

华中科技大学
Huazhong University of Science and Technology

Flex: High-Availability Datacenters With Zero Reserved Power

报告人：廖子逸
时间： 2021.12.24

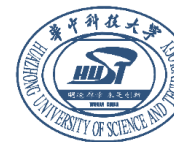


01 背景与动机

02 Flex系统设计

03 Flex系统评估

Motivation



华中科技大学
Huazhong University of Science and Technology

云服务高可用性需求 (e.g., IaaS)

数据中心冗余结构

预留部分电力浪费

充分利用预留电力
Zero Reserved Power

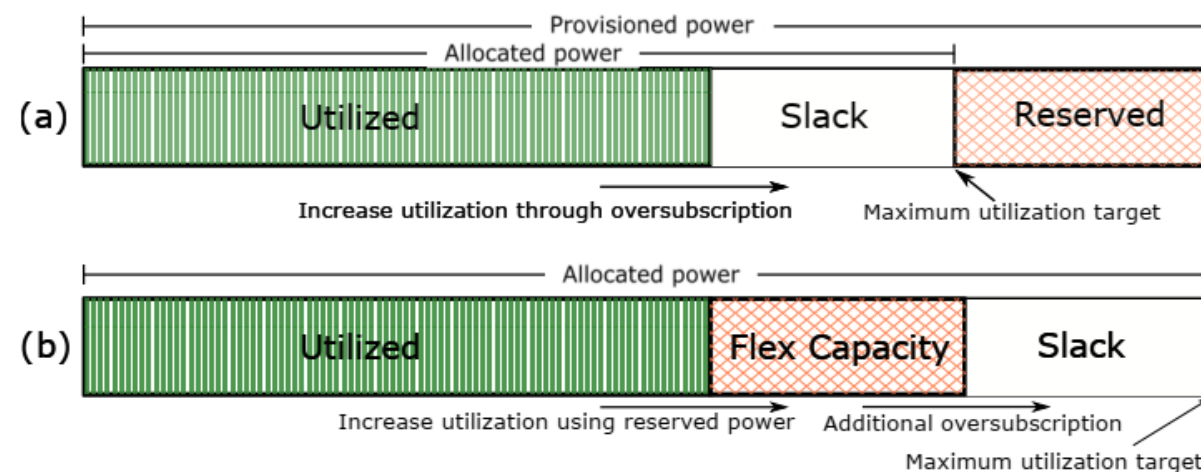
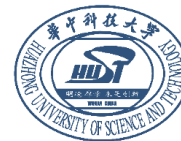


Fig. 1. Difference between traditional oversubscription (a) and Flex (b).



Three key observations

云服务的电力利用率低，很少达到峰值。且功率峰值和设备故障同时出现的概率低。

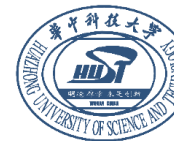
存在软件冗余负载（e.g., Web search）

有的负载可以容忍轻微性能限制（e.g., first-party VMs）

- 出现故障时，将负载快速转移。
- 应安全要求，负载需要在当前数据中心。
- 在UPS容差范围内完成负载调度。

- 提出了高可用、零预留功率数据中心的概念。
- Flex的故障期间调节方法：关闭软件冗余机架和节流可容忍性能限制机架。
- 实验证明Flex可用性。
- 通过部分实际部署，得到数据中心架构改进意见。

Distributed redundant power infrastructure



$$UPS_Allocation_Limit = UPS_capacity \times \frac{y}{x}$$

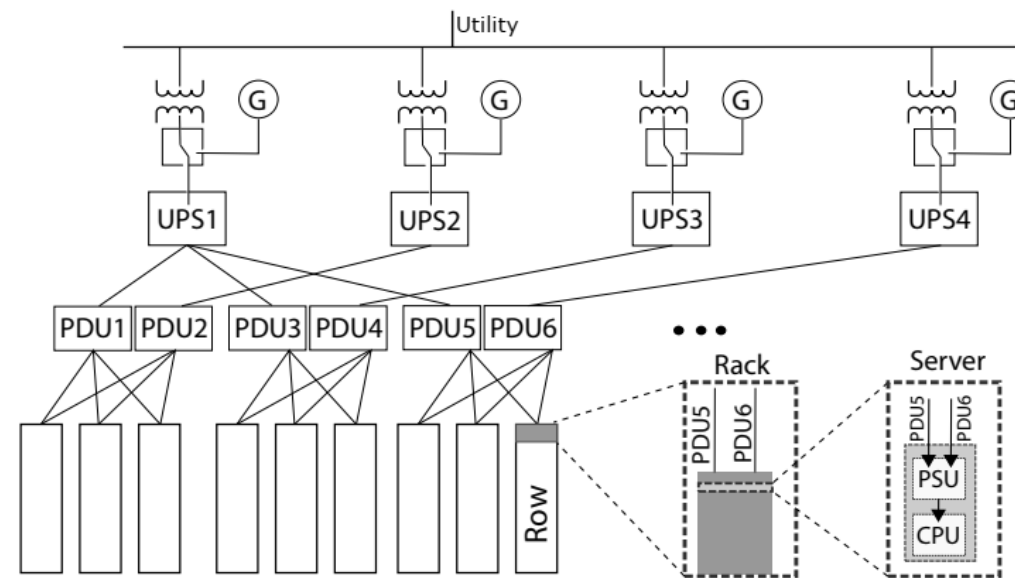
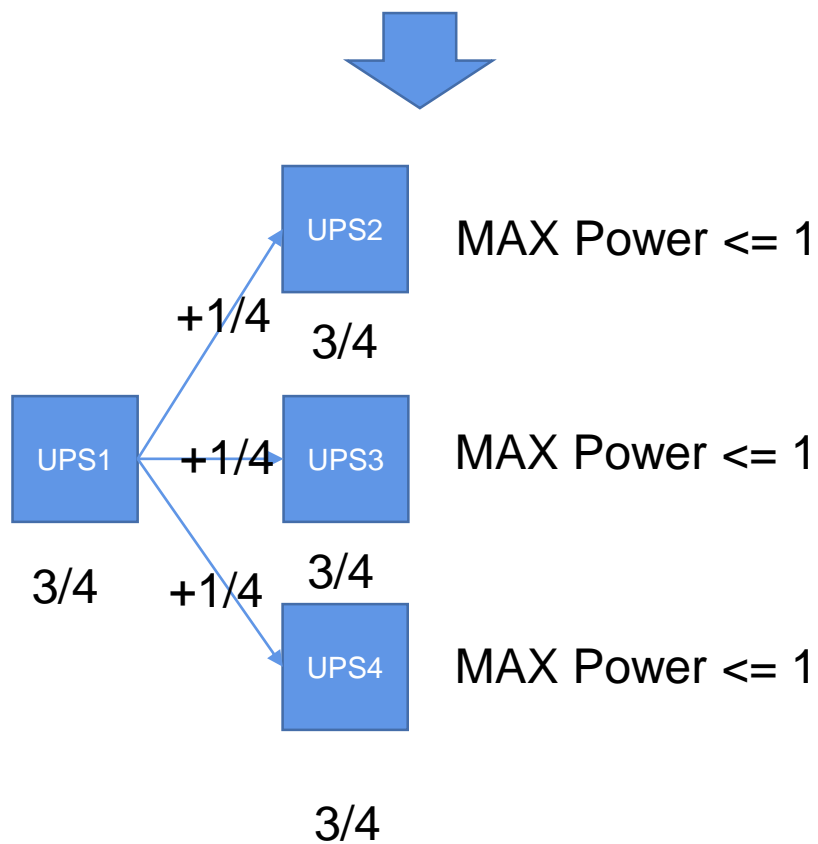


Fig. 2. 4N/3 power hierarchy design. G = generator; UPS = Uninterruptible Power Supply; PDU = Power Delivery Unit; PSU = Power Supply Unit.

如图所示是4N/3电源冗余架构

当一个UPS故障，其上负载转移给其他三个UPS

Workload分为两类：软件级别冗余（e.g. SaaS），非冗余

Hardware分为两类：硬件功率可限制，硬件功率不可限制

总结出三类负载： *software-redundant*, *non-redundant but power-capable*, and *non-redundant and non-cap-able*.

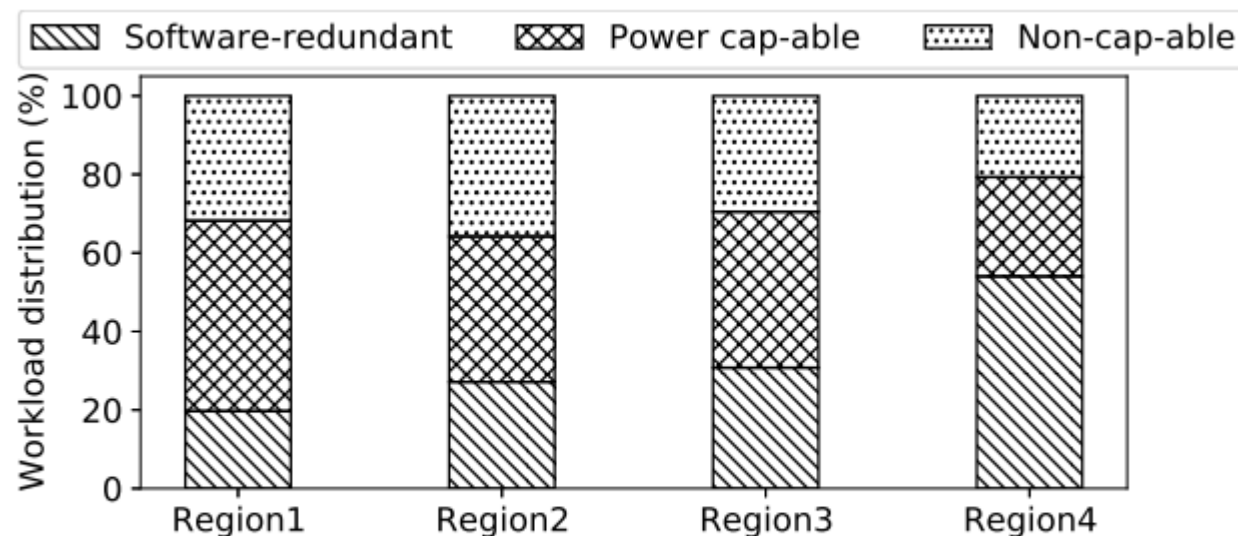
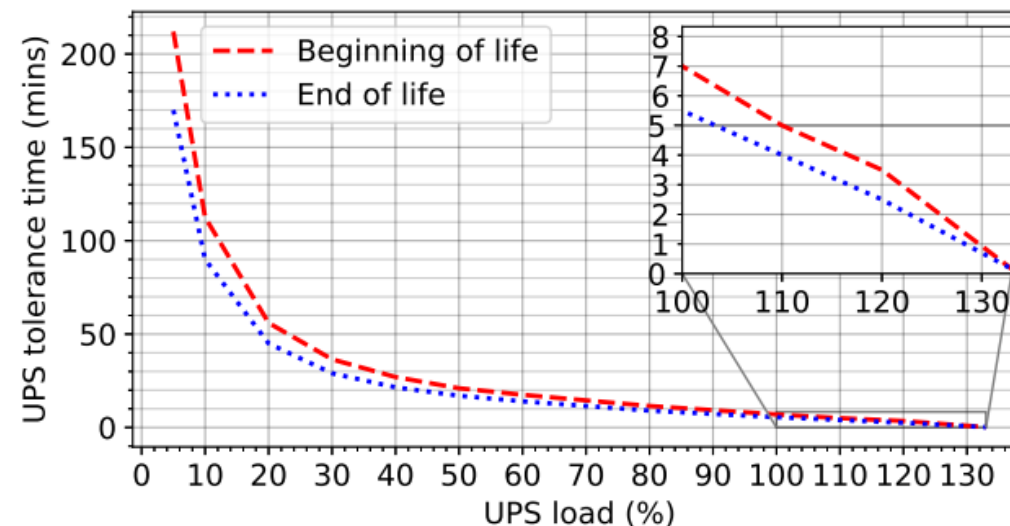


Fig. 3. Workload distribution across 4 Microsoft regions.

在部署Flex之前，集群计划内维护为 **1hour/year**

99.99%数据中心机房运行时间无需采取纠正措施。

UPS的容差曲线显示系统最大允许调度时间为
10s





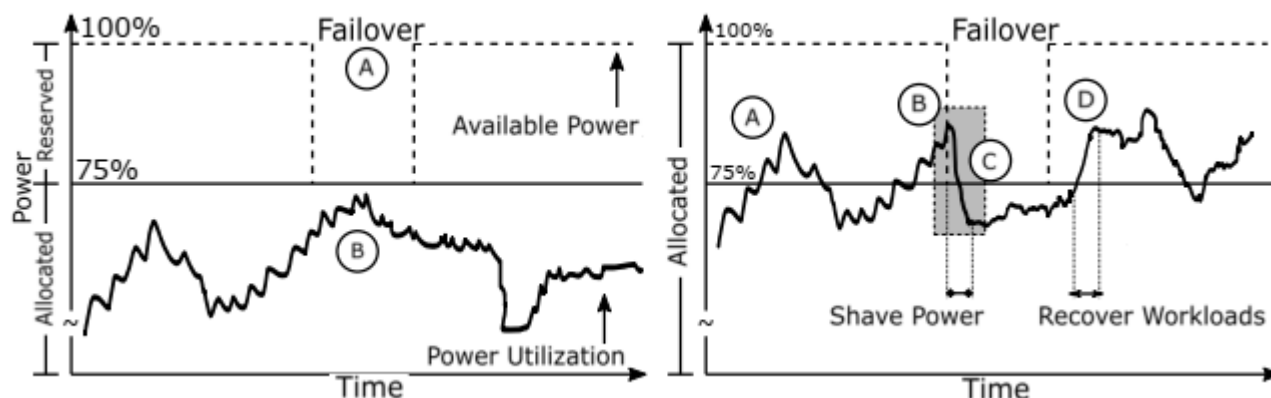
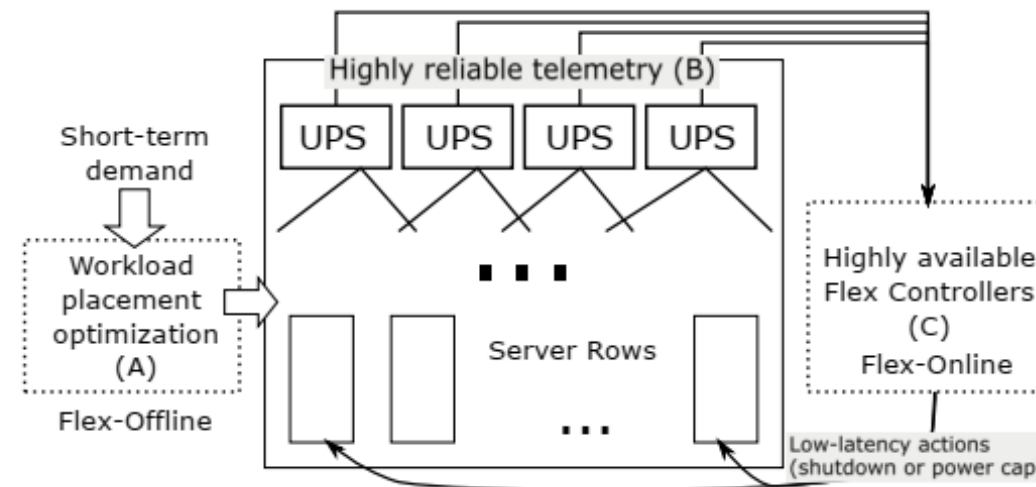
01 背景与动机

02 Flex系统设计

03 Flex系统评估

Design of Flex: Overview

- 1) 在保证安全的情况下充分利用预留电力
- 2) 在容差时间范围内采取有效行动
- 3) 调节过程对当前工作负载影响降至最低



A: Flex-Offline 负载放置优化

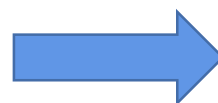
B+C: Flex-Online

高精度实时监控系统

高可用性控制器

Flex-Offline: Workload placement optimization

负载放置优化问题：如何充分利用预留电力的同时保证在故障调节阶段有足够的功率削减空间



整数线性规划问题

1. 一个负载同一时间至多部署在一个电源下。

$$\forall d \in \hat{d}, \sum_{p \in \hat{p}} P_{d,p} \leq 1 \quad (1)$$

2. 正常情况下，每个UPS的工作电力是所有连接在该UPS的负载一半，总和小于UPS额定功率。

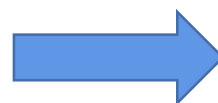
$$\forall u \in \hat{u}, \left[Load_u = \sum_{\substack{d \in \hat{d} \\ p \in \hat{p}}} 0.5 \times P_{d,p} \times Pow_d | u \in Map(p) \right] \leq Capacity_u \quad (2)$$

3. CapPow_d表示该负载所需可在故障期间最低降到多少功率。

$$CapPow_d = \begin{cases} 0 & d \in \text{Software-redundant} \\ FlexPow_d & d \in \text{Non-redundant, cap-able} \\ Pow_d & d \in \text{Non-redundant, non-cap-able} \end{cases} \quad (3)$$

Flex-Offline: Workload placement optimization

负载放置优化问题：如何充分利用预留电力的同时保证在故障调节阶段有足够的功率削减空间



整数线性规划问题

4.故障情况下，其他机器故障转移来的负载+原负载小于额定功率。

$$\forall f \in \hat{u}, u \in \hat{u}, f \neq u, \\ \left(\sum_{\substack{d \in \hat{d} \\ p \in \hat{p}}} 0.5 \times P_{d,p} \times CapPow_d | u \in Map(p) \right) + \\ \left(\sum_{\substack{d \in \hat{d} \\ p \in \hat{p}}} 0.5 \times P_{d,p} \times CapPow_d | u \in Map(p), f \in Map(p) \right) \leq Capacity_u \quad (4)$$

5.总的目标是预留电力最小化。

$$StrandedPow = \sum_{u \in \hat{u}} Capacity_u - Load_u \quad (5)$$

Highly available power telemetry pipelinee

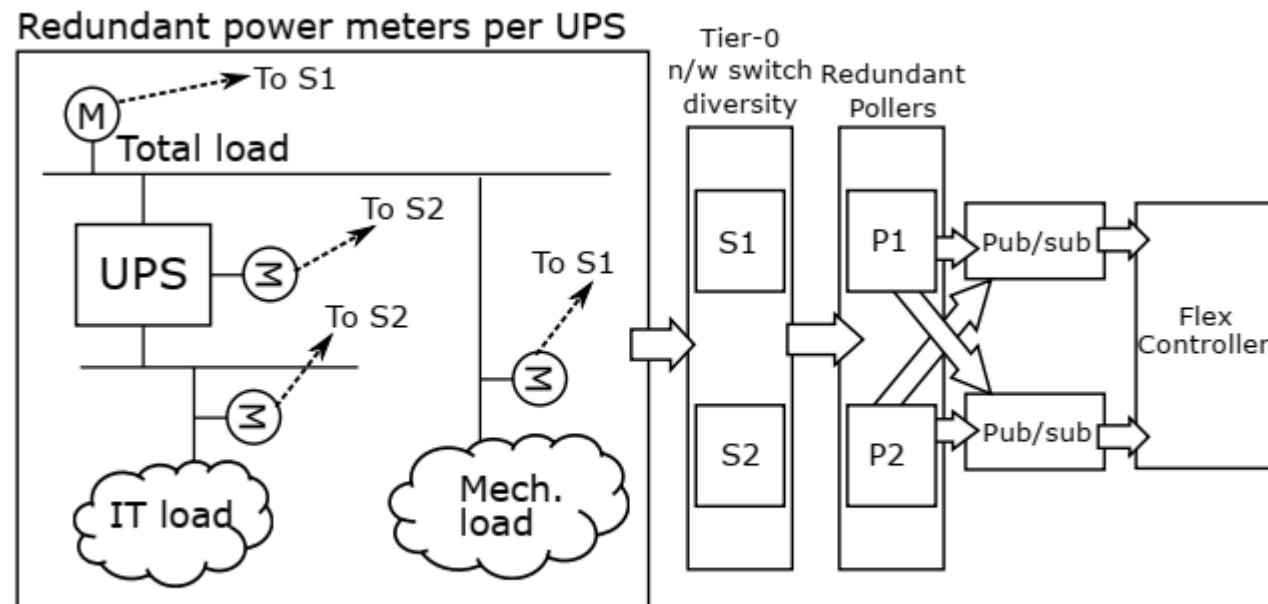


Fig. 7. Highly reliable telemetry pipeline.

高度冗余和数据通路多样性来保证监控系统 **高可靠，高精度**

Flex-Online: Runtime decision-making

Algorithm 1 Online decision policy based on impact functions

```

1:  $Actions \leftarrow \{ \}$ 
2:  $P_u \leftarrow \{UPSPowerSnapshot_u\}$ 
3:  $P_r \leftarrow \{RackPowerSnapshot_r\}$ 
4: while  $any(P_u > Limit_u - buffer)$  do
5:    $C \leftarrow \{ \}$ 
6:   for  $w \in workloads$  do
7:      $r \leftarrow PickRack(w)$ 
8:      $A_r \leftarrow \begin{cases} Shutdown & w \in \text{Software-redundant} \\ Throttle & w \in \text{Non-redundant, cap-able} \end{cases}$ 
9:      $R_r \leftarrow \begin{cases} P_r & w \in \text{Software-redundant} \\ P_r - FlexPow_r & w \in \text{Non-redundant, cap-able} \end{cases}$ 
10:     $I_w \leftarrow Impact(w, Actions \cup (r, I_w, A_r, R_r))$ 
11:     $C \leftarrow C \cup (r, I_w, A_r, R_r)$ 
12:   end for
13:    $a \leftarrow argmin_{c \in C} \{I_c\}$ 
14:    $Actions \leftarrow Actions \cup a$ 
15:    $P_u \leftarrow P_u - R_a^u$ 
16: end while
17: return  $Actions$ 

```

目标:

选择对工作负载影响最小的调节方式



操作:

循环遍历机架，计算影响值，返回影响值最小的操作集合

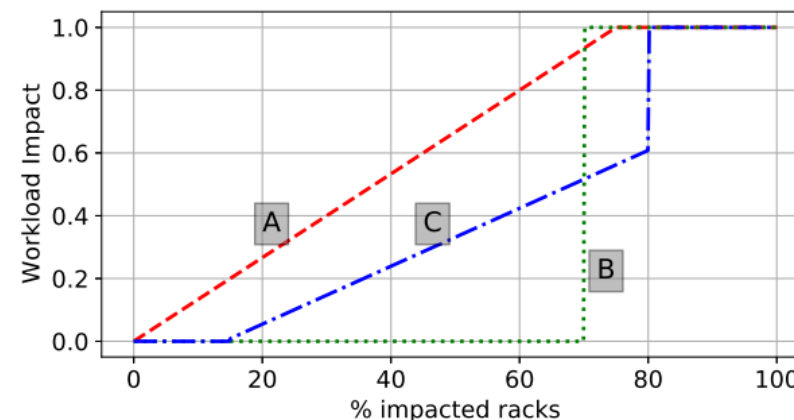
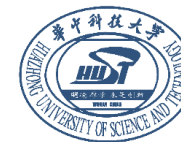


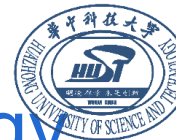
Fig. 8. Impact functions for different workloads.



01 背景与动机

02 Flex系统设计

03 Flex系统评估



Workload placement with Flex-Offline: Methodology

Workload: 4N/3电源架构, 模拟微软典型部署情况

对比策略: 随机放置、平衡轮循机制、Flex-Offline

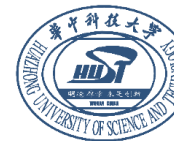


Flex-Offline-Short (33%)

Flex-Offline-Long (66%)

Flex-Offline-Oracle (entire)

Workload placement with Flex-Offline: Results



功率方面:

1. 所有搁浅功率都小于**10%**。
2. 所有Flex方案**平均搁浅功率更小**,
搁浅功率分布更集中。

节流稳定性方面:

Flex的稳定性**均值**明显比随机和轮训第

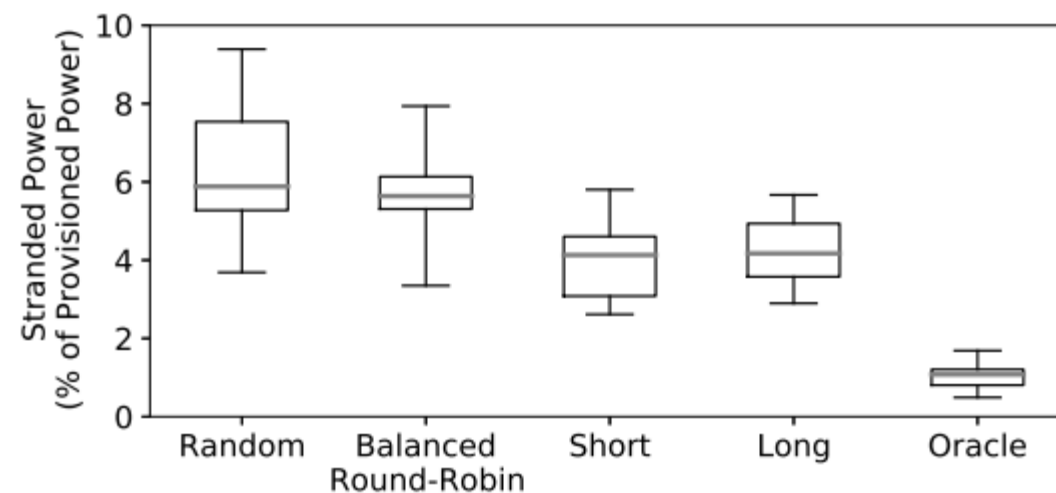


Fig. 9. Impact of placement policy on stranded power.

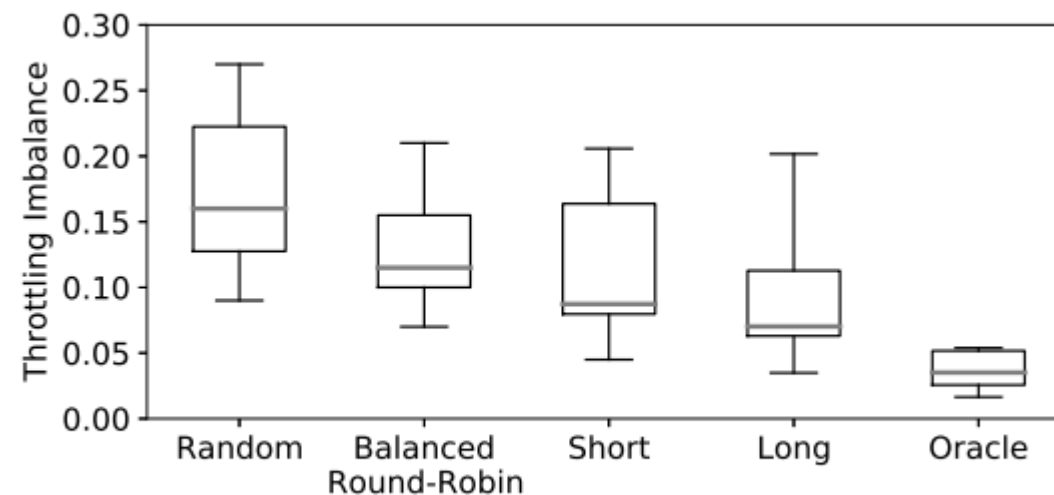


Fig. 10. Impact of placement policy on throttling imbalance.

两种极端情况：

A: 只关闭软件冗余机架

B: 只节流非冗余但可限制功率机架

两种真实情况：

C, D

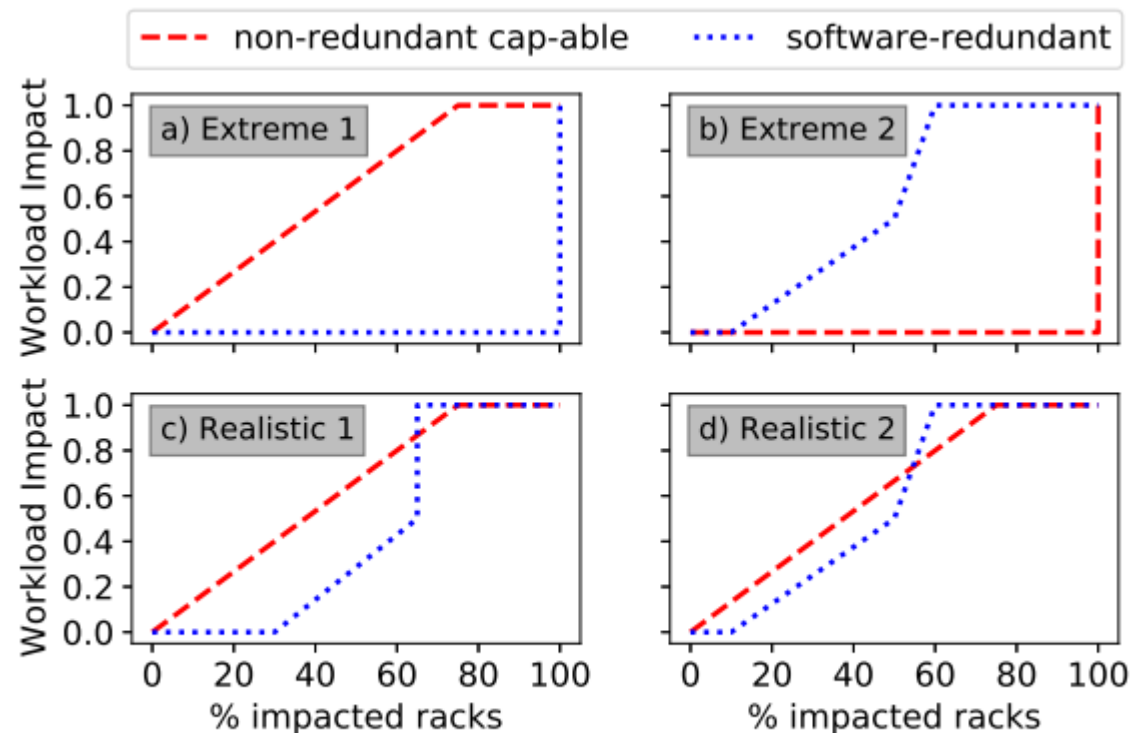
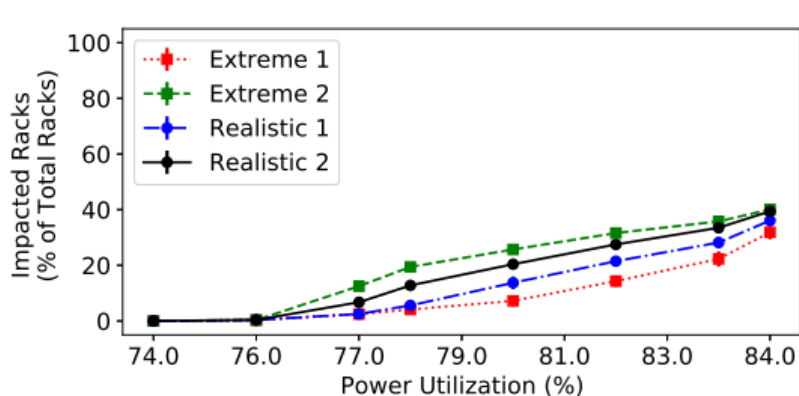
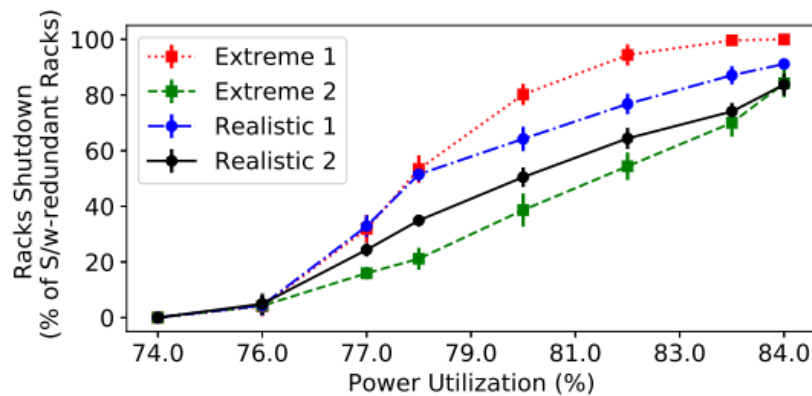


Fig. 11. Impact functions for workloads.

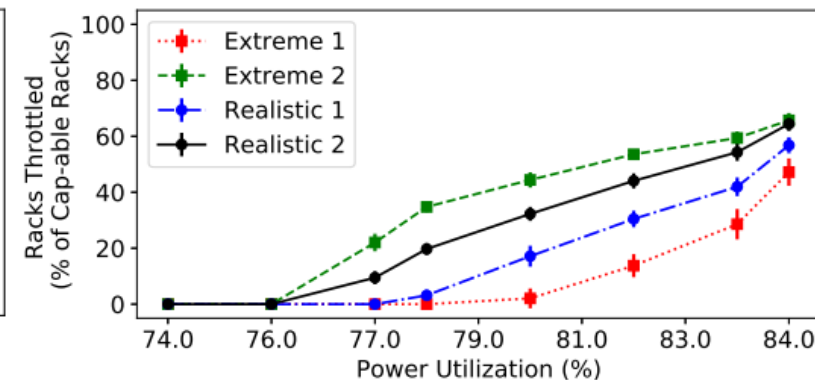
Runtime decisions with Flex-Online: Result



(a) Percent of impacted racks.



(b) Percent of racks shut down.



(c) Percent of racks throttled.

软件冗余机架关闭越多

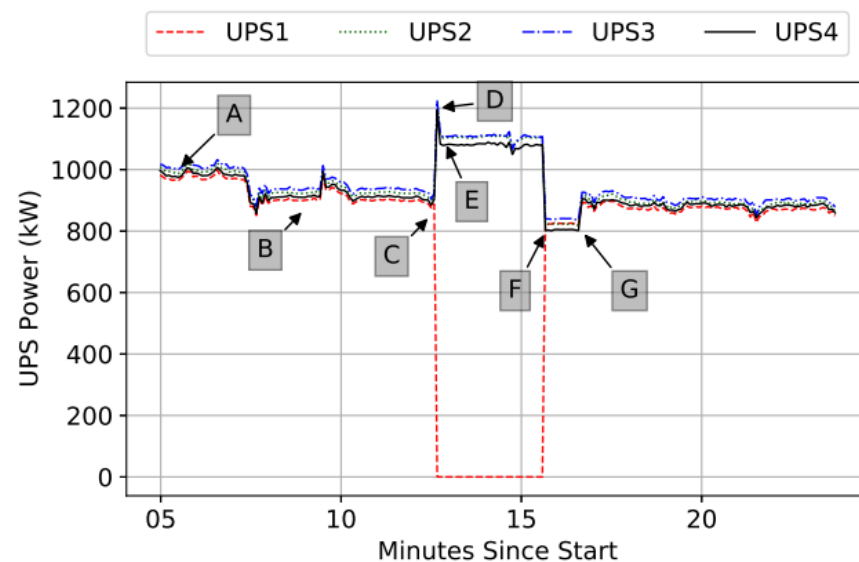
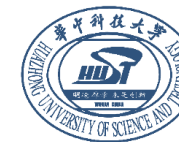
需要节流的机器越少

节流机架越多

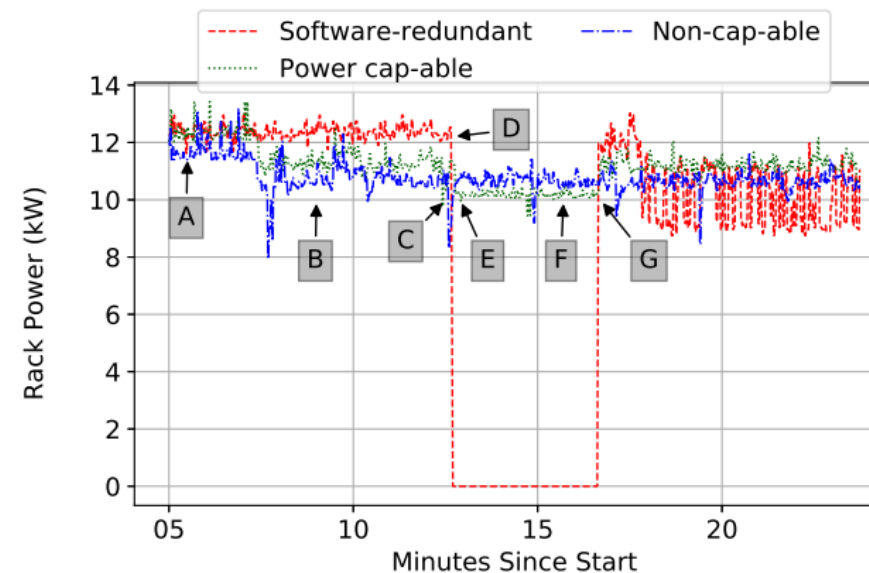
需要关闭的软件冗余机架越少

可以满足不同
负载的需求

End-to-end Flex-Online performance



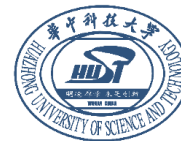
(a) UPS Power



(b) Rack Power

1200KW为额定功率。C点为故障点。

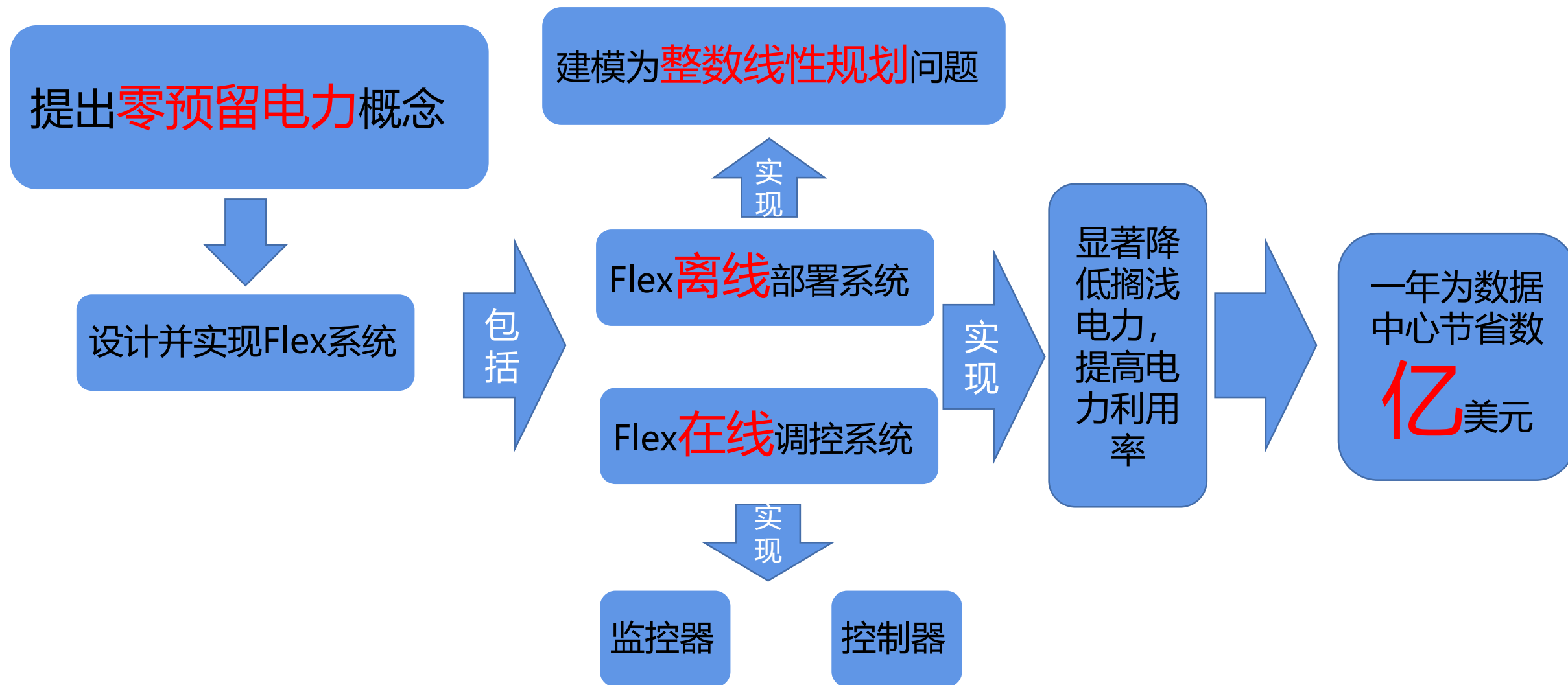
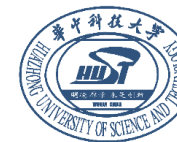
正常运行、故障发生、故障转移、恢复的过程显示Flex系统在故障下的有效调节作用

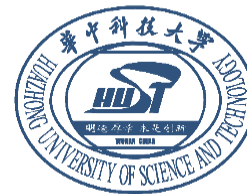


Lessons from production deployment

1. **基础设施升级**：上游设备(如中压变压器、馈线和发电机)的持续负载能力**不能支持全部负载**；提高UPS容差时间。
2. **对冷却基础设施的影响**：Flex的有效负载转移特点可以**容许冷却系统故障**。
3. **改进功率计**：现有的UPS功率计精度有限，需要设计使用**Flex专用高精度测量仪**。
4. **固件和网络状态**：关闭或节流机架依赖机架管理器(RM)和底板管理控制器(BMC)，需要**单独服务器来监控**，保证高可用性。
5. **性能特点**：从监测到故障到采取纠正措施的**尾延迟 (99.99%) : 1.5s**；DC实际部署的10 MW功率房间中测试的**尾延迟 (99.99%) : 2s**；端到端延迟：**3.5s**，均小于UPS容差时间 (**10s左右**) 。
6. **针对低可用性工作负载的财务激励措施**：正在开发新的收费模式，以**放宽性能**和**可用性要求**来**降低收费**。

Summary of the full text





華中科技大學
Huazhong University of Science and Technology

THANKS!

Question and answer session

报告人：廖子逸

时间： 2021.12.24

