Mohamed Taher¹, Kris Gaj², Tarek El-Ghazawi¹, and Nikitas Alexandridis¹
¹ The George Washington University
² George Mason University

Job Management System Extension to Support SLAAC-1V Reconfigurable Hardware
Problem:

- Reconfigurable resources expensive and unutilized

- Many of these resources available over the network

- It is desirable to leverage networked reconfigurable resources to help other users within the same organization
Approach of Job Management System
The LSF Suite (Load Sharing Facility) is an industry-standard set of integrated products that manage distributed computing resources and workloads.

LSF is a layer of software services on top of UNIX and Windows NT operating systems.

LSF creates a single system image on a network of heterogeneous computers so that the whole network of computing resources can be utilized effectively and managed easily.
SLAAC1-V

- Full-sized 64-bit PCI card.

- Three Xilinx XCV1000-6 compute FPGAs

- The memories are 36x256K ZBT SRAMs.
SLAAC1-V Top Level Block Diagram

PE datapath ports:
X : XBAR
L : LEFT
R : RIGHT

64/66 PCI
Approach

• Extend LSF to recognize and utilize SLAAC1-V reconfigurable resources
  - add new dynamic resources
  - configure scheduling to be based on these new resources
Networked Reconfigurable Resource Management System

Submission Host

Master Host

Tasks 1, 2, 3

Task 1

Task 2

Task 3

Execution Host 1

Execution Host 2

Execution Host 3

FPGA boards
LSF Structure

User Server

Job Scheduler

Resource Manager

Resource Monitor

Job Dispatcher

scheduling policies

resource requirements

available resources

resource allocation and job execution

jobs & their requirements
General Architecture of LSF

Submission host

Master host

Execution host

Batch API

MLIM

MBD

LIM

SBD

Child SBD

RES

User job

bsub app

queue

Load information

other hosts

other hosts

LIM – Load Information Manager
MLIM – Master LIM
MBD – Master Batch Daemon
SBD – Slave Batch Daemon
RES – Remote Execution Server
Extension of LSF to reconfigurable hardware

Master host

Remote user

MLIM

MBD

other hosts

Execution host

Load information

other hosts

LIM

SBD

Child SBD

RES

RUser job

SLAAC API

FPGA board

ELIM

SLAAC API

Status of the board

Remote user

bsub

queue

P27

11
Try to open the board in an exclusive mode

Yes

Status = board available

Close the board

Succeed

No

Status = board unavailable

Report the status of the board to LIM

Wait for \( N \) sec
Experimental verification
Testbed used in the experiments

Submission & Master Host
Linux RH7.0 – PIII
450 MHz, 512 MB RAM

Execution Host 1
Windows 2000 – PIV
1.3 GHz, 256 MB RAM

Execution Host 2
Windows 2000 – PIII
450 MHz, 128 MB RAM
## Parameters of Performed Experiments

<table>
<thead>
<tr>
<th>Exp. No.</th>
<th>No. of execution hosts</th>
<th>Number and execution times of jobs</th>
<th>Delay between job submissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>40 x 20 s</td>
<td>5 s</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>8 x 20 s, 8 x 30 s, 8 x 40 s, 8 x 50 s, 8 x 60 s</td>
<td>5 s</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>40 x 120 s</td>
<td>5 s</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>40 x 300 s</td>
<td>5 s</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3 x 20 s, 3 x 40 s, 3 x 60 s, 3 x 80 s, 3 x 100 s, 3 x 120 s, 3 x 140 s, 3 x 160 s, 3 x 180 s, 3 x 200 s, 3 x 220 s, 3 x 240 s, 3 x 260 s, 3 x 280 s, 3 x 300 s</td>
<td>5 s</td>
</tr>
</tbody>
</table>
Utilization of machines in Experiment 1, Iteration 1

Board availability according to:
MLIM

ELIM

Jobs executed on :hpcI6

Board availability according to:
MLIM

ELIM

Jobs executed on :hpcI3

Job Submissions

Time (sec)
Utilization of machines in Experiment 2, Iteration 1

Board availability according to:
- MLIM
- ELIM

Jobs executed on :hpcl6

Board availability according to:
- MLIM
- ELIM

Jobs executed on :hpcl3

Job Submissions

Time (sec)
Utilization of machines in Experiment 3, Iteration 1

Board availability according to:
- MLIM

- ELIM

Jobs executed on :hpcl6

Board availability according to:
- MLIM

- ELIM

Jobs executed on :hpcl3

Job Submissions

Time (sec)
Utilization of machines in Experiment 4, Iteration 1

Board availability according to:
- MLIM
- ELIM

Jobs executed on :hpcl6

Board availability according to:
- MLIM
- ELIM

Jobs executed on :hpcl3

Job Submissions
Utilization of machines in Experiment 5, Iteration 1

Board availability according to:
MLIM

ELIM

Jobs executed on :hpc16

Board availability according to:
MLIM

ELIM

Jobs executed on :hpc13

Job Submissions
## Result

<table>
<thead>
<tr>
<th>Experiment No.</th>
<th>Utilization [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iteration 1</td>
</tr>
<tr>
<td>1</td>
<td>65.9</td>
</tr>
<tr>
<td>2</td>
<td>75.6</td>
</tr>
<tr>
<td>3</td>
<td>90.8</td>
</tr>
<tr>
<td>4</td>
<td>95</td>
</tr>
<tr>
<td>5</td>
<td>91.9</td>
</tr>
</tbody>
</table>
Conclusion

• An extension of LSF, supporting SLAAC-1V FPGA accelerator boards was developed and experimentally tested

• Full implementation and verification of the middleware for SLAAC-1V boards

• The architecture was verified experimentally

• The utilization of the idle boards was demonstrated to reach up to 95% in our experimental setting which include Linux and Windows NT workstations