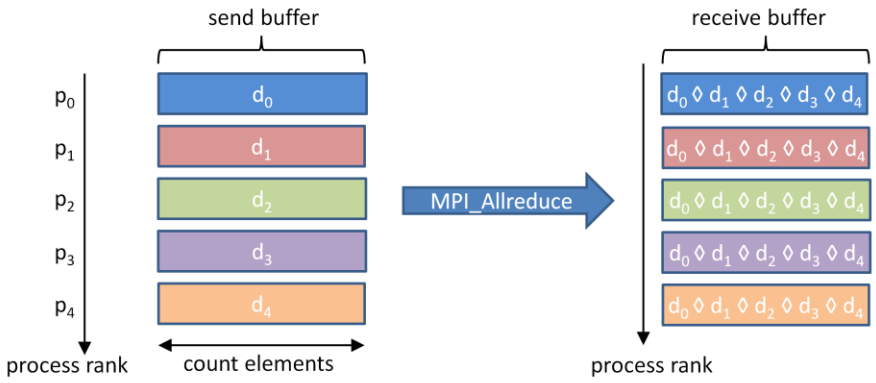
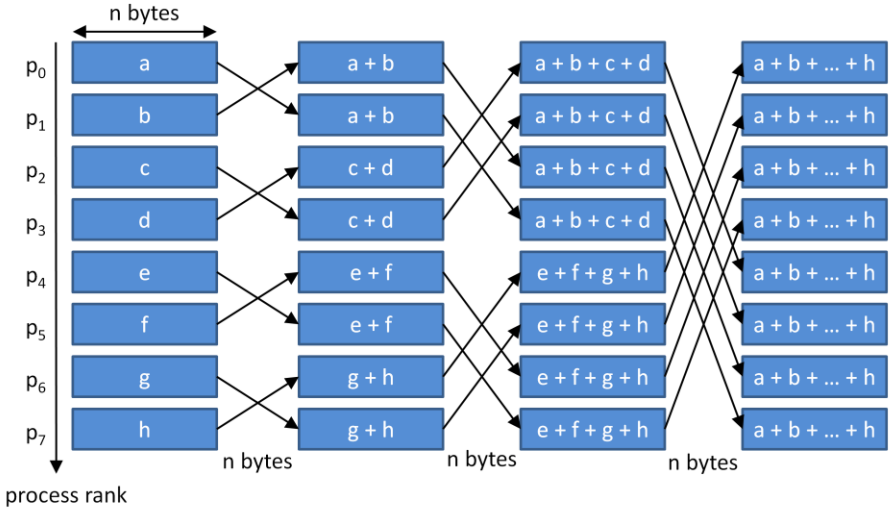
Allreduce algorithm

```
MPI_Allreduce(void *sendbuf, void *recvbuf, int count,
              MPI_Datatype dataType, MPI_Op op, MPI_Comm comm)
```

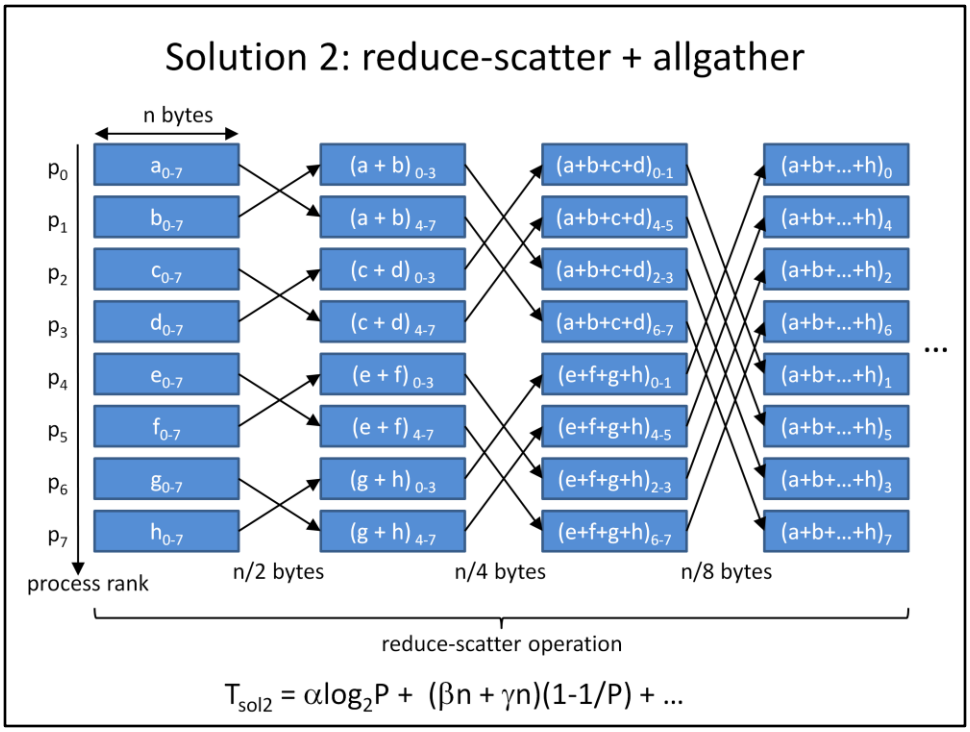


Q: Develop an algorithm for the allreduce routine. Derive the time complexity for that algorithm. Assume a fully non-blocking (duplex) network.

Solution 1: butterfly communication scheme

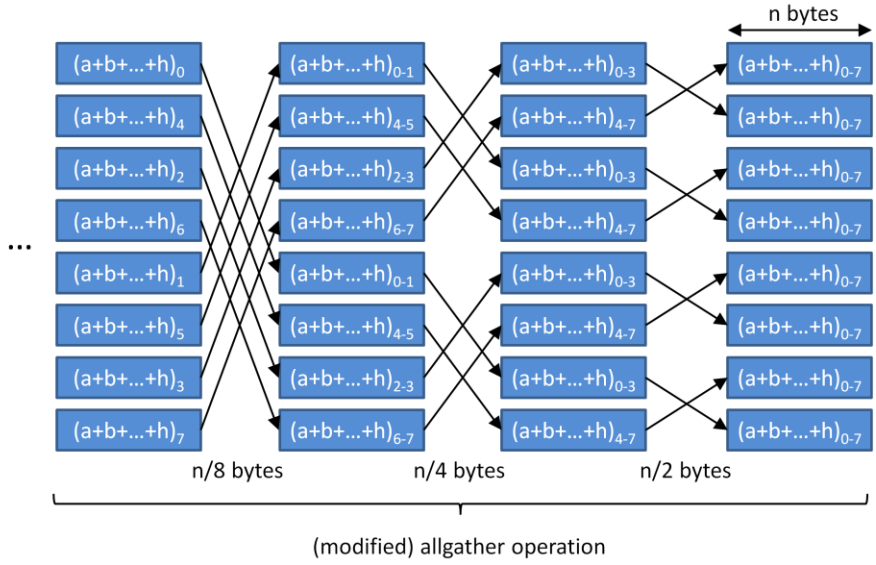


$T_{sol1} = (\alpha + \beta n + \gamma n) \log_2 P$ (assume $P = 2^i$)



Warning: box sizes are not drawn to scale.

Solution 2: reduce-scatter + allgather (cont'd)



$T_{sol2} = 2\alpha \log_2 P + (2\beta n + \gamma n)(1-1/P)$ (assume $P = 2^i$)