Best Practices and Lessons Learned Developing Linux Products

Old: OS noise and HPC Application

Christoph Lameter, Ph.D.
Technical Advisory Board
Linux Foundation

Overview

Practices: How does Open Source Work

What is OS noise?

Disturbances of a single thread

Disturbances through SLAB queue processing.

Increased latencies through kernel bloat.

Some techniques to limit noise

Best Practices

- Open Exchange of Information
- Review of source code (in many ways can replace Q and A)
- End user able to influence, fix or implement what is going wrong with software.
- Different paradigm from proprietary packaged source code.

Lessons

- Effective open source projects depend on having an interested audience.
- Organization using websites, collaboration sites, mailing lists and IRC channels.
- Biggest problem is usually communication.
  - Miscommunication.
  - People get offended.
  - Interaction only via Internet.
  - Need to listen and understand one another.

Maintain influence in Open Source project

- Typically organization gets involved in a project with a certain agenda.
- Organization may withdraw after the issue has been fixed.
- Open source project continues to develop and fulfill the agenda of other participants.
- Feature wanted may disappear.

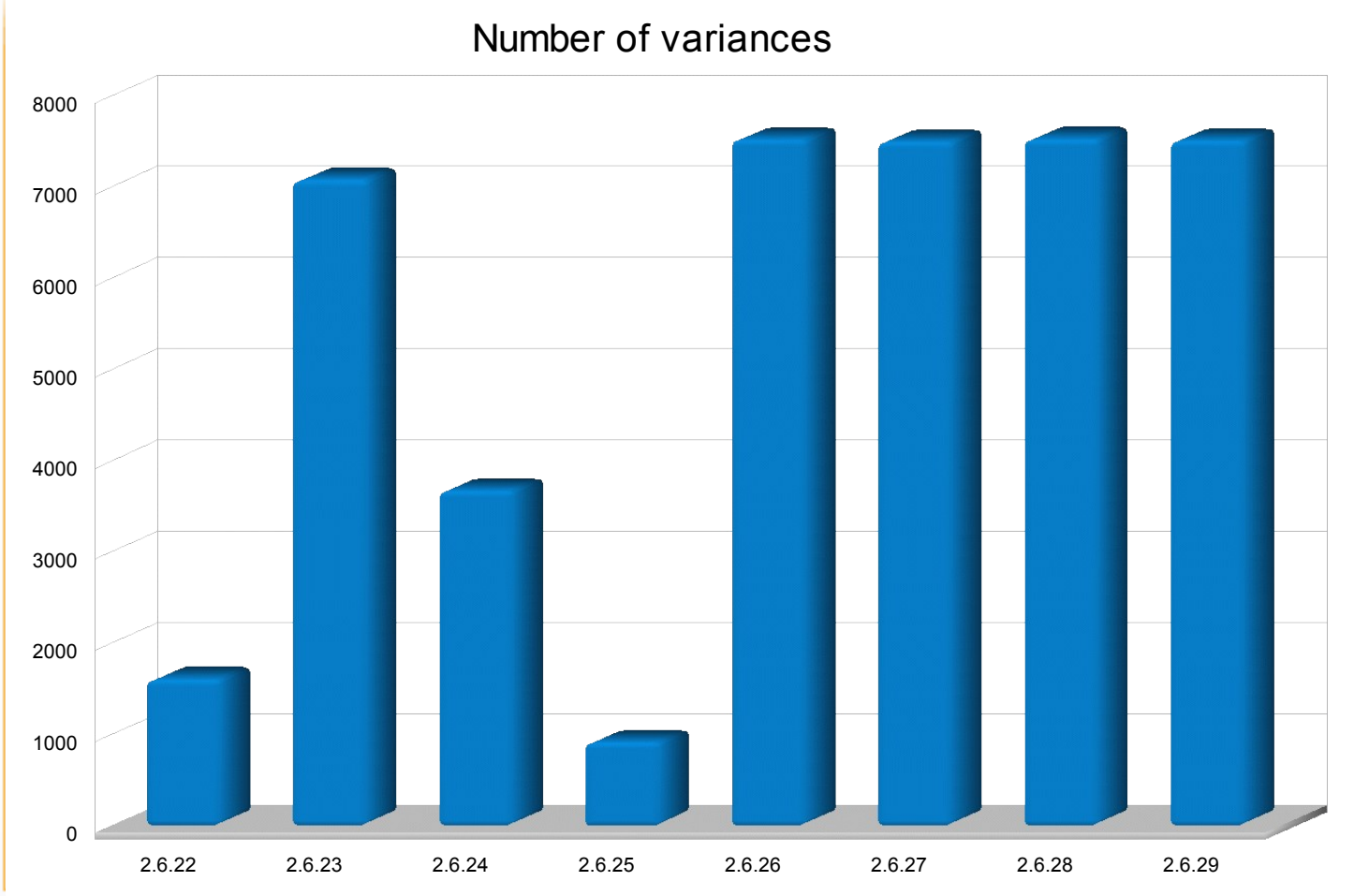
Kernel Latency Regressions

- OS use for low latency uses
- Must use old kernel since newer kernel add bloat and increase latencies.
- HPC, Gaming, financial industry is affected by this in particular.
- Cut off from newer kernel features.

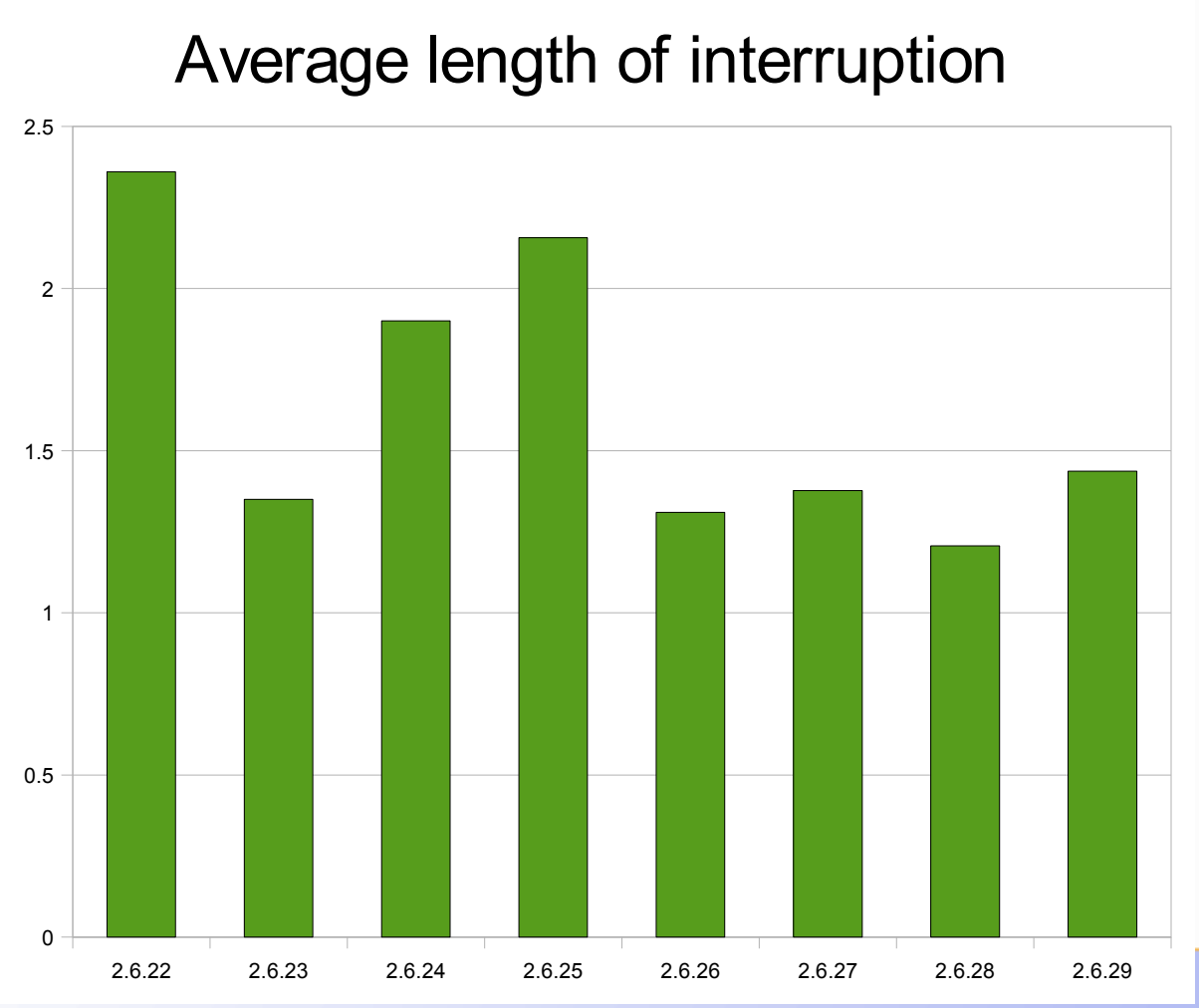
OS Noise

- Application experiences random delays due to Operating system actions.
- On the same cpu that the application is running the OS may
  - Schedule OS threads
  - Hardware interrupts
  - Run timers
- Disturbances increases with higher scheduling frequency.
- Lower scheduling frequency makes the delays longer.

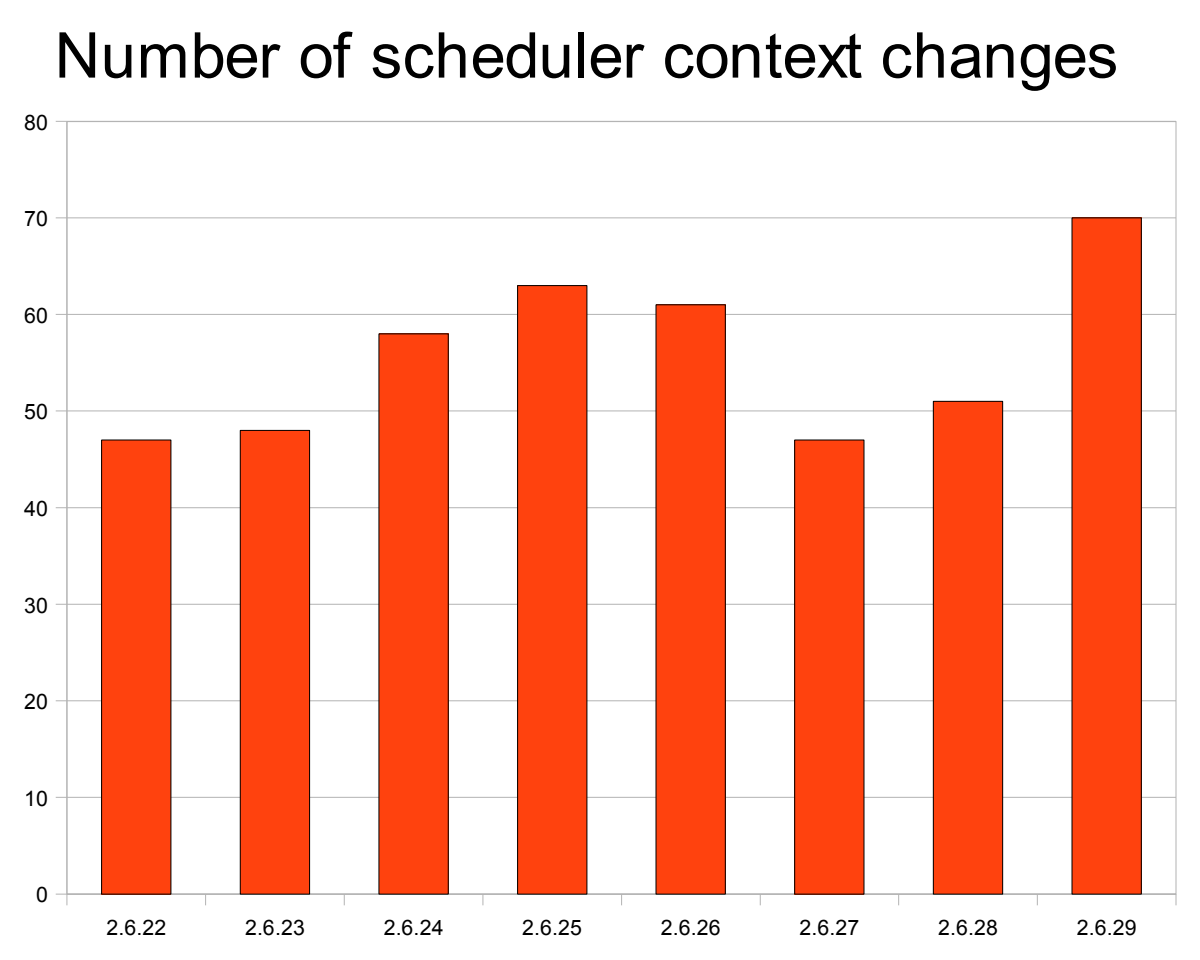
Noise created by the Linux OS



Length of Noise periods (microseconds)



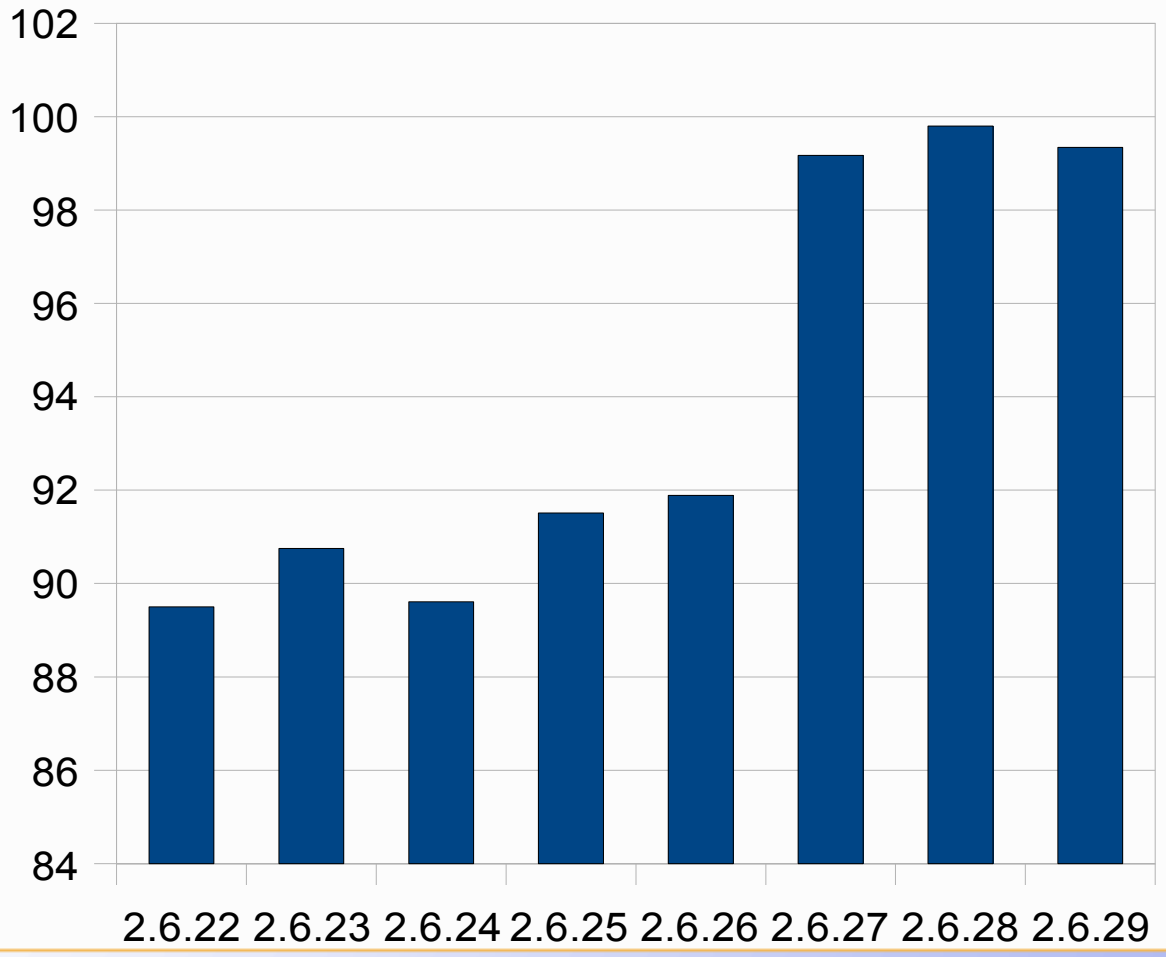
Scheduler interventions



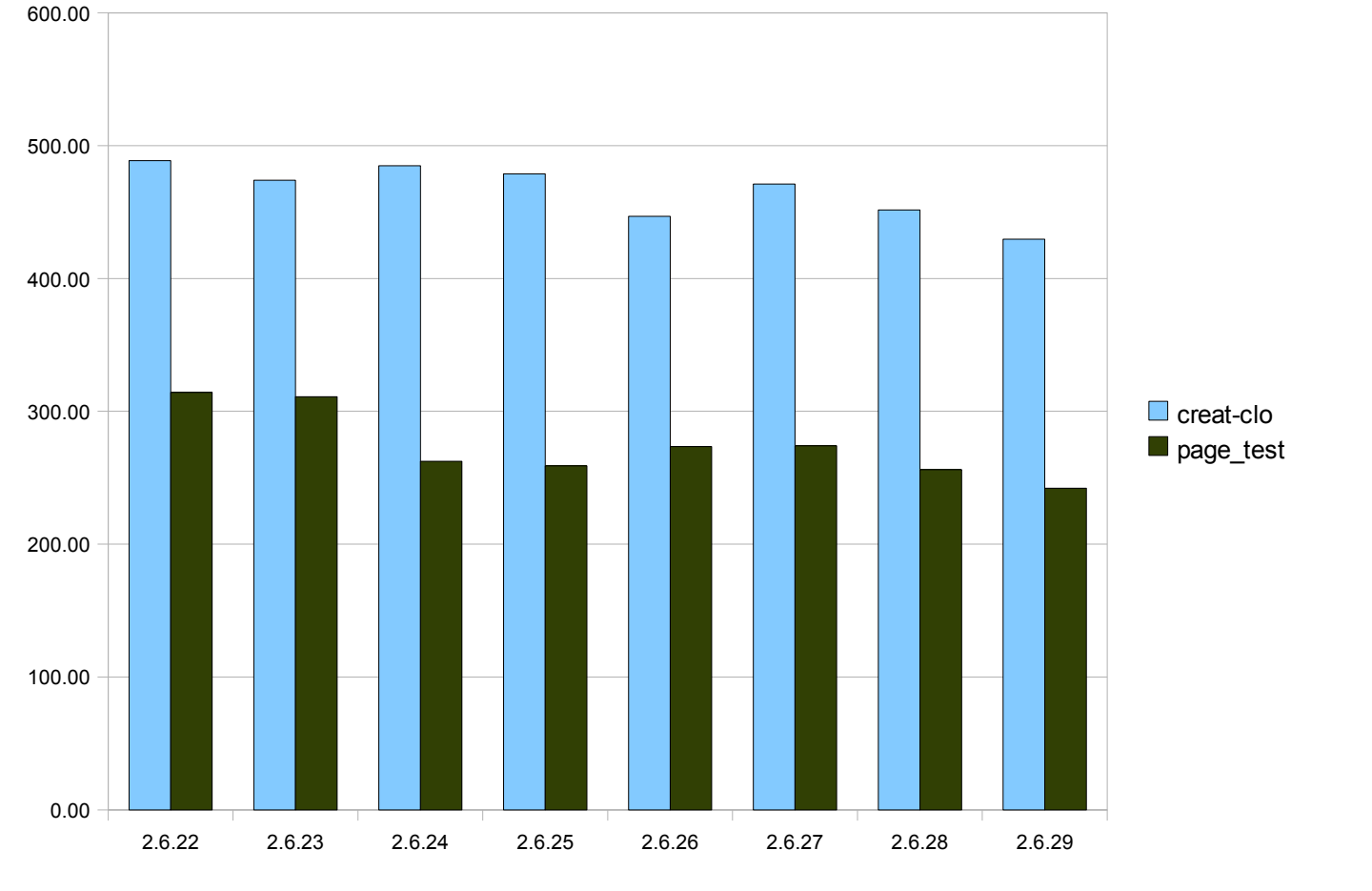
Low Latency tools (gentwo.org/ll)

- **latencytest:** An OS noise measurement tool
  - Number of OS reschedules
  - Number of Faults
  - Holdoffs and their frequency
- **udpping:** Measure minimum communication latencies.
  - Histogram of UDP ping pong traffic
  - Serialized or streaming modes

UDP ping pong times (microseconds)



AIM9 regressions



Measures to reduce OS noice

Process pinning: taskset

Realtime priorities: chrt

Prefaulting pages

Cache prepopulation

OS features off

Smaller cache footprint

OS should not defer processing.

Things to do

Establish better tools to measure OS noise.

Feedback to OS developers re OS noise

Establish latencies for critical OS paths and benchmark newly released kernels.