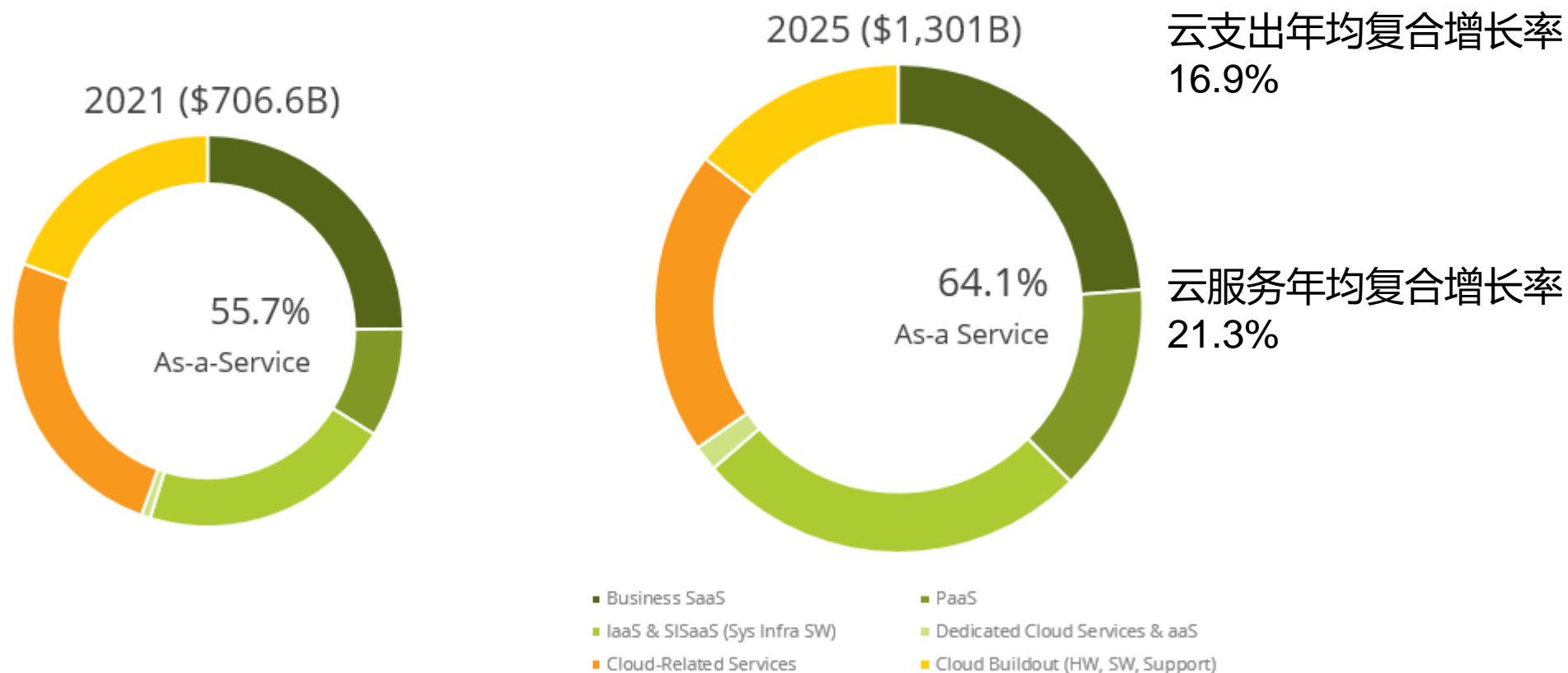


云数据中心的性能优化

郭健美

2022年秋

Worldwide Whole Cloud Outlook



Source: IDC Whole Cloud Forecast 2021-2025: The Path Ahead for Cloud in a Digital-first World, IDC #US47397521, Sept 2021



SOLE

系统优化实验室
华东师范大学

《软件系统优化》本科生课程材料

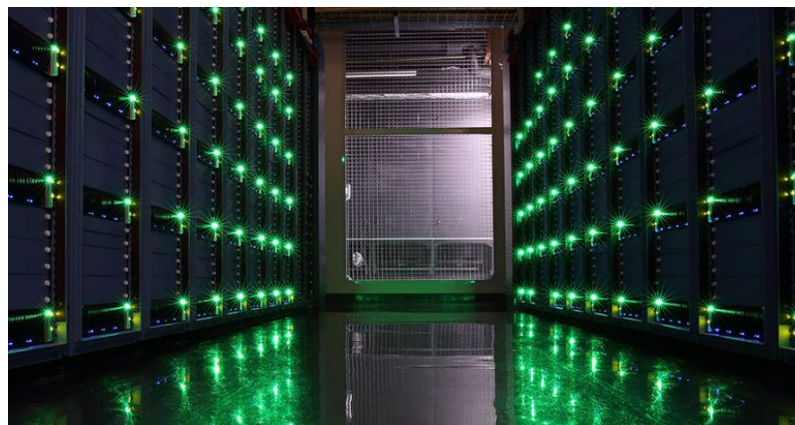
© IDC



http://www.xinhuanet.com/2020-07/14/c_1126233382.htm



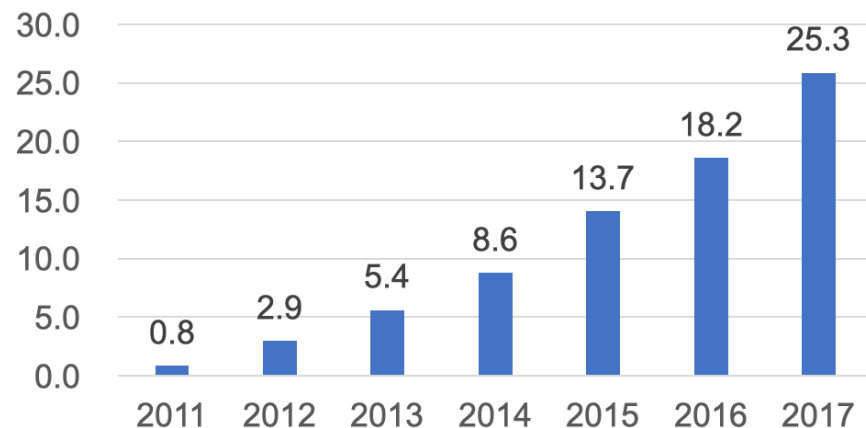
阿里巴巴张北数据中心总投资180亿元
建设规模30万台服务器
2017年阿里巴巴“双十一”的成交额1682亿元
张北阿里云基地承担了40%的业务量
支付峰值达到每秒25.6万笔



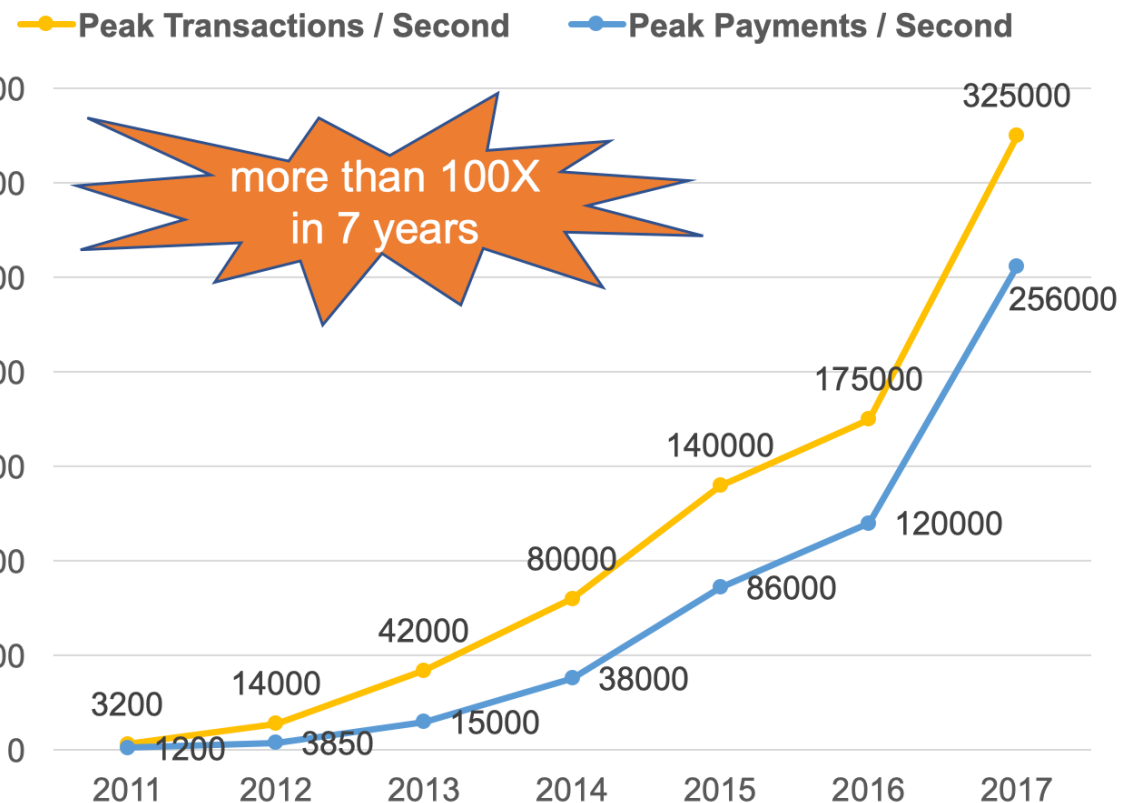
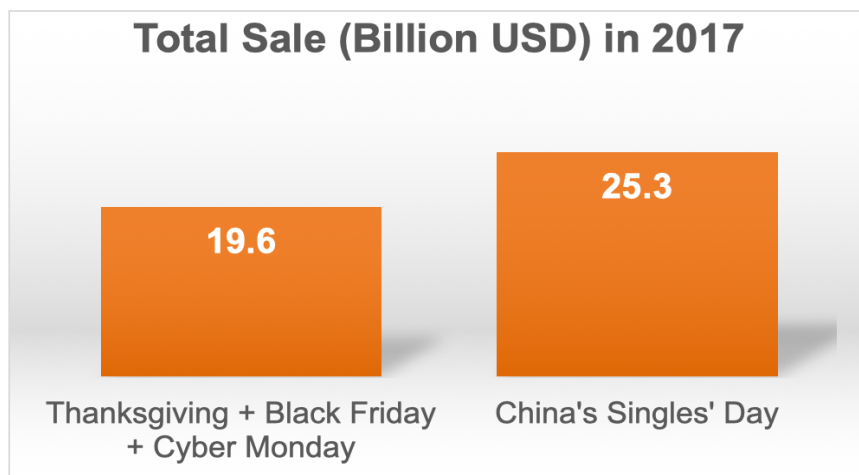
https://blog.csdn.net/weixin_34174322/article/details/90590041

http://www.gov.cn/xinwen/2018-02/28/content_5269382.htm

Total Sales (Billion USD)



Total Sale (Billion USD) in 2017



GreenTea JUG Java Meetup, 上海 2018

中国数据中心耗电量超过三峡大坝发电量



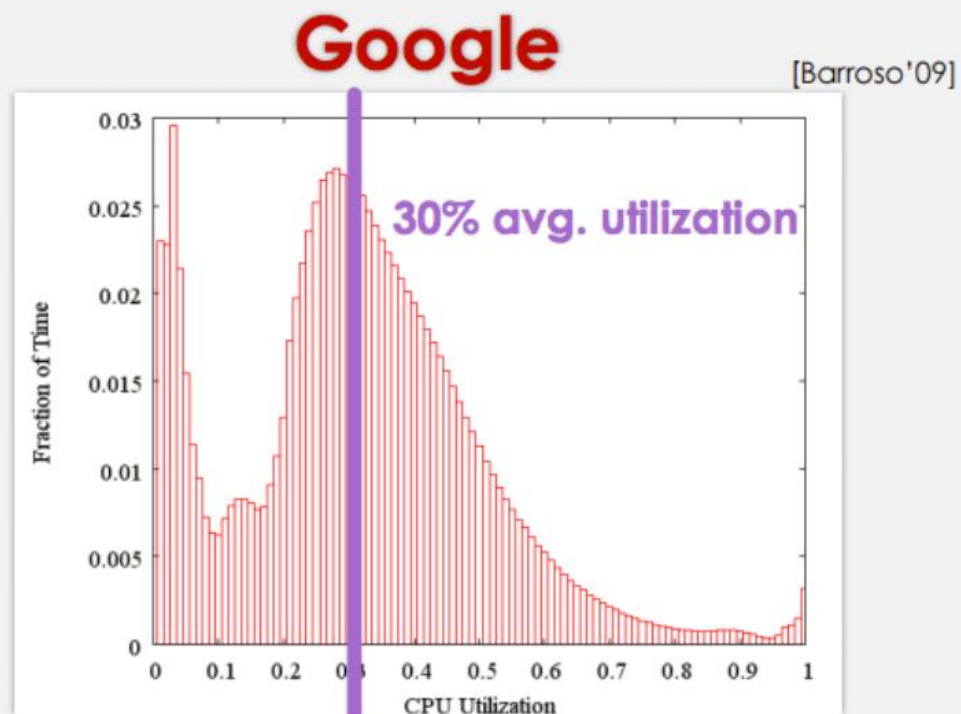
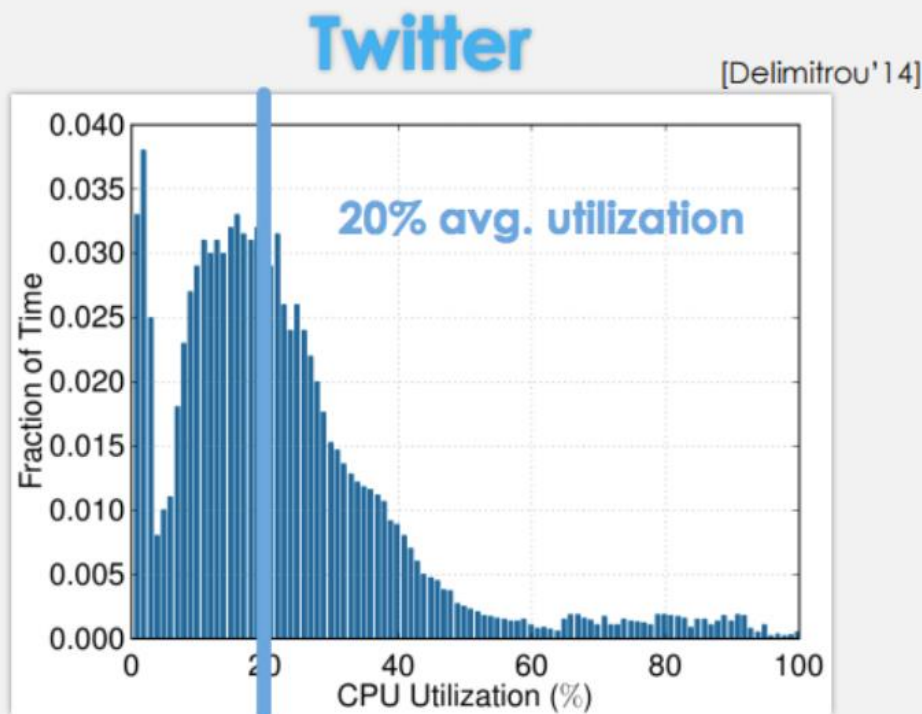
2020年中国数据中心耗电量为**2045亿千瓦时**
预计2022年将达到2700亿千瓦时，2030年达到
4000亿千瓦时

<https://www.163.com/dy/article/H8ITTKGG0512B07B.html>



2020年三峡电站全年累计生产清洁电能**1118亿千瓦时**

But the datacenters are poorly utilized!



- Low utilization in large-scale clouds, even with automated management systems

David Lo Oral Defense April 7, 2015

“软件+硬件+数据” 系统优化技术

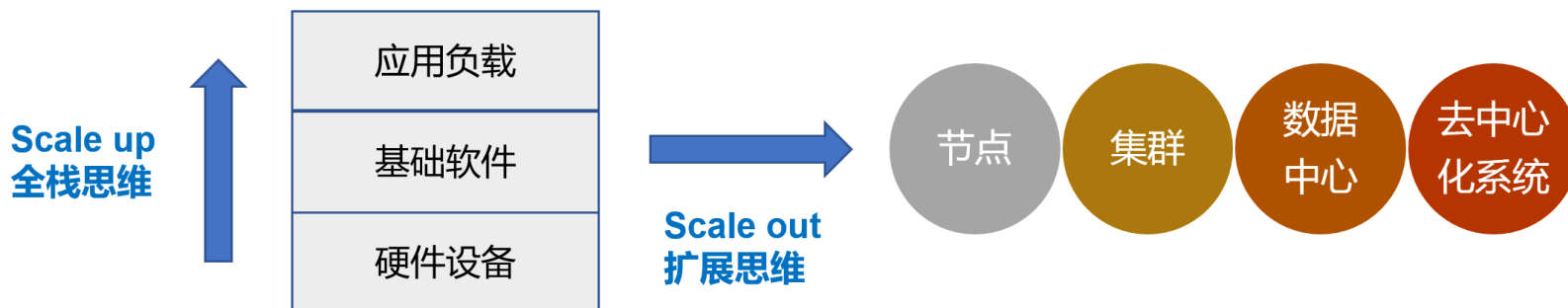
价值体现

- 性能提升
- 能耗降低
- 存储减少
- 资源利用率提升
- 运营成本降低
- 架构和系统创新



性能分析方法论

- Scale up
 - Hennessy & Patterson, “Computer Architecture: A Quantitative Approach”: **CPU time’s Iron Law & CPI breakdown method**, 1990
 - Ahmad Yasin, “A Top-Down Method for Performance Analysis and Counters Architecture”: **TMAM**, ISPASS 2014
 - Brendan Gregg, “Systems Performance: Enterprise and the Cloud”: **USE method**, 2013
- Scale out
 - Google: “**Google-Wide Profiling (GWP)**, a continuous profiling infrastructure for data centers”, IEEE Micro, 2010
 - Google: “**WSMeter**: A Performance Evaluation Methodology for Google's Production Warehouse-Scale Computers”, ASPLOS 2018
 - 数据中心混部系统的性能分析, CCF全国高性能计算学术年会HPC性能建模论坛, 2021



Google WSMeter

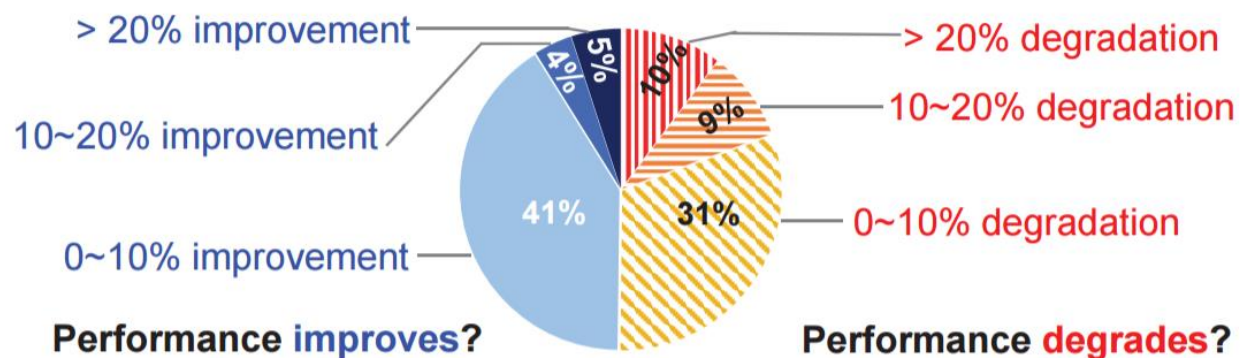
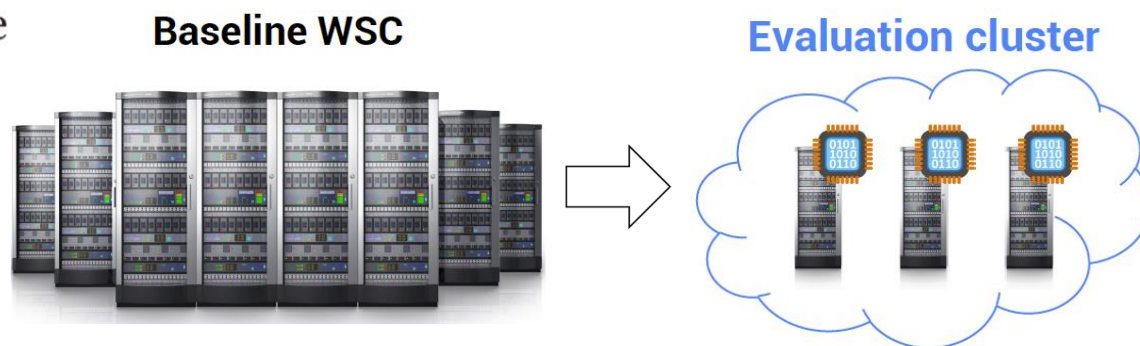


Figure 1. Performance impact of a new DVFS policy on 1000+ production jobs. The pies represent the portion of the jobs falling into the category.

$$WSMeter = \sum_i Weight_i \times IPC_i$$



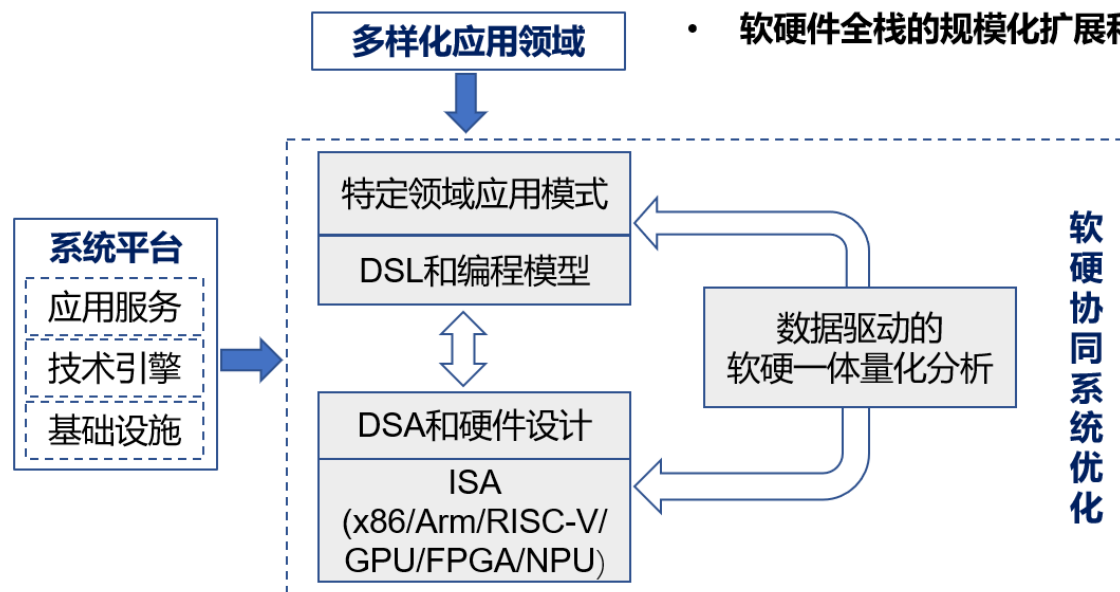
[Jaewon Lee, Changkyu Kim, Kun Lin, Liqun Cheng, Rama Govindaraju, Jangwoo Kim: **WSMeter: A Performance Evaluation Methodology for Google's Production Warehouse-Scale Computers**. *ASPLOS* 2018: 549-563]

Talks

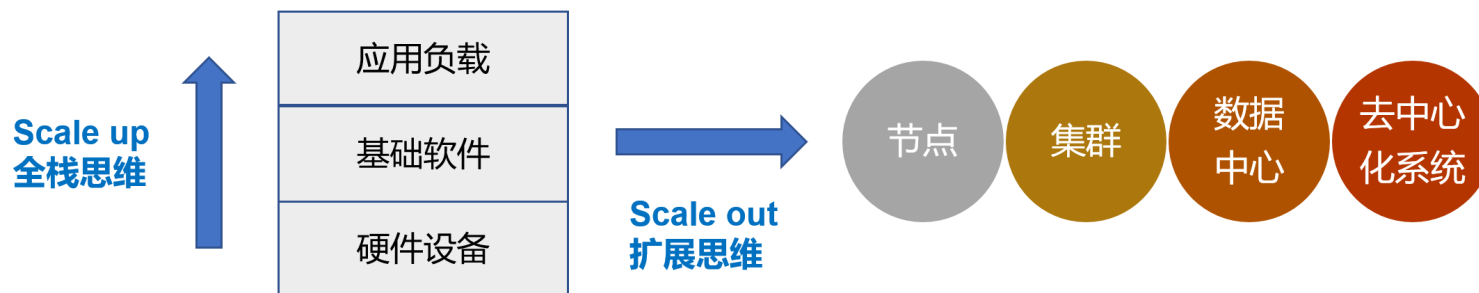
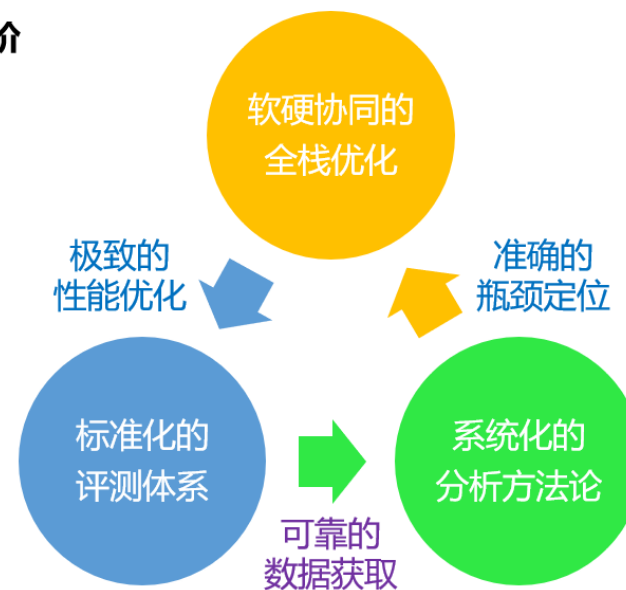
- From SPEC Benchmarking to Online Performance Evaluation in Data Centers, LTB@ICPE 2022 Keynote
<https://www.youtube.com/watch?v=dOnpJUZ4mrs&t=295s>
- 数据中心混部系统的性能分析, 2021 CCF全国高性能计算学术年会HPC性能建模论坛特邀报告
- SPEED: 大规模数据中心的性能分析平台. 2019 年杭州云栖大会阿里云系统软件开发专场报告
<https://yunqi.youku.com/2019/hangzhou/review>

Summary: 数据驱动的系统优化

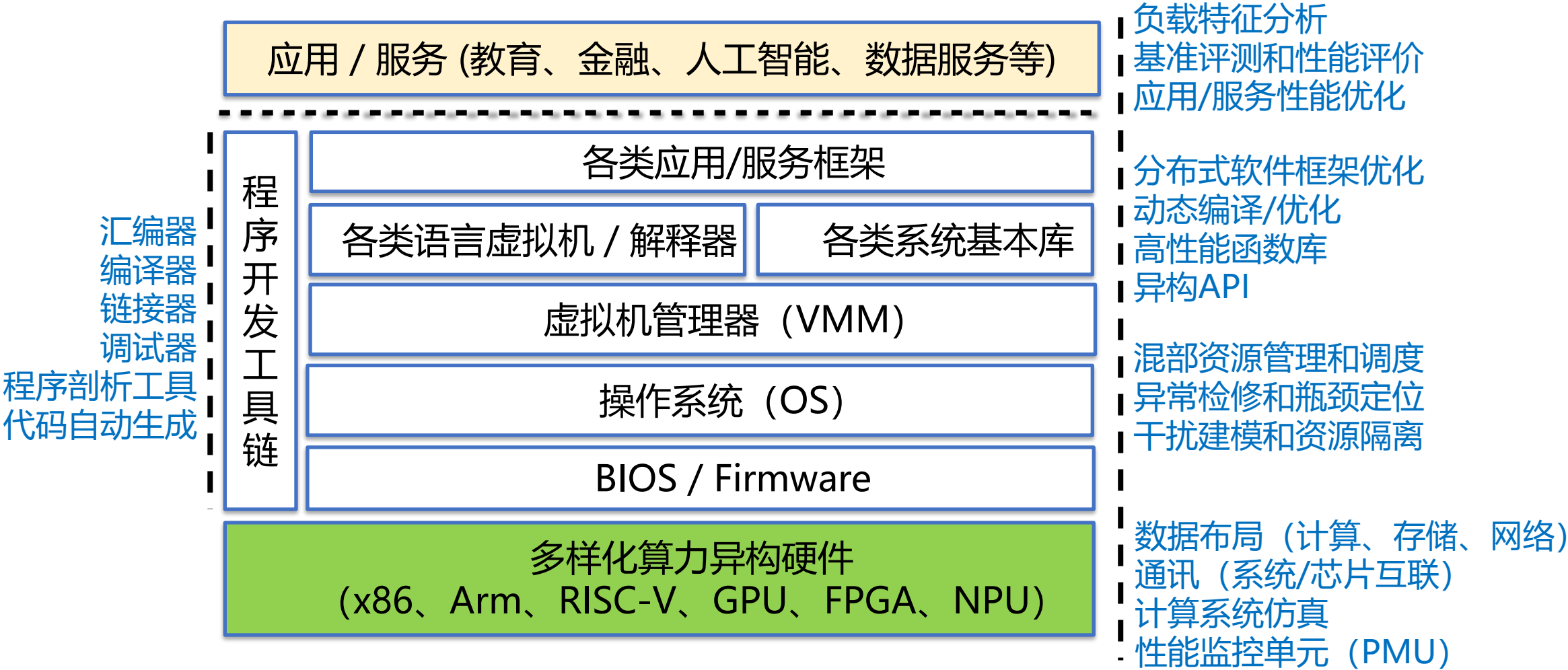
- 异构软件应用的高效编译和平滑迁移
- 异构硬件资源的统一分析和协同优化
- 软硬件全栈的规模化扩展和全局评价



“软件+硬件+数据”



Scale up: 软硬件系统全栈



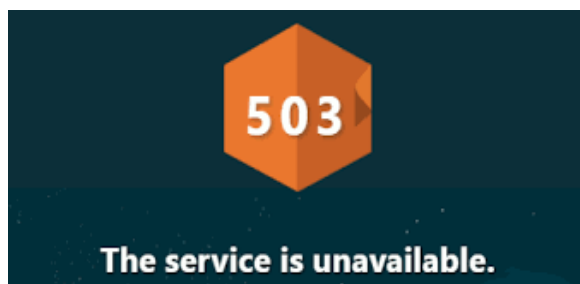
Scale out: 云数据中心全域



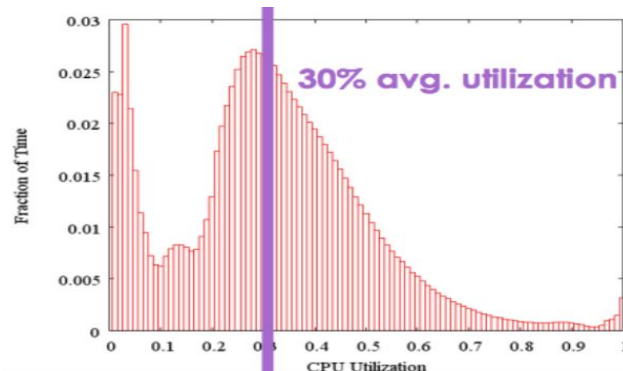
大规模异构云数据中心



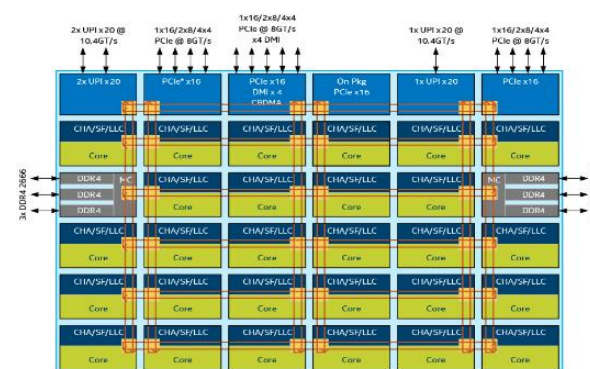
缺少全局性能评价方法



宕机和服务中断事件频发



集群资源利用率普遍低下



众核服务器上混部应用竞争加剧