目 录

- 为什么自主开发?
- 我们的系统
- 架构设计
- 特点
- 未来计划

目前开源监控系统存在的问题

- 只有“监”的功能，没有“控”的功能
- 当出现故障时，往往会产生”报警风暴”
- 监控数据都是孤立的，无法产生更多的价值
- 缺少根源分析
- 效率低(性能、配置的工作量)
- 。。。。。。

- 在50万-100万台设备规模的环境下，上面这些问题尤其突出。我们需要在一个平台中能够支持这么多设备的监和控。

目标

- 眼：
  - 使得该监控系统能够很好地成为我们的“眼”，帮我们7*24小时监控我们整个系统
- 脑：
  - 使得该系统成为我们的“脑”，能够根据历史的经验和知识库给我们诊断的建议和历史处理该问题的经验，方便我们快速的诊断和解决问题。
  - 为我们提供问题根源分析。
- 手：
  - 使得该系统成为我们的“手”，方便我们快速地进行操作
  - 将运维的方式由黑屏变白屏，成为我们的运维操作平台。

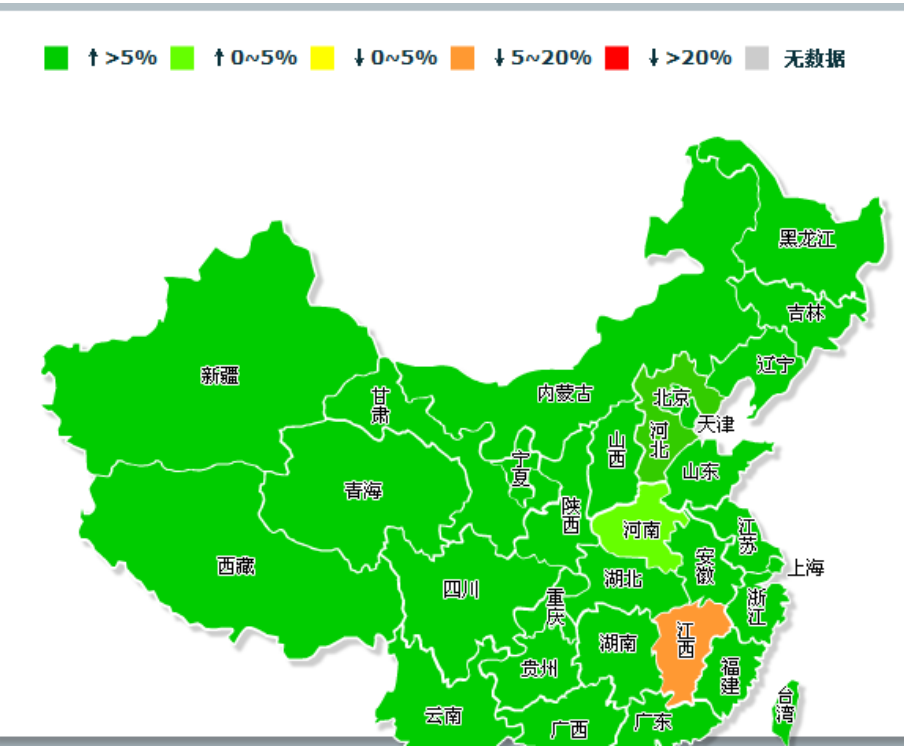
目 录

- 为什么自主开发?
- 我们的系统
- 架构设计
- 特点
- 未来计划

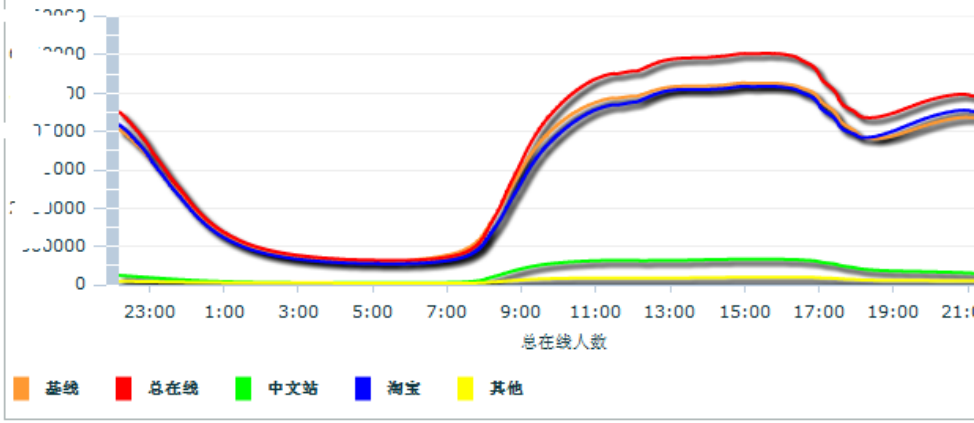
我的报警 | 所有报警 | 盯屏

集群名	监控项		状态	状态信息	持续时间	预警次数	Warning阀值	Critical阀值	最后检测时间
AY21	check_[illegible]an_		Critical	total=67, ok=64, exception=3, exception_close=0	6h 5m 35s	626/3	3	3	10-09-26 18:19:1

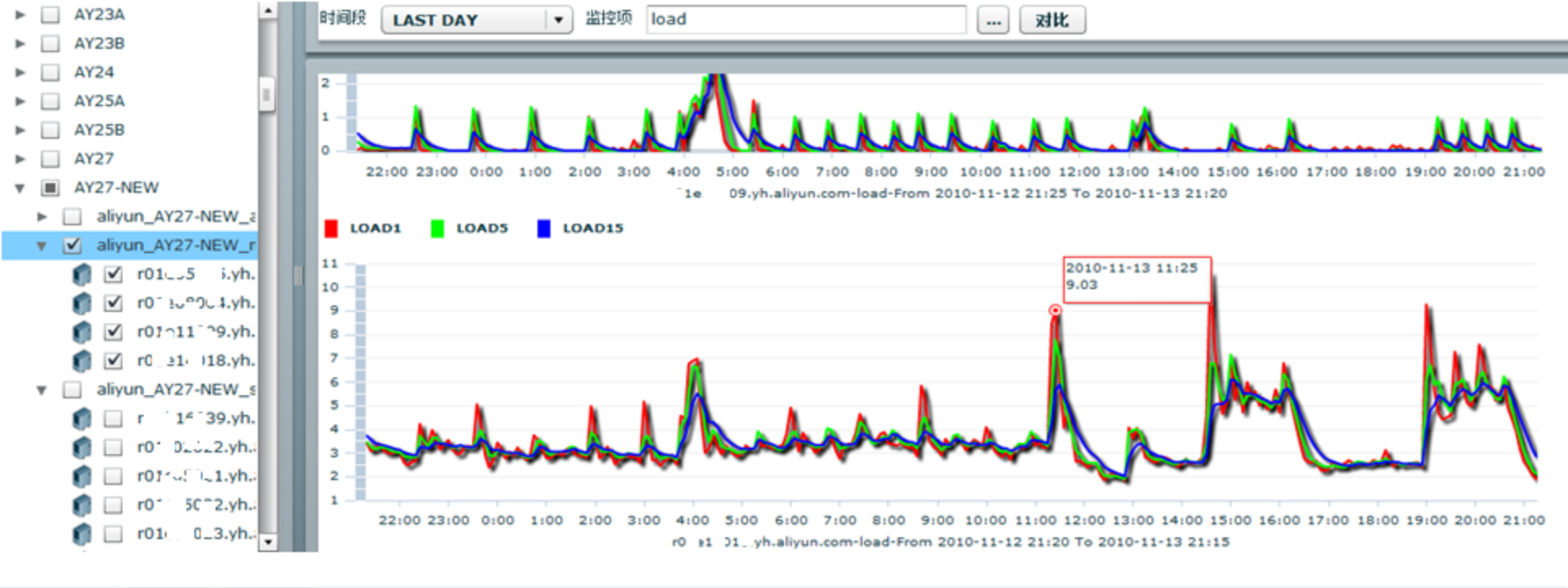
主机名	IP	设备组	监控项		状态	状态信息	持续时间	预警次数	最后检测时间
r0[illegible]019.yh.aliy	[illegible]82.[illegible]	aliyun_ay21_s	check_feitian_		Warning	Warning - disk usage: 81%	6h 5m 14s	73/2	10-09-26 18:14:4
dw-ctr1.growing.a	[illegible]3.121	aliyun_growin	check_logwatc		Warning	Warning - Err|130228|betterman.chenf@alibaba-inc.com|BEGIN x	14h 59m 9s	1/1	10-09-26 04:20:1
r0[illegible].yh.aliy	[illegible]82.[illegible]	aliyun_ay21_s	check_feitian_		Warning	Warning - disk usage: 82%	2d 10h 29m 31s	702/2	10-09-26 18:17:2
csw-lab-c6509.sd	[illegible]62.[illegible]	aliyun_idcnet_	Check_Csw_C		Warning	WARNING - CPU used 75%	5d 0.h 1.m 45s	2405/2	10-09-26 18:16:3
r0[illegible]5.yh.aliy	[illegible].69.[illegible]	aliyun_ay30_a	check_logwatc		Warning	warning - daogou_dump_workflow_alert: dump question failed (	6d 4h 22m 2s	7/1	10-09-24 13:47:5
r0[illegible]21.yh.aliy	[illegible]9.82.[illegible]	aliyun_ay21_s	check_feitian_		Warning	Warning - disk usage: 83%	10d 1h 10m 34s	2894/2	10-09-26 18:16:1
r0[illegible]0.yh.aliy	[illegible]71.[illegible]	AT-YHHOUYI_s	get_hwconfig		Miss		93d 6h 57m 24s	76/3	10-07-01 18:42:1
r[illegible]041.yh.aliy	[illegible]9.106	aliyun_ay08-n	check_feitian_		Miss		166d 23h 14m 28s	10/2	10-04-15 19:07:5

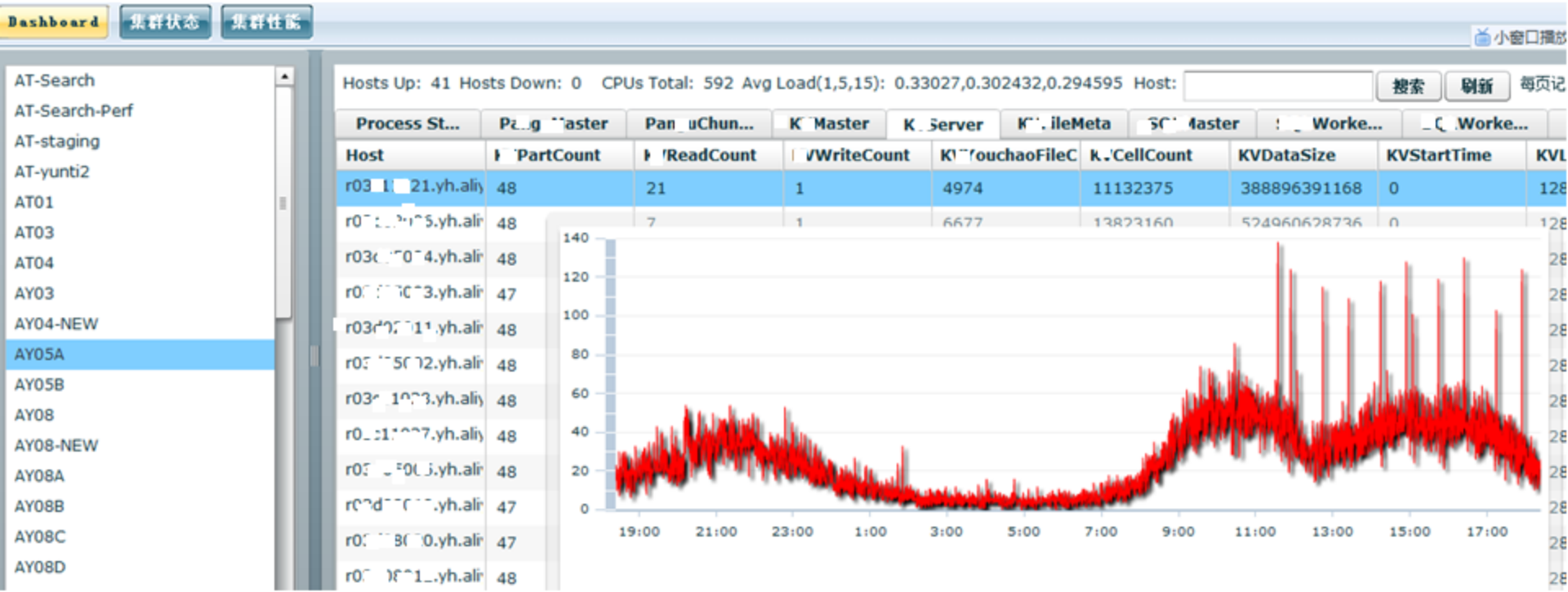


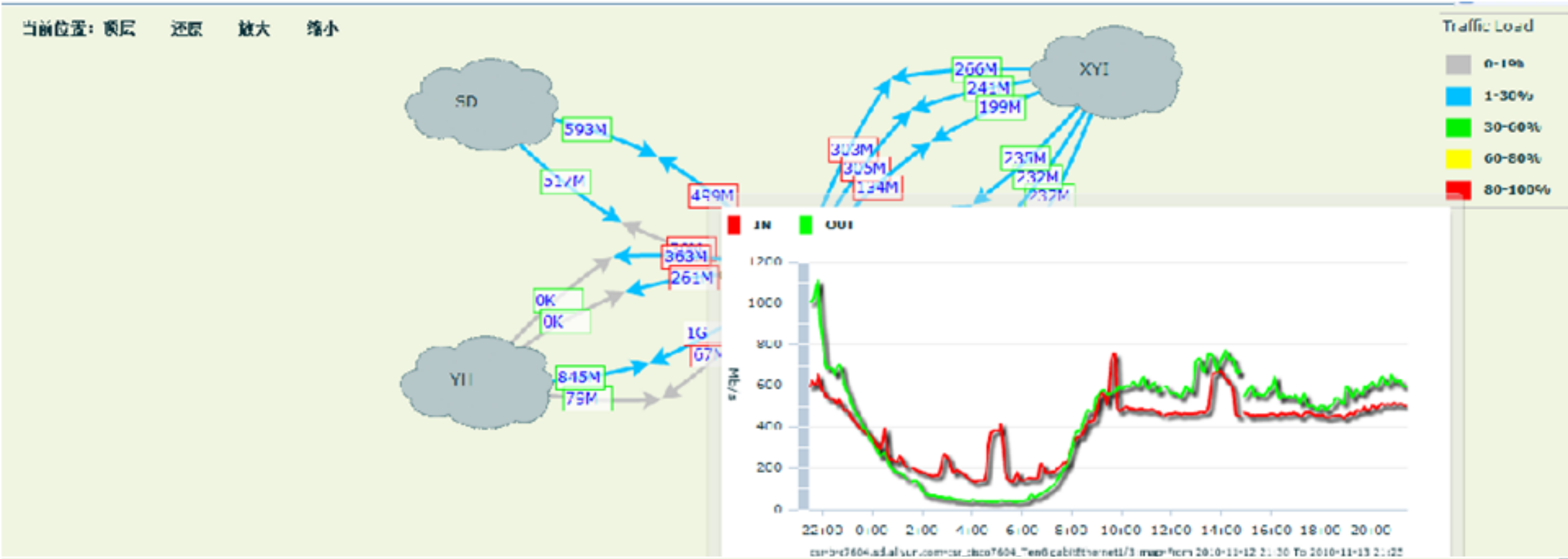
一天在线数实时曲线 | 旺旺在线数实时曲线 | 旺旺登录登出数曲线

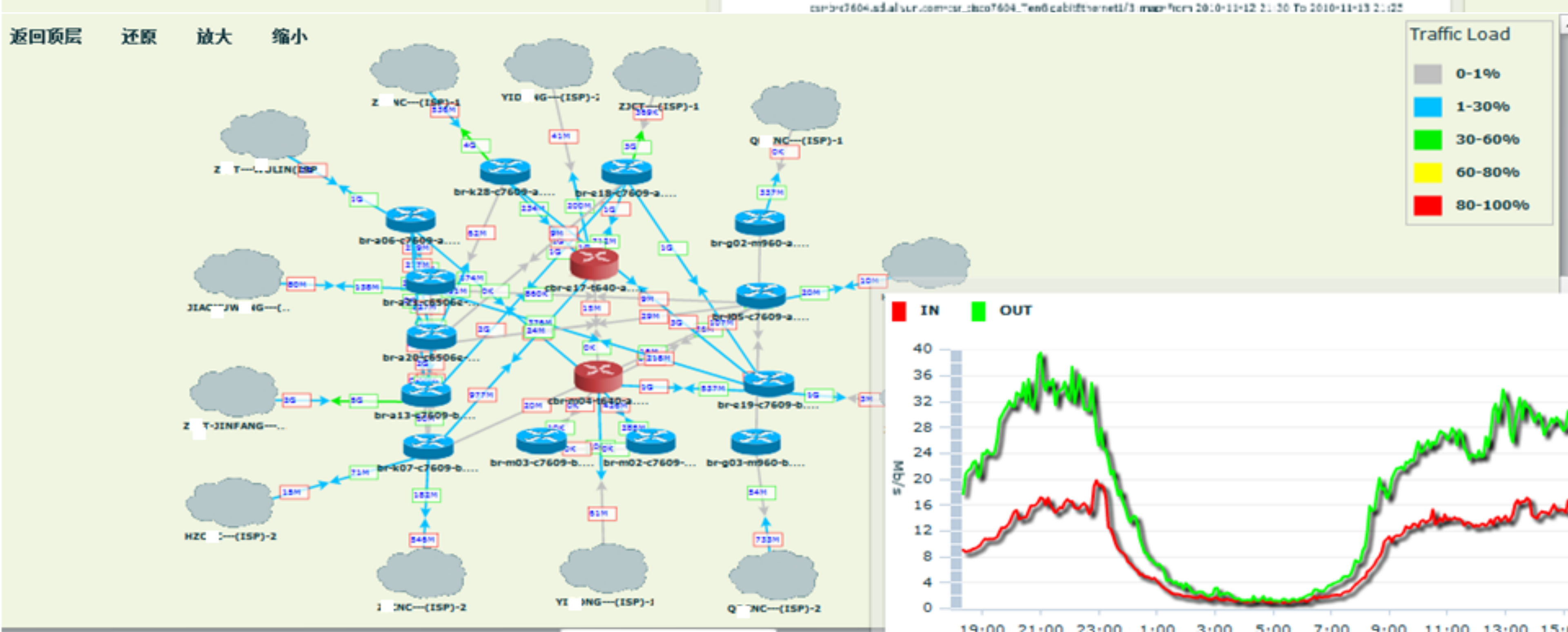


城市	基线值	总在线 ▼	中文站	淘宝	其他	更新时间
历史最高人数		[illegible]0	[illegible]0	[illegible]8	[illegible]42	2010-[illegible]5:
手机在线	[illegible]1	[illegible]2	[illegible]5	[illegible].19	[illegible]08	2010-12-07 22:
总在线人数	[illegible]'57	[illegible]19	[illegible].386	[illegible]53	[illegible]70	2010-12-07 22:
广东	[illegible]34	[illegible]52	[illegible]4	[illegible]3	[illegible]6	2010-12-07 22:
浙江	[illegible]81	[illegible]10	[illegible]3	[illegible]7	150	2010-12-07 22:


AY23A
AY23B
AY24
AY25A
AY25B
AY27
AY27-NEW
aliyun_AY27-NEW_a
aliyun_AY27-NEW_r
aliyun_AY27-NEW_s
时间段
LAST DAY
监控项
load
...
对比
09.yh.aliyun.com-load-From 2010-11-12 21:25 To 2010-11-13 21:20
LOAD1
LOAD5
LOAD15
2010-11-13 11:25
9.03
yh.aliyun.com-load-From 2010-11-12 21:20 To 2010-11-13 21:15


Dashboard
集群状态
集群性能
小窗口播放
AT-Search
AT-Search-Perf
AT-staging
AT-yunti2
AT01
AT03
AT04
AY03
AY04-NEW
AY05A
AY05B
AY08
AY08-NEW
AY08A
AY08B
AY08C
AY08D
Hosts Up: 41 Hosts Down: 0 CPUs Total: 592 Avg Load(1,5,15): 0.33027,0.302432,0.294595 Host:
搜索
刷新
每页记
Process St...
Master
Server
Worke...
Host
PartCount
ReadCount
WriteCount
CellCount
KVDataSize
KVStartTime

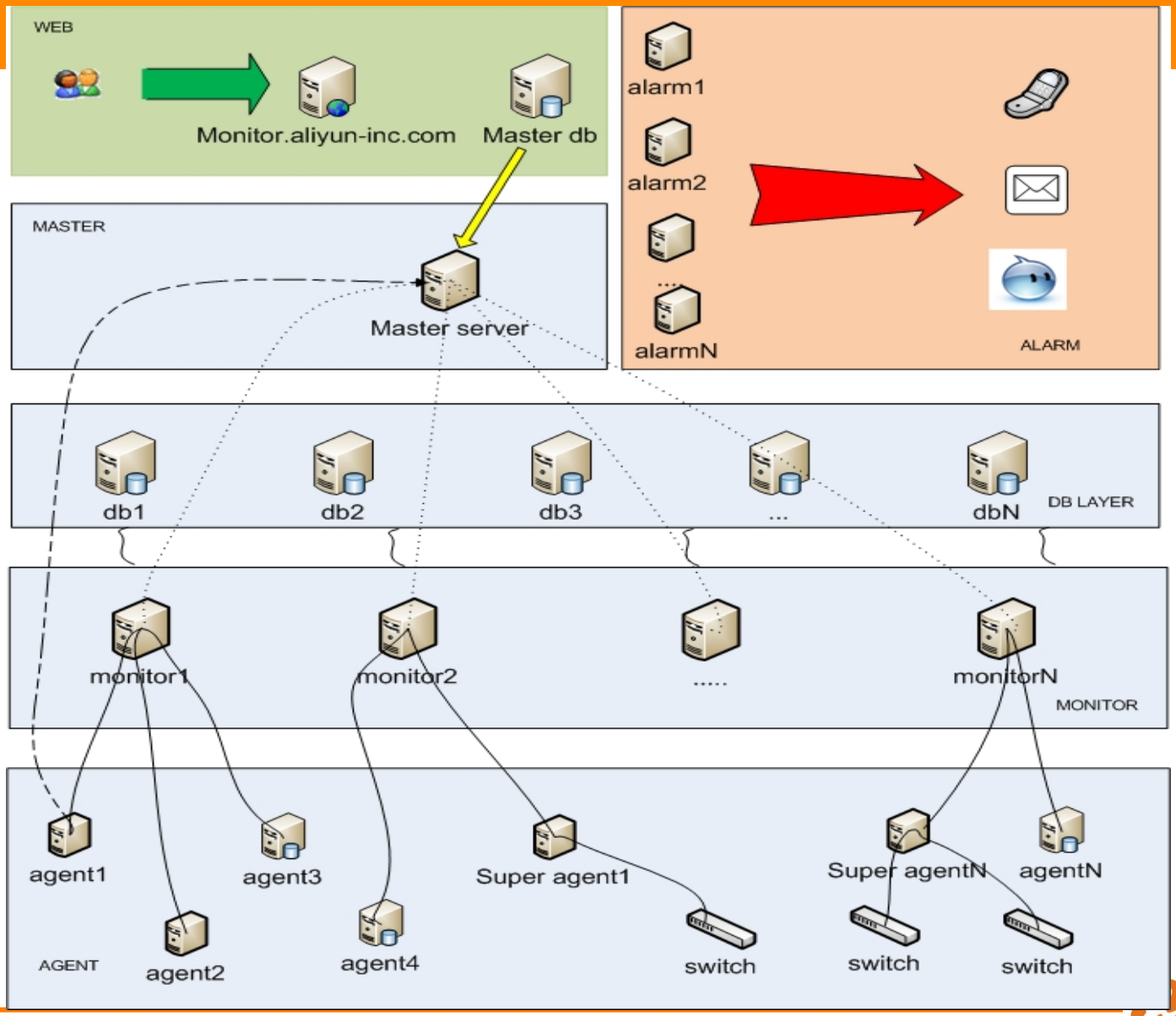

当前位置：顶层
还原
放大
缩小
Traffic Load
0-1%
1-30%
30-60%
60-80%
80-100%
SD
XYI
YII
593M
517M
499M
266M
241M
199M
303M
305M
134M
235M
232M
363M
261M
0K
0K
1G
845M
79M
IN
OUT
Mb/s
1200
1000
800
600
400
200
0
22:00
0:00
2:00
4:00
6:00
8:00
10:00
12:00
14:00
16:00
18:00
20:00


返回顶层
还原
放大
缩小
Traffic Load
0-1%
1-30%
30-60%
60-80%
80-100%
br-k28-c7609-a....
br-e18-c7609-a....
br-a06-c7609-a....
br-g02-m960-a....
br-i05-c7609-a....
br-e19-c7609-b....
br-a13-c7609-b....
br-k07-c7609-b....
br-m03-c7609-b....
br-m02-c7609-....
br-g03-m960-b....
IN
OUT
Mb/s
40
36
32
28
24
20
16
12
8
4
0
19:00
21:00
23:00
1:00
3:00
5:00
7:00
9:00
11:00
13:00

目 录

- 为什么自主开发?
- 我们的系统
- 架构设计
- 特点
- 未来计划

鷹眼架构



整体的架构设计

- master 负责monitor, agent 注册并监控Master DB表的变化并将这个变化传播给monitor。
- monitor 分为多个进程，利用共享内存交互，负责接收agent传递上来的数据，根据状态进行报警，并把异常事件、当前状态、性能数据入库。
- agent 主要负责调度本地的插件，并把插件执行的结果传递给 monitor 。
- db layer 存放所有的元数据和监控数据，并保证agent 切换到其他 monitor 之后数据还是连续的。Agent与DB是多对一的关系，并保持不变。
- alarm 主要是从master 的异常表抢占式的取出异常数据，并报警

监控设计的关键点

- **去单点， 可扩展，自我监控**

- Master一组，热备，自动切换。
- Monitor 单台，出现了问题，可以快速迁移agent到另外的monitor。
- Master 能够发现monitor 机器上任意进程的问题，通过心跳来实现。发现5次心跳错误就报警。
- Super agent单台，出现了问题，可以快速迁移agent到另外的super agent。
- monitor 会发现super agent 的问题，报警。
- DB采用自主研发的HA Aurora做热备。如果lister进程发现db有问题，数据先记录到文件。
- Monitor 发现 master的故障，报警，继续自己与agent之间的工作，同时启线程连master。受影响的只有新连接的agent

目 录

- 为什么自主开发?
- 我们的系统
- 架构设计
- 特点
- 未来计划

监控范围广

- 监控面广泛，从数据中心最低层到最高层应用
  - 机房环境，如电力，空调,能耗，湿度
  - 网络设备
  - 服务器
  - 重要基础设施，DNS，VIP，DB等等
  - 应用监控
  - 集群监控，如集群的各种性能指标，集群的监控状态

高性能及良好的扩展性

- 监控效率非常高，每台监控机可以监控20万监控点，采用主动+被动
- 集成了轻量级的icmp，ntp，snmp，ipmi协议，避免fork进程
- 采用epoll+半异步/半同步的线程池
- 分布式架构，支持100万台设备的规模

监控自动部署

- 定义监控模板，监控模板包括监控项，报警策略，报警人
- 与CMDB打通，所有被监控设备以及设备的组织和关系均来自于CMDB。
- 将设备组与监控模板关联。n:n的关系。所有在该设备组内的机器都具有监控模板里面的监控项。 用户只需要操作CMDB即可。
- 允许增加个性化的监控项(不在监控模板中)

根源分析

- 以service监控入手，将与某个service相关联的各种设备和监控项层层建立服务依赖（目前人工建立这种依赖关系，这种服务依赖不是nagios中报警依赖）。
- 将这种依赖关系以拓朴的形式展现在一张图上。
- 当某个服务出现问题时，通过这种依赖，快速的定位到相关联的监控项或者设备组。

控

- 重启物理机
- 增加安全清洗策略
- 重启VM
- 重启云计算OS
- 执行各种事先好的命令

提高报警的准确性

- 合并报警，抑制报警风暴
  - 单位时间内某个报警接收人的报警超过6条时，程序会采取报警合并，将设备组、监控项、状态相同的报警合并成一条报给用户
  - 合并后的旺旺、短信将告知有多少条相似的报警被合并，而邮件中将列出这些报警的详细信息。
- 关联报警
  - 定义一些规则，实现关联报警，比如当load超过某个阈值时，所有监控项的timeout不要报警，但异常信息还是要报。

采用数据库替换RRD文件

- RRD文件具有两个缺点：
  - 一张性能图一个小文件，文件数多
  - 小IO，对IO性能要求较高

- 我们采用JAVA实现RRD算法，用MySQL数据库来代替RRD的存储，大大提升性能，数据得以集中管理，也很容易多张RRD合并成一张

文件形式接口

- 一个开放的接口。它的工作原理是：
  鹰眼Agent进程会实时检测一个固定文件的行增量变化，并把增量消息发送给前端monitor。

- 这个日志文件每一行内容的格式如下：
  - Critical-xxxx
  - Warning-xxxx
  - Ok-xxxx

- 如果行头是“Warning -”，“Critical -”的消息，则agent会将这些消息报给logwatch指定的报警组。

报警处理case库和历史报警

- 可以查看每一个监控项的历史报警和历史操作
- 可以查看每一个监控项的历史报警的原因、处理方法，形成case库，帮助处理人员快速定位问题和解决问题

鹰眼Agent升级和插件升级

- 鹰眼Agent可以自动升级。
- 插件也可以自动升级。当插件变更，或者增加新的插件时，通过插件升级机制可以将插件推到所有的机器上面。

监控数据分析

- 由于元数据和异常数据的集中、状态数据和性能数据库化存储，使得监控数据分析非常方便
- 报警分析、异常数据按集群分析、每个监控项的可用率的自动计算、集群平均性能指标等等
- ……

目 录

- 为什么自主开发?
- 我们的系统
- 架构设计
- 特点
- 未来计划

未来计划

- 应用接口级别监控
- 网络设备监控的自动部署
- 取代crontab
- 配置管理的Agent，取代cfengine
- .....

Q/A?