TOW2025 Operational Data Transport in the IVS (part 1)

JASON SOOHOO MIT HAYSTACK OBSERVATORY MAY 5TH – MAY 8^{TH,} 2025

Outline – Part 1

- Overview of data transport
- Networks
 - Topology
 - Performance
 - Network stack and protocols
 - Software tools
 - Operational data transport procedure e-Transfer



With IVS operations how do we get the data from our stations to the Correlators for processing?



- Data recorded to Data Recorder unit
- Modules are pulled from the recorder and brought to shipping
- Shipments can take days/weeks to arrive
- Correlator centers then process the modules

Overview – e-Transfer data



- Data recorded to Data recording unit
- The data on the recorder or server are prepared for network transfers
- Transfer of data is initiated and sent to Correlator data servers over the network
 - Transfer rates will vary (discussed later on)
- Correlator centers then process the files





Networks

Network Topology





Local network





Local router and firewall





Uplink to edge network



Edge network to Internet backbone



Everyone Else!



Resource availability and bottlenecks!

Networks

Resource availability and bottlenecks

- Networks are a shared resource, using more than available will impact performance creating a bottle across the network for everyone using it.
- Correlator data volumes are shared with other stations. We need to be sure there is enough data resources available for your data.



session data size <= data resource available



Networks

Resource availability and bottlenecks

- Networks are a shared resource, using more than available will impact performance creating a bottle across the network for everyone using it.
- Correlator data volumes are shared with other stations. We need to be sure there is enough data resources available for your data.

Input streams <= total throughput



10Gbps total throughput

session data size <= data resource available



Total data volume 100TB

Network Speed vs Transfer Time

Network Speed	900GBytes	2 TeraBytes	4 TeraBytes
100Mbps	~20hrs	~45hrs	~55.5hrs
1Gbps	~2hrs	~4.5hrs	~5.5hrs
10Gbps	~12min	~26min	~53min

However network speeds will vary depending on factors of optimization.

(Transport protocol, frame size, routing, etc.)



Network/Protocol Stack

The network/protocol stack is a conceptual model for splitting up the communication over a network into layers.

Protocols and **TCP/IP Model OSI Model** Services 7 HTTP, HTTPS, FTP, Presentation 6 DHCP, PNG Session 5 Transport TCP, UDP 4 Network Internet IP, ARP, ICMP 3 Datalink 2 Link Ethernet, Wi-Fi **Physical** 1

Network Transport Protocols

TCP - Transmission Control Protocol

- Established connection/handshaking
- Data sequenced
- Data retransmission/ Successful delivery
- Slow but complete data transmitted
- Information/File application where all bits matter

UDP – User Datagram Protocol

- Connection not needed
- Does not sequence data
- No retransmission of data
- Fast but with risk
- Streaming application where loss is acceptable

UDT - UDP-based Data Transfer Protocol

- Application layer over UDP
- Connection oriented
- Data sequenced
- Data retransmitted
- Faster than TCP and more reliable than UDP alone

Software

Data Transferring Tools

- Linux utilities
 - ftp/sftp, rsync, scp, etc.
- Tsunami transfer software
 - http://tsunami-udp.sourceforge.net/
- jive5AB/m5copy
 - https://github.com/jive-vlbi/jive5ab
- etransfer etc/etd client
 - https://github.com/jive-vlbi/etransfer

Network Testing Tools

- Linux utilities
 - ping, traceroute, etc
- Iperf3
 - https://iperf.fr



Part 2