



360  
互联网专业委员会

# 测试之美 360技术嘉年华

第九季

# 极致diff测试-铺设高效测试之路

分享人：孙聚





# 个人简介

- + 2019年毕业于大连理工大学
- + 先后负责移动网页搜索、引擎测试
- + 搜索事业部测试开发高级工程师

# CONTENTS

01

业务测试痛点

02

Diff如何解决业  
务测试痛点

03

Diff在搜索质量保障  
中的拓展与探索

# 业务测试痛点







## 基础自动化体系

持续集成；  
接口自动化；  
性能自动化；  
前端页面功能自动化；  
前端打点自动化；



## 工具平台化

项目管理平台.；  
性能自动化平台；  
接口功能自动化平台；



## 线上监控&数据挖掘

全链路业务监控平台；  
数据分析平台；

# 360搜索测试中的难点：前端样式繁多

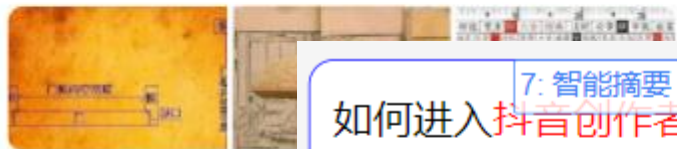
炸大虾

1: 智能摘要 (mso-baike)

鲁班尺又称为鲁班尺,尺寸大小决定出房间风水的大小

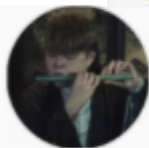
发帖时间: 2018年11月26日

一般来说,在定门窗之前就要确定好这个风水尺寸,再给厂商做,这里要注意框的内空高度... 对于普遍的...



360doc个人图书馆

伦桑 高歌曲下载



酷狗音乐

烟雨行舟 吉他

发帖时间: 2019年

烟雨行舟的吉他谱  
论·艺人:伦桑·专辑  
伦桑(乐谱: 12 粉丝

jitashe.org

一起作业网手机版下载 一起作业网学习端安卓版



更新时间: 4月26日

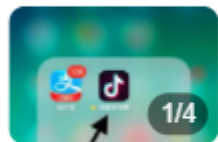
系统: android

下载

太平洋下载中心

1: 智能摘要 (mso-app-download)

如何进入抖音创作者服务中心进行学习



首先点击手机桌 然后点击屏幕右 接着点击屏幕右 面中的抖音。 下方的我的。 上方的三条。

百度经验 2020年2月15日

7: 智能摘要 (mso-jingyan-step)

离婚有方: 专业律师帮您打离婚官司



简介: 《离婚有方专业律师帮您打离婚官司》是2011年法律出版社出版的图书, 作者是杨晓林, 贾明军。内容简介幸福...

3: 智能摘要 (mso-baike-special)

饺子抱蛋的做法步骤图,饺子抱蛋怎么做好吃 燕子燕 8



材料: 饺子 看个人用量 葱花少...  
做法: 1、忘了拍步骤:先将不粘锅烧热放入少许油,油热后加入...

9: 智能摘要 (mso-cookbook)

三五图库大全 天涯博客 天涯社区

乖乖图库 三五图库大全

发表时间: 2011年5月5日

三五图库大全 《三枪》首映礼很平常 阵容不容小看

发表时间: 2011年5月4日

港妹图库 城市跨界越野车 MG 3SW 带你畅游十一

发表时间: 2011年5月4日

tk35tk353.blog.tianya.cn

6: 智能摘要 (mso-blog-home)

精神疾病的一种表现, 患者一般以自心, 富有表演色彩, 情感丰富多变, 现自己, 成年以后伴有不同的精神症状和...



接口上万个字段验证

底层海量数据正确性

链路中数十个服务回归



# Diff如何解决业务测试痛点

- 样式差异diff
- 接口差异diff

# 前端样式差异diff



前端回  
归类型:

智能摘要类型**100+**

box类型**500+**



圆角化



手工测试

抽样测试;  
覆盖率低;  
漏测易导致线上事故;



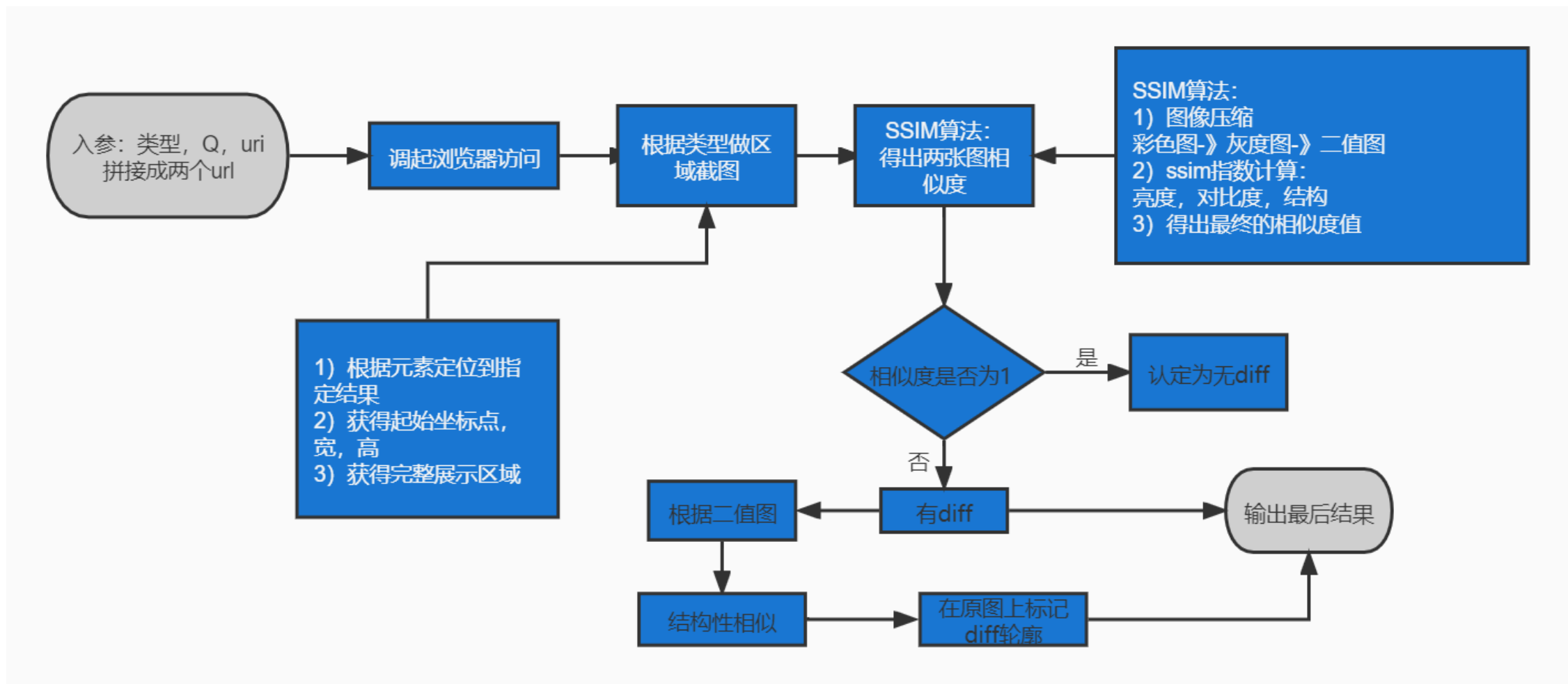
自动化回归

编码成本高;  
耗时久, 回归效率较低;



前端样式diff

功能覆盖率  
可达**100%**;  
保证服务正  
确性;

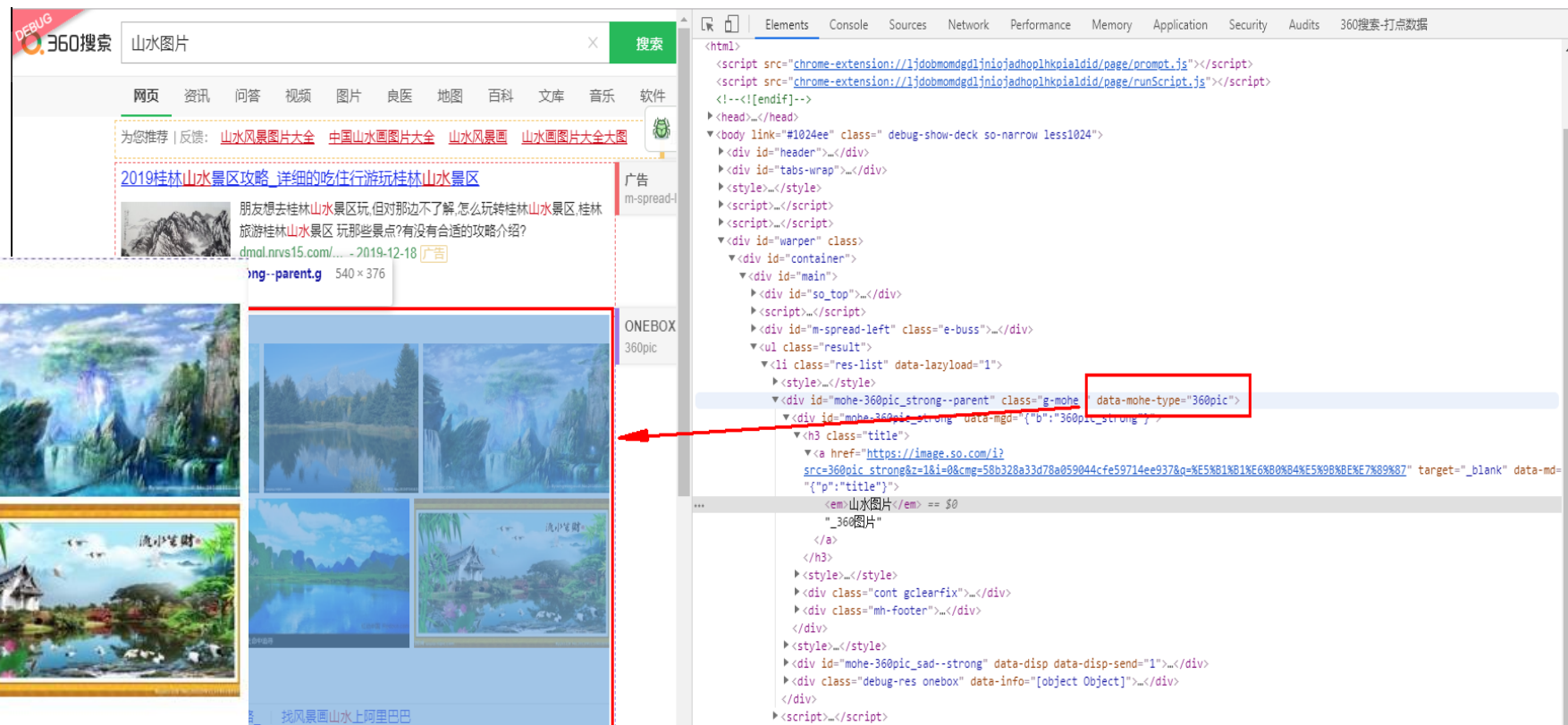


[山水图片](#) [360图片](#)



image.so.com

2019桂林山水攻略 | 找山水画技法上阿里巴巴



- 通过**selenium**无头浏览器截图，确保网页完全加载，通过**JS**像素点逐步加载来实现
- 提高**diff**效率，减少**diff**误差，仅对结果展示指定区域截图



## SSIM: 图像压缩

彩色图  
(三通道)



灰度图  
(单通道)



二值图  
(单通道二值)



## SSIM:图像对比

亮度对比函数



$$L(X, Y) = \frac{2u_X u_Y + C_1}{u_X^2 + u_Y^2 + C_1}$$

对比度对比函数



$$C(X, Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}$$

结构对比函数



$$S(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3}$$

$$SSIM(X, Y) = L(X, Y) * C(X, Y) * S(X, Y)$$

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

[W 2012高考满分作文](#) 浏览2090次

[W 《谁欠谁的幸福》.2012高考满分作文](#) 浏览1012次

[W 2012高考满分作文之三](#) 浏览9851次

[W 安徽2012高考满分作文800字](#) 浏览9780次

[ishare.iask.sina.com.cn](#)

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

14页 [W 2012高考满分作文600字](#) 浏览991次

9页 [W 2012高考英语满分作文范文](#) 浏览180次

13页 [W 2012高考英语满分作文范文](#) 浏览47次

17页 [W 2012高考满分作文\\_免费下载](#) 浏览219次

[查看更多文档>>ishare.iask.sina.com.cn](#)

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

[W 2012高考满分作文](#) 浏览2090次

[W 《谁欠谁的幸福》.2012高考满分作文](#) 浏览1012次

[W 2012高考满分作文之三](#) 浏览9851次

[W 安徽2012高考满分作文800字](#) 浏览9780次



[查看更多文档>>](#)

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

14页 [W 2012高考满分作文600字](#) 浏览991次

9页 [W 2012高考英语满分作文范文](#) 浏览180次

13页 [W 2012高考英语满分作文范文](#) 浏览47次

17页 [W 2012高考满分作文\\_免费下载](#) 浏览219次

[查看更多文档>>](#)

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

[W 2012高考满分作文](#) 浏览2090次

[W 《谁欠谁的幸福》.2012高考满分作文](#) 浏览1012次

[W 2012高考满分作文之三](#) 浏览9851次

[W 安徽2012高考满分作文800字](#) 浏览9780次

[ishare.iask.sina.com.cn](#)

[W 2012高考语文满分作文. 爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

14页 [W 2012高考满分作文600字](#) 浏览991次

9页 [W 2012高考英语满分作文范文](#) 浏览180次

13页 [W 2012高考英语满分作文范文](#) 浏览47次

17页 [W 2012高考满分作文\\_免费下载](#) 浏览219次

[查看更多文档>>ishare.iask.sina.com.cn](#)

[查看更多文档>>](#)



# 前端样式差异diff： 平台化结果展示

task_id	任务名称	类型	query数	diff率	任务状态	开始时间		
normal_1631004427	pic_diff	通用	49	95%	执行完成	2021-09-07T16:47:07.513083		
normal_1631003910	pic_diff	通用	4	75%	执行完成	2021-09-07T16:38:30.622145		
normal_1631003117	pic_diff	通用	4	0%	执行完成	2021-09-07T16:25:17.758899		
normal_1631002799	pic_diff	通用	4	0%	执行完成	2021-09-07T16:19:59.861193		
normal_1631002517	pic_diff	通用	4	0%	执行完成	2021-09-07T16:15:17.882005		
normal_1627477743	pic_diff	通用	4	660%	执行完成	2021-07-28T21:00:02.017594		
		query					终端	Diff结果
normal_1627476355	pic_diff							
normal_1627455801	pic_diff	新商盟					pc	失败:0.345
normal_1627455363	pic_diff	4399小游戏大全					pc	失败:0.399
pic_1627455257	pic_diff	钉钉					pc	失败:0.837
		58同城					pc	失败:0.899
		网易云音乐					pc	失败:0.579

## 一眼并找不到哪里有diff



W [2012高考语文满分作文](#) [爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

W [2012高考满分作文](#) 浏览2090次

W [《谁欠谁的幸福》](#) [2012高考满分作文](#) 浏览1012次

W [2012高考满分作文之三](#) 浏览9851次

W [安徽2012高考满分作文800字](#) 浏览9780次

[ishare.iask.sina.com.cn](#)

W [2012高考语文满分作文](#) [爱问共享资料](#)

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

14页 W [2012高考满分作文600字](#) 浏览991次

21页

9页 W [2012高考英语满分作文范文](#) 浏览180次

5页

13页 W [2012高考英语满分作文范文](#) 浏览47次

91页

17页 W [2012高考满分作文免费下载](#) 浏览219次

30页

[查看更多文档>>ishare.iask.sina.com.cn](#)

[查看更多文档>>](#)

## 2012高考语文满分作文. 爱问共享资料

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

2012高考满分作文 浏览2090次

《谁欠谁的幸福》,2012高考满分作文 浏览1012次

2012高考满分作文之三 浏览9851次

安徽2012高考满分作文800字 浏览9780次

ishare.iask.sina.com.cn

查看更多文档>>ishare.iask.sina.com.cn

## 2012高考语文满分作文. 爱问共享资料

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

14页 2012高考满分作文600字 浏览991次

21页

9页 2012高考英语满分作文范文 浏览180次

5页

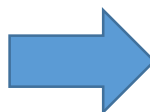
13页 2012高考英语满分作文范文 浏览47次

91页

17页 2012高考满分作文\_免费下载 浏览219次

30页

查看更多文档>>



## 2012高考语文满分作文. 爱问共享资料

3页 - 浏览640次

爱问共享资料拥有大量关于2012高考满分作文的实用类文档资料，所有文档由知名合作机构以及专业作者提供，线上总资料超过两个亿，保证满足您的需求。

2012高考满分作文600字 浏览991次

21页

2012高考英语满分作文范文 浏览180次

5页

2012高考英语满分作文范文 浏览47次

91页

2012高考满分作文\_免费下载 浏览219次

30页

ishare.iask.sina.com.cn

查看更多文档>>

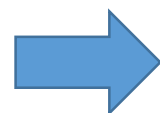


前端样式迭代

底层数据优化

前端技术升级

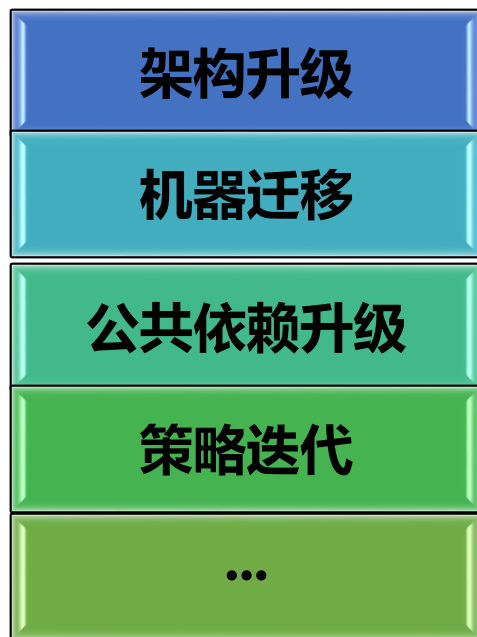
...



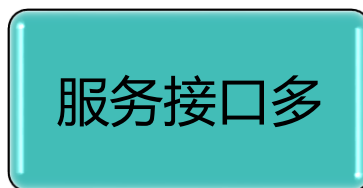
批量回归case, 覆盖可达100%

大型迭代项目一键回归数据

# 接口差异diff



整个搜索涉及服务模块数十个



每个服务涉及数十个接口

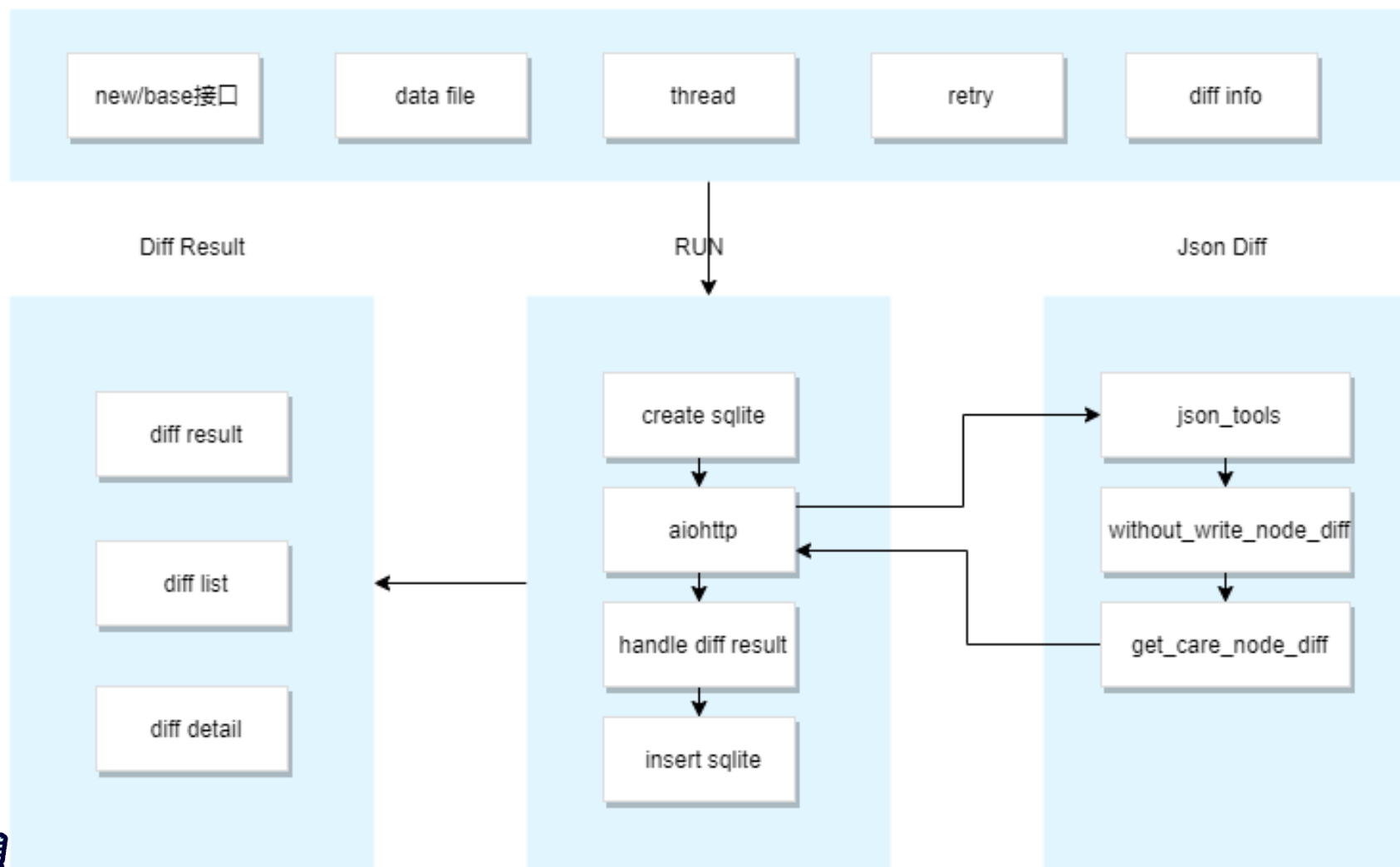


每个接口有多个参数逻辑



以上笛卡尔乘积结果，对应超多回归项





## Json diff: 获得全部diff数据

- 1) 字段缺失
- 2) 字段值不一致
- 3) 字段新增

```
def get_all_diff(self):
    all_diff = json_tools.diff(self.base_json, self.new_json)
    # 处理diff结果
    for i in range(len(all_diff)):
        # 如果diff类型为字段值不一致
        diff_dict = all_diff[i]
        if diff_dict.get("replace"):
            diff_dict["type"] = "replace"
            key = diff_dict.get("replace")
            base_value = diff_dict.get("prev")
            new_value = diff_dict.get("value")
            diff_dict["key"] = key
            diff_dict["info"] = "{key} 字段不一致, Base值为 {base}, New值为 {new}"

        # 如果 diff类型为字段缺失
        elif diff_dict.get("remove"):
            diff_dict["type"] = "remove"
            key = diff_dict.get("remove")
            diff_dict["key"] = key
            diff_dict["info"] = "{key} 字段缺失".format(key=key)

        # 如果 diff类型为新增字段
        elif diff_dict.get("add"):
            diff_dict["type"] = "add"
            key = diff_dict.get("add")
            diff_dict["key"] = key
            diff_dict["info"] = "{key} 新增字段".format(key=key)

        # 识别不了的diff类型
        else:
            diff_dict["type"] = "other"
            diff_dict["key"] = "none_key"
            diff_dict["info"] = "无法识别的diff类型"

    return all_diff
```

## Json diff: diff结果整理

1) 剔除白名单的diff内容

2) 提取关心节点的diff内容

```
def get_without_write_node_diff(self):
    """
    获取除了白名单产生的diff 的diff结果
    :return:
    """
    # 白名单节点造成diff的列表
    write_node_diff_list = []
    for write_node in self.write_node_list:
        for diff_dict in self.all_diff:
            key = diff_dict.get("key")
            if key.startswith(write_node):
                write_node_diff_list.append(diff_dict)
    without_write_node_diff = [item for item in self.all_diff if item not in write_node_diff_list]
    return without_write_node_diff

def get_care_node_diff(self):
    """获取关心节点diff(已经去掉了白名单产生的diff)"""
    care_node_diff_list = []
    for care_node in self.care_node_list:
        for diff_dict in self.without_write_node_diff:
            key = diff_dict.get("key")
            if key.startswith(care_node):
                care_node_diff_list.append(diff_dict)
    return care_node_diff_list
```

Q: diff不稳定, 排查未发现diff原因

diff优化: 去除噪点

- 1) 获取全部diff数据
- 2) 相同环境二次diff, 作为噪点
- 3) 相减得到最终diff结果





Q: diff呈块状, 且排查发现只是位置有变化, 内容相同

## Diff优化: 识别结果rank进行diff

### 1) 按格式进行解析diff

### 2) 遍历对比, 获取相同结果的排序位置

```
# 开始 new-diff
for index_new in range(new_length):
    # 相同位置内容全相等
    if index_new < base_length and rank_new[index_new].get("teipe_md5") == rank_base[index_new].get("teipe_md5"):
        # 当前数据不存在diff
        rank_new[index_new]["diff_rank"] = index_new + 1
        rank_base[index_new]["diff_rank"] = index_new + 1
        continue
    # 相同位置内容不全相对
    result_dict["has_diff"] = "diff"

if index_new < base_length and rank_new[index_new].get("teipe_md5") != rank_base[index_new].get("teipe_md5"):
    for key in key_list:
        if rank_new[index_new].get(key, "") != rank_base[index_new].get(key, ""):
            result_dict["diff_type"] = key
# 相同位置对方无结果, 遍历不同位置
if index_new >= base_length:
    for index_base in range(base_length):
        if rank_new[index_new].get("teipe_md5") == rank_base[index_base].get("teipe_md5"):
            result_dict["diff_type"] = "rank"
            rank_new[index_new]["diff_rank"] = index_base + 1
            if rank_base[index_base]["diff_rank"] != rank_base[index_base]["rank"]:
                rank_base[index_base]["diff_rank"] = index_new + 1
```

Q: diff呈块状，且排查发现只是位置有变化，内容相同

## Diff优化：识别结果rank进行diff

1) 按格式进行解析diff

2) 遍历对比，获取相同结果的排序位置

5

[下肢静脉曲张 39健康问答](#)

下肢静脉曲张?39快速问医生为您在线解答下肢静脉曲张! 下肢静脉曲张 手术费用各地医院都是不一样的, 所以没有准确的数据, 但是一般来说手术要上千元。静脉曲张是...

<http://wapask.39.net/keywords/631953.html>

from:2 erro:0 source:2 dspcnt:640 total:338000

6 7

[上肢静脉曲张 家庭医生在线](#)

家庭医生在线健康宝典(是中国最专业的健康热点资讯答疑平台, 问题涵盖面广, 最新的'上肢静脉曲张'书写'相关资讯。最专业的'上肢静脉曲张'书写'就医指导。创建千家...

<https://baodian.familydoctor.com.cn/4361873>

from:2 erro:0 source:2 dspcnt:640 total:338000

7 8

[血栓闭塞性脉管炎 快速问医生 免费咨询](#)

血栓闭塞性脉管炎 医生回复 呵呵, 这个和它的病格式是一样的, 只是把详情检查 辅助检查的结果 按实际情况写就可以了。

<https://m.120ask.com/qgywy/waikelipuwaik/482499.htm>

from:2 erro:0 source:2 dspcnt:640 total:338000

8 9

[血栓闭塞性脉管炎 快速问医生 免费咨询](#)

5

[下肢静脉曲张 39健康问答](#)

下肢静脉曲张?39快速问医生为您在线解答下肢静脉曲张! 下肢静脉曲张 手术费用各地医院都是不一样的, 所以没有准确的数据, 但是一般来说手术要上千元。静脉曲张是...

<http://wapask.39.net/keywords/631953.html>

from:2 erro:0 source:2 dspcnt:640 total:338000

6 无

[血栓性脉管炎 39问医生 39健康网](#)

血栓性脉管炎 发病时间不清楚 小腿肌肉疼痛 打完篮球后小腿肿胀 包块疼痛难忍, 自以为自己肌肉发达没在意 医院检查没说出病因 从小腿处抽出20毫升血 肌肉松...

<http://ask.39.net/question/36160384.html>

from:2 erro:0 source:2 dspcnt:640 total:338000

7 6

[上肢静脉曲张 家庭医生在线](#)

家庭医生在线健康宝典(是中国最专业的健康热点资讯答疑平台, 问题涵盖面广, 最新的'上肢静脉曲张'书写'相关资讯。最专业的'上肢静脉曲张'书写'就医指导。创建千家...

<https://baodian.familydoctor.com.cn/4361873>

from:2 erro:0 source:2 dspcnt:640 total:338000

8 7

[血栓闭塞性脉管炎 快速问医生 免费咨询](#)

Diff 测试

★ 接口 Diff

★ 排序 Diff

★ 图片 Diff

★ 代码 Diff

★ Diff 工具

首页

接口 Diff

task\_id

task\_name

Author

任务状态

查询

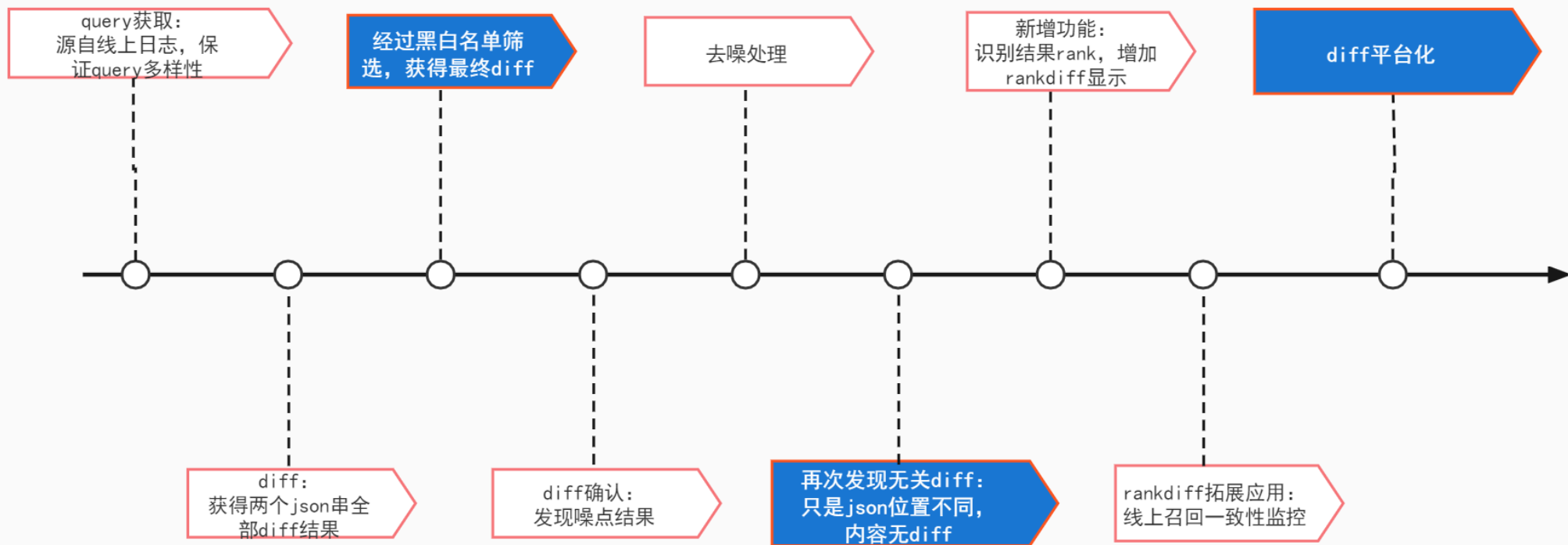
我的任务

新建任务

更多列

任务名称	任务进度	query数	diff率	开始时间	负责人	操作
gaia_suggest_diff	100%	10000	0 %	2021-10-25T20:23:58		复制 Stop Delete
gaia_suggest_diff	100%	10000	0 %	2021-10-24T17:57:18		复制 Stop Delete
gaia_suggest_diff	40%	10000	0 %	2021-10-24T17:36:44		复制 Stop Delete
gaia_suggest_diff	100%	10000	0 %	2021-10-24T16:31:58		复制 Stop Delete
gaia_suggest_diff	16%	10000	100 %	2021-10-24T16:09:57		复制 Stop Delete

jsondiff



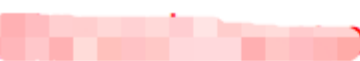


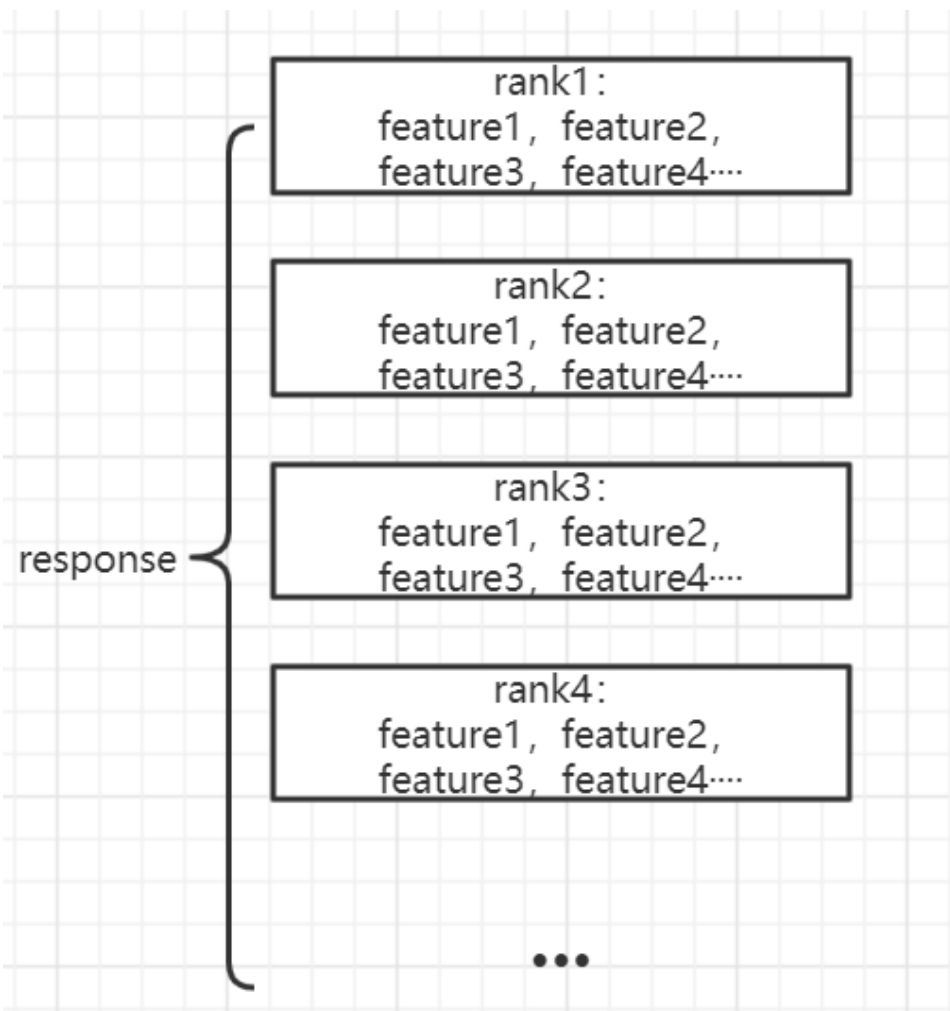
# Diff在搜索质量中的拓展与探索

+ review board: 问题发现前置: 提高提测质量

+ package diff: package大小; 配置文件list; 配置文件内容diff

Revision	New Change
49 lines hidden [Expand]	
50 label=_("Enable search"),	50 label=_("Enable search"),
51 help_text=_("Provides a search field for quickly	51 help_text=_("Provides a search field for quickly
searching through "	searching through "
52 "review requests."),	52 "review requests."),
53 required=False)	53 required=False)
54	54
55 search_index_file = forms.CharField(	55 search_index_dir = forms.CharField(
56 label=_("Search index file"),	56 label=_("Search index directory"),
57 help_text=_("The file that search index data should be	57 help_text=_("The directory that search index data files
stored in."))	will be "
	58 "stored in. This defaults to \"search-index\"
	in your "
	59 "site directory."))
58 required=False,	60 required=False,
59 widget=forms.TextInput(attrs={'size': '50'}))	61 widget=forms.TextInput(attrs={'size': '50'}))
60	62
61 auth_backend = forms.ChoiceField(	63 auth_backend = forms.ChoiceField(
62 label=_("Authentication Method"),	64 label=_("Authentication Method"),
78 lines hidden [Expand]	
141 ', '.join(self.siteconfig.get('auth_custom_backends'))	143
142	144
143 can_enable_search, reason = get_can_enable_search()	145
144 if not can_enable_search:	146
145 self.disabled_fields['search_enable'] = True	147
146 self.disabled_fields['search_index_file'] = True	148
147 self.disabled_reasons['search_enable'] = _(reason)	149
148	150
149 can_enable_ldap, reason = get_can_enable_ldap()	151
150 if not can_enable_ldap:	152
151 self.disabled_fields['auth_ldap_uri'] = True	153
6 lines hidden [Expand]	
158 self.disabled_fields['auth_ldap_anon_bind_password']	160
= True	
159 self.disabled_reasons['auth_ldap_uri'] = _(reason)	161
160	162
161 super(GenericSettingsForm, self).load()	163
162	164

- 不仅  r.so变动, 请到回归测试子job确认
- 所有集群配置文件均发生变更。请确认
- pass all regress check (see run\_info for detail)
- 服务文件发生变动, 请确认



预期：

**baseRank1(feature) > Rank2(feature)**

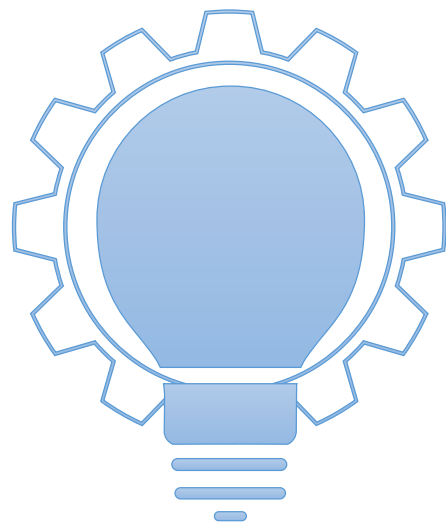
逆序率应用：

- + **new & base**环境逆序率diff，确定服务稳定性&正确性
- + 热门**query-badcase**优化特征值计算流程

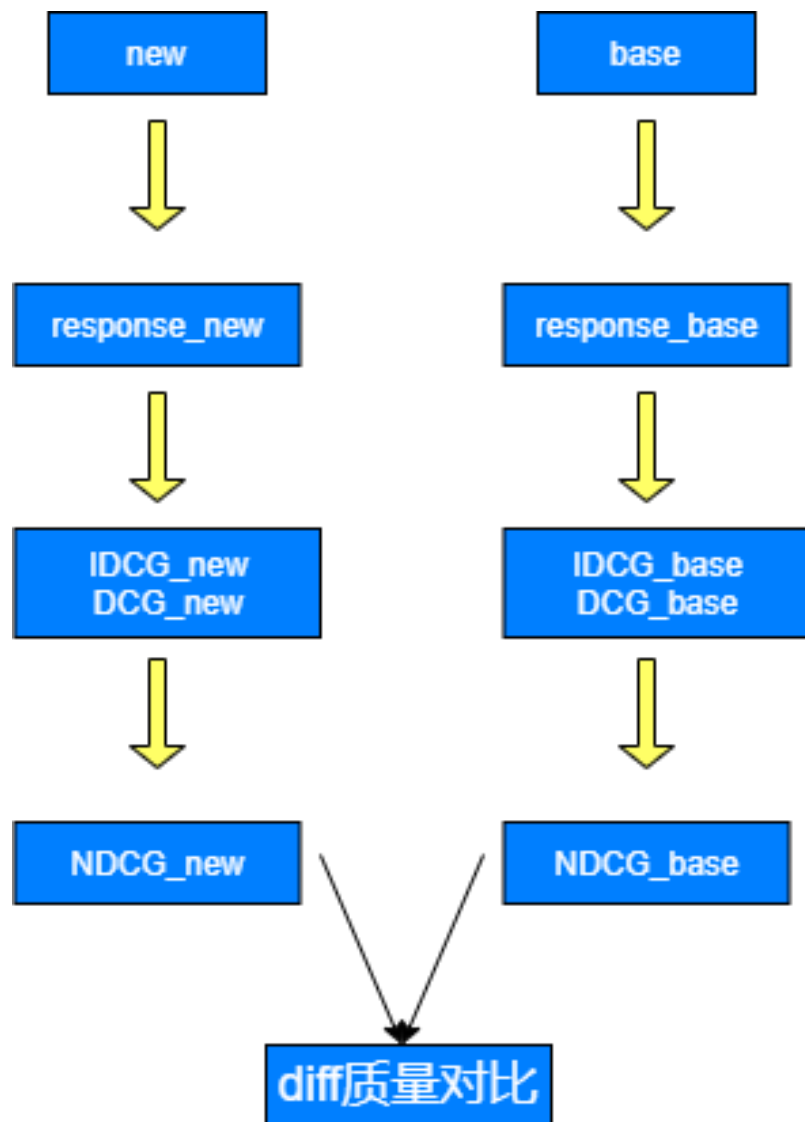
## 接口diff拓展应用：线上rank diff质量监控







**深入挖掘rank diff如何产生：**  
排序过程中的每一个决策树的得分进行一次diff展示，帮助策略优化，提效diff排查



注:

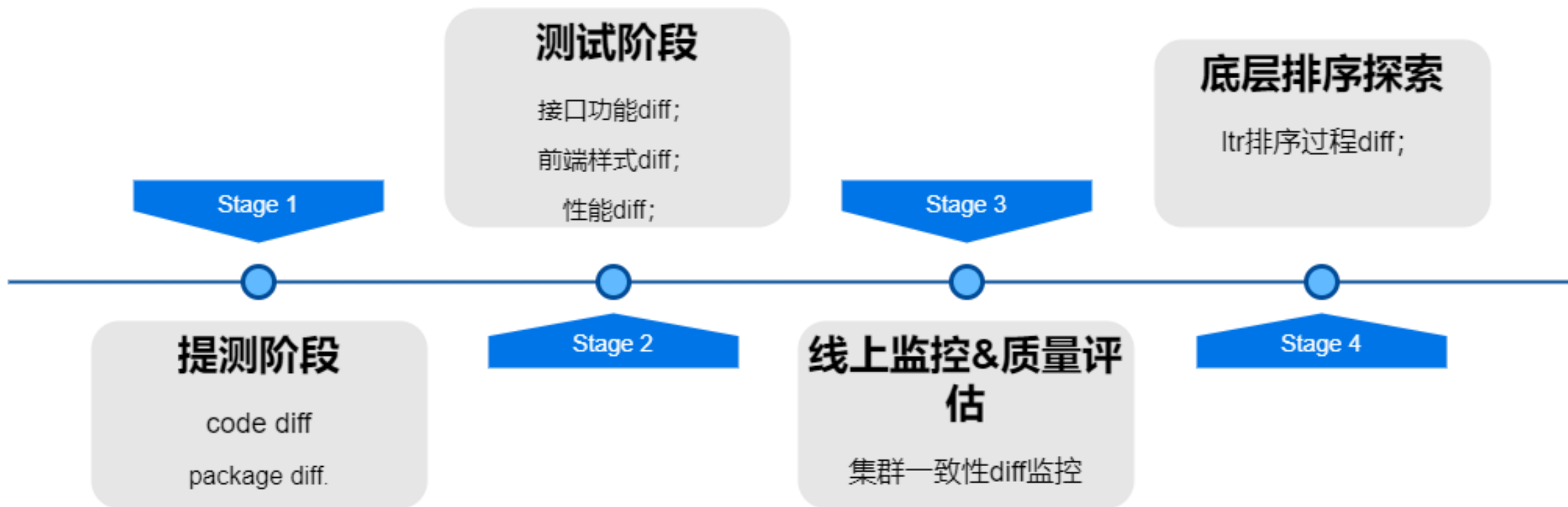
$nDCG = DCG / IDCG$ ;

**IDCG**: 人工标注的最佳排序

**DCG**: 位置加权后的排序得分

**nDCG**: 归一化得分, 可以做分支对比而不需要考虑每次召回结果个数不同的问题

# diff总结





格局打开  
拓宽思路  
万物皆可diff

# THANKS



360技术



360质量效能