

他趣大数据应用实践

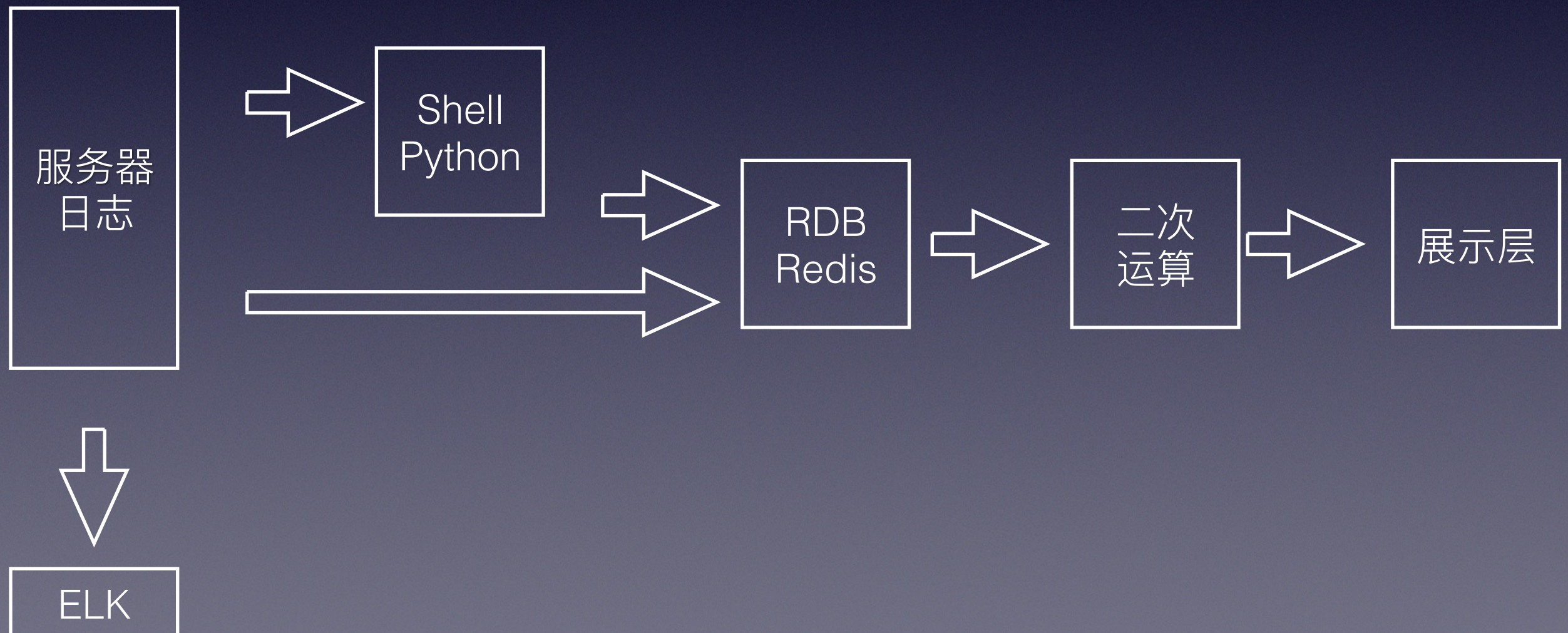
by 张嵩

目录

- 业务倒逼，基础选型
- 发展历程，步步简化
- 日志实践，少走弯路

初始

- 小数据? RDB、Redis、Shell、Python、ELK



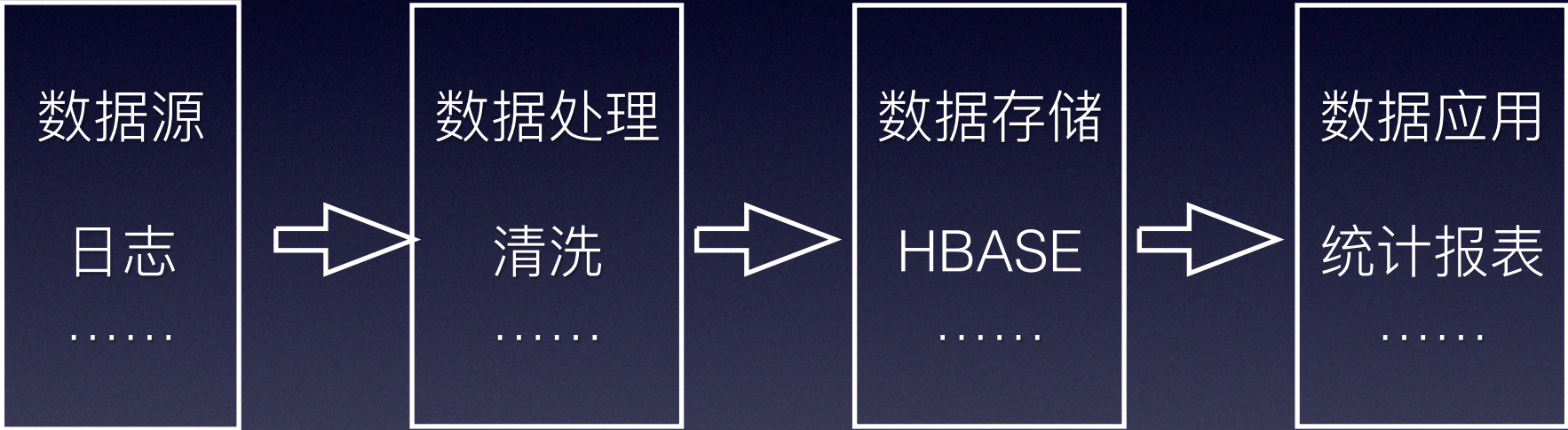
问题：

- 直接统计出结果数据集过于庞大，新增不易
- 二次运算复杂度取决于空间的维度
- 可接受的时间复杂度
- 集群维护复杂度

例如：

查询来自android，渠道为xiaomi，版本为6443，
时间粒度为5分钟级别

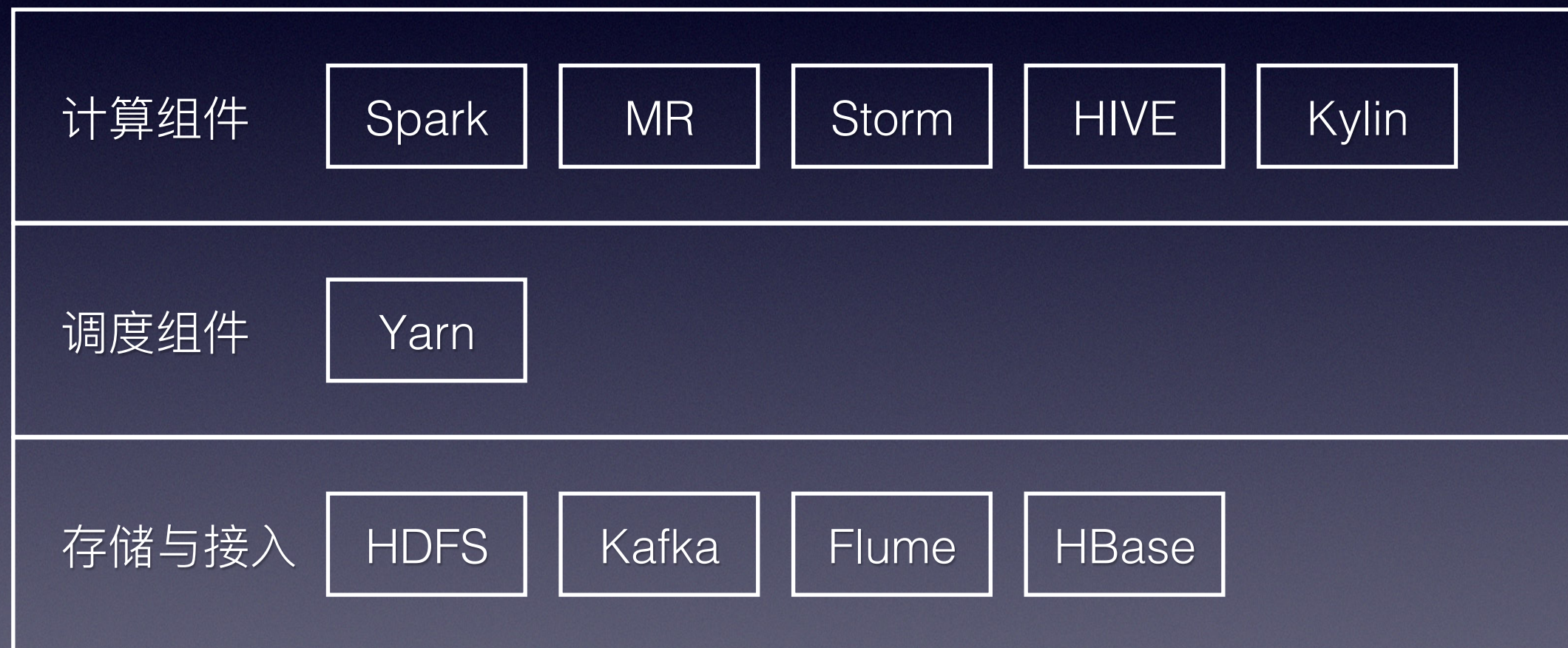
整理需求



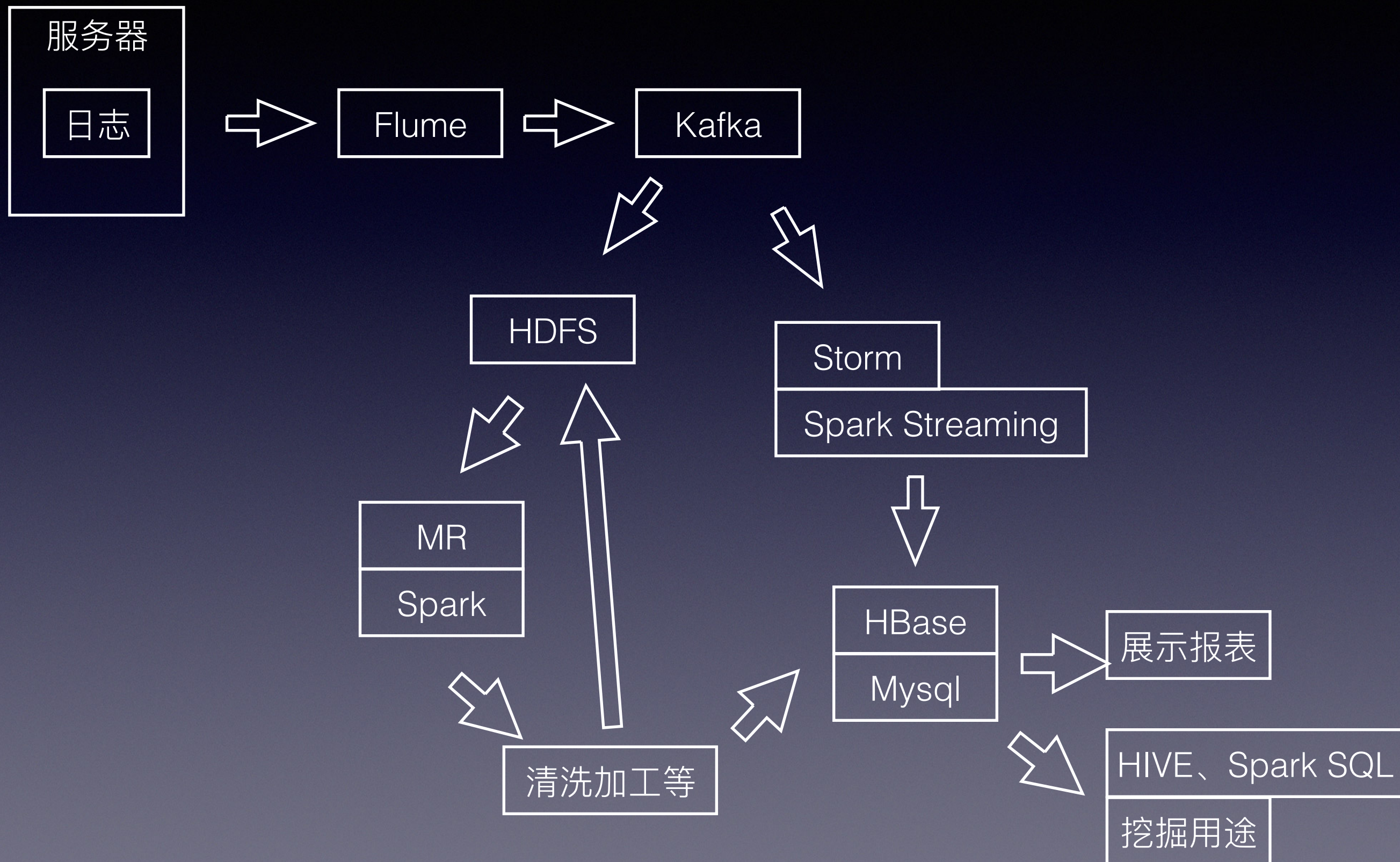
需求细化



- 一代结构



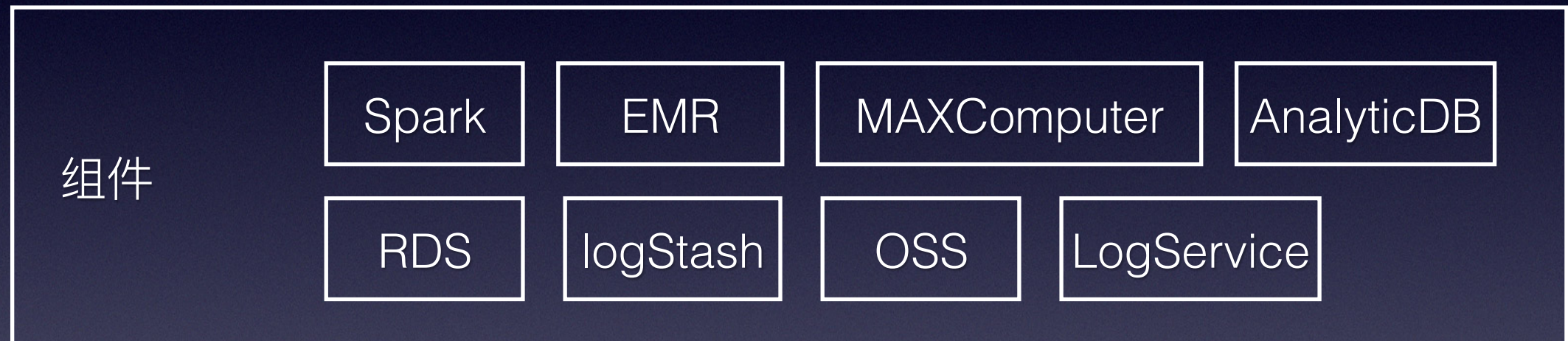
• 一代流程图



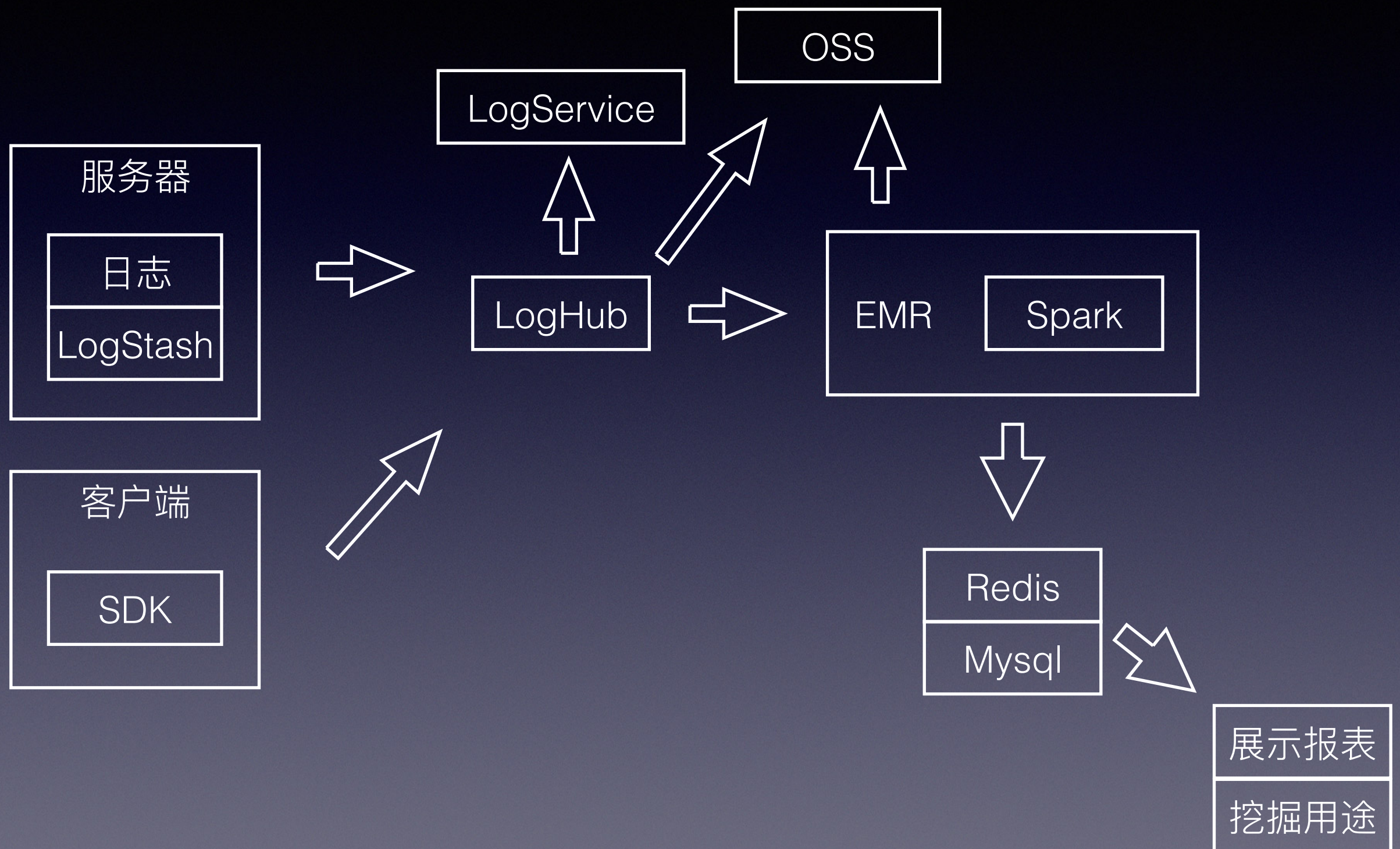
一些小经验

- 不建议业务小型时自行搭建
- 基础组件搭配兼容问题不少
- 费用及人力成本与收益长期(扩容)来看较差
- 搭起来、跑起来仅仅是个开始
- batchjob、streaming资源、核心非核心资源的竞争
- 作业的管理，特别是依赖
- 多维计算可以上Kylin，速度会有惊喜
- FlumeCenter的扩展性要自己做好

- 二代结构



- 二代流程图



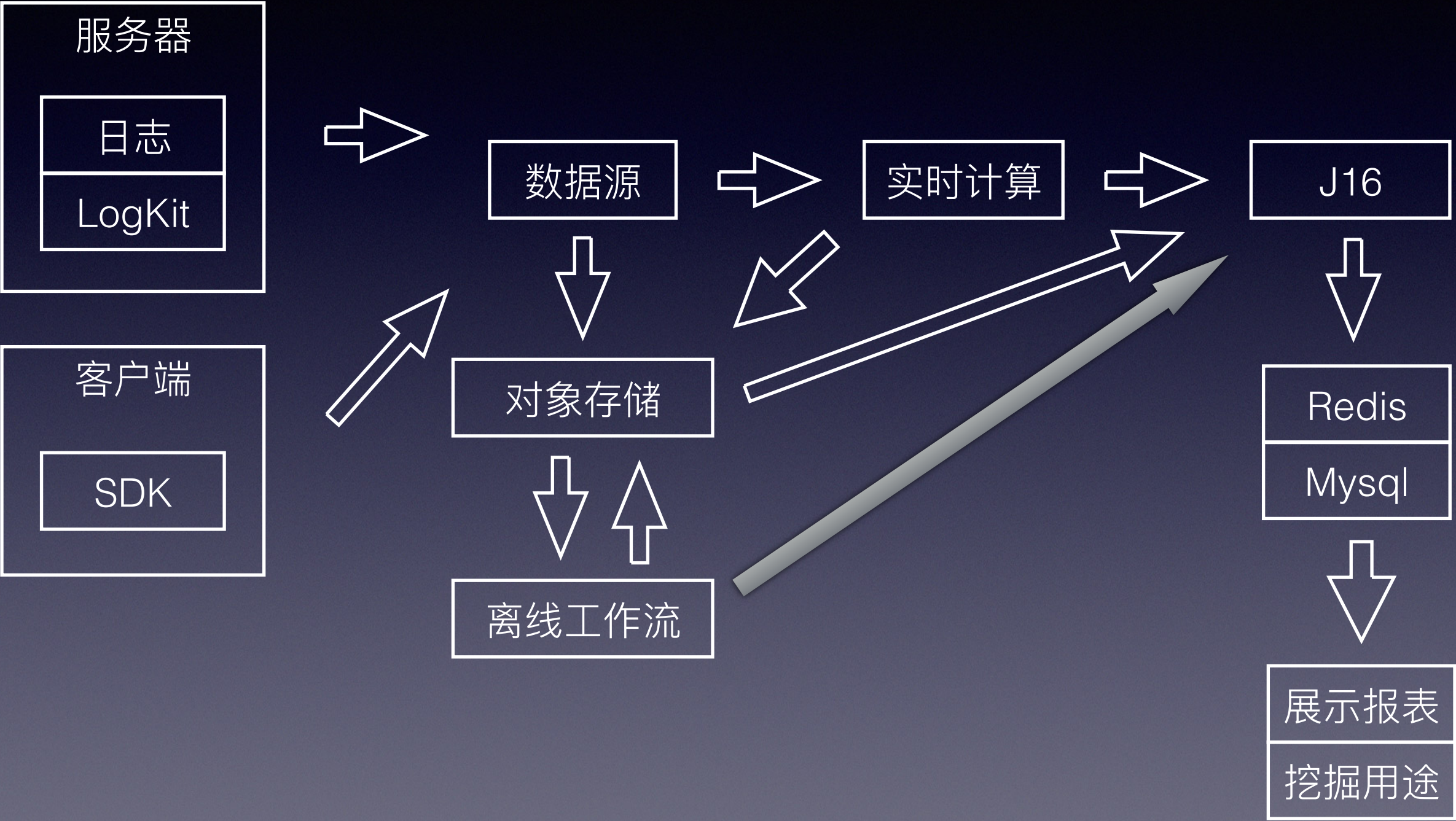
一些小经验

- 日志服务比自己搭建ELK简易得多，索引较贵
- 不同组件间集成异常繁琐
- 扩容方便，但人力并未得到释放
- 内存型快速计算准实时出数据价格偏高
- 资源竞争问题依然存在
- 产品变更太快

- 最新结构(未完全迁移完成)



- 最新流程图



一些小经验

- 迁移还是比较累的，最好一开始就直接选定一种
- 人力接近完全解放
- 组件不存在集成及兼容问题
- 资源可以达到隔离的目的，但是还未能scala

当前情况

- 日均500GB数据增量
- 日均几十亿条日志数据
- 客户端SLA、数据库审计数据等未迁移

当前情况

- 商城
- 社区
- 直播
- 运维

- 分析
- 统计
- 监控
- 告警

日志实践-数据源

- nginx
- tomcat
- fpm
- 网关
- 业务
- 数据库日志
- 上报异常
- sla日志等客户端统计数据

日志实践-nginx

```
{
  "appVersion": "6445",
  "channel": "OPP0",
  "execTime": 0.036,
  "highlight": {
    "system": [
      "@pandora_logdb_span_slive_api@pandora_logdb_span_e"
    ]
  },
  "logOrigin": "rest",
  "logType": "rest-call",
  "method": "POST",
  "mtracerId": "-",
  "offsetTimes": 0,
  "platform": "android",
  "privateIp": "117.136.4.246, 10.19.251.6",
  "query": "/v1/Gift/getForumGiftList?xx=api-l&distinctRequestId=1f688d745aa240585914db23a6e537cb",
  "responseTime": 0.036,
  "retCode": 200,
  "retSize": 921,
  "serverIp": "10.19.10.201",
  "sourceIp": "10.19.128.65",
  "spanId": "0.0",
  "system": "live_api",
  "time": "2017-07-25 17:01:49.000000+08:00",
  "timestamp": "1500973309.095",
  "token": "fc49528f7fb832c513d129ff8b73ce1e",
  "tracerId": "1f688d745aa240585914db23a6e537cb",
  "ua": "Taqu/6445/android/fc49528f7fb832c513d129ff8b73ce1e",
  "upstream": "unix:/dev/shm/default.socket",
  "uri": "v1/Gift/getForumGiftList"
```

日志实践-service

```
{
  "access": "WIFI",
  "appVersion": "841",
  "durationTimes": 0,
  "highlight": {
    "system": [
      "@pandora_logdb_span_slive_api@pandora_logdb_span_e"
    ]
  },
  "interface": "Gift",
  "level": "info",
  "logOrigin": "rest",
  "logType": "",
  "message": "[interface-request] ",
  "mtracerId": "-",
  "offsetTimes": 5,
  "pid": 5137,
  "platform": "iphone",
  "privateIp": "220.115.23.11",
  "serverIp": "10.19.5.213",
  "service": "getForumGiftList",
  "spanId": "0.1",
  "system": "live_api",
  "time": "2017-07-25 17:00:09.000000+08:00",
  "timestamp": "1500973209.78",
  "token": "A9087439-F968-4B32-9888-290209BA9497",
  "tracerId": "eb2e0ec9bd164530371c7541e2fdc004",
  "version": "1"
}
```

日志实践-查询



☒ 基本查询 ☐ 自定义查询

日志库: touch-nginx uri: tracerId: ea3b65a414e18ab226 日志类型: rs日志 时间(默认最近1小时): 2017-07-25 16:13: 至 2017-07-25 17:13: 查询

日志列表 (基于性能考虑,请尽量缩短查询时间)

	时间	uri	设备信息 说明	其他内容 说明	日志内容	更多
1	2017-07-25 17:13:22	v1/Account	OPPO android 6445 token: 21e7e07e07c803f01ce0aaec237c43d1 privateIp: 124.237.68.70, 10.19.251.6	live_api rest-call rest 0.473 200.0 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	/v1/Account/getLiveOverview?xx=api-l&distinctRequestId=ea3b65a414e18ab2263e73b0cb6b918b	查看更多
2	2017-07-25 17:13:22	v1/info	oppo android 6445 token: 21e7e07e07c803f01ce0aaec237c43d1 privateIp: 10.255.152.134	j2 soa-call rest 0.002 200.0 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	/tq-account/api/?service=info&distinctRequestId=ea3b65a414e18ab2263e73b0cb6b918b&env=0&log_origin=rest&serial_sequence=0.1-2.0-3.0&offset_time=163&token=21e7e07e07c803f01ce0aaec237c43d1&platform_name=android&platform_id=1&appcode=1&app_version=6445&channel=oppo&access	查看更多
3	2017-07-25 17:13:22	v5/Live	- - - token: - privateIp: 10.19.5.213	bbs soa-call rest 0.162 200.0 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	/v5/Soa/yarService?distinctRequestId=ea3b65a414e18ab2263e73b0cb6b918b&env=0&log_origin=rest&serial_sequence=0.1-2.0&offset_time=158&token=21e7e07e07c803f01ce0aaec237c43d1&platform_name=android&platform_id=1&appcode=1&app_version=6445&channel=oppo&access=wifitamp=1	查看更多
4	2017-07-25 17:13:22	v1/info	oppo android 6445 token: 21e7e07e07c803f01ce0aaec237c43d1 privateIp: 10.19.5.213	j2 soa-call rest 0.002 200.0 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	/tq-account/api/?service=info&distinctRequestId=ea3b65a414e18ab2263e73b0cb6b918b&env=0&log_origin=rest&serial_sequence=0.1-1.0&offset_time=4&token=21e7e07e07c803f01ce0aaec237c43d1&platform_name=android&platform_id=1&appcode=1&app_version=6445&channel=oppo&access=wifit	查看更多

☒ 基本查询 ☐ 自定义查询

日志库: touch-service tracerId: ea3b65a414e18ab226 日志来源: 全部 系统名称: 日志级别: 全部 日志内容: 时间(默认最近1小时): 2017-07-25 16:13: 至 2017-07-25 17:13: 查询

日志列表 (基于性能考虑,请尽量缩短查询时间)

	时间	uri	设备信息 说明	其他内容 说明	日志内容	更多
1	2017-07-25 17:13:22	v1/info	android wifi 6445 token: 21e7e07e07c803f01ce0aaec237c43d1	warning j2 rest 0.1-2.0-3.0 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	登陆已过期, 请重新登陆~(错误码:a1003 `expire - 请求参数:["account_level","sex_type","uid"])	查看更多
2	2017-07-25 17:13:22	v1/info	android wifi 6445 token: 21e7e07e07c803f01ce0aaec237c43d1	warning j2 rest 0.1-2.0-3.2 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	Handler execution resulted in exception	查看更多
3	2017-07-25 17:13:22	v5/Live	android wifi 6445 token: 21e7e07e07c803f01ce0aaec237c43d1	info bbs rest 0.1-2.1 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	[service-request]	查看更多
4	2017-07-25 17:13:21	v1/Account	android wifi 6445 token: 21e7e07e07c803f01ce0aaec237c43d1	info live_api rest 0.1 tracerId: ea3b65a414e18ab2263e73b0cb6b918b	[interface-request]	查看更多

日志实践-时序图

系统名称	来源 / 级别	uri	span	
live_api	rest	v1/Account/getLive	0.0	
live_api	rest / info	v1/Account/getLive	0.1	0s : interface-request
j2	rest	v1/info/g	0.1-1.0	0.002s : soa-call
bbs	rest	v5/Live/g	0.1-2.0	
j2	rest / warning	v1/info/get	0.1-2.0-3.0	
j2	rest	v1/info/getI	0.1-2.0-3.0	
j2	rest / warning	v1/info/getInf	0.1-2.0-3.2	
bbs	rest / info	v5/Live/get	0.1-2.1	



日志实践



				   	
API状态码统计		SOA调用状态码统计		L7状态码统计	
retCode	Count ▾	retCode	Count ▾	retCode	Count ▾
200	176476621	200	1118751624	200	2188072484
401	11525			404	23523
302	5638			401	10598
404	3993			302	4556
304	495			408	3454
500	12			400	2128
206	2			304	65
				502	18
				504	3
				405	3

日志实践

<

Zoom Out

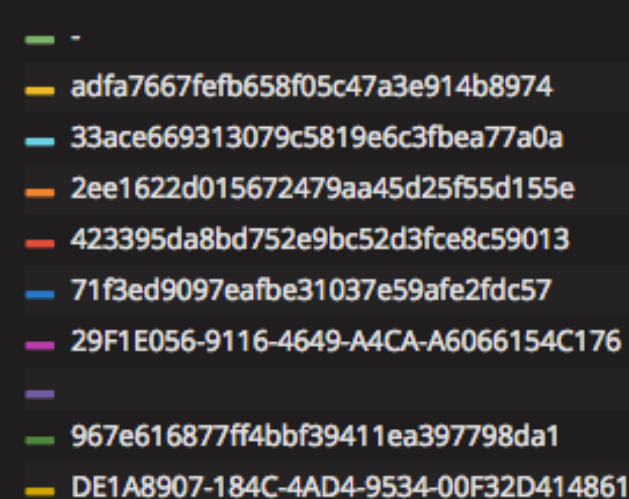
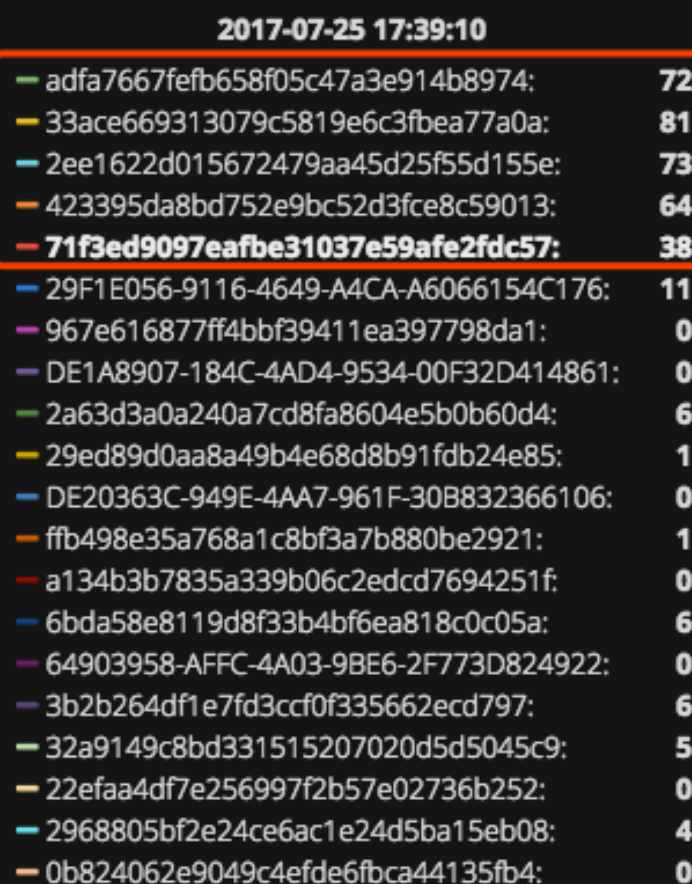
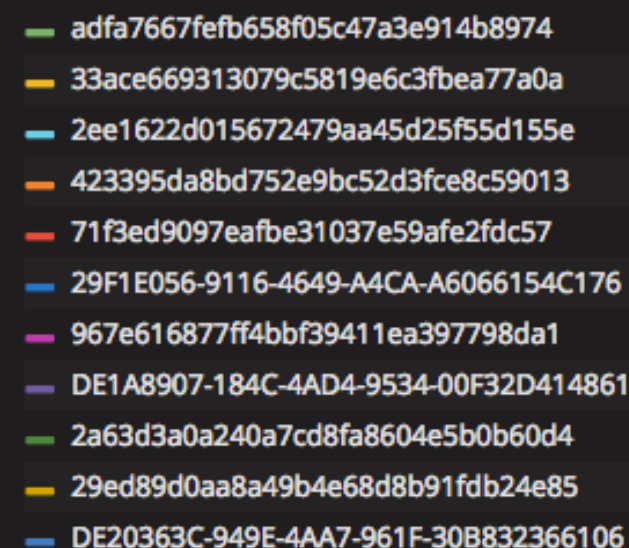
>

🕒 Yesterday

🔄

1秒到2秒内的接口量		2秒以上的接口量	
uri	Count ▾	uri	Count ▾
v5/Info/getAc	515293	v2/Ord	542
v5/Review/addF	52261	v2/Order/che	343
v1/Order/sendQ	17152	v1/Room/e	288
v1/AccountFollow/	13371	v2/Order/	239
v5/Wet/likeRevie	6330	v1/Follow	190
v2/Order/che	5498	v1/-/-	108
v5/Reply/add	3288	v4/Tag/get	104
v5/Wet/lik	2373	v5/Info/get	95
v1/PirateShip	1883	v2/IM/getI	92
v5/Account/g	1586	v1/CallbackAr	44

◀ Zoom Out ▶ ⌚ Last 5 minutes Refresh every 5s ↺



日志实践

布

☆

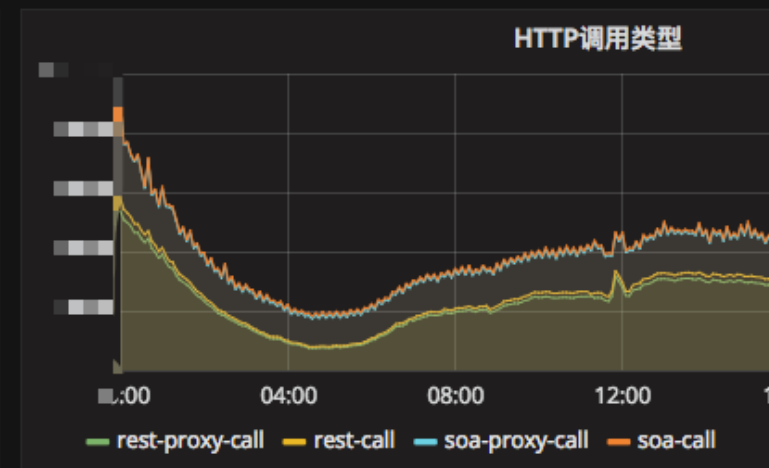
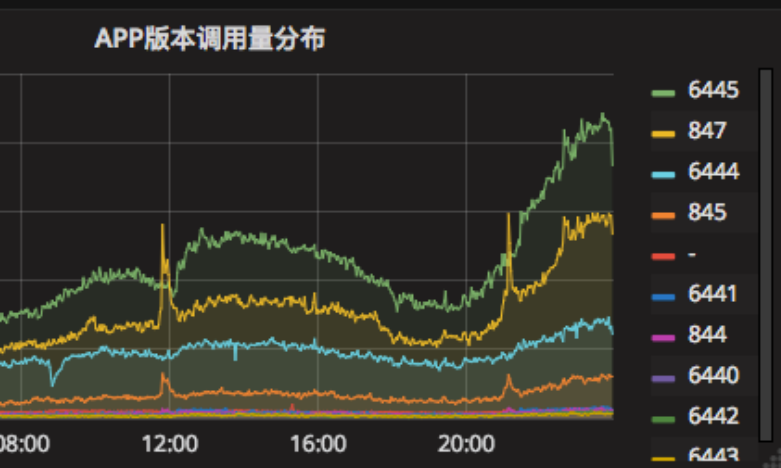
🔗

📁

⚙️

🔍

Zoom



接口版本调用分布		
version	Count	
5	1	
1	1	
2	1	
1	1	
3	1	
5	1	
2	1	
2	1	

服务版本调用分布		
system	version	Count
tqmq	2	1
live_api	1	1
bbs	5	1
live_trade	1	1
tqmq	1	1
trade	2	1
promotion	2	1
time	1	1

谢谢