CCTools
New Capabilities

2012-2013
CCTools Suite Overview

Scientist’s Laptop

Demo Running Locally

Local Filesystem

task 1

task 2

task ...

task n
Local to High Throughput

Task 1
Task 2
Task...
Task n

Demo Running Locally
Local Filesystem

Campus grids and clusters
Opportunistic resources
Cloud
Our Philosophy

✤ Harness all the resources that are available: desktops, clusters, clouds, and grids.

✤ Make it easy to scale up from one desktop to national scale infrastructure.

✤ Provide familiar interfaces that make it easy to connect existing apps together.

✤ Allow portability across operating systems, storage systems, middleware…

✤ Make simple things easy, and complex things possible.

✤ No special privileges required.
CCTools Suite

Makeflow

portable workflow manager
run this task when this other task is completed
make syntax

Tuesday, October 15, 13
CCTools Suite

Makeflow
portable workflow manager
run this task when this other task is completed
make syntax

local  work queue  condor  sge  ...

Tuesday, October 15, 13
CCTools Suite

AllPairs

- specialized execution engine
- apply the operation to all combinations
- cross products

Wavefront

- specialized execution engine
- run this operation following a wave pattern
- dynamic programming
CCTools Suite

Work Queue

lightweight distributed execution system
run this task in some available resource
master-worker architecture
CCTools Suite

Work Queue
lightweight distributed execution system
run this task in some available resource
master-worker architecture

Makeflow
Work Queue API
C Python Perl

cloud
campus grids and clusters
opportunistic resources
CCTools Suite

Chirp
user-level distributed filesystem
file sharing for scientific workflows
mount volumes without root access

Parrot
personal user-level virtual filesystem
access remote files as if they were local
system call interposition agent

at some_sever.xyz
chirp_server -r /home/user/my_files

at the local machine
parrot_run ls /chirp/some_server.xyz/*/*.txt
parrot_run cat /http/www.nd.edu/index.html
parrot_run ls /anonftp/ftp.gnu.org/pub
parrot_run cd /cvmfs/cms.cern.ch
CCTools Suite

* Open source, GNU General Public License 2.

* Compiles in 1-2 minutes, installs in $HOME.

* All tools may be used independently or in conjunction.

* Runs on Linux, Solaris, OSX, Cygwin, FreeBSD, …

* Interoperates with many distributed computing systems.

  * Condor, SGE, Torque, Globus, iRODS, Hadoop…
New Capabilities
1. Limited scalability: ~1000 (number of open file descriptors)
2. Limited throughput: One task per worker
3. No authentication: Any worker can pull data from master
4. Big file transfers over WAN
Work Queue Hierarchy

- Master
  - Foreman
    -Thousands of Workers in SGE
  - Foreman
    -Thousands of Workers in FutureGrid
  - Foreman
    -Thousands of Workers in Condor
  - Foreman
    -Thousands of Workers in Stampede
  - Foreman
    -Thousands of Workers in Lonestar
Work Queue Hierarchy

Master

Foreman 1

Foreman 2

Worker

Worker

Worker

Foreman

Tuesday, October 15, 13
Work Queue Hierarchy

```
work_queue_specify_name(q, master_name);

makeflow -N master_name
```

Tuesday, October 15, 13
Work Queue Hierarchy

```
work_queue_specify_name(q, master_name);
makeflow -N master_name

work_queue_worker -f -M master_name -N foreman_2
```
Work Queue Hierarchy

- **Master**
  - `makeflow -N master_name`
  - `work_queue_specify_name(q, master_name);

- **Foreman 1**
  - `work_queue_worker -M foreman_1`

- **Foreman 2**
  - `work_queue_worker -f -M master_name -N foreman_2`

- **Worker**

---

Tuesday, October 15, 13
Work Queue Hierarchy

Foremen cache files

- Master
- Foreman
- big file A
- Foreman
- $$
- Worker
- Worker
- Worker
- Worker
Work Queue Hierarchy

- Foremen cache files

Diagram:
- Master
- Foreman
- Worker
  - big file A
  - $$$
  - big file A
  - $$$
  - Worker
  - Worker
  - Worker
  - Worker
Work Queue Hierarchy

- Foremen cache files
Work Queue Hierarchy

- Foremen cache files

- Master
- Foreman
  - big file A
- Foreman
  - big file A
- Worker
  - big file A
- Worker
  - big file A
- Worker
  - big file A
- Worker
Work Queue Hierarchy

* Foremen cache files

- Master
  - big file A
  - big file B
- Foreman
  - $$ big file A $$
- Foreman
  - $$ big file A $$
- Worker
  - big file A
- Worker
  - big file A
- Worker
  - big file B
- Worker
  - big file B
- Worker
  - big file B
Work Queue Hierarchy

- Foreman running on head nodes
## Work Queue Status

<table>
<thead>
<tr>
<th>PROJECT</th>
<th>HOST</th>
<th>PORT</th>
<th>WAITING</th>
<th>BUSY</th>
<th>COMPLETE</th>
<th>WORKERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>wq_chief</td>
<td>cclws16.cse.nd.edu</td>
<td>1024</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;wq-foreman-a</td>
<td>cclws03.cse.nd.edu</td>
<td>1024</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-&gt;wq-foreman-a1</td>
<td>cclws02.cse.nd.edu</td>
<td>1024</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-&gt;wq-foreman-a2</td>
<td>cclws17.cse.nd.edu</td>
<td>1028</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;wq-foreman-b</td>
<td>cclws15.cse.nd.edu</td>
<td>1025</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>arab_cnvrnt</td>
<td>crcfe01.crc.nd.edu</td>
<td>9551</td>
<td>0</td>
<td>88</td>
<td>902</td>
<td>213</td>
</tr>
<tr>
<td>forcebalance</td>
<td>fire-4-0-ext.slac.stan</td>
<td>9793</td>
<td>1418</td>
<td>18</td>
<td>43187</td>
<td>18</td>
</tr>
<tr>
<td>alldat_T_7</td>
<td>fugu.biostat.wisc.edu</td>
<td>9140</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>dihedral</td>
<td>leeping.Stanford.EDU</td>
<td>7329</td>
<td>524</td>
<td>0</td>
<td>77343</td>
<td>856</td>
</tr>
<tr>
<td>quad_cnvrnt</td>
<td>newcell.crc.nd.edu</td>
<td>9552</td>
<td>481</td>
<td>153</td>
<td>1608</td>
<td>153</td>
</tr>
<tr>
<td>cclosdc18</td>
<td>workspace.crc.nd.edu</td>
<td>9907</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cclosdc1</td>
<td>workspace.crc.nd.edu</td>
<td>9908</td>
<td>0</td>
<td>0</td>
<td>354</td>
<td>6</td>
</tr>
</tbody>
</table>
# Work Queue Status

<table>
<thead>
<tr>
<th>PROJECT</th>
<th>HOST</th>
<th>PORT</th>
<th>WAITING</th>
<th>BUSY</th>
<th>COMPLETE</th>
<th>WORKERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>wq_chief</td>
<td>cclws16.cse.nd.edu</td>
<td>1024</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;wq-foreman-a</td>
<td>cclws03.cse.nd.edu</td>
<td>1024</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-&gt;wq-foreman-a1</td>
<td>cclws02.cse.nd.edu</td>
<td>1024</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-&gt;wq-foreman-a2</td>
<td>cclws17.cse.nd.edu</td>
<td>1028</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;wq-foreman-b</td>
<td>cclws15.cse.nd.edu</td>
<td>1025</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>arab_cnvrtn</td>
<td>crcfe01.crc.nd.edu</td>
<td>9551</td>
<td>0</td>
<td>88</td>
<td>902</td>
<td>213</td>
</tr>
<tr>
<td>forcebalance</td>
<td>fire-4-0-ext.slac.stan</td>
<td>9793</td>
<td>1418</td>
<td>18</td>
<td>43187</td>
<td>18</td>
</tr>
<tr>
<td>alldat_T_7</td>
<td>fugu.biostat.wisc.edu</td>
<td>9140</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>dihedral</td>
<td>leeping.Stanford.EDU</td>
<td>7329</td>
<td>524</td>
<td>0</td>
<td>77343</td>
<td>856</td>
</tr>
<tr>
<td>quad_cnvrtn</td>
<td>newcell.crc.nd.edu</td>
<td>9552</td>
<td>481</td>
<td>153</td>
<td>1608</td>
<td>153</td>
</tr>
<tr>
<td>cclosdc18</td>
<td>workspace.crc.nd.edu</td>
<td>9907</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cclosdc1</td>
<td>workspace.crc.nd.edu</td>
<td>9908</td>
<td>0</td>
<td>0</td>
<td>354</td>
<td>6</td>
</tr>
</tbody>
</table>
Work Queue Multi-slot Workers
Work Queue Multi-slot Workers

BEFORE

Worker

$\$$_{}$

Single Task

Tuesday, October 15, 13
Work Queue Multi-slot Workers

BEFORE

Worker

Single Task

NOW

Multi-slot Worker

Multiple Tasks
Work Queue Multi-slot Workers

1. Worker can have multiple slots for running tasks depending on available resources
2. Simultaneously runs multiple tasks if they fit within available resources
3. Tasks must be specified with their resource requirements
Work Queue Multi-slot Workers

Master

Task A
Task needs:
2 cores
1 GB Memory
3 GB Disk

Worker 1
Worker has:
1 core

Worker 2
Worker has:
4 cores
8 GB Memory
300 GB Disk

Task A

Tuesday, October 15, 13
Work Queue Mutli-slot Workers

Master

- work_queue_task Specify_cores(q, 2);
- work_queue_task Specify_memory(q, 1024);
- work_queue_task Specify_disk(q, 3072);

Task A

- work_queue_task Specify_cores(q, 2);
- work_queue_task Specify_memory(q, 1024);
- work_queue_task Specify_disk(q, 3072);

Worker 1

- work_queue_worker --cores 1
- 1 core

Worker 2

- work_queue_worker --cores 4 --disk 300000 -mem...
- 4 cores
- 8 GB Memory
- 300 GB Disk

Worker has:

- 4 cores
- 8 GB Memory
- 300 GB Disk
Work Queue Multi-slot Workers
Work Queue Multi-slot Workers

- Master

```cpp
work_queue_task_specify_cores(q, 2);
```

- Worker

```cpp
work_queue_worker --cores 64
```

- Task 1
- Task 2
- Task ...
- Task n
Work Queue Status

```bash
work_queue_status -R
```

```
MASTER        CORES  MEMORY        DISK
arab_cnvrts   196    1815041       163063132
quad_cnvrts   167    3004155       67479902
forcebalance  ???    ???           ???
alldat_T_7    14     1830483       5321591
force_balance 14     1830483       5321591
ccloosdc18    0      0             0
ccloosdc1     6      14549         137146
```
WQ Using Project Names
WQ Using Project Names

“myproject”

Work Queue (port 9037)
WQ Using Project Names

“myproject”

Work Queue (port 9037)
advertise

Catalog
WQ Using Project Names

“myproject”

Work Queue (port 9037)

advertise

Catalog

“myproject” is at opteron:9037

Tuesday, October 15, 13
WQ Using Project Names

```
work_queue_worker
-N myproject
```

```
"myproject"
```

```
Work Queue (port 9037)
```

```
Worker
```

```
Catalog
```

```
"adver1se"
```

```
"myproject"
is at opteron:9037
```

Tuesday, October 15, 13
WQ Using Project Names

```
 Using Project Names

 "myproject"

 Work Queue (port 9037)

 advertise

 Catalog

 "myproject"

 is at opteron:9037

 Worker

 work_queue_worker
 -N myproject
```

Tuesday, October 15, 13
WQ Using Project Names

work_queue_worker
–N myproject

“myproject”

Work Queue (port 9037)

connect to opteron:9037

Worker

advertise

Catalog

query

“myproject” is at opteron:9037

Tuesday, October 15, 13
WQ Using Project Names

```
$ work_queue_worker -N myproject
```

```
$ connect to opteron:9037
```

```
$ work_queue_status
```

```
"myproject" is at opteron:9037
```

```
$ advertise
```

```
$ query
```

```
$ query
```

```
$ connect to opteron:9037
```

```
$ work_queue_status
```

```
"myproject"
```

```
is at opteron:9037
```

Tuesday, October 15, 13
Find master location in *work_queue_status*

<table>
<thead>
<tr>
<th>PROJECT</th>
<th>NAME</th>
<th>PORT</th>
<th>WAITING</th>
<th>BUSY</th>
<th>COMPLETE</th>
<th>WORKERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>awe-fip35</td>
<td>fahnd04.crc.nd.edu</td>
<td>1024</td>
<td>719</td>
<td>1882</td>
<td>1206967</td>
<td>1882</td>
</tr>
<tr>
<td>hfeng-gromacs-10ps</td>
<td>lclsstor01.crc.nd.edu</td>
<td>1024</td>
<td>4980</td>
<td>0</td>
<td>1280240</td>
<td>111</td>
</tr>
<tr>
<td>hfeng2-ala5</td>
<td>lclsstor01.crc.nd.edu</td>
<td>1025</td>
<td>2404</td>
<td>140</td>
<td>1234514</td>
<td>140</td>
</tr>
<tr>
<td><strong>forcebalance</strong></td>
<td><strong>leeping.Stanford.EDU</strong></td>
<td><strong>5817</strong></td>
<td>1082</td>
<td>26</td>
<td>822</td>
<td>26</td>
</tr>
<tr>
<td>forcebalance</td>
<td>leeping.Stanford.EDU</td>
<td>9230</td>
<td>0</td>
<td>3</td>
<td>147</td>
<td>3</td>
</tr>
<tr>
<td>fg-tutorial</td>
<td>login1.futuregrid.tacc</td>
<td>1024</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Tuesday, October 15, 13*
Work Queue Authentication

Master
My password is secret

Good Worker
I want to worker for Master and the password is secret

Rogue Worker
I want to worker for Master

Tuesday, October 15, 13
Work Queue Authentication

Master

My password is secret

Good Worker

I want to work for Master and the password is secret

Rogue Worker

I want to work for Master

Tuesday, October 15, 13
Work Queue Authentication

Master

```
work_queue_specify_password_file(q, file);
My password is secret
```

Good Worker

```
work_queue_worker --password file
I want to worker for Master
Master and the password is secret
```

Rogue Worker

```
I want to worker for Master
```

Tuesday, October 15, 13
Work Queue Partial Files

Master

Worker 1
1st chunk

Worker 2
2nd chunk

Worker 3
3rd chunk

1st chunk
2nd chunk
big file
3rd chunk
Work Queue Partial Files

**Master**

- **1st chunk**
  - `work_queue_task_specify_file_piece`
  - (... path_local, name_remote, start, end, ...);

- **2nd chunk**
  - big file

- **3rd chunk**

**Worker 1**

- **1st chunk**
  - `stat name_remote`

**Worker 2**

- **2nd chunk**

**Worker 3**

- **3rd chunk**

Tuesday, October 15, 13
Resource Monitor

Summary, with total maximum values

<table>
<thead>
<tr>
<th>Command</th>
<th>Limits over</th>
<th>Exit status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall time</td>
<td># processes</td>
<td>CPU time</td>
</tr>
<tr>
<td>Memory</td>
<td>Disk</td>
<td>IO</td>
</tr>
</tbody>
</table>

Time series, with total current values

| Wall time | # processes | CPU time |
| Memory | Disk | IO |

List of opened files

<table>
<thead>
<tr>
<th>Filename</th>
<th>Original size</th>
<th>Final size</th>
</tr>
</thead>
<tbody>
<tr>
<td># reads</td>
<td># writes</td>
<td></td>
</tr>
</tbody>
</table>
Resource Monitor and Makeflow

makeflow

makeflow --monitor dir

dir/

rule 01 {summary, series, files}

rule ... {summary, series, files}

rule n {summary, series, files}

Batch system

Resource Monitor

Task

Tuesday, October 15, 13
Resource Monitor and WQ

```
work_queue_enable_monitoring(q, "summaries.txt")
```
Resource Monitor Visualizer

dir/

- task 01 {summary, series}
- task ... {summary, series}
- task n {summary, series}

visualizer
Resource Monitor Visualizer

- dir/
  - task 01 {summary, series}
  - task ... {summary, series}
  - task n {summary, series}

resource_monitor_visualizer  logs  output  title
Makeflow Task Categories

Makeflow file

CATEGORY=preprocess
DISK=100
MEMORY=10
CORES=2
rule: ...

CATEGORY=analysis
...

Resource monitor as watchdog

Work Queue task requirements

Condor ClassAd requirements

Tuesday, October 15, 13
Makeflow Remote Renaming

Makeflow file

output_local->output_remote: input_local->input_remote cmd_local-> cmd_remote
cmd_remote --in input_remote --out output_remote

Local filesystem
- output_local
- input_local
- cmd_local

Remote filesystem
- output_remote
- input_remote
- cmd_remote
Makeflow Linker - Coming Soon

Makeflow file

Linker

makeflow
Chirp: Unix Authentication

Diagram:
- **Client**
- **Chirp Server**
- **Local Disk**
- **File**

Connections:
- **UNIX**
- **create**
- **stat**

Username: **unix:pdonnel3**

Date: Tuesday, October 15, 13
Chirp: Hostname/Address Authentication

hostname:ws01.nd.edu

Chirp Server

Local Disk

DNS

HOSTNAME

Client
Chirp: Kerberos Authentication

REQUIRES ROOT!

Host Kerberos Credential

sudo Chirp Server

Local Disk

Kerberos Auth Server

KERBEROS

Kerberos Ticket

Client

ws01.nd.edu

kerberos:PDONNEL3@ND.EDU
Chirp: Globus Authentication

Chirp Server

Local Disk

globus:/O=Cooperative_Computing_Lab/CN=Patrick_Donnelly

Globus Ticket

GLOBUS

Client

Host Certificate

ws01.nd.edu

chirp.nd.edu

Tuesday, October 15, 13
Chirp: Challenges

Problems to solve:

• Low value (low risk) credential to ship with jobs.

• Disposable & Time Limited.

• Works across multiple infrastructures.
Chirp: Ticket Authentication

1. Create Ticket

2. Register Ticket

3. Set ACL Limits

Since CCTools 3.4.0
Chirp: Using a Ticket

chirp -a ticket -i my.ticket chirp.nd.edu ls /data

or

parrot_run -a ticket -i my.ticket /bin/ls /chirp/chirp.nd.edu/data
Chirp/Parrot: File Extended Attributes

parrot_run getfattr -n user.instrument sensor.dat
Parrot and CVMFS

cvmfs repo 1  cvmfs repo ...  cvmfs repo n

parrot

athenahelloworld.sh
Parrot and CVMFS

parrot

parrot_run --cvmfs-repo-switching -r "*.cern.ch: ...

cvmfs repo 1  cvmfs repo ...  cvmfs repo n

athenahelloworld.sh
**Parrot Search System Call**

- `ls /dir/*.xyz`
- `stat(...)`
- `stat(...)`
- `stat(...)`
- ...
- `remote filesystem`

- `parrot_run`
- `parrot_search '/dir/*xyz'`
- `search(...)`
- `remote filesystem`
Development in Github

https://github.com/cooperative-computing-lab/cctools
git clone https://github.com:cooperative-computing-lab/cctools.git
Thanks to All of the Contributors!

Michal Albrecht
DeVonte Applewhite
Neil Best
Brian Bockelman
Dan Bradley
Peter Bui
Iheanyi Ekechukwu
Patrick Donelly
Brian Du Sell
Brenden Kokoszka
Kyle Mulholland
Francesco Prelz
Dinesh Rajan
Casey Robinson
Peter Sempolinski
Douglas Thain
Andrew Thrasher
Benjamin Tovar
Li Yu