TECHNISCHE UNIVERSITÄT DRESDEN

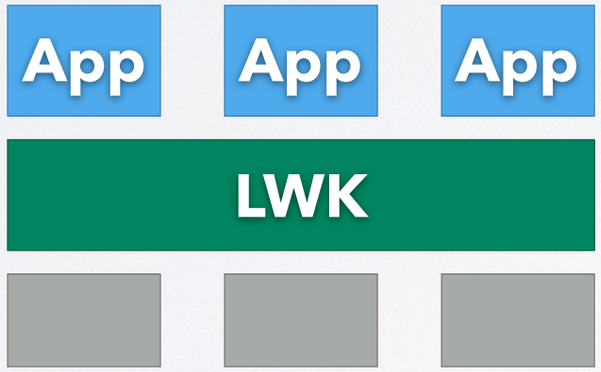
DECOUPLED: LOW-EFFORT NOISE-FREE EXECUTION ON COMMODITY SYSTEMS

ADAM LACKORZYNSKI, CARSTEN WEINHOLD, HERMANN HÄRTIG
TU DRESDEN, GERMANY

PROBLEM

- Execution-time jitter / OS noise
- Bulk-synchronous programming codes

APPROACHES



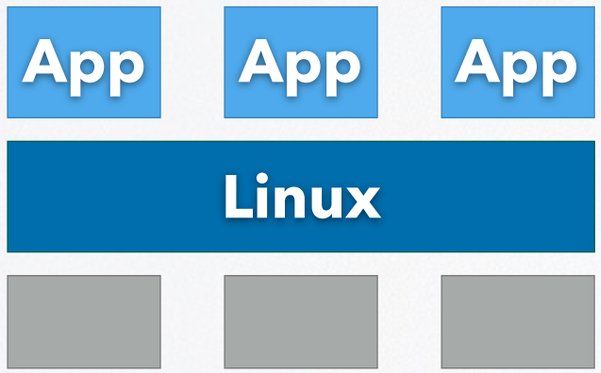
LWK

⊕ No Noise

⊖ Compatibility

⊖ Features

CNK



Linux (tweaked)

⊙ Low Noise

⊕ Compatibility

⊕ Features

⊖ Fast moving target

Cray

Argo

APPROACHES

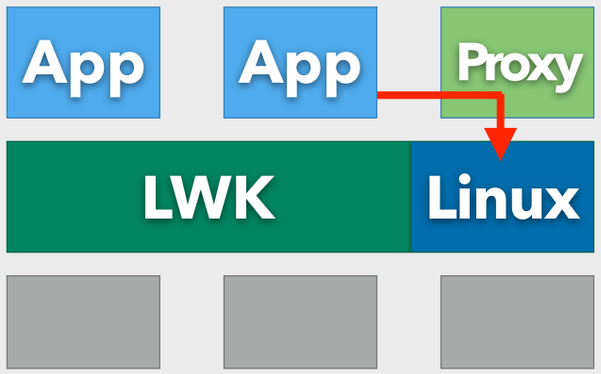
LWK

⊕ No Noise

⊖ Compatibility

⊖ Features

CNK



LWK + Linux

⊕ No Noise

⊕ Compatibility

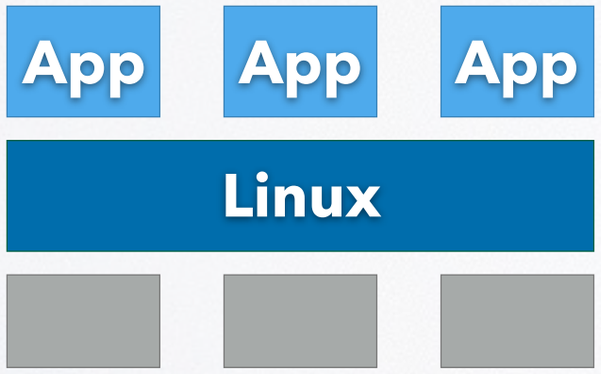
⊕ Features

⊖ **Much effort**

mOS

McKernel

Hobbes/Kitten



Linux (tweaked)

⊙ Low Noise

⊕ Compatibility

⊕ Features

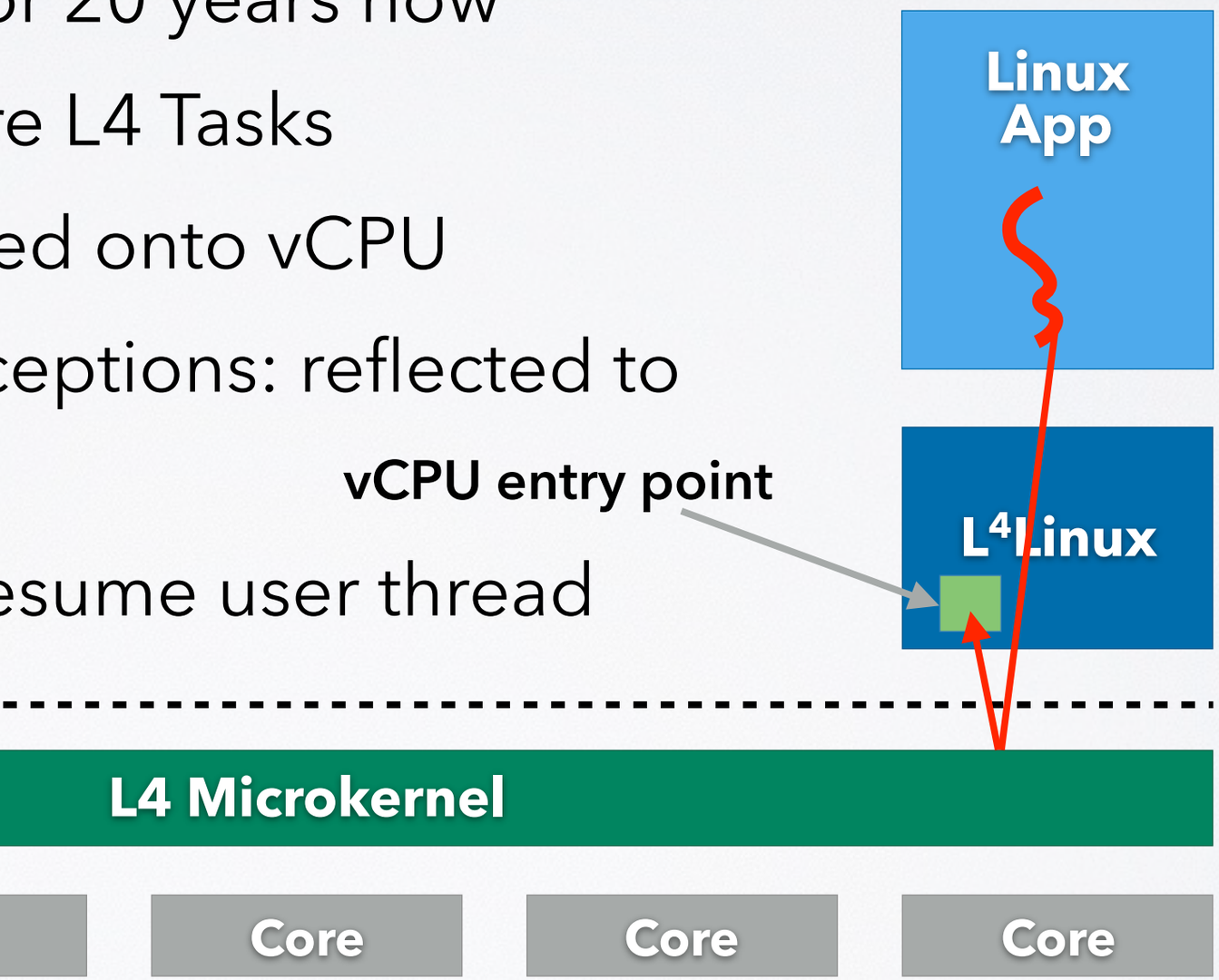
⊖ Fast moving target

Cray

Argo

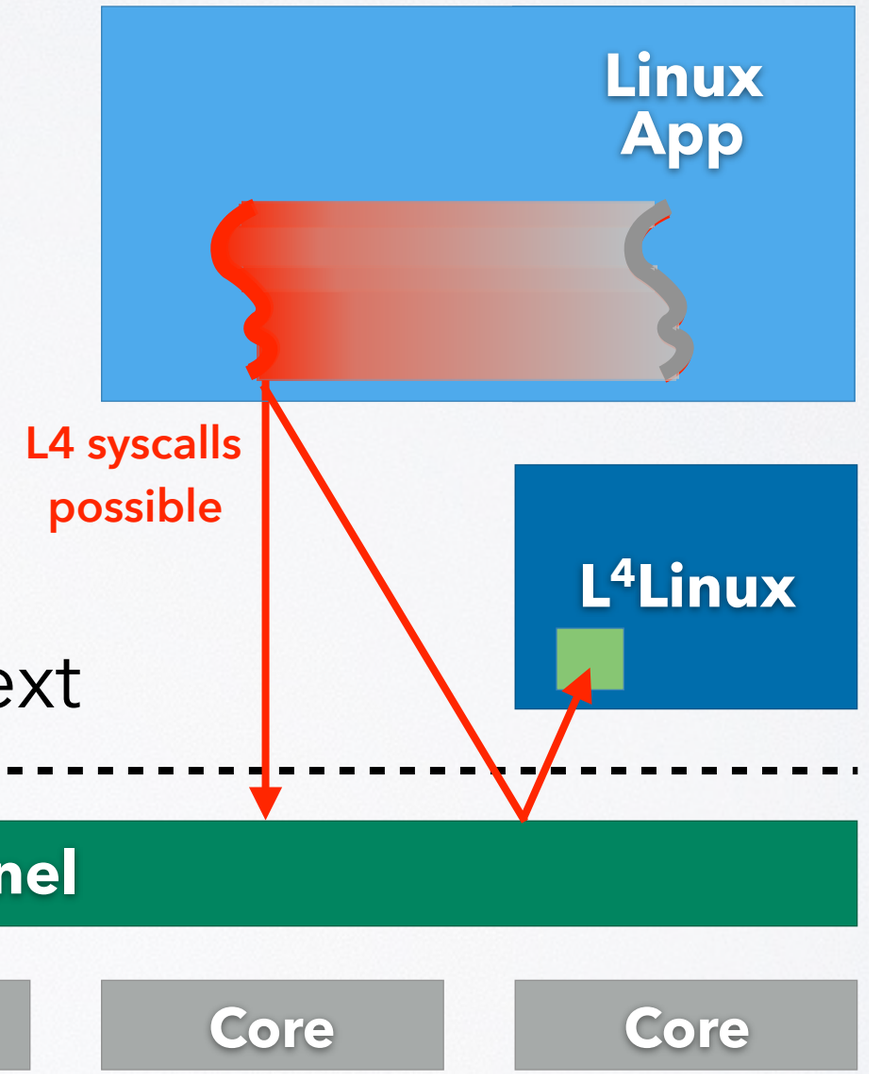
L4 + L⁴LINUX

- Paravirtualized L⁴Linux: **`arch/l4`**
- Well maintained for 20 years now
- Linux processes are L4 Tasks
- Threads multiplexed onto vCPU
- Linux syscalls / exceptions: reflected to vCPU entry point
- Handle syscall + resume user thread



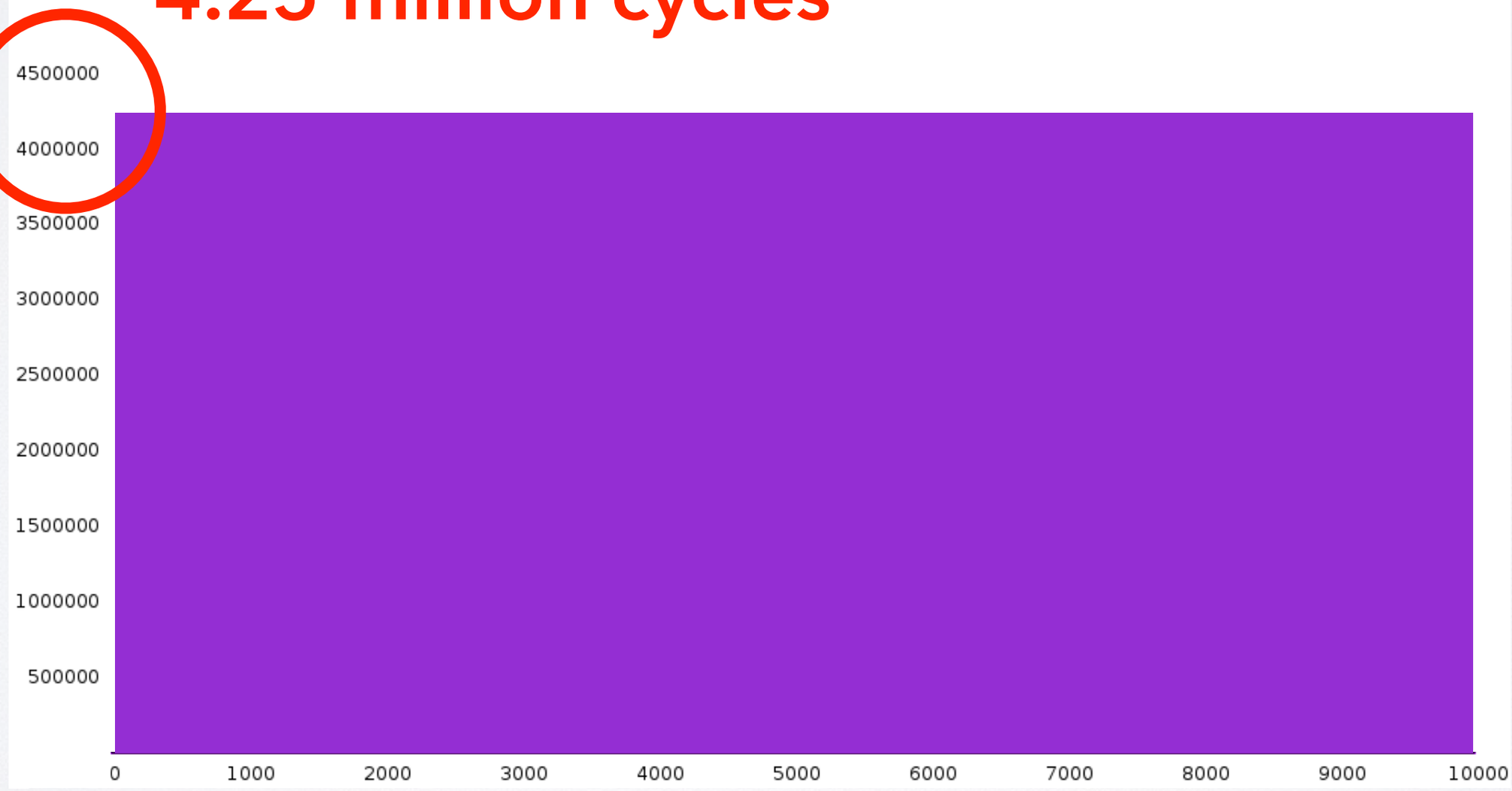
DECOUPLED EXECUTION

- Decoupling:
  - Create new L4 thread on dedicated core
  - Mark Linux thread context uninterruptible
- Linux syscall:
  - Forward to vCPU entry point
  - Reactivate Linux thread context

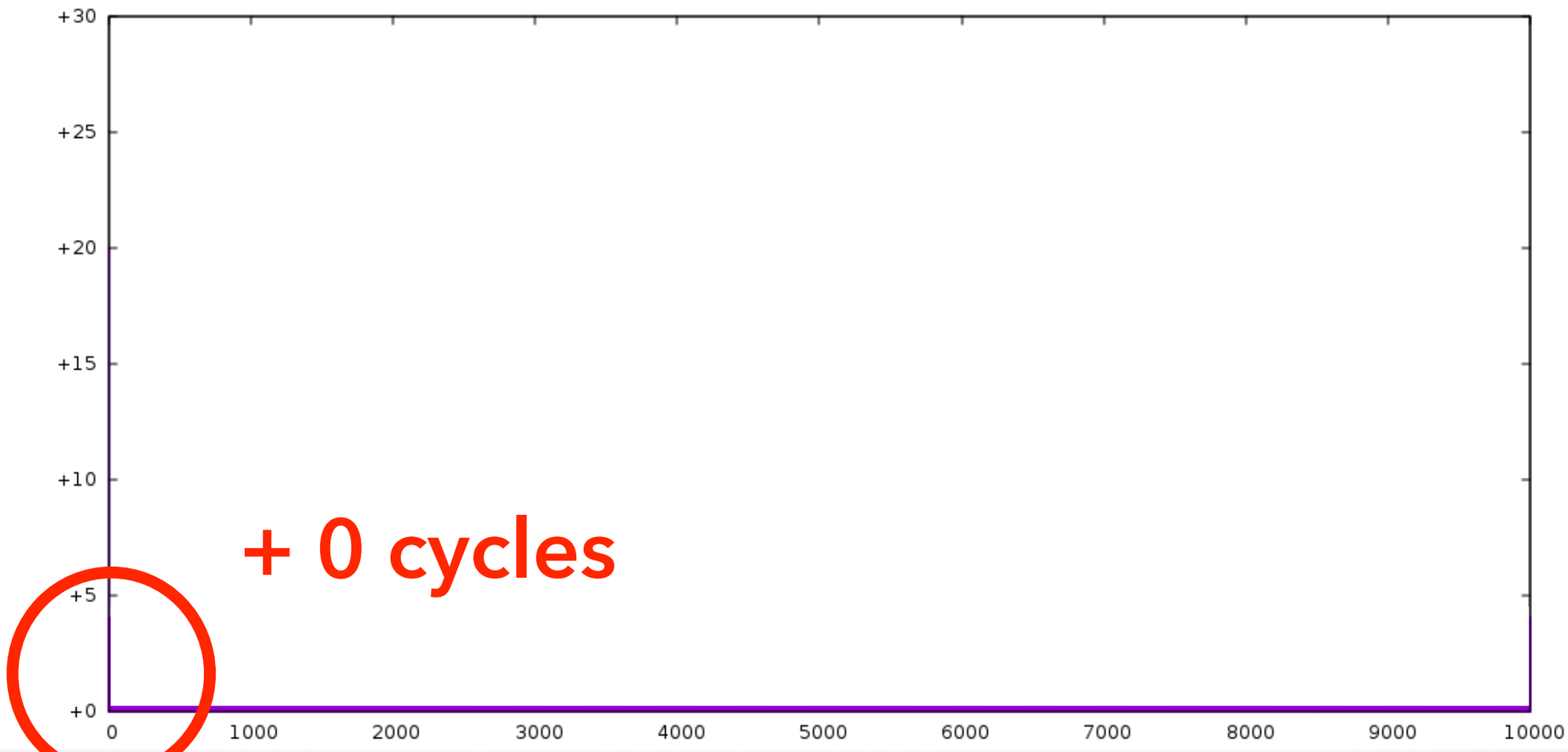


FWQ BENCHMARK

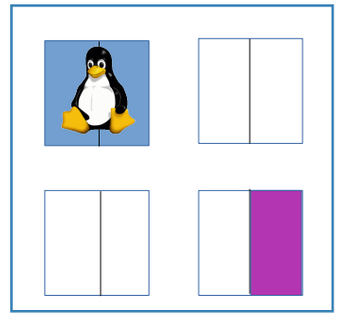
4.25 million cycles



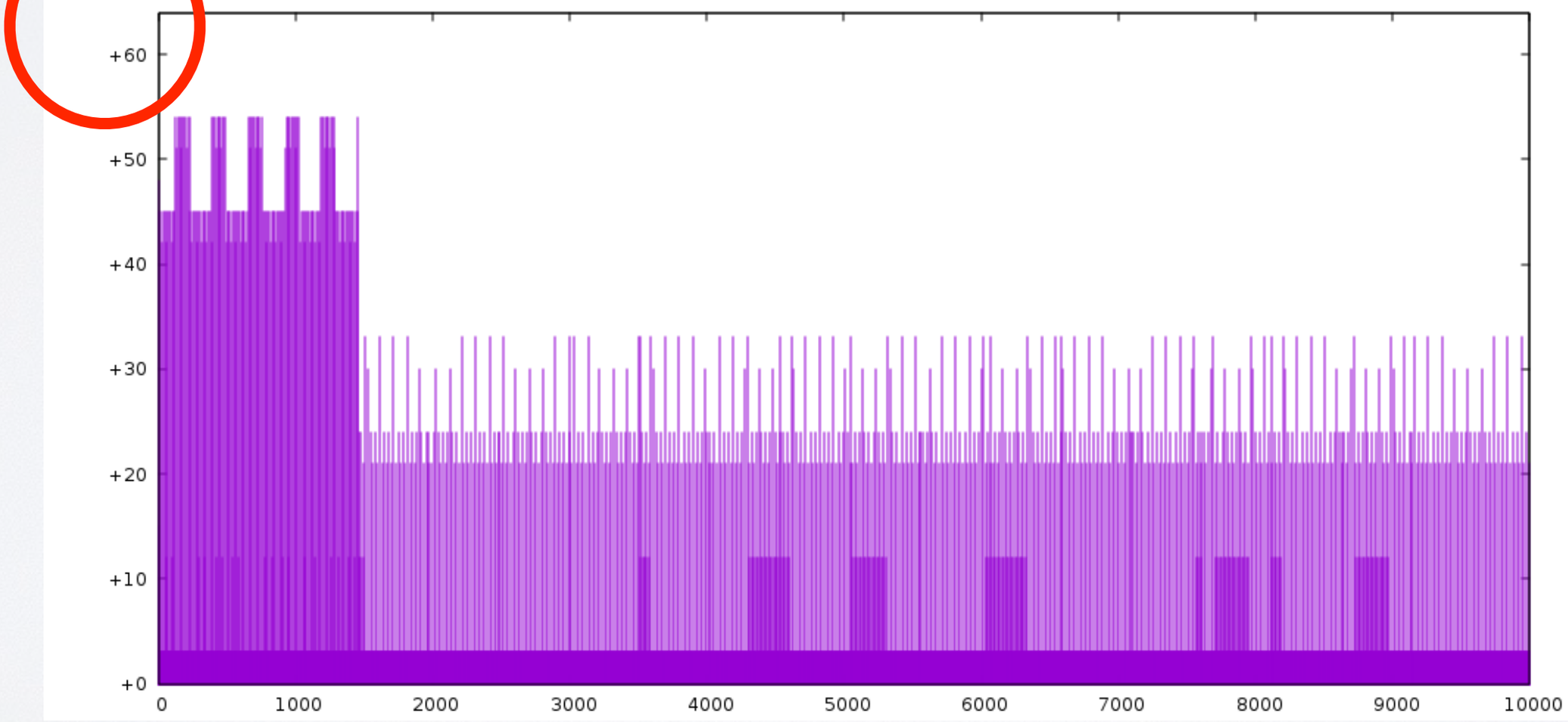
Ideal: zero extra cycles



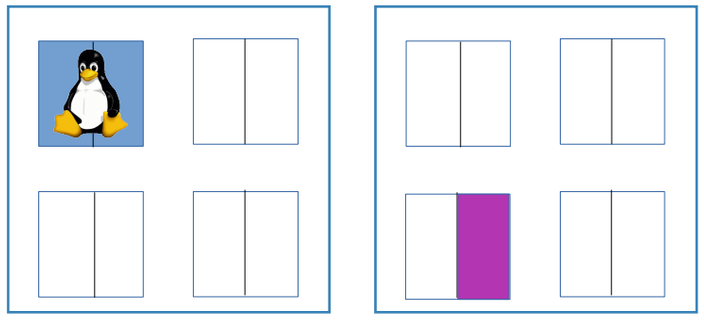
Decoupled Linux thread

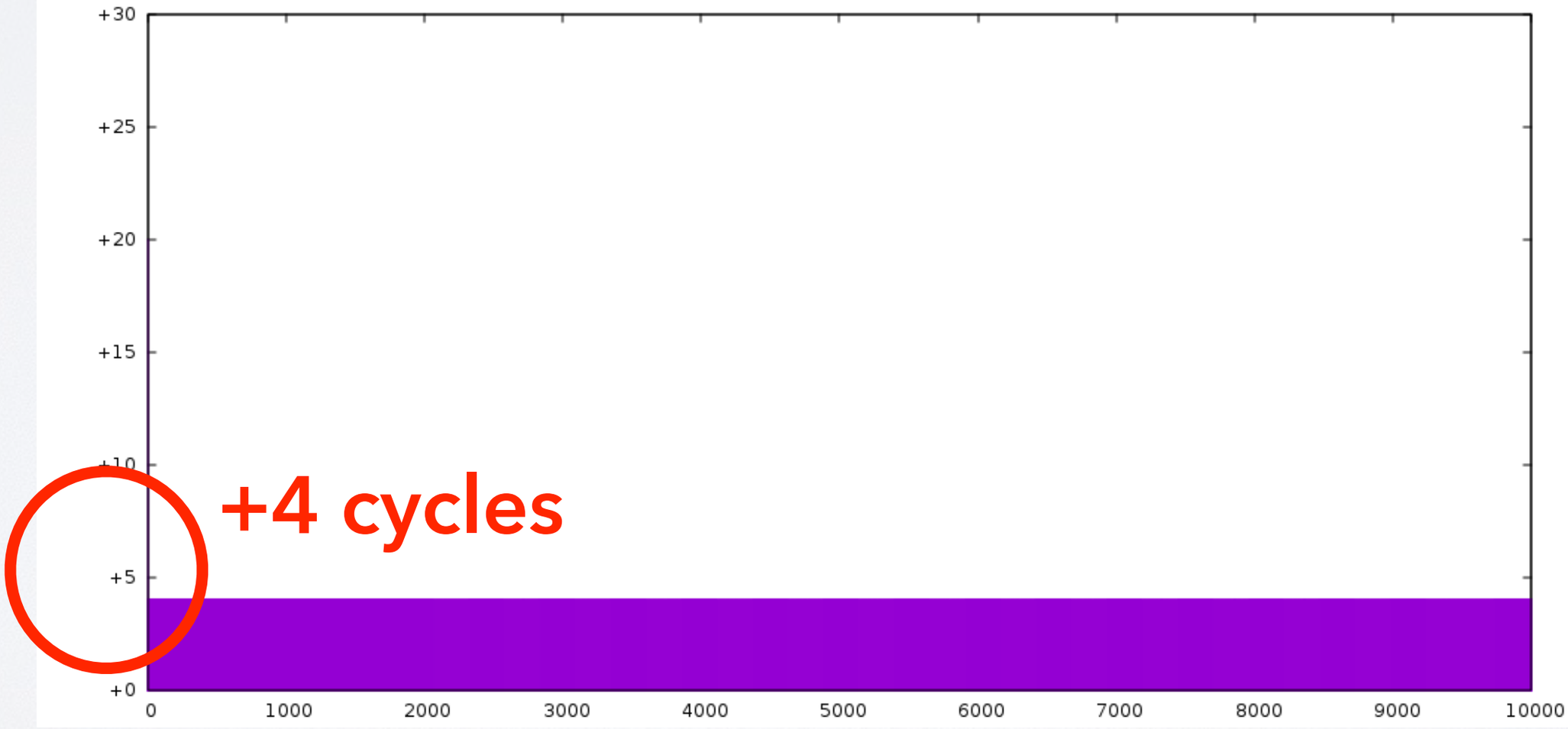


+60 cycles



Decoupled Linux thread





StartSync:

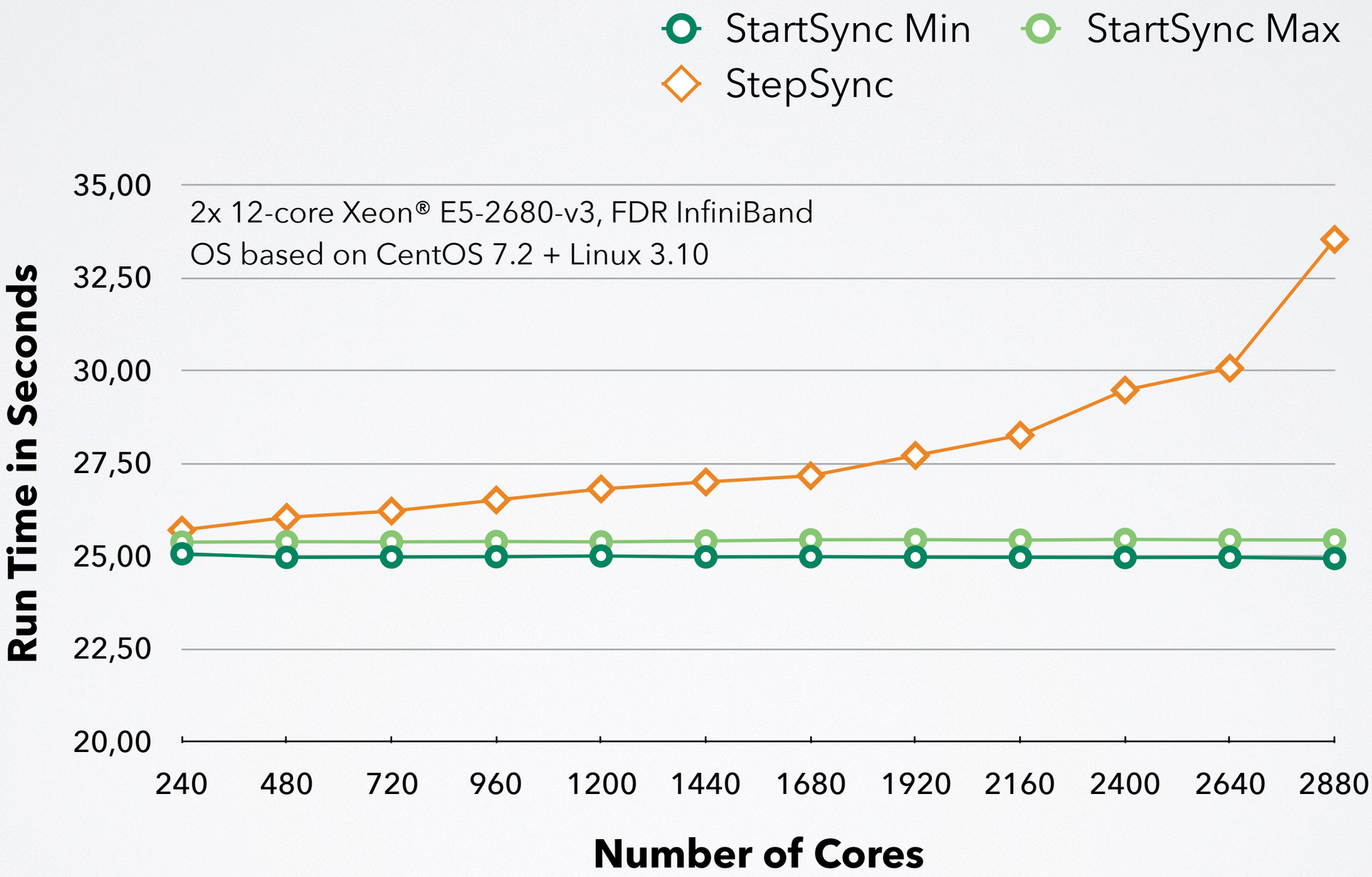
Behavior: **embarrassingly parallel**

StepSync:

Behavior: **bulk-synchronous**

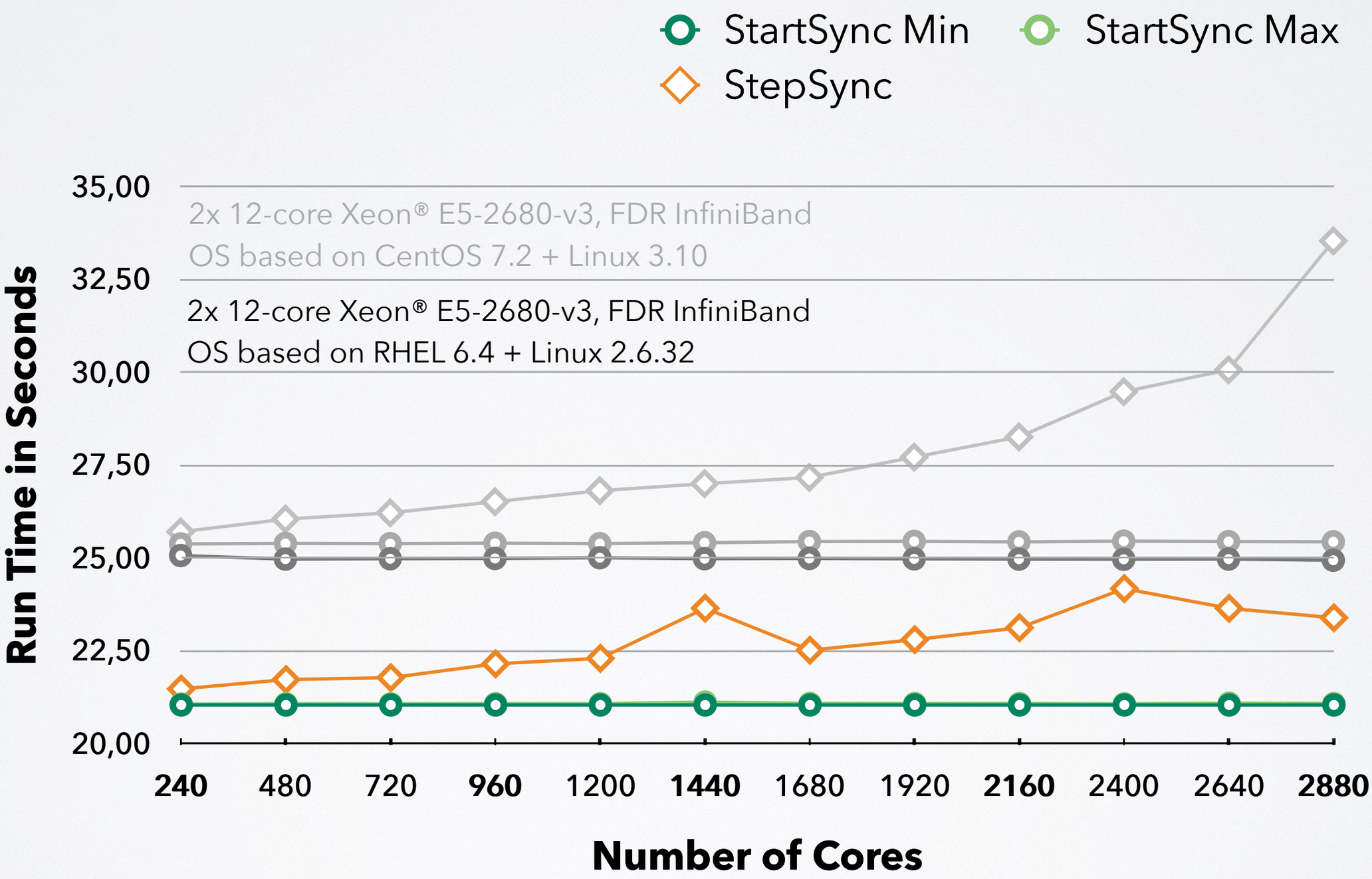
TECHNISCHE UNIVERSITÄT DRESDEN

VENDOR-PROVIDED OS [JURECA]



TECHNISCHE UNIVERSITÄT DRESDEN

VENDOR-PROVIDED OS [TAURUS]

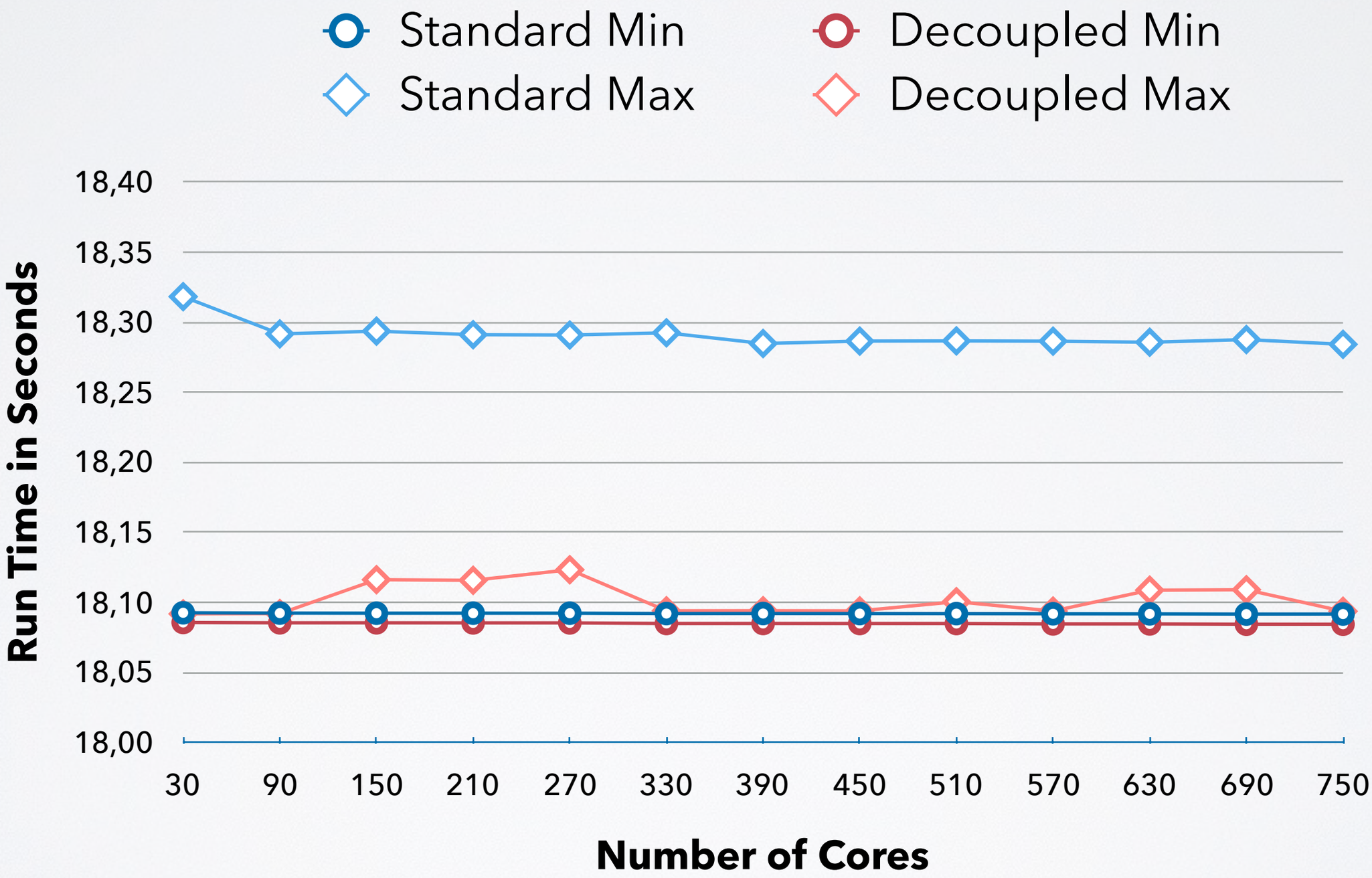


EVALUATION SETUP

- **Bare-metal access** to Taurus:
  - Little time
  - Fewer cores
  - Different type of nodes
- Vendor OS: **Linux 2.6.32** or **3.10** …
- Decoupled threads: **L⁴Linux 4.4**
- Custom Linux distribution

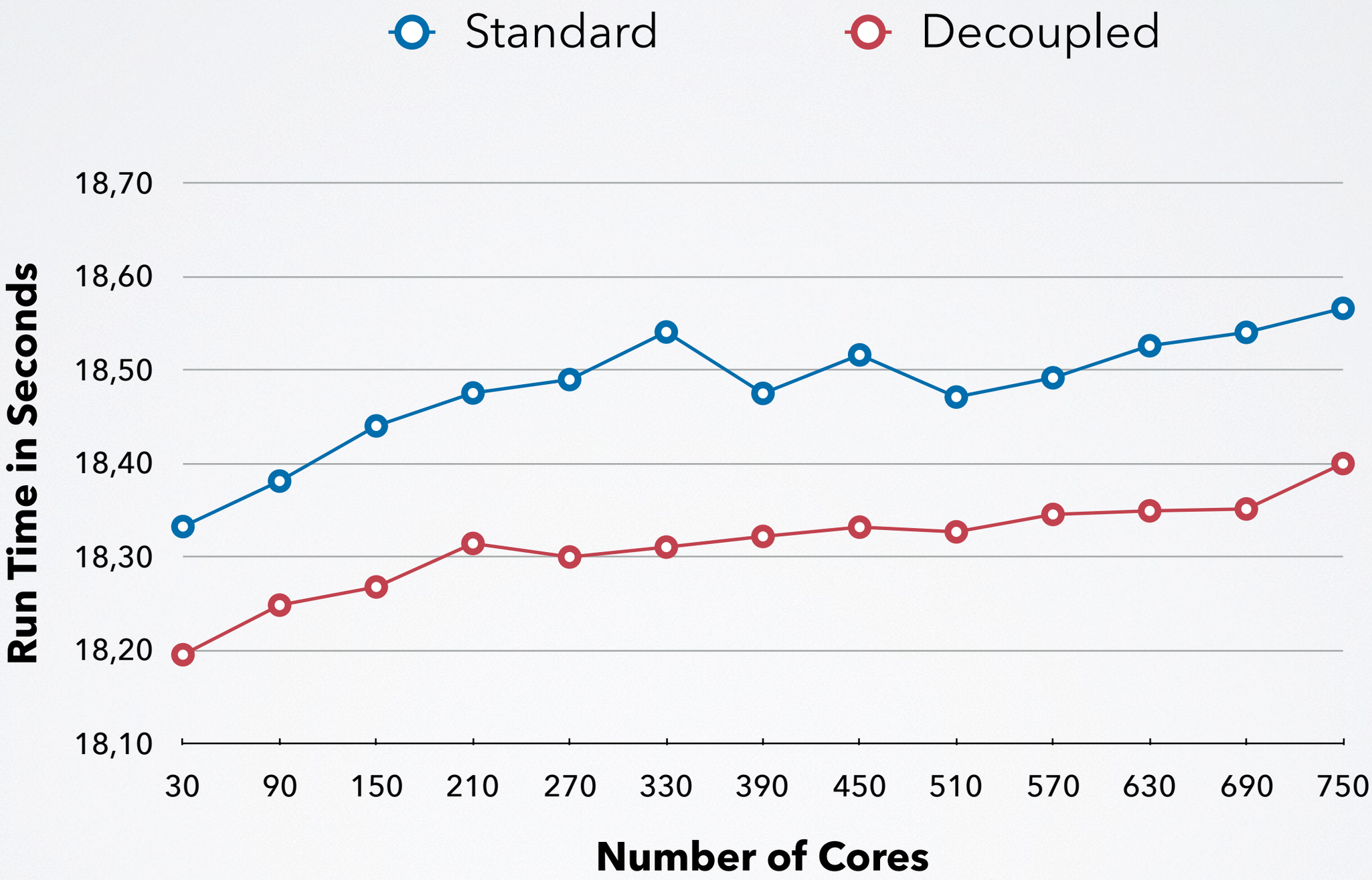
TECHNISCHE UNIVERSITÄT DRESDEN

L4LINUX+DECOUPLING: STARTSYNC



TECHNISCHE UNIVERSITÄT DRESDEN

L4LINUX+DECOUPLING: STEPSYNC



CONCLUSION & FUTURE WORK

- **Decoupled threads:** reduced noise
- **Virtualization:** run unmodified HPC codes
- Reuse existing components:
  L4 microkernel + **L⁴Linux**
- **Low effort:** developed within 2 weeks(*)
- Next steps: more nodes, more workloads

DFG Deutsche Forschungsgemeinschaft
SPPEXA German Priority Programme 1648
Software for Exascale Computing