Implementing the Bulk Synchronous Parallel Model in CIEL

Valentin Dalibard
Bulk Synchronous Parallel Model

- A computation is a sequence of iterations
- Each iteration is called a superstep
- User defined function computed at each vertex in parallel
CIEL

• Allows iterative and recursive task-parallel algorithms
• Dynamically builds the DAG
## Comparison to Pregel

<table>
<thead>
<tr>
<th>What I will implement</th>
<th>What I won’t implement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master-Worker Model</td>
<td>Fault Tolerance</td>
</tr>
<tr>
<td>Vote to halt</td>
<td>Aggregators</td>
</tr>
<tr>
<td>Topology mutation</td>
<td>Combiners</td>
</tr>
</tbody>
</table>
Evaluation against Giraph

- Open source of Pregel
- Runs on top of Hadoop

**Thread architecture**

Map-only job in Hadoop

Map 0  Map 1  Map 2  Map 3  Map 4

Thread assignment in Giraph

Master  Worker  ZooKeeper