PDES in an Adaptive Runtime System

UIUC: Eric Mikida, Nikhil Jain, Joshua Lew, Laxmikant Kale
RPI: Chris Carothers, Elsa Gonsiorowski
LLNL: David Jefferson, Peter Barnes
What is PDES?

● Simulation of events at discrete points in virtual time
  ○ Contrary to time-stepped, work is not uniform/dense in time
  ○ Events must be executed in increasing order of virtual time
  ○ Traffic, Supercomputer Interconnects, Circuits, Battle Sims

● Events executed by Logical Processes (LPs)
● Synchronization required to maintain event order
● Focus is on ROSS and the Time Warp protocol
Basic Terminology

- Logical Process (LP)
- Total Executed Events
- Committed Events
- Rolled-back Events
- Global Virtual Time (GVT)
- Fossil Collection
- Event Efficiency
- Event Rate
A (Small) Example

LP0

Current Time: 1

1 6 13 15 19

Total Events: 0
Total Rollbacks: 0
Committed Events: 0

LP1

Current Time: 1

1 2 4 10 15 21
A (Small) Example

Current Time: 6

LP0

Current Time: 2

LP1

Total Events: 2
Total Rollbacks: 0
Committed Events: 0
A (Small) Example

Current Time: 13

LP0
1 6
13 15 19

Current Time: 4

LP1
1 2
4 10 15 21

Total Events: 4
Total Rollbacks: 0
Committed Events: 0
A (Small) Example

Current Time: 15
Total Events: 6
Total Rollbacks: 0
Committed Events: 0

Current Time: 10

LP0

1 6 13
5

LP1

1 2 4
10 15 21
A (Small) Example

Current Time: 13
Total Events: 6
Total Rollbacks: 1
Committed Events: 0

Current Time: 10

LP0
1 6 13 15 19

LP1
1 2 4 10 15 21

5
A (Small) Example

Total Events: 6
Total Rollbacks: 2
Committed Events: 0
A (Small) Example

Current Time: 5

LP0

1 5 6 13 15 19

Total Events: 6
Total Rollbacks: 2
Committed Events: 0

Current Time: 10

LP1

1 2 4 10 15 21

10
A (Small) Example

Current Time: 6

LP0
1 5 6 13 15 19

Total Events: 8
Total Rollbacks: 2
Committed Events: 0

Current Time: 15

LP1
1 2 4 10 15 21

11
A (Small) Example

GVT COMPUTATION: Find the minimum time

LP0

1 5

6 13 15 19

Current Time: 6

Total Events: 8
Total Rollbacks: 2
Committed Events: 0

LP1

1 2 4 10

15 21

Current Time: 15

12
A (Small) Example

Total Events: 8
Total Rollbacks: 2
Committed Events: 5

FOSSIL COLLECTION: Commit and free events

Current Time: 6
Current Time: 15
A (Small) Example

Current Time: 6
LP0
6 13 15 19

Current Time: 15
LP1
10 15 21

Total Events: 8
Total Rollbacks: 2
Committed Events: 5

Event Efficiency: \((E_C - E_R)/E_C\)

\((5-2)/5 = 60\%\)
A (Small) Example

Current Time: 6

LP0

6
13
15
19

Total Events: 8
Total Rollbacks: 2
Committed Events: 5

Event Efficiency: \( \frac{E_C - E_R}{E_C} \)

\( \frac{5-2}{5} = 60\% \)

Current Time: 15

LP1

10
15
21

Event Rate: \( \frac{E_C}{t} \)
Previous Work

- Implemented a version of ROSS on top of Charm++
  - LPs are chares: adaptive overlap, migration, location mgmt.
  - Good match to programming model: 7,277 \( \Rightarrow \) 3,991 SLOC

- Improved performance on PHOLD and Dragonfly models
  - Up to 40% higher event rate for PHOLD
  - Up to 5x higher event rate for Dragonfly model
New Work - Load Balancing

- LPs are now migratable by the runtime system
- Load of each chare measured by RTS
- Three different load balancing strategies:
  - GreedyLB
  - DistributedLB
  - HybridLB
Load Balancing - PHOLD

- Basic PHOLD micro-benchmark to test comm loads
  - Each event causes a new event to be created/sent
  - New events are remote sends with probability $P$
  - Results in a uniform and balanced execution

- Added two sources of imbalance
  - Work imbalance: some LPs take longer per event
  - Event imbalance: some LPs receive more events

- All runs on 64 nodes of Vesta (BG/Q machine at ANL)
Load Balancing - PHOLD

Event Rate for PHOLD Work

Event Counts for PHOLD Work
Load Balancing - PHOLD

Event Rate for PHOLD Event

- No LB
- GreedyLB
- HybridLB
- DistributedLB

Event Counts for PHOLD Event

- No LB
- GreedyLB
- HybridLB
- DistributedLB

Event Rate (million events/s)

Billions of Events

Committed Events
Total Events
Rolled Back Events
Load Balancing - PHOLD

Event Rate for PHOLD Combo

- No LB: 60 million events/s
- GreedyLB: 100 million events/s
- HybridLB: 80 million events/s
- DistributedLB: 60 million events/s

Event Counts for PHOLD Combo

- No LB: 3.05 billion
- GreedyLB: 25 billion
- HybridLB: 20 billion
- DistributedLB: 8 billion

Efficiency: 82% for GreedyLB, 97% for DistributedLB

Total Events:
- No LB: 20.2 billion
- GreedyLB: 17.7 billion
- HybridLB: 19.8 billion
- DistributedLB: 19.8 billion

Rollbacks:
- No LB: 3.05 billion
- GreedyLB: 0.51 billion
- HybridLB: 0.51 billion
- DistributedLB: 0.51 billion
Load Balancing - Overhead

- Data collection, decision cost, migration cost vary
- Reasons to use lower overhead strategies
  - Shorter duration simulations
  - Simulations that require frequent load balancing
  - Simulations with high memory consumption

<table>
<thead>
<tr>
<th></th>
<th>No LB</th>
<th>GreedyLB</th>
<th>HybridLB</th>
<th>DistributedLB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Runtime</td>
<td>311 seconds</td>
<td>179 seconds</td>
<td>247 seconds</td>
<td>321 seconds</td>
</tr>
<tr>
<td>LB Time</td>
<td>N/A</td>
<td>8 seconds</td>
<td>0.9 seconds</td>
<td>0.02 seconds</td>
</tr>
</tbody>
</table>
Load Balancing - Traffic

- LPs are intersections arranged in a grid
  - Events are cars arriving, departing, changing lane
  - Cars travel from source to destination
  - Roads have a capacity, cars wait until a road is free

- Two imbalanced configurations
  - Source congestion - many cars start from one place
  - Destination congestion - many cars going to one place
Load Balancing - Traffic

Event Rate for Traffic Src

Event Rate for Traffic Dest

Event Rate (million events/s)
Future Work

- Look deeper into DistributedLB
- Look into other strategies (communication aware)
- More interesting load metrics focused on event efficiency
- Joining GVT improvements with load balancing