FlexMem: Proactive Memory Deduplication for Qcow2-Based VMs with Virtual Persistent Memory

Weinan Liu, Zhihao Zhang, Xiangrong Liu and Yiming Zhang*
wnliu@stu.xmu.edu.cn

厦門大學

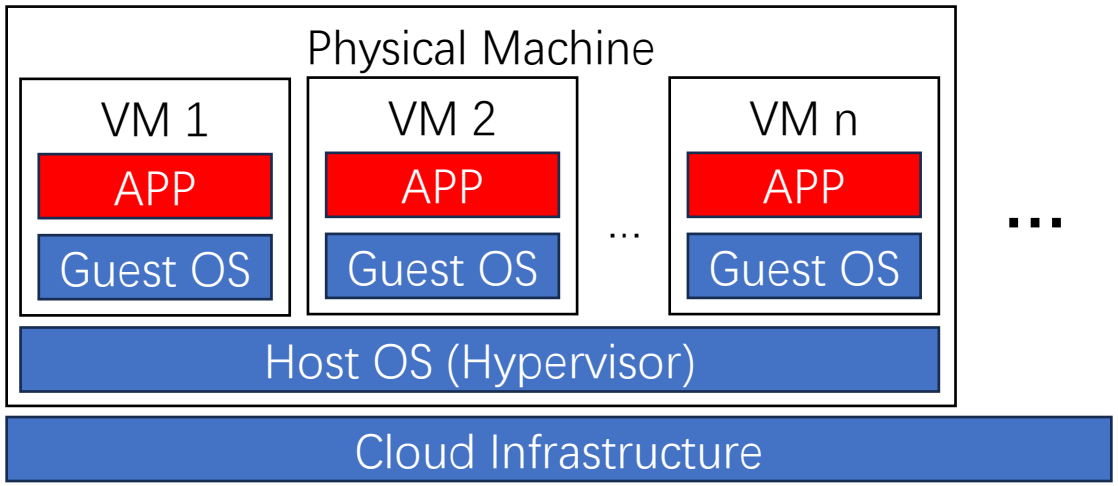
×

NICEXLAB
Networked Intelligent Computing at EXascale

*corresponding author

Background: Virtualization-based Cloud

Commodity cloud providers virtualization-based cloud services to tenants.

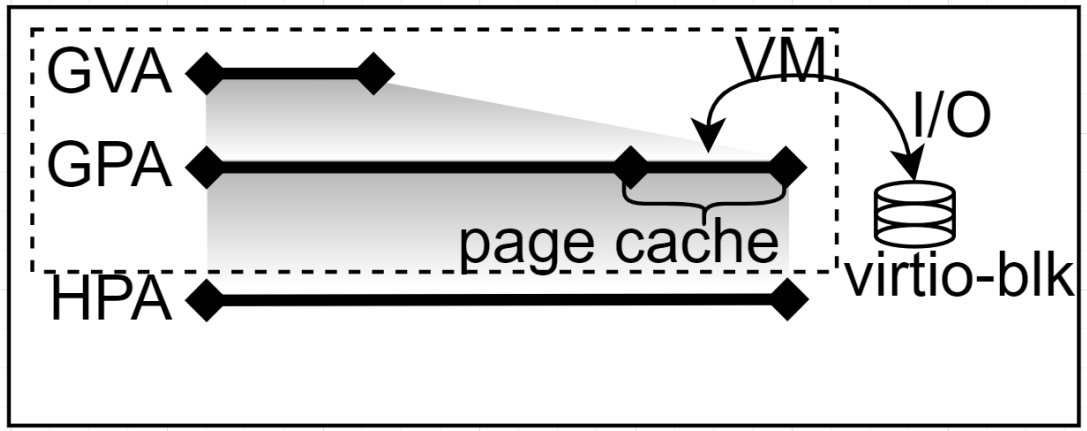


The density of VMs on a physical machine could be extremely high.

Background: Hardware Virtualization

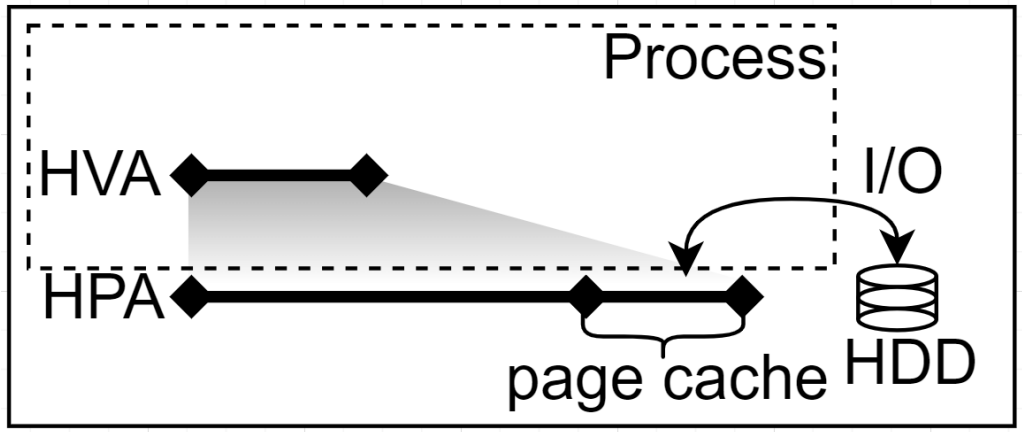
A VM closely resembles a Process.

For a VM



- GPA maps to HPA
- From *host supervisor mode* to enter *guest mode*
- To *host supervisor mode* when interrupted

For a normal user program



- HVA maps to HPA
- From *host supervisor mode* to enter *user mode*
- To *host supervisor mode* when interrupted

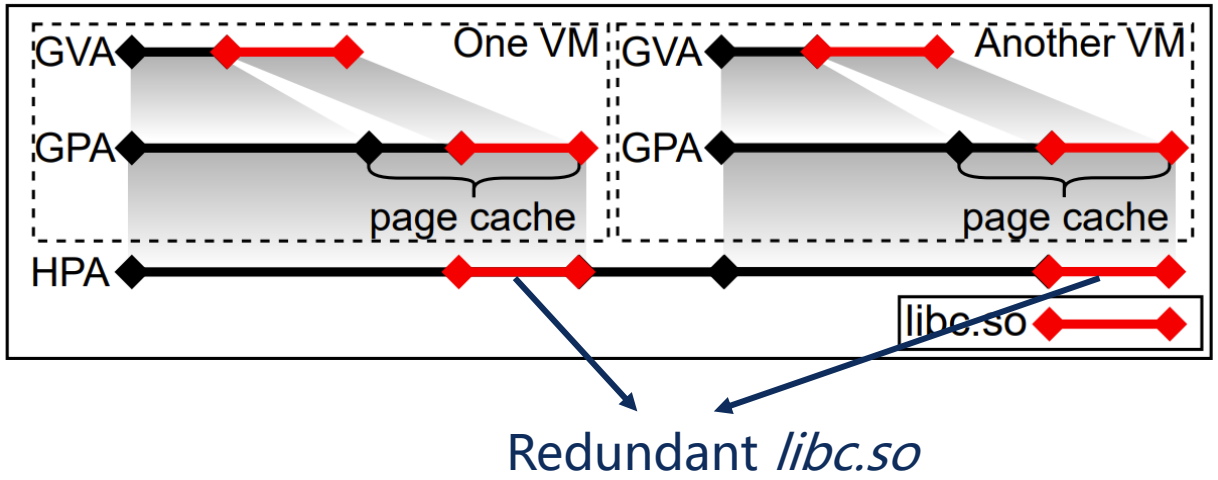
$$\begin{cases} \text{GPA} \approx \text{HVA} \\ \textit{guest mode} \approx \textit{host user mode} \end{cases}$$

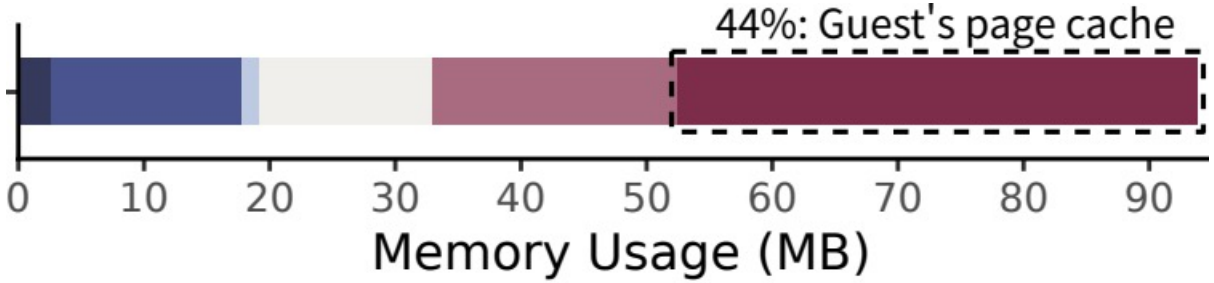
VM has near-native performance and system-isolation.

Motivation: Pagecache Redundancy

Multiple duplicates of read-only data coexist on the host machine.

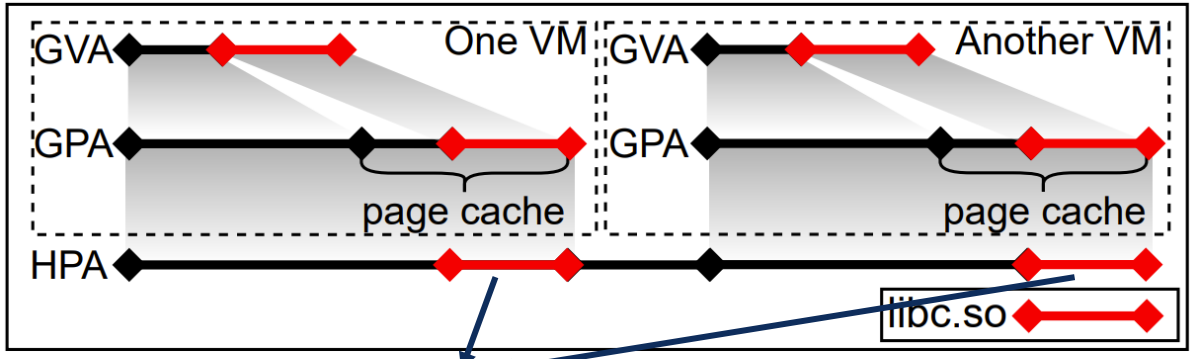
- Vast APPs base on the identical image due to their similarity.
- The greedy nature of page cache.





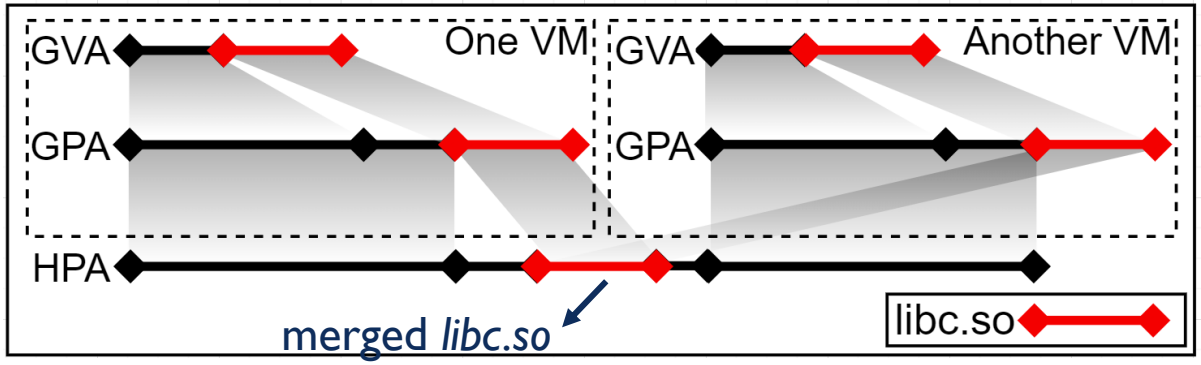
Substantial guest pagecache redundancy on the high VM-density machine.

Our Goal: RO Data Deduplication



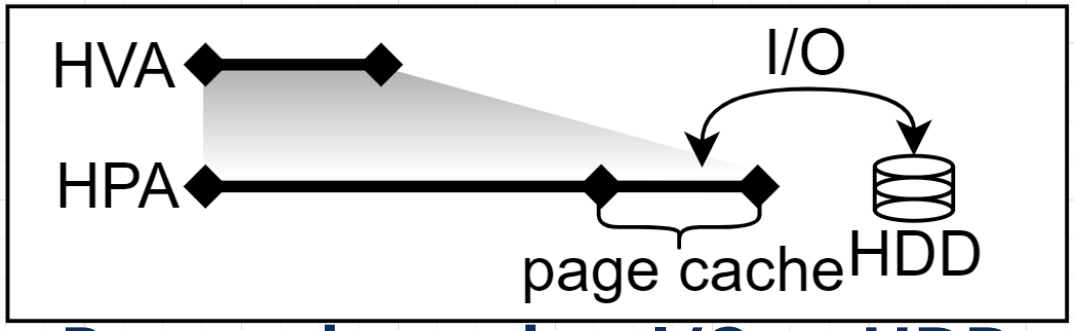
Traditional: Duplicated libc.so in the HPA space

If we can make this



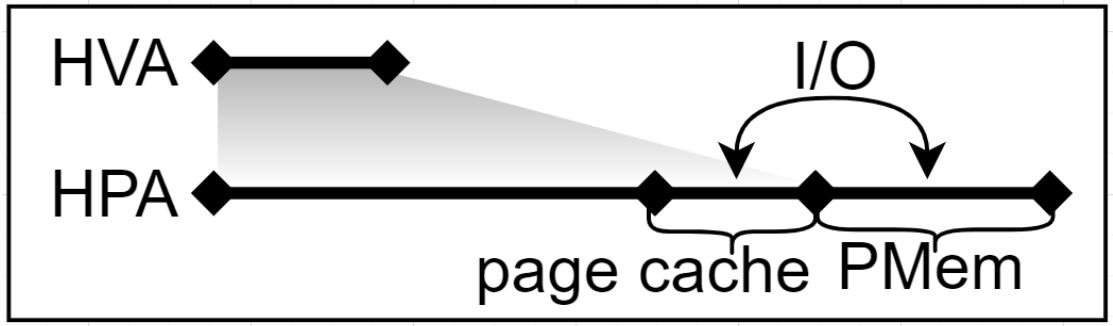
The duplicated *libc.so* in the guest pagecache will be merged.

Our Chance: PMem and DAX



Pagecache caches I/O to HDD

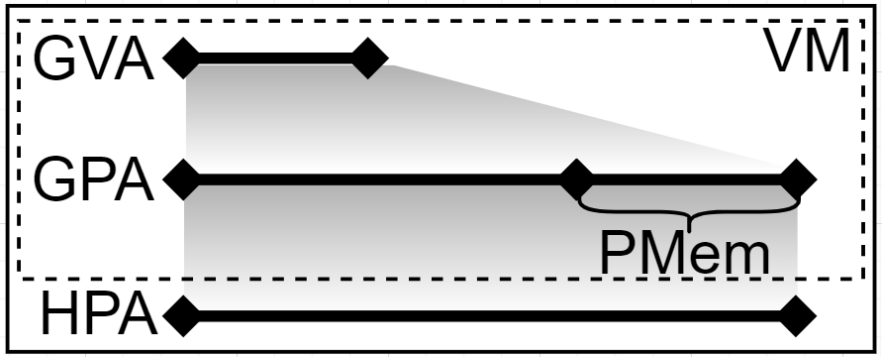
1



PMem (a block device accessed by HPA)

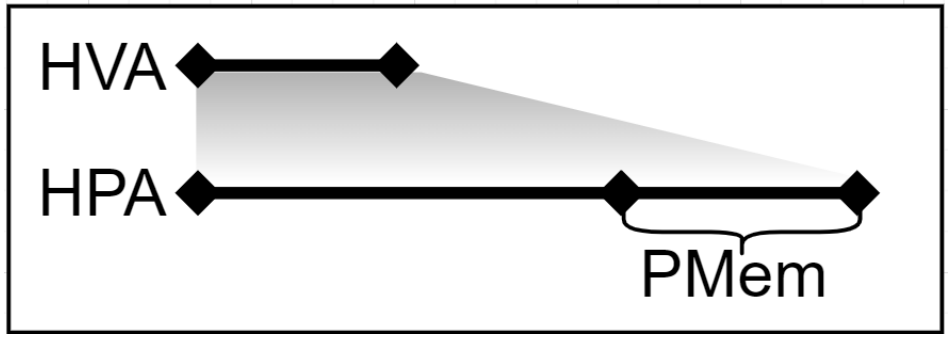
2

4



Virtual PMem with guest DAX

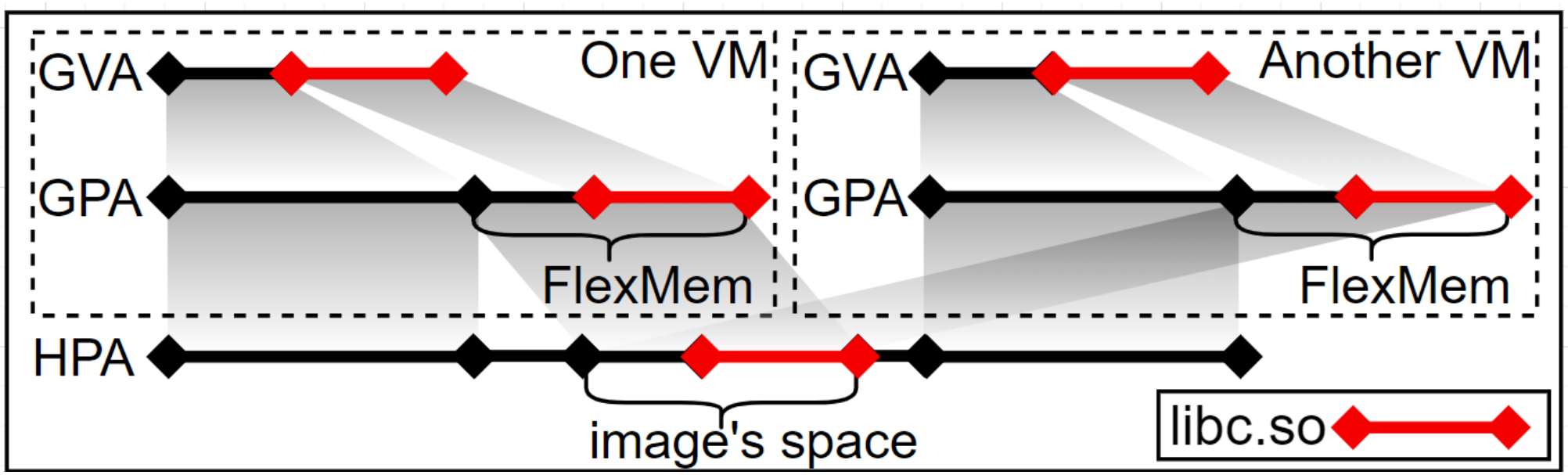
3



PMem with DAX (disable page cache)

FlexMem: A Virtual PMem

FlexMem is a virtual PMem.
Provide the identical QCOW2 image to VMs in the form of FlexMem.

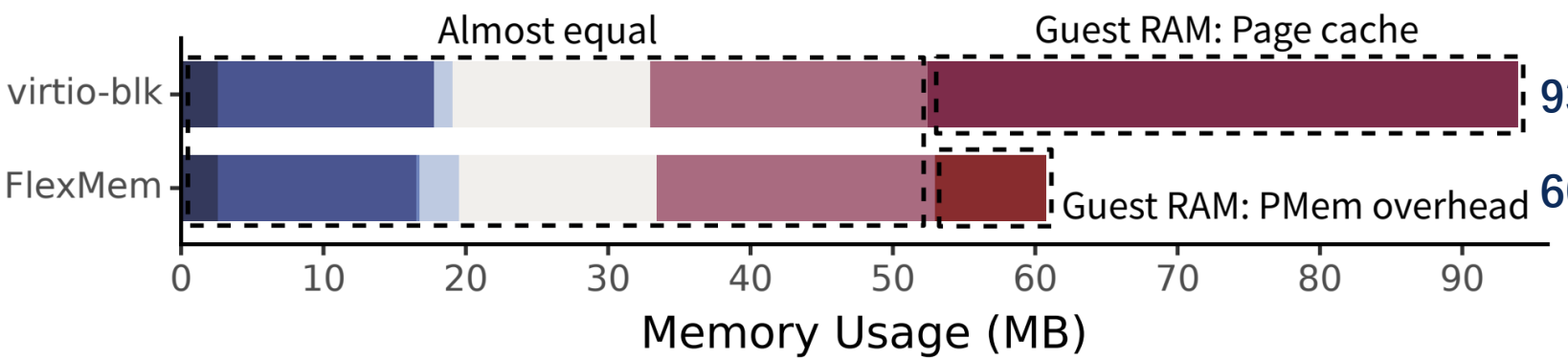


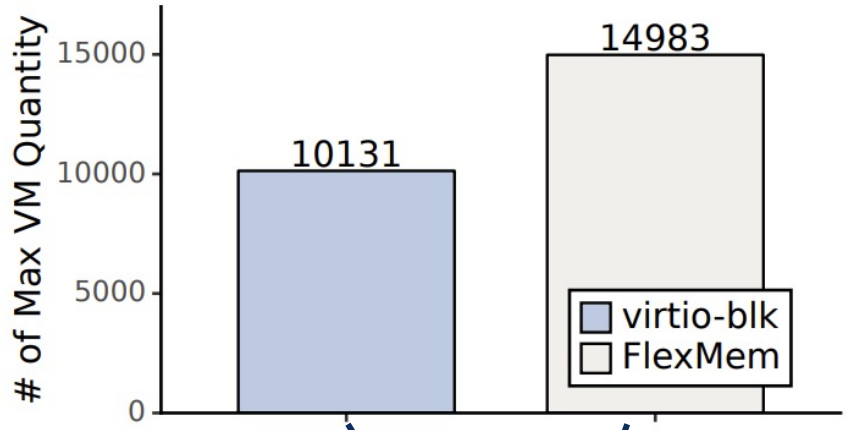
To merge the duplicated *libc.so* in the guest pagecache.

FlexMem: VM Memory Usage

Does FlexMem eliminate the pagecache redundancy?

- VM memory breakdown.
- VM Quantity.





VM Quantity Improvement

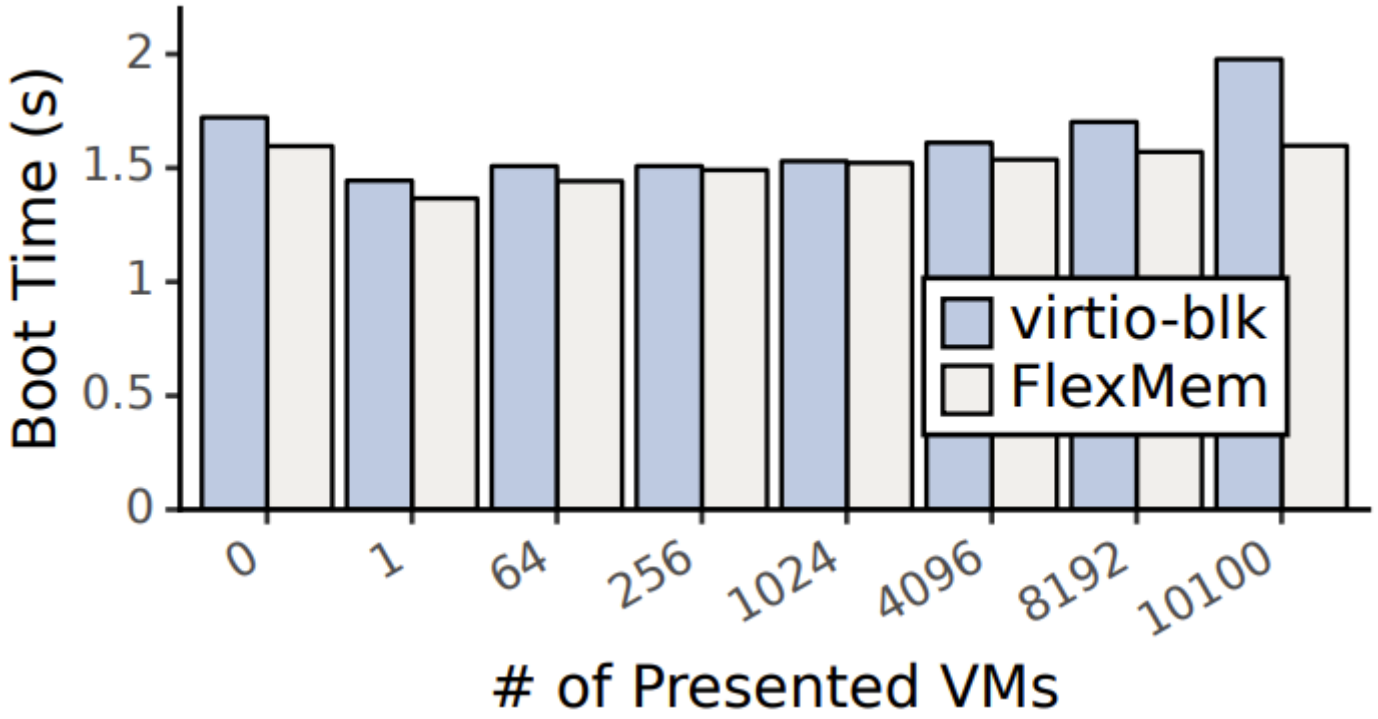
93 MB → $\frac{93MB}{60MB} \approx \frac{14983}{10131}$

60 MB →

The pagecache redundancy is eliminated, but FlexMem incurres overhead.
The breakdown is precise.

FlexMem: Boot Time Accelerating

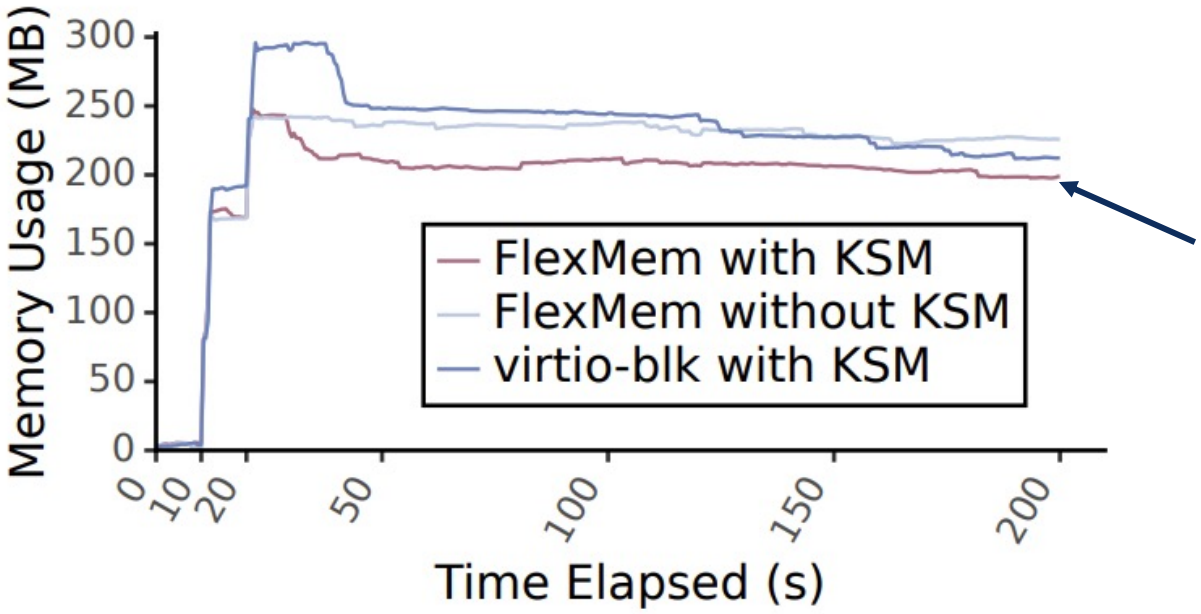
Does FlexMem impove I/O performance?



FlexMem can accelerate boot time under the same VM density.

FlexMem: Compared to KSM

Does FlexMem eliminate all the redundancy?



Memory Usage Changes over Time with or without KSM

Compared to non-KSM, *FlexMem with KSM* can merge more pages.
There still has been redundancy other than pagecache, like the linux kernel.

Conclusion

FlexMem pro:

- Can eliminate pagecache redundancy.
- Can accelerate VM boot time.

Future works:

- To eliminate the overhead of PMem.
- Kernel sharing, to eliminate redundancy other than pagecache.

Thanks & QA

FlexMem: Proactive Memory Deduplication for Qcow2-Based VMs with Virtual Persistent Memory

Weinan Liu, Zhihao Zhang, Xiangrong Liu and Yiming Zhang*
wnliu@stu.xmu.edu.cn

厦門大學

×

NICEXLAB
Networked Intelligent Computing at EXascale

*corresponding author