

# AI 驱动 软件研发 全面进入数字化时代

中国·北京 08.18-19

AI+  
software  
Development  
Digital  
summit



# Linux全栈智能调优的好帮手

胡玉溪 KeenTune (轻豚)

# 科技生态圈峰会 + 深度研习 ——1000+ 技术团队的选择



2023K+  
全球软件研发行业创新峰会  
上海站

会议时间 | 06.09-10



2023K+  
全球软件研发行业创新峰会  
北京站

会议时间 | 07.21-22



2024K+  
全球软件研发行业创新峰会  
深圳站

会议时间 | 05.17-18



K+峰会详情



会议时间 | 08.18-19

AiDD AI+软件研发数字峰会  
北京站



会议时间 | 11.17-18

AiDD AI+软件研发数字峰会  
深圳站



AiDD峰会详情

# ▶ 演讲嘉宾



## 胡玉溪

KeenTune产品发起人、负责人；阿里云高级技术专家

- 龙蜥社区KeenTune SIG发起人、Maintainer
- 阿里云高级技术专家、钻石布道师
- 前华为测试专家、测试经理
- 发表论文： ICSE/FSE'22 《Industry Practice of Configuration Auto-tuning for Cloud Applications and Services》 ; ACM/ISSTA'23 《KeenTune: Automated Tuning Tool for Cloud Application Performance Testing and Optimization》

# 目录

## CONTENTS

1. 性能调优的重要性和难点
2. KeenTune产品介绍及特点
3. KeenTune性能调优最佳实践
4. 总结展望 & 开源共建

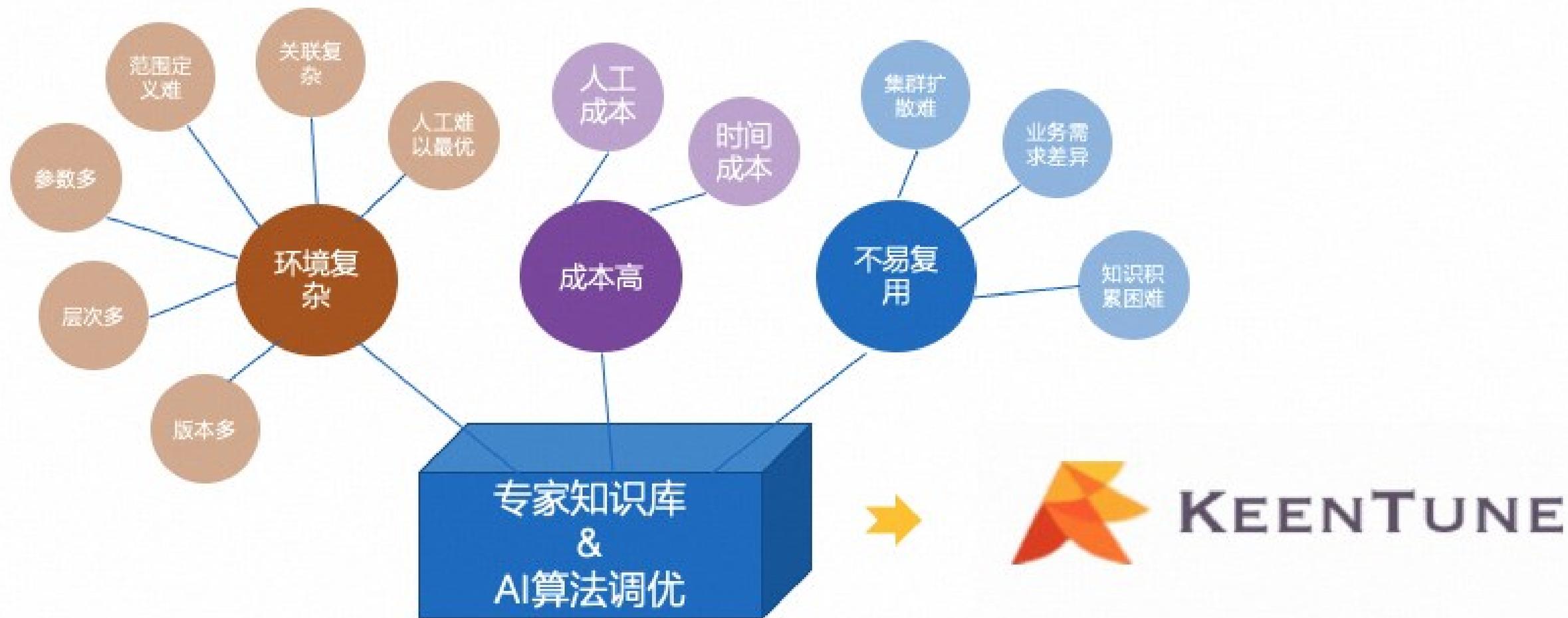
# PART 01

# 性能调优的重要性和难点

问题·策略



# ▶ 性能调优的问题 & 解法

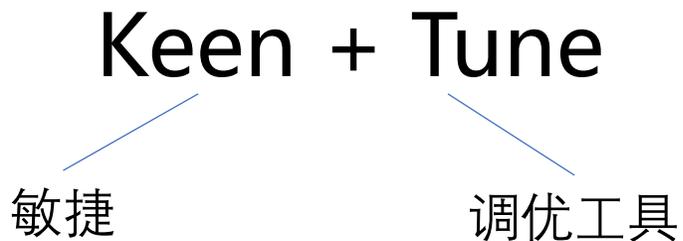


## PART 02

# KeenTune产品介绍及特点

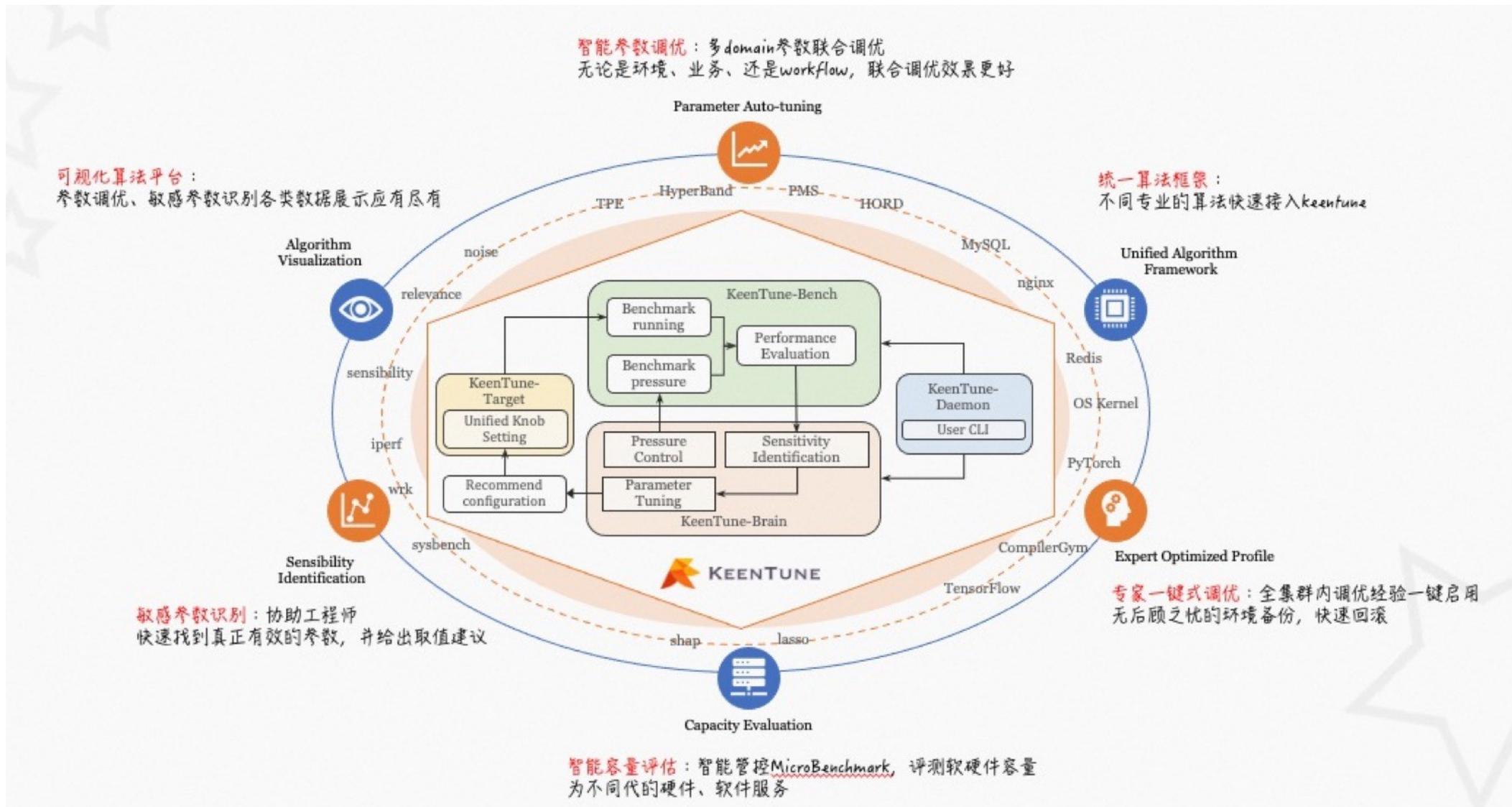
产品·特点

# ▶ KeenTune产品介绍

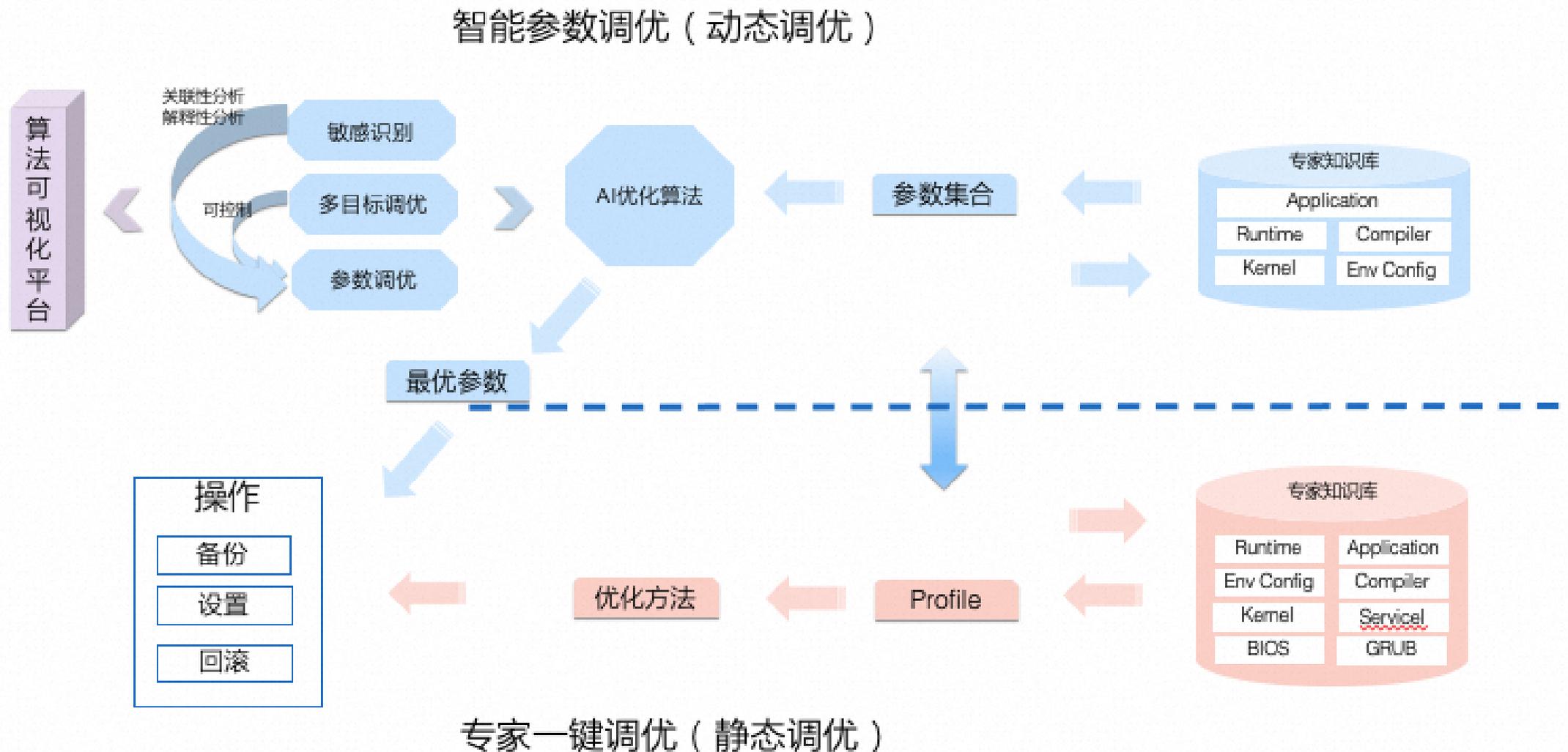


KeenTune 是一款AI算法与专家知识库双轮驱动的操作系統全栈式智能优化产品，为主流的操作系统提供轻量化、跨平台的一键式性能调优，让应用在智能定制的运行环境发挥最优性能。

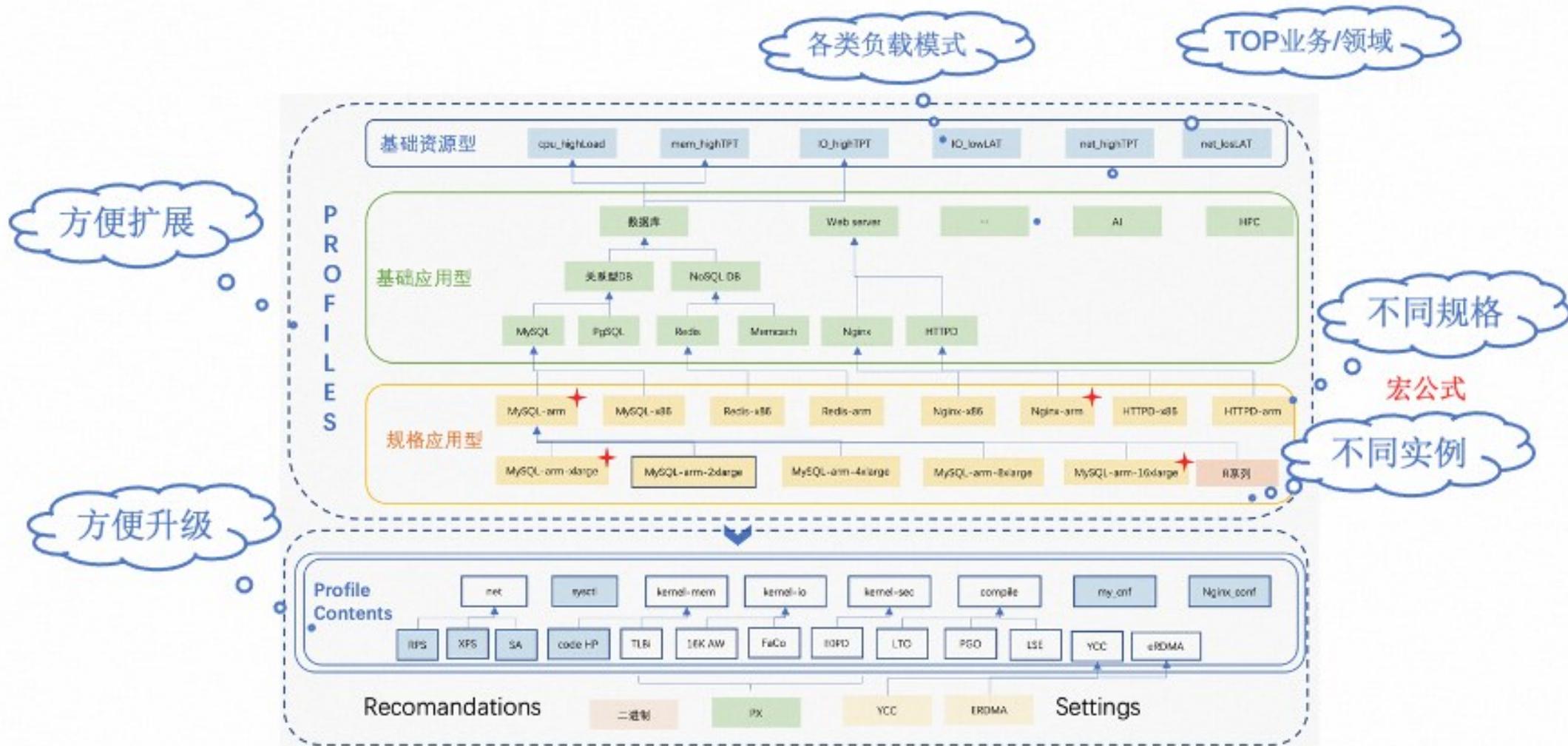
# ▶ KeenTune (轻豚) : 三大特性, 三个方向



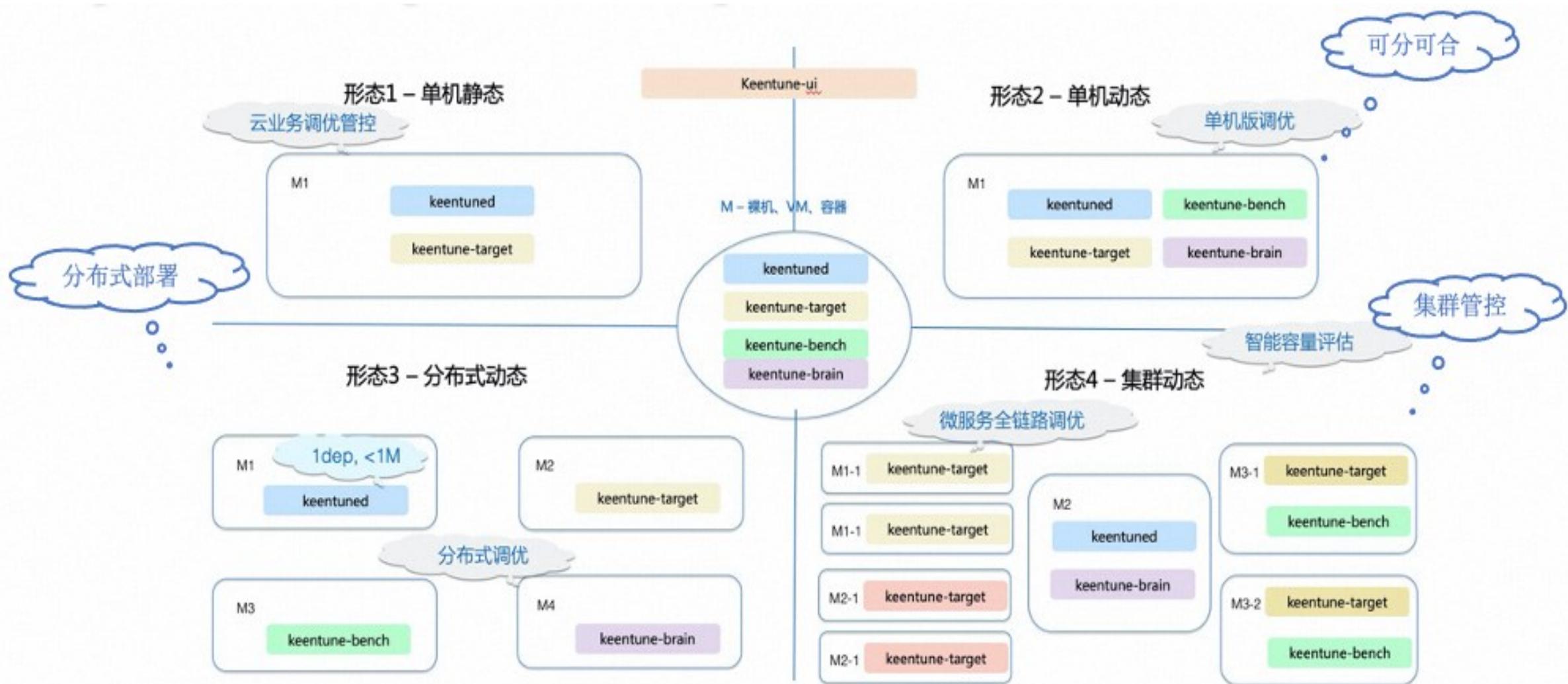
# ▶ 特点1. 动静态协同调优



# ▶ 特点2. 完备的全栈调优专家知识库



# ▶ 特点3. 模块化&集群管控

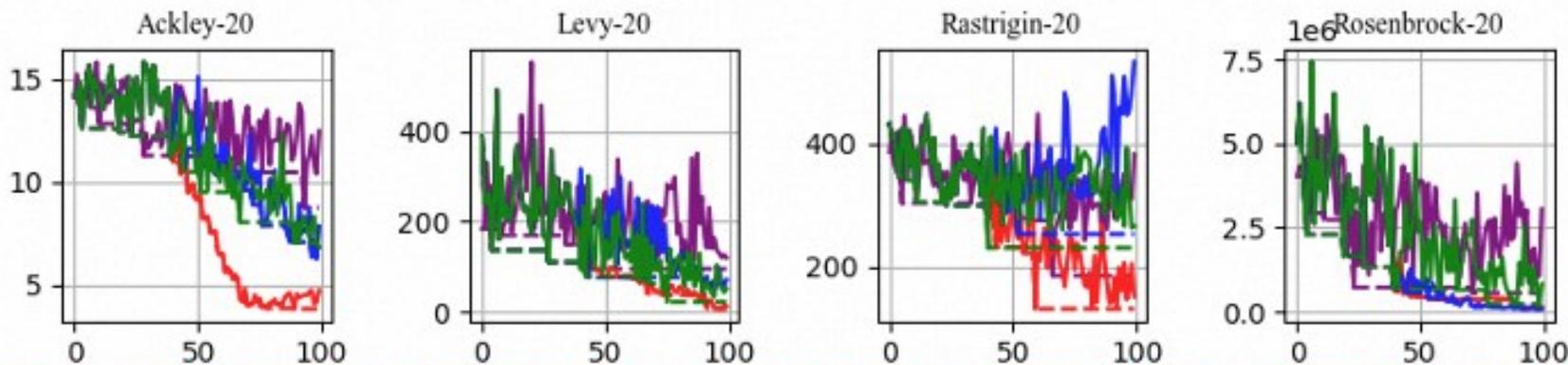


## ▶ 特点4. 多领域算法统一框架

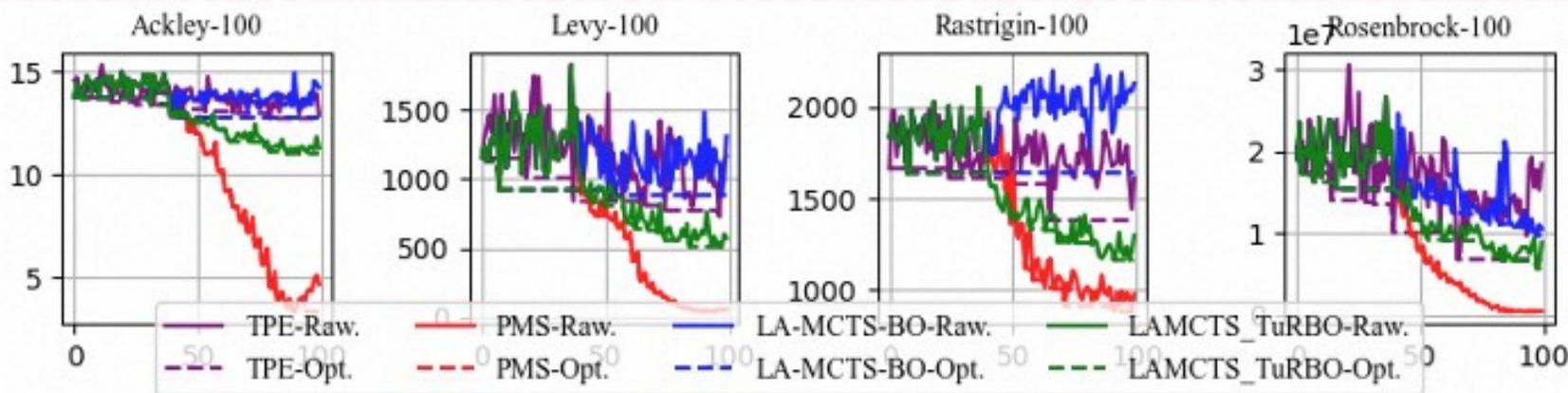


# ▶ 特点5. 自研高效算法

低维



高维



业务常用100  
维以上

# ▶ 特点6. 管控 & 算法可视化



性能调优工程师

## 性能调优 & 算法优化

工具 >>> 平台

**一键式专家调优**  
通过多种场景下海量任务的专家知识，根据业务类型对系统进行一键式调优。

**智能参数调优**  
提供多种高级算法，根据业务运行状态智能调整参数，使业务运行在最佳环境中。

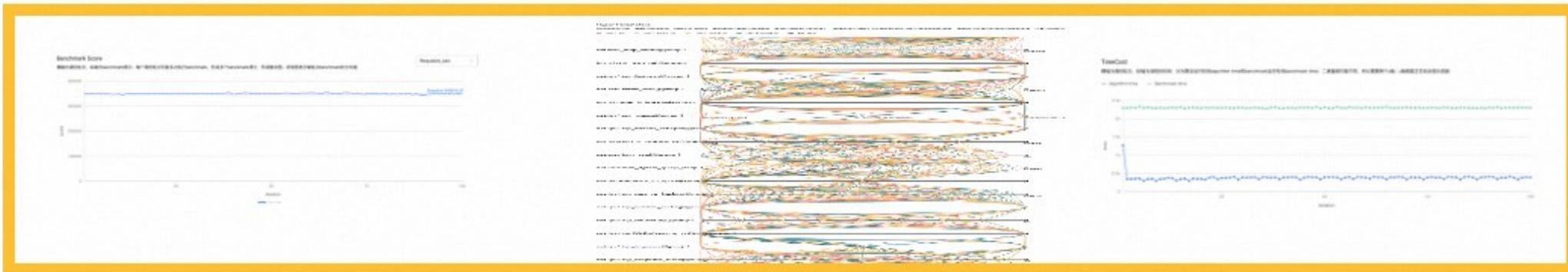
**敏感参数识别**  
对参数进行敏感性分析，识别出对性能影响最大的参数，并优化其取值范围，辅助参数可解释性调优。

请前往官网访问: <http://keentune.kylinos.com>

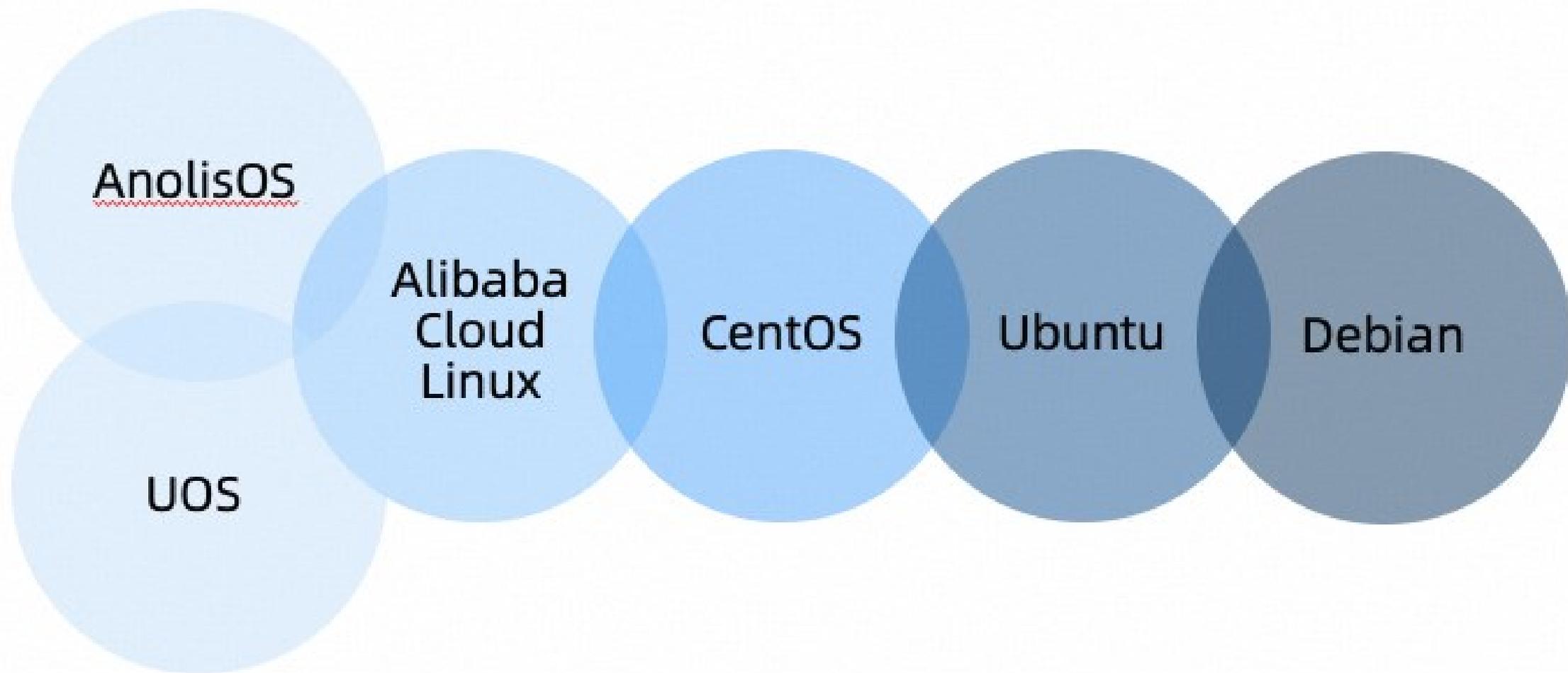
算法工程师 (超参调优)



算法工程师 (参数可解释)



## ▶ 特点7. 跨语言 & 适配主流OS



## PART 03

# KeenTune性能调优最佳实践

背景·拓扑·效果

# ▶ 场景1. 阿里云 ECS 倚天 “应用加速” 功能



## 多AMI优化不同场景

- 场景化：SQL、TensorFlow、Pytorch、HVM
- 不足：维护人力多，难以覆盖更多场景

### ▼ 应用程序和操作系统映像(Amazon Machine Image) 信息



#### Amazon Linux 2 AMI (HVM) - Kernel 4.14, SSD Volume Type

ami-07251f912d2a831a3 (64 位(x86)) / ami-093dde316d91535f5 (64 位(ARM))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs

#### Amazon Linux 2 LTS with SQL Server 2017 Standard

ami-0bea0be466a642cf0 (64 位(x86))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs

#### Amazon Linux 2 with .NET 6, PowerShell, Mono, and MATE Desktop Envi

ami-07c7efcf843300087 (64 位(x86))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs

#### Amazon Linux 2 LTS with SQL Server 2019 Standard

ami-0a4f85fc401a53c3d (64 位(x86))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs Amazon Linux 2 LT

#### Deep Learning AMI GPU PyTorch 1.12.0 (Amazon Linux 2) 20220630

ami-0b28559abcc7fe7a9 (64 位(x86))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs

#### Deep Learning AMI GPU TensorFlow 2.9.1 (Amazon Linux 2) 20220628

ami-09a62090c73f45ec9 (64 位(x86))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs

#### Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type

ami-02d1e544b84bf7502 (64 位(x86)) / ami-03e57de632660544c (64 位(ARM))  
虚拟化: hvm 已启用 ENA: true 根设备类型: ebs



## 单OS镜像+优化选项

- Web、DB、cache、Media、密集计算
- 方案：单OS镜像+场景化配置+打包优化版RPM

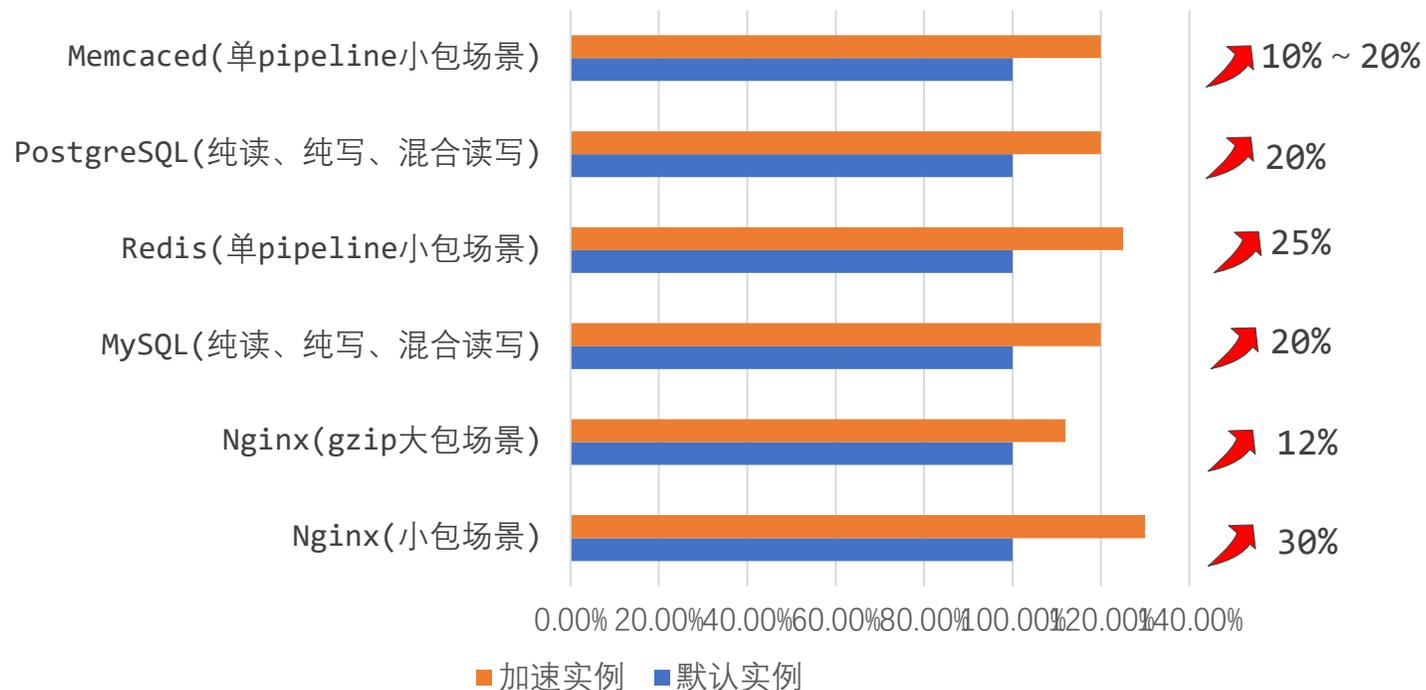


### 实例创建时，选择OS场景 (Cloud-init增加选项)

- 需求1：cloud-init如 `cloud-init` 可通过云助手、Cloudinit等通道下发
- ✓ Web优化场景：Nginx、Apache、NodeJS、WordPress (默认配置XPS、E0PD、BHB；eg. Nginx收益23%)
- ✓ DB优化场景：MYSQL、PG (中断绑核策略、进程Migation策略、BHB优化；eg.mysql +10%)
- ✓ 缓存场景：Redis、Memcached (RFS、XPS, eg. Redis收益20%)
- ✓ Media场景：Fmpeg (载入编译优化rpm, x264-opt, 性能+20%)
- ✓ 集群计算密集：TensorFlow、Pytorch、Spark (eRDMA技术, +20%)

# ▶ 场景1. 阿里云 ECS 倚天 “应用加速” 功能

已上线应用的调优效果

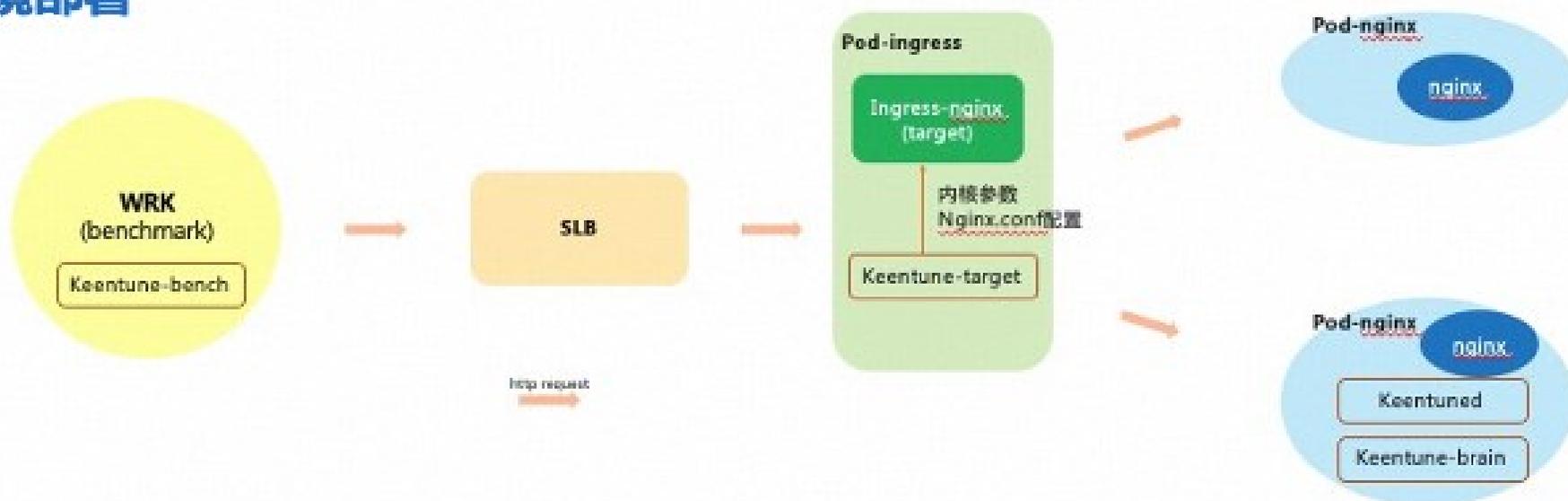


# ▶ 场景2. 云原生Ingress-Nginx全栈调优

## 背景介绍

- **业务背景**：阿里云K8S容器大客户中ingress-nginx业务占比高达90%+
- **调优场景**：http(s)长/短连接，在不使时延变长的情况下，提升吞吐量

## 环境部署

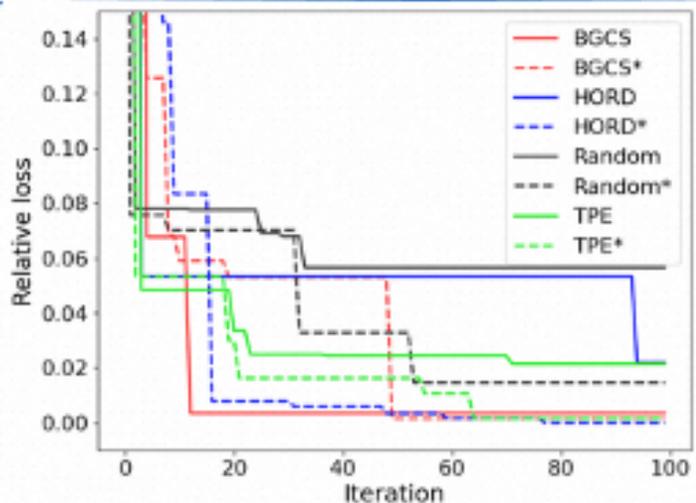
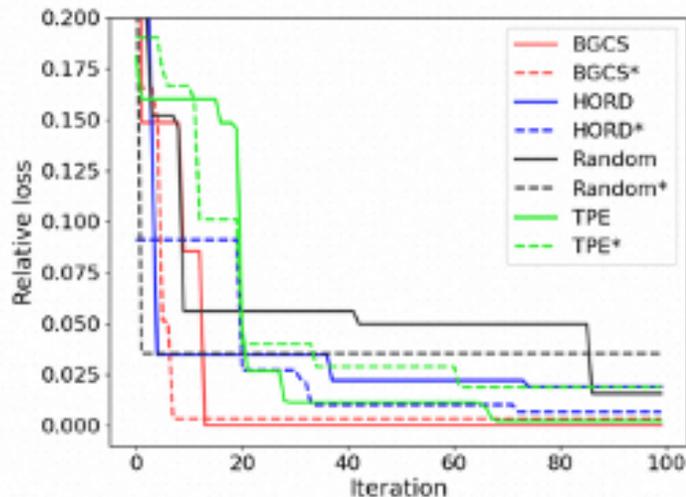


# ▶ 场景2. 云原生Ingress-Nginx全栈调优

## 调优效果

调优场景	吞吐调优比例	时延调优比例
http短链接	44.47%	15.36%
https短链接	13.97%	61.40%
http长连接	6.05%	24.79%
https长连接	3.47%	3.26%

举一个例子



# **PART 04**

# **总结展望 & 开源共建**

重要性·问题·策略

# 思考 & 展望



## 思考？我们的价值是什么

- **全栈调优专家知识库**，拉通软硬件领域的多个团队，形成针对TOP云业务的全栈调优解决方案，相比AWS镜像仓解决方案人力和运维成本减少90%+
- **统一算法框架**，引入多领域算法解决业务问题，尤其针对客户需要的多目标、参数可解释、环境波动等问题找到了比较好的解决方案，并取得了比较好的落地成果
- **高效有效的自研算法**，更加适用于业务调优的高维场景，相比传统Bayes算法调优需要轮次减少65%，算法运行时间减少90%+



## 展望，未来哪里能做的更好

- **扩展专家知识库，服务更多业务**，向下对接更多软硬件调优的特性，向上形成更多业务的调优解决方案
- **构建算法模型库，复用已有场景的调优知识**，根据已经完成的调优场景，形成模型库，在相似场景上复用，能够有效减少运行轮次
- **形成业务画像能力，提供更加智能的调优服务**，向下对接更多软硬件调优的特性，向上形成更多业务的调优解决方案

# ▶ 开源共建

Main Page ( 升级中 ) :

<http://keetune.io>

OpenAnolis SIG :

<https://openanolis.cn/sig/keentune>

Gitee 源码仓 :

<https://gitee.com/anolis/keentuned>

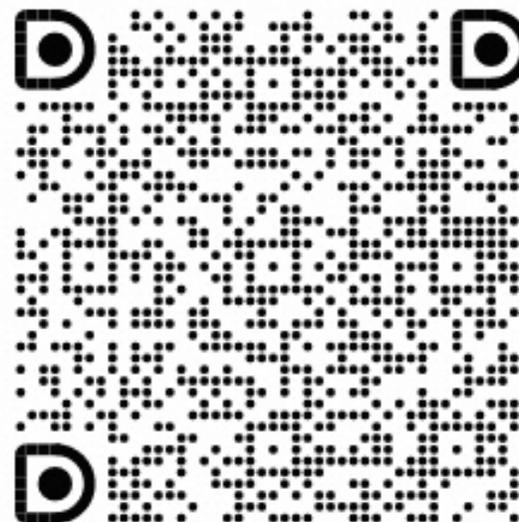
[https://gitee.com/anolis/keentune\\_target](https://gitee.com/anolis/keentune_target)

[https://gitee.com/anolis/keentune\\_brain](https://gitee.com/anolis/keentune_brain)

[https://gitee.com/anolis/keentune\\_bench](https://gitee.com/anolis/keentune_bench)

[https://gitee.com/anolis/keentune\\_ui](https://gitee.com/anolis/keentune_ui)

<https://gitee.com/anolis/keenopt>



KeenTune SIG 钉钉群 :  
42497140

# 感谢聆听

