Performance Analysis and Evaluation of LANL's PaScalBB IO nodes using QDR Infiniband and Multiple 10-Gigabit Ethernets

Mentors:
Hsing-bung Chen
Alfred Torrez
Parks Fields

Instructor:
Andree Jacobson

Students:
Rocio Perez-Medina
Juan C. Franco
Daniel Illescas

Outline

A. Abstract
B. Why this project is important?
C. Tests
   a. Back-to-back Multiple 10-Gigabit Bonding
      - Results
   b. IB/QDR + Multiple 10-Gigabit Bonding
      - Results
D. Conclusion
E. Questions

Abstract

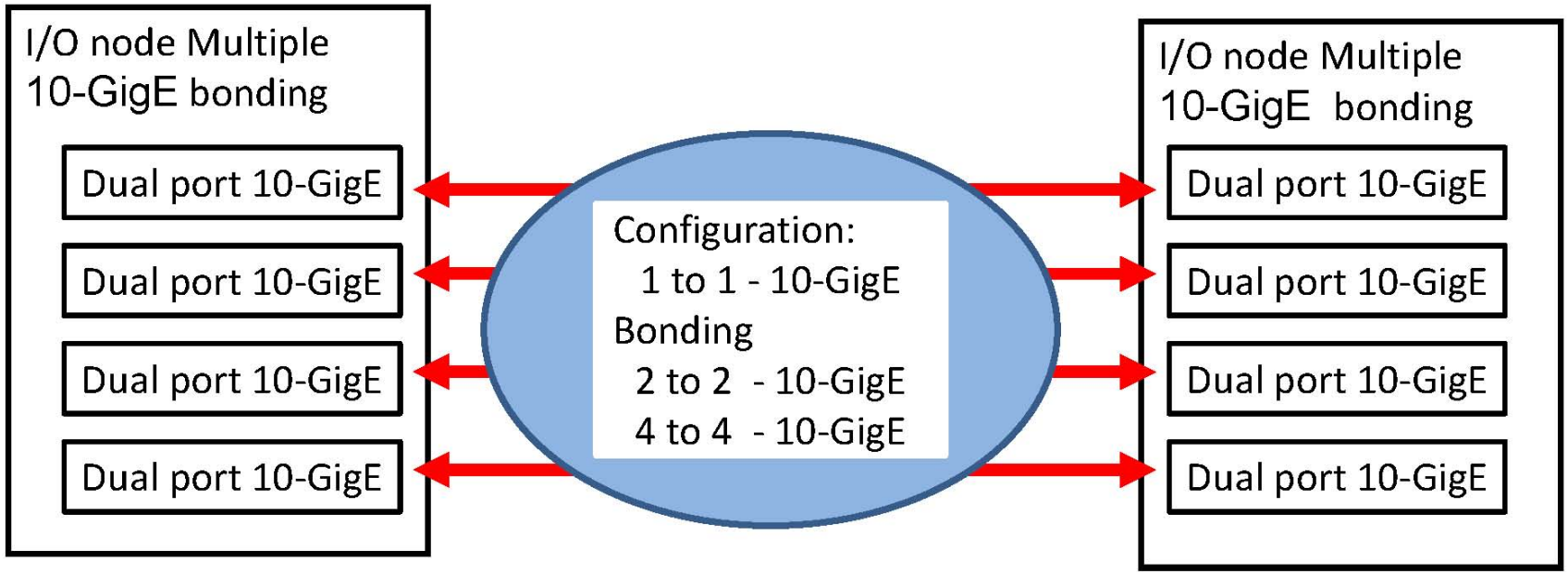
- I/O nodes are carried data traffic between backend compute nodes and global scratch file systems.
- Data Movement using Quad-Data Rate (QDR) Infiniband and multiple 10-Gigabit Ethernet bonding
- Small-scale PaScalBB test bed and conduct a sequence of I/O node performance tests.
- Discovery of enhanced I/O node network configuration

Why this Project is important?

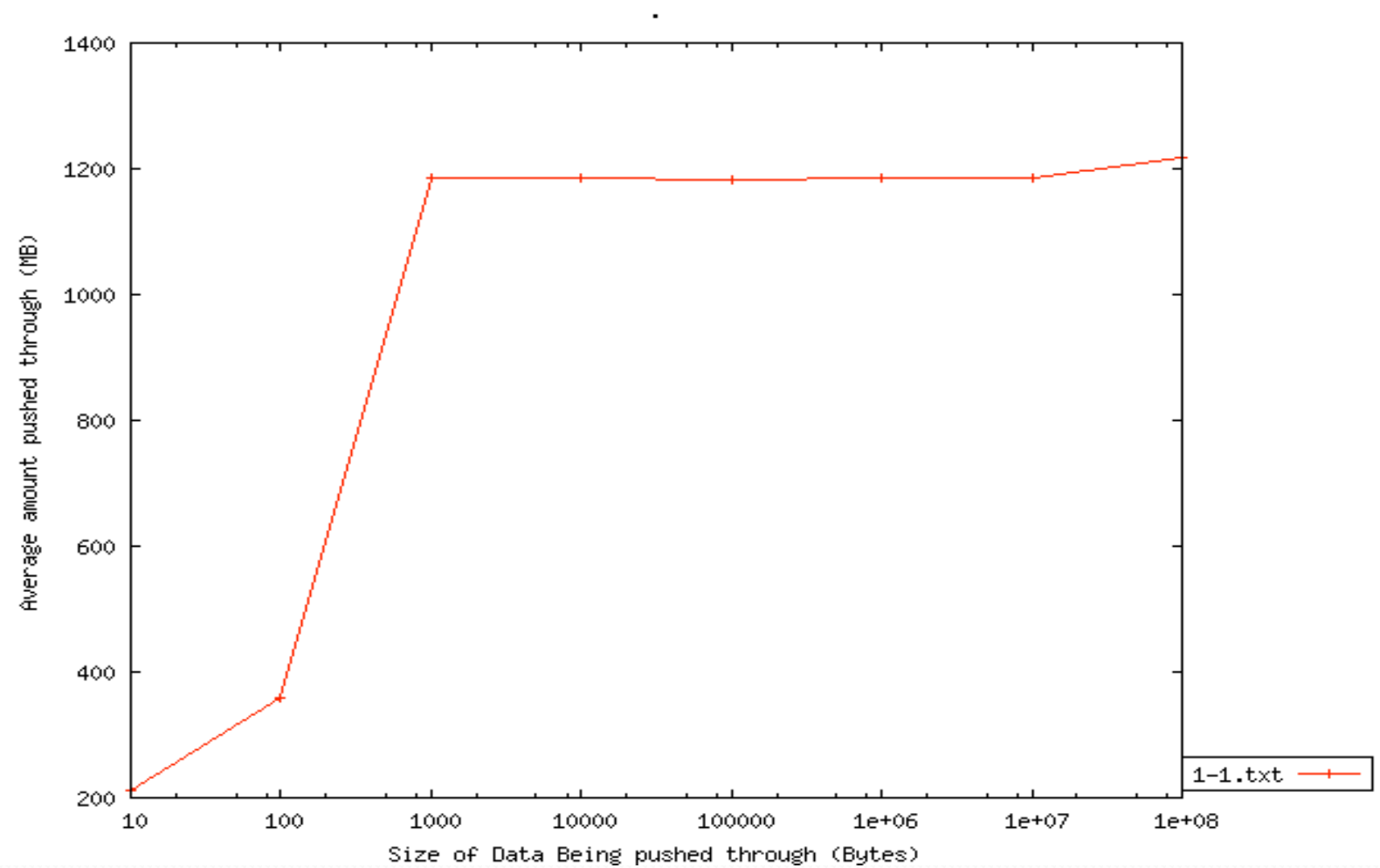
Multiple 10-Gigabit bonding testing - using back-to-back & 10-gigabit Ethernet switch

- Using back-to-back node connection
- Using Arista 48-port 10-Gigabit switch
- Using Linux bonding Mode 0 (round-robin) and Mode 5 (load sharing) for load balancing testing
- Using IB/QDR and 10-GigE Ethernet TCP Tuning Parameters from Mellanox, Myricom, LBL TCP tuning

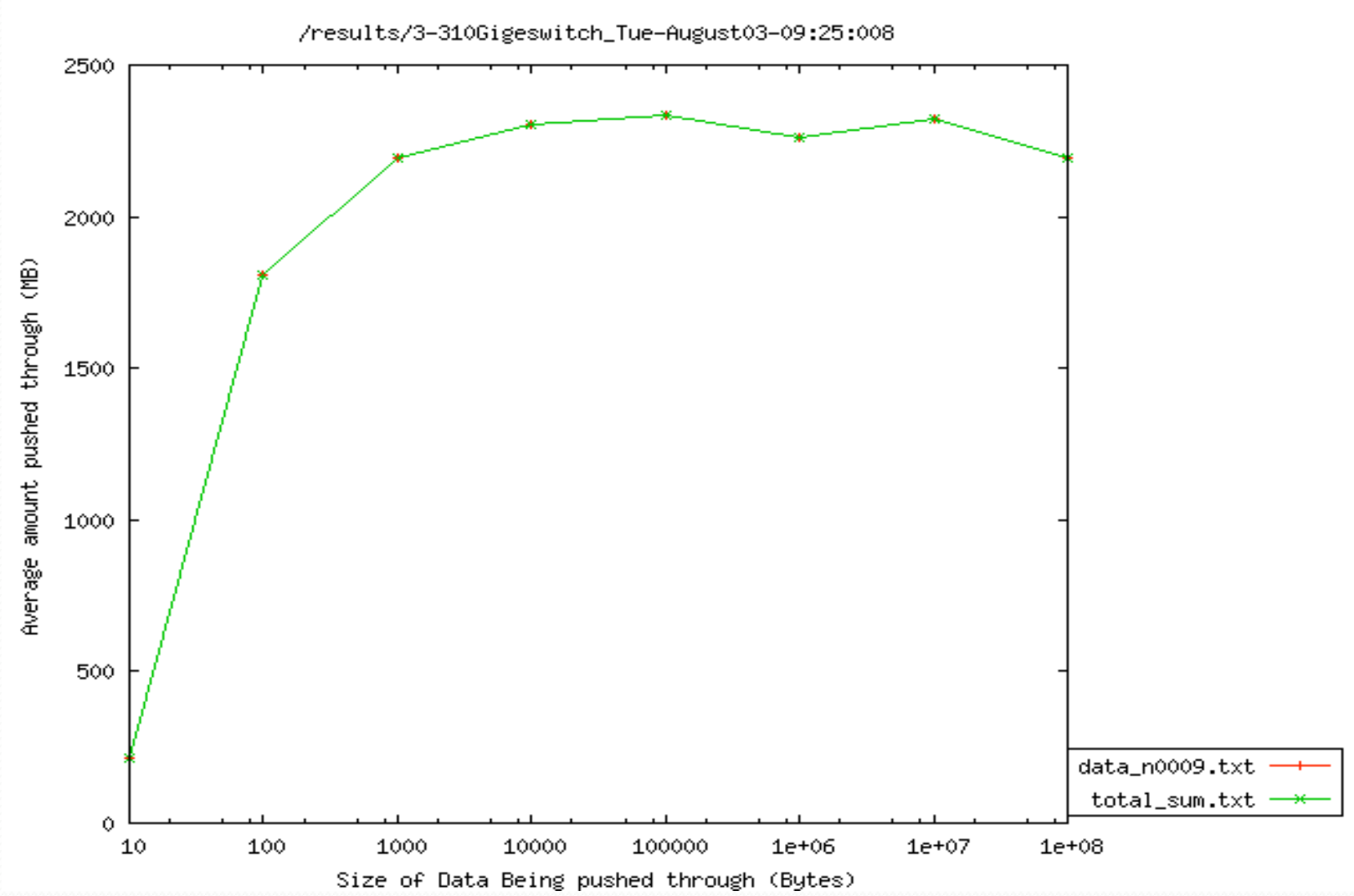
Back-to-Back Multiple 10-Gigabit bonding



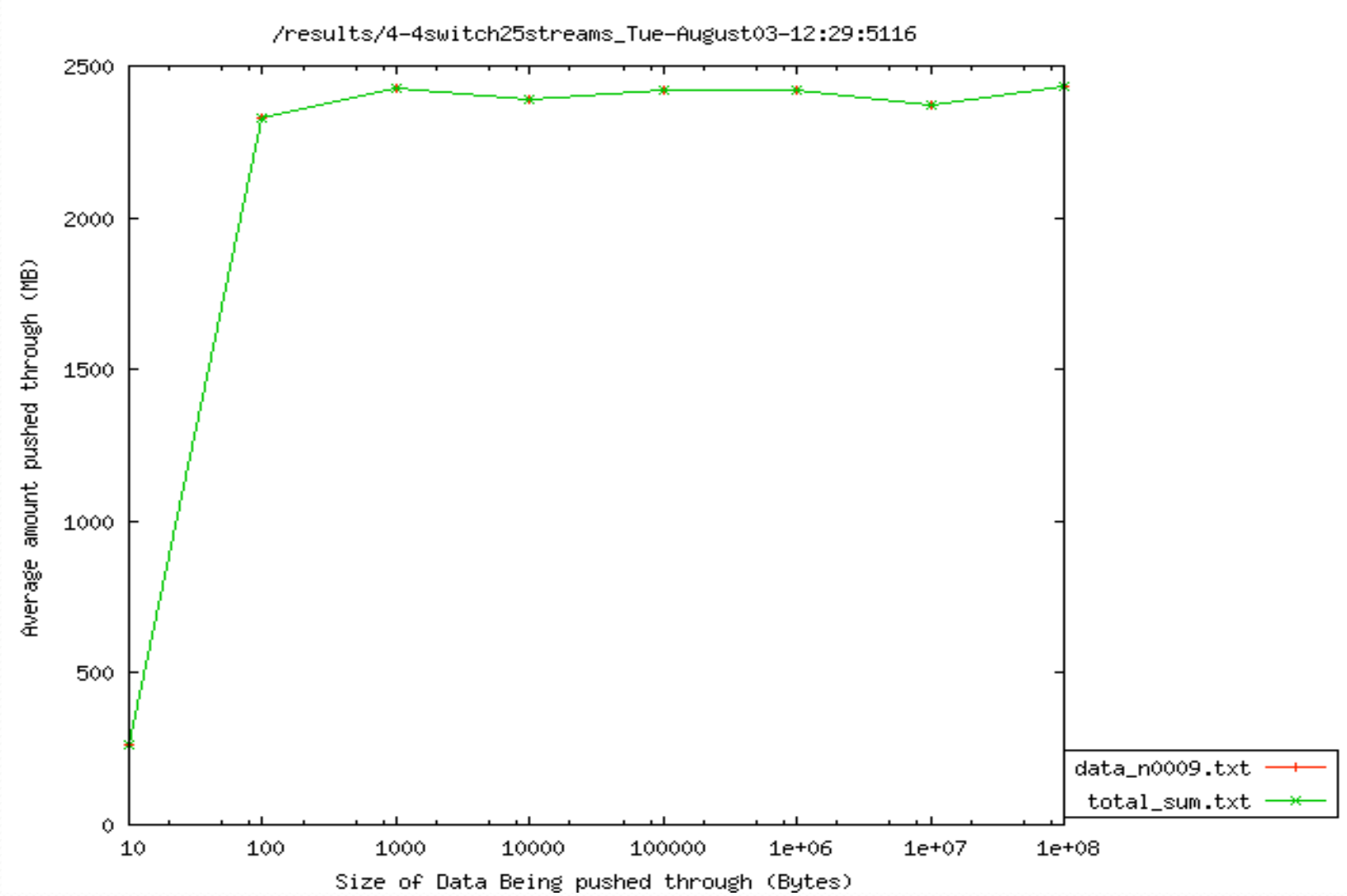
1 to 1


1400
1200
1000
800
600
400
200
Average amount pushed through (MB)
10
100
1000
10000
100000
1e+06
1e+07
1e+08
Size of Data Being pushed through (Bytes)
1-1.txt

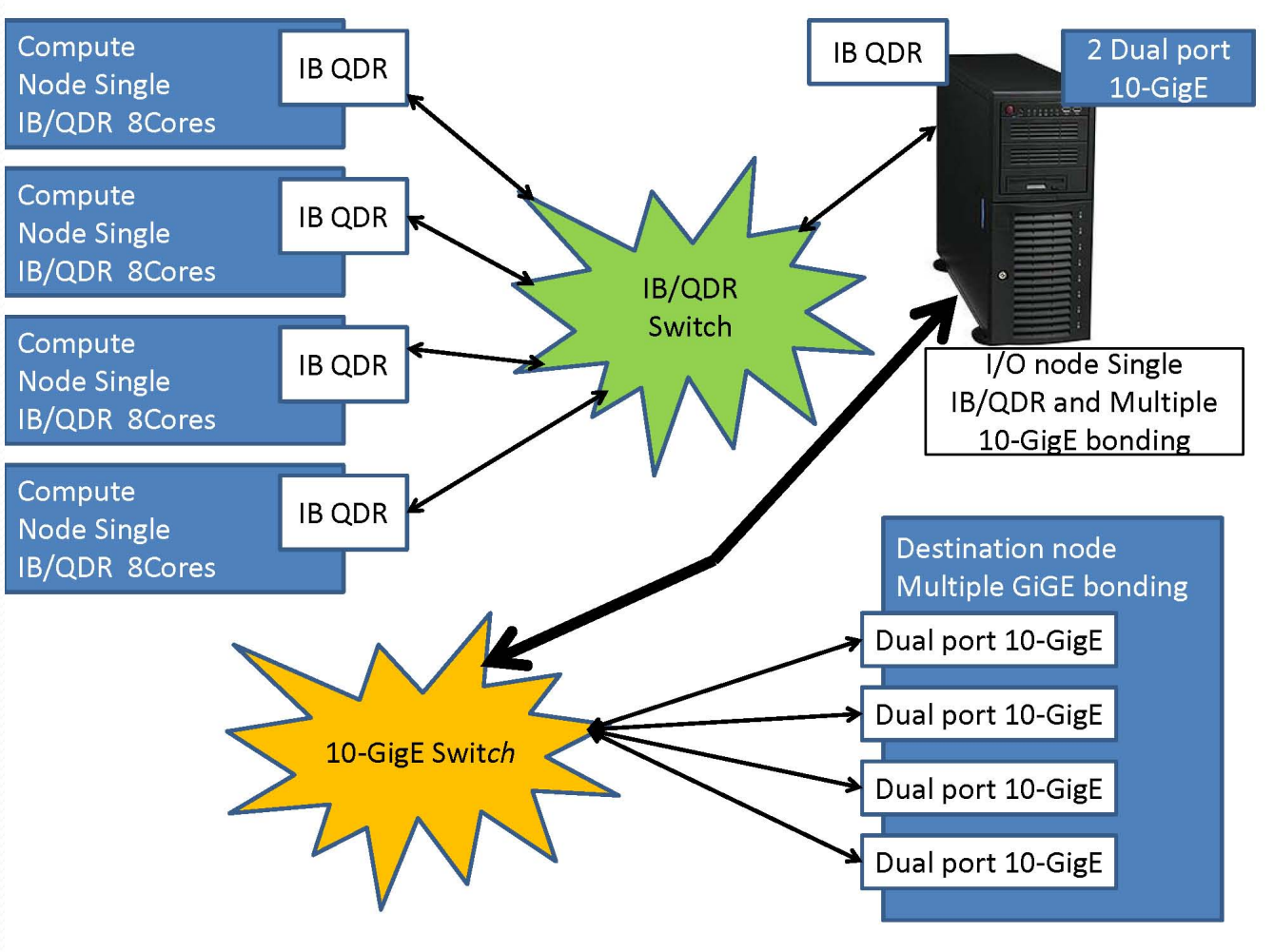
3 to 3


/results/3-310Gigeswitch_Tue-August03-09:25:008
Average amount pushed through (MB)
2500
2000
1500
1000
500
0
10
100
1000
10000
100000
1e+06
1e+07
1e+08
Size of Data Being pushed through (Bytes)
data_n0009.txt
total_sum.txt

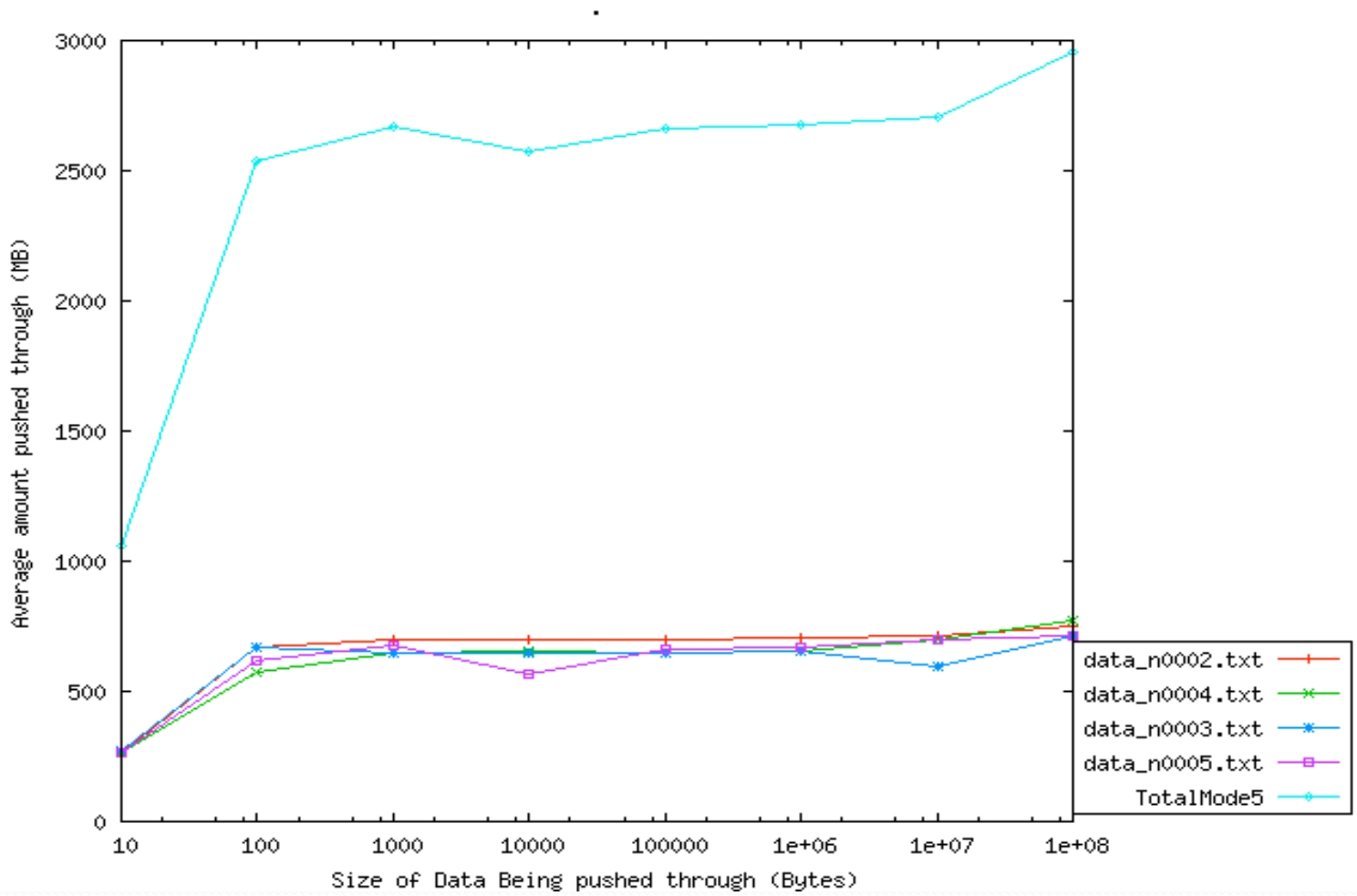
4 to 4


/results/4-4switch25streams_Tue-August03-12:29:5116
Average amount pushed through (MB)
2500
2000
1500
1000
500
0
10
100
1000
10000
100000
1e+06
1e+07
1e+08
Size of Data Being pushed through (Bytes)
data_n0009.txt
total_sum.txt

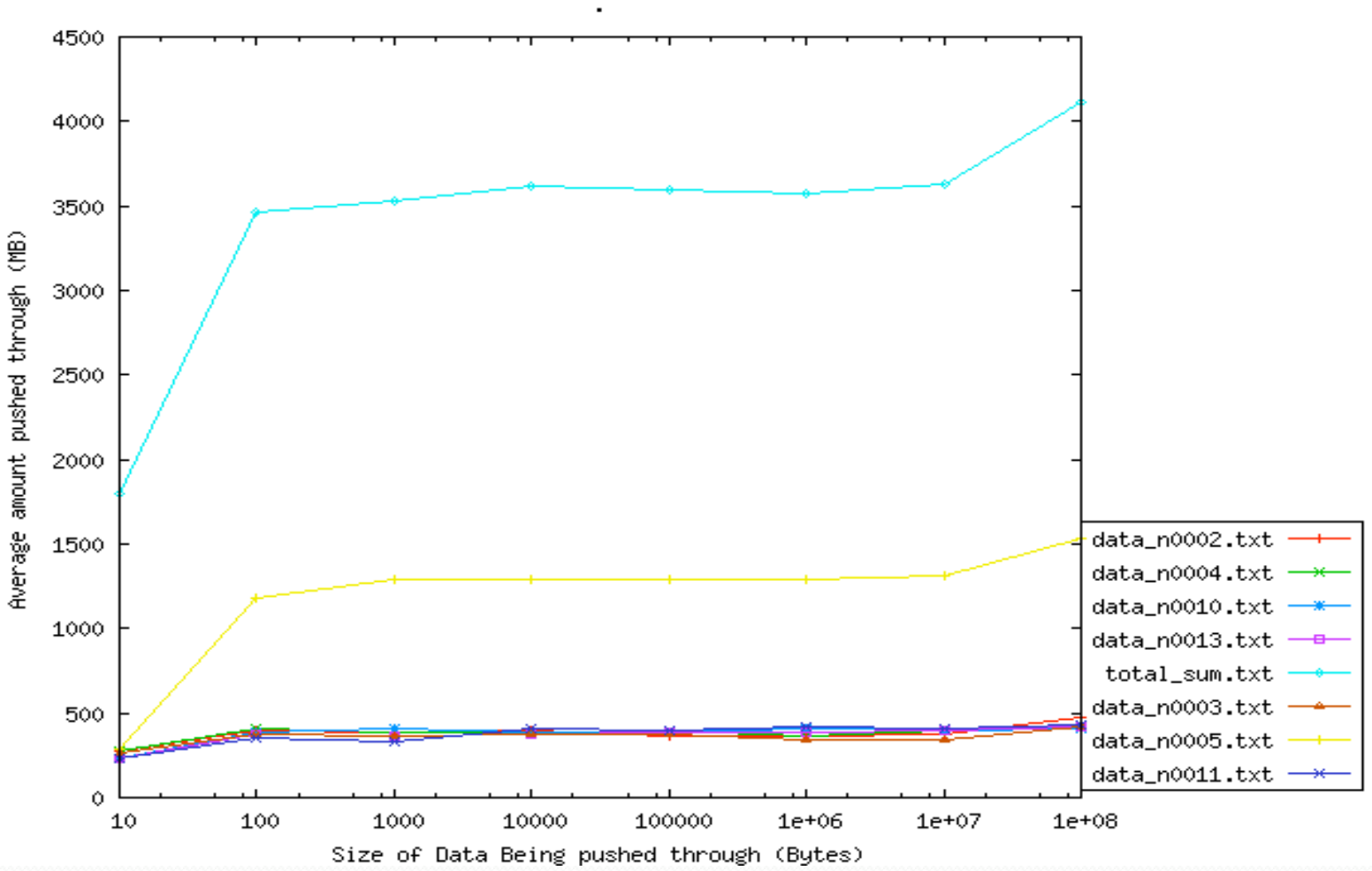
IB/QDR + Multiple 10-Gigabit bonding



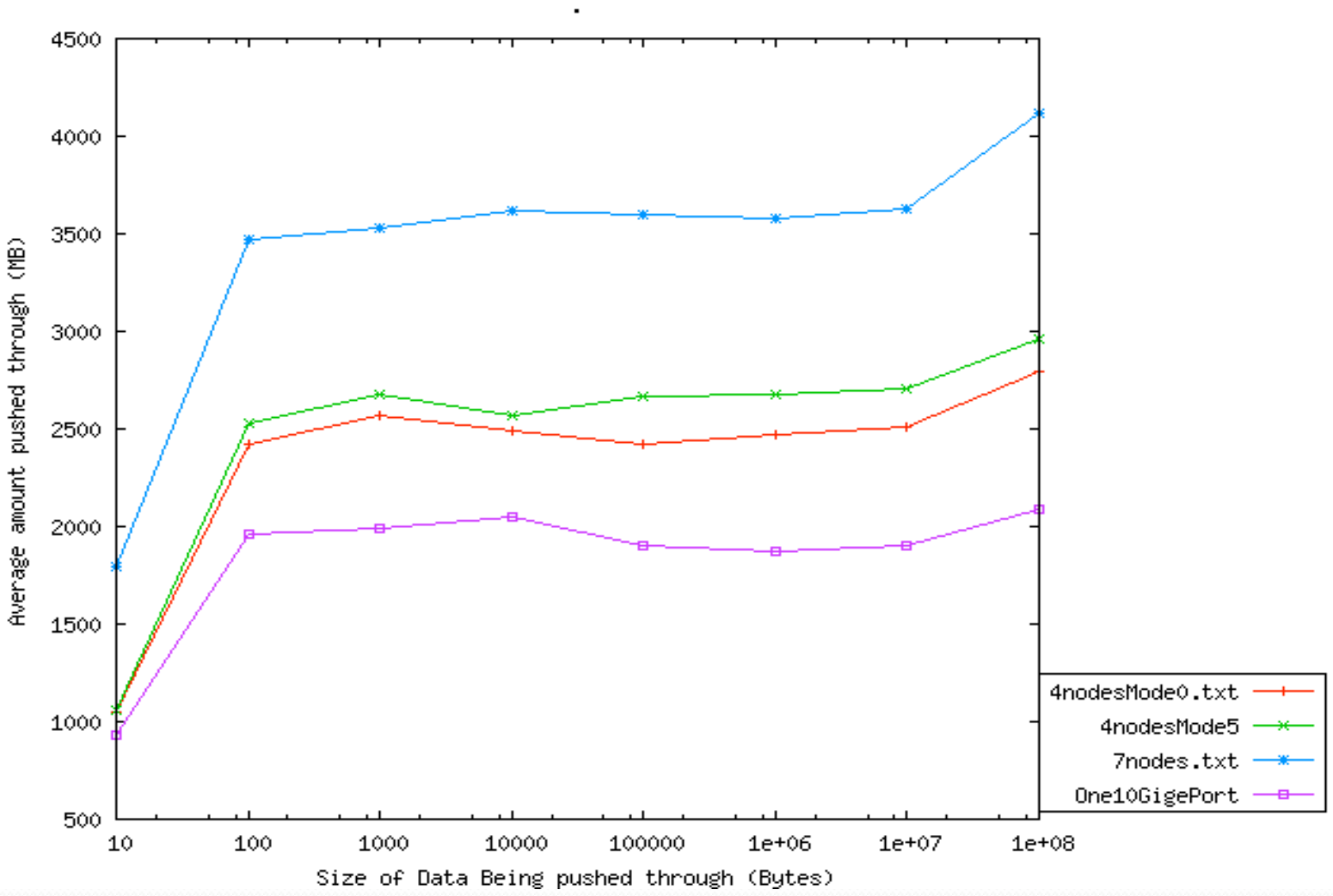
PaScalBB using Load Balancing in 4 compute nodes



PaScalBB using Load Balancing in 7 compute nodes



7 compute nodes , 4 compute nodes, and Dual Port 10-Gigabit card using Load Balancing



Conclusion

LANL's Today System: (Example)

12 I/O nodes + **12** DDR Cards + **12** 10-Gigabit

12 x .917 GB = 11.004 GB

Our System: (Example)

6 I/O nodes + **6** QDR Cards + **12** 10-Gigabit

6 x 1.97 GB = 11.4 GB

3 I/O nodes + **3** QDR Cards + **2** 10-Gigabit

3 x 3.5 GB = 10.5 GB

Questions?