

# Matrix

吕毅@百度基础架构部

# About Myself

- 天津大学
- 2007 加入百度
- 2009 Noah监控&机器管理
- 2012 Matrix
- 微博：毅LV(Lvyi@baidu.com)



# OutLines

- 面临的问题
- 解决思路
- 架构
- 经验总结
- 现状和收益

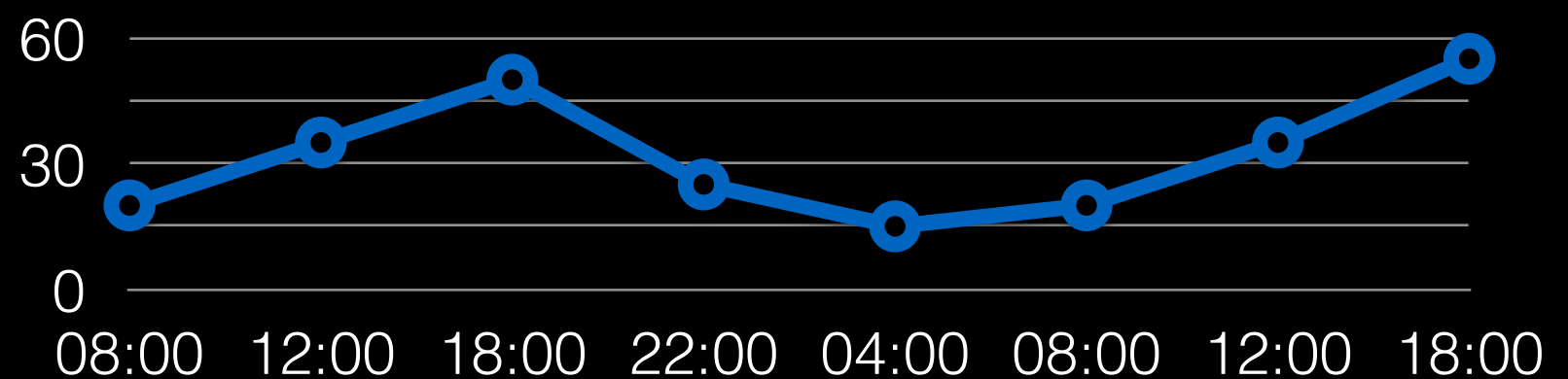
- 比肩Google
- 国内领先

# 面临的问题

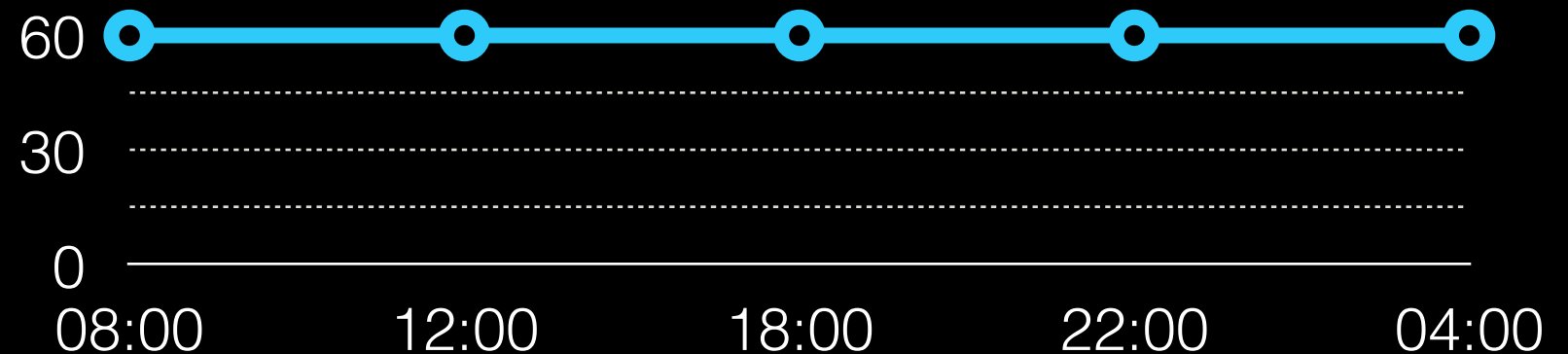
- 资源利用不均衡
- 故障处理效率低
- 预算/调配复杂

# 问题1：资源利用不均衡

- 在线业务



- 离线计算



- Buffer资源

# 问题2: 故障处理效率低

- 故障规模大，种类多
- 耦合多，维修困难

# 问题3: 预算调配复杂

- 独立业务预算
- 业务独享机器，备机
- 独立物理资产归属

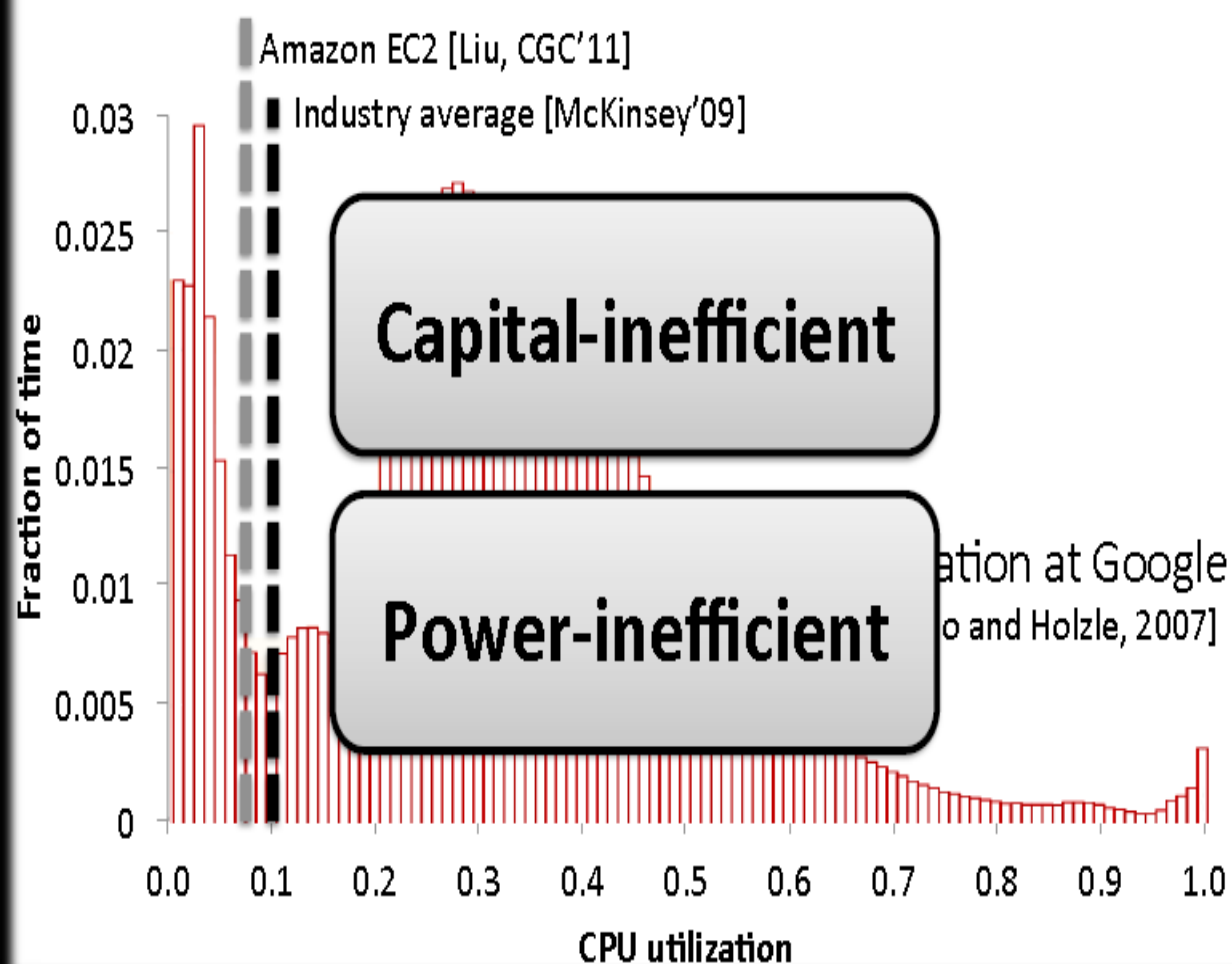


# 解决思路

- 混部：在线和离线计算混部 —>资源利用效率
- 解耦：应用和物理资源解耦 —>故障处理
- 统一：机器资源统一管理调度—>故障处理&预算

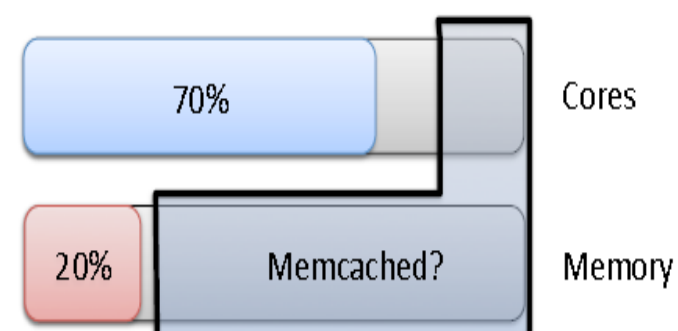
# 核心问题1：服务质量vs利用率

## Server utilization is low

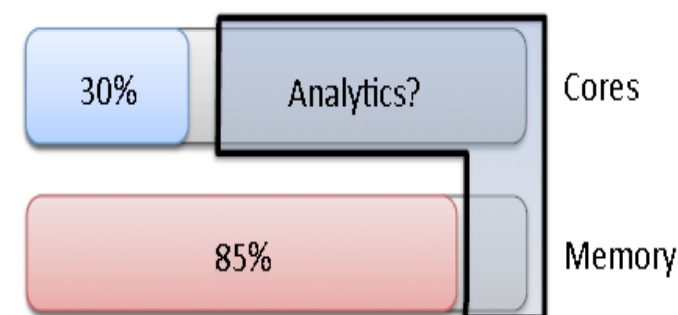


## Two consolidation examples

- Analytics cluster with unused memory

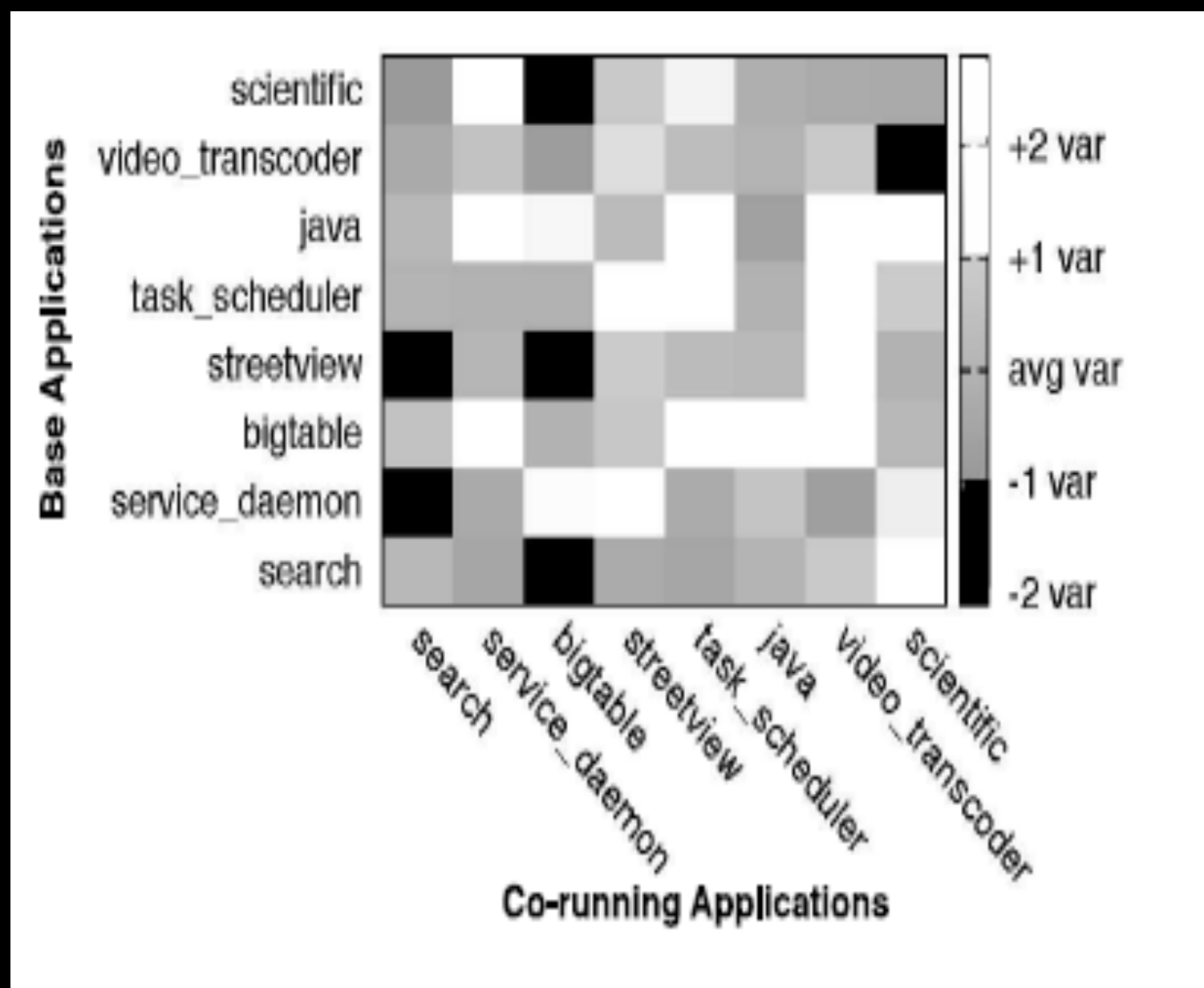


- Memcached cluster with unused CPU



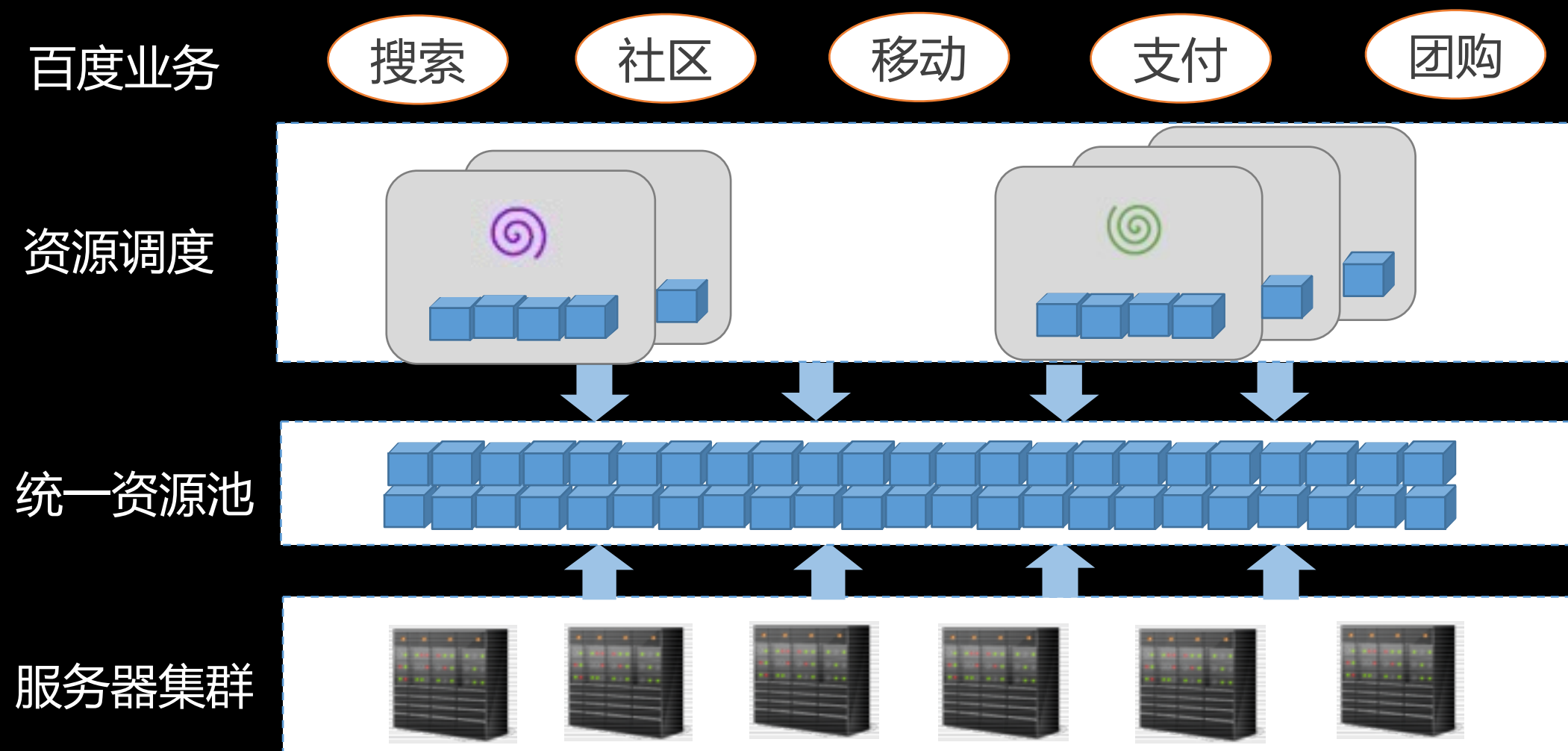
权衡两者取关键

# 核心问题2：哪些业务适合混部



结合业务和实现的混部梳理

# 总体架构



# 经验1：混部

- 混部技术
  - 混部/解耦带来的技术问题
  - 互相影响：container技术
- 关键组件
  - 服务发现：naming机制

# 经验2：弹性

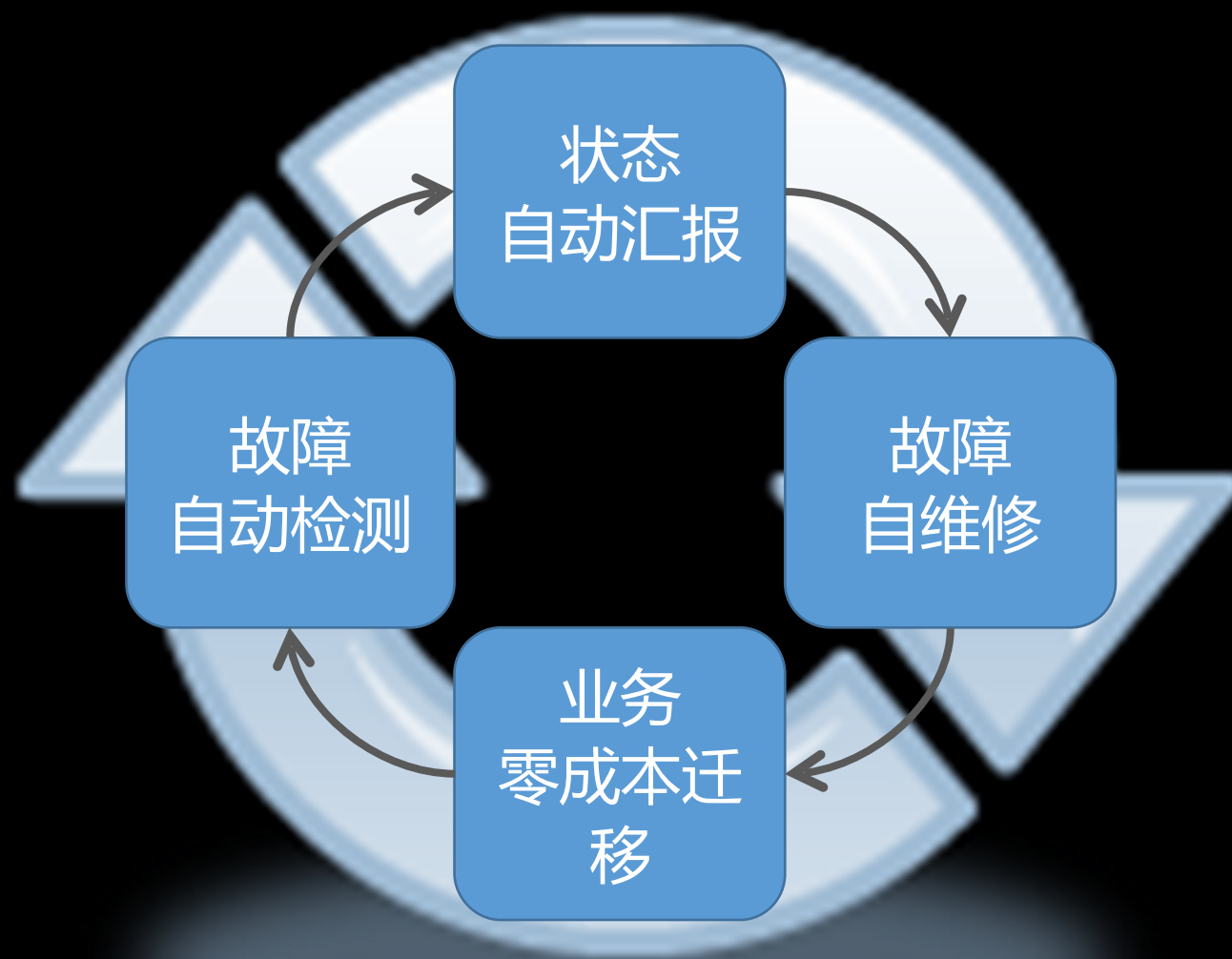
- 资源化
  - 资源替代机器的模式转变
  - 构建资源模型并应用
  - 配套的自动化系统和工具
- 自动化部署

# 经验3：应用迁移

- 自动化系统构建
- 降低业务迁移成本
  - 旧自动化系统兼容改造
  - 调度器对接成本低

# 经验4:机器自动流转

- 基础环境一致化
- 处理方式简化
- 处理流程统一

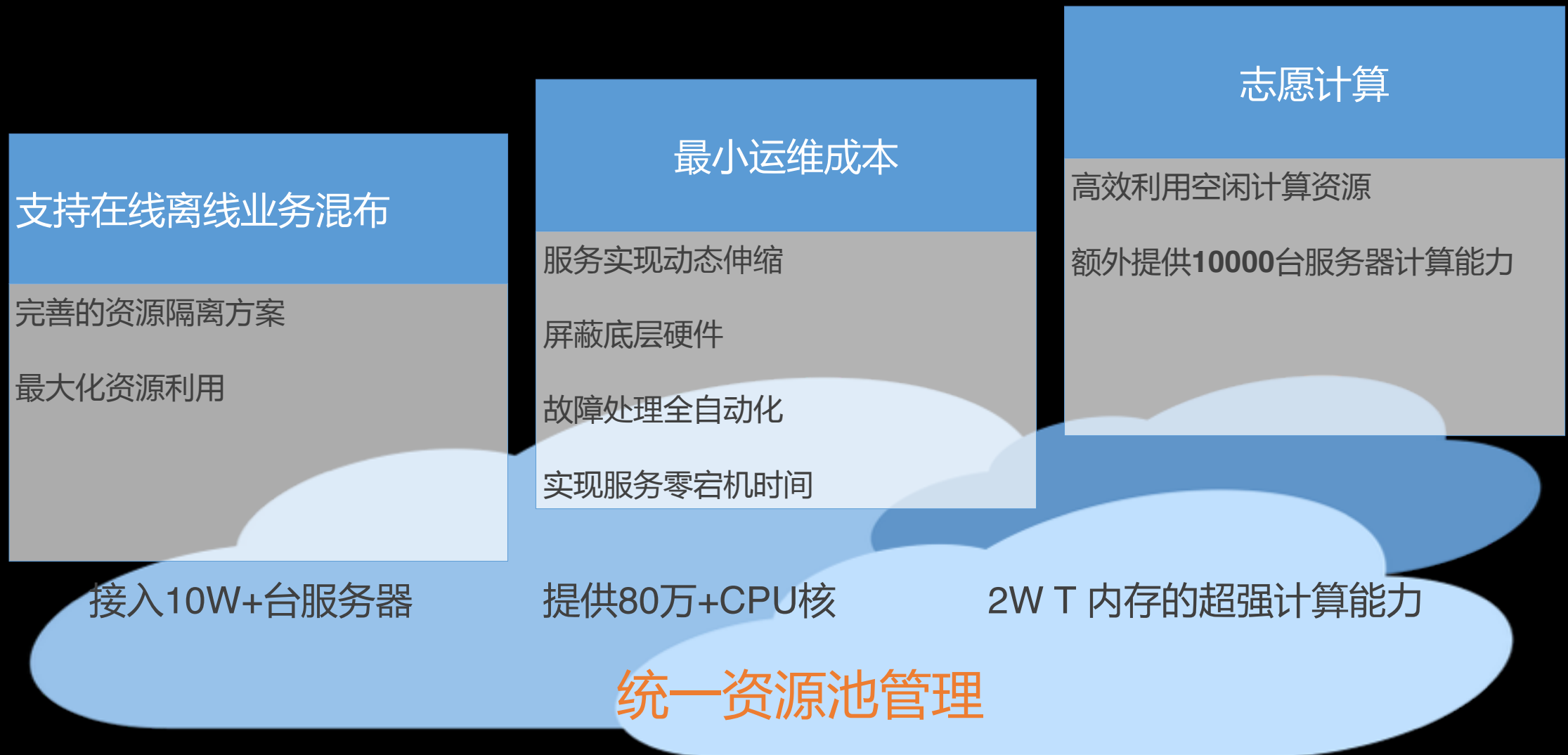


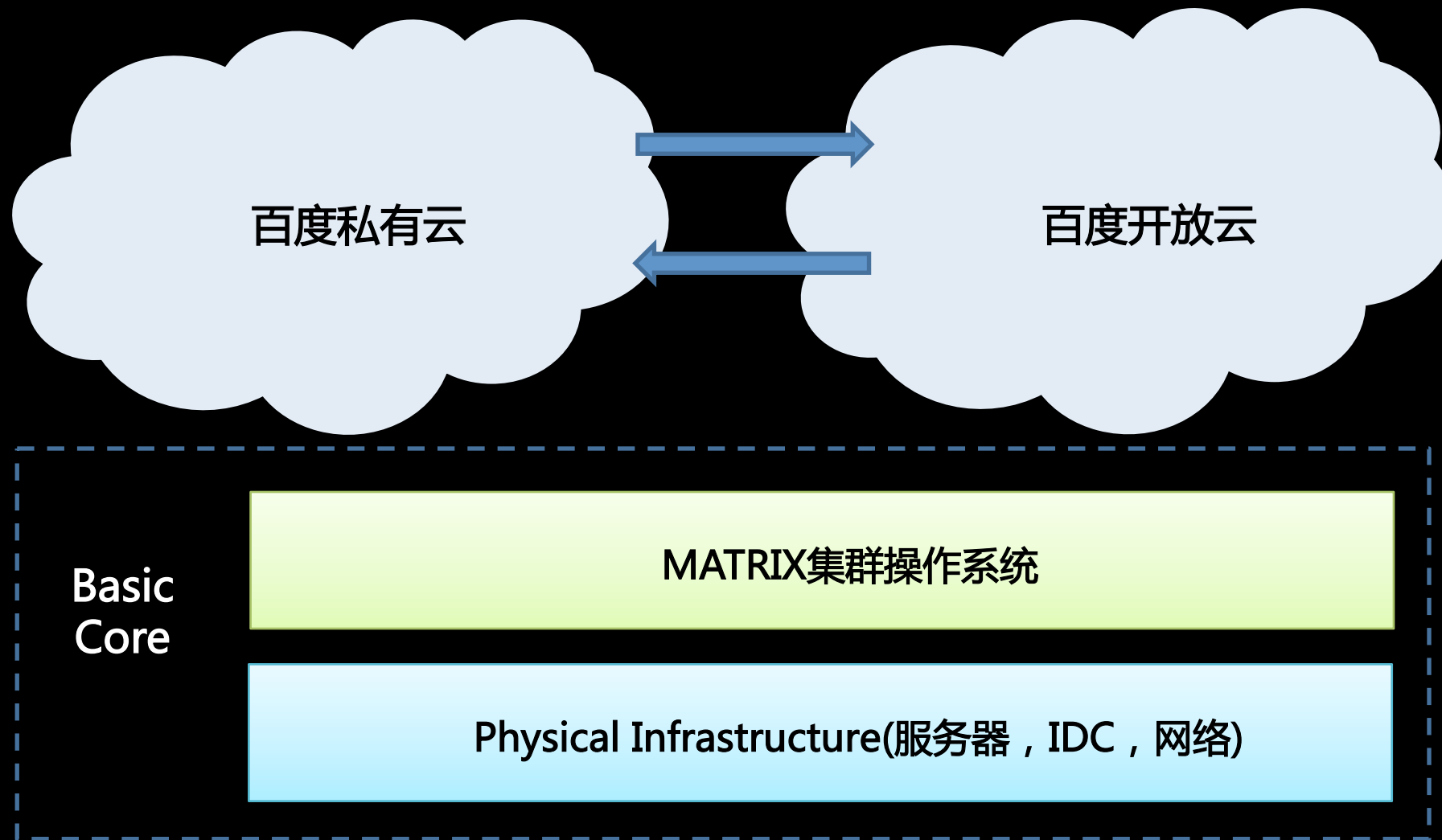


# 经验5：改进

- 设计新预算模式
- 设计新运维模式
- 业务分级和全局序

# 特性和收益





# 平台的平台

谢谢大家！

*-Lvyi@baidu.com*