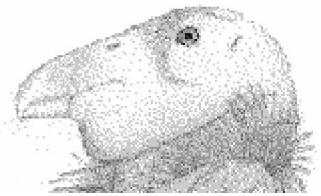
Introduction to Condor
4/19/02

C. Walter

http://www.cs.wisc.edu/condor



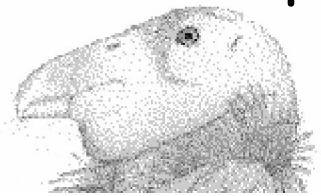
Condor
High Throughput Computing

What is Condor?

- Condor is a system designed for "High Throughput Computing".
- It manages large numbers of jobs for users on lots of machines automatically. Especially usefull for running lots of jobs over a long period of time.
- Developed and supported by University of Wisconsin (> 30 full time faculty and staff). Over 300 pools worldwide running on over 8500 CPUs.

Some advantages of Condor:

- Balances requests for people, groups, and machines dynamically.
- You can manage a very large set of similar jobs with just one file. Condor runs them all sends you email when done etc.
- If job is linked specially, jobs can restart after computer crashes etc. at the exact point they left off, also they can move among computers and even run on computers where you don't have an account.
- Remote Condor pools can be linked together.



Condor
High Throughput Computing

How to run a job (A simple submit file):

```
# condor job control file
Executable     = my_prog
universe       = vanilla
getenv         = true

log            = condor−out.log

Arguments    =  file1.dat

Output         = output.log
Error          = output.log

Queue 1
```

Use command: condor_submit condor_job

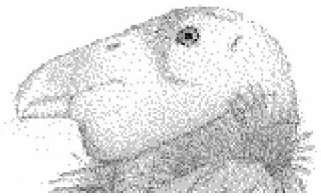
How to process many files in one job:

```
# condor job control file
Executable     = my_prog
universe       = vanilla
getenv         = true

log            = condor−out.log

Output         = output−01.log
Error          = output−01.log
Arguments      =  run−01.dat
Queue 1

Output         = output−02.log
Error          = output−02.log
Arguments      =  file−02.dat
Queue 1
```



Condor
High Throughput Computing

How to use macro variables:

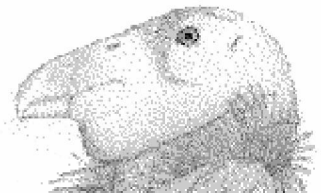
```
# condor job control file
Executable     = my_prog
universe       = vanilla
getenv         = true

log            = condor−out.log
Output         = output−$(fileNumber).log
Error          = output−$(fileNumber).log

Arguments    =  file−$(fileNumber).dat

fileNumber = 01
Queue 1

fileNumber = 02
Queue 1
```



Condor
High Throughput Computing

Sometime you can use $(Process):

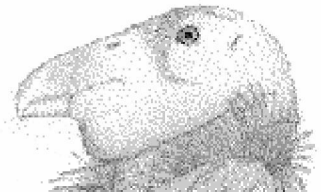
```
# condor job control file
Executable     = my_prog
universe       = vanilla
getenv         = true

log            = condor−out.log
Output         = output−$(Process).log
Error          = output−$(Process).log

Arguments    =  file−$(Process).dat

Queue 100
```

Please look at condor_submit manual for many more commands.



Condor
High Throughput Computing

Managing your jobs

- condor_submit (submit your job)
- condor_q (status of all your jobs)
- condor_q –r (status of running jobs [what machines])
- condor_status (What machines are available/busy)
- condor_rm (Remove a job)
- condor_history (Show the history of a previous job)

Clusters and Processes

```
> condor_q −r


−− Submitter: sukap02.suketto.icrr.u−tokyo.ac.jp : <10.10.10.102:62501> :
sukap02.suketto.icrr.u−tokyo.ac.jp
 ID      OWNER            SUBMITTED     RUN_TIME HOST(S)
 217.0   walter          4/18 19:18   0+00:00:11 sukap58
 217.1   walter          4/18 19:18   0+00:00:09 sukap94
 217.2   walter          4/18 19:18   0+00:00:07 sukap58
>
```

Looking at the history of the jobs

```
> condor_history 190
(ClusterId == 190)
 ID     OWNER          SUBMITTED    CPU_USAGE ST   COMPLETED CMD

 190.1   walter         4/14 17:56   0+00:56:08 C   4/14 18:52 /home/walter/k2
 190.0   walter         4/14 17:56   0+01:37:36 C   4/14 19:34 /home/walter/k2
 190.2   walter         4/14 17:56   0+03:45:45 C   4/14 21:43 /home/walter/k2
>
```

Our Condor Pool

- 200 CPUs (sukap00–99)
- Interactive processes/Condor on 00–09
- Condor Only on sukap10–99
  - atmpd priority on 40–59
  - lowe priority on 60–79
  - upmu priority on 80–99

Where to get more info

http://www-sk/~condor (local pool status)

For more info and documentation look at the main Condor Site:
http://www.cs.wisc.edu/condor

http://www.cs.wisc.edu/condor/quick-start.html

http://www.cs.wisc.edu/condor/manual/v6.2/
(users manual and man pages)

The Condor team has been very helpful to us. We have made a special agreement so that Fujitsu can access the source code and be treated as an academic group so they can support our system.

Condor
High Throughput Computing