pt 京航空航天大学计算机新技术研究所

The Institute of Advanced Computing Technology

2020-2021学年研究生选修课程 云计算技术原理

Cloud Computing: Principles and Technologies

教学组: 胡春明,李建欣,沃天宇,林学练

2021年4月13日下午14:00-15:35

北京航空航天大学计算机新技术研究所

The Institute of Advanced Computing Technology

第六讲 数据中心资源管理





内容提要

$\Delta \leq$

- 数据中心
- 数据中心的能耗
- 互联网规模的资源管理与调度



数据中心

$\Delta \leq 1$

- 数据中心是云计算的物质基础,是计算基础设施的载体
- 能耗大户
 - 美国自然资源保护委员会 (NRDC)2019年发布的调查报告
 - 到2020年,全球数据中心用电量预计将增加到每年约1400亿千瓦时,约占全球能源消耗的5%
- 云计算加速数据中心大型化、专业化、集中化、运维自动化、绿色



产业界动向

Δ

- Microsoft
 - 芝加哥数据中心(Azure)
 - 约6.5万平方米
 - -30兆瓦能耗
 - -30 x10^6 W
 - -3万kW
- 大亚湾核电站
 - -90万千瓦



数据中心的评价指标 (1)

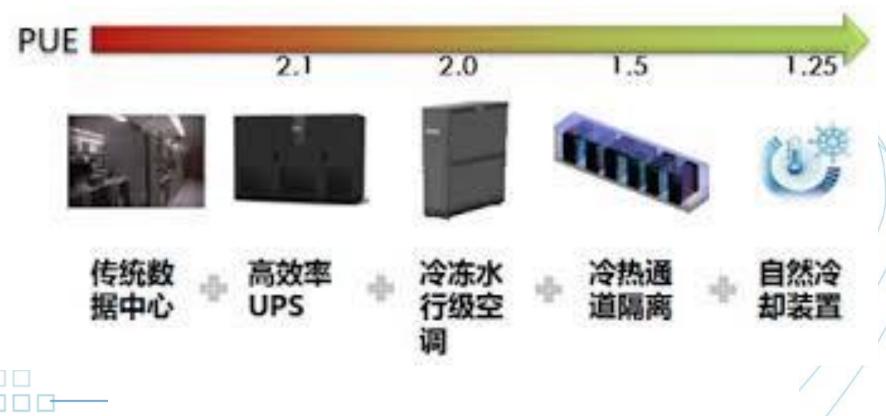
$\Delta \leq 1$

- PUE: 电能利用效率
- PUE = 数据中心总耗电/IT设备耗电
- 2013年中国云计算技术发展报告
 - 我国大部分DC的PUE > 2.5
 - -世界范围,平均水平1.9
 - 经过改进的云计算数据中心PUE 1.6
 - -设计较好的云计算数据中心可以降到1.2以下

数据中心的评价指标(1)

 $\wedge \leq 1$

• PUE: 电能利用效率





数据中心的评价指标 (2)

Δ

- 每千瓦建设成本
 - 成本性价比
 - 我国: >= 30000 RMB
 - -世界范围: 平均20000 RMB
 - 经过改进的云计算数据中心: 1-2万元
 - -途径:定制化、高性价比服务器、存储和网络设备



数据中心的评价指标 (3)

\.<

- 资源利用率
 - CPU UE:
 - 30-40%, 改进后可达到60%, 最佳实践可达到80%
 - I/O使用率:平均I/O吞吐率占设计总容量的比例
- 数据中心集群规模
 - -1000节点,经过良好设计可达到1-2万



数据中心的评价指标 (4)

\.<

- 数据中心自动化管理能力评价指标
 - 每1000节点所需的管理人员数量
 - 目前, 5-10人; 经过良好设计3-5人
 - 优秀的云计算数据中心, 1人以下

- 途径:
 - 自动化监测和管理工具, 软硬件容错方案

数据中心的评价指标 (5)

\<

- 数据中心服务能力评价指标
 - 多指标综合考核
 - 不同业务指标差别很大

- 高数据流量类: 流媒体
- 中等数据流量类: 搜索引擎
- 低数据流量类:数据备份类业务

数据中心的评价指标 (6)

_\\<1

- 数据中心选址评价指标
 - 适宜的气温条件,便于采用自然冷源(空气、水)
 - -丰富的电能和低廉的电价
 - Google 2007年在哥伦比亚河边的DC.10万kW低价 电力接入
 - 互联网接入能力
 - -低廉的土地成本
 - 便利的交通条件
 - -人力资源条件

数据中心厂商

- IBM: '卓越'数据中心(Leadership Data Center)
 - -安全的、高效率的、生命周期成本低的、节能的数据中心,此外,还需要使用相应的设备技术手段实现最有效的运维管理
- HP
 - 推出模块化"蝴蝶" 数据中心设计
 - PUE=1.18





模块化数据中心

 $\Lambda \leq 1$

· 2007年10月9日, Sun BlackBox问世, 开 启集装箱数据中心先河





数据中心建设方针

\wedge <

- 高密度机房
 - ▶思考:是密度越高越好么?会有哪些问题?
- 冷却方式
 - 风冷、水冷、自然冷却
 - 合理追求高PUE值
- 总规模
 - 10MW 20MW的规模
 - -3万-8万台的规模
- 建设数据中心地点:电价、通信、...

From: ALIBABA

科学/技术问题?

 $\wedge \leq$

• 在大规模的数据中心环境下:

- 如何保持系统集成能力的可持续发展?
- 如何保持性能线形增长?(Scale out的问题)
- 如何保持可靠性/可用性不下降?
- 如何实现灵活高效的管理?

需要可比较的实验环境,进行感性认识和深入分析



内容提要

$\Delta \leq$

- 数据中心
- 数据中心的能耗
- 互联网规模的资源管理与调度



Managing today's infrastructure has its challenges



Utilization
Average 15-20% per system*



Agility
Weeks to months per task



Power
40% of data center cost*



Space

>50% of enterprises see shortages by 2012*

*Industry sources



数据中心能耗

 $\Delta \leq$

- PUE只是一个方面,可靠性和节能是矛盾的两面
 - EnergyStar EPA/LBNL: 服务器节能
 - GreenCloud联盟: PUE指标
 - -绿色建筑理事会:数据中心建设
 - 美国采暖、制冷和空调工程师学会, TC9.9 数据中心空调效能
 - 国际绿色数字中心评估体系与组织 TIA942



新技术的替代:综合能耗降低

_\\\

- 高压直流供电
 - 替代方案
 - 380V交流-240V直流进服务器(93%)-服务器 转低压直流给主板供电(92%),总效率85.5%。
 - 其他优化:服务器出入风口精细化控制,服务器功耗线性;冷热通道一致,保证热场均匀
 - 阿里杭州机房: 1200平米, IT电力1500kW, 制冷900kW。单机架功率5.7kW。实测PUE 1.55-1.6
 - 行业标准: 通信用240V直流供电系统, YD/T 2378-2011



新技术的替代:综合能耗降低

/\\

- 空调的能耗替代
 - 自然冷空调系统 vs 水冷空调:能效比提高8-10倍。
 - 和地域相关:哈尔滨2/3时间可做自然冷;杭州62.4%。
 - -控制系统:水冷+自然冷+空调系统制冷。
- 模块化
 - -微模块,提高机架功率密度,提供精确制冷, 降低空调导致的PUE

۸-۷



- Facebook
 - -6-7万台服务器(1亿用户、400亿照片, 2000 new/s)
 - \$1 Million A Month in Power Costs

- · 2011年4月,公布了其新数据中心设计 Open Compute Project
- http://opencompute.org



Δ

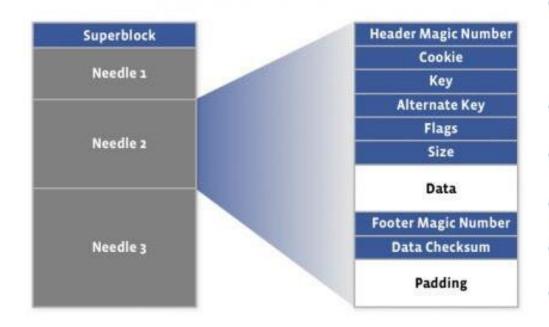
- Data center 相关技术分为四个部分
 - Electrical
 - Mechanical
 - Battery Cabinet
 - Triplet Racks



另外采用两项软件技术实现节能

¥Haystack: 分布式文件系统

★相比于一般工业级存储系统,Haystack帮助减少 20% 的能耗



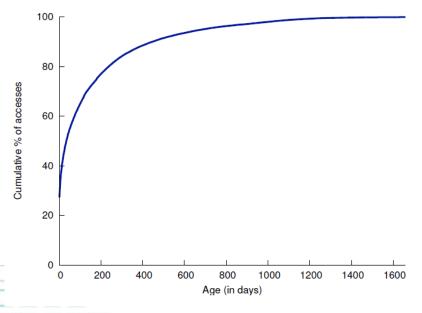


he Institute of A

Haystack: 小文件存储

\.<

- 创新点1: 元数据操作 > 通过内存访问替代磁盘访问
- 创新点2:面向小照片/缩略图等小图片处理的检索优化与容错,利用文件命名的编码与哈希机制,减少数据检索的开销



Operations	Daily Counts
Photos Uploaded	∼120 Million
Haystack Photos Written	∼1.44 Billion
Photos Viewed	80-100 Billion
[Thumbnails]	10.2 %
[Small]	84.4 %
[Medium]	0.2 %
[Large]	5.2 %
Haystack Photos Read	10 Billion

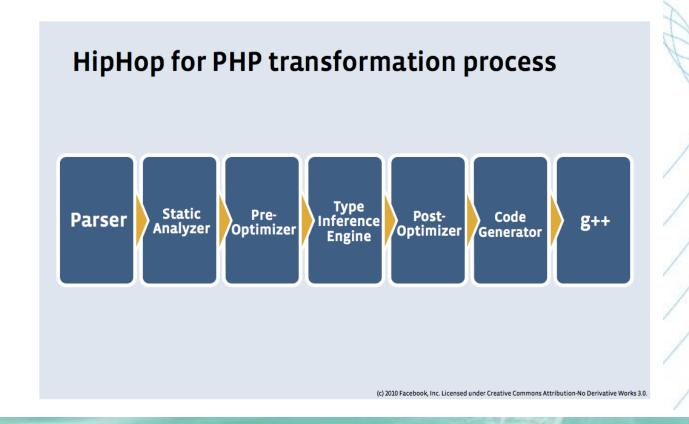
Table 3: Volume of daily photo traffic.

OSDI 2010, Finding a needle in Haystack: Facebook's photo storage

Doug Beaver, Sanjeev Kumar, Harry C. Li, Jason Sobel, Peter Vajgel,

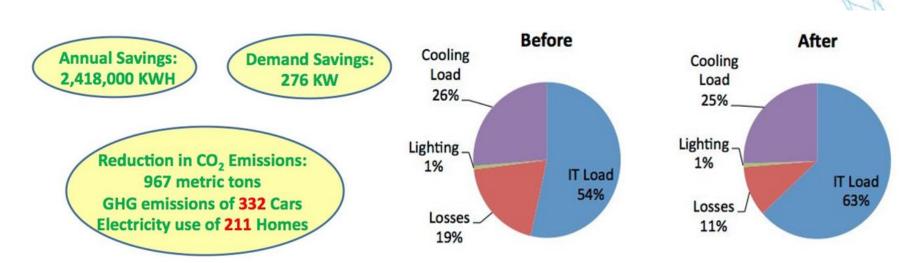
另外采用两项软件技术实现节能

HipHop for PHP: 通过将Php代码到C++的转换和编译,提高执行效率



$\wedge \leq$

 相比于传统的数据中心, Open Compute 项目包含了很多新的技术, 这些技术可以减少能源的耗费,增加项目的可靠性,可以使得操作简 化并增加耐用性,其目标是减少24%花费的同时效率提高38%。



还有什么办法提高PUE?

或者,还有什么办法提高 能效 Energy Efficiency





数据中心能耗

_\\<`

- 数据中心研究方向
 - -能耗
 - 网络
 - -灵活性
 - 数据分布

SIGCOMM Review.2009. The Cost of a Cloud Research Problems in Data Center Networks. Microsoft



什么造成能耗问题如此严重

• 客观现实1: 计算系统负载时空分布不均匀

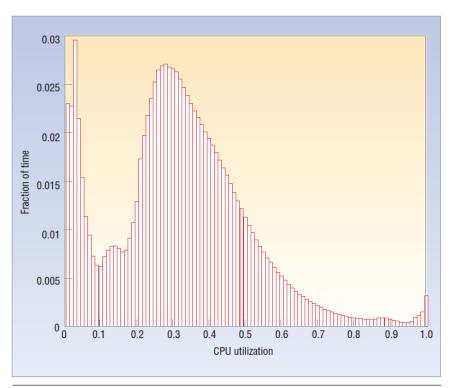
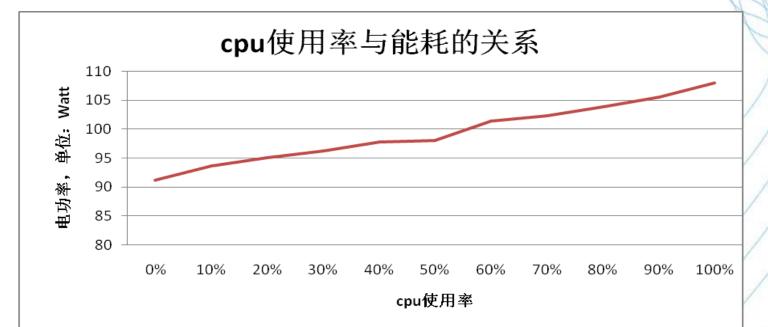


Figure 1. Average CPU utilization of more than 5,000 servers during a six-month period. Servers are rarely completely idle and seldom operate near their maximum utilization, instead operating most of the time at between 10 and 50 percent of their maximum utilization levels.

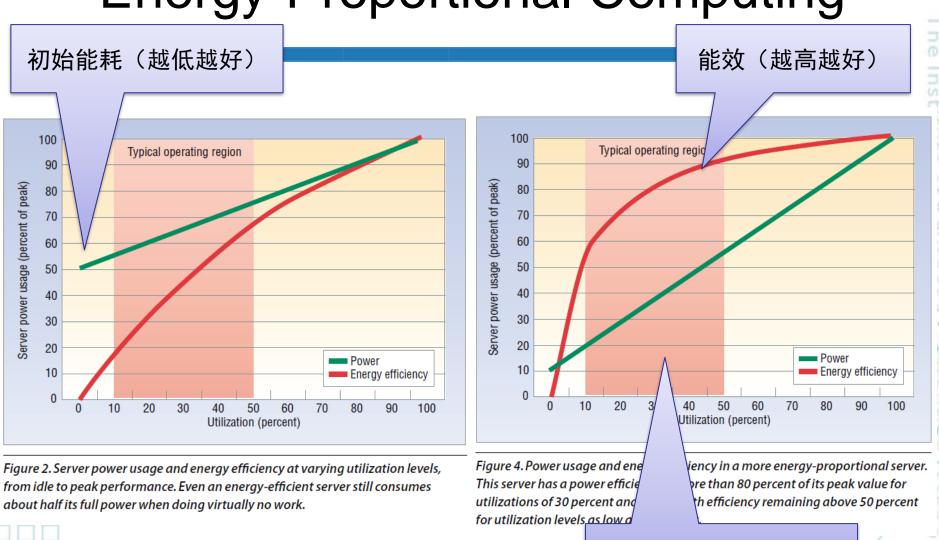


什么造成能耗问题如此严重

- _/_
- 客观现实2: 计算系统能耗与负载不成比例 变化
 - 我们的DELL PC实际测试



Energy-Proportional Computing



服务器通常工作区间

Barroso and U. H. olzle, "The case for energy-proportional computing," IEEE Computer, Jan 2007.



细致分析

 $\Delta \leq 1$

• 能量消耗不随负载等比例变化





相关工作

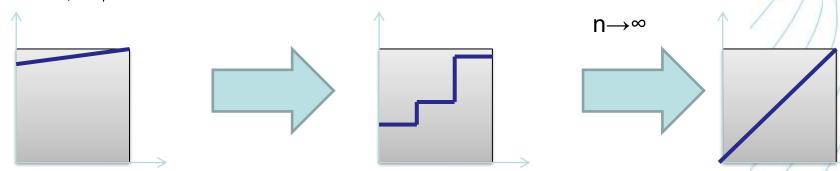
$\Delta \leq$

- 单个计算系统的节能技术
 - 器件: CPU、内存、磁盘...
 - 节能技术和系统的基础
- 数据中心资源调度与能耗管理
 - 负载聚合
 - 能耗感知的调度
- 新Input
 - -虚拟化、多能耗可控系统



单个计算系统的节能技术

- _^<1
- 器件: CPU、内存、磁盘...
 - DVFS: Intel SpeedStep/AMD PowerNow!
 - -磁盘降速,风扇自动调速
 - 仍有一些设备难以Energy-Proportional
 - 结果:





数据中心资源调度与能耗管理

• 数据中心能耗评估

- CPU based [ISCA2007]
- 数据中心组件的能效改进
 - 网络系统
 - [Sigcomm2003] [ISCA2010] tion by d
 [NSDI2010] [INFOCOMM.2009]
 - 电源系统 [ASPLOS2009]
- 负载集中策略
 - Machine Learning Based [e-Energy.2010]

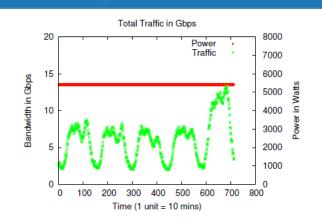


Figure 2: E-commerce website: 292 production web servers over 5 days. Traffic varies by day/weekend, power doesn't.



数据中心资源调度与能耗管理

- Load Concentration—"请随手关灯"
 - -集中负载,关闭空闲节点



DPM 在要求降 低时关闭服务器 DPM 在工作负 载增加时恢复服 务器在线状态

- -问题1:为什么要空闲?空闲也是一种浪费
 - 阿里云: 提高集群的应用适应能力
- 问题2: 关了再开,延迟怎么办?来得及么?
 - 减小能量状态迁移开销与延迟

nstitute

矛盾: 节能 vs 服务质量

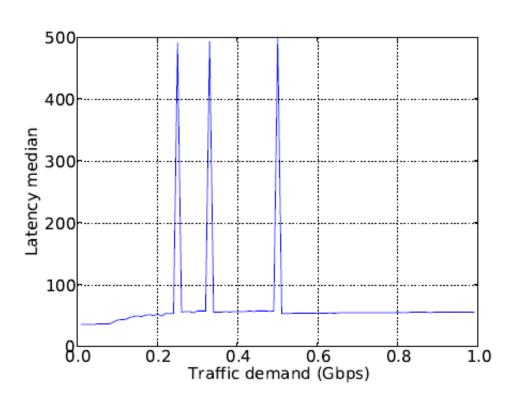


Figure 14: Latency vs demand, with uniform traffic.



内容提要

$\Delta \leq$

- 数据中心
- 数据中心的能耗
- 互联网规模的资源管理与调度
- 混合云与云际计算



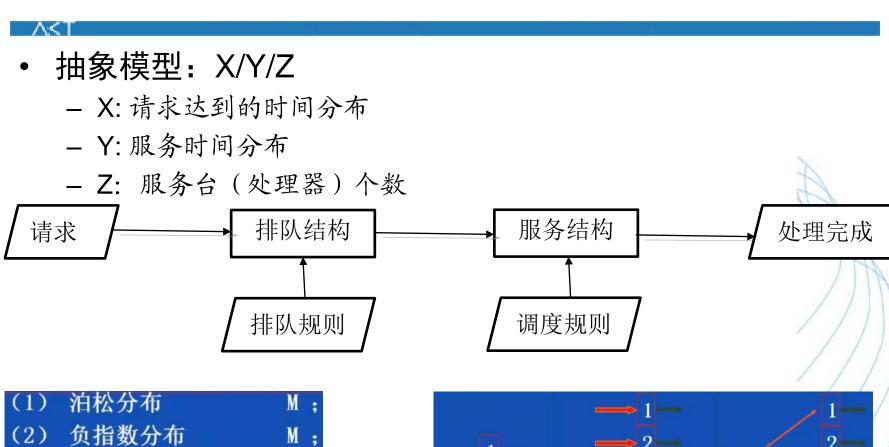
调度是做什么?

$\Delta \leq$

- 资源管理
 - 资源组织
 - 资源状态监控
- 任务组织
 - -接纳与分配任务:队列、优先级、任务特征
 - 已分配任务的执行情况监控
- 调度系统与调度策略
 - 完成任务到资源的匹配

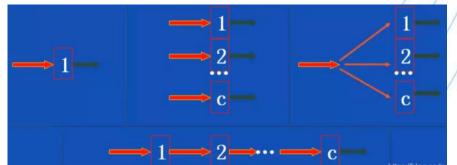


调度是做什么?



	11211674 11	
(2)	负指数分布	М ;
(3)	k阶爱尔朗分布	E _k ;
(1)	确空刑	n

(5) 一般服务时间分布 G;



调度系统的基本问题

\^<

• 资源的特征

- 同构、异构; 跨域、不跨域
- 抽象粒度与隔离机制

• 作业/任务的特征

- 同构、异构;任务依赖关系
- 作业/任务级容错支持

• 调度特征

- 数据局部性(任务分配是否要考虑数据位置)
- 抢占式与非抢占式
- 一分配策略(All-or-nothing, Reserve-based, Allocation on-demand)
- 安全性、公平性: 饿死(Starvation)、死锁(deadlock)



常见调度策略

$\Delta \leq$

- Random
- Round-Robin, Weighted Round-Robin
- FIFO
- Fair Scheduler (Hadoop): DRF
- Capacity Scheduler (Hadoop): User-based
- Deadline Constrain Scheduling:
- •

调度系统的评价指标

$\Delta \leq$

- 站在全局资源使用角度
 - 资源利用率
- 站在作业/任务的角度
 - 公平性
 - 平均/最大等待时间;平均/最大完成时间(Make span)
 - 作业的宏观衡量指标: 如吞吐率等
- 站在用户的角度
 - 公平性, QoS保证(SLA违约)
 - 抢占式调度下的"恢复现场时间"



互联网数据中心的特点(1)

• 数据中心资源规模化

- 数据中心规模庞大,单一调度集群节点数 快速增长
 - 2013年: Yahoo公司年在Yarn的系统研发基础之上支持4000个计算节点规模
 - 2014年:阿里云飞天计算平台实现并发布了 单集群5000台调度与管理技术

• 负载类型多样化

- 计算类负载:
 - 离线计算: MapReduce任务
 - 在线计算:交易处理与数据查询、流计算、图计算、Web负载、VM负载等
- 存储类负载:分布式KV存储、分布式文件 系统等

・并发用户访问和请求规模

- 2012年Taobao平均每天6000万访客,浏览 近8亿件商品:
- 2013年9月Amazon独立访问量近1亿,日 图片浏览量超过20亿;
- 2014年双十一支付宝支付的峰值则突破**了每** 分钟285万笔
- 整个3亿笔交易和清算在2小时内完成

海量数据处理规模

- 2008年Google平均每天处理9100万次查询, 每天处理多达20PB的数据量,
- 2013年Baidu每天处理数据超过100PB
- 2014年阿里云开放数据处理服务(ODPS)的 生产系统中离线数据处理任务每天将会处理数 万个PB级的离线处理任务
- 2015 ODPS单集群作业每日提交量超过十万,超过18,5000个任务以及超过4200万个任务实体



传统HPC/网格任务: 预先分配, 生命周期中不调整, **单一与静态**



互联网任务特征:弹性调整,**动**

态与多样

互联网数据中心的特点(2)

• 任务偏好多样化和动态化

- 任务可以提出"任务特定"的偏好和约束条件(e.g., 对处理器的要求、数据本地性需求、执行节点偏好)
- 任务可以在执行过程中实时的动态弹性调整资源

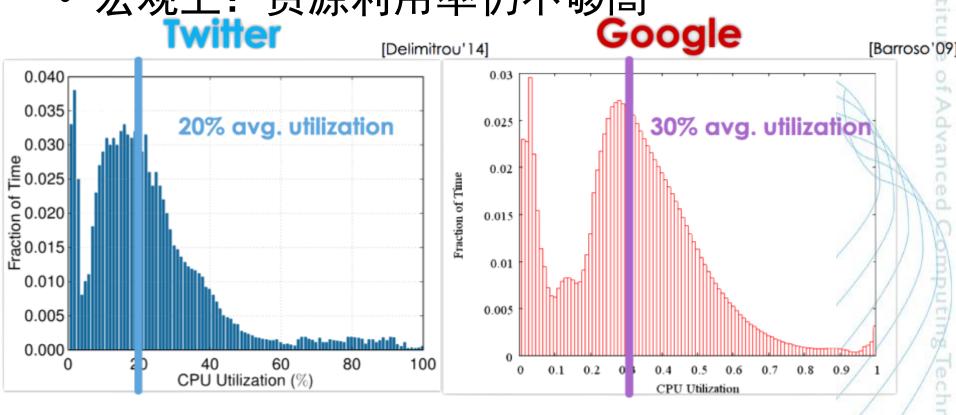
• 资源需求多样化

- 可分为CPU密集型、IO密集型、Mem密集型等
- 产生大量资源竞争,导致负载之间性能干扰
- 这些干扰将产生大量额外的能量消耗,造成资源的浪费

d

核心指标之一: 资源利用率

• 宏观上: 资源利用率仍不够高



处于世界先进水平的云系统资源利用率大致在20%~30%

Twitter和Google



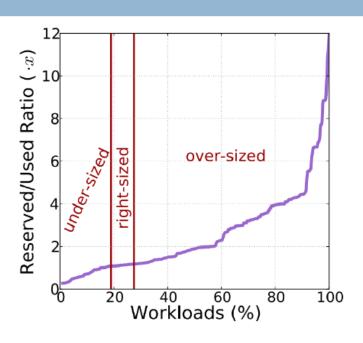
• 原因1: 资源利用率依负载变化

	数据来源	利用率	负载类型
1	基于2008年亚马逊汇总2万台 服务器数据	数据中心服务器的利 用率仅为6%	混合负载利用率
2	亚马逊AWS	平均CPU利用率7- 17%	混合负载利用率
3	5000节点集群,连续6个月负 载平均	平均约为20-30%	混合负载利用率
4	阿里云	???	混合负载利用率
5	腾讯游戏	在线业务10%	单一负载
6	2013年1-3月,Google数据中 心Trace数据	批处理作业的资源利 用率约为75%	单一负载

_/\< |

• 原因2: 资源预留不靠谱

Reserved vs. Used Resources

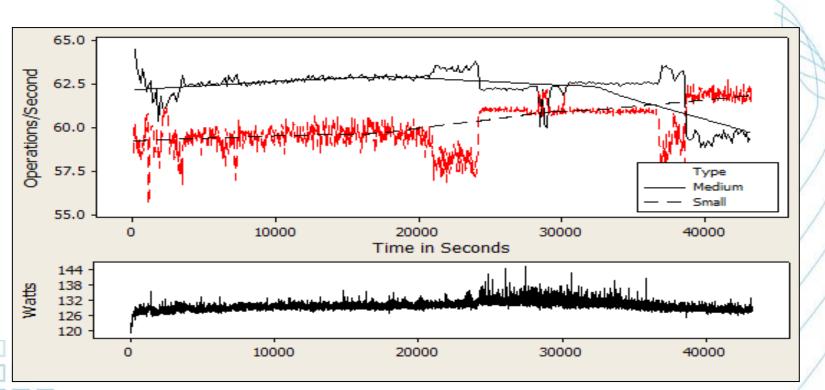


20% of job under-sized, ~70% of jobs over-sized



_\<

- 原因3: 各种计算框架中存在资源的无效利用现象
- 例如: 多个共享负载的性能干扰问题



- 原因3: 各种计算框架中存在资源的无效利用现象
- 例如:离线批处理作业的长尾慢任务问题

TABLE A CLASSIFICATION FOR STRAGGLER ROOT CAUSE

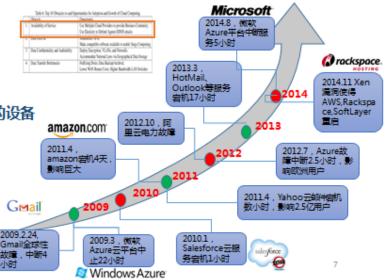
TABLE 4 CLASSIFICATION FOR STRAGGLER ROOT-CAUSE.						
Type	Category	Specified Description	Occurrence frequency			
1	High CPU utilization	Low time-slice sharing and process scheduling due to certain bad user-defined worker logic, unbalanced workload aggregation etc.	30%			
2	High disk utilization	Local disk read and write conflicts, unbalanced tasks aggregation, disk faults etc.	23%			
3	Unhandled operational access request	Distributed file system request surging(usually read request) and overpass the capability of request handling.	23%			
4	Network package loss	Network traffic package loss, resulting in repeating intermediate file and data transmission.	13%			
5	Hardware faults	Server timing-out, hang etc.	7%			
长屋任冬的原因分析		block input resulting in data skew.	3%			

K甩钉为的你凶儿们

- 异构的节点容量、网络拥塞、网络丢包、资 源竞争、CPU利用率、磁盘IO忙
- 任务或服务器内部出现的各种软硬件故障

$\Delta \leq$

- 原因3: 各种计算框架中存在资源的无效利用现象
- 例如: 为保障高可用而进行的多副本冗余执行
- 例如: 由于系统或组件故障导致的任务重调度
 - · 系统和服务的高可用性、可靠性
 - 2009,伯克利白皮书
 - 十大关键问题之首
 - · 互联网计算平台管理大量低可靠的设备
 - Amazon, 150万台服务器
 - Google存储节点失效率 2% 4%
 - Google网络节点 1.1%-11.4%
 - 系统规模增大复杂性
 - 系统中组件个数大幅度增加
 - 软硬件故障组合暴增



资源利用率

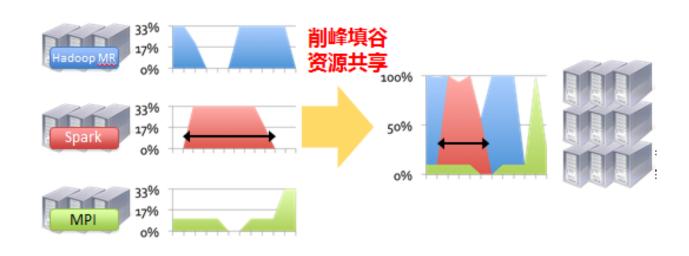


有效计算 资源利用率

本质问题

$-\Delta \leq$

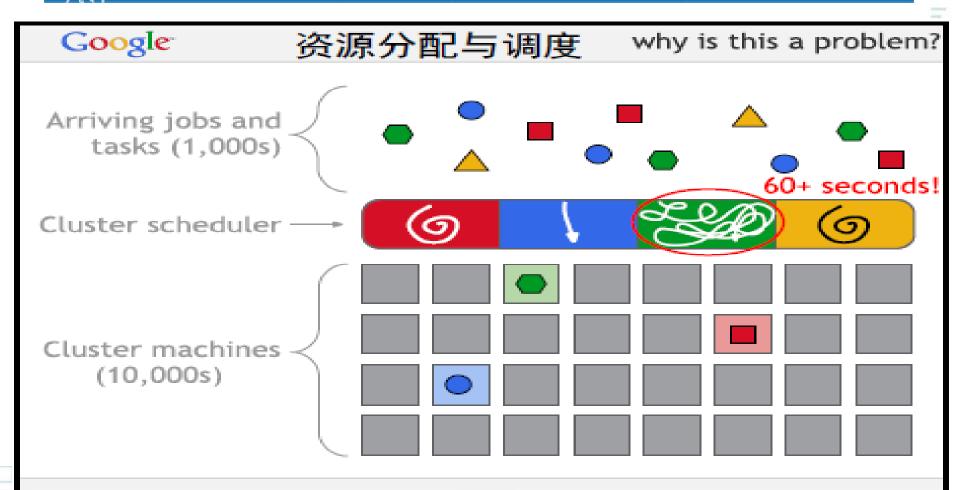
- 平衡共享与隔离的矛盾
 - 需要更好的资源抽象与调度设计



- 适应规模化(故障、伸缩性)
 - 需要更好的弹性与容错设计
- 适应资源、请求的不确定性
 - 一需要数据建模理解不确定性,并提供控制机制



调度的本原问题



4

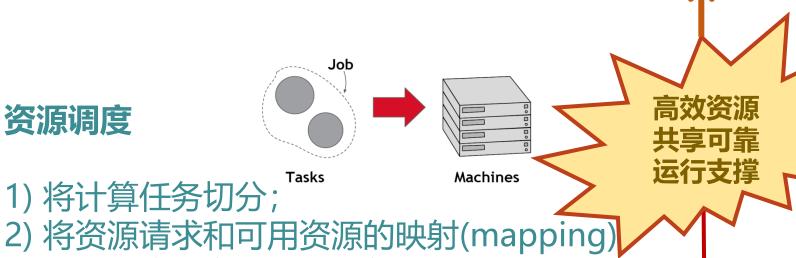
EuroSys 2013



调度的本原问题

资源调度

互联网新型应用 (服务、计算框架)



3) 利用对应的资源进行任务执行

互联网海量资源



统一资源调度框架

$\Delta \leq 1$

Mesos

- UCB, NSDI 2011
- http://mesos.apache.org
- Mesos: a platform for fine-grained resource sharing in the data center

YARN

Apache Hadoop v2, v3

Corona

Facebook

OMEGA

- Google, eurosys 2013
- Omega: flexible, scalable Schedulers for large compute cluster

Quasar

Eurosys 2014

BORG

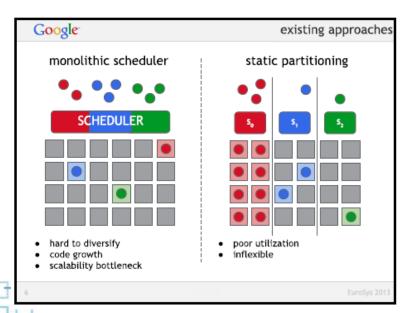
- Google, Eurosys 2015
- Large-scale cluster management at Google with Borg

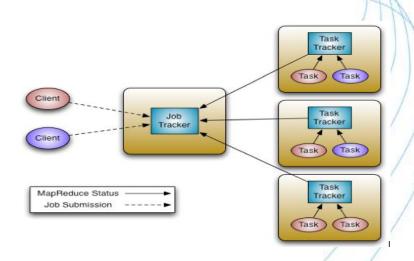
调度系统的发展(1)

_/\\

• 集中调度器

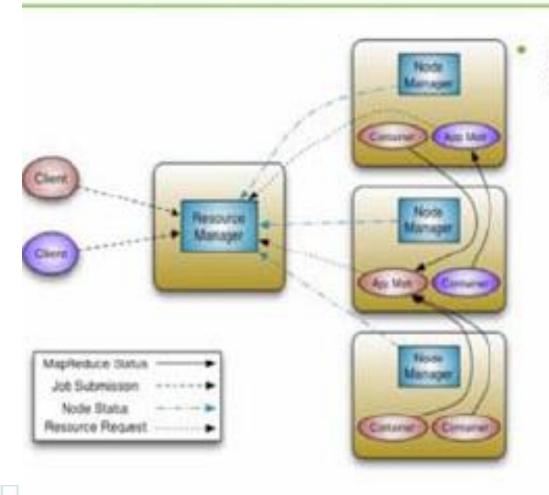
- 单独的Master中负责所有的调度决策、状态监控与更新 (e.g., Hadoop 1.0中的JobTracker)
- 问题: 进程过载, 极易出现单点故障
- 只能支持单一的计算模式与模型 (e.g., Map or Reduce slot)







YARN - Architectural Overview

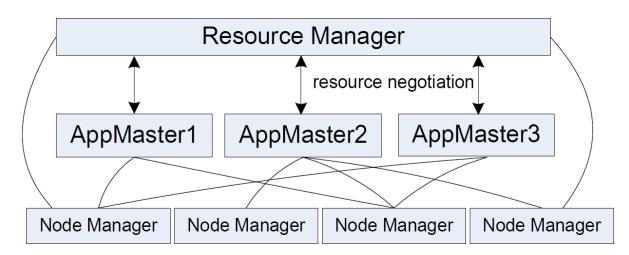


- Scalability Clusters of 6,000-10,000 machines
 - Each machine with 16 cores, 48G/96G RAM, 24TB/36TB disks
 - 100,000+ concurrent tasks
 - 10,000 concurrent jobs

调度系统的发展(2)

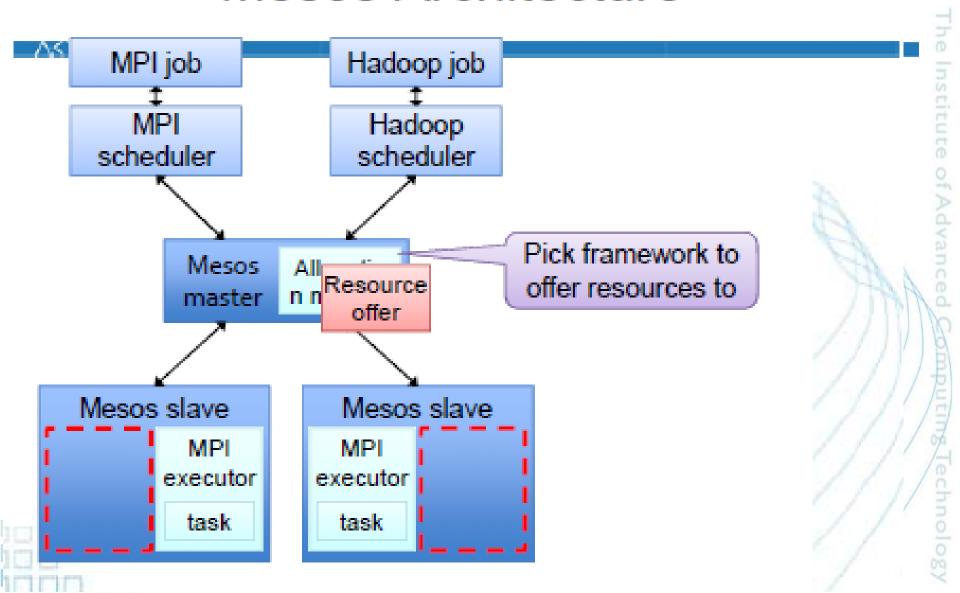
_\\<`

- 二层调度系统
 - 将资源管理与任务调度解耦合
 - 中心资源管理器(Resource Manager)负责整个集群资源在多租户应用之间进行资源协调与分配,每个应用控制器(AppMaster)则控制内部的逻辑划分、资源分发与绑定



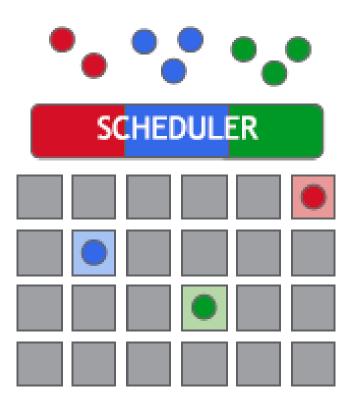
d

Mesos Architecture



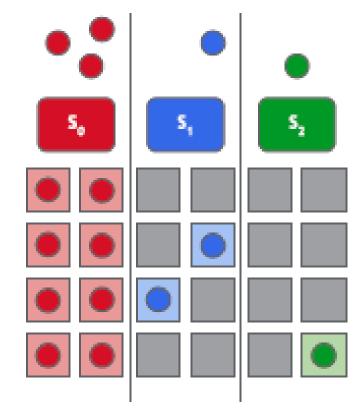


monolithic scheduler



- hard to diversify
- code growth
- scalability bottleneck

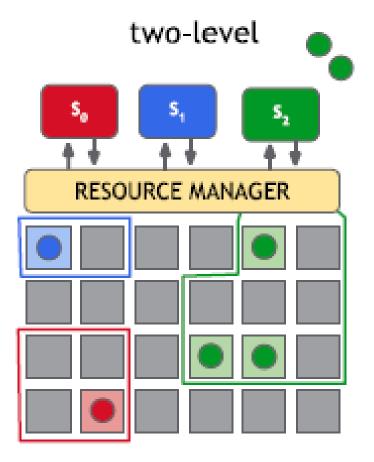
static partitioning



- poor utilization
- inflexible

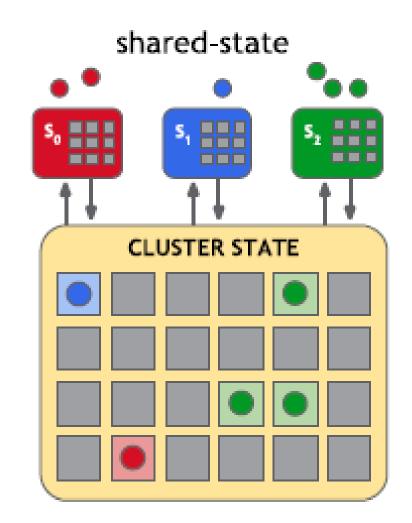






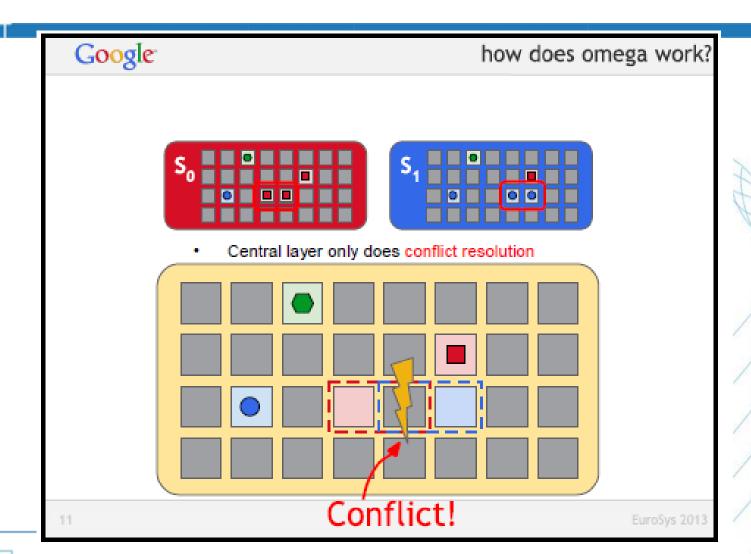
- hoarding
- information hiding

e.g. UCB Mesos [NSD] 2011]





OMEGA

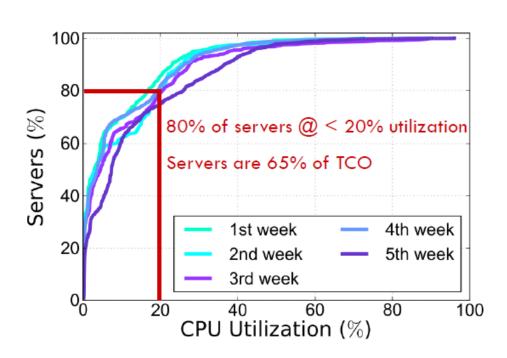


QUASAR: profiling based

۸۲۱

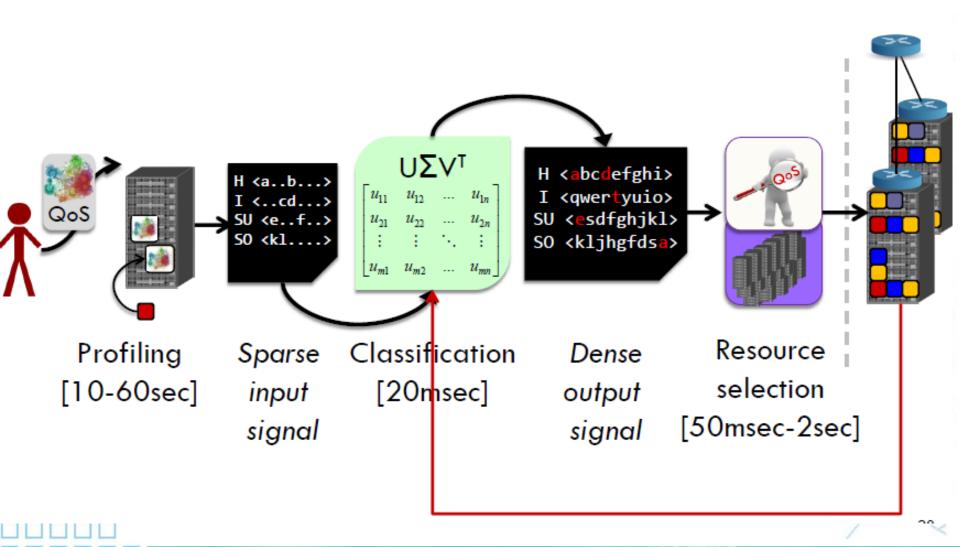
Datacenter Underutilization

- A few thousand server cluster at Twitter managed by Mesos
- Running mostly latency-critical, user-facing apps





Quasar Overview



Quasar: 典型的混合负载调度 (ASPLOS 2014)

$\Delta \leq 1$

- > 多负载混合调度
- ▶ 基于Profilling的方法,通过机器学习判断任务特征

Utilization: Fig. 7 shows the per-server CPU utilization (average across all cores) over time in the form of a heatmap. Utilization is sampled every 5 sec. In addition to improving individual job performance, Quasar increases utilization, achieving 62% on average versus 34% with the individual framework schedulers (right heatmap). Because performance is now higher the whole experiment completes faster. Workloads after t = 14400 are mostly best-effort jobs that take longer than the main analytics workloads to complete.

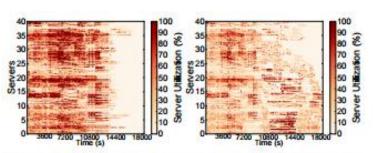


Figure 7: Cluster utilization with Quasar (left) and the framework schedulers (right).

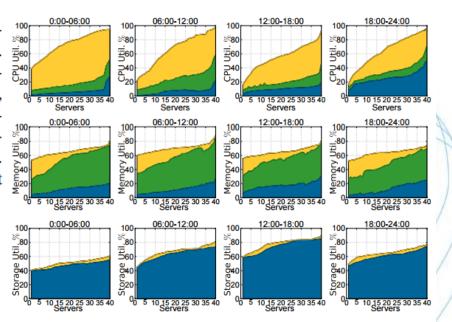


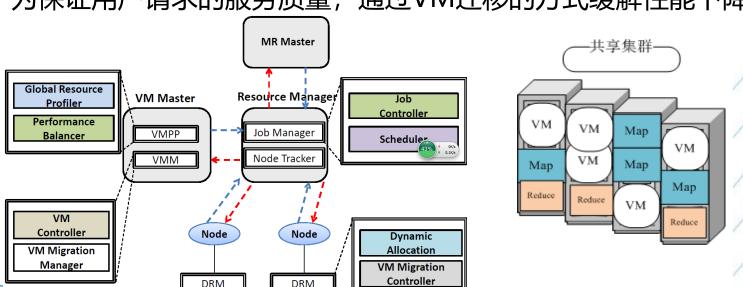
Figure 10: Average resource usage across all servers for four 6-hour snapshots. The cluster is running memcached (green), Cassandra (blue), and best-effort tasks (yellow) and is managed by Quasar.



在线离线任务的混合调度

ΔC

- 基于Yarn实现VM与MR任务的弹性混合
- 设计支持VM作业生命周期管理的协议,通过任务请求合并以减少并发请求的响应时间。
- VM资源实时动态绑定,基于动态绑定的方式减少over-provision造成的资源利用率低的问题
- 为保证用户请求的服务质量,通过VM迁移的方式缓解性能下降



Y.Wang, R. Yang, T. Wo, W. Jiang and C. Hu. Improving utilization through dynamic VM resource allocation in hybrid cloud environment. *IEEE ICPADS 2014*



Borg

\.\<`

- Eurosys 2015. Google.
 - Hide detail of resource management & fail over
 - High Reliability & availability
 - Scalability: running on a large scale (5000-20000+)
- Multi-workload merging
 - Mixed workload scheduling
 - Job/Task abstraction
 - Heterogeneous Machine capability (big/small cells)
 - QoS gurarntee

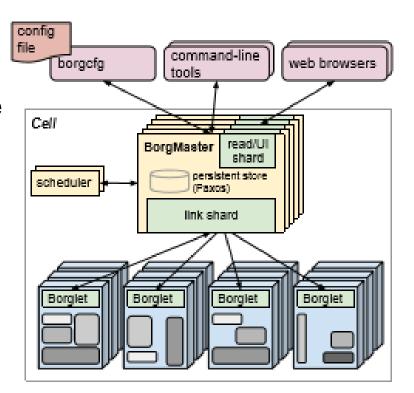
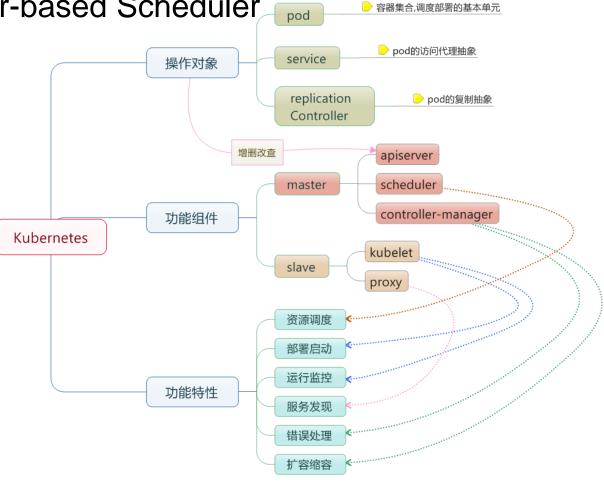


Figure 1: The high-level architecture of Borg. Only a tiny fraction of the thousands of worker nodes are shown.

Kubernets

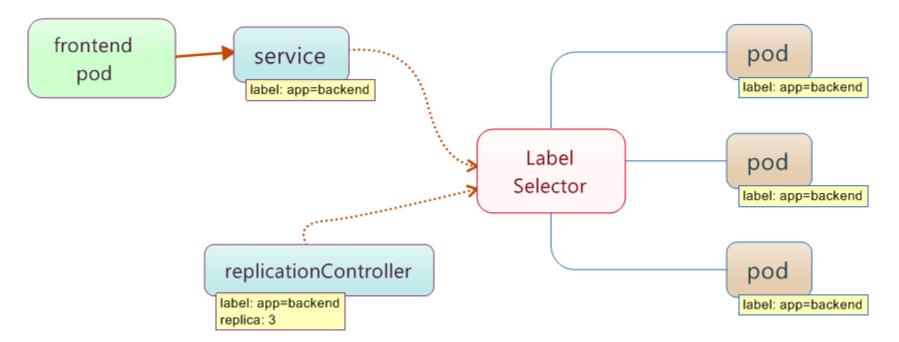




Kubernets

\.<

Container-based Scheduler



he institute o

调度系统的发展(3)

• 分布式调度框架

 多个调度器副本(多进程或多线程的形式),每一个 调度器都能根据本地缓存的状态或全局共享状态来独 立并发地处理资源请求

· 本地状态副本,Master进行协调

Distributed Scheduler

Distributed Scheduler

· 共享状态(shared states), 所有状态被所有调度器可见

Node

Node

coordinator

Distributed

Scheduler

Node

Node

无状态的完全分布式调度

Runtime

Runtime

Runtime

Runtime

- ▶ 为了实现低延迟,可采用基于采样的节点探测(Sampling-based probing)
- 每一个调度器进程能从随机m个server中自发探测那些拥有最少队列长度的 服务器,作为等待任务的候选目标节点

Sparrow: 分布式调度器 (OSDI 2013)

- 在面临高时效性计算任务时,调度的挑战
 - Bring response in 100ms

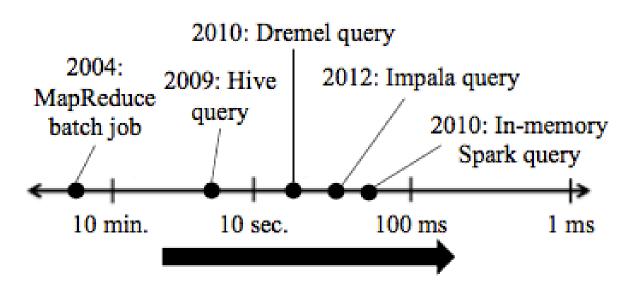


Figure 1: Data analytics frameworks can analyze large volumes of data with ever lower latency.



Sparrow: 分布式调度器

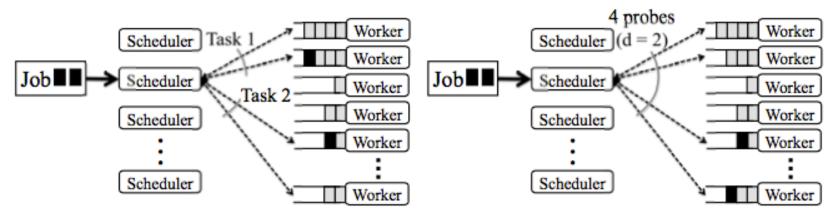
\<`

- 在面临高时效性计算任务时,调度的挑战
 - Bring response in 100ms
- 基本思想: 分而治之
 - 长时间运行的批作业 (mins, hours, days)
 - 短时运行的小作业 (<=100ms)
 - 变化1: 需要在局部进行调度决策
 - 变化2: 需要在亚秒级完成状态同步
 - 去中心,decentralized scheduler

Sparrow: 分布式调度器

/\≾

Batch Sampling + Late Binding



- (a) Per-task sampling selects queues of length 1 and 3.
- (b) Batch sampling selects queues of length 1 and 2.

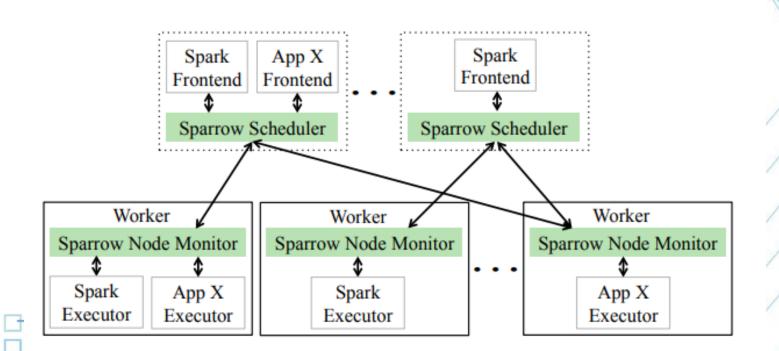
Figure 2: Placing a parallel, two-task job. Batch sampling outperforms per-task sampling because tasks are placed in the least loaded of the entire batch of sampled queues.



Sparrow: 分布式调度器

 $-\Delta \leq$

- 纯分布式的调度器结构
 - Client维护Scheduler列表,对Scheduler的心 跳检测,故障后任务重提交





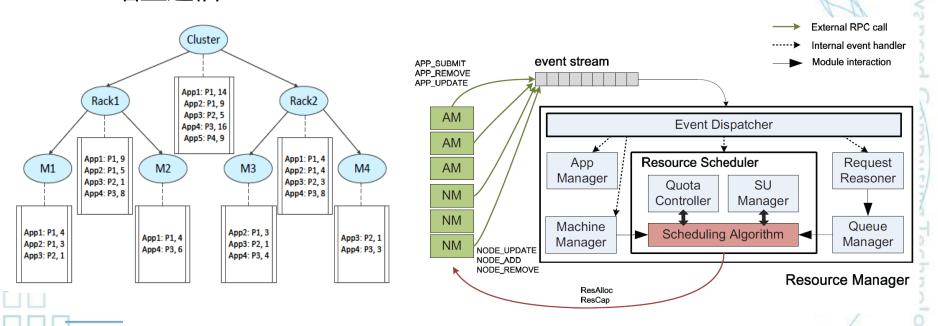
Aliyun Fuxi

$\Delta \leq$

• 重点解决"规模化"带来的调度决策可伸缩问题 . 和大规模集群消息通信可伸缩问题

Institu

- 基于locality-tree的资源请求队列
- 增量通信





北京航空航天大学计算机新技术研究所

The Institute of Advanced Computing Technology

技术挑战

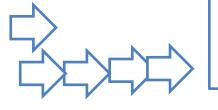
■ 如何在保障在线(延迟敏感)任务服务质量的同时, 提升集群资源利用率

在线任务

(延迟敏感)

Web请求、 虚拟机负载、 数据库查询…

负载波动剧烈



在线 任务 集群



平均资源 利用率<10%

物理隔离

资源预留

离线任务

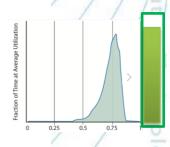
(延迟不敏感)

后台统计任务 机器学习训练...



离线 任务 集群







北京航空航美大学计算机新技术研究所

The Institute of Advanced Computing Technology

技术 挑战 如何在保障在线(延迟敏感)任务服务质量的同时, 提升集群资源利用率

在线任务

(延迟敏感)

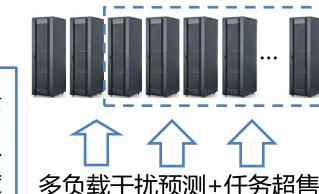
Web请求、 虚拟机负载、 数据库查询.

在线任务 负载较低



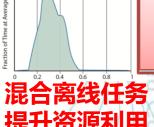
在同一个 物理集群上 完成 在离线任务

在离 线 混合 调度 集群

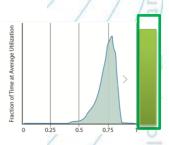


多负载干扰预测+任务超售 将超售任务推送至部分节点





提升资源利用



离线任务

(延迟不敏感)

后台统计任务 机器学习训练...

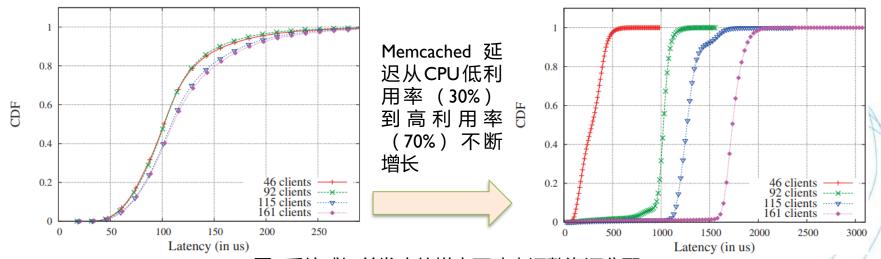


北京航空航天大学计算机新技术研究所

The Institute of Advanced Computing Technology

在线、离线任务混合部署提高资源利用率,然而

- 离线任务资源频繁被抢占造成Makespan增加
- 性能干扰造成在线任务SLA违约,如下图所示



系统感知并发度的增高而动态调整资源分配

高资源利用率



计算有效性



混部的合理性



北京航空航天大学计算机新技术研究所

The Institute of Advanced Computing Technology

技术 挑战 ■ 如何在保障在线(延迟敏感)任务服务质量的同时, 提升集群资源利用率

在线任务

(延迟敏感)

Web请求、 虚拟机负载、 数据库查询… 在线任务 负载较高



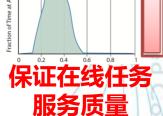
在同一个 物理集群上 完成 在离线任务 在线合度群



多维资源隔离+调度避让









离线任务

(延迟不敏感)

后台统计任务 机器学习训练...





基于资源超售的任务调度优化

\wedge

负载类型异构问题

- 如何满足不同计算框架的不同业 务需求?
- 如何保障批处理的端到端时间和 延迟敏感应用的服务质量?

集群资源利用率低问题

- 如何高效利用系统中的碎片和全 闲资源?
- 如何最大限度地精准启动和执行 超卖任务?
- 如何在节点进行生命周期与资源 隔离?

核心思想: 去中心化调度框架 + 资源超售技术 调度层的多阶段多维度机器筛选机制 + 基于任务升降级的 资源分配与调度 + 多维资源隔离机制

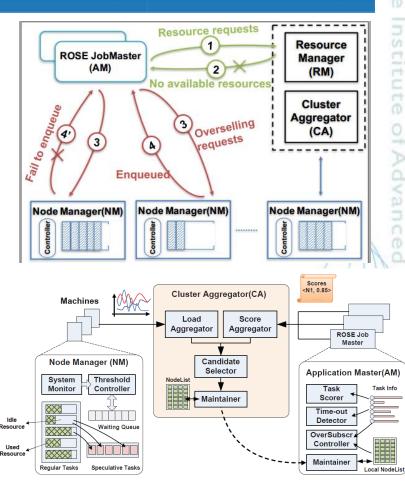
基于资源超售的调度系统框架

精准超卖: 多阶段节点筛选、机器列表增量维护

- 根据任务运行状态(抢占、crashed、长尾等)对其所在机器的进程惩罚式评分
- 根据历史信息的多维度资源预测,基于多维度 资源向量、实时负载、队列状态的筛选
- 多阶段机器筛选算法: 负载预测、异常排除、 多轮竞选
- 机器节点管理:细粒度机器监控和队列管理
 - 实时多维负载监控: Disk等
 - 队列管理:根据实时负载决定超卖Worker是否 启动、抢占等

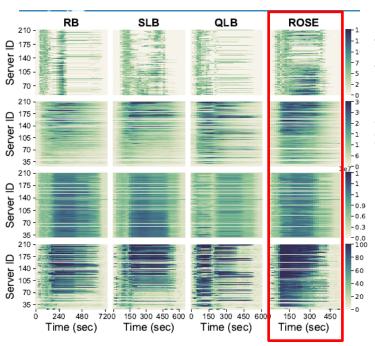
作业端基于超卖的作业调度机制

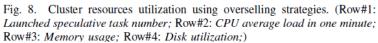
- 超卖仟务调度逻辑:根据自身DAG关系和资源 占有情况,决定是否分发超卖任务
- 超时重调度: 使用计时方式对饥饿等待的超卖任 务给予重调度

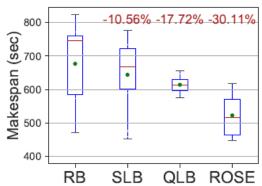


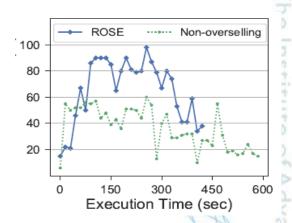


基于资源超售的资源调度系统性能分析









USAGE STATISTICS

Avg CPU Load 9.739 10.704 9.263 13	.176
	3.902
Mem Usage(GB) 10.040 9.610 7.294 10	0.457
Disk Util(%) 32.727 33.870 31.814 50	0.956

提升约 18.23% 的磁盘利用率缩减约 30.11% 的作业完成时间



在离线混合调度的工程实现

 $\wedge \leq$

• 某在离线混合调度系统



内容提要

$\Delta \leq$

- 数据中心
- 数据中心的能耗
- 互联网规模的资源管理与调度

86



小结

 $\Delta \leq$

- 数据中心
 - 能耗管
- 云计算理
 - -混合负
 - 更好的



- -与计算框架的紧密集成
- 调度的任务与挑战
 - 如何更有效的利用资源
 - 如何从框架层面容忍错误



参考文献

$\Delta \leq$

- Quasar: 混合负载调度
 - Quasar: Resource-Efficient and QoS-aware
 Cluster Management, ASPLOS 2014, Stanford
- Sparrow: 分布式调度器
 - Sparrow: Distributed, Low Latency Scheduling, SOSP 2013, UCBerkeley
- Fuxi: 大规模二层调度与容错
 - Fuxi: a Fault-Tolerant Resource Management and Job Scheduling System at Internet Scale, VLDB 2014, Alibaba & Beihang

タヒネネルタネルメメサ計算机新技术研究所

The Institute of Advanced Computing Technology

谢谢!

办公室:

新主楼G1119, 电话: 8233 9679

新主楼G506, 电话: 8233 9274

电子邮件:

hucm@buaa.edu.cn woty@act.buaa.edu.cn

