TECHNISCHE UNIVERSITÄT DRESDEN

ZIB

The Hebrew University of Jerusalem

FFMK – A FAST AND FAULT-TOLERANT MICROKERNEL-BASED SYSTEM FOR EXASCALE COMPUTING

Amnon Barak	Hebrew University Jerusalem (HUJI)
Hermann Härtig	TU Dresden, Operating Systems Group (TUDOS)
Wolfgang Nagel	TU Dresden, Center for Information Services and HPC (ZIH)
Alexander Reinefeld	Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

HERMANN HÄRTIG, TU DRESDEN

2015NOVEMBER17RUNTIME&OS BOF SC15

MOTIVATION/APPROACH

Node OS

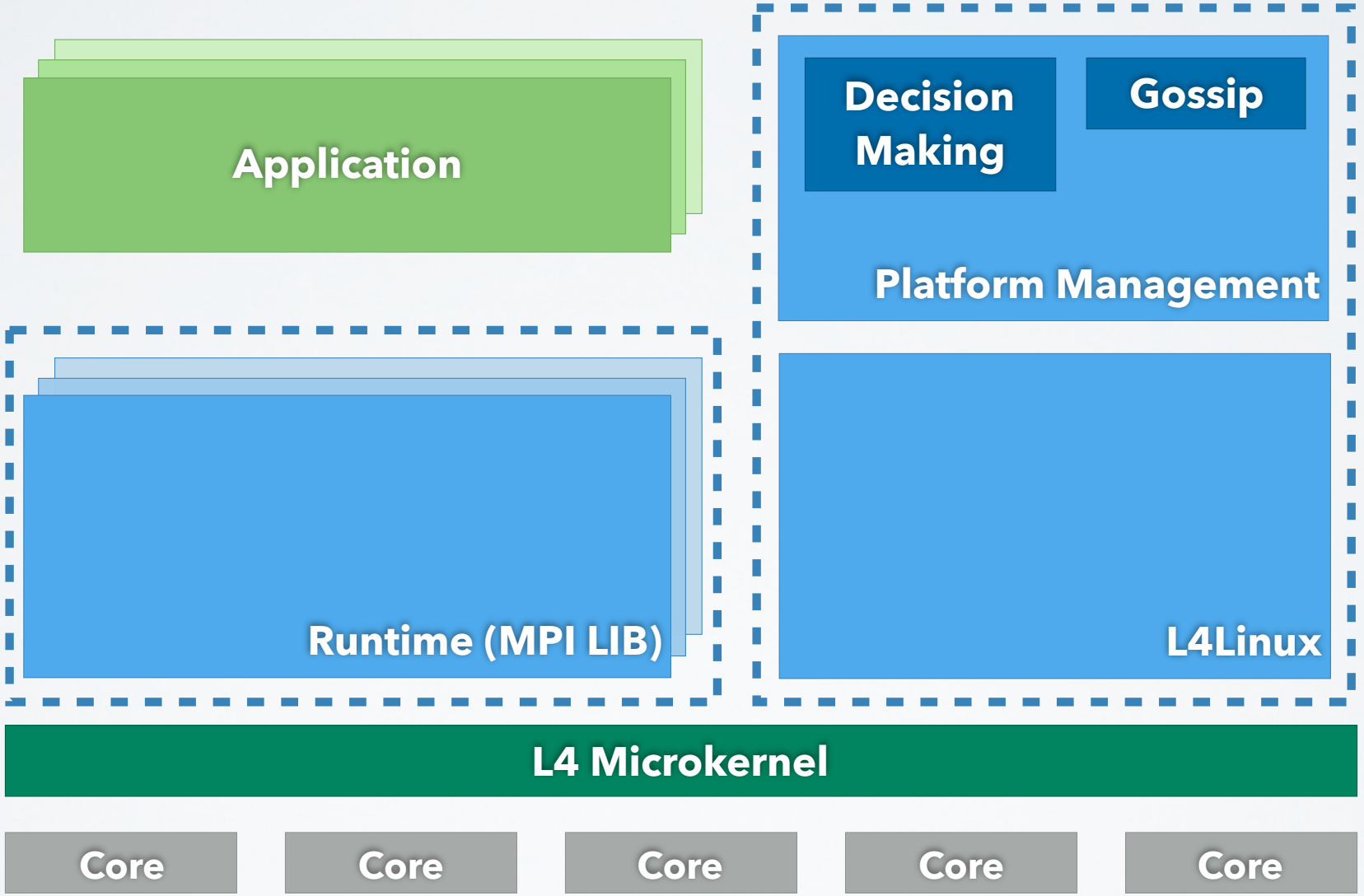
- load management thru OS mechanisms
  - full control over node platform
  - predictability (determinism)
  - OS scheduling (blocking instead of polling)
  - software partitioning (schedulers & capabilities)

MOTIVATION/APPROACH

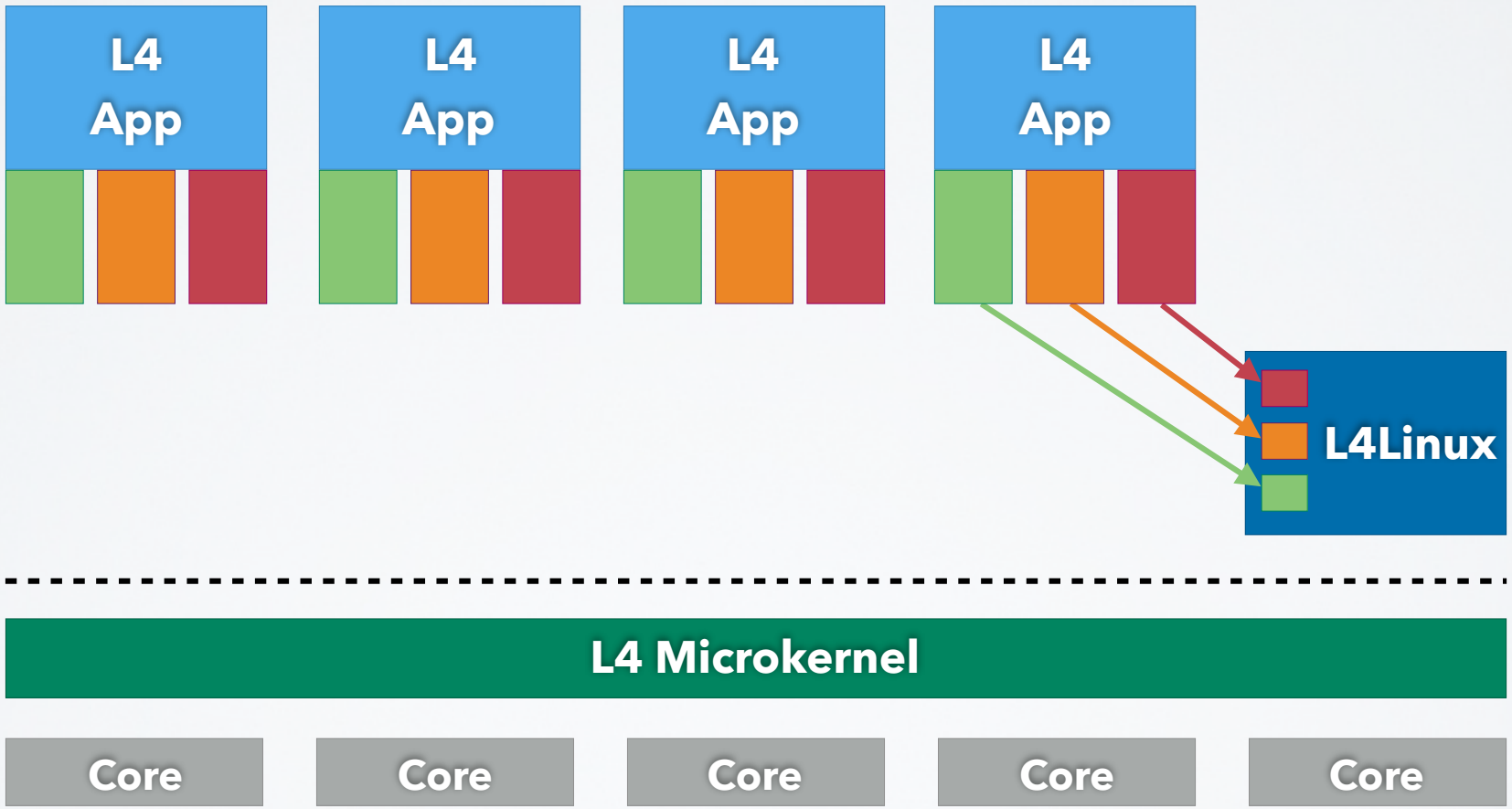
Node OS

- exploit Linux & take it out of the way
  - use microkernel & virtualization
  - specialized LibOS

NODE ARCHITECTURE



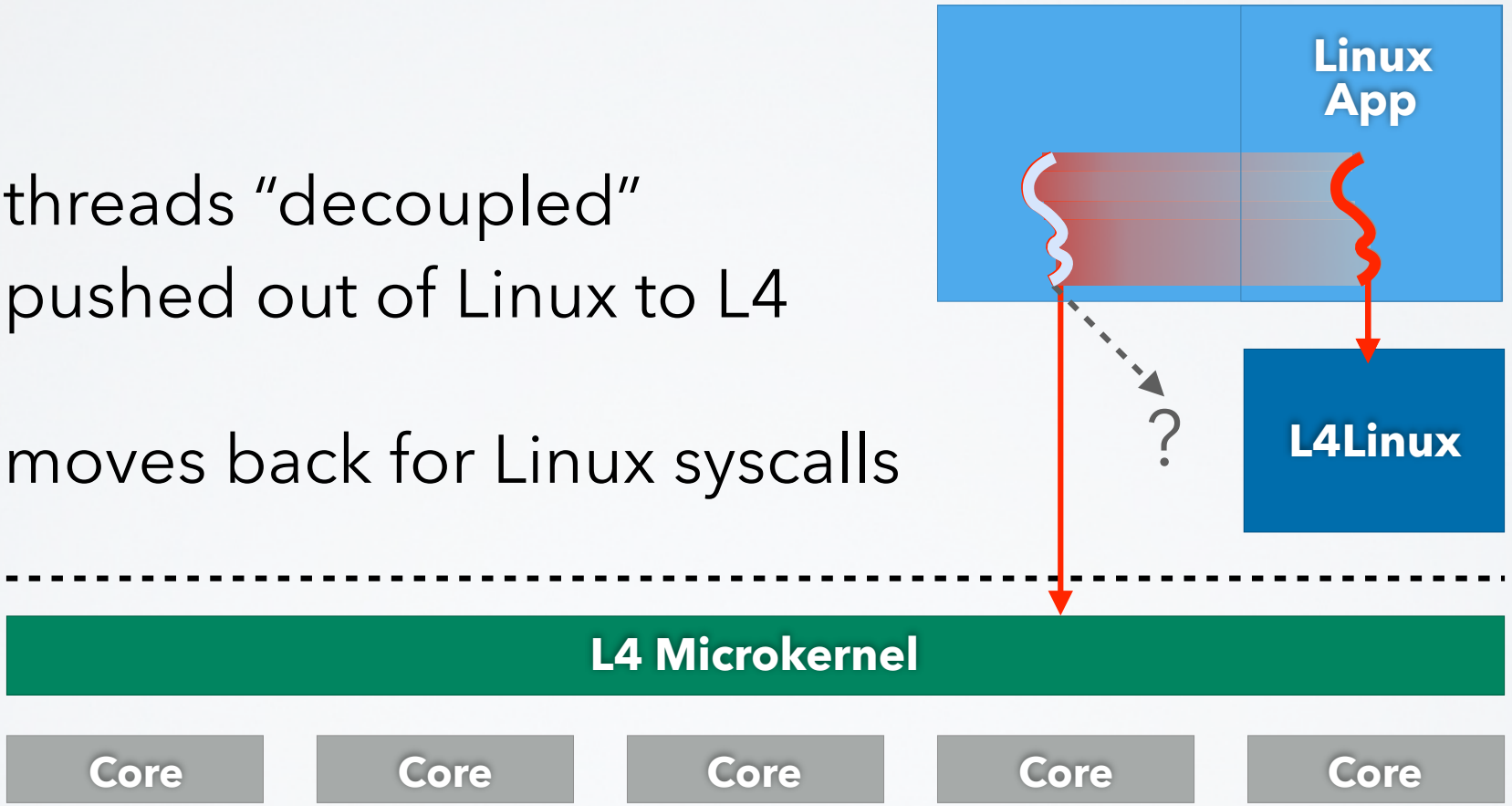
EXECUTION MODEL (1)



EXECUTION MODEL(2)

threads “decoupled”
pushed out of Linux to L4

moves back for Linux syscalls



MORE

global platform control:

- elastic partitions
- Amnon Barak & team:
  gossip and hints
- Alexander Reinefeld:
  checkpointing into on-node storage

LICENSE

- GPL v2:
  L4/Fiasco & L4Linux & L4Re
  caveat: HPC implementation not yet available (immature)
- visit: l4re.org

ACKNOWLEDGMENT

Funded by DFG
Priority Program “Software for Exa-Scale”