Using Work Queue Inside and Outside the Classroom

Peter Bui

University of Wisconsin - Eau Claire
Motivation
Parallel and distributed computing are becoming increasingly important.
Solution: Work Queue

Use Work Queue to introduce undergraduates to parallel and distributed computing in the context of:

- Class Assignments
- Independent Study
- Undergraduate Research
Background
Teaching and Research @ UWEC

Liberal arts, undergraduate focused university.

Emphasize teaching and faculty/student collaborative research.
Distributed Computing @ UWEC

CENTER FOR HIGH THROUGHPUT COMPUTING
Teaching
Teaching Approach

● **Core**
  ○ Computer Organization and Design

● **Service**
  ○ Introduction to Programming in C++
  ○ Computing for the Sciences and Mathematics

● **Elective**
  ○ Cloud Computing
  ○ Unix Systems Programming
Work Queue MapReduce

Name: wqmr-buipj
Port: 9001
Work Directory: ./wqmr-buipj

Map: [==============================================] 100.00%
Reduce: [==============================================] 100.00%

Workers: I: 0 R: 12 B: 0 J: 12 Q: 0
Tasks: W: 0 R: 0 U: 0 D: 981 C: 981
Data: S: 237.21MB R: 706.65KB

Elapsed Time: 58
Elapsed Work Time: 45

Last Event: Task r0000 returned with exit status 0
Brute-force Password Cracking

Using **Go**, students implemented a brute-force password cracker:

- **Serial** version
- **Parallel** version using **CSP**
- **Distributed** version using **Work Queue**
Research
Research Approach

- Focus on high-level applications
- Take advantage of frameworks
- Keep students engaged and motivated
- Reach out and take advantage of resources

*Work Queue* is a great framework for enabling novice users to explore parallel and distributed computing.
Distributed Animation Rendering
Distributed Animation Rendering

![Graph showing the average number of seconds to complete rendering as a function of the number of submitted workers, with lines for different scenes and frame counts. The graph demonstrates how rendering time decreases with an increasing number of workers.]
Photo Processing Pipeline

1. Transfer photos from cameras to laptops
2. Upload photos to Dropbox
3. Monitor Dropbox for incoming photos
4. Dispatch processing tasks to Work Queue Pool
5. Archive photos and generated artifacts
6. Access processed photos via Web browser
Photo Processing Pipeline

Num of Batches 448
Max Batch Size 1385
Min Batch Size 1
Avg Batch Size 16.4
Num of Tasks Submitted 7372
Num of Tasks Failed 104
Scalable Image Transcoding

1. Process input images
2. Distribute tasks and inputs to remote workers
3. Perform transcoding and return outputs
# Scalable Image Transcoding

<table>
<thead>
<tr>
<th>File Size</th>
<th>Set Size</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>15KB</td>
<td>10</td>
<td>1x</td>
<td>1.47x</td>
<td>1.56x</td>
<td>2.13x</td>
<td>1.85x</td>
<td>2.00x</td>
<td>2.40x</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1x</td>
<td>1.60x</td>
<td>2.80x</td>
<td>4.43x</td>
<td>5.96x</td>
<td>6.42x</td>
<td>6.44x</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1x</td>
<td>1.65x</td>
<td>3.12x</td>
<td>5.02x</td>
<td>7.97x</td>
<td>9.27x</td>
<td>9.31x</td>
</tr>
<tr>
<td>1MB</td>
<td>10</td>
<td>1x</td>
<td>1.65x</td>
<td>2.40x</td>
<td>2.78x</td>
<td>3.05x</td>
<td>3.73x</td>
<td>3.87x</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1x</td>
<td>2.10x</td>
<td>3.87x</td>
<td>6.55x</td>
<td><strong>9.56x</strong></td>
<td>7.65x</td>
<td>8.27x</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1x</td>
<td>2.17x</td>
<td>4.28x</td>
<td>7.75x</td>
<td>11.2x</td>
<td>10.5x</td>
<td><strong>12.12x</strong></td>
</tr>
<tr>
<td>10MB</td>
<td>10</td>
<td>1x</td>
<td>1.84x</td>
<td>2.46x</td>
<td>2.88x</td>
<td><strong>4.48x</strong></td>
<td>3.43x</td>
<td>3.27x</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1x</td>
<td>1.98x</td>
<td>3.90x</td>
<td>4.95x</td>
<td><strong>7.34x</strong></td>
<td>4.61x</td>
<td>4.76x</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1x</td>
<td>1.74x</td>
<td>3.97x</td>
<td>5.63x</td>
<td><strong>6.26x</strong></td>
<td>4.75x</td>
<td>4.93x</td>
</tr>
</tbody>
</table>
Final Thoughts
Summary

- **Work Queue** is easy to use.
- **Work Queue** is flexible.
- **Work Queue** is portable.
- **Work Queue** is extensible.

**Work Queue** is a great way to introduce **PDC** to undergraduate students!
Future Work

Teaching

- Incorporate into Computational Science course
- Introduce in Computer Science systems course

Research

- Web portal for art students to utilize DSABR
- Visualization and monitoring of Work Queue
Acknowledgements

- **Students**
  Travis Boettcher, Nick Jaeger, Jeffrey Westphal

- **ORSP**
  Travel funding and student stipends

- **CHTC**
  HTCondor flocking
Questions?

Peter Bui

EMail: buipj@uwec.edu

WWW: http://cs.uwec.edu/~buipj