

Multipathing with MPTCP and OpenFlow

RoN Meeting

Utrecht

8 June 2012

Ronald van der Pol

rvdp@sara.nl

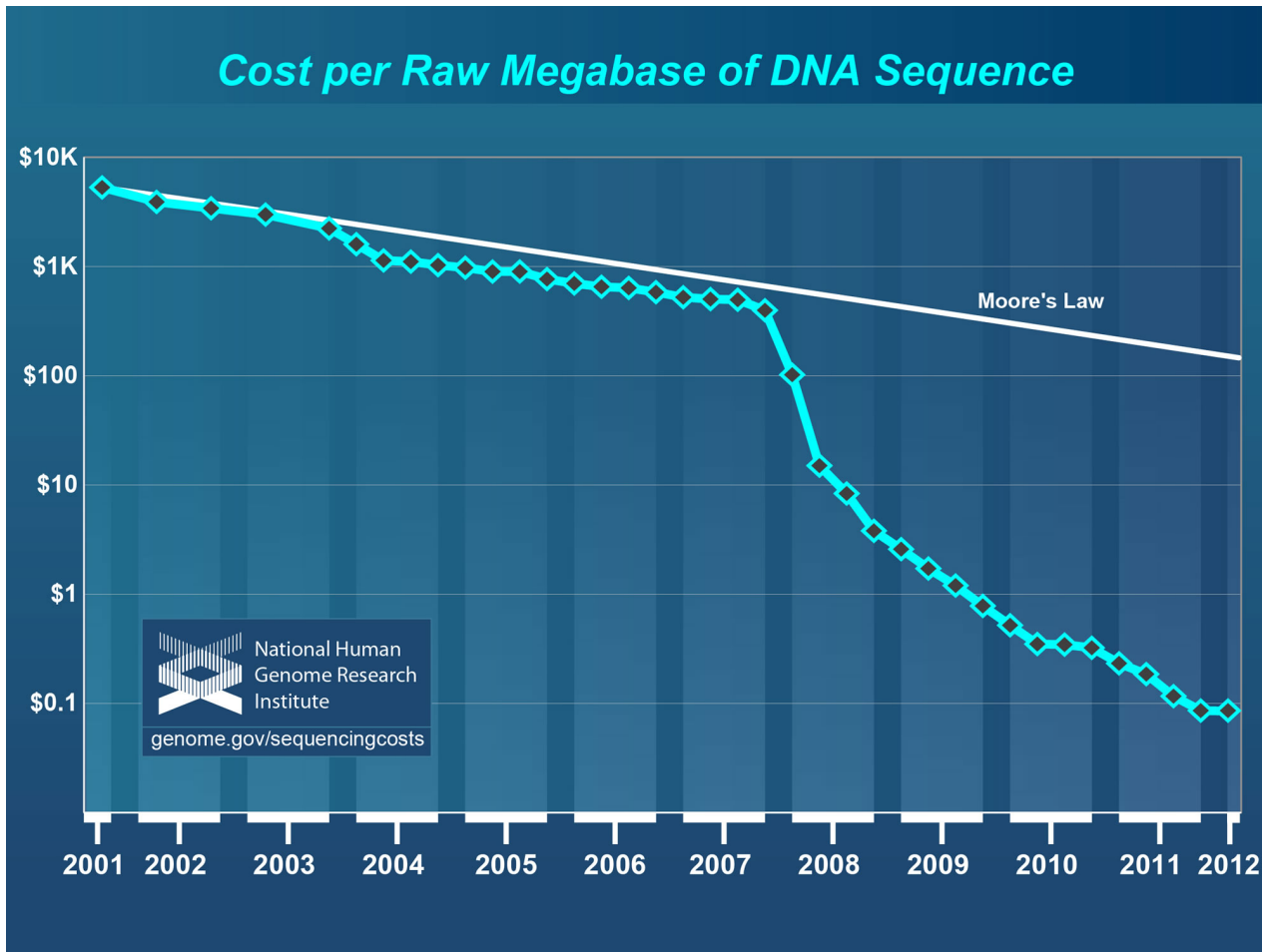
Demonstration at GLIF and SC12

- Partners:
 - Caltech
 - iCAIR
 - SURFnet/SARA
- International OpenFlow testbed
 - OpenFlow
 - Multipath TCP
 - Ethernet OAM (IEEE 802.ag) monitoring

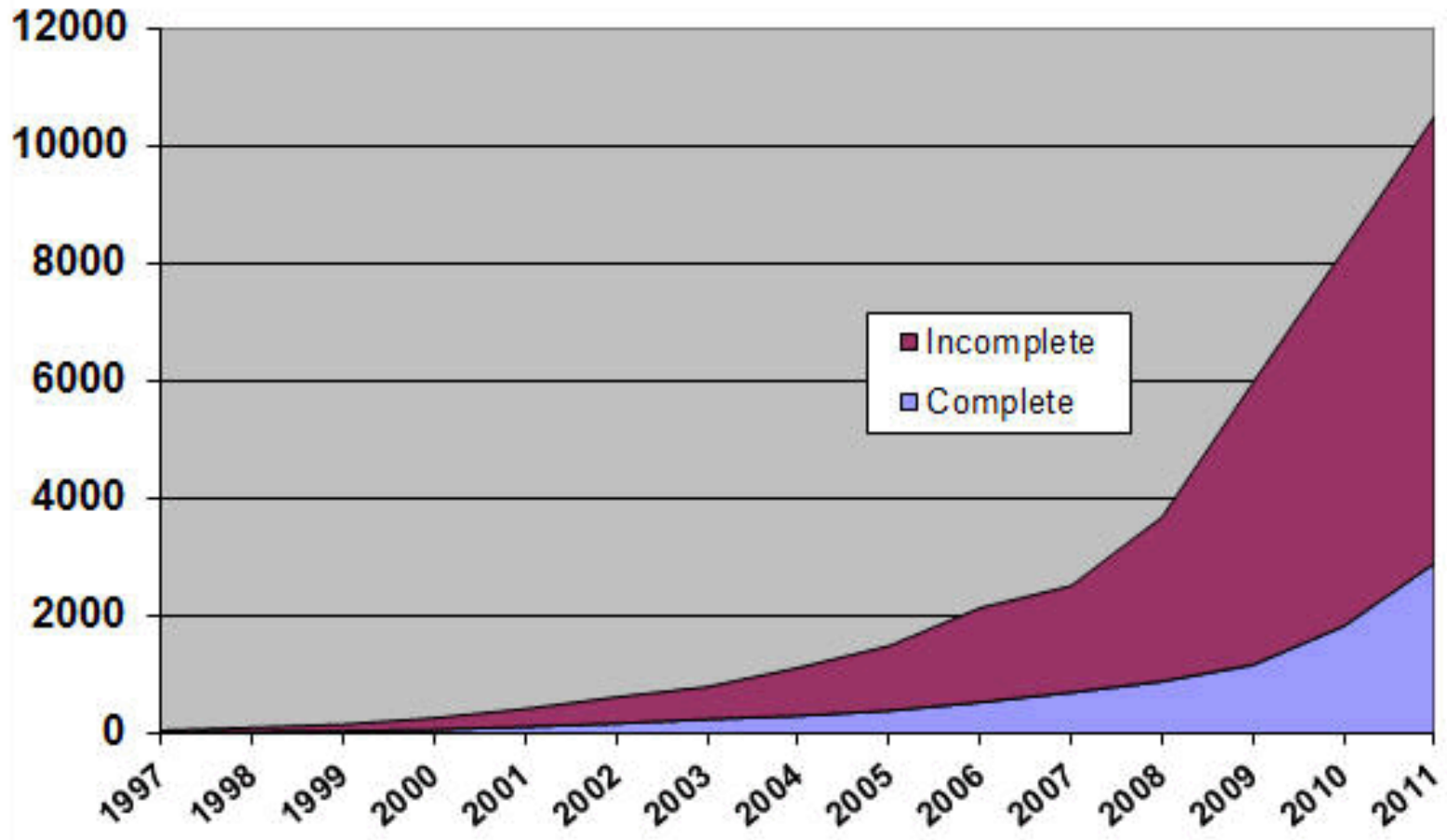
Why multipathing?

- Network needs to keep up with data tsunami
 - Exponential increase in e-science data and data transfers
 - Need for a world-wide e-science infrastructure
 - Huge data streams between data centres
- Reaching the limit of cost effective use of fiber capacity
- Just the next step in parallelism
 - Storage → RAID
 - Computing → multi-core
 - Servers → Grid & Cloud

Price Drop in Data Sequencing



Genome Sequencing Data Explosion



Genome Online Database, DOE, USA

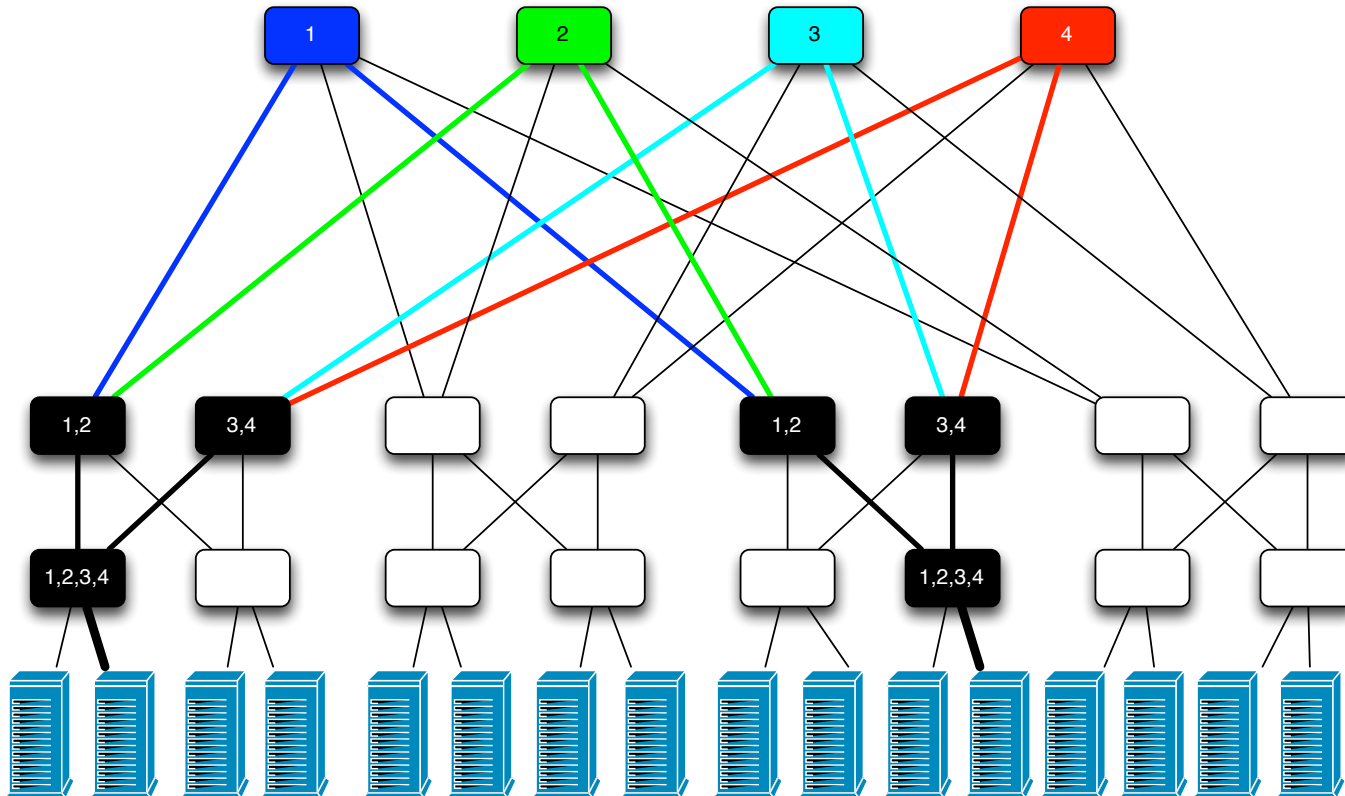
E-Science Data Centre Infrastructure

- We need to work on a world-wide e-science storage, computing & visualisation infrastructure
- Economy of scale (like Google, Amazon, Yahoo!, etc.)
- Cloud services like Amazon, but with the scientist in control over the data
 - Scientist decides where (which country) to store the data
 - Scientist decides if data is transferred to computing or VMs are moved to the data
- Big data streams between the data centres only
- Data transfers towards end-user limited by number of pixels of screen (everything in the cloud, only results are show on the scientist screen)
- Jointly operated as one operational domain (KISS!)
 - Centralised OpenFlow based Traffic Engineering (like Google)

Data Centres and Multipathing

- Big data centres (Google, Amazon, Yahoo!) already moved to clos/fat-tree topologies with massive multipathing
 - Clouds, Hadoop clusters, etc with large data streams between the servers
- Protocols for data centre multipath topologies
 - IETF TRILL & IEEE 802.1aq (SPB)
 - No spanning tree used
 - Topology exchange via IS-IS
 - Simultaneous use of multiple links
 - Cisco FabricPath
 - Juniper QFabric
 - Brocade VCS
 - NVGRE, VXLAN

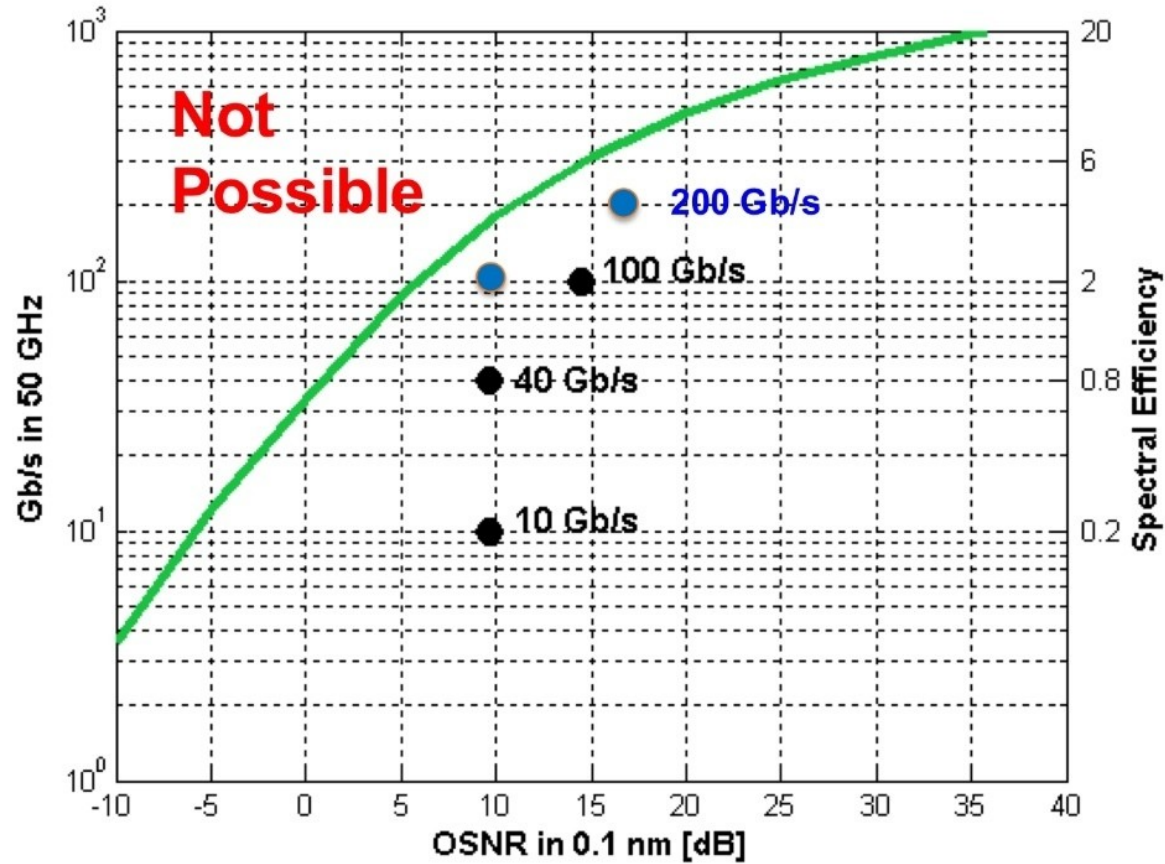
4 paths between server pairs



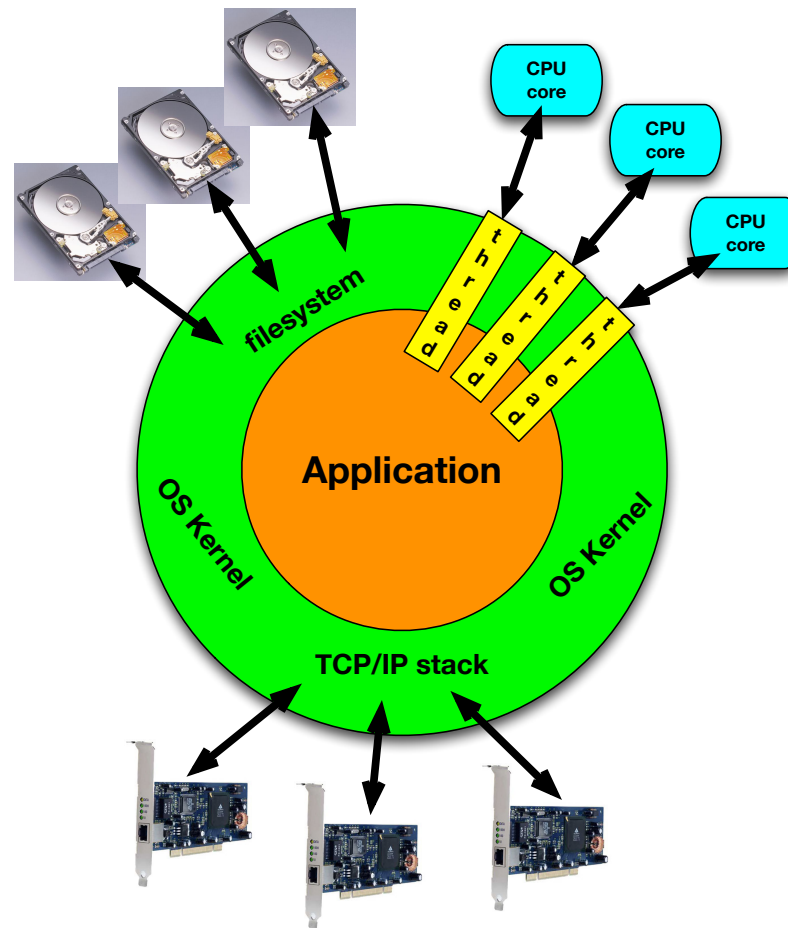
Fiber Capacity

- Reaching the limit of fiber capacity
 - Modern modulation technologies reach the Shannon limit
 - Tradeoff between bandwidth, reach, spectral efficiency and cost

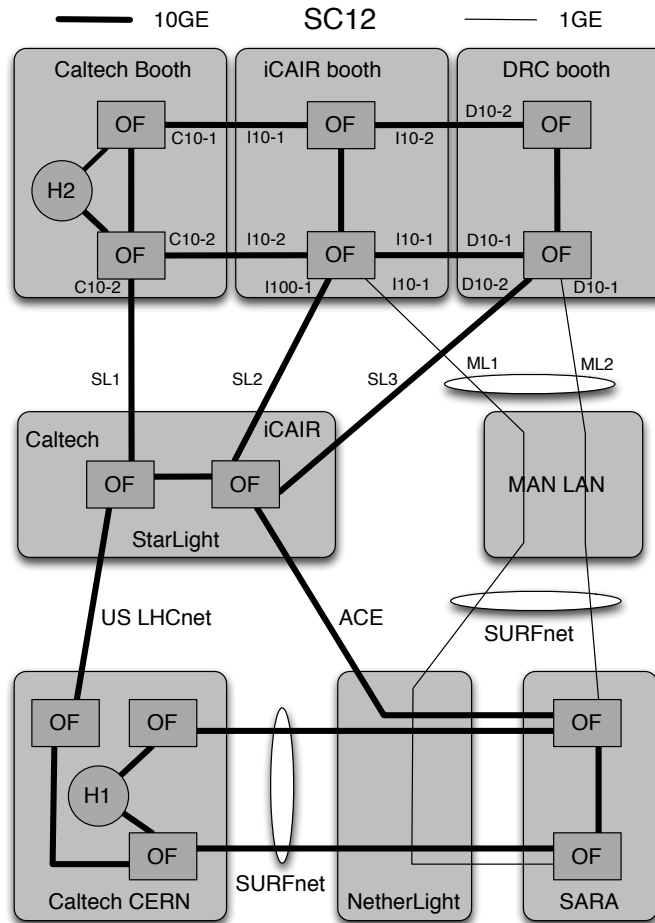
Shannon Limit

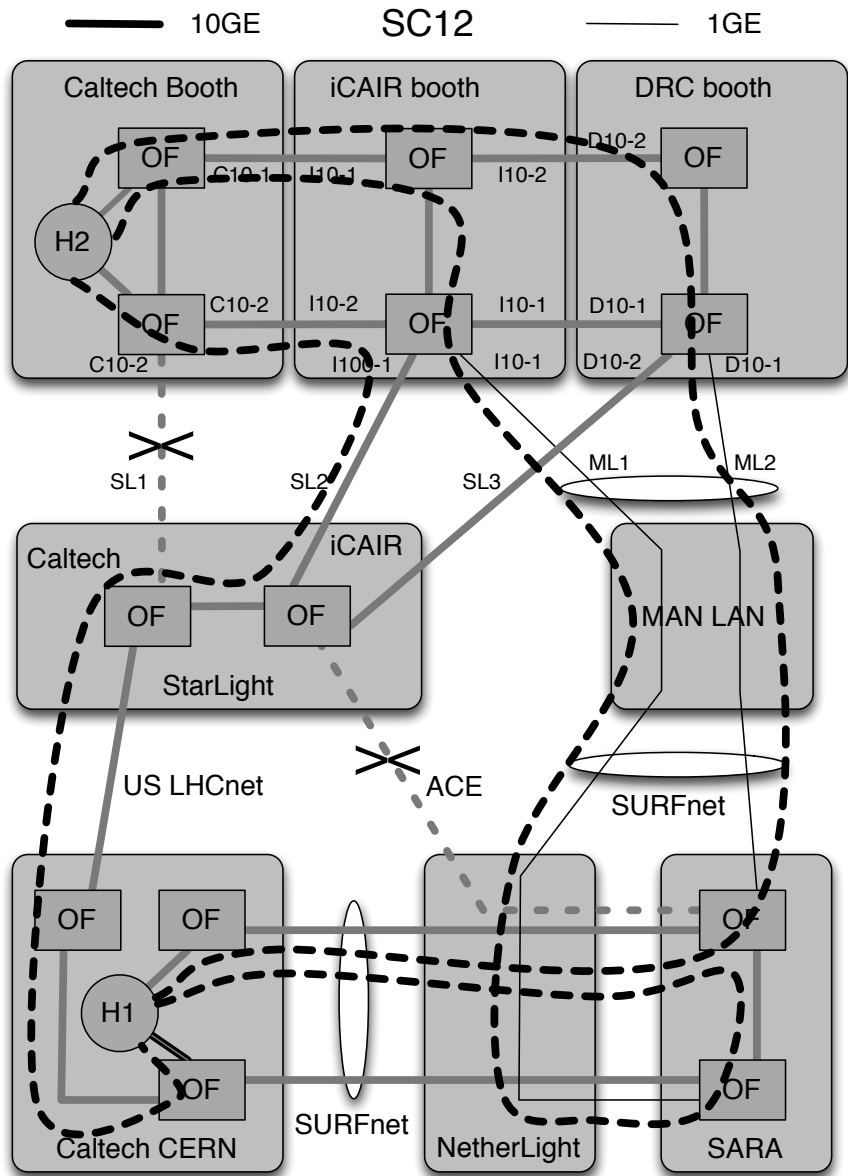


Multiple Disks, Cores & NICs

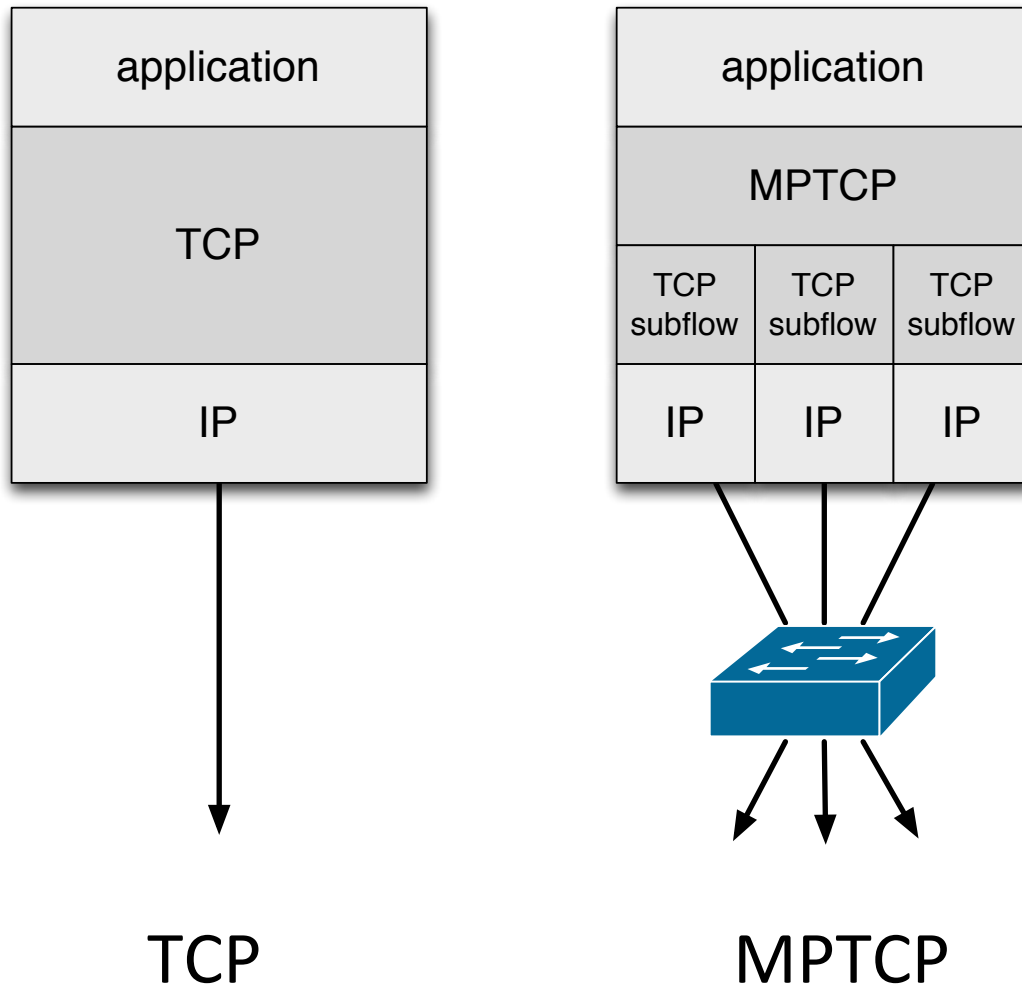


OpenFlow Topology





MPTCP Network Stack



MPTCP Main Functions

- Path management
 - IPv4 & IPv6 address negotiation
 - Setup and tear down of subflows
- Packet scheduling
 - Split application byte stream into segments
 - Transmit segments on subflows
- Congestion control
 - Spread load across subflows
 - Move traffic from one subflow to another in case of congestion or failure

MPTCP Handshakes

