AIOps
Challenge
智能运维挑战赛

落地生根：AIOps路线图

2017/11/17

清华大学 裴丹

清華大學
NetMan实验室

运维的重要性

- 在非完美的软硬件之上，保障业务可靠、高速、高效、安全运转；
- 直接影响业务收益和成本

互联网
金融
物联网
智能智造
电信
电力网络
政府
...

到2020年，全球将有500亿设备、无数上层软件及服务

故障发现

故障止损

故障修复

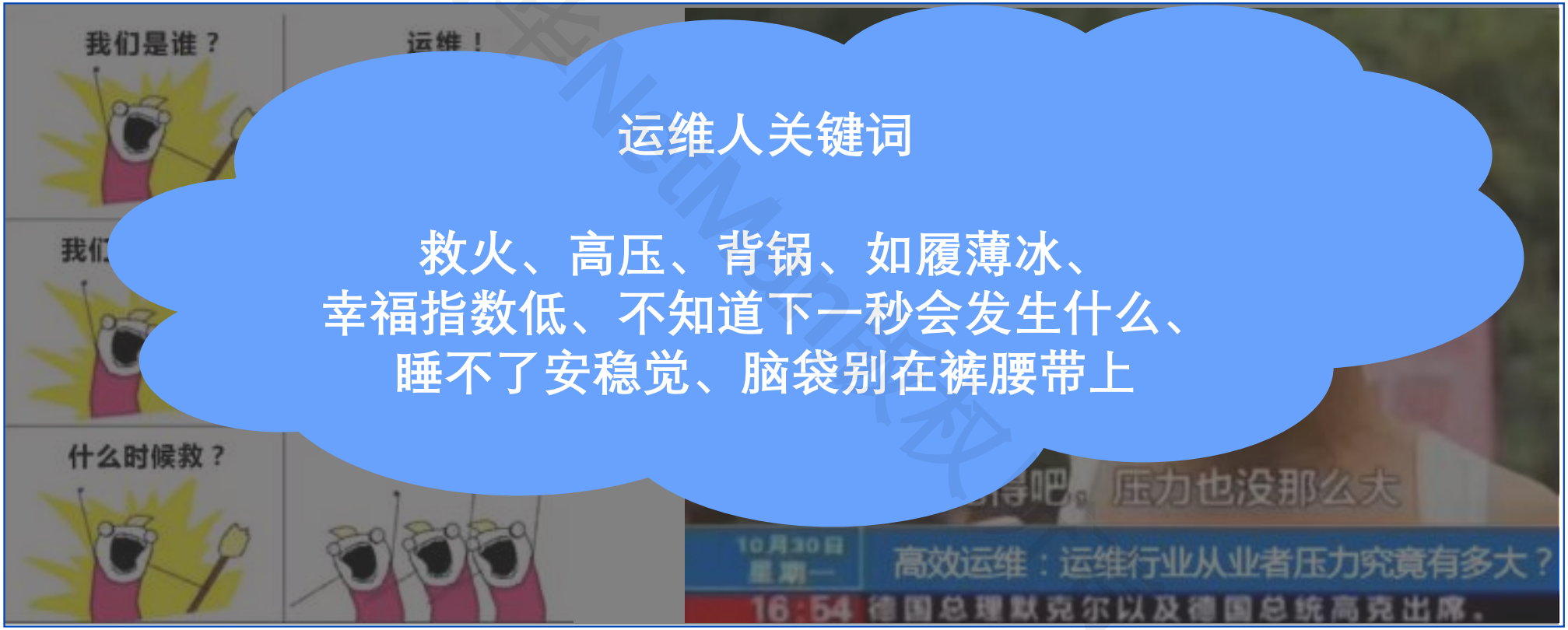
故障规避

运维人的痛点


我们是谁？
运维！
我们的工作是？
救火！
什么时候救？
7×24！


CCAV1
高某
北京望京某互联网公司
25岁系统运维工程师
其实我觉得吧，压力也没那么大
10月30日
星期一
高效运维：运维行业从业者压力究竟有多大？
16:54 德国总理默克尔以及德国总统高克出席。

运维人的痛点


我们是谁？
运维！
什么时候救？
运维人关键词
救火、高压、背锅、如履薄冰、
幸福指数低、不知道下一秒会发生什么、
睡不了安稳觉、脑袋别在裤腰带上
压力也没那么大
10月30日
星期一
高效运维：运维行业从业者压力究竟有多大？
16:54 德国总理默克尔以及德国总统高克出席。

运维现状：面对突发故障，仍大量依赖于人力分析决策， 效率低、不准确、不及时

手工运维

自动运维

人力分析决策

故障发现 → 故障止损 → 故障修复 → 故障规避

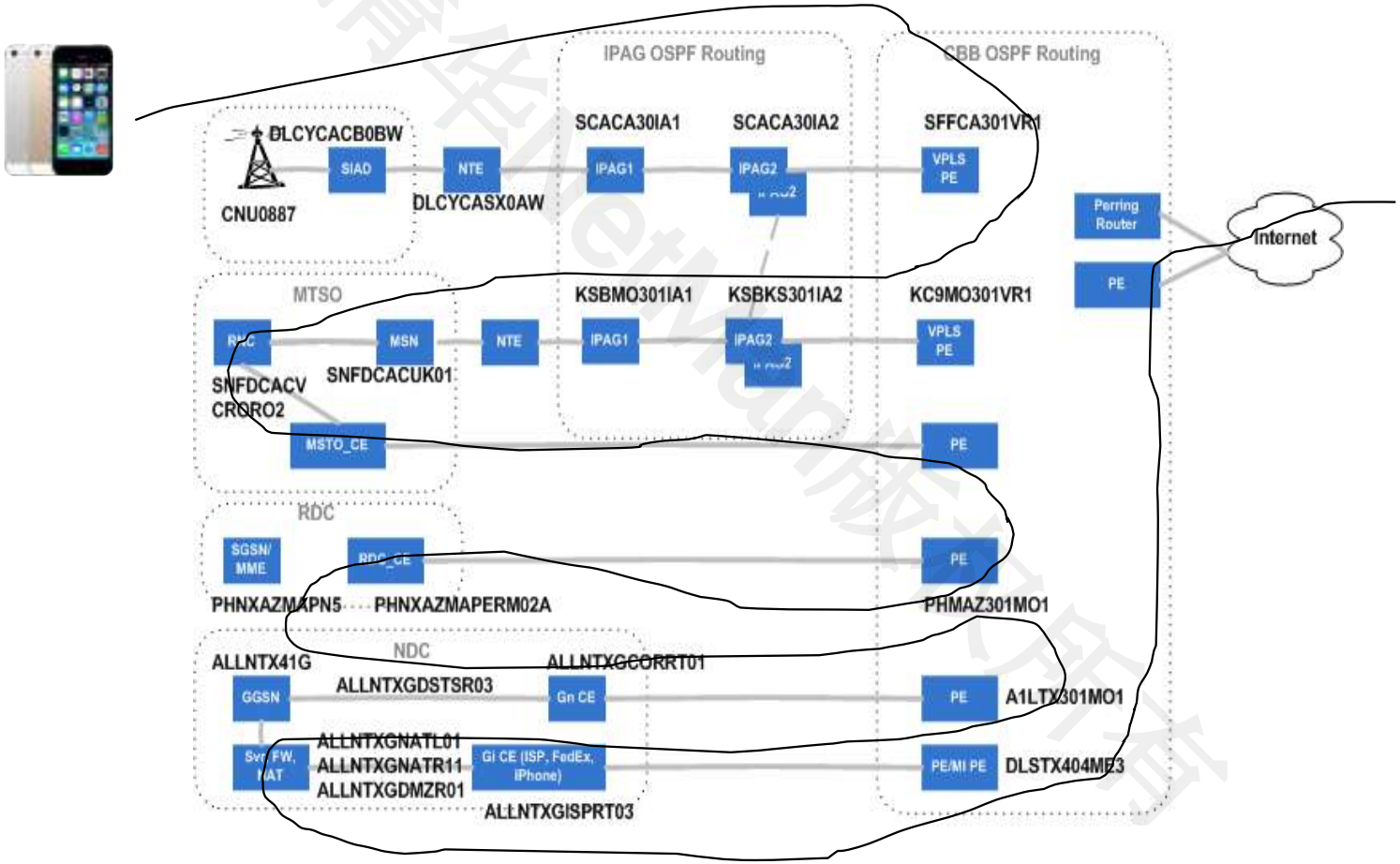
监测

（海量日志）

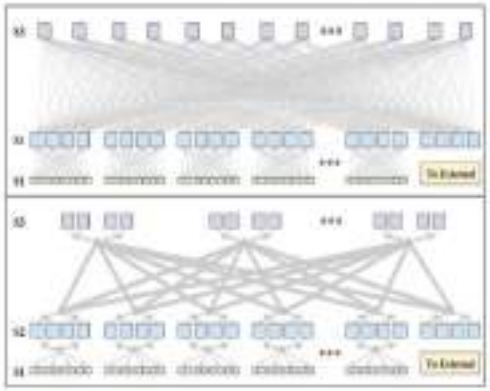
控制

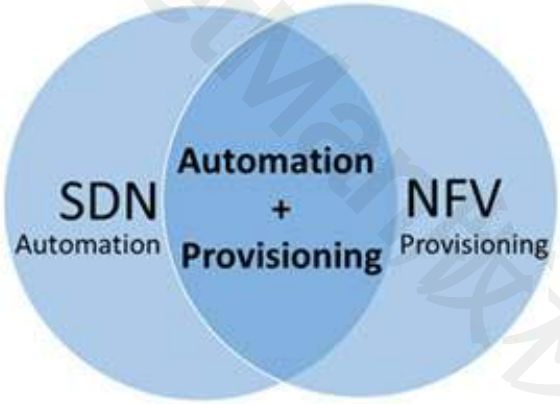
（自动化脚本）

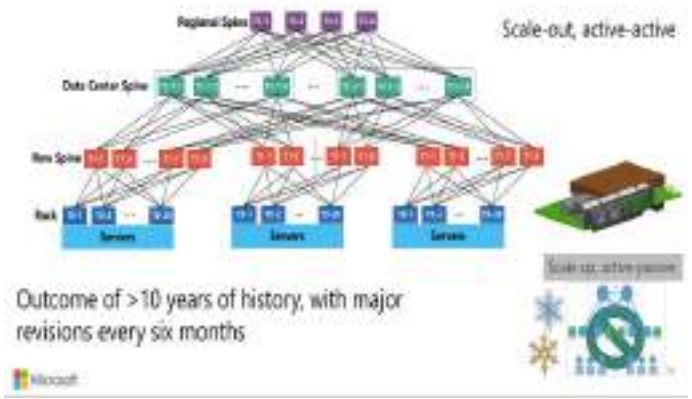
挑战： 广域网络越来越复杂

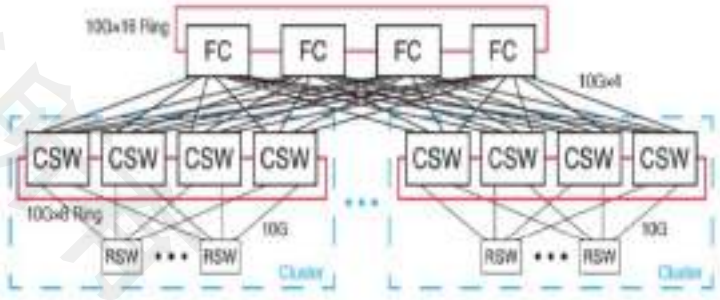


挑战：系统不断演进，规模、复杂度、变更频率增大、技术更新

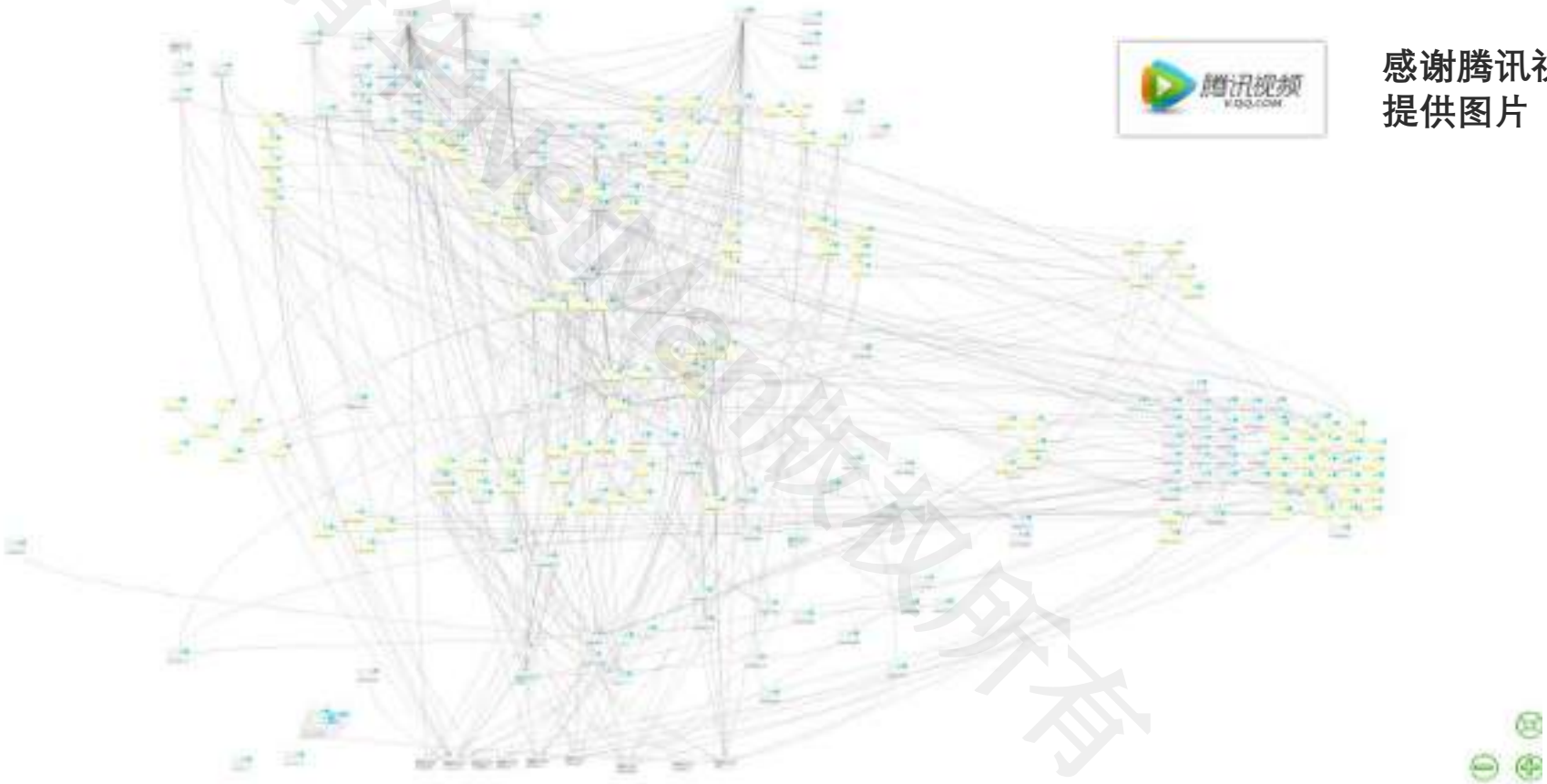






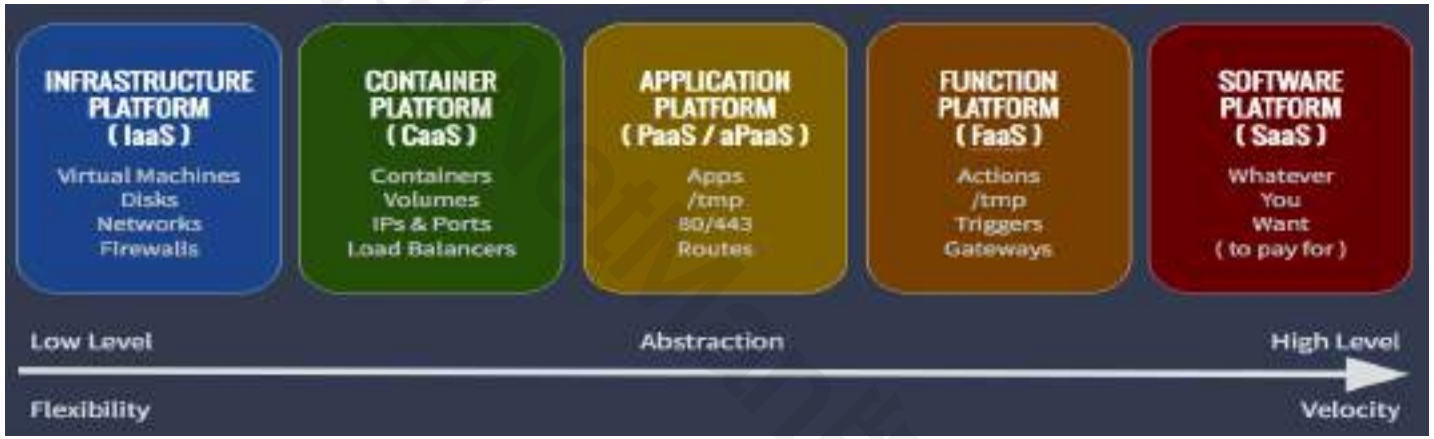


挑战： 软件规模、调用关系复杂度、变更频率逐渐增大



感谢腾讯视频
提供图片

挑战：软件架构、工程方法不断演进

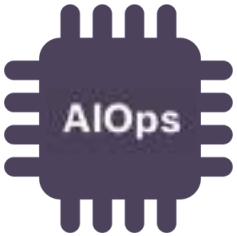


必然走向： 基于机器学习的智能运维(AIOps)

- 庞大、复杂、多变的软硬件系统故障难以避免
- 保障业务可靠、高速、高效、安全运转，如何快速准确决策？
- 运维规则复杂、多变，无法人力维护
- 海量、高速、多样、高价值的监控数据

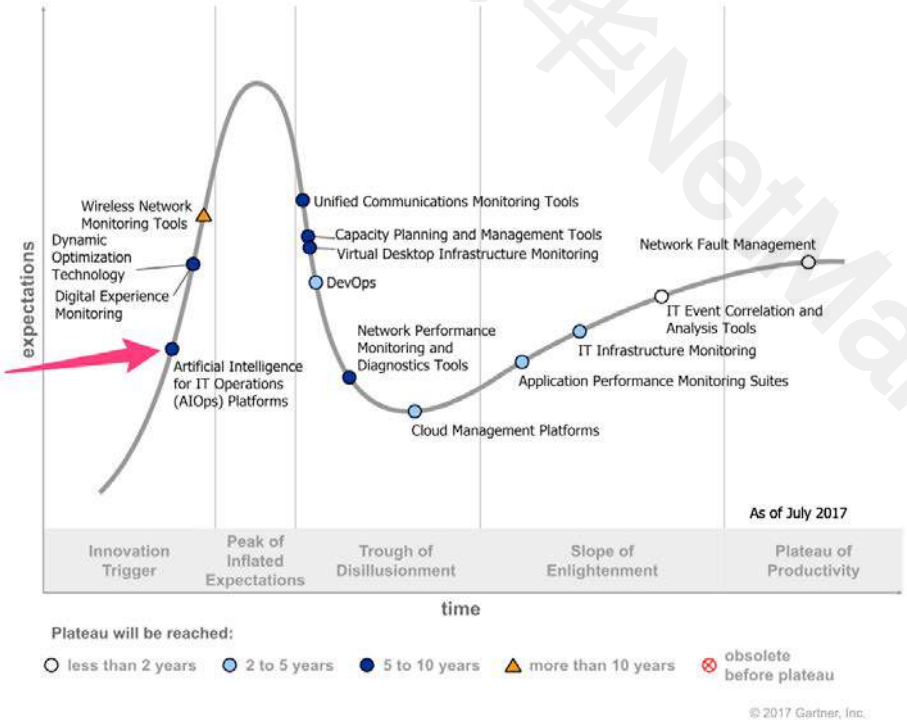
通过AIOps 大幅提升运维效率

自动分析决策



手工运维 → 自动运维 → AIOps

Gartner Report： AI for IT Operations



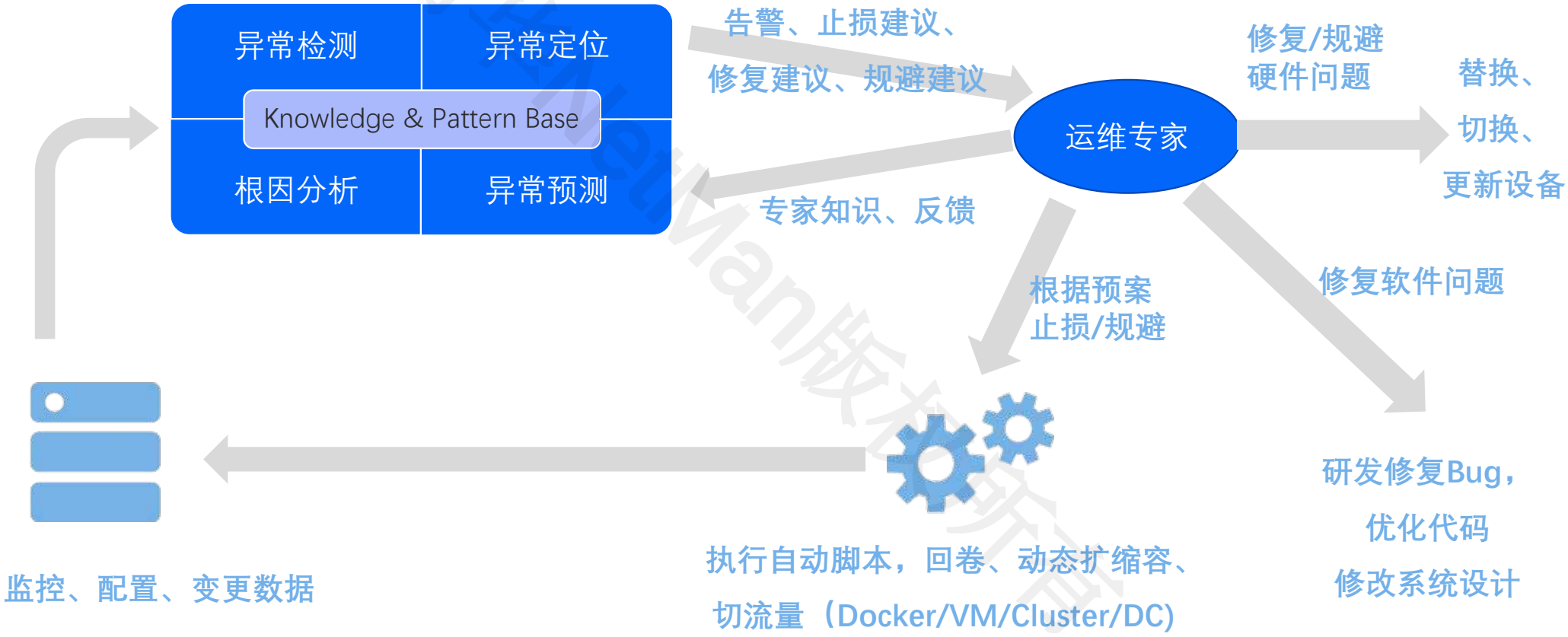
Source: Gartner (July 2017)

全球部署率

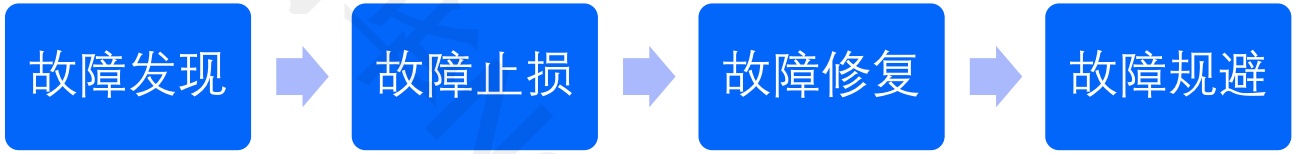
10% in 2017

50% in 2020

AIOps：少量运维专家+运维机器人



AIOPs现状

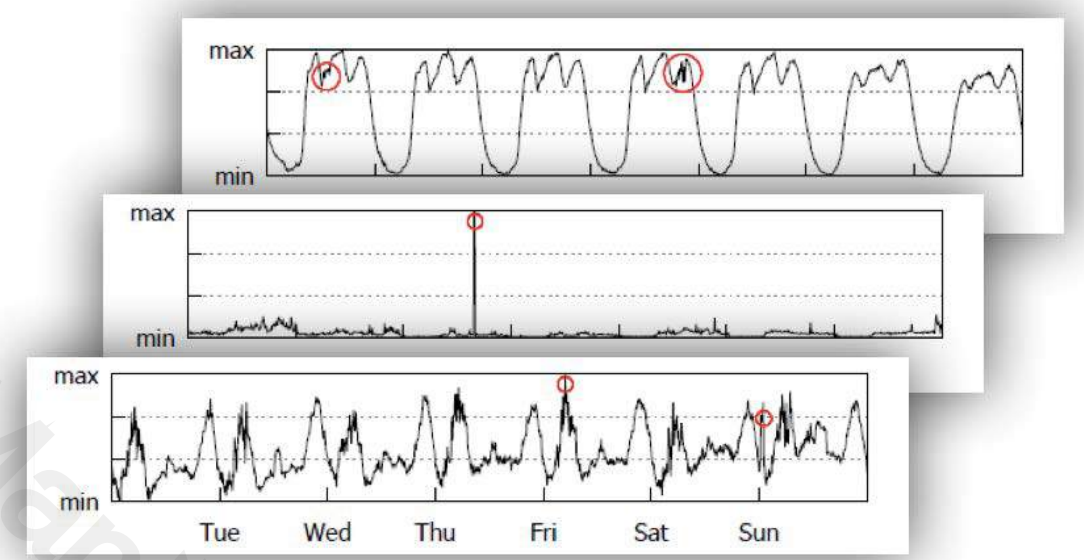


鸿沟：漫无头绪、无从下手、零星尝试、浅尝辄止

机器学习算法

ARIMA，时序数据分解，Holt-Winters, CUSUM，奇异谱变换（SST），DiD，DBSCAN，Pearson 关联分析，J-Measure，Two-sample test，Apriori，FP-Growth，K-medoids，CLARIONS，Granger 因果性分析，狄利克雷过程，逻辑回归、关联关系挖掘（事件-事件、事件-时序数据、时序数据-时序数据）、聚类、层次聚类，聚类树，决策树、随机森林、支持向量机、蒙特卡洛树搜索、马尔科夫链，隐式马尔科夫、多示例学习、迁移学习、卷积神经网络，递归神经网络（RNN），变分自动编码（VAE），自然语言处理技术...

困境举例：KPI故障发现



- **现状：漏报多误报多，故障发现不及时**
- **挑战：**
  - 静态阈值不工作不能满足复杂的检测需求
  - 多种时序算法，适用场景不明确，选择困难
    - 算法有大量可选参数，含义不直观，无从选择
    - 算法受数据异常、缺失影响，准确率低
  - 无漏报、误报反馈，无法评估算法效果
    - 根据个别case拍脑袋修改算法，丢西瓜捡芝麻
  - 大量差异性KPI需适配算法，开销巨大
  - 变更导致KPI剧变，已有算法失效，大量误报
  - 异常标注难以批量获得，无法有效支撑有监督的异常检测
    - 只有零星的case
    - 运维人员既不愿意标，也往往标不好
  - 既包括“KPI反常”，也包括“KPI（如CPU利用率）取极限值导致上层业务受影响”
- **导致：任何一个已有算法都无法同时解决上述所有挑战**

AI擅长解决的问题

人工智能在解决以下类型问题时，不管问题多么复杂，都可能做到甚至超过人类的水平。这类问题的特点是：

（1）有充足的数据或知识

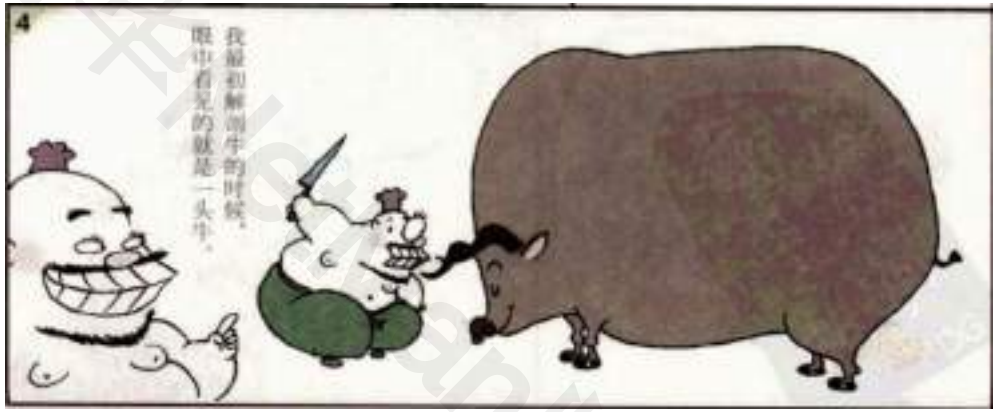
（2）完全信息

（3）确定性（well-defined)

（4）单领域

——张钹 2017年5月

解决思路：庖丁解牛

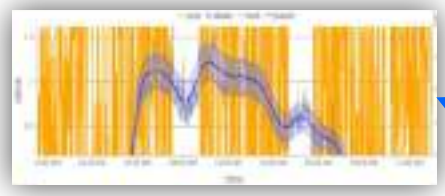


目无全牛
官止神行

批郤[xì]导窾[kuǎn]
切中肯綮[qìng]

游刃有余
踌躇满志

自适应异常检测：分解为多个AI擅长解决的问题，各个击破

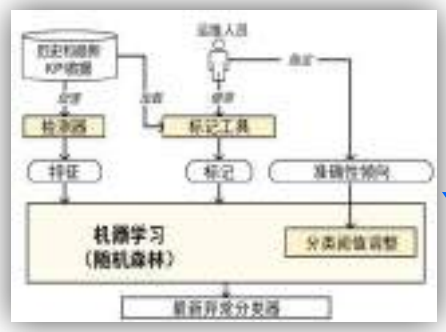


极限阈值自动配置

曲线波动关联、回归分析等，无需人工配置

普适的无监督异常检测

适用更多差异性KPI，
无需人工配置，开箱即用

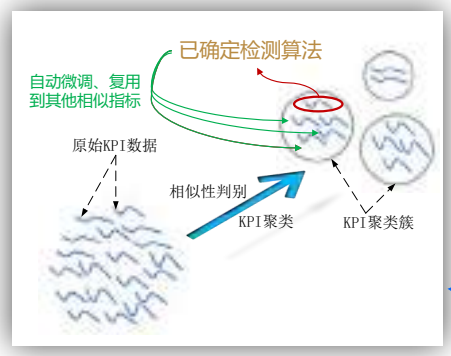


高效的误报、漏报反馈工具

针对每一次反馈，自动搜索相似情况，举一反百

自动算法、参数选择与优化

基于有监督机器学习，根据简单的反馈
自动选择最准确的算法和参数组合，自我优化

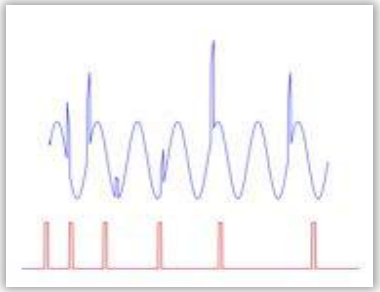


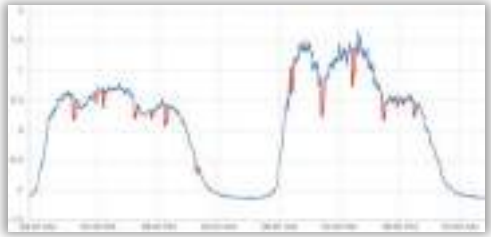
快速适应KPI剧变

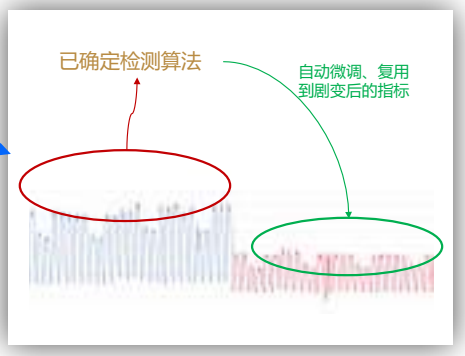
利用迁移学习，快速调节历史检测算法，适应最新KPI

多KPI自动分类并适配检测算法

识别不同KPI相似程度，自动为每类KPI适配算法

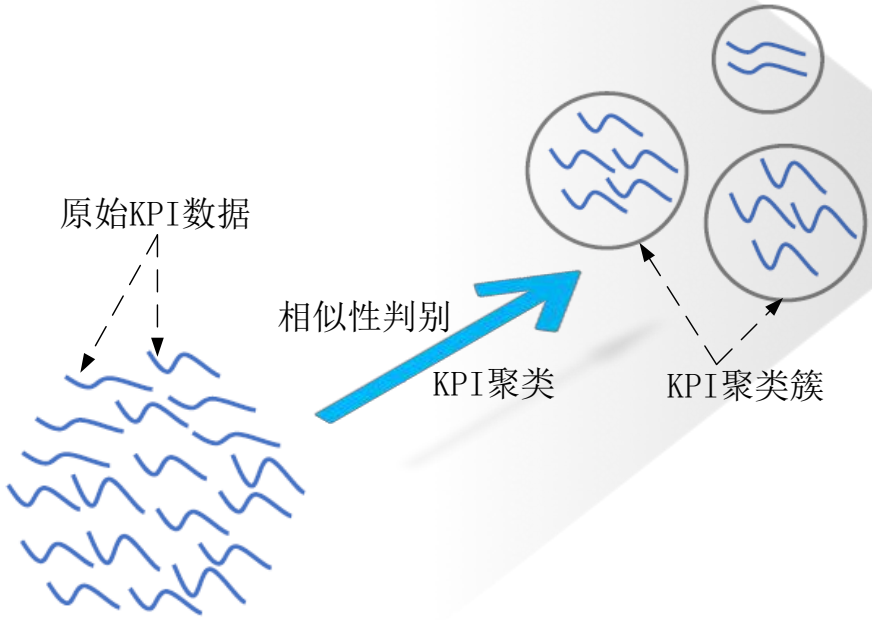






故障发现： 自适应的异常检测系统

Well-defined problem



应用层：自适应的异常检测系统

智能运维技术层

无监督异常检测	有监督异常检测	算法自动选择和调参	指标辅助标注	多指标分类	巨变适配	极限阈值配置

智能运维算法层

指标聚类	指标分布预测	相似曲线识别	异常区间对比

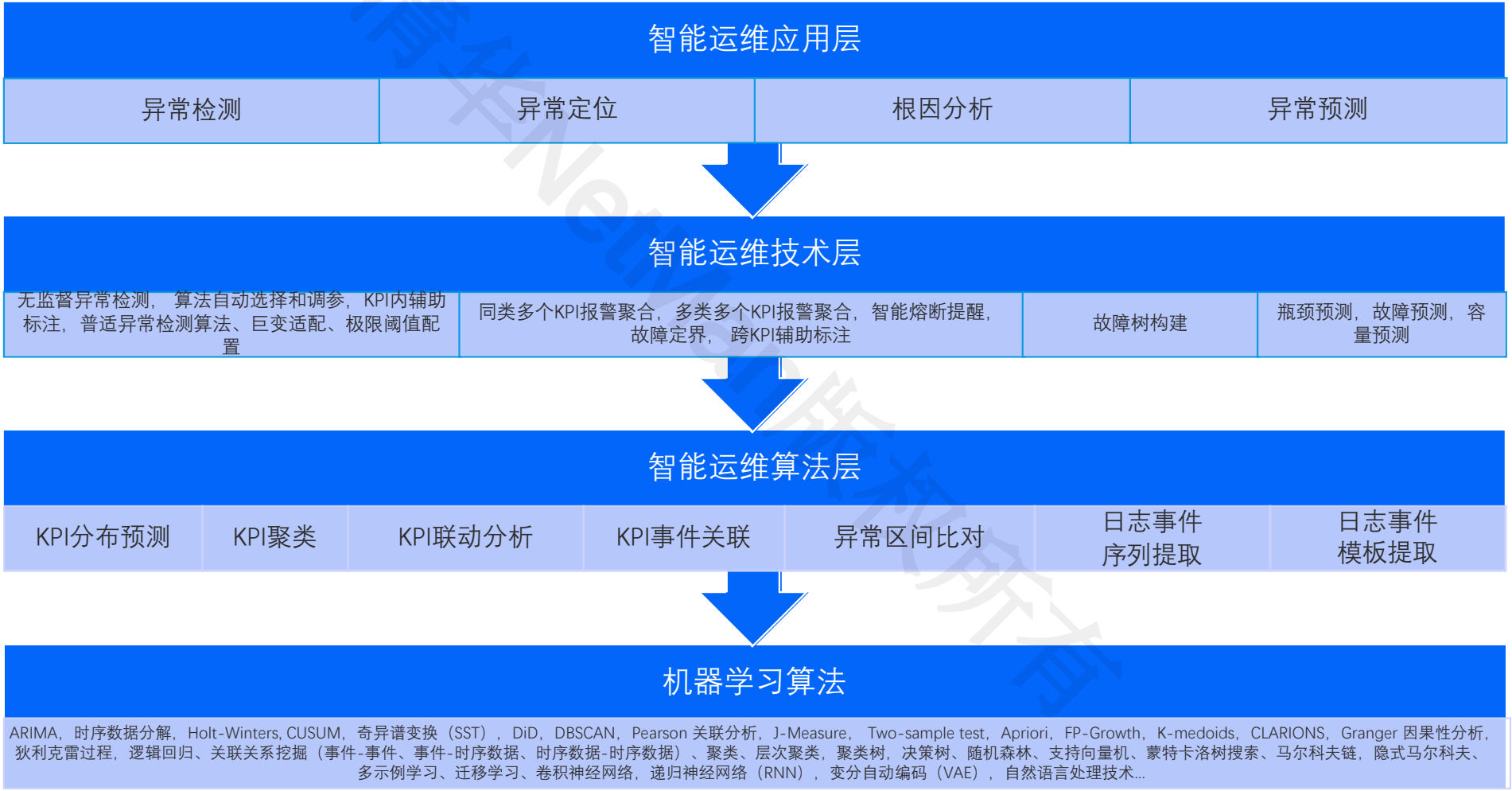
机器学习算法

ARIMA，时序数据分解，Holt-Winters, CUSUM，奇异谱变换（SST），DBSCAN，Pearson 关联分析，J-Measure，层次聚类，聚类树，随机森林、支持向量机，迁移学习、卷积神经网络，递归神经网络（RNN），变分自动编码（VAE）……

AIOps落地技术路线图

清华大学
NetMan实验室

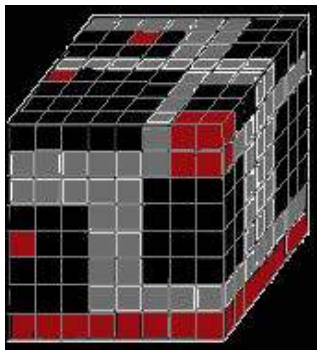
AIOps技术路线图


智能运维应用层
异常检测
异常定位
根因分析
异常预测
智能运维技术层
无监督异常检测，算法自动选择和调参，KPI内辅助标注，普适异常检测算法、巨变适配、极限阈值配置
同类多个KPI报警聚合，多类多个KPI报警聚合，智能熔断提醒，故障定界，跨KPI辅助标注
故障树构建
瓶颈预测，故障预测，容量预测
智能运维算法层
KPI分布预测
KPI聚类
KPI联动分析
KPI事件关联
异常区间比对
日志事件序列提取
日志事件模板提取
机器学习算法
ARIMA，时序数据分解，Holt-Winters, CUSUM，奇异谱变换（SST），DiD，DBSCAN，Pearson 关联分析，J-Measure，Two-sample test，Apriori，FP-Growth，K-medoids，CLARIONS，Granger 因果性分析，狄利克雷过程，逻辑回归、关联关系挖掘（事件-事件、事件-时序数据、时序数据-时序数据）、聚类、层次聚类，聚类树，决策树、随机森林、支持向量机、蒙特卡洛树搜索、马尔科夫链，隐式马尔科夫、多示例学习、迁移学习、卷积神经网络，递归神经网络（RNN），变分自动编码（VAE），自然语言处理技术...

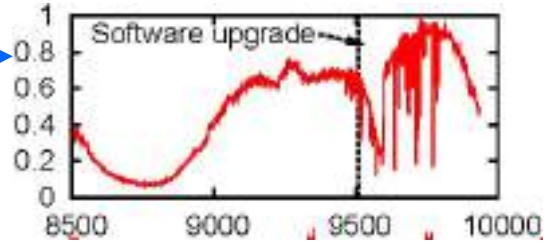
故障止损：定位粒度足以实施修复预案即可

执行自动脚本，回卷、动态扩缩容、
切流量（Docker/VM/Cluster/DC)

故障止损：异常定位的粒度足够采用止损预案即可



智能熔断提醒

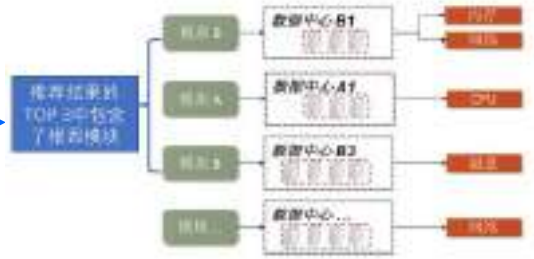


业务故障是否由某个变更导致？

单指标多维度告警聚合

自动聚合到根因维度组合

指标波动聚类

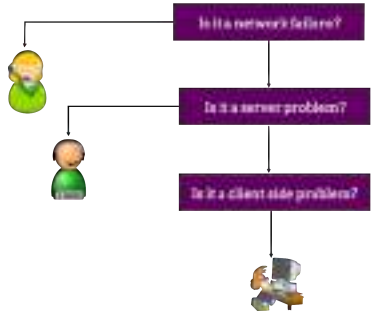


找到与业务一起波动的指标，聚类后推荐Top N

粗粒度根因分析

依据非完整故障树做粗粒度根因分析

粗粒度故障定界

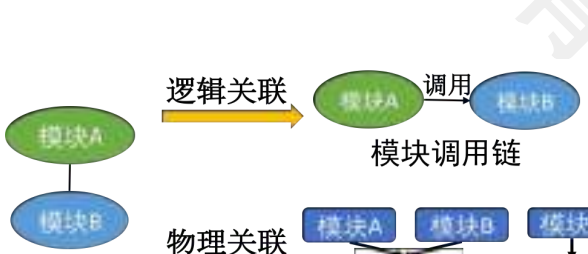


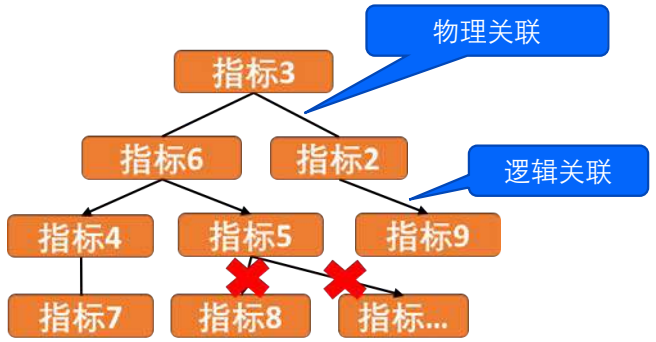
定界到服务器、网络、数据库、用户

动态扩容提醒

确定故障根因是否为容量不足

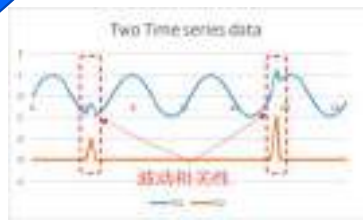
故障修复：根因分析（异常检测 + 故障树构建）





可能的故障传播关系

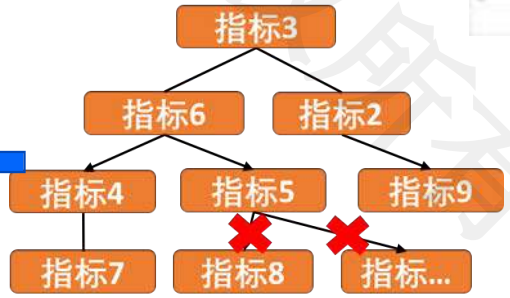
剪枝



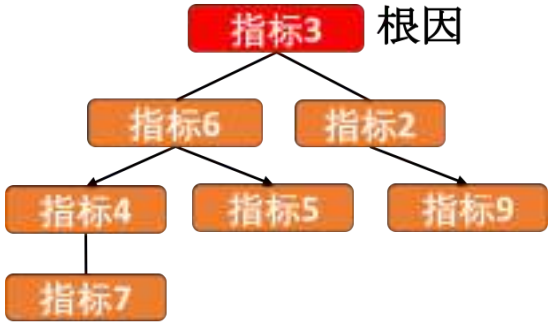
故障	异常
故障1	e1, e2, e3, e4
故障2	e1, e2, e5, e6
故障3	e1, e2, e7
故障4	e1, e3, e5, e7, e9
...	...

- 联动分析

异常关联分析

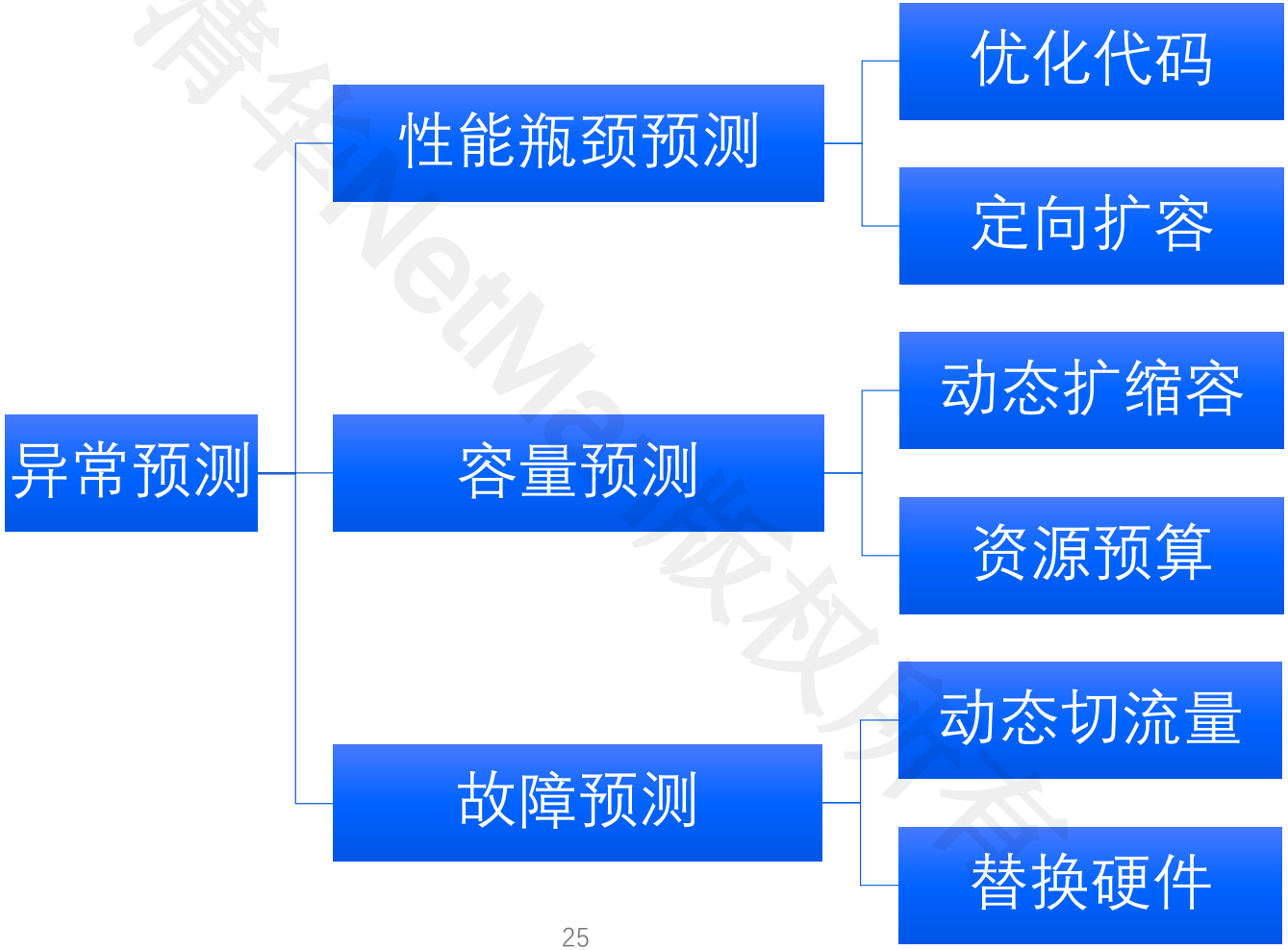


可能的故障传播关系



最终的故障传播关系

故障规避：异常预测



AIOps落地战略路线图

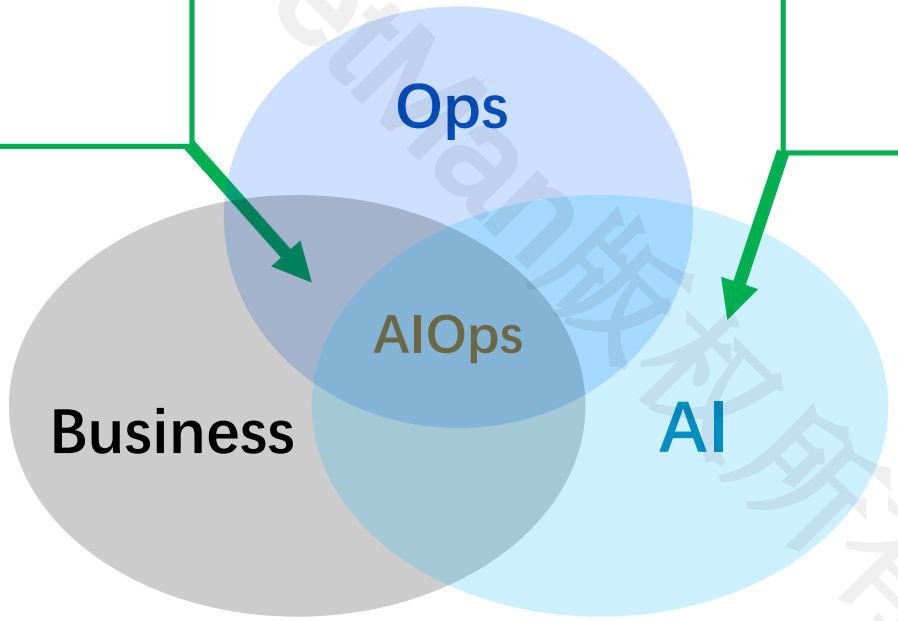
通过社区力量解决AIOps难题

清华大学
NetMan实验室

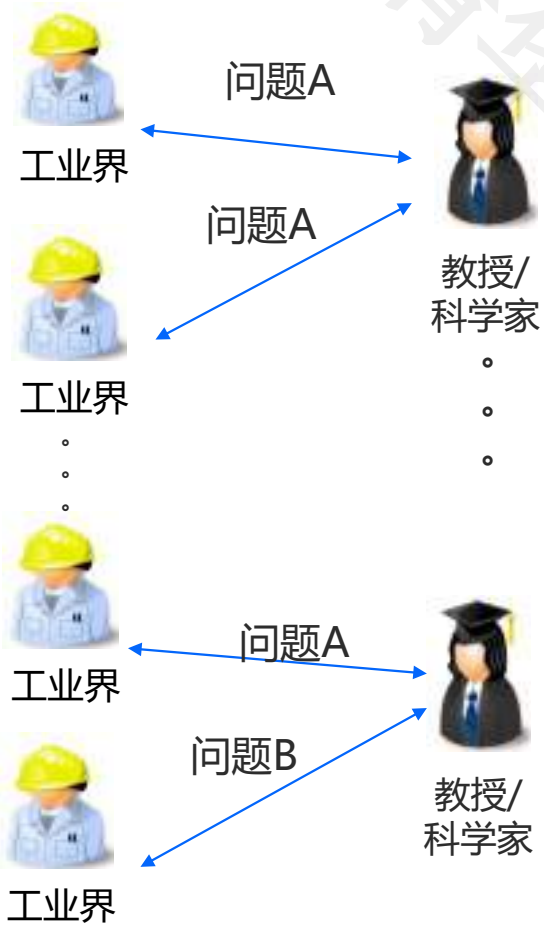
AIOps 落地需要运维界-算法界密切合作

- 工业界
  - 熟悉行业和运维场景
  - 熟悉生产实践中的难题
  - 有数据
  - 不熟悉AI

- 算法界
  - 不熟悉运维领域知识
  - 没有数据
  - 熟悉算法
  - 不了解AIOps



工业界-科学家合作 1.0： 一对一交流合作



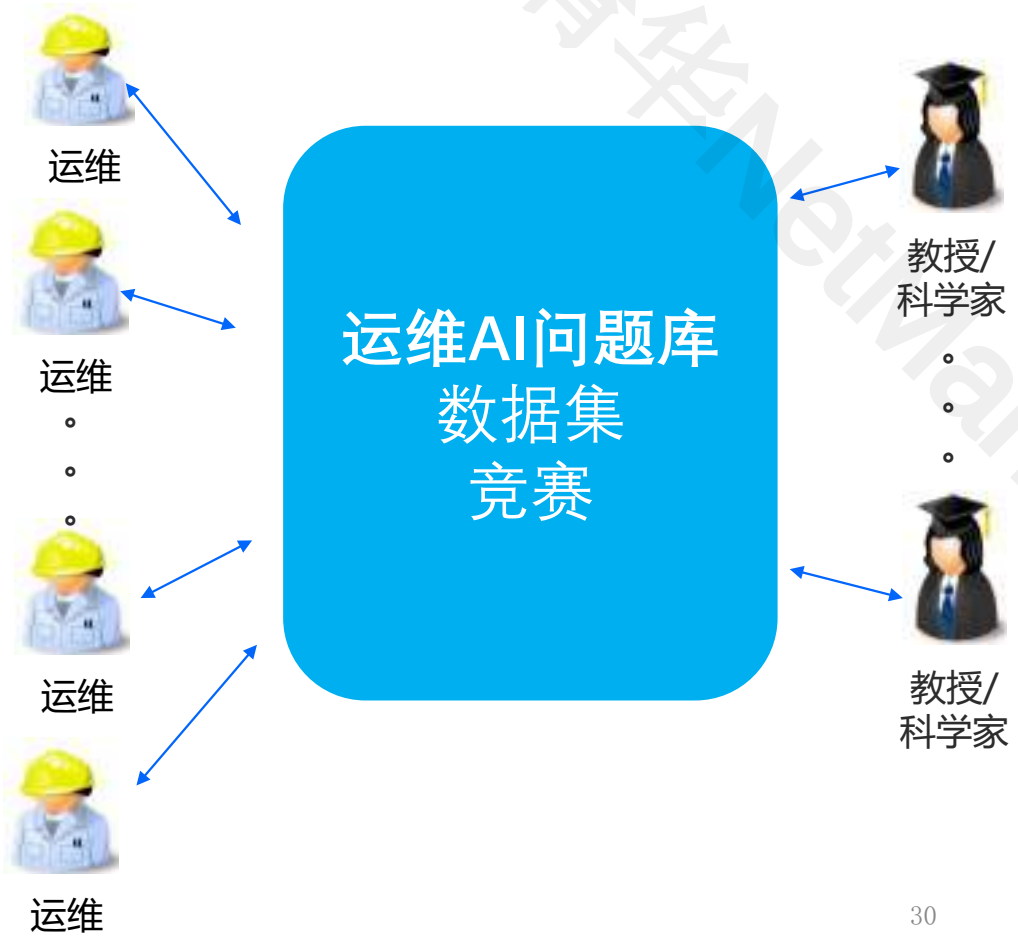
- 重复进行问题定义
- 各自为战
- 没有公认的Benchmark数据

受 ImageNet启发，工业界-算法界合作 2.0（社区合作）

“毕业于普林斯顿大学的华裔女科学家李飞飞，在不被看好的情况下，带领团队创健了名为 ImageNet 的数据集和人工智能挑战赛。

ImageNet 大赛不但带动了人工智能（AI）的高速发展，更为Google、facebook 等企业培养了一众优秀的 AI 专家，重新定义了人们研究人工智能的思考方式，推动了如今如火如荼的人工智能浪潮。”

AIOps Challenge：通过社区力量解决运维难题



解决AIOps落地的如下挑战：

- 数据
- 算法
- 人才

智能运维挑战赛官网 http://iops.ai

科研问题示例：异常检测

背景介绍

随着互联网，特别是移动互

要。

Web服务的稳定性主要靠

题。图1中展示了一个KPI异

这些KPI大致分为两种类型

反映机器（服务器、路由器

KPI异常检测指的是通过算

1 异常发生的频率很低。在

2 异常种类的多样性。因为

3 KPI的多样性。KPI有表现

数据集举例

表1 训练KPI数据集例

Timestamp	Value	Label
1503831000	10.8	0
1503831060	12.3	1

评估指标

准确率（precision），召回率（recall）和 F-score计算方式：

评价指标示例：

标注	0	0	1	1	1	0	0	1	1	1
算法输出	1	0	0	1	1	1	0	0	0	1
调整后的算法输出	1	0	1	1	1	1	0	0	0	0

第一届智能运维挑战赛

/ 赛题

来自搜狗、腾讯、eBay、百度的多条真实KPI曲线，有异常标注。
预赛：基于预赛数据集训练一个统一的异常检测算法，对多条多种类的KPI曲线进行异常检测。
决赛：预赛排名靠前的队伍进入决赛，在指定计算资源上，对决赛数据集进行训练并检测异常，决定决赛排名。

/ 奖金

亚军
8,000人民币
颁发获奖证书

冠军
80,000人民币
颁发获奖证书

季军
4,000人民币
颁发获奖证书

/ 时间点

2017.11.17	2017.12.01	2018.01.01	2018.04.01	2018.04 中旬	2018.04.下旬
官网试运行 发布样本数据	预赛开始 发布预赛数据	预赛报名截止 榜单排名开始	排名靠前队伍 进入决赛	决赛选拔 现场答辩队伍	现场答辩 决定最终名次

数据赞助

Sogou 搜狗

腾讯游戏 Tencent Games

ebay

Baidu 百度

网站建设

腾讯游戏 Tencent Games

协办

高效运维社区 GreatOPS Community

诚邀参与！

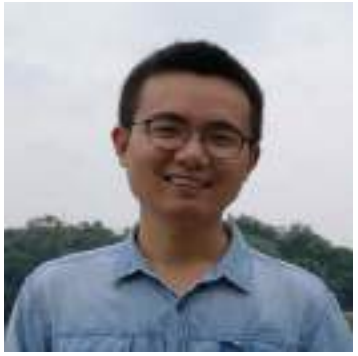
运维界

- **付出：**
  - 赞助脱敏数据
  - 资金赞助
  - 建议新问题
  - 社区讨论
- **收获：**
  - AIOps算法
  - 算法合作者
  - 人才招聘
  - 影响力

算法界

- **付出：**
  - 参赛
  - 社区讨论
  - 定义问题
- **收获：**
  - 前沿问题
  - 生产数据
  - 工业界合作机会
  - 算法影响力
  - 同行切磋
  - 推动AIOps进展

挑战赛组织者



张圣林

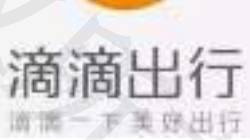
讲师
南开软件学院

孟媛

博士生
清华计算机系

全体NetMan实验室成员

致谢NetMan科研合作方



总结

- AIOps前景光明，“顶天立地，既是前沿的，又是现实的”
  - 具有丰富的数据和应用场景
  - 将极大提高运维领域的生产力
  - 是AI领域尚未充分开采的金库和低垂果实

- 通过社区努力推动AIOps落地
  - 技术路线图
    - 一系列Well-defined 的算法、技术、应用，及其依赖关系
    - 每一个都是清晰定义的AI擅于解决的问题
  - 运维界提供脱敏数据作为benchmark
  - 算法科学家贡献算法

谢谢！

挑战赛微信群

http://iops.ai

peidanwechat

清華大學

NetMan实验室