

阿里大数据运维经验分享

柯旻



自我介绍：

柯旻--花名大舞

阿里集团基础架构事业群-大数据SRE-资深技术专家

2007年加入阿里，目前负责阿里大数据平台(Hadoop/odps(5k)/OTS/Hbase/Spark/Jstorm等)运维工作,专注于大规模海量数据分布式计算平台运维工作

Email : dawu@taobao.com

微博 : @大舞-ukl

大纲



- 面临的变化挑战
- 需要关注的那些事
- Devops
- 自动化运维平台
- 运维数据化&运营化
- 演进变化历程

SQL

MR

迭代计算

图计算

流计算

MaxCompute : 统一的计算引擎

飞天分布式操作系统 : 一台大计算机

10000台



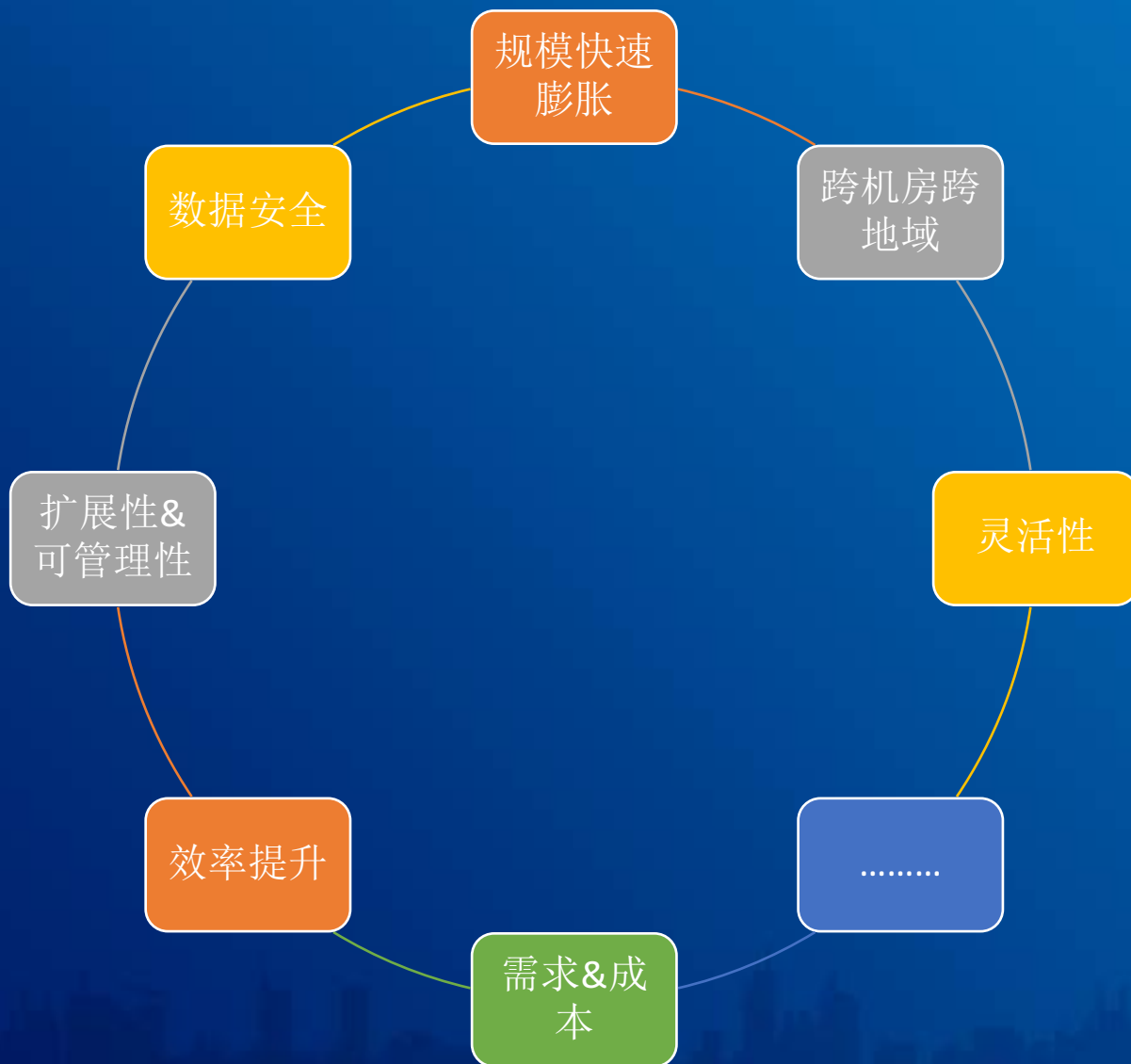
10000台



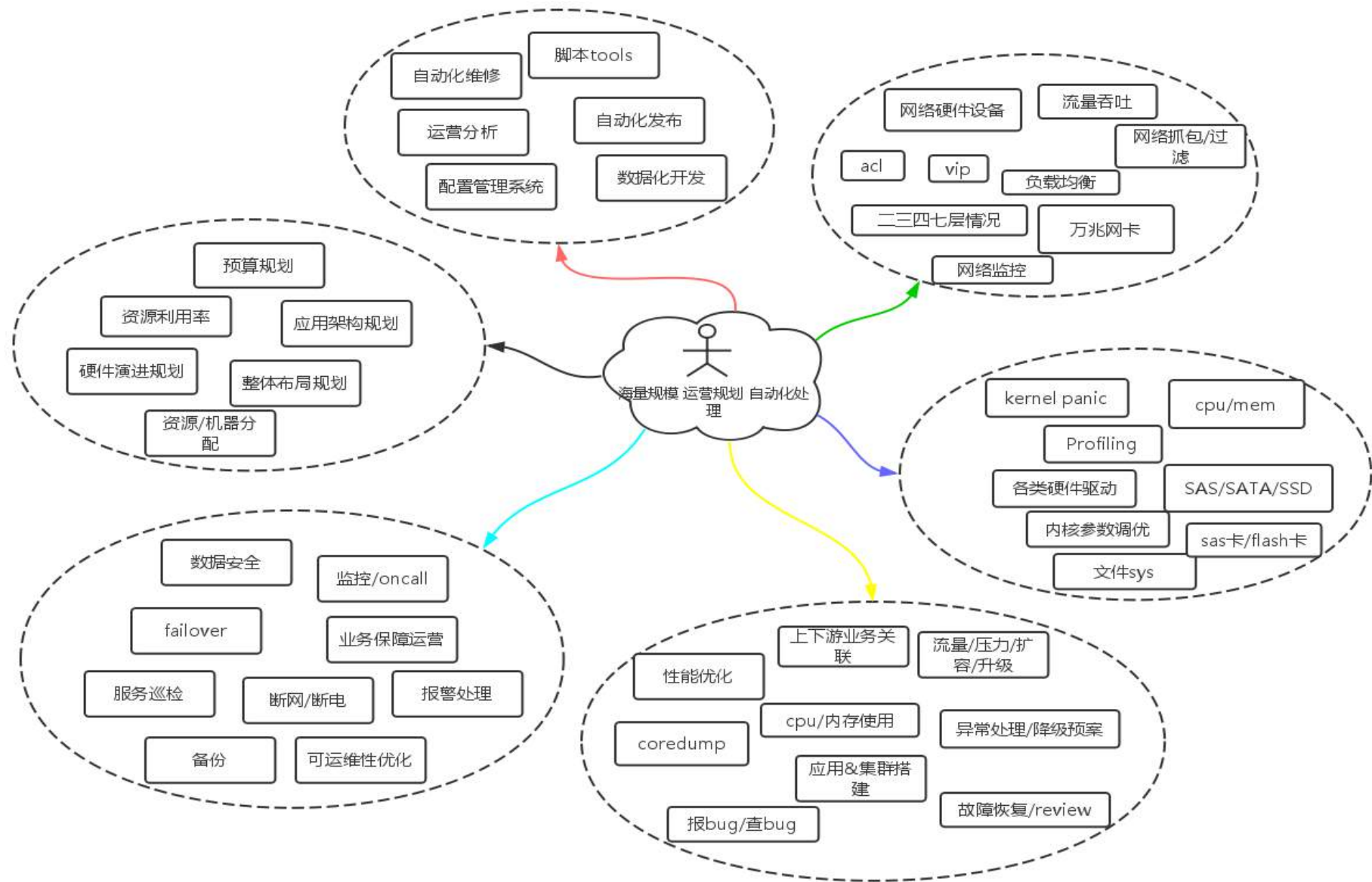
10000台



面临的变化挑战

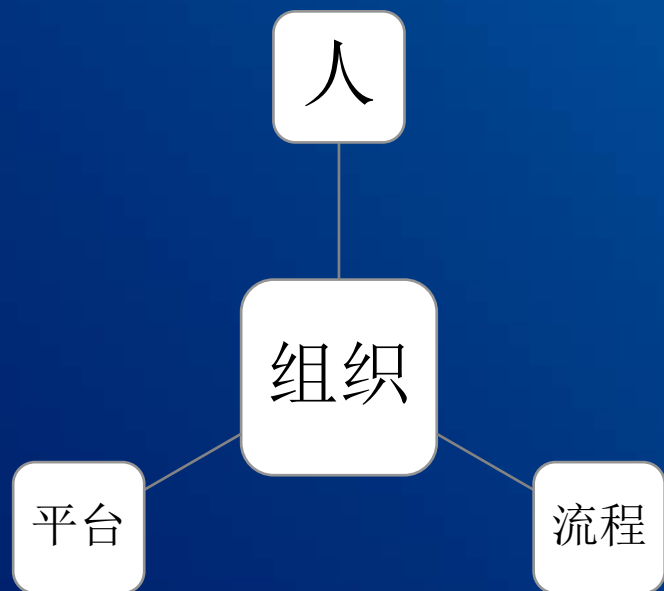


需要关注的那些事



传统需要变革





人、平台和流程需要有效的结合，才能发挥组织的力量。

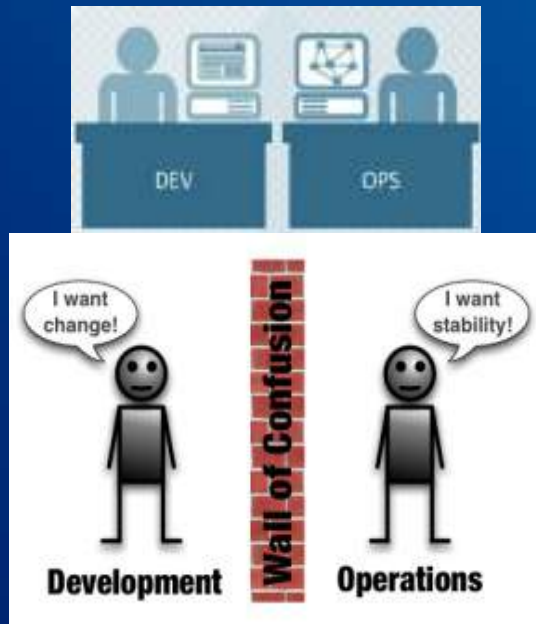
没有最好的，只有最合适的。

DevOps误区：

无论是研发至上还是运维至上，都是极端的。随着IT成熟度的不同，会有不同的表现。

强行用一种模式去套用，可能适得其反。代价可能是很高。

因此，DevOps看似一个平台工具的革命，实际上它是整个组织的革命，需要人、平台和流程有效的配合，甚至文化的驱动。每个公司需要摸索一套合适的落地方案和计划。



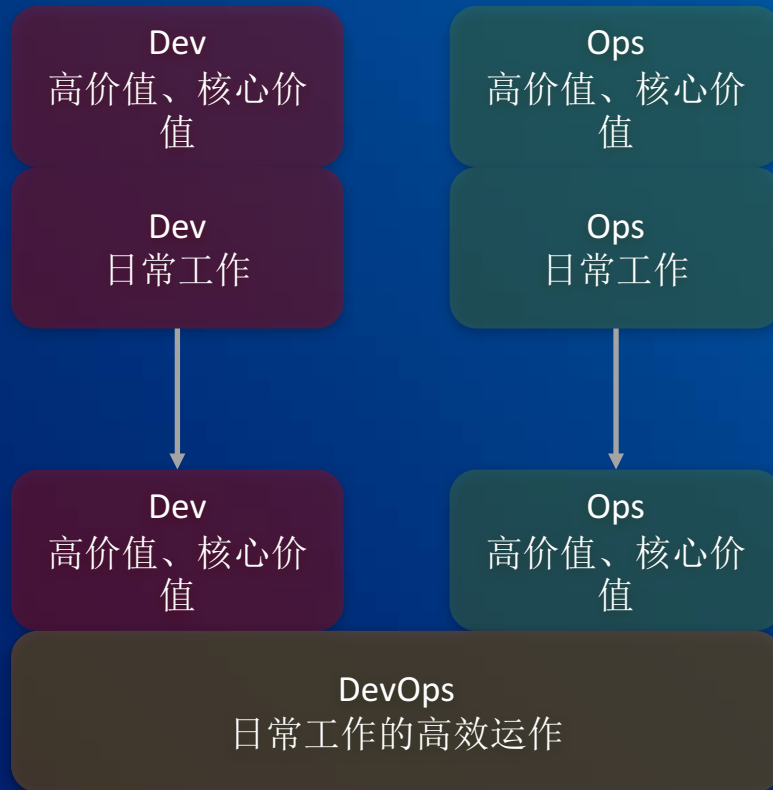
为运维“平反”

善意但却容易误解的价值传递：

“要革运维的命！”

“以后都没有运维了！”

如果运维的核心价值仅仅看作是**发布、变更和部署**等工作，那是对运维核心价值的狭义理解。



日常工作的高效运作也
华丽转身成高价值的工作

Dev 和 Ops 的相互核心价值部分是不可替代的。

EG :

- 开发对网络、操作系统、服务器和安全等方向的理解会弱于运营
- 运维对代码和软件架构的设计会弱于研发

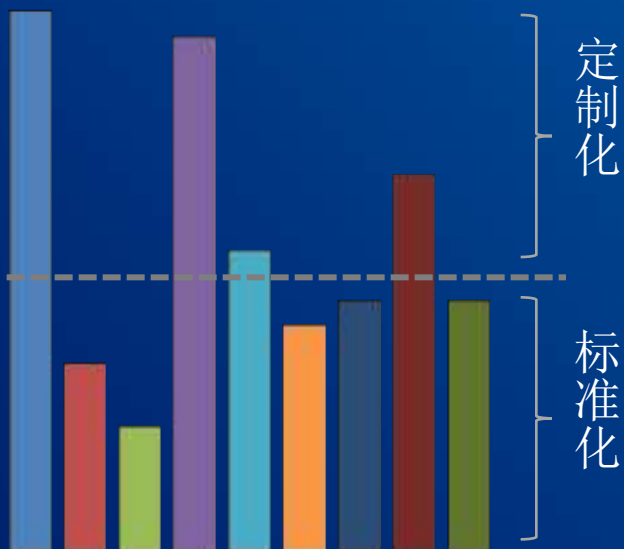
术业有专攻

各自找准价值方向，持续构建核心竞争力
相互融合，而非替代

注：适度的竞争有利于发展，完全的计划经济并不一定高效，需要动态平衡；

自动化运维平台

产品的架构和运维成熟度



定制化

标准化

定制化和标准化的分界点，也是经常会变化的。

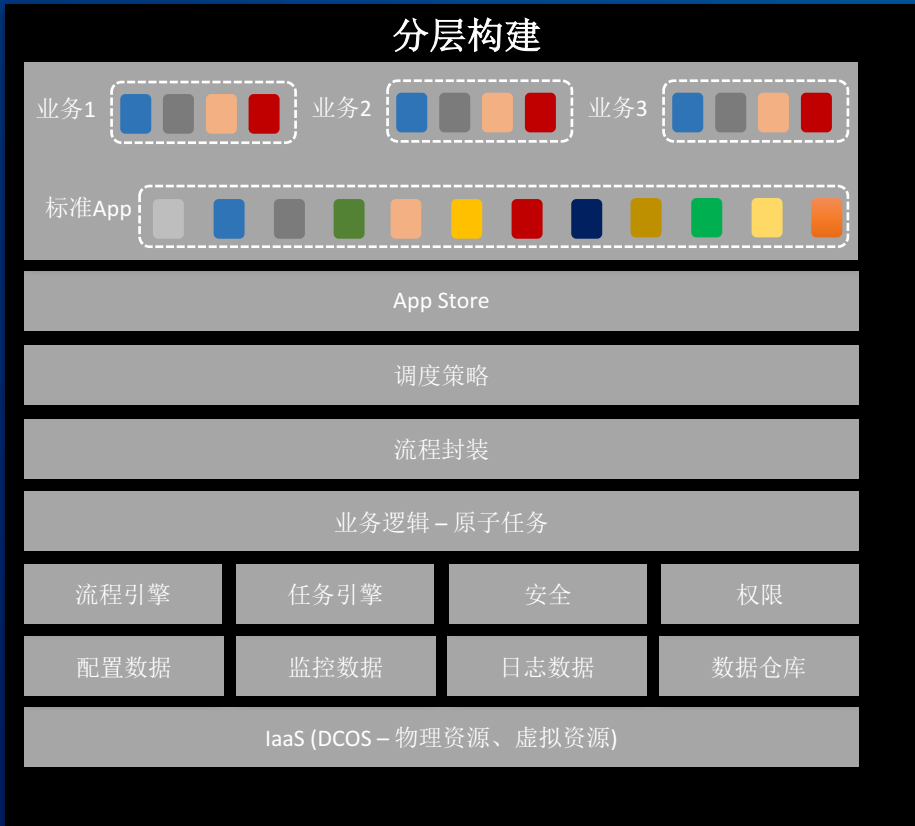
定制化：对单个产品来说高效、灵活

标准化：对整体效率有提升、统一

平台的构建需要充分考虑**扩展性和兼容性**

业务的架构不是一开始就完全定稿的。而是随业务发展而不断变复杂、稳定、成熟和高效。

分层构建



运维定制

优势互补，打好组合拳，
相互发挥出核心价值。

统一平台

运维的同学们拥有一套高效的引擎、规则引擎和各类已经封装好标准原子任务，其实运维的同学们可以快速的定制自己想要的调度逻辑和门户。这样，DevOps的推动会更能**高效和敏捷**。

分层式样架构设计



标准化一站式变更，单工作流程细节掌控

Tesla

执行
22598 (CF编号:207469)

当前进度 100%

准备

已完成
提交
changefree

执行人:

人工操作: 设置 odps_g_checkpermission_timeout

只可查看

workflow: Plan Update

状态 **执行成功**

影响集群

只可查看

[查看执行详情](#)

已结单



批量设置手动 批量设置自动 ? 批量设置报警方式: 旺旺 短信 电话

第 1 步 **执行成功** 报警方式: 旺旺-短信-电话 超时时间(秒): 900, 剩余自动重试次数: 0, 重试间隔: 0 秒 第 1 次执行, 1秒

命令

```
发送旺旺(集群 ay58 开始 PlanUpdate0.1 升级, 升级内容: OdpsService, release1( ))给用户组: odps_notify, 用户列表: 68424
```

[查看输出](#)

第 2 步 **执行成功** 报警方式: 旺旺-短信-电话 超时时间(秒): 900, 剩余自动重试次数: 0, 重试间隔: 0 秒 第 1 次执行, 1秒

命令

```
发送短信(集群 ay58 开始 PlanUpdate0.1 升级, 升级内容: OdpsService, release1( ))给用户组: odps_notify, 用户列表: 68424
```

[查看输出](#)

第 3 步 **执行成功** 报警方式: 旺旺-短信-电话 超时时间(秒): 120, 剩余自动重试次数: 0, 重试间隔: 0 秒 第 1 次执行, 5秒

命令 注释: 推送配置文件service_pan.ini

```
推送配置文件(fs-odps_pe:internal/global/service_plan.ini) ----> :/tmp/odps_tesla_config/service_plan.ini
```

[查看输出](#)

第 4 步 **执行成功** 报警方式: 旺旺-短信-电话 超时时间(秒): 900, 剩余自动重试次数: 0, 重试间隔: 0 秒 第 1 次执行, 15秒

命令 注释: get_service_plan -S \$service_name -F \$Suffix

```
[admin@ $ /usr/local/odps/module/PlanManage/get_service_plan(-S OdpsService -F release1)
```

[查看输出](#)

第 5 步 **执行成功** 报警方式: 旺旺-短信-电话 超时时间(秒): 900, 剩余自动重试次数: 0, 重试间隔: 0 秒 第 1 次执行, 10秒

命令 注释: change_service_plan -S \$service_name -F \$Suffix

应用基础环境配置管理rpm, 文件, 目录

- 低部署成本, 简单通用
- 灰度发布, 低消高效
- 规则多样, 文件模板
- 多级绑定, 灵活组合



ID	任务名称	当前执行人	任务状态	优先级	创建时间	最后执行人	最后执行时间	操作
411	更新yum	运维01	成功	高	2016-09-07 11:03:01	运维01	2016-09-07 10:09:00	[-] [E]
412	安装yum	运维01	成功	高	2016-09-10 10:47:00	运维01	2016-09-10 10:04:20	[-] [E]
417	更新yum	运维01	成功	高	2016-09-10 10:00:01	运维01	2016-09-10 09:01:11	[-] [E]
416	更新yum	运维01	成功	高	2016-09-10 09:00:01	运维01	2016-09-10 08:01:01	[-] [E]
415	更新yum	运维01	成功	高	2016-09-10 08:00:04	运维01	2016-09-10 07:04:41	[-] [E]
414	更新yum	运维01	成功	高	2016-09-10 07:00:01	运维01	2016-09-10 06:00:00	[-] [E]
413	更新yum	运维01	成功	高	2016-09-10 06:00:01	运维01	2016-09-10 05:00:00	[-] [E]
412	更新yum	运维01	成功	高	2016-09-10 05:00:01	运维01	2016-09-10 04:00:00	[-] [E]
406	更新yum	运维01	成功	高	2016-09-09 10:00:01	运维01	2016-09-09 11:00:01	[-] [E]
405	更新yum	运维01	成功	高	2016-09-09 10:00:01	运维01	2016-09-09 11:00:01	[-] [E]

集中管理线上服务器定时调度任务

- Crond管理散乱
- 执行结果丢失
- 异常通知机制
- 积累日常任务效果数据

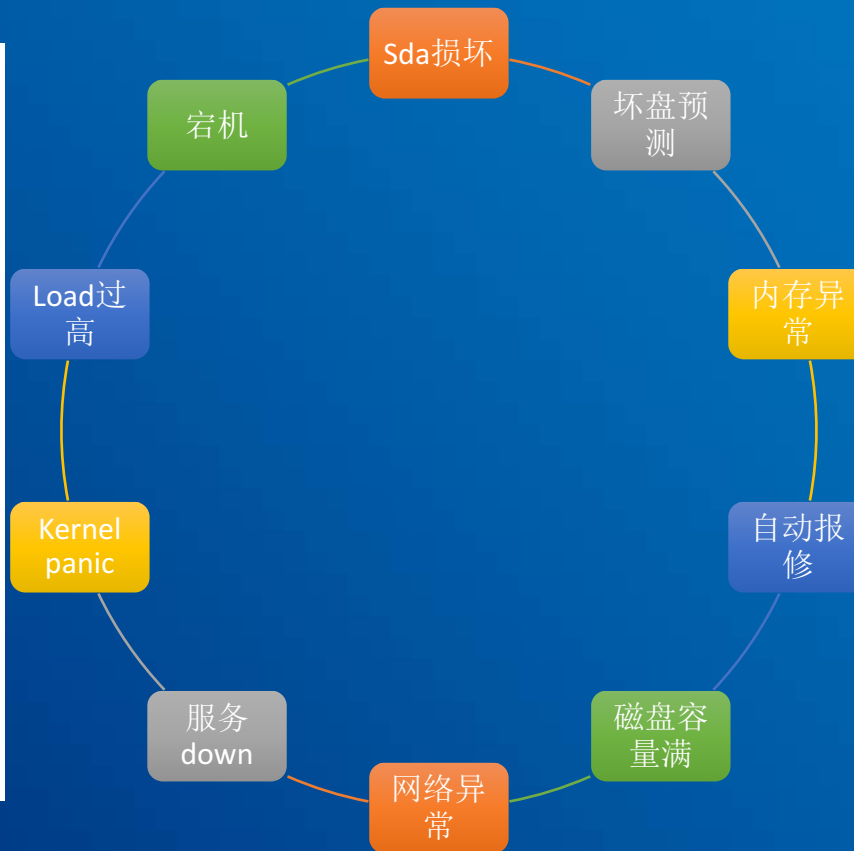
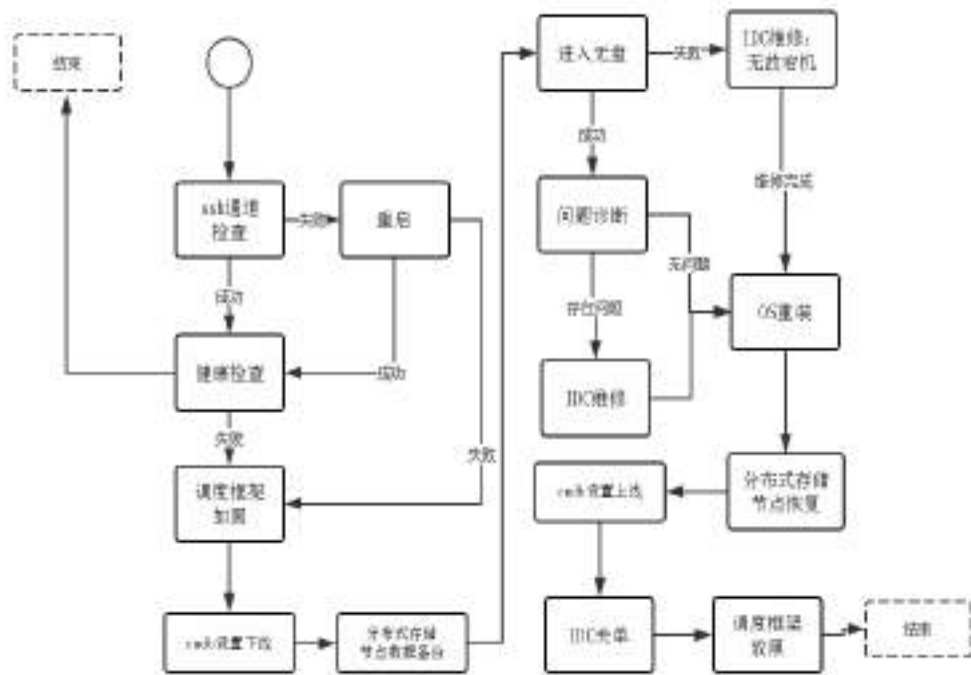


任务名称	ID	用户	任务状态	cron	操作
update yum	101461111	admin	成功	* * * * *	[-] [E]
update yum	101461111	admin	成功	* * * * *	[-] [E]
AT task	101461111	admin	成功	* * * * *	[-] [E]
update yum	101461111	admin	成功	* * * * *	[-] [E]

```

[message] {"id": "101461111", "user": "admin", "status": "成功"}
[message] {"id": "101461111", "user": "admin", "status": "成功"}
    
```

自动化故障监测和处理系统



运维数据化&运营化



- 1.传统自动化运维实现了末端流程的沉淀，但末端流程需要通过人或第三方系统连续，对人依赖大
- 2.传统运维分析注重被动分析，固化分析，核心技术往往在第三方工具，展现能力弱，关联度低
- 3.对运维产品的运营能力低下，劳动密集型，低附加值



- 1.加固末端产品能力
- 2.强调系统分工，构建通用连接产品
- 3.创建问题排查系统，分析决断执行过程

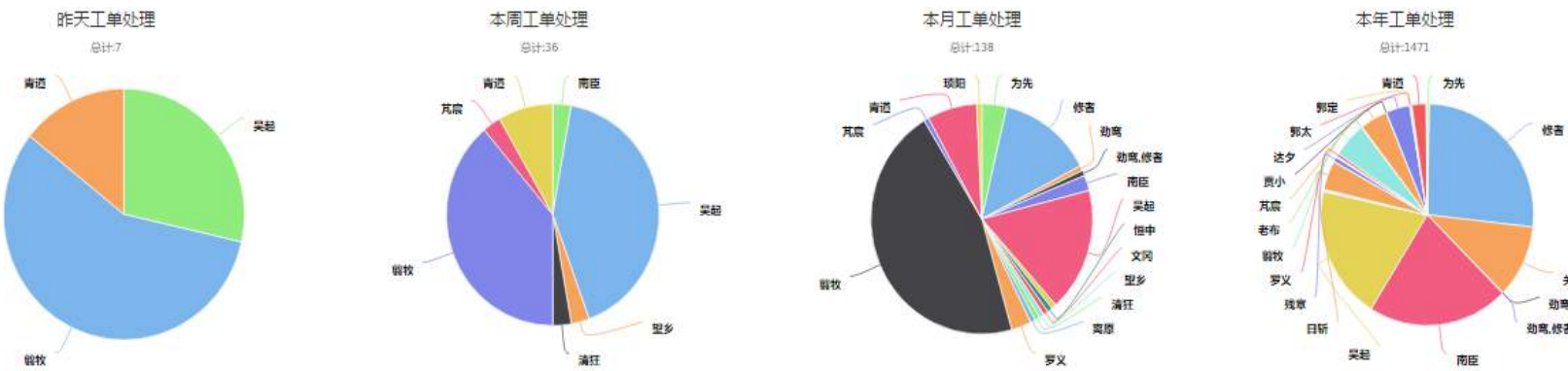
数据化运营



工单概览

 352 待处理	 618 超SLA工单	 ↑ 112% 工单量(较上周)	 ↑ 44% 工单量(较上月)	 8 处理时长(小时/本周)	 ↑ 14% 处理时长(较上周)
----------------	-------------------	------------------------	-----------------------	----------------------	------------------------

工单处理



业务概览

 1987 项目成员	 2251 登陆用户	 27 新用户(本周)	 1086 项目	 4 新项目(本周)	 254 服务部门
------------------	------------------	-------------------	----------------	------------------	-----------------

业务发展

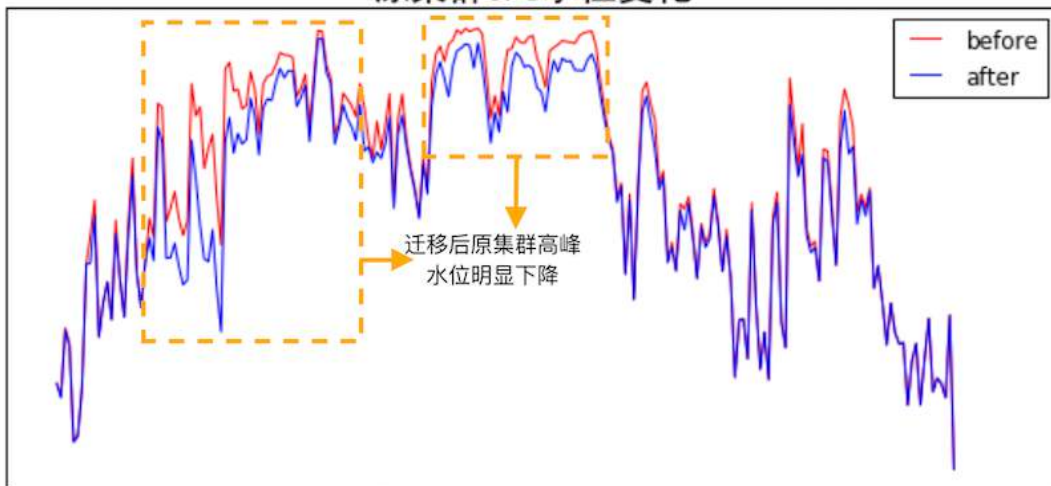
数据化分析

1.通过设定的个性化集群水位迁移阈值计算各个集群的需迁移量及可接收量

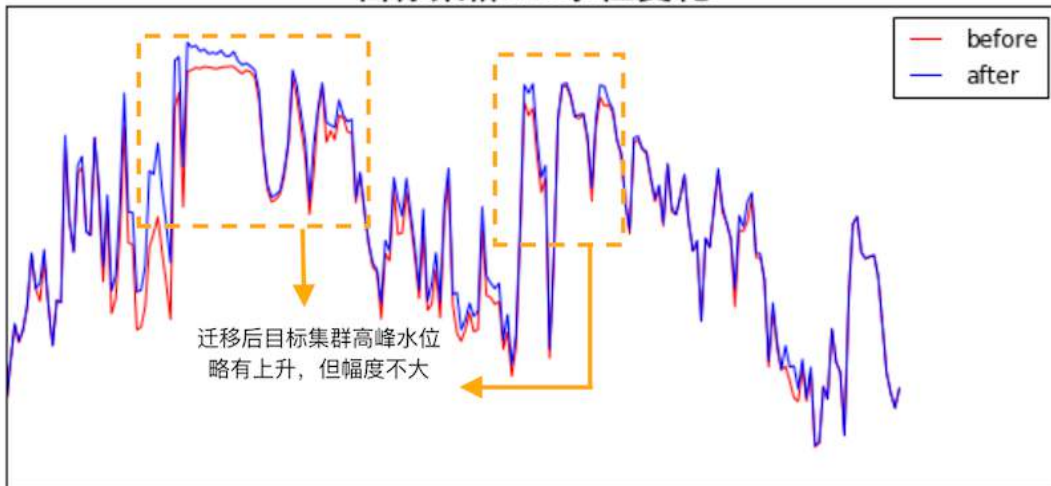
2.利用优化算法产生特定约束下的最优迁移策略
(线性优化规划算法)

3.量化执行该迁移策略后对集群资源水位及跨机房流量的影响

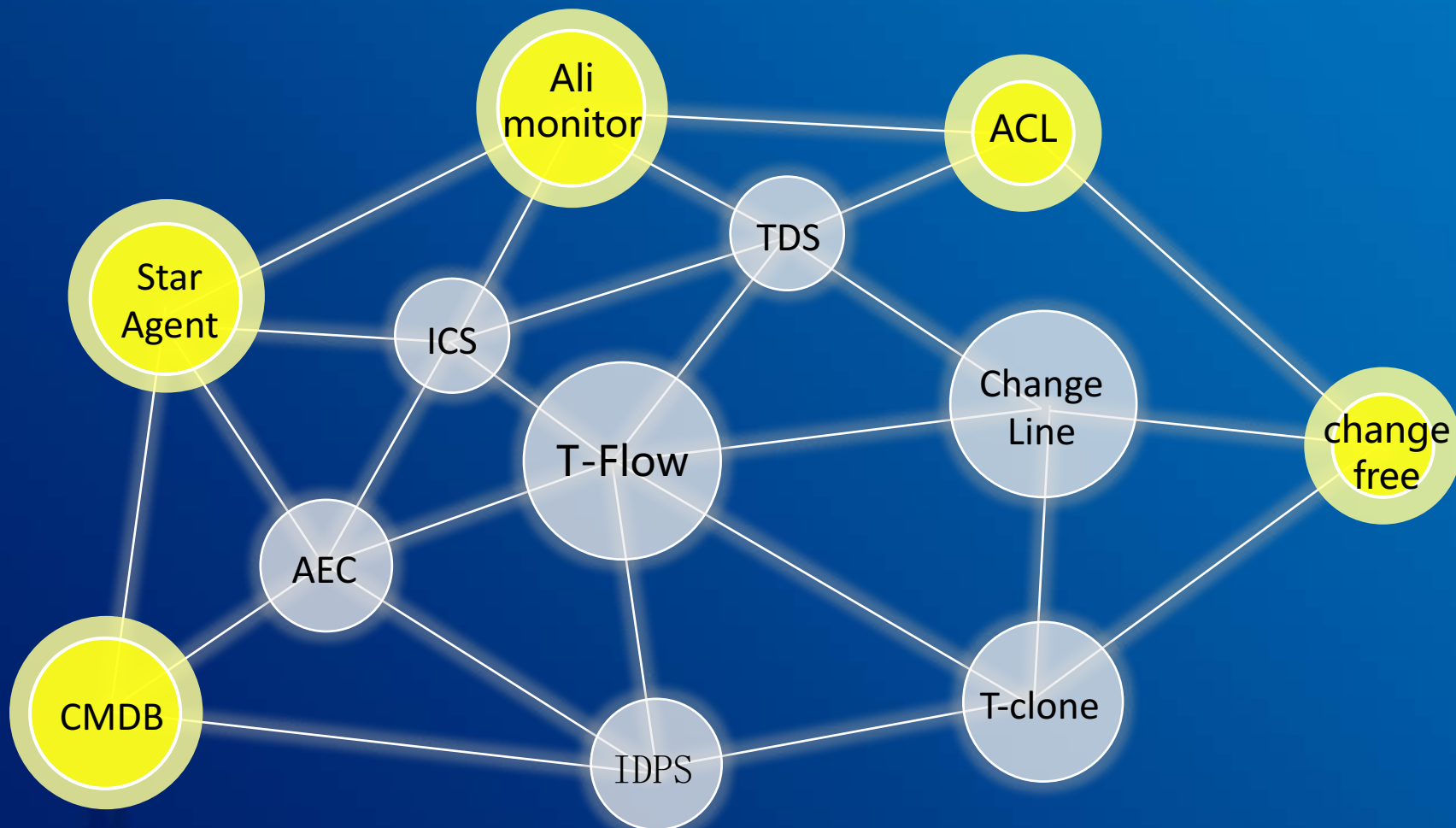
原集群CPU水位变化



目标集群CPU水位变化



面向连接的智能化运维体系

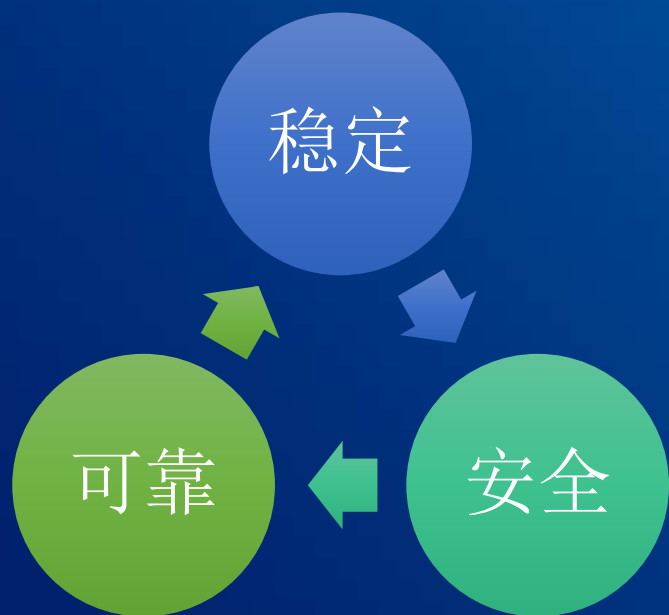


演进变化历程

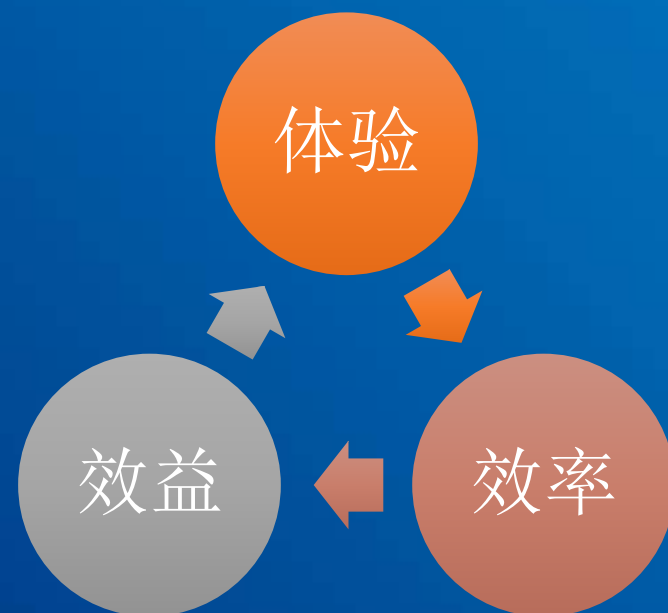


从实战中总结、提炼、构想、设计、建设出来的一套体系，提高效率，稳定性，准确率，提供深度价值的服务

运维--维护



运营--经营



THANKS