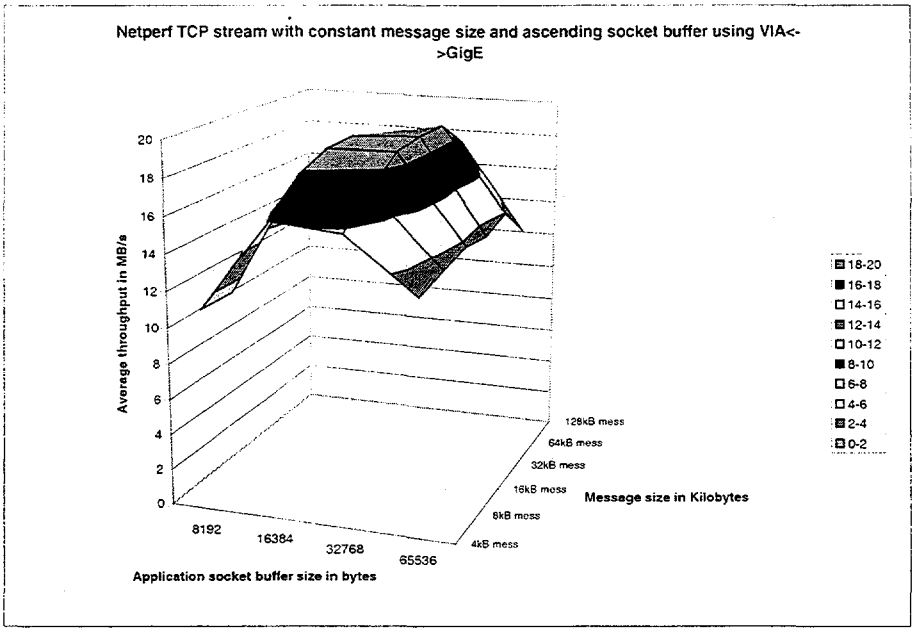

Figure 22 TCP combination results

For a more detailed look, Netpipe [7] was used to provide the signature graphs of the NDIS services over the Giganet VIA and Packet Engines Gigabit Ethernet hardware. Two Hamachi, a second generation Gigabit Ethernet NIC, were installed in placed of the first generation cards. From figure 23 the Gigabit Ethernet (GigE) had significant less bandwidth performance although the theoretical line speed was equal. This means that buffering and device tuning would be necessary for such a gateway to function effectively.

To examine the TCP/IP transport stack delay, the Giganet VIA latency and the Netpipe TCP stream tests were used. The following assumption were made during the analysis:

1. The latency introduced by address translation between VIA and TCP/IP is assumed to be negligible compared to the TCP/IP processing. For a real gateway, the lookup would be done via a hash table and set up a priori.
2. Transferring the data from the VIA to the TCP/IP stack would be done through a DMA copy from the VIA NIC to user memory, followed by a memory copy into the TCP/IP stack. The DMA transfer time is assumed to be negligible compared to the memory copy. The latency of the memory copy is included as part of the TCP/IP processing time.

The VIA latency test provides a baseline of the inbound VIA. Both the Giganet and the Netpipe application were using the same hardware setup. The time difference between the two provides an insight as to the latency added by the TCP/IP processing. The graph shows the time difference between polling and non-polling VIA latency versus the TCP/IP processing.

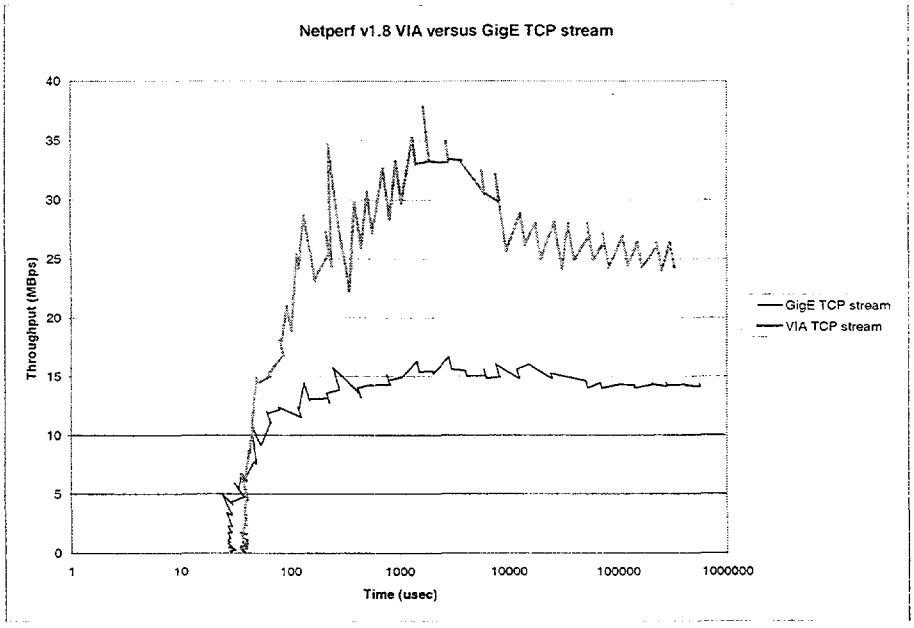

Figure 23 Netperf signature graph

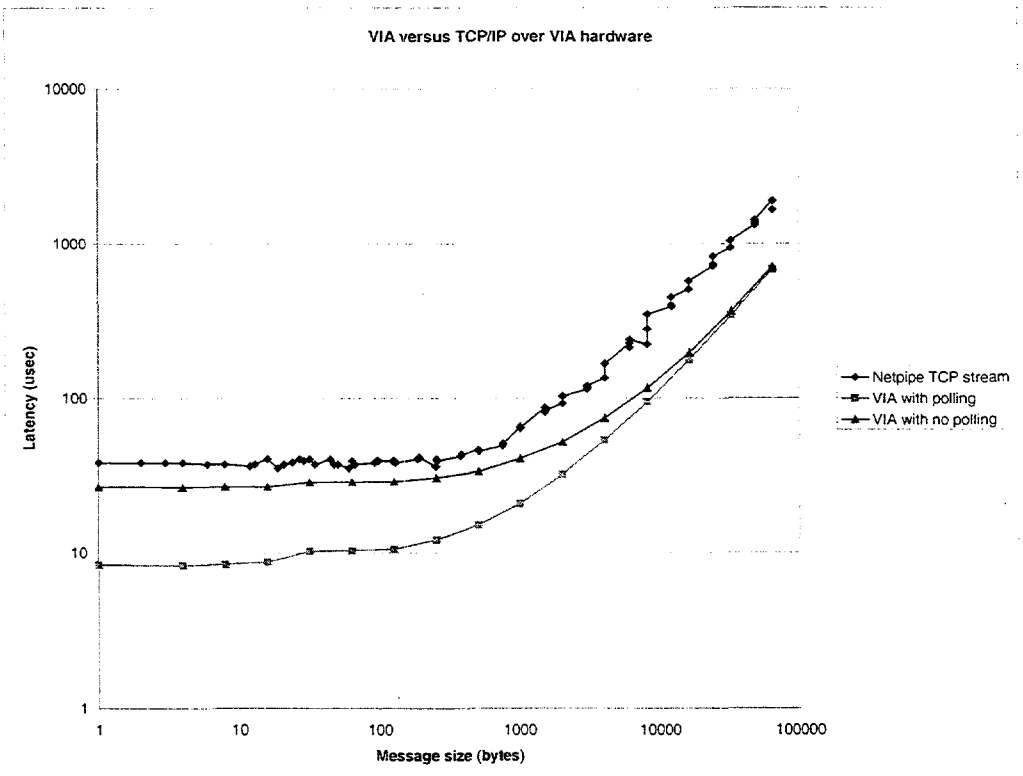

Figure 24 VIA versus TCP/IP latency over the same VIA hardware

From both the Netpipe and latency graphs we can see that significant buffering would be needed to compensate for the lower Gigabit Ethernet's throughput and higher TCP/IP stack processing time. Therefore to make the gateway viable for connecting the cluster to the legacy or production network, a gateway with specialized software and hardware needs to be developed.

CONCLUSION

To evaluate emerging communications models and the requirements for the computer cluster, a test-bed was constructed. The PC hardware was interconnected with both 100BaseT and VIA implementations from Giganet. In the course of the evaluation, it was found that a standardized VIA bandwidth and latency test was needed. The comparison of the different interconnect technology with the NAS NPB benchmark shows that the performance of the cluster interconnect was improved going from TCP/IP over 100BaseT Ethernet to VIA. Between the equivalent high speed interconnect of the Terraflop, Cplant, and VIA cluster, however, VIA has less of an impact on computation compared to the processor performance and the number of processors. The evaluation also showed that although the VIA implementation provides high bandwidth and low latency, significant work remains to be done in the areas of reliability, reparability, and scalability. In addition, special gateways have to be developed to connect the VIA cluster to standards based networks.

There are issues that arise as part of the evaluation that are not addressed by the work done to date. They should be considered for future work and are listed here for completeness:

1. Expansion of VIA tests to include a number of simultaneous connections.
2. Better software management tools to handle the cluster updates.
3. Special I/O gateway.
4. Evaluation of additional VIA products.

References

1. *The Virtual Interface Architecture Proof-of-Concept Performance Results*, FrankBerry, Ellen Deleganes, Anne Marie Merritt. Server Systems Technology. Intel Corporation.
2. *An Implementation and Analysis of the Virtual Interface Architecture*, Philip Buonadonna, Andrew Geweke, and David E. Culler. Department of Electrical Engineering and Computer Science. UC Berkeley.
3. The Virtual Interface Architecture Specification 1.0, 1997.
4. M-VIA: A High Performance Modular VIA for Linux. http://www.nersc.gov/research/FTG/via
5. Intel and VI Architecture Developer tools. http://www.intel.com/design/servers/vi/developer/ia_imp_guide.htm
6. NetPerf http://www.cup.hp.com/netperf/NetperfPage.html
7. *NETPIPE: A Network Protocol Independent Performance Evaluator*, Quinn O. Snell, Armin R. Mikler and John L. Gustafson. Ames Laboratory/Scalable Computing Lab, Ames, Iowa 50011, USA.
8. *Interconnection Networks: An Engineering Approach*, Jose Duato, Sudhakar Yalamanchili, Lionel Ni. IEEE Computer Society Press ISBN 0-8186-7800-3.
9. *A comparison of Large-Scale Software Installation Methods on NT and UNIX*, Michail Gomberg, Remy Evard, and Craig Stacey, Proceedings of the Large Installation System Administration of Windows NT Conference, 1998.
10. *Scalable, Remote Administration of Windows NT*, Michail Gomberg, Craig Stacey, and Janet Syre, Proceedings of the 2nd Large Installation System Administration of Windows NT Conference, 1997.
11. *A Networked Machine Management System*, Dave Roth, Proceedings of the 2nd Large Installation System Administration of Windows NT Conference, 1997.
12. *Lights-Out Operations Guide for Microsoft NT Server*, Microsoft Corporation, 1998.
13. *Microsoft Windows NT Workstation Deployment Guide: Automating Windows NT Setup*, Microsoft Corporation, 1997.

DISTRIBUTION:

20	MS 0806	L. Stans,4616
10	MS 0806	R. Hu, 4616
10	MS 0449	T. D. Tarman, 4616
1	MS 0806	B. Dean, 4616
5	MS 0806	J. Brenkosh, 4616
1	MS 0806	J. Eldridge, 4616
1	MS 0806	M. J. Ernest, 4616
1	MS 0806	S. Gossage, 4616
1	MS 0806	J Hudson, 4616
1	MS 0806	L.G. Martinez ,4616
5	MS 0806	M. Miller, 4616
1	MS 0806	J. Naegle, 46126
1	MS 0806	L.G. Pierson, 4616
1	MS 0806	T. Pratt, 4616
1	MS 0806	J.A. Schutt, 4616
1	MS 0806	L.F. Tolendino, 4616
1	MS 1393	J. F. Jones, 9000
5	MS 0806	L. G. Pierson, 4616
1	MS 0806	M. R. Sjulin, 4614
1	MS 0801	M. O. Vahle, 4600
5	MS 0806	E. L. Witzke, 4616
1	MS 1110	R. Riesen, 09223
1	MS 0321	A. Hale, 09224
1	MS 0807	J. Noe, 4618
1	MS 9217	R. Clay, 8920
1	MS 0449	R. L. Hutchinson, 6236
1	MS 9003	D. L. Crawford, 9900
1	MS 9003	K. E. Washington, 8900
1	MS 9011	B. V. Hess, 8910
10	MS 9011	H. Y. Chen, 8910
5	MS 9011	P. S. Wyckoff, 8910
1	MS 9011	P. W. Dean, 8903
1	MS 9011	P. E. Nielan, 8920
1	MS 9012	S. C. Gray, 8930
1	MS 9037	J. C. Berry, 8930-1
1	MS 9019	B. A. Maxwell, 8940
1	MS 9217	J. C. Meza, 8950
1	MS 9019	J. A. Larson, 8970
1	MS 9012	K. R. Hughes, 8990
1	MS 9001	M. E. John, 8000
		Attn: R. C. Wayne, 2200
		J. Vitko, 8100
		W. J. Mclean, 8300

D. Henson, 8400
P. N. Smith, 8500
T. M Dyer, 8700

1	MS 0188	LDRD Office, 4001
1	MS 9018	Central Technical File, 8940-2
2	MS 0899	Technical Library, 4916
1	MS 0612	Review and Approval Desk, 4912 For DOE/OSTI