Run-Time

Susan Coghlan, Ron Minnich, Pete Beckman

Burning OS Issues

- **Stop** asking if Linus will accept the changes!! Just Stop it!
  - Understand Open Source
- **Stop** talking about the perceived HUGE human overhead of reintegrating changes
  - Who keeps making these claims?
- **Stop** talking about how the next commodity thing the industry pre-announces will solve everything.
  - How long is your memory?

More Burning OS Issues

- Lots of low-level msging layers exist, because it is still real real hard
  - Scalability cannot be an afterthought!
  - (Portals, GM, FM, AM, new-NSCA-thing, etc)
- Reliability needs more thought
- Linux is not just for 32-node babywulfs: Cplant and HPTi have shown large working systems

YMBOSI

- Linux for TeraScale machines will need kernel mods…
  - I'm not sure we even know what those are!
  - What happens when 6000 asynchronous interrupt-driven nodes want to act like they are tightly coupled?
- The OS will need to be more fault tolerant, and borrow technology from the high-availability marketplace.
- No, we don't need nano-kernels (cougar)

YMBOSI++

- The real burning issue: Will success hurt Linux?
- Benchmarks can hurt Linux
  - Linux, meet the real world of ZDnet and “independent” benchmarking
- USA Today: Linux runs 31% of web servers. Supercomputing is on the same slope, if not faster

MPI-2

- They lost the big Mo
- Standards are great, and Apps folks will port to the standard when they see it provides a TANGIBLE benefit, and is available EVERYWHERE
- To be honest, Open Source is king. Until Argonne releases full MPI-2, for every platform, people will continue to say huh?
- Will one-sided MPI-2 operations be popular?

YM MPI-2

- Does VIA, ST, Infiniband change MPI put/get semantics?
  - We don't even understand send/recv on smart nics…
  - What about MPI put/get on smart nics?
    - CRC where?
    - Reliability where?
    - Buffer management where?
    - Connection Setup?
    - Security?
- A word about "commodity networking"….

Dynamic Everything

- If “Da Grid” really becomes something other than RSH and Qstat….
- App folks may actually write dynamic stuff. Some already are (Cumulvs, PAWS)
- Process creation, client/server, externally coupled parallel apps…
  - Kiss that nano-kernel goodbye
  - Embrace full Linux (with lots o mods and fixes)
  - Embrace TCP/IP everywhere

Debugging is still a challenge

- Use printf()
  - An MPI company: “printf is not part of the MPI standard”
- Use TotalView
  - As much fun as a Voodoo curse
  - Does not scale (painful for 3 nodes)
  - Always a tool of last resort
  - GUIs don’t scale, and they are uniformly hated
  - Don’t improve TotalView. Start over in Open Source
- Future Research?
  - Anomaly detection
  - Advanced rule-based debugging

Did I mention....

Extreme Linux Conference #3

www.extremelinux.org

www.usenix.org

Call for papers

USENIX ALS Conference in October 2000

Participate now!