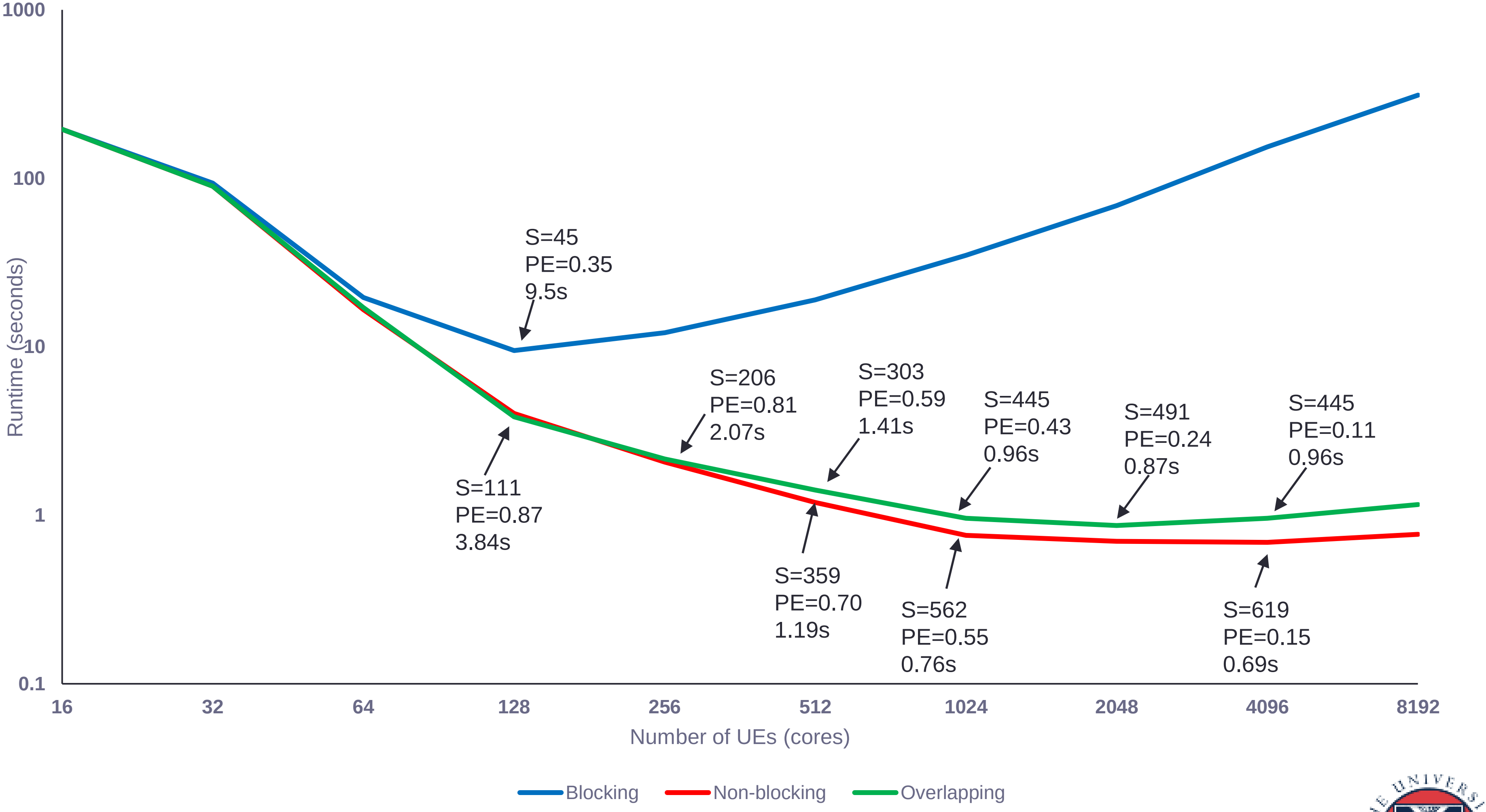
Practical two

Wrap up

Strong scaling



ARCHER2 vs ARCHER

Cores	ARCHER	ARCHER2
1	922.58	427.27
16	125.26	195.4
32	62.84	93.77
64	33.99	19.64
128	16.23	9.5
256	16.5	12.11

blocking

Cores	ARCHER	ARCHER2
1	922.58	427.27
16	124.63	195.26
32	61.01	89.94
64	29.62	16.61
128	8.82	4.02
256	5.22	2.07
512	4.45	1.19
1024	3.92	0.76

Non-blocking

ARCHER2 vs ARCHER

Cores	ARCHER	ARCHER2	ARCHER2 widely
1	922.58	427.27	-
16	125.26	195.4	34.91
32	62.84	93.77	29.11
64	33.99	19.64	13.82
128	16.23	9.5	13.37
256	16.5	12.11	-

blocking

Cores	ARCHER	ARCHER2	ARCHER2 widely
1	922.58	427.27	-
16	124.63	195.26	33.18
32	61.01	89.94	26.87
64	29.62	16.61	9.67
128	8.82	4.02	4.24
256	5.22	2.07	-
512	4.45	1.19	-
1024	3.92	0.76	-

Non-blocking

--distribution=block:block will fill up each NUMA region one by one

So will allocate to core 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15 etc….

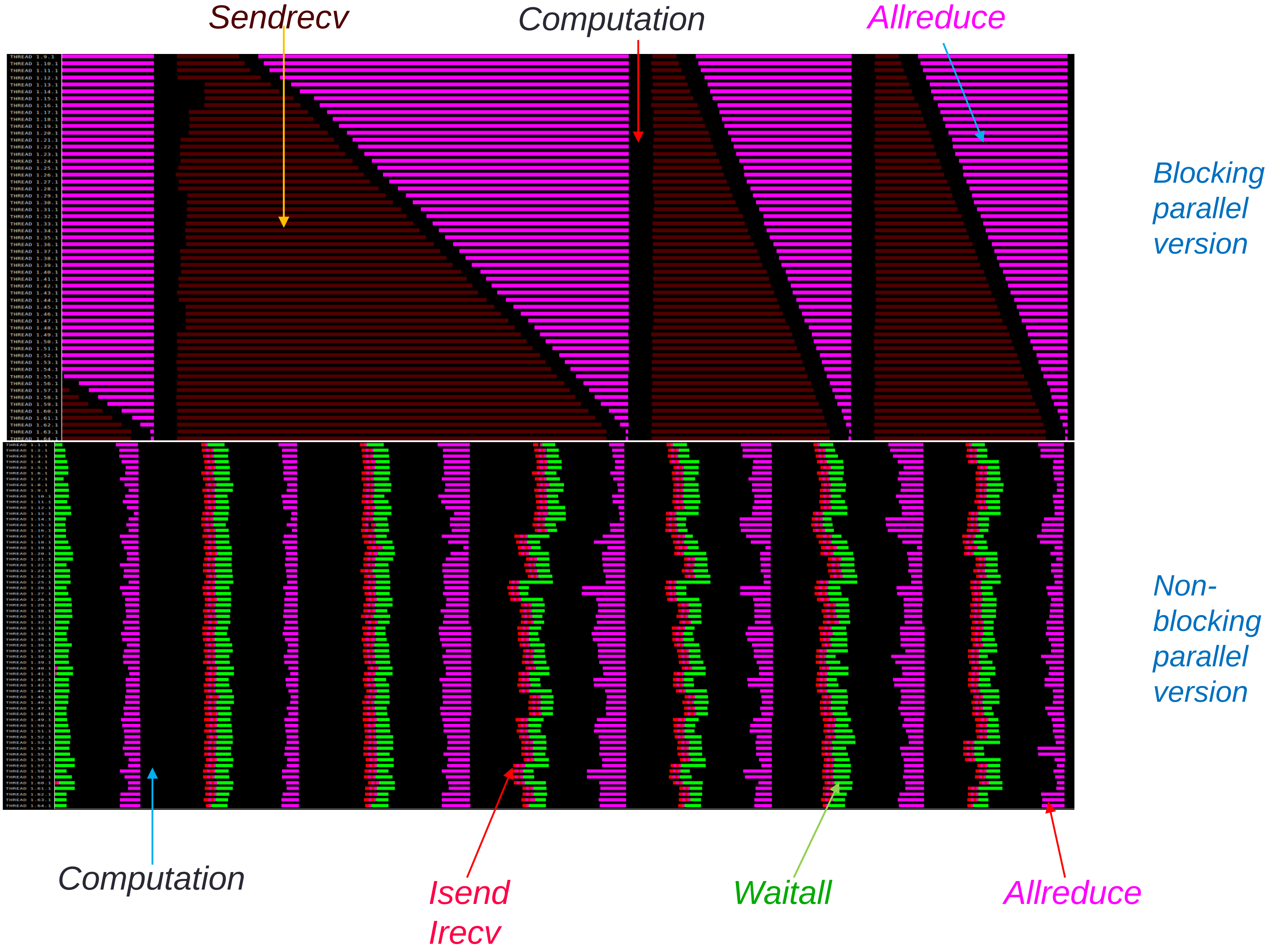
If we remove it will allocate widely

So will allocate to core 0,16,32,48,64,80,96,112,1,17,33,49,65,81,97,113 etc….

Therefore for small numbers of cores this is better as we use more NUMA regions and hence memory links (improve our memory bandwidth)

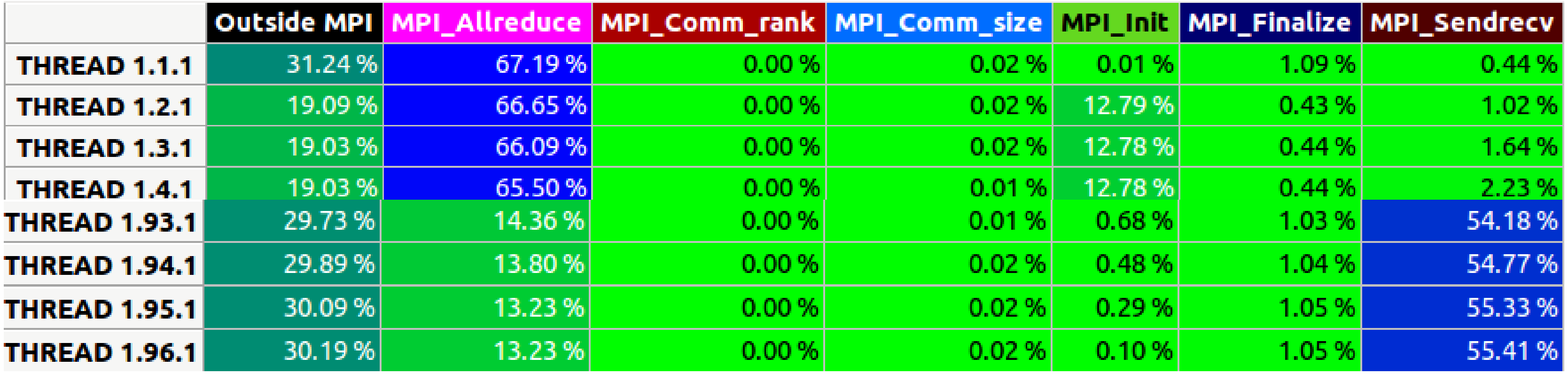
But when we use all the cores on the CPU then there is lots of cross NUMA communication (16 cores per NUMA region, so works badly here for us) and results in overhead. Also groups of 4 cores share an L3 cache.

Likely could get slightly higher performance with 32 and 64 cores if we pinned those in the queue submission file due to these reasons



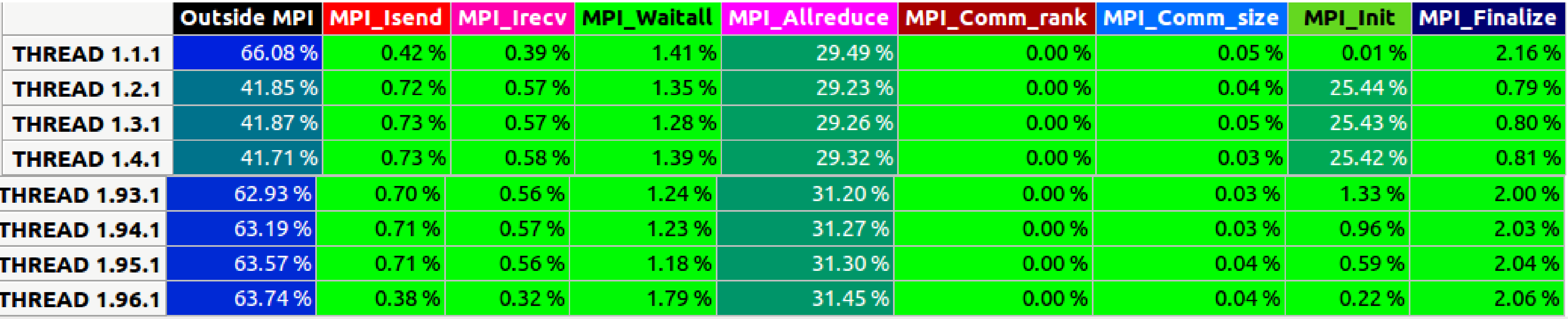
Blocking parallel version

Non-blocking parallel version



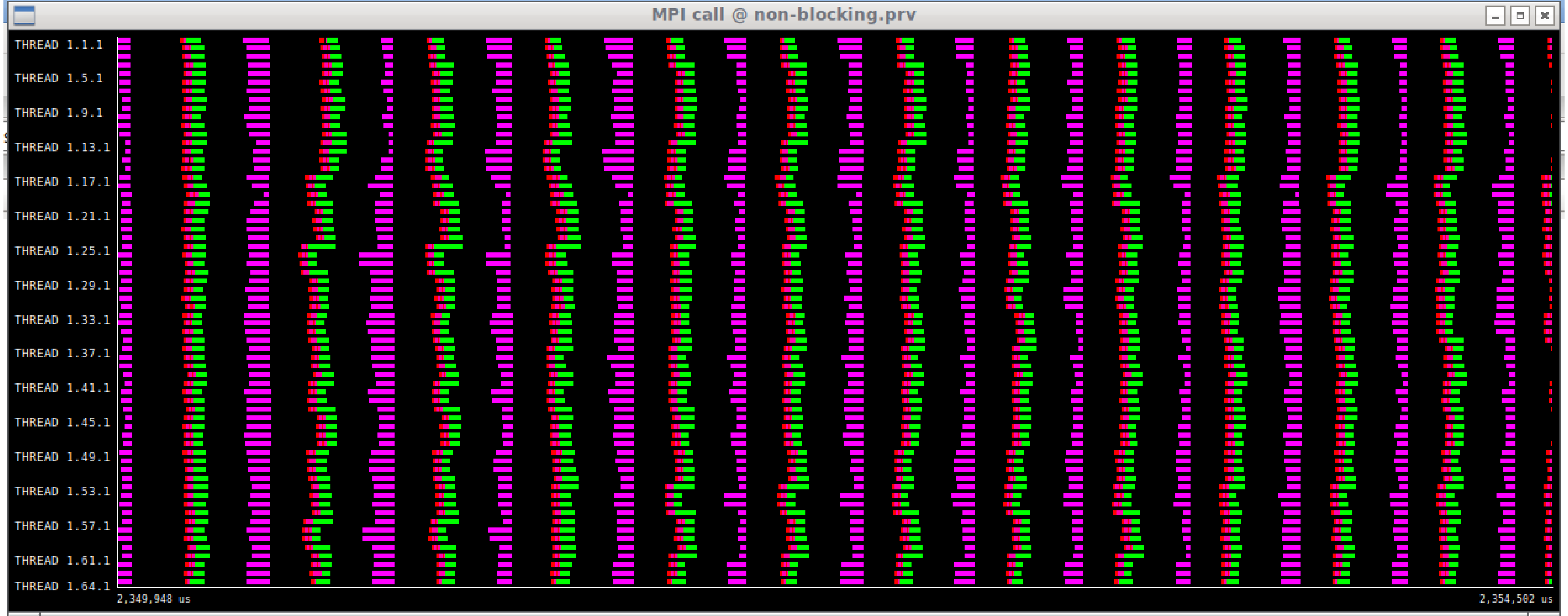
	Outside MPI	MPI_Allreduce	MPI_Comm_rank	MPI_Comm_size	MPI_Init	MPI_Finalize	MPI_Sendrecv
THREAD 1.1.1	31.24 %	67.19 %	0.00 %	0.02 %	0.01 %	1.09 %	0.44 %
THREAD 1.2.1	19.09 %	66.65 %	0.00 %	0.02 %	12.79 %	0.43 %	1.02 %
THREAD 1.3.1	19.03 %	66.09 %	0.00 %	0.02 %	12.78 %	0.44 %	1.64 %
THREAD 1.4.1	19.03 %	65.50 %	0.00 %	0.01 %	12.78 %	0.44 %	2.23 %
THREAD 1.93.1	29.73 %	14.36 %	0.00 %	0.01 %	0.68 %	1.03 %	54.18 %
THREAD 1.94.1	29.89 %	13.80 %	0.00 %	0.02 %	0.48 %	1.04 %	54.77 %
THREAD 1.95.1	30.09 %	13.23 %	0.00 %	0.02 %	0.29 %	1.05 %	55.33 %
THREAD 1.96.1	30.19 %	13.23 %	0.00 %	0.02 %	0.10 %	1.05 %	55.41 %

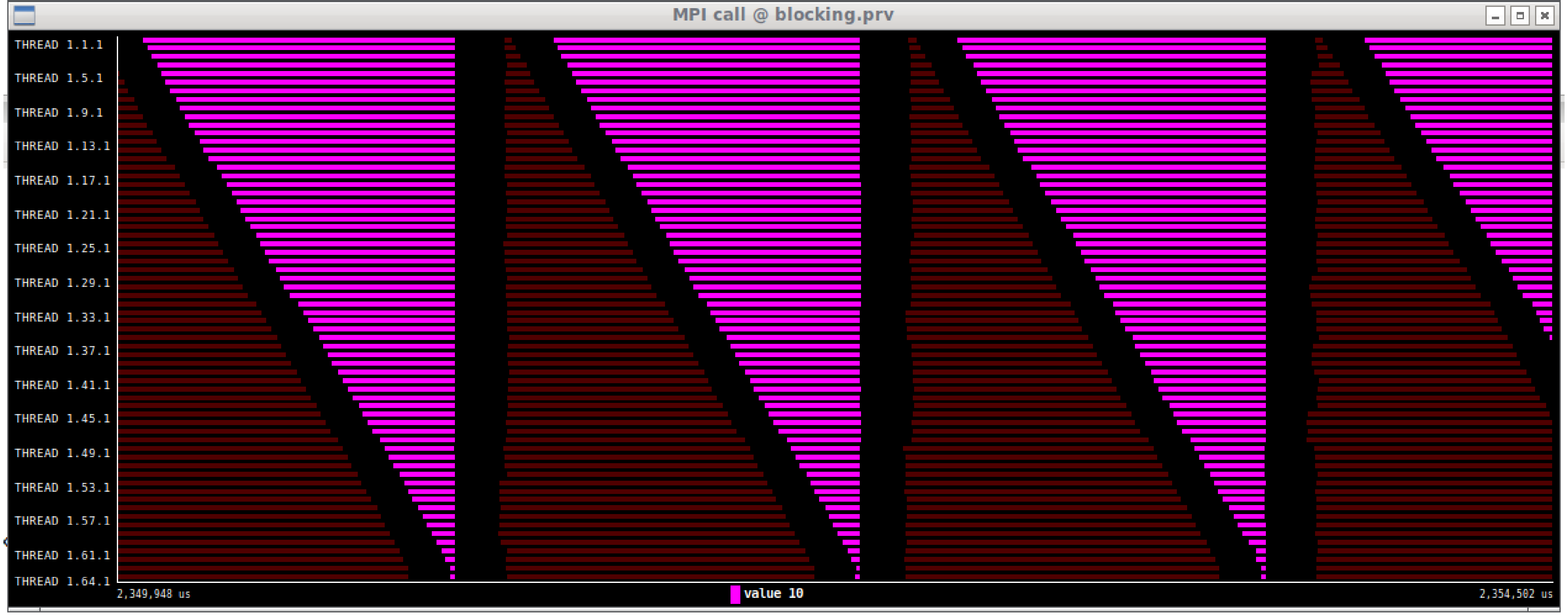
Blocking parallel version



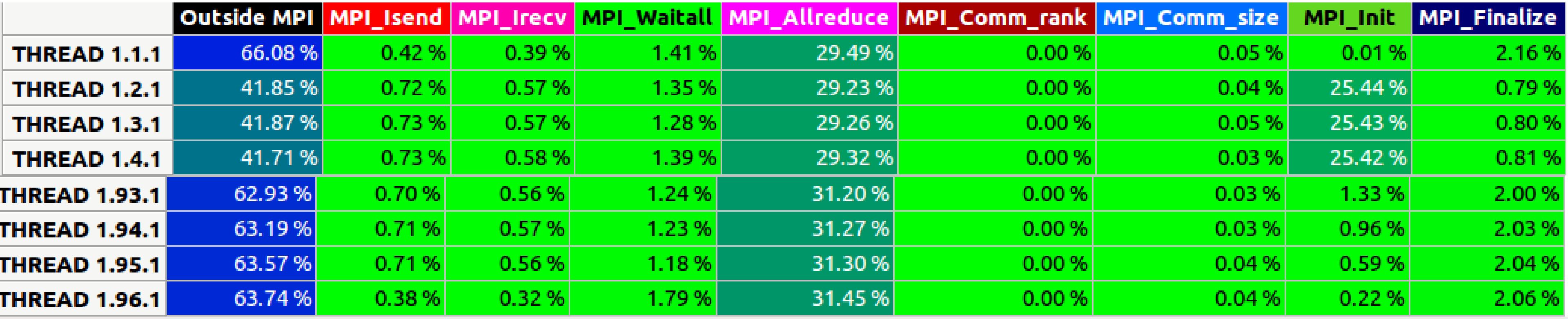
	Outside MPI	MPI_Isend	MPI_Irecv	MPI_Waitall	MPI_Allreduce	MPI_Comm_rank	MPI_Comm_size	MPI_Init	MPI_Finalize
THREAD 1.1.1	66.08 %	0.42 %	0.39 %	1.41 %	29.49 %	0.00 %	0.05 %	0.01 %	2.16 %
THREAD 1.2.1	41.85 %	0.72 %	0.57 %	1.35 %	29.23 %	0.00 %	0.04 %	25.44 %	0.79 %
THREAD 1.3.1	41.87 %	0.73 %	0.57 %	1.28 %	29.26 %	0.00 %	0.05 %	25.43 %	0.80 %
THREAD 1.4.1	41.71 %	0.73 %	0.58 %	1.39 %	29.32 %	0.00 %	0.03 %	25.42 %	0.81 %
THREAD 1.93.1	62.93 %	0.70 %	0.56 %	1.24 %	31.20 %	0.00 %	0.03 %	1.33 %	2.00 %
THREAD 1.94.1	63.19 %	0.71 %	0.57 %	1.23 %	31.27 %	0.00 %	0.03 %	0.96 %	2.03 %
THREAD 1.95.1	63.57 %	0.71 %	0.56 %	1.18 %	31.30 %	0.00 %	0.04 %	0.59 %	2.04 %
THREAD 1.96.1	63.74 %	0.38 %	0.32 %	1.79 %	31.45 %	0.00 %	0.04 %	0.22 %	2.06 %

Non-blocking parallel version


MPI call @ non-blocking.prv
THREAD 1.1.1
THREAD 1.5.1
THREAD 1.9.1
THREAD 1.13.1
THREAD 1.17.1
THREAD 1.21.1
THREAD 1.25.1
THREAD 1.29.1
THREAD 1.33.1
THREAD 1.37.1
THREAD 1.41.1
THREAD 1.45.1
THREAD 1.49.1
THREAD 1.53.1
THREAD 1.57.1
THREAD 1.61.1
THREAD 1.64.1
2,349,948 us
2,354,502 us

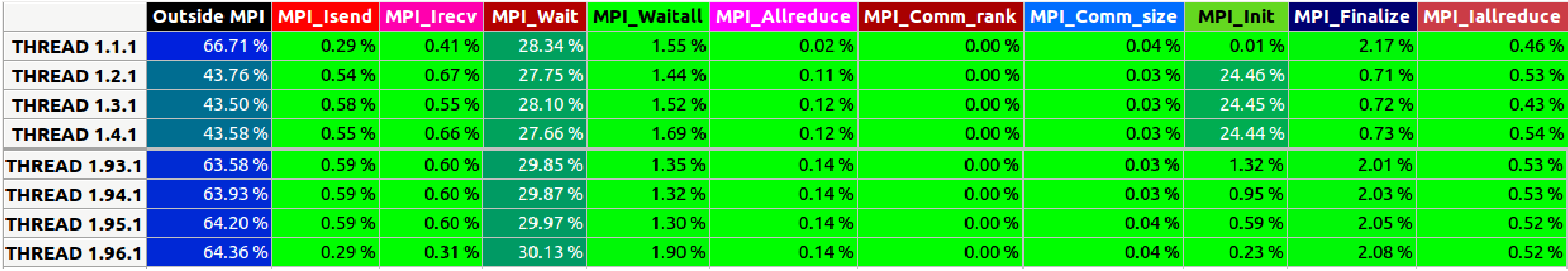

MPI call @ blocking.prv
THREAD 1.1.1
THREAD 1.5.1
THREAD 1.9.1
THREAD 1.13.1
THREAD 1.17.1
THREAD 1.21.1
THREAD 1.25.1
THREAD 1.29.1
THREAD 1.33.1
THREAD 1.37.1
THREAD 1.41.1
THREAD 1.45.1
THREAD 1.49.1
THREAD 1.53.1
THREAD 1.57.1
THREAD 1.61.1
THREAD 1.64.1
2,349,948 us
value 10
2,354,502 us

Non-blocking vs overlapping



	Outside MPI	MPI_Isend	MPI_Irecv	MPI_Waitall	MPI_Allreduce	MPI_Comm_rank	MPI_Comm_size	MPI_Init	MPI_Finalize
THREAD 1.1.1	66.08 %	0.42 %	0.39 %	1.41 %	29.49 %	0.00 %	0.05 %	0.01 %	2.16 %
THREAD 1.2.1	41.85 %	0.72 %	0.57 %	1.35 %	29.23 %	0.00 %	0.04 %	25.44 %	0.79 %
THREAD 1.3.1	41.87 %	0.73 %	0.57 %	1.28 %	29.26 %	0.00 %	0.05 %	25.43 %	0.80 %
THREAD 1.4.1	41.71 %	0.73 %	0.58 %	1.39 %	29.32 %	0.00 %	0.03 %	25.42 %	0.81 %
THREAD 1.93.1	62.93 %	0.70 %	0.56 %	1.24 %	31.20 %	0.00 %	0.03 %	1.33 %	2.00 %
THREAD 1.94.1	63.19 %	0.71 %	0.57 %	1.23 %	31.27 %	0.00 %	0.03 %	0.96 %	2.03 %
THREAD 1.95.1	63.57 %	0.71 %	0.56 %	1.18 %	31.30 %	0.00 %	0.04 %	0.59 %	2.04 %
THREAD 1.96.1	63.74 %	0.38 %	0.32 %	1.79 %	31.45 %	0.00 %	0.04 %	0.22 %	2.06 %

Non-blocking parallel version



	Outside MPI	MPI_Isend	MPI_Irecv	MPI_Wait	MPI_Waitall	MPI_Allreduce	MPI_Comm_rank	MPI_Comm_size	MPI_Init	MPI_Finalize	MPI_Iallreduce
THREAD 1.1.1	66.71 %	0.29 %	0.41 %	28.34 %	1.55 %	0.02 %	0.00 %	0.04 %	0.01 %	2.17 %	0.46 %
THREAD 1.2.1	43.76 %	0.54 %	0.67 %	27.75 %	1.44 %	0.11 %	0.00 %	0.03 %	24.46 %	0.71 %	0.53 %
THREAD 1.3.1	43.50 %	0.58 %	0.55 %	28.10 %	1.52 %	0.12 %	0.00 %	0.03 %	24.45 %	0.72 %	0.43 %
THREAD 1.4.1	43.58 %	0.55 %	0.66 %	27.66 %	1.69 %	0.12 %	0.00 %	0.03 %	24.44 %	0.73 %	0.54 %
THREAD 1.93.1	63.58 %	0.59 %	0.60 %	29.85 %	1.35 %	0.14 %	0.00 %	0.03 %	1.32 %	2.01 %	0.53 %
THREAD 1.94.1	63.93 %	0.59 %	0.60 %	29.87 %	1.32 %	0.14 %	0.00 %	0.03 %	0.95 %	2.03 %	0.53 %
THREAD 1.95.1	64.20 %	0.59 %	0.60 %	29.97 %	1.30 %	0.14 %	0.00 %	0.04 %	0.59 %	2.05 %	0.52 %
THREAD 1.96.1	64.36 %	0.29 %	0.31 %	30.13 %	1.90 %	0.14 %	0.00 %	0.04 %	0.23 %	2.08 %	0.52 %

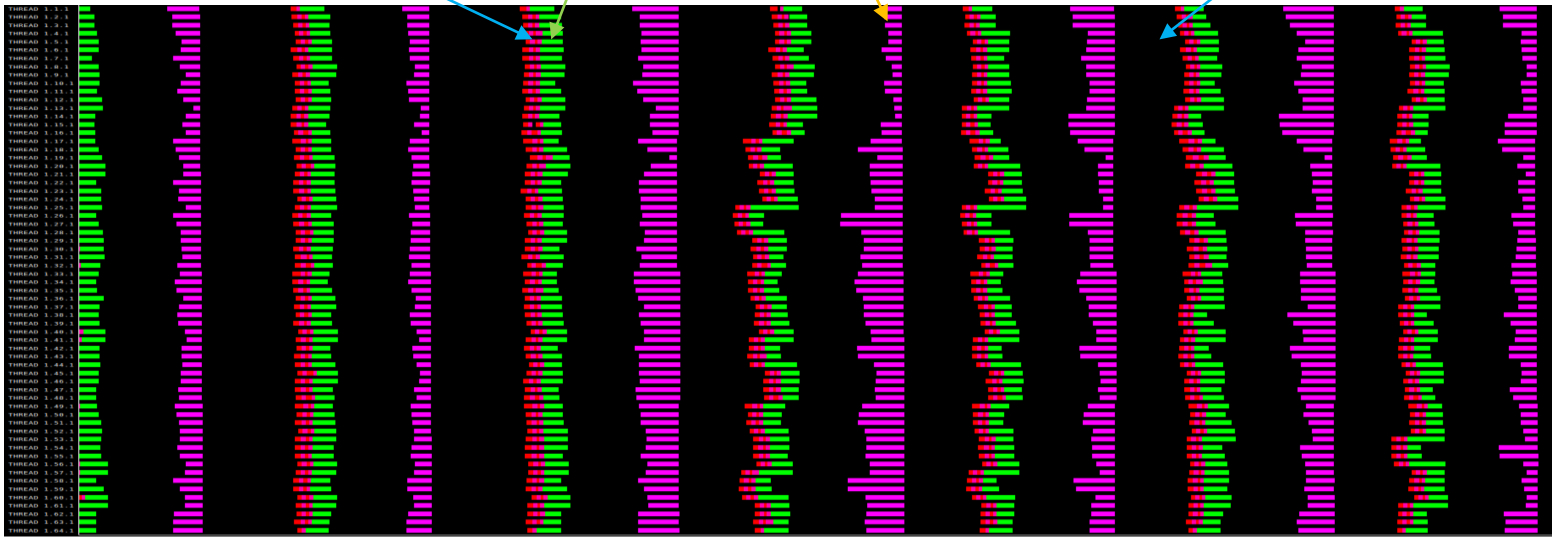
Overlapping parallel version

Isend
Irecv

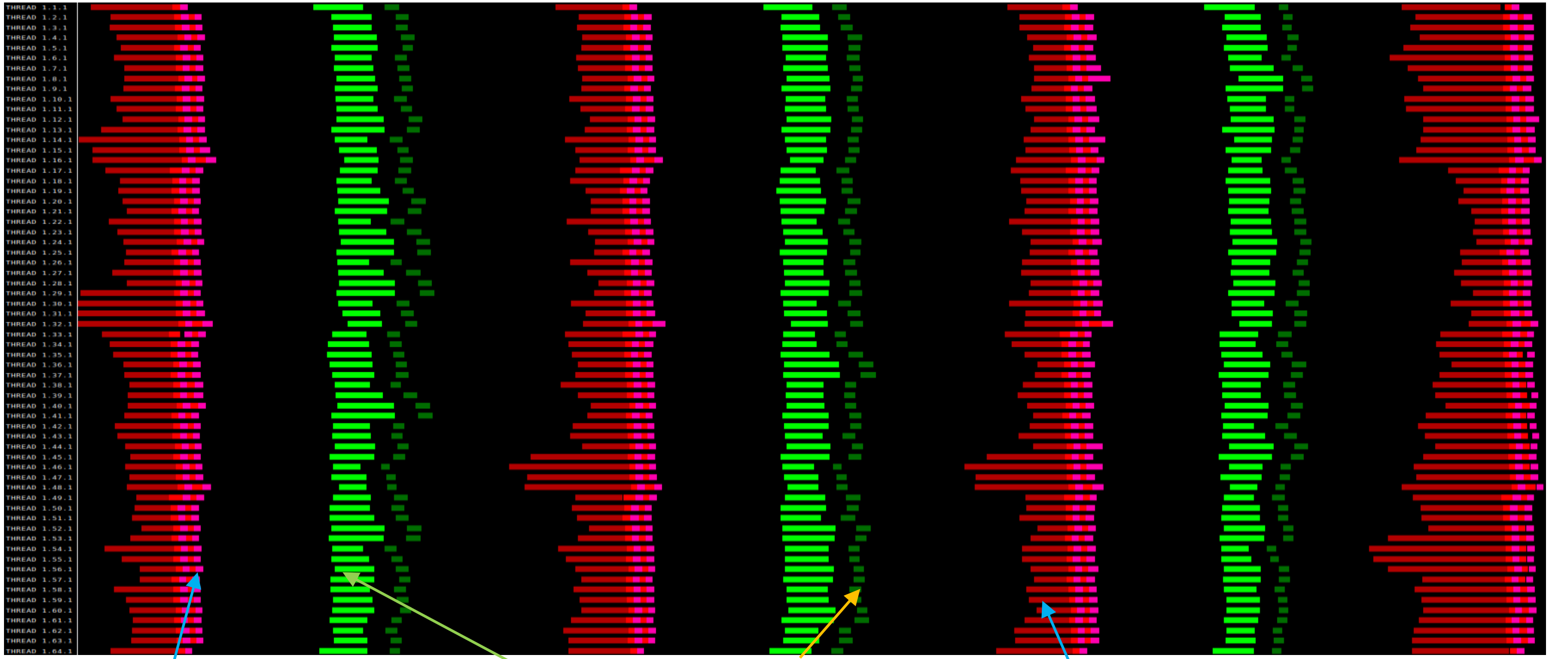
Waitall

Allreduce

Computation



Non-blocking parallel version



Overlapping parallel version

Isend and Irecv

Waitall

Iallreduce

Wait