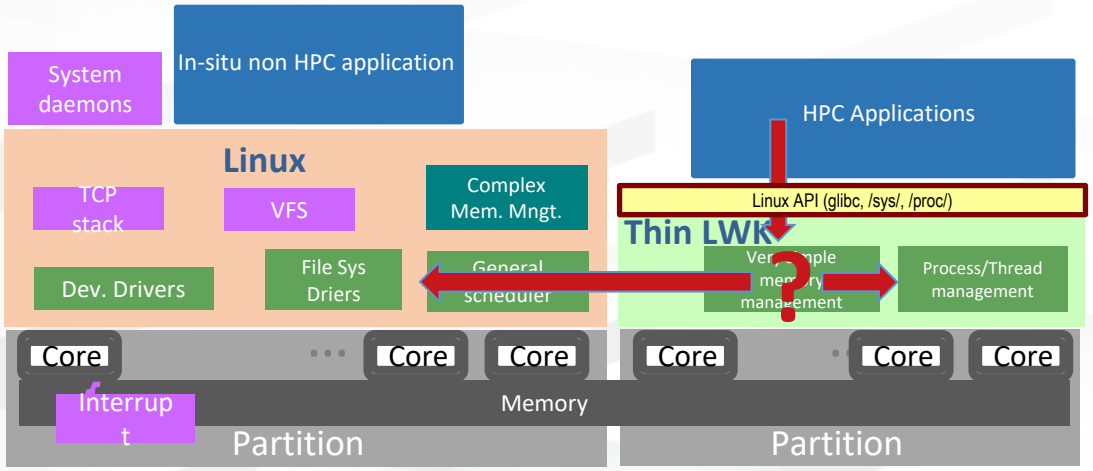
McKernel:
Light-weight Kernel for HPC Applications with Linux

Operating System on Compute Node

Linux kernel on 2 or 4 cores
System daemons and in-situ non HPC applications
Device drivers
Light-weight kernel(LWK), McKernel on other cores
HPC applications

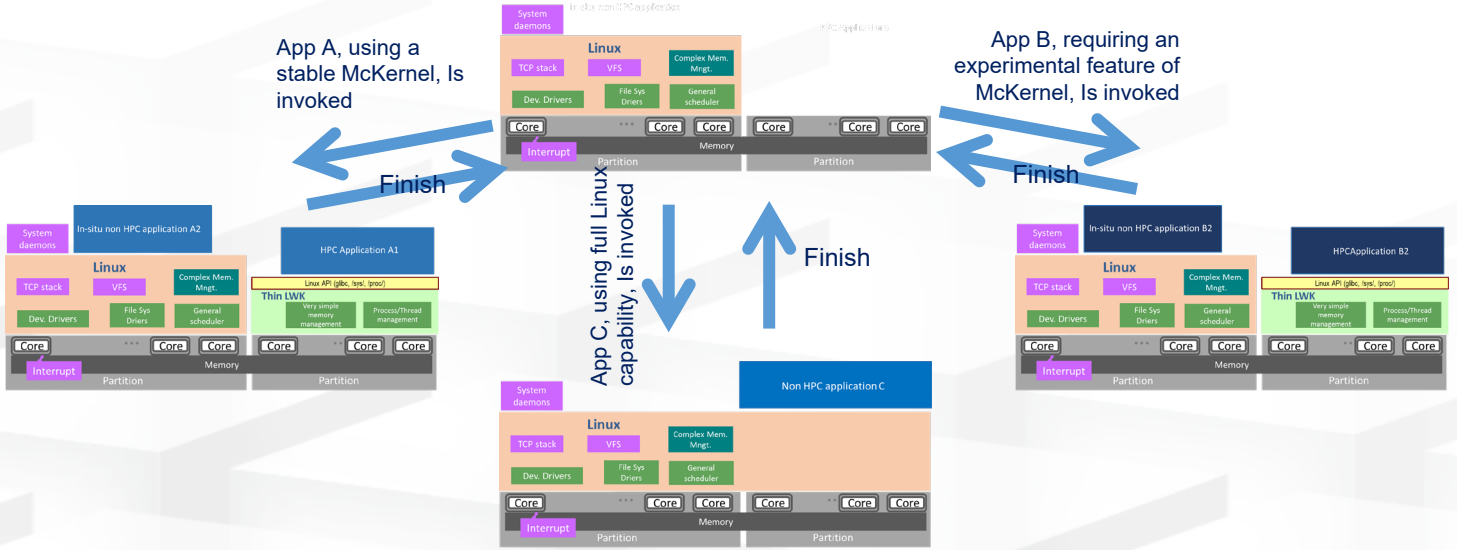


- **McKernel**
  - Executes the same binary of Linux without any recompilation
  - One of advantages is that McKernel provides much larger page sizes
    - Applications, accessing a huge memory area randomly, may benefit
  - User may select one of McKernel configurations without rebooting

	McKernel (4K)	*McKernel (64K)*	*Linux*
.text	*4K*	*64K*	*64K*
.data	*64K,2M,32M, 1G*	*2M, 512M*	*2M*
.bss	*64K,2M,32M, 1G*	*2M, 512M*	*2M*
Stack	*64K,2M,32M, 1G*	*2M, 512M*	*2M*
malloc	*64K,2M,32M, 1G*	*2M, 512M*	*2M*
thread stack	*64K,2M,32M, 1G*	*2M, 512M*	*2M*
System V IPC	*64K,2M,32M, 1G*	*2M, 512M*	*64K*
POSIX shm	*4K*	*64K*	*64K*
XPMEM	*64K,2M,32M, 1G*	*2M, 512M*	*64K*

How to deploy IHK/McKernel

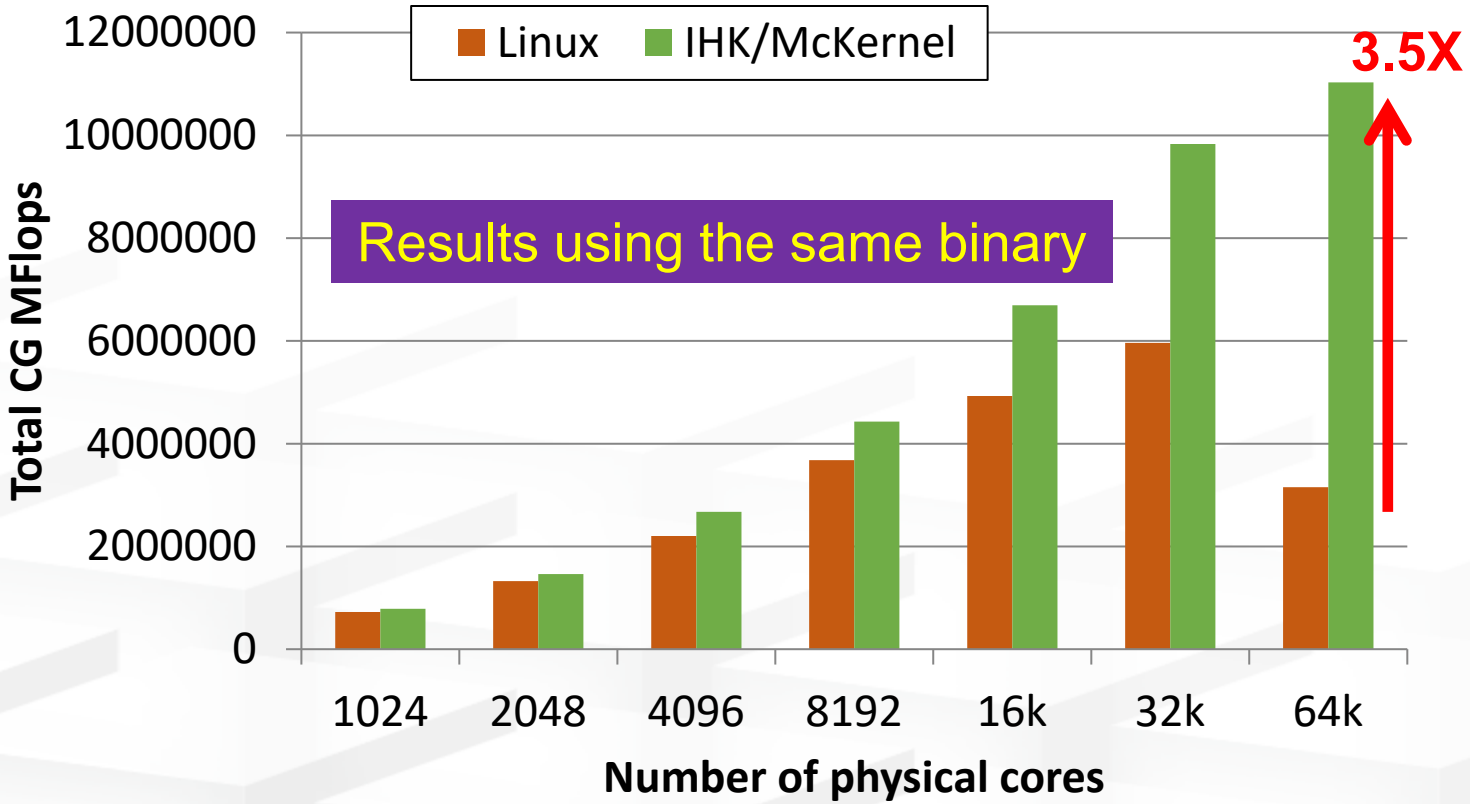
- Linux Kernel with IHK kernel module is resident
  - daemons for job scheduler and etc. run on Linux
- McKernel is dynamically reloaded (rebooted) by IHK for each application
  - No hardware reboot



miniFE (CORAL benchmark suite)

- Conjugate gradient - strong scaling
- Up to 3.5X improvement (Linux falls over..)

Oakforest-PACS supercomputer, 25 PF in peak, at JCAHPC organized by U. of Tsukuba and U. of Tokyo

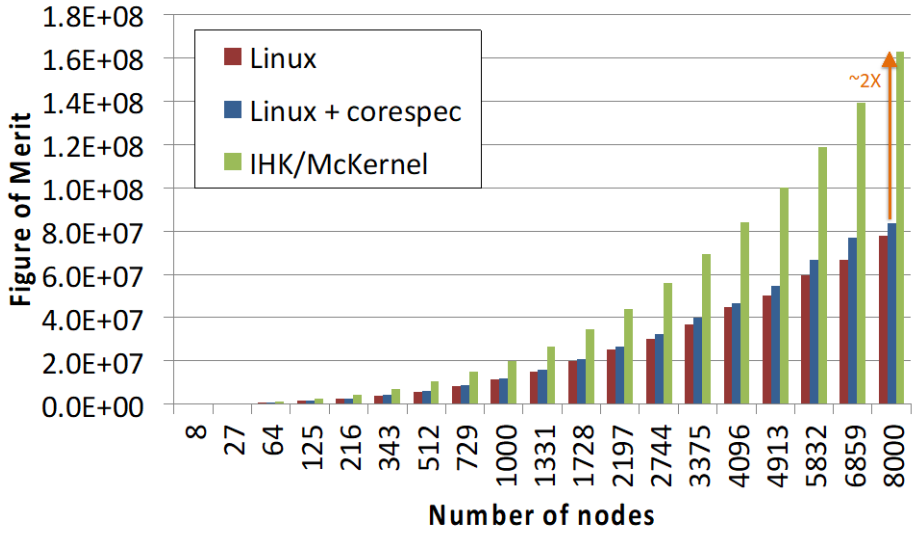


Balazs Gerofi, Rolf Riesen, Robert W. Wisniewski and Yutaka Ishikawa: “Toward Full Specialization of the HPC System Software Stack: Reconciling Application Containers and Lightweight Multi-kernels”, International Workshop on Runtime and Operating Systems for Supercomputers (ROSS), 2017

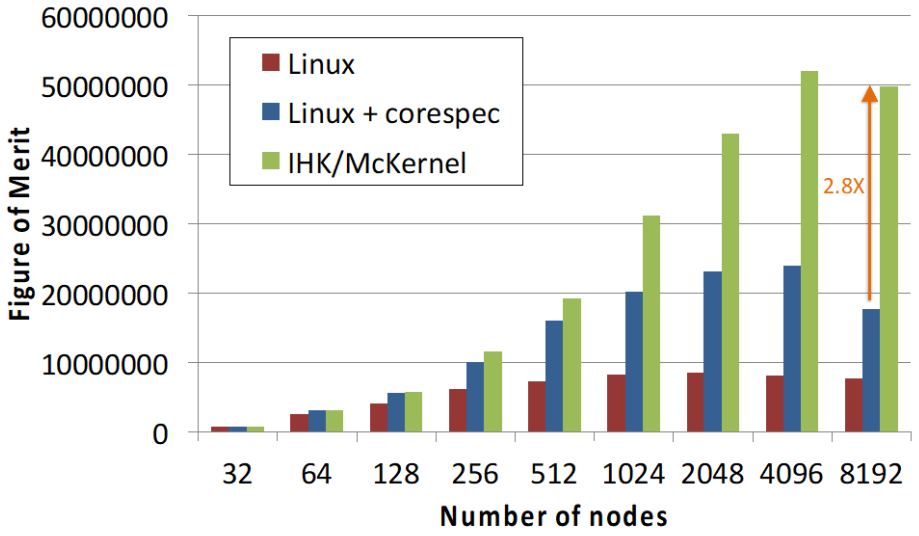
Lulesh and MiniFE (CORAL benchmark suite)

- **BSP applications and stencils seem to benefit from the jitterless environment of the LWK**

Oakforest-PACS supercomputer, 25 PF in peak, at JCAHPC organized by U. of Tsukuba and U. of Tokyo



Lulesh (CORAL), weak-scaled



MiniFE (CORAL), strong-scaled

Results using the same binary as on Linux

Balazs Gerofi, Rolf Riesen, Masamichi Takagi, Taisuke Boku, Yutaka Ishikawa, Robert W. Wisniewski: "**Performance and Scalability of Lightweight Multi-Kernel based Operating Systems**", *IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, 2018