Linux at the HPCf

Humberto Ortiz Zuazaga
University of Puerto Rico
High Performance Computing facility
humberto@hpcf.upr.edu

What is the HPCf?

- A shared resource for the 11 UPR campuses, providing technical support for scientific computing.

- Through the PR-BRIN program also support 7 private institutions with regard to biomedical research.

- Our services include numerical supercomputing, high bandwidth Internet2 connectivity, scientific visualization, consulting, videoconferencing, training.

Linux at the HPCf

- All Internet services (HTTP, FTP, SMTP, POP, IMAP, NTP, SNMP)
- Software development (GCC, python, perl, lisp)
- Scientific computing (BLAST, MPI, EMBOSS)
- Databases (Oracle, MySQL, Postgress)
- Experimental networking protocols (jabber, Access Grid, bittorrent)

Access Grid (15 Mbps video conference)

Why linux?

- Cost, Cost, Cost (and licensing)
- Complete development environment (compilers, debuggers, IDE, documentation tools)
- Pace of development (Opteron, Itanium, IPV6)
- Flexibility (mix and match kernels, windowing systems, networking stacks, modular devices)

Reason: cost, licencing

- HPCf mission: scientific computing, Internet services are unfunded.
- The BiRC is building a linux cluster, since we use linux, every cent was spent on hardware
- Since we don't need to license linux, we don't have to worry about BSA audits.
- We don't need to wait for licenses before deploying new services.

Reason: pace of development

- Internet operating system, developed in Internet time.
- Tracking down IDE bug in 1992, I tested 2 kernels a day for over a week.
- Linux booted on the Itanium on a simulator in Intel labs, before the chips were built!

Reason: flexibility

- Fixing CDRW bug, recompiled and replaced cdrom drivers in a running linux system (without rebooting)!
- On this laptop I can choose between 20 window managers and 3 desktop environments.
- Two different IPV6 implementations.
- Several schedulers, memory allocators, device drivers.

Why not linux?

- Linux doesn't run our core routers, limitations on line speeds that standard PC hardware can handle. Linux runs our NAT router, and linux routers are available that can handle multiple gigabit Ethernet and OC-12 lines.

- Linux doesn't run our supercomputers (Origin 2000, Origin 300). The SGI Altix is an Itanium2 based supercomputer that runs linux. Linux clusters have become a major player in the TOP500 list of fastest supercomputers in the world.

- We are awaiting delivery of a linux cluster of 170 2.4 GHz Xeon processors on multiple gigabit Ethernet backplane, this will become the fastest machine at the HPCf.

Caveats and Limitations

- The HPCf has extensive linux experience in-house, and recruits and trains linux hackers.

- We use Microsoft Office and Adobe products (on MacOS X) for office and graphical design tasks. Using linux for these tasks is possible, but a bit painful for novices or casual users.

- Development of drivers for new hardware depends on the availability of code/documentation from the manufacturer. In many cases, this is not available (winmodems, 3D graphics hardware, laptops).