

SACC 2014中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2014

发现架构之美

腾讯云数据中心操作系统Gaia（盖娅） ——大数据应用坚实基础



罗韩梅(mavisluo)

海量用户产生海量数据



月活跃用户8.3亿，最高同时在线2.1亿；
在线人际关系链超X000亿；



月活跃4.4亿；
日均消息量超X0亿；



月活跃用户数6.5亿；
日均相册上传超过X亿，日写操作总数过X0亿；



腾讯游戏月活跃用户超X亿；
手机游戏月活跃用户超X亿；



日均pv超X亿，手机侧近超X亿；
日均uv超X千万，手机侧超X千万；

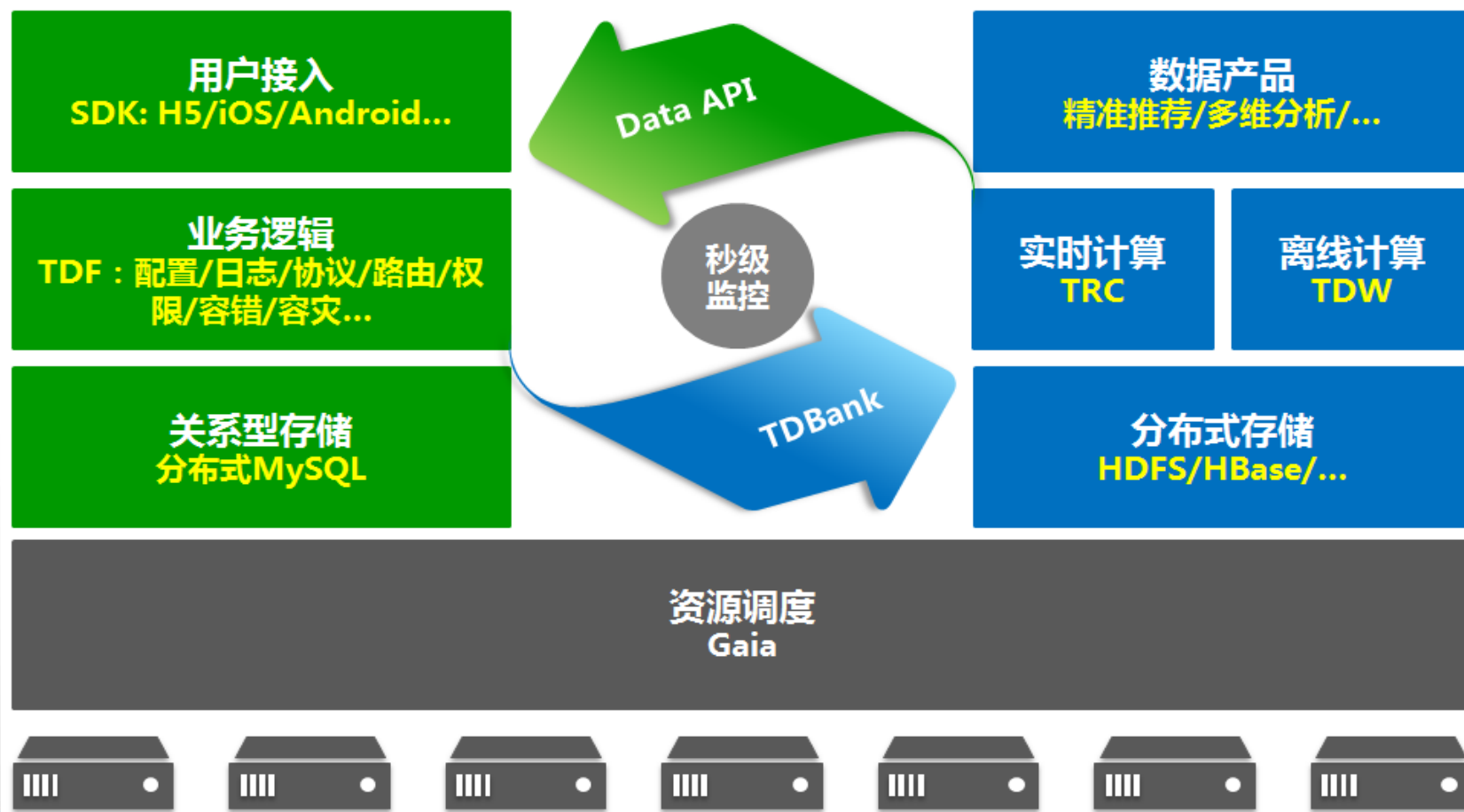
部分数据来自腾讯2014第二季度综合业绩报告

海量数据造就强大技术支撑



- 腾讯拥有业界顶级的分布式计算集群
 - 实时接入数据**万亿**条/天
 - 单集群规模最大**6000**台，日均Job数**120w+**，支持多存储引擎和多并行计算框架
 - 统一资源管理和调度，磁盘、CPU、网络等利用率**85%+**

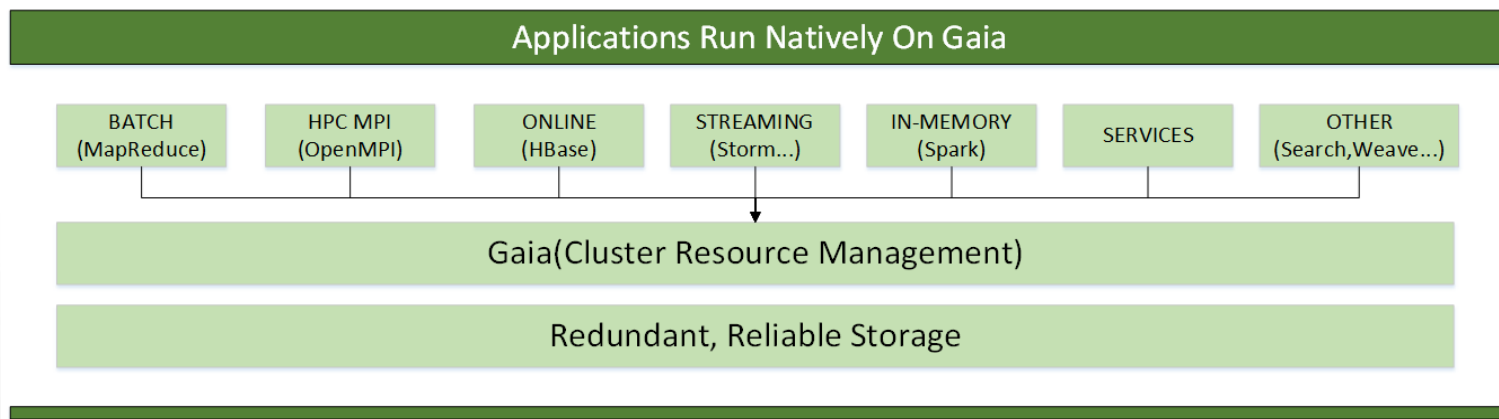
腾讯数据平台整体架构



Gaia as Cluster Operation System

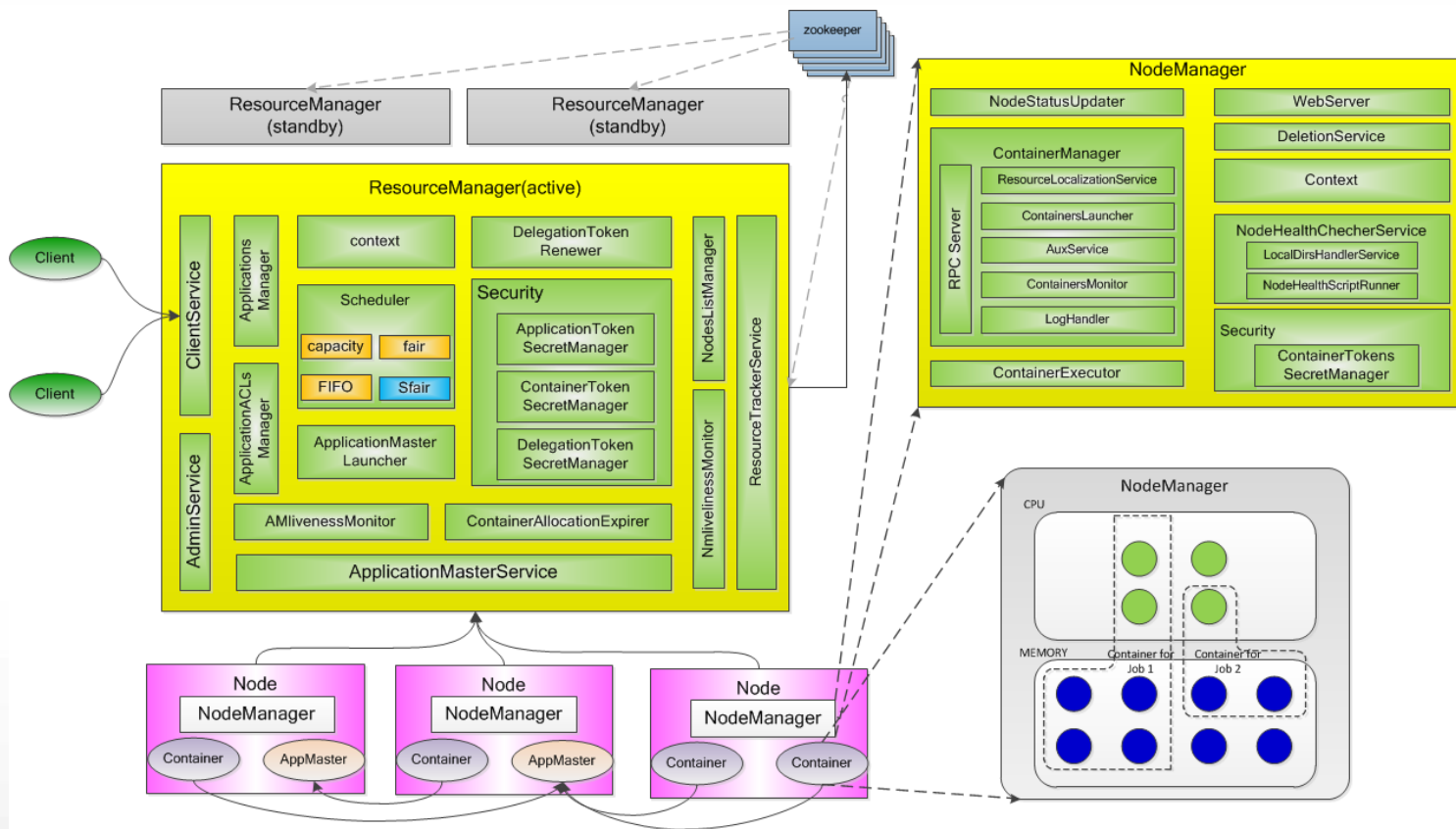


Gaia (盖娅)：希腊神话中的大地之神，是众神之母，所有神灵中德高望重的显赫之神。Gaia以后可以承载各种编程框架、各种应用，是个统一的资源管理调度系统——各种业务都植根于“大地”之上。



Gaia — 基于Yarn的通用资源调度平台，提供高并发任务调度和资源管理，实现集群资源共享，可伸缩性和可靠性，**不仅可以为MR等离线业务提供服务，还可以支持实时计算，甚至在线service业务。**

为什么选择Yarn



可扩展、高可用、低耦合、开放性

Gaia自研调度器sfair scheduler

◆ 挑战一. 可扩展性差 **4000 vs 8800 !**

◆ 挑战二. 高作业并发量，但调度吞吐率低

tdw集群要承担公司各个业务线的数据任务，任务量并发高（4k），且持续增长中。

◆ 挑战三. HA考虑不周，影响可用性

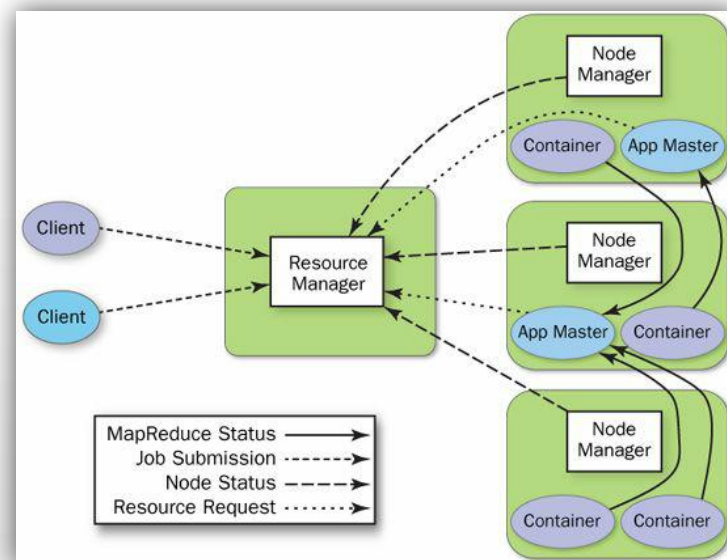
AM意外挂掉，整个application attempt失败。

◆ 挑战四. 资源利用率低

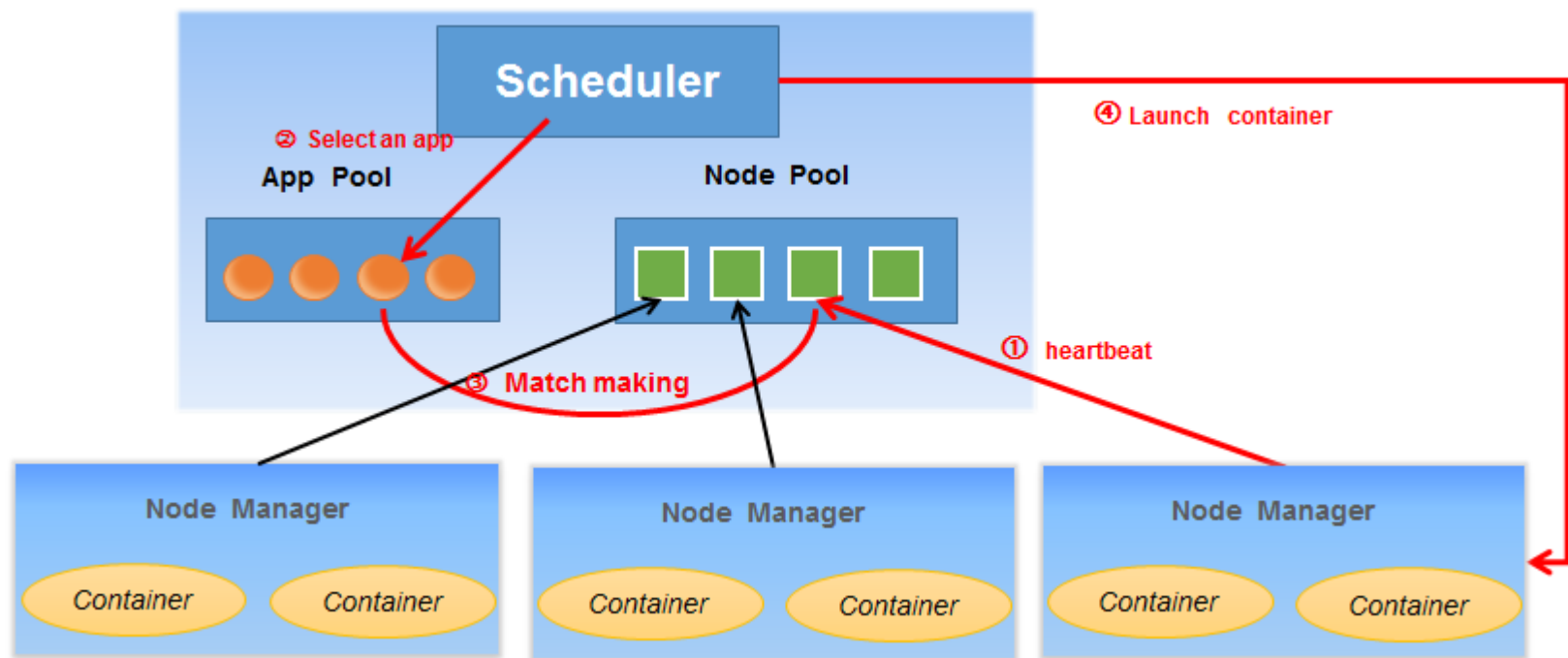
资源利用率低，提高了成本。

◆ 挑战五. Yarn不成熟，bug较多

Sfair
Scalable Fair Scheduler

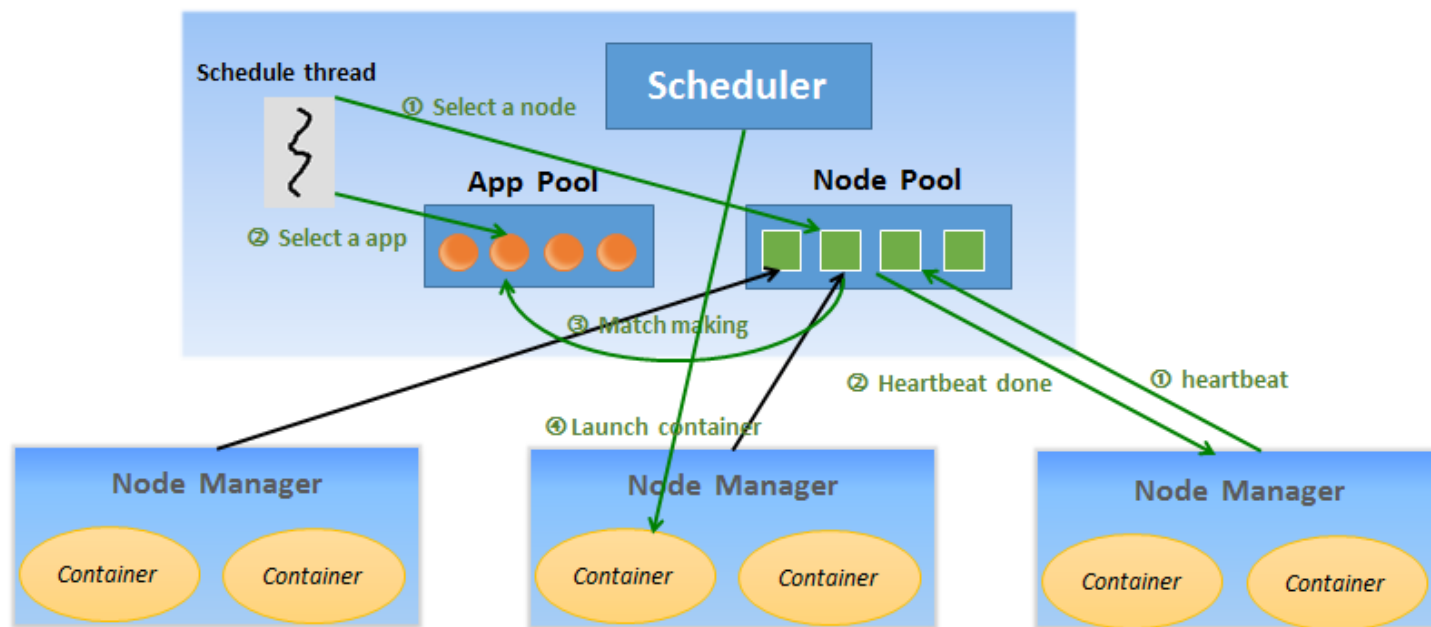


Gaia自研调度器sfair scheduler



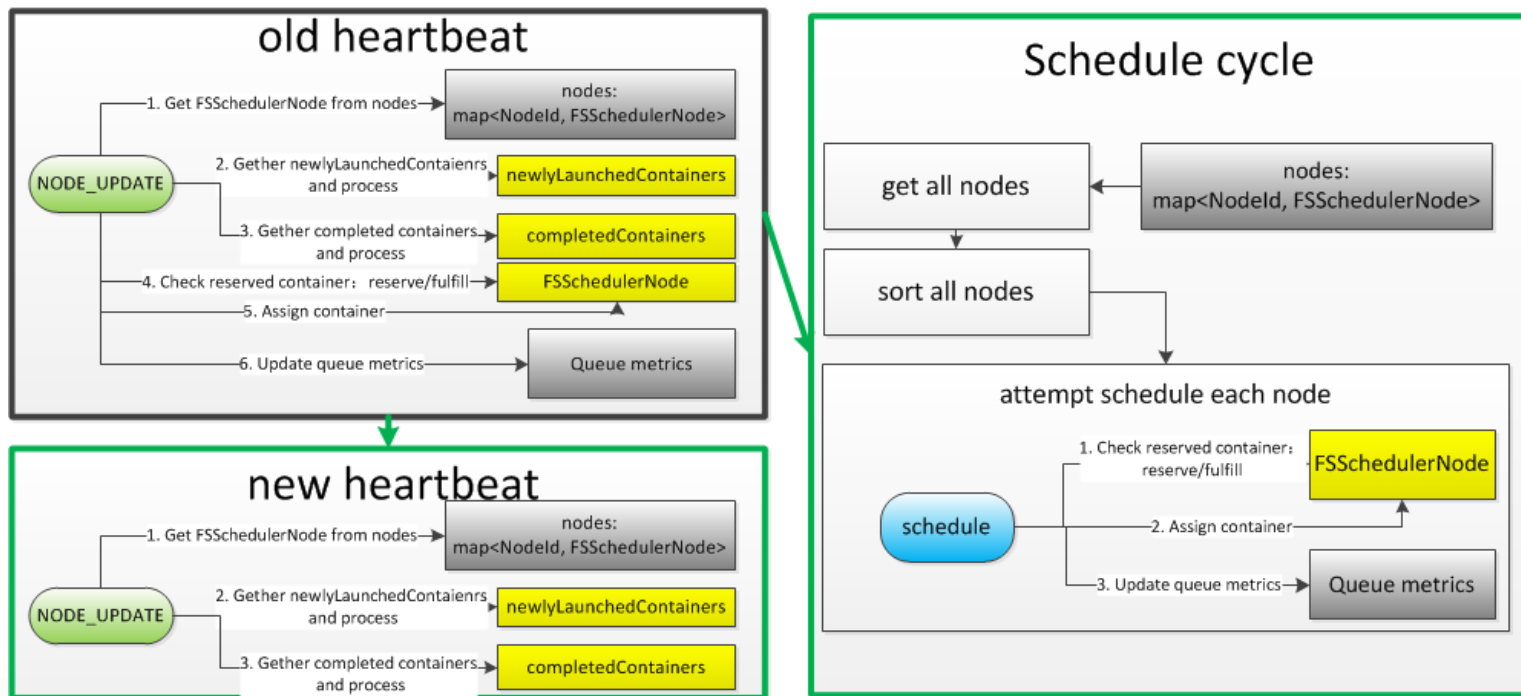
心跳处理过重，严重影响Scalability；
调度吞吐率过低，无法保证高并发作业；
调度没有对集群整体的考虑，资源分配不均衡；

Gaia自研调度器sfair scheduler



- ① 调度不再由心跳触发；
- ② 引入单独线程做周期性调度；
- ③ 心跳中只做RM和NM的信息交互等轻量处理；
- ④ 调度效率对集群规模不那么敏感

Gaia自研调度器sfair scheduler



解耦和后带来的优势：

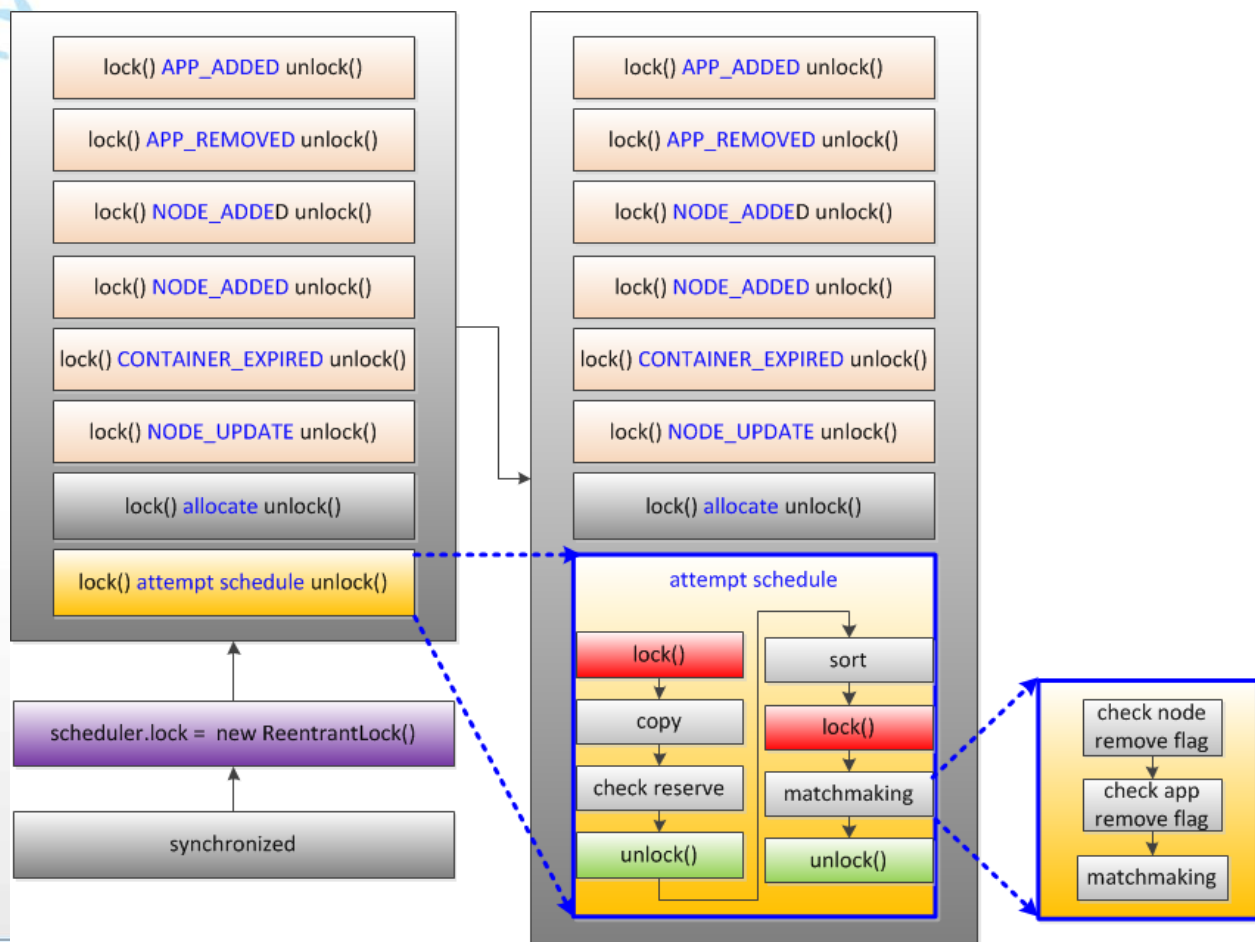
- 心跳处理变得轻量，不再是大规模集群可扩展性的瓶颈: 10ms→1ms
- 调度不被动的等待心跳的到来，在调度周期内执行批量调度，提高了调度的吞吐。

解耦和后带来的问题：

- 调度周期内会长期持有大锁，造成eventDispatcher中其他事件处理耗时增加：
- **3349ms→329328 ms** (add 1000 app)

Gaia自研调度器sfair scheduler

◆优化多线程间同步



◆优化多线程同步：

- ① 优化attempt schedule 中的锁：sort是针对于copy之后app，而sort占用调度的时间又过久，去掉sort过程的持锁
- ② 增加removed 标记位
- ③ 329328 ms-> 3926ms

Gaia自研调度器sfair scheduler

◆优化队列和作业排序方式

$$O(n * \lg n) \rightarrow O(n + k * \lg n)$$

其中:

n:app/queue个数

k:一次调度平均要遍历的次数

K(n=1000)	SortAndWalk		HeapifyAndWalk	
	time	compare	time	compare
1000	136	123565	46	43280
100	130	123665	24	21224
10	134	123439	20	18875
2	141	123631	21	18828

Gaia自研调度器sfair scheduler

◆ 降低调度开销

- 增加对queue/app的资源需求统计;
- 执行调度时, 没有资源需求的queue/app不参与调度。

◆ 提升调度公平性

- 对node也做全局考虑, 不再依赖心跳的随机性。

◆ 减少不必要的对象创建

- 避免每次比较两个app时都要分配新对象, 优化后一次调度所花的时间大概减少了5%。另外, 对GC的压力减小了50%。

◆ 优化调度参数的计算

- 将遍历和累加, 改为变量维护, 需要时直接读取。这类调度参数, 每秒需要读取千万次。

Gaia自研调度器sfair scheduler

◆ 心跳处理与调度解耦和，提升可扩展性

- 现网：将单集群规模扩展至**6k**（年底**8800**）个节点（业内最大集群），作业最大并发已达**4k**，日运行作业数**120w**，task数**7500w**，调度**1250**个资源池；

◆ 控制多线程间的同步

- 优化线程间锁，将调度线程持锁时间减少**70%**；

◆ 优化队列和作业排序方式

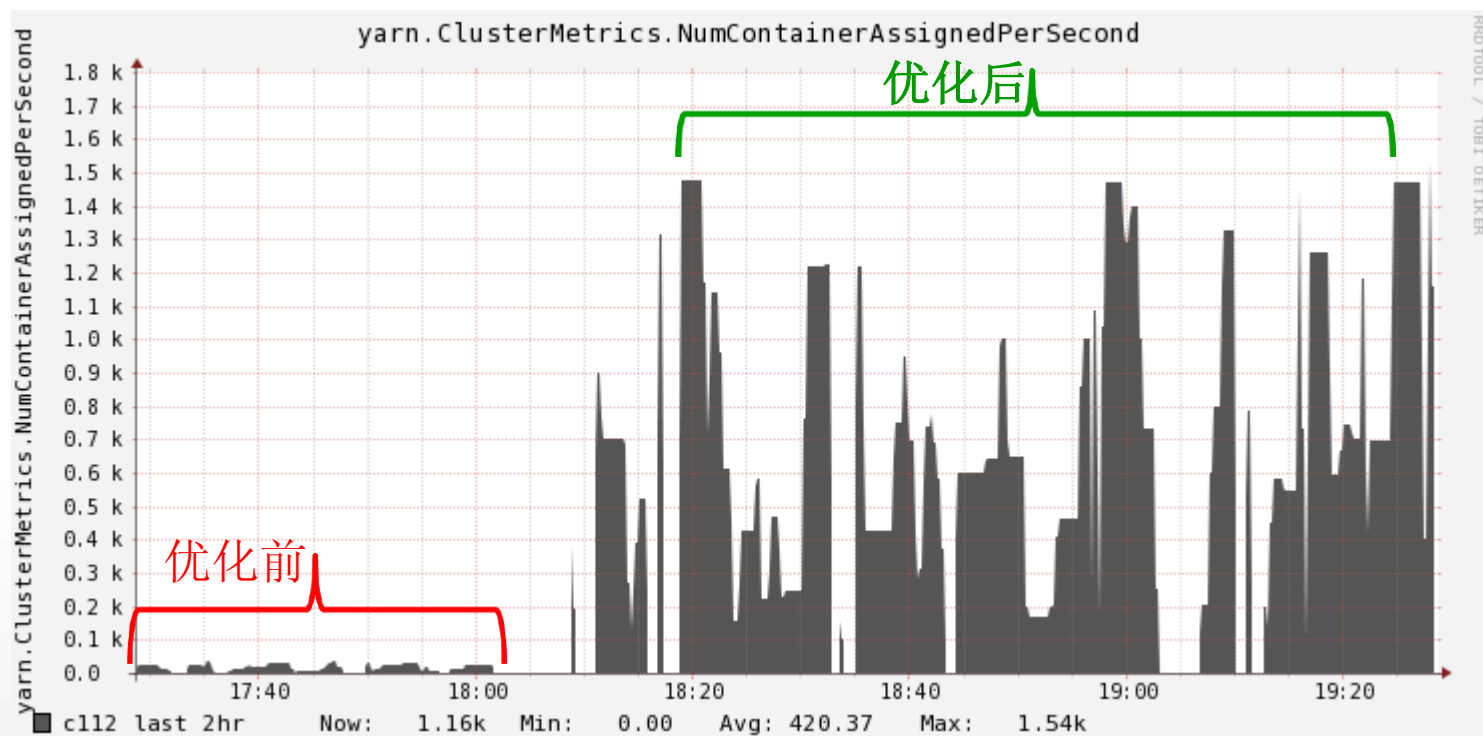
- 取消全排序，采用堆排序,减少调度器**80%**的cpu时间；

◆ 降低调度开销

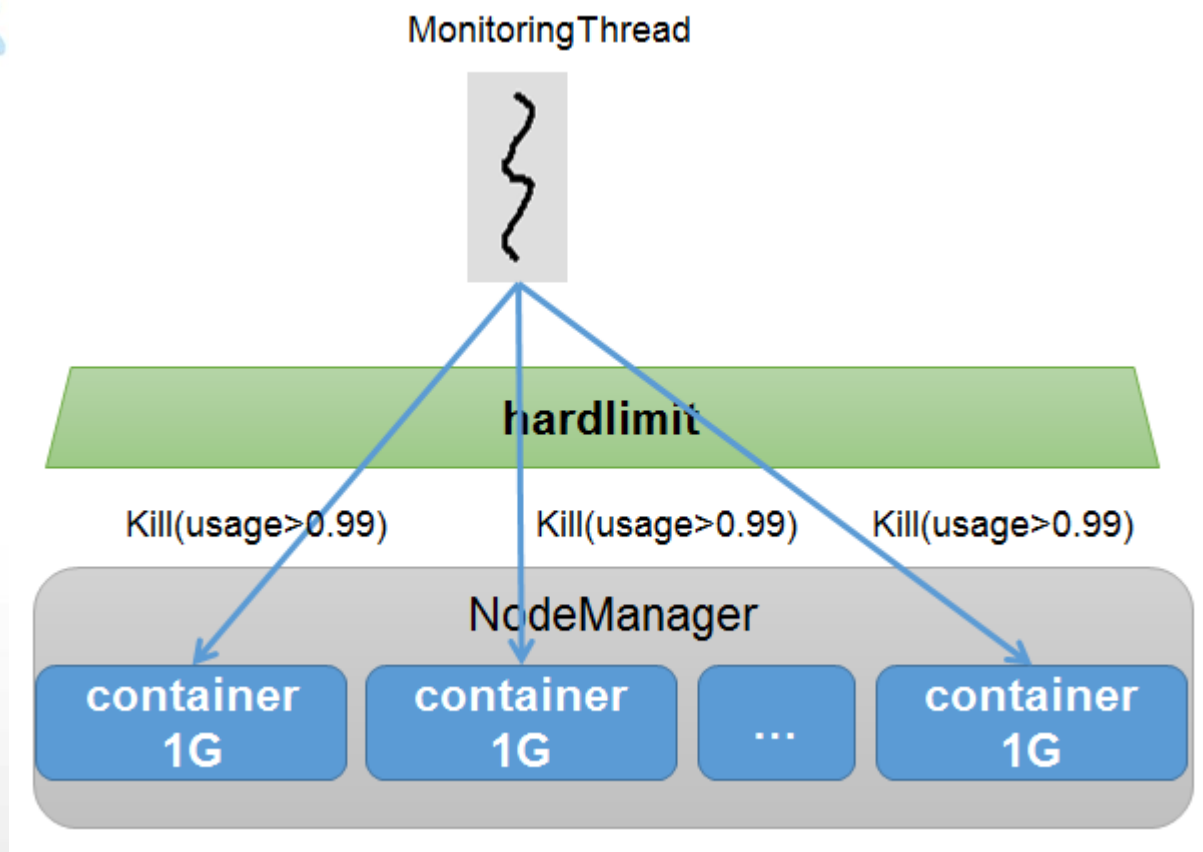
- 统计app/queue的资源需求，对无资源需求的app/queue,禁止其参与调度，现网中降低了近**80%**的调度开销。

Gaia自研调度器sfair scheduler

经过优化，在大规模环境下，调度器的平均吞吐率从20提高到1000，提高**50**倍。

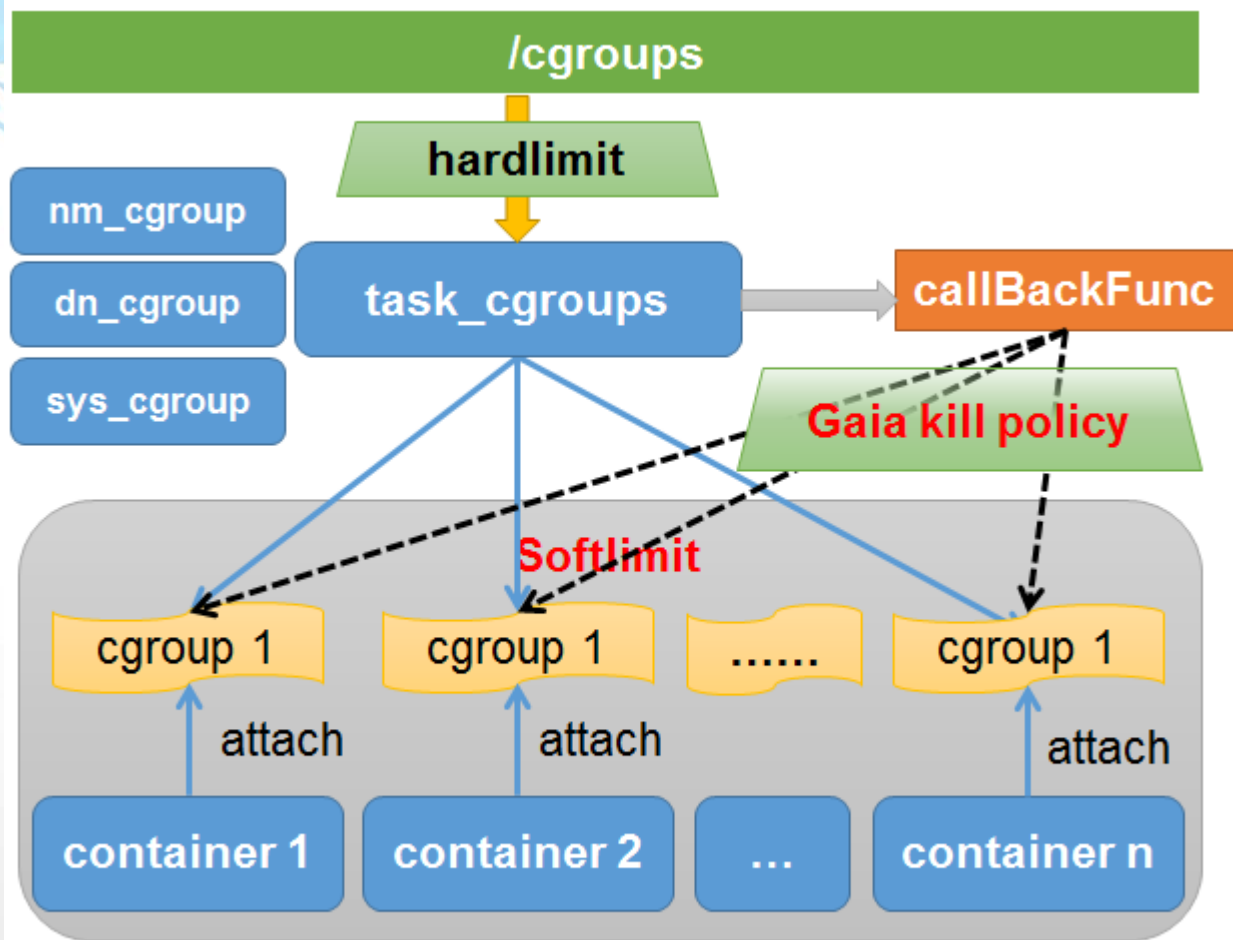


Gaia资源管理——内存



- 1) **系统不可控**: 周期性监控可能触发系统oom kill。
- 2) **资源浪费**: 无法利用机器的整体资源。
- 3) **资源利用率低**: 按照峰值分配container。
- 4) **失败率高**: hardlimit导致container容易被kill。
- 5) **资源需求不易评估**

Gaia资源管理——内存



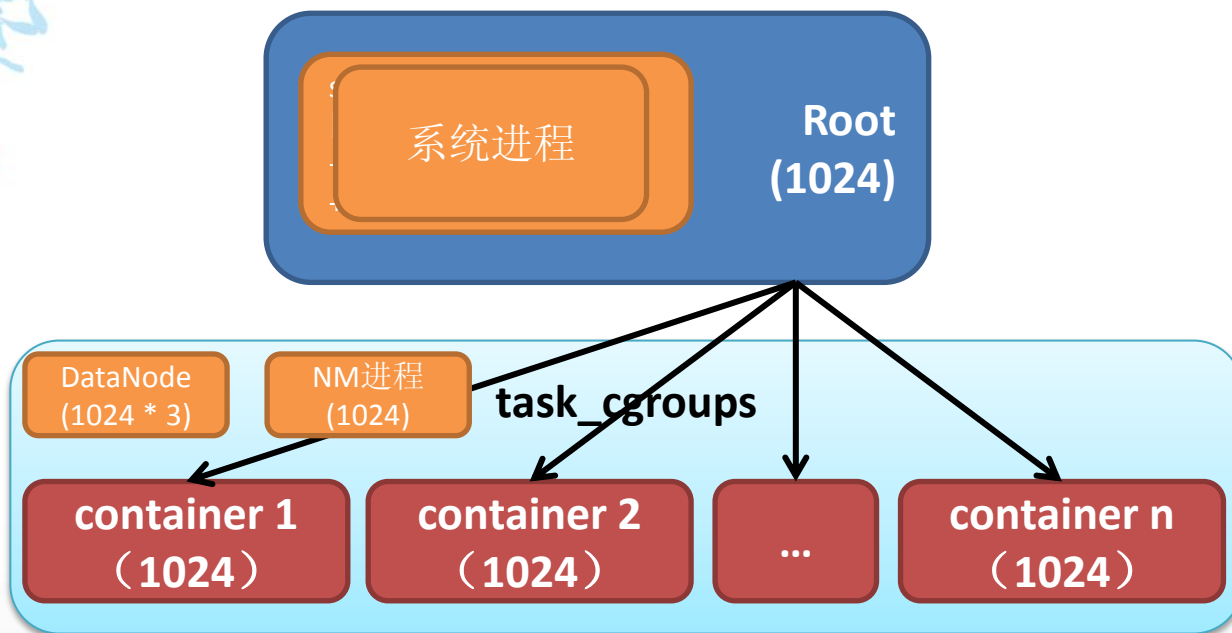
$$\sum \text{task softlimit} \leq \text{task cgroup hardlimit}$$

EMC——业内首创

Elastic Memory Control
(弹性内存控制)

- 1) 不会触发系统oom kill: 使用了container机制, 且task_cgroup是hardlimit。
- 2) 可以容纳更多container: 可按照平均值分配container。
- 3) 作业失败率大大降低: container之间是softlimit机制。
- 4) 对用户资源评估能力要求降低

Gaia资源管理——CPU

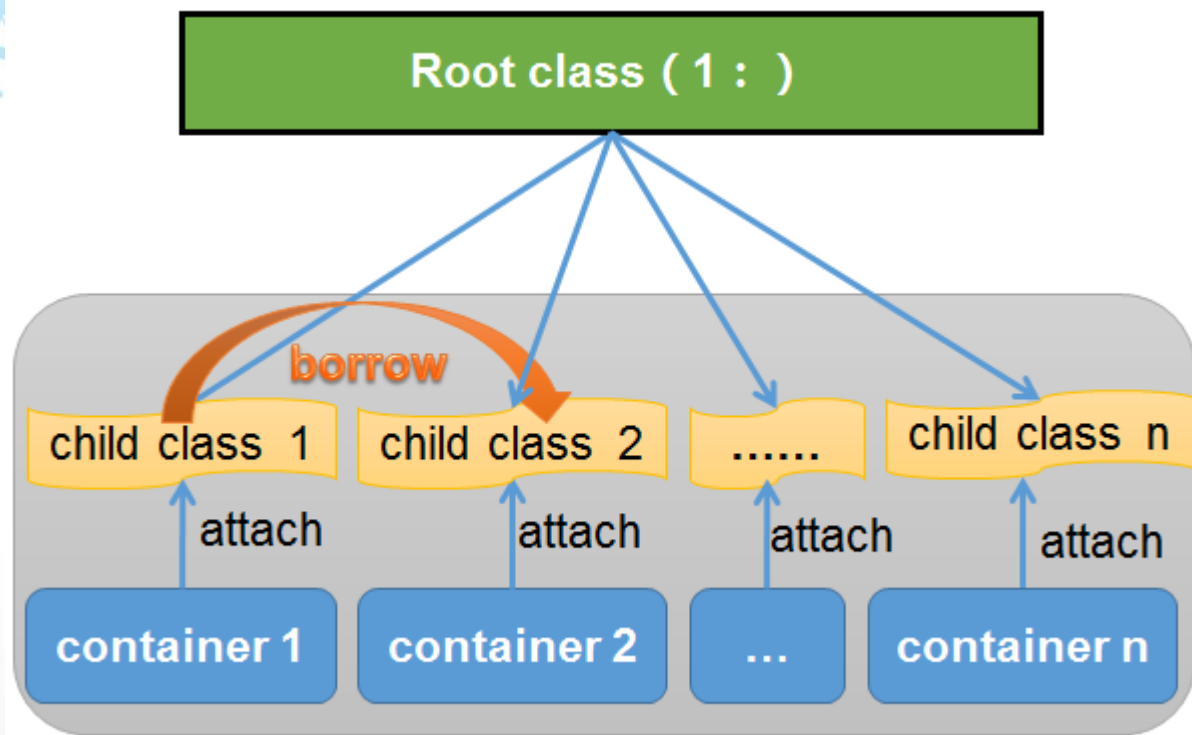


雪崩效应

——最终整个集群瘫痪

- 1) cpu share + cpuset结合的机制;
- 2) 为系统进程留足资源;
- 3) 将NM与DataNode进程纳入container管理;

Gaia资源管理——网络出带宽



1) TC+cgroups相结合的方式控制。

2) container之间的网络带宽可以相互borrow，可以充分利用网络资源。

Future work



规模更大

业务更多

场景更加复杂

可靠

精细化运营

多维监控

稳定

RM/NM HA

AM Retain

经济

建设大集群

提升实际资源使用率

易用

丰富api

作业可视化

Hive on Gaia

MR on Gaia

Pig on Gaia

Spark on Gaia

Storm on Gaia

Service on Gaia

Gaia期待您的加入

欢迎投递简历:

mavisluo@tencent.com

Q&A

THANKS

SequeMedia
盛拓传媒

IT168.com
www.it168.com

ChinaUnix

ITPUB