Summary of last lecture

- One-to-all Broadcast
- All-to-all broadcast
A2A broadcast algorithm

**Procedure** ALL_TO_ALL_BC_HCUBE($d$, $my_id$, $my_msg$, $result$)

begin

$\text{result} := my\_msg$;

for $I := 0$ to $d - 1$ do

begin

$partner := my\_id \text{ XOR } 2^i$

send $result$ to $partner$;

receive $msg$ from $partner$;

$result := result \cup msg$;

end;

end ALL_TO_ALL_BC_HCUBE
A2A Broadcast with CT routing

For A2A broadcast the CT case does not improve over the S&F

- note the $mt_w(p - 1)$ term appearing in the S&F expressions for all the topologies
- the same term is a lower bound for the communication time on parallel computers where processors can communicate on one port at a time
One-to-all personalized communication

- A.k.a single-node scatter
  - $p$ messages of size $m$ on root processor - each one sent to a different processor
  - different from one-to-all broadcast because data is different for each destination

- Dual operation: single-node gather
  - a single node collects a unique message from each other processor
  - different from accumulation operation because there is no combination of data
One-to-all personalized on hypercube

- Duration of step $i$: $t_s + mt_w(p/2^i)$
- Total duration
  - $T_{one\_to\_all\_pers} = t_s \log p + mt_w(p - 1)$
  - same as time required for all-to-all broadcast
All-to-all personalized communication

- **A.k.a total exchange**
  - each processor sends a distinct message to of size $m$ to every other processor
A2A personalized communication: ring (S&F)

\[ T_{all\_to\_all\_pers} = \sum_{i=1}^{p-1} (t_s + mt_w(p-i)) = (t_s + mpt_w/2)(p - 1) \]
A2A personalized communication: mesh (S&F)

Data distribution at the beginning of first phase

- \( T_{all\_to\_all\_pers} = (2t_s + mpt_w)(\sqrt{p} - 1) \)
A2A personalized communication: hypercube (S&F)

- $T_{all\_to\_all\_pers} = (t_s + mpt_w/2)(\log p)$
Table 4.2  MPI names of the various operations discussed in this chapter.

<table>
<thead>
<tr>
<th>Operation</th>
<th>MPI Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-to-all broadcast</td>
<td>MPI_Bcast</td>
</tr>
<tr>
<td>All-to-one reduction</td>
<td>MPI_Reduce</td>
</tr>
<tr>
<td>All-to-all broadcast</td>
<td>MPI_Allgather</td>
</tr>
<tr>
<td>All-to-all reduction</td>
<td>MPI_Reduce_scatter</td>
</tr>
<tr>
<td>All-reduce</td>
<td>MPI_Allreduce</td>
</tr>
<tr>
<td>Gather</td>
<td>MPI_Gather</td>
</tr>
<tr>
<td>Scatter</td>
<td>MPI_Scatter</td>
</tr>
<tr>
<td>All-to-all personalized</td>
<td>MPI_Alltoall</td>
</tr>
</tbody>
</table>