

京东基于openEuler的混部能力共建

汇报人：京东云事业部 PaaS产研部 张伟

京东科技信息技术有限公司

2023年12月



目录

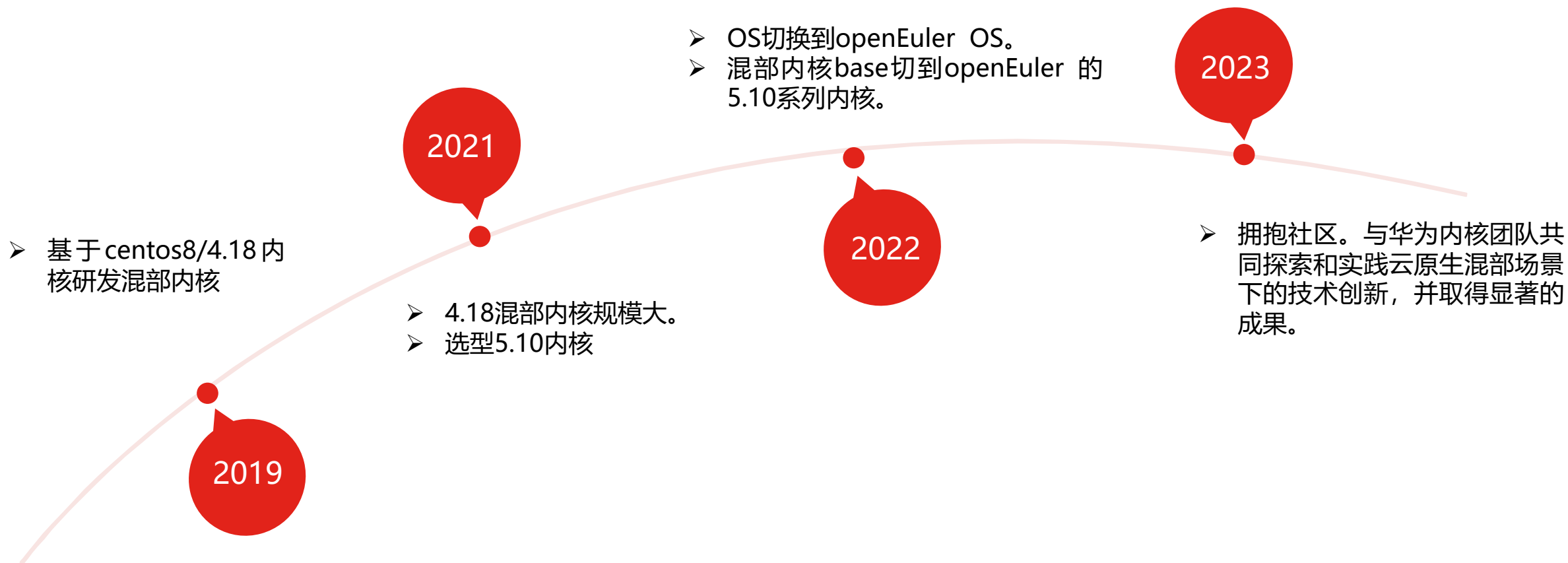
一、京东混部内核历程

二、社区共建-容器内存潮汐

三、社区共建-容器SMT驱逐

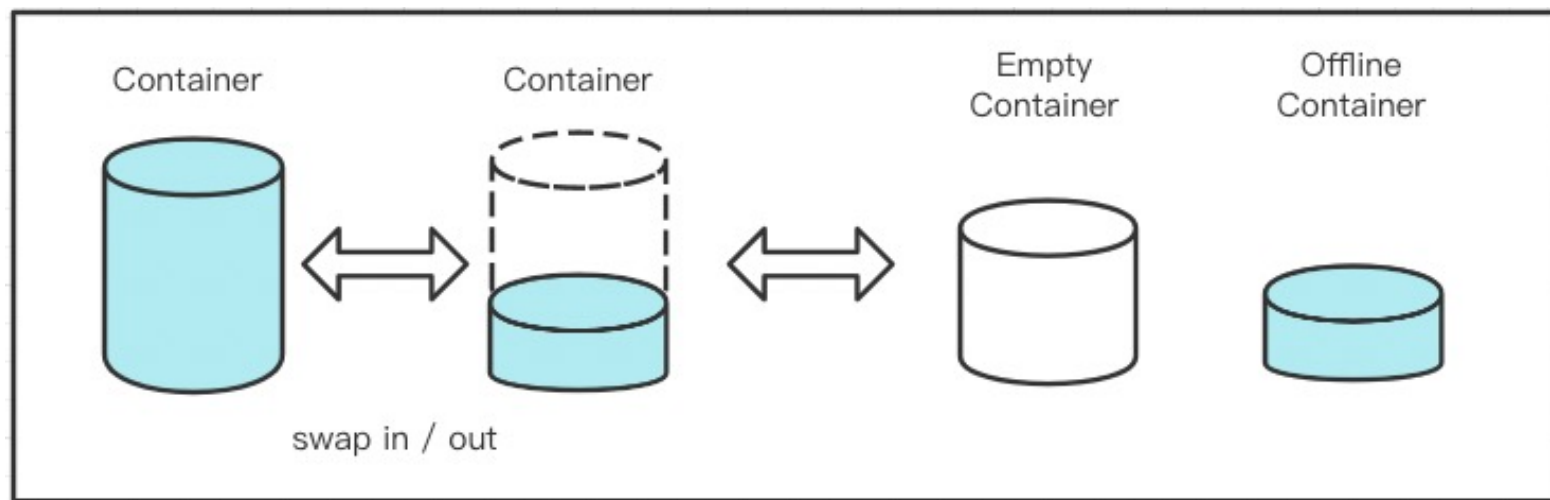
四、社区共建-潮汐affinity

京东混部内核历程



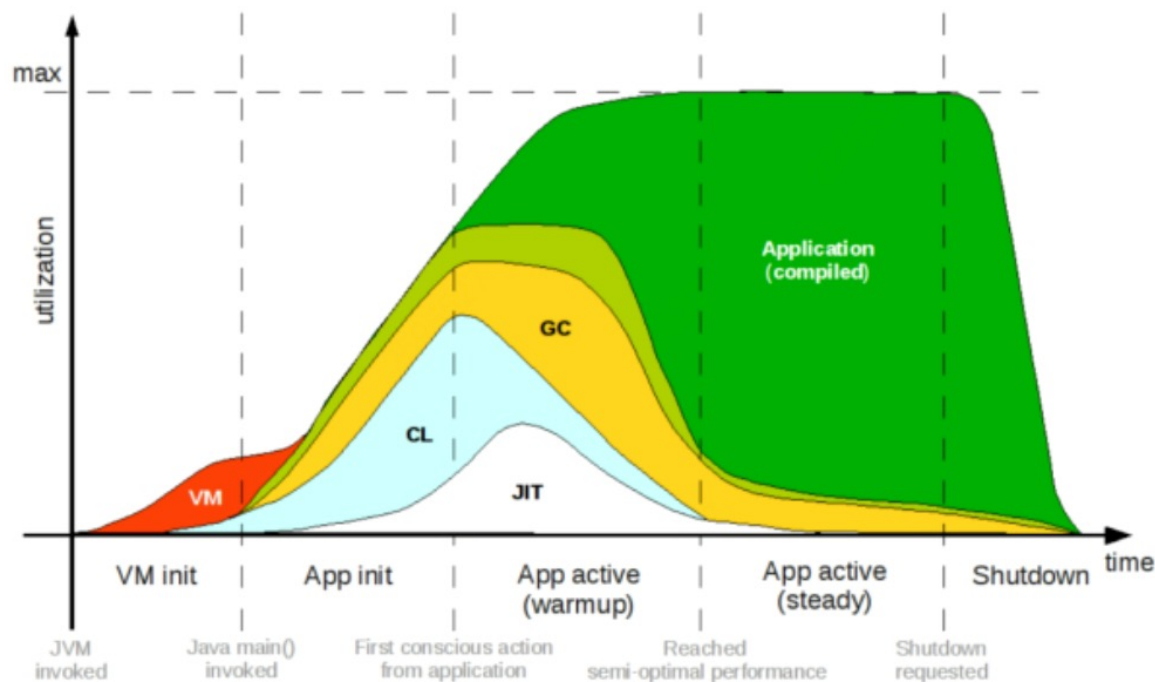
• 场景

- 降低夜间应用成本
- 黄金流服务在两次大促之间可保有资源
- 大促前做好热备（提前创建好容器并降成低规格），需要时可快速恢复

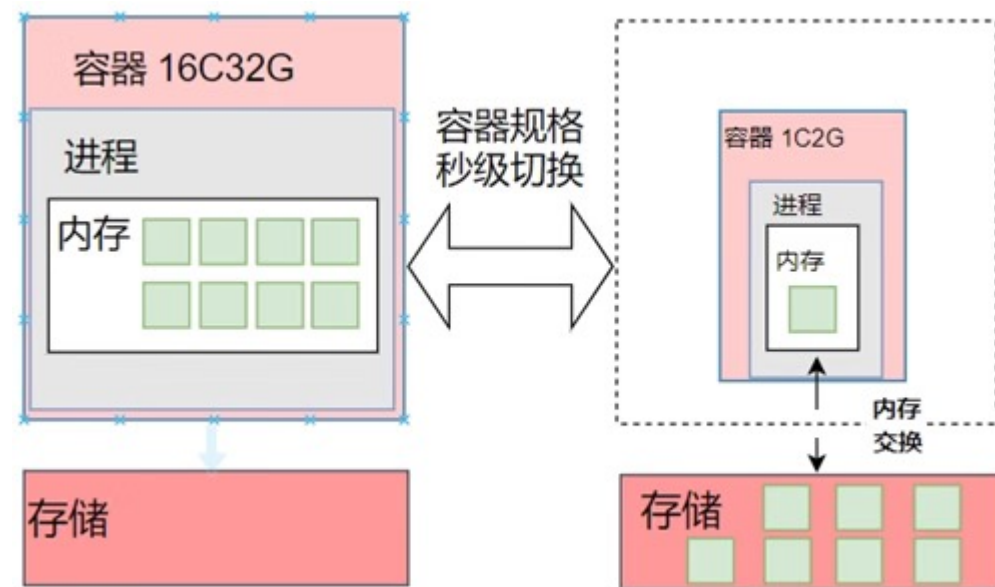
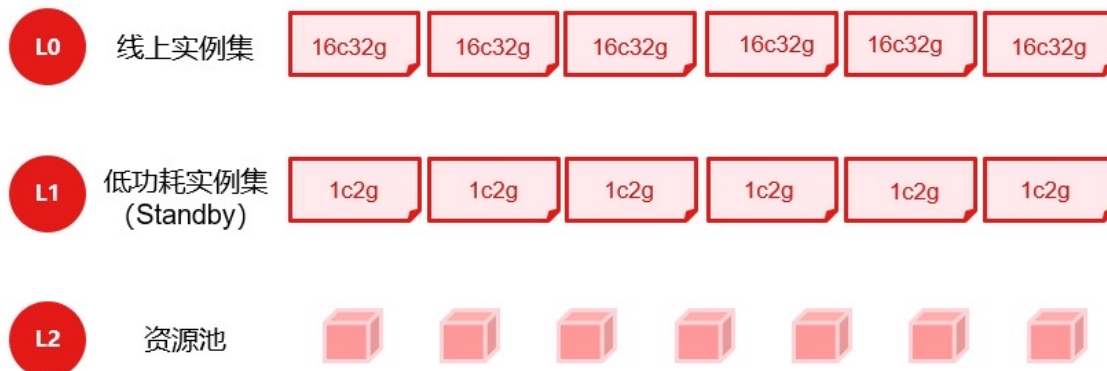


openEuler社区共建：容器内存潮汐 (2)

- ◆ 低规格热备，快速规格切换
- ◆ 核心技术
 - 容器级内存换入换出(memcg swap qos)



典型Java应用生命周期

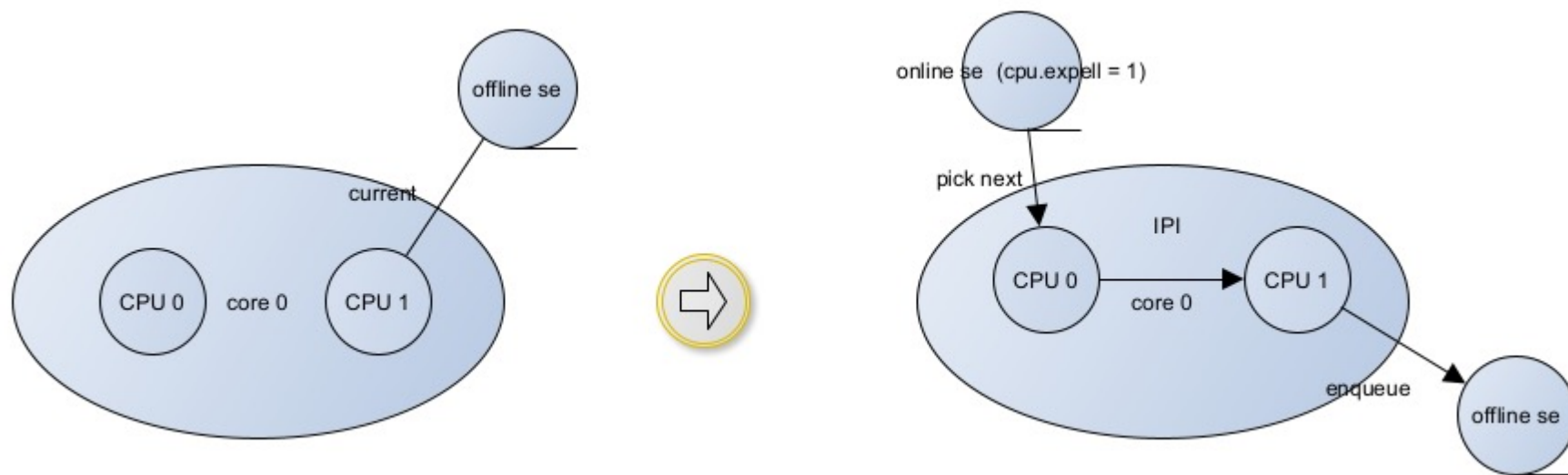


• 场景

- 离线和在线任务在同core的不同cpu上运行时，离线和在线之间会产生竞争。在线业务需要具有SMT驱逐能力。
- 从实际需求看，并不是所有的在线任务都需要启用驱逐能力。按需启用的同时，也能提升整机利用率。

openEuler Cpu cgroup SMT 驱逐特性是基于 qos sched 框架，进一步实现了容器级的超线程隔离调度，保障特定高优先级业务不会受到来自同core 的低优先级任务干扰。

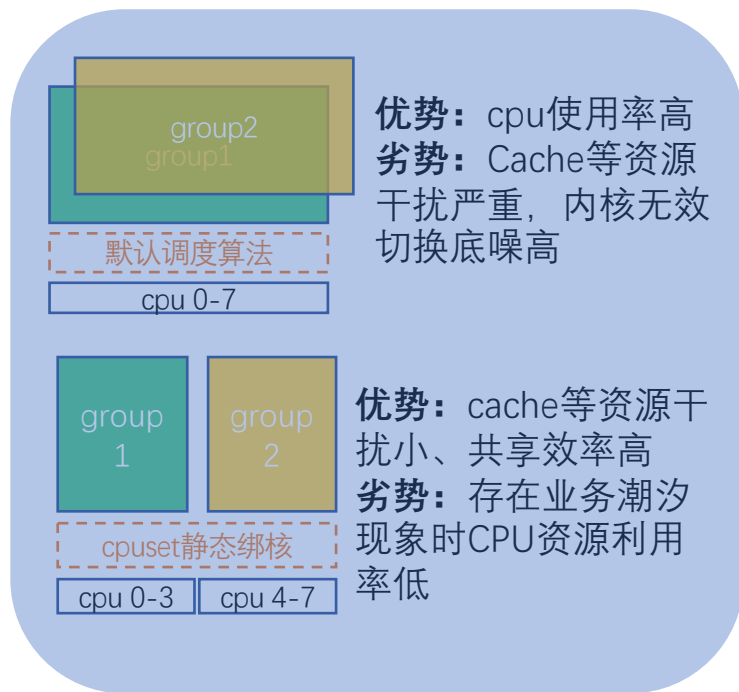
主动驱逐示意图如下：



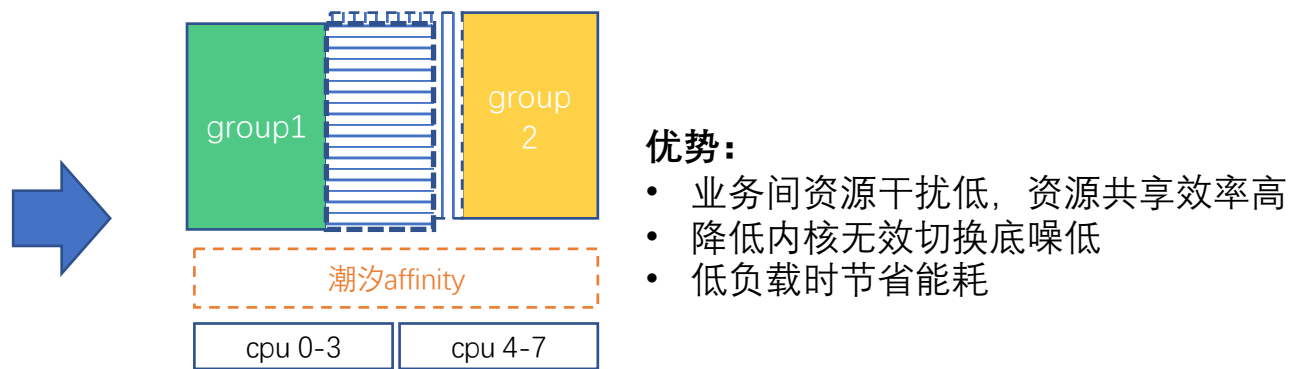
场景

感知业务负载动态调整业务CPU亲和性，当业务负载低时，约束在指定cpu子集上处理，增强资源的局部性；当业务负载高时，突破范围限制，通过增加cpu核的供给提高业务的QOS。

传统方案缺陷



技术方案



谢 谢!

谢 谢!

谢 谢!