The use of Condor at CCLRC Daresbury Laboratory

John Kewley

October 4, 2004

1 Introduction

In common with many large institutions and companies, CCLRC has considerable amounts of under-utilised computational resources (mostly personal desktops) along with a number of highly-used resources (typically servers or clusters).

It was felt that we should investigate using middleware such as Condor to combine these under-utilised workstations into a resource capable of performing some of our scientist's computations, freeing up their major computational resources for those jobs for which it would be impractical or impossible to run under Condor.

Additionally, within the the CCLRC e-Science Centre, we have a responsibility to obtain expertise in the all the major technologies to facilitate our role within the Grid Support Centre. Cultivating some expertise in Condor was therefore considered essential.

Unlike many large institutions, we do not have large quantities of training room or public access workstations and have therefore had to grow our pool slowly whilst we gained the trust of the resource owners. In the meantime, our understanding of Condor and its capabilities has grown and we are now encouraging our scientists to try out the Condor pool for themselves and see how they can utilise Condor to exploit these untapped resources.

2 Status

Internal Pool

- 24 processors on 19 machines:
  - 3x Windows XP Professional
  - 2x Windows 2000 Professional
  - 1x Windows NT 4.0 Workstation
  - 6x SuSE Linux 9.0
  - 3x Red Hat Linux 9
  - 5x White Box Enterprise Linux 3.0
  - 2x SuSE Linux 8.0
  - 1x Mandrake Linux 10.0
  - 1x Gentoo Linux 1.4
- New users will be asked to contribute at least one their own machines to the pool.
- One of the machines is setup as dual-boot Condor so that regardless whether it is booted as Linux or Windows XP, it will be available for Condor jobs. The inspiration for this is for training labs which may switch operating system from one week to the next.

- Currently supporting initial users from:

  CCP1 : "The Electronic Structure of Molecules" using GAMESS-UK (for further information, see `http://cse-www.dl.ac.uk/qcg/gamess-uk/`, [GvLK+96] and [Gv98])

  CCP4 : "Protein Crystallography", e-HTPX project (for further information, see `http://www.e-science.clrc.ac.uk/web/projects/e-HTPX/`, [AKM+04] and [KMWW04])

External Pool

- 6 processors on 5 machines:
  - 2x Red Hat Linux 7.3
  - 4x White Box Enterprise Linux 3.0
- The head node is currently setup to allow secure flocking to/from Glasgow, Cambridge, Belfast and the OMII in Southampton for the ETF remote build and test testbed. `http://polaris.ecs.soton.ac.uk:65000/`.

3 Future

- The Internal Pool will grow as more users add their own machines, and as workstation owners gain more confidence in the security of their machines and the use of the pool.
- We anticipate adding some training room machines to the pool if and when they are available.
- We hope to add a Mac to the External Pool for ETF/OMII testing/building.
- Once the DL pool is showing significant use, a pool will be set up at RAL. Our intention is to allow flocking between these pools.

References

[AKM+04] Rob Allan, Ronan Keegan, David Meredith, Martin Winn, Graeme Winter, Jonathon Diprose, Chris Mayo, Ludovic Launer, Joel Fillon, and Paul Young. HPC and Grid Applications in High Throughput Protein Crystallography. In Simon J. Cox, editor, *Proceedings of the UK e-Science All Hands Meeting 2004*, pages 187–93. EPSRC, September 2004.

[Gv98] Martin F. Guest and Huub J. J. van Damm. Algorithms, Developments and Applications in Molecular Modelling: The GAMESS-UK *Ab Initio* Code. In M.R. Johnson, G.J. Kearley, and H.G. Büttner, editors, *Neutrons and numerical methods - N2M,*, pages 9–18. American Institute of Physics (AIP), 1998.

[GvLK+96] Martin F. Guest, J.H. van Lenthe, J. Kendrick, K. Schoeffel, and Paul Sherwood. GAMESS-UK: User's Guide and Reference Manual. Technical report, Daresbury Laboratory, 1996.

[KMWW04] Ronan Keegan, David Meredith, Graeme Winter, and Martin Winn. HPC and Grid Applications in High Throughput Protein Crystallography. In Simon J. Cox, editor, *Proceedings of the UK e-Science All Hands Meeting 2004*, pages 1034–38. EPSRC, September 2004.