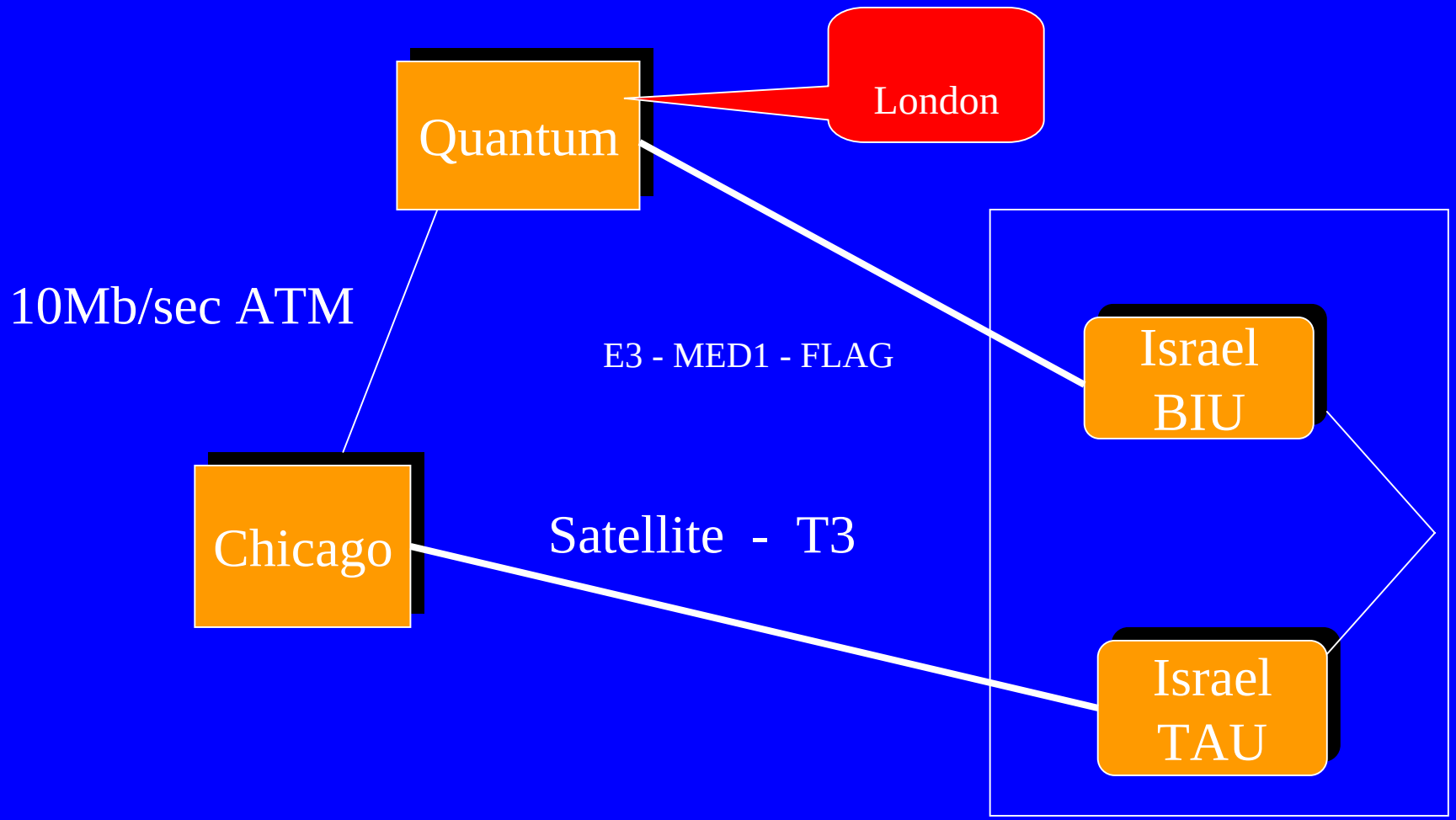
The Asymmetry of Internet-2

Hank Nussbacher

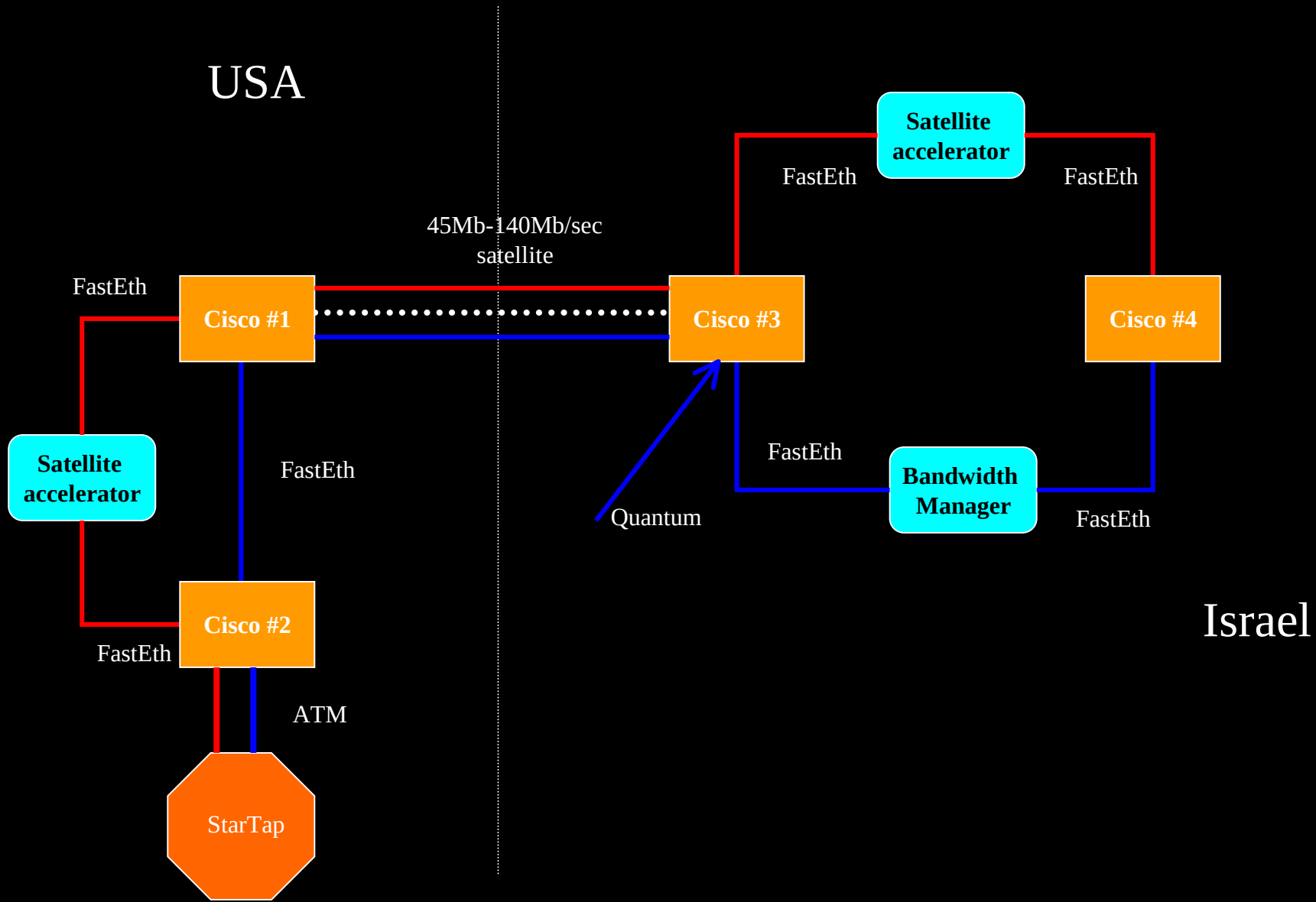
IUCC

March 2000

Israeli Internet-2 Design Summary



GigaPOP Design for Differentiated Services



Satellite issues and QoS

- **TCP streams are limited to 936kb/sec**
  - **RFC2488 - Enhancing TCP Over Satellite Channels using Standard Mechanisms**
    - **Path MTU - RFC1191**
    - **Large windows - RFC1323 (default is 64KB)**
    - **Large socket buffers - bandwidth*delay = 45Mb*600ms = 3.3Mbytes**
    - **TCP Selective Ack (SACK) - RFC2018**
- **UDP unaffected**

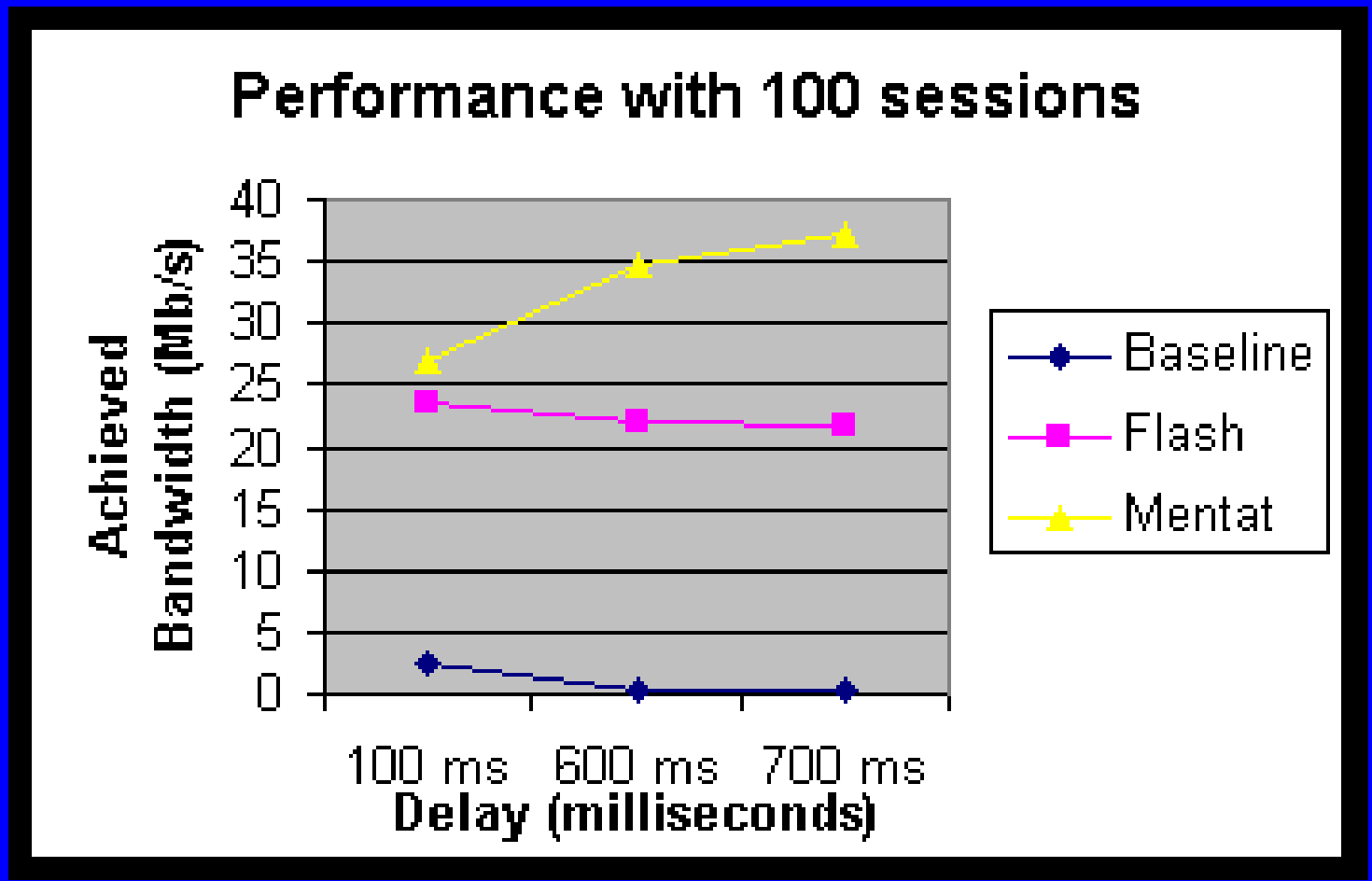
Satellite issues and QoS

- **Thruput = window size * RTT**
  - **64K / 560ms = 117,027 bytes/sec (936kb/sec)**
    - **64K is maximum default - W98 is 8K**
  - **1M * 30ms = 33Mb/sec (Abilene TCP limit?)**
- **Enabling High Performance Data Transfers**
  - **http://www.psc.edu/networking/perf_tune.html**

Satellite black box testing

- **Testing performed in April at Intelsat lab**
  - **Flash Networks and Mentat**
    - **results located at: www.internet-2.org.il/satellite-testing.html**

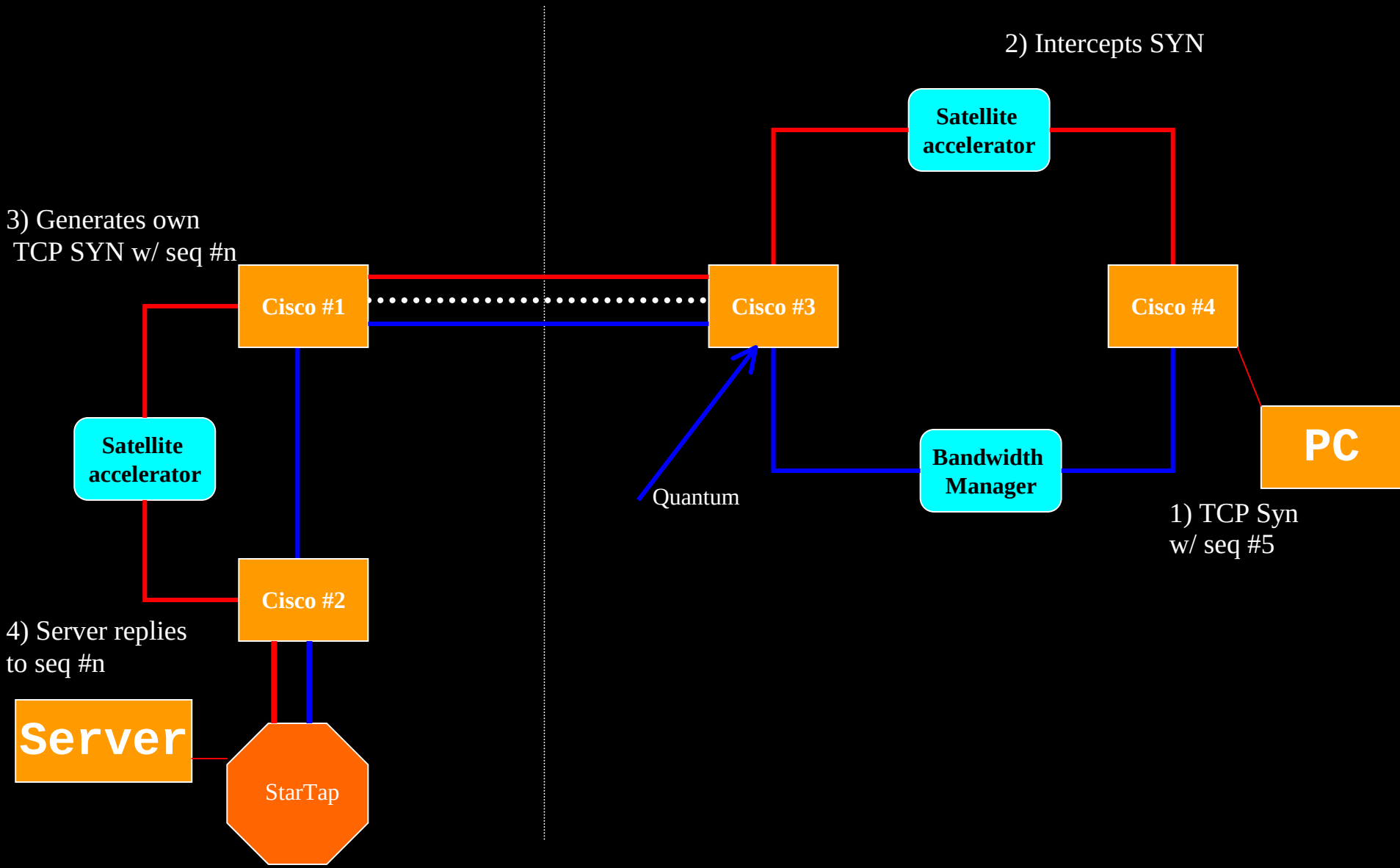
Satellite results



Mentat SkyX

- **Only affects TCP - UDP and ICMP is bypassed (as well as Ipsec)**
  - **requires symmetric routing**
- **Intercepts TCP connections and replaces it over satellite with “SkyX protocol”**
  - **uses NACKS to request again lost data packets**
  - **unlimited window size**
  - **no slow start over satellite link**
  - **streamlined TCP handshake on initial connection**
  - **TCP rate control over satellite link**

Asymmetry breaks Mentat SkyX boxes

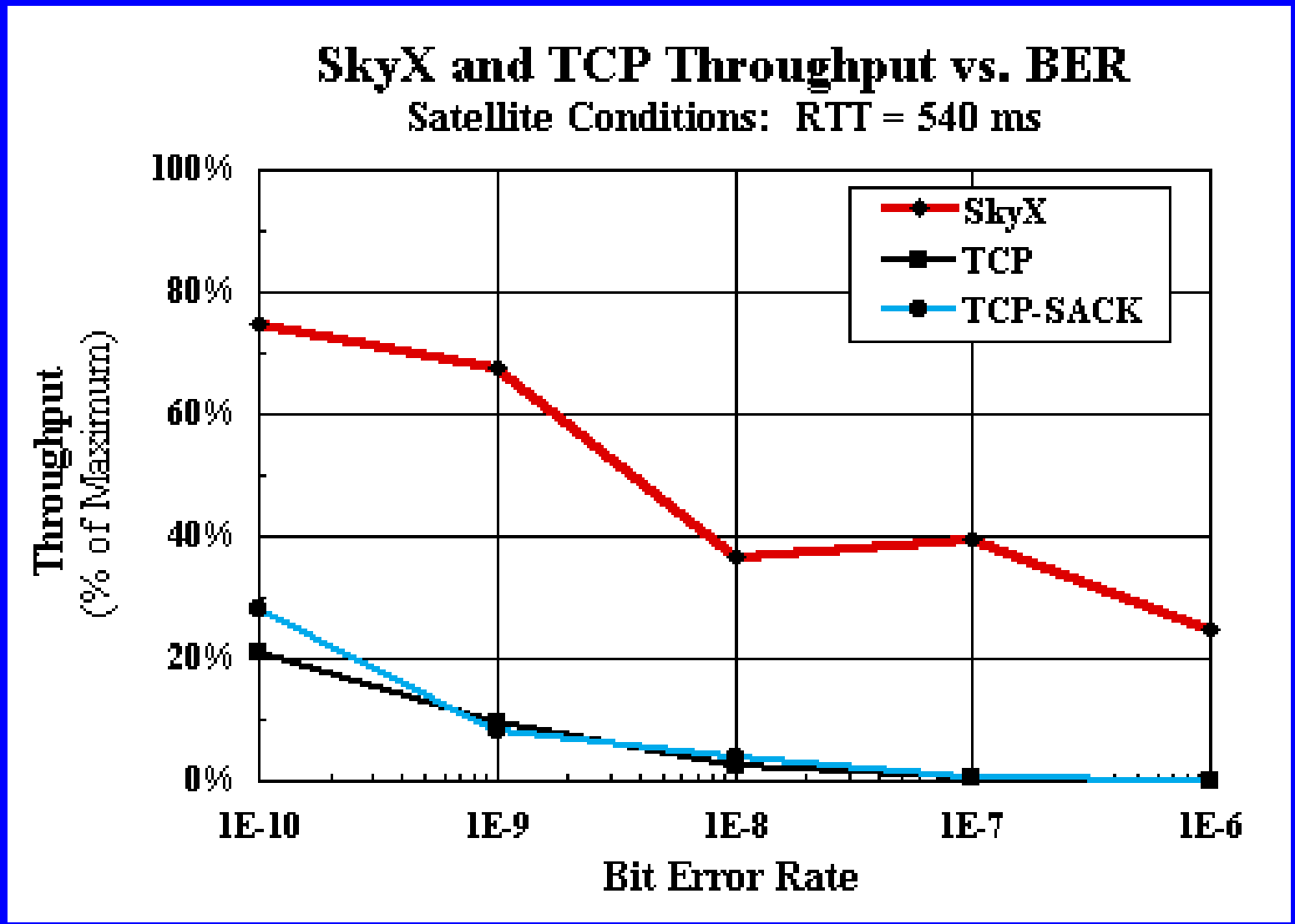


NASA testing

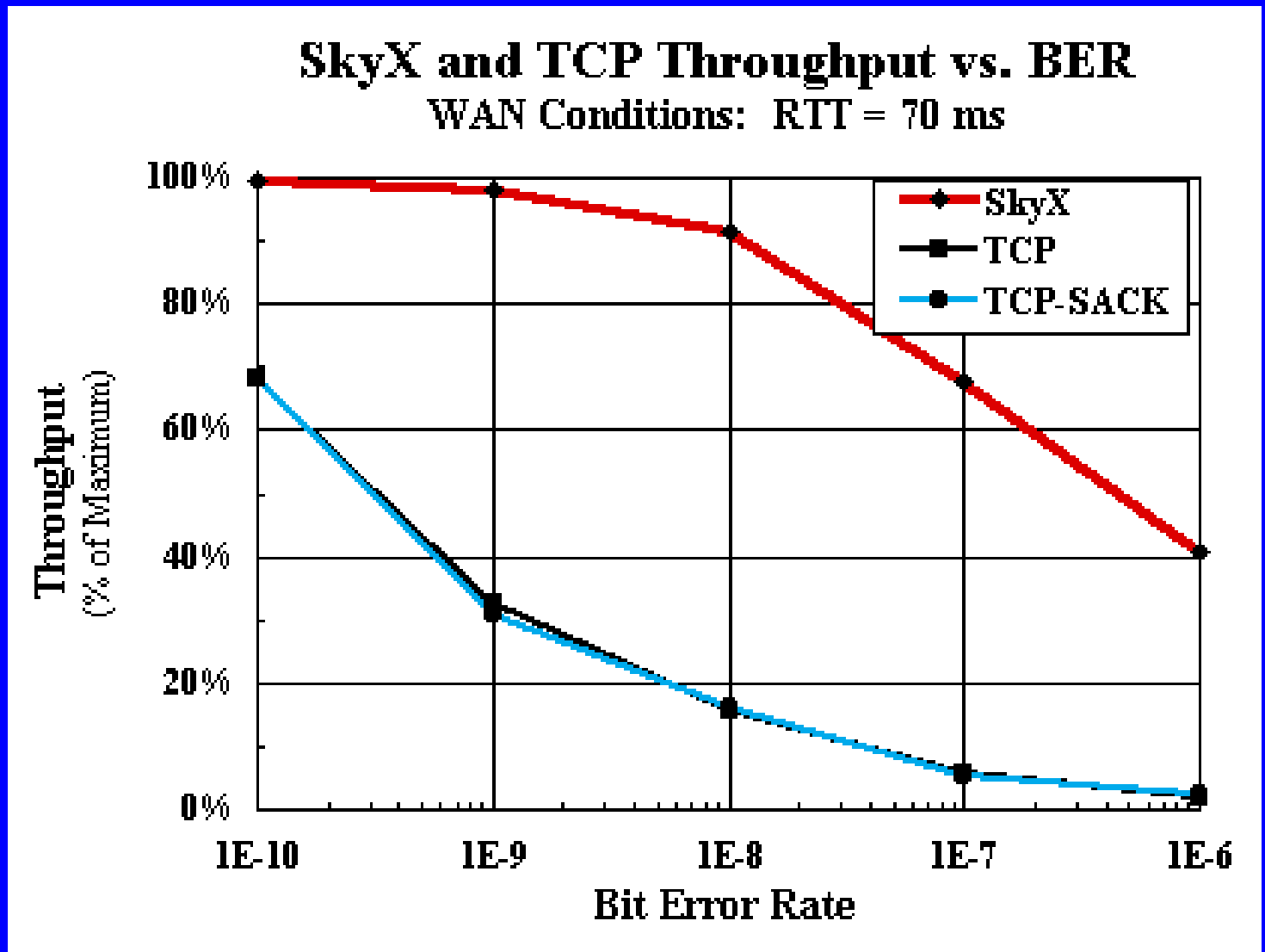
- **OC-3 speed testing in a lab**
- **Details located at:**

http://www.mentat.com/skyx/skyx-nasa.html

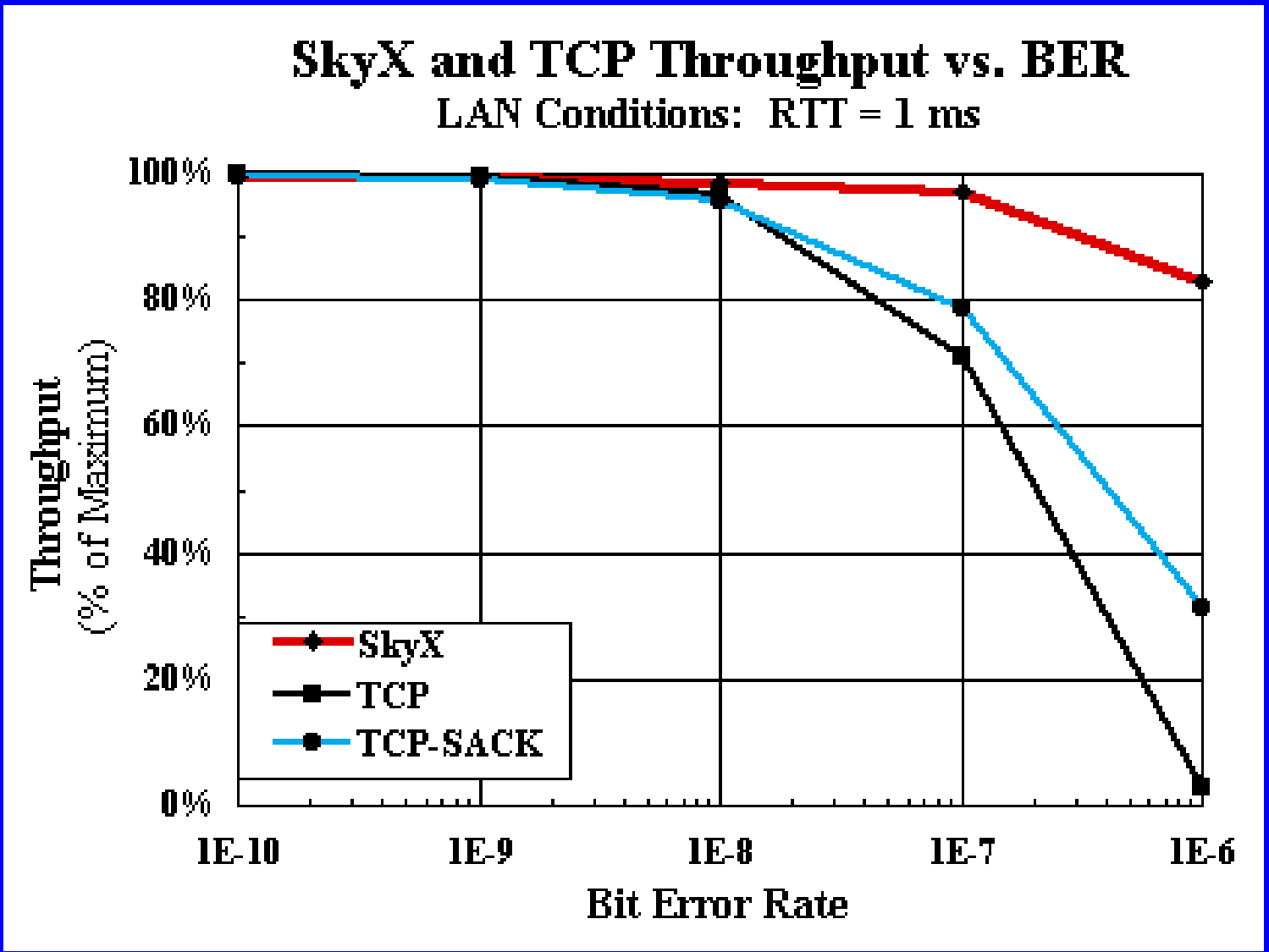
NASA tests of SkyX - #1



NASA tests of SkyX - #2



NASA tests of SkyX - #3



Israeli Mentat results (Jan 2000)

- **30Mb/sec pipe, iperf to U of Oregon**
- **No SkyX (560ms RTT)**
  - **8Kbyte TCP window- 118kbit/sec**
  - **64Kbyte TCP window - 646kbit/sec**
  - **500Kbyte TCP window - 2.9Mbit/sec**
- **With SkyX (560ms RTT)**
  - **8Kbyte TCP window - 19.5Mbit/sec**
  - **64Kbyte TCP window - 18.0Mbit/sec**
  - **500Kbyte TCP window - 18.5Mbit/sec**

Asymmetry

- **SkyX no longer functions**
- **Feb 2: NIH announced via vBNS at StarTap - but NIH not routing back to us via vBNS**
- **Feb 15: we find all peers - Abilene, vBNS, StarTap, Ca*net3 have some nets not routing back to us properly**

Examples 2 & 3

- **ESnet - Sandia National Lab - no solution**
  - **132.175.249.6, 134.253.26.1, 198.51.238.10**
- **Feb 22: Korea - Seoul National University**
  - **147.46.0.0/16 announced via Startap**
  - **internal routing problem with routers not able to handle OSPF or RIP2**
  - **Xylan Omni9X is the problem - once upgraded - problem will go away**

Example 4

- **Feb 24: Arizona State University**
  - **129.219.0.0/16, 149.169.0.0/16 - some subnets will route outward via Internet-1 some won't**
  - **129.219.13.81 goes via Internet-1; 129.219.253.14 goes out via Abilene**
  - **upgrade in progress**

Example 5

- **Feb 25: Virginia Tech - reverse problem discovered**
  - **208.35.69.0/24 sending data out via Abilene - but not being announced by Abilene**
  - **Belongs to Virginia Polytechnic Institute**
  - **Resolved: Abilene missing some routes on inbound BGP acl**

Example 6

- **Mar 1: University of Mich.**
  - **141.215.10.0/24 announced via Abilene - but routed back to us via Internet-1**
  - **Same problem as before - internal networks are disjoint and can’t always route back to Abilene**

Example 7

- **Mar 7: Extensive amount of networks sending data via Internet-2 but not being announced - examples:**
  - **131.109.0.0/16 - Brown University**
  - **209.100.64.0/24 - Northwestern University**
  - **153.106.0.0/16 - Calvin College & Seminary**
  - **137.48.0.0/16 - University of Nebraska at Omaha**
  - **137.197.0.0/16 - University of Nebraska Medical Center**

Monitoring

- **http://noc.ilan.net.il/LG/**
  - **select: tau-gp1.ilan.net.il**
  - **select: rip**
  - **if FastEthernet1/0.21 appears - data comes to us via Internet-2 peer - but is not being announced as an Internet-2 network**
  - **if FastEthernet0/0 appears - network announced by I2 - but all return data is via commodity Internet**
  - **3 hours and RIP entry removed**

Conclusions

- **30% of Internet-2 is asymmetric**
- **Gigapops having problems segregating Internet-2 and Internet-1 traffic**
- **Abilene to start unicast RPF checks in the near future**
- **Serious thruput problems when data is asymmetric and not using Internet-2 for both paths**