

HEXO: Offloading HPC Compute-Intensive Workloads on Low-Cost, Low-Power Embedded Systems

Pierre Olivier*, A K M Fazla Mehrab*, Stefan Lankes⁺,
Mohamed Lamine Karaoui*, Robert Lyerly*, Binoy Ravindran*

*Virginia Tech, ⁺RWTH Aachen University

{ polivier | mehrab | karaoui | rlyerly | binoy }@vt.edu, slankes@eonerc.rwth-aachen.de

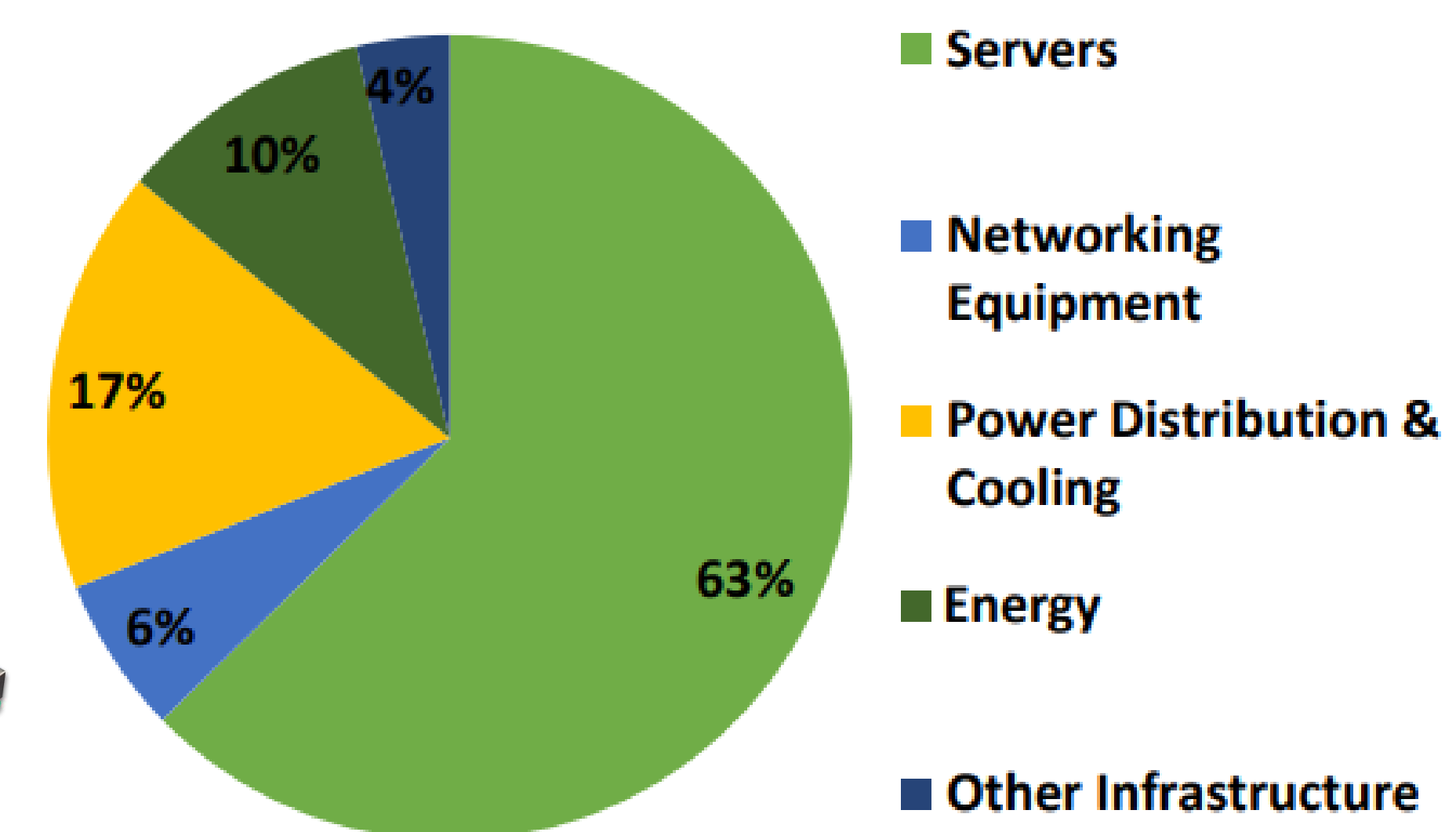
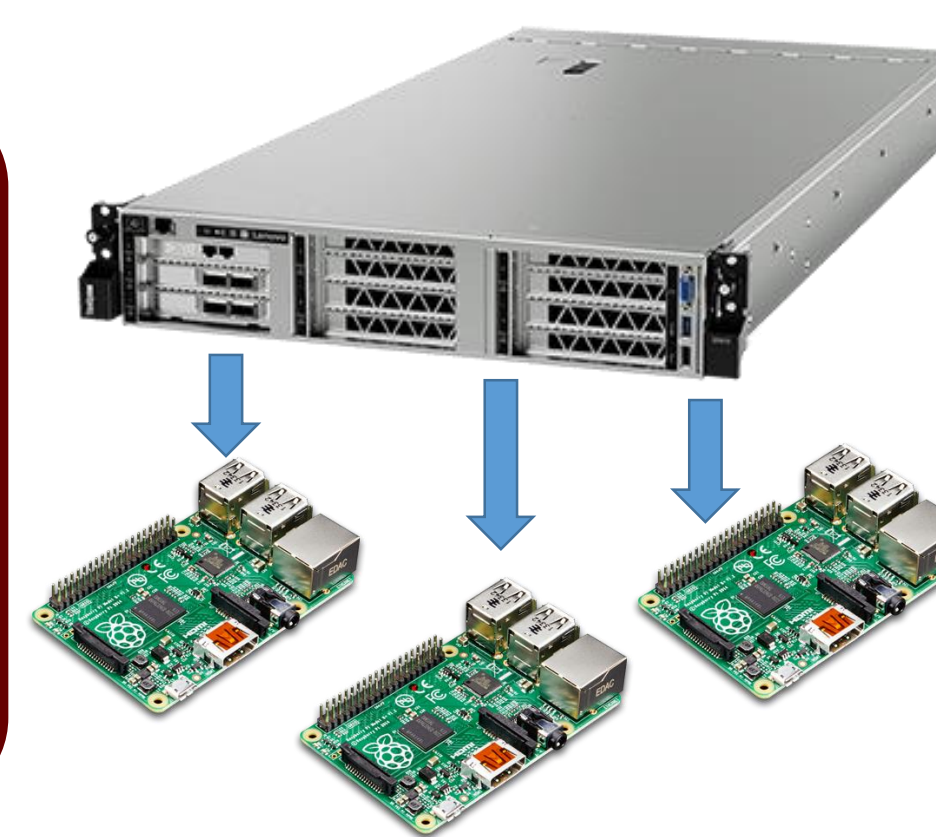
Supported by:



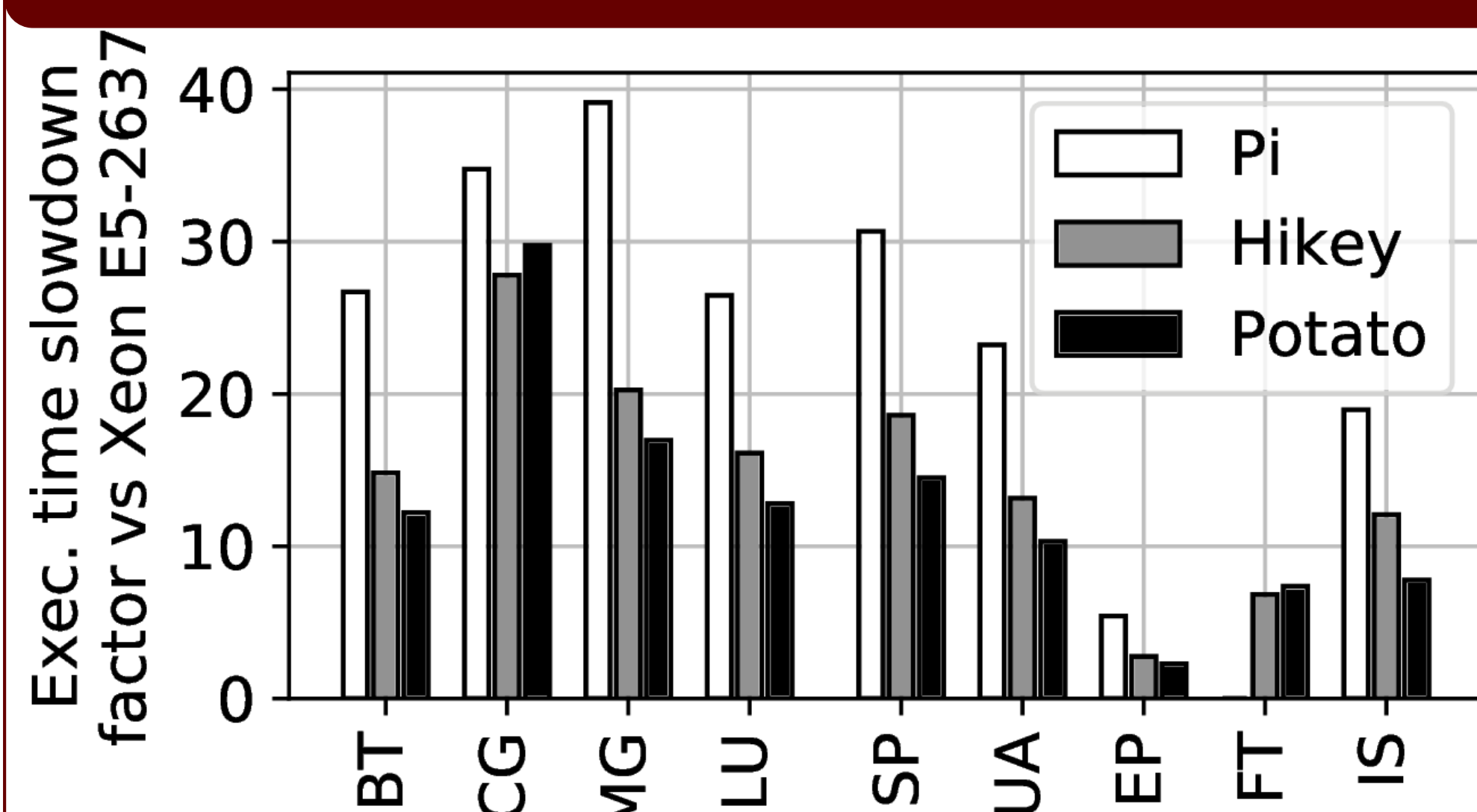
Context & Idea

- Major costs in the datacenter are driven by **server acquisition** and **power consumption** [1]
- Today manufacturers can produce **embedded systems** with a **very low power consumption**, at an **extremely low price point**

How can we integrate embedded systems within the datacenter?
Heterogeneous Execution Offloading (HEXO) selectively offloads at runtime datacenter/HPC workloads from servers to embedded systems for consolidation purposes



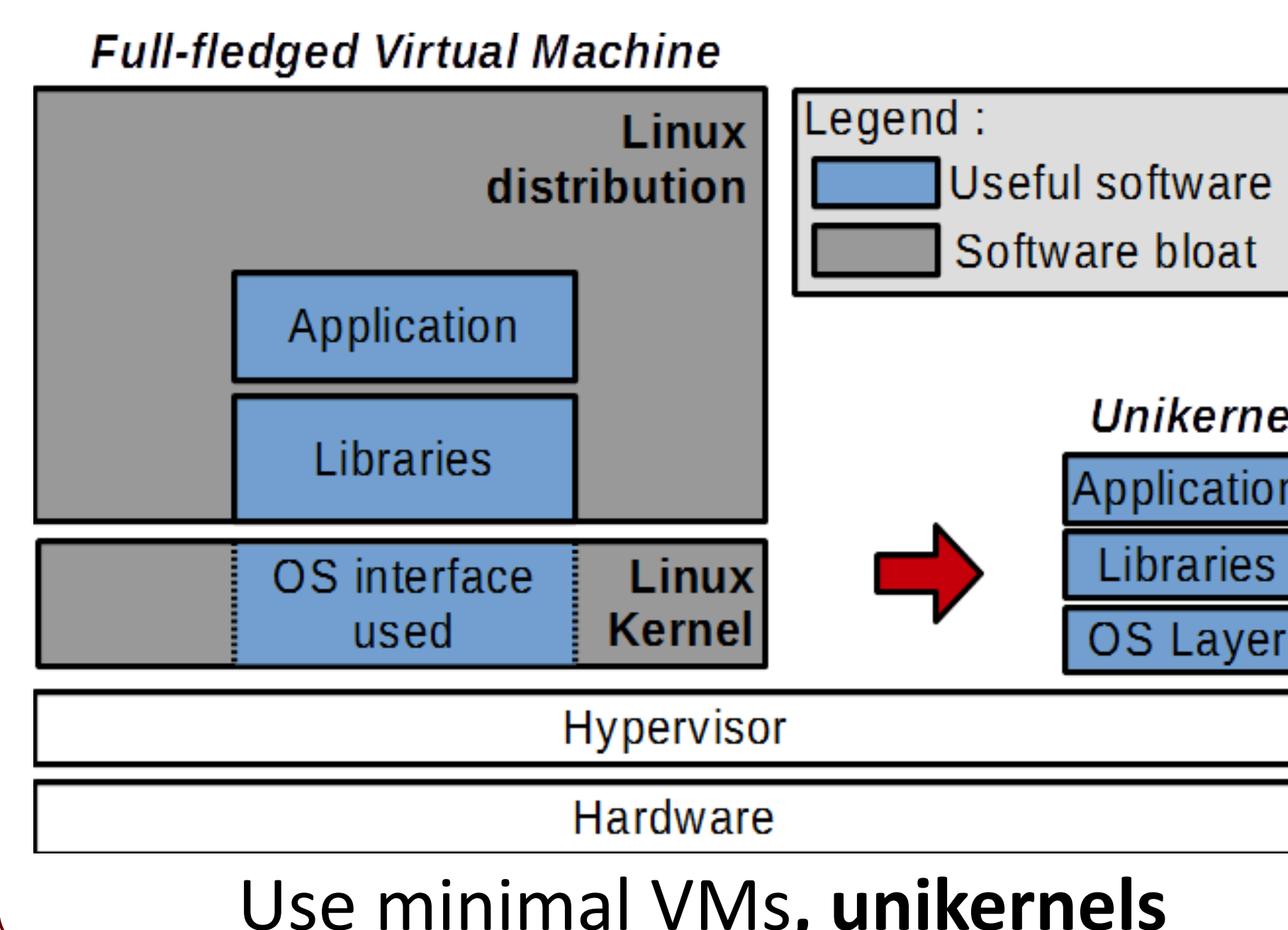
Motivation



Compared to servers, single-board computers are *not as slow* (2x – 40x) *as they are cheap* (100x)

Challenges

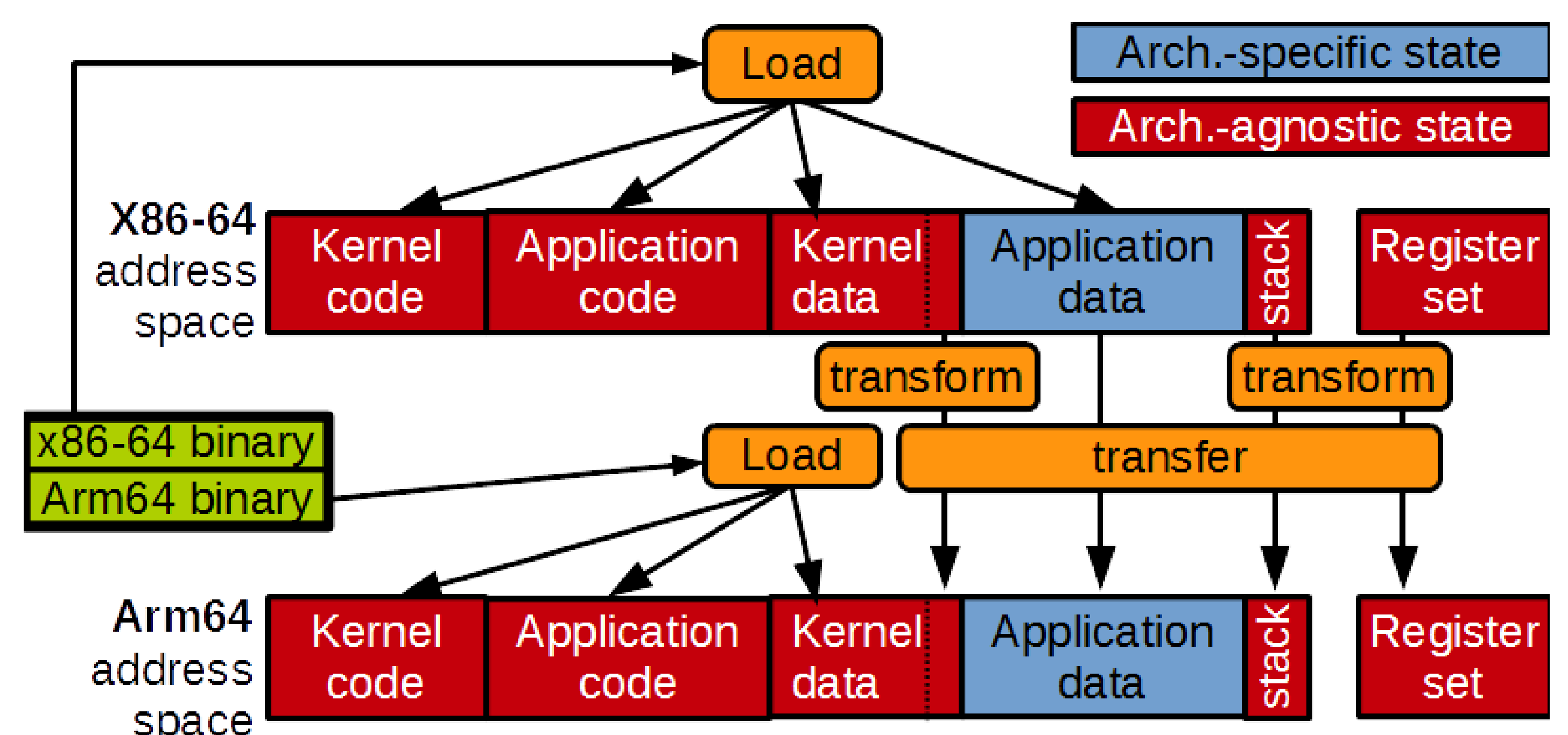
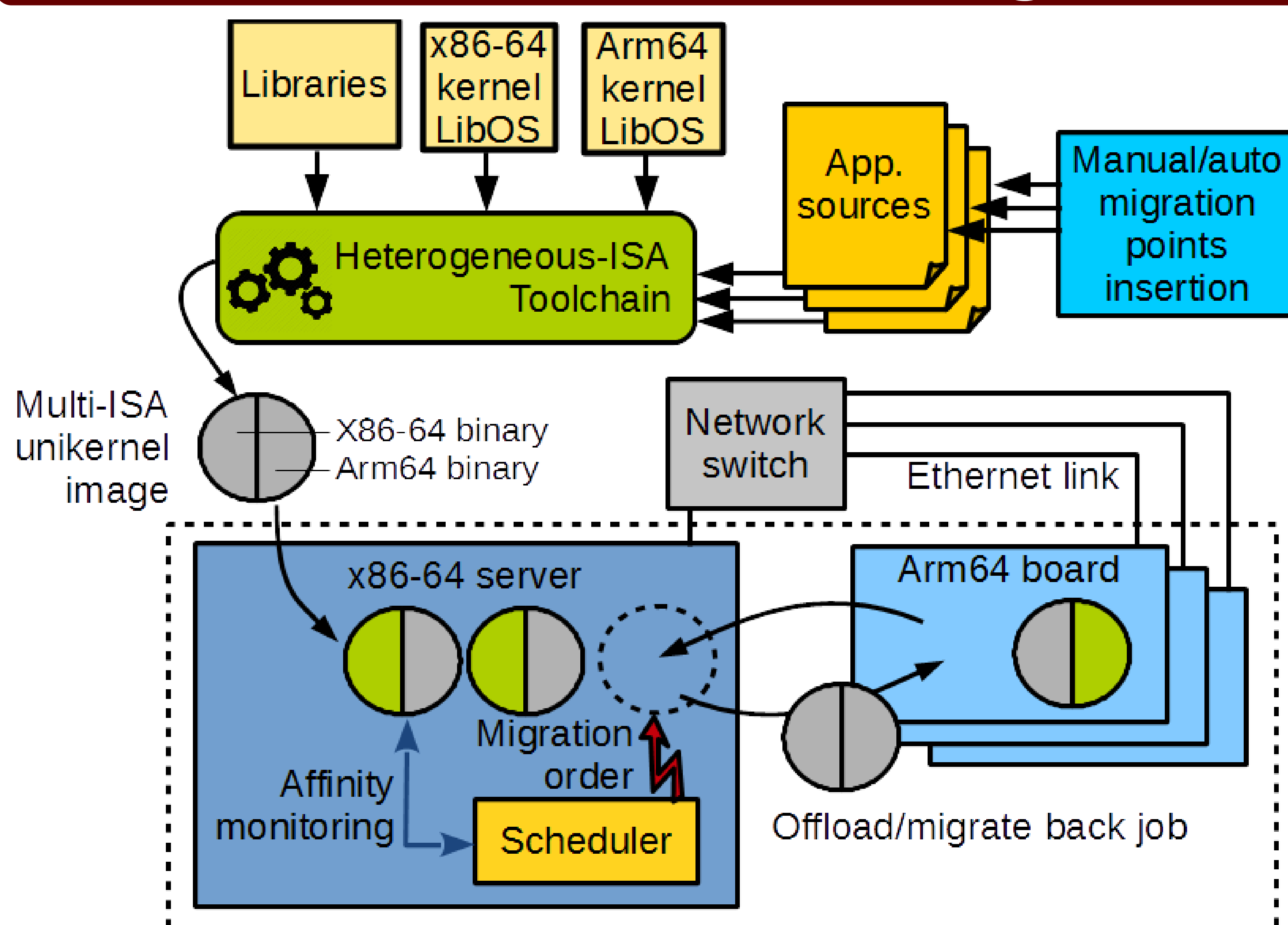
Low amount of resources



ISA difference

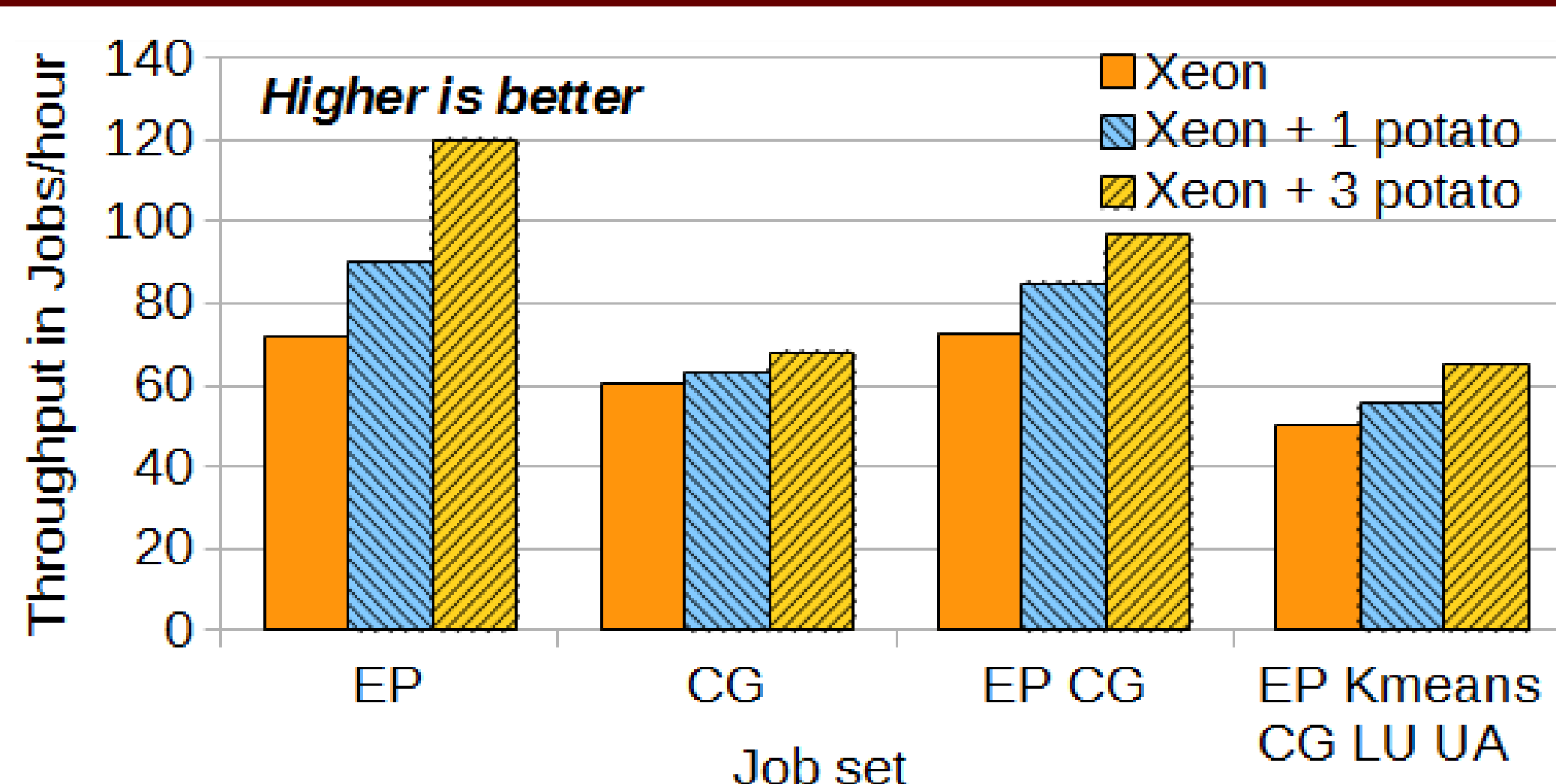
- Adapt **cross-ISA state transformation techniques** from Popcorn Linux [2]
- Create **mostly common address space**:
 - Assume common data format (endianness, types sizes/alignments)
 - Align functions & global variables at the same addresses in both ISAs
- Transform architecture-specific state**:
 - Stack and register set

HEXO: Heterogeneous EXecution Offloading



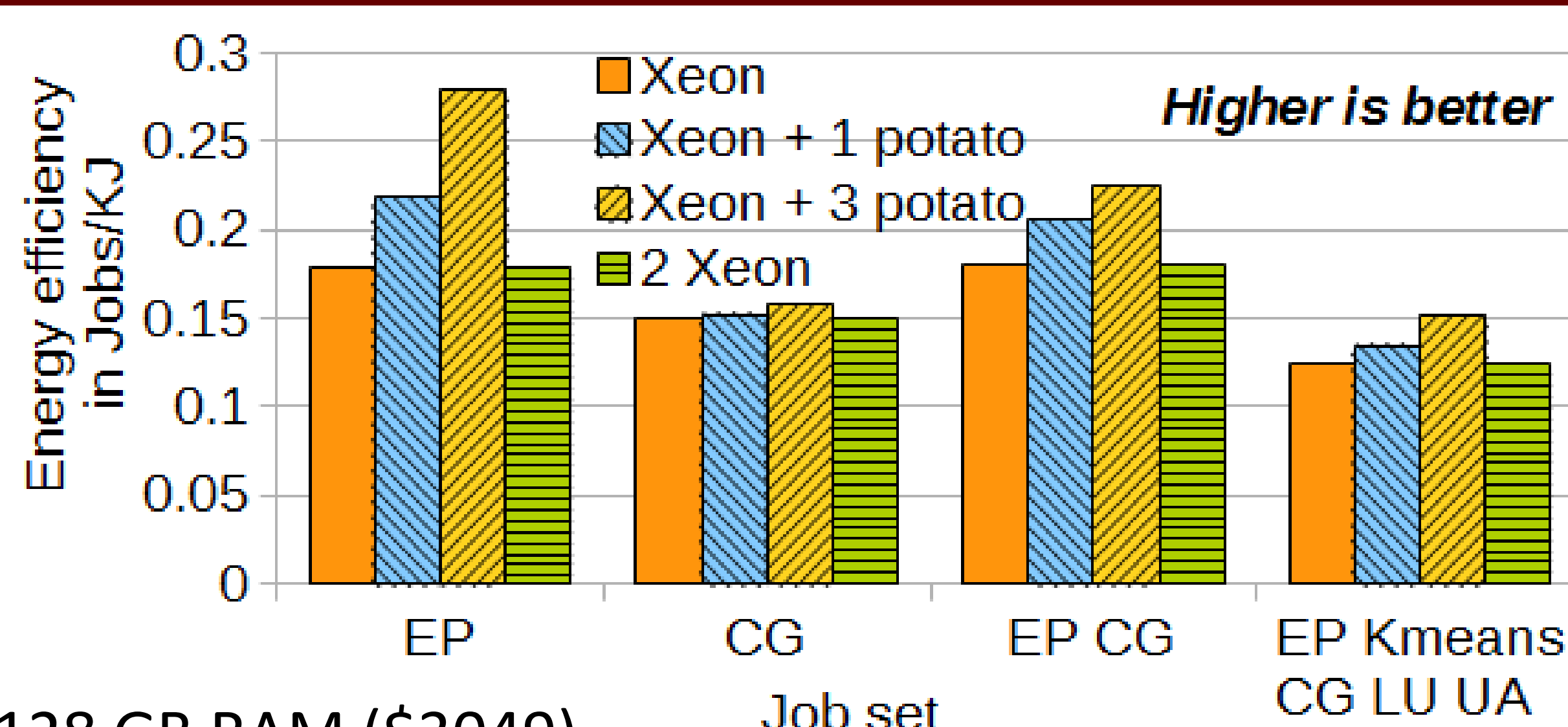
- Custom toolchain** to create multi-ISA unikernel images
 - Instrumented for migration between x86-64 and ARM64
 - OS Based on HermitCore [3] (ported to ARM64)
- Semantic VM migration** to transfer and potentially transform application's state
- Scheduler** offloads jobs based on slowdown estimation

Evaluation



Server: Colfax x86-64 Xeon E5-2637 @3.5 GHz + 128 GB RAM (\$3049)

Embedded systems: LibreComputer LePotato ARM64 Amlogic S905X @1.5 GHz + 2 GB RAM (\$45)



References

- [1] R. Bianchini. 2017. **Improving Datacenter Efficiency**. ASPLOS keynote, 2017.
- [2] A. Barbalace, R. Lyerly, C. Jelesnianski, A. Carno, H. Chuang, V. Legout, and B. Ravindran, **Breaking the Boundaries in Heterogeneous-ISA Datacenters**, ASPLOS, 2017.
- [3] S. Lankes, S. Pickartz, and J. Breitbart, **HermitCore: a Unikernel for Extreme Scale Computing**, ACM ROSS workshop, 2016.

HEXO is open source and available at <http://www.popcornlinux.org/index.php/hexo>