EPICS, Linux & NUMA I/O

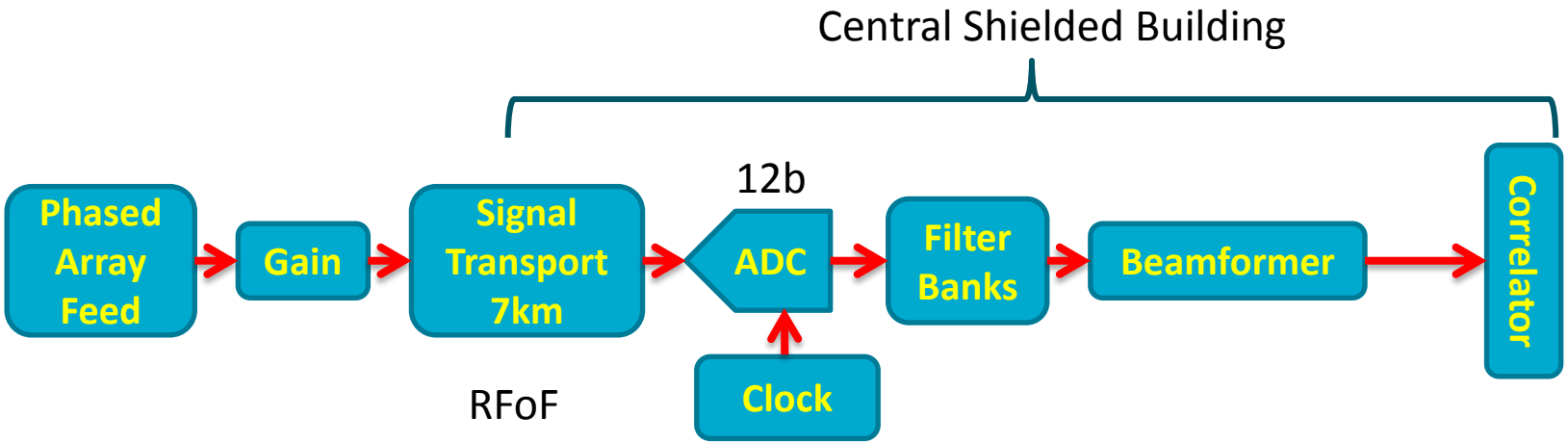
An example

Euan Troup

29 October, 2015

CSIRO ASTRONOMY AND SPACE SCIENCE
www.csiro.au

What are we doing?



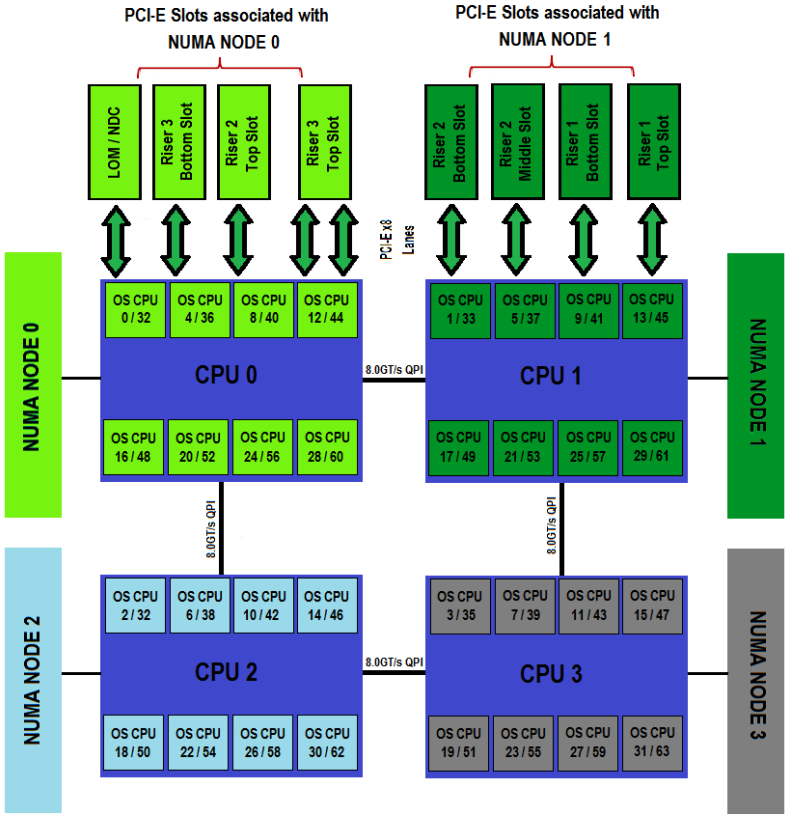
- system up to correlator custom data transmission
  - Correlator is 8x12-card system UDP to control/monitor computer
  - 2628x36x216 32-bit complex numbers each 5s 12x1GE staggered on 10GE
  - Reorder into frequency/product order & upload via 10GE staggered
  - 2 x dual-port 10GE on server

What are we doing with EPICS?

- Soft IOCs controlling custom hardware & firmware
  - Asyn port driver derived objects using ADE common library
  - EPICS PVs provide parameters for common library methods (RPC)
    - disconnect/connect
    - startup/shutdown
    - set frequency mapping
    - specialised hardware setups
    - …
  - Data download triggered via asyn port driver event driver
  - Upload to ingest using UDP and ASKAP/asyn port driver

What are we doing on Linux?

- Dell R820 4 8-core CPUs 128GB 2 x 2-port 10GE
- Debian wheezy minimal system



What happens when it's all combined?

- Data streaming can't be sustained
  - Only 4-6 cards worth of data can be processed reliably
    - Large network buffers used
    - Network card driver parameters tuned
  - As above diagram shows, compute load can be distributed symmetrically
  - NUMA system is not symmetric
    - Use numactl to schedule IOC on 1 node – no difference
  - I/O load is not symmetric
    - default scheduler setup is all interrupt service on CPU 0
    - irqbalance also doesn't help

What can be done to fix the problem?

- Don’t use irqbalance
  - Tie irqs to appropriate NUMA node
  - Spread load across cores in the node
    - cat /proc/irq/[nnn]/node
    - echo [cpu mask] > /proc/irq[nnn]/smp_affinity
  - irq numbers may change across reboots/device restart (eg ifup/ifdown)
  - Tie EPICS app to the same NUMA node
    - Use sched_setaffinity() in iocMain
    - All threads will start on the selected cores of the node

References

- http://events.linuxfoundation.org/sites/events/files/eeus13_shelton.pdf
- https://queue.acm.org/detail.cfm?id=2513149

Thank you

CSIRO Astronomy & Space onomy
Euan Troup

t +61 2 9372 4660
E Euan.Troup@csiro.au
w www.csiro.au/en/Research/Astr