

容器环境下基于APM的海量日志全链路跟踪分析

苕程

杭州谐云科技 VP

QCon

全球软件开发大会

成为软件技术专家 的必经之路

[北京站] 2018

2018年4月20-22日 北京·国际会议中心

7折

购票中, 每张立减2040元

团购享受更多优惠



识别二维码了解更多

AiCon

全球人工智能与机器学习技术大会

助力人工智能落地

2018.1.13 - 1.14 北京国际会议中心



扫描关注大会官网



极客时间

重拾极客精神·提升技术认知

下载极客时间App

获取有声IT新闻、技术产品专栏，每日更新



扫一扫下载极客时间App

谐云科技具有强大的研发实力

谐云科技核心团队来自于浙江大学SEL实验室

在开源Kubernetes上面贡献代码全国第一

#	Company	Lines of code
1	Google	83675740
	*independent	51779800
2	Red Hat	33498821
3	FathomDB	19490829
4	Zhejiang University	13538507
5	Fujitsu	2795423
6	SAP	2192870
7	Intel	1144411
8	HarmonyCloud	943710
9	IBM	853169

Showing 1 to 10 of 105 entries

SPEAKER INTRODUCE



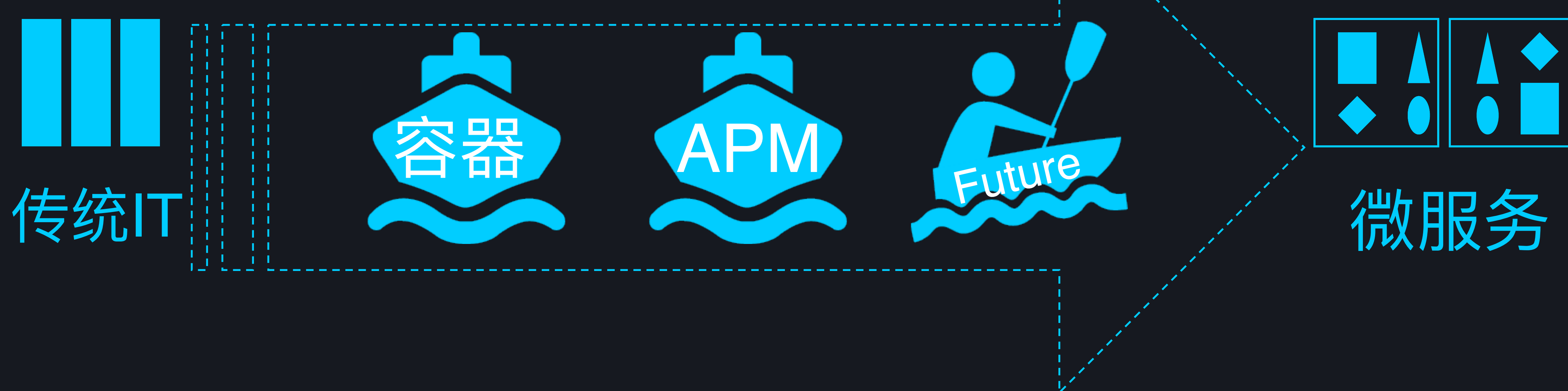
苒程

谐云科技 VP

谐云科技联合创始人，专注于分布式系统容错，大数据处理和分析，主导设计了谐云APM产品。

未来的IT架构一定是微服务化的

容器支持服务，APM定位故障

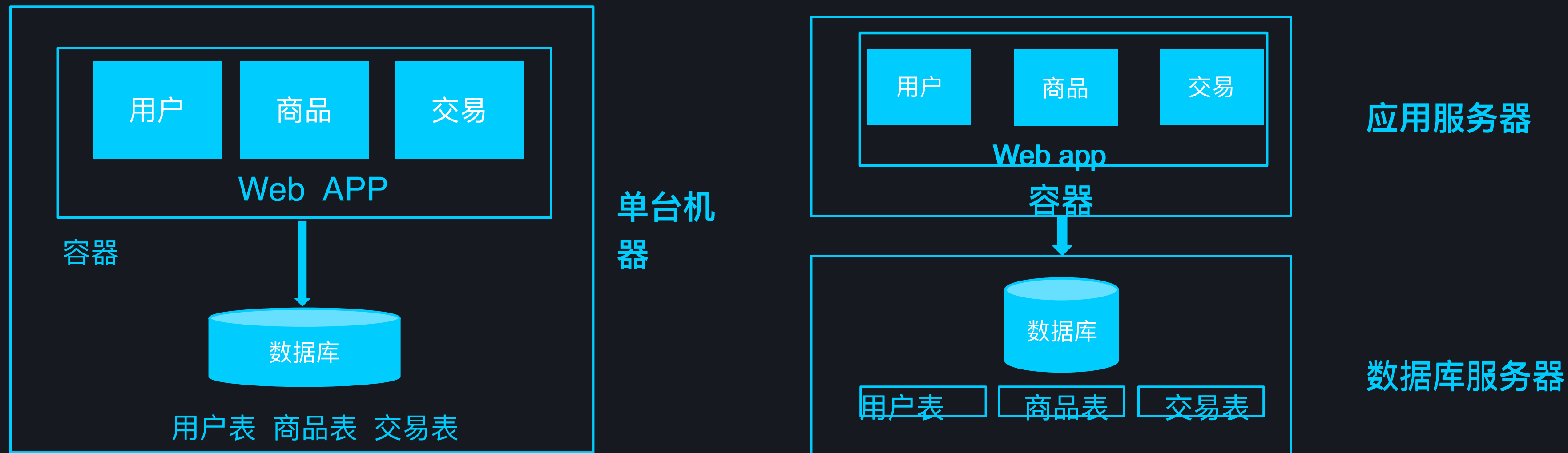


谐云拥有这两款产品，且有很好的结合点

TABLE OF CONTENTS 大纲

- 为什么需要全链路的日志排查
- 什么是全链路
- 代码中如何实现全链路
- 如何基于代码增强的方式实现日志增强
- 容器云带来的挑战

简单应用架构的日志排查过程



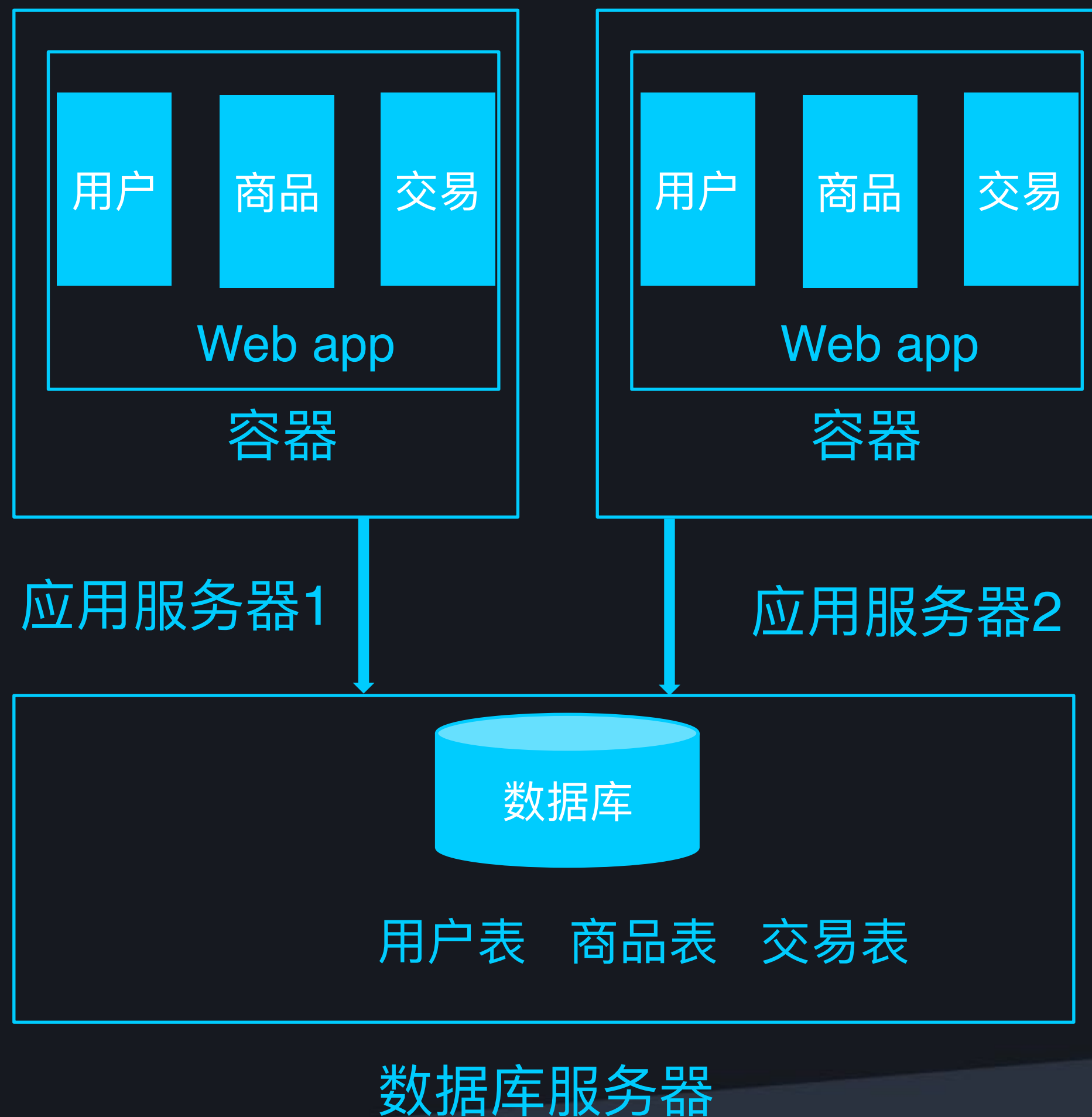
当应用架构较简单的时候

只有一个日志文件，直接查看文件即可

EASY



稍微复杂一点应用架构的排查过程



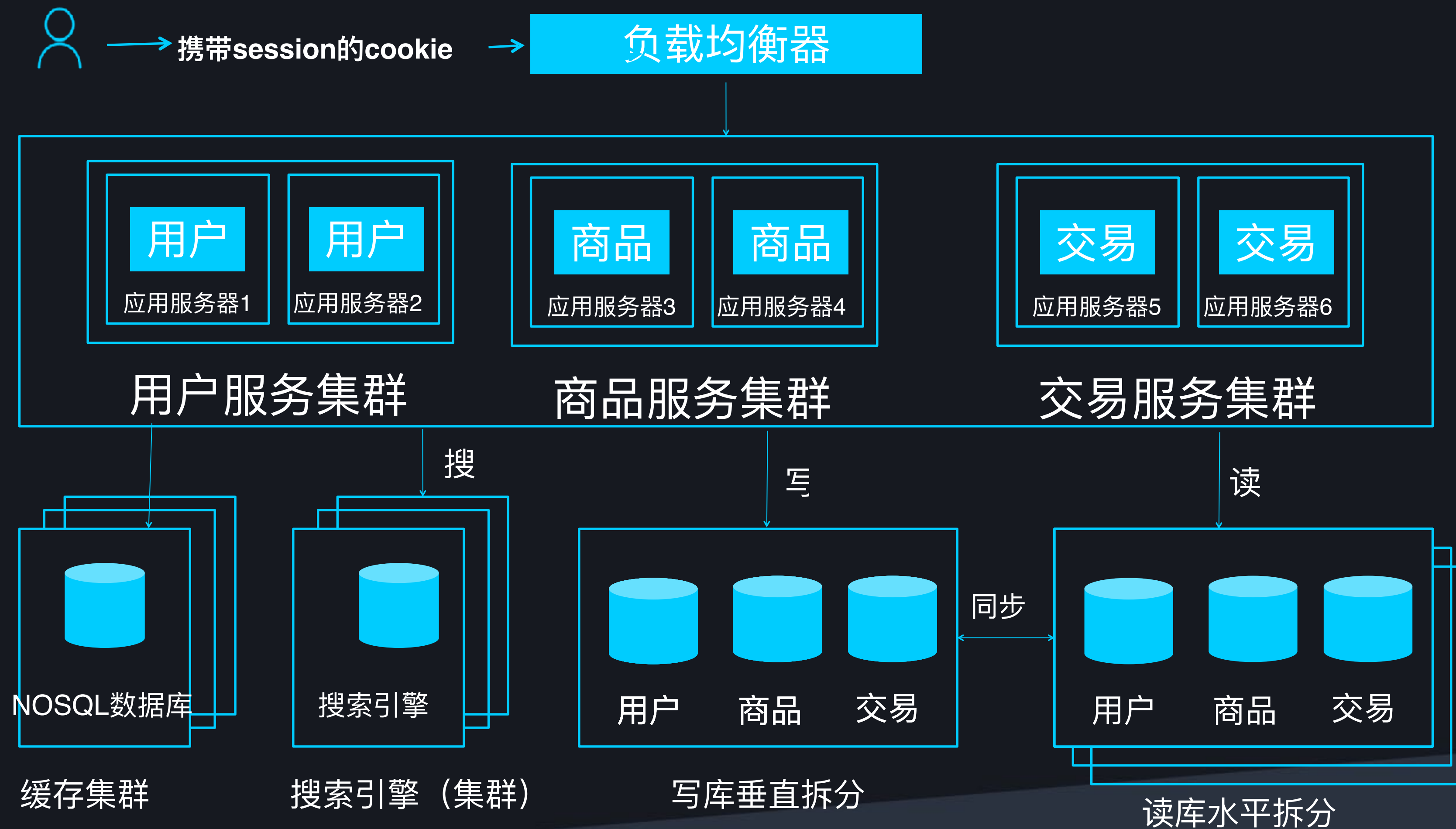
一个集群里面有限的几个应用日志

多花点时间，直接查看文件即可

有点烦，但还能接受



较复杂的应用架构的日志排查过程



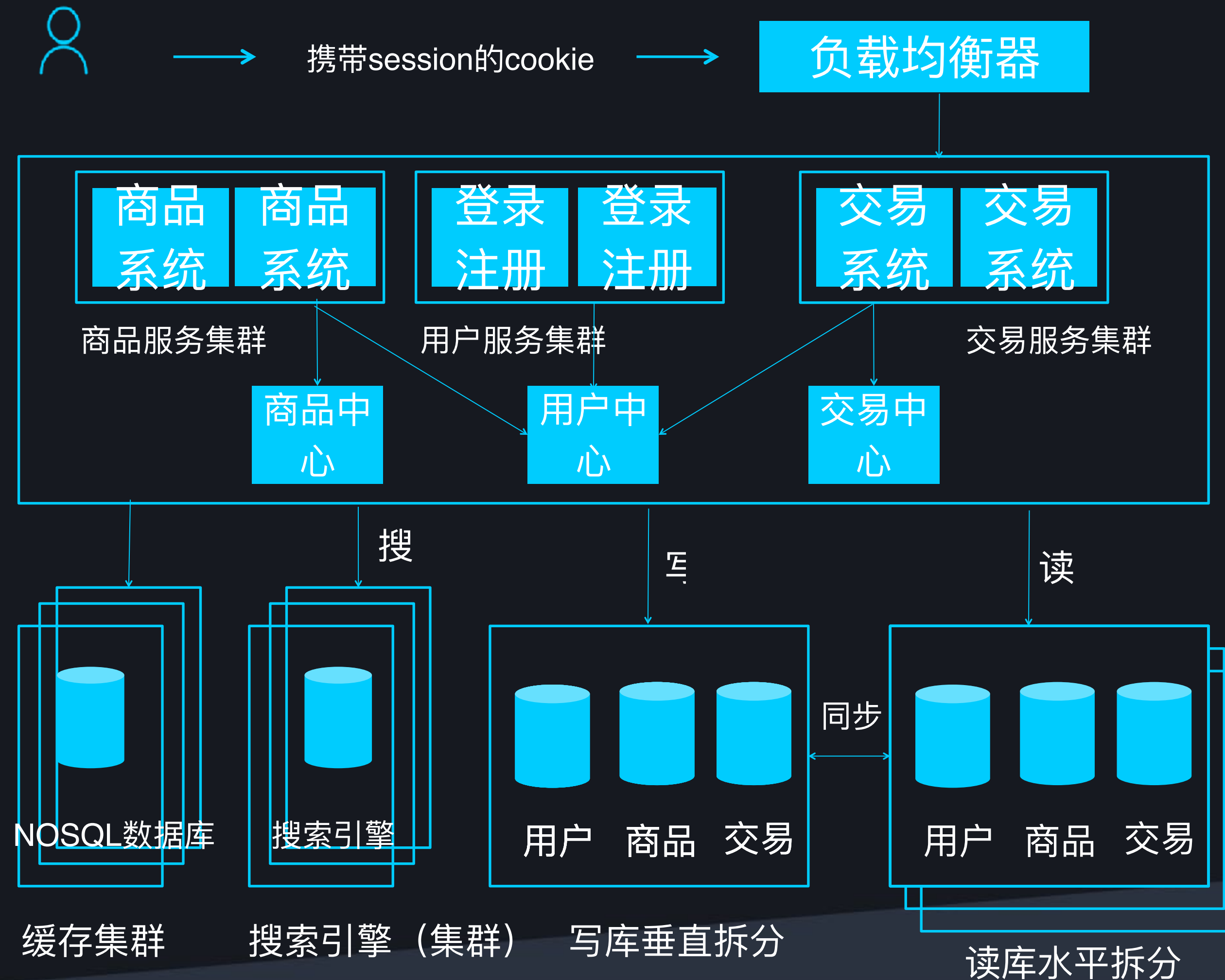
基于ELK
的统一日志

- 搜索
- 文件导出



在统一日志支撑下
还能统一接受

典型基于微服务调用的日志排查过程



容器环境的特点

弹性伸缩

2

相同的服务实例数很多几十个，甚至上百个

统一日志为啥不能完全解决问题？

1

统一日志的使用方式是：搜索

2

在微服务的整条链路上的应用都需要关心

3

结论：搜索出来的非相关性信息太多

- 比如搜索：Error，那意味整个系统所有的微服务的Error日志可能都需要看
- 运气好找到一个相关性的日志之后，再去关联其它服务节点上的日志很复杂，无从下手

全链路的日志含义

■ 可以将用户一次访问相关的日志信息经过所有服务不同实例完全找出来，并且能够按照实际响应顺序显示出来

Mar 02, 2016 10:12:28 AM org.apache.coyote.http11.Http11Protocol pause
INFO: Pausing Coyote HTTP/1.1 on http-8080
Mar 02, 2016 10:12:29 AM org.apache.catalina.core.StandardService stop
INFO: Stopping service Catalina
Mar 02, 2016 10:12:29 AM org.apache.coyote.http11.Http11Protocol destroy
INFO: Stopping Coyote HTTP/1.1 on http-8080

A服务处
理此次访
问的日志
信息

a1

B服务处
理此次访
问的日志
信息

b2

C服务处
理此次访
问的日志
信息

c4

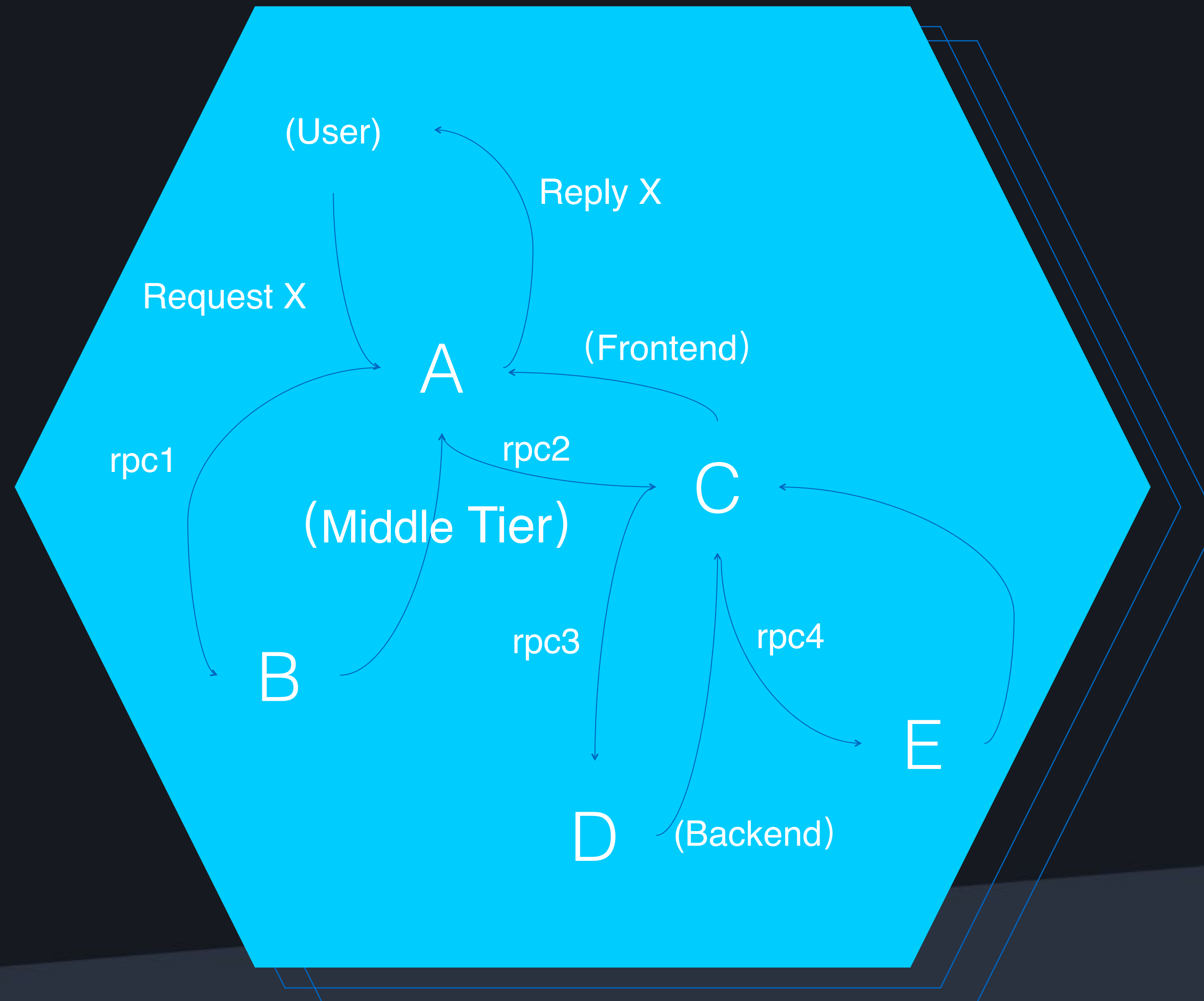
D服务处
理此次访
问的日志
信息

d9

TABLE OF CONTENTS 大纲



- 为什么需要全链路的日志排查
- 全链路的日志排查
- 代码中如何实现全链路
- 如何基于代码增强的方式实现日志增强
- 容器云带来的挑战

什么是全链路? Google Dapper的全链路



为什么要全链路?

快速交付与快速排查



VIP用户

张总, 想买50万的理财产品失败, 登陆不上。。。

老板

李导, 我马上联系相关人, 周末请你喝酒 (抱拳状)



老板

50万的理财产品登陆不上! 火速解决!!!

前台

我没问题, 问问服务报错了吗?

中台

是你没调过来吧, 我连日志都没打!

后台

胡说什么? 我这有反应, 你传错了!



老板

50万的理财产品登陆不上! 火速解决!!!

运维

进程都在, 也没看见ZB报错呀!

DBA

什么叫, 连数据都没连上!

基础

傻X, 磁盘满了, 都不知道在叫啥。。。

老板

张三说李四, 李四说王五, 说了半天都两个小时了 (叹气状), 就不能解决地快点吗?

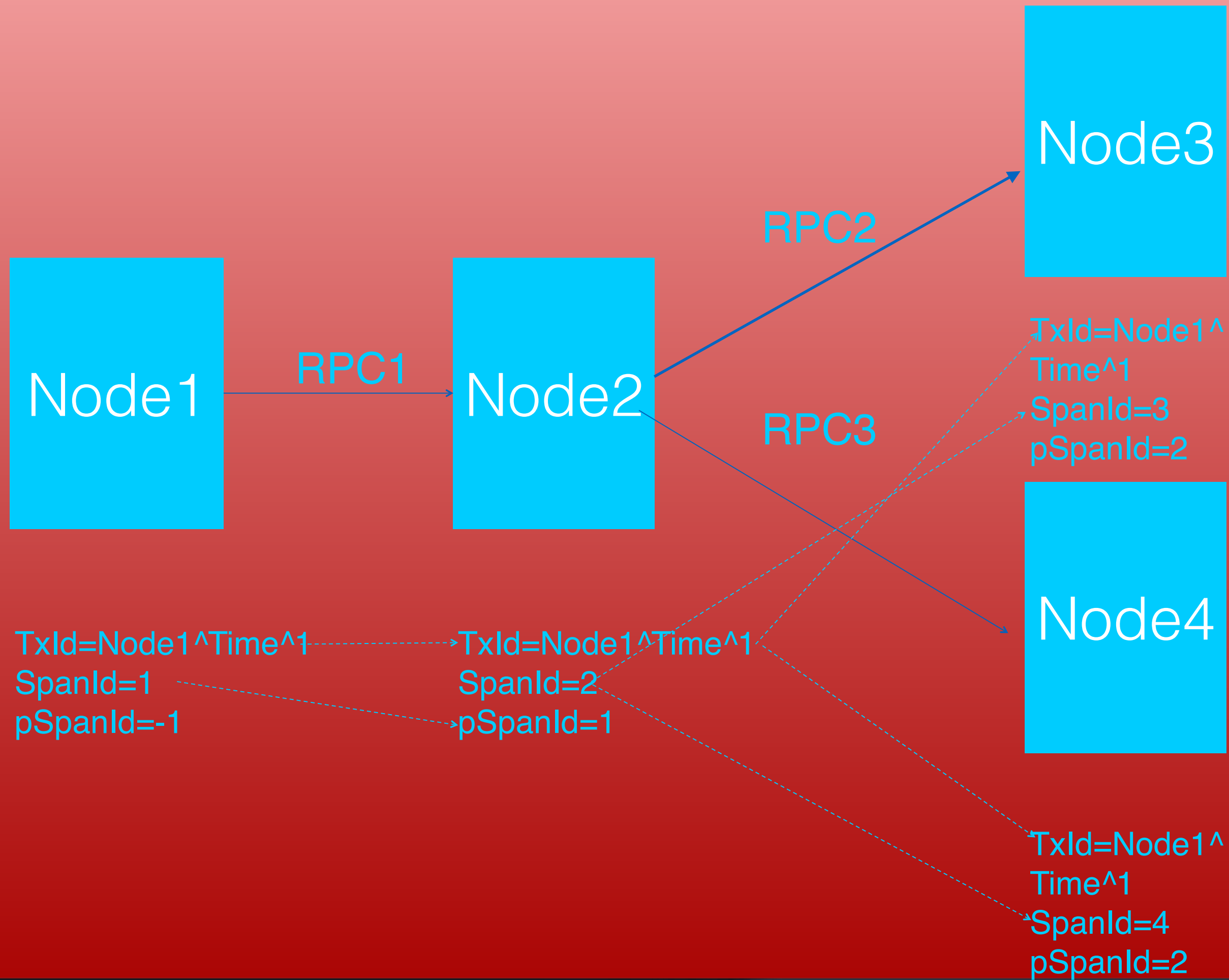
员工

谁不想快啊? 但也要快的起来呀! 老板, 我们都没偷懒啊 (无奈状)。。。。。。

TABLE OF CONTENTS 大纲

- 为什么需要全链路的日志排查
- 全链路的日志排查
- 代码中如何实现全链路
- 如何基于代码增强的方式实现日志增强
- 容器云带来的挑战

全链路的实现



TraceID : 识别用户一次请求, 所有全链路上的节点共用一个TraceID

SpanID : 正在处理用户请求的节点

ParentSpanID : 正在处理用户请求节点的上一个节点

Opentracing VS 自动化插装

Opentracing的优劣势

①

适用于任何代码框架

②

对代码开发团队需要有较强的要求，测试团队需要配合一起测试

VS

自动化插装的优劣势

①

无需代码开发

②

支持的框架有限制

TABLE OF CONTENTS 大纲

- 为什么需要全链路的日志排查
- 全链路的日志排查
- 代码中如何实现全链路
- 如何基于代码增强的方式实现日志增强
- 容器云带来的挑战

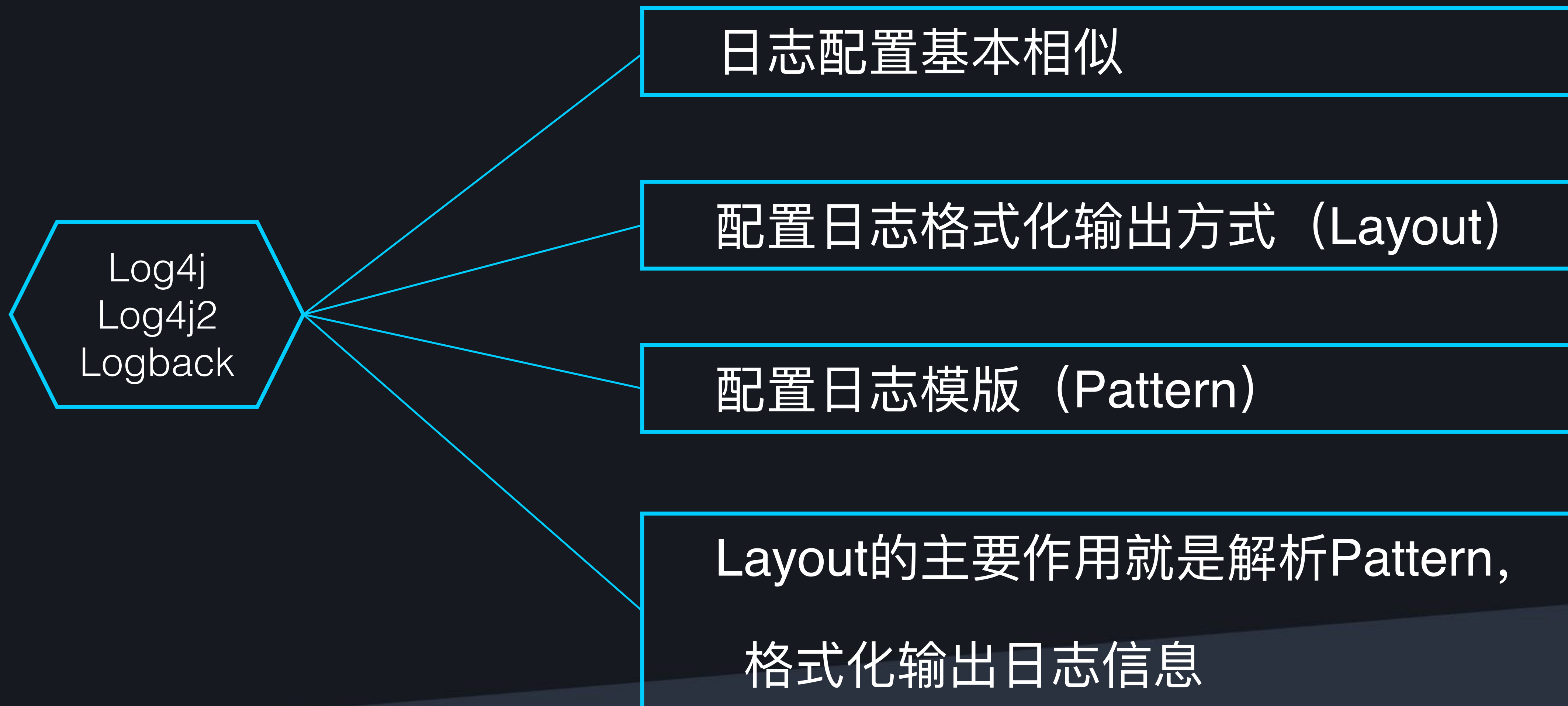
全链路日志示例

[2017-12-06 18:53:36] [ERROR] (AccountAction:206) - AccountAction.prepare... , transactionId = test@jeeshopclient-3177904429-875h5:8080^1512552181462^74

[2017-12-06 18:53:36] [ERROR] (AccountAction:678) - toLogin... , transactionId = test@jeeshopclient-3177904429-875h5:8080^1512552181462^74

[2017-12-06 18:53:36] [ERROR] (FrontInterceptor:28) - CommonInterceptor.intercept... , transactionId = test@jeeshopclient-3177904429-875h5:8080^1512552181462^74

Java常用的日志框架配置



日志常见配置 (Log4J为例)

配置日志输出方式

`log4j.appender.stdout=org.apache.log4j.ConsoleAppender`

配置日志路径

`log4j.appender.stdout.Target=System.out`

配置日志格式化方式

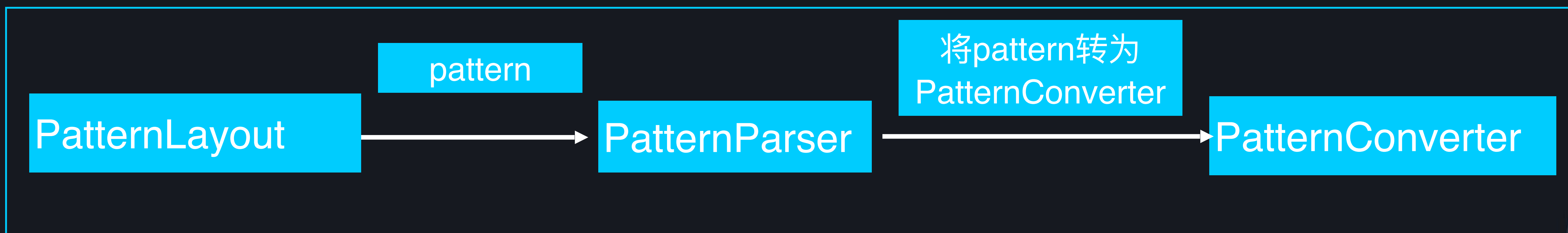
`log4j.appender.stdout.layout=org.apache.log4j.PatternLayout`

配置具体格式, 通过
layout解析pattern

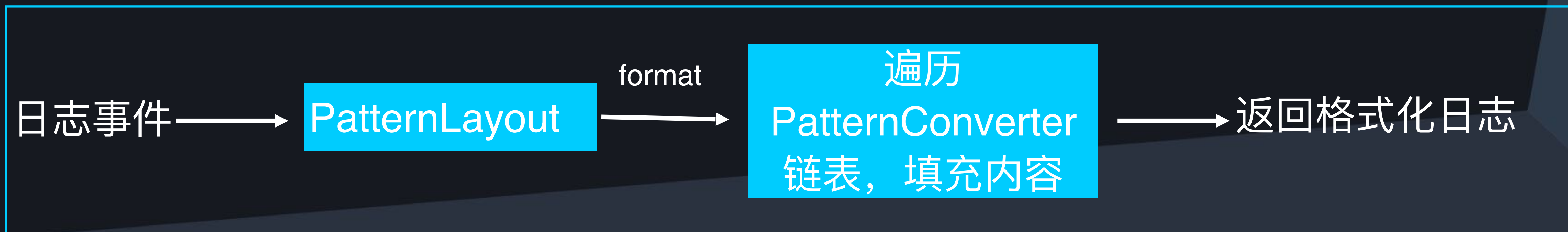
`log4j.appender.stdout.layout.ConversionPattern
=
%-d{yyyy-MM-dd HH:mm:ss} %m%n`

日志解析过程

- 根据配置文件，实例化PatternLayout，解析pattern
PatternConverter对象实质就是一个链表，把pattern中的参数拆解



- 每当日志事件到来，调用PatternLayout的format获取格式化日志



日志增强输出全链路的traceid

自定义开发

PatternLayout、
PatternParser、
PatternConverter

在配置文件中需要将layout参数指定为自定义的PatternLayout

log4j.appender.CONSOLE.layout=org.xx.TracePatternLayout

log4j.appender.CONSOLE.layout.ConversionPattern=%d [%T]
%-5p %c{1}:%L - %m%n

修改pattern，添加%T。“%T”是能够被PatternConverter识别的，可以将其转化为TraceID

基于APM日志增强方式

无需额外配置就可以开启日志增强

拦截某方法，追加Pattern格式，加入%T（只执行一次）

通过每次打印日志，都会调用format方法，按照pattern打印具体信息。通过拦截format，注入%T的值。

TABLE OF CONTENTS 大纲

- 为什么需要全链路的日志排查
- 全链路的日志排查
- 代码中如何实现全链路
- 如何基于代码增强的方式实现日志增强
- 容器云带来的挑战

Kubernetes 中IP不固定带来的问题



IP不再能唯一的标识一个应用了

10: 00 AM : 实际产生日志时的链路信息

10.10.101.100(A应用)->10.10.101.105(B应用)-> 10.10.101.119(C应用)

13: 00 PM: 分析日志时, 日志所记录的IP对应的服务已经不再是实际产生日志的服务

10.10.101.100(C应用)->10.10.101.105(D应用)-> 10.10.101.119(E应用)

如何解决?

优点

实现简单

方案一

记录日志时，同时多记录一个字段
，对应的是应用名称

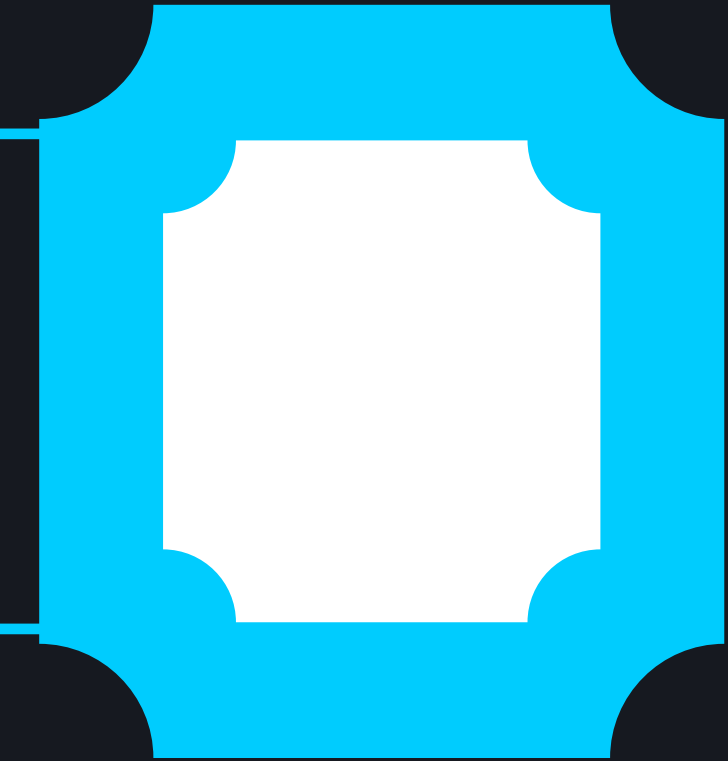
How?

缺点

浪费存储

通过Pod Name来唯一的标识应用

采集日志数据时，按照pod name进行应用标识



APM的agent也按照pod name进行应用标识



HarmonyCloud - Google Chrome

HarmonyCloud × 网上商店系统(jeesh ×

10.100.100.147:8887/#/overview/overviewAll

Apps httpd myself Welcome to Gern Bugzilla GL mijiang / k8s | Git cfets K8S cluster tool python sdn docker New Tab 代码对比/归并 qbview.url.cn/ge 从 Firefox 导入 Other bookmarks

用户 admin 租户 sjdemo

谐云科技HARMONY CLOUD

总览

集群

租户管理

应用中心

交付中心

CICD

日志管理

告警中心

系统设置

10/10集群组件

1集群

3租户

7主机

14应用

cluster1

查看详情

可用配额

0%CPU0/8 Core

0%内存0/10.9 GB

已用资源

10%CPU0.8/8 Core

62%内存6.8/ 10.9 GB

21%磁盘61.4/ 288 GB

节点 (7)

10.10.102.148

4CoreCPU

7.5G内存

196GB磁盘

10.10.102.149

2CoreCPU

3.6G内存

96GB磁盘

10.10.102.150

2CoreCPU

3.6G内存

96GB磁盘

租户 (3)

demo

CPU0.6 / 4.9 Core12%

内存768 MB/5.6 GB13%

sjsdemo

CPU3 / 22 Core14%

内存3.75 GB/34 GB11%

huitest

CPU0 / 1 Core0%

内存0 GB/1 GB0%

应用 (14)

错误运行

50%

1小时

错误4

1小时

信息12

3小时

错误0

3小时

信息13

7小时

错误0

7小时

信息7

1天

错误0

1天

信息0

THANK YOU

如有需求，欢迎至 [讲师交流会议室] 与我们

ArchSummit
全球架构师峰会 2017



苙程 