

# 产品简介

故障自愈是行业领先的"故障自动化处理"解决方案，提升企业的服务可用性和降低故障处理的人力投入，实现故障自愈从"人工处理"到"无人值守"的变革！

{% video %}media/fta\_brand\_video.mp4{% endvideo %}

通过自动化处理节省人力投入，通过预定的恢复流程让恢复过程更可靠，通过并行分析达到更快的故障定位和恢复。

一句话概括：实时发现告警，预诊断分析，自动恢复故障，并打通周边系统实现整个流程的闭环。

## 问题反馈

- QQ: 800802001
- 邮件: contactus\_bk@tencent.com
- [我给故障自愈提需求](#)

## 术语解释

### 自愈套餐

故障自愈收到告警后，执行的动作，比如作业套餐。

自愈套餐		自愈套餐说明
套餐类型 作业平台	套餐命名* 脚本采集2:5分钟负载-单台	套餐是业务运维设计制作的一套恢复故障的方案，可以复用于不同的告警，也可作为原子套餐用于制作组合套餐。 详情： <a href="#">【自愈套餐大全】</a>
业务： 蓝鲸	作业名称： 自愈:脚本采集2:5分钟负载-单台 (3116)	
<input type="checkbox"/> 作业失败重试 <input type="checkbox"/> 向作业平台自定义参数 <input type="checkbox"/> 从作业中获取参数 <input checked="" type="checkbox"/> 用告警IP替代作业执行IP		
<button>保存自愈套餐</button>		

## 自愈方案

关联告警 和 自愈套餐的策略

自愈场景

告警类型 \* [主机监控] 5分钟平均负载 ▾ 按内容筛选 使用正则表达式匹配告警, 不填为不过滤  
集群 默认全选 模块 默认全选

自愈处理

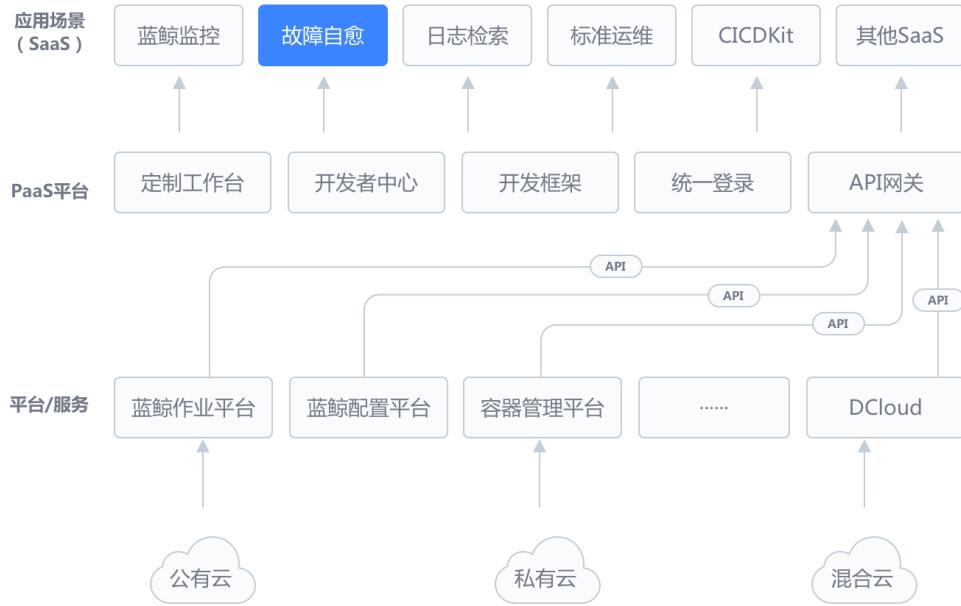
自愈套餐 『快捷』发送CPU使用率TOP10的进程(微信) ▾ 查看 +

## 产品架构图

故障自愈集成行业主流开源监控产品或以 REST API 方式获取企业监控产品的告警，匹配告警设置的**自愈套餐**，同时通过**收敛防御** 实现在安全的前提下完成告警处理的**无人值守**。

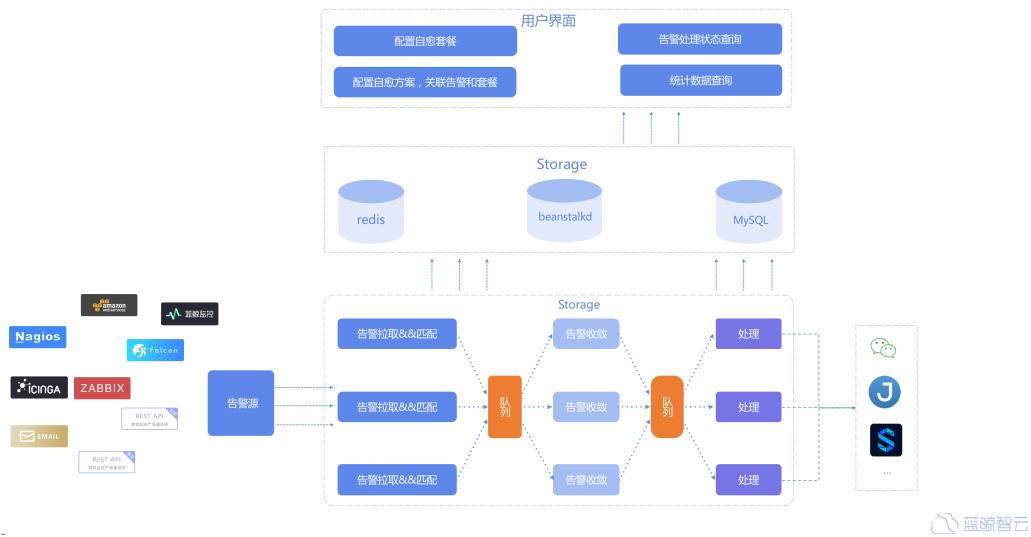
### 故障自愈在蓝鲸中的位置

在蓝鲸的体系架构中，故障自愈位于**应用场景 SaaS** 层，通过蓝鲸 PaaS 平台的 ESB(API 网关)调用配置平台(CMDB)、作业平台、标准运维等产品的 API，实现告警的自动化处理。

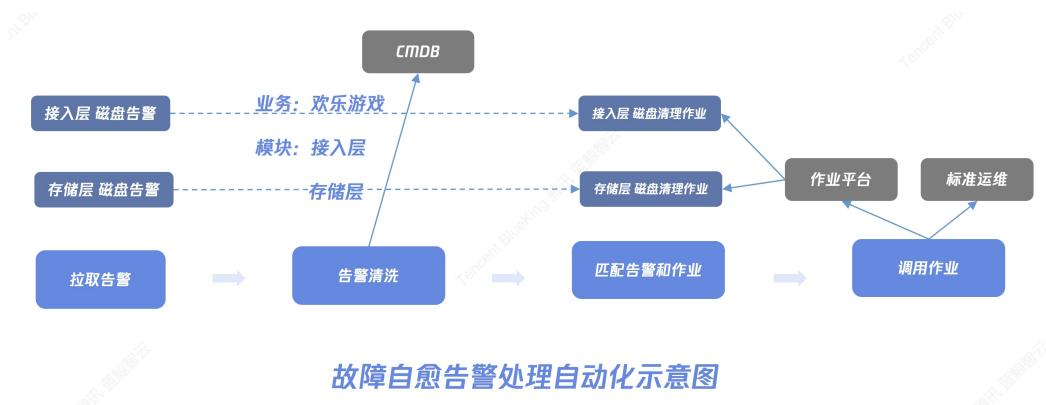


### 故障自愈产品架构图

故障自愈的告警拉取模块，周期性或通过 回调 方式从各大监控系统中获取告警，通过调用蓝鲸 CMDB 的 API 解析 告警所属的业务、模块，查询该业务、模块在故障自愈配置的处理动作，经过 告警收敛 模块的过滤以确认没有大批量相同属性(如同业务、机房等)，最后执行对应的 处理动作，告警恢复，业务恢复正常。



为了便于理解，下图为欢乐游戏业务的接入层、存储层发生磁盘使用率告警的自动化处理示意图。



将人的处理经验沉淀到自动化工具上，故障自愈，重新定义告警的处理方式！



故障自愈和传统告警处理方式对比

## 设计理念

以产品设计理念剖析企业建设故障自动化处理方案的思路

人工处理告警，一直是运维心中的痛。大年初一拜年、结婚、和老婆孩子外出过周末等美好时光，作为运维的你，好像一直心系 IT 系统，保持与笔记本的安全距离。

为什么这么多年过去了，还是这么苦逼，不是说运维行业转 AIOps 了，我竟然还在手工处理告警，我该怎么办？

本文介绍故障自愈攻克的 3 个技术点，以及 献上开箱即用的方案。

### 基本流程

自动化的要点是什么？把人的经验抽象、固化为程序处理，工业(第 3 次工业革命)或互联网都是如此。

举个例子，磁盘出现告警，运维首先想到的是登陆服务器清理磁盘。

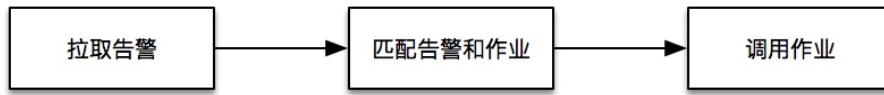


接下来，我们拆解背后的逻辑。

### 抽象告警处理流程

- 拉取磁盘告警
- 编写磁盘清理的脚本或作业任务
- 设计关联模块：把拉取到的磁盘告警，与调用脚本的模块串起来

流程图如下：

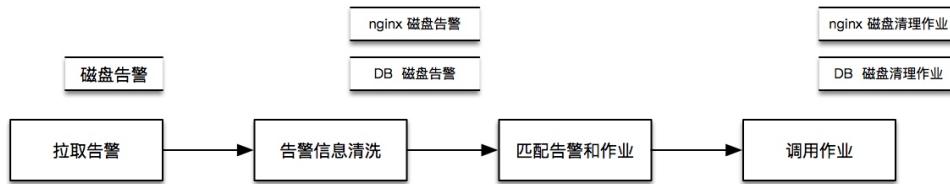


## 通过 CMDB 做资源清洗

不同模块的磁盘清理方案不一样，如何解决呢？

这时需要引入 CMDB(设备、人、业务的映射关系)，通过 CMDB 把 IP 清洗为 模块，这样就解决了接入层 和 逻辑层、存储层的 告警使用对应的磁盘清理方案。

流程图如下：



## 对接企业内部网关

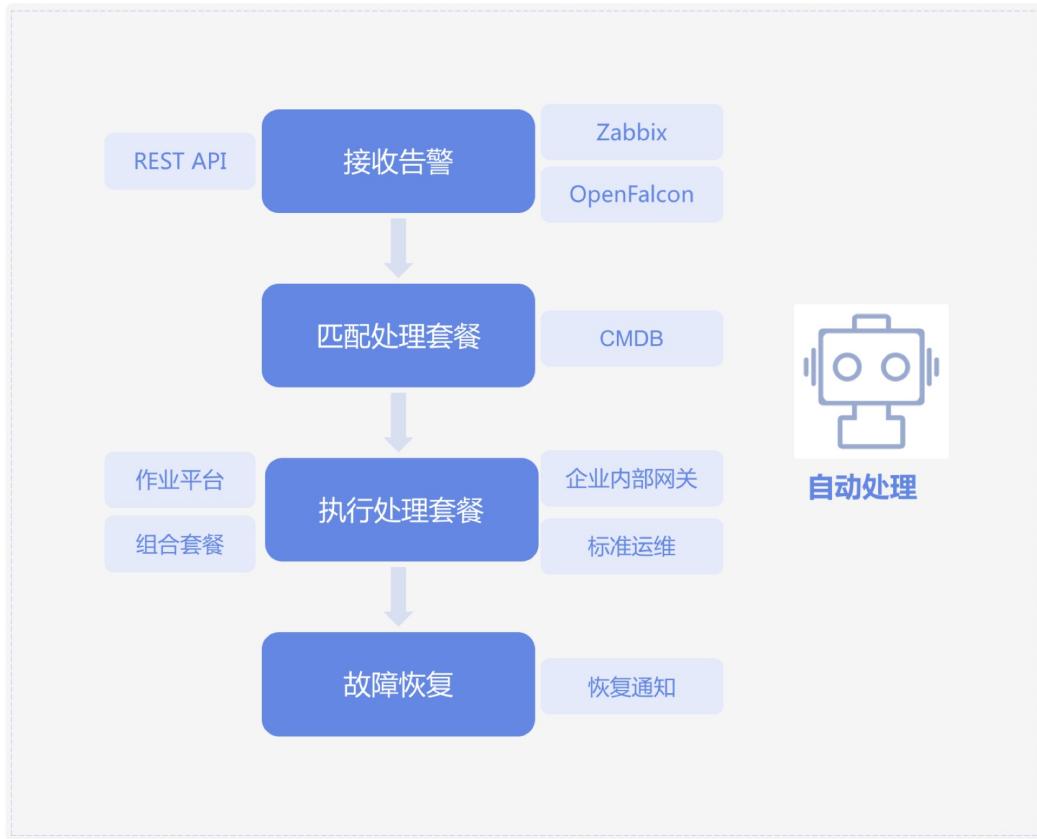
故障自愈可能会处理失败，这时需要通知用户。故障自愈的处理方式除了调用作业外，还可能需要调用企业内部的网关，比如服务器重启、申请服务器等。

使用 PaaS 层的 ESB 是一种解决思路，通过 ESB 封装企业内部网关，解决权限校验、频率控制、访问统计、路由分发以及自助接入等功能，不要直接调用裸接口了。

下图为故障自愈调用邮件和微信通知网关：



经过这一轮的探索，故障自动处理的流程如下：



### 对接企业内部监控产品

以 Zabbix、Open-Falcon 为例，介绍如何对接企业内部的监控产品，。



### 对接 ZABBIX

《当 Zabbix 遇见故障自愈》介绍了拉取 Zabbix 告警的方案，通过 ActionScript 调用脚本，把 Zabbix 告警推送至自愈的告警拉取模块。

推送(或叫回调)可以保证告警拉取的实时性。

下图为 Zabbix 推送告警示例：

The screenshot shows the Zabbix Action log interface. A single alert entry is displayed:

- Time:** 2017-09-21 12:29:37
- Action:** FTA\_Act
- Type:** FTA\_Event\_Handler
- Recipient(s):** FTA\_Mgr (FTA\_Mgr FTA\_Mgr)  
FTA\_Mgr
- Message:**

```
*****
#001
HOSTCONN: 19.0.4.3
HOSTHOST: Agent3
HOSTNAME: Agent3
ITEM.ID: 20796
ITEM.KEY: system.swap.size[free]
TRIGGER.NAME: Lack of free swap space on Agent3: PROBLEM
TRIGGER.EXPRESSION: (Agent3) system.swap.size[free].last()<50
TRIGGER.DESCRIPTION: It probably means that the system requires more physical memory.
TRIGGER.URL:
TRIGGER.STATUS: PROBLEM
EVENTID: 180
EVENTTYPE: 36
EVENTDATE: 2017-09-21
EVENTTYPEVALUE: 1
ACTION.ID: 7
ACTION.NAME: FTA_Art
*****
```
- Status:** Sent

实际上是 Zabbix 调用推送告警的脚本：

The screenshot shows the Zabbix Media types configuration page. A new media type entry is shown:

- Name:** FTA\_Event\_Handler
- Type:** Script
- Script name:** zabbix\_da\_alarm.py
- Script parameters:** Parameter [ALERTMESSAGE] (Action)
- Enabled:** checked

对接 Zabbix 的落地案例可以参考陈亮撰写的 [那些年我们想做的无人值守](#)。

除 Zabbix 外，Open-Falcon 在国内的社区热度也不错，所以也介绍拉取其告警的方案。

## 对接 OPEN-FALCON

方案类似 Zabbix，不过 Open-falcon 直接提供了 callback 功能，简化了流程。

下图为在 Open-Falcon 告警模板中，将故障自愈接收告警的地址，录入至 **Callback 地址** 中。

The screenshot shows the Open-Falcon alarm template configuration page. The **Callback** field contains the URL:

```
Callback: http://paas.../api/c/compapi/fta/event/open-falcon/...
```

The payload section shows the JSON structure for the callback request:

```
{
  "endpoint": event.Endpoint(),
  "metric": event.Metric(),
  "status": event.Status(),
  "step": event.CurrentStep(),
  "priority": event.Priority(),
  "time": event.FormattedTime(),
  "tp_id": event.TpId(),
  "exp_id": event.ExpressionId(),
  "stra_id": event.StrategyId(),
  "tags": "srv=falcon,mount=sda"
}
```

At the bottom, there are checkboxes for handling messages before and after recovery.

收到了 Open-Falcon 推送的告警后，解析对应的字段即可。

如果企业内部的 CMDB 以 IP 来标识主机，需要再做一层转换，因为 Open-Falcon 的资源标识

`endpoint` 默认是主机名，那么就需要使用 CMDB 的自动发现功能自动上报主机名，同时提供把主机名清洗为 IP 的功能。

下面是 Nginx 模块磁盘告警的自愈示例，匹配 Nginx 模块的磁盘清理套餐，清理 Nginx 模块的日志文件，整个过程不到 30 秒。



告警自动处理看似如此简单，然而并非如此。

## 两面性

故障自动处理就像一把刀，有其两面性。

因为要确保告警的真实性，一旦把假告警也自动处理了，就很悲催了…

举个例子。网络波动，批量出现 PING 告警。实际上服务器运行正常，这时你把服务器都重启了，那就 GG 了。

如何解决呢？分析事物的规律。

批量出现告警，那可以在告警拉取模块后面，增加一个收敛模块：

- 比如，在 X 时间内出现 Y 个告警，打电话给运维审批。
- X 时间内同一主机出现使用相同套餐的告警，则收敛时间窗口中后面的告警则跳过，比如同时收到进程告警 和 端口告警，就不用拉 2 次进程了。

还有就是，原有监控系统没有收敛能力，那么可以借用这个功能来做告警汇总，因为收敛逻辑一样，只是收敛的处理方式有差异。

2017-09-22 - 2017-09-22 11:12

● 1 次收效事件 ● 共 7 次自愈

刷新 导出IP 请输入IP 搜索

类型	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果
收到后处理	今天 00:56:43						服务器可能存在批量停电, 请关注 影响范围: Zabbixtest
Ping检查(icmpping*)	今天 00:45:30	23秒	Zabbixtest	3点2	10.0.0.3	✓	执行Job作业成功[1634]
Ping检查(icmpping*)	今天 00:45:30	21秒	Zabbixtest	3点4	10.0.0.5	✓	执行Job作业成功[1635]
Ping检查(icmpping*)	今天 00:45:30	21秒	Zabbixtest	3点2	10.0.0.7	✓	执行Job作业成功[1636]
Ping检查(icmpping*)	今天 00:45:30	0秒	Zabbixtest	3点4	10.0.0.12	⌚	被收敛: 对于(Ping检查(icmpping*))告警类型, 在(5)分钟...
Ping检查(icmpping*)	今天 00:45:30	0秒	Zabbixtest	3点0	10.0.0.15	⌚	被收敛: 对于(Ping检查(icmpping*))告警类型, 在(5)分钟...
Ping检查(icmpping*)	今天 00:45:30	0秒	Zabbixtest	3点0	10.0.0.41	⌚	被收敛: 对于(Ping检查(icmpping*))告警类型, 在(5)分钟...
Ping检查(icmpping*)	今天 00:45:30	0秒	Zabbixtest	3点2	10.0.0.48	⌚	被收敛: 对于(Ping检查(icmpping*))告警类型, 在(5)分钟...

« 1 »

解决了安全和基本的告警自动处理诉求后，你可能还想处理复杂的故障处理场景。

## 复杂告警的处理方案 - 组合套餐

举个例子，A 模块是重要模块，出现 PING 不可达告警，首先要校验 A 模块是否真的故障，如果真的故障，接下来是从资源池中获取备机 ... 故障替换等等，期间每个环节都有可能出错，那就要考虑异常分支的场景。

树结构可以解决该问题，二叉树足以满足大部分场景(成功、失败两种分支)。



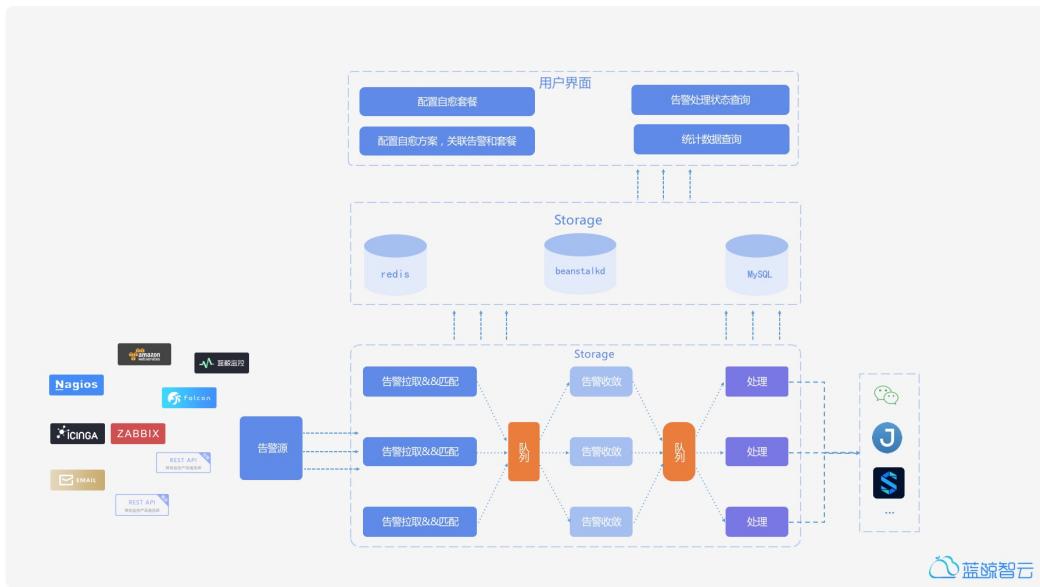
上面这张图，是一个自愈处理方案，可以称之为组合套餐。

这里同时引入了原子的概念，通过组装原子来满足各种需求场景，和 [资源编排](#) 说的是同一个理儿。

注：如果你想使用三叉树，其实可以把组合套餐也作为一个原子套餐(节点)。

## 技术架构

经过前面对 故障自愈的处理流程、故障自愈的两面性、复杂的故障处理方案 的层层梳理，我们有了一张故障自愈的技术架构图。



故障自愈的告警拉取模块，周期性或通过 回调 方式从各大监控系统中获取告警，通过调用蓝鲸 CMDB 的 API 解析 告警所属的业务、模块，查询该业务、模块在故障自愈配置的处理动作，经过 告警收敛 模块的过滤以确认没有大批量相同属性(如同业务、机房等)，最后执行对应的 处理 动作，告警恢复，业务恢复正常。

相信这次以经行业验证的故障自愈做技术剖析，能对大家建设企业内部的故障自动处理方案提供参考思路。

## 特点及优势



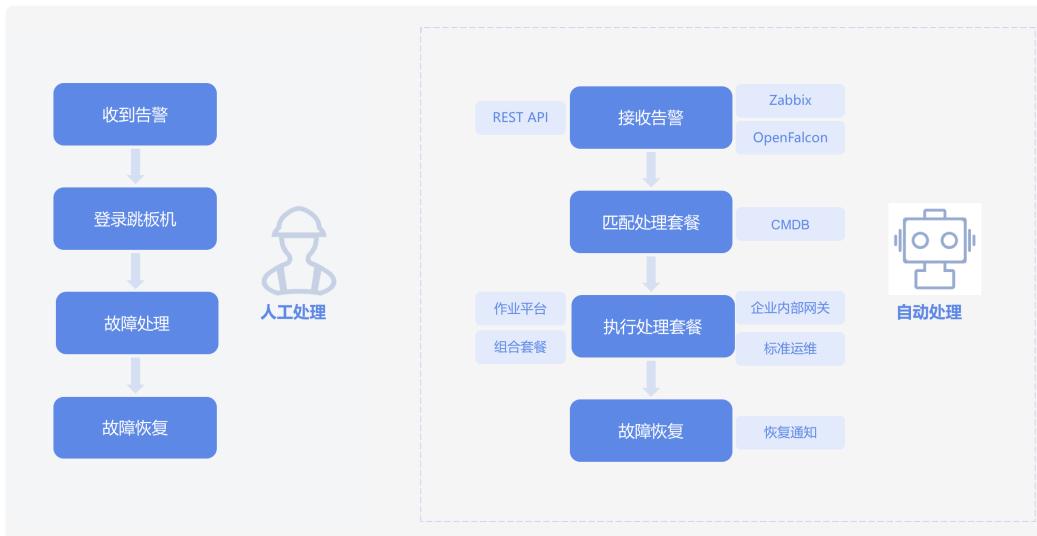
1. 集成主流监控产品：告警源集成蓝鲸监控、4款主流开源监控产品 Zabbix、Open-Falcon、

Nagios、Icinga，及 AWS、邮件的告警接入，更能通过 REST API 拉取、推送告警。

2. 丰富的处理套餐：除支持作业平台、标准运维外，还支持快捷套餐类(磁盘清理、汇总、检测 CPU 使用率 TOP10 等)、组合套餐类(获取故障机备机、通知、审批等)。
3. 告警收敛和防御：系统预定收敛和防御规则，对异常告警事件进行收敛，更能通过收敛审批功能对异常的执行做审批。
4. 组合套餐：把自定义自愈套餐通过 FTA(故障树分析)处理流程，组装成解决复杂场景的组合套餐。
5. 健康诊断：根据系统内置的健康诊断策略，周期性回溯异常事件，并通过邮件方式推送出来。
6. 预警自愈：是健康诊断功能的延伸，把健康诊断发现的问题通过自愈方案解决，完成异常事件的闭环。
7. 操作审计：感知故障自愈的每一次改动，确保运营安全，问题可回溯。

## 引领行业故障处理新潮流

故障自愈重新定义故障处理流程，在运维领域系较早提出故障自动化理念并落地为产品。



## 事件处理流程引擎，实现无人值守自愈

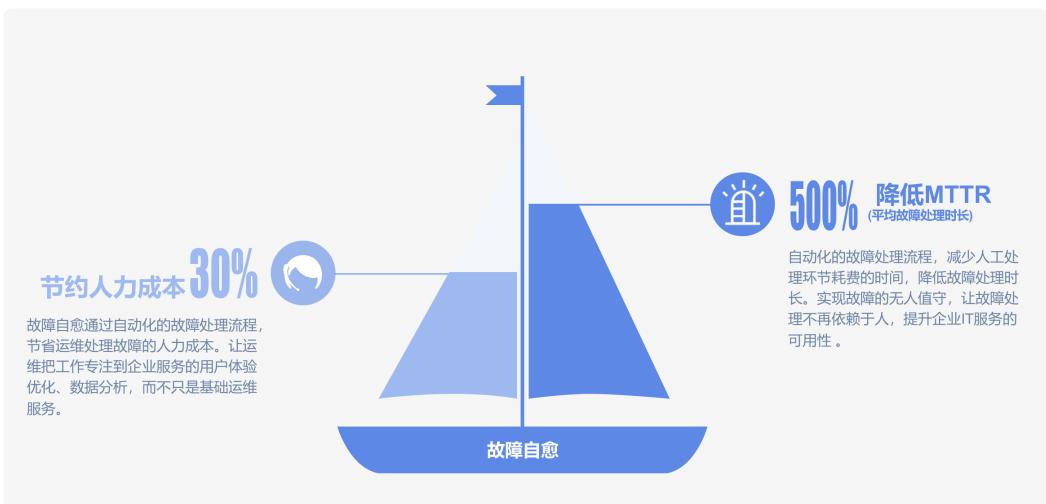
获取监控告警发现异常，预诊断分析，调用预定义的处理流程，实现故障无干预自动处理。



## 为企业节省人力及降低 MTTR

故障自愈通过自动化的故障处理流程，节省运维处理故障的人力成本。让运维把工作专注到企业服务的用户体验优化、数据分析，而不只是基础运维服务。

自动化的故障处理流程，减少人工处理环节耗费的时间，降低故障处理时长。实现故障的无人值守，让故障处理不再依赖于人，提升企业 IT 服务的可用性。



## 告警自动处理

将告警接入自愈套餐后，告警将匹配配置的处理套餐自动执行，无需人工干预。

下图为 Zabbix 的磁盘容量告警接入故障自愈。

The screenshot shows the 'Fault Recovery' module in the BlueKing platform. On the left is a sidebar with navigation links: 首页 (Home), 自愈详情 (Recovery Details), 接入自愈 (Integrate Recovery), 按入白名单 (White List), 客餐管理 (Catering Management), 管理告警源 (Manage Alert Sources), 高级配置 (Advanced Configuration), 统计报表 (Statistics Report), and 管理员 (Administrator). The main content area has three tabs: '自愈场景' (Recovery Scenario), '自愈处理' (Recovery Processing), and '其他信息' (Other Information). In the '自愈场景' tab, there are dropdowns for '告警类型' (Alert Type) set to '磁盘容量(vfs.fs.\*)' and '按内容筛选' (Content Filter) with the placeholder '使用正则表达式匹配告警, 不填为不过滤'. Below these are sections for '平台' (Platform), '集群' (Cluster), and '模块' (Module). The '自愈处理' tab contains settings for '自愈套餐' (Recovery Package) '接入层日志清理' (Access Layer Log Cleaning), '通知方式' (Notification Methods) for start, success, and failure, and '通知人员' (Notify Persons). The '其他信息' tab includes fields for '超时' (Timeout) set to 40, '自愈方案名称' (Recovery Plan Name) '接入层日志清理', and '是否启用' (Enabled) with a radio button for '是' (Yes). At the bottom are buttons for '保存自愈策略' (Save Recovery Strategy) and '本策略拷贝至' (Copy this strategy to).

在【自愈详情】菜单中，可以找到每一次故障自愈处理的记录。

The screenshot shows the 'Fault Recovery' module's 'Recovery Details' page. It displays a table of recovery events from July 1, 2019, to July 31, 2019. The columns include: 类型 (Type), 产生时间 (Generated Time), 自愈耗时 (Recovery Duration), 平台 (Platform), 集群 (Cluster), 模块 (Module), IP, 状态 (Status), and 自愈结果 (Recovery Result). The table lists various events such as '进程端口' (Port Process) and '磁盘容量' (Disk Capacity) alerts, along with their corresponding recovery actions like '执行job作业成功' (Job execution successful). A search bar at the top right allows users to search by IP.

点击其中一个自愈记录，可以查看到告警处理详情，下图为 Zabbix 的磁盘使用率告警的自愈详情。

This screenshot shows the 'Fault Recovery' module's alert handling details for a Zabbix disk usage alarm. It includes sections for '告警详情' (Alert Details), '处理过程' (Handling Process), '处理状态' (Handling Status), and '操作' (Operations). The alert details mention a host '10.0.4.128' with a disk space issue. The handling process shows steps from '07:40' to '08:00' involving log cleaning and job execution. The status is marked as '成功' (Successful) with the message '执行Job作业成功[337]'. The operations section includes a button to '重试整个流程' (Retry the entire process).

下图为 蓝鲸监控进程端口告警的自动化处理记录。

告警详情 业务 欢乐游戏(demo) 的主机 10.0.4.29 在 2019-07-31 15:22:00 +0800 发生 [进程端口](#) : 进程端口: 当前进程(MariaDB)不存在 [查看原始告警](#)

处理过程 15:24:00 开始处理套餐[启动MariaDB]  
15:24:19 #0 启动MariaDB | 成功: 执行Job作业成功[375]

处理状态 成功 执行Job作业成功[375]

操作 重试整个流程

#0 [44] 启动MariaDB

告警自动处理，如此简单！

## 告警收敛

针对满足收敛条件的告警，汇总为一个告警事件，或进行异常防御审批。

下图为集中收到 4 条 REST API 告警，被收敛成 1 条告警的示例。

类型	产生时间	自愈耗时	平台	集群	模块	IP	状态	自愈结果
进程端口	07-31 15:22:00 +0800	21秒	Android_Weixin	一区	存储层	10.0.4.29	<span style="color: green;">✓</span>	执行Job作业成功[375]
异常防御需审批	07-31 14:32:50 +0800							影响范围：一区
REST默认分类	07-31 14:31:39 +0800	19分钟	Android_Weixin	一区	存储层	10.0.4.29	<span style="color: green;">✓</span>	[系统]驳回了审批[脚印]
REST默认分类	07-31 14:31:39 +0800	23秒	Android_Weixin	一区	存储层	10.0.4.29	<span style="color: green;">✓</span>	执行Job作业成功[372]
REST默认分类	07-31 14:31:39 +0800	22秒	Android_Weixin	一区	存储层	10.0.4.29	<span style="color: green;">✓</span>	执行Job作业成功[373]
REST默认分类	07-31 14:31:39 +0800	19分钟	Android_Weixin	一区	存储层	10.0.4.29	<span style="color: green;">✓</span>	[系统]驳回了审批[脚印]

原因是在【告警收敛】菜单中针对 REST API 告警设置了 **相同业务 5 分钟内 收到 2 条** 以上的收敛策略。

针对告警类型	在一定条件下	触发频次
Ping检测(icmping*) Ping	主机: 相同	5分钟内 1条以上
Ping检测(icmping*) Ping	主机: 相同 自愈套餐: 相同	5分钟内 1条以上
Ping检测(icmping*)	业务: 相同	5分钟内 2条以上
REST默认分类	业务: 相同	5分钟内 2条以上

新建收敛规则

告警类型: REST默认分类

条件: 相同业务

触发频次: 5 分 2 条 以上

收敛方式: 异常防御需审批

备注: 疑似网络波动, 请审核.

保存

当然，故障自愈内置常见的收敛规则，例如包含几种监控产品的 PING 告警收敛策略，以防因网络波动造成假告警导致服务器真正被重启。

告警收敛，在保障业务安全的前提下做告警的无人值守。

“

收敛审批是通过企业微信实现，请参考 [微信审批接入流程](#)。

”

## 健康诊断

依托于腾讯故障处理的经验，集成常见故障隐患的专家配置库，回溯过往发生的告警单据来提前发现问题。

每天早上 8 点回溯自愈处理过的告警，处理分析出的潜在风险。下图为 1 天内产生 2 次磁盘容量告警，故障自愈将提示： - 请检查当前的磁盘清理策略确实是否需要调整 - 确认该模块当前机型的硬盘空间是否合理



The screenshot shows the 'Health Diagnosis' section of the Fault Recovery interface. It displays a single告警记录 (alarm record) for a disk capacity issue. The record details: 告警类型 (Alarm Type: 磁盘容量), 发现时间 (Discovery Time: 2019-07-30), 发生频次 (Occurrence Frequency: 一天内2次), 可疑IP (Suspicious IP: 10.0.4.128), 平台 (Platform: --), 集群 (Cluster: 一区), 模块 (Module: 接入层), 处理类型 (Handling Type: 建议), 处理方案 (Handling Plan: 1. 测量直连...; 2. 确认读写...), 状态 (Status: 警告), and 事件详情 (Event Details: 查看详情).

除了针对风险的文本建议，还可以在【接入预警】页面配置自动处理的动作。

下图的预警自愈策略为：当 1 个月内同 1 台主机产生了 5 次 Ping 告警，则故障自愈认为该主机存在异常(可能主板、内存故障)，直接将主机在 CMDB 中移到“故障机”模块。当然，你也可以设置其他处理动作。



The screenshot shows the 'Access Warning' section of the Fault Recovery interface. It displays a configuration for a self-healing rule. The rule details: 告警类型 (Alarm Type: Ping检测), 告警数量 (Alarm Count: 0), 察考时长(天) (Observation Duration (Days): 30), 察考阈值(次) (Observation Threshold (Times): 5), (所有) (All), 模块 (Module: 空闲机), 处理方案 (Handling Plan: [快进] CCP移到...), 处理类型 (Handling Type: 告警), 是否启用 (Enabled: 是), 方案来源 (Plan Source: 系统内置), and 操作 (Operation: A green toggle switch icon).

健康诊断，在告警中提前发现业务隐患。

## 附录 1： Django 后台调整健康诊断策略

用 PaaS 管理员访问 `/o/bk_fta_solutions/admin/`，进入故障自愈的 Django 后台。

动作	执行	18 个 0 个被选
<a href="#">Advice defs</a>	<a href="#">定义 健康诊断</a>	<a href="#"></a> <a href="#"></a>
<a href="#">Advice fta defs</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Advices</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm defs</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm instance archives</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm instance backups</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm instance logs</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm instances</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Alarm types</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Approve callbacks</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Biz confs</a>		<a href="#"></a> <a href="#"></a>
<a href="#">Contexts</a>		<a href="#"></a> <a href="#"></a>

点击 `Advice_defs` 进入健康诊断策略配置页面，可以看到健康诊断建议，可以依据实际情况修改。

ID	英文名称	是否启用	考察对象	考察数据	详细数据类型	考察时长(天)	考察间隔	描述	建议类型	具体建议
18	host-down-monthly	是	主机	告警	NAGIOS-ping	30	5	30天内 同一主机 产生5+条严重告警，应产生硬件待优化事件进行后续跟踪。	硬件待优化化	确认主机是否存在硬件故障，需要做下线或者替换操作
17	host-down-monthly	是	主机	告警	ZABBIX-icmpping*	30	5	30天内 同一主机 产生5+条严重告警，应产生硬件待优化事件进行后续跟踪。	硬件待优化化	确认主机是否存在硬件故障，需要做下线或者替换操作
16	host-down-daily	是	主机	告警	NAGIOS-ping	1	2	1天内 同一主机 产生2+条严重告警，应产生硬件待优化事件进行后续跟踪。	硬件待优化化	确认主机是否存在硬件故障，需要做下线或者替换操作
15	host-down-daily	是	主机	告警	ZABBIX-icmpping*	1	2	1天内 同一主机 产生2+条严重告警，应产生硬件待优化事件进行后续跟踪。	硬件待优化化	确认主机是否存在硬件故障，需要做下线或者替换操作
14	disk-readonly-daily	是	主机	告警	BASE_ALARM_3	1	2	1天内 同一主机 产生2+条硬盘只读告警，应产生硬件待优化事件进行人工确认。	硬件待优化化	简单请现场确认主机硬件是否故障
13	disk-readonly-monthly	是	主机	告警	BASE_ALARM_3	30	3	30天内 同一主机 产生3+条硬盘只读告警，应产生硬件待优化事件进行人工确认。	硬件待优化化	简单请现场确认主机硬件是否故障
12	disk-full-daily	是	主机	告警	BASE_ALARM_6	1	2	1天内 同一主机 产生2+条磁盘满告警，应产生硬件待优化事件进行后续跟踪。	运维待优化化	1. 检查当前的磁盘清理策略是否需要调整 2. 确认该模块当前机型的硬盘空间是否合理
11	disk-full-daily	是	主机	告警	ZABBIX-vfs.fs.*	1	2	1天内 同一主机 产生2+条磁盘满告警，应产生硬件待优化事件进行后续跟踪。	运维待优化化	1. 检查当前的磁盘清理策略是否需要调整 2. 确认该模块当前机型的硬盘空间是否合理

集成主流监控产品

告警源集成蓝鲸监控、4 款主流开源监控产品 Zabbix、Open-Falcon、Nagios、Icinga，及 AWS、邮件的告警接入，更能通过 REST API 拉取、推送告警。

管理告警源 对接企业正在使用的监控产品，迈向无人值守的第一步

已启用监控产品

- Icinga: 告警: 0, 最后接收告警: --
- ZABBIX: 告警: 0, 最后接收告警: --
- REST API: 告警: 49, 最后接收告警: 18/08/13 20:06:59 (暂选)
- 蓝鲸监控: 告警: 15505, 最后接收告警: 18/08/22 18:02:52

未启用监控产品

- falcon: 集成当前正在使用的 Open-Falcon, 启用
- Nagios: 集成当前正在使用的 Nagios, 启用
- REST API: 其他监控产品请选择, 启用
- amazon webservices: 从AWS获取告警, 启用
- EMAIL: 从邮件中获取告警, 启用

以下为部分监控产品的故障自愈执行记录：

#### • Zabbix

类型	产生时间	自愈耗时	平台	集群	模块	IP	状态	自愈结果
磁盘容量(/vfat,fs)	07-29 20:07:34 +0800	23秒	Android_Weixin	一区	接入层	10.0.4.128	<span>正常</span>	执行job作业成功[337]
磁盘容量(/vfat,fs)	07-29 19:52:34 +0800	29秒	Android_Weixin	一区	接入层	10.0.4.128	<span>正常</span>	执行job作业成功[336]

#### • 蓝鲸监控

类型	产生时间	自愈耗时	平台	集群	模块	IP	状态	自愈结果
进控端口	07-31 15:22:00 +0800	21秒	Android_Weixin	一区	存储层	10.0.4.29	<span>正常</span>	执行job作业成功[375]
异常防御策略批	07-31 14:32:50 +0800	2秒	Android_Weixin	一区	存储层	10.0.4.29	<span>正常</span>	新策略范围：一区
RESTAPI分类	07-31 13:11:28 +0800	22秒	Android_Weixin	一区	存储层	10.0.4.29	<span>正常</span>	执行job作业成功[371]

#### • REST API

- Open-Falcon

告警详情 业务 蓝鲸 的主机 1\* 在 2018-05-07 20:13:00 +0800 发生 磁盘容量(df.\*) : PROBLEM(df.statistics.used.percent)  
 处理过程 ✅ 20:13:36 开始处理套餐[清理nginx日志]  
 ✅ 20:13:57 #0 清理nginx日志 | 成功: 执行Job作业成功[82]  
 处理状态 成功 执行Job作业成功[82]

操作 重试整个流程

#0 [42] 清理nginx  
日志

无论是主流的监控产品，还是自研监控产品，故障自愈总能与其对接。

## 组合套餐

依托于故障树分析(FTA : Fault Tree Analysis)理念，将单个原子套餐组装为组合套餐，根据父节点的执行结果(成功还是失败)来确定子节点的执行分支，以解决复杂场景的故障处理和分析。

下图为服务器故障替换的场景案例，将人做故障替换的经验沉淀在故障自愈的组合套餐中，在保障安全的前提下，高效地完成故障替换。



组合套餐，实现复杂的运维故障处理场景。

# 蓝鲸监控告警自动处理

## 情景

故障处理是运维的职能之一，人工登录服务器处理告警，存在 2 个问题：**故障处理效率低** 和 **操作疏忽时可能影响生产环境**，例如删除文件输入绝对路径时，在根目录和日志目录间误敲空格，导致根目录删除。

接下来通过“蓝鲸监控的进程告警接入故障自愈”这个案例，来了解故障自愈是如何解决这 2 个痛点。

## 前提条件

- 蓝鲸配置平台纳管了主机
- 蓝鲸配置平台纳管了进程
- 作业平台新建一个作业

## 术语解释

- **自愈套餐**：告警的处理动作，如拉起进程的作业；
- **自愈方案**：关联 告警 和 处理动作的一个组合；

## 操作步骤

1. 启用蓝鲸监控告警源
2. 接入自愈方案
3. 自愈测试

### 启用蓝鲸监控告警源

在菜单 **[接入自愈] -> [管理告警源]** 中，启用 **蓝鲸监控**。

管理告警源 对接企业正在使用的监控产品，迈向无人值守的第一步

已启用监控产品

已处理告警: 10 未处理告警: 0  
最后接收告警: 19/07/25 16:07:45

ZABBIX  
已处理告警: 2 未处理告警: 1  
最后接收告警: 19/07/29 20:07:38

未启用监控产品

REST API 其他监控产品请选择  
集成当前正在使用的蓝鲸监控  
启用

ZABBIX  
集成当前正在使用的 Zabbix  
启用

falcon  
集成当前正在使用的 Open-Falcon  
启用

Nagios  
集成当前正在使用的 Nagios  
启用

REST API 其他监控产品请选择  
从企业告警API中获取告警  
启用

Icinga  
集成当前正在使用的 Icinga 2  
启用

AWS CloudWatch Metrics  
从AWS获取告警  
启用

EMAIL  
从邮件中获取告警  
启用

Copyright © 2012-2019 Tencent BlueKing.  
All Rights Reserved.

## 接入自愈方案

在菜单 [接入自愈] 中, 点击 **接入自愈**, 告警类型选择 **[主机监控] 进程端口**, 模块选择 **存储层**, 因为不同类型服务器拉起进程的作业不一样。

点击新建 **自愈套餐** 的按钮, 准备新建拉起进程的作业。

自愈场景

告警类型: [主机监控] 进程端口 1. 将什么告警, 做自动化处理  
平台: 默认全选 集群: 默认全选  
模块: 存储层 2. 将什么类型的服务器, 做自动化处理

自愈处理

自愈套餐: 请选择处理告警的自愈套餐  
通知方式: 开始时: 微信, 邮件, 短信, 电话  
成功时: 微信, 邮件, 短信, 电话  
失败时: 微信, 邮件, 短信, 电话  
通知人员: 业务运维: blueking.admin  
更多通知人  
注: 需要提前在蓝鲸控制台的【用户管理】填写联系方式

3. 告警自动化处理的动作

在套餐中, 套餐类型选择 **作业平台**, 新建 **启动MariaDB的作业**。

点击 **新建作业的按钮** 后，跳转至作业平台，在菜单 **[作业执行] -> [新建作业]** 中，新建如下作业：

```
# Check
ps -ef | grep -i mysqld
netstat -ntlp | grep -i 3306

# Start MariaDB
systemctl start mariadb || job_fail "start mariadb fail"

# Check
ps -ef | grep -i mysqld
netstat -ntlp | grep -i 3306
netstat -ntlp | grep -i 3306 || job_fail "mariadb not listen 3306"

job_success "start mariadb succ"
```

保存 **启动MariaDB** 的套餐后，自动回到接入自愈的页面，保存自愈方案即可。

The screenshot shows the '故障自愈' (Fault Recovery) module in the BlueKing platform. In the '自愈场景' (Recovery Scenario) section, a new scenario is being configured for a '进程端口' (Process Port) alert. The '告警类型' (Alert Type) is set to '主机监控 进程端口'. Under '自愈处理' (Recovery Processing), a recovery step named '启动MariaDB' (Start MariaDB) is selected. The '通知方式' (Notification Method) includes WeChat, Email, and SMS. The '通知人员' (Notify Person) field contains '业务运维: blueking:admin'. A note at the bottom states: '注: 需要在蓝鲸控制台的【用户管理】填写联系方式' (Note: You need to fill in contact information in the User Management section of the BlueKing Control Panel).

回到接入自愈列表，在列表中可以找到刚刚创建的自愈方案。

The screenshot shows the '接入自愈' (Access Recovery) list. It displays four recovery schemes: 1. REST默认分类 (REST Default Category) -告警数量: 10, 平台: (所有), 集群: (所有), 模块: (所有), 自愈套餐: 默认测试套餐, 告警源: REST API监控, 方案来源: REST API测试, 启用: 是. 2. 磁盘容量(\%fs.) -告警数量: 2, 平台: (所有), 集群: (所有), 模块: (所有), 自愈套餐: 接入层日志清理, 告警源: Zabbix监控, 方案来源: 接入层日志清理, 启用: 是. 3. 进程端口 -告警数量: 0, 平台: (所有), 集群: (所有), 模块: 存储层, 自愈套餐: 自动MariaDB8, 告警源: 蓝鲸监控, 方案来源: 蓝鲸监控, 启用: 是. 4. 5分钟平均负载 -告警数量: 0, 平台: (所有), 集群: (所有), 模块: (所有), 自愈套餐: [快速]发送C..., 告警源: 蓝鲸监控, 方案来源: 5分钟平均负载... 系统推荐, 启用: 否. 5. CPU单核使用率 -告警数量: 0, 平台: (所有), 集群: (所有), 模块: (所有), 自愈套餐: [快速]发送C..., 告警源: 蓝鲸监控, 方案来源: CPU单核使用率... 系统推荐, 启用: 否.

## 自愈测试

接下来将停止 MariaDB 进程，来验证是否可以自动启动进程，以恢复 DB 服务。

```
# ps -ef | grep -i mysqld
mysql      926      1  0 10:47 ?        00:00:00 /bin/sh /usr/bin/mysqld_safe --basedir= /usr
mysql     1159    926  0 10:47 ?        00:00:09 /usr/libexec/mysqld --basedir=/usr --datadir=/var/lib/mysql --plugin-dir=/usr/lib64/mysql/plugin --log-error=/var/log/mariadb/mariadb.log --pid-file=/var/run/mariadb/mariadb.pid --socket=/var/lib/mysql/mysql.sock
root     16763   7429  0 19:40 pts/1    00:00:00 grep --color=auto -i mysqld
# systemctl stop mariadb
```

稍等片刻，收到蓝鲸监控的进程端口告警。

事件中心(列表)

当前业务: [4]欢乐游戏(demo)

列表模式 日历模式

起止时间: 2019-07-31 00:00 至 2019-07-31 23:59 监控类型: 全部 告警级别: 全部

告警内容: 请输入告警内容 告警状态: 全部 告警ID: 请输入告警ID

查询 重置

告警ID	告警时间	恢复时间	监控名称	监控类型	级别	告警内容	状态
B41635	2019-07-31 15:22:00	-	进程端口	进程端口	●	维度信息: 开发商ID(0)-协议(tcp)-平台(Android_Weixin)-业务(欢乐游戏(demo))-云区域ID(0)-进程显示名称(MariaDB)-云没有监听的端口, 多个用, 分隔(3306)-集群(-区)-模块(存储层)-biz_id(4)-采集器IP地址(10.0.4.29).4.29)-绑定的IP地址(10.0.4.29) 触发条件: 主机[10.0.4.29]-当前进程(MariaDB)不存在	未恢复

在故障自愈的自愈详情中，找到了该条告警的自愈记录，耗时 21 秒。

故障自愈

欢乐游戏(demo)

1 次收故障件 6 次自愈

类型	产生时间	自愈耗时	平台	集群	模块	IP	状态	自愈结果
进程端口	07-31 15:22:00 +0800	21秒	Android_Weixin	一区	存储层	10.0.4.29	<span>成功</span>	执行job作业成功[375]
异常防御需求批	07-31 14:32:50 +0800	-	-	-	-	-	-	疑似网络波动, 请审核。影响范围: 一区
REST默认分类	07-31 13:11:28 +0800	22秒	Android_Weixin	一区	存储层	10.0.4.29	<span>成功</span>	执行job作业成功[371]

跳转到作业平台的执行历史，可以看到 MariaDB 已经启动成功。

蓝鲸作业平台

当前业务: 欢乐游戏(demo)

作业执行 >

节点名称	节点类型	执行脚本	服务器账户	总时间(s)
start_mariadb.sh	执行脚本	/root/start_mariadb.sh	root	3.731

节点状态: 执行成功 开始时间: 2019-07-31 15:24:01 +0800 结束时间: 2019-07-31 15:24:05 +0800

执行成功[start mariadb succ(1)]

输入搜索内容

IP过滤: 完整日志

云区域名称	IP	返回码	耗时(s)
default area	10.0.4.29	0	2.296

[2019-07-31 15:24:01][PID:32023] job\_start  
mysql 32071 1 0 15:24:01 00:00:00 /bin/sh /user/bin/mysqld\_safe --basedir=/usr  
mysql 32245 32071 1 15:24:01 00:00:00 /usr/libexec/mysqld --basedir=/usr --datadir=/var/lib/mysql --plugin-dir=/usr/lib/b64/mysql/plugin --log-error=/var/log/mariadb/mariadb.log --pid-file=/var/run/mariadb/mariadb.pid --socket=/var/lib/mysql/mysqld.sock  
tcp 0 0 10.0.4.29:3306 0.0.0.0\* LISTEN 32245/mysqld  
tcp 0 0 10.0.4.29:3306 0.0.0.0\* LISTEN 32245/mysqld  
[2019-07-31 15:24:03][PID:32023] job\_success:[start mariadb succ]

回到监控的事件中心，可以看到告警已经恢复。

告警自动处理，如此简单。

以上为主机监控的告警自动化处理，其他类型告警请参考对应文档：[组件监控的告警自动化处理](#)、[自定义采集的告警自动化处理](#)。

故障自动处理是把双刃剑，需要考虑因为网络波动等场景导致的假告警，这时可以用到故障自愈的[异常防御需审批](#)功能。具体请参照[故障自愈的收敛防护](#)。

故障自愈，在安全的前提下完成告警的自动化处理。

## Zabbix 告警自动处理

### 情景

故障处理是运维的职能之一，Zabbix自带ActionScript虽然可以实现告警自动处理，但存在2个问题：[无法集中管理自动处理的脚本](#)、[没有收敛防护，安全性无法保障](#)。

接下来我们通过将“[Zabbix 中磁盘使用率（vfs.fs.\\*）告警接入故障自愈](#)”这个案例，来了解故障自愈是如何解决这2个痛点。

### 前提条件

- 蓝鲸配置平台纳管了Zabbix监控的对象
- 拥有Zabbix管理员账号，用于注册Zabbix Action

### 术语解释

- **自愈套餐**：告警的处理动作，等同于Zabbix的Action；
- **自愈方案**：关联告警和处理动作的一个组合；

## 操作步骤

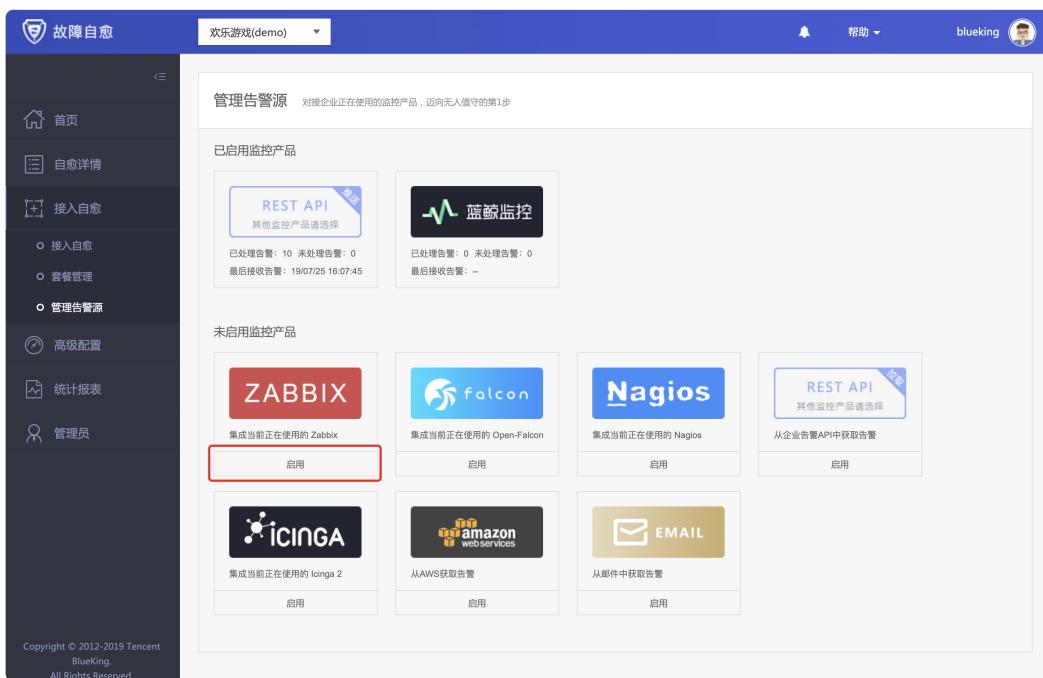
1. 接入 Zabbix 告警源
2. 接入自愈方案
3. 自愈测试

## 视频教程

{% video %}media/zabbix\_fta.mp4{% endvideo %}

### 接入 Zabbix 告警源

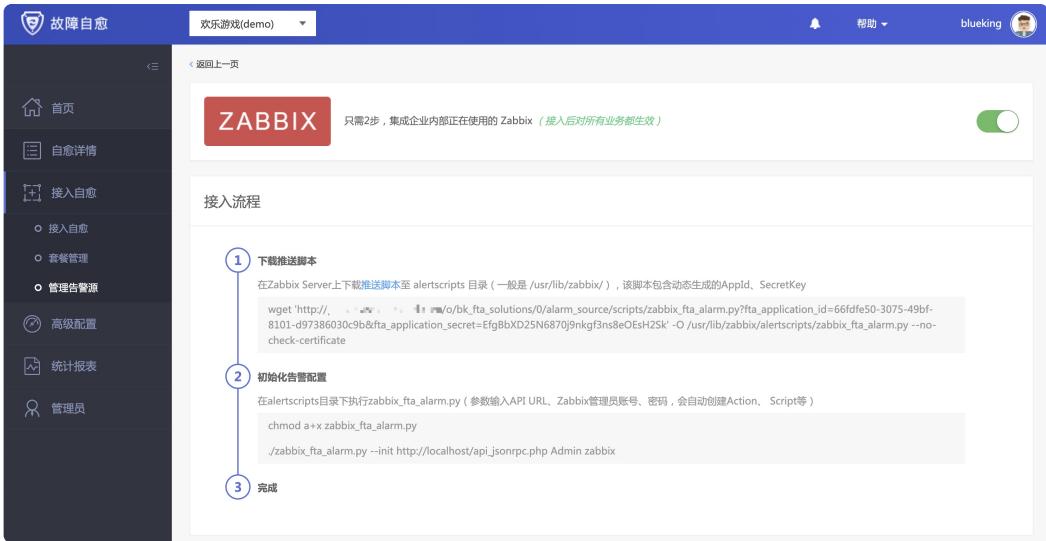
在菜单 [接入自愈] -> [管理告警源] 中，点击 启用 Zabbix。



The screenshot shows the 'Management of Alert Sources' section. It lists various monitoring products with their current status and configuration options:

- 已启用监控产品 (Enabled Monitoring Products):**
  - REST API (其他监控产品请选择) (Other monitoring products please select)
  - 蓝鲸监控 (BlueWhale Monitoring)
- 未启用监控产品 (Disabled Monitoring Products):**
  - ZABBIX (集成当前正在使用的 Zabbix)  
启用了 (Enabled) **启用 (Enable)**
  - falcon (集成当前正在使用的 Open-Falcon)  
启用了 (Enabled)
  - Nagios (集成当前正在使用的 Nagios)  
启用了 (Enabled)
  - REST API (其他监控产品请选择) (Other monitoring products please select)  
从企业告警API中获取告警 (Get alerts from enterprise alarm API)  
启用了 (Enabled)
  - Icinga 2 (集成当前正在使用的 Icinga 2)  
启用了 (Enabled)
  - amazon web services (从AWS获取告警) (Get alerts from AWS)  
启用了 (Enabled)
  - EMAIL (从邮件中获取告警) (Get alerts from email)  
启用了 (Enabled)

[跳转到接入流程页面](#)



Zabbix 接入故障自愈的逻辑是，告警产生时，执行 `Action`，将告警推送至故障自愈接收告警的回调接口。

接下来，我们下载并初始化该 `Action`。

### 下载初始化脚本

参照上图，进入 Zabbix Action 的目录 `/usr/lib/zabbix/alertscripts`，下载初始化脚本

`zabbix_fta_alarm.py`。

```
[root@37ae504b6646 alertscripts]# wget 'http://${PaaS_Host}/o/bk_fta_solutions/0/alarm_source/scripts/zabbix_fta_alarm.py?fta_application_id=66fdfe50-3075-49bf-8101-d97386030c9b&fta_application_secret=EfgBbXD25N6870j9nkgf3ns8eOEsH2Sk' -O /usr/lib/zabbix/alertscripts/zabbix_fta_alarm.py --no-check-certificate
```

注：请直接复制故障自愈页面的命令，其中包含故障自愈的页面 URL 以及账号信息。

### 初始化 ZABBIX 告警配置

执行初始化 Zabbix 告警配置脚本 `zabbix_fta_alarm.py`，参数依次为 `--init`、`Zabbix API URL`、`Zabbix账号`、`Zabbix密码`

```
[root@37ae504b6646 alertscripts]# chmod +x zabbix_fta_alarm.py
[root@37ae504b6646 alertscripts]# ./zabbix_fta_alarm.py --init http://${Zabbix_Host}/api_jsonrpc.php Admin zabbix
[2019-07-30 10:51:45,374] INFO fta: get auth token: 136b14f3b8fe226bc02bc5eb4dfd7ac6
[2019-07-30 10:51:45,455] INFO fta: action_get success: {u'jsonrpc': u'2.0', u'result': [{u'actionid': u'8'}], u'id': 1}
[2019-07-30 10:51:45,572] INFO fta: action_delete success: {u'jsonrpc': u'2.0', u'result': {u'actionids': [u'8']}, u'id': 1}
[2019-07-30 10:51:45,640] INFO fta: user_get success: {u'jsonrpc': u'2.0', u'result': [
```

```

{u'userid': u'6'}], u'id': 1}
[2019-07-30 10:51:45,809] INFO fta: user_delete success: {u'jsonrpc': u'2.0', u'result':
: {u'userids': [u'6']}, u'id': 1}
[2019-07-30 10:51:45,902] INFO fta: mediatype_get success: {u'jsonrpc': u'2.0', u'result':
: [{u'mediatypeid': u'7'}], u'id': 1}
[2019-07-30 10:51:45,984] INFO fta: mediatype_delete success: {u'jsonrpc': u'2.0', u'result':
: {u'mediatypeids': [u'7']}, u'id': 1}
[2019-07-30 10:51:46,077] INFO fta: mediatype_create success: {u'jsonrpc': u'2.0', u'result':
: {u'mediatypeids': [u'8']}, u'id': 1}
[2019-07-30 10:51:46,174] INFO fta: user_create success: {u'jsonrpc': u'2.0', u'result':
: {u'userids': [u'7']}, u'id': 1}
[2019-07-30 10:51:46,274] INFO fta: action_create success: {u'jsonrpc': u'2.0', u'result':
: {u'actionids': [9]}, u'id': 1}

```

该脚本会创建一个名为 **FTA\_Event\_Handler** 的 Media Type, 名为 **FTA\_Act** 的 Action, 名为 **FTA\_Mgr** 的用户。

Name	Type	Status	Used in actions	Details	Action
Email	Email	Enabled		SMTP server: "smtp.163.com", SMTP helo: "smtp.163.com", SMTP email: "bluekingtest@163.com"	<a href="#">Test</a>
FTA_Event_Handler	Script	Enabled		Script name: "zabbix_fta_alarm.py"	<a href="#">Test</a>
Jabber	Jabber	Enabled		Jabber identifier: "jabber@example.com"	<a href="#">Test</a>
SMS	SMS	Enabled		GSM modem: "idevitySD"	<a href="#">Test</a>

## 接入自愈方案

Zabbix 告警源 接入成功后, 接下来关联告警 和 告警的处理动作。

将 Zabbix 中磁盘容量告警关联一个磁盘清理的处理动作。

选择菜单 [接入自愈] -> [接入自愈], 点击接入自愈

告警类型	告警数量	平台	集群	模块	自愈套餐	告警源	自愈方案	方案来源	启用	操作
REST默认分类	10	(所有)	(所有)	(所有)	<a href="#">默认测试套餐</a>	REST API监控	REST API测试	人工配置	<input checked="" type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
磁盘容量(vfs.fs.*)	2	(所有)	(所有)	接入层	<a href="#">接入层日志...</a>	Zabbix监控	接入层日志...	人工配置	<input checked="" type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
5分钟平均负载	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	5分钟平均负载...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
CPU单核使用率	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	CPU单核使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
CPU总使用率	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	CPU总使用...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
交换分区使用量	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	交换分区使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
可用物理内存	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	可用物理内...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
应用内存使用率	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	应用内存使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
应用内存使用量	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	应用内存使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
物理内存使用率	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	物理内存使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>
物理内存使用量	0	(所有)	(所有)	(所有)	<a href="#">[快速] 发...</a>	蓝鲸监控	物理内存使...	系统推荐	<input type="checkbox"/>	<a href="#">编辑</a> <a href="#">删除</a>

进入接入自愈页面, 告警类型选择 **磁盘容量(vfs.fs.\*)**, 自愈套餐点击 **+号** 新建 磁盘清理自愈

套餐。

The screenshot shows the 'Fault Recovery' interface. In the 'Recovery Scenario' section, the 'Alert Type' is set to '磁盘容量(vfs.fs.\*.)'. The 'Recovery Action' section shows three rows for '开始时': WeChat (checked), Email (checked), SMS (unchecked), Phone (unchecked). For '成功时' and '失败时', the configurations are identical. Under '通知人员', '业务运维: blueking:admin' is checked. A note at the bottom states: '注: 需要在蓝鲸控制台的【用户管理】填写联系方式'.

跳转至磁盘清理的自愈套餐页面。

The screenshot shows the 'Recovery Plan' configuration page. The 'Plan Name' is '接入层日志清理'. The 'Disk Cleaning Input Hint' section contains a detailed note about allowed characters and paths.

保存自愈套餐后，自动回到接入自愈页面。添加完自愈方案后，在接入自愈列表页可以找到刚才创建的自愈方案。

The screenshot shows the 'Access Recovery' configuration page. It lists several alert types and their recovery plans:

告警类型	告警数量	平台	集群	模块	自愈套餐	告警源	自愈方案	方案来源	启用	操作	
REST默认分类	10	(所有)	(所有)	(所有)	默认测试套餐	REST API监控	REST API测试	人工配置	是		
磁盘容量(vfs.fs.*.)	0	(所有)	(所有)	接入层	接入层磁盘...	Zabbix监控	接入层磁盘...	人工配置	是		
5分钟平均负载	0	(所有)	(所有)	(所有)	【快排】发...	蓝鲸监控	5分钟平均...	系统推荐	否		
CPU单核使用率	0	(所有)	(所有)	(所有)	【快排】发...	蓝鲸监控	CPU单核使...	系统推荐	否		
CPU总使用率	0	(所有)	(所有)	(所有)	【快排】发...	蓝鲸监控	CPU总使用...	系统推荐	否		

## 自愈测试

生成一个大文件，使磁盘剩余空间低于 20% ( Zabbix 中默认设定的 Trigger 是<20% )

```
[root@access_layer_breaking gse]# pwd
/data/logs/gse
[root@access_layer_breaking gse]# touch -d "10 days ago" agent-20190726-00001.log
[root@access_layer_breaking gse]# ll
总用量 3906980
-rw-r--r-- 1 root root 4000000000 7月 20 11:38 agent-20190726-00001.log
-rw-r--r-- 1 root root 196795 7月 26 23:59 agent-20190726-00002.log
-rw-r--r-- 1 root root 194952 7月 27 23:59 agent-20190727-00003.log
-rw-r--r-- 1 root root 198532 7月 28 23:59 agent-20190728-00004.log
-rw-r--r-- 1 root root 142948 7月 29 17:33 agent-20190729-00005.log
[root@access_layer_breaking gse]# dd if=/dev/zero of=4gb.log bs=1GB count=4
记录了4+0 的读入
记录了4+0 的写出
4000000000字节(4.0 GB)已复制, 34.0365 秒, 118 MB/秒
```

稍等片刻，收到 Zabbix 邮件告警，以及故障自愈的处理通知。



在故障自愈的 [自愈详情] 菜单中也可以找到自愈记录。

A screenshot of the Blueking fault recovery detail page. The title bar says '故障自愈' and '欢乐游戏(demo)'. The left sidebar has options like '首页', '自愈详情' (selected), '接入自愈', '高级配置', '统计报表', and '管理员'. The main area shows a table of recovery logs for July 29, 2019, at 20:07:34. There are two entries, both for '磁盘容量(vfs.fs.\*)' on IP 10.0.4.128, with status '执行job作业成功[337]'.

可以查看详情执行记录。

A screenshot of the fault recovery execution log details. It shows the following information:

- 告警详情: 业务 欢乐游戏(demo) 的主机 10.0.4.128 在 2019-07-29 20:07:34 +0800 发生 磁盘容量(vfs.fs.\*) : Free disk space is less than 20% on volume /()
- 处理过程: 20:07:40 开始处理套餐[接入层日志清理]  
20:08:00 #0 接入层日志清理 | 成功: 执行Job作业成功[337]
- 处理状态: 成功 执行Job作业成功[337]
- 操作: 重试整个流程

At the bottom, there is a green button labeled '#0 [43] 接入层日志清理'.

从告警产生到处理结束，耗时 30 秒。

**告警自动处理，如此简单，不要再手动登录到服务器上处理告警了。**

## 扩展阅读

### 告警收敛确保安全的告警自动处理

选择菜单 **[高级配置] -> [告警收敛]**，新增高危告警的收敛规则，比如 Ping 告警。

一般网络波动时，可能触发假的 Ping 告警，这时不能直接重启服务器，可以通过告警收敛，让运维审批。

具体执行记录，请参照 [故障自愈的收敛防护](#)

## 健康度日报

如果服务器频繁触发自愈，那么需要去思考本质问题是什么，是磁盘使用率的告警策略不合适，还是主板故障，亦或是程序运行异常。

这时可以使用 预警功能。

我们在做该教程的时候，在同一天触发了 2 条磁盘使用率的自愈记录，所以产生了一条健康诊断记录。

The screenshot shows the BlueKing Fault Recovery interface. On the left is a sidebar with navigation links: 首页 (Home), 自愈详情 (Recovery Details), 接入自愈 (Access Recovery), 高级配置 (Advanced Configuration) (selected), 告警收敛 (Alert Convergence), 告警详情 (Alert Details), 接入预警 (Access Alert), 自愈小助手 (Recovery Assistant), 统计报表 (Statistical Reports), and 管理员 (Administrator). The main content area is titled '健康诊断' (Health Diagnosis) and displays a table of results for July 24, 2019, to July 30, 2019. The table has columns: 告警类型 (Alert Type), 发现时间 (Discovery Time), 发生频次 (Frequency), 可疑IP (Suspect IP), 平台 (Platform), 集群 (Cluster), 模块 (Module), 处理类型 (Handling Type), 处理方案 (Handling Plan), 状态 (Status), and 事件详情 (Event Details). One row is shown: 磁盘容量(v... 07-30 一天内2次 10.0.4.128 ... 一区 接入层 建议 1. 检查磁盘容量... 2. 确认该模块... 看看详情. The status column shows '1次未处理' (1 unhandled) and '0次处理失败' (0 failed handling).

在邮件中也可以找到。

The email subject is '[自愈通知] 健康度日报 欢乐游戏(demo) 07月30天 1个新风险点 ☆'. The header includes: 发件人: breaking <bluekingtest@163.com>, 时 间: 2019年7月30日(星期二) 上午8:00, 收件人: Pythoning <pythoning@qq.com>; Pythoning <pythoning@qq.com>. The body starts with a logo and the text '故障自愈 · 你的无人值守伙伴!'. It says: '早上好，亲爱的欢乐游戏(demo)运维帝：恭喜你，当你收到这封邮件，说明你的业务享受到了故障自愈的深度服务，以下是自愈今天最新的分析报告。' Below this is a section about new risks: '以下是今天自愈新发现的可疑故障点，建议在故障前尽早消除，降低业务损失的风险。猛击 [欢乐游戏\(demo\) 健康度完整报告](#)，查看所有的可疑故障点，全面整顿！' A red warning message follows: 'host 10.0.4.128 一天内 2次 磁盘容量(vfs.fs.\*)'. Below it is a detailed analysis: '集群: 一区 模块: 接入层 建议: 1、请检查当前的磁盘清理策略是否需要调整 2、确认该模块当前机型的硬盘空间是否合理'. At the bottom is the copyright notice: 'Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved'.

可在菜单 [高级配置] -> [接入预警] 中设置。

## 集成 REST API 推送

### 情景

故障处理是运维的职能之一，人工登录服务器处理告警，存在 2 个问题：**故障处理效率低** 和 **操作疏忽时可能影响生产环境**，例如删除文件输入绝对路径时，在根目录和日志目录间误敲空格，导致根目录删除。

前面介绍了**蓝鲸监控**、**Zabbix** 告警的自动处理，接下来通过“**REST API 告警接入故障自愈**”这个案例，来了解故障自愈如何集成第 3 方监控系统。

## 前提条件

- 蓝鲸配置平台纳管了主机
- 作业平台新建一个作业

## 术语解释

- **自愈套餐**：告警的处理动作，比如清理日志的作业
- **自愈方案**：关联 告警 和 处理动作的一个组合

## 操作步骤

1. 启用 REST API(推送)告警源
2. 接入自愈方案
3. 自愈测试

### 启用 REST API(推送)告警源

第 3 方监控系统调用 REST API(推送)接口，故障自愈收到告警后立即做自动化处理。



第3方监控系统 告警自动处理 流程图

首先，启用告警源。

在菜单 [**接入自愈**] -> [**管理告警源**] 中，启用 **REST API(推送)**。

已启用监控产品

ZABBIX  
已处理告警: 2 未处理告警: 1  
最后接收告警: 19/07/29 20:07:38

蓝鲸监控

未启用监控产品

falcon  
集成当前正在使用的 Open-Falcon

Nagios  
集成当前正在使用的 Nagios

REST API  
其他监控产品请选择  
实时推送告警给自愈  
从企业告警API中获取告警

ICINGA 2  
集成当前正在使用的 Icinga 2

Amazon Web Services  
从AWS日志获取告警

EMAIL  
从邮件中获取告警

## 接入自愈方案

在菜单 [接入自愈] 中，点击 **接入自愈**，告警类型选择 **REST默认分类**。

点击新建 **自愈套餐** 的按钮，准备一个告警的处理动作。

自愈场景

告警类型 **REST默认分类** 1. 将什么告警做成自动处理?

平台 默认全选 集群 默认全选

模块 默认全选

自愈处理

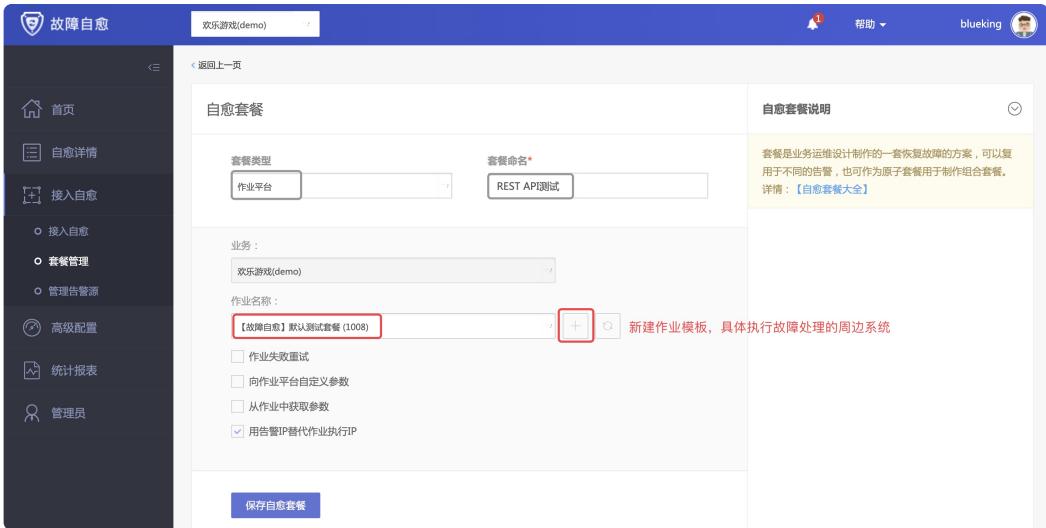
自愈套餐 请选择处理告警的自愈套餐 + 2. 告警处理的动作是什么?

通知方式  
开始时  微信  邮件  短信  电话  
成功时  微信  邮件  短信  电话  
失败时  微信  邮件  短信  电话

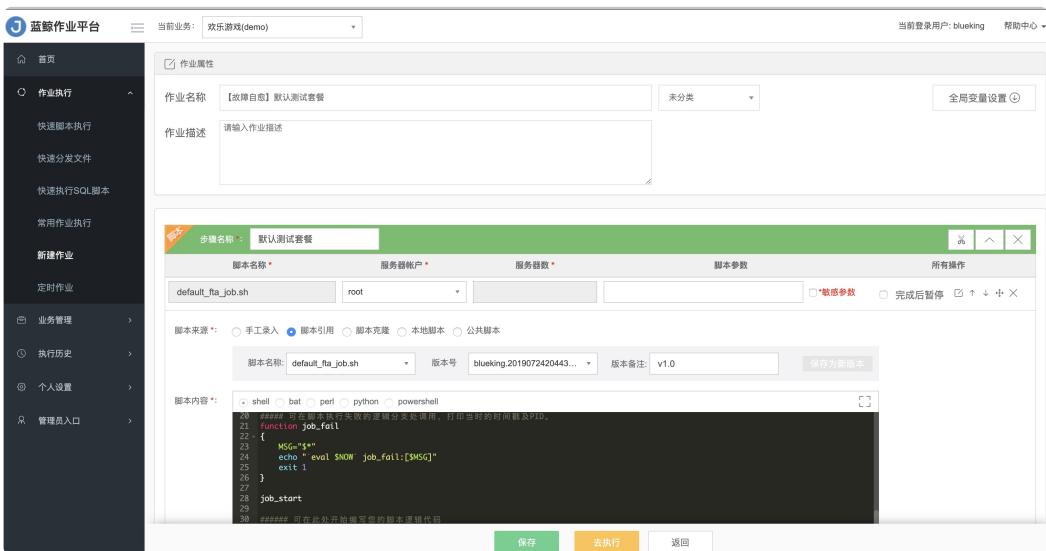
通知人员  业务运维: blueking.admin  
 更多通知人

注: 需要提前在蓝鲸控制台的【用户管理】填写联系方式

在自愈套餐页面，套餐类型选择 **作业平台**，点击作业名称右侧的加号，新建一个测试的作业模板。



点击 **新建作业的按钮** 后，跳转至作业平台，在菜单 **[作业执行] -> [新建作业]** 中，新建一个默认的作业即可。



保存 **REST API 测试** 的自愈套餐后，自动回到接入自愈的页面，保存自愈方案即可。

The screenshot shows the '故障自愈' (Fault Recovery) configuration interface. On the left sidebar, there are several sections: 首页 (Home), 自愈详情 (Recovery Details), 接入自愈 (Integrate Recovery), 套餐管理 (Plan Management), 管理告警源 (Manage Alert Sources), 高级配置 (Advanced Configuration), 统计报表 (Statistics Report), and 管理员 (Administrator). The main content area is titled '自愈场景' (Recovery Scenario). It includes a dropdown for '告警类型' (Alert Type) set to 'REST默认分类' (Default REST Category), and filters for '平台' (Platform) and '集群' (Cluster). Below this is the '自愈处理' (Recovery Processing) section, which contains a dropdown for '自愈套餐' (Recovery Plan) set to 'REST API测试' (REST API Test), and settings for '通知方式' (Notification Method) and '通知人员' (Notify Personnel). A note at the bottom states: '注：需要提前在蓝鲸控制台的【用户管理】填写联系方式' (Note: You need to fill in contact information in the User Management section of the BlueKing Control Panel in advance).

回到接入自愈列表，在列表中可以找到刚刚创建的自愈方案。

The screenshot shows the '接入自愈' (Integrate Recovery) list page. The table has columns: 告警类型 (Alert Type), 告警数量 (Alert Count), 平台 (Platform), 集群 (Cluster), 模块 (Module), 自愈套餐 (Recovery Plan), 告警源 (Alert Source), 自愈方案 (Recovery Scheme), 方案来源 (Source of Scheme), 启用 (Enabled), and 操作 (Operations). One row is highlighted with a red border, showing 'REST默认分类' (Default REST Category) as the alert type, 0 as the count, '(所有)' as the platform, '(所有)' as the cluster, '(所有)' as the module, 'REST API测试' (REST API Test) as the recovery plan, 'Zabbix监控' (Zabbix Monitoring) as the alert source, '人工配置' (Manual Configuration) as the recovery scheme, '是' (Yes) as enabled, and edit and delete icons.

## 自愈测试

在 **REST API(推送)** 的告警源管理页面，复制调用实例。

The screenshot shows the '告警源管理' (Alert Source Management) page for 'REST API(推送)'. The left sidebar includes sections: 首页 (Home), 自愈详情 (Recovery Details), 接入自愈 (Integrate Recovery), 告警源管理 (Alert Source Management), 告警源 (Alert Source), 管理告警源 (Manage Alert Sources), 高级配置 (Advanced Configuration), 统计报表 (Statistics Report), and 管理员 (Administrator). The main content area includes a table for告警源 (Alert Source) with columns: alarm\_type, alarm\_content, and alarm\_context. To the right is a '调用示例' (Call Example) section with a curl command example and a Secret key. At the bottom is the '编辑告警类型' (Edit Alert Type) section, which lists 'REST默认分类' (Default REST Category) with rule 'api\_default' and matching mode '字符串' (String).

将示例中的 IP 替换给该业务下任意一个 IP，然后贴到终端下执行。

```

[{"PANEL": "Logs", "LOG": "curl -d \"[{\"ip\": \"10.0.4.29\", \"source_id\": \"1004290088\", \"source_time\": \"2019-07-31 13:11:28+00:00\", \"alarm_type\": \"api_Default\", \"alarm_content\": \"FAILURE for production HTTP on machine 10.0.0.1\", \"alarm_context\": \"key1\":\"value1\", \"key2\":\"value2\"}]\" -X POST https://10.0.0.1/api/v1/cesp/api/fta/event/api/v1/5b9ec390-fb24-4755-9e04-8819017dc74/ HTTP/1.1\r\nHost: post-class.o.qcloud.com\r\nUser-Agent: curl/7.54.0\r\nAccept: */*\r\nProxy-Connection: Keep-Alive\r\nX-Secret-DCT1jKz3ZxNzKsCtUElYtDzSBjW0s\r\nContent-Length: 236\r\nContent-Type: application/x-www-form-urlencoded\r\n\r\n* upload completely sent off: 236 out of 236 bytes\r\n< HTTP/1.1 200 OK\r\n< Date: Mon, 01 Aug 2019 05:14:16 GMT\r\n< Content-Type: application/json; charset=utf-8\r\n< Vary: Accept-Language, Cookie\r\n< X-Frame-Options: SAMEORIGIN\r\n< Content-Language: zh\r\n< Content-Security-Policy: default-src 'self'; script-src 'self' 'unsafe-eval'; style-src 'self' 'unsafe-style-src'; font-src 'self'; img-src 'self'; frame-src 'self'; object-src 'self'; media-src 'self'; child-src 'self';\r\n< X-Cache-Lookup: MISS from SK-SQUIDWEB-00:8080\r\n< Transfer-Encoding: chunked\r\n< Connection: keep-alive\r\n< Connection: keep-alive\r\n\r\n* Connection #0 to host 10.0.0.1 left intact\r\n{"message": "[push restart event success]", "code": 1200, "data": {}, "result": true, "request_id": "45009db107c4fdf0b5d8362d2534f7"}", "TIME": "2019-07-31T13:11:28.456Z", "LEVEL": "INFO", "TAG": "curl", "FILE": "curl.log", "LINE": 1}
    
```



如果多次调试，请保证 `source_id` 唯一(重复会丢弃)，`source_time` 和服务器时间一致  
(过长会丢弃)。



在 [自愈详情](#) 页面可以找到自愈记录。

类型	产生时间	自愈耗时	平台	集群	模块	IP	状态	自愈结果
REST默认分类	07-31 13:11:28 +0800	22秒	Android_Weixin	一区	存储层	10.0.4.29	<span>成功</span>	执行job作业成功[371]

点开 `状态` 按钮，可以查看详情。

报警详情 业务 欢乐游戏(demo) 的主机 10.0.4.29 在 2019-07-31 13:11:28 +0800 发生 REST默认分类 : FAILURE for production HTTP on machine 10.0.0.1

处理过程 13:14:19 开始处理套餐[REST API测试]  
13:14:39 #0 REST API测试 | 成功: 执行Job作业成功[371]

处理状态 成功 执行Job作业成功[371]

操作 [重试整个流程](#)

#0 [46] REST API 测试

点击 `详情` 中的作业执行 ID，可查看执行的作业。

The screenshot shows the BlueKing Job Platform interface. At the top, there's a navigation bar with '当前业务: 欢乐游戏(demo)' and user information '当前登录用户: blueking 帮助中心'. Below the navigation is a sidebar with links like '首页', '作业执行', '业务管理', '执行历史', '个人设置', and '管理员入口'. The main content area displays a table for a completed job execution:

节点名称	default_fta_job.sh	节点类型	执行脚本	服务器账户	root	总时间(s)	0.118
节点状态	执行成功	开始时间	2019-07-31 13:14:20 +0800	结束时间	2019-07-31 13:14:21 +0800		

Below the table, there's a button labeled '执行成功(1)'. To the right, there are buttons for '查看步骤详情', '搜索日志', and '导出执行日志'. A search bar with placeholder '输入搜索内容' is also present. The bottom of the page shows a table of logs with columns: 云区域名称, IP, 返回码, 耗时(s). One entry is shown: default area, 10.0.4.29, 0, 0.164. Navigation buttons like '<', '<<', '>', and '>' are at the bottom, along with a '复制IP' button.

至此，一次模拟告警的故障自愈演示完毕。

REST API(推动)的场景在于，如果你使用的监控系统故障自愈默认未集成，则可以通过回调 REST API 的方式，将告警推送至故障自愈，故障自愈执行对应的处理动作，完成告警的自动处理。

故障自愈，如此简单。

## 对接 Open-Falcon

对接 Open-Falcon 的机制：Open-Falcon 模板中包含 CallBack（回调）功能，在回调地址中填写故障自愈分配给 Open-Falcon 的告警接收地址。

### 对接 Open-Falcon 告警源

在【接入自愈】下找到【管理告警源】菜单，点击【启用】Open-Falcon 告警源。

The screenshot shows the BlueKing Fault Recovery interface. On the left sidebar, there are several menu items: 首页 (Home), 自愈详情 (Recovery Details), 接入自愈 (Integrate Recovery), 套餐管理 (Plan Management), 管理告警源 (Manage Alert Sources), 高级配置 (Advanced Configuration), 统计报表 (Statistics Reports), and 管理员 (Administrator). The main content area is titled '管理告警源' (Manage Alert Sources) and displays a section for '已启用监控产品' (Enabled Monitoring Products) and '未启用监控产品' (Disabled Monitoring Products). Under '已启用监控产品', there are cards for REST API (其他监控产品请选择), ZABBIX (显示处理告警数: 15, 最后接收告警: 19/07/31 14:32:56), and 蓝鲸监控 (显示处理告警数: 123, 最后接收告警: 19/07/29 20:07:38). Under '未启用监控产品', there are cards for Falcon (集成当前正在使用的 Open-Falcon, 启用按钮被选中), Nagios (集成当前正在使用的 Nagios, 启用), REST API (其他监控产品请选择, 启用), Icinga (集成当前正在使用的 Icinga 2, 启用), and Amazon Web Services (从AWS获取告警, 启用). At the bottom left, there is a copyright notice: Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved.

可以找到故障自愈为 Open-Falcon 生成的全局回调地址，针对所有业务生效。

The screenshot shows the BlueKing Fault Recovery interface. The sidebar includes the same menu items as the previous screenshot. The main content area has a heading '接入流程' (Integration Process) and a sub-section for 'Falcon'. It shows a configuration step: '只需2步，集成企业内部正在使用的 Open-Falcon (接入后对所有业务都生效)' with a green toggle switch. Below this, there is a '回调地址' (Callback Address) input field containing 'callback: http://paas /api/c/compapi/ifa/event/open-falcon/'. There is also a note: '回调之前发提醒短信' (Send reminder SMS before callback), with several checkboxes for different types of callbacks. The bottom section is titled '编辑告警类型' (Edit Alert Types) and lists various alert rules with matching patterns and operations. A green button '新增告警类型' (Add New Alert Type) is visible at the top right of this section. At the bottom left, there is a copyright notice: Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved.

将该回调地址录入至 Open-Falcon 告警模板的回调地址中。

模板报警配置，对模板中的所有策略生效

报警接收组: bk.ops

修改报警组内成员

```
Callback: http://paas.m.../api/c/compapi/fta/event/open-falcon-plus/strategy/api/v1/events/execute?tenantId=123456789&strategyId=123456789&step=1&time=2019-01-01T12:00:00Z&operator=123456789&status=1&priority=1&metric=1&tags=srv=falcon,mount=sda
```

```
#callback request payload:
{
    "endpoint": event.Endpoint,
    "metric": event.Metric(),
    "status": event.Status,
    "step": event.CurrentStep,
    "priority": event.Priority(),
    "time": event.FormattedTime(),
    "tpId": event.TpId(),
    "expId": event.ExpressionId(),
    "straId": event.StrategyId(),
    "tags": "srv=falcon,mount=sda"
}

# for more infomation, look the code https://github.com/open-falcon/falcon-plus/blob/master/common/model/event.go
```

回调之前发提醒短信  回调之前发提醒邮件  回调之后发结束短信  回调之后发结束邮件

故障自愈对接完告警源后，接下来 **创建自愈方案** 和 **自愈套餐**。

## 创建自愈方案

选择【接入自愈】菜单，点击【接入自愈】，在接入自愈页面选择 Open-Falcon 的告警类型，比如磁盘容量，在【自愈套餐】一栏点击【+】，新建自愈套餐。

故障自愈

欢乐游戏(demo)

帮助 blueking

自愈场景

告警类型: 磁盘容量(df.\*)

平台: 默认全选

模块: 接入层

如果不同模块的处理方式有差异，请选择此项。

自愈处理

自愈套餐: 接入层日志清理

通知方式:

- 开始时: 微信, 邮件
- 成功时: 微信, 邮件
- 失败时: 微信, 邮件

通知人员: 业务运维: blueking.ops.a

其他信息

超时: 40 分以上按失败处理

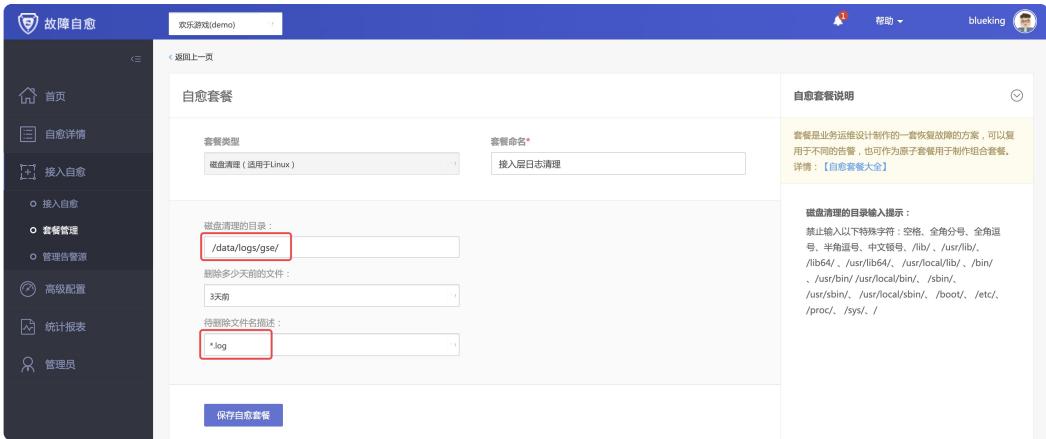
自愈方案名称: 清理 Nginx 日志

是否启用: 是

添加自愈策略

Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved.

在【新建自愈套餐】页面，选择清理的磁盘目录、时间以及文件名描述。



点击【保存自愈套餐】，回到接入自愈界面，点击【添加自愈策略】，至此 自愈方案 接入成功。

接下来，在 Open-Falcon 中模拟告警，以触发故障自愈。

## Open-falcon 告警自动处理

在 Open-Falcon 中触发告警后，在下图中可以看到接入层模块磁盘告警的自愈示例，匹配接入层的磁盘清理套餐，清理其日志文件，整个过程不到 30 秒。

告警详情 业务 蓝鲸 的主机 10.\*\*\* 在 2018-05-07 20:13:00 +0800 发生 磁盘容量(df.\*) : PROBLEM(df.statistics.used.percent)

处理过程  20:13:36 开始处理套餐清理nginx日志  
 20:13:57 #0 清理nginx日志 | 成功: 执行Job作业成功[82]

处理状态 成功 执行Job作业成功[82]

操作 重试整个流程

#0 [42] 清理nginx 日志

Open-Falcon 的资源标识 `endpoint` 默认是主机名，于是故障自愈将蓝鲸 CMDB 自动上报的主机名转换为 IP，然后在做匹配、告警自动处理。

## 集成 Icinga 2

两步完成 Icinga 2 接入自愈。

## 启用 Icinga 2 告警源

在【管理告警源】菜单中，【启用】Icinga 2 告警源。

管理告警源 对接企业正在使用的监控产品，迈向无人值守的第一步

已启用监控产品

Falcon 已处理告警: 0 未处理告警: 0 最后接收告警: --	REST API 其他监控产品请选择	ZABBIX 已处理告警: 15 未处理告警: 0 最后接收告警: 19/07/31 14:32:56	蓝鲸监控 已处理告警: 123 未处理告警: 62493 最后接收告警: 19/07/31 15:23:59
---	-----------------------	---	--

未启用监控产品

Nagios 集成当前正在使用的 Nagios 启用	REST API 从企业告警API中获取告警 启用	Icinga 2 集成当前正在使用的 Icinga 2 <b>启用</b>	AWS 从AWS获取告警 启用	EMAIL 从邮件中获取告警 启用
----------------------------------	---------------------------------	---	-----------------------	-------------------------

Copyright © 2012-2019 Tencent  
BlueKing.  
All Rights Reserved.

参照 Icinga 2 的接入流程，完成告警源的接入。

只需3步，集成企业内部正在使用的 Icinga 2 ( 绿入后对所有业务都生效 )

开关状态：开启

接入流程

- 1 下载推送脚本  
在Icinga2 Server上下载[推送脚本](#)(脚本中包含动态生成的AppId、SecretKey)  
wget 'http://pass-952dfe8&fta\_application\_secret=' .com/o/bk\_fta\_solutions/0/alarm\_source/scripts/icinga2\_fta\_pusher.py?fta\_application\_id=' -O icinga2\_fta\_pusher.py --no-check-certificate  
chmod +x icinga2\_fta\_pusher.py
- 2 安装告警配置  
指定Icinga2根目录进行安装(如/etc/icinga2)  
.icinga2\_fta\_pusher.py -c install /etc/icinga2
- 3 重新加载Icinga2  
.etc/init.d/icinga2 reload
- 4 完成

编辑告警类型 自定义告警类型，让告警处理更加明确

新增告警类型

名称	规则	匹配模式	操作
----	----	------	----

配置信息

Copyright © 2012-2019 Tencent  
BlueKing.  
All Rights Reserved.

## 创建自愈方案

参照 [对接 Open-falcon](#)，完成自愈的接入。

# 集成 Nagios

两步完成 Nagios 接入自愈。

## 启用 Nagios 告警源

在【管理告警源】菜单中，【启用】Nagios 告警源。

The screenshot shows the BlueKing interface for managing alert sources. On the left sidebar, under '故障自愈' (Fault Recovery), there is a '管理告警源' (Manage Alert Sources) section. In the main content area, there are two sections: '已启用监控产品' (Enabled Monitoring Products) and '未启用监控产品' (Disabled Monitoring Products). Under '已启用监控产品', there are four cards: Icinga, Falcon, REST API, and Zabbix. Under '未启用监控产品', there are four cards: Nagios, REST API, Amazon Web Services, and Email. The 'Nagios' card has a red box around its '启用' (Enable) button, indicating it is the target for this step.

参照 Nagios 的接入流程，完成告警源的接入。

The screenshot shows the BlueKing interface for integrating Nagios. On the left sidebar, under '故障自愈' (Fault Recovery), there is a '接入自愈' (Integration Recovery) section. In the main content area, there is a '接入流程' (Integration Flow) section with five numbered steps: 1. 下载推送到本机 /usr/local/nagios/libexec/eventhandlers/ 目录 (将本地命令生成的AppId, SecretKey) 2. 修改告警配置 在配置文件 /var/local/nagios/etc/objects/commands.cfg 中添加以下配置： 3. 声明全局 event handler 在配置文件 /var/local/nagios/etc/nagios.cfg 中添加如下配置 4. 重新加载Nagios /etc/init.d/nagios reload 5. 完成. Below the flowchart is a table titled '编辑告警类型' (Edit Alert Type) with two rows: 'HTTP(http)' and 'CPU(cpu)'. The table includes columns for '名称' (Name), '规则' (Rule), '匹配模式' (Match Mode), and '操作' (Operation).

## 创建自愈方案

参照 对接 Open-falcon，完成自愈的接入。

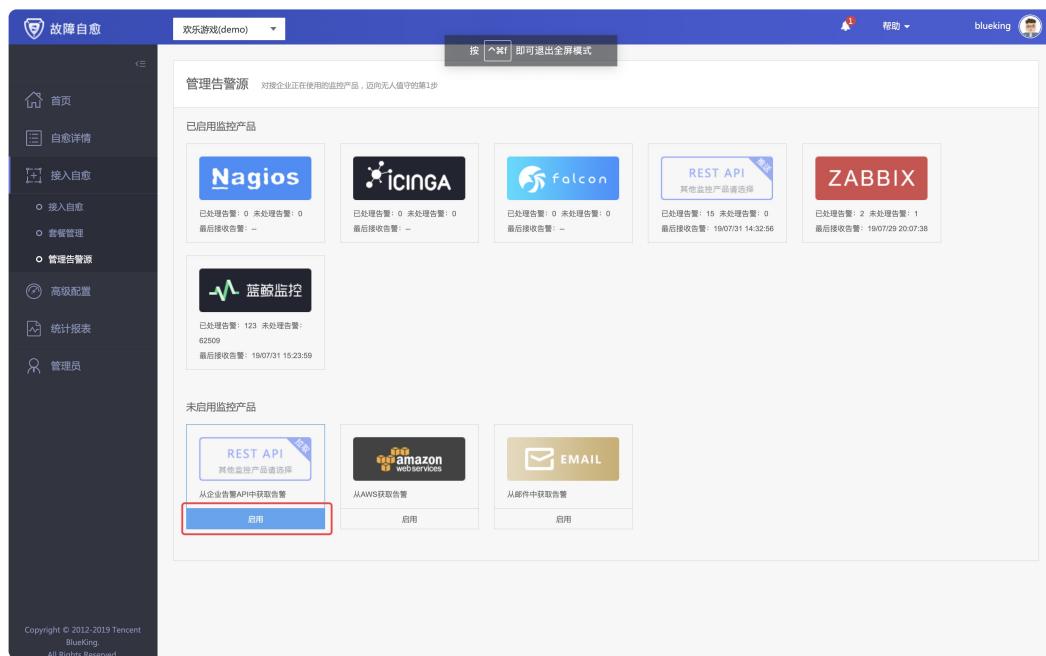
# 集成 REST API 拉取

如果企业使用的监控产品故障自愈未集成，可以把监控产品的告警使用 **REST API 推送** 至故障自愈，或故障自愈定期从监控产品通过 **REST API 拉取** 的方式获取告警。

本文介绍如何周期性拉取监控产品告警。

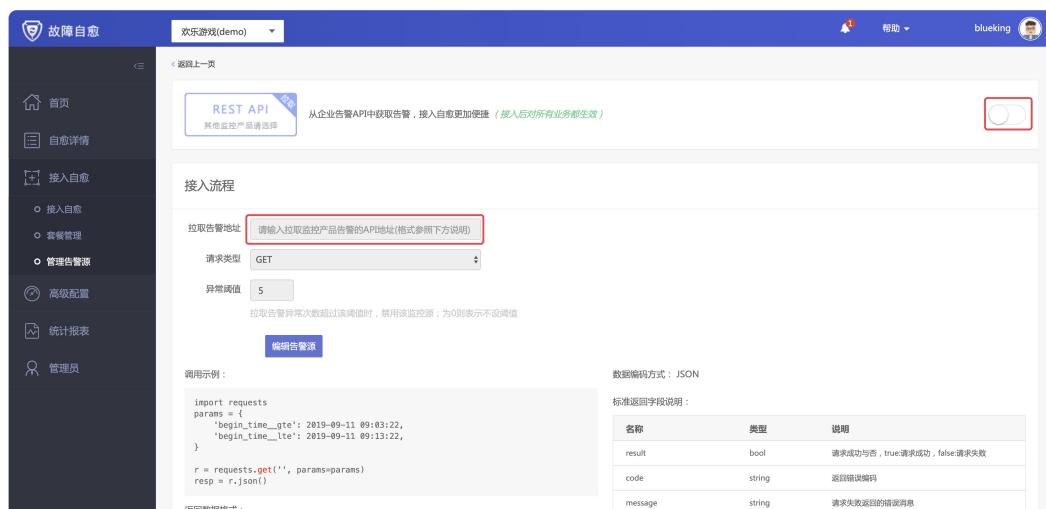
## 启用 REST API 拉取告警源

在【管理告警源】菜单中，【启用】REST API 拉取 告警源。



The screenshot shows the BlueKing Fault Recovery interface. On the left sidebar, under '接入自愈' (Integrate Recovery), there's a '管理告警源' (Manage Alert Sources) section. It displays a grid of monitoring products. The 'REST API' row is highlighted with a red box around its '启用' (Enable) button. Other products shown include Nagios, Icinga, Falcon, and Zabbix. Below the grid, there's a section for '未启用监控产品' (Unenabled Monitoring Products) with three options: '从企业告警API中获取告警' (Get alerts from enterprise alert API), '从AWS获取告警' (Get alerts from AWS), and '从邮件中获取告警' (Get alerts from email).

参照 REST API 拉取 的接入流程，完成告警源的接入。



The screenshot shows the 'Integrate Recovery' section of the BlueKing Fault Recovery interface. It includes a 'REST API' configuration panel with fields for '取告警地址' (Get alert address) and '请求类型' (Request type: GET). There's also a '异常阈值' (Exception threshold: 5) input field. Below these, there's a '调用示例' (Call example) code block and a '标准返回字段说明' (Standard return field description) table. The table has columns for '名称' (Name), '类型' (Type), and '说明' (Description). It includes rows for 'result' (bool, true:请求成功, false:请求失败), 'code' (string, 返回错误编码), and 'message' (string, 请求失败返回的错误消息).

The screenshot shows a user interface for self-healing configuration. On the left, there's a sidebar with options like '接入自愈' (Access Self-healing), '高级配置' (Advanced Configuration), '统计报表' (Statistics Reports), and '管理员' (Administrator). The main area has three tabs: 'data', 'list', and '请求成功返回的数据' (Data returned successfully). The 'data' tab shows a JSON response:

```
{
  "result": true,
  "message": '',
  "data": [
    {
      "ip": "10.0.0.1",
      "source_id": "123456",
      "source_time": "2017-04-06 16:51:00",
      "alarm_type": "default",
      "alarm_content": "FAILURE for production/HTTP on machine 10.0.0.1",
    },
    {
      "ip": "10.0.0.1",
      "source_id": "123457",
      "source_time": "2017-04-06 16:50:00",
      "alarm_type": "default",
      "alarm_content": "FAILURE for production/HTTP on machine 10.0.0.1",
    }
  ]
}
```

The 'list' tab contains a table of parameters:

参数	必须	备注
IP	Y	告警源IP
source_id	Y	告警源的告警ID，全局唯一
source_time	Y	告警发生的时间，格式：YYYY-MM-DD HH:mm:ssZZ
alarm_type	N	告警类型， <a href="#">点击添加告警类型</a>
alarm_content	N	告警详情

The '请求成功返回的数据' tab shows the JSON response again.

**自助接入说明**

如您的API格式不满足上述要求，您可以按照如下方式自助接入：

- 在故障自愈后台项目的 \$proj\_base\_dir/project/poll\_alarm/custom\_monitor.py 文件中修改拉取告警、清洗告警信息的方法
- 修改完成后，重启故障自愈后台服务，重启步骤如下：

```
# 在部署故障自愈后台的服务器上执行
workon fta && bin/fta.sh restart
```

**编辑告警类型** 自定义告警类型，让告警处理更加明细

**新增告警类型**

名称	规则	匹配模式	操作
默认分类	default	字符串	

Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved.

## 创建自愈方案

参照 [对接 Open-falcon](#)，完成自愈的接入。

# 创建磁盘清理自愈套餐和方案

磁盘使用率告警的自动化是一个经典的场景，本文通过这个案例介绍故障自愈中 **自愈套餐**、**自愈方案** 的概念，以及故障自愈是如何实现告警的自动化处理。

首先，创建自愈套餐。

## 创建磁盘清理自愈套餐

磁盘告警来了，执行什么动作来清理磁盘。

{% video %}1.1\_fta\_create\_fta\_solutions\_disk.mp4{% endvideo %}

接入流程：

依次选择 【接入自愈】 → 【套餐管理】 → 【创建自愈套餐】

按照 【磁盘清理(适用于 Linux)】套餐页面的提示，输入【套餐命名】、【磁盘清理的目录】，选【删除多少天的文件】和【待删除文件名描述】，然后保存自愈套餐即可。

自愈套餐

套餐类型：磁盘清理(适用于Linux)

套餐命名\*：请输入

磁盘清理的目录：

删除多少天前的文件：

待删除文件名描述：

自愈套餐说明

套餐是业务运维设计制作的一套恢复故障的方案，可以复用于不同的告警，也可作为原子套餐用于制作组合套餐。详情：[【自愈套餐大全】](#)

磁盘清理的目录输入提示：

禁止输入以下特殊字符：空格、全角分号、全角逗号、半角逗号、中文顿号、/lib/、/usr/lib/、/lib64/、/usr/lib64/、/usr/local/lib/、/bin/、/usr/bin/、/usr/local/bin/、/sbin/、/usr/sbin/、/usr/local/sbin/、/boot/、/etc/、/proc/、/sys/、/

**保存自愈套餐**

提示：该套餐实现出现磁盘使用率告警时，找出 `/data/log/` 目录下 `3天前` 以 `.log` 结尾的文件并删除。

接下来我们需要把 **磁盘使用率** 告警接入刚刚创建的磁盘清理套餐。

## 接入磁盘清理自愈方案

{% video %}1.2\_fta\_bind\_fta\_solutions\_disk.mp4{% endvideo %}

在【接入自愈】菜单中点击【接入自愈】。

接入自愈

告警类型	告警数量	生效集群	生效模块	自愈套餐	告警源	备注	是否启用	操作
CPU使用率(system.cpu.*)	31	(所有)	(所有)	Zabbix_test	Zabbix监控	zabbix&nbsp;test	是	
CPU利用率	21	(所有)	(所有)	【快速】发送CPU使用率TOP10进程列表(适用于	腾讯云监控		是	
jiayuantest	36	(所有)	(所有)	cc相关套餐_jiayuan	REST API		否	
ping不可达	126	(所有)	(所有)	【快速】腾讯云CVM重启	腾讯云监控	腾讯云服务器重启的自愈方	是	
Ping检查(icmping*)	148	(所有)	(所有)	组合套餐-审批和JOB	Zabbix监控	Zabbix每日测试自愈方案	是	
REST默认分类	76	(所有)	(所有)	【快速】发送内存使用率TOP10进程列表(适用于	REST API		是	

进入【接入自愈】页面，做如下配置：

**自愈场景**

● 把什么告警接入故障自愈?

告警类型 \* 磁盘利用率

按内容筛选 (只筛选内容匹配的那部分告警, 使用正则表达式匹配, 不填为不过滤)

● 一般在自定义告警下使用, 匹配告警内容

● 选择自愈套餐生效的集群和模块

集群:

模块:

**自愈处理**

是否启用  是  否 ● 勾选是, 标识启用该自愈套餐

自愈套餐 /data/log/目录的磁盘清理套餐 ● 重要: 选择这个告警类型对应的自愈套餐

通知方式 **开始时**  微信  邮件  短信  电话 ● 自愈执行结果的通知方式

**成功时**  微信  邮件  短信  电话

**失败时**  微信  邮件  短信  电话

通知人员  业务运维  项目组  客户支持  第三方集成  其他  选择自愈执行时的通知接收人

额外通知人  请选择额外通知人

注: 需要提前在蓝鲸桌面的“个人中心”填写联系方式

**额外信息**

超时 \* 40 分 以上按失败处理 ● 如果自愈执行超过指定时间, 则按失败处理, 默认40分钟

备注 清理 /data/log/目录

如此, 完成磁盘清理告警接入故障自愈。

/data/log/目录的磁盘清理套餐	600	1	磁盘清理 (适用于 Linux )
---------------------	-----	---	-------------------

## 告警通知渠道

故障自愈是蓝鲸 PaaS 上一个 SaaS, 通知渠道使用 PaaS 的通知 ESB 组件, 在蓝鲸的独立部署版本(企业版、社区版)需要在开发者中心后台设置。

### 在通知 ESB 组件中配置通知渠道

在菜单【开发者中心】->【API 网关】->【通道管理】中, 过滤 [CMS] 蓝鲸消息管理, 如配置邮件网关, 点击【[CMS]发送邮件】即可。

组件通道列表

#	通道名称	请求路径	更新时间	是否开启	API地址
<input type="checkbox"/>	[CMISI] 查询消息发送类型	/cmisi/get_msg_type/	2019-09-10	<input checked="" type="radio"/>	API地址
<input type="checkbox"/>	[CMISI] 发送邮件	/cmisi/send_mail/	2019-09-10	<input checked="" type="radio"/>	API地址
<input type="checkbox"/>	[CMISI] 通用消息发送	/cmisi/send_msg/	2019-09-10	<input checked="" type="radio"/>	API地址
<input type="checkbox"/>	[CMISI] 发送短信	/cmisi/send_sms/	2019-09-10	<input checked="" type="radio"/>	API地址
<input type="checkbox"/>	[CMISI] 公共语音通知	/cmisi/send_voice_msg/	2019-09-10	<input checked="" type="radio"/>	API地址
<input type="checkbox"/>	[CMISI] 发送微信消息	/cmisi/send_weixin/	2019-09-10	<input checked="" type="radio"/>	API地址

+ 添加通道 删除

帮助

- 通道默认是开启的；如果关闭，则会在通道名称前提示已关闭

更多详情请查看 [使用指南](#)

进入邮件网关的配置页面，按提示完成邮件网关的配置。

修改通道

通道名称	发送邮件 *														
通道名称，长度限制为64字符，例如“查询服务器列表”															
通道路径	/cmisi/send_mail/ *														
通道路径，以斜杠开头，只能包含斜杠、字母、数字、下划线(_)、连接符(-)，一般设置为“/system_name/component_name/”，例如“/host/get_host_list/”；通道路径需唯一															
所属系统	[CMISI] 蓝鲸消息管理														
对应组件代号	generic.cmisi.send_mail *														
组件代号，只能包含小写字母、数字、下划线或点号，由三部分组成：“前缀(generic) 系统名小写.组件类名小写”，例如“generic.host.get_host_list”															
API类型	执行 API *														
超时时间	单位秒，未设置时以所属系统超时长为准														
组件配置	<table border="1"> <thead> <tr> <th>变量名</th> <th>变量值</th> </tr> </thead> <tbody> <tr> <td>dest_url</td> <td></td> </tr> <tr> <td>smtp_host</td> <td></td> </tr> <tr> <td>smtp_port</td> <td>25</td> </tr> <tr> <td>smtp_user</td> <td>blueking</td> </tr> <tr> <td>smtp_pwd</td> <td></td> </tr> <tr> <td>smtp_usessl</td> <td>False</td> </tr> </tbody> </table>	变量名	变量值	dest_url		smtp_host		smtp_port	25	smtp_user	blueking	smtp_pwd		smtp_usessl	False
变量名	变量值														
dest_url															
smtp_host															
smtp_port	25														
smtp_user	blueking														
smtp_pwd															
smtp_usessl	False														

更多细节，请参考以下 2 篇文档：

- 如何配置通知渠道，如邮件、微信、短信等？
- 经验分享：测试邮件服务是否正常

## 告警通知效果

通知渠道有 4 种：微信、电话、邮件、短信，以下为部分通道的通知效果：

- 邮件通知

【自愈通知】磁盘使用率 蓝鲸 【自愈开始】 ☆

发件人:  圈

时间: 2019年9月11日(星期三) 中午11:00

收件人:  圈

纯文本 |    

自愈场景

蓝鲸  
10.0.4.140 (--)  
10:58:00 发生  
告警: 磁盘使用率 (磁盘使用率: 当前指标值(80.61%) >= (80.0%))

自愈结果

【自愈开始】

套餐: 蓝鲸服务器清理日志  
详细: 无

自愈过程

Copyright © 2012-2019 Tencent BlueKing. All Rights Reserved

- 微信通知



## 故障自愈套餐大全

故障自愈套餐主要有三类组合套餐: 快捷套餐、周边系统、组合套餐, 以下为每个套餐的详细说明。

## 快捷套餐

每个业务默认都会有一批可以选择的套餐，用户几乎不需要对其做任何配置可以直接使用。对于这种套餐，我们用称为『快捷』套餐，部分『快捷』套餐只能在组合套餐中选用。

### 常规快捷套餐

#### 磁盘清理(适用于 LINUX)

背后对应一个作业平台套餐，调用作业平台内置的一个跨业务作业，执行磁盘清理。

更多请参考 [创建磁盘清理自愈套餐和方案](#)。

### 汇总

- 具体操作: 根据配置按一定维度收敛后，发送通知给用户。
- 使用场景: 对于一些大量出现的告警，如单机性能告警、流量告警等，可以用此套餐收敛后发送通知。
- 参数说明: 时间段和告警数必须填一个。
  - 按时间段汇总: 从收到第一条告警开始计时，到规定时间后结束。
  - 按告警数汇总: 从收到第一条告警开始计数，到规定数量后结束。
  - 少于多少条不通知: 选择时间段汇总时，如果到时间后，告警数量少于指定数量，就不发通知。

#### 【快捷】发送 CPU 使用率 TOP10 的进程(微信)

当产生 CPU 使用率时，调用作业平台的作业，通过微信发送 占用 CPU 的 TOP 10 进程列表，辅助运维分析原因。

该套餐实际是一个组合套餐，第 1 步执行作业，第 2 步通过微信发送第 1 步的结果，更多细节请参考 [上下文传参](#) 的介绍。

#### 【快捷】发送内存使用率 TOP10 的进程(微信)

同上。

#### 【快捷】CC 移到“故障机”模块

调用蓝鲸配置平台的接口，将故障机移动至当前业务的故障机模块。

### 组合套餐中的快捷套餐

该分类下套餐仅能在组合套餐中使用。

#### 【快捷】配置平台拷贝故障机属性到备机

“

(需先调用 获取故障机备机 套餐)

”

- 具体操作:

- 调用配置平台替换接口，将替换机转移到与故障机相同的模块下(包括属于故障机多个模块的情况)。
  - 拷贝故障机的标准属性、自定义属性到替换机。

- 使用场景: 故障机切换后拷贝故障机 CC 信息到备机;

## 【快捷】后续处理对象故障机与备机互换

“

(需先调用 获取故障机备机 套餐)

”

- 具体操作: 调用此套餐后，在组合套餐的后续流程中，操作对象将会替换。流程默认都是对故障机进行操作，调用一次后将会默认对备机进行操作。再调用一次，则恢复原状。
- 使用场景: 例如当获取选定的备机后，想对备机进行初始化操作，那么这时候就要先调用一次此套餐，再调用初始化。初始化后，又想把故障机移动到配置平台的故障机模块，那么就又要调用一次此套餐，然后再调用配置平台移动模块。

## 周边系统

故障自愈自身并没有恢复告警的执行能力，而是通过蓝鲸 PaaS 的 ESB 模块调用作业平台、标准运维的能力来实现。

### 作业平台

调用蓝鲸作业平台，作业平台脚本需要按照一定规范来。

- 具体操作: 调用作业平台作业。
- 使用场景: 希望在机器上执行脚本的故障恢复操作，如拉起进程等。
- 参数说明:
  - 通过选择框选择作业平台作业。
  - 自定义传入作业平台的参数(不填默认传入故障 IP)。
  - 通过作业平台脚本获取自愈变量来在其他套餐中使用。

## 标准运维流程

同上，区别在于调用的是标准运维。

## HTTP 回调

裸调企业内部的恢复告警的处理接口，如重启服务器，不经过蓝鲸 PaaS 的 ESB。

一般不推荐裸调接口。

## 组合套餐类

### 组合套餐

组合套餐，顾名思义就是把该业务下的套餐和官方通用套餐组合起来使用。

- 具体操作: 按照配置的流程顺序执行套餐
- 使用场景: 当单一的套餐无法满足故障恢复流程时，将套餐串起来用。

### 获取故障机备机

根据配置平台的配置属性获取备机许多故障替换操作都依赖于此套餐。

除了通过这个套餐获取备机外，也可以通过作业平台来获取备机，只要使用作业平台套餐获取变量为 ip\_bak 的参数即可。

- 具体操作: 根据配置，在配置平台指定的 Set 与 AppModule 中，查找指定属性与故障机相同的机器。将所有符合条件的机器 IP 通过微信让运维审批选择。此后备机 IP 就能通过变量在其他套餐中获取。
- 使用场景: 在组合套餐中调用故障替换的操作套餐前必须先调用过此套餐。

更多变量请参考 [套餐内置变量](#)。

### 审批

- 具体操作: 调用不同的接口发送审批(发送微信失败会改为短信，发送短信失败则发微信，邮件通知接口调用失败将不做处理)。
- 使用场景: 一般用于组合套餐中；审批，执行服务器故障替换流程时，让运维审批是否执行替换，降低风险。

### 通知

- 具体操作: 调用不同的接口发送通知(发送微信失败会改为短信，发送短信失败则发微信，邮件通知接口调用失败将不做处理)。

- 使用场景: 一般用于组合套餐中, 通知, 自定义消息发送到微信等平台。

## 暂停等待

- 具体操作: 暂停一定时间。
- 使用场景: 重启操作后 Agent 可能还没准备好, 可以等待一定时间后再继续后面的流程。

## 自定义收敛防御

- 具体操作: 在指定时间内, 出现过指定告警数量的相同 IP, 相同告警类型, 将会收敛次条告警。
- 使用场景: 想添加某些异常防御条件的时候, 大部分需求故障自愈的收敛功能就能满足。

# 进程告警

故障自愈除了能处理单机性能告警外, 还能处理服务类的告警, 比如进程告警。

比如 Nginx 进程挂掉了, 你需要拉起 Nginx 进程。

下面以 Nginx 进程告警接入自愈为例。

### 编写拉起 Nginx 进程的作业

在作业平台编写拉起 Nginx 进程的脚本。

“

脚本中除了拉起进程, 你还可以考虑增加进程检测的逻辑, 保证拉起进程这个过程无误。

”

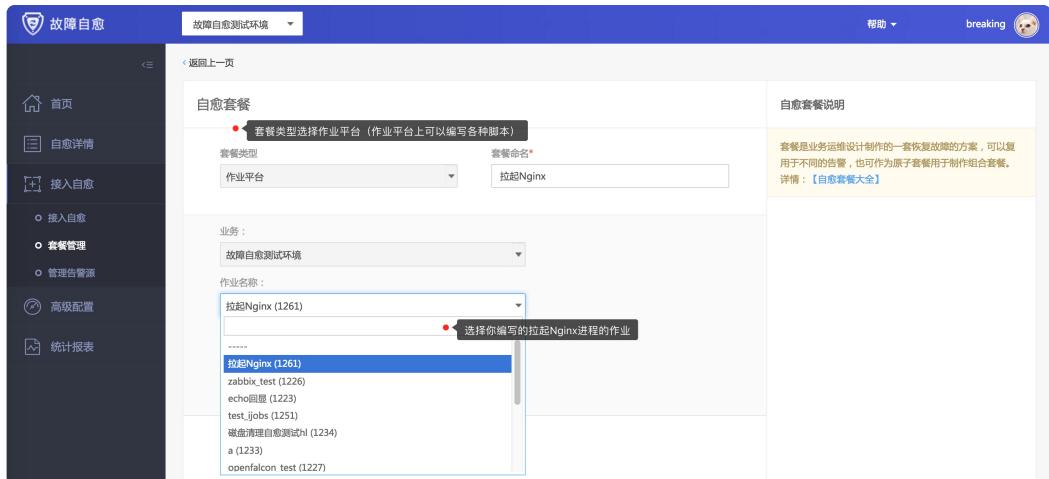
The screenshot shows the Bluewhale Job Platform interface. On the left is a sidebar with navigation items: 首页 (Home), 作业执行 (Job Execution) with a dropdown, 业务管理 (Business Management), 执行历史 (Execution History), and 个人设置 (Personal Settings). The main area has a title bar with the current business: 故障自愈测试环境 (Fault Recovery Test Environment). Below this is a search bar with the text start\_nginx.sh, a dropdown for root, a green plus button, and a status bar indicating 共1台 (1 available). A large text area contains a shell script named start\_nginx.sh:

```
start_nginx.sh
root
+ 共1台

脚本来源 *: 手工录入 脚本克隆 本地脚本 公共脚本
脚本内容 *:
12 ##### 可在脚本执行成功的逻辑分支处调用, 打印当时的时间戳及PID。
13 function job_success
14 {
15     MSG=$*
16     echo "eval $NOW job_success:$MSG"
17     exit 0
18 }
19
20 ##### 可在脚本执行失败的逻辑分支处调用, 打印当时的时间戳及PID。
21 function job_fail
22 {
23     MSG=$*
24     echo "eval $NOW job_fail:$MSG"
25     exit 1
26 }
27
28 job_start
29
30 ##### 可在此处开始编写您的脚本逻辑代码。
31 ##### 作业平台中执行脚本成功和失败的标准只取决于脚本最后一条执行语句的返回值
32 ##### 如果返回值为0, 则认为此脚本执行成功, 如果非0, 则认为脚本执行失败
33 service nginx restart
```

## 创建拉起 Nginx 的自愈套餐

在【套餐管理】中新建作业平台套餐，选择刚刚在作业平台中创建的【拉起 Nginx】作业。



## 接入自愈

在【接入自愈】页面将【进程告警】关联【拉起 Nginx】套餐，自愈范围选中【nginx】模块。

自愈场景

告警类型：进程告警

按内容筛选：使用正则表达式匹配告警，不填为不过滤

集群：默认全选

模块：nginx • 只处理CMDB中的nginx模块

自愈处理

自愈套餐：拉起Nginx

至此，进程告警的自愈处理方案配置完毕。

更多进程告警的配置细节请参考 [蓝鲸监控告警自动处理](#)。

## CPU 使用率告警

一般出现 CPU 使用率告警后，我们很想知道是哪个进程造成的？

于是，故障自愈内置一个分析 CPU 使用率的快捷套餐 **『快捷』发送 CPU 使用率 TOP 10 进程列表 (适用于Linux)**。

### 接入 CPU 使用率告警套餐

选择【接入自愈】菜单，点击【接入自愈】，告警类型选择“CPU 使用率”，自愈套餐为“『快捷』发送 CPU 使用率 TOP 10 进程(微信)”

故障自愈

欢乐游戏(demo)

自愈场景

告警类型：[主机监控] CPU使用率

按内容筛选：使用正则表达式匹配告警，不填为不过滤

平台：默认全选

集群：默认全选

模块：默认全选

自愈处理

自愈套餐：『快捷』发送CPU使用率TOP10的进程(微信)

通知方式：

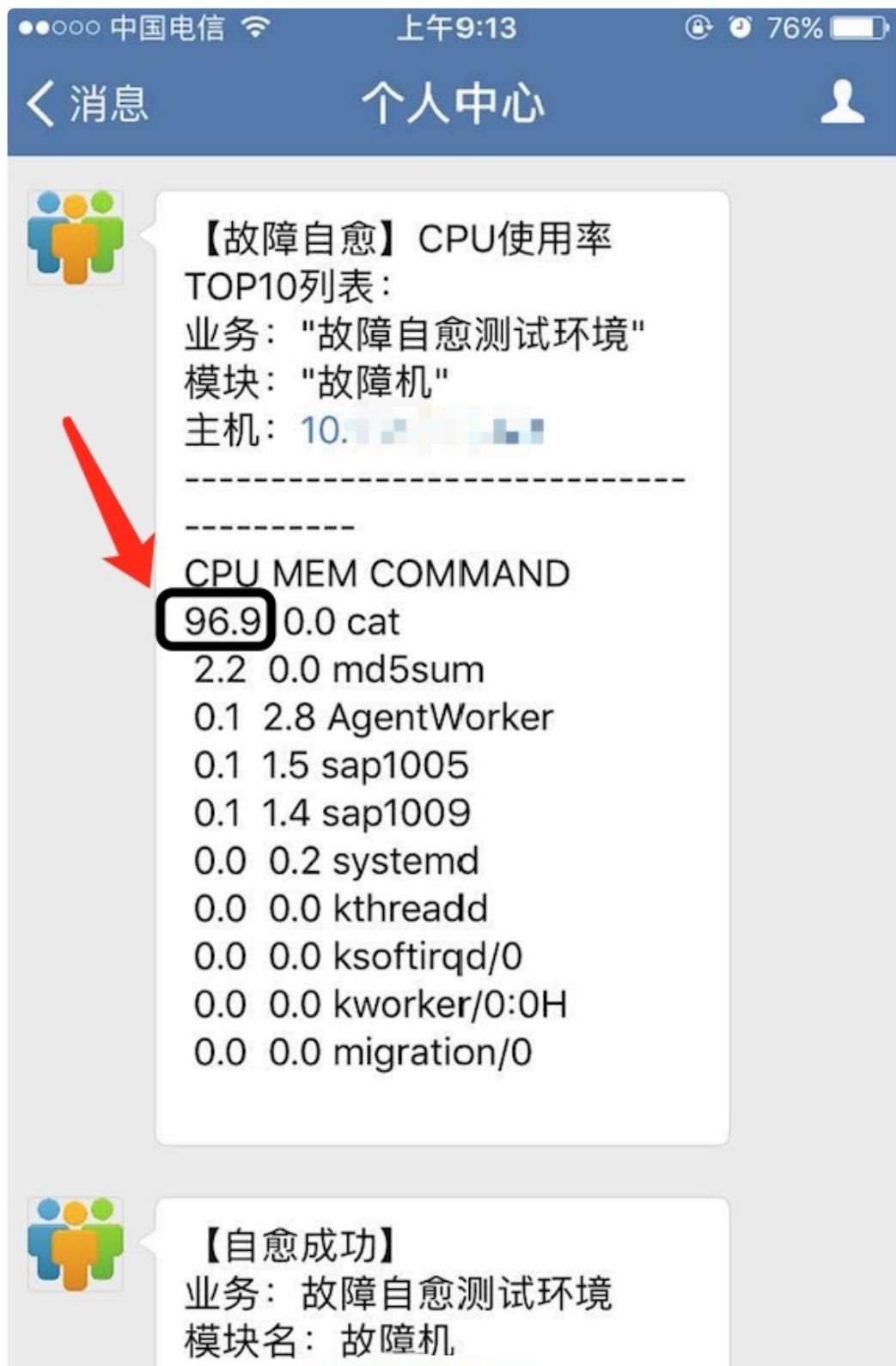
开始时	<input checked="" type="checkbox"/> 微信	<input checked="" type="checkbox"/> 邮件	<input type="checkbox"/> 短信	<input type="checkbox"/> 电话
成功时	<input checked="" type="checkbox"/> 微信	<input checked="" type="checkbox"/> 邮件	<input type="checkbox"/> 短信	<input type="checkbox"/> 电话
失败时	<input checked="" type="checkbox"/> 微信	<input type="checkbox"/> 邮件	<input type="checkbox"/> 短信	<input type="checkbox"/> 电话

通知人员： 业务运维: blueking:ops\_a  
 更多通知人

### 自愈测试

在监控系统中产生 CPU 使用率，触发自愈。

下图是一条 CPU 使用率告警的自愈记录，微信消息显示占用 CPU 使用率最高的进程是我们测试的进程 `cat`。





关于 CPU 使用率套餐的技术细节，请参考 [上下文传参](#)。

## 内存使用率告警

一般出现 [内存使用率告警](#) 后，我们很想知道是哪个进程造成的？

于是，故障自愈内置一个分析内存使用率的快捷套餐 [『快捷』发送内存使用率 TOP 10 进程列表\(微信\)](#)。

### 接入内存使用率告警套餐

选择【接入自愈】菜单，点击【接入自愈】，告警类型选择“内存使用率”，自愈套餐为“『快捷』发送内存使用率 TOP 10 进程(微信)”。

告警类型: [主机监控] 应用内存使用率

平台: 默认全选

模块: 默认全选

集群: 默认全选

自愈套餐: 『快捷』发送内存使用率TOP10的进程(微信)

通知方式:

- 开始时:  微信  邮件  短信  电话
- 成功时:  微信  邮件  短信  电话
- 失败时:  微信  邮件  短信  电话

通知人员:  业务运维: blueking;ops\_a  
 更多通知人

注: 需要提前在蓝鲸控制台的【用户管理】填写联系方式

### 自愈测试

在监控系统中产生内存使用率，触发自愈。

下图是一条内存使用率告警的自愈记录，微信消息显示占用内存使用率最高的进程是 [AgentWorker](#)。



## 【故障自愈】内存使用率TOP10

列表：

业务： "故障自愈测试环境"

模块： "故障机"

主机： 10.10.5.100.50

CPU MEM COMMAND

0.1 2.8 AgentWorker

0.0 2.1 secu-tcs-agent

0.1 1.5 sap1005

0.1 1.4 sap1009

0.0 1.3 tuned

0.0 1.2 barad\_agent

0.0 0.9 barad\_agent

0.0 0.6 barad\_agent

0.0 0.3 basereport

0.0 0.2 rsyslogd

关于内存使用率套餐的技术细节，请参考 [上下文传参](#)。

## 告警收敛

我们很有可能收到重复的告警，所以故障自愈推出告警收敛功能。

满足一定规则后，执行对应的收敛方式。

## 单业务设置

点击右上角 新建收敛规则，按下图灰色框中内容添加一条规则。

后台分析收敛规则库 (按优先级从高到低)						
针对告警类型	在一定条件下	触发频次	收敛方式	备注	生效范围	操作
ping不可达, 机器重启, 磁盘只读	告警类型: ping不可达, 机器重启,磁盘只读	5分钟内 1条以上	成功后跳过	一系列单机异常类告警 (同一起事件造成的多个告警)	<input type="checkbox"/> 全业务	
PrivateBandwidthOut, 外网出带宽, 内存利用率, 故障	主机: 相同 自愈套餐: 相同	1分钟内 1条以上	成功后跳过	一系列处理套餐相同的告警 (同一起事件造成的多个告警)	<input type="checkbox"/> 全业务	
ping不可达, 机器重启	告警类型: ping不可达, 机器重启 业务: 相同	2分钟内 3条以上	异常防御需审批	一系列同业务的单机异常告警 (可能服务器批量掉电)	<input type="checkbox"/> 全业务	
CPU使用率(load.*)	告警类型: 相同	10分钟内 3条以上	超出后汇总		<input type="checkbox"/> 当前业务	
CPU使用率 (system.cpu.*), Ping检查 <b>icmping*</b>	主机: 相同	5分钟内 3条以上	异常防御需审批	Zabbix异常防御收敛审批	<input type="checkbox"/> 当前业务	

图 1. 告警收敛

下图中命中了上述最后一条规则(在同一台主机上，5分钟内出现3条以上告警，由于没有进行审批动作，于是20分钟后超时了)。

2017-05-18 - 2017-05-18		● 1次收敛事件	● 10次重要警告	● 0次预警	刷新	导出IP	请输入IP	🔍
*所有收敛	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果	
□ 异常防御 10	05-18 17:50:05		Zabbix	异常防御收敛审批 (#45)			影响范围: 空闲机池	
□ Ping检查(icmping*)	05-18 17:49:11	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:14	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:16	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:17	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:19	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:21	20分钟	空闲机池	故障机	10.135.181.30		执行组合套餐(组合套餐-审批和IOPS)失败	
□ Ping检查(icmping*)	05-18 17:49:55	19分钟	空闲机池	故障机	10.135.181.30		[系统]驳回了审批[超时]	
□ Ping检查(icmping*)	05-18 17:49:57	19分钟	空闲机池	故障机	10.135.181.30		[系统]驳回了审批[超时]	
□ Ping检查(icmping*)	05-18 17:49:58	19分钟	空闲机池	故障机	10.135.181.30		[系统]驳回了审批[超时]	
□ Ping检查(icmping*)	05-18 17:50:00	19分钟	空闲机池	故障机	10.135.181.30		[系统]驳回了审批[超时]	

图 2. 告警收敛结果

## 设置全局收敛规则

默认手动添加的收敛规则的生效范围是 **当前业务**，如果希望在全业务下都生效，可以在 Django 后台设置。

任意找一个业务设置收敛规则

图 3. 添加收敛规则



图 4. 收敛规则添加成功

## 在 Django 后台修改生效范围

使用管理员角色访问以下地址

[http://\\${PaaS\\_URL}/o/bk\\_fta\\_solutions/admin/fta\\_solutions\\_app/incidentdef/](http://${PaaS_URL}/o/bk_fta_solutions/admin/fta_solutions_app/incidentdef/)

ID	英文名或代号	规则简介	优先级	是否启用	业务编码
2	same_solution	一系列处理套餐相同的告警（同一起事件造成的多个告警）	50	✓	0
5	ping_and_agent_time_out	一系列单机异常类告警（同一起事件造成的多个告警）	30	✓	0
6	collect_alarm	请求	100	✓	0
7	349f5cd68559a4b86a0c3d5601c26337	应用内存使用率	100	✓	0
8	560045526eedebc4f856e26ac17fa64	测试	100	✓	5
10	breaking-test		100	✓	2

图 5. 在 Django 后台修改生效范围

找到刚添加的收敛规则，将其 **业务编码** 修改为 0 (即对所有业务生效.)

访问 **告警收敛** 可以发现 **规则来源** 从 **当前业务** 换成了 **系统内置**



图 6. 收敛规则来源修改成功

在不同业务下测试，均生效

类型	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果
成功后跳过	09-05 12:20:10 +0800		测试				影响范围：数据服务模块
breaking	09-05 12:14:39 +0800	21秒	数据服务模块	monitor	10.178.106.31	✓	执行Job作业成功[8833]
breaking	09-05 12:14:39 +0800	1秒	数据服务模块	monitor	10.178.106.1*	⟳	被收敛：对于(8c85a17ed7734f96a99a...)
breaking	09-05 12:14:39 +0800	0秒	数据服务模块	monitor	10.178.106.1*	⟳	被收敛：对于(8c85a17ed7734f96a99a...)
breaking	09-05 12:20:39 +0800	20秒	数据服务模块	monitor	10.178.106.1*	✓	执行Job作业成功[8834]
breaking	09-05 12:20:40 +0800	0秒	数据服务模块	monitor	10.178.106.1*	⟳	被收敛：对于(8c85a17ed7734f96a99a...)

图 7. 在 A 业务下测试收敛规则

类型	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果
成功后跳过	09-05 13:12:07 +0800		测试				影响范围：idle pool
breaking	09-05 13:08:40 +0800	22秒	idle pool	Idle machine	10.178.106.1*	✓	Execute combination solution[ 'Short...
breaking	09-05 13:08:41 +0800	1分钟	idle pool	Idle machine	10.178.106.1*	⟳	Converged: For (8c85a17ed7734f96a99a... af97533493f85) alarm type, if (5) (1) alarms appear within (8c85a17ed7734f96a99a... af97533493f85)告警类型 and Same主机) minute(s), (Skip after success).
breaking	09-05 13:08:42 +0800	0秒	idle pool	Idle machine	10.178.106.1*	⟳	Converged: For (8c85a17ed7734f96a99a... af97533493f85) alarm type, if (5) (1) alarms appear within (8c85a17ed7734f96a99a... af97533493f85)告警类型 and Same主机) minute(s), (Skip after success).

图 8. 在 B 业务下测试收敛规则

## 企业微信审批接入流程

### 情景

在故障替换等高危场景中，微信审批功能可以把控风险。

### 前提条件

- 企业微信号一个，[点击注册](#)。
- 蓝鲸故障自愈 APP 已经正常运行。
- 外网域名一个，能代理访问到故障自愈 APP weixin api(可通过 nginx,apache 等)，下面以 mycompany.com 做示例。
- 若是已有企业微信，注意需要企业微信管理员才能进入企业微信后台。

### 操作步骤

- 

#### 1. 配置企业微信

-

## 1. 配置故障自愈

•

### 1. 自愈测试

微信审批分两种，一种是回复指令审批，适合组合套装中的【通知或审批】套餐，还有一种需要在微信页面审批(通过菜单进入)，适合告警收敛的【异常防御需审批】。这些都需要一个企业微信应用为载体，下面以故障自愈应用为例，指引接入审批流程，相关参数可按照配置自行修改。

## 配置企业微信

### 新建应用

- 入口: 企业微信后台管理页面 -> 应用中心 -> 自建应用-> 新建应用
- 操作步骤: 点击创建应用，选择消息型应用，上传自己 LOGO，填写应用名称，请根据人员选择可见范围，点击提交完成。
  - 名称填写: 故障自愈
  - 功能介绍: 一套帮助业务全自动的发现告警、分析告警、自动恢复故障的服务。

截图示例如下：

## 创建新应用



建议上传图片尺寸为640x640，大小不超过5M。

应用名称

故障自愈

功能介绍

一套帮助业务全自动的发现告警、分析告警、自动恢复故障的服务。 |

应用可见范围 ①



提交

注意：可见范围，请务必保证审批管理员，业务管理员可见。

### 配置回调参数

此模式适用指令回复审批

- 入口： 应用中心 -> 故障自愈 -> 模式选择|回调模式

- 操作步骤：在对应栏填写下面的默认值即可
  - URL: http://mycompany.com/o/bk\_fta\_solutions/wechat/entry/
  - Token: FTAToken
  - EncodingAESKey: FTAEncodingAESKeyFTAEencodingAESKey923456781

截图示例如下：

### 故障自愈-回调模式

开启回调模式，请先填写接口配置信息  
请填写接口配置信息，此信息需要你拥有自己的服务器资源。  
填写的URL需要正确响应微信验证 URL的请求，具体说明请阅读[接口文档](#)。

URL

Token  
 [随机获取](#)

EncodingAESKey  
 [随机获取](#)

[保存](#)

“

注意：上面 TOKEN 和 EncodingAESKey 都是默认参数，请在 APP 配置后，务必修改 Token 和 EncodingAESKey

”

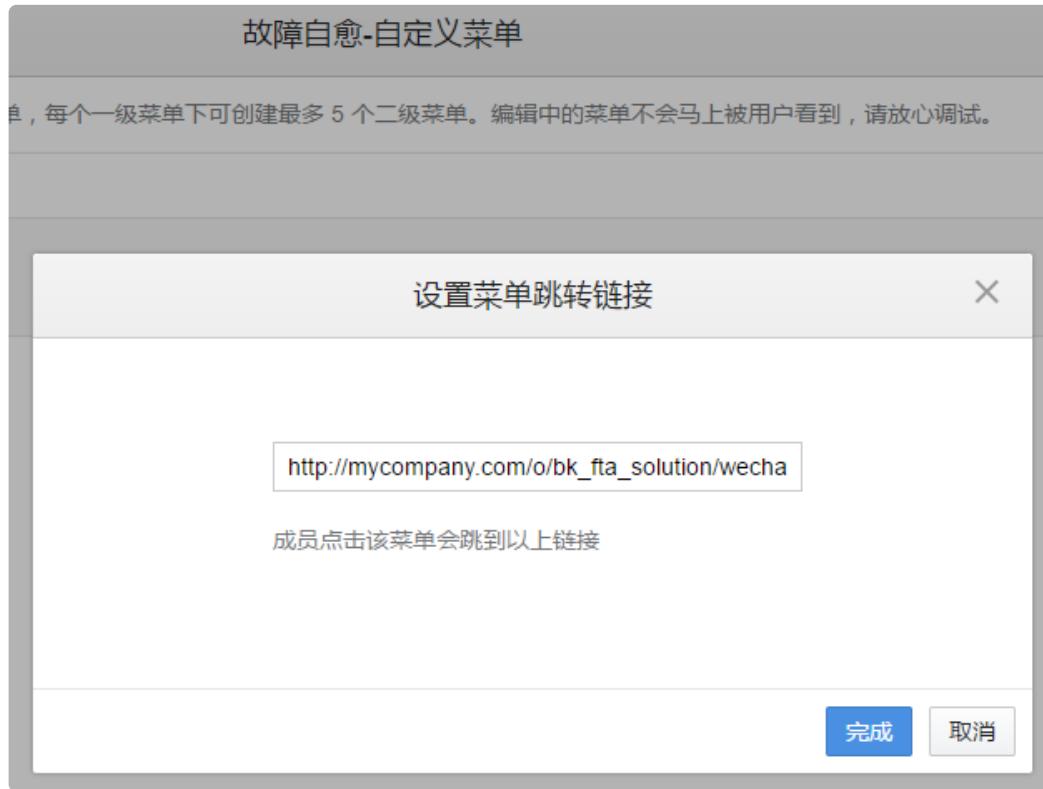
## 配置审批菜单

此模式适合收敛审批

- 入口：应用中心 -> 故障自愈 -> 模式选择|回调模式 -> 自定义菜单|设置
- 操作步骤：

- 点击右边+号按钮，输入名称后，事件类型为调整到网页，输入跳转地址完成，
- 点下右下方保存按钮后，发布即可，菜单会在 5 分钟内生效。
- 菜单名称：审批列表
- 跳转链接: [http://mycompany.com/o/bk\\_fta\\_solutions/wechat/todo/](http://mycompany.com/o/bk_fta_solutions/wechat/todo/)

截图示例如下：



完成后，企业微信接入就完成了，下一步需要把生成的 token 等配置到自愈 APP 中。

## 配置故障自愈

配置完 APP 后，才能发送审批消息，对于一些特殊处理的静态资源路径，API 路径，也需要在这里配置。

- 入口：admin 页面：[http://mycompany.com/o/bk\\_fta\\_solutions/doc/wechat\\_config/](http://mycompany.com/o/bk_fta_solutions/doc/wechat_config/)
- 也可以通过后台 admin 页面右上角->微信审批配置进入。
- 配置页面示例如下：

故障自愈后台管理

### 微信审批配置

微信端地址(外网可访问)	<input type="text" value="http://[REDACTED]/wechat/"/>
微信端静态资源地址(外网可访问)	<input type="text" value="/static/wechat/[REDACTED]"/>
TOKEN	<input type="text" value="[REDACTED]"/>
EncodingAESKey	<input type="text" value="[REDACTED]"/>
微信企业号ID	<input type="text" value="[REDACTED]"/>
微信企业号Secret	<input type="text" value="[REDACTED]"/>
发送消息的AGENT_ID	<input type="text" value="14"/>
审批管理员	<input type="text" value=""/>

下面对每个配置项详解。

- 微信端地址(外网可访问): 填写外网能访问的域名, url 到 wechat/结束, 如上面的域名应该填写:

“

http://mycompany.com/o/bk\_fta\_solutions/wechat/

”

- 微信端静态资源地址(外网可访问): 默认即可, 如果 nginx 做了路径映射, 或者使用 CDN, 需要填写绝对路径, 如:

“

/static/wechat/ (默认) http://mycompany.com/o/bk\_fta\_solutions/static/wechat/ (绝对路径, 适合 nginx 做了路径映射, 或者 CDN 场景)

”

- TOKEN 和 EncodingAESKey: TOKEN 对应第二步中，配置回调参数中的 Token。注意，如果这里修改，在上面配置也需要同步修改 EncodingAESKey 对应第二步中，配置回调参数中的 EncodingAESKey，注意，如果这里修改，上面配置也需要同步修改，长度固定为 43 个字符
- CorpID 和 Secret 需要在企业号中获取，入口在设置->权限管理中，如果没有，新建一个管理组即可。

权限管理		
管理组名	故障自愈	修改
管理员	暂未设置管理员	修改
应用权限	具有发消息和管理权限 故障自愈 ( ONLINE ) 故障自愈 故障自愈-企业版 故障自愈测试	修改
通讯录权限	具有管理权限 	修改
CorpID	<input type="text" value=""/>	
Secret	<input type="text" value=""/>	重置

删除分组

- 微信消息的 Agent\_ID: Agent\_ID 在创建完企业号应用就可以获取到，在应用中心->故障自愈，进入即可看到

故障自愈

---

应用Logo



[修改](#)

---

应用名称	故障自愈	<a href="#">修改</a>
应用类型	<input checked="" type="checkbox"/> 消息型应用	
应用ID	14	
应用介绍	一套帮助业务全自动的发现告警、分析告警、自动恢复故障的服务。	<a href="#">修改</a>
可见范围	<input checked="" type="checkbox"/> 腾讯云版 <input type="checkbox"/> fta	<a href="#">修改</a>
应用管理组	故障自愈	
PaaS管理组		

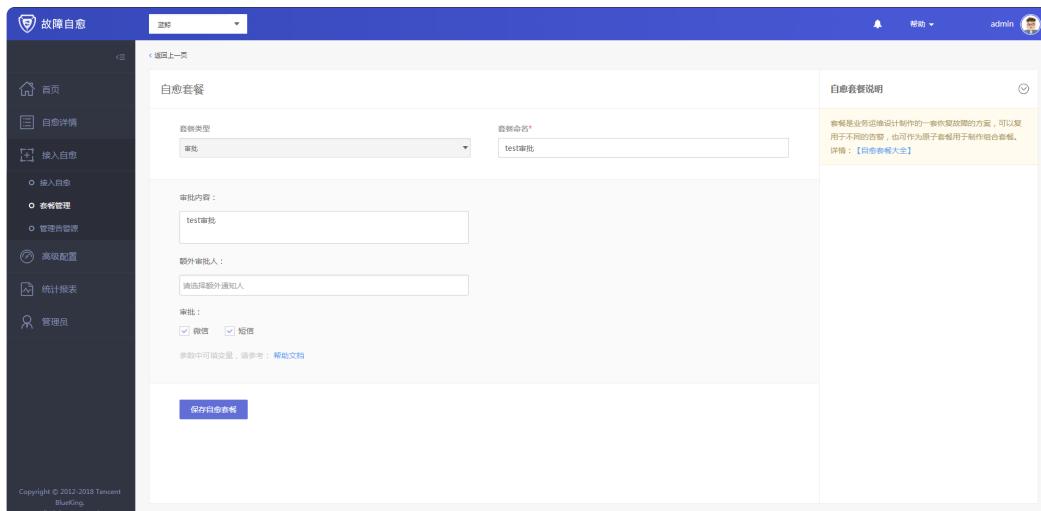
---

- 审批管理员 审批管理员是一个组超级用户，可以接受到任意审批消息，也可以审批任意的收敛审批。填写对应的名称，以逗号分隔即可。

“ 注意，名称是已经在企业微信注册的用户。 ”

## 自愈测试

- 创建审批套餐



故障自愈

自愈套餐

自愈套餐说明

套餐类型：重发  
套餐命名：

审批内容：

额外审批人：

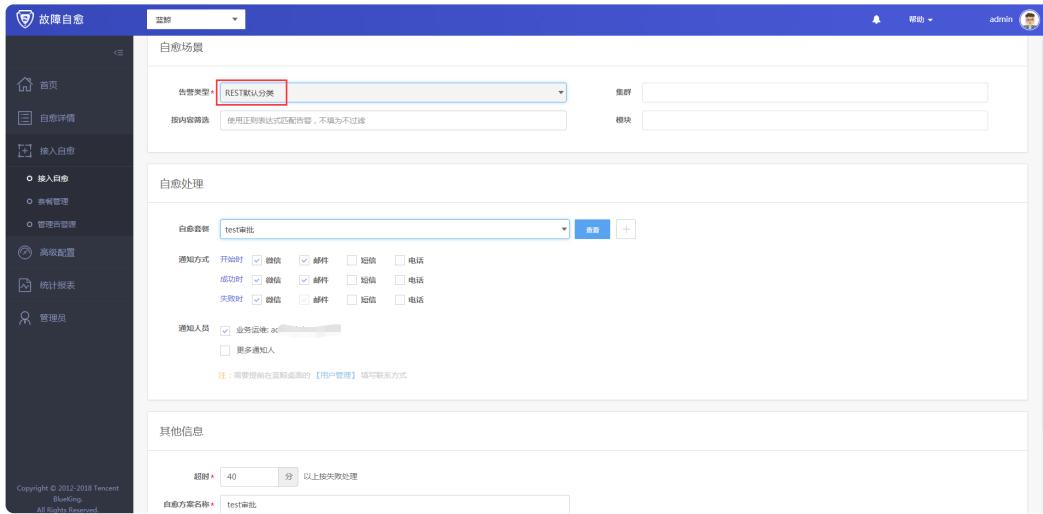
审批：  
 微信  短信

参数中可设置变量，参见：[帮助文档](#)

保存自愈套餐

Copyright © 2012-2018 Tencent BlueKing. All Rights Reserved.

- 在接入自愈流程中，告警类型选择 REST 默认分类，这里选择 REST 默认分类是为了方便触发告警，实际使用过程中，请根据自己实际需求选择告警类型。



- 触发告警

完整流程请参照 [REST API 推送](#)。

- 审批套餐执行详情

2018-12-07 - 2018-12-07		0 次收敛事件	共 4 次自愈	刷新	导出IP	请输入IP	
类型	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果
test3	12-07 20:57:10 +0800	48秒	空闲机池	故障机	10.184	<span style="color: green;">(✓)</span>	同意 (微信审批)
test3	12-07 19:44:10 +0800	21分钟	空闲机池	故障机	10.184	<span style="color: red;">(✗)</span>	审批超时
test3	12-07 19:33:10 +0800	20分钟	空闲机池	故障机	10.184	<span style="color: red;">(✗)</span>	审批超时
test3	12-07 19:31:10 +0800	20分钟	空闲机池	故障机	10.184	<span style="color: red;">(✗)</span>	审批超时

微信审批，降低高危告警处理的风险。

## 带审批 Ping 告警组合套餐接入流程

### 情景

产生 Ping 告警，服务器故障，首先要做故障检查(再次 Ping 一次，或者调用一次)，发送审核通知(包含故障检查的内容)，用户审核即可重启。(适用保守的传统企业)

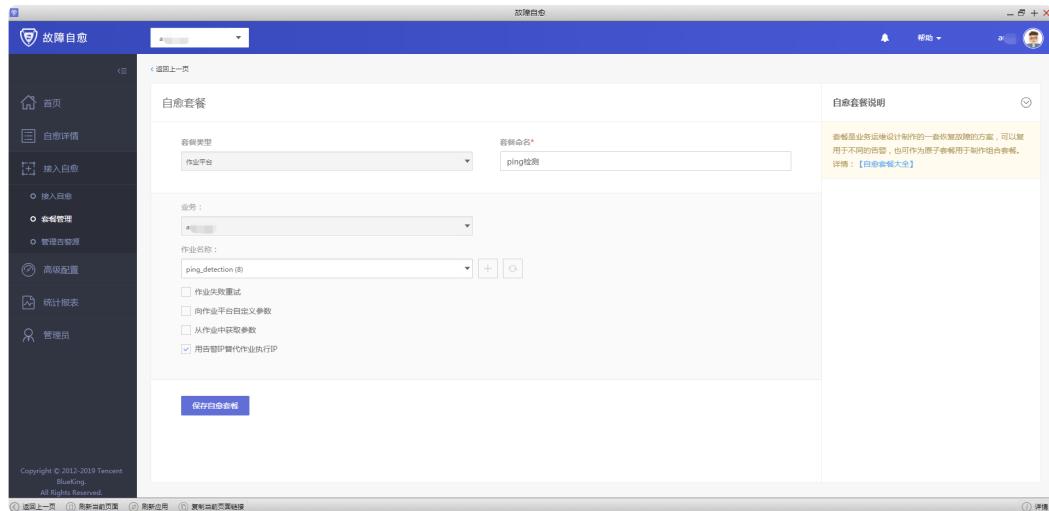
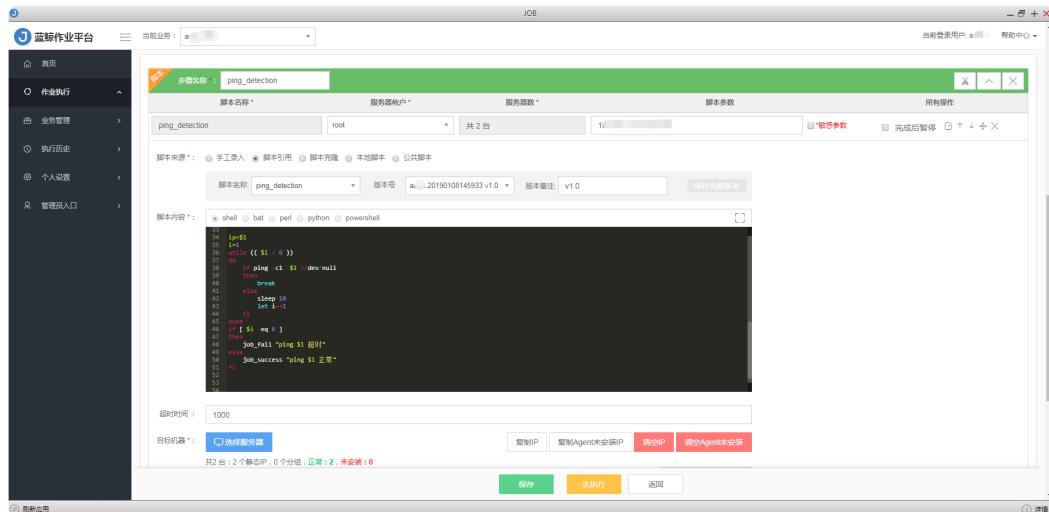
### 前提条件

- 需要先配置企业微信，注册链接：[企业微信首页](#)，注意：开启微信端口 80443
- 蓝鲸故障自愈 APP 已经正常运行 创建故障自愈 APP 请参照[微信审批接入流程](#)。

## 准备好了组合套餐中每个原子(节点)的套餐

### 配置 Ping 检测的原子套餐

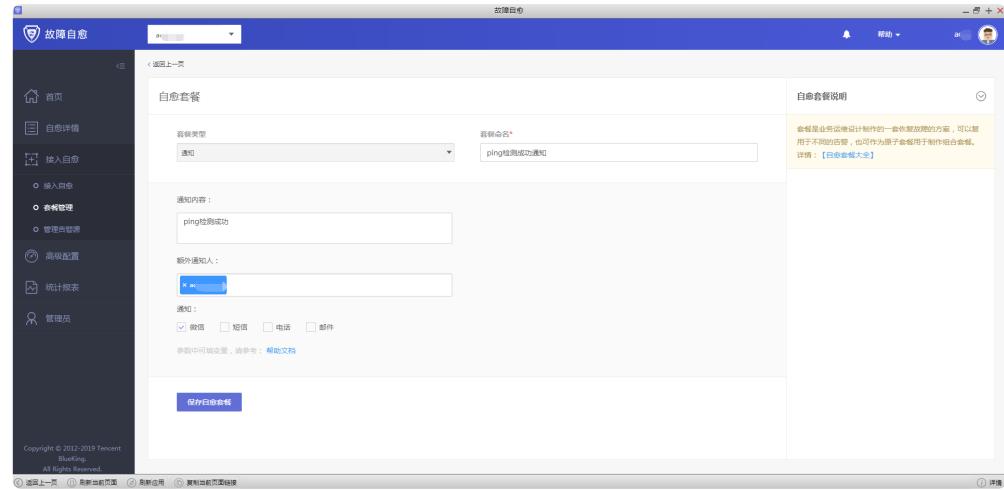
在作业平台写个简单的 Ping 检测脚本，再去故障自愈中配置 Ping 检测的自愈套餐。



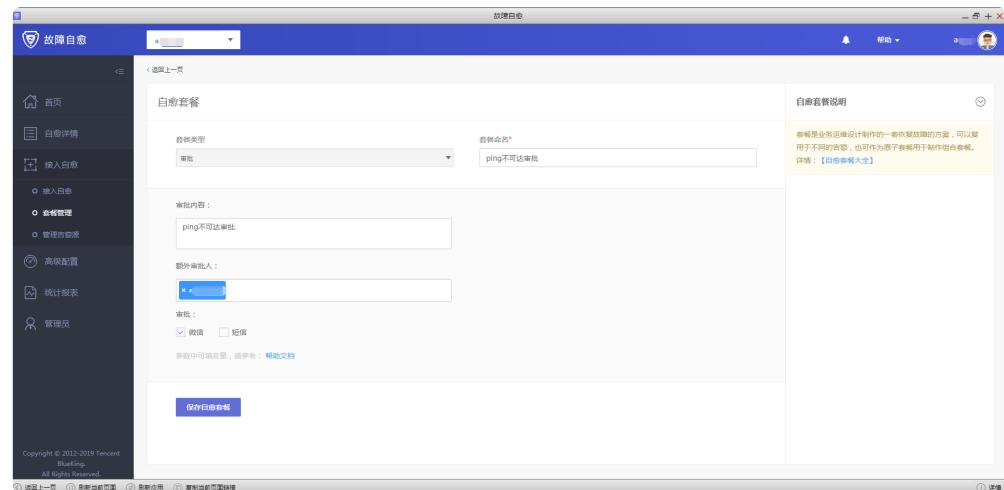
### Ping 检测没有异常，则发送正常通知

如 Ping 检测异常，则发送审批通知，用户审批通过即可重启，审批如果不通过，则发送审批失败通知。

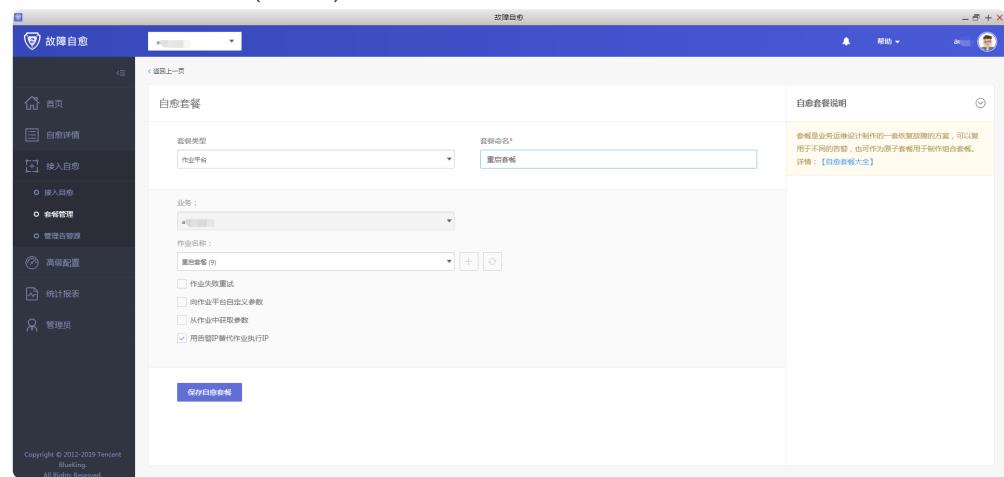
- 配置 Ping 检测正常通知



- 配置审批套餐

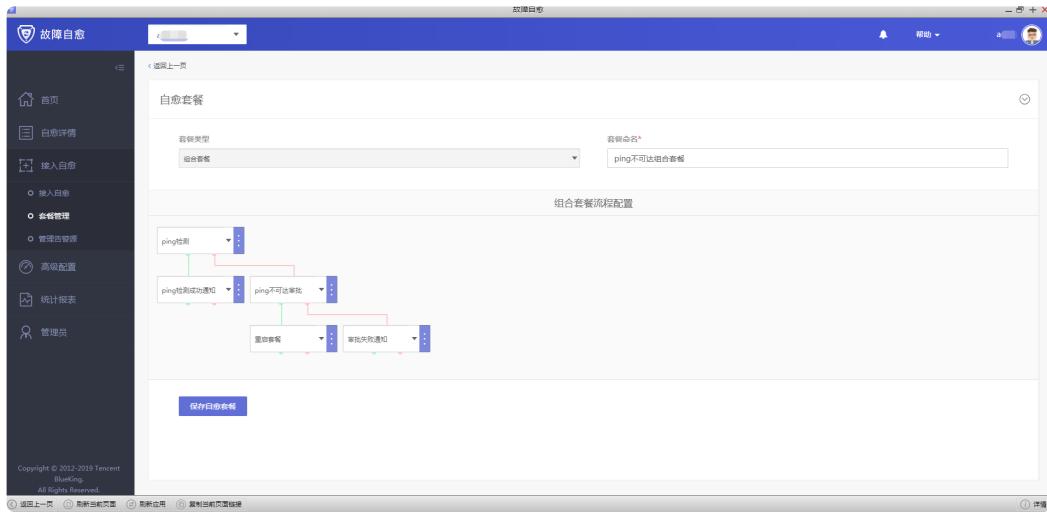


- 配置重启套餐，因为是简单模拟，所以只是 echo 了一下重启成功，实际应用时要调用在 ESB 上注册重启服务器接口(物理机)，如果是虚拟机，需要自己写脚本来调用虚拟机的接口

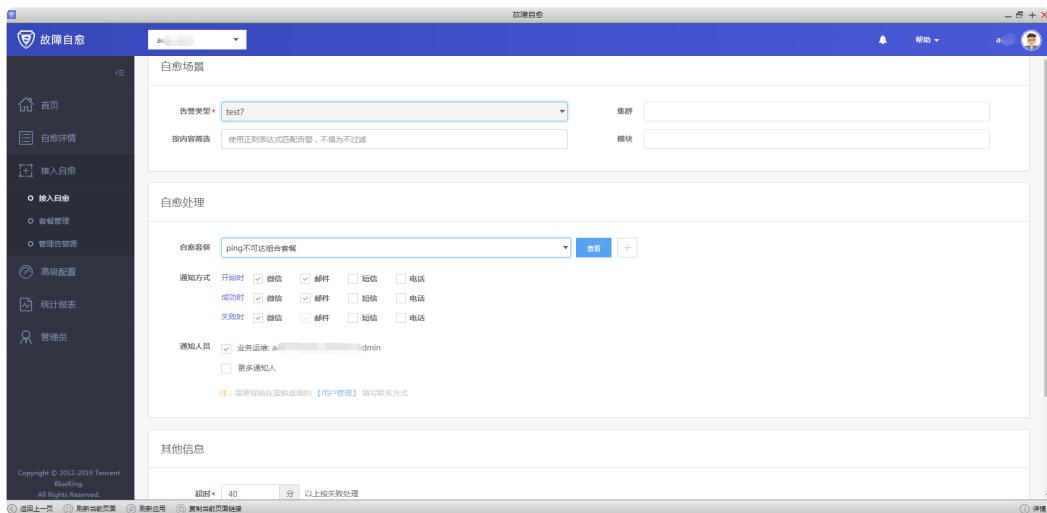


## 配置组合套餐，并接入故障自愈

接入故障自愈这里选择 REST 默认分类是为了方便触发告警，实际应用选择 Ping 不可达告警类型。



(Ping 告警组合套餐)



(接入自愈)

## 触发告警，完成自愈

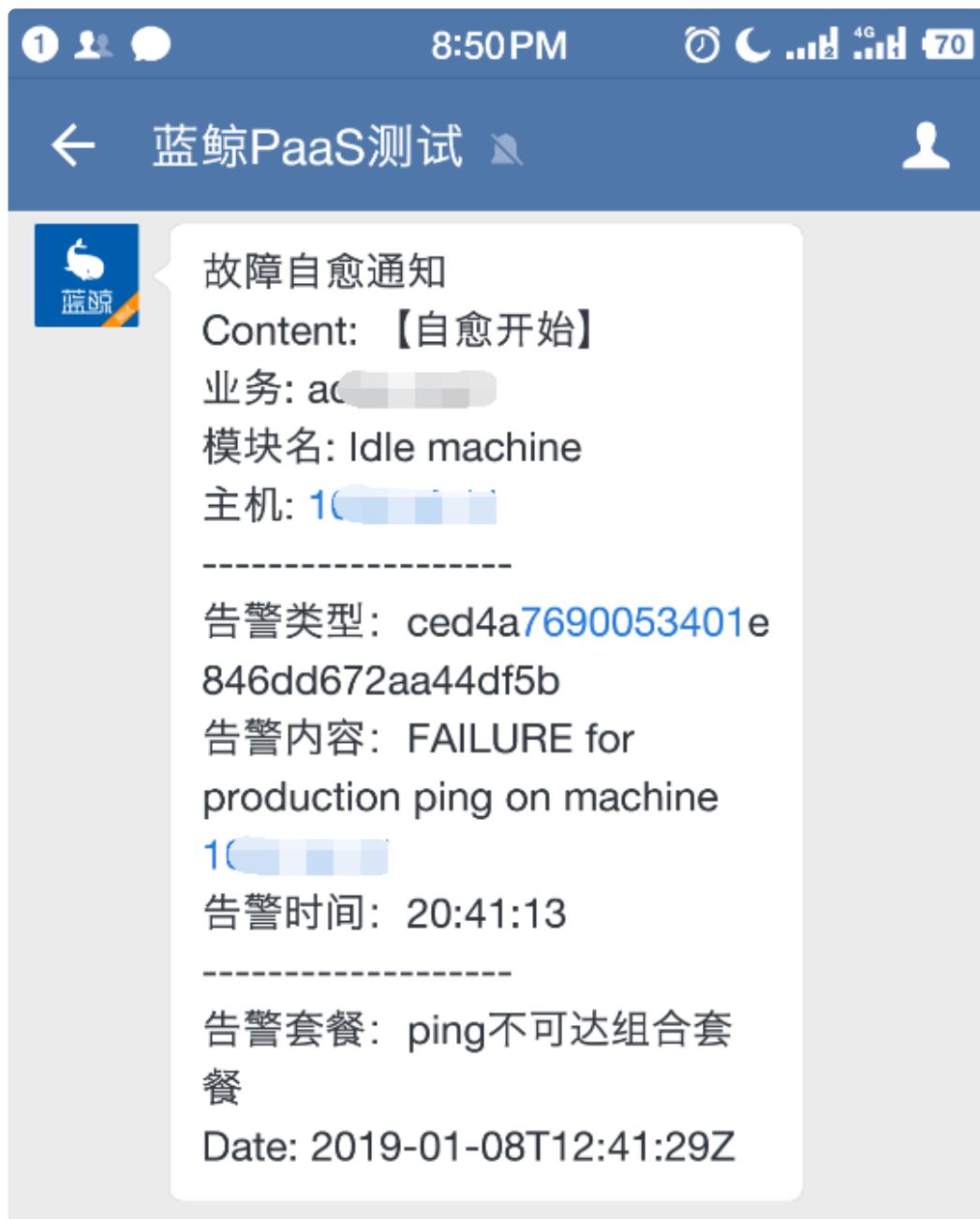
### 触发告警

由于这里是做测试，就不拿生产环境了，用 REST API 可以更方便的产生告警，完整流程请参照 [REST API 推送](#)。

### 审核

Ping 检测没有异常，则发送正常通知。如 Ping 检测异常，则发送审批通知，用户审批通过即可重启，审批如果不通过，则发送审批失败通知。

- Ping 检测异常，发送审批通知，用户审批通过即可重启





### 故障自愈通知

Content: 【故障自愈】即将执行 ping不可达审批,请审核!

同意请回复: TY [2774997\\_2](#)

驳回请回复: BH [2774997\\_2](#)

Date: 2019-01-08T12:42:31Z

下午 8:47

[TY 2774997\\_2](#)



### 故障自愈通知

【故障自愈】审批成功

Date: 2019-01-08T12:47:40Z



## 故障自愈通知

Content: 【自愈成功】

业务: a[REDACTED]

模块名: Idle machine

主机: 1[REDACTED]

-----  
告警类型: ced4a7690053401e

846dd672aa44df5b

告警内容: FAILURE for  
production ping on machine

1[REDACTED]

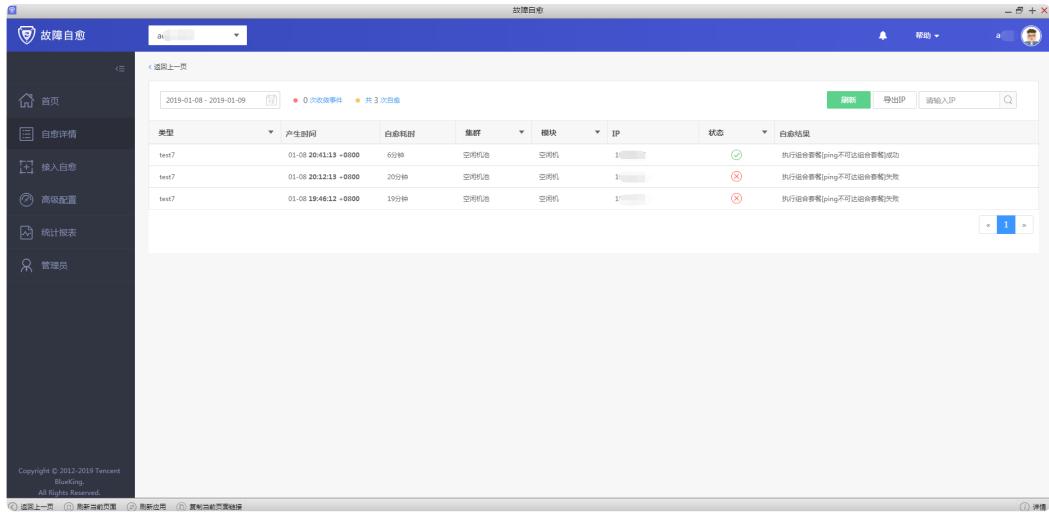
告警时间: 20:41:13

-----  
告警套餐: ping不可达组合套餐

Date: 2019-01-08T12:48:15Z



- 回到故障自愈中，查看自愈详情



## 组合套餐在故障替换场景中的应用

### 情景

场景：A 模块是重要模块，出现 Ping 不可达告警，首先要校验 A 模块是否真的故障，如果真的故障，接下来是从资源池中获取备机 ... 故障替换等等，期间每个环节都有可能出错，那就要考虑异常分支的场景。

### 前提条件

- 需要先配置企业微信，注册链接：[企业微信首页](#)，注意：开启微信端口 80443
- 蓝鲸故障自愈 APP 已经正常运行 创建故障自愈 APP 请参照 [微信审批接入流程](#)。

### 准备好组合套餐中每个原子(节点)的套餐

#### 配置 Ping 检测的原子套餐

在作业平台写个简单的 Ping 检测脚本，再去故障自愈中配置 Ping 检测的自愈套餐。

The screenshot shows the BlueKing Job interface. A step named 'ping\_detection' is being configured. The step details include:

- 脚本名称:** ping\_detection
- 服务账户:** root
- 服务端数:** 共 2 台
- 脚本参数:** 无
- 脚本来源:** 手工录入 (selected)
- 脚本内容:**

```

13 #!/bin/sh
14
15 #!/bin/sh
16 while (( $i < 6 ))
17 do
18     ping -c 1 $1 >/dev/null
19     if [ $? -eq 0 ]
20     then
21         break
22     else
23         sleep 10
24         let i=$i+1
25     fi
26 done
27 if [ $i -eq 6 ]
28 then
29     job_fail "ping $1 失败"
30 else
31     job_success "ping $1 正常"
32 fi
33

```
- 超时时间:** 1000
- 目标机器:** 选择部署列表 (selected)
- 操作按钮:** 保存, 执行, 返回

The screenshot shows the BlueKing Self-healing interface. A new recipe is being created with the following details:

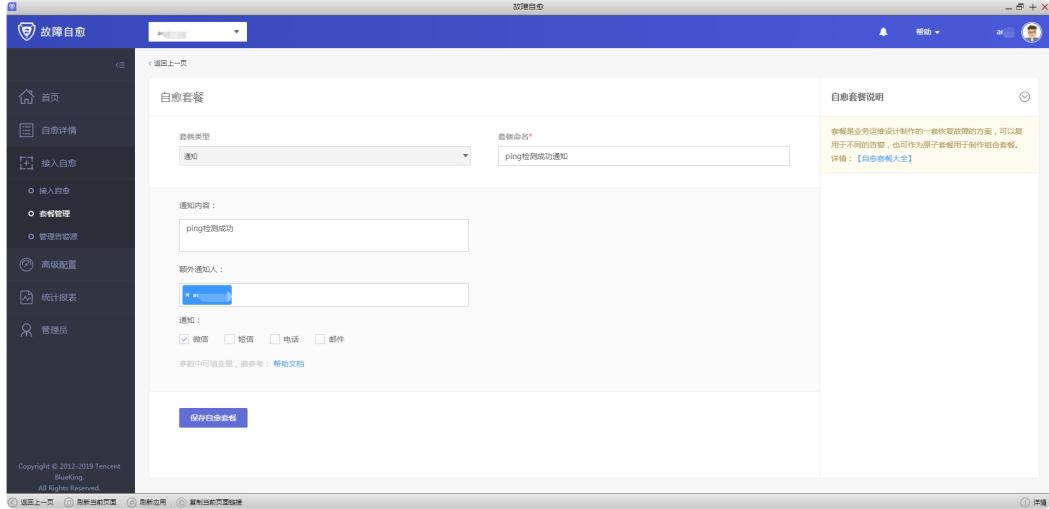
- 套餐类型:** 作业平台
- 套餐命名:** ping检测
- 业务:** a... (selected)
- 作业名称:** ping\_detection (3) (selected)
- 高级配置:**
  - 作业失败重试
  - 向作业平台自定义参数
  - 从作业中获取参数
  - 用备机IP替代作业执行IP

**保存白名单套餐**

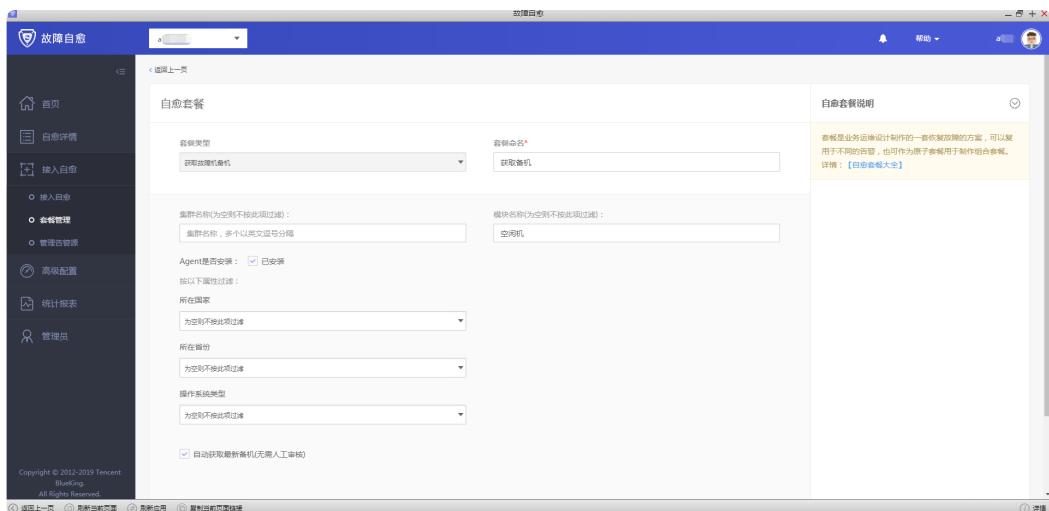
## 通知和获取备机

Ping 检测没有异常，则发送正常通知。如 Ping 检测异常，则使用获取备机套餐，自动获取备机，前提是空闲机池中有空闲机。

- 配置 Ping 检测正常通知



### ● 配置自动获取备机套餐



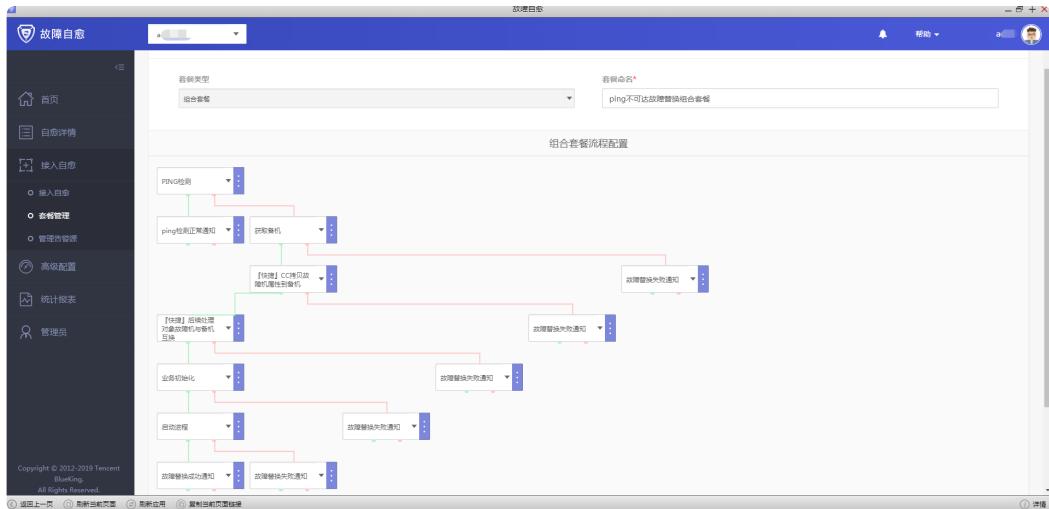
### 拷贝故障机属性到备机

成功获取备机后，拷贝故障机属性到备机，后续处理对象故障机与备机互换，然后初始化业务，启动进程通知故障替换成功，以上步骤失败都加一个失败通知

- 『快捷』 CC 拷贝故障机属性到备机、『快捷』后续处理对象故障机与备机互换，都是快捷套餐，只要选择就好，这里就不展开了。后面初始化业务请根据企业的初始化流程来配置初始化套餐，启动进程也是一样，因为这里只是模拟所以仅用通知代替。

## 配置组合套餐，并接入故障自愈

接入故障自愈这里选择 REST 默认分类是为了方便触发告警，实际应用选择 Ping 不可达告警类型。



## 触发告警，完成自愈

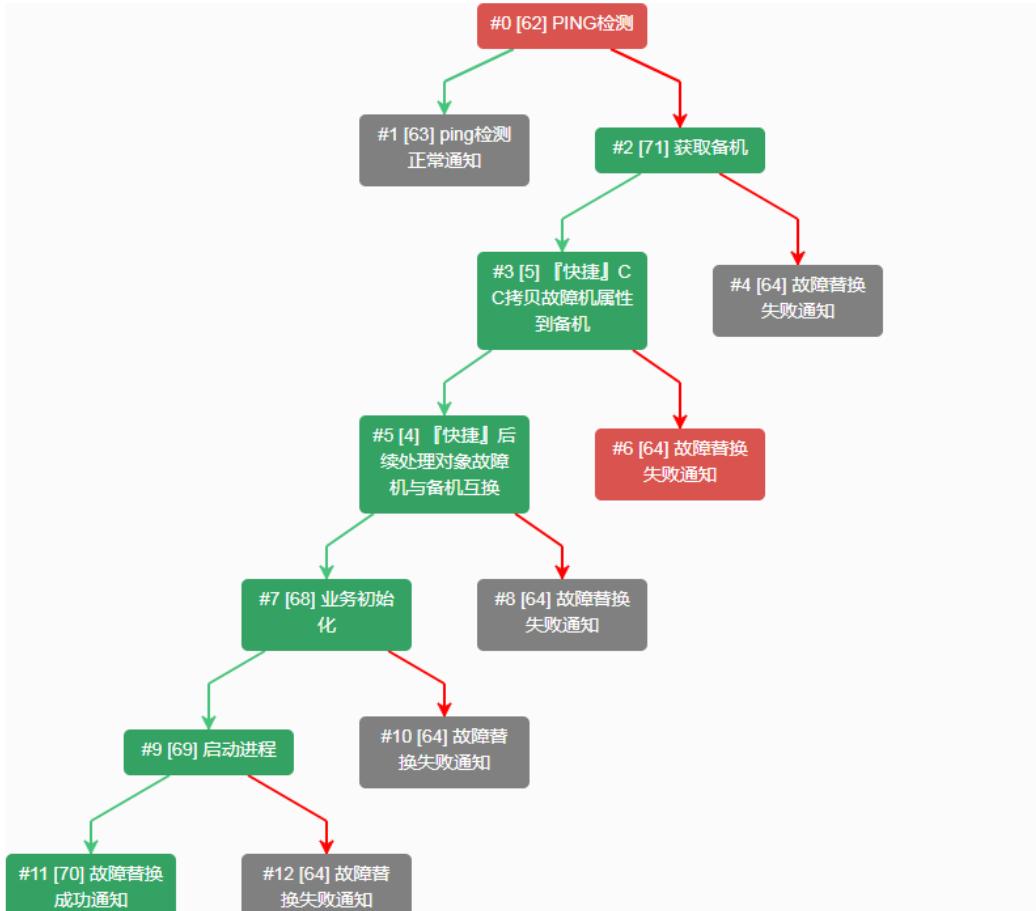
- 触发告警，由于这里是做测试，就不拿生产环境了，用 REST API 可以更方便的产生告警，完整流程请参照[REST API 推送](#)。
- 回到故障自愈中，查看自愈详情，也可以点击状态，查看执行详情

```

12:03:07 #0 PING检测 | 失败: 执行Job作业失败[10]
12:03:09 #2 获取备机 | 成功: 成功获取备机1
12:03:11 #3 『快捷』CC拷贝故障机属性到备机 | 成功: 执行蓝鲸cc组件clone_host_property任务成功
12:03:13 #5 『快捷』后续处理对象故障机与备机互换 | 成功: 替换操作对象为备机
12:03:15 #7 [REDACTED] 业务初始化 | 成功: 发送通知成功
12:03:38 #9 启动进程 | 成功: 执行Job作业成功[1C]
12:03:40 #11 故障替换成功通知 | 成功: 发送通知成功

处理状态 成功 执行组合套餐[ping不可达故障替换组合套餐]成功
操作 重试整个流程

```



## 上下文传参

上下文传参一般用于组合套餐中，将上一个节点的输出作为下一个节点的输入。

“

目前支持将作业平台套餐的输出作为任意下一个节点(作业平台、通知、标准运维等)的输入。

”

## 使用场景

故障自愈的默认套餐 **『快捷』发送CPU使用率TOP10的进程(微信)**、**『快捷』发送内存使用率TOP10的进程(微信)** 实际上是一个使用上下文传参的组合套餐。

- 第 1 步：创建输出 CPU、内存使用率的 TOP10 的作业
- 第 2 步：创建作业平台套餐
- 第 3 步：创建 **通知** 套餐，引用上一步输出的结果，以微信的方式发送出来
- 第 4 步：使用组合套餐，将两个套餐串起来

## 在作业平台中定义传递的参数

在作业平台的作业中按照如下格式将传递的参数定义好

```
echo "FTAARGV 变量:值"
```

以 **『快捷』发送内存使用率TOP10的进程(微信)** 套餐中的第 1 步的脚本为例：

```
#!/bin/bash
#
#           USAGE: ./get_top_proc_mem.sh <cpu|mem> <number>#
#   DESCRIPTION: 输出系统当前占用资源(cpu、内存)最多的TopN进程
#=====
set -o nounset                                # Treat unset variables as an error

fields="pcpu,pmem,comm"

usage() {
    cat <<EOF
    get_top_proc_in_oneline.sh <cpu|mem> <number of top proc>
EOF
    exit 1
}

join() {
    local IFS="$1"
    shift; echo "$*"
}

if (( $# < 1 )) || (( $# > 2 )); then
    usage
fi

case $1 in
    cpu) sort_field=pcpu ;;
    mem) sort_field=rss ;;
    *) usage ;;
esac

if [[ $# -eq 2 ]]; then
    top_n=$2
else
    top_n=6
fi

return _mem=$(ps -eo "$fields" --sort=-"$sort_field" | head -$(( top_n + 1 )) | awk 'NR=
```

```
=1 { gsub(/%/,"") } {printf "%s\\n", $0 }')

echo "FTAARGV return_mem:${return_mem}"
```

“

上面示例中，`return_mem` 为传递给下一个原子节点的变量， `${return_mem}` 为变量的值。

”

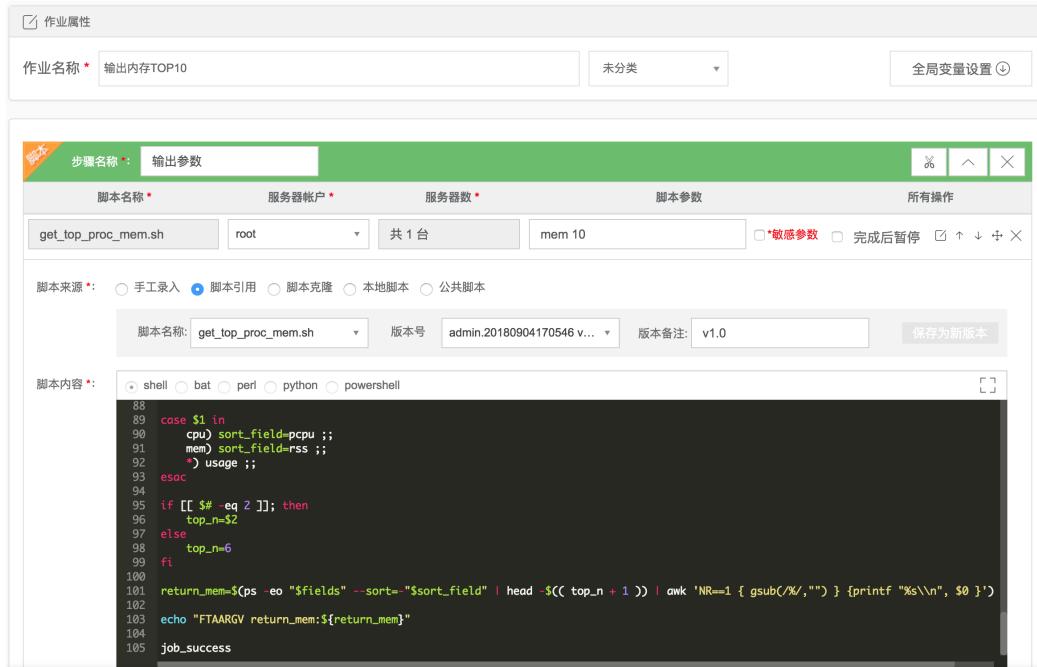


图 1. 创建输出传参(传递给下一个步骤的参数)的作业

## 创建作业平台套餐

创建自愈套餐，套餐类型选择作业平台，选择上一步创建的作，并勾选 [从作业中获取参数](#)

自愈套餐

套餐类型	作业平台	套餐命名*	1.输出内存TOP10
业务：	蓝鲸		
作业名称：	输出内存TOP10 (1026)	+	
<input type="checkbox"/> 作业失败重试			
<input type="checkbox"/> 向作业平台自定义参数			
<input checked="" type="checkbox"/> 从作业中获取参数			

( 在作业脚本中 , 将变量单独输出一行 , 详细的说明请参考 : [获取作业执行结果说明](#) )

用告警IP替代作业执行IP

图 2. 创建输出传参(传递给下一个步骤的参数)的套餐

## 创建使用上一步输出参数的套餐

创建通知套餐或作业平台套餐来使用上一步输出的参数。

“

如果你没配置通知套餐，也可以通过作业平台套餐来测试.

”

自愈套餐

套餐类型	套餐命名*
<input type="text" value="通知"/>	<input type="text" value="2.内存使用率TOP10微信通知"/>
通知内容： ● return_mem为变量， 其他为固定格式 <code> \${bpm_context jobs_return_mem}</code> ● 引用上一步输出的变量	
额外通知人： <input type="text" value="请选择额外通知人"/>	
通知： <input checked="" type="checkbox"/> 微信 <input type="checkbox"/> 短信 <input type="checkbox"/> 电话 <input type="checkbox"/> 邮件	
参数中可填变量，请参考： <a href="#">帮助文档</a>	

图 3. 创建通知套餐，引用上一步输出的参数

```

28 job_start
29
30 ##### 可在此处开始编写您的脚本逻辑代码
31 ##### 作业平台中执行脚本成功和失败的标准只取决于脚本最后一条执行语句的返回值
32 ##### 如果返回值为0，则认为此脚本执行成功，如果非0，则认为脚本执行失败
33
34
35 echo $1
  ● 打印作业平台套餐中传递给该步骤的参数

```

图 4. 创建作业，引用作业平台套餐传递的参数

自愈套餐

套餐类型	套餐命名*
作业平台	3.打印上一步获取的参数
业务：	蓝鲸
作业名称：	打印上一步获取的参数 (1027)
<input type="checkbox"/> 作业失败重试 <input checked="" type="checkbox"/> 向作业平台自定义参数	
<span style="color: red;">●</span> return_mem 为变量，其他为固定格式	
echo_fta_arg.sh    "\${bpm_context jjobs_return_mem}"	
<span style="color: red;">●</span> 引用上一步输出的变量	
<input type="checkbox"/> 从作业中获取参数 <input checked="" type="checkbox"/> 用告警IP替代作业执行IP	

图 5. 创建作业平台套餐，引用上一步输出的参数

## 创建组合套餐

通过组合套餐，将输出参数 和 使用参数 的原子套餐串起来。

自愈套餐

套餐类型	套餐命名*
组合套餐	4.输出内存TOP10
组合套餐流程配置	
1.输出内存TOP10 2.内存使用率TOP1 0微信通知	

图 6. 创建组合套餐

## 测试

如果故障自愈不是很方便获取监控的告警，可以使用 **REST API 推送** 方式，来验证故障自愈的执行。

## 作业间传参效果

类型	产生时间	自愈耗时	集群	模块	IP	状态	自愈结果
fta_mix_2	09-05 10:36:15 +0800	44秒	公共组件故障自愈	consul,rabbitmq	192.168.1.10	成功	执行组合套餐[5.打印上一步获取的参数]...

图 7. 模拟自愈，执行上一步创建的组合套餐(包含上下文传参特性)



图 8. 该自愈的执行详情

This screenshot shows the execution log for step #0. At the top, it says '执行成功(1)' (Execution successful) and '查看步骤详情' (View step details). Below is a search bar with '输入搜索内容' (Input search content) and buttons for '搜索日志' (Search log) and '导出执行日志' (Export execution log). The main area shows a table for IP filtering and a log output. The log output shows three log entries: '[2018-09-05 10:39:24][PID:5283] job\_start', '[2018-09-05 10:39:24][PID:5283] job\_start', and '[2018-09-05 10:39:24][PID:5283] job\_start'. The log also includes system information like 'FTAARGV return\_mem:CPU MEM COMMAND\n 1.3 3.1 influxd\n 0.2 1.2 systemd-journal\n 0.0 0.7 uwsgi\n 0.0 0.7 uwsgi'. At the bottom, it says '[2018-09-05 10:39:24][PID:5283] job\_success:[]'.

图 9. #0 步 输出参数

This screenshot shows the execution log for step #1. At the top, it says '输入搜索内容' (Input search content) and '查看步骤详情' (View step details). Below is a search bar with '输入搜索内容' (Input search content) and buttons for '搜索日志' (Search log) and '导出执行日志' (Export execution log). The main area shows a table for IP filtering and a log output. The log output shows three log entries: '[2018-09-05 10:39:46][PID:5811] job\_start', 'CPU MEM COMMAND\n 1.3 3.1 influxd\n 0.2 1.2 systemd-journal\n 0.0 0.7 uwsgi\n 0.0 0.7 uwsgi\n 0.0 0.7 uwsgi\n 0.0 0.7 uwsgi', and '[2018-09-05 10:39:46][PID:5811] job\_start'. At the bottom, it says '[2018-09-05 10:39:46][PID:5811] job\_start'.

图 9. #1 步 获取参数

This screenshot shows the configuration for step #1. It includes a code editor with the following content:

```
33
34
35 echo $1
```

Below the code editor, there is a section for '脚本参数:' (Script parameters) containing the command 'echo \$1'. A note states: '● 脚本参数的值，就是上一个步骤输出的结果' (The value of the script parameter is the result of the previous step's output). At the bottom, there is a '超时时间：' (Timeout) field set to '1000'.

图 10. #2 步 作业执行时传递参数内容

作业 + 通知 组合套餐的输出结果

【故障自愈】内存使用率TOP10  
列表：  
业务： "故障自愈测试环境"  
模块： "故障机"  
主机： 10.10.5.100.50  
-----  
CPU MEM COMMAND  
0.1 2.8 AgentWorker  
0.0 2.1 secu-tcs-agent  
0.1 1.5 sap1005  
0.1 1.4 sap1009  
0.0 1.3 tuned  
0.0 1.2 barad\_agent  
0.0 0.9 barad\_agent  
0.0 0.6 barad\_agent  
0.0 0.3 basereport  
0.0 0.2 rsyslogd

图 11. 作业+通知 组合套餐的输出结果

注：如果参数仅在作业平台中传递，可以使用 [作业平台的上下文传参功能](#)。

# 套餐内置变量

自愈的很多套餐里都需要传入参数，但有些参数是套餐执行时根据故障机信息进行动态获取的。

目前自愈里很多套餐类型都支持了这种变量参数，比如通知审批等。

## 场景示例

B 机器是 VMM 的管理端，A 机器是 VMM 中的应用服务器。当 A 机器上发生告警时，可以通过调用 B 机器上管理端的接口来完成自愈

The screenshot shows the '愈套餐' (Recovery Recipe) configuration page. It includes fields for '套餐类型' (Recipe Type) set to '作业平台' (Job Platform), '套餐命名\*' (Recipe Name) set to 'bk-机器A告警，去B机器执行', '业务' (Business) set to '蓝鲸', '作业名称' (Job Name) set to 'bk-机器A告警，去B机器执行 (3087)', and checkboxes for '向作业平台自定义参数' (Customize parameters for job platform) and '从作业中获取参数' (Get parameters from job). A note '记得取消勾选' (Remember to uncheck) is present next to the '用告警IP替代作业执行IP' (Replace job execution IP with alarm IP) checkbox. The variable \${ip} is highlighted with a red arrow pointing to the '故障机的IP' (Fault machine's IP) placeholder in the 'a.sh' script field.

图 1. 套餐中引用变量

## 目前可选的变量有(可能为空)

- \${ip}: 告警的 IP
- \${raw}: 告警的字符内容
- \${alarm\_type}: 告警类型

- \${source\_time}: 告警时间(格式如 2014-01-01)
- \${cc\_biz\_id}: CMDB 的业务 ID
- \${operator}: 业务负责人的第一个

## CMDB 主机属性和 SET 属性的变量

格式为: \${cc|属性名字}, 如:

- \${cc|OuterIP}: 故障主机的外网 IP
- \${cc|AssetID}: 故障主机的固资编号 具体属性名请在 CMDB 上查询。

## CMDB 变量支持五个参数

- all、set、custom、alarm\_ci\_name、ip\_bak
- all: 当有多个参数的时候, 将返回通过逗号间隔的字符串。如有多个主机名称的时候
- \${cc|HostName|all}: 返回"hostname1,hostname2,hostnameN", 不添加默认返回第一个 set:  
查询 Set 属性。如
- \${cc|SetName|set}: 故障主机的 Set 名称 custom: 查询自定义属性。如查询一个名为 IDC 的  
Set 属性
- \${cc|IDC|set|custom}
- alarm\_ci\_name: 指定查询故障机的 CMDB 属性
- ip\_bak: 指定查询备机的 CMDB 属性

三个参数能任意组合, 如以下两个写法是等价的:

- \${cc|IDC|set|custom|all}
- \${cc|IDC|all|set|custom}

## 故障机替换时备机

故障自愈现在有两种获取备机的 IP, 通过获取备机套餐在 CC 中寻找符合要求的机器或者通过作业平台脚本来获取。获取到的备机参数为: \${bpm\_context|ip\_bak}

与 IP 有关的变量有三个:

- \${ip} 当前流程处理的 IP, 默认是故障机 IP, 可以被替换操作对象的套餐改为备机 IP
- \${bpm\_context|ip\_bak} 备机 IP

- \${bpm\_context|alarm\_ci\_name} 故障机 IP

## 常见案例

- 1.根据告警传入告警 IP
  - \${ip}
- 2.根据告警 IP 传入外网 IP
  - \${cc|OuterIP}
- 3.根据告警 IP 传入自定义主机属性
  - \${cc|gametype|custom}
- 4.根据告警 IP 传入 set 名称
  - \${cc|SetName|set}
- 5.根据告警 IP 传入自定义 set 属性
  - \${cc|openstate|set|custom}
- 6.跟进告警 IP 传入组合属性(如：1 区\_虎啸谷)
  - \${cc|SetWorldID|set}\_\\${cc|SetChnName|set}

## 注意事项

- 注意大小写

# 收敛规则大全

介绍自愈目前收敛的几种模式，并用现有的一些通用收敛规则举例说明其用法。希望收敛规则能满足你各式各样的需求，并且在配置好后，让你会忘了它的存在。

## 异常防御需审批

触发规则后，会打电话通知用户，让用户审批决定是否收敛。如果超时未审批则会收敛跳过，不处理。如果 30 分钟内发生相同规则的异常防御事件，会被汇集到同一个收敛事件中。

可用于防御大规模告警的异常，如发布未屏蔽，网络问题，机房故障等等。通过人工判断大量的告警是否需要处理。

- #13 一系列同业务的进程端口告警(可能发布变更未屏蔽):

同业务 5 分钟内出现 3 条以上进程端口则认为是异常事件需要审批(默认不开启)

- #14 一系列同机房的单机异常告警:

同机房 2 分钟内出现 10 条以上的 ping/agent/restart 告警则认为是异常事件需要审批

- #15 一系列同业务的单机异常告警:

同业务 2 分钟内出现 3 条以上的 ping/agent/restart 告警则认为是异常事件需要审批

## 触发通知

触发规则后，不影响处理，发送通知。

可用于配置阀值告知。

## 汇集相关事件

触发规则后，不影响处理，只是把满足收敛规则的告警汇集在一起展示为同一个事件。

在界面上把相关的告警汇集在一起展示，能更好自定义告警间的关联性。

- #9 掉线可能因为大区内基础告警:

对于在线告警，汇总相同大区内的基础告警

- #18 归纳同主机的一系列告警:

汇总相同主机 5 分钟内出现的告警

- #20 归纳同主机的一系列进程端口告警:

汇总相同主机 30 分钟内出现的进程端口告警

## 收敛后处理

与其他收敛规则相反。未触发规则时，配置的告警类型不处理。触发规则后，才开始处理。

可以等告警数量超过一定阀值后才处理告警。或者一定时间内同时出现 A 告警和 B 告警的时候再开始处理。

## 执行中跳过

触发规则后，如果有满足规则的其他告警正在自愈，或刚结束自愈 5 分钟，则跳过当前告警。

可用于避免重复处理。

- #12 Ping/Agent 引发的一系列告警：

对于由 ping/agent 告警引发的上层告警，如果 ping/agent 告警在处理中，则收敛

## 执行中等待

触发规则后，如果有满足规则的其他告警正在自愈，则等其他告警自愈完成后再继续处理当前告警。

可用用户互斥的告警处理，或有先后顺序依赖的告警处理。

## 成功后跳过

触发规则后，如果有满足规则的其他告警自愈成功，则跳过当前告警。失败的话则继续自愈处理。

可用于实现失败重试。

- #2 一系列单机异常类告警：

对于相同主机的[ping 超时,上报超时,系统重启,磁盘只读]告警，5 分钟内出现过且自愈成功则收敛

- #5 一系列处理套餐相同的告警：

对于相同主机相同处理的告警，5 分钟内出现过且自愈成功则收敛

## 成功后跳过，失败时审批

触发规则后，如果有满足规则的其他告警自愈成功，则跳过当前告警。失败的话则发送审批由用户判断是否继续执行自愈处理。

## 超出后汇总

触发规则后，超出数量的告警将会收敛不处理，并发送汇总通知

如果告警在一定时间内不断出现，超过某个阀值可以认为其有异常，不再自愈，触发通知。

# 用户分享案例

## 社区用户分享案例

1. 一次模拟进程宕掉的蓝鲸故障自愈测试分享
2. 蓝鲸监控服务拨测测试文档
3. 【社区案例】组合套餐：在保障业务安全的前提下实现故障处理自动化

## 蓝鲸公众号分享案例

1. 某银行的故障自愈落地案例
2. 那些年我们想做的无人值守
3. 当 Zabbix 遇见故障自愈
4. 心得分享-故障自愈：清理最早一天日志分享

# APP 部署失败

## 测试环境 APP 部署失败

后台只有正式环境，所以 APP 不支持测试环境部署

### 常见问题及排错

◆ 自愈的后台只部署了正式环境，所以APP也只支持正式环境部署

测试部署失败

```
for chunk in iterable:  
    File "/cache/.bk/env/lib/python2.7/json/encoder.py", line 434, in _iterencode  
        for chunk in _iterencode_dict(o, _current_indent_level):  
    File "/cache/.bk/env/lib/python2.7/json/encoder.py", line 408, in _iterencode_dict  
        for chunk in chunks:  
    File "/cache/.bk/env/lib/python2.7/json/encoder.py", line 408, in _iterencode_dict  
        for chunk in chunks:  
    File "/cache/.bk/env/lib/python2.7/site-packages/django/core/serializers/json.py", line 115, in default  
        return super(DjangoJSONEncoder, self).default(o)  
    File "/cache/.bk/env/lib/python2.7/json/encoder.py", line 184, in default  
        raise TypeError(repr(o) + " is not JSON serializable")  
TypeError: <djang.utils.functional._proxy__object at 0x7fdfee778a90> is not JSON serializable  
-----FAILURE: Migrate Database-----
```



图 1. 测试环境部署故障自愈 APP 失败

## 正式环境 APP 部署失败

部署后台的时候没有执行初始化数据的步骤：

```
./bkcec initdata fta //社区版
```

### 常见问题及排错



图 2. 正式环境部署故障自愈 APP 失败

## 查看依赖组件运行状况

当故障自愈出现异常时，可以访问运行状态页面，呈现故障自愈自身以及依赖的第 3 方组件的运行状态。



图 1. 故障自愈运行状态入口

“

或直接访问 URL: [https:// \\${PAAS\\_URL} /o/bk\\_fta\\_solutions/healthz/](https:// ${PAAS_URL} /o/bk_fta_solutions/healthz/)

”

以下是运行状态的部分截图

```
{
  "fta_backend": {
    "cc:details": {
      },
      "ok": true,
      "cc:error": "",
      "mysql:details": {
        },
        "redis:details": {
          "can_set": true,
          "can_expire": true
        },
        "supervisor:error": "",
        "supervisor:details": Object{...},
        "redis:error": "",
        "result": true,
        "request_id": "e89788d4d07d45b1afddf4266e80851a",
        "beanstalk:error": "",
        "beanstalk:details": {
          "10.235.46.11:6380": {
            "FTA_ALARMS_TO_SCHEDULER": {
              "current-jobs-delayed": 0,
              "pause": 0,
              "name": "FTA_ALARMS_TO_SCHEDULER",
              "cmd-pause-tube": 0,
              "current-jobs-buried": 0,
              "cmd-delete": 0,
              "pause-time-left": 0,
              "current-waiting": 3,
              "current-jobs-ready": 0,
              "total-jobs": 0,
              "current-watching": 142,
              "current-jobs-reserved": 0,
              "current-using": 142,
            }
          }
        }
      }
    }
  }
}
```

图 2. 运行状态效果

## 故障自愈异常时运行状态效果

当故障自愈自身或依赖的第 3 方组件异常时，会在帮助菜单上呈现异常指标(值为 false)的数量。



图 3. 异常时运行状态提示

以下是故障自愈依赖的组件异常时的效果

```
{
  "fta_backend": {
    "message": "Component request third-party system [FTA] interface [fta_status_process] error: HTTPConnectionPool[host='fta.service.consul', port=13031]: Max retries exceeded with url: /fta/status/process/ (Caused by NewConnectionError('craigslist.packages.urllib3.connection.HTTPConnection object at 0x7fd3c2c6410<: Failed to establish a new connection: [Errno 111] Connection refused')), please try again later or contact component developer to handle this",
    "result": false,
    "request_id": "e82a5ea9c14949aa9d71ab0bd83d4dfc"
  }
}
```

图 4. 异常时运行状态效果

## 故障自愈是监控系统吗

故障自愈本身不是监控系统，它获取监控系统的告警，然后执行对应的处理动作，实现故障处理的无人值守。

详细请访问故障自愈的 [产品架构](#)。

# 故障自愈依赖哪些周边系统

故障自愈依赖蓝鲸的 [配置平台](#)、[作业平台](#)

你需要提前在 [配置平台](#) 上创建一个业务，并且把你的服务器录入到配置平台中

## 故障自愈首页指标计算方法



- 1 和 5, 自愈成功次数 和 自愈趋势: 每半小时更新一次
- 2, 健康诊断: 每天早上 8 点刷新
- 3, 自愈小助手: 实时刷新
- 4, 人力节省, 单个告警的人力节省: 15 分钟减去自愈耗时, 此处是最近 30 天内所有自愈成功告警的人力节省之和