

# 大规模检索系统的DevOps实践分享

## DevOpsDays Shenzhen

2018 中国·深圳 刘文俊 百度资深开发测试工程师



## 主要内容

- 1 百度检索系统的现状
- 2 凤巢检索系统的DevOps选型
- 3 凤巢检索系统的DevOps实践
- 4 凤巢检索系统的DevOps应用效果

# 百度搜索系统的现状





搜索：立足之本

新闻 hao123 地图 视频

Baidu 百度



百度一下



凤巢：发展之匙

百度

巧克力



百度一下

【京东商城】好时HERSHEY'S-全新推出万圣节限量礼盒



购万圣节限量礼盒赢酒吧入场券, pop club携  
百大DJ阵容, 好时万圣节电音派对等你来嗨! 立  
即访问好时京东旗舰店, 了解更多活动信息.

mall.jd.com 2018-10 - 评价 - 广告

【天猫】-淘宝巧克力 大牌“惠”聚 理想生活上天猫



「天猫」TMALL.COM-淘宝巧克力, 8.9万品牌旗舰店, 品质护航, 支付放心, 购物无忧! 天猫淘宝商城, 引领生活新趋势!

www.tmall.com 2018-10 - 评价 - 广告

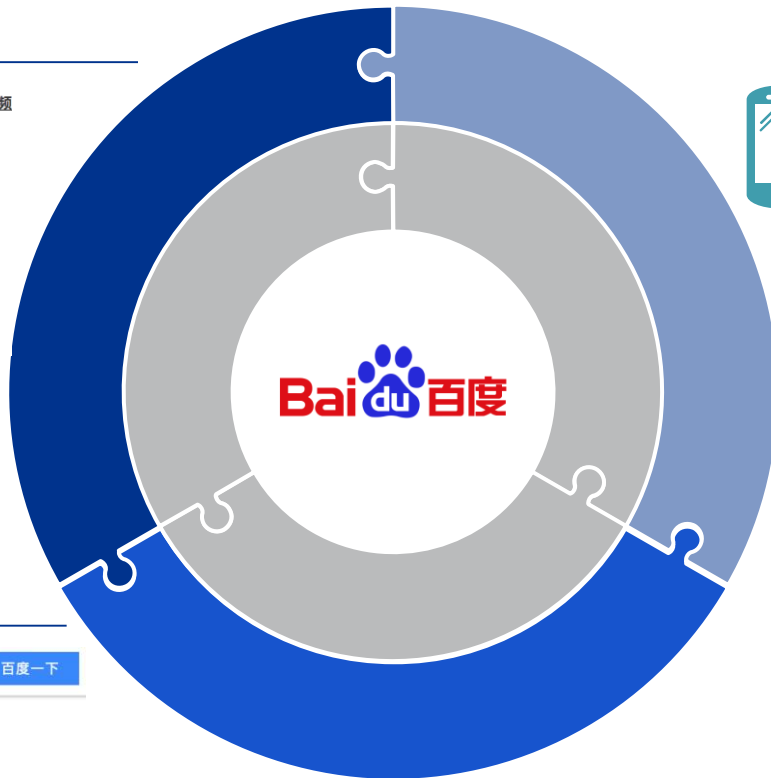
苏宁易购 日本白色恋人巧克力 苏宁超市 品类齐全 极速送达



苏宁易购, 日本白色恋人巧克力, 苏宁超市, 海量钜惠, 应有尽有, 百万商品  
促不停! 苏宁易购, 日本白色恋人巧克力, 超值低价, 优品钜惠, ...

休闲零食 生奶饮品

www.suning.com 2018-10 - 评价 - 广告



手百：时代之趋

中国移动 4G

10:05

87%

设置城市  
展现天气



Baidu 百度

搜索或输入网址



关注

推荐

新时代

热榜



习近平点出东北振兴3项新要素

置顶 新华网

《平“语”近人》：治国有常民为本

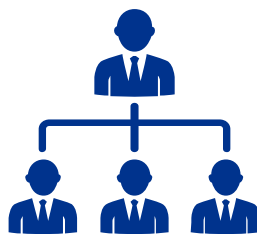
置顶 中国日报网

# 百度凤巢检索系统的内部挑战



业务复杂

线上服务模块200+  
代码量10w+/周



团队庞大

大商业团队400+人



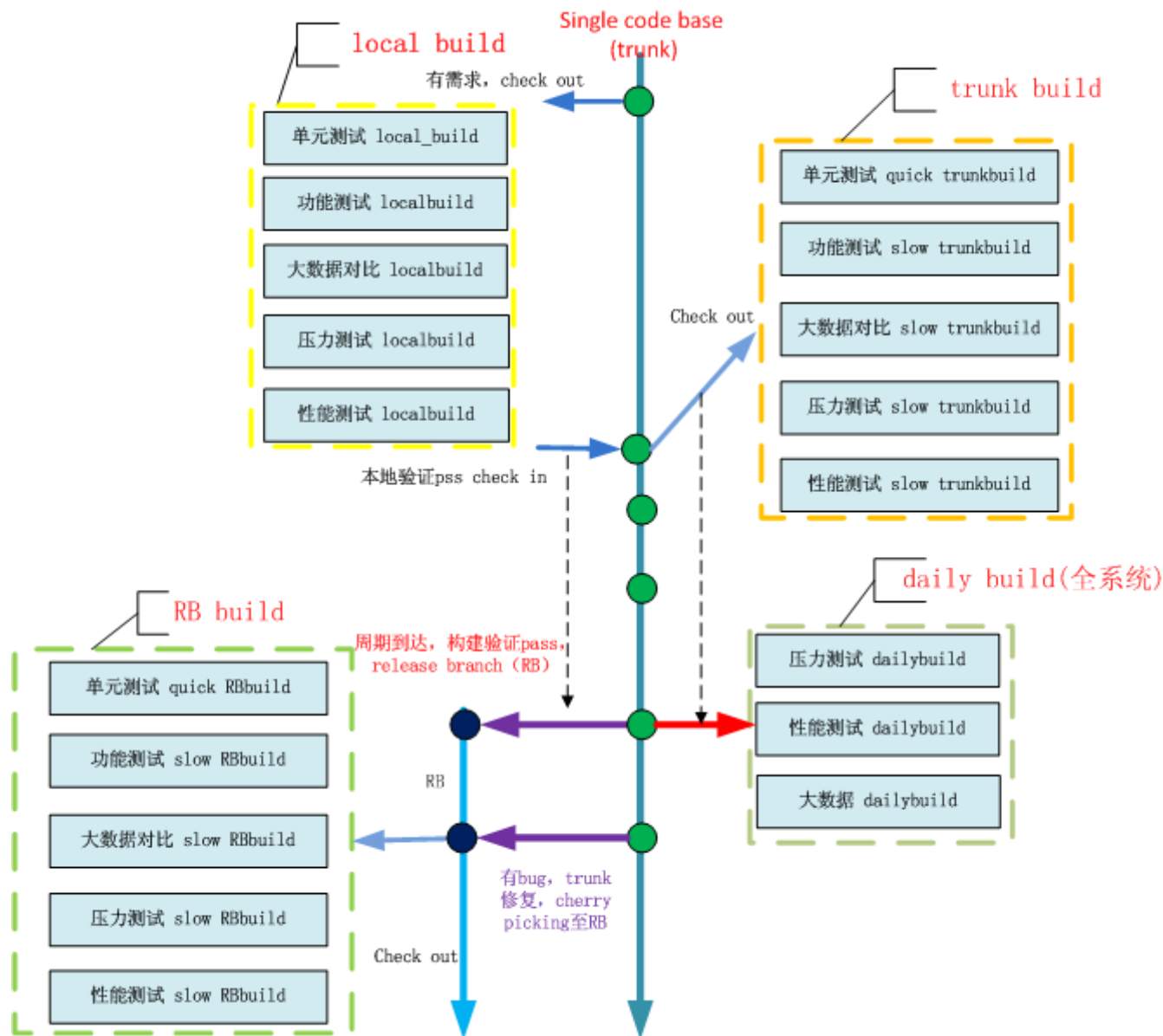
迭代快速

核心模块要求每周2次上线  
具有紧急上线的能力

# 凤巢检索系统的DevOps选型



# 凤巢检索系统的DevOps选型——Single branch



## 三大特性：

### ✓ Single Code Base

产品源代码，只保持最新的版本，每次编译都基于该最新版本，每次发布二进制产品也基于该最新版本。

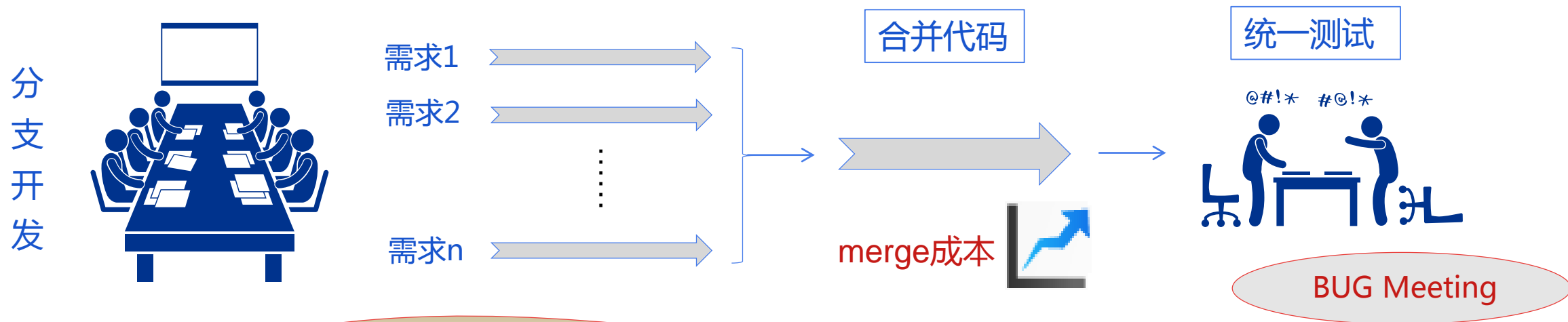
### ✓ Continuous Build Verification

对该base的源代码，维护一套自动化测试的系统，对该base上源代码的质量进行持续编译，测试，再编译，再测试的过程。

### ✓ Regular Release Process

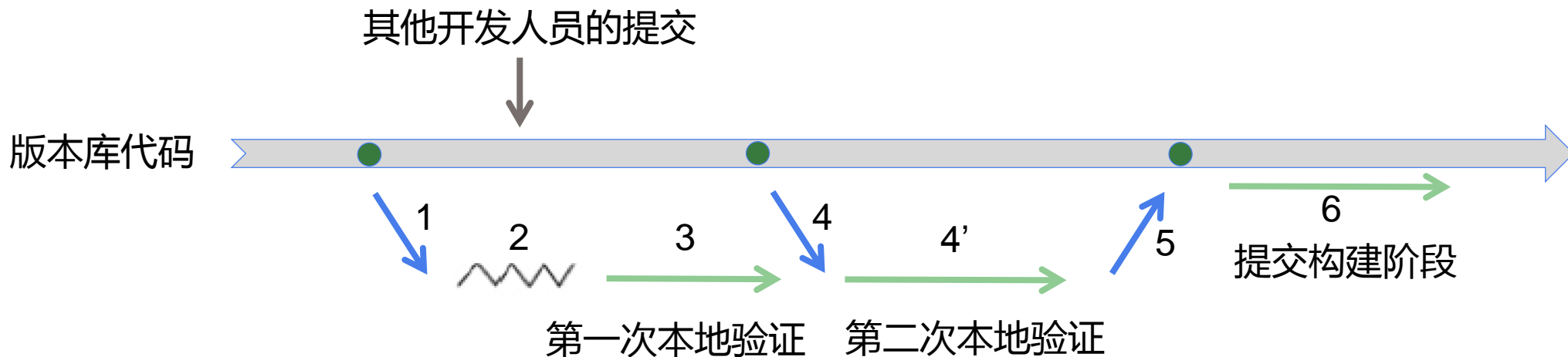
软件产品的发布是定期的。该周期不依赖于某一个特性是否研发完毕而决定的，只要到了时间，就发布。

Single branch相对Muti branch development之优势：



- ✓ 解决merge回主干困难，回归成本高问题。
- ✓ 解决随着项目增多，分支增多，越来越难管理问题。
- ✓ 修复bug，可以达到修改主干，多处都可以fix的效果。
- ✓ 解决在实验过程中，难以比较两个实验间的相互影响问题。
- ✓ 解决QA只保证单上线分支质量，忽视merge后质量的问题。





图例： ● 构建成功的版本 | 代码修改 | 运行测试 | 代码流向

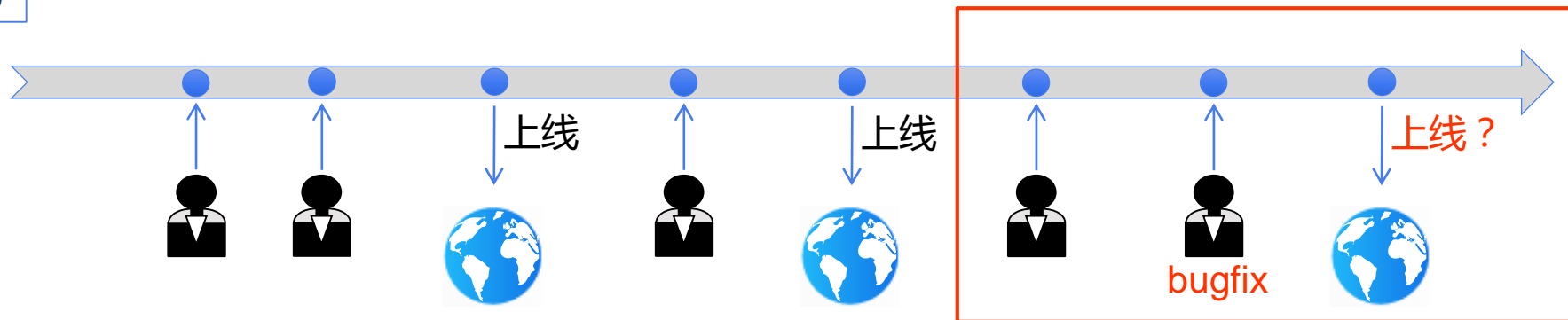
## 六步走

- ✓ 将代码从版本库检出到本地
- ✓ 为新功能或者bug修改代码，或开发测试代码
- ✓ 第一次运行本地验证（localbuild）
- ✓ 与版本库中的代码合并，再次执行本地验证(localbuild)
- ✓ 提交代码到版本库（提交前需要发起CodeReview，并通过）
- ✓ 构建任务被自动触发

# 凤巢检索系统的DevOps选型——分支上线

## 上线模式A：主干上线

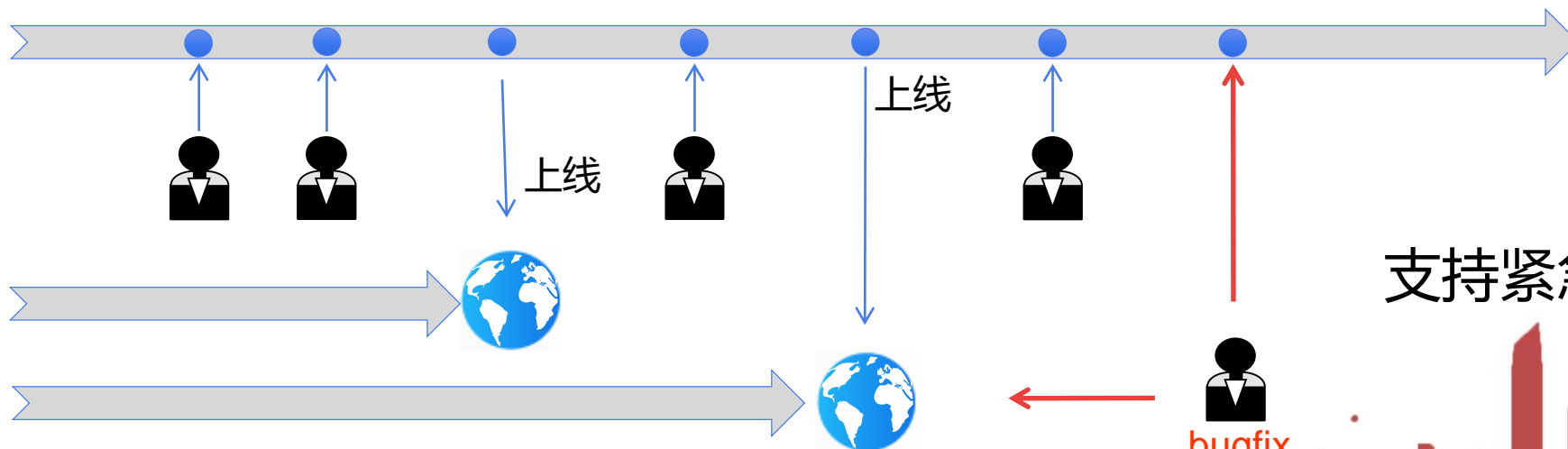
master/trunk



线上代码独立于主干之外

## 上线模式B：分支上线

master/trunk



支持紧急上线

bugfix

# 凤巢检索系统的DevOps实践



# 凤巢检索系统的DevOps实践——概述



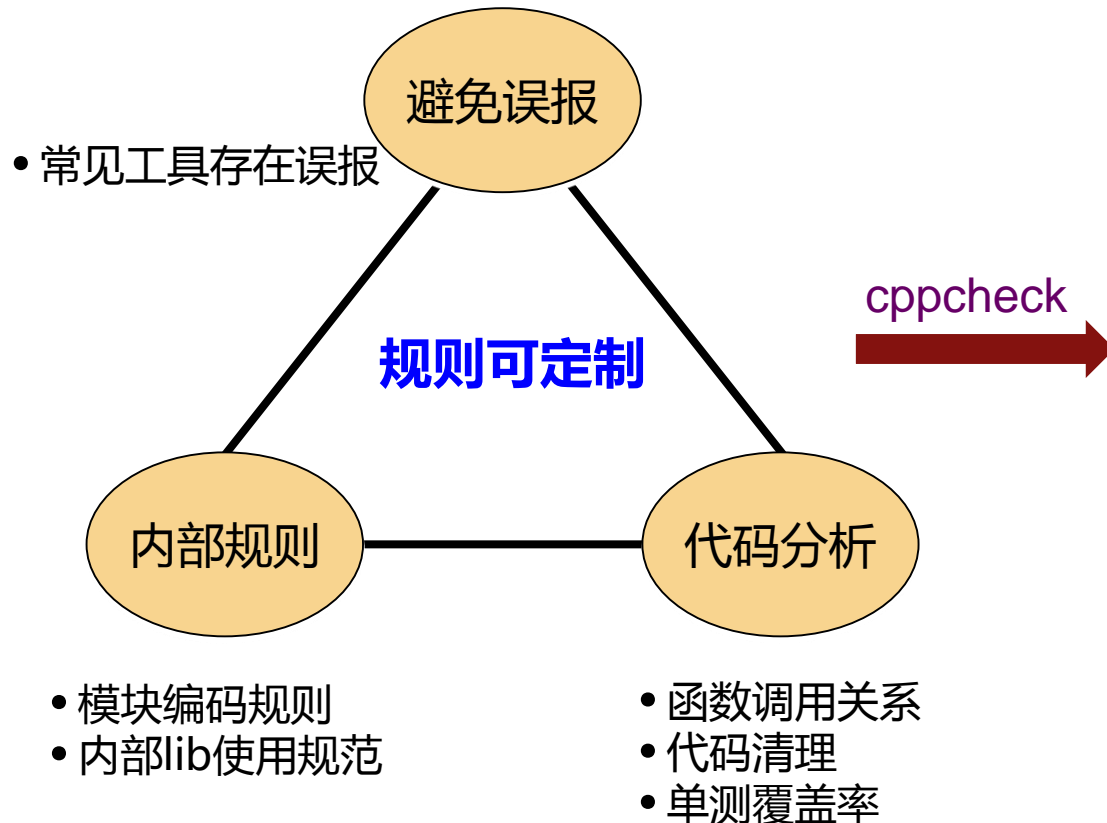
## 需求管理全流程

|  | 项目启动+需求评审  | 详细设计  | 开发  | 开发完成                       | 测试       | 测试完成      | 上线       |
|--|--|---|---|----------------------------|----------|-----------|----------|
|  涉及角色   | PM   | RD  | RD&QA   | RD                         | QA       | QA        | PM RD QA |
|  动作     | 1、建卡片（story或task）<br>2、上传MRD及UE图（或在卡片正文描述需求）<br>3、修改MRD或UE图并上传，卡片正文写出变更点<br>4、勾选邮件通知全组同学 | 1、上传详设文档<br>2、紧急上线等不需要PM参与的feature，由RD建卡片，正文描述升级点，升级方案，测试关注点等<br>3、勾选邮件通知全组同学 | 1、RD拆分卡片<br>2、每个卡片写明实现的功能及涉及的模块<br>3、RD在每个卡片正文列举需重点关注的要点<br>4、RD标注RD负责人、QA标注QA负责人<br>5、若实现方式或逻辑有变动，RD在卡片正文描述最新设计方案及实现方法<br>6、勾选邮件通知相关同学 | 1、卡片移动至待测试<br>2、勾选邮件通知相关同学 | 卡片移动至测试中 | 卡片移动至测试完成 | 卡片移动至已上线 |
|  卡片位置 | 需求池  | 待开发   | 开发中   | 待测试                        | 测试中      | 测试完成      | 已上线      |

# When ? Why ?

前置测试：代码提交准入阶段（绑定代码库hook）

低成本：不需要case、不需要环境、不需要数据



|                                 |  |
|---------------------------------|--|
| 资源泄露风险                          | framework/src/bespafeedbs_process_module.cpp:335 |
| 下标越界风险                          | framework/src/interact_with_feedbs.cpp:1834      |
| string类型字符串长度获取直接使用str.length() | framework/src/interact_with_feedbs.cpp:2893      |

## 47. string类型字符串长度获取

原理：目前有些RD在获取string类型字符串长度的时候，会先转换成char数组，直接使用string.length()接口。

代码：check\_strlen\_cstr

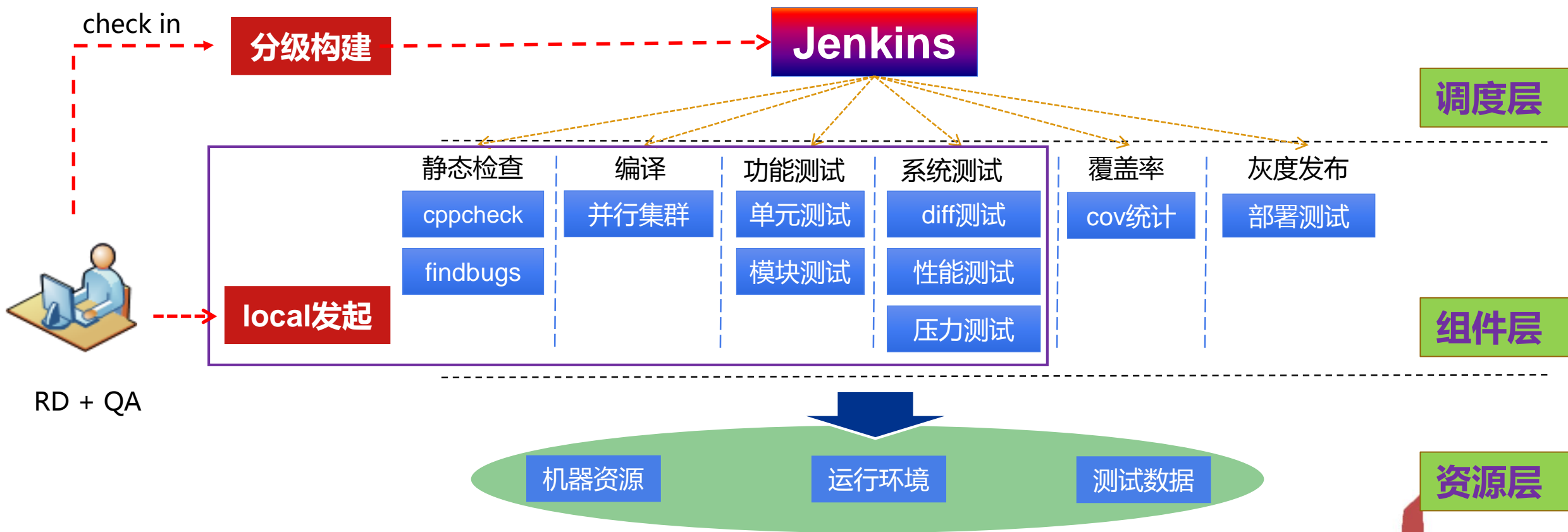
id：CheckStrlenCstr

覆盖：全部代码

tag：common

规则设计说明：string类型字符串长度获取

代码提交后，进行测试任务分级构建





## 功能测试和系统测试的对比

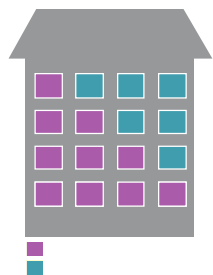
| 功能测试：<br>验证某功能点是否符合预期 |         | 系统测试：<br>本次升级对整个服务的影响（diff、性能、压力） |
|-----------------------|---------|-----------------------------------|
| 范围                    | 单函数、单模块 | 整个系统：多模块、多服务                      |
| 环境                    | 轻量级     | 仿真线上                              |
| 数据                    | 单条、多条   | 万/十万级别                            |
| 下游                    | mock    | 真实下游                              |
| 运行时长                  | 秒级别     | 分钟/小时级别                           |

### 系统测试：构建系统中最复杂的环节

- 1、环境仿真、数据量大、任务执行时间长
- 2、涉及范围广：多模块、多服务



## 系统测试三大挑战



**资源**  
机器&数据需求

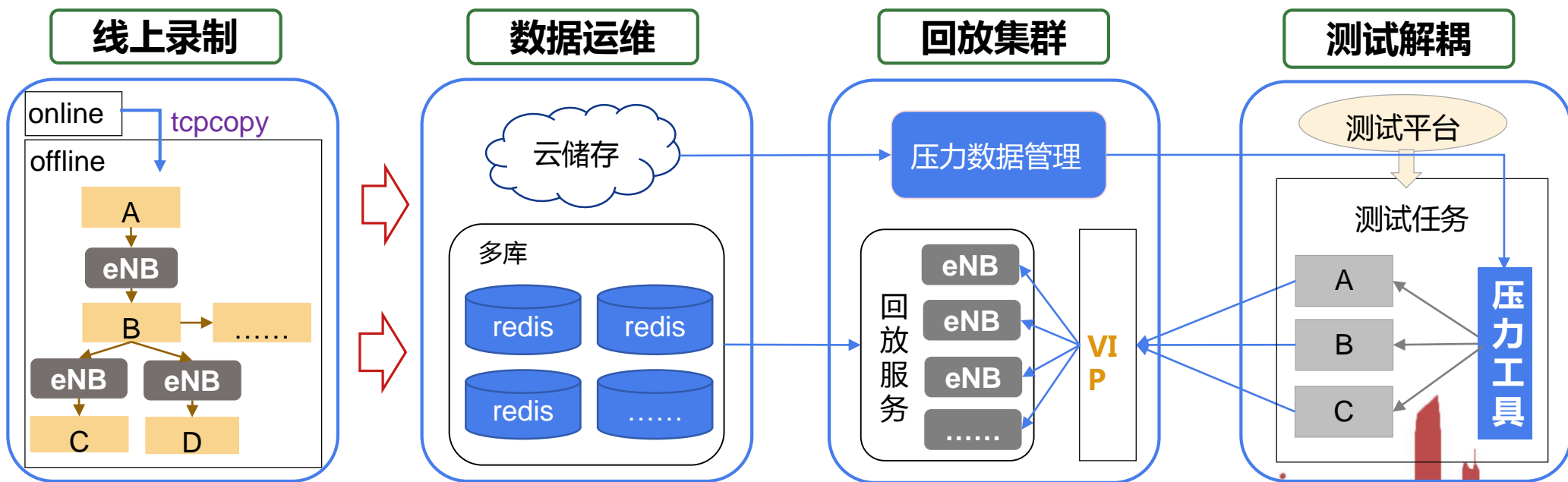


**效率**  
环境&任务时间成本

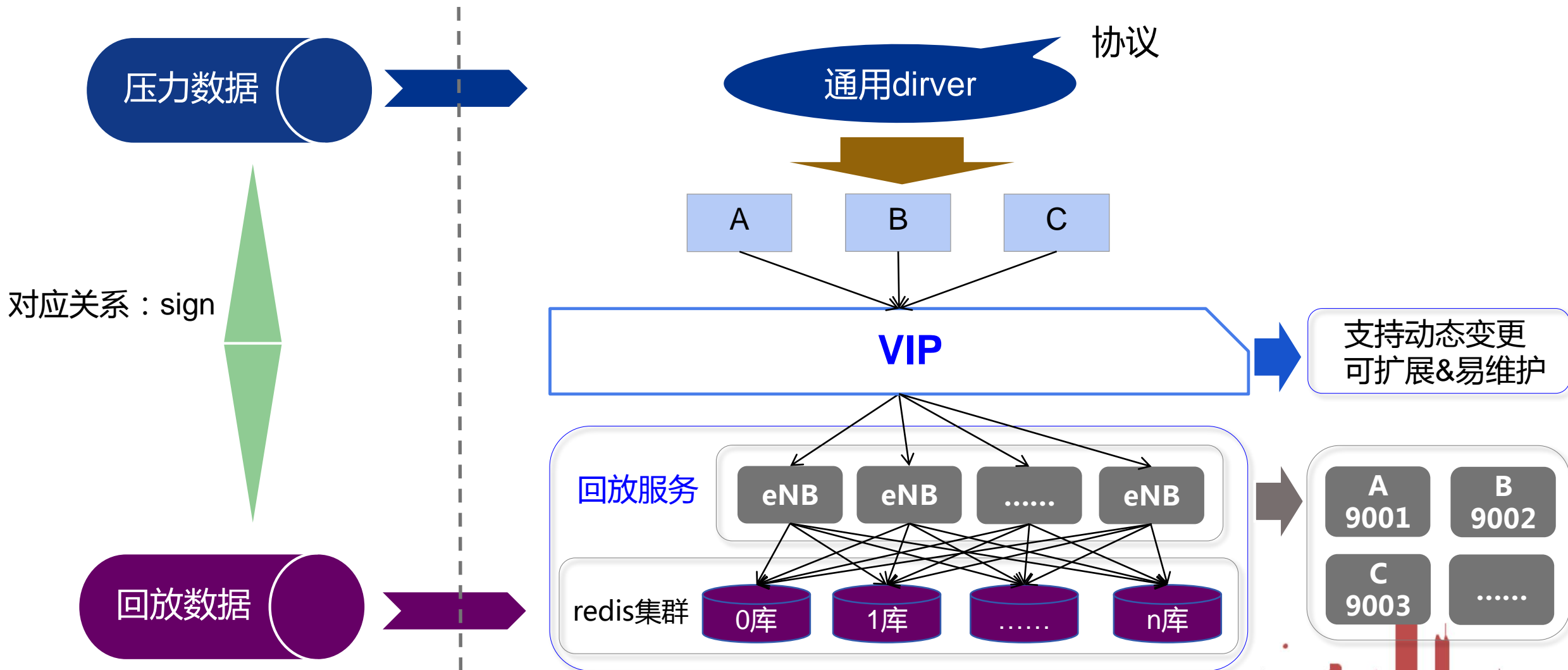


**准确**  
百万测试数据真实性

线上录制  
线下回放



## 测试解耦过程

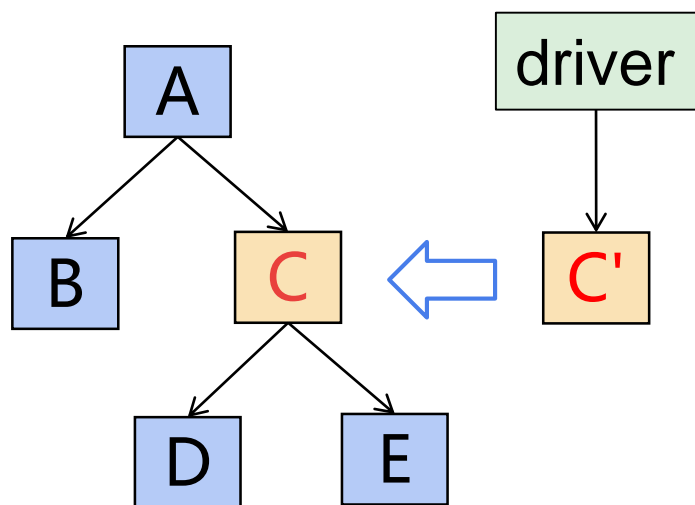




灰度发布：避免线上环境不一致&配置冲突导致的问题

上线前的最后一道闸门：环境仿真&配置一致

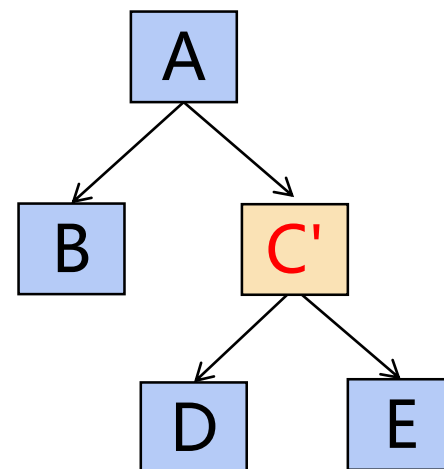
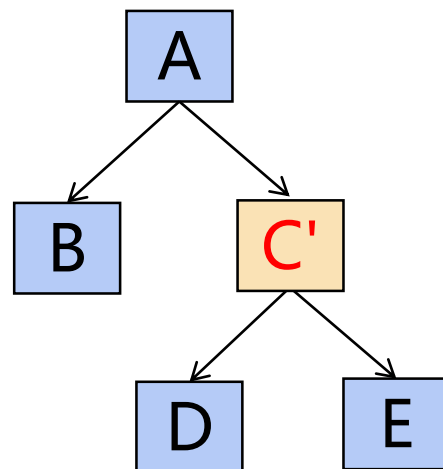
DEPTEST



启动？  
崩溃？

PASS

ONLINE

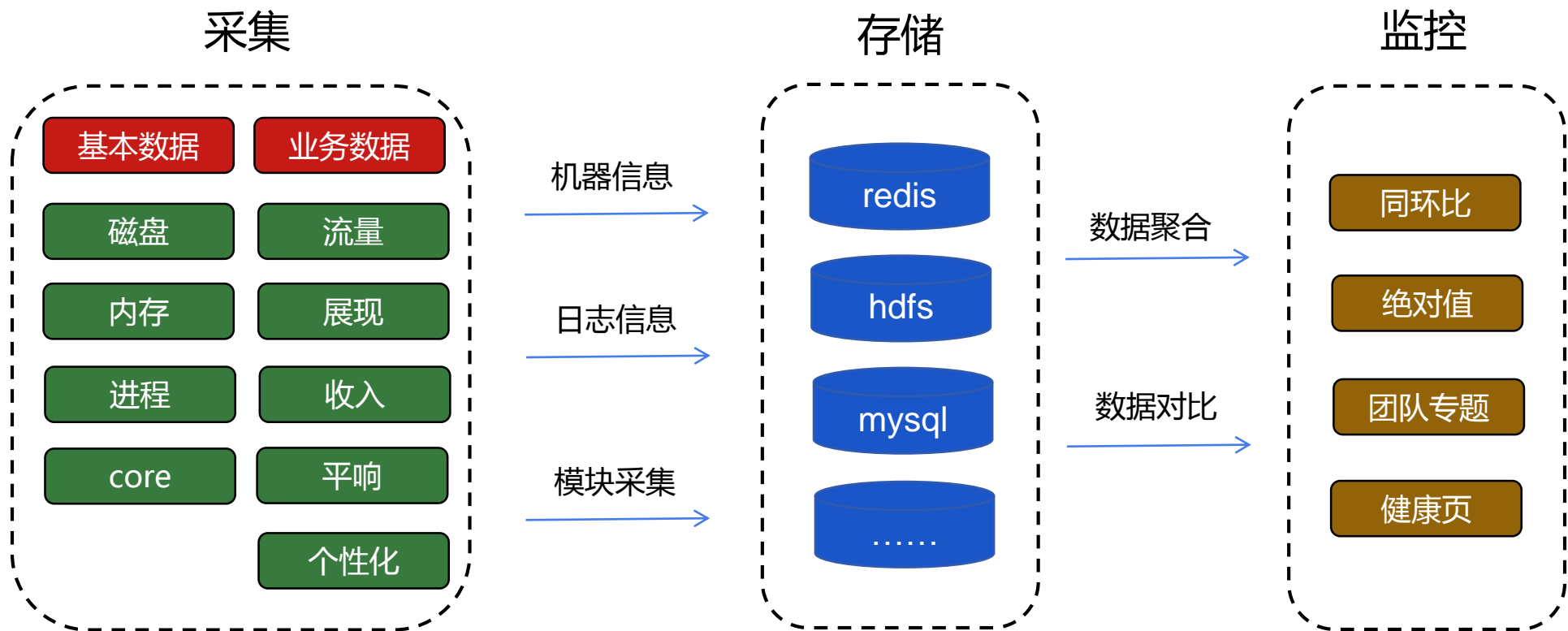


北京  
上海  
深圳  
杭州  
.....



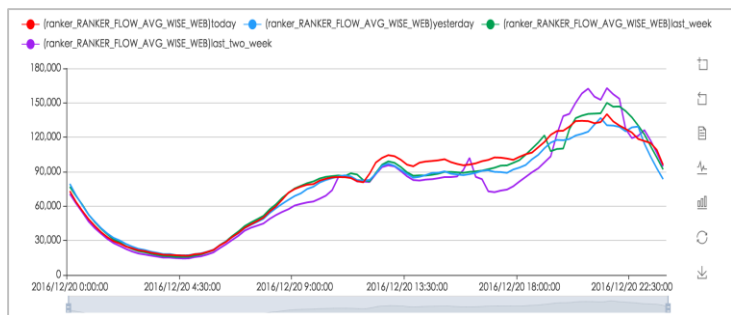
不进行业务逻辑检查

目的：对线上服务了如指掌



## 支持监控图表类型：

- ✓ 趋势图：同环比、多维对比
- ✓ 表格：报表展现等
- ✓ 其他个性化图表：分布图，健康页等



## 主要报警类型：

- ✓ 同环比报警, 绝对值报警, 小流量报警(支持指标所有维度报警巡查)
- ✓ 多种个性化定制: 节假日, 闲时设定, 连续报警方式等
- ✓ 三类通知方式: 短信, 邮件, hi群(报警截图)

### 机器人



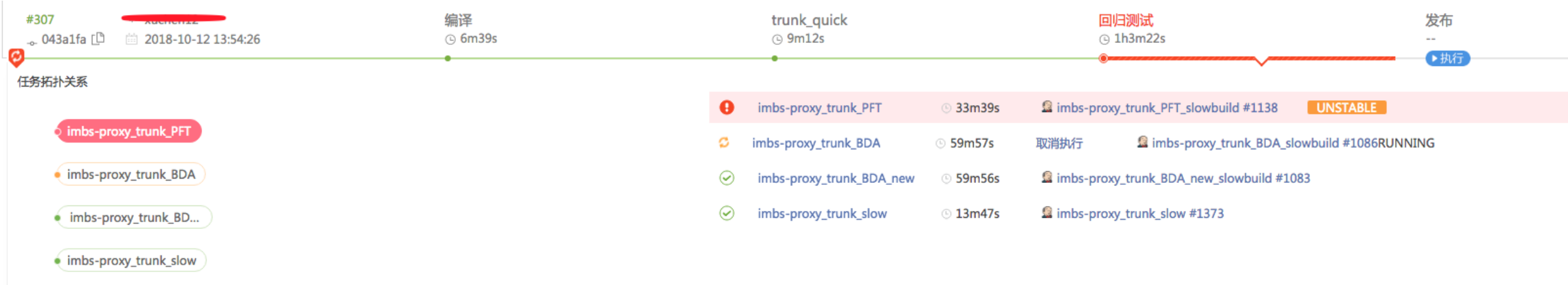
### 机器人

[【ANTI】检索劫持生效过滤量主备DIFF](2016-12-20 22:00:00),绝对值max 137785.0000 > 2000.0000, url: [http://huhang.baidu.com/charts/gui\\_new?ChartId=58452b9c2bfa7b0fb352d46&start\\_str=2016122000000&end\\_str=20161220235959&Product=team\\_special](http://huhang.baidu.com/charts/gui_new?ChartId=58452b9c2bfa7b0fb352d46&start_str=2016122000000&end_str=20161220235959&Product=team_special)

# 凤巢检索系统的DevOps应用效果



## 测试流水线：分级构建



## 需求管理

|            |  |          |           |            |
|------------|--|----------|-----------|------------|
| mixer-3624 |  | features | CoreRe... | 2018-10-15 |
| mixer-3623 |  | features | 待安排       | 2018-10-12 |
| mixer-3622 |  | features | RB测试      | 2018-10-12 |
| mixer-3621 |  | features | 待安排       | 2018-10-11 |
| mixer-3620 |  | features | 已安排       | 2018-10-11 |
| mixer-3618 |  | features | 已发布       | 2018-10-10 |



## 工程能力地图





# THANKS

Website :  
[chinadevopsdays.org/](http://chinadevopsdays.org/)

Global Website:  
[www.devopsdays.org/events/2018-shanghai/](http://www.devopsdays.org/events/2018-shanghai/)

Official Email:  
[organizers-shanghai-2018@devopsdays.org](mailto:organizers-shanghai-2018@devopsdays.org)



Official Wechat

