GOPS2016 Shanghai

GOPS

2016 全球运维大会

重新定义运维

上海站

会议时间： 9月23日-9月24日

会议地点： 上海·雅悦新天地大酒店

主办单位： 开放运维联盟 OOPSA Open OPS Alliance　高效运维社区 GreatOPS Community

指导单位： 数据中心联盟 Data Center Alliance

《IaaS平台运营的数据之魅》

-----浅谈数学模型和工具在业务运营和运维中的应用

•许杨毅 Ucloud运营平台部总监

•2016年09月

IaaS平台运营的挑战：

- 与传统的IDC托管、与自建私有云平台相比（也包括自建idc），公有云IaaS平台的运营属于3.0版本。原因如下：

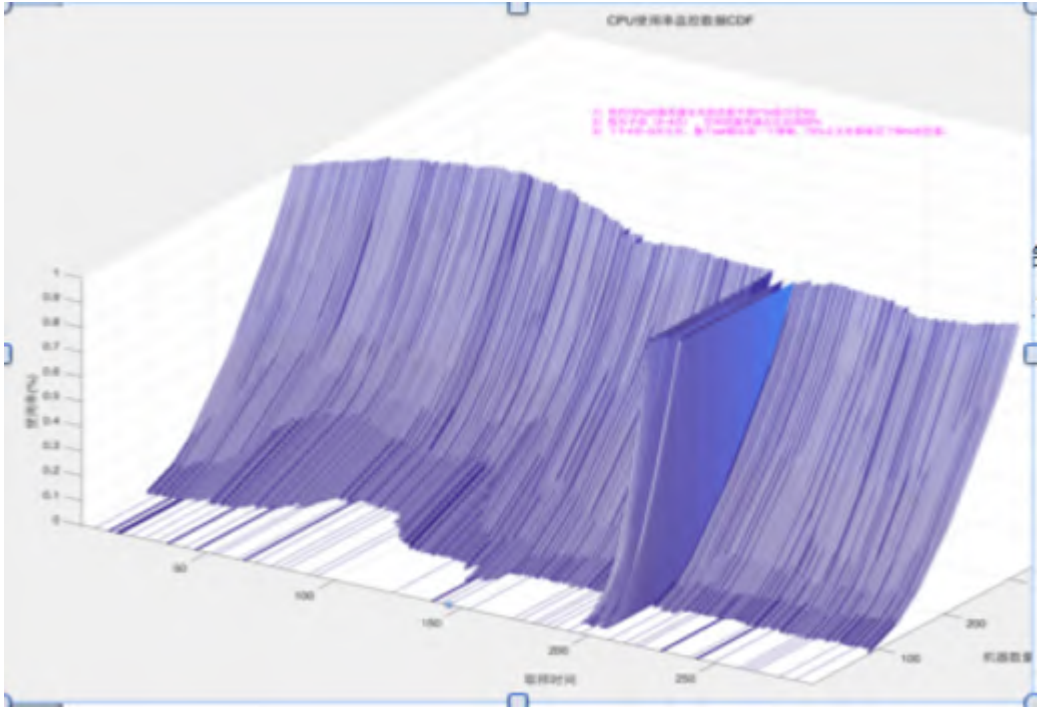
演进	平台	租户	业务	资源共享特点
ver1.0	传统 IDC	单租户	单一业务	主机资源独享 带宽资源独享 or 共享
ver2.0	私有虚拟化平台	单一租户	多业务	基于 vmware 或者 Opentack 技术自建的 共享资源 *单一组织内，自建的私有云平台面临的资源隔离和管理问题反而比 1.0 时代突出
ver3.0	公有 IaaS 平台	多租户	多业务	多业务，多租户的特点将公有云 IAAS 平台的管理复杂性带上了新的高度。

IaaS平台运营的挑战：

- 但是实际运营过程中间，基于云主机资源运行态的调度工作变成了必须要解决的大问题。（本文讨论的不仅仅限于具体KVM/Qemu级别的资源调度方法）也即是：需要在一个更高的层次上，循着｛物理机>单set > 整个idc >AZ可用区｝的层次，进行IaaS平台级别业务运行态的观测和特征识别。
- 在观察和识别的基础上，再辅助以资源调度方法来维持整个IaaS平台的动态平稳过程。保证每个用户的云主机稳定运行，关键性能-*如IO能力，CPU能力，包吞吐量，网络吞吐带宽*-不被其他用户云主机抢占。而资源调度包括了人工判断，逻辑化策略或则两者的组合。
-
- 最近2年流行起来的mesos和kubernetes，在资源管理上，实际上是达不到商用稳定性SLA的。因为其开源和开箱即用式的调度算法，调度逻辑上必然是通用而粗糙的。
- 大家可以去stackoverflow上看看大家对类似问题的关注，fairness scheduler以及task scheduler是这类框架在调度上的设计原则，因此在调度设计目标上，和公有云IAAS完全是两个方向。

发现与识别 – IaaS运营数据的特征工程：

- 我们先看看运行态指标的几种不太常见的形式。
- 采用3D的surface图来表示在整个IDC的某set中，一天的时间内，CPU，MEM，出入包量这几个常见指标负载的情况，我们是以CDF（累计概率密度函数）来表示的。
- 我们内部基于plotly.js和d3.js实现了这几类数据的实时web仪表台显示，可以帮助运维和运营人员实现交互可视化的业务运行情况分析。CPU负载的CDF概率曲面，其形态犹如《冰火之歌》的北境之墙
- 可以得到的结论如图红字：

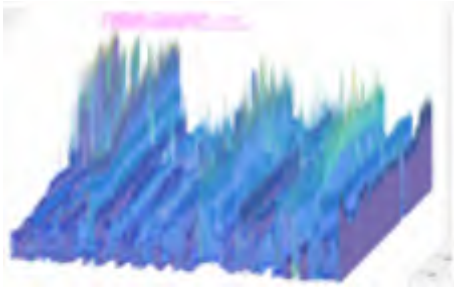


发现与识别 – IaaS运营数据的特征工程：

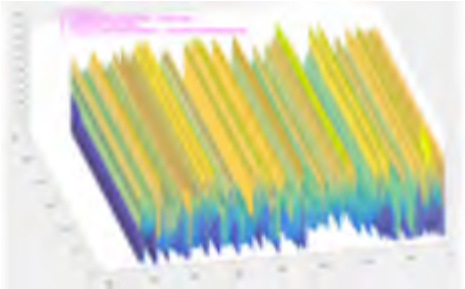
- 我们先看看运行态指标的几种不太常见的形式。

另外一类是metric 3D-surface,描绘了整个set内部，一天时间内某个指标的时间序列情况。下面几张图从上往下分别是：

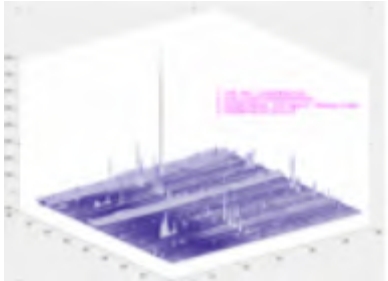
- cpu负载mesh



内存子系统负载mesh



网卡包量mesh



发现与识别 – IaaS运营数据的特征工程：

我们定义的特征：运行态指标的某种模式。

因为大部分数据都具备典型的周期性特征，例如{ day to day的类似度很高； 一天之内也有明显的峰谷特征}，采用傅里叶变换就成了我们很自然的选择

目的：目的是对宿主机进行分类，而分类的目的是提取宿主机的各种特征{CPU负载特征，网络IO特征，存储IO特征}，特征可以被应用在如下表格中的场景：

分类可以依赖的数据有{性能指标，宿主机的硬件类型，宿主机的uhost类型，其上的用户客户ID，宿主机上uhost的配置类型，宿主机的状态（锁定，待修复，专区等）}，其中隐含的一些假设是：

1. *uhost的配置类型（x核xGB nTB）可能隐含了用户的应用场景假设，内存8GB以上的可能用来进行部署DB或者MC；REDIS -内存吞吐密集型；2核4G的可能用来作为前端 - CPU密集型/包量密集型；而200GB以上存储的可能是持久化应用DB，或则文件存储 - 因此应该是IO密集型； --总之我们要把配置类型作为特征放入考虑范围，具体的实现办法是整理所有的配置，假设是50种，做出标注。*
2. *用户ID： 某些用户可能申请云主机进行特定类型的任务，比如作为爬虫，作为秒杀器，作为外挂 ，甚至作为肉鸡，因此用户ID+配置类型这个组合必然可以提取到特征。*
3. *Uhost类型： 是指 普通，SSD，高包量，大数据这些类型，由于特征空间不到10个，作为分类特征可能是表层的，也可以考虑进来。但是未必是重要的*
4. *锁定可能是由于某种原因保留的宿主机，可能目前低负载的那些机器是锁定状态。*

发现与识别 – IaaS运营数据的特征工程：

最终我们是需要解释这种分类结果的，例如｛某一类主机在夜间的某几个时间段，运行在cpu util 70%~80%同时内存消耗在50%~60%的位置｝

我们采用了两个方法来侧面印证这种分类解释性：

1） 序列原始数据的比照

2）高斯随机过程的非参回归方法

原始数据的比照是人工方法，这部分类似图像识别中的人工标注部分。

尽管高斯过程也是机器学习的一个重要方法，可是由于我们的方向是解析型的，过程中的每一步都有人工解释性的特点，所以自然不能算机器学习这类黑盒方法。

发现与识别 – IaaS运营数据的特征工程：

非参高斯回归过程（Nonparametric Gaussian Regression process ）

是我们采用的另外一个重要方法。如果要找到你感兴趣数据的特征，例如寻找一个模型来匹配数据，非参高斯过程是和SVM的常用而典型的监督学习方法。这里也可见其和机器学习的不同。

这和实际情况也是吻合的，运行态指标其实往往具备显而易见的特征（双峰，单峰，浪涌，高或低），我们寻求数学模型的目的在于，通过数学形式确定特征指标，将人类可以获得的感性认识转化为可定量/确定公式化的内容，这是将判断方法程序化的必要步骤。

一个高斯过程由其平均函数 u(x)=E(F(X))和核函数（或者协方差函数）k(x,x‘)=cov(f(x),f(x’))完全确定 。通常的情况是假设均值为零，因为求解一个未知平均函数的边界可以等效为求解一个均值为零的新核函数的高斯过程。核的结构捕获了未知函数f的高层属性，在顺序上这决定了模型怎样产生出新的数据。

FYI: http://www.gaussianprocess.org/#code

发现与识别 – IaaS运营数据的特征工程：

- 数据处理的pipeline：

1. 获取到重要的云主机关键运行态指标序列（我们按照day,week,month的尺度分别获取序列）
2. 对某个指标序列进行DFT（离散傅里叶变换），得到其傅立叶特征，因此这个时候这份数据已经从时域转换为频域的离散数据，但是只要是特征，数据就有价值，在数学过程中，中间步骤的物理含义往往无需纠结。
3. DFT后的傅立叶特征值，仍然是一个序列（频域序列），我们利用这份序列数据进行分类。
4. 分类上我们采用了一些特别的方法，首先2个sigma以外的数据会被作为异常值处理，利用pandas插值为均值数据，例如能在mesh上观察到的异常的尖峰（这实际上是ddos-out或者ddos-in的场景）。
5. 然后我们采用了一种被称为DTW的方法进行预处理，预处理实际完成的是将频域上离散的多维数据压缩为一维的一个lin值。
6. 采用DBSCAN这种空间聚类方法以这个值计算多个不同序列间的距离，从而得到分类结果。

数学模型的奥卡姆剃刀准则：

高斯过程回归方法中，通过计算共轭梯度，采用贪心算法计算核函数组合对于模型的拟合度。这一过程中间往往依据残差服从正态分布的先验假设，使用贝叶斯信息准则（Bayesian information criterion, BIC）来判断自变量对于回归模型的重要性。

如前所述，业务特征往往具备周期特征，这时候选择周期函数，例如傅里叶作为关键核函数就很顺理成章，这也体现了奥卡姆剃刀准则，例：如非必要，我们绝不将WN（高斯白噪音）作为重要的核函数，而是将其作为填补模型处理可理解为残差一些次要特征 。

之前我们提到的一些被处理为异常值的指标点（往往表示DDOS-IN,DDOS-OUT的事件，IaaS平台这类事情每时每刻都会发生），则采用Changepoint作为核函数来处理。

采用梯度下降得到的模型参数，其评价采用了贝叶斯信息准则计算。限于时间不再展开，参考如下：

Jefferys, William H.; Berger, James O. (199"Ockham's Razor and Bayesian Statistics (preprint available as "Sharpening Occam's Razor on a Bayesian Strop"]" (PDF). American Scientist. 80: 64 – 72.

$$BIC(M)=-2\log p(D|M)+p\log n$$

谢谢大家!

- 我们正在组建强大的OpsDev团队
- 我们在实现云计算运维的工程控制方法

- *Soft define engineering process @IaaS*
- *软件定义工程化*

- 请用简历砸晕我 jim.xu@ucloud.cn

Thanks

高效运维社区

开放运维联盟

荣誉出品

想第一时间看到高效运维公众号的好文章么？

请打开高效运维公众号，点击右上角小人，并如右侧所示设置即可：

GOPS2016 全球运维大会更多精彩

GOPS2016

全球运维大会·北京站

2016 年12月16日-17日

北京国际会议中心