



Lecture Objectives:

- 1) Explain how MPI programs deal with I/O and the side effects that may occur from using simply IO in an MPI program.
- 2) Explain the concept of tree structured communications.
- 3) Construct an MPI application using MPI reduce.
- 4) Explain the concept of a broadcast message.
- 5) Explain the concept of scattering and gathering data.
- 6) Explain the concept of an MPI derived datatype.

- What is the limiting factor in effective MPI implementation?

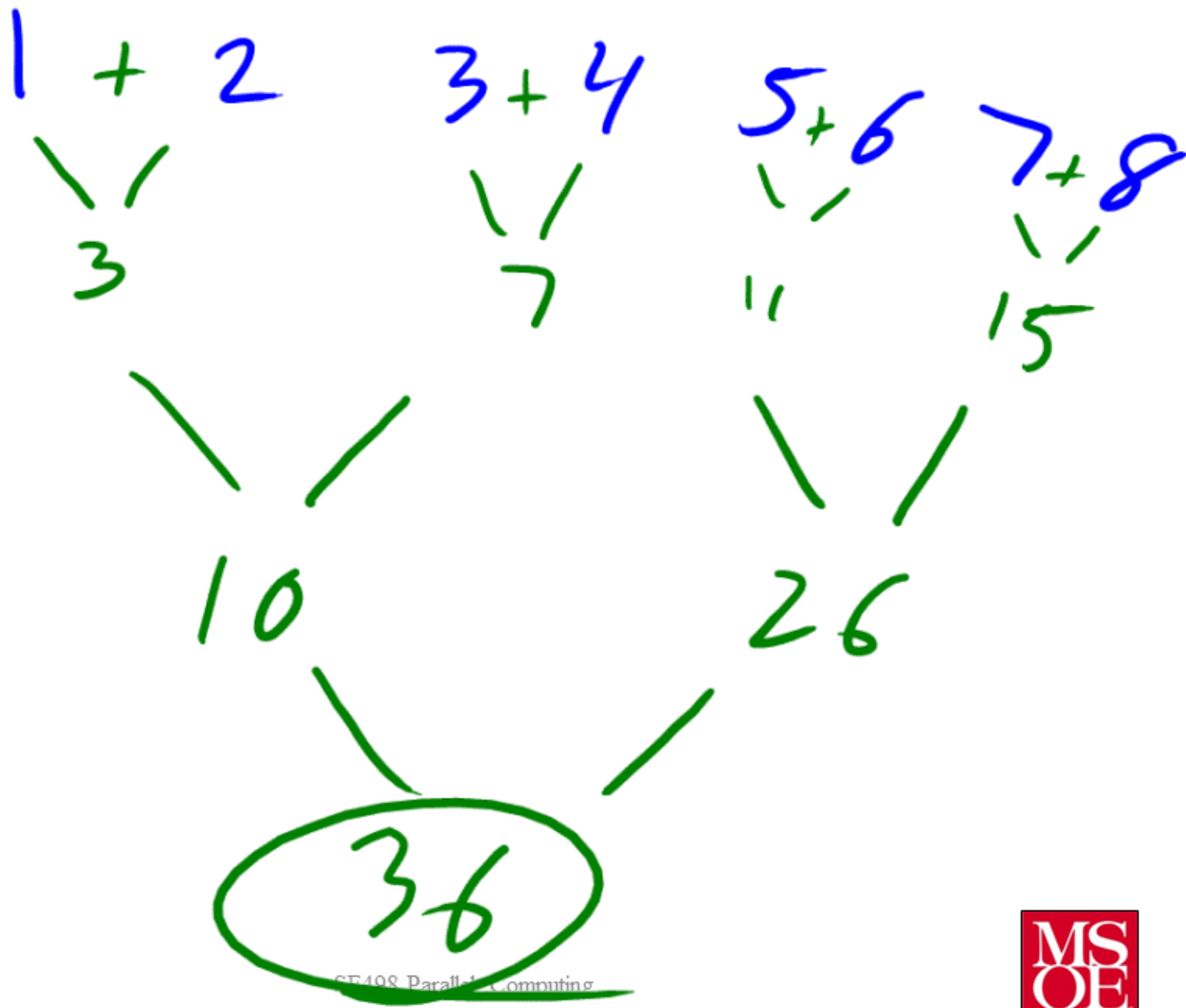
Can not do a
low level \Rightarrow high level
method.

Architecture: Requests messages
+ b. sent.

Discussion

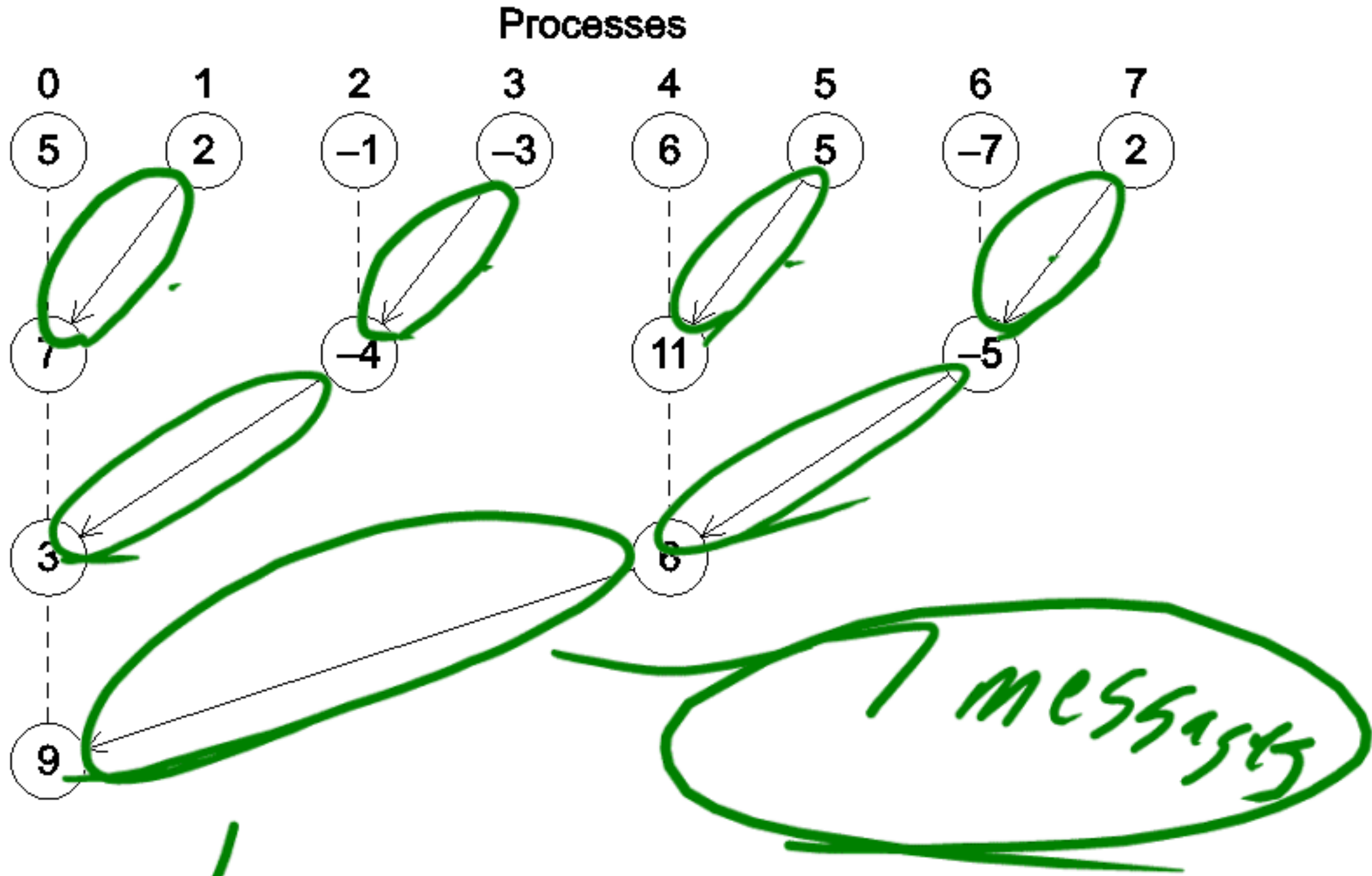
A Tree structured global

sum



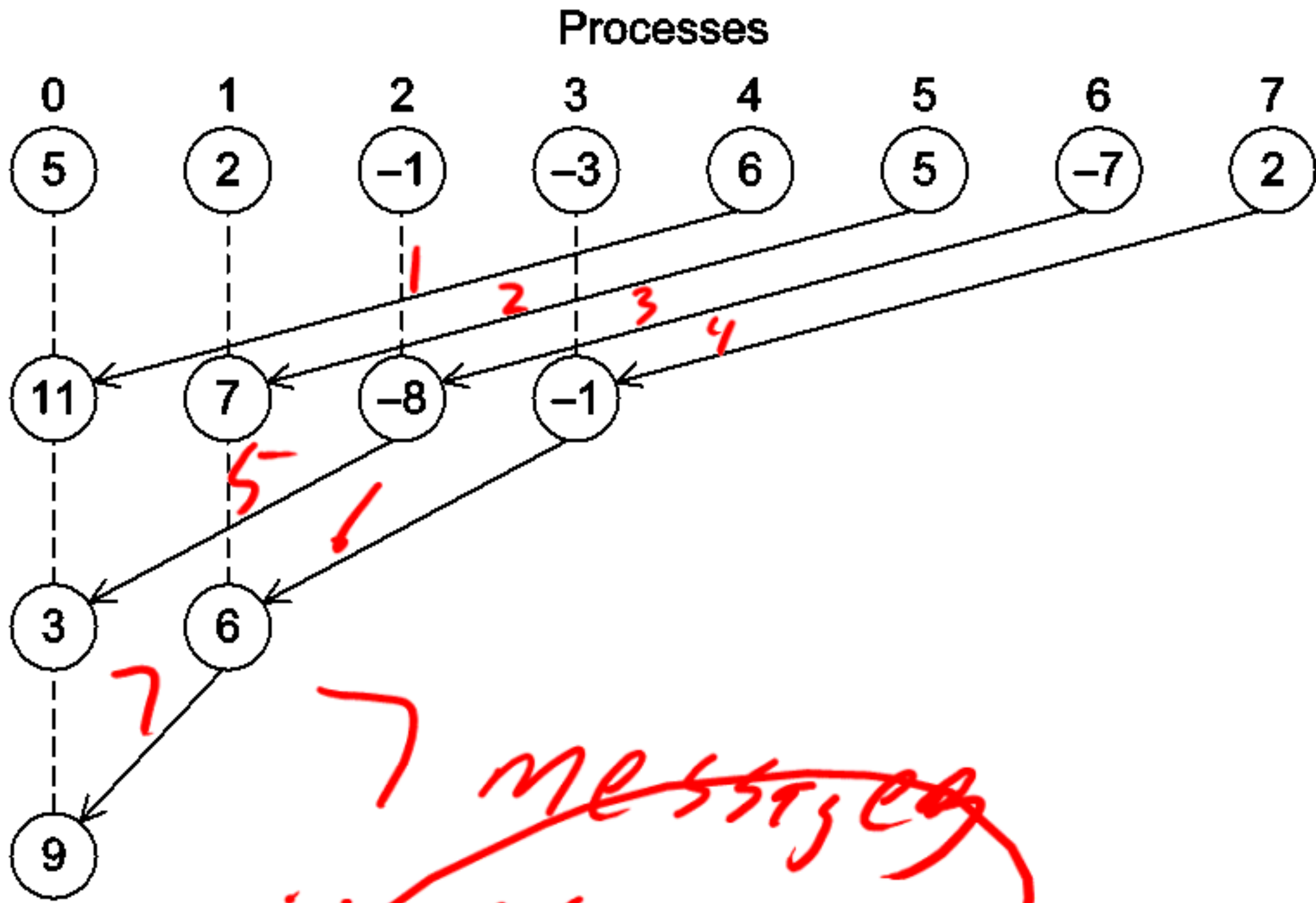
A Tree structured global

sum



w/m MPI, each Arrow
is a message.

An alternative global sum



7 messages
No gain

What is the problem here?

Message take time
to send. \therefore We
want to minimize
when possible the
number of msgs
being sent.

- Point to Point Communications
 - Communications in which two processes in a communicator communicate

- MPI_Send
- MPI_Receive



*Socket
Read/Write*

- Collective Communications
 - A communication function that involves all processes in a communicator

- MPI_Reduce
- MPI_Bcast
- MPI_Scatter
- MPI_Gather

*No examples
it thus far.*

Solution: MPI global sum

reduction methods

http://www.open-mpi.org/doc/v1.4/man3/MPI_Reduce.3.php

```
int MPI_Reduce(void *sendbuf, void *recvbuf, int count,  
              MPI_Datatype datatype, MPI_Op op, int root,  
              MPI_Comm comm)
```

- Sendbuf — *ptr*
→ Address of send buffer (choice).
- Count —
— Number of elements in send buffer (integer).
- Datatype .
— Data type of elements of send buffer (handle).
- Op .
— Reduce operation (handle).
- Root
— Rank of root process (integer).
- Comm
— Communicator (handle).

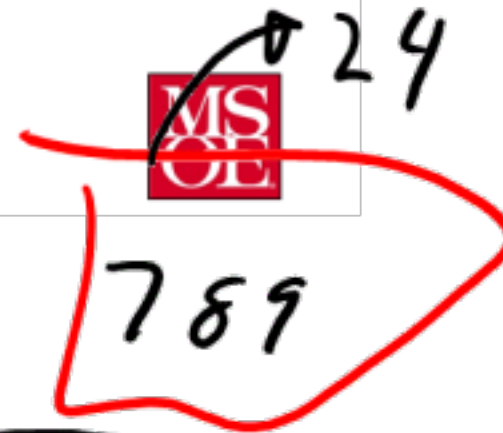
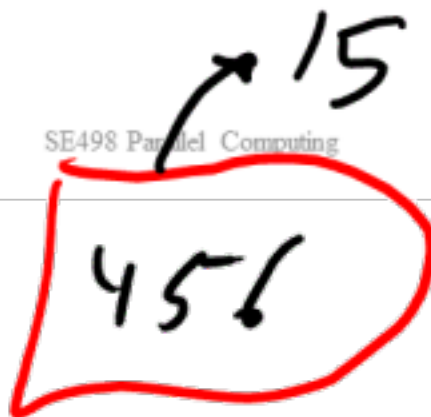
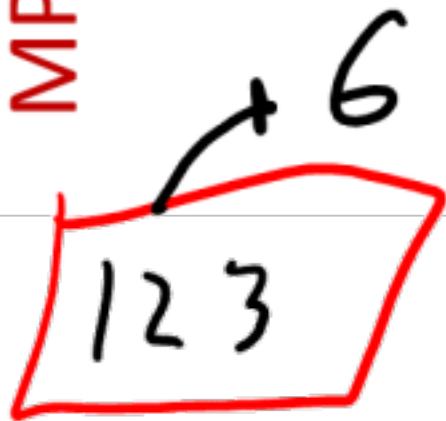
↑ what reduction op to do.

what to do w/ the data response.

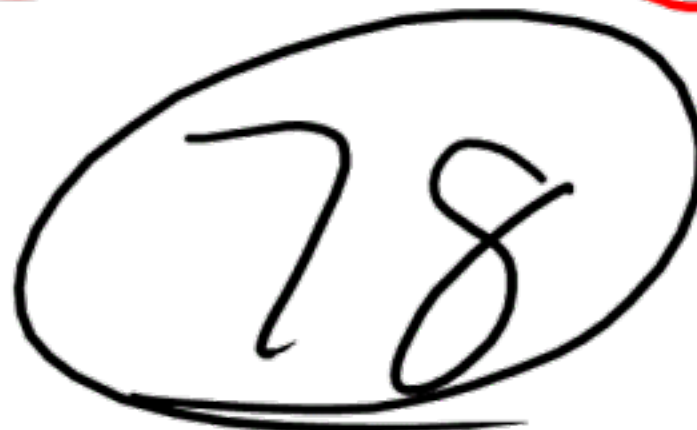


MPI Reduction Operators

Operation Value	Meaning
MPI_MAX	Maximum
MPI_MIN	Minimum
MPI_SUM	Sum
MPI_PROD	Product
MPI_LAND	Logical and
MPI_BAND	Bitwise and
MPI_LOR	Logical or
MPI_BOR	Bitwise or
MPI_LXOR	Logical exclusive or
MPI_BXOR	Bitwise exclusive or
MPI_MAXLOC	Maximum and location of maximum
MPI_MINLOC	Minimum and location of minimum



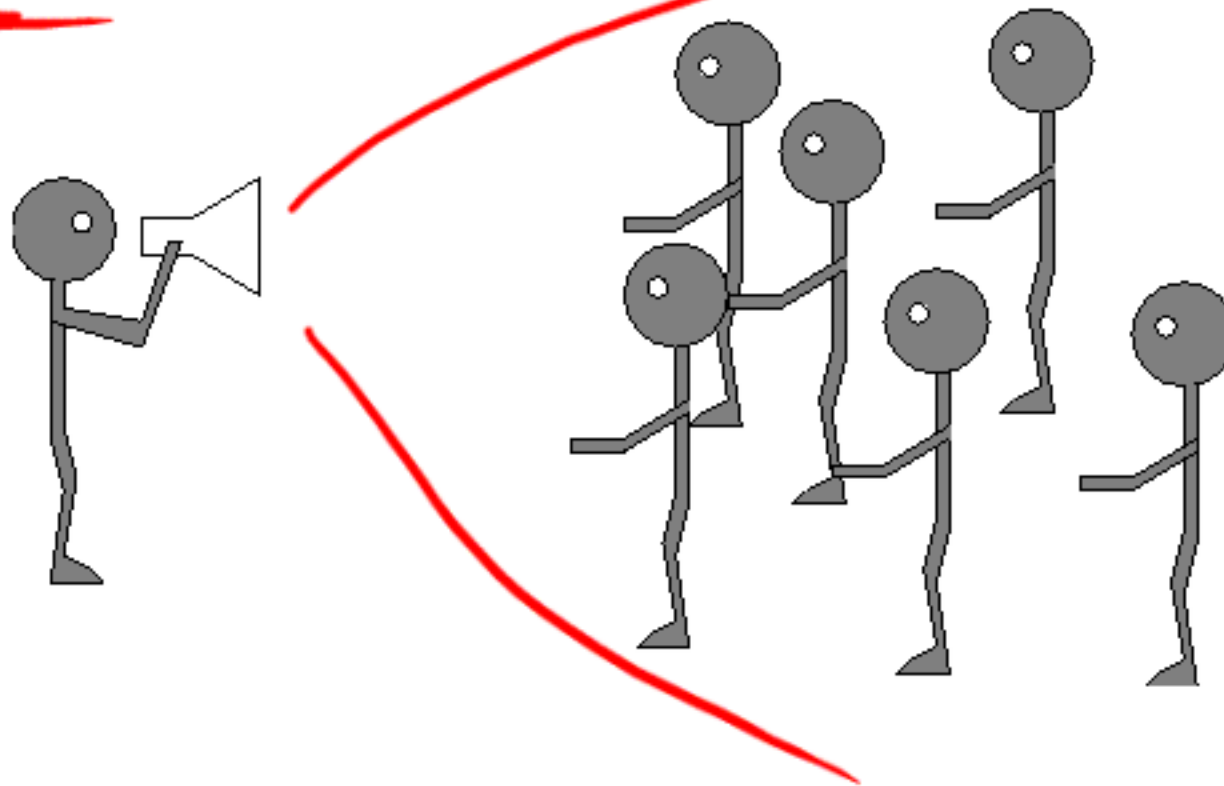
MPI Reduce:



Lets look at our pi Example using an MPI Reduce

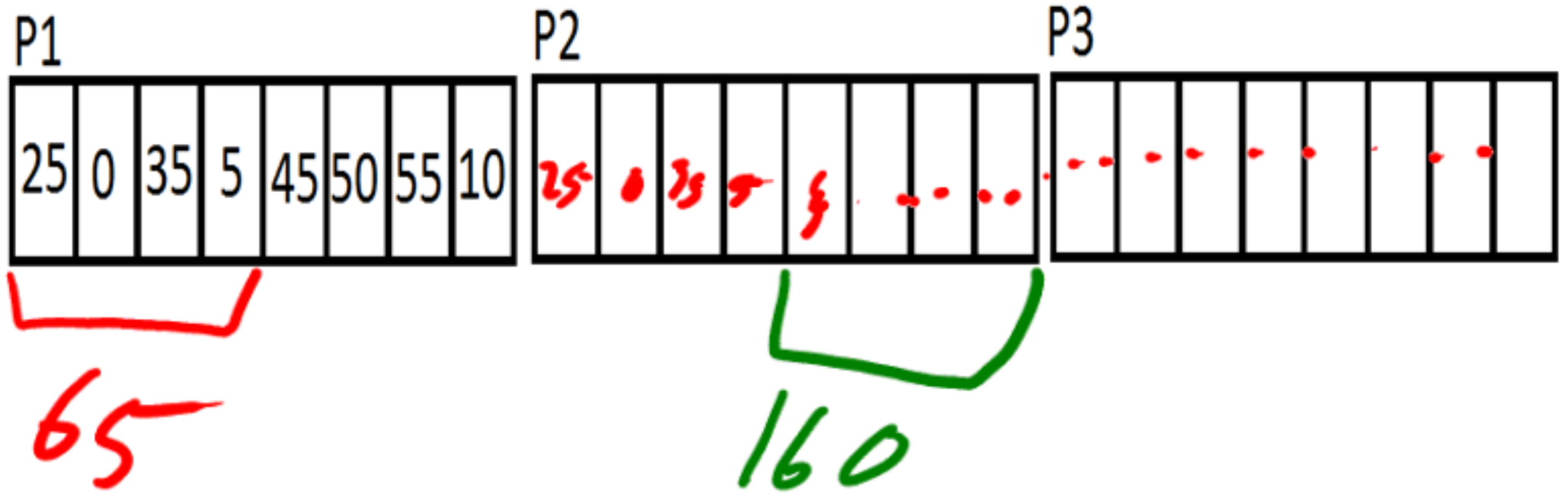
MPI Broadcast

- Data belonging to a single process is sent to all other processes.



Everyone has
a copy.

MPI Broadcast



MPI Broadcast Example

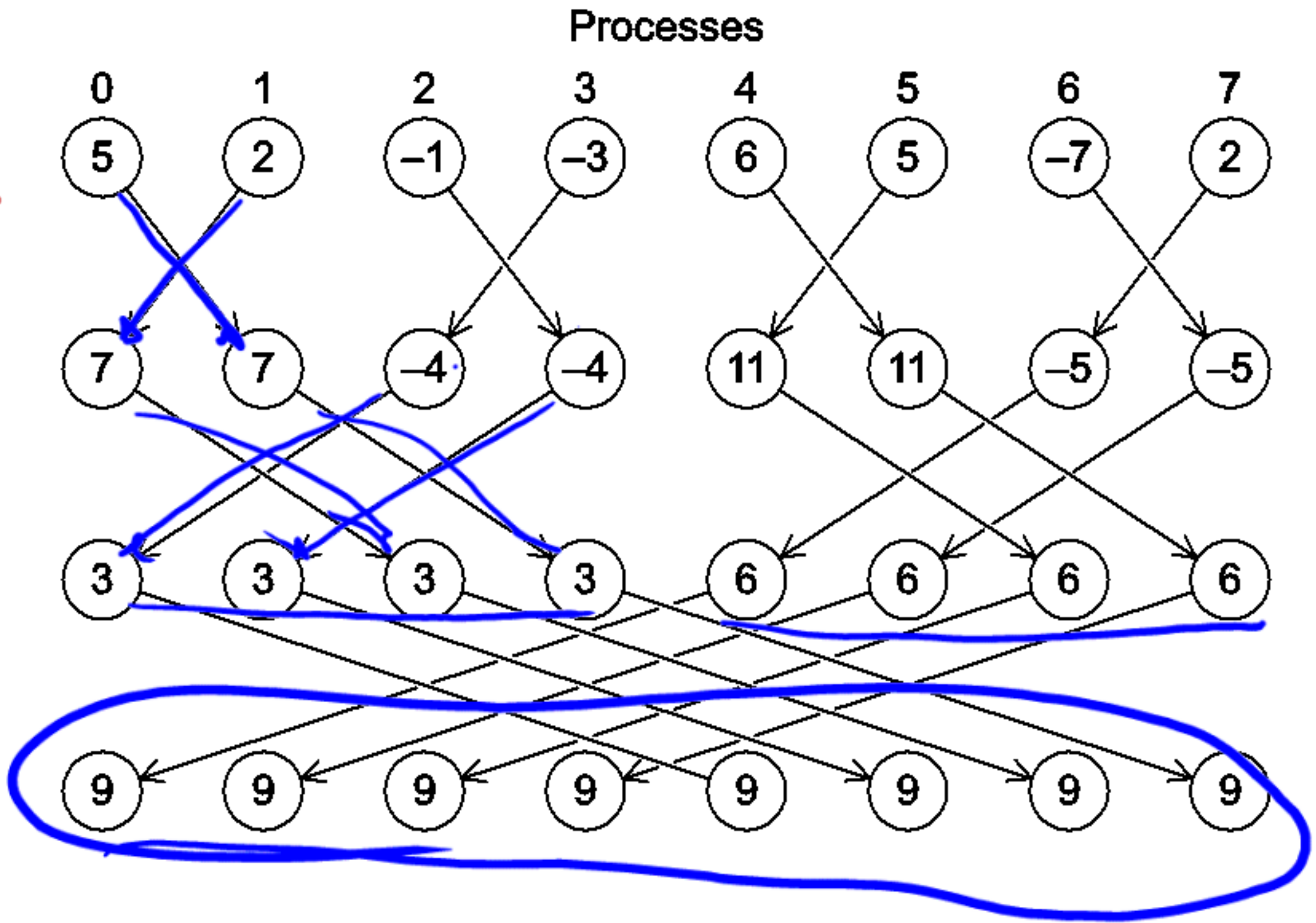
- Broadcast Example

- What is MPI_Allreduce?

It is a puzzlement

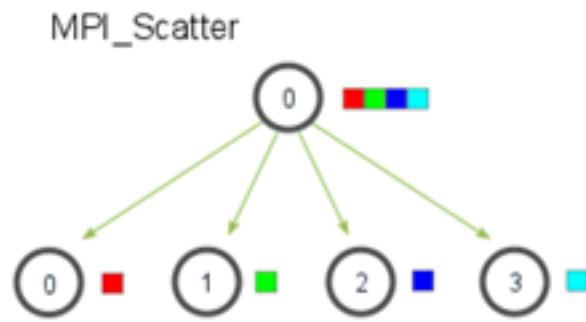
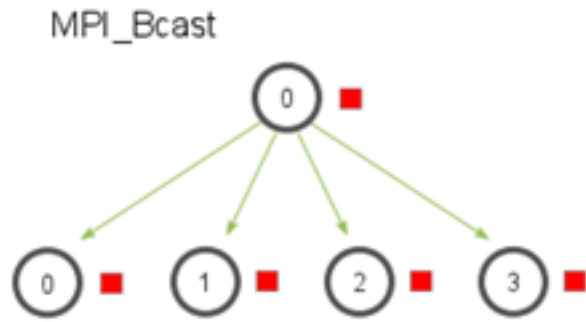
Reduced
the results and
populated the result
to all nodes.

MPI All Reduce example



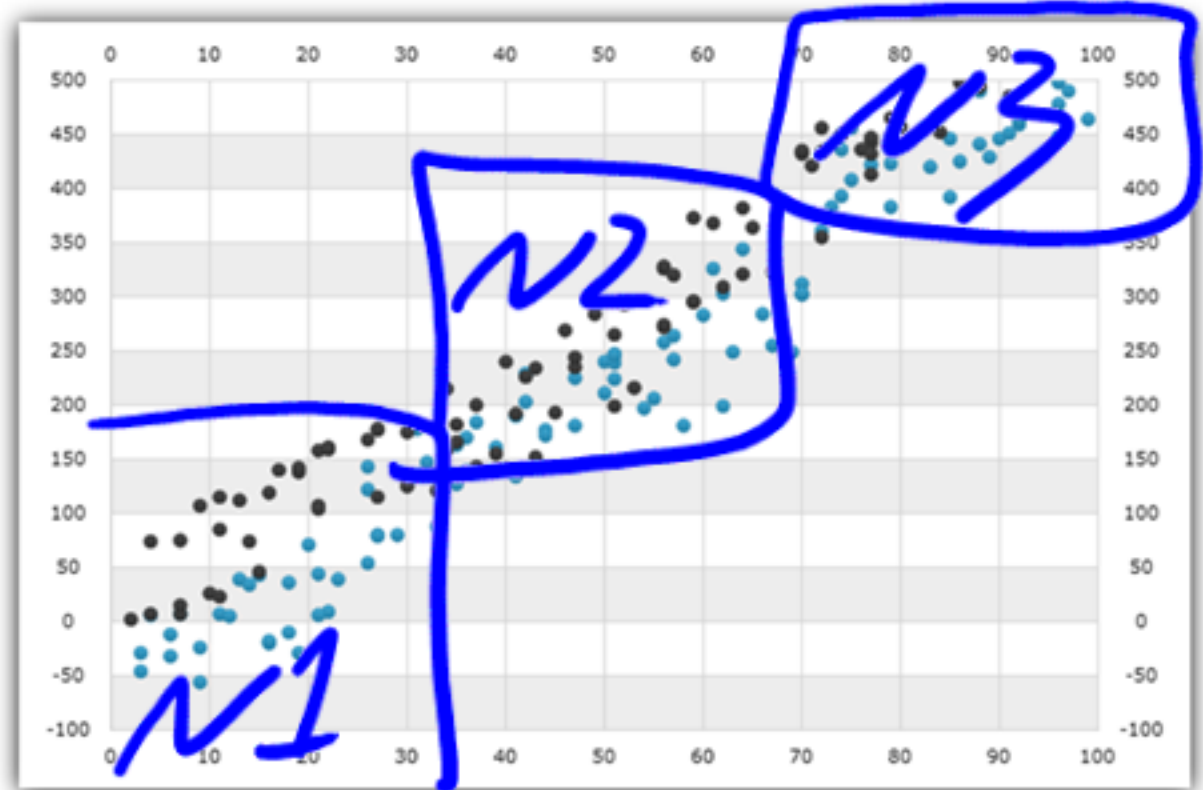
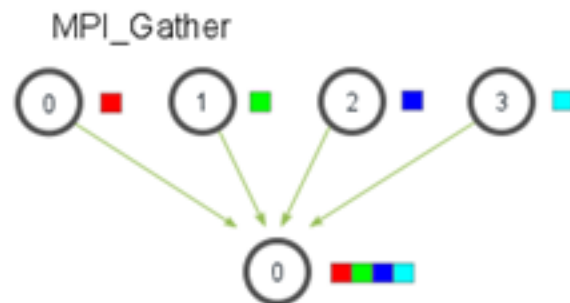
- Scatter – Gather

- Scatter sends a portion of the data to each process



- Gather

- Combines the results of the scatter back together



Scatter-Gather Averaging

Hopefully
will work -1

A Better scatter gather

example

- I want to figure out the distribution of random numbers between 0 and 1
 - Where do they fall overall?