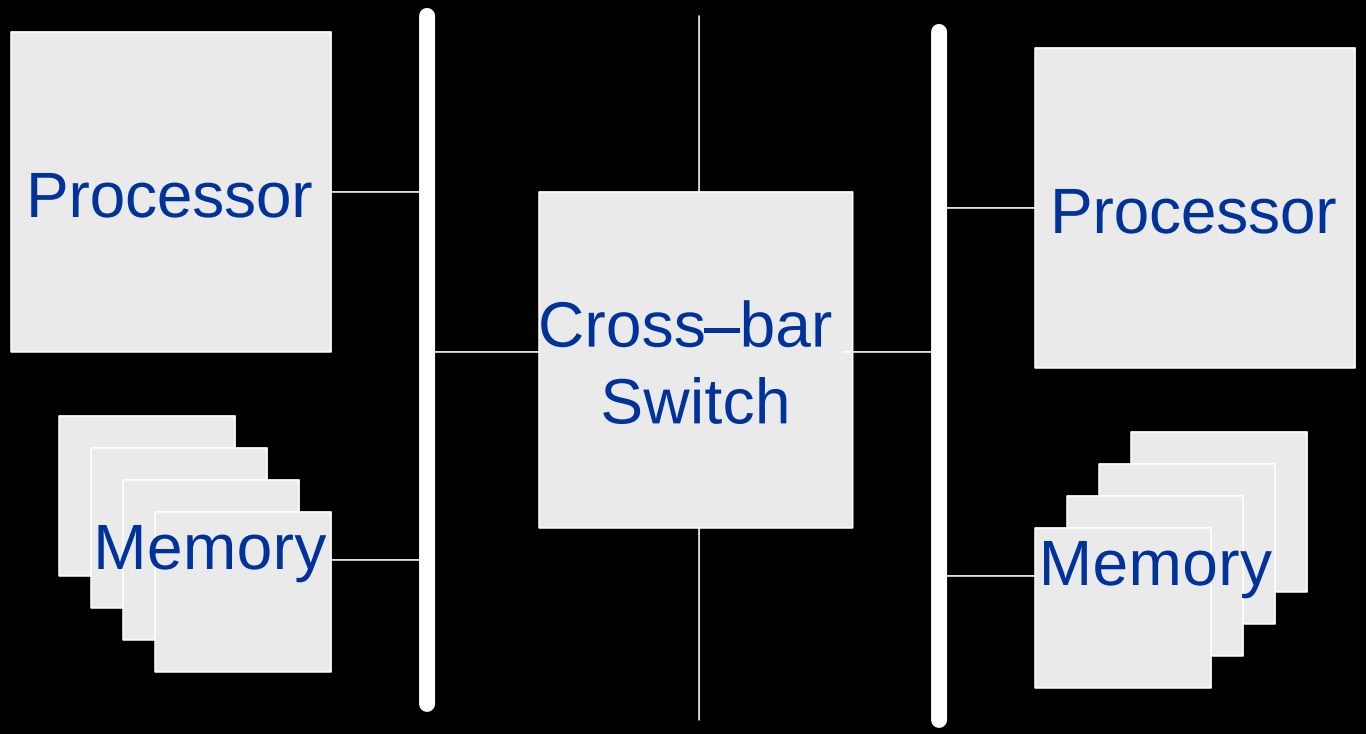
Performance of ORBs on Switched Fabric Transports

Victor Giddings
Objective Interface Systems
victor.giddings@ois.com

ORB*express*™

Switched Fabrics

- **High-speed interconnects**
  - High-bandwidth, low latency switched circuits
  - Adaptive routing through alternate paths
  - DMA transfers between memories of processors



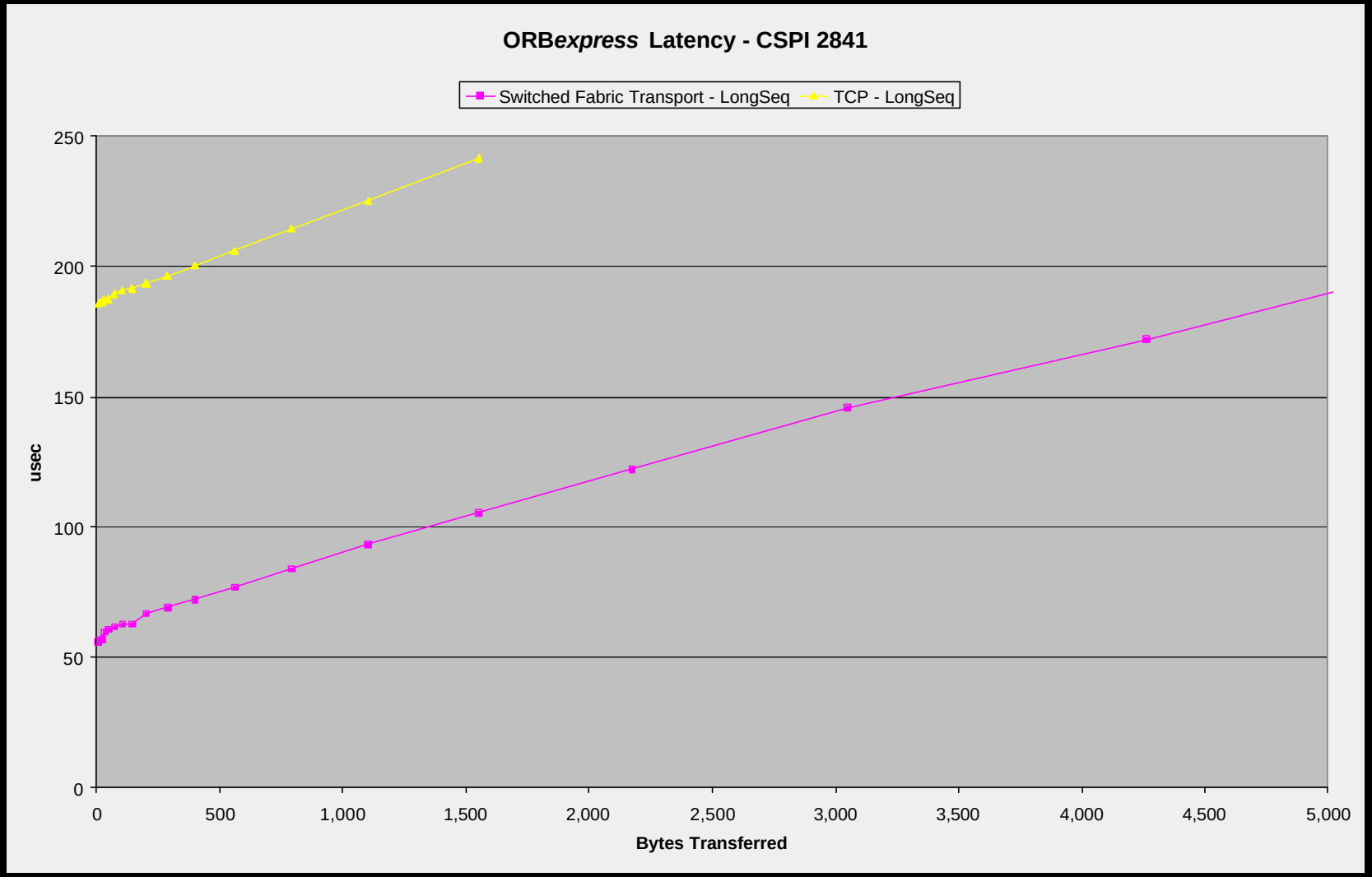
Motivation: ORB Performance on Ethernet

- **Most ORB performance studies have used TCP over Ethernet**
  - Most common use of CORBA
  - Well-known performance
- **Prediction of an ORBs performance using TCP over Ethernet**
  - CPU speed is largest determinant of performance
    - Startup latency is dominated by processing in the protocol stack
    - Data throughput is dominated by marshalling (copy) time
  - Extrapolating performance is a matter of scaling CPU speeds
- **Problem: how to predict ORB performance on switched fabric transports**

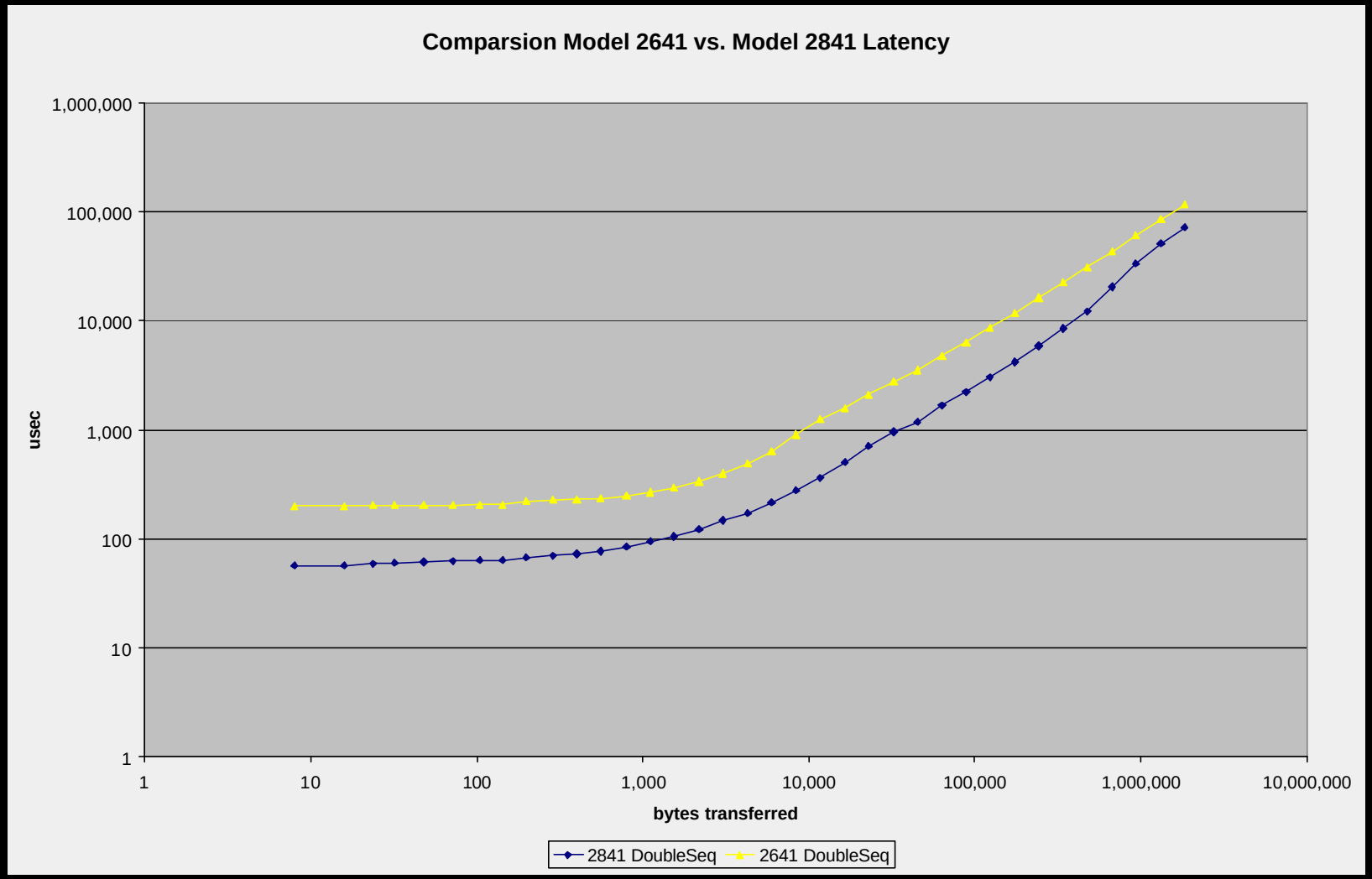
Context: ORB Switched Fabric Transports

- **ORB*express* transports developed for two different switched fabric technologies**
  - Mercury Computing's RACEway
    - Joint development with Mercury Computing
  - Myrinet (CSPI & Myricom)
- **Performance results shows**
  - Extremely low latency
  - Low variability

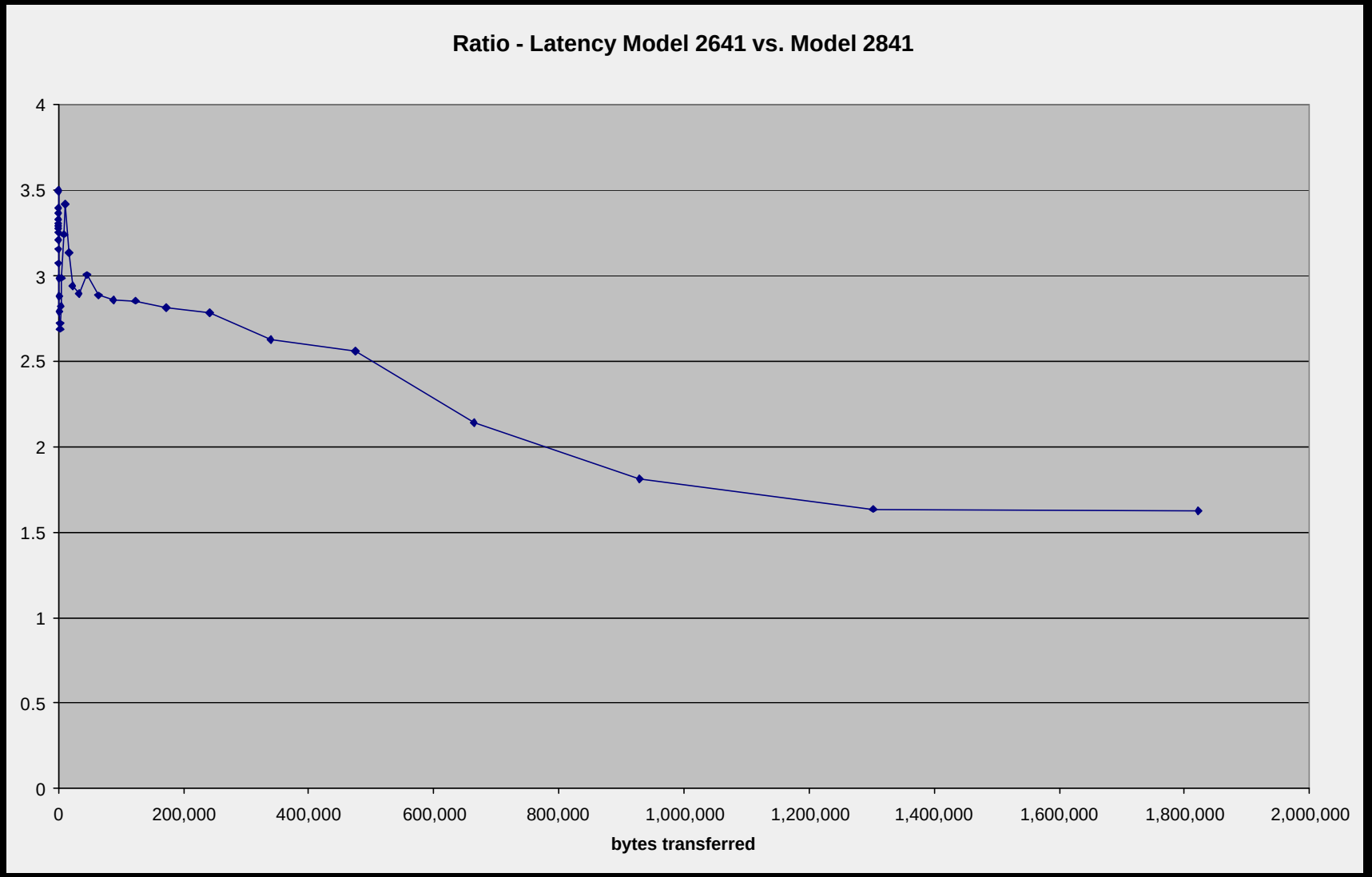
Example Latency Results CSPI 2841



Comparison - Latency



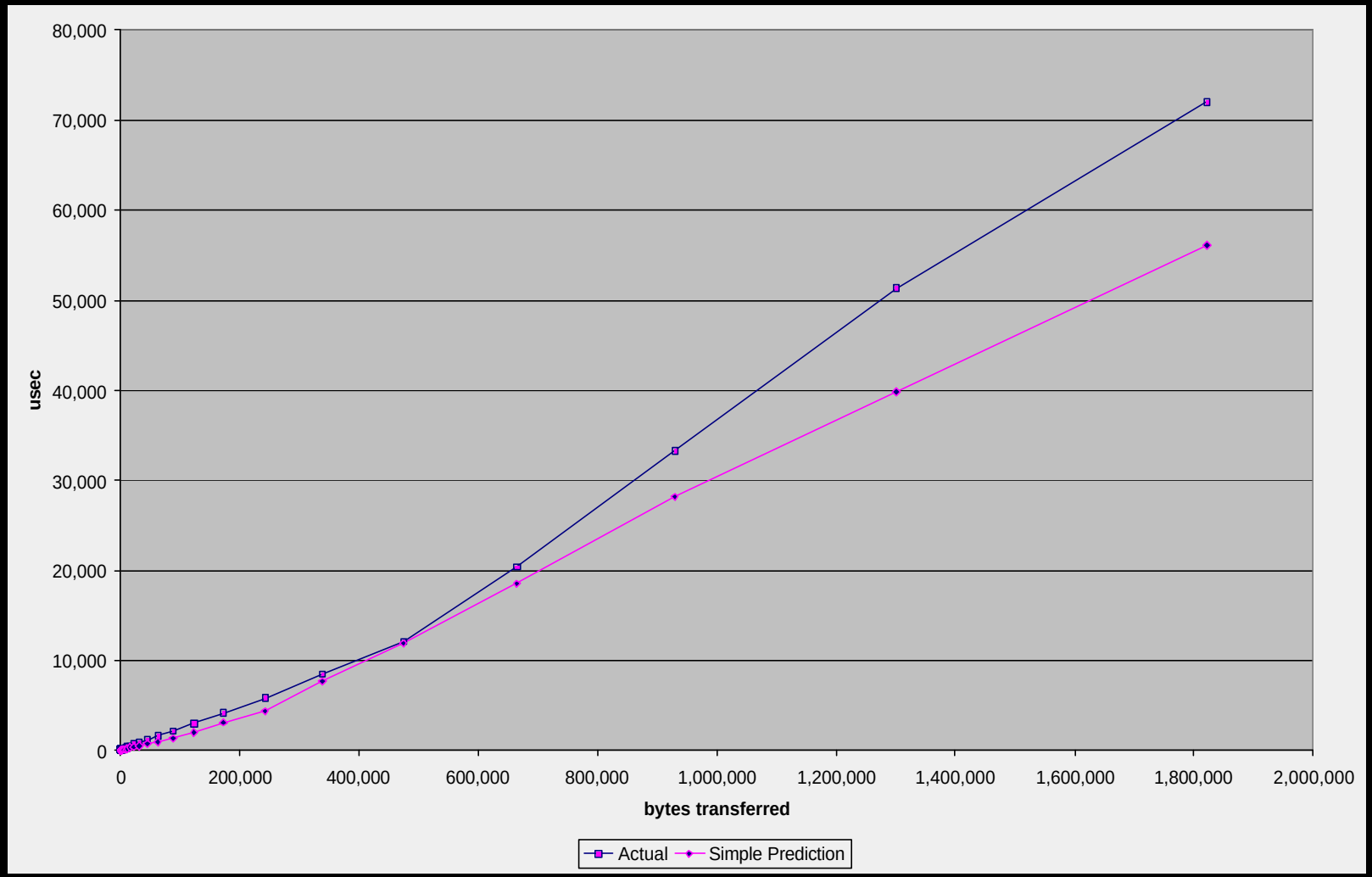
Ratio of Latencies



Performance Model – First Attempt

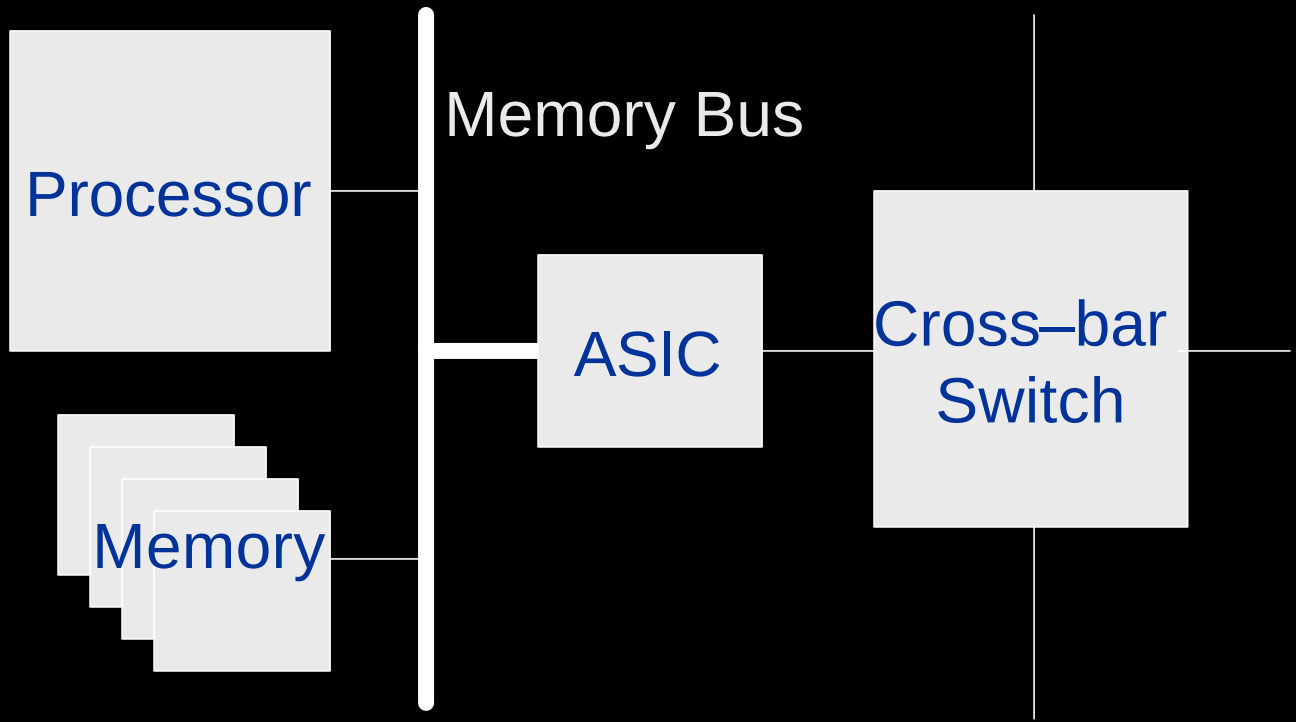
- **Use "Ethernet" model**
  - Latency = ORB "overhead" + Transport Propagation delay
- **ORB*express* provides a "mirror" transport**
  - "Reflects" requests back to collocated sender
  - Directly measures ORB overhead including marshalling
- **Simple model:**
  - Latency = ORB latency
    + Number of bytes * Propagation per byte

Performance Model – Result of First Attempt



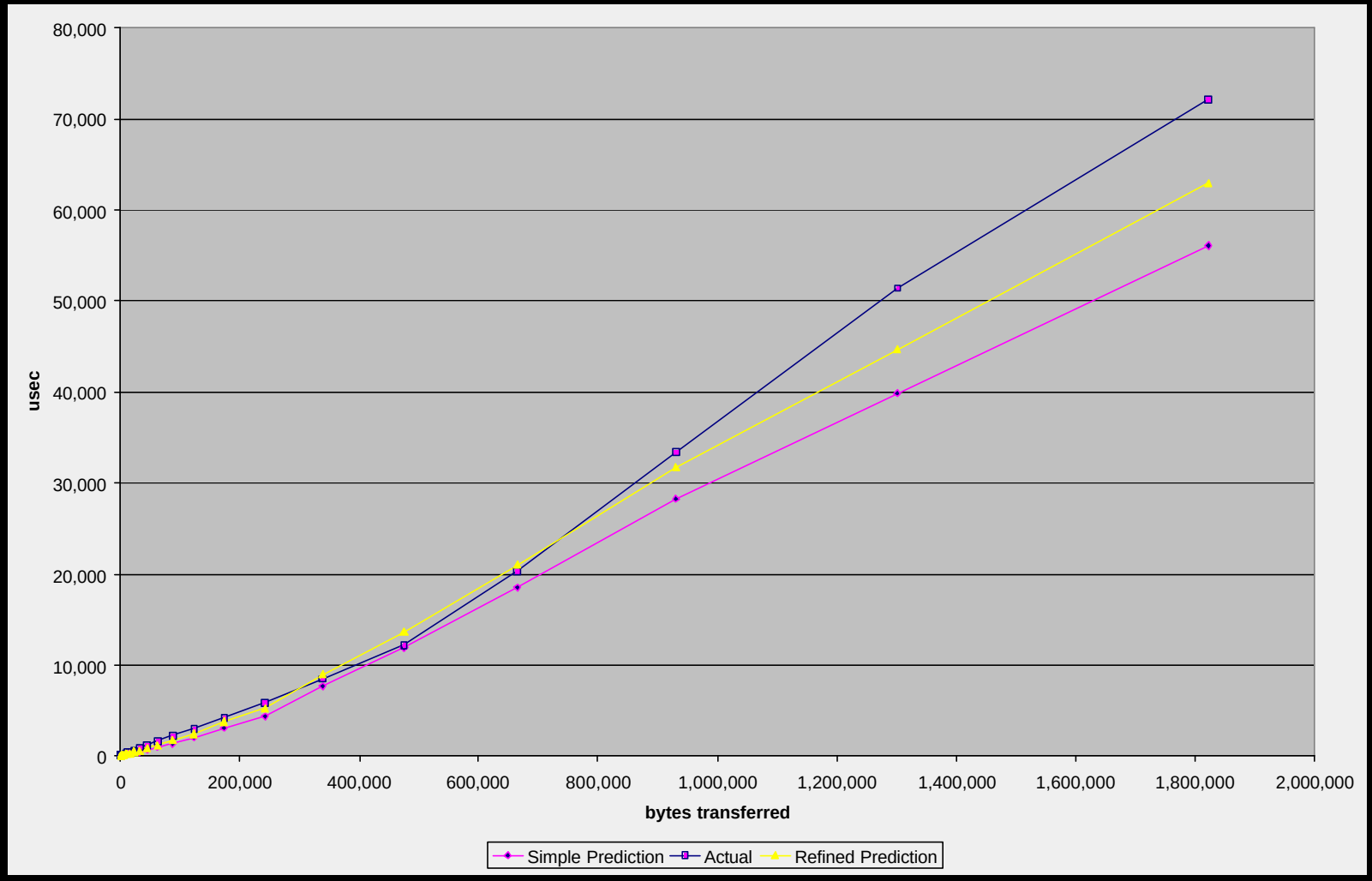
Performance Model - Refinement

- **Refined hardware block diagram**

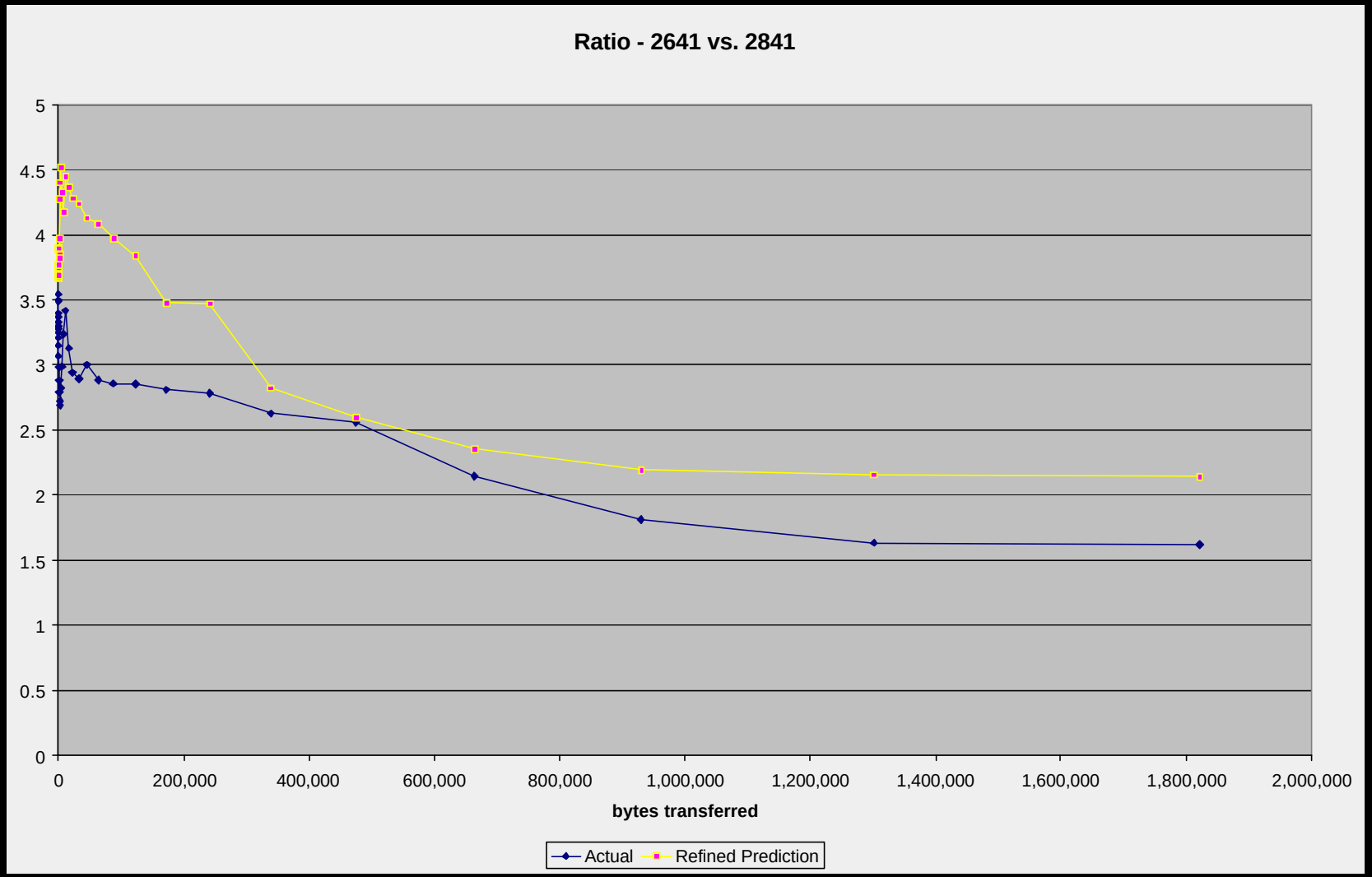


- **Account for propagation delay on memory bus**

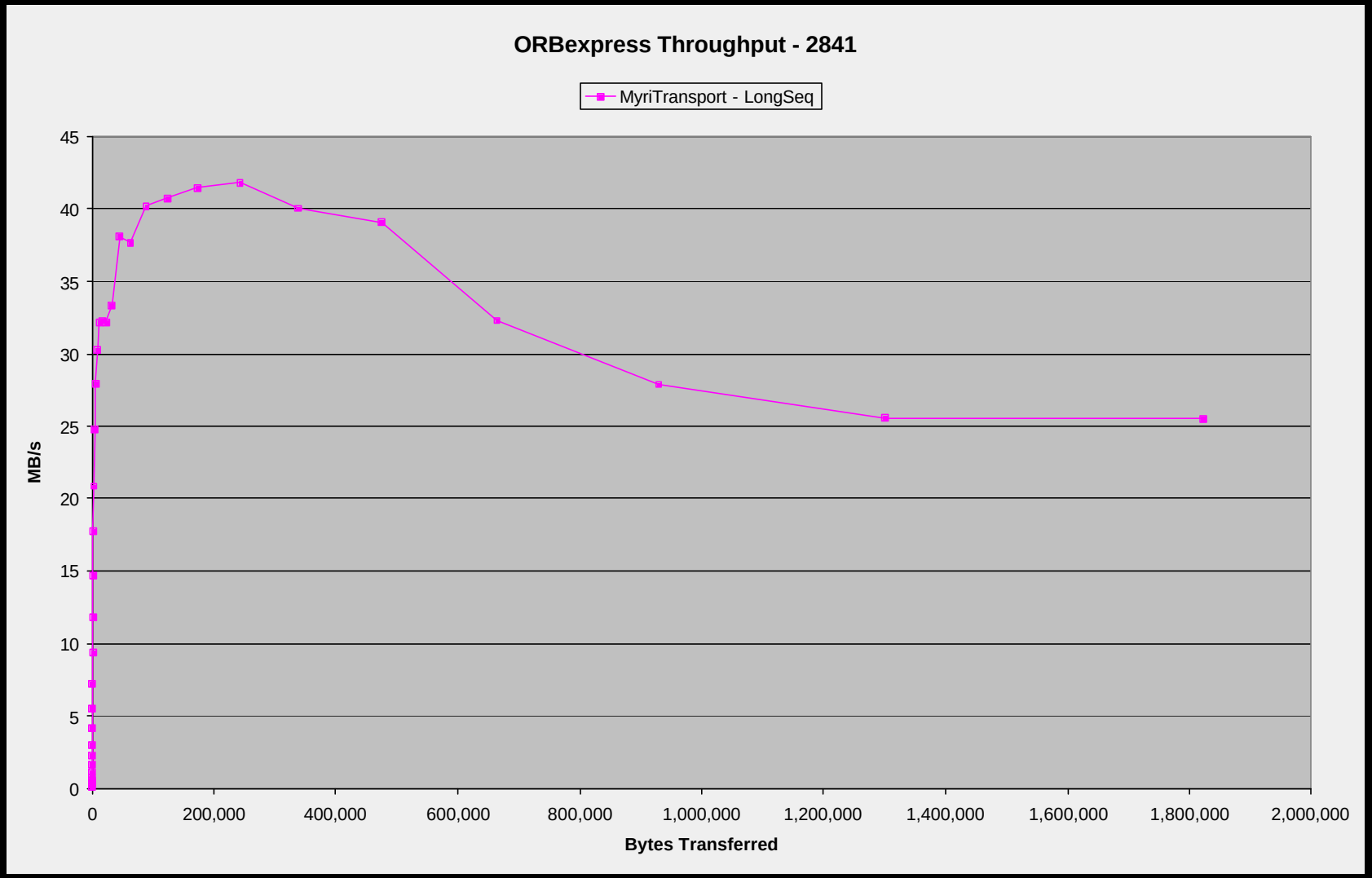
Performance Model – Result of Refinement



Performance Model – Prediction of Ratios



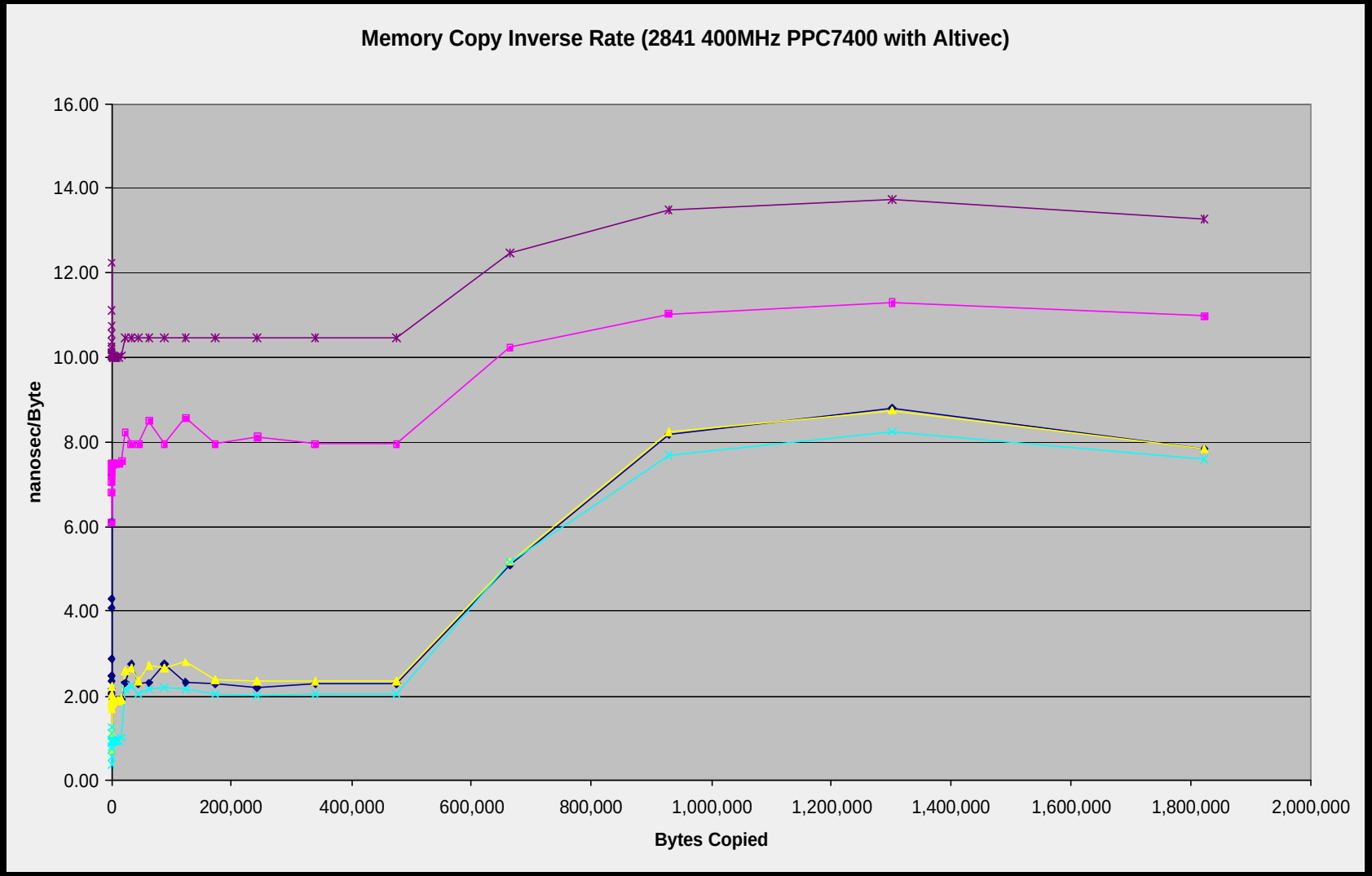
Bandwidth – ORB over Switched Fabric Transport



Bandwidth

- **Attained bandwidth is small part of available transport bandwidth**
- **Mirror transport bandwidth offers clue to cause**
  - Since inverse bandwidths add (additional time per byte adds)
  - Transport bandwidth must be combined with “ORB bandwidth”
- **ORB bandwidth factors**
  - Startup latency – insignificant for significant byte counts
  - Memory copies

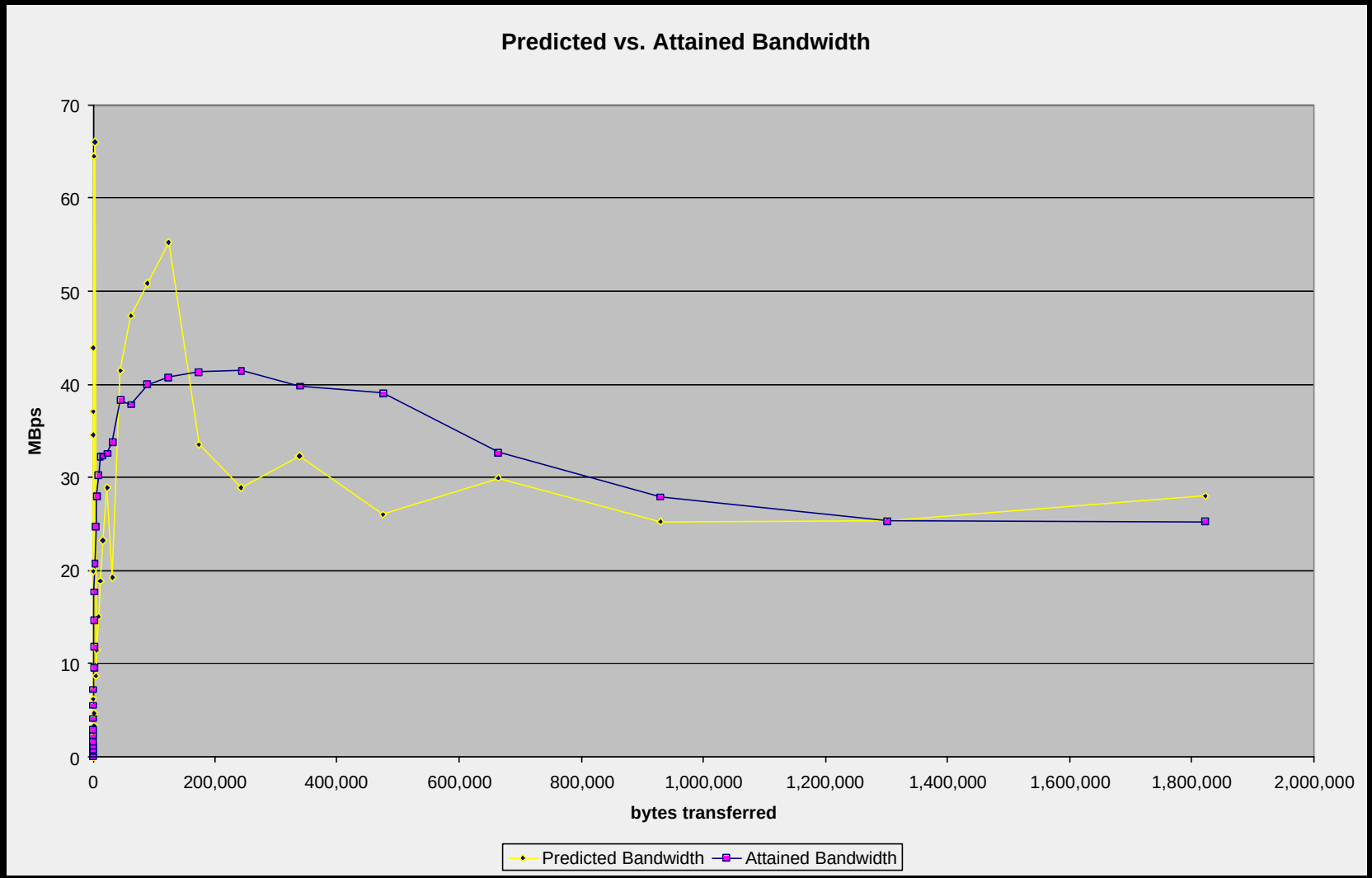
Memory Bandwidth – Copy Inverse Rate



ORB Bandwidth

- **ORB Bandwidth**
  - Measures the ability of the ORB to transfer volumes of data
- **ORB bandwidth is dominated by**
  - The number of copies
  - And the memory bandwidth of the processor
- **Increasing an ORBs bandwidth**
  - Requires elimination of copying
  - Motivation for “High Performance Enablers” RFP

Bandwidth - Predicted



Summary

- **Examined two aspects of ORB performance over switched fabric transports**
- **Latency**
  - Prediction is more complicated than for TCP/IP over Ethernet or Loopback
  - More complex model needed
- **Bandwidth**
  - Prediction is more straightforward
  - Introduced concept of “ORB Bandwidth”
  - ORB bandwidth is dependent on the number of copies