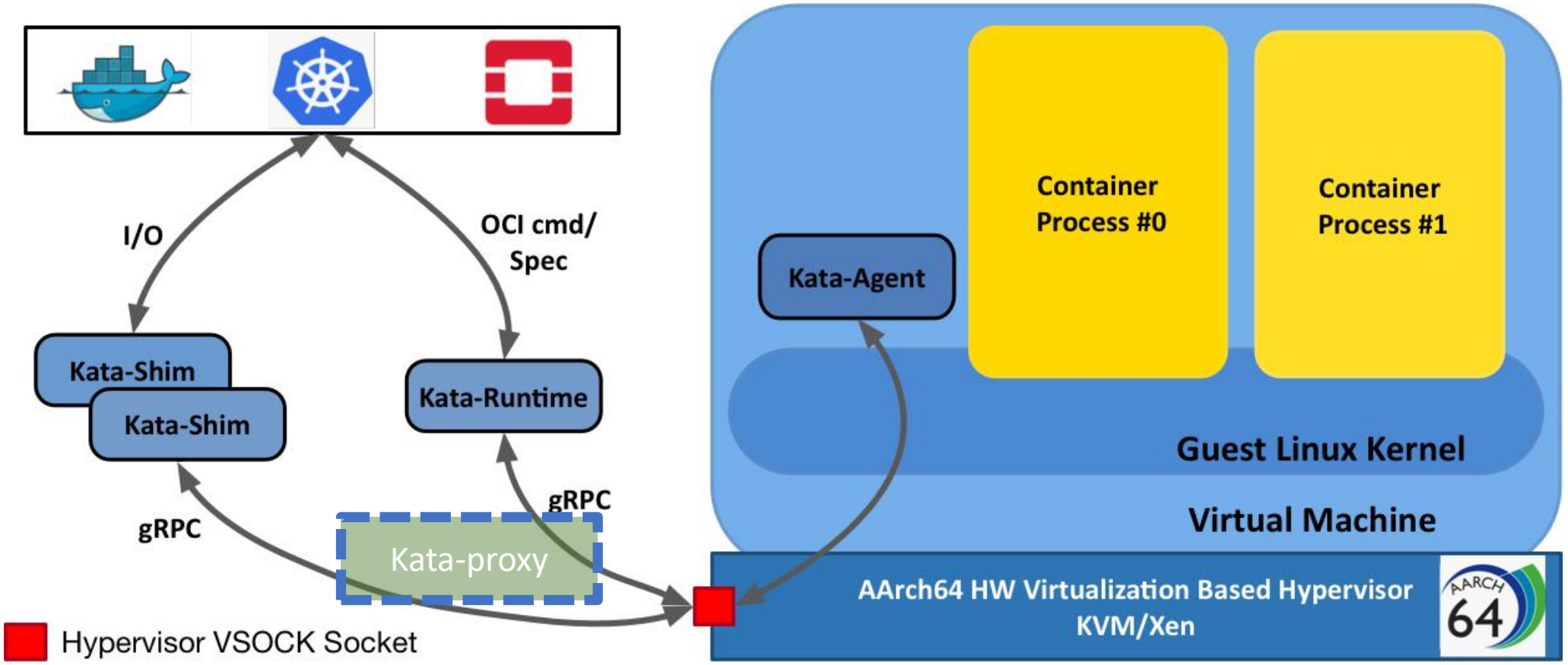
Kata containers performance evaluation and optimization on arm64

Jia He justin.he@arm.com
Arm software engineer

Agenda

- What is Kata containers
- Status update on arm64
- Performance evaluation
- User stories

Kata Containers Architecture Design



Kata containers status on arm64

- Brief summary on overall status of arm64 kata containers
  - **Run smoothly on arm64**
  - **Install kata on arm64**
    - sudo snap install kata-containers
    - Build from source code
  - **Run kata by ctr**
    - sudo ctr image pull docker.io/library/busybox:latest
    - sudo ctr run --runtime io.containerd.run.kata.v2 -t --rm docker.io/library/busybox:latest hello sh
  - **Even supported on Raspberry Pi 4**

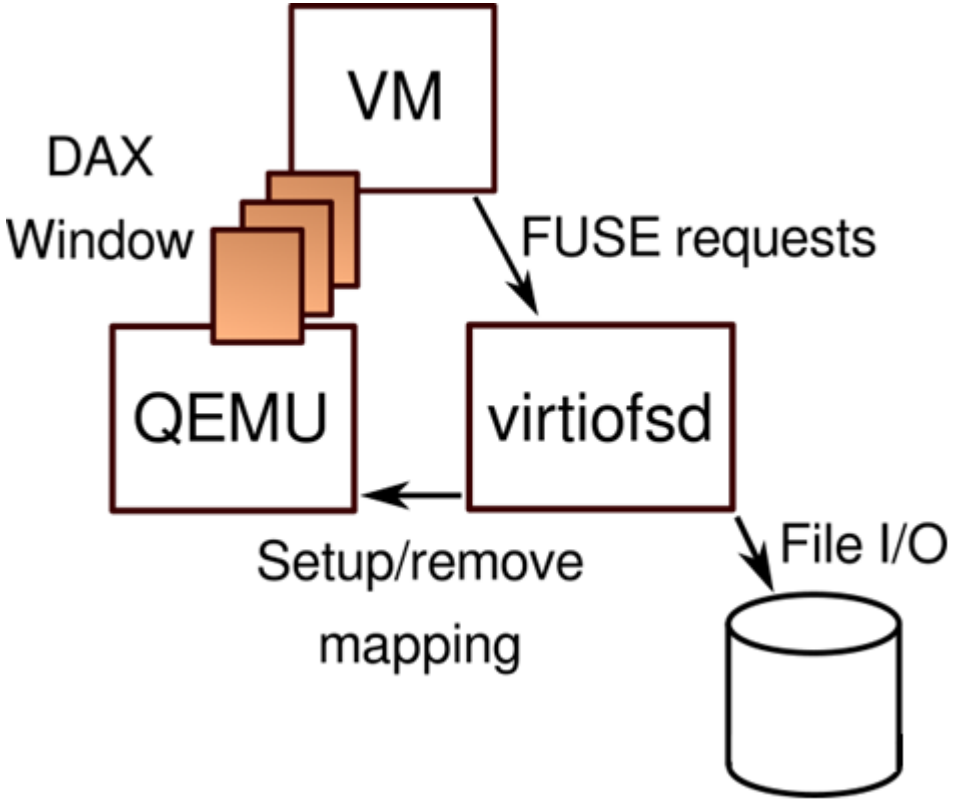
status on arm64 – features comparison

Features	Status on X86	Status on arm64
hypervisor	qemu/fc/clh/acrn	qemu/fc/clh
nvdimm(dax)	Y	Y
virtiofs	qemu/clh	qemu*/clh
vm template	Y	Y
Rust-agent	Y	Y
memory hotplug	Y	Y*
vcpu hotplug	Y	N
Nested kvm	Y	N

* Supported by upstream only, not in Kata repo

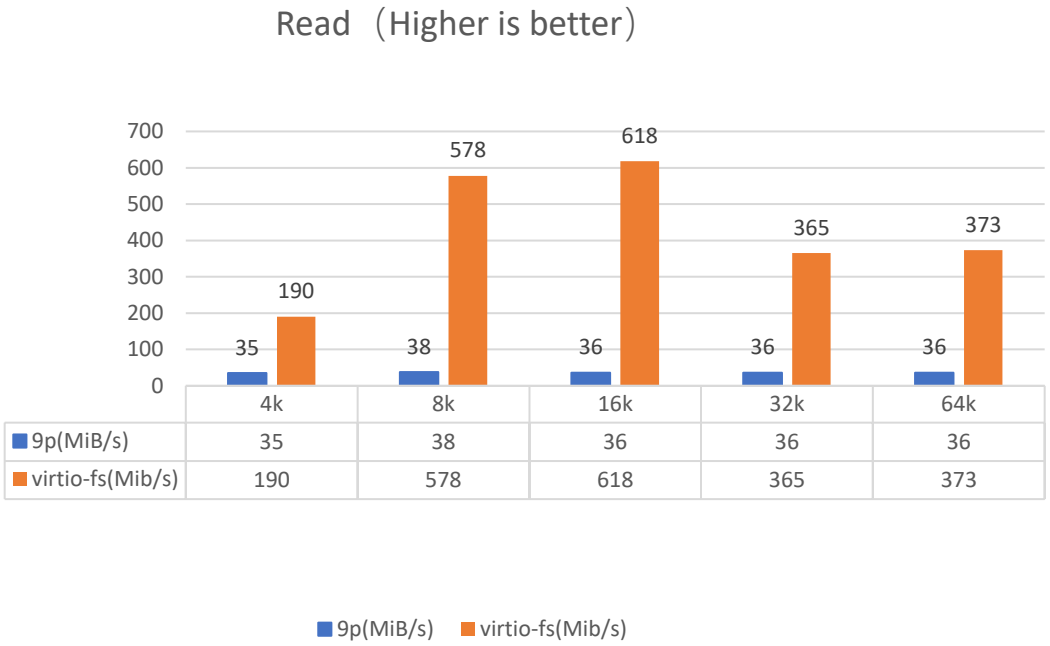
virtio-fs

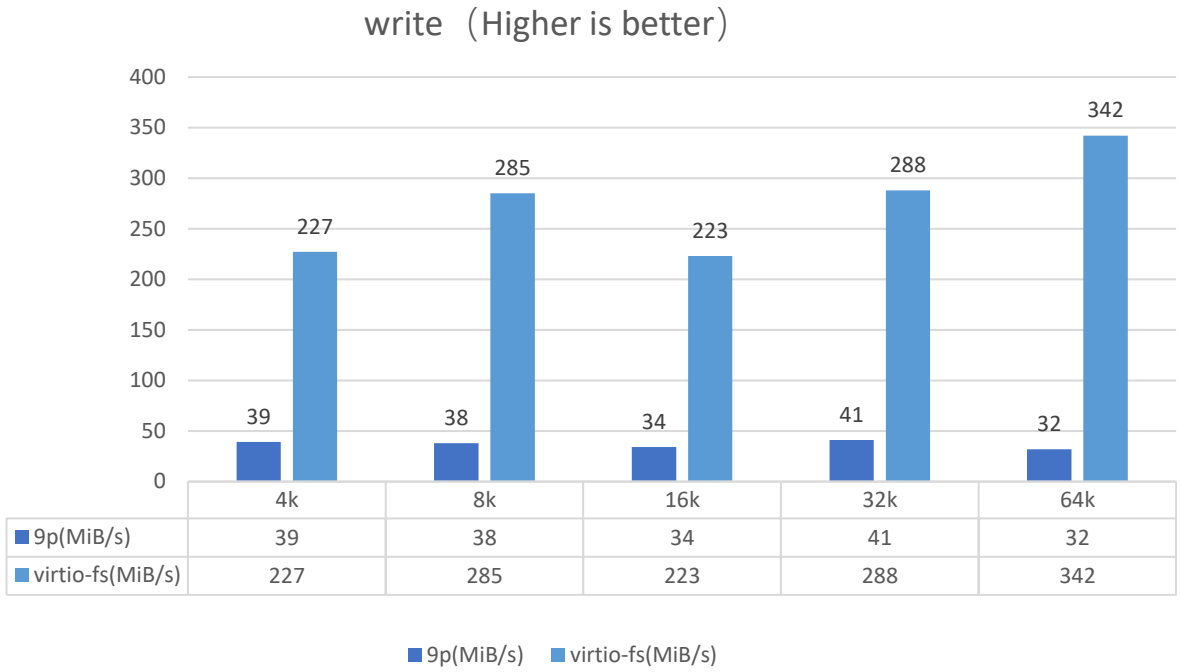
- Virtio-fs is a shared file system that lets virtual machines access a directory tree on the host. Unlike existing approaches, it is designed to offer local file system semantics and performance
  - Base on fuse
  - Independent userspace daemon Virtiofsd
  - DAX, avoid unnecessary VM exit.
- **Limitation**



virtio-fs

- performance test of Virtio-fs compared with 9pfs on Arm (dax enabled)





- mount -t 9p -o trans=virtio,version=9p2000.L,msize=4k hostshare /tmp/host_files
- mount -t virtiofs myfs /mnt

Functional features development

- what Arm container team has done

- Enable the runtime/rust-agent for arm64
- CI maintenance for Kata arm64
- Firecracker/Cloud hypervisor arm64 support on arm64
- Kubernetes integration test

Feature todo

- Memory hotplug
- Cpu hotplug
- Nested virtualization

Performance comparison

- **Boot time**
- **Binary code size**
- **Memory footprint**
- **Hardware/Software setup:**

  Arm64 host: ThunderX2 5.3.0-rc4+ kernel, ubuntu 18.04.4

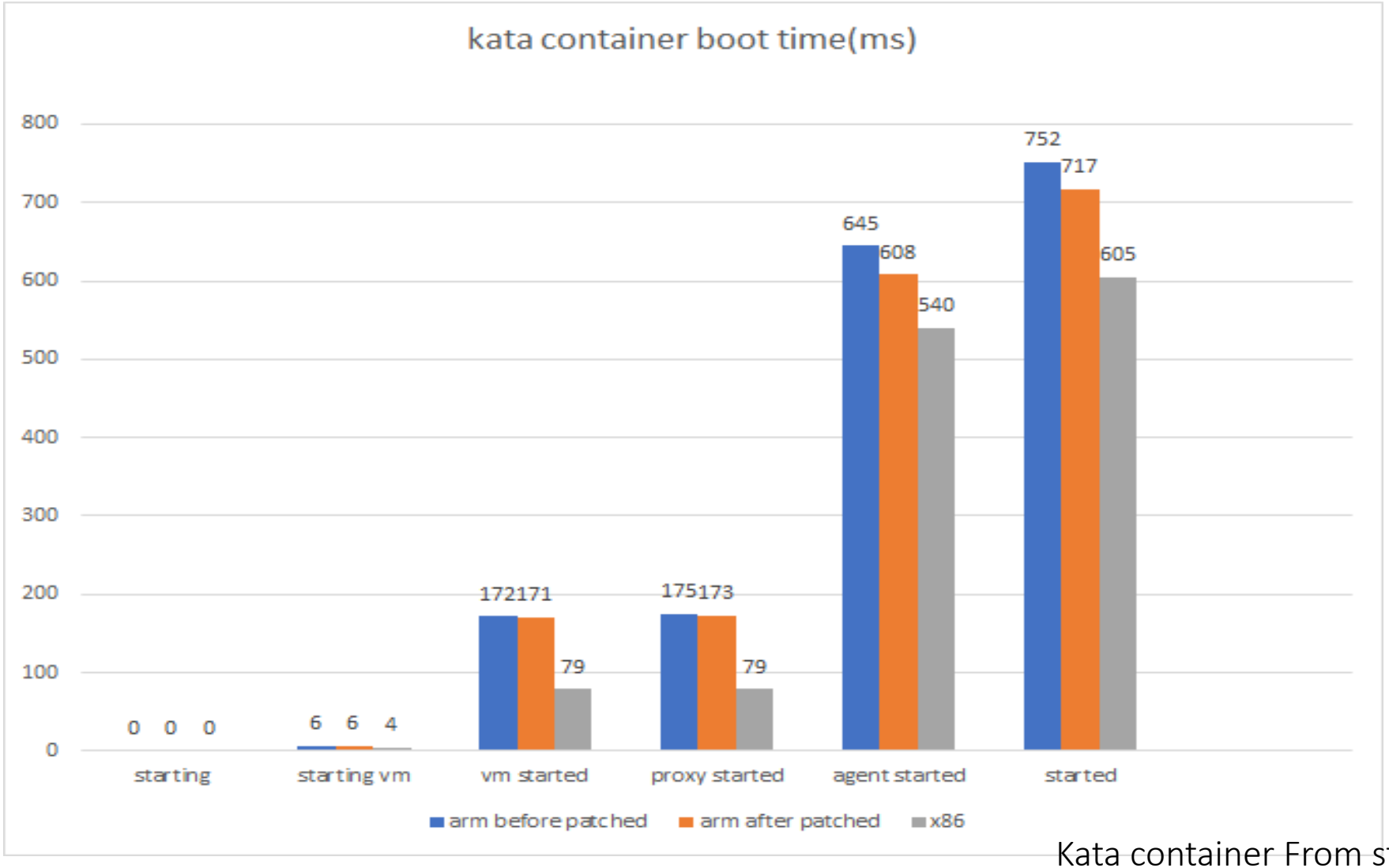
  Qemu: 4.1 from kata upstream

  X86 host: desktop Intel(R) Core(TM) i7-9700 CPU + Ubuntu 19.10

  Qemu from Ubuntu 19.10

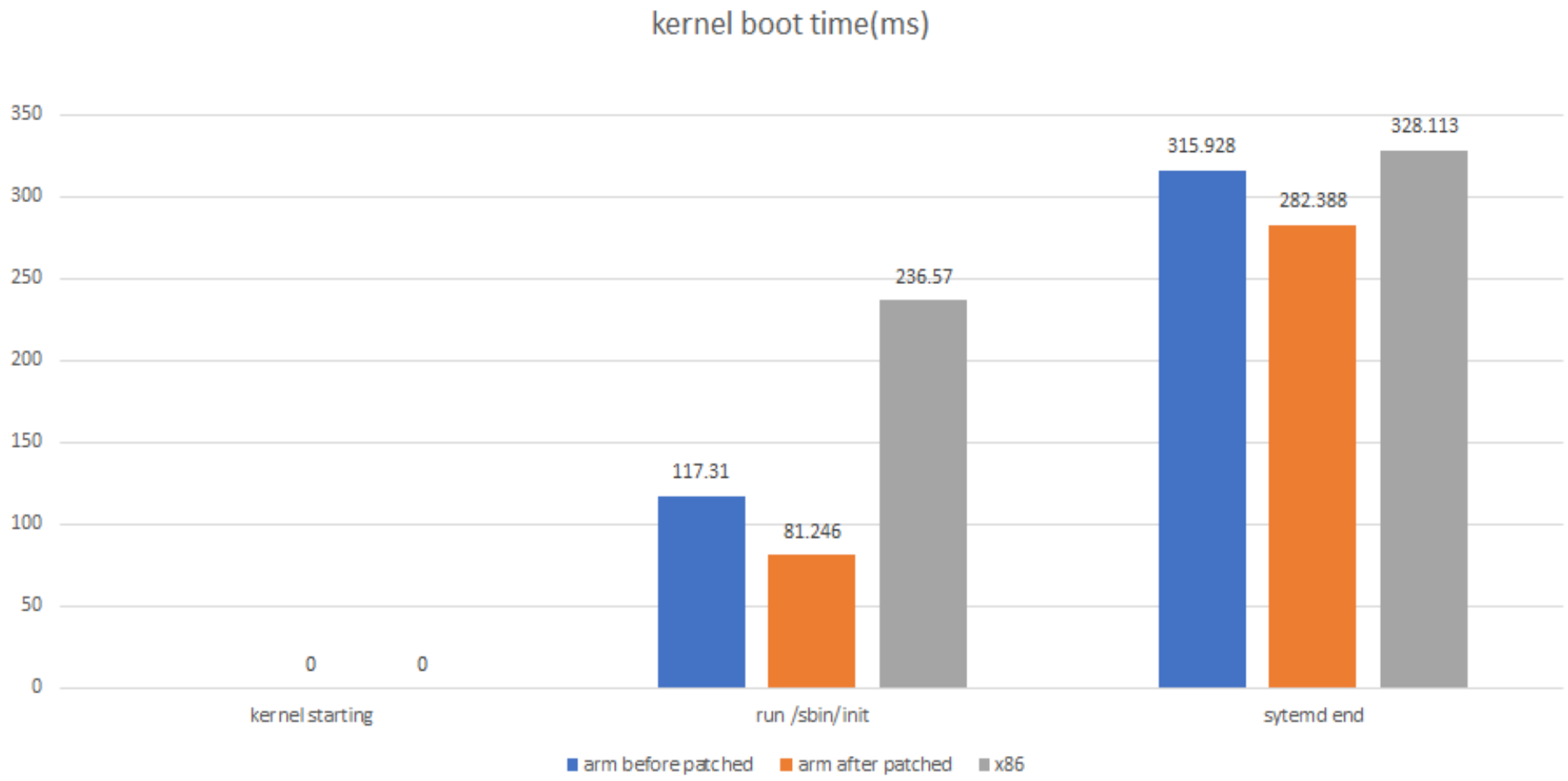
  Kata: latest (2020-July) with default configuration

Performance comparison – **container** boot time



Kata container From starting to being started

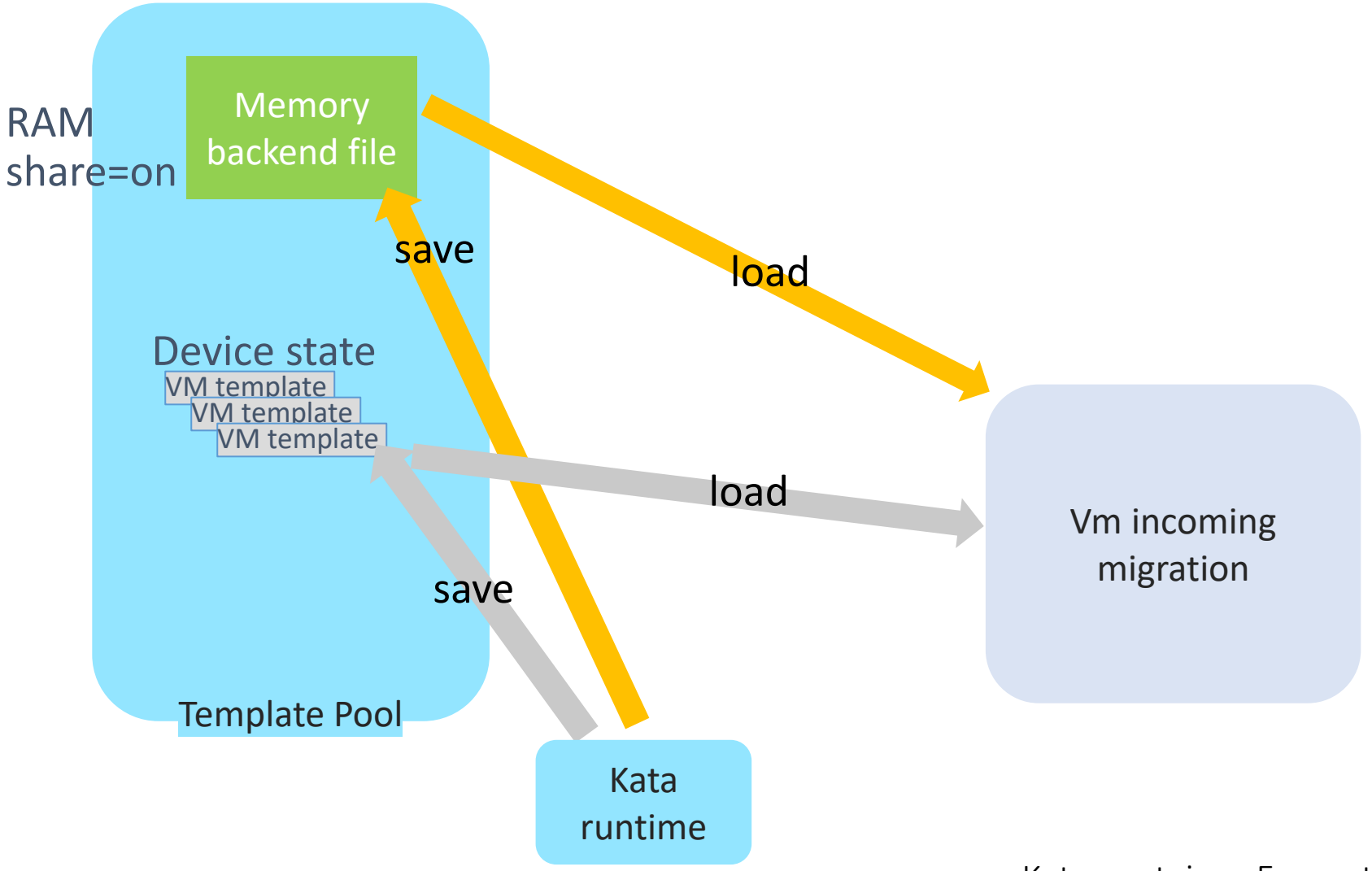
Guest kernel boot time



Tuning items (boot time)

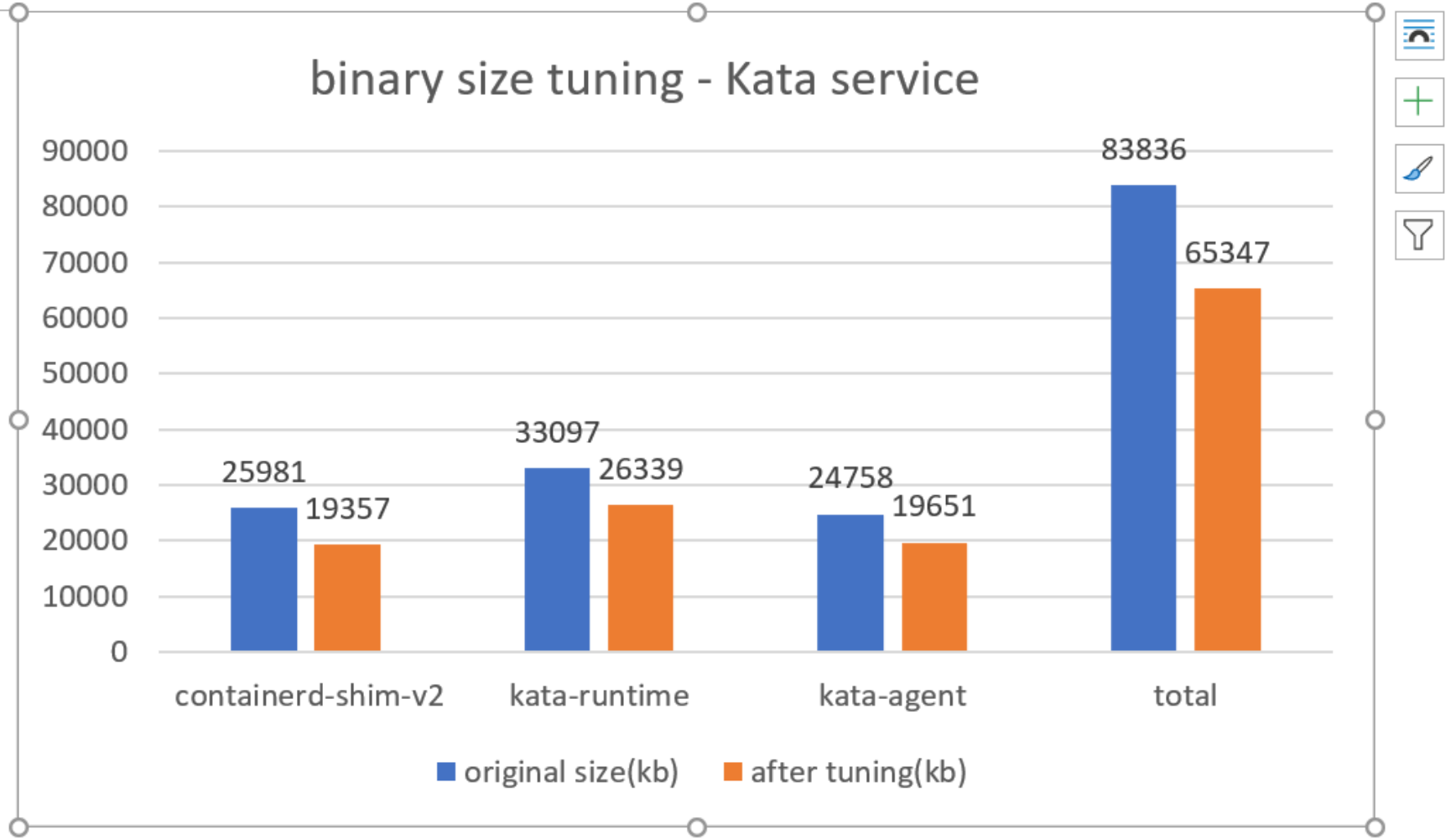
- **Workitems:**
- pmu=off
- scsi scan none
- virtio mmio disable
- **More aggressive: VM template**

VM template



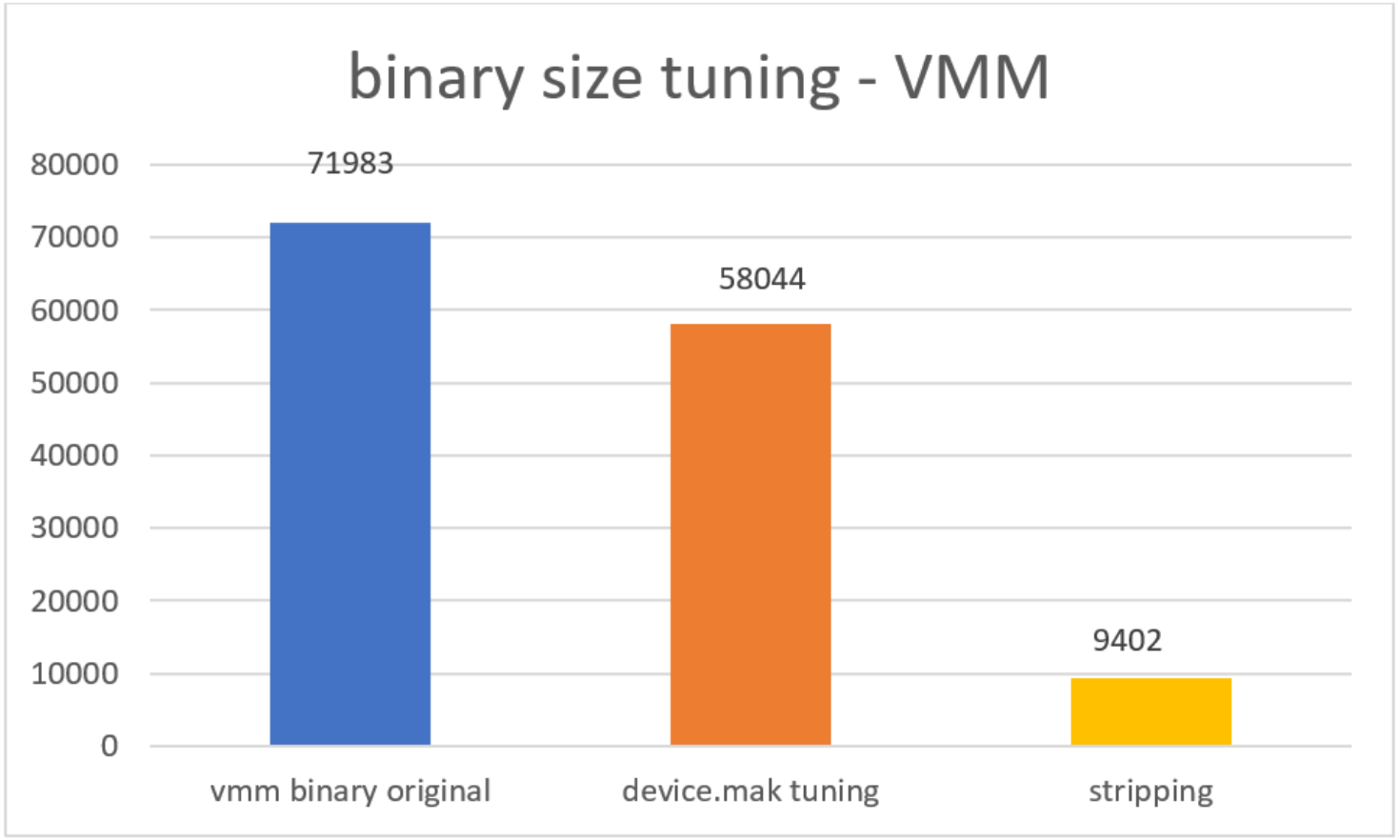
Kata container From starting to being started

Performance comparison – binary size



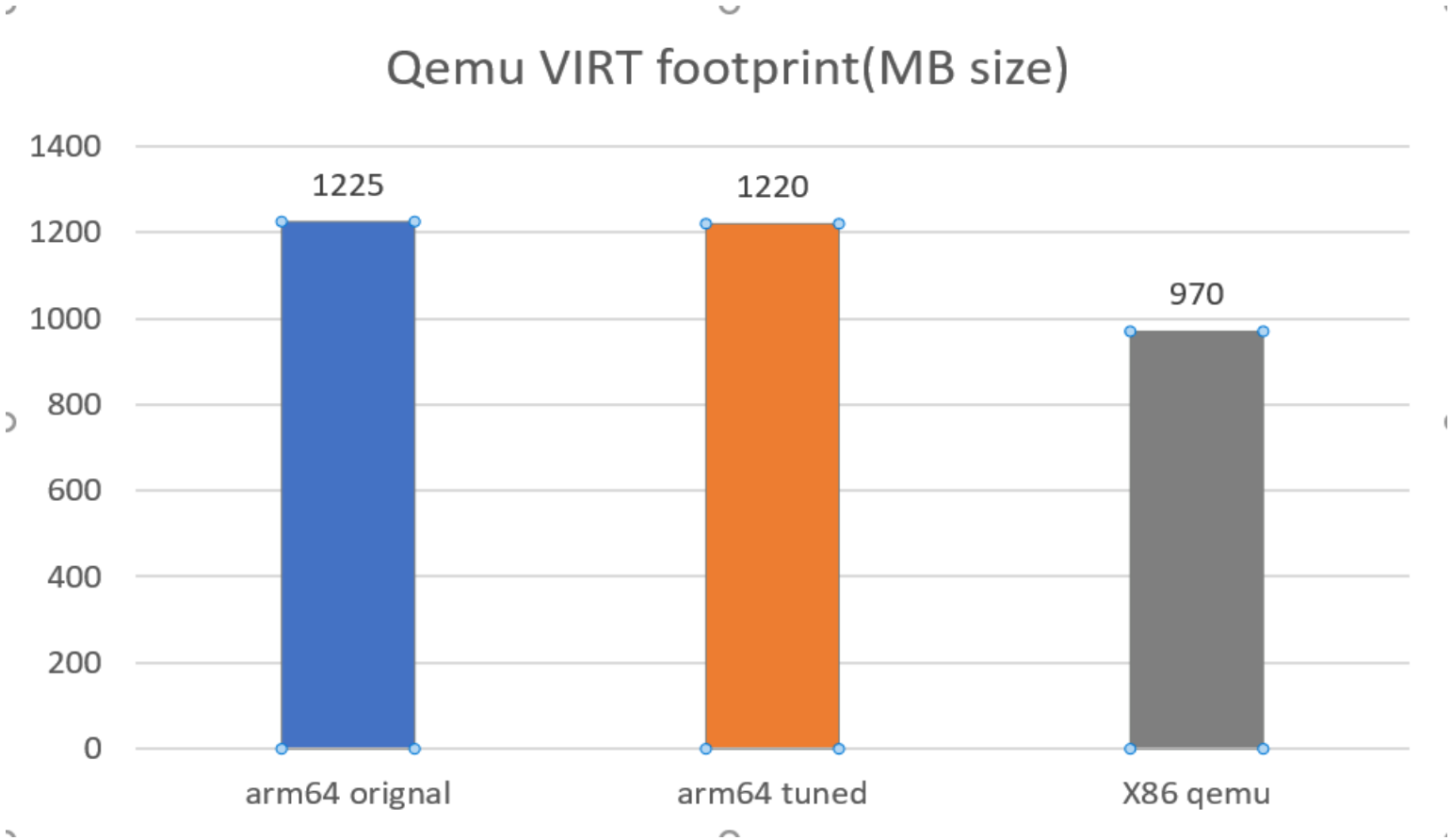
Summary: The total binary size was reduced from 83.8Mbytes to 65.3Mbytes (-22%)

Performance comparison - memory footprint

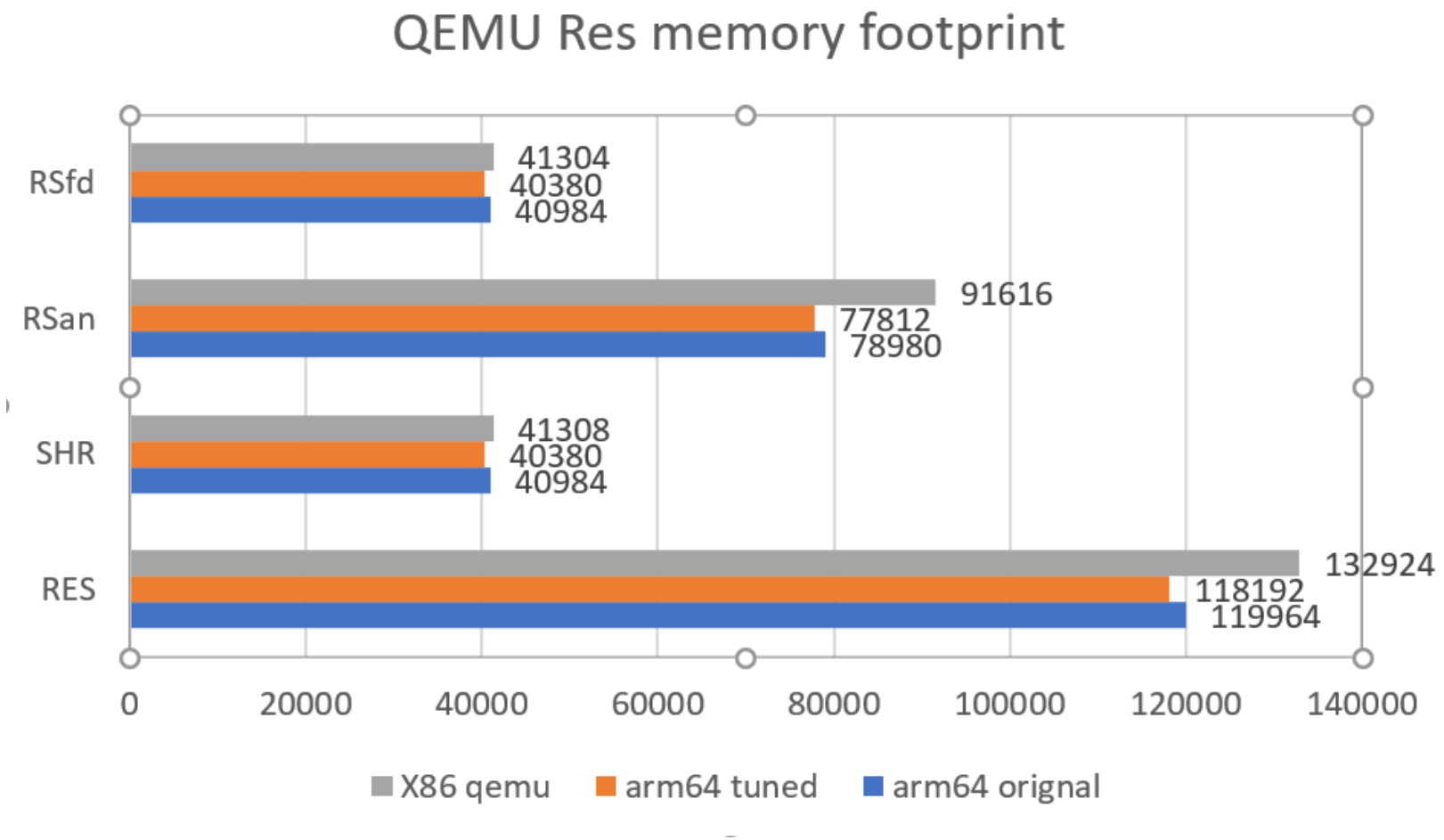


Summary: The vmm binary size was reduced from 71.9 Mbytes to 9.4 Mbytes (-86%)

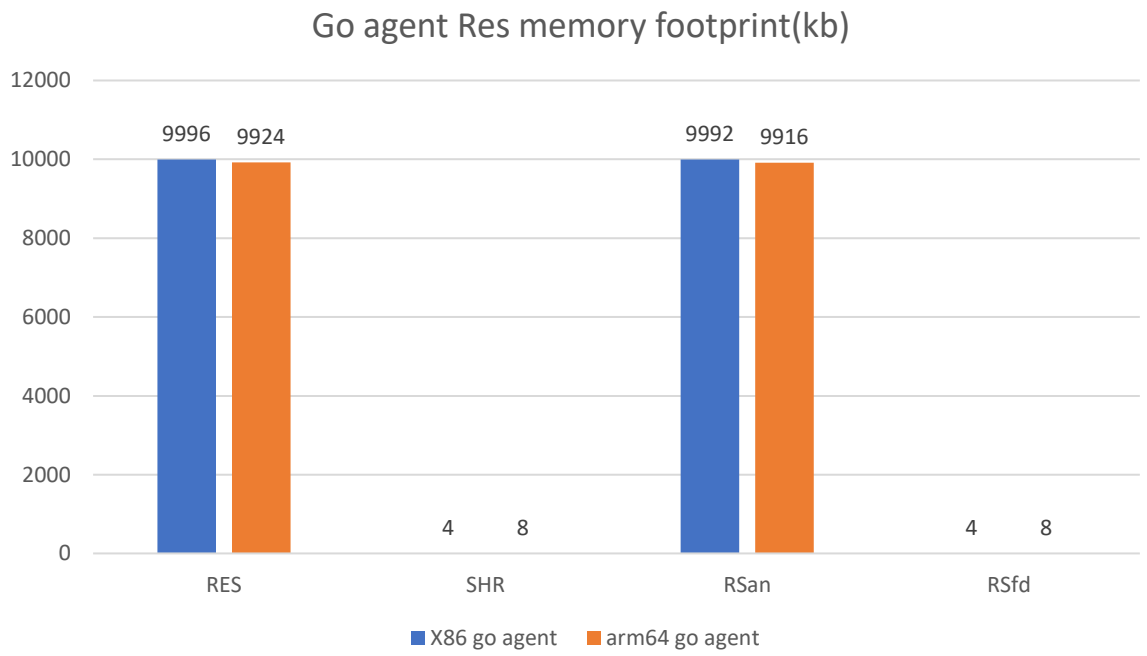
Performance comparison - memory footprint

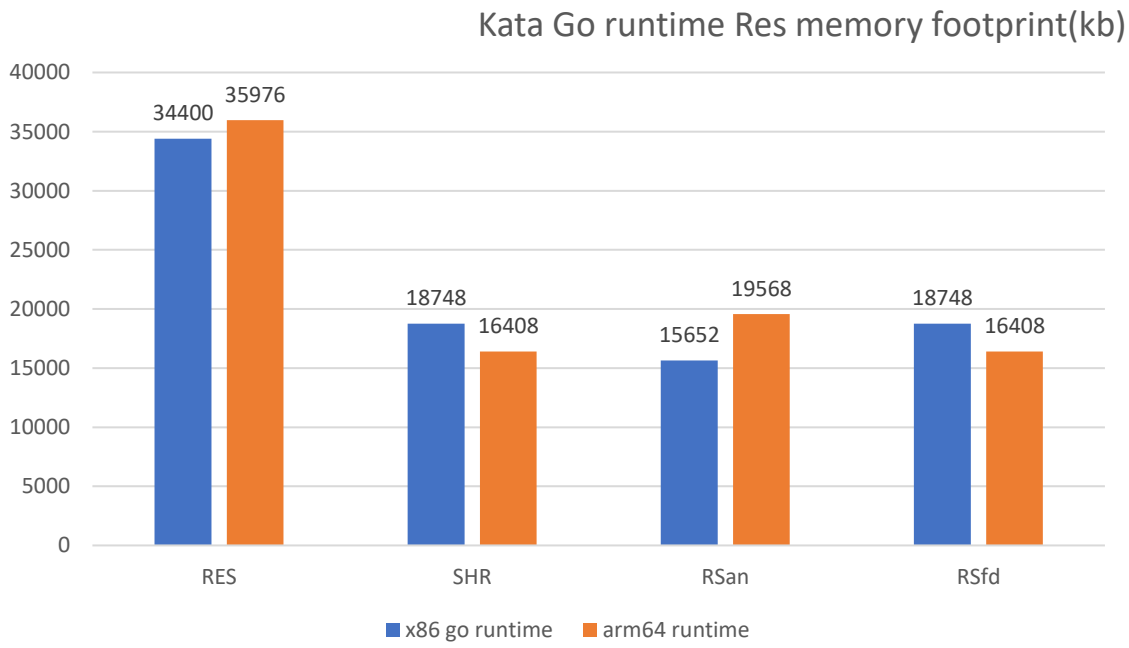


Performance comparison - memory footprint



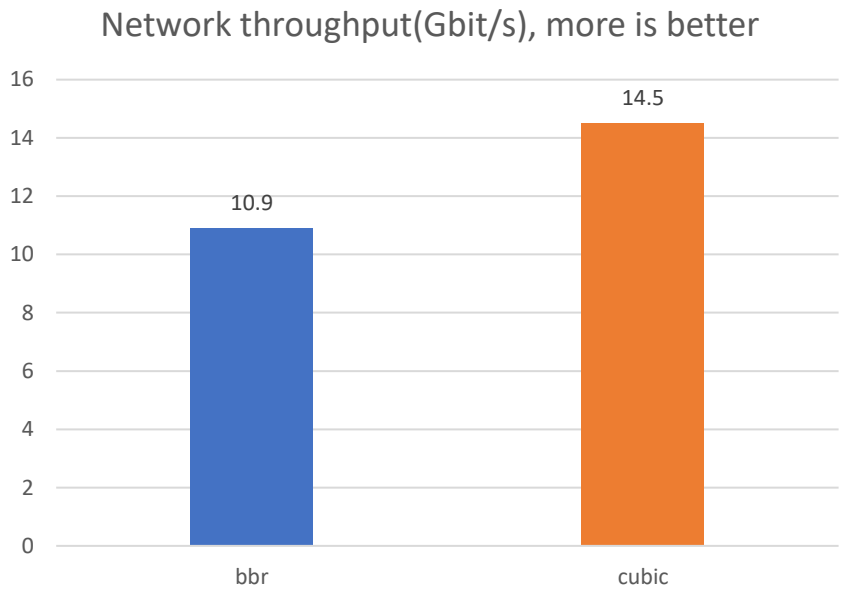
Kata go binary memory footprint comparison





Network throughput

what Arm container team has done to tune the performance



- Server: iperf3 -s
- Client: iperf3 -c $ip -i 2 -t 30

Performance tuning items

what Arm container team has done to tune the performance

- VM template support for arm64
- virtiofs/dax enablement and bugfix
- pmem (nvdimm) support and bugfix
- Change the algorithm to Cubic for Kata

User Stories - Kata in Chinese internet giants

- Baidu's journey to offer AI Cloud and Edge Computing services at massive scale by taking advantage of Kubernetes, Kata Containers
- ECS Bare Mental Instances + Kubernetes as Serverless infrastructure, with Kata Containers as container runtime in Alibaba cloud

KEEP CLOUD NATIVE
EVERYWHERE
KubeCon
CloudNativeCon
North America 2020
Virtual
HELM
KV