



十年架构 成长之路

SACC 第十届中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2018

2018年10月17-10月21日 北京海淀永泰福朋喜来登酒店



瓜子IM智能客服数据架构设计

封宇



SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018



目录

CONTENTS



1

项目背景

2

系统架构

3

存储架构

4

Q&A



SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018



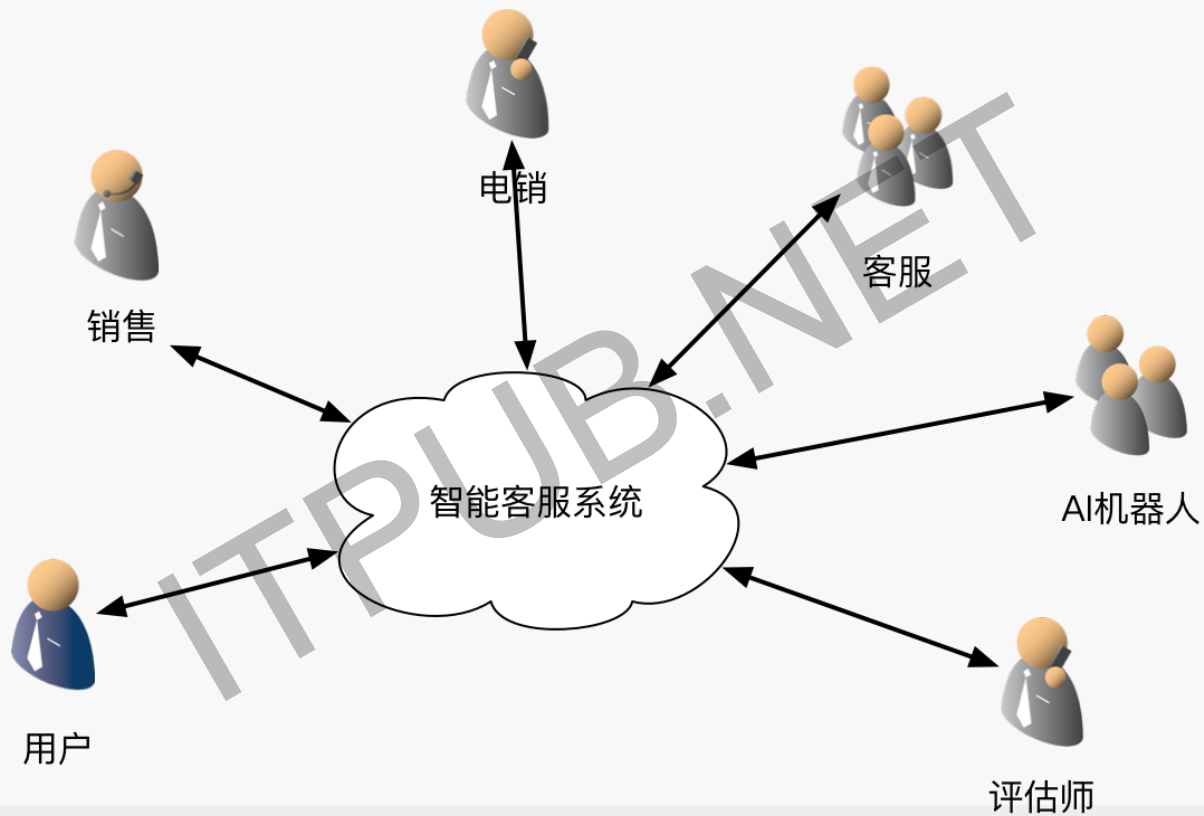
项目背景



十年架构 成长之路

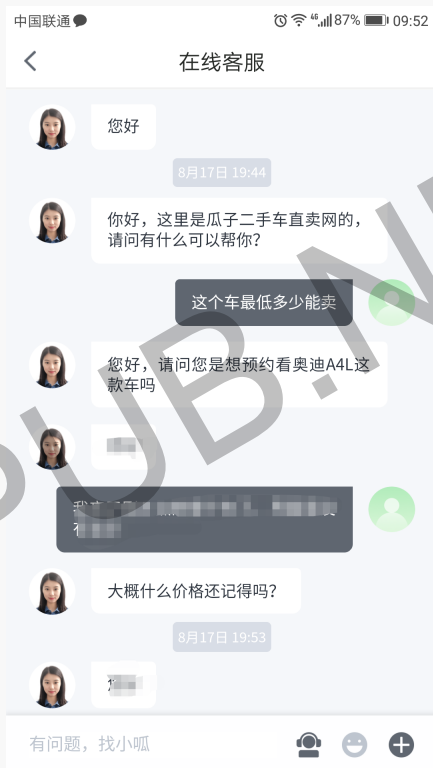


项目背景—业务关系



十年架构 成长之路

项目背景—系统界面



十年架构 成长之路



目录

CONTENTS



1

项目背景

2

系统架构

3

存储架构

4

Q&A

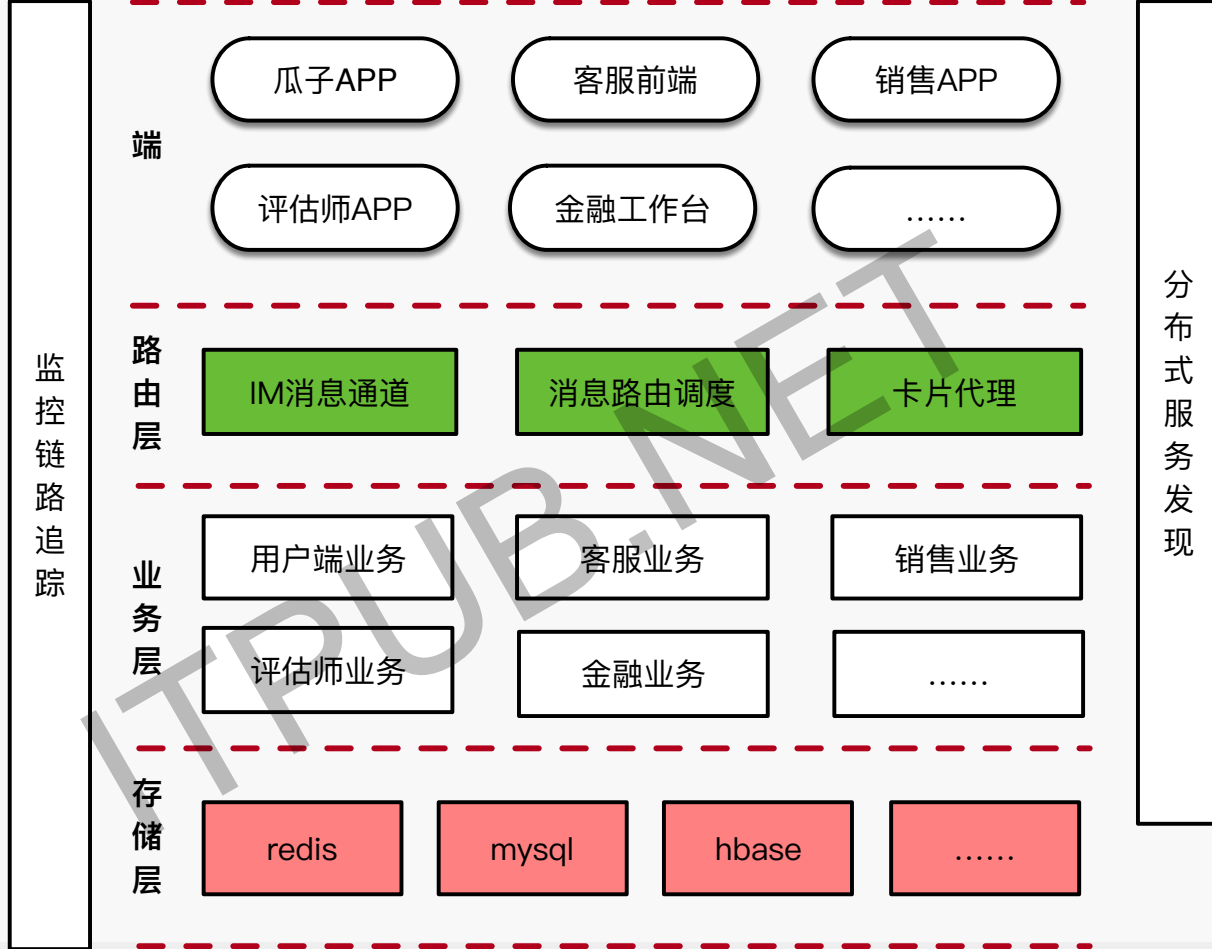


SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018



系统架构



十年架构 成长之路



目录

CONTENTS

1

项目背景

2

系统架构

3

存储架构

4

Q&A



SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018



系统架构

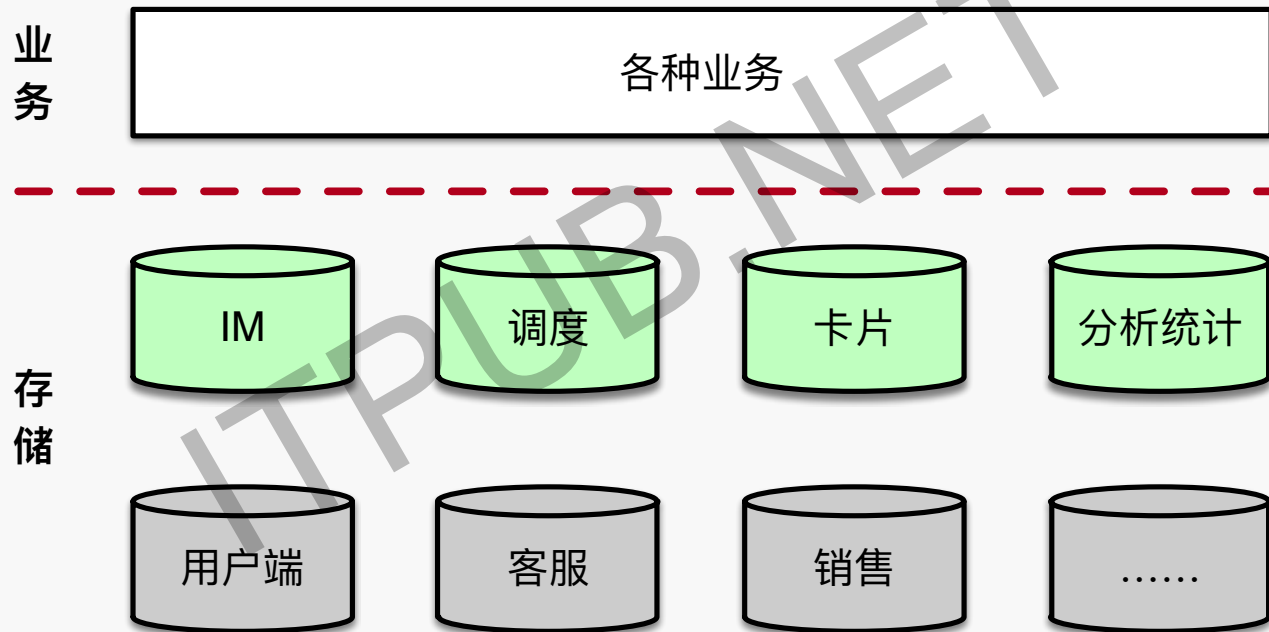


十年架构 成长之路



存储架构—垂直切分

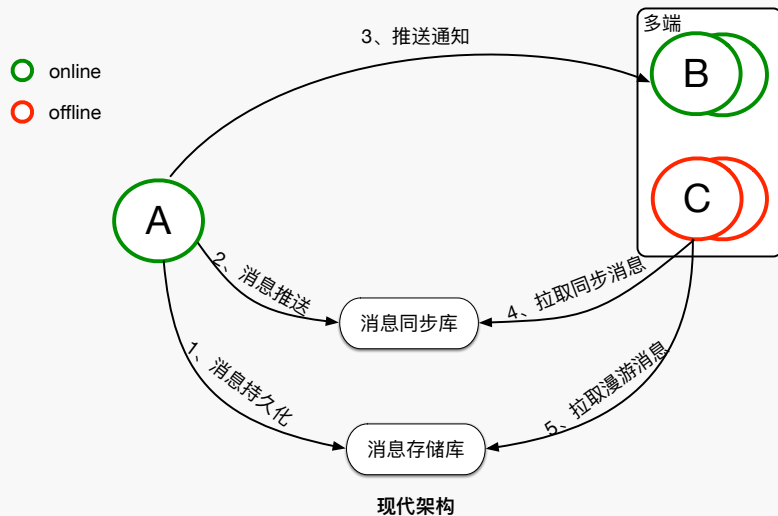
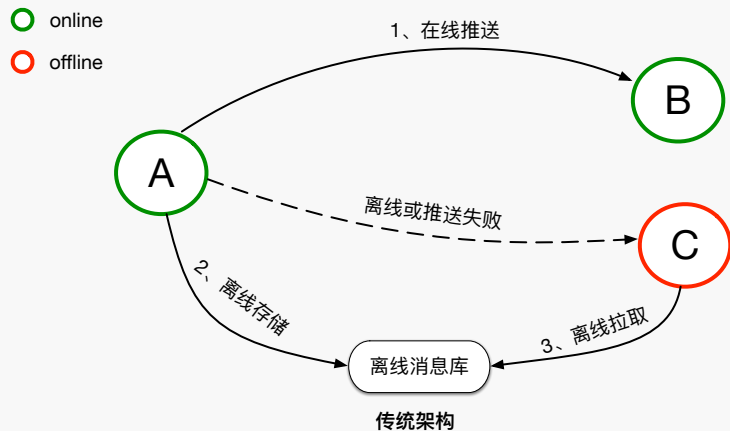
按照业务把存储拆成若干个数据库



十年架构 成长之路



存储架构—消息同步



十年架构 成长之路



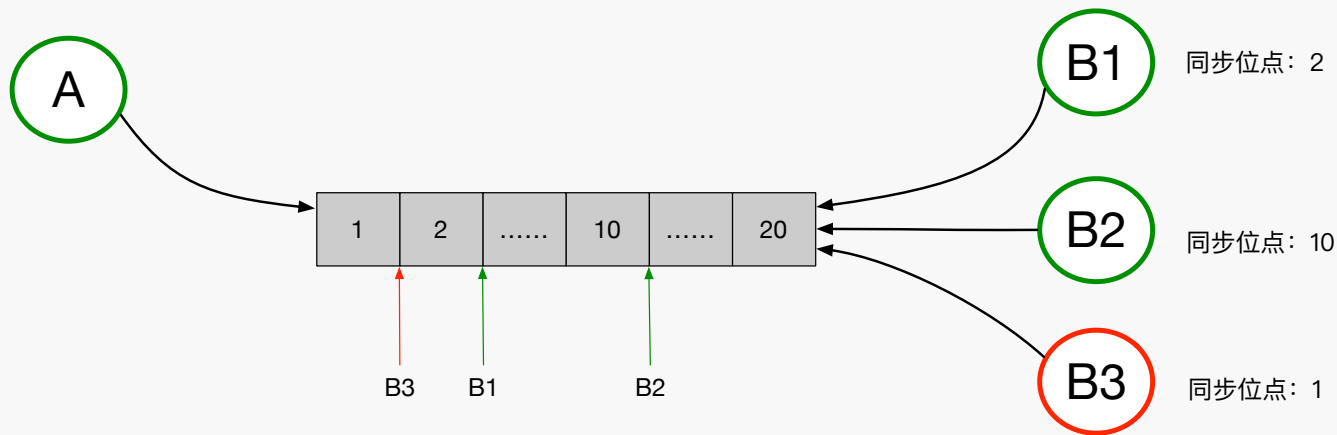
存储架构—消息同步

怎么存储同步消息

扩散写、扩散读

三个版本号的存储

采用Redis cluster



十年架构 成长之路



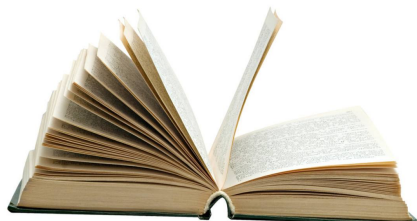
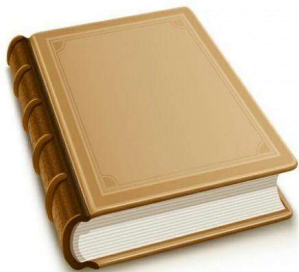
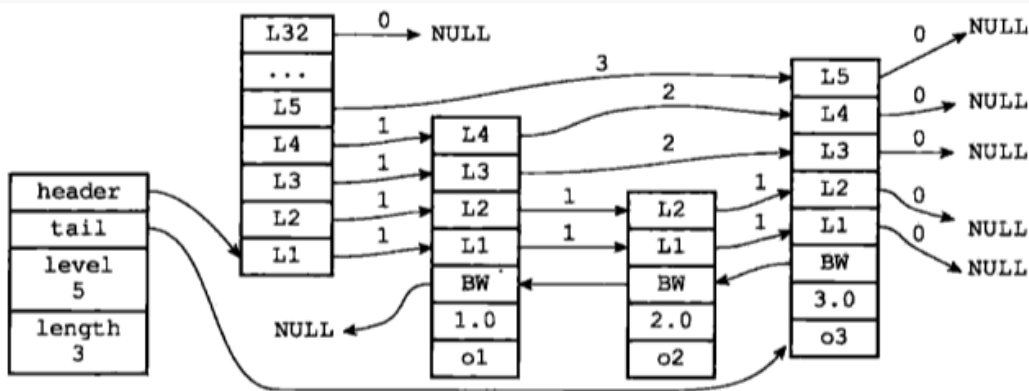
存储架构—消息同步

SortedSet结构

跳表结构

ZADD $O(M \cdot \log(N))$, N 是有序集的基数, M 为成功添加的新成员的数量。

ZRANGEBYSCORE
 $O(\log(N) + M)$, N 为有序集的基数, M 为被结果集的基数



十年架构 成长之路



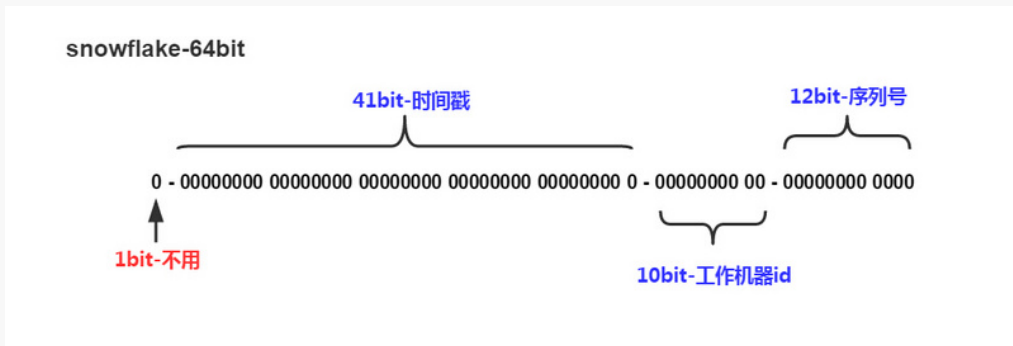
存储架构—消息同步

遇到的问题

按照会话维度存储。大群把单聊消息弄丢

web端，本地缓存不好用，服务端需要记录未读数

Sortedset丢失精度问题



1bit (符号位)	11bit (指数位)	52bit (位数位)
------------	-------------	-------------



十年架构 成长之路



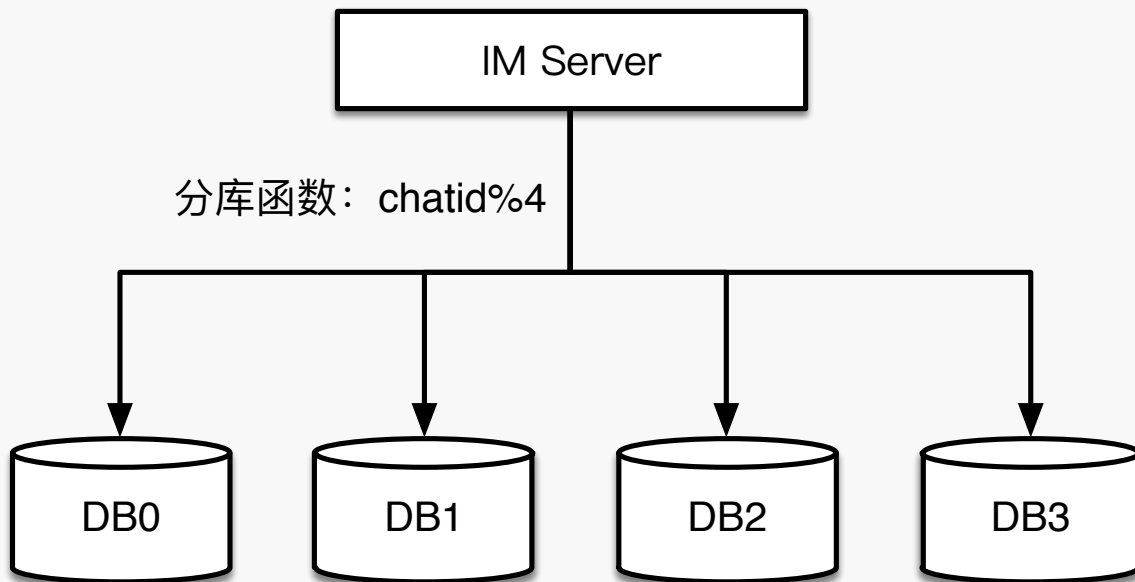
存储架构—消息存储

水平分库

Chatid代表一个对话

Chatid生成算法，snowflake

防止分布不均



十年架构 成长之路

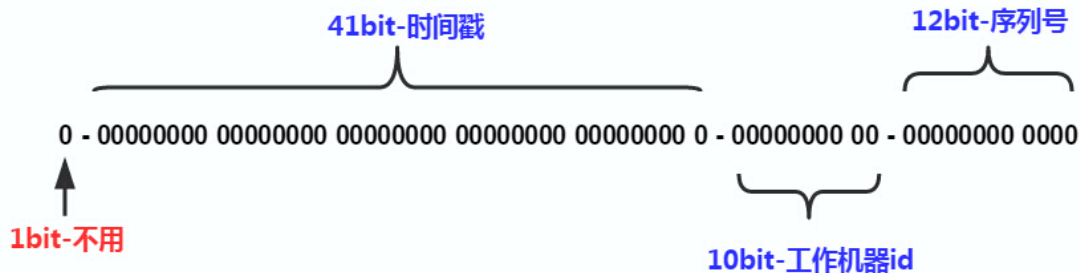


存储架构—消息存储

防止ID分布不均

利用时间戳的后8位 (bit) 作为分库随机数。
支持128个库的水平分库

snowflake-64bit



```
else {  
    //this.sequence = 0;  
    this.sequence = 127 & timestamp; //利用时间戳解决id取模平均分配问题  
}  
if (timestamp < this.lastTimestamp) {  
    throw new RuntimeException(String  
        .format("Clock moved backwards. Refusing to generate id for %d milliseconds",  
            (this.lastTimestamp - timestamp)));  
}  
  
this.lastTimestamp = timestamp;  
return timestamp - twepoch << timestampLeftShift | workerId << workerIdShift  
    | this.sequence;
```



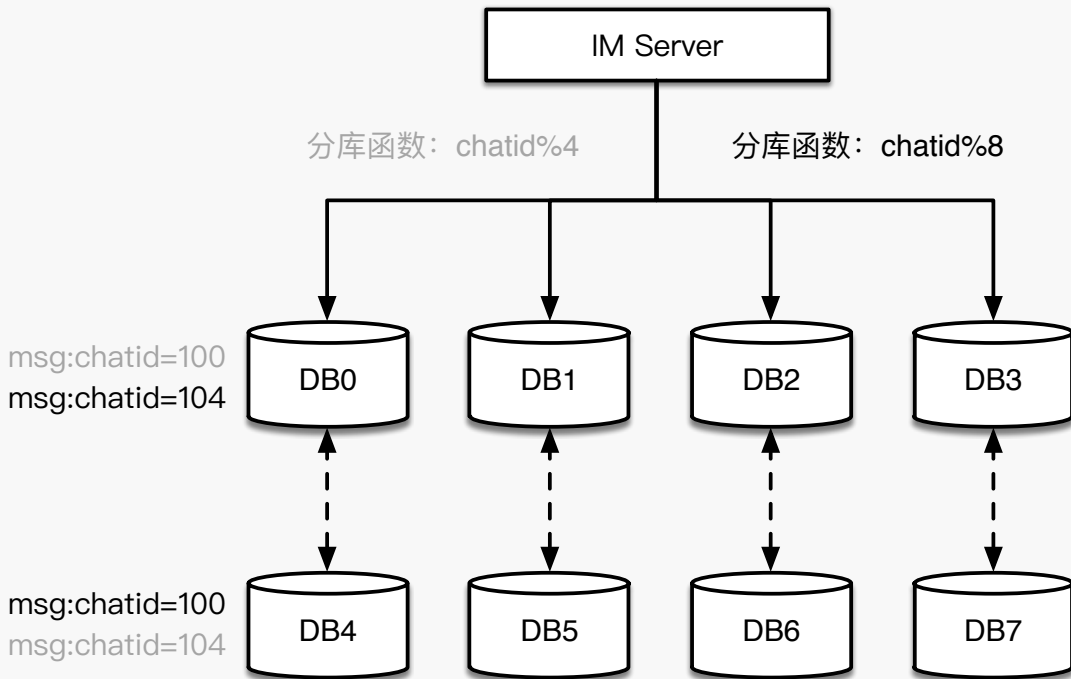
十年架构 成长之路



存储架构—消息存储

怎么扩容

- 1、DB4作为DB0从库，同步数据
- 2、断开同步关系
- 3、修改分库规则



十年架构 成长之路



存储架构—消息存储

Hbase , Rowkey怎么设计？

OpenTSDB中的key模式是：[元数据类型][时间戳]

系统采用这个模式就好，
[chatid][Long.max-msgid]

Chatid可以分散region，msgid
时间有序用于范围检索

12 systems in ranking, September 2018

Rank			DBMS	Database Model	Score		
Sep 2018	Aug 2018	Sep 2017			Sep 2018	Aug 2018	Sep 2017
1.	1.	1.	Cassandra +	Wide column store	119.55	-0.02	-6.65
2.	2.	2.	HBase +	Wide column store	58.47	-0.33	-5.87
3.	3.	3.	Microsoft Azure Cosmos DB +	Multi-model i	19.18	-0.35	+7.95
4.	4.		Datastax Enterprise +	Multi-model i	7.76	+0.46	
5.	5.	↓ 4.	Accumulo	Wide column store	3.73	-0.05	-0.15
6.	6.	↓ 5.	Microsoft Azure Table Storage	Wide column store	3.67	+0.65	+0.41
7.	7.	↓ 6.	Google Cloud Bigtable +	Wide column store	1.31	+0.46	+0.49
8.	8.	↑ 9.	ScyllaDB	Wide column store	0.97	+0.23	+0.62
9.	9.	↓ 8.	MapR-DB	Multi-model i	0.62	+0.01	+0.14
10.	10.	↓ 7.	Sqrrl	Multi-model i	0.34	-0.00	-0.17
11.	11.		YugaByte DB	Multi-model i	0.19	-0.02	
12.	12.	↓ 10.	Alibaba Cloud Table Store	Wide column store	0.06	+0.01	+0.06



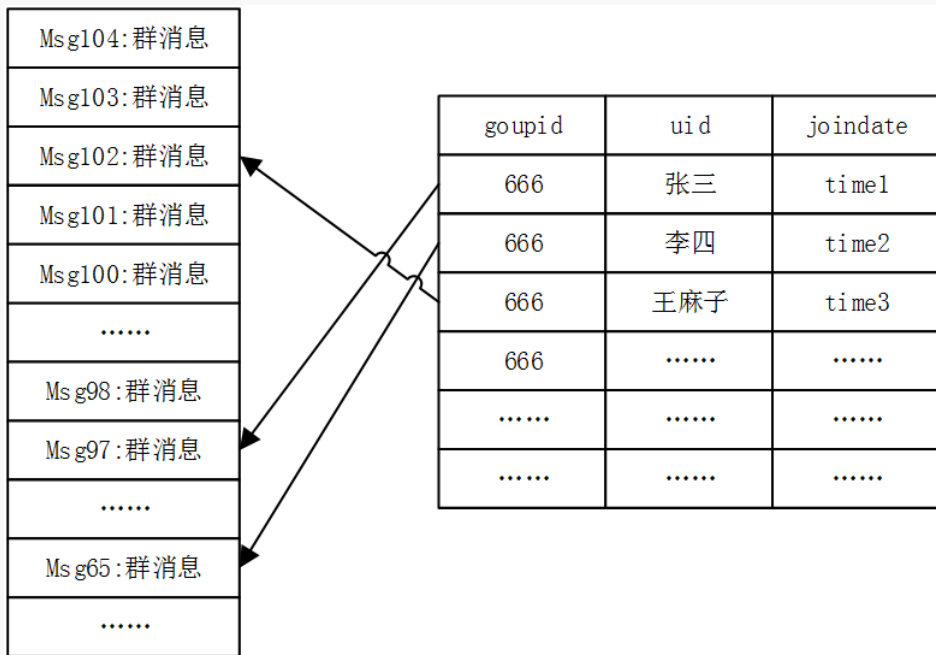
十年架构 成长之路



存储架构一群模型

一些要点

- 1、减少存储量
- 2、查看最近X条
- 3、会话排序
- 4、已读未读
- 5、消息撤回



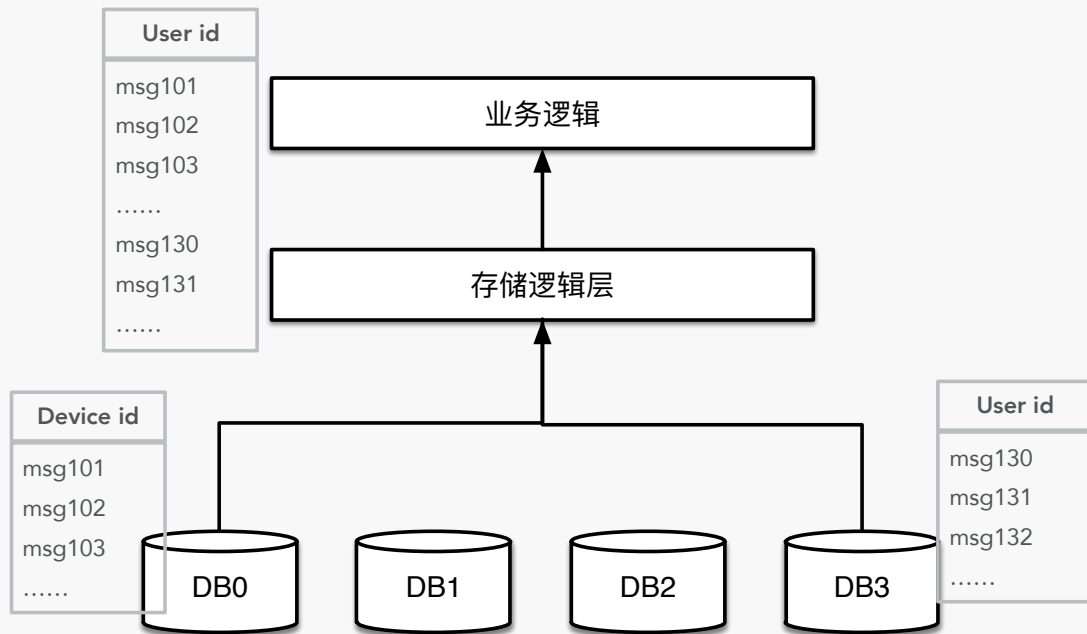
十年架构 成长之路



存储架构—存储逻辑层

匿名实名

- 1、两个uid，匿名用deviceid
- 2、消息进行Merge



十年架构 成长之路



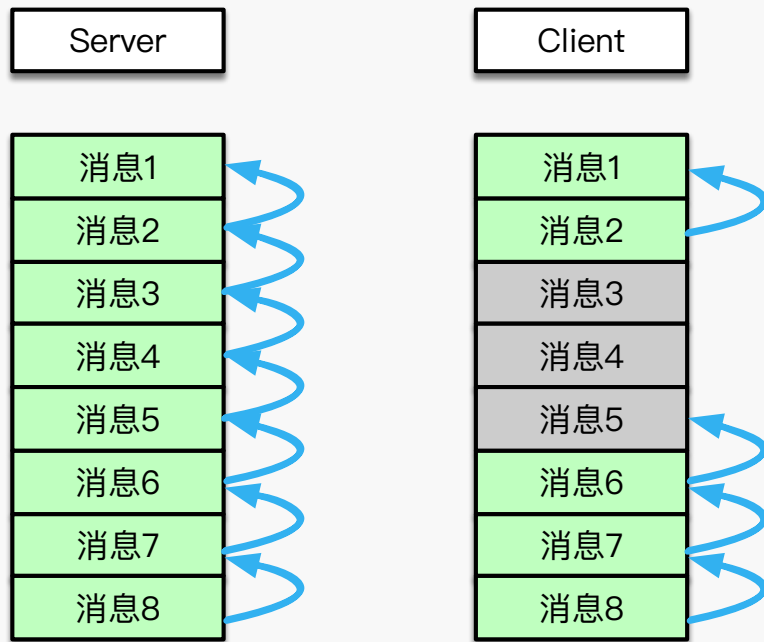
存储架构—存储逻辑层

消息同步其实很复杂

- 1、消息太多感觉卡
- 2、换着手机用
- 3、同步会话和未读数
- 4、每个chat是一个TimeLine

msgid:snowflake

msgindex:1,2,3



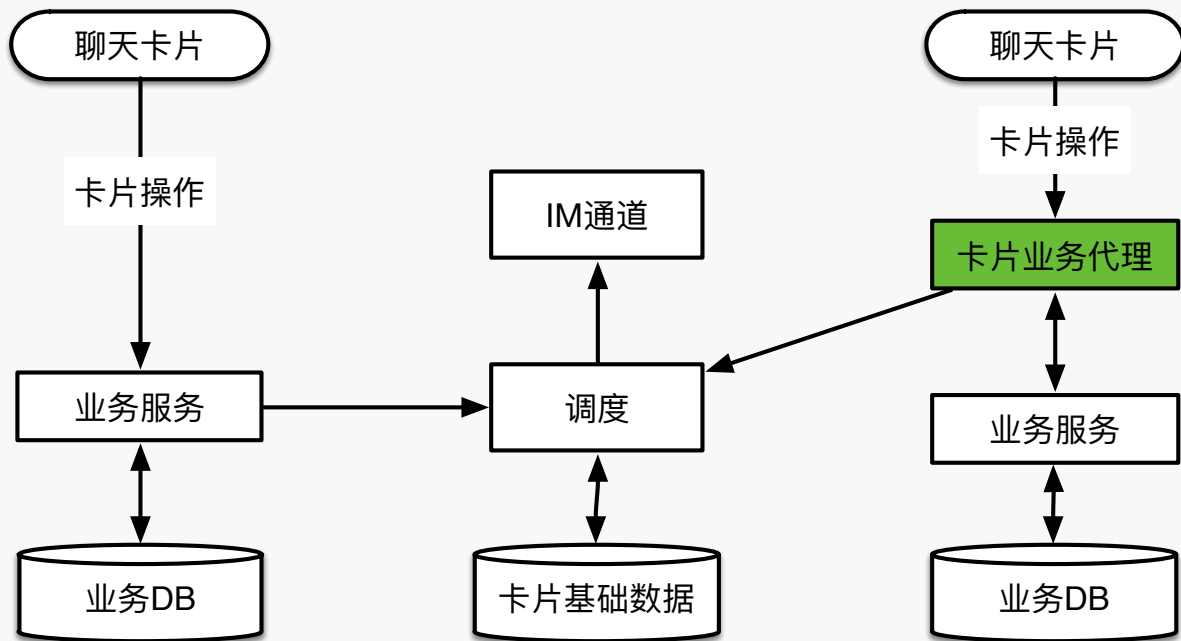
十年架构 成长之路



存储架构—业务卡片

卡片业务要点

- 1、兼容原有业务
- 2、不增加调用链长度



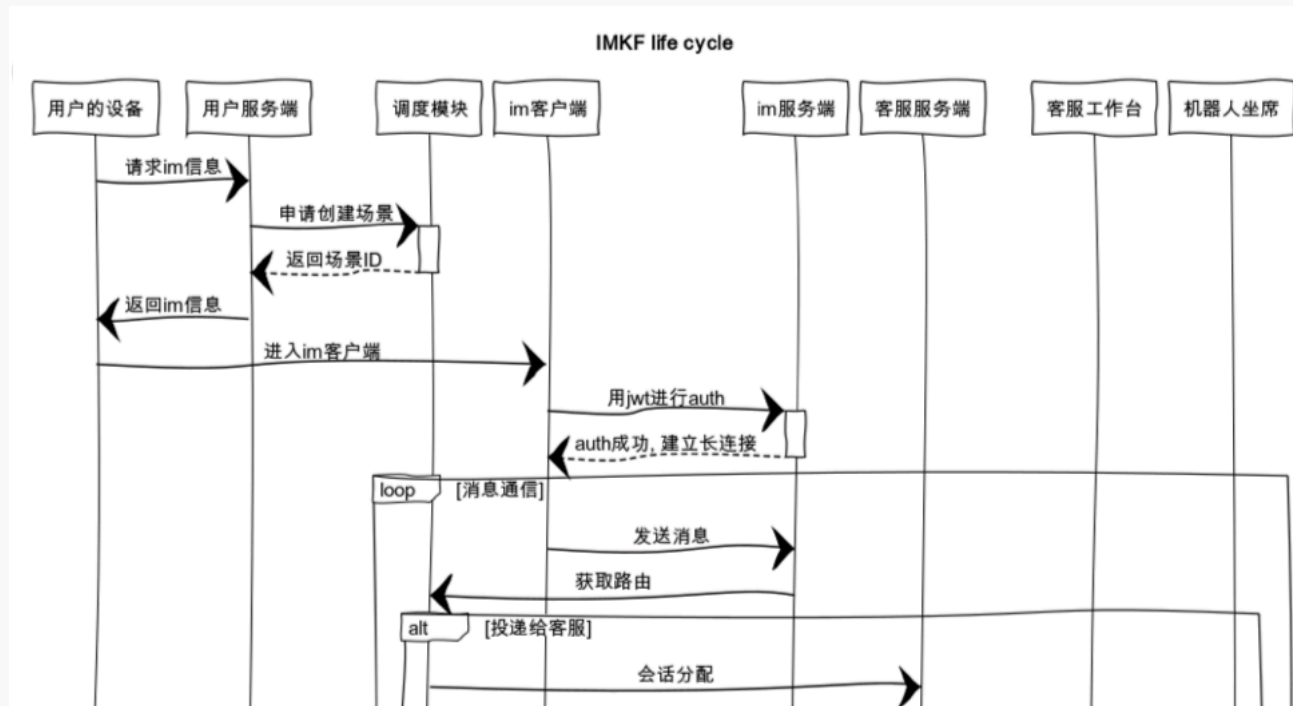
十年架构 成长之路



存储架构—业务卡片

调度是怎么回事

- 1、入口写场景
- 2、结合用户其他信息调度



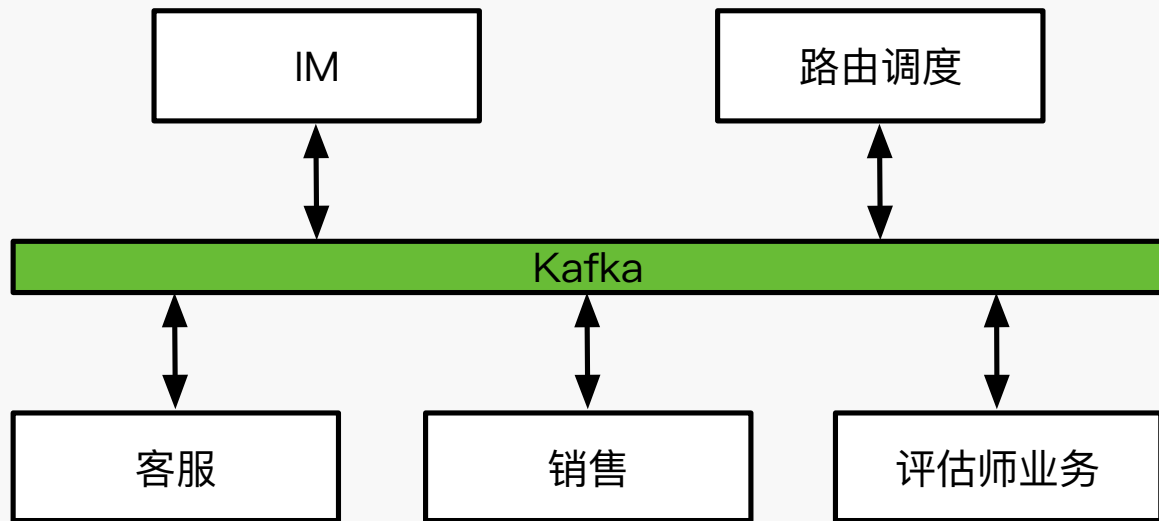
十年架构 成长之路



存储架构—分析统计

遇到了什么问题

- 1、太重了不行，成本太高
- 2、系统异步交互核心是kafka，能不能利用kafka来统计



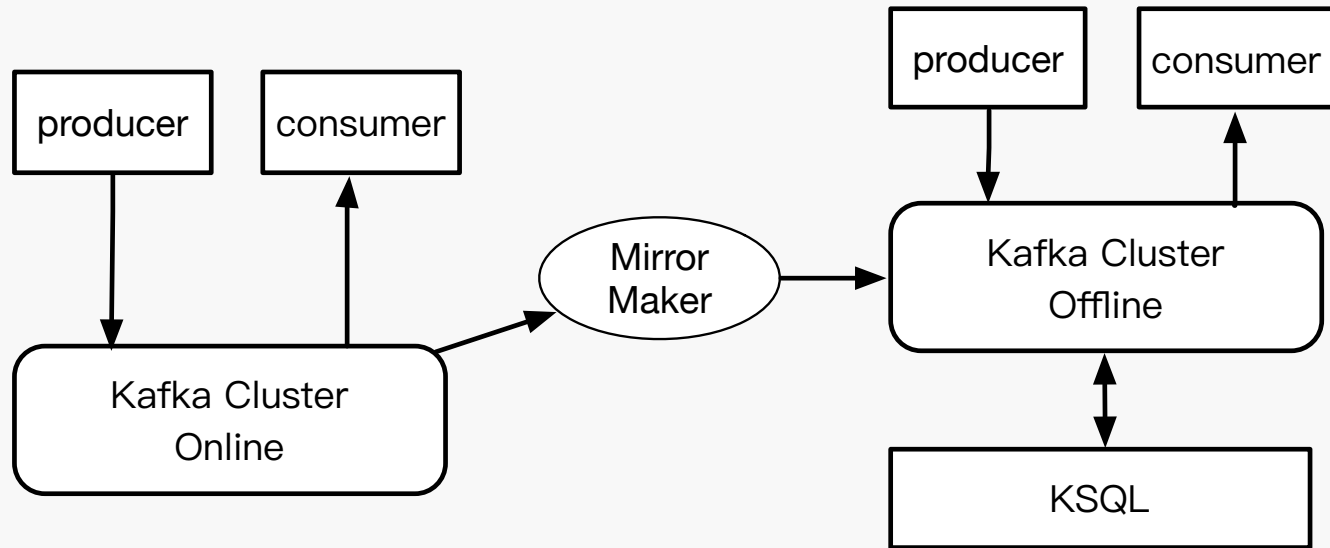
十年架构 成长之路



存储架构—分析统计

方案要点

- 1、online保证消息链路可靠
- 2、offline用于流式数据库
- 3、ksql用于分析计算



十年架构 成长之路



目录

CONTENTS

1

项目背景

2

系统架构

3

存储架构

4

Q&A



SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018





Q
&
A

10

SACC

第十届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2018

IT168.com

ChinaUnix

ITPUB



THANKS