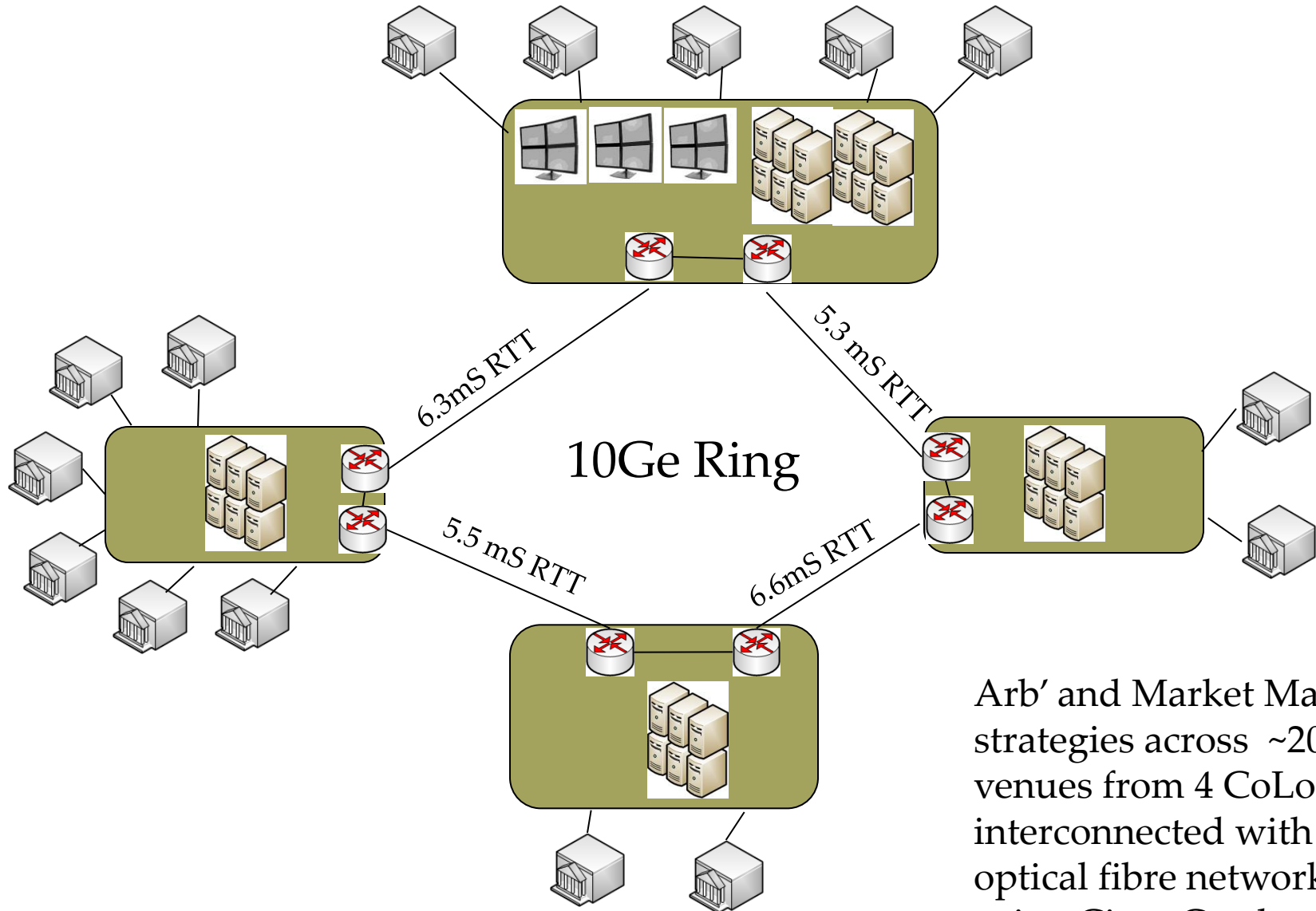
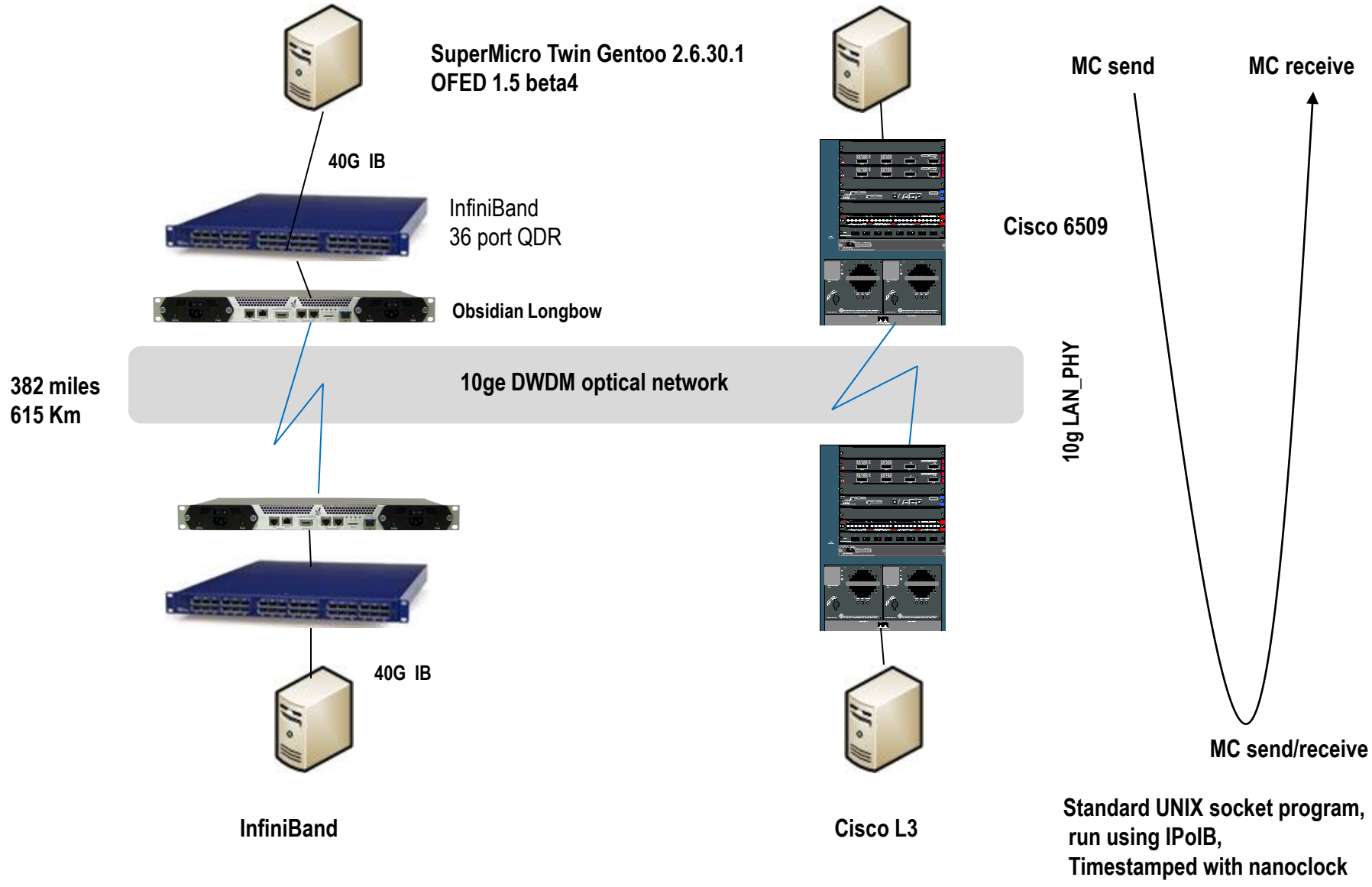


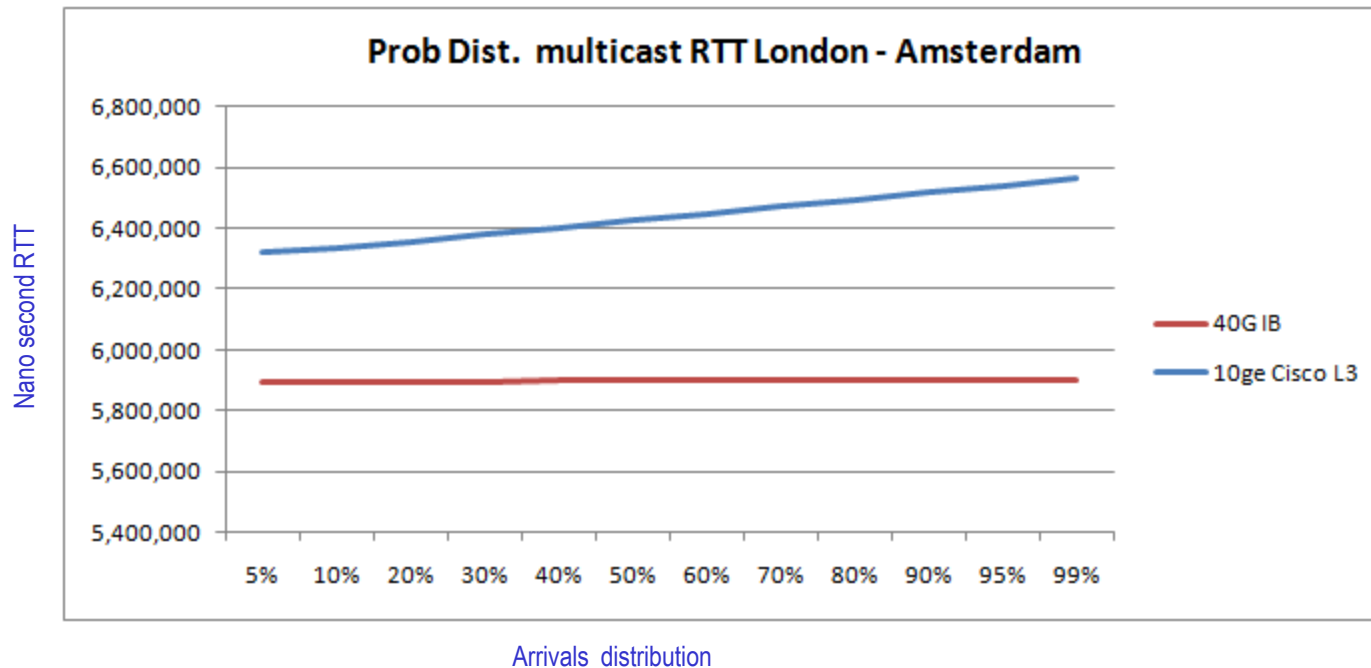
# **TRADING WITH LONG HAUL INFINIBAND**

- Richard Croucher – Chief Architect and Founder of Informatix Solutions
- Former Chief Architect at Sun Microsystems and Principle Architect at Microsoft (Live)
- ~20 years working with Financial Services companies
- Specialize in discovering and exploiting new technologies
  - Including SANs, distributed computing, high performance networking, Grid computing, virtualization and Cloud computing
  - Full life cycle - problem analysis, solution design, selection and POC, operability, observability and diagnosability
  - Working with – Architects, Infrastructure SME's and application developers
- Currently concentrating on low latency trading systems for both buy and sell sides
- [Richard.Croucher@informatix-sol.com](mailto:Richard.Croucher@informatix-sol.com)
  - Based in the UK but a frequent visitor to New York
  - See [www.informatix-sol.com](http://www.informatix-sol.com) for more info



Arb' and Market Making strategies across ~20 venues from 4 CoLo's interconnected with 10G optical fibre network using Cisco Catalyst 6500 routers

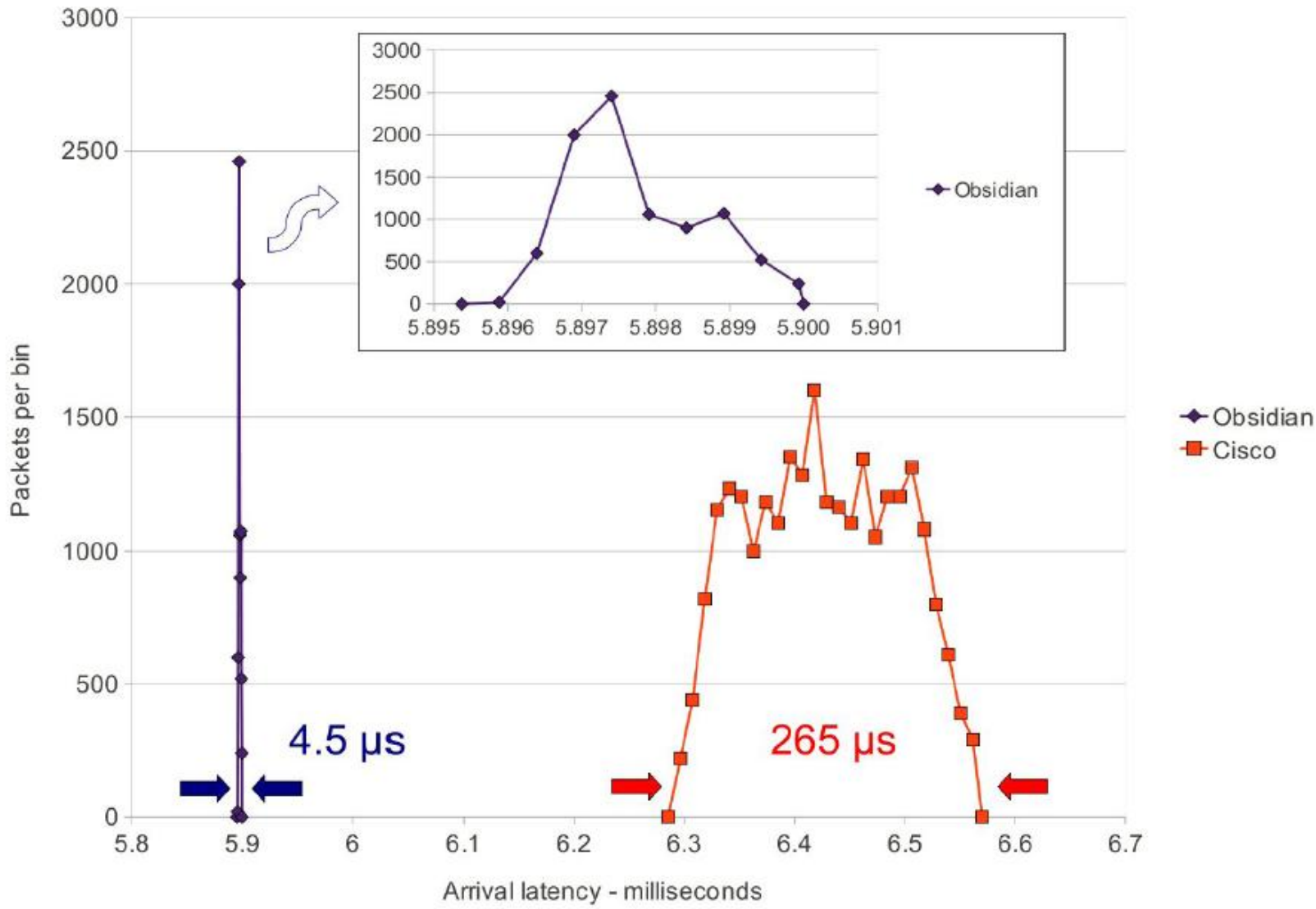




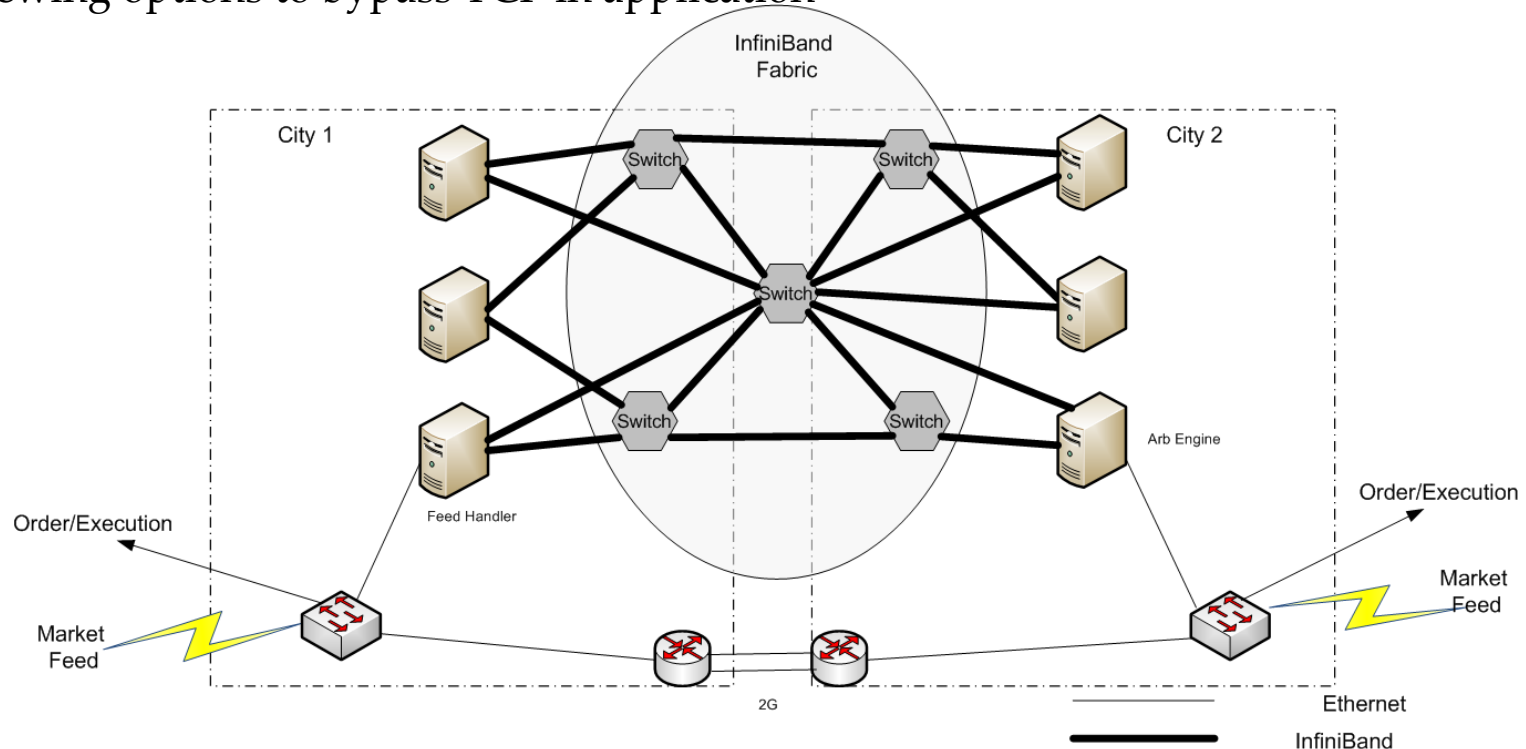
Measurement made using same Amsterdam to London long haul fibre path  
Compares existing Cisco Catalyst network using Layer 3 routing to InfiniBand running IPoIB  
Overall 526µS improvement on round trip time, providing 8% lower latency

100,000 multicast UDP packets sent:  
InfiniBand 0 dropped 0 duplicated  
Cisco 0 dropped 325 duplicated (0.32%)

### RTT Arrival latency histogram



- Local InfiniBand clusters interconnected across long haul using Obsidian Longbows
- Single InfiniBand Subnet spanning all locations using Host based OpenSM 3.3.2 with weighted path selection running on Management server at each location
- Pod design scalable to 96 local servers using 6x 36-port switches, total of 8u + 1u cabling tray
- Using Longbow 1G Ethernet “side channels” to preserve L3 network addressing to ease migration
- Evaluating Ethernet/InfiniBand gateways to allow remote access to raw feeds across InfiniBand
- Reviewing options to bypass TCP in application



Solution	Budgetary Cost	Strengths	Weakness
Cisco Catalyst 6500	€1,437K	<p>Widest installed</p> <p>Proven technology</p> <p>Risk adverse</p>	<p>Poor Bandwidth usage</p> <p>Complexity of configuration</p> <p>Costly given provided functionality</p> <p>Same as everybody else – no latency advantage</p>
Nortel ERS8600	€ 919K	<p>Well proven</p> <p>High B/W usage through Active:Active L2 links</p> <p>Simpler L2 management than Cisco</p> <p>Better POP scalability through multipoint support (SW upgrade in 2009)</p> <p>Risk Neutral</p> <p>Lowest cost solution</p>	<p>Different to Cisco so small learning curve</p>
InfiniBand	€1,330K	<p>Lowest latency solution</p> <p>High B/W usage through Active:Active</p> <p>First to deploy pan-Market low latency solution in Europe</p> <p>Includes 20gb/s server attach providing additional application performance benefits</p>	<p>Learning curve of new technology</p> <p>First installation in Financial Services for Europe, for this distributed IB fabric</p> <p>Could be considered bleeding edge solution and therefore highest risk</p>



- Long distance InfiniBand works with minimal changes and is reliable
- Large Layer2 subnets simplify many network functions e.g. MC routing
- InfiniBand switch based SM's are not flexible enough - constrained by GUI configuration
- IPoIB was 100% compatible with programs run over Ethernet
- At the time we evaluated them, Ethernet-InfiniBand gateways were not ready for production deployment
- Reliable packet delivery is not free – requires careful QoS configuration to avoid congestion
- Mesh topologies can suffer from loops during reconfiguration events, needs careful design to avoid problems
- Standard IB vendor management tools need supplementing with “start of day” and diagnosability scripts
- Existing Network teams require substantial training – allow minimum of 2 days plus hands-on