



Ruzhu Chen

PMB-2.2.1 Benchmarking on POWER4+ Platforms p655+ and p690+

Problem

The benchmark performance on the POWER4+™ platforms p690+ and p655+ platforms was evaluated using the PMB-2.2.1 benchmark written by Pallas.

Proposed solution

Pallas has written a comprehensive set of MPI benchmarks, known as PMB, which has the following objectives:

- ▶ Providing a concise set of benchmarks for measuring the MPI functions of point-to-point message-passing, global data movement and computation routines, one-sided communications, file I/O
- ▶ Establishing precise benchmark procedures, including run rules, a set of required results, repetition factors and message lengths
- ▶ Avoiding interpretation of the measured results: execution time, throughput, global operations performance

For a complete explanation and interpretation of PMB benchmark results, refer to the PMB-MPI1.pdf and to the PMB-MPI2.pdf. For detailed results and output logs, refer to the output and log files in the directory PMB2.2.1-mpi, available at:

<http://www.pallas.com/e/products/pmb/index.htm>

System configuration

The PMB2.2.1 benchmark was tested on the IBM® POWER4+ platforms p690+ and p655+.

Table 1 lists the details of the configurations of these platforms as used in this benchmark.

Table 1 System and hardware configurations

Configurations		P690+	P655+
Processor		1.7 GHz Power4+	1.5GHz POWER4+
Processors/node		32	
Memory/node		128 GB (8-card)	16 GB (2-card)
Mem(GB)/processor			
Caches	L1	64/32 KB (1-way/2-way)	64 / 32 KB (1-way/2-way)
	L2	1.5 MB/card (4-way)	1.5 MB/card (4-way)
	L3	128 MB	128 MB
OS		AIX® 5.1.0.0	AIX 5.1.0.0
AIX Kernel		64-bit	64-bit
File system(s)		Local or gpfs	Local or gpfs
FORTRAN compiler		XLF 8.1	XLF 8.1
C/C++ compiler		VAC 6.0	VAC 6.0

Measurement and results

Our testing gave the following results.

Example 1 Compilation

```

MPI_HOME    = /usr/lpp/ppe.poe/
MPI_INCLUDE = $(MPI_HOME)/include
LIBS        = -bmaxdata:0x70000000 -bmaxstack:0x10000000 -lm
CC          = mpcc_r
CLINKER     = mpcc_r
OPTFLAGS    =
CPPFLAGS    = -DnoCHECK

```

Example 2 Run script

```

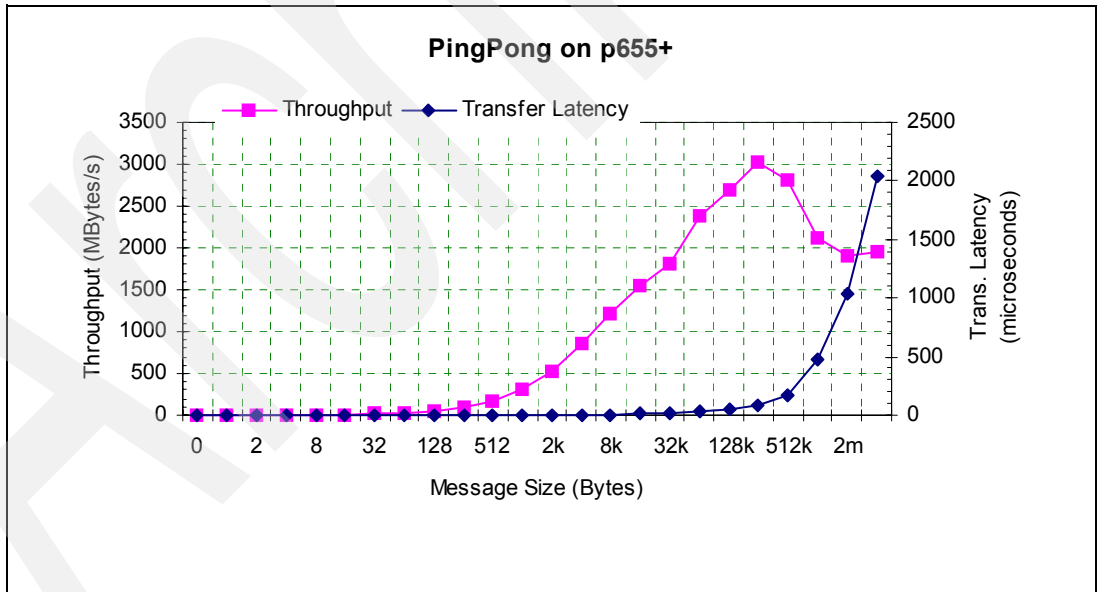
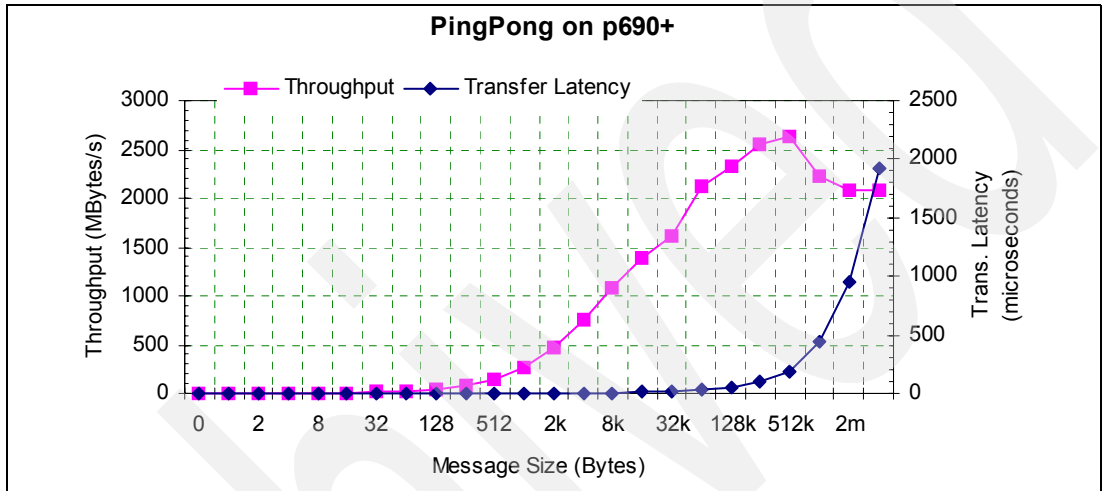
export MP_EULIB=us
export MP_EUIDEVICE=csss
export MP_INFOLEVEL=0
export MP_SHARED_MEMORY=yes
export MP_STDINMODE=none
export MP_EAGER_LIMIT=65536 #(try this to see if performance can be )
export MP_BUFFER_MEM=67108864 #(set this when MP_EAGER_LIMIT is set)
export MP_WAIT_MODE=poll #(need to set this when MP_EULIB=ip )
export MP_HOSTFILE=host.list
export MP_PROCS=$1
PMB-MPI1 (or PMB-IO, PMB-EXT)

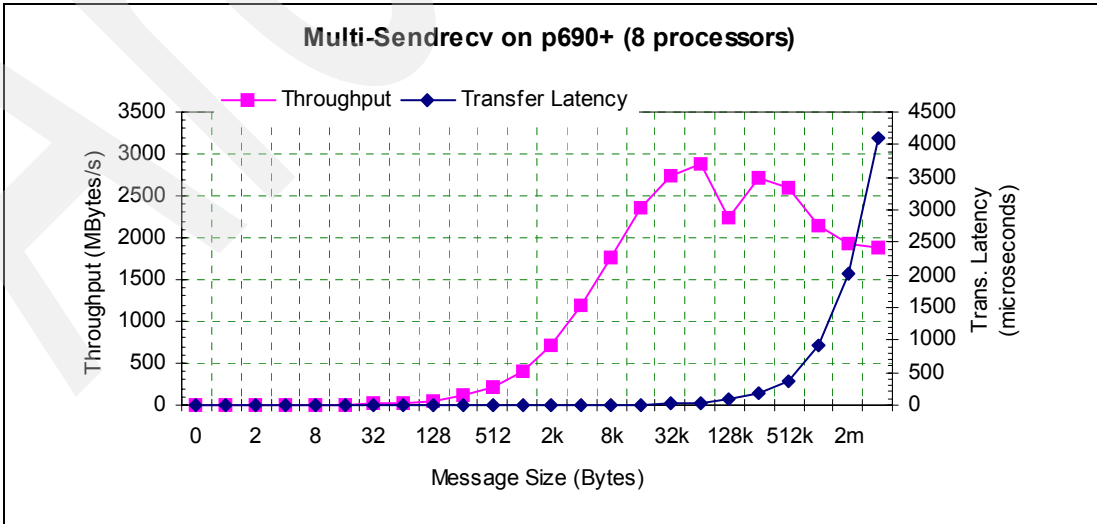
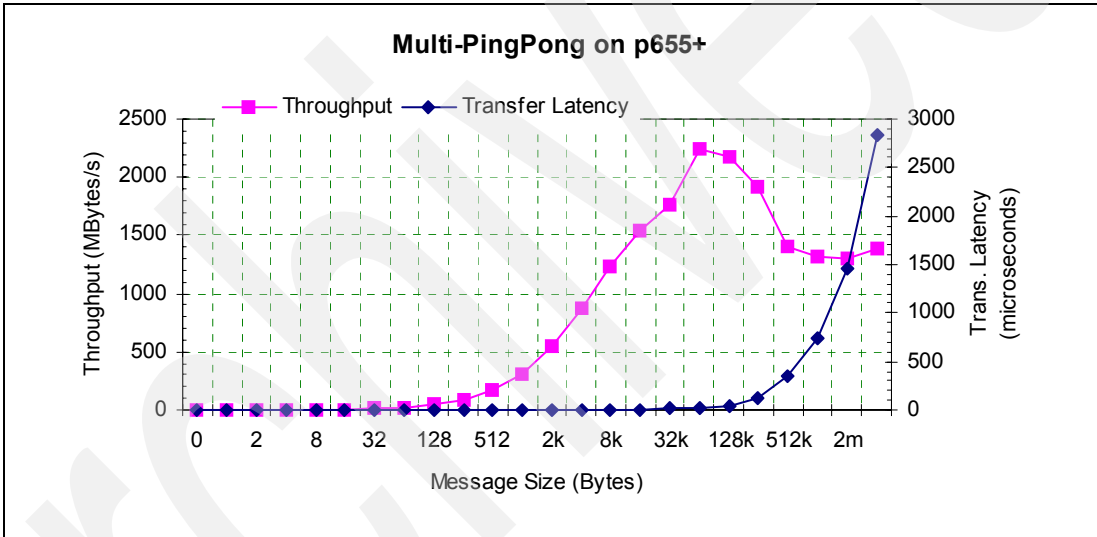
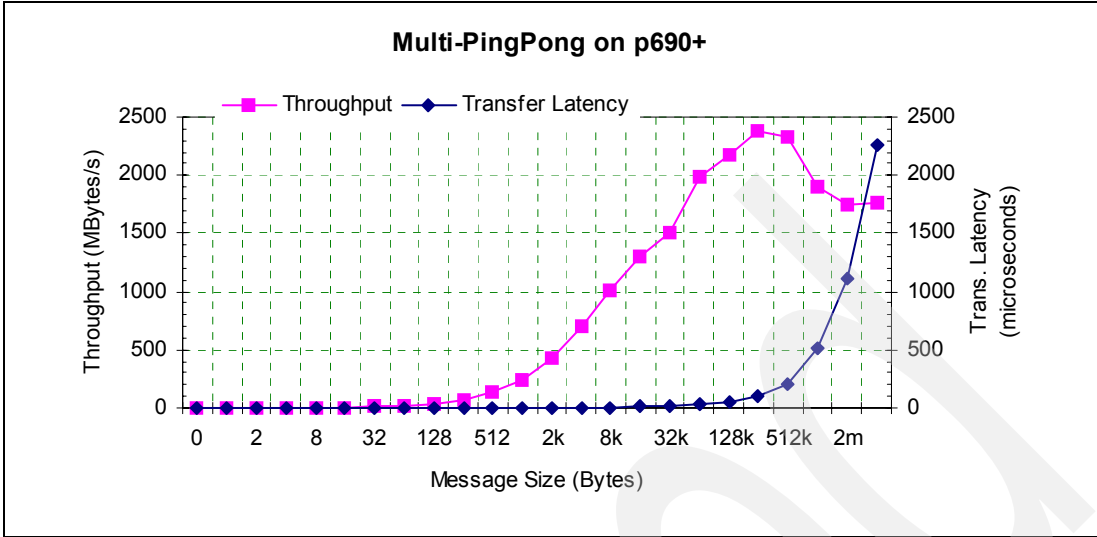
```

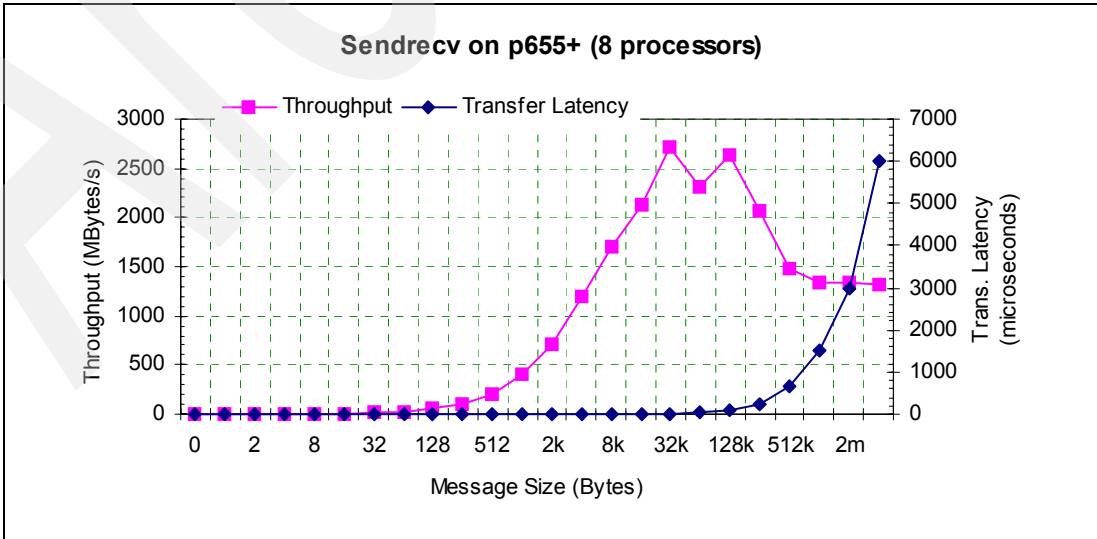
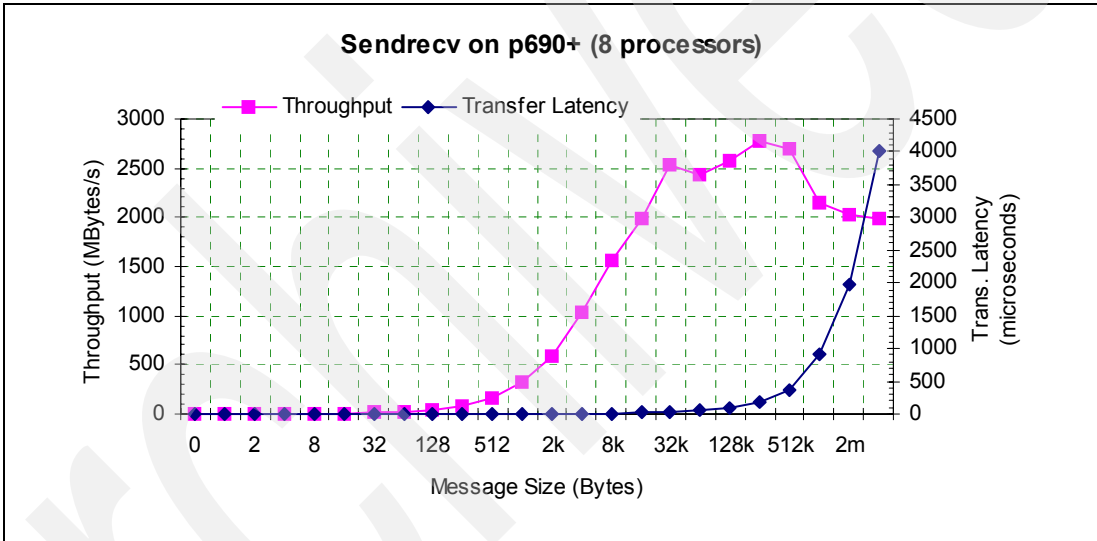
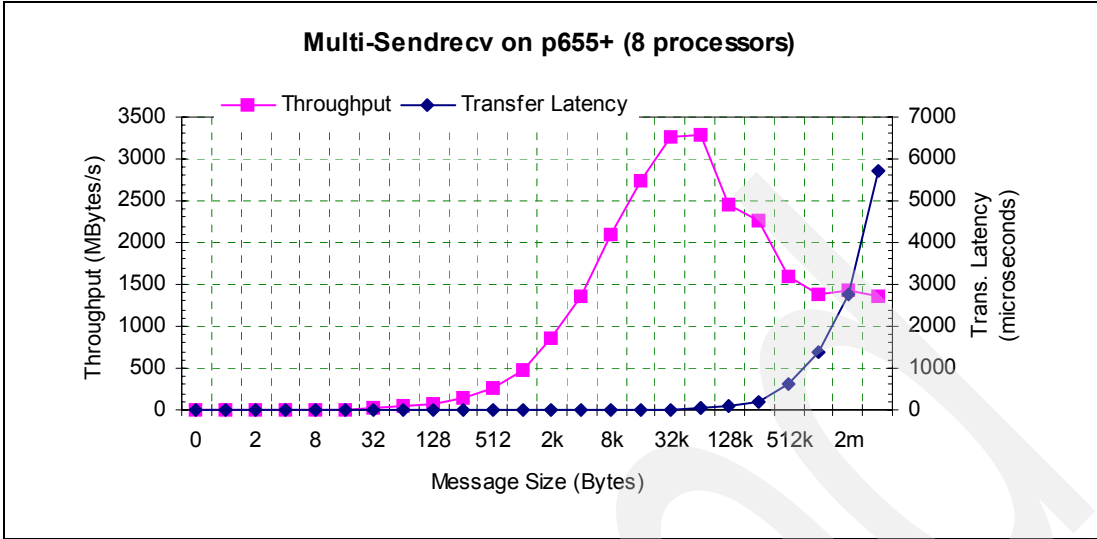
Point-to-point performance

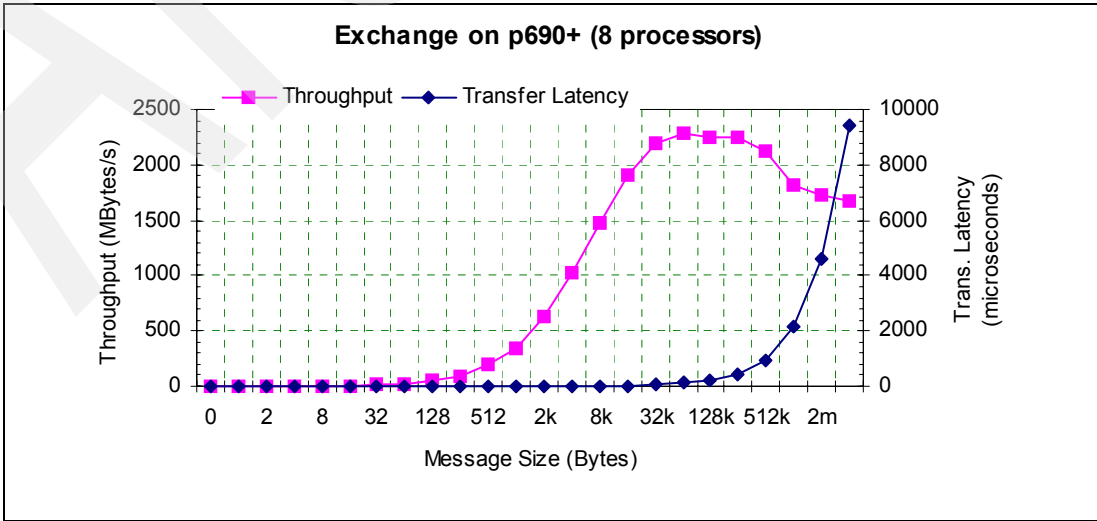
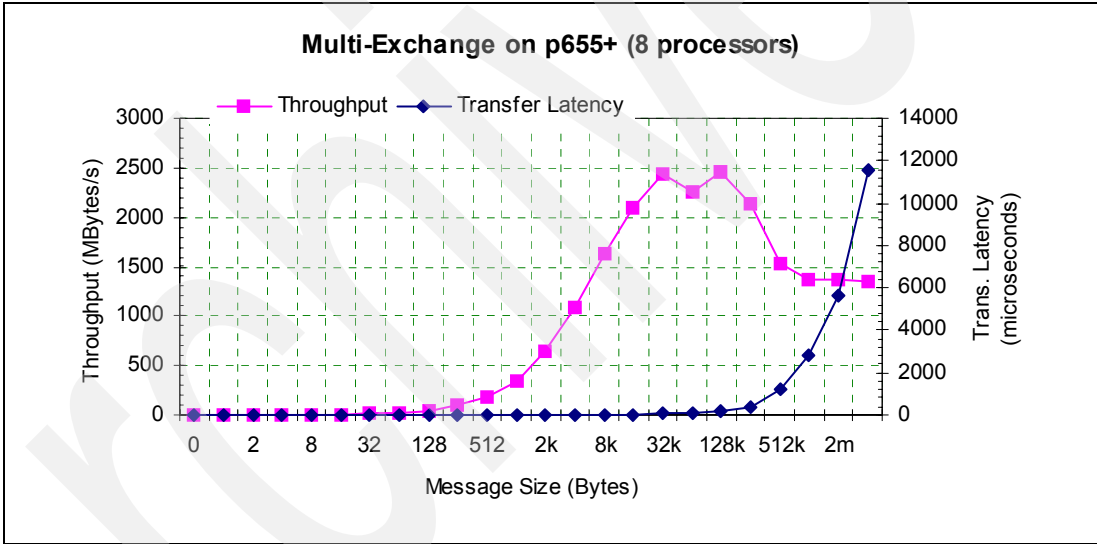
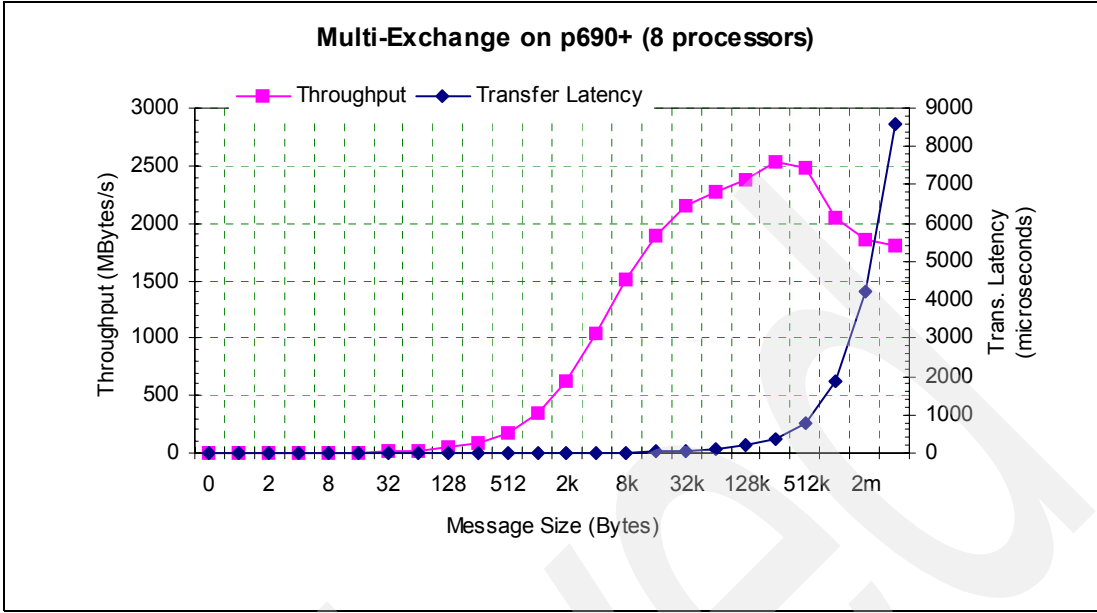
Point-to-point performance is measured between two processes within the same node (memory performance), or between two nodes (network performance). The performance is measured in MBytes/s per process (send+recv) in units of microseconds.

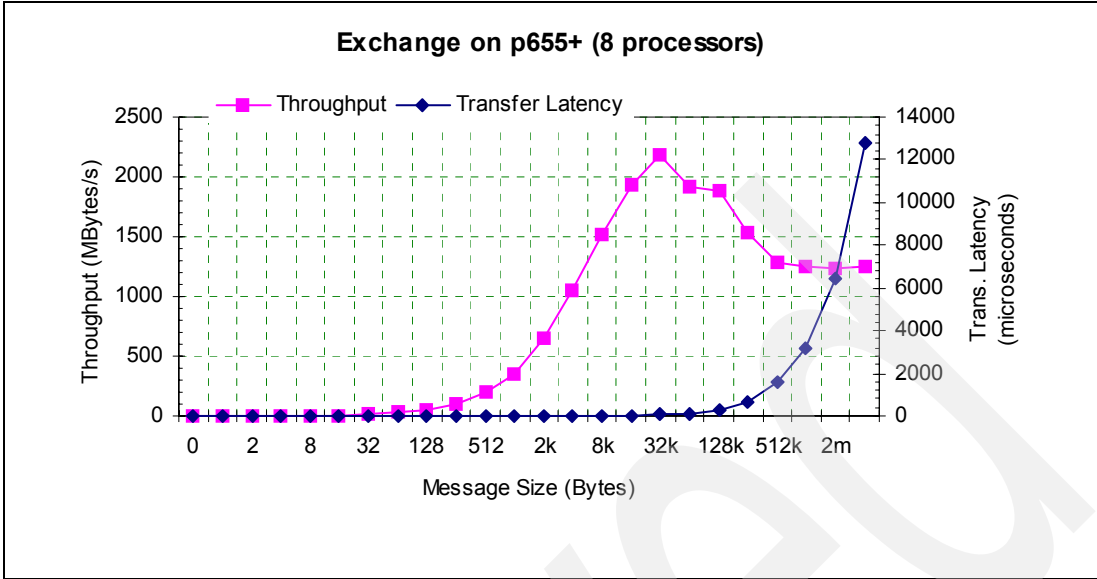
The following series of graphs illustrate the performance of PingPong, Multi-PingPong, Multi-Sendrecv, Sendrecv, Multi-Exchange and Exchange on p690+ and on p655+.







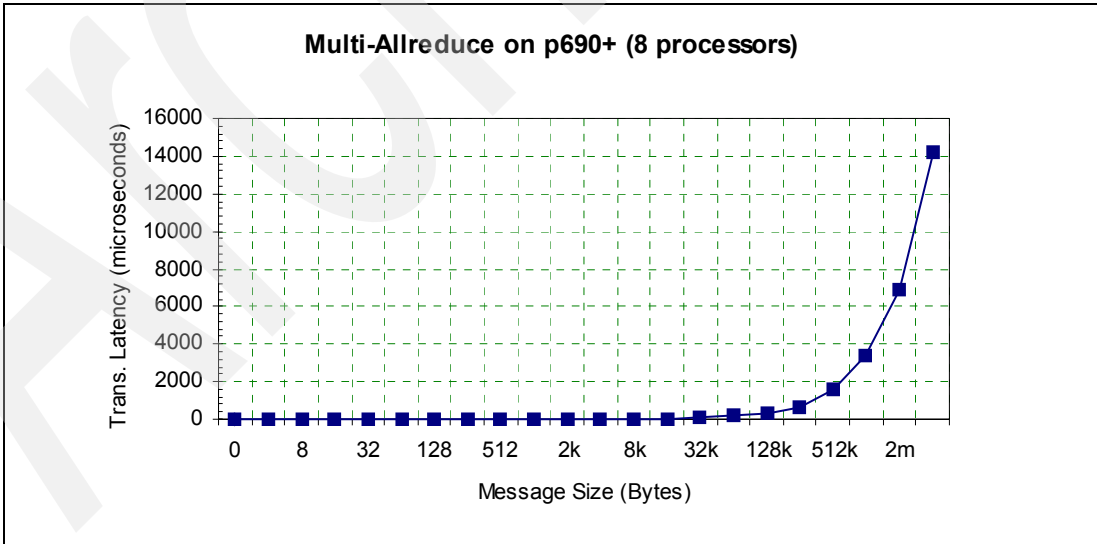


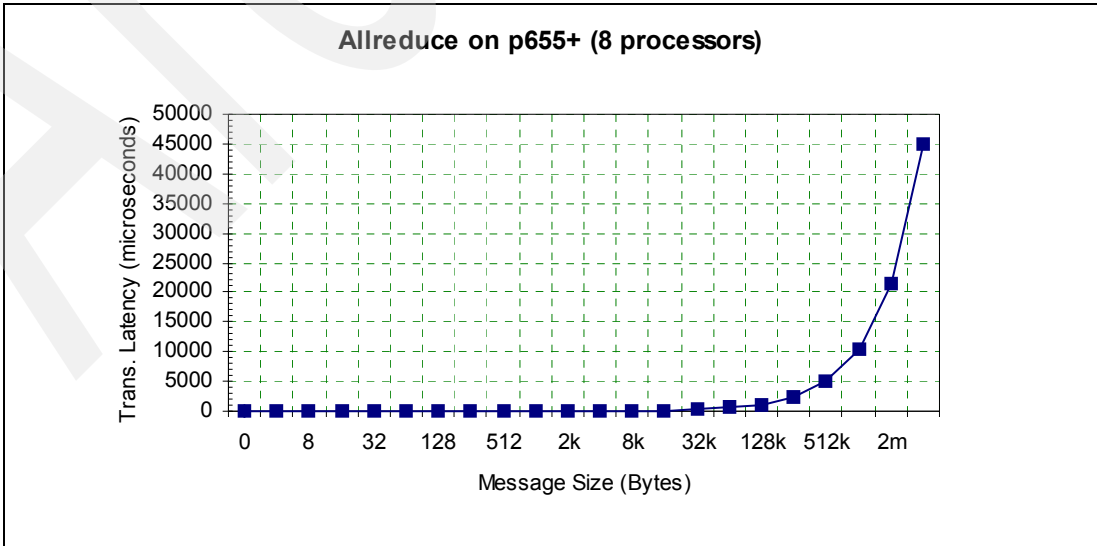
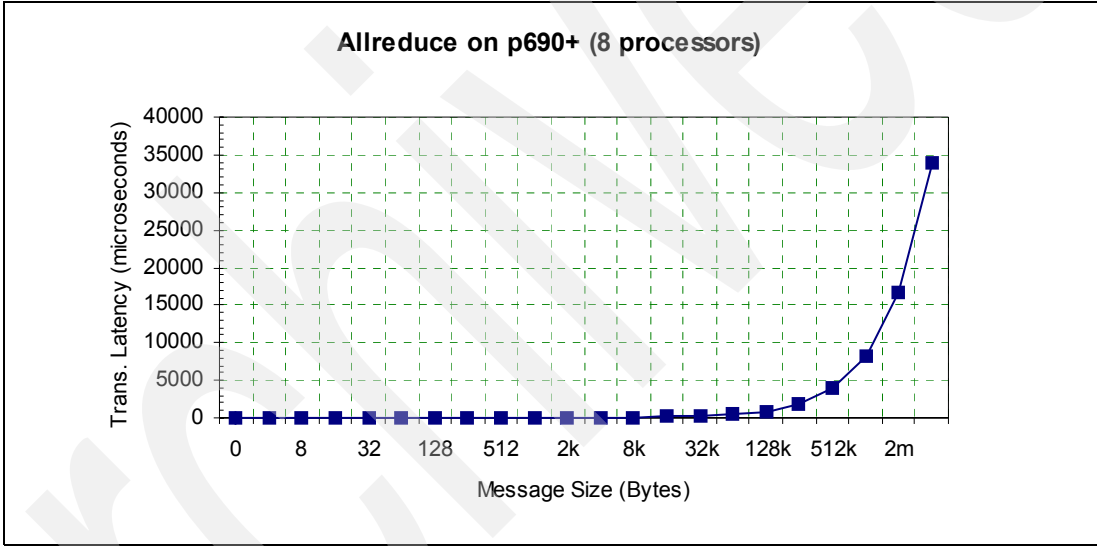
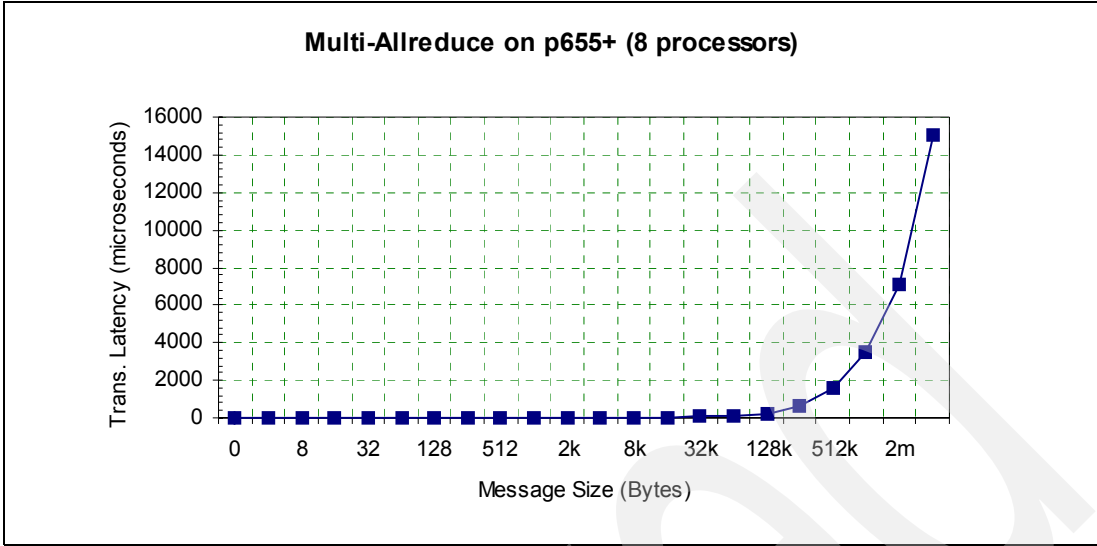


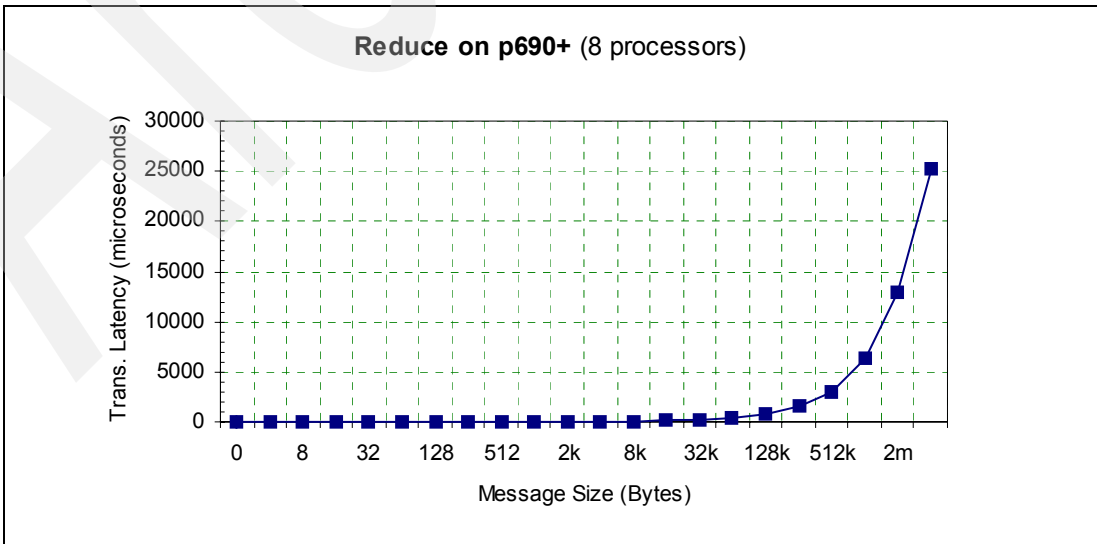
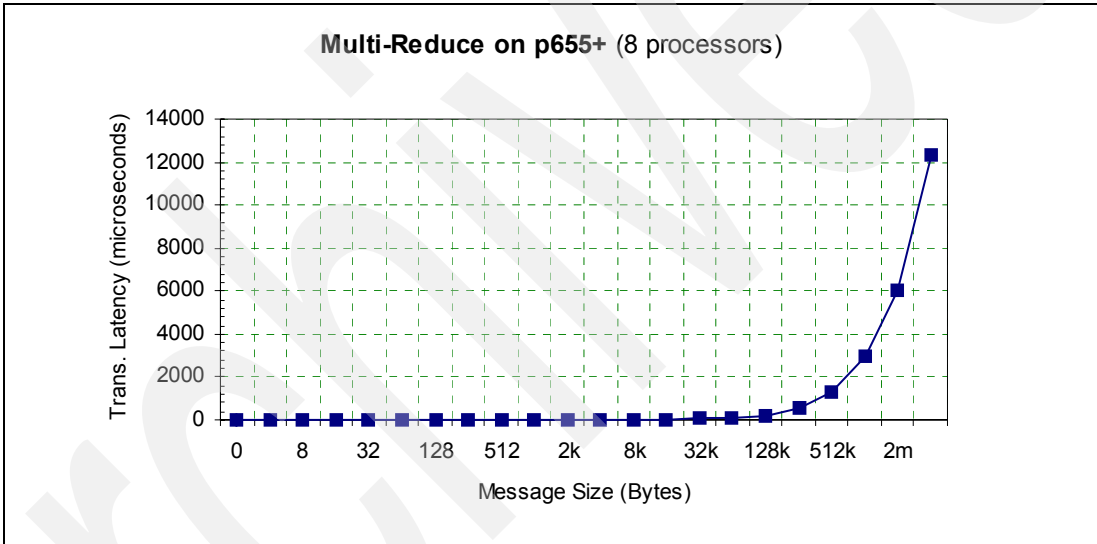
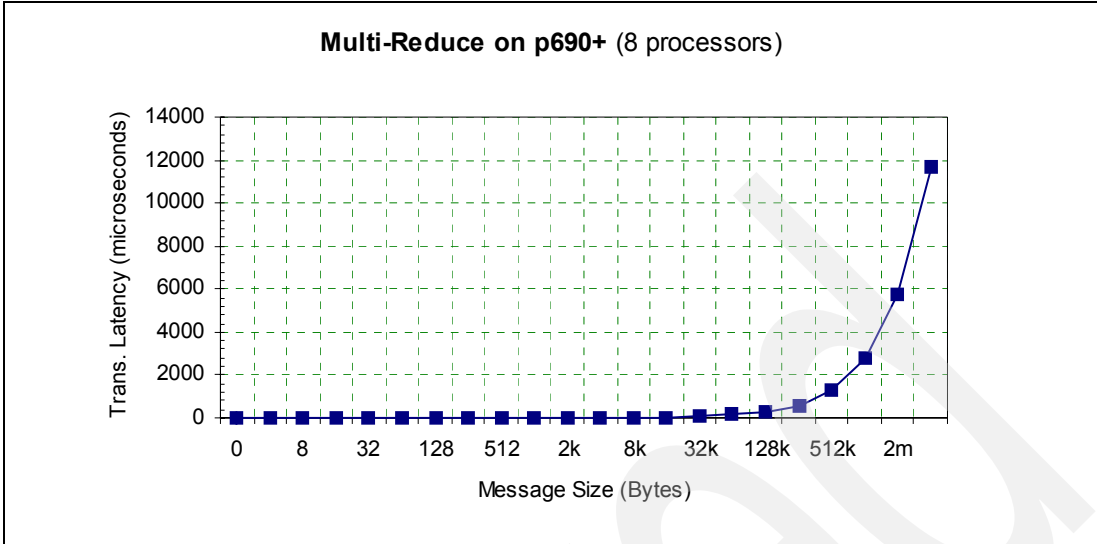
Collective benchmarks

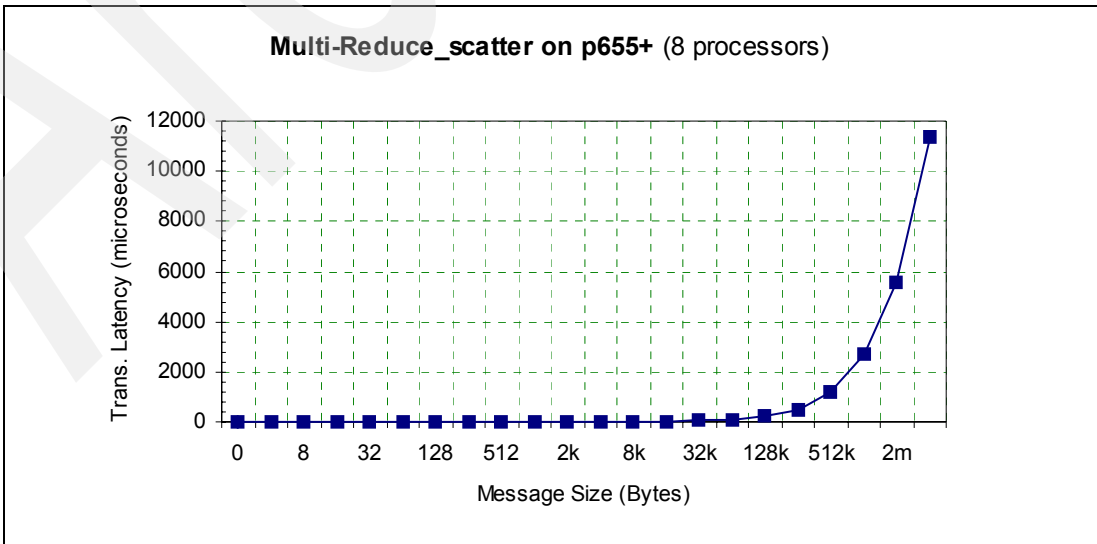
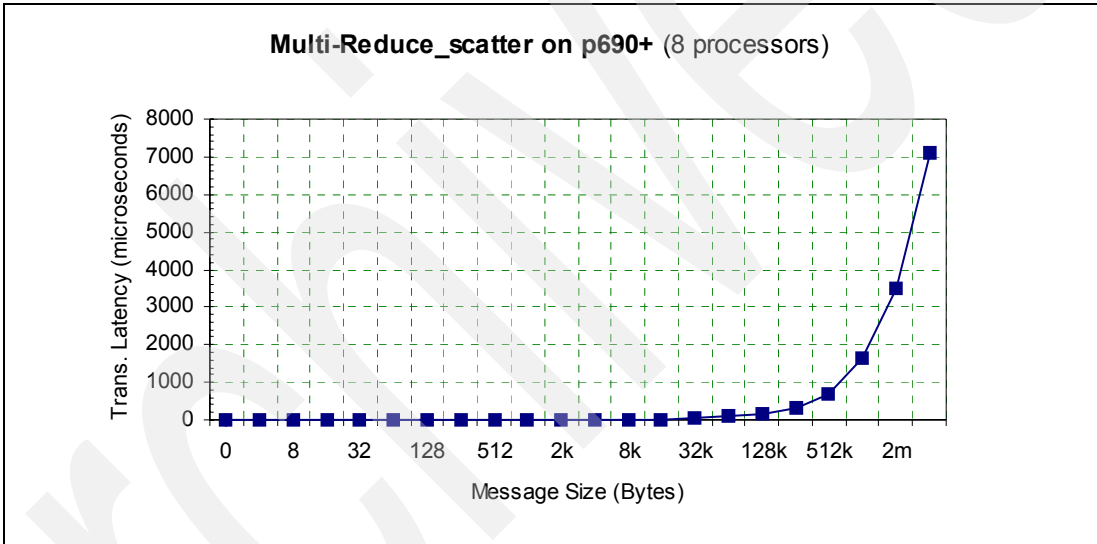
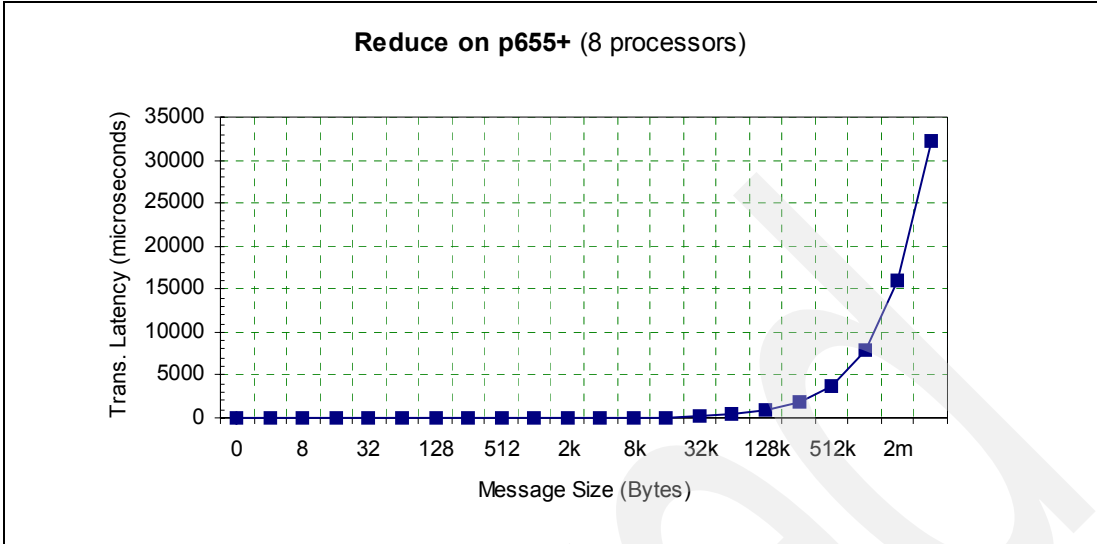
Collective or system-wide interconnect performance is measured between all or a subset of the nodes in the system. All collective benchmarks are measured in Microseconds transfer latency.

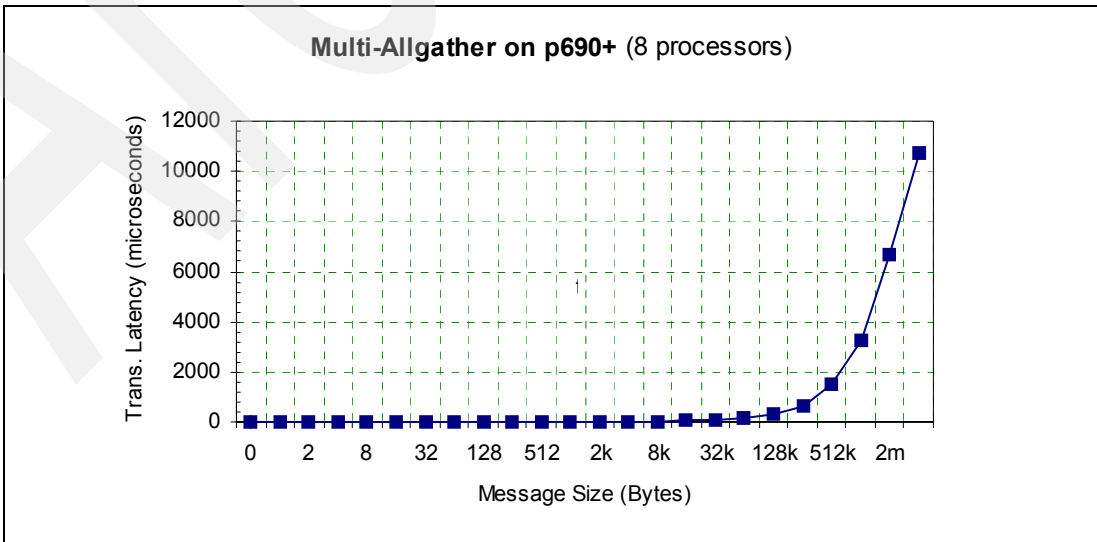
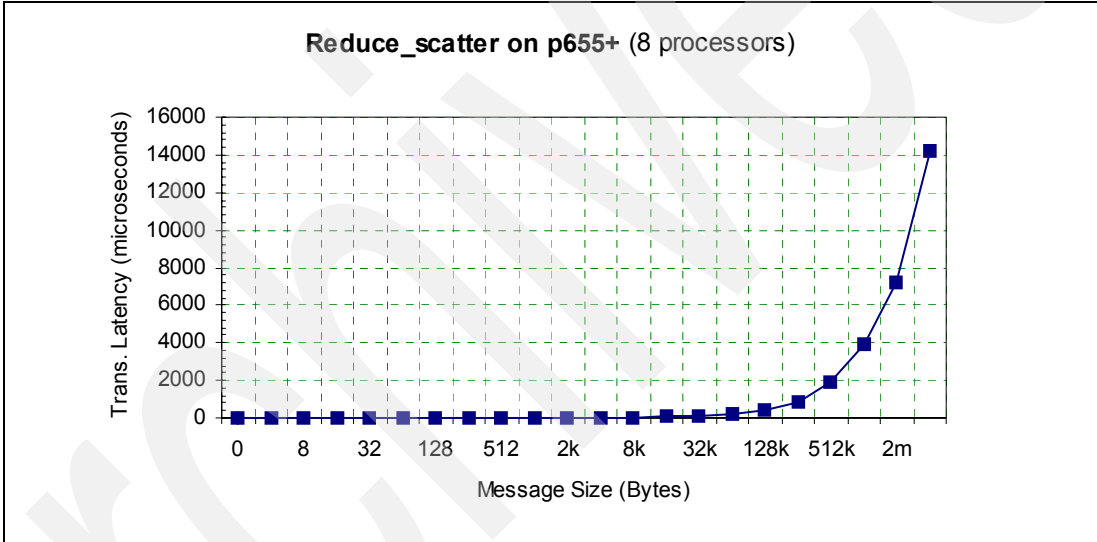
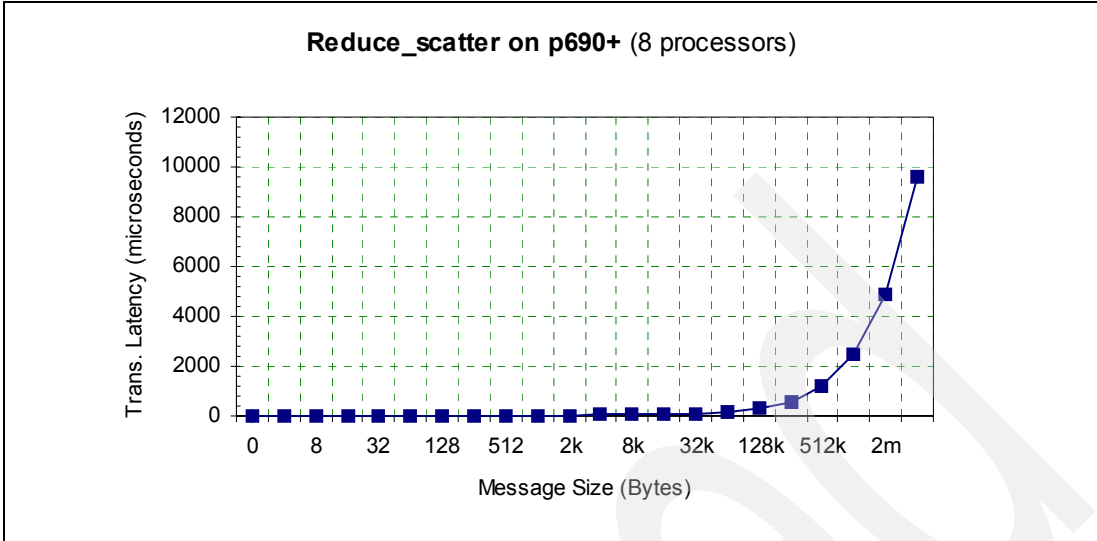
The following series of graphs illustrate the performance of Multi-All reduce, Allreduce, Multi-Reduce, Reduce, Multi-Reduce_scatter, Reduce_scatter, Multi-Allgather, Allgather, Multi-Allgatherv, Allgatherv, Multi-Alltoall, Alltoall, Multi-Bcast, and Bcast on p690+ and on p655+.

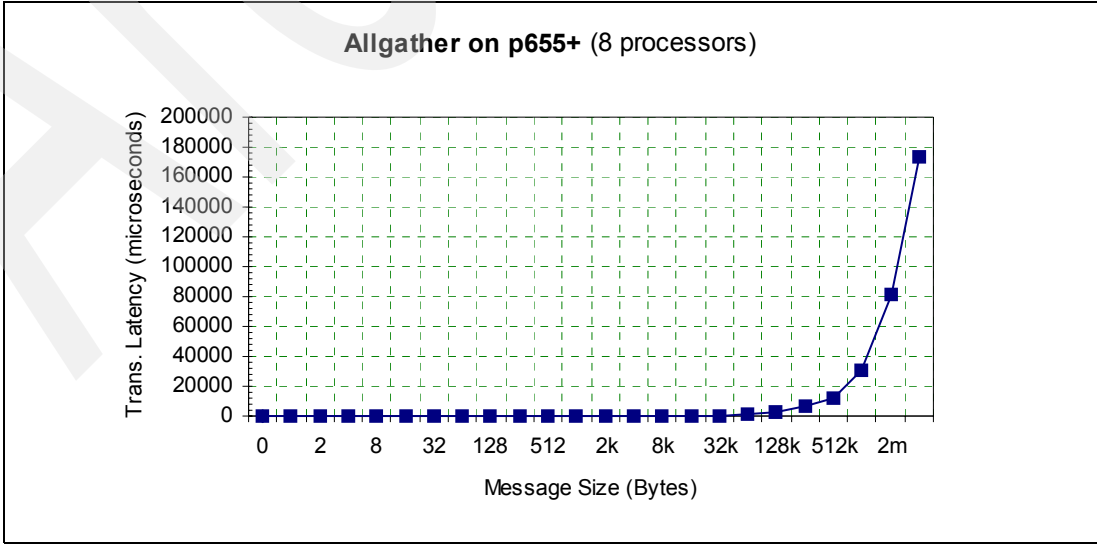
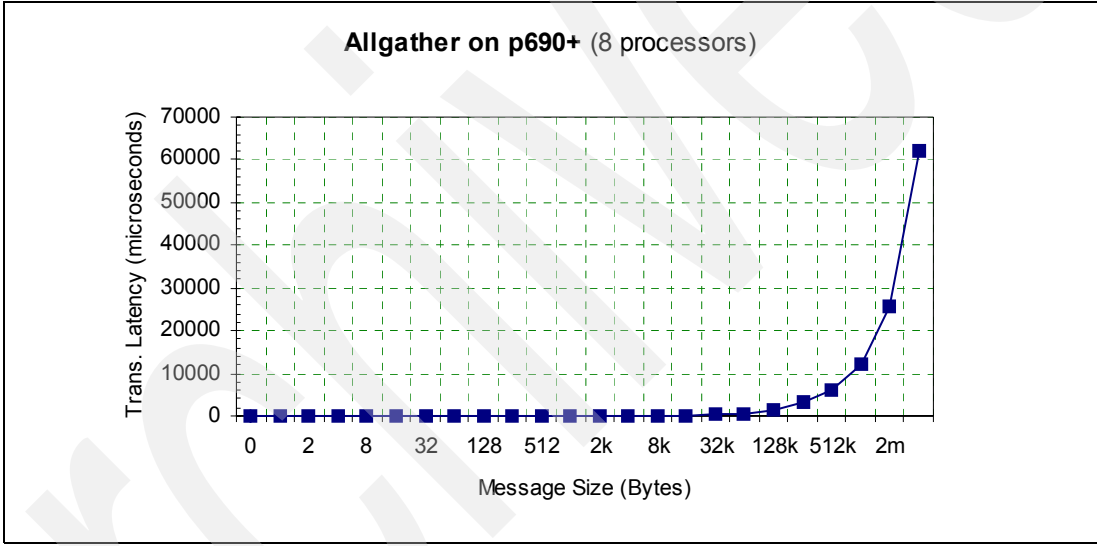
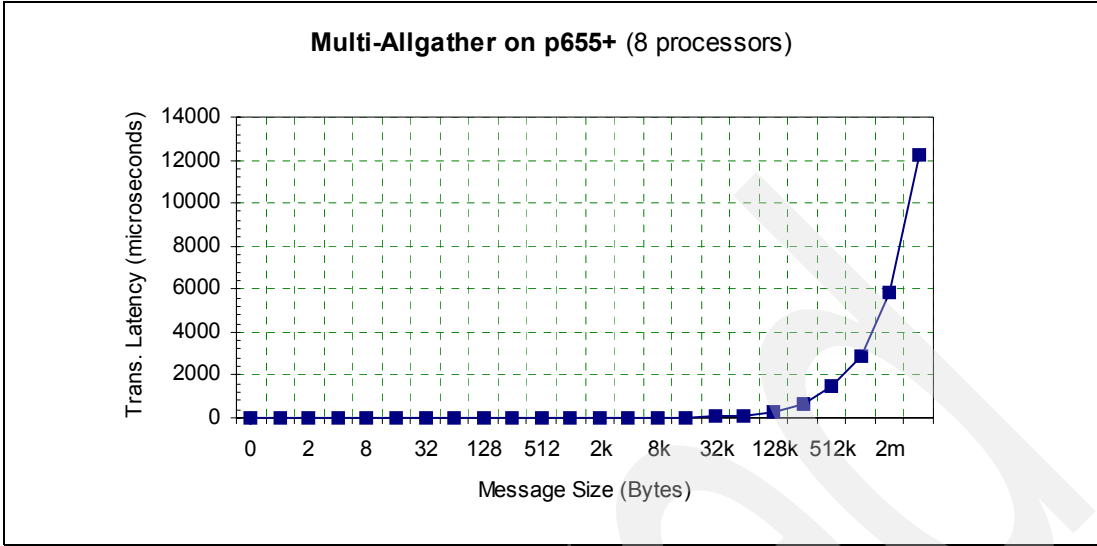


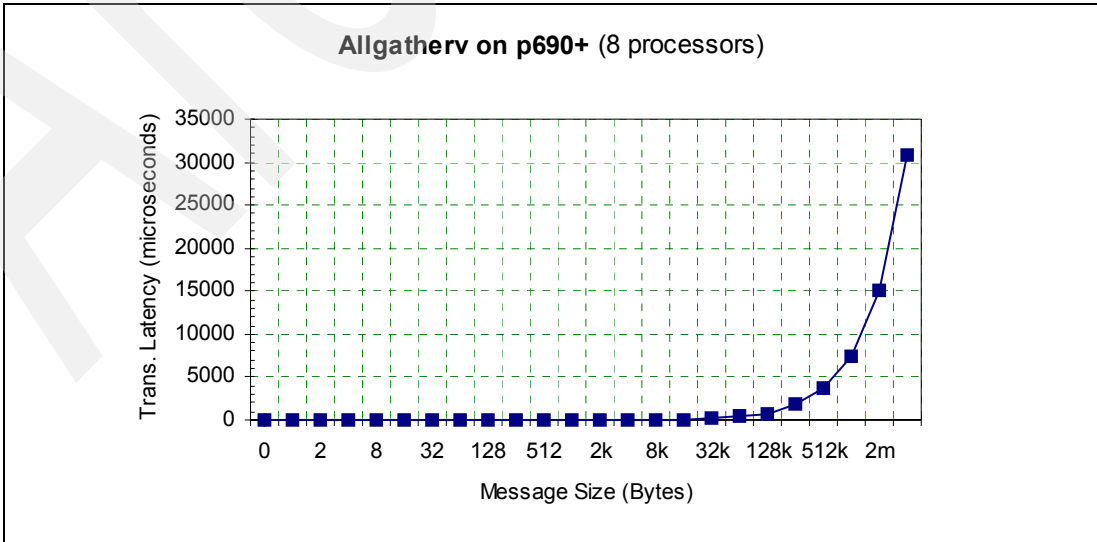
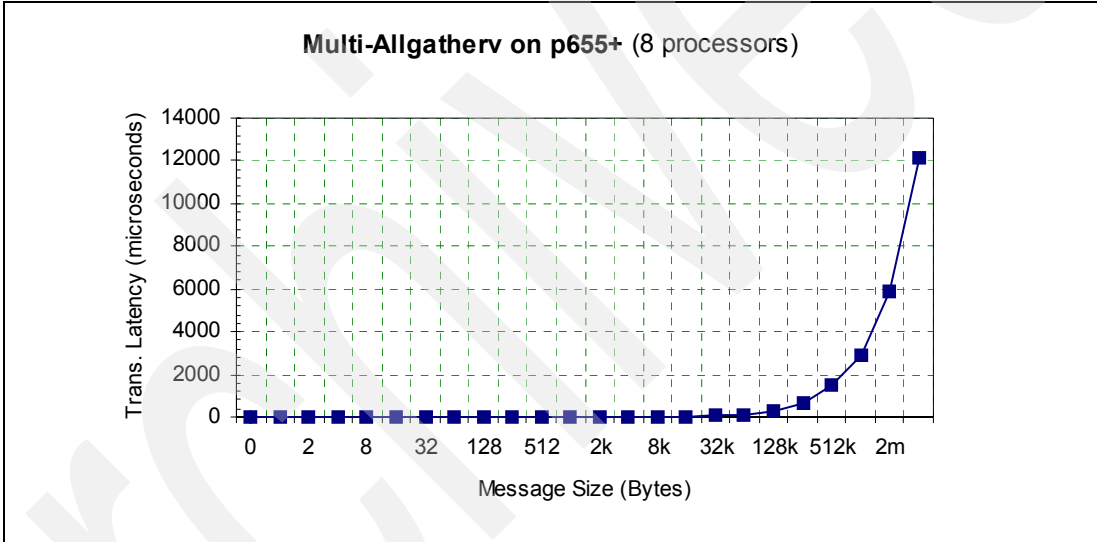
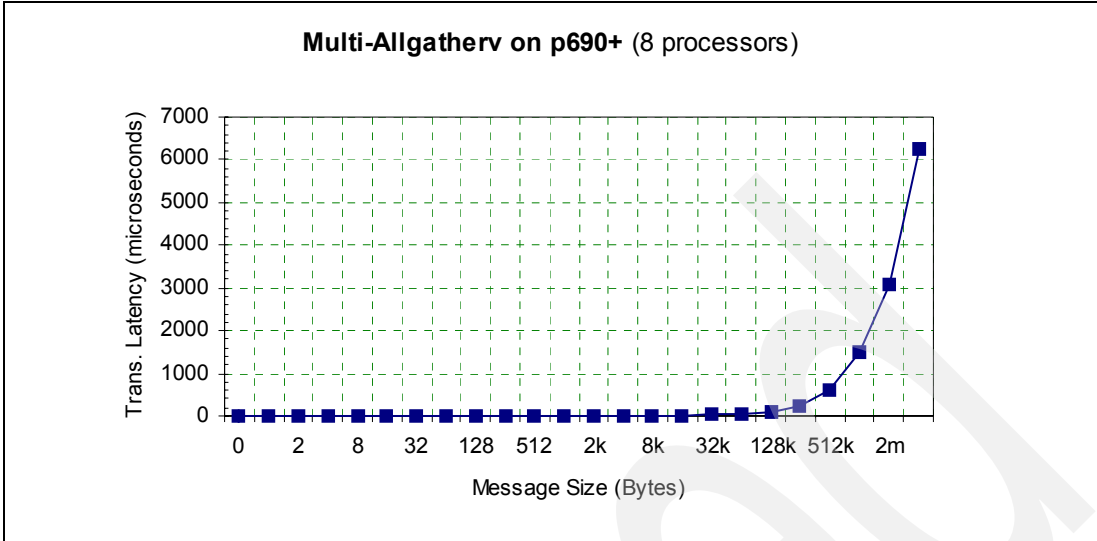


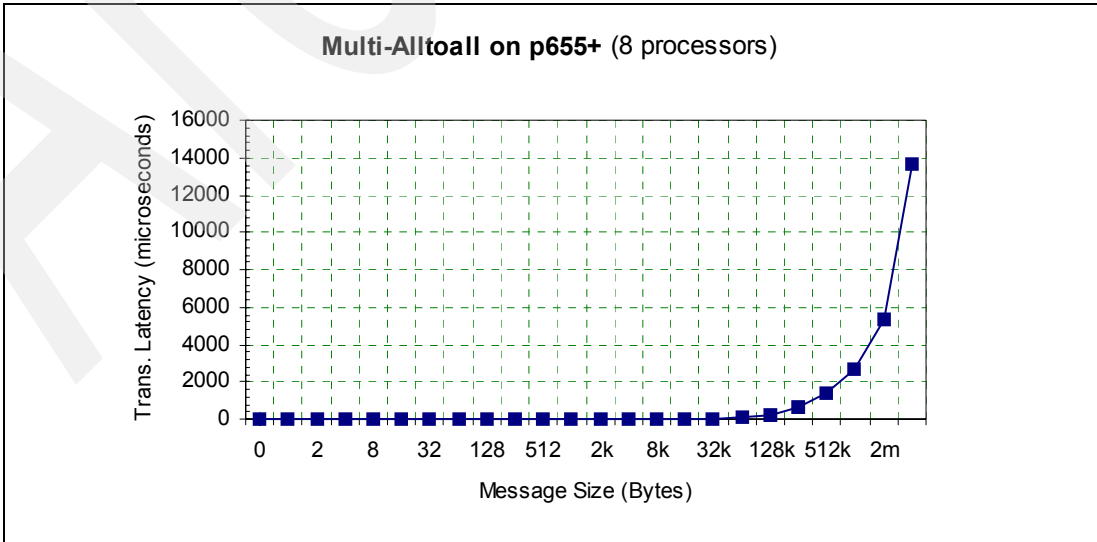
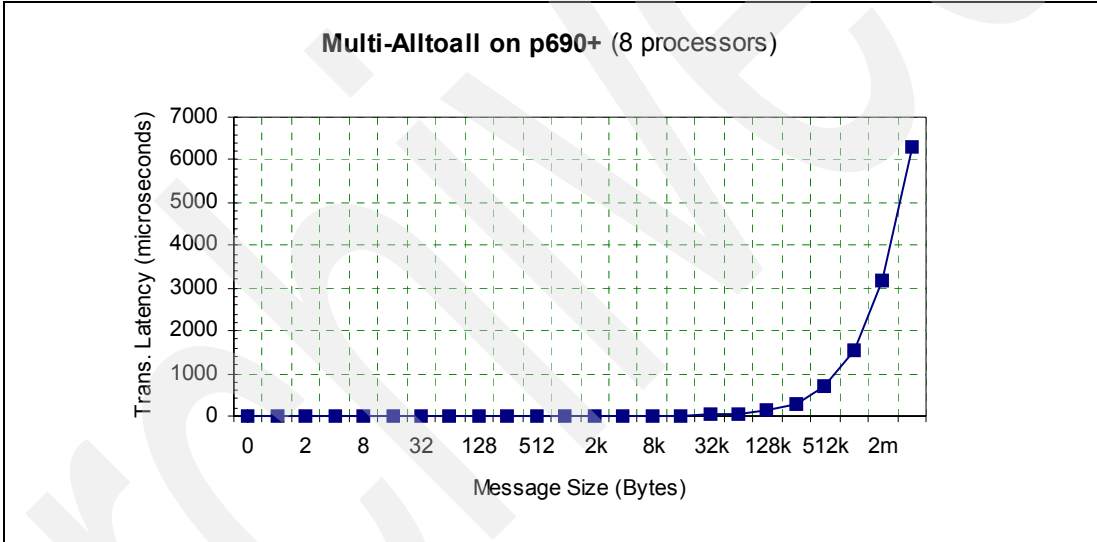
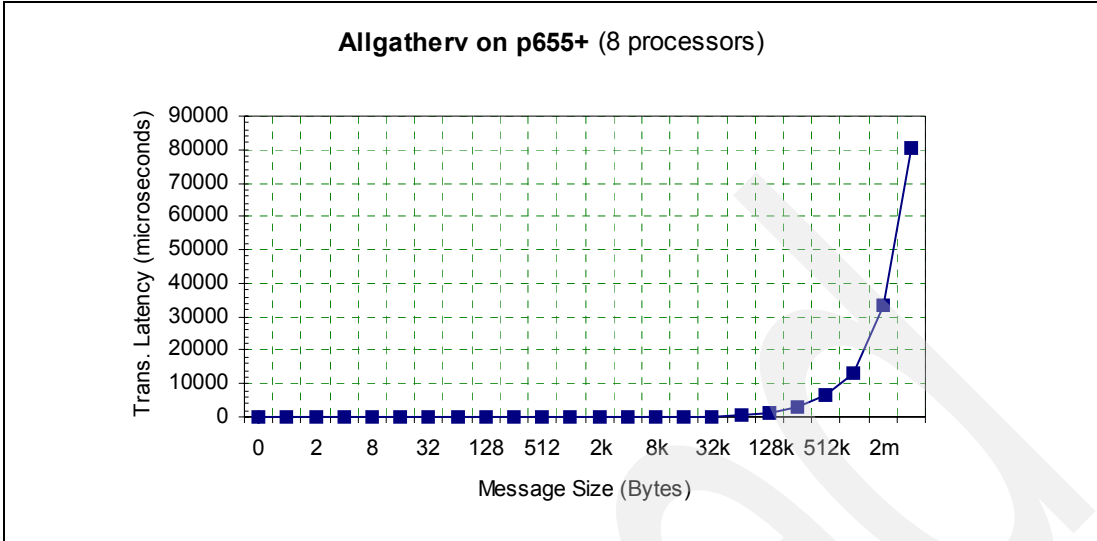


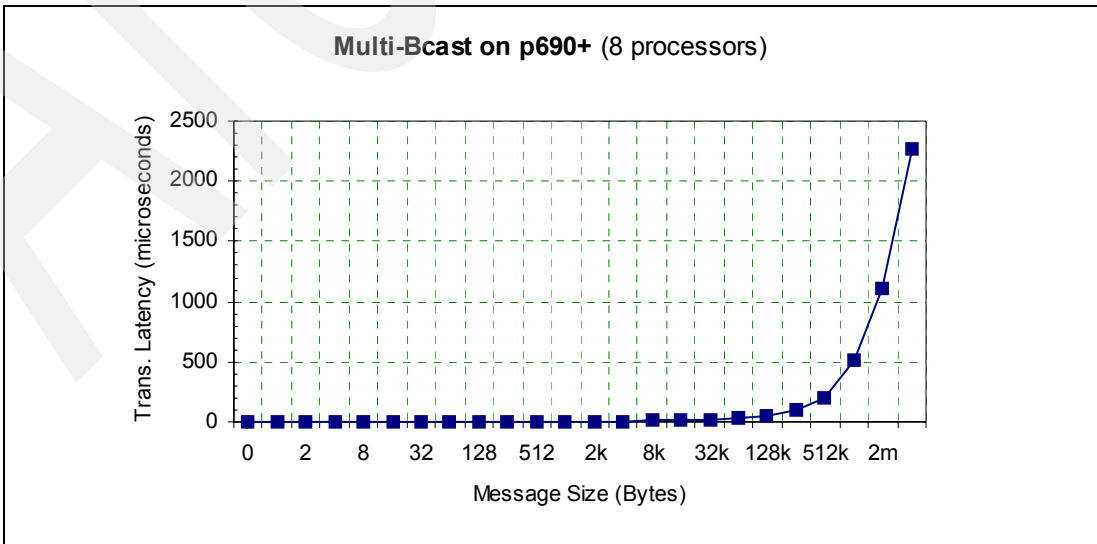
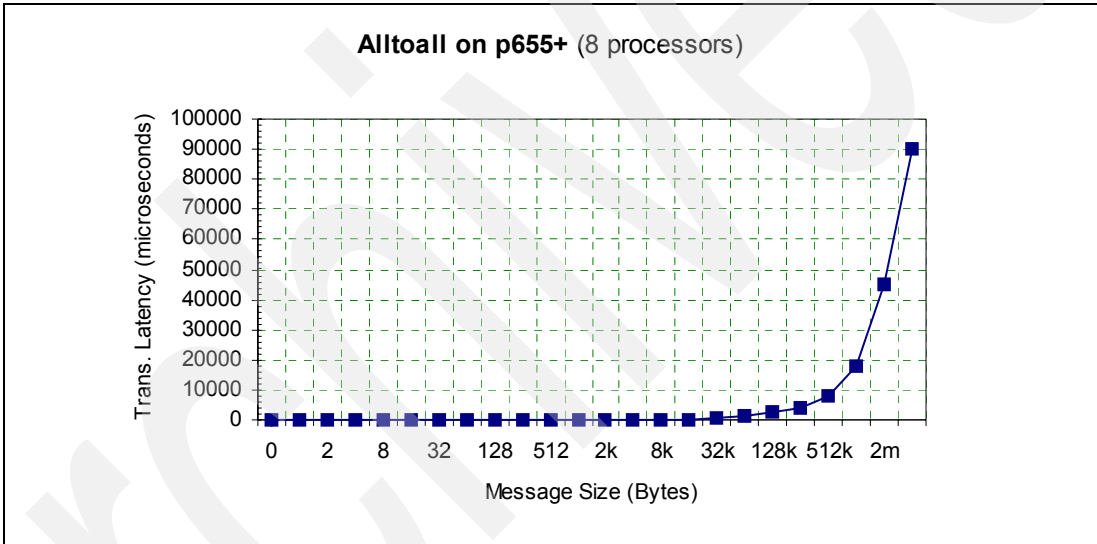
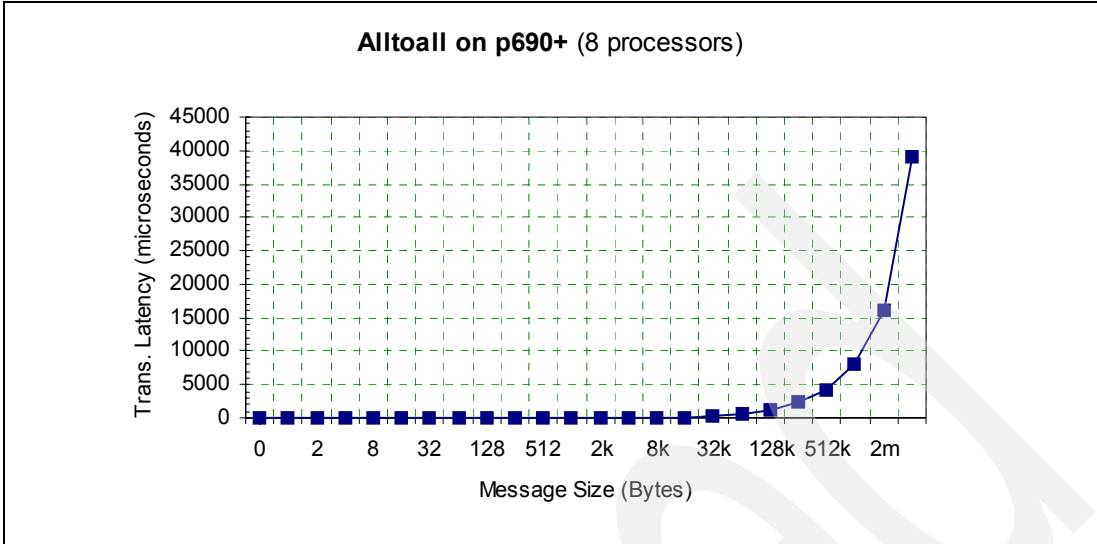


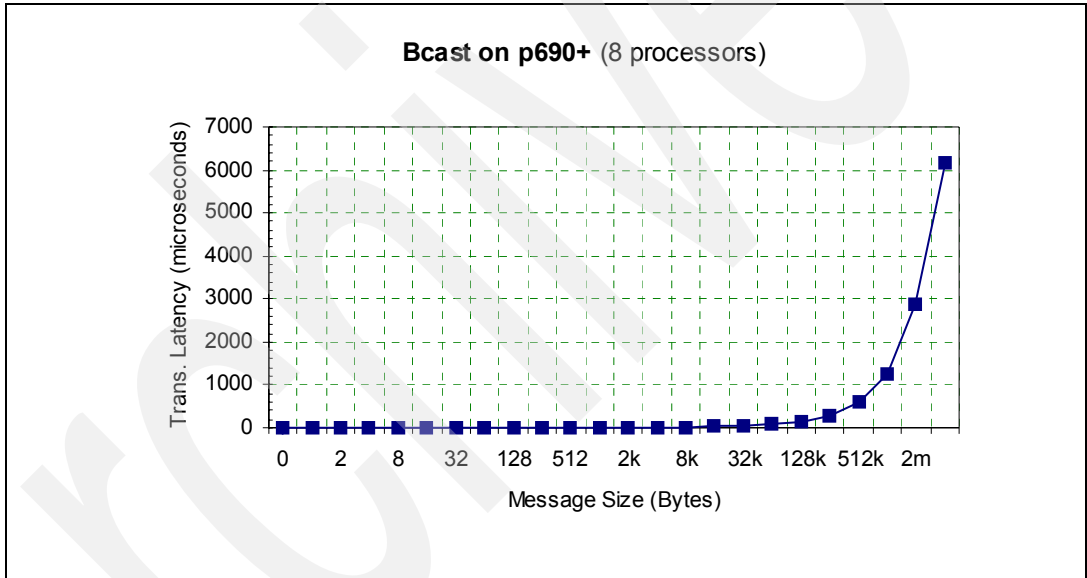
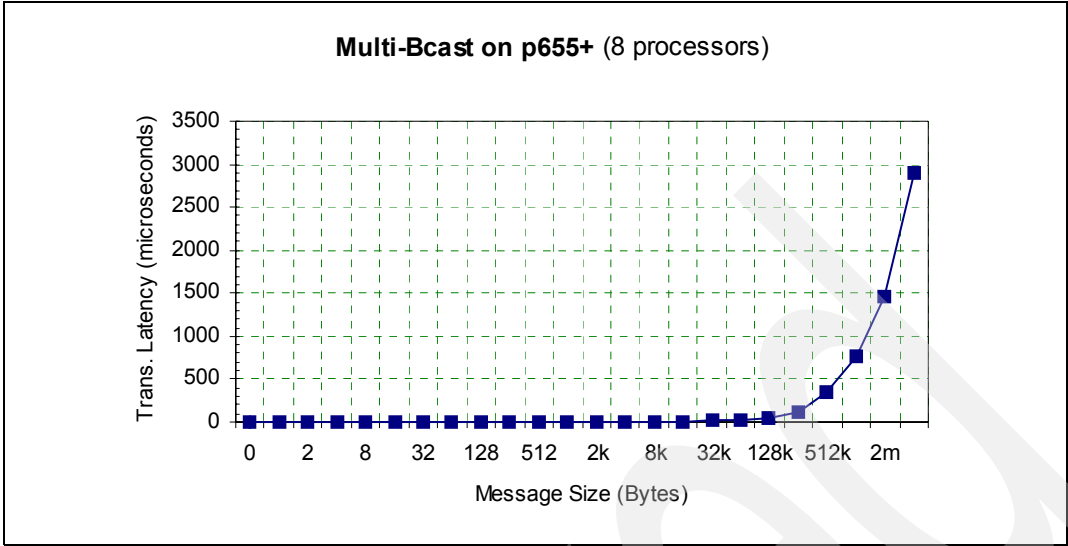


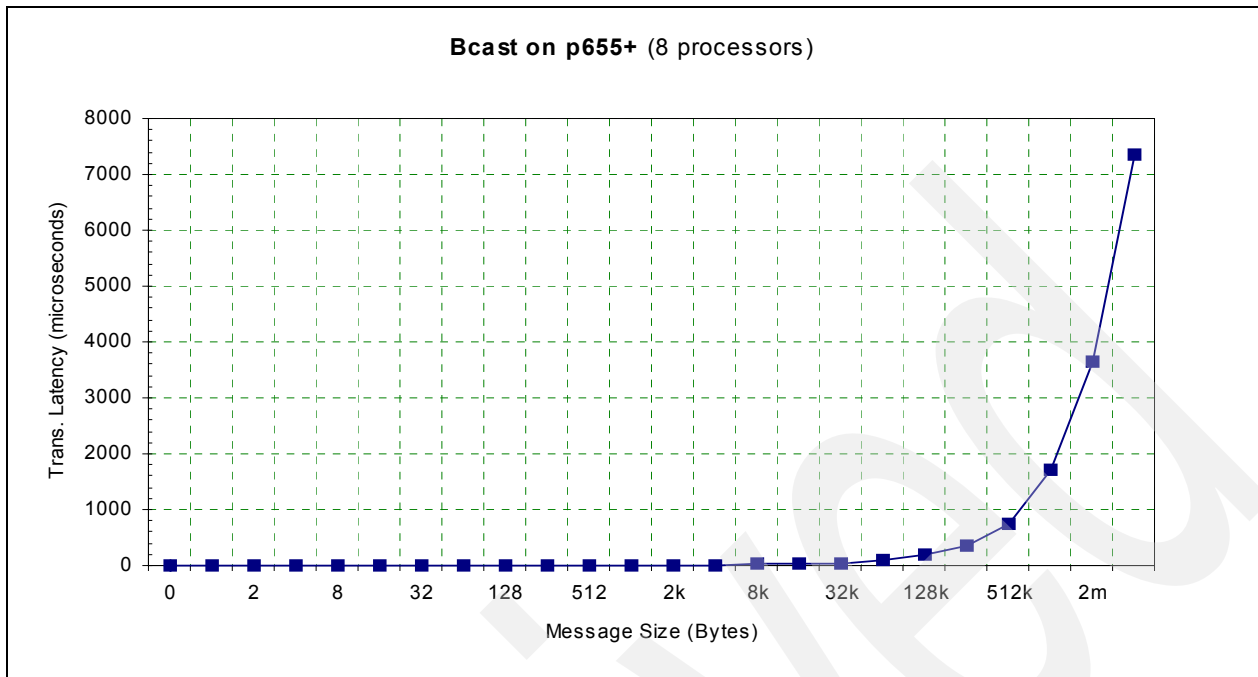












MPI_Barrier

Table 2 MPI_Barrier () function benchmark

Test	Barrier (microseconds)				
	8	16	32	64	128
P690+	3.64	8.48	14.97	22.35	29.96
P655+	3.27	6.23	10.69	78.57	125.43

Summary

The PMB-2.2.1 benchmark was completed on POWER4+ platforms p690+ and p655+. The MPI-I output results, without modification, are shown in graphical format in this report.

Author

Ruzhu Chen (ruzhuchen@us.ibm.com)

pSeries® & HPC Benchmark Center, IBM Poughkeepsie, NY

Reference

- ▶ PMB2.2.1-mpi:
<http://www.pallas.com/e/products/pmb/index.htm>

Archived

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.



Send us your comments in one of the following ways:

- ▶ Use the online **Contact us** review redbook form found at:
ibm.com/redbooks
- ▶ Send your comments in an Internet note to:
redbook@us.ibm.com
- ▶ Mail your comments to:
IBM Corporation, International Technical Support Organization
Dept. HYJ Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
IBM®
POWER4+™

pSeries®
Redbooks™
Redbooks (logo)™

Redbooks (logo) ™

Other company, product, and service names may be trademarks or service marks of others.