



DTCC

2016中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2016

数据定义未来

SequeMedia
盛拓传媒

IT168.com

ChinaUnix

ITPUB

搜索实时更新 架构实践

周航@1号店



DTCC

2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

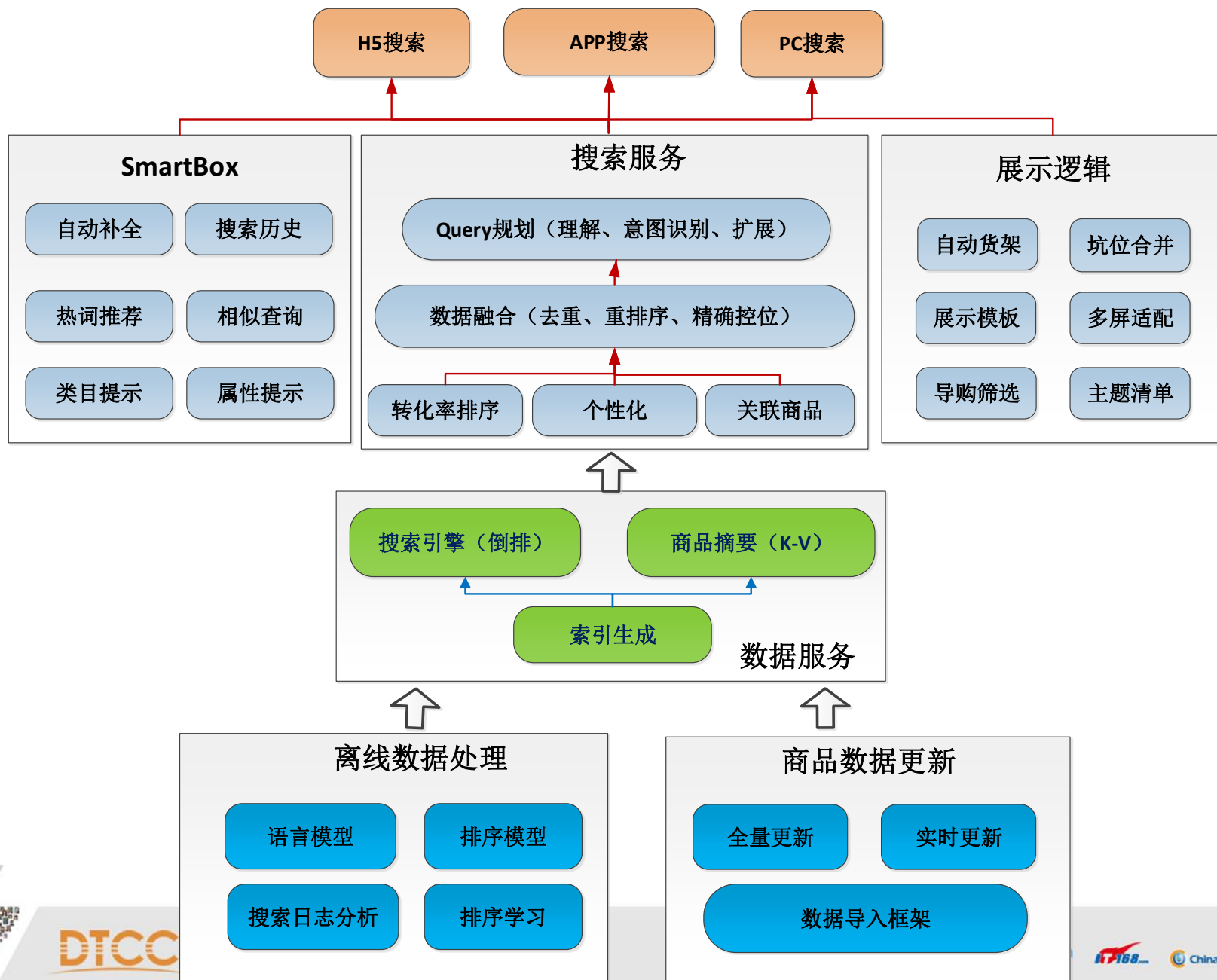
SequeMedia
数据传媒

IT 168

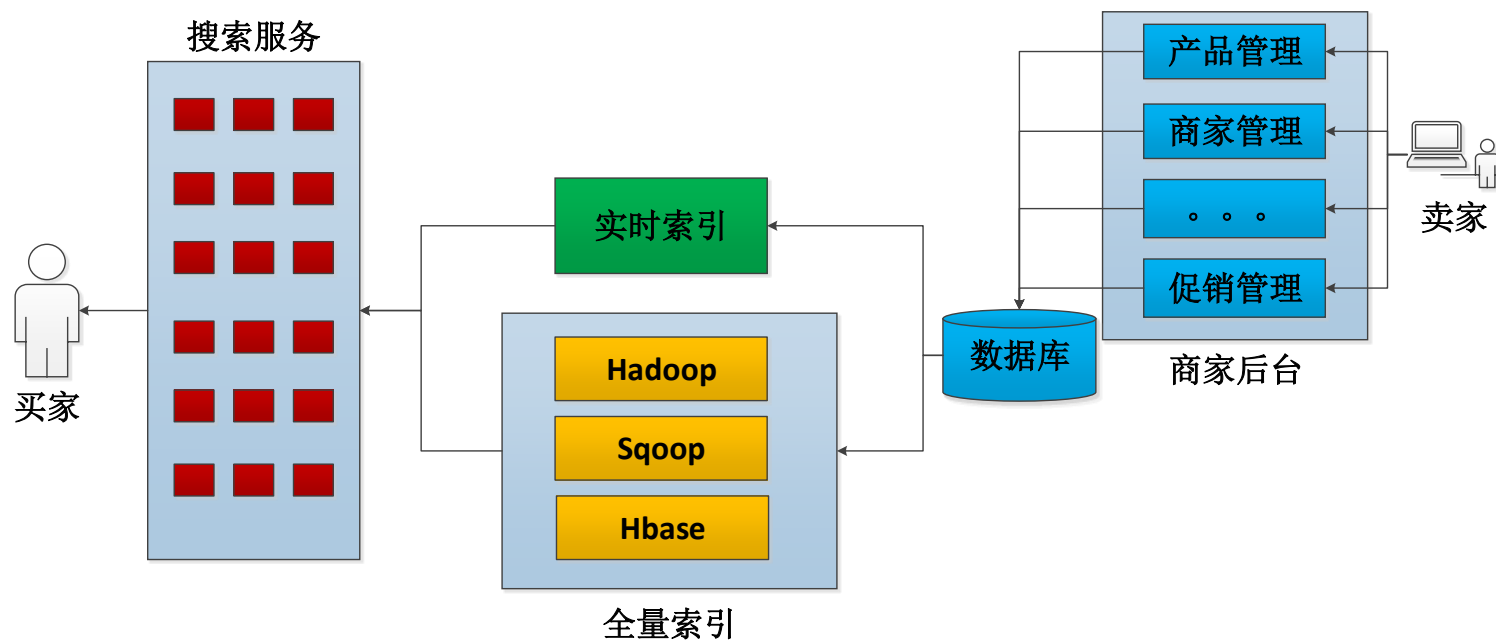
ChinaUnix

IT PUB

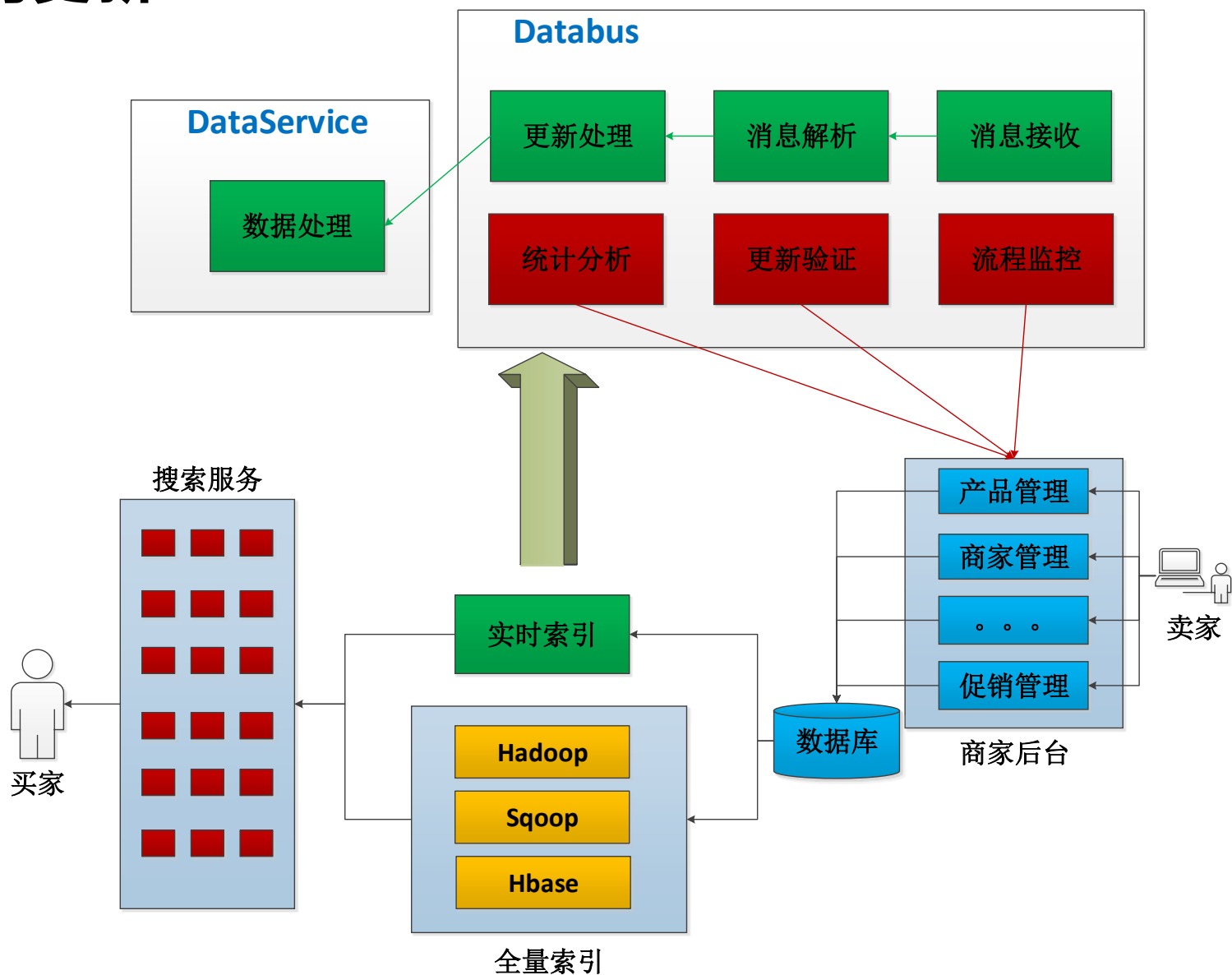
搜索架构



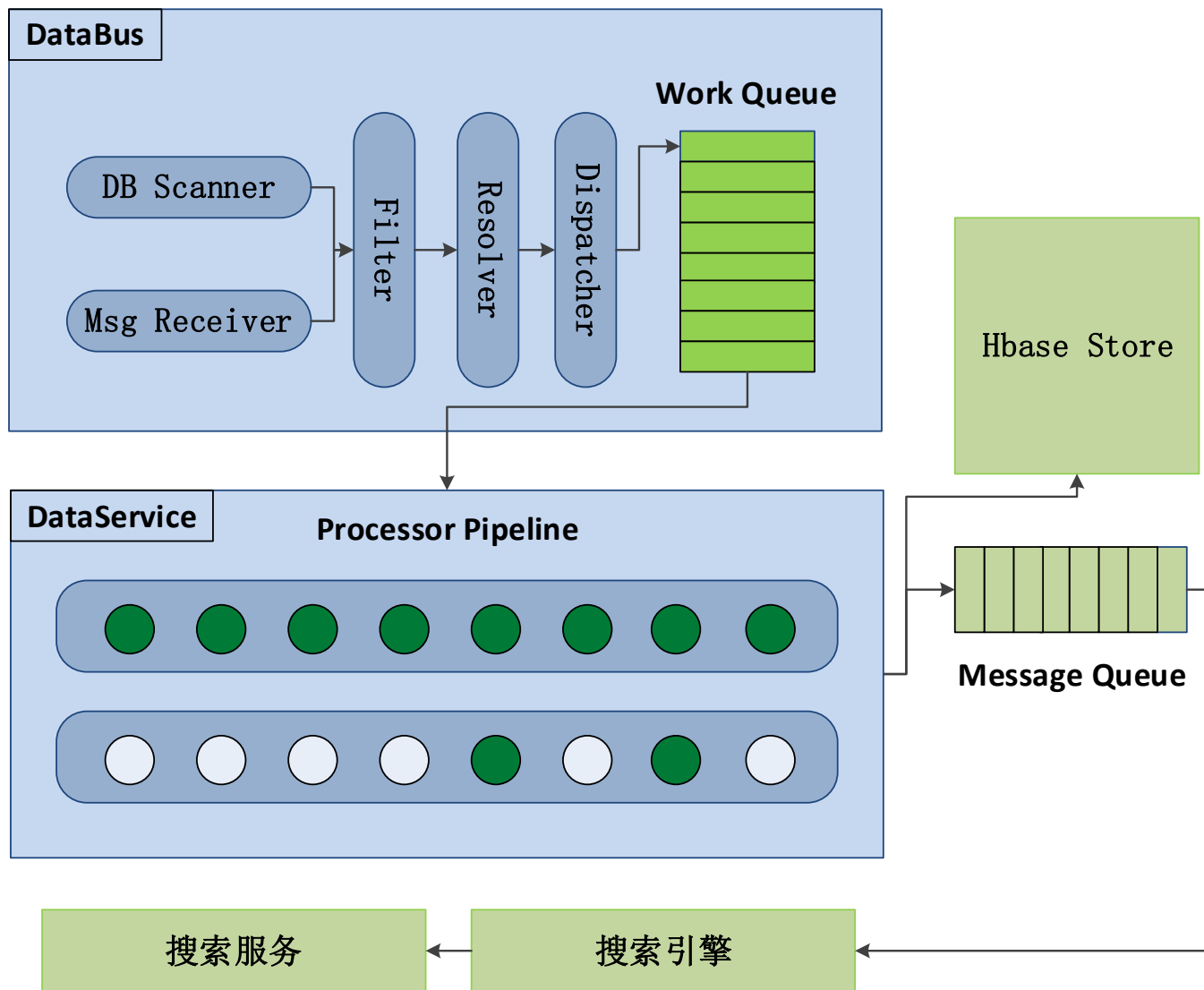
数据更新



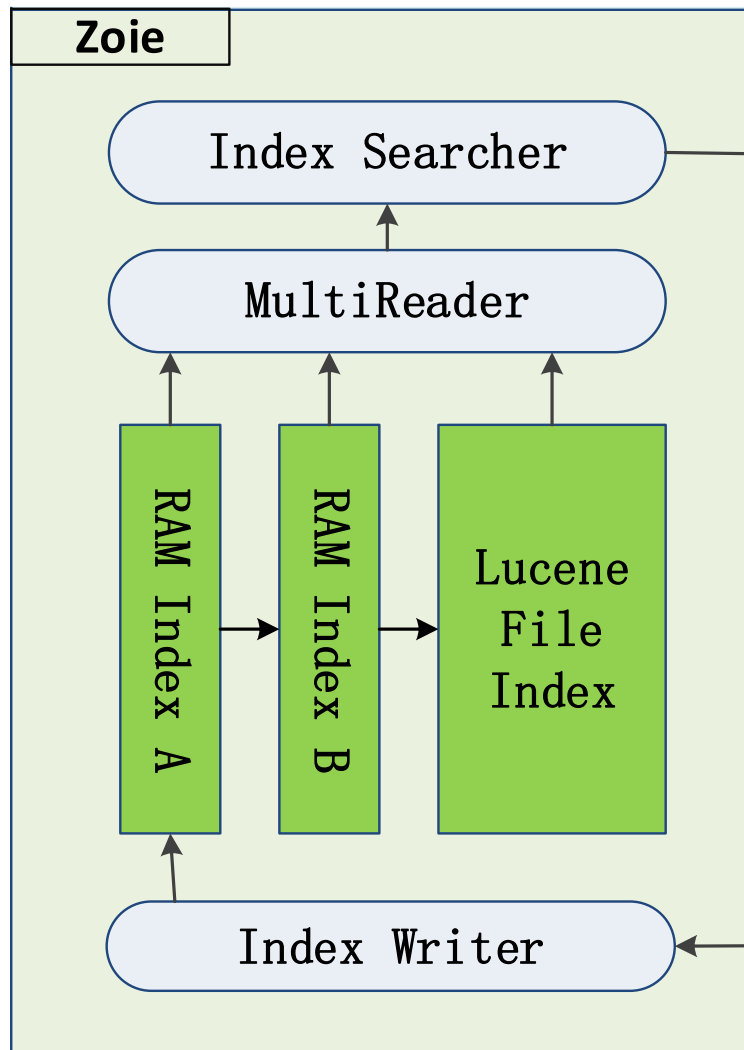
实时更新



实时更新



索引更新



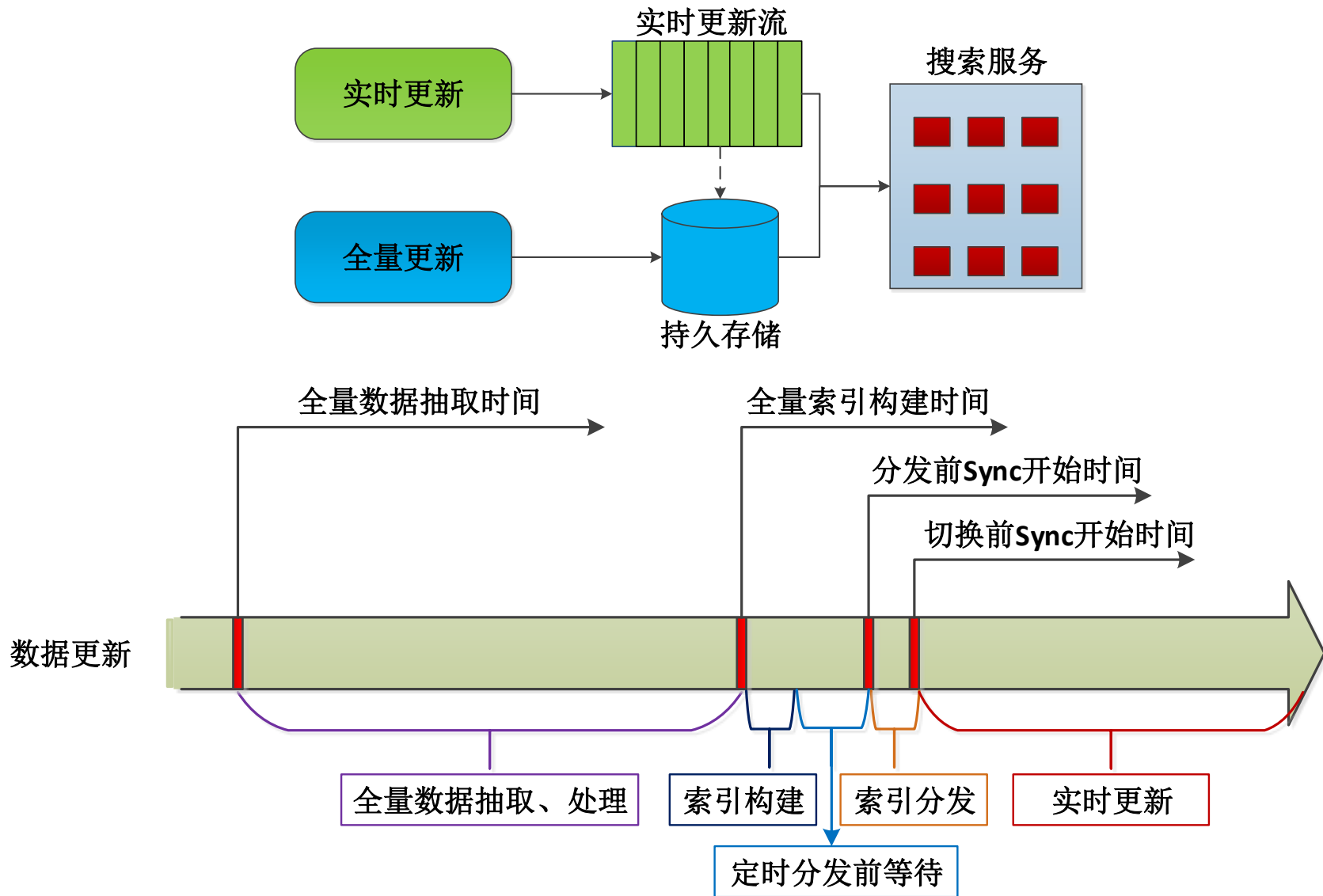
SoftCommit

HardCommit

缓存

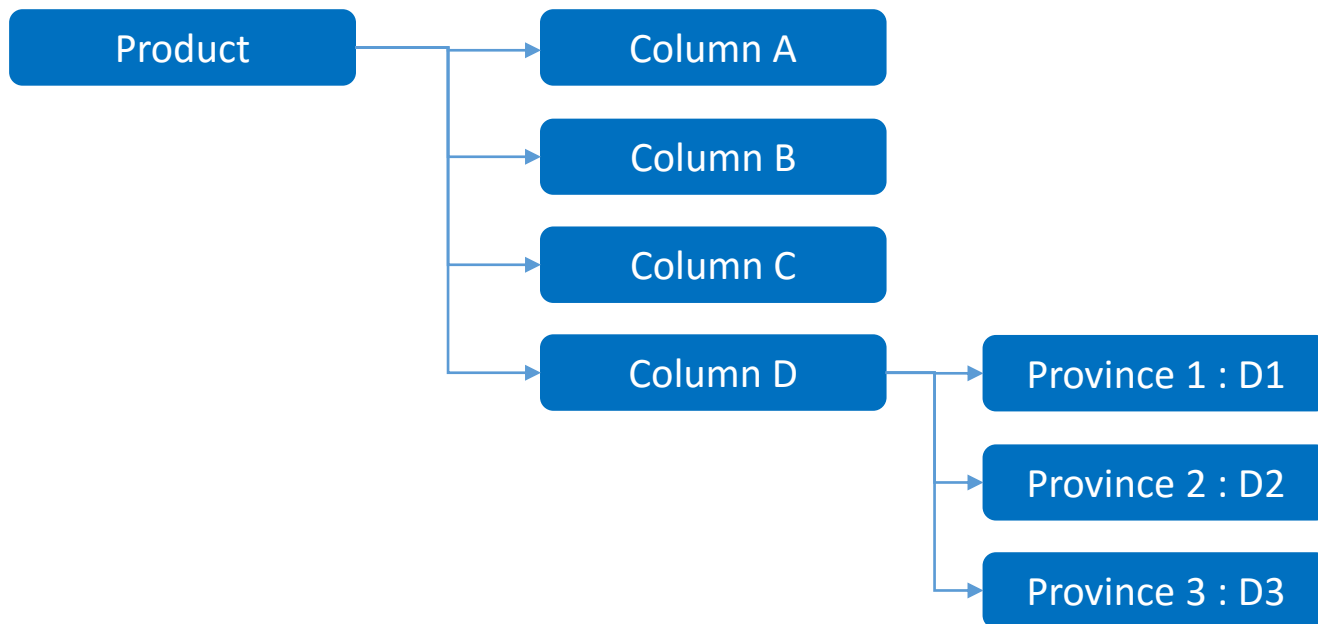


一致性：全量+实时



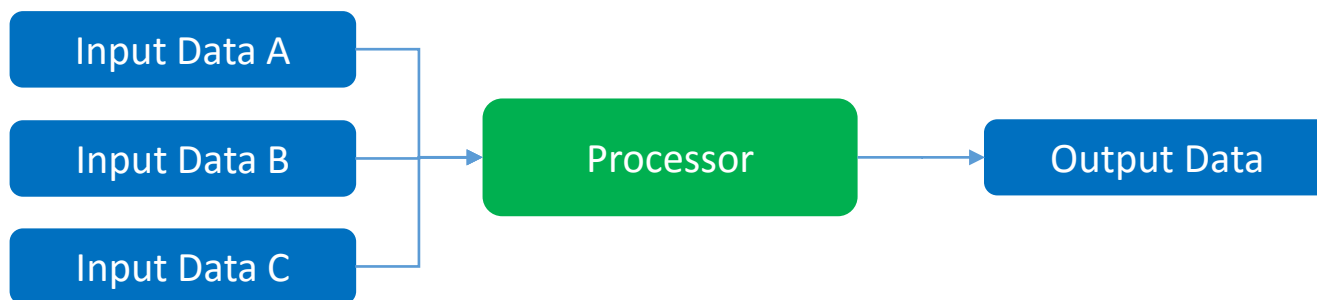
数据模型

- (1) 将产品数据分割为有意义的列，抽象化。
- (2) 数据按照列的粒度来操作、重用。
- (3) 控制复杂度，数据可能有几百上千维，列保持在几十维。



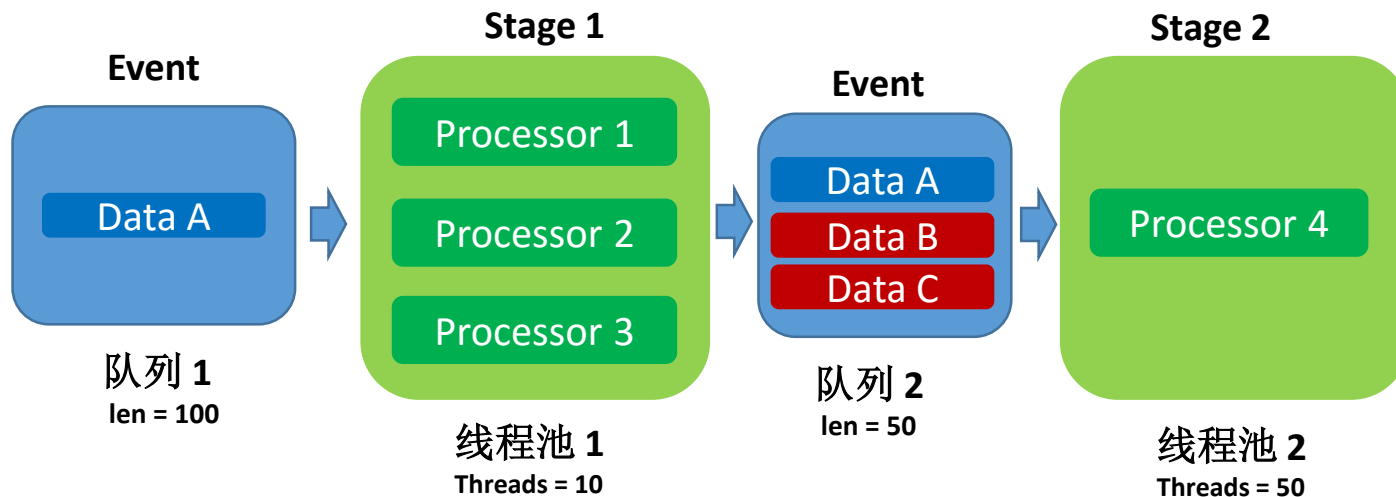
计算模型

(1) 定义输入和输出，将逻辑划分到Processor内



(2) 根据逻辑复杂度，给Processor分配资源

SEDA (Staged Event Driver Architecture)



DTCC

2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia

IT68

ChinaUnix

ITPUB

计算模型

(3) 根据数据的生成和使用关系，计算Processor间的依赖

DataColumn /Processor	MerchantInfo	PriceStock	Promotion
覆盖省份	生成	使用	使用
价格库存		生成	
促销			生成

(4) 根据依赖生成DAG，构建Processor Pipeline



(5) 全量更新，使用MR封装Processor Pipeline

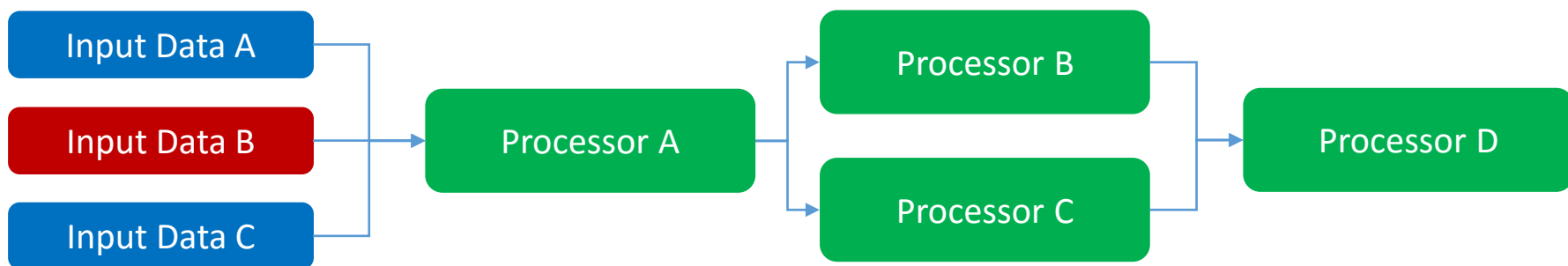


实时更新

含义：某一个**输入数据**发生了改变

(1) 根据定义，找到**使用**这个数据的Processor

(2) 根据**DAG**，构建Processor Pipeline



(3) 更新Pipeline中所有的**输出数据**



数据更新

全量原始数据

产品数据

05-26 17:30

促销数据

05-26 17:30

价格库存

05-26 18:00

全量更新 (Hadoop) 05-27 00:00

产品更新

促销更新

价格库存更新

数据存储

产品数据

05-26 17:30

促销数据

05-27 10:30

05-26 17:30

价格库存

05-26 18:00

索引数据 05-27 00:00

产品数据

05-26 17:30

促销数据

05-26 17:30

价格库存

05-26 18:00

增量更新 05-27 10:30

促销更新

索引数据 05-27 10:45

产品数据

05-26 17:30

促销数据

05-27 10:30

价格库存

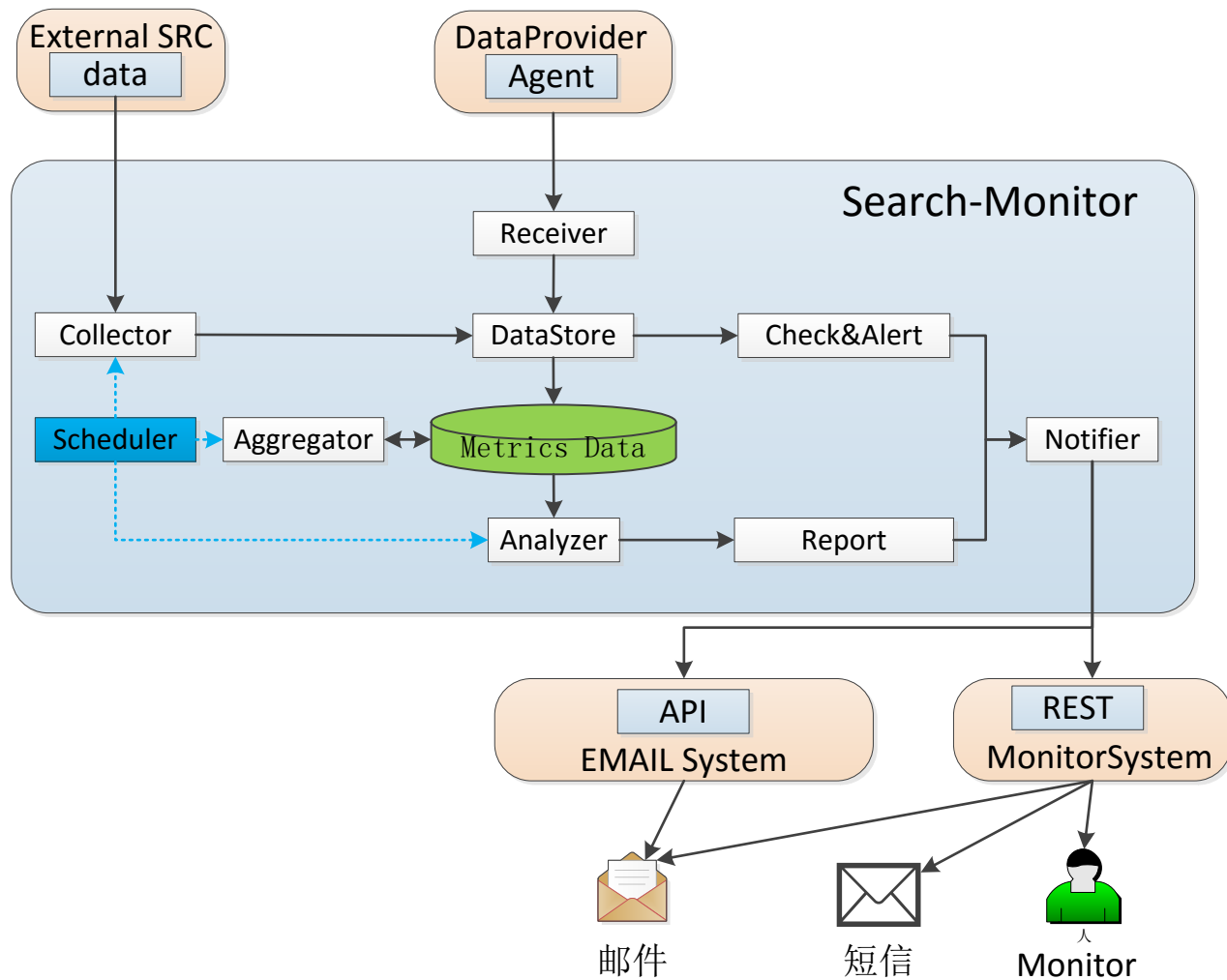
05-26 18:00



实时更新监控

指标：正确性，及时性。

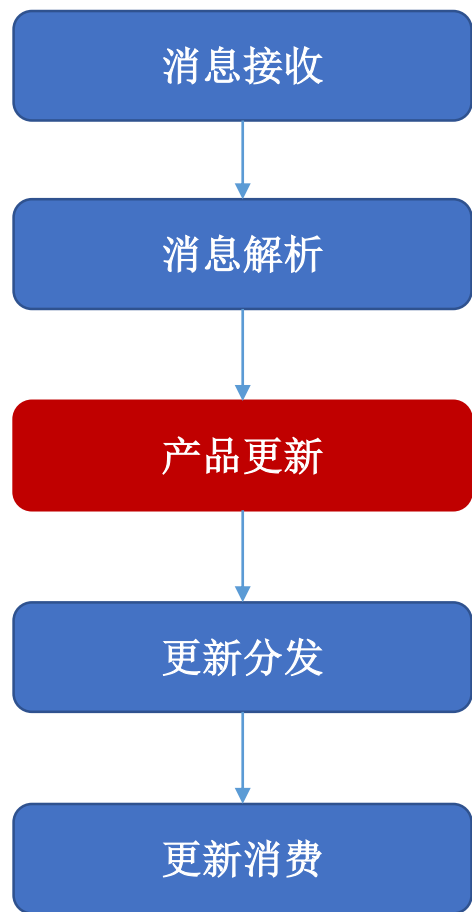
(1) 与系统的集成



(2) 与人的集成



可跟踪性设计



DataChangeMessage

Topic_ReceiveTime_Sequence_ReceiverIP

ProductChange

<DCM Id>_ResolveTime_Sequence

ProductUpdateJob

<PC Id>_ProcessTime_Sequence

UpdateRequest

IndexName_ReceiveTime_ReceiverIP

<PC Id>_ProcessTime_Sequence

IndexRequest

ShardId_IndexName_ReceiveTime_ReceiverIP



监控Dashboard



更新延迟

更新延迟总体统计

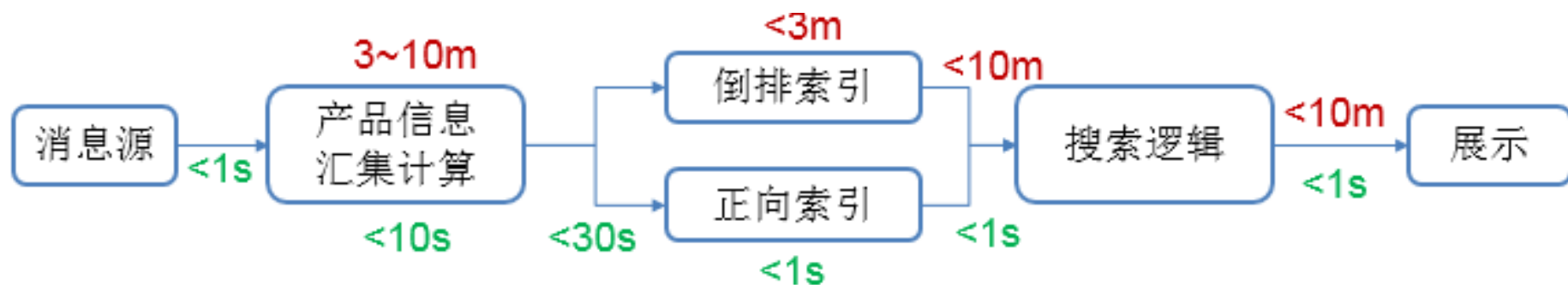
时间段(分钟)	消息数	消息数占比
[0,15)	289,932	100%
[15,30)	3	0%
[30,45)	0	0%
[45,60)	0	0%
[60,75)	0	0%
[75,90)	0	0%
[90,105)	0	0%
[105,120)	0	0%
>=120	0	0%

更新延迟分topic统计(单位: 分钟)

序号	topic	消息数	平均延迟	[0,15)	[15,30)	[30,45)	[45,60)	[60,75)	[75,90)	[90,105)	[105,120)	[120,)
0	backend_market_promo_topic	99	11分30秒	96.97%	3.03%	0%	0%	0%	0%	0%	0%	0%
1	mm_search_punish	125	10分2秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
2	DBChecker-Promotion	135	8分6秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
3	pss_dxx_change	107,494	5分4秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
4	myyhd_change_point_product	324	3分51秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
5	can_show_sale	2,854	3分32秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
6	inshopProd_saleStatus_change	33,846	3分32秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
7	pss_merchant_area_change	227	3分15秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
8	gss_stock_soldout_monitor	224	3分13秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
9	inshop_m_delivery_range	14	3分10秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
10	groupon_search	48	3分7秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
11	product_change	127	3分5秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
12	pminfo_changed	1,484	3分5秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
13	sales_dxx_change	48,353	3分5秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
14	gss_stock_refill_monitor	145	3分5秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
15	inshopProd_normProd_change	102	3分4秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
16	gss_stock_pro_soldout	18,251	3分4秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
17	gss_stock_pro_refill	14,299	3分4秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
18	inshopProd_baseInfo_change	61,647	3分4秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
19	inshop_p_delivery_range	137	3分2秒	100%	0%	0%	0%	0%	0%	0%	0%	0%
--	汇总	289,935	3分53秒	--	--	--	--	--	--	--	--	--



更新延迟



更新失败

- 分阶段统计(粒度为DataChangeMessage)

阶段	消息总数	失败消息数	未完成消息数	成功率
DataChangeMessage	289,945	0	0	100%
ProductChange	289,945	3	0	99.99%
ProductMessage	289,945	0	0	100%
UpdateRequest	289,945	0	0	100%
IndexRequest	289,945	0	0	100%

- 失败消息明细

消息Id	消息时间	消息是否成功	详细链接
1462579801578_gss_stock_soldout_monitor_10.4.6.146_0	2016/05/07 08:10:01	false	detail
1462608601317_gss_stock_soldout_monitor_10.4.4.97_0	2016/05/07 16:10:01	false	detail
1462608990419_sales_dxx_change_10.4.4.97_0	2016/05/07 16:16:30	false	detail

- 阶段明细(消息链路: DataChangeMessage->ProductChange->ProductMessage->UpdateRequest->IndexRequest)

phaseName	startTs	endTs	期望接收条数	实际接收条数	裂变成功条数	裂变失败条数	succeed	数据源链接
DataChangeMessage	2016/05/07 08:10:01	2016/05/07 08:10:01	1	1	1	0	true	sourceData
ProductChange	2016/05/07 08:13:02	2016/05/07 08:13:02	1	1	0	1	false	sourceData
ProductMessage	1970/01/01 08:00:00	1970/01/01 08:00:00	0	0	0	0	true	sourceData
UpdateRequest	1970/01/01 08:00:00	1970/01/01 08:00:00	0	0	0	0	true	sourceData
IndexRequest	1970/01/01 08:00:00	1970/01/01 08:00:00	0	0	0	0	true	sourceData

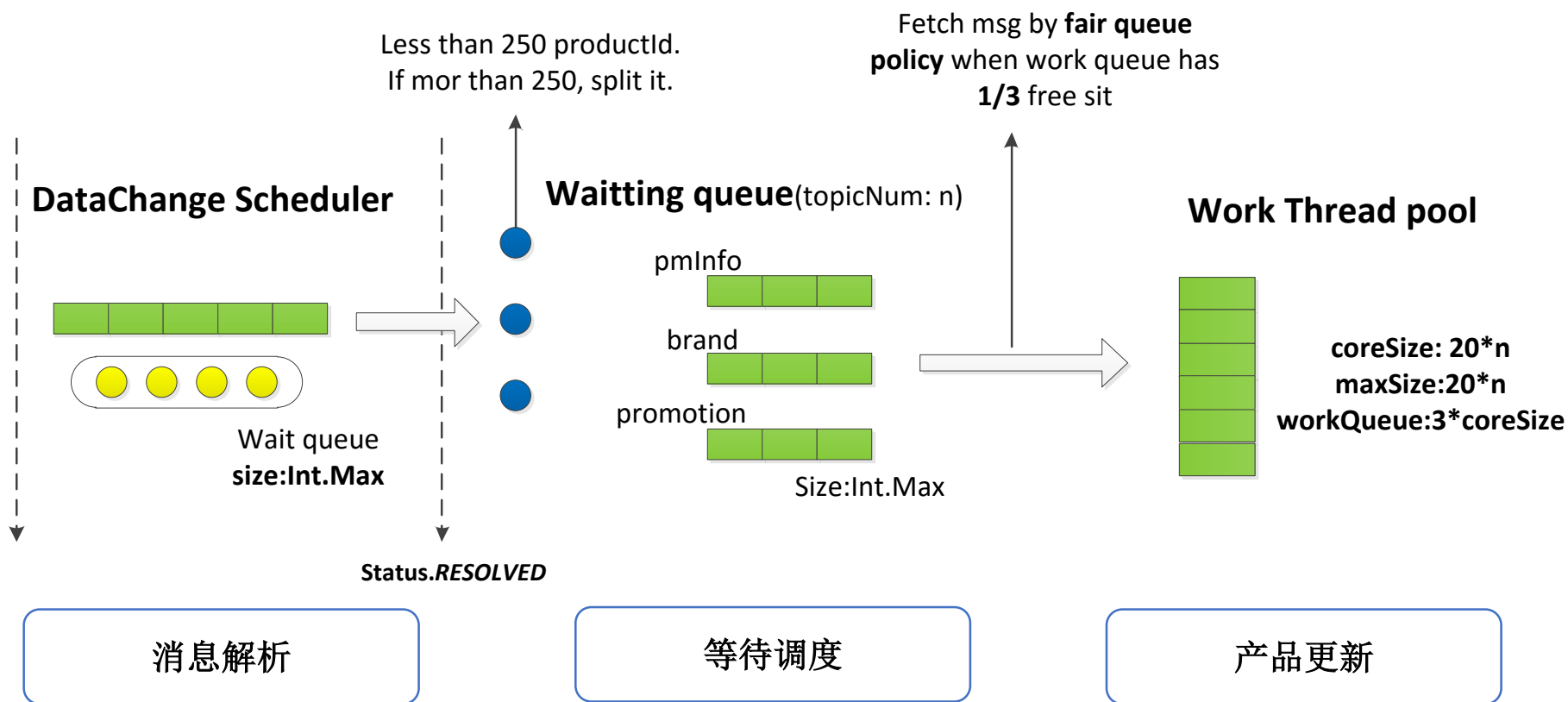


多源更新调度

问题：各个消息源更新频率和数量不均，服务器的资源利用率不高。

方案：使用加权公平队列（WFQ），按照有数据的消息源平均分配处理资源。

效果：40秒完成更新2万条促销消息，优化前需要20分钟。



DTCC

2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia

IT68

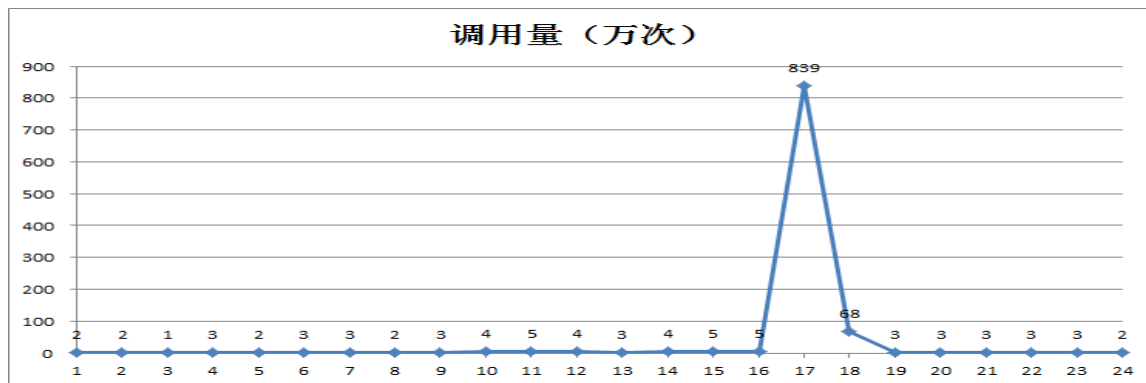
ChinaUnix

ITPUB

进行中的工作

(1) 代码一致性

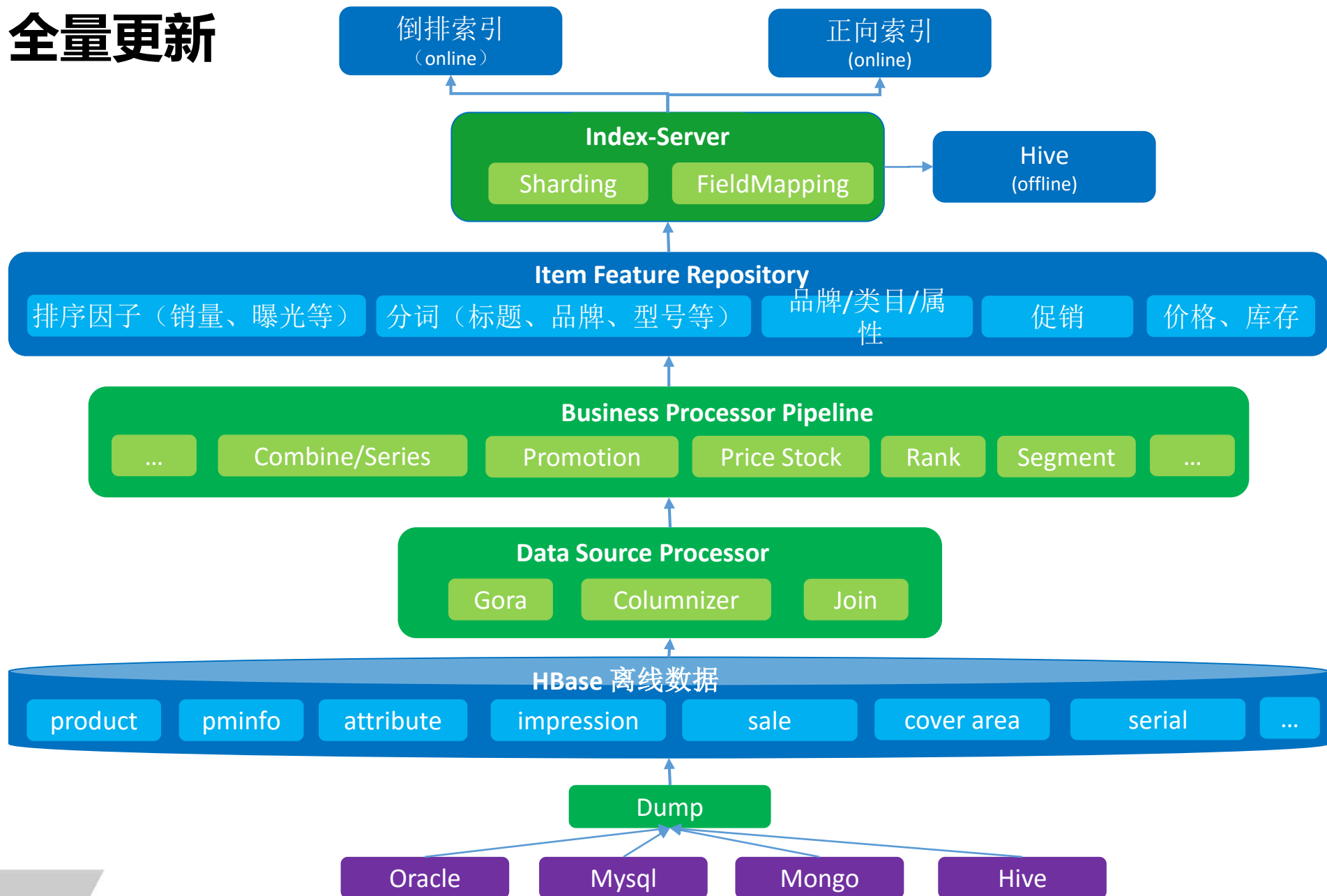
- 架构：业务逻辑 = 20:80
- 业务逻辑代码100%复用
- 降低代码开发和维护的成本
- 缩短测试时间，使用实时更新做功能测试



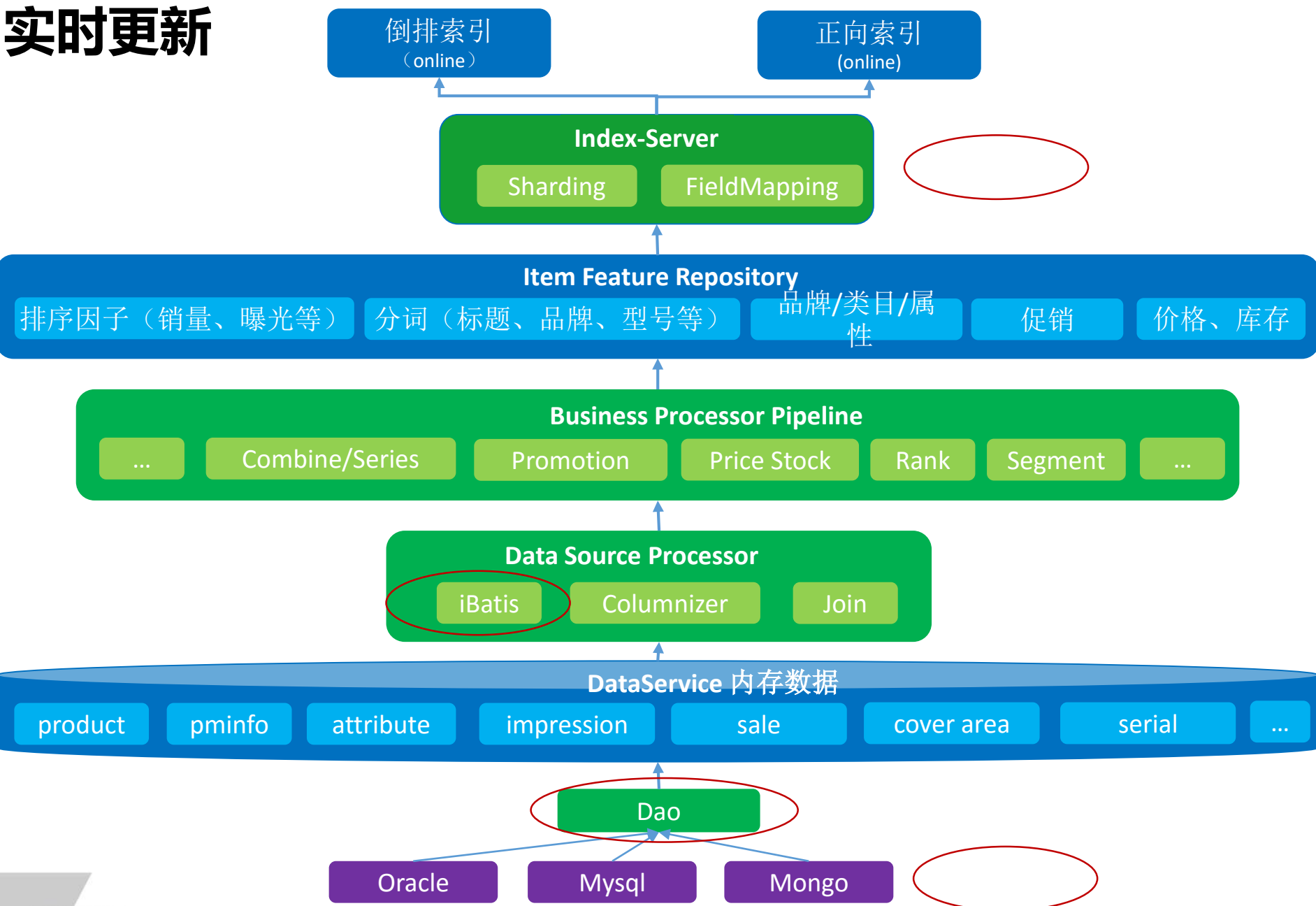
(2) 降低全量构建的频率

- 数据一致性检查，计算全量构建的收益

全量更新



实时更新



We are hiring

上海/武汉

1. 搜索架构师/工程师

- Lucene/Solr/ElasticSearch
- Hadoop/HBase
- Storm/Spark
- 高并发、大数据

2. 算法专家

- 大规模数据挖掘
- 自然语言处理、文本分类
- 排序算法、推荐算法等



zhouhang@yhd.com



DTCC

2016年中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2015

SequeMedia
数据传媒

IT680

ChinaUnix

ITPUB

SequeMedia
盛拓传媒