Karlsruhe Institute of Technology

Benchmarks on Juqueen

Michael Axtmann
19. January 2016

Institute of Theoretical Informatics – Algorithmics

Reduce

Hypercube Exchange

Bcast

Juqueen

Architecture

- 458.752 cores
- Node: 16 cores
- IBM PowerPC A2, 1.6 GHz
- Overall peak performance: 5.9 Petaflops

Network

- Linpack: 5.0 Petaflops
- 5D Torus — 40 GBps; 2.5 50 µs latency
- Collective network

Reduce-/ Bcast Configuration

Karlsruhe Institute of Technology

Architecture

- 8192 cores
- Node: 512

Parallelism

- 16 Processes / Node
- 1 Process / Node
- 1 Process / Node + 16 Threads

Input

- Constant input size per process

Reduce: (Non-) Shared Memory

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node
4	0.00002	0.00007	0.00001	0.00005
8	0.00002	0.00007	0.00001	0.00005
16	0.00003	0.00011	0.00001	0.00007
32	0.00003	0.00011	0.00001	0.00007
64	0.00004	0.00012	0.00001	0.00009
...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248
32768	0.00144	0.00650	0.00016	0.00482
65536	0.00211	0.01279	0.00031	0.00951
131072	0.00458	0.02541	0.00068	0.01696
262144	0.01338	0.05201	0.00135	0.03380
524288	0.02886	0.09430	0.01110	0.06746
1048576	0.06111	0.18845	0.02187	0.13479

Reduce: (Non-) Shared Memory

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node
4	0.00002	0.00007	0.00001	0.00005
8	0.00002	0.00007	0.00001	0.00005
16	0.00003	0.00011	0.00001	0.00007
32	0.00003	0.00011	0.00001	0.00007
64	0.00004	0.00012	0.00001	0.00009
...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248
32768	0.00144	0.00650	0.00016	0.00482
65536	0.00211	0.01279	0.00031	0.00951
131072	0.00458	0.02541	0.00068	0.01696
262144	0.01338	0.05201	0.00135	0.03380
524288	0.02886	0.09430	0.01110	0.06746
1048576	0.06111	0.18845	0.02187	0.13479

Reduce: (Non-) Shared Memory

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node
4	0.00002	0.00007	0.00001	0.00005
8	0.00002	0.00007	0.00001	0.00005
16	0.00003	0.00011	0.00001	0.00007
32	0.00003	0.00011	0.00001	0.00007
64	0.00004	0.00012	0.00001	0.00009
...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248
32768	0.00144	0.00650	0.00016	0.00482
65536	0.00211	0.01279	0.00031	0.00951
131072	0.00458	0.02541	0.00068	0.01696
262144	0.01338	0.05201	0.00135	0.03380
524288	0.02886	0.09430	0.01110	0.06746
1048576	0.06111	0.18845	0.02187	0.13479

Reduce: (Non-) Shared Memory

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node
4	0.00002	0.00007	0.00001	0.00005
8	0.00002	0.00007	0.00001	0.00005
16	0.00003	0.00011	0.00001	0.00007
32	0.00003	0.00011	0.00001	0.00007
64	0.00004	0.00012	0.00001	0.00009
...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248
32768	0.00144	0.00650	0.00016	0.00482
65536	0.00211	0.01279	0.00031	0.00951
131072	0.00458	0.02541	0.00068	0.01696
262144	0.01338	0.05201	0.00135	0.03380
524288	0.02886	0.09430	0.01110	0.06746
1048576	0.06111	0.18845	0.02187	0.13479

Reduce: (Non-) Shared Memory

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node	Binomial + OpenMP
4	0.00002	0.00007	0.00001	0.00005	0.00028
8	0.00002	0.00007	0.00001	0.00005	0.00027
16	0.00003	0.00011	0.00001	0.00007	0.00028
32	0.00003	0.00011	0.00001	0.00007	0.00028
64	0.00004	0.00012	0.00001	0.00009	0.00030
...	...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248	0.00122
32768	0.00144	0.00650	0.00016	0.00482	0.00200
65536	0.00211	0.01279	0.00031	0.00951	0.00360
131072	0.00458	0.02541	0.00068	0.01696	0.00670
262144	0.01338	0.05201	0.00135	0.03380	0.01300
524288	0.02886	0.09430	0.01110	0.06746	0.02561
1048576	0.06111	0.18845	0.02187	0.13479	0.05084

Reduce: OpenMP

N / Process	MPI 16P/Node	Binomial 16P/Node	MPI 1P/Node	Binomial 1P/Node	Binomial + OpenMP
4	0.00002	0.00007	0.00001	0.00005	0.00028
8	0.00002	0.00007	0.00001	0.00005	0.00027
16	0.00003	0.00011	0.00001	0.00007	0.00028
32	0.00003	0.00011	0.00001	0.00007	0.00028
64	0.00004	0.00012	0.00001	0.00009	0.00030
...	...	...	...	...	...
16384	0.00104	0.00336	0.00009	0.00248	0.00122
32768	0.00144	0.00650	0.00016	0.00482	0.00200
65536	0.00211	0.01279	0.00031	0.00951	0.00360
131072	0.00458	0.02541	0.00068	0.01696	0.00670
262144	0.01338	0.05201	0.00135	0.03380	0.01300
524288	0.02886	0.09430	0.01110	0.06746	0.02561
1048576	0.06111	0.18845	0.02187	0.13479	0.05084

Reduce and Bcast Conclusion

Karlsruhe Institute of Technology

Drawbacks

- Just constant number of elements per process
- OpenMP boost not applicable to Bcast

Hypercube Configuration

Karlsruhe Institute of Technology

Architecture

- 8192 cores
- Node: 512

Parallelism

- 16 Processes / Node
- 1 Process / Node
- 1 Process / Node + 16 Threads

Input

- Constant input size per core

Hypercube Benchmark

N / Core	16P/Node	1P/Node	Slowdown
1	0.00023	0.00022	0.94
2	0.00022	0.00022	0.97
4	0.00023	0.00024	1.07
8	0.00023	0.00025	1.11
16	0.00037	0.00027	0.75
32	0.00037	0.00030	0.80
64	0.00039	0.00023	0.59
...	...	...	...
65536	0.02948	0.04266	1.44
131072	0.04931	0.08573	1.73
262144	0.04607	0.18441	4.00
524288	0.25197	0.36071	1.43
1048576	0.46242	0.70452	1.52
2097152	0.32798	1.40116	4.27

Hypercube Conclusion

Karlsruhe Institute of Technology

16 Processes per Node (P16): Pros and Cons

- Less processes (messages)
- Less communication phases
- More data per process

One process per node is not able to saturate the network!?