
SACC
SACC
2021
2021 中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2021
数字转型 架构重塑
ITT168.com
ChinaUnix.net
ITPUB
云上会议 网络直播 | 2021.5.20-2021.5.22

使用夜莺构建混合云场景的监控

重新定义国产开源监控

秦晓辉 18612185520（微信同号）

先后履职于百度、小米、金山云、滴滴，十年开发、运维从业经验，互联网监控解决方案Open-Falcon、Nightingale开源发起者之一、开源PaaS平台DINP作者，在滴滴负责产业云技术中心，推动滴滴内部平台能力商业化输出

个人公众号：运维散兵

大纲

- **需求来源**：监控的原始需求来自业务稳定性
- **产品要求**：端上、链路、资源、组件、应用多维度跨云监控
- **当下现状**：细数现有的监控体系，横评对比
- **夜莺介绍**：国产开源、云原生、大规模、好用到爆
- **功能拆解**：详细拆解夜莺各个方面的能力
- **社区介绍**：介绍社区情况，欢迎大家共建国产监控社区

需求来源

监控的原始需求来自业务稳定性

如果贵司的业务强依赖IT技术，IT故障会直接影响营业收入，稳定性体系一定要重视起来，而监控，就是稳定性体系中至关重要的一环

需求来源 监控的原始需求来自业务稳定性

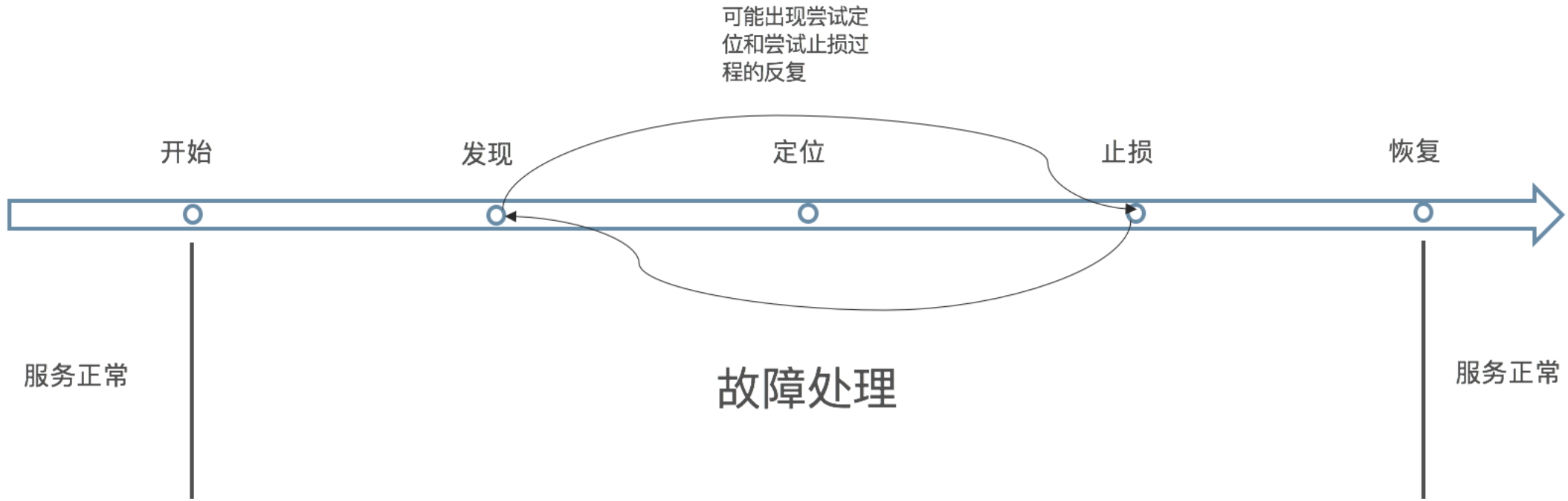
左图是2013年的一个新闻，讲Google宕机的影响。2020年也出现过aws大规模宕机的情况，影响不止是55万美元，直接影响大半个互联网！

2018年有美国调研机构指出，如果服务器宕机1分钟，银行会损失27万美元，制造业会损失42万美元

如果美团故障？滴滴故障？

需求来源 监控的原始需求来自业务稳定性

如何减少服务停摆导致的经济损失？尽快发现故障并止损！故障处理过程中，监控是『发现』和『定位』两个环节的关键工具。故障处理过程的首要原则是『止损』因此过程中的『发现』和『定位』都是面向尽快『止损』来实现

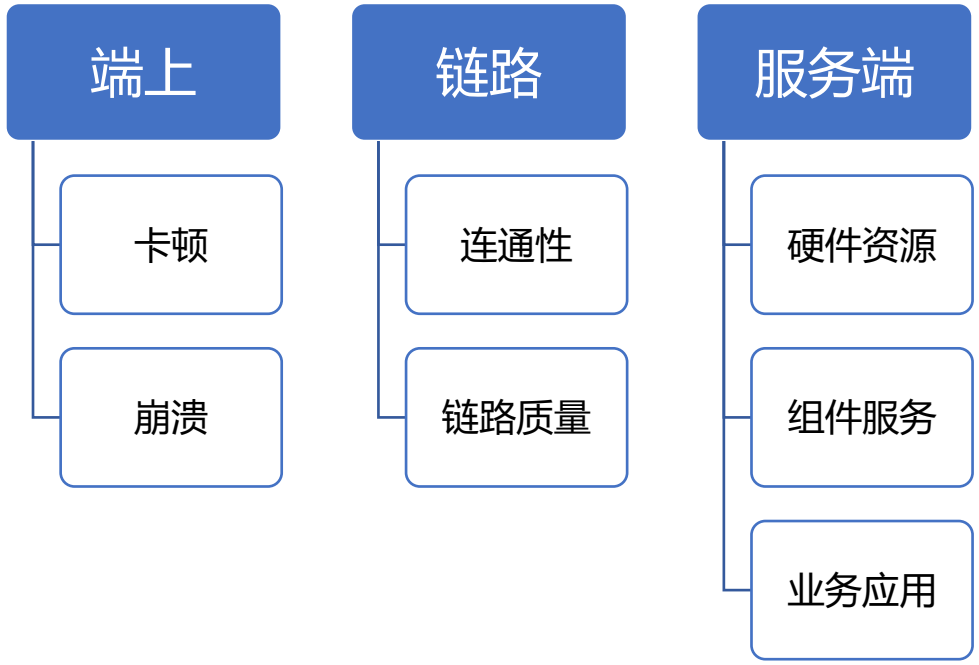


产品要求

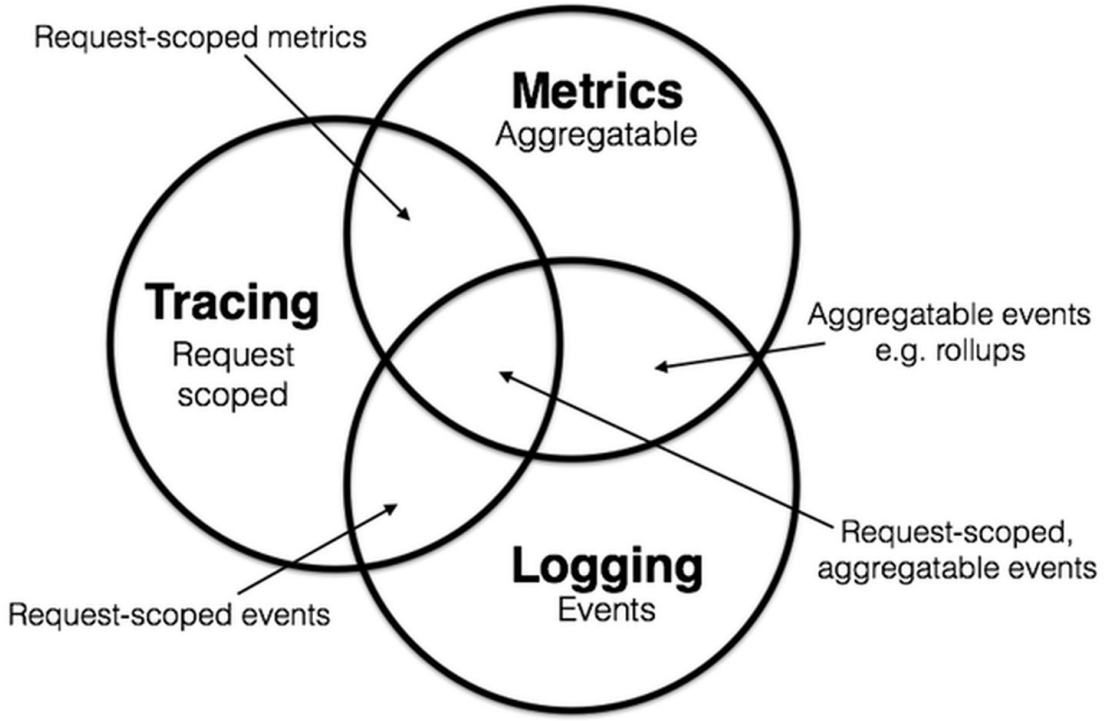
端上、链路、资源、组件、应用多维度跨云监控

我们希望有个银弹，不管哪个环节出问题，都能及时监控感知，同时搞定公有云、私有云、混合云场景。总结两个词：完备性、跨云

产品要求 端上、链路、资源、组件、应用多维度跨云监控



监控目标要完备



监控手段要完备

产品要求 端上、链路、资源、组件、应用多维度跨云监控

大部分企业都在上云，大部分企业都不止用一家云。

公有云的监控偏封闭、私有云的监控偏简单，我们迫切需要一个跨多云的、中立的、功能完备的可观测性体系

当下现状

细数现有的监控体系，横评对比

开源社区在metrics、logging、tracing不同方向都有对应的产品来解决，没有开箱即用的组合方案，当然，这也是因为受限于各公司的研发体系标准不一

当下现状 细数现有的监控体系，横评对比

ZIPKIN

Skywalking

ZABBIX

Thanos

Nagios

falcon

graphite

elastic

fluentd

Grafana loki

当下现状 细数现有的监控体系，横评对比 – metrics监控之时序数据存储

metrics监控通常是首要搭建的系统，存储成本低，效果显著，metrics监控的难点是时序数据的存储，特别是移动互联网时代的到来，APM层面的监控需求暴涨，时序数据暴增，行业里涌现出了众多解决方案，如果贵司有基础架构人力，基于这些时序库开发metrics监控，也不失为一个好选择



当然，更好的选择是基于某个开源metrics监控做改造，比如Open-Falcon开源之后，就有众多厂商基于它做了二次开发，接下来要介绍的Nightingale，可以看做是Open-Falcon的下一代，拥抱新的时序库，拥抱Prom生态，强烈建议大家试用一下

夜莺介绍

国产开源、云原生、大规模、好用到爆

国产开源监控产品相对比较匮乏，夜莺希望重新定义国产开源监控，支持云原生监控，经受了滴滴大规模生产检验，另外，好用到爆！

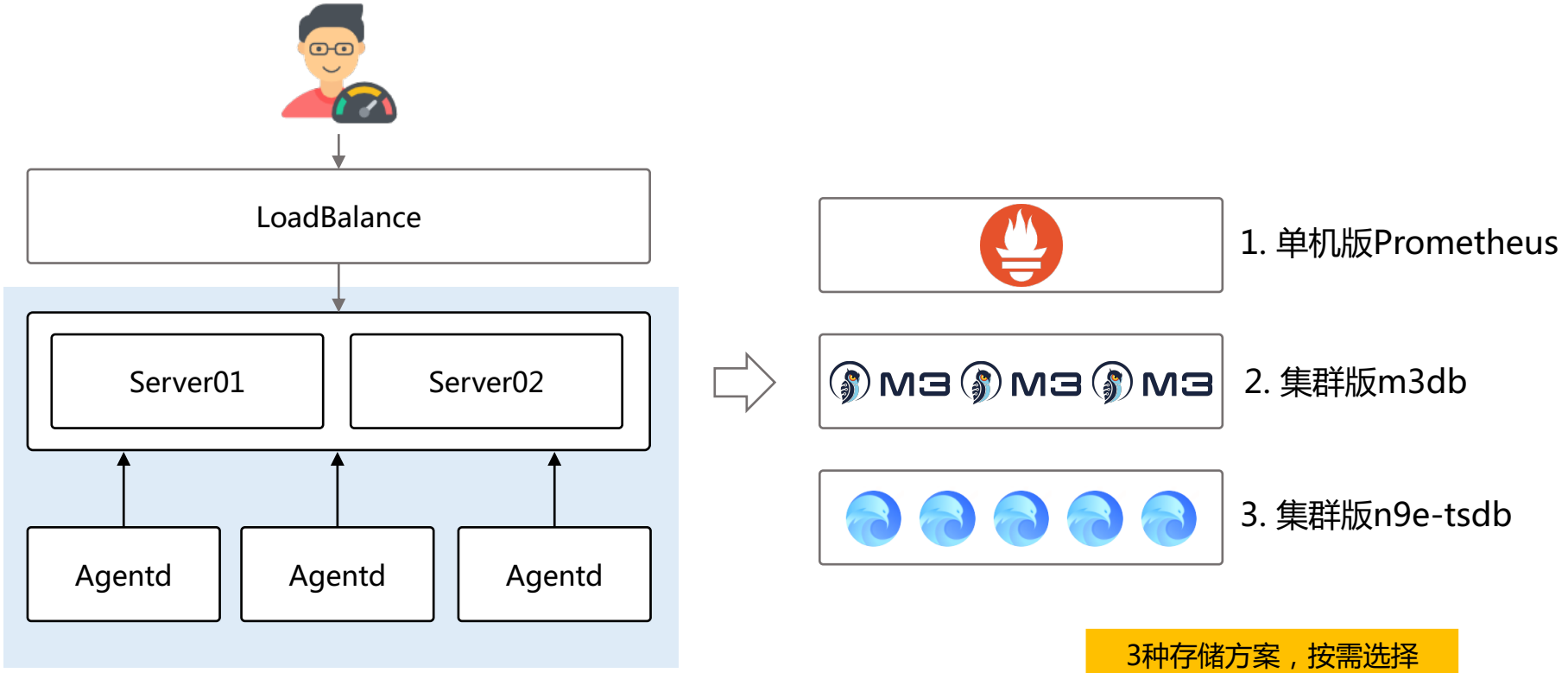
Nightingale

夜莺是**新一代国产智能监控平台**，既可以解决传统物理机虚拟机的场景，也可以解决容器的场景。衍生自Open-Falcon和滴滴Odin监控，经受了包括小米、美团、滴滴在内的数百家企业的生产环境验证，**简单可依赖，好用到爆！**

3000+	600+	500+
star	issue	fork

项目：https://github.com/didi/nightingale 官网：https://n9e.didiyun.com/

夜莺介绍 架构概览，核心组件只有2个，存储后端可插拔式设计



功能拆解

详细拆解夜莺各个方面的能力

监控系统的核心功能，是数据采集、存储、分析、展示，

难点在服务端，完备性看采集能力，是否能够兼容并包，

纳入更多生态的能力，至关重要

夜莺功能拆解 监控数据采集，All in one的Agentd

- 支持在web上配置采集策略，不同的采集可以指定不同的探针机器、目标机器。便于管理和知识传承
- 独创在端上流式读取日志，根据正则提取指标的机制，轻量易用，无业务侵入性
- 内置集成了多种数据库中间件的采集以及网络设备的采集，复用telegraf和datadog-agent的能力
- 支持statsd的udp协议，用于业务应用的apm监控分析

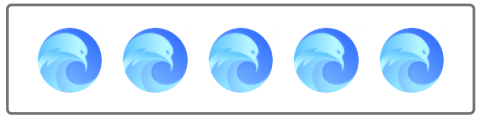
夜莺功能拆解 监控数据采集，轻松复用Prometheus生态各exporter

把Prometheus当做采集器，通过remote_write接口，写入M3，把Nightingale当做数据消费方，共用M3存储，nightingale-server中内置promql引擎。通过Prometheus生态采集的数据使用promql消费，通过Nightingale生态采集的数据使用内置引擎（或promql）消费

夜莺功能拆解 监控数据存储，支持多种后端，可插拔式设计

需要保存原始精确数据，对容灾没有太高要求，单机快速测试；对于小规模场景通常够用，500台设备以内的公司，用这种方式非常简单，把数据存储目录置于网盘上，容灾能力也有了一定的保障

需要保存原始精确数据，并且对容灾有要求，可以选择集群版本的m3db，扩容方便，有副本机制，挂掉一个存储节点，不影响整体服务。当然， 既然保存原始数据，对硬盘容量消耗会大，也没法直接查看年级别的数据，返回数据量太大会把浏览器打崩



不需要保存原始精确数据，想查看年级别的采样数据，分析历史趋势，又对硬盘存储空间较为敏感，可以选择n9e-tsdb，底层存储使用rrdtool，环形数据库，自动降采样

夜莺功能拆解 两套告警规则引擎，满足不同层次的用户需求

- 基于资源标识和指标标签的简单匹配，指标预计算，直观易于上手，能解决大部分场景的问题
- 提供资源分组机制，分组支持前缀匹配，特别适合处理传统物理机虚拟机时代服务混部的场景
- 弱化告警事件的处理逻辑，未来会拆出单独的产品来解决告警事件的问题，类似omnibus、bigpanda

- 内置了promql引擎，可以复用prom的各类计算函数，指标后计算，提供更大的灵活性
- 查询时即时聚合会带来内存开销的问题，高基数大查询可能会直接导致进程OOM，需自行把握
- 对于有基础架构研发投入的公司，建议对大量数据的聚合计算采用流式计算方式来处理

夜莺功能拆解 何为“提供资源分组机制，分组支持前缀匹配”

资源分组1：n9e.server

下面挂载2台机器资源：

c3-sre-n9e-server01.bj

c3-sre-n9e-server02.bj

资源分组2：n9e.prober

下面挂载2台机器资源：

c3-sre-n9e-prober01.bj

c3-sre-n9e-prober02.bj

为n9e.server和n9e.prober两个分组分别配置进程、端口监控采集，看起来很顺畅

如果要为所有n9e的机器配置一个插件脚本呢？就需要每个分组单独配置，如果要查询n9e的所有机器呢？要分别查询两个分组的机器并做merge :-(

资源分组3：n9e. 名称是资源分组1和2的前缀部分

下面挂载0台机器资源：

新建一个空组，不挂任何资源，名称有讲究，是资源分组1和2的前缀。于是：

统一的监控插件就可以配置到n9e.这个空组上，前缀匹配的方式生效到下面所有的子组上，获取机器资源列表，也是前缀匹配的方式获取。其实，组名就是一个树形结构！

夜莺功能拆解 为何要“弱化告警事件的处理逻辑”

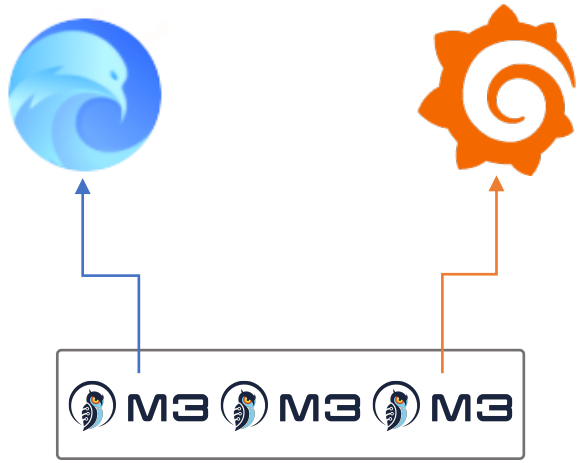
告警合并	告警认领	告警升级	排班处理	关联分析
告警过滤	事件规整	统计报表	全局工作台	处理工作流

- 告警事件的处理实际是一个单独的领域，很重要很通用，值得当做一个单独的产品来对待处理
- 不止夜莺会产生告警事件，其他监控系统也会产生，ipmi、snmp trap等也会产生，这是一个共性的事情
- 从全局分析的角度，只有全部事件汇聚在一起分析的价值才大，单系统的事件分析略显单薄
- 实际上，各个公司，很少有只使用一套监控系统包打天下的，一定会有多套，事件源一定会是多个

夜莺功能拆解 怎么做到支持所有通知媒介

- 告警事件的通知，夜莺还是会保留这个能力，毕竟不是所有的公司都有一套完备的告警事件处理平台
- 怎么做到支持所有通知媒介？放弃原来的版本的很多内置处理逻辑，抽出一个单独的脚本，具备最大的灵活性的，一定是代码，用户可以轻易定制这个发送脚本（官方提供默认版本）

夜莺功能拆解 展示方面可以使用内置的看图，也可以复用Grafana



- 对于Prometheus或者M3，都可以直接对接Grafana，这个版本的夜莺不再单独开发Grafana的DataSource。Prometheus生态上报的数据可以直接查看，夜莺生态上报的预计算的数据也可以走promql查看
- 夜莺自身也会提供看图的能力，内置监控大盘、即时看图等相关功能，样式没那么好看，胜在看大量数据时性能更好

社区介绍

介绍社区情况，欢迎大家共建国产监控社区

夜莺社区当前差不多有1000家企业在尝试，部分已经上到生产环境，数十个外部贡献者，欢迎大家共建，把国产开源监控这个事情，咱们一起做到极致

夜莺社区情况 众多企业已上生产，共同打磨夜莺

中泰证券 ZHONGTAI SECURITIES
Envision
广西民族大学 Guangxi University for Nationalities
PANSHI 盘石 全球新经济平台
GPU Cloud
河北工业大学 HEBEI UNIVERSITY
魏 WEIJIA
AsiaInfo 亚信科技
EAST CHINA NORMAL UNIVERSITY 华东师范大学
掌阅 iReader
iPi
网联清算
跟谁学 好课就在跟谁学
闪送
SiSYPHE 西西弗书店 PARK BOOKS & UP COFFEE
N@XIN 农信数据
Joybos 佳帮手 专注自领家居生活
funstory.ai
海豚新媒体 DOLPHIN NEW MEDIA
招采猫 BIDCAT.CN
LONGFOR 龙湖
浙江日报报业集团 ZHEJIANG DAILY PRESS GROUP
转转 二手交易网
朗森特科技 LANCET TECHNOLOGY
友谊时光 FRIENDTIMES
微餐 WCT 时代
阅信云
途游游戏
易车
由创科技 UCREATER
China unicom 中国联通
无锡恒鼎超级计算有限公司
U-MOOCS
SANGFOR 深信服科技
赞播智店


SACC
THANKS