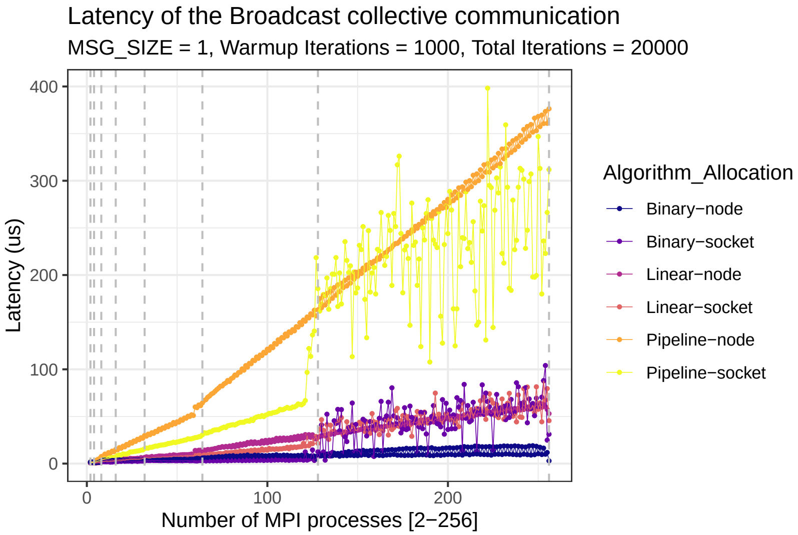
Modeling the Latency of MPI Collective Communication Algorithms

Introduction

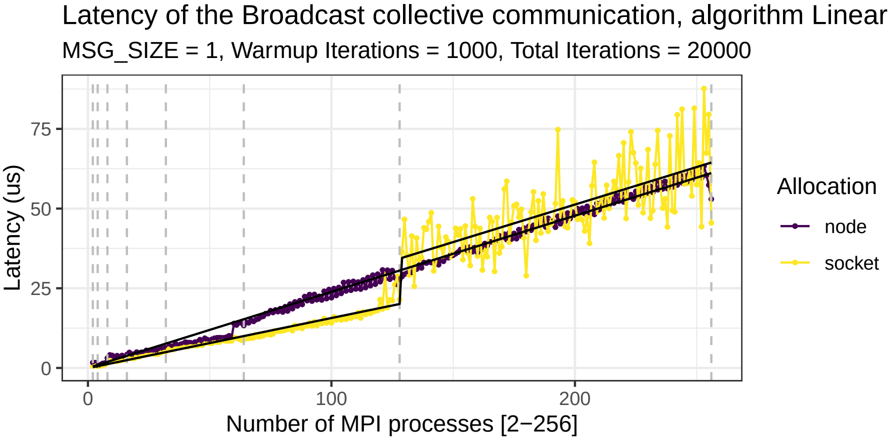
A number of aspects to consider:

- Different algorithms [flat tree, pipeline, binary tree]
- Number of MPI processes [up to 256]
- Size of the buffer [up to 1MB]
- Topology of the nodes and allocation of the computing resources
- Parameters of the benchmark [warmup iterations, total iterations]
- Possible interactions between these factors

Broadcast, Latency and Number of Processes



Broadcast, Linear Algorithm



Flat Tree, Results of the Linear Model

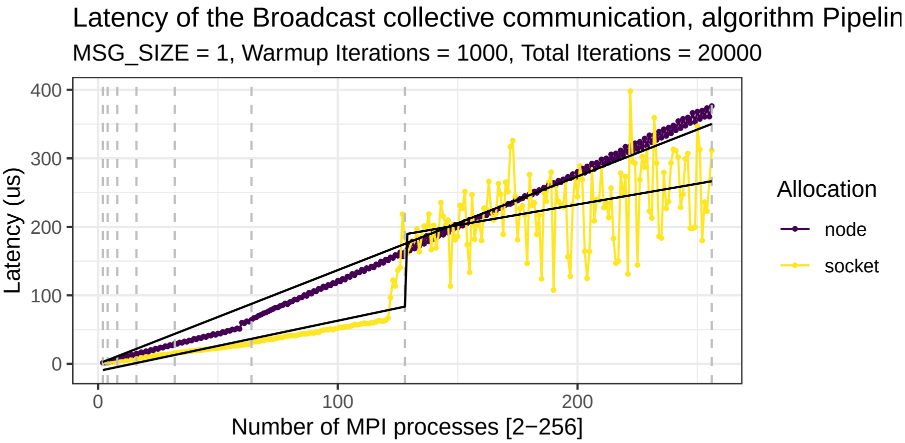
(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.1572	0.0151	10.44
MPI_Processes > 128	4.1839	2.9183	1.43
MPI_Processes : MPI_Processes > 128	0.0781	0.0212	3.69

(b) Allocation by Node

	Estimate	Std. Error	t value
MPI_Processes	0.2388	0.0007	360.94

Broadcast, Pipeline Algorithm



Pipeline, Results of the Linear Model

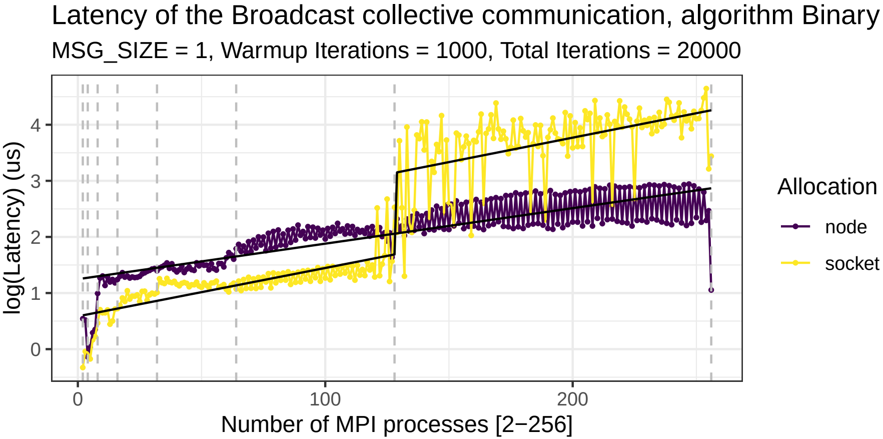
(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.7352	0.0918	8.01
MPI_Processes $>$ 128	111.6247	17.7808	6.28
MPI_Processes : MPI_Processes $>$ 128	-0.1298	0.1290	-1.01

(b) Allocation by Node

	Estimate	Std. Error	t value
MPI_Processes	1.3682	0.0060	229.15

Broadcast, Binary Algorithm



Binary Tree Tree, Results of the Linear Model

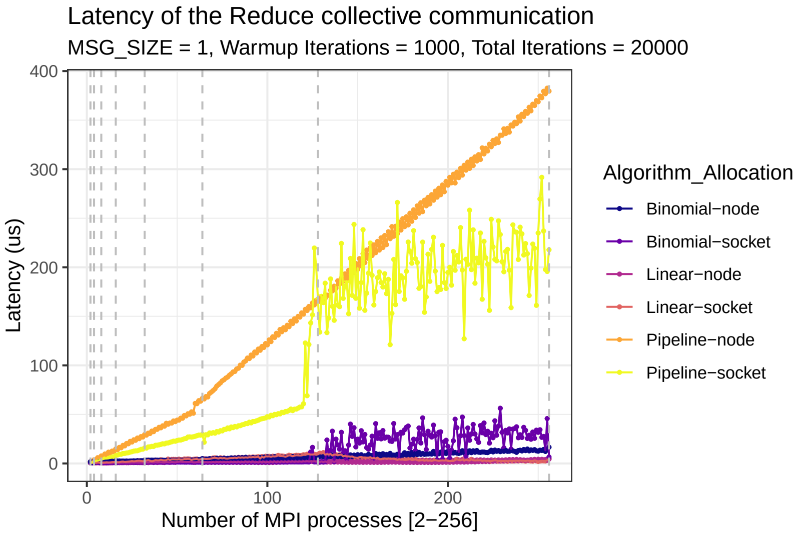
(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.0086	0.0010	8.39
MPI_Processes > 128	2.0269	0.1989	10.19
MPI_Processes : MPI_Processes >	0.0001	0.0014	0.07

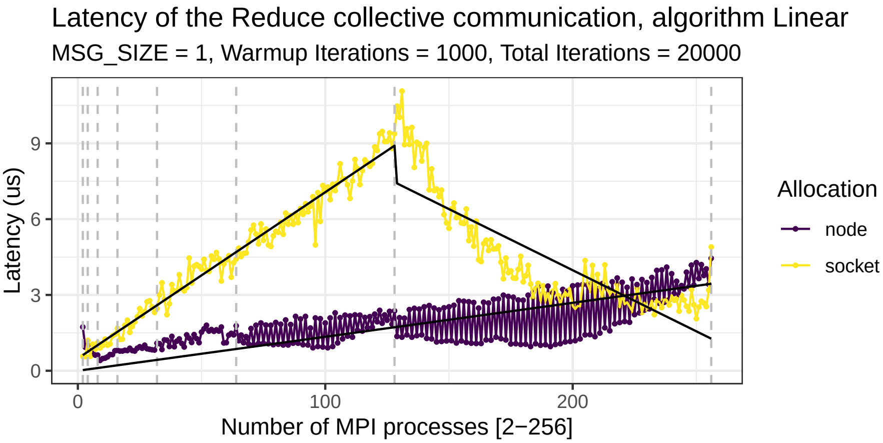
(b) Allocation by Node

	Estimate	Std. Error	t value
(Intercept)	1.2472	0.0396	31.50
MPI_Processes	0.0063	0.0003	23.74

Reduce, Latency and Number of Processes



Reduce, Linear Algorithm



Flat Tree, Results of the Linear Model

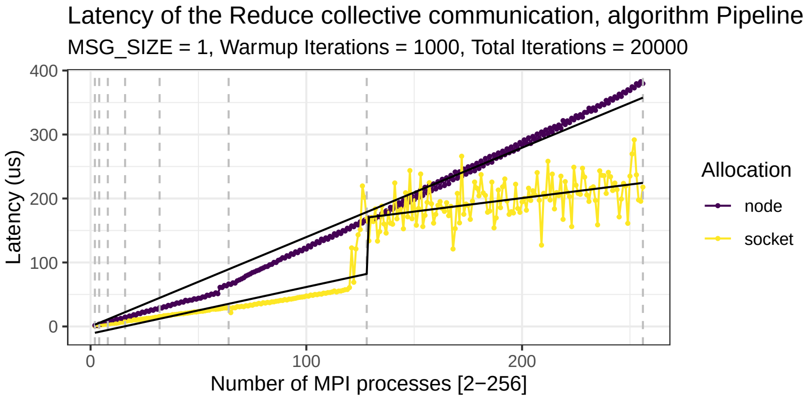
(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.0658	0.0021	30.66
MPI_Processes > 128	13.6566	0.4157	32.85
MPI_Processes : MPI_Processes > 128	-0.1142	0.0030	-37.85

(b) Allocation by Node

	Estimate	Std. Error	t value
MPI_Processes	0.0134	0.0003	43.48

Reduce, Pipeline Algorithm



Pipeline, Results of the Linear Model

(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.7301	0.0606	12.05
MPI_Processes > 128	116.0453	11.7366	9.89
MPI_Processes : MPI_Processes > 128	-0.3068	0.0852	-3.60

(b) Allocation by Node

	Estimate	Std. Error	t value
MPI_Processes	1.3969	0.0064	217.55

Reduce, Binary Algorithm

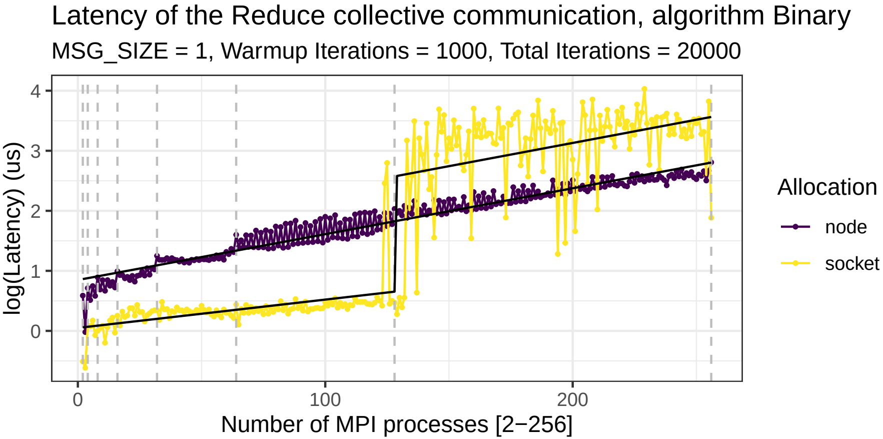


Figure: Reduce, Binary Algorithm

Binary Tree, Results of the Linear Model

(a) Allocation by Socket

	Estimate	Std. Error	t value
MPI_Processes	0.0047	0.0013	3.70
MPI_Processes > 128	1.5831	0.2469	6.41
MPI_Processes : MPI_Processes > 128	0.0030	0.0018	1.68

(b) Allocation by Node

	Estimate	Std. Error	t value
(Intercept)	0.8476	0.0184	46.01
MPI_Processes	0.0076	0.0001	61.58

Latency and Size of the Message

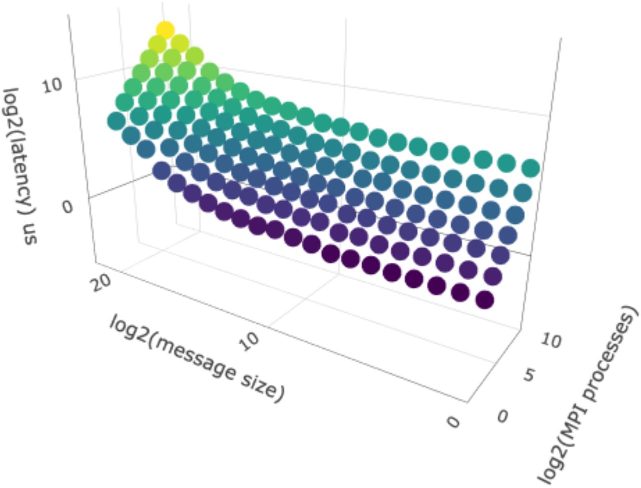


Figure: Pipeline Algorithm, allocation by Node

Latency and Size of the Message

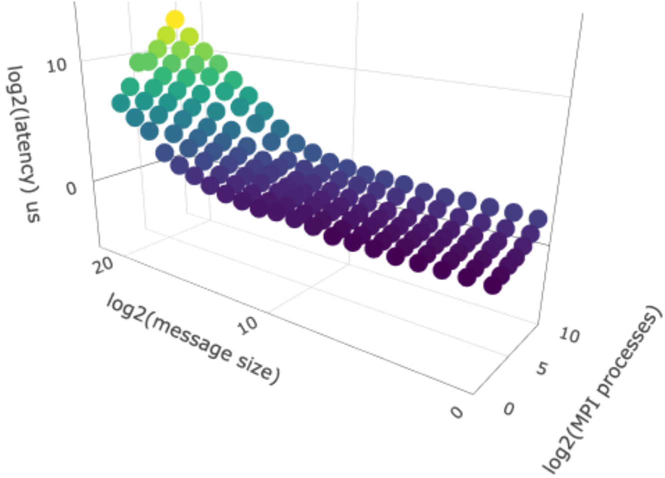


Figure: Binary Algorithm, allocation by Node

Broadcast, Results of the Linear Model

Table: Summary of the Linear Model for the Broadcast Communication

	Estimate	Std. Error	t value
log2(MPI_Processes) : Binary-node	0.6440	0.0308	20.94
log2(MPI_Processes) : Binary-socket	0.5938	0.0308	19.31
log2(MPI_Processes) : Linear-node	0.8181	0.0308	26.60
log2(MPI_Processes) : Linear-socket	0.8728	0.0308	28.38
log2(MPI_Processes) : Pipeline-node	1.1559	0.0308	37.59
log2(MPI_Processes) : Pipeline-socket	1.1131	0.0308	36.20
Message_Size : Binary-node	8.131e-06	5.205e-07	15.62
Message_Size : Binary-socket	9.099e-06	5.205e-07	17.48
Message_Size : Linear-node	7.871e-06	5.205e-07	15.12
Message_Size : Linear-socket	9.166e-06	5.205e-07	17.61
Message_Size : Pipeline-node	5.947e-06	5.205e-07	11.43
Message_Size : Pipeline-socket	8.130e-06	5.205e-07	15.62

Reduce, Results of the Linear Model

	Estimate	Std. Error	t value
log2(MPI_Processes) : Binary-node	0.6532	0.0342	19.09
log2(MPI_Processes) : Binary-socket	0.4583	0.0342	13.39
log2(MPI_Processes) : Linear-node	0.3954	0.0342	11.55
log2(MPI_Processes) : Linear-socket	0.6488	0.0342	18.96
log2(MPI_Processes) : Pipeline-node	1.1630	0.0342	33.98
log2(MPI_Processes) : Pipeline-socket	1.1310	0.0342	33.05
Message_Size : Binary-node	9.016e-06	5.793e-07	15.56
Message_Size : Binary-socket	8.876e-06	5.793e-07	15.32
Message_Size : Linear-node	1.077e-05	5.793e-07	18.59
Message_Size : Linear-socket	1.001e-05	5.793e-07	17.29
Message_Size : Pipeline-node	6.658e-06	5.793e-07	11.49
Message_Size : Pipeline-socket	8.280e-06	5.793e-07	14.29