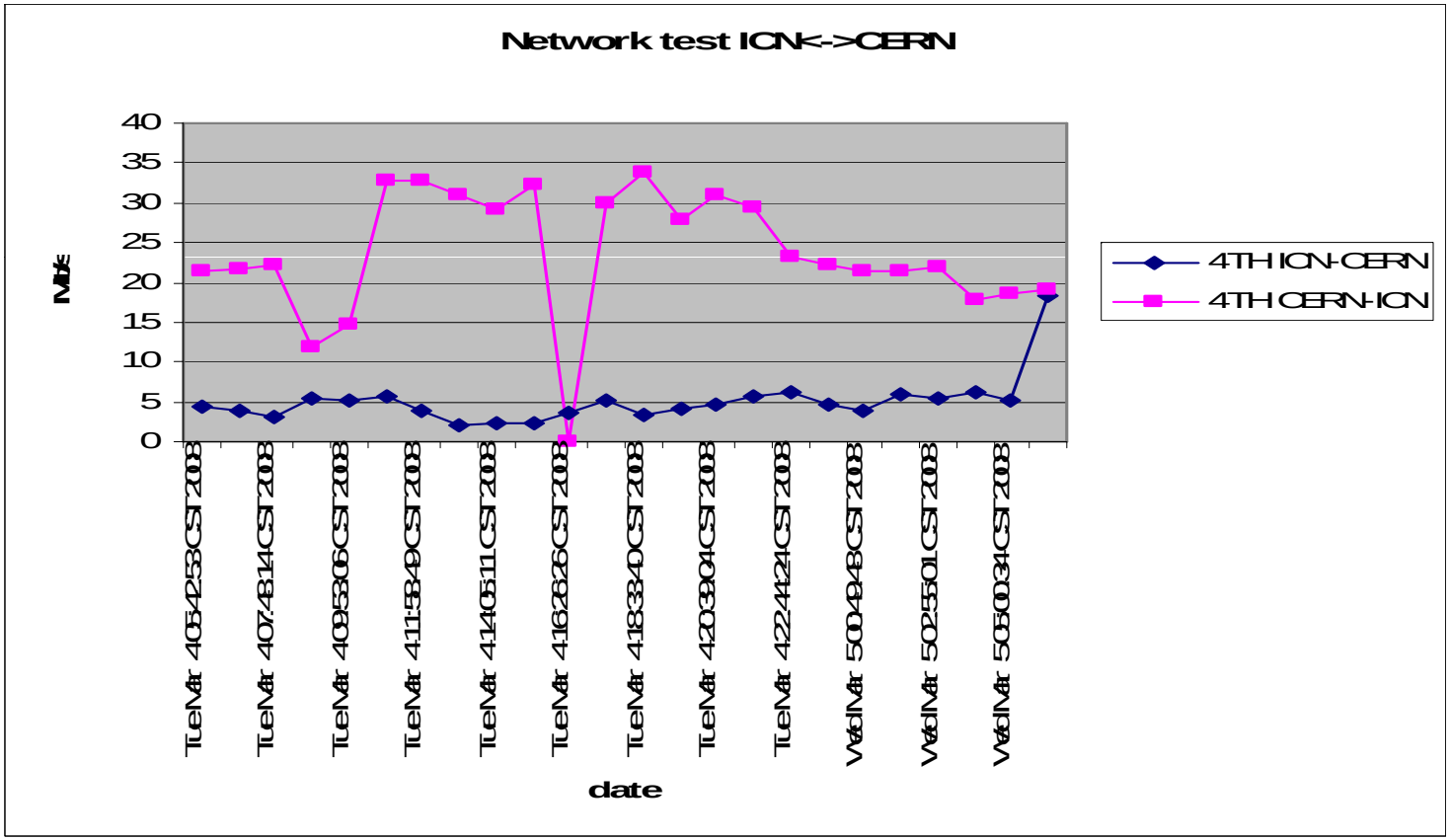
ICN->CERN network route

```
traceroute to hufsa.cern.ch (192.91.244.189), 30 hops max, 38 byte packets
 1  132.247.250.190 (132.247.250.190)  0.349 ms  0.297 ms  0.283 ms
 2  132.247.251.5 (132.247.251.5)  0.260 ms  0.240 ms  0.224 ms
 3  132.247.251.193 (132.247.251.193)  0.327 ms  0.289 ms  0.259 ms
 4  132.247.251.202 (132.247.251.202)  0.417 ms  0.358 ms  0.336 ms
 5  132.247.255.221 (132.247.255.221)  0.494 ms  0.420 ms  0.430 ms
 6  * * *
 7  * * *
 8  * * *
 9  * * *
10  * * *
11  * * *
12  * * *
13  e513-e-rci76-2-te8.cern.ch (192.91.246.110)  219.900 ms  219.297 ms  219.984 ms
14  e513-e-rci65-3-ne4.cern.ch (192.65.184.38)  219.726 ms  219.759 ms  219.576 ms
15  hufsa.cern.ch (192.91.244.189)  219.620 ms  220.471 ms  219.628 ms
```

CERN->ICN network route

```
traceroute to citli.nucleares.unam.mx (132.248.29.126), 30 hops max, 40 byte packets
 1  e513-e-rci65-3-ia4 (192.91.244.129)  1.534 ms  155.351 ms  0.480 ms
 2  e513-e-rci76-1-ne2 (192.65.184.77)  0.241 ms  0.212 ms  0.239 ms
 3  e513-e-rci76-2-ne1 (192.65.184.54)  0.243 ms  0.221 ms  0.240 ms
 4  fe-0-3-0.400.ar3.ZRH1.gblx.net (208.48.236.117)  9.865 ms  9.198 ms  9.109 ms
 5  te7-2-0GE.ar4.LAX1.gblx.net (67.17.109.86)  180.771 ms  165.505 ms  166.653 ms
 6  64.213.78.22 (64.213.78.22)  173.024 ms  173.365 ms  172.896 ms
 7  bb-mex-nextengo-25-pos12-1.uninet.net.mx (200.38.193.2)  214.254 ms  214.199 ms  214.494
ms
 8  * * *
 9  customer-201-117-71-133.uninet-ide.com.mx (201.117.71.133)  219.521 ms  219.945 ms
219.991 ms
10  132.247.251.233 (132.247.251.233)  219.740 ms  219.721 ms  219.493 ms
11  132.247.251.194 (132.247.251.194)  219.494 ms  220.223 ms  219.502 ms
12  132.247.251.6 (132.247.251.6)  219.484 ms  219.467 ms  219.495 ms
13  citli.nucleares.unam.mx (132.248.29.126)  219.125 ms  219.328 ms  219.494 ms
```

Graphic



TCP Windows size = 512KB, 4 threads per transfer

TEST ID: 6

DATE: Fri Mar 14 11:26:32 CST 2008

Network routing fixed to I2

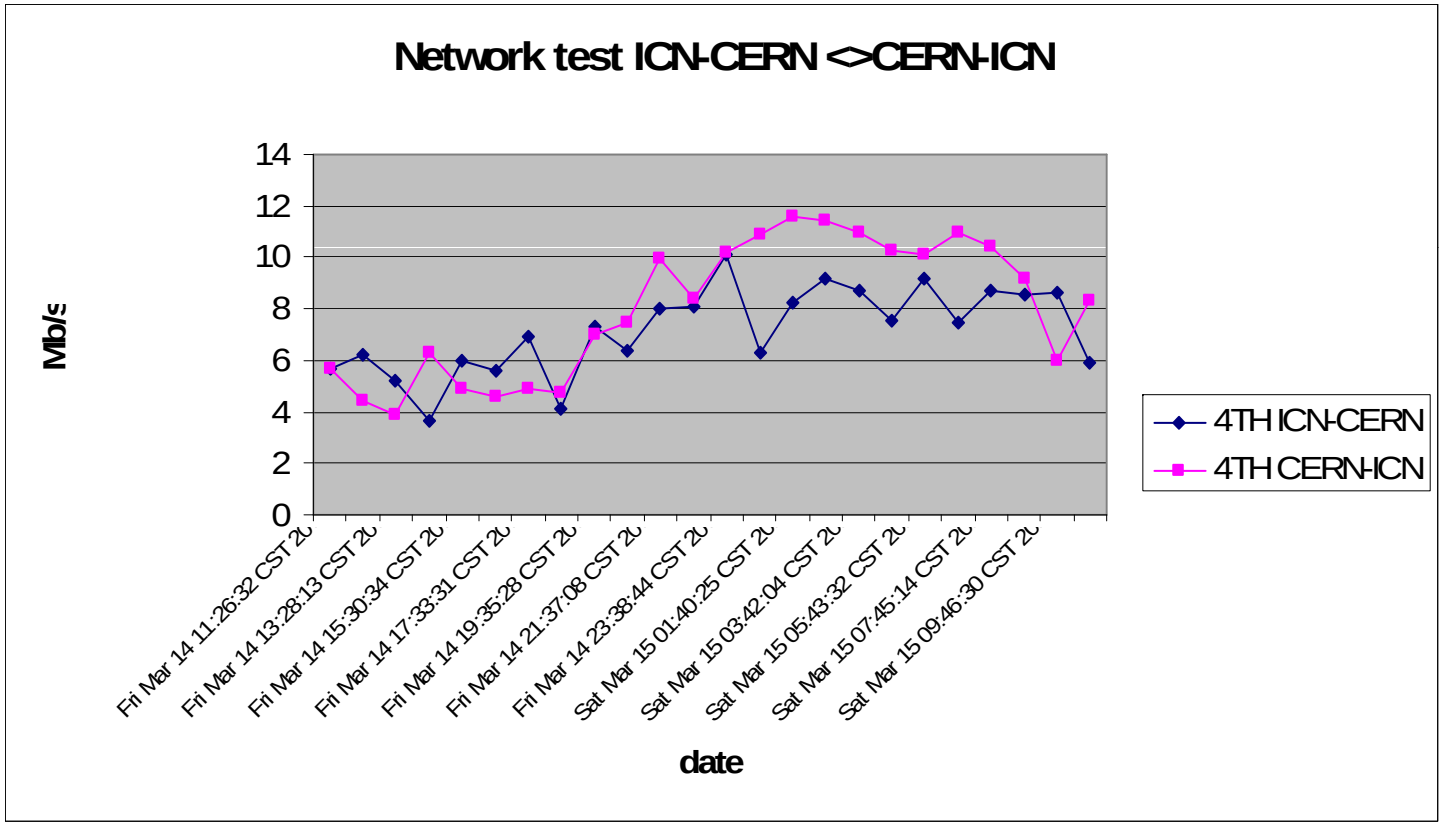
ICN->CERN network route

```
traceroute to hufsa.cern.ch (192.91.244.189), 30 hops max, 38 byte packets

 1  132.247.250.190 (132.247.250.190)  0.344 ms  0.306 ms  0.442 ms
 2  132.247.251.5 (132.247.251.5)  0.437 ms  0.250 ms  0.270 ms
 3  132.247.251.193 (132.247.251.193)  0.374 ms  0.303 ms  0.282 ms
 4  132.247.251.202 (132.247.251.202)  0.572 ms  0.365 ms  0.358 ms
 5  132.247.255.221 (132.247.255.221)  0.571 ms  0.537 ms  0.862 ms
 6  mexico7200-unam.core.cudi.edu.mx (200.23.60.193)  4.678 ms  3.647 ms  5.412 ms
 7  gdl-mexico7200.core.cudi.edu.mx (200.23.60.226)  14.798 ms  13.296 ms  18.457 ms
 8  tijuana-gdl.core.cudi.edu.mx (200.23.60.241)  40.263 ms  43.106 ms  42.749 ms
 9  nlr-1-is-jmb-776.sttlwa.pacificwave.net (207.231.241.14)  74.164 ms  72.899 ms  73.535
ms
10  denv-seat-58.layer3.nlr.net (216.24.186.6)  98.710 ms  99.493 ms  98.617 ms
11  chic-denv-36.layer3.nlr.net (216.24.186.4)  132.510 ms  132.020 ms  121.659 ms
12  216.24.184.22 (216.24.184.22)  129.169 ms  125.631 ms  127.194 ms
13  e513-e-rci76-2-te8.cern.ch (192.91.246.110)  221.046 ms  217.733 ms  218.038 ms
14  e513-e-rci65-3-ne4.cern.ch (192.65.184.38)  682.922 ms  217.793 ms  221.984 ms
15  hufsa.cern.ch (192.91.244.189)  218.194 ms  218.487 ms  221.224 ms
```

CERN->ICN network route

```
traceroute to citli.nucleares.unam.mx (132.248.29.126), 30 hops max, 40 byte packets

 1  e513-e-rci65-3-ia4 (192.91.244.129)  0.328 ms  0.175 ms  0.237 ms
 2  e513-e-rci76-1-ne2 (192.65.184.77)  0.867 ms  0.595 ms  27.478 ms
 3  e513-e-rci76-2-ne1 (192.65.184.54)  0.362 ms  0.223 ms  0.237 ms
 4  abilene-cern-s1 (192.91.246.125)  333.701 ms  314.797 ms  322.312 ms
 5  so-4-3-0.0.rtr.kans.net.internet2.edu (64.57.28.36)  139.795 ms  123.648 ms  123.797 ms
 6  so-0-1-0.0.rtr.hous.net.internet2.edu (64.57.28.57)  137.670 ms  137.648 ms  137.664 ms
 7  200.23.60.121 (200.23.60.121)  155.156 ms  155.129 ms  155.008 ms
 8  200.23.60.145 (200.23.60.145)  168.157 ms  168.193 ms  168.020 ms
 9  200.23.60.234 (200.23.60.234)  203.836 ms  204.222 ms  203.862 ms
10  200.23.60.225 (200.23.60.225)  215.377 ms  216.917 ms  215.734 ms
11  200.23.60.194 (200.23.60.194)  219.491 ms  221.998 ms  215.978 ms
12  ve77-zc-core.ge.unam.mx (132.247.255.222)  217.854 ms  218.925 ms  224.969 ms
13  132.247.251.201 (132.247.251.201)  219.599 ms  240.055 ms  216.708 ms
14  132.247.251.194 (132.247.251.194)  216.364 ms  219.937 ms  220.733 ms
15  132.247.251.6 (132.247.251.6)  222.212 ms  217.577 ms  226.208 ms
16  citli.nucleares.unam.mx (132.248.29.126)  228.492 ms  221.819 ms  222.714 ms
```

Graphic



TCP Windows size = 512KB, 4 threads per transfer

TEST ID: 8

DATE:Mon Mar 24 01:43:16 CST 2008

I2 Network routing

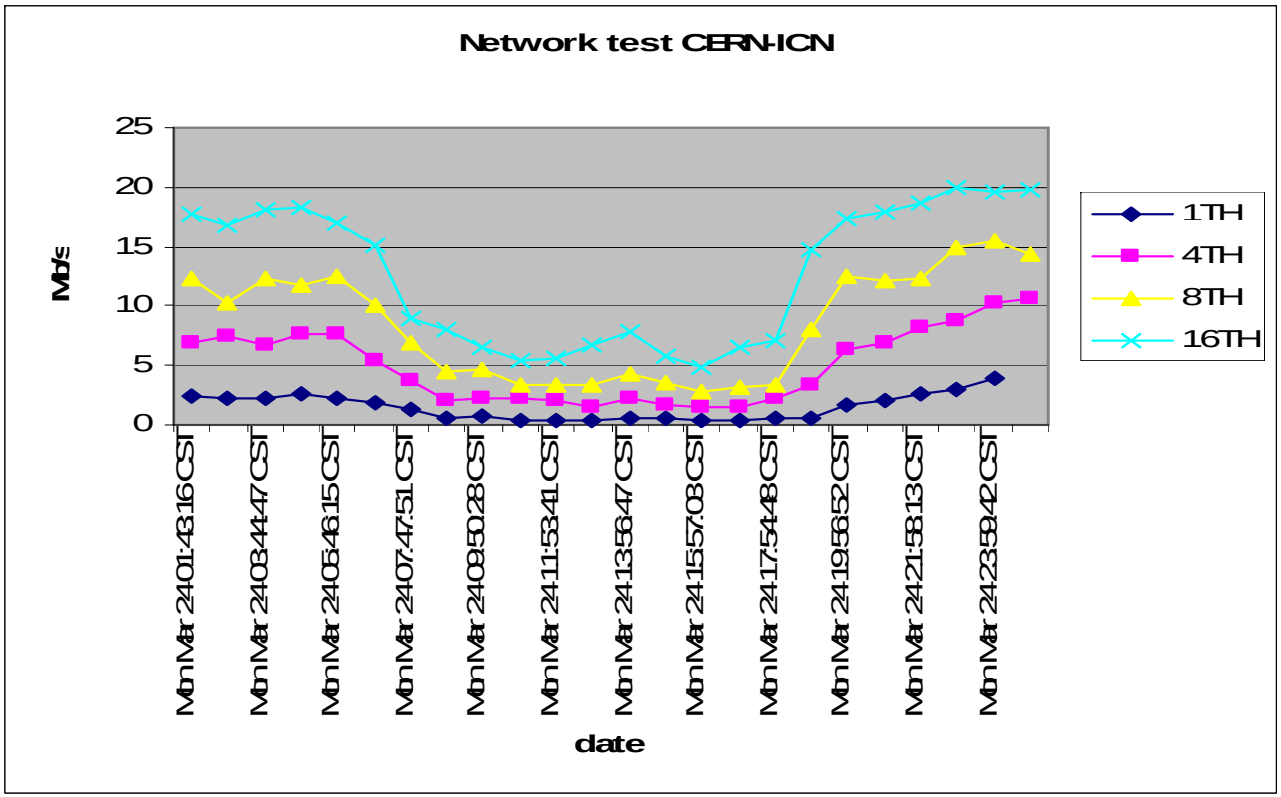
ICN->CERN network route

```
traceroute to hufsa.cern.ch (192.91.244.189), 30 hops max, 38 byte packets

 1  132.247.250.190 (132.247.250.190)  0.296 ms  0.258 ms  0.241 ms
 2  132.247.251.5 (132.247.251.5)  0.268 ms  0.213 ms  0.230 ms
 3  132.247.251.193 (132.247.251.193)  0.387 ms  0.340 ms  1.060 ms
 4  132.247.255.221 (132.247.255.221)  0.448 ms  0.384 ms  0.409 ms
 5  mexico7200-unam.core.cudi.edu.mx (200.23.60.193)  2.665 ms  2.494 ms  2.191 ms
 6  gdl-mexico7200.core.cudi.edu.mx (200.23.60.226)  14.942 ms  13.213 ms  14.941 ms
 7  tijuana-gdl.core.cudi.edu.mx (200.23.60.241)  39.731 ms  39.488 ms  40.703 ms
 8  nlr-1-is-jmb-776.sttlwa.pacificwave.net (207.231.241.14)  71.397 ms  71.040 ms  72.996 ms
 9  denv-seat-58.layer3.nlr.net (216.24.186.6)  97.617 ms  96.424 ms  96.689 ms
10  chic-denv-36.layer3.nlr.net (216.24.186.4)  122.087 ms  121.076 ms  121.169 ms
11  216.24.184.22 (216.24.184.22)  120.017 ms  126.744 ms  126.402 ms
12  e513-e-rci76-2-te8.cern.ch (192.91.246.110)  354.642 ms  397.858 ms  406.810 ms
13  e513-e-rci65-3-ne4.cern.ch (192.65.184.38)  346.720 ms  217.184 ms  216.316 ms
14  hufsa.cern.ch (192.91.244.189)  216.158 ms  216.170 ms  216.180 ms
```

CERN->ICN network route

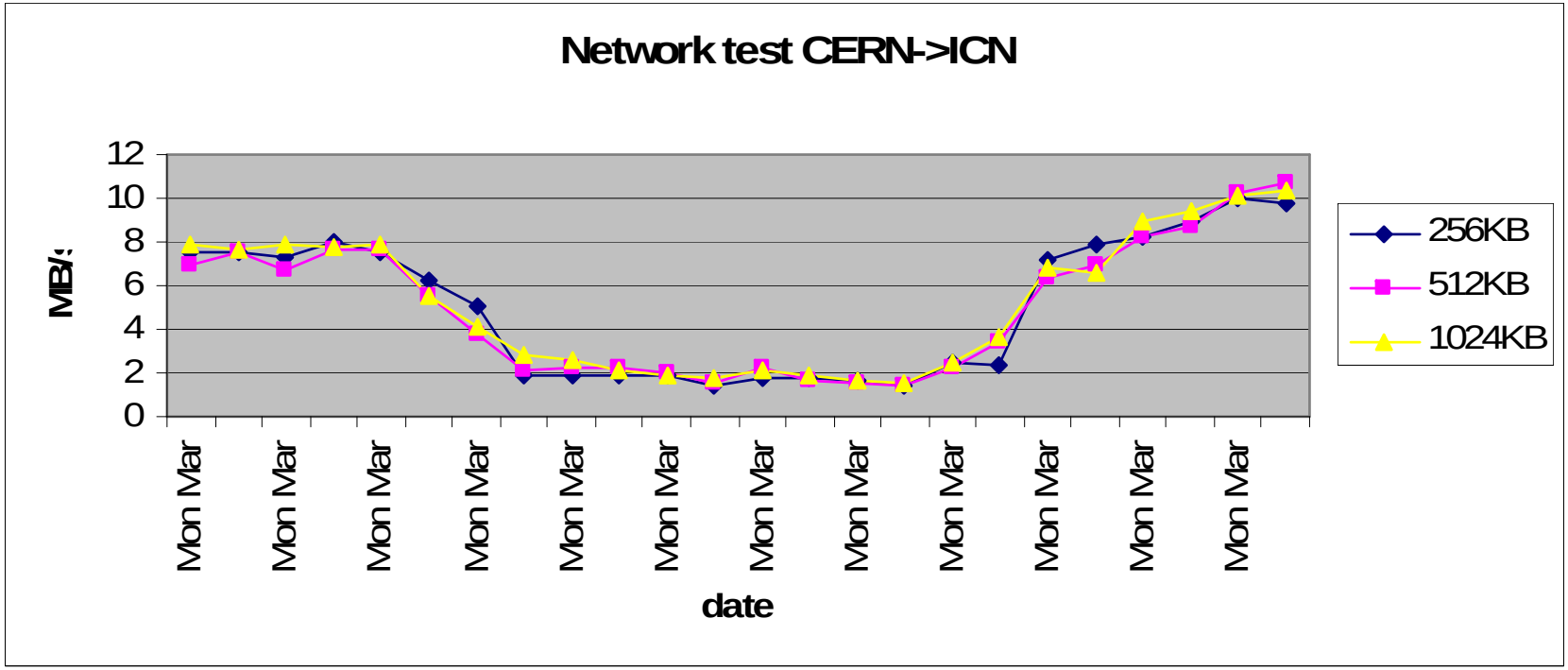
```
traceroute to citli.nucleares.unam.mx (132.248.29.126), 30 hops max, 40 byte packets

 1  e513-e-rci65-3-ia4 (192.91.244.129)  0.415 ms  0.283 ms  0.229 ms
 2  e513-e-rci76-1-ne2 (192.65.184.77)  0.369 ms  0.217 ms  0.239 ms
 3  e513-e-rci76-2-ne1 (192.65.184.54)  0.246 ms  0.203 ms  0.232 ms
 4  abilene-cern-s1 (192.91.246.125)  125.697 ms  113.501 ms  113.287 ms
 5  so-4-3-0.0.rtr.kans.net.internet2.edu (64.57.28.36)  130.519 ms  123.755 ms  123.641 ms
 6  so-0-1-0.0.rtr.hous.net.internet2.edu (64.57.28.57)  137.913 ms  137.856 ms  137.757 ms
 7  200.23.60.121 (200.23.60.121)  155.018 ms  155.159 ms  154.867 ms
 8  200.23.60.145 (200.23.60.145)  168.507 ms  168.023 ms  168.477 ms
 9  200.23.60.234 (200.23.60.234)  203.858 ms  203.551 ms  203.954 ms
10  200.23.60.225 (200.23.60.225)  215.348 ms  215.073 ms  215.089 ms
11  200.23.60.194 (200.23.60.194)  218.323 ms  220.834 ms  221.579 ms
12  ve77-zc-core.ge.unam.mx (132.247.255.222)  217.604 ms  219.708 ms  218.722 ms
13  132.247.251.194 (132.247.251.194)  222.079 ms  216.830 ms  216.847 ms
14  132.247.251.6 (132.247.251.6)  218.212 ms  222.436 ms  218.340 ms
15  citli.nucleares.unam.mx (132.248.29.126)  221.214 ms  226.941 ms  226.838 ms
```

Graphic, CERN->ICN



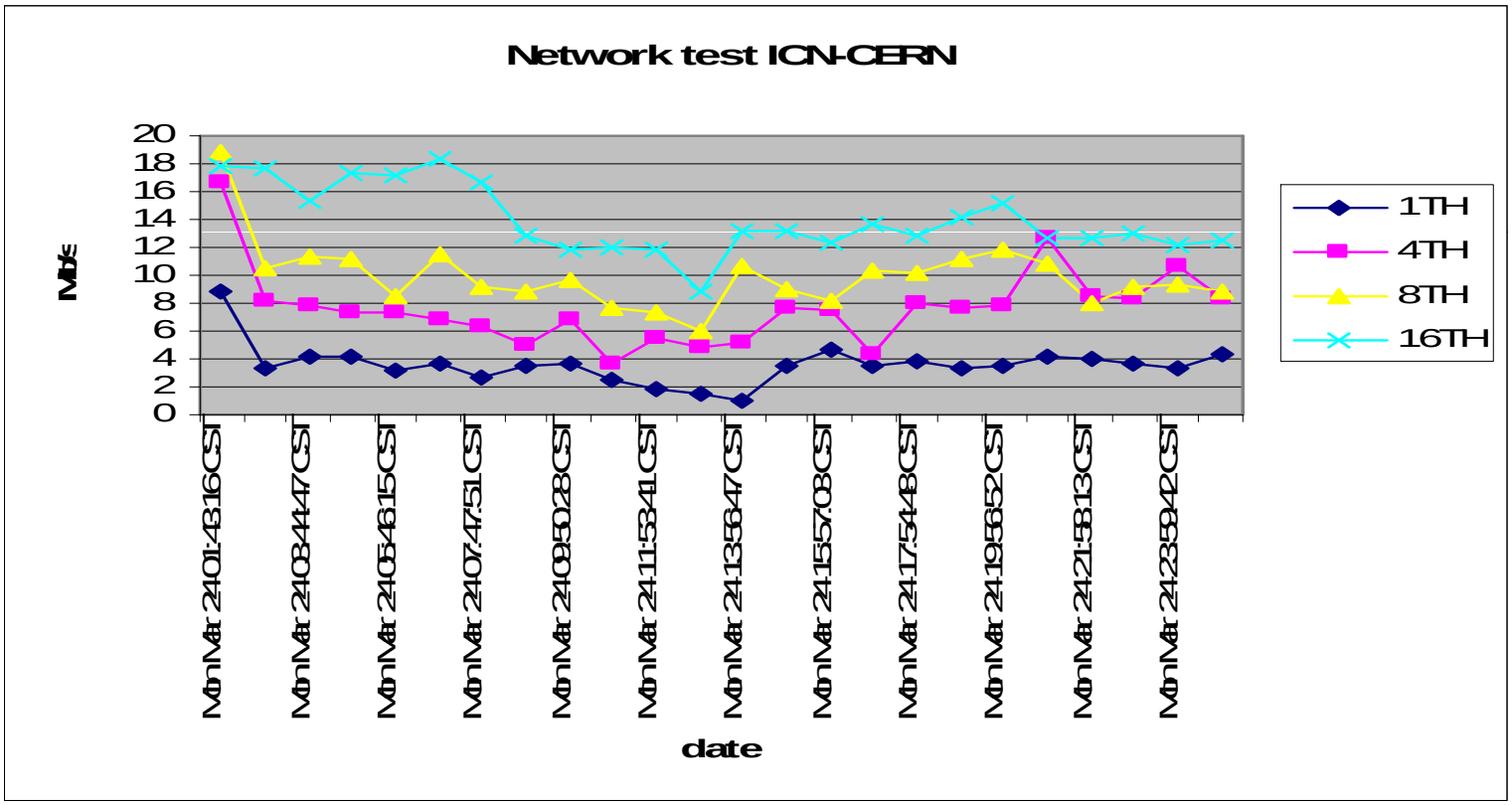
TCP Windows size = 512KB. 1,4,8,16 threads per transfer

Graphic, CERN->ICN



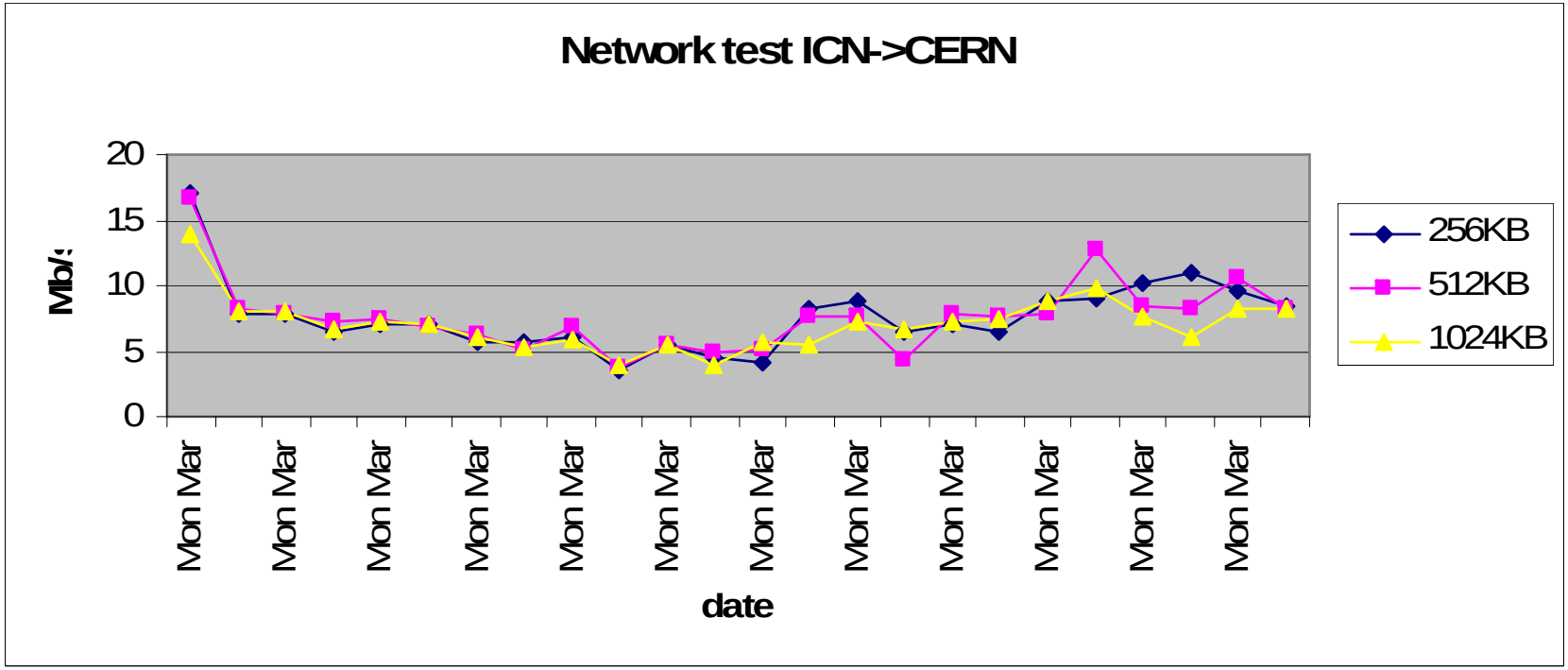
TCP Windows size = 256, 512, 1024KB. 4 threads per transfer

Graphic, ICN->CERN



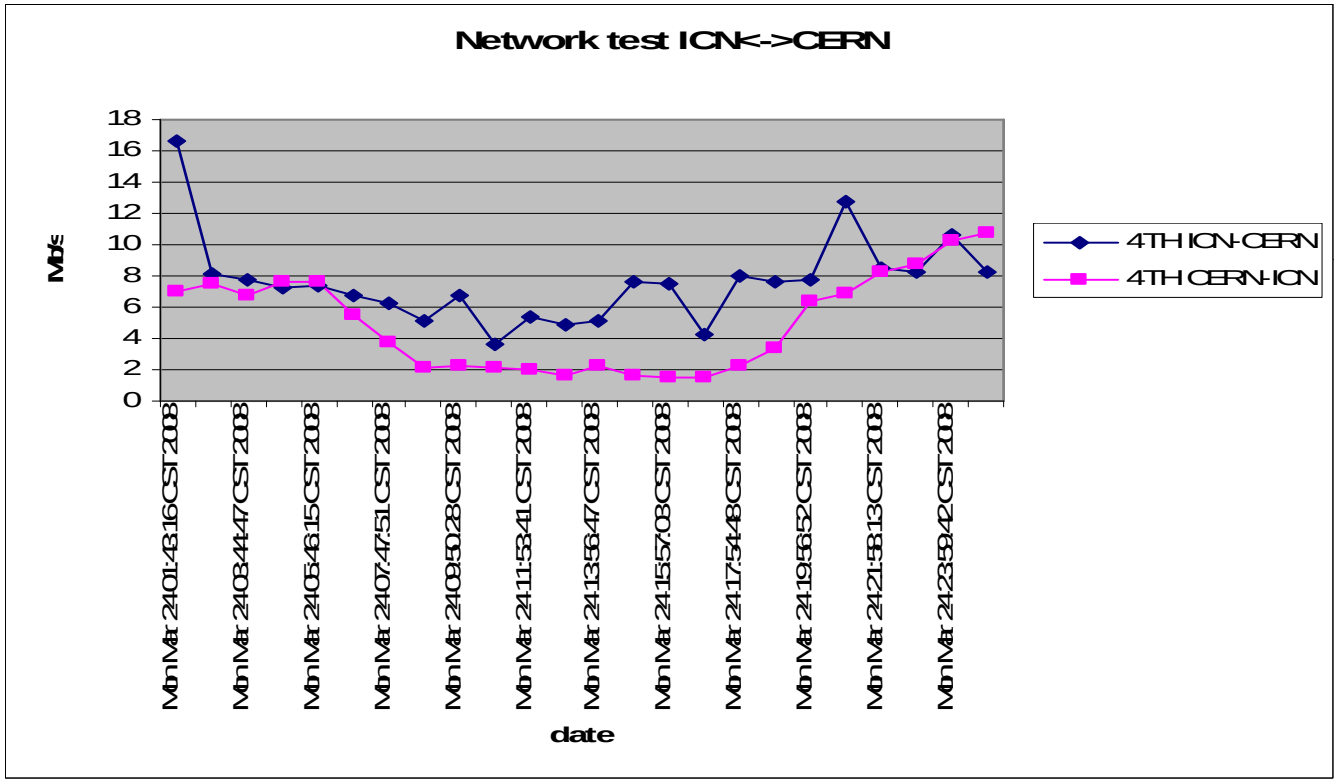
TCP Windows size = 512KB. 1,4,8,16 threads per transfer

Graphic, ICN->CERN



TCP Windows size = 256, 512, 1024 KB. 4 threads per transfer

Graphic, ICN<->CERN

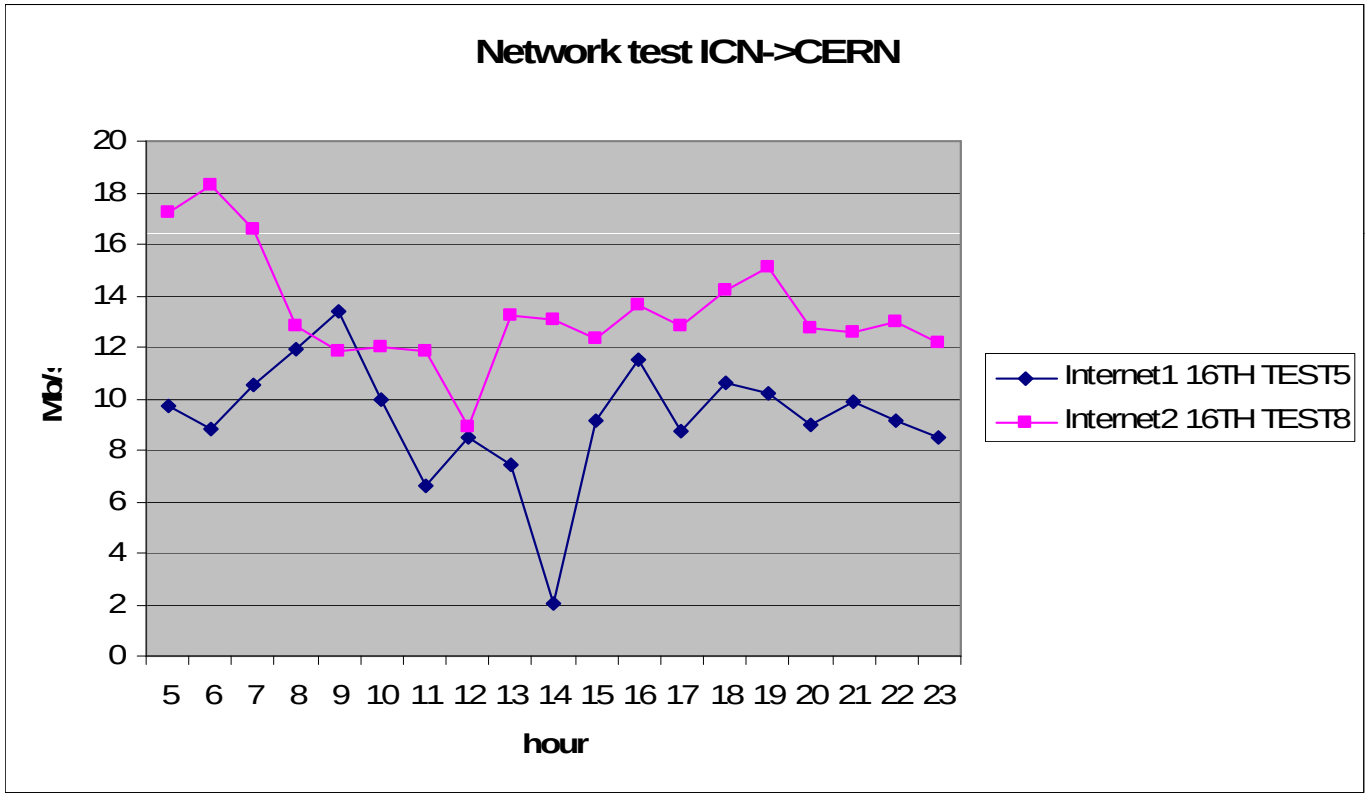


TCP Windows size = 512KB. 4 threads per transfer

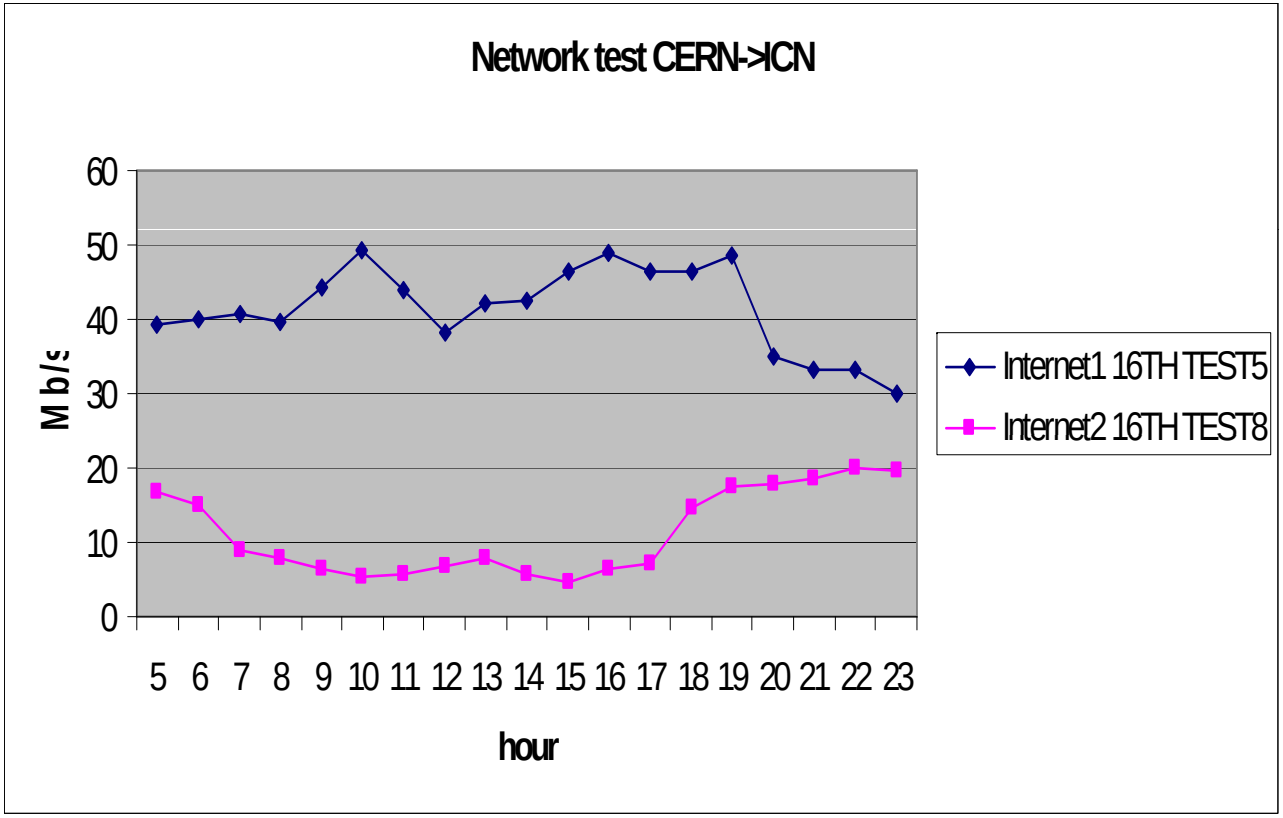
UNAM case:

Public Internet VS Internet 2 (CUDI)

- Data sources: test 5 and 8
- TCP Window Size = 512KB
- Number of threads 16, in order to get maximum BW
- UNAM Internet2 BW = 35 Mb/s
- UNAM Public Internet BW= 600 Mb/s



TCP Windows size = 512KB. 16 threads per transfer



TCP Windows size = 512KB. 16 threads per transfer

Conclusion

- Better performance with 4 threads per transfer.
- Network not tuned for large file transfers. From the computer to International links.
- Is Internet2 in Mexico (CUDI) ready for Grid??

ALICE node at ICN- UNAM

Current status

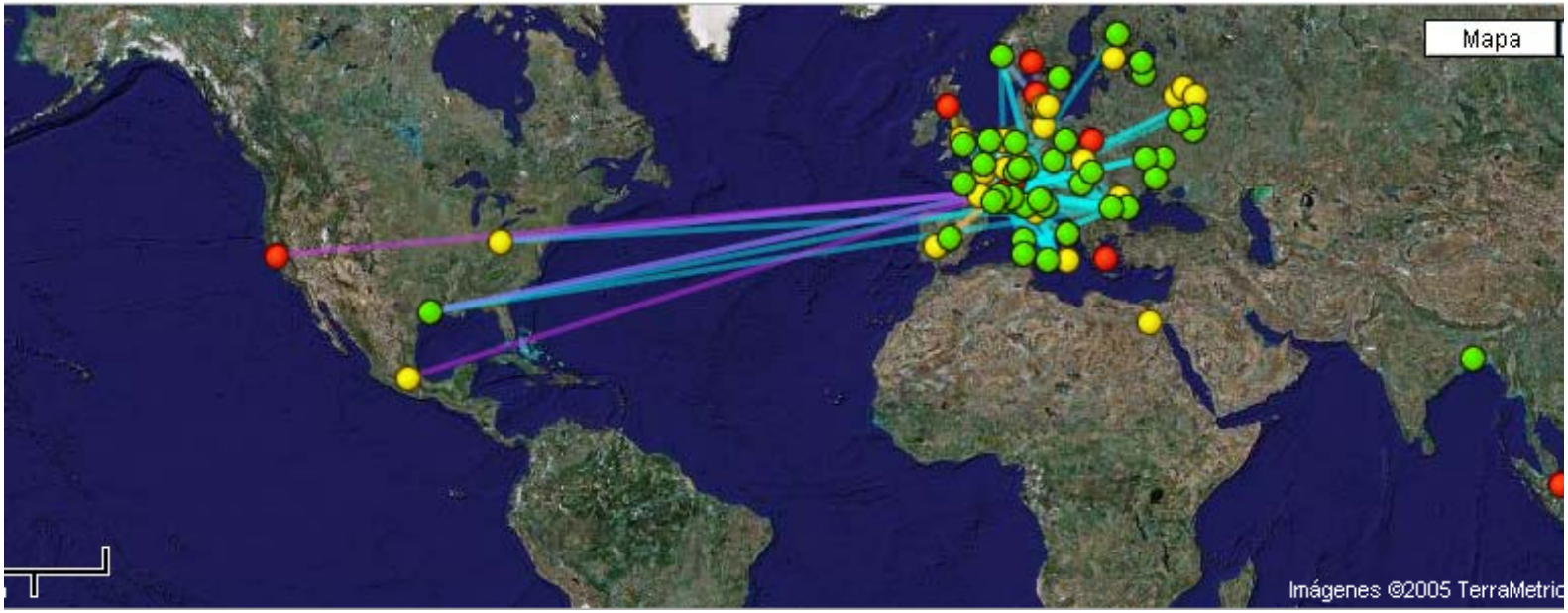
Resources

- Grid services installed
  - CE-Computer Element
  - SE-Storage Element
  - RB- Resource Broker
  - LFC –Local File Catalog
  - ALICE VoBOX
  - MonBOX- Monitoring BOX
  - UI-User interface
  - WN- Worker nodes
- CPU's
  - 14 WN's / 28 Xeon CPU's available
- Storage
  - 15 TB of storage just for ALICE

Resources shared

- The same resources are shared with EELA and AUGER projects.
- About storage: 15 TB are available just for EELA, AUGER and local users.

ALICE-MONALISA



http://pcalimonitor.cern.ch/map.jsp

Improvements necessary

Better communixcation of CUDI with users when special events: opera abierta

Inclusion of the possibility for jumbo frames transfers DGSCA-CUDI

Increase in the bandwidth (we cannot advance with the present BW and auxilliary services.

ICN BW required

Based on information from the ALICE data challenge a typical job in a Xeon 2.5GHz processor produce 800MB per 8 hours.

Using this information just the ICN cluster will consume 11 MB/s approximately. 1/3 of the BW available for UNAM Internet2.

Further plans

- Add CPU resources to the local site.
- Install new sites to the Grid infrastructure:
  - EELA- IBT-UNAM
  - ALICE, AUGER- BUAP
- Network testing between ICN and University of Houston

Conclusion

In order to be prepared for:

- The ALICE data processing/analysis.
- The addition/creation of new sites.
- The support of Grid users: EELA, ALICE, AUGER
- Storage demand
- The develop of new grid applications

Is necessary to improve the network connectivity before we stay “isolated” from the experiments, with available CPU’s and storage resources, a lot of users but without the possibility to connect the pieces: data experiment, resources, users.