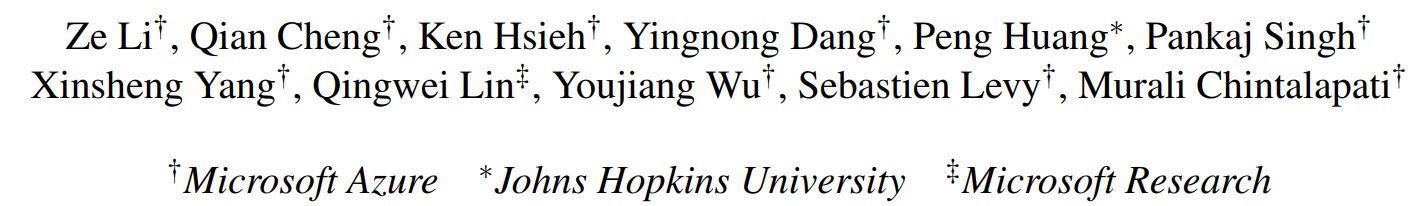
**Gandalf: An Intelligent, End-To-End Analytics Service for**

**Safe Deployment in Cloud-Scale Infrastructure**



**一、问题：**

在云规模系统基础设施中，不同的团队需要频繁地在代码和配置上进行软件更改。但是此类基础设施的规模大和复杂性高，即使在更新组件时出现一个小问题，也可能导致广泛的故障，从而对客户造成重大影响。而且由于测试和生产环境在各个方面存在差异，准确评估云系统中部署的影响具有挑战性。主要表现在：

假阴性：遗漏一些潜在的问题。

假警报：误导监视器，错误的警报导致无害的发布停止，并阻止应用及时更改。

**二、解决方案：**

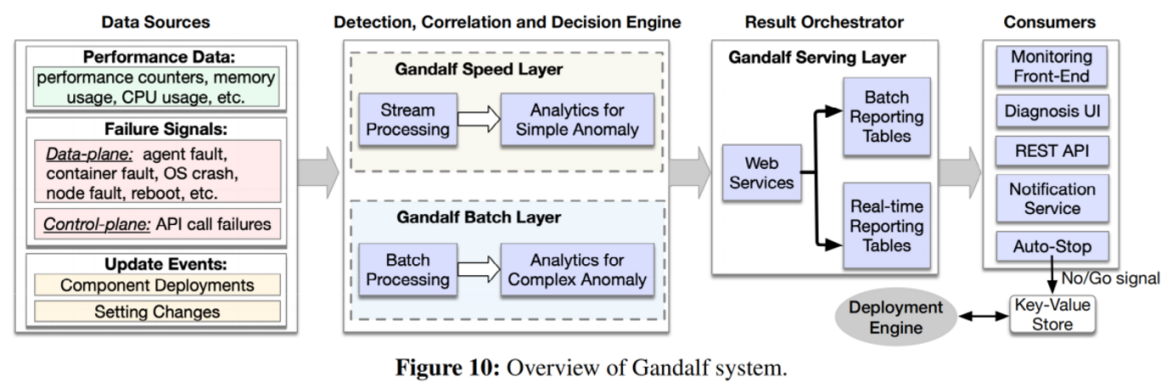
Gandalf：一种用于安全部署云基础设施的端到端分析服务。

Gandalf采用自上而下的方法来全面评估部署的影响。当检测到系统异常时，Gandalf分析它是否是由部署引起的。如果发现一个不好的部署，便会阻止它。

Gandalf的核心决策逻辑是一个由异常检测、相关分析和影响评估组成的新模型。使用lambda体系结构提供实时和批量部署监控，还提供自动部署决策、通知服务，详细的支持证据和交互式前端。

**三、具体实现：**

**3.1系统结构：**



1) Data Source：

Gandalf使用来自各种数据源的综合信号，执行预处理以解析原始数据并提取失败签名，并根据它们的时间戳、节点id和服务类型进行聚合。

2) Stream and Batch Processing：

为了平衡速度和覆盖率，Gandalf采用lambda体系结构[6]，同时具有流式和批处理分析引擎。speed layer使用来自Microsoft Kusto[7]的数据，batch layer使用来自Cosmos的数据。lambda架构使Gandalf快速决策和高覆盖率。

1. Result Orchestration and Actions：

使用Azure服务结构框架[8]实现一个高可靠性和可伸缩的web服务。

存储来自speed layer和batch layer的结果在两个单独的报告表中。

各种DevOps应用程序从报告表中提取结果。

通知服务。

1. Monitoring and Diagnosis Front-End：

提供了一个web前端，为发布经理和开发人员提供了实时的发布监控和问题诊断支持。

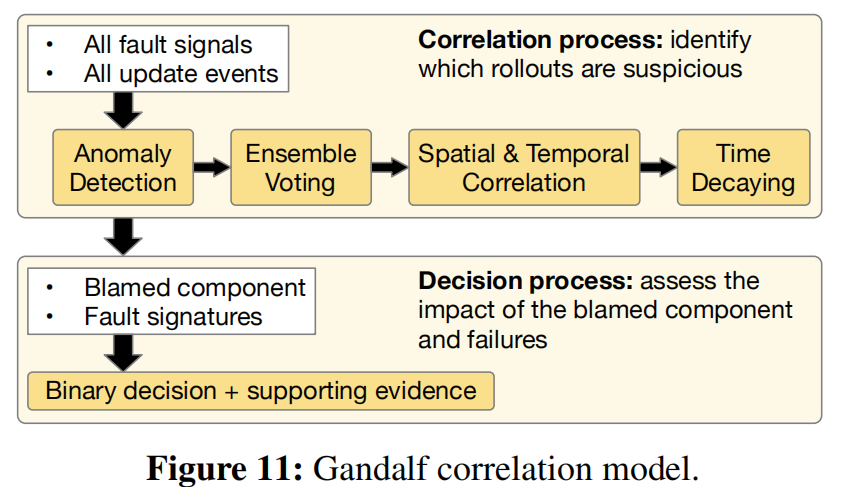
**3.2算法设计：**

相关性模型的步骤：

(1) 异常检测从原始遥测数据中检测系统级故障。

(2) 相关性分析确定在多个rollouts中对检测到的故障负责的组件；又分为即集体投票、时空相关性和指数时间衰减。

(3) 决策步骤评估受影响的范围，并决定是否应停止部署。



**四、评估：**

Gandalf已经在Azure产品上运行了18个多月。平均每天处理270K个平台事件，高峰日处理77000个事件，每天在控制平面中记录约6亿个API调用，包括2000多种故障类型。每天分析的总数据量超过20TB。对于每一个部署，Gandalf可以在5分钟左右的时间内做出决定，将整个生产机队的部署时间缩短了一半以上。

Gandalf在进入生产前阻止了99.2%的不良推广(rollouts)。对于data-plane rollouts，Gandalf实现了92.4%的准确率和100%的召回率。对于control-plane rollouts, ，Gandalf实现了94.9%的精确度和99.8%的召回率。

**五、优缺点：**

优点：效果好，考虑全面。

缺点：可移植性是否良好未知。这篇文章主要使对Azure上部署，不知道在其他平台上部署是否也能达到这么好的效果。

报告人：黄家晏

2020.09.26