

智能化大数据平台打造实践

腾讯云大数据基础团队
邹建平

new trend
new technology
new application

Cloud + community
Developer conference

腾讯云大数据产品现状

01

智能化运维的概念和价值

02

智能化大数据平台最佳实践

03

未来展望

04

- 企业级数仓
- 金融风控
- 政务数据
- 用户画像
- 电商推荐
- 点击流分析
-



01

腾讯云大数据产品现状

02

智能化运维的概念和价值

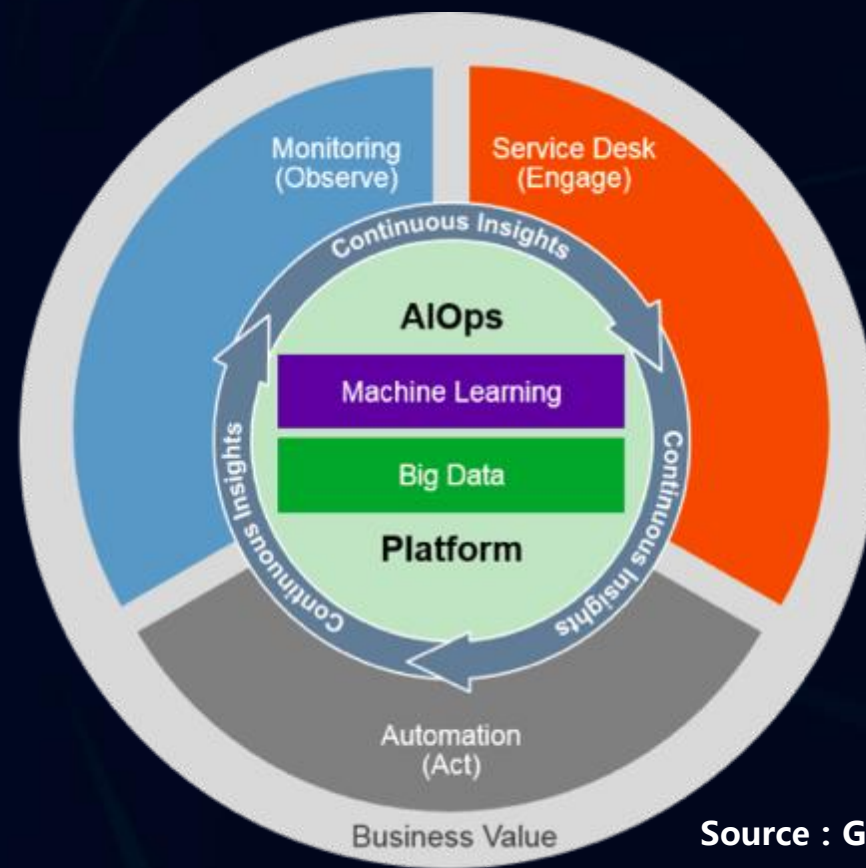
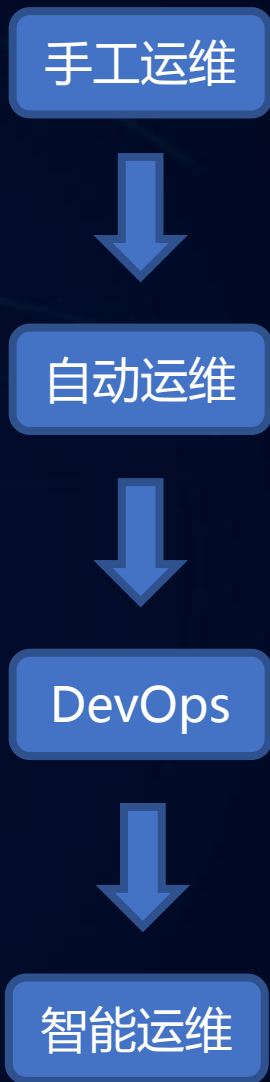
03

智能化大数据平台最佳实践

04

未来展望

智能化运维是什么



Source : Gartner 2017

优化问题：智能调参

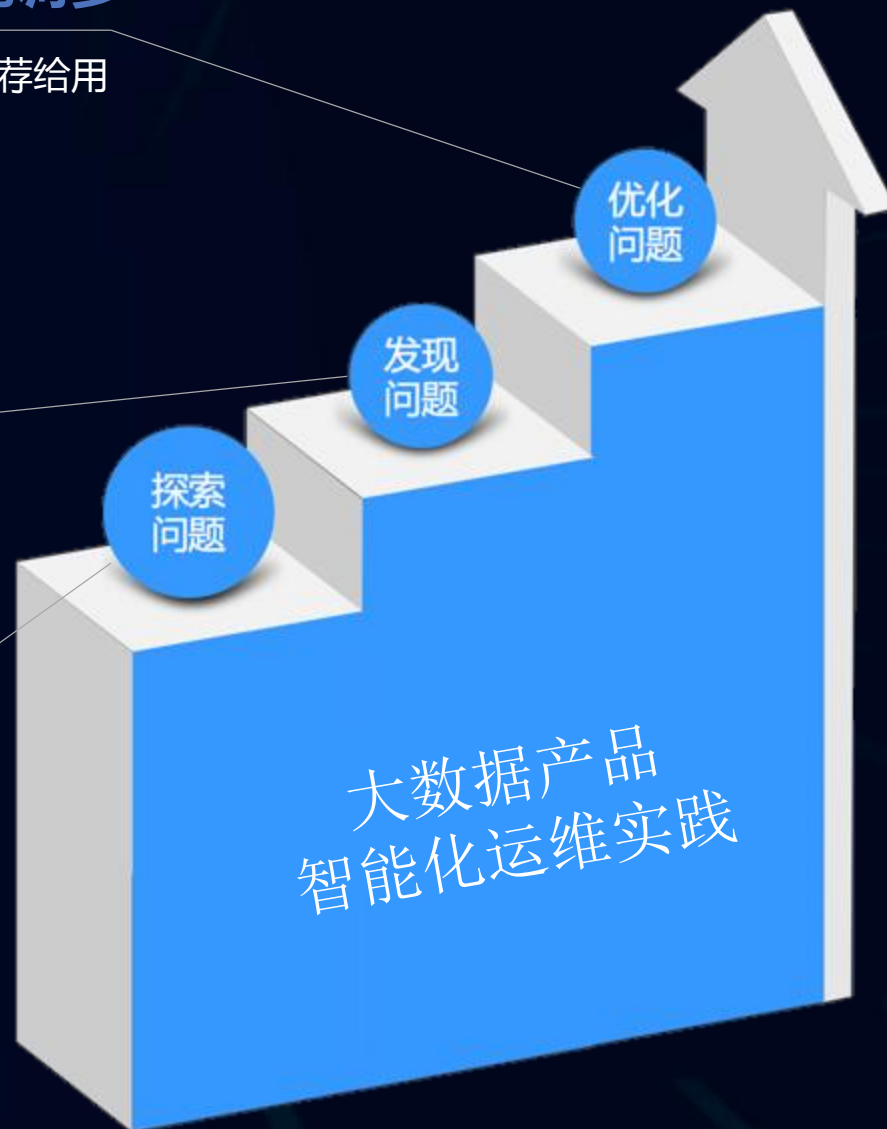
针对腾讯云大数据产品，使用AI算法推荐给用户最优参数

发现问题：异常检测

使用基于统计和机器学习的算法代替传统的人的经验，及时迅速地对各类时间序列做出异常判断

探索问题：日志汇聚

采用ElasticSearch、Jaeger、Dr.elephant等产品将大数据产品的日志和指标汇总，采用可视化的方式呈现服务内部流程，方便复盘问题和追查原因



腾讯云大数据产品现状

01

智能化运维的概念和价值

02

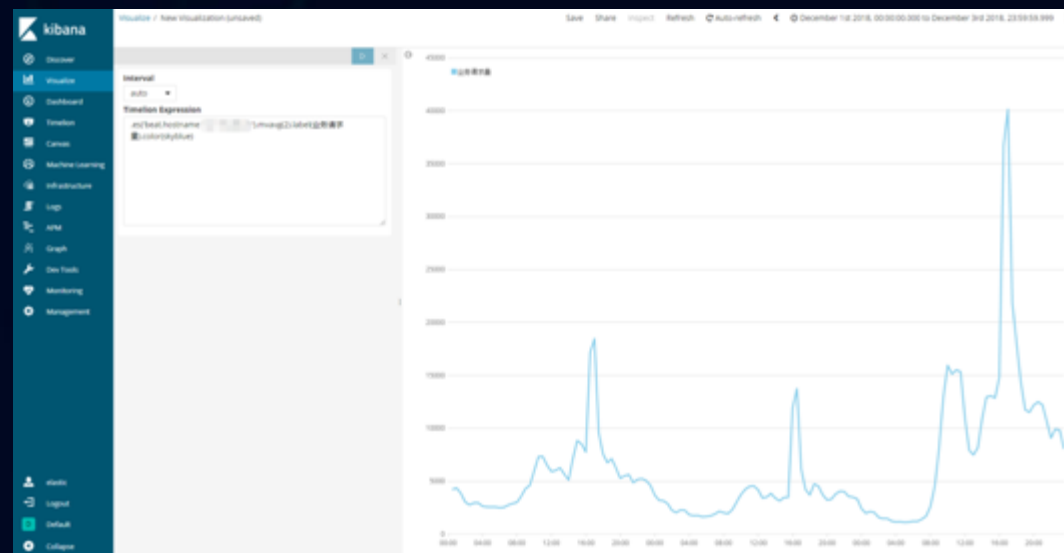
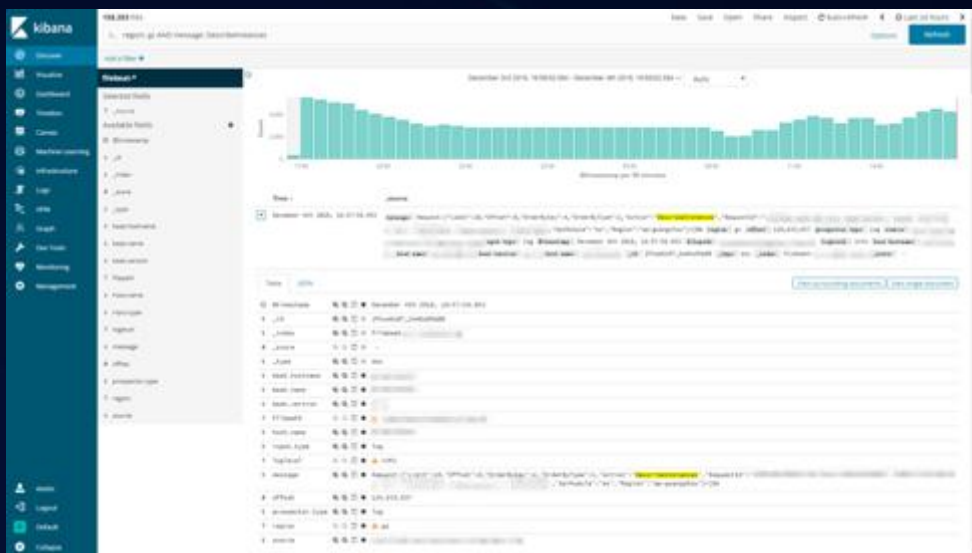
智能化大数据平台最佳实践

03

未来展望

04

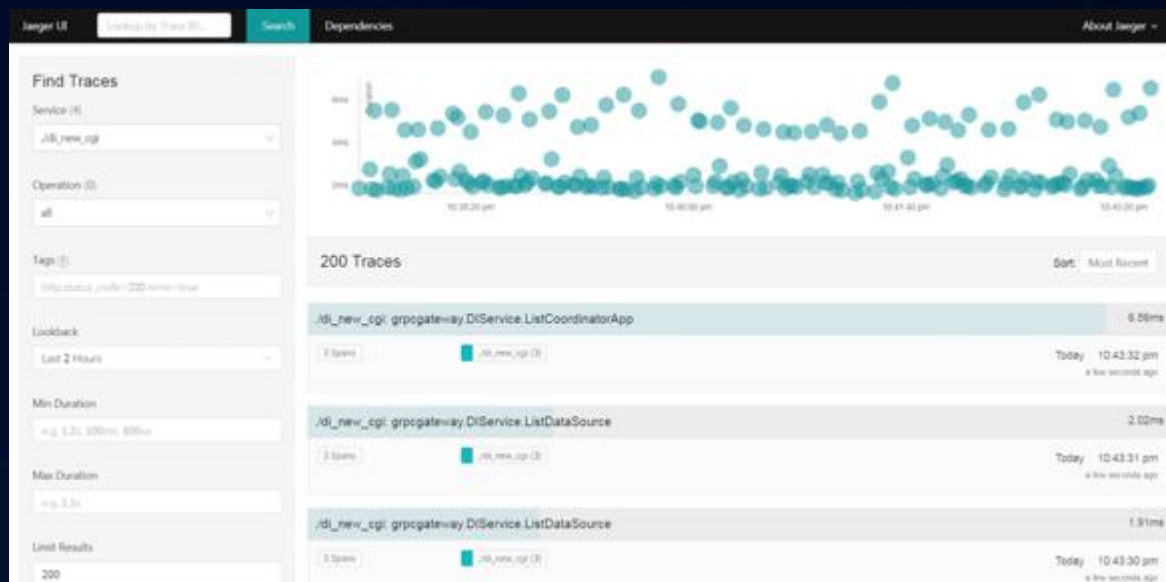
探索问题 — 业务日志收集



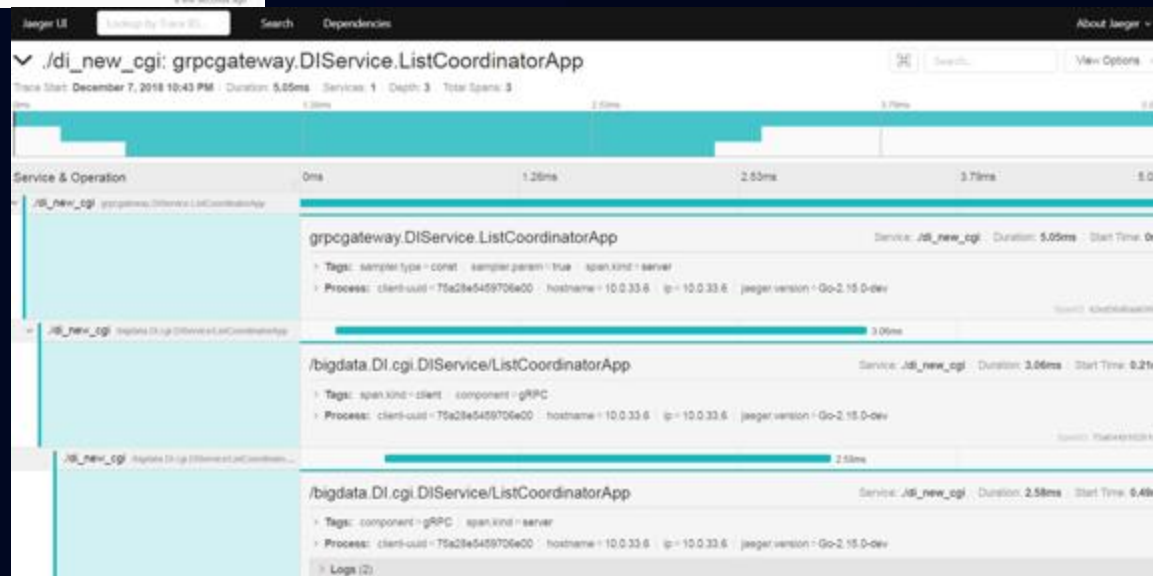
ElasticSearch
Kibana



探索问题 — 模块调用链分析



Jaeger



探索问题 — 作业性能监控与分析

Hello there, I've been busy!
I looked through 58841 jobs today.
About 12425 of them could use some tuning.
About 3666 of them need some serious attention!

Latest jobs

- [User] [Jobtype] application_134343_3434 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
48.774 GB Hours 10.37% 0:10:48 2.52%
- [User] [Hadoop.java] application_148004886726_130804 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
3.280 GB Hours 18.88% 0:00:37 1.52%
- [User] [Pig] application_148004886726_140534 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
6.000 GB Hours 28.88% 0:00:46 27.84%
- [User] [Pig] application_148004886726_140208 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
6.000 GB Hours 18.88% 0:00:36 25.31%
- [User] [Hadoop.java] application_148004886726_140227 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog

Search Results

Job/App ID: [Job ID]

Flow Execution URL/ID: [Flow Execution URL/ID]

User: [User]

Job Type: [Job Type]

Results:

- [User] [Pig] application_148004886726_1405412 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
18.873 GB Hours 9% 0:00:44 1.34%
- [User] [Pig] application_148004886726_1402995 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
18.720 GB Hours 23.76% 0:01:26 27.27%
- [User] [Jobtype] application_134343_3434 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
48.774 GB Hours 10.37% 0:10:48 2.52%
- [User] [Pig] application_148004886726_1402992 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog
18.873 GB Hours 18.88% 0:00:36 1.34%
- [User] [Pig] application_148004886726_1402997 Thu Jul 24 2016 22:25 GMT+08:00 (PDT)
PigLatinSampleLog

Dr. Elephant

Job History Results: <https://dev.advisor.9000manager/analysis-project/analysis-flow-flow-sample-job-cats>

Performance Score (Lower the better)

2k
1.5k
1k
500
0

2016-Mar-01 2016-Mar-03 2016-Mar-05 2016-Mar-07 2016-Mar-09 2016-Mar-11 2016-Mar-13 2016-Mar-15

Task Mar 15 2016 01:52:00 GMT+0800 (PDT)
Job Score = 701

Score Distribution

- Stage 7 36.45%
- Stage 5 13.53%
- Stage 8 12.87%

application_140298322143_2993007
Mapper Memory

- Avg Physical Memory (MB) 308
- Avg task runtime 3 sec
- Avg Virtual Memory (MB) 4860
- Max Physical Memory (MB) 300
- Min Physical Memory (MB) 300
- Number of tasks 1
- Requested Container Memory 4 GB

Job Executions	Stage 1	Stage 4	Stage 5	Stage 6	Stage 7
Mar 15, 2016 01:52 AM	●●●●●	●●●●●	●●●●●	●●●●●	●●●●●
Mar 14, 2016 02:20 AM	●●●●●	●●●●●	●●●●●	●●●●●	●●●●●
Mar 13, 2016 03:24 AM	●●●●●	●●●●●	●●●●●	●●●●●	●●●●●

Search Results

Job/App ID: [Job ID]

Flow Execution URL/ID: [Flow Execution URL/ID]

User: [User]

Queue: [Queue]

Job Type: [Job Type]

Severity: [Severity]

Job Finish Date: [Job Finish Date]

Results:

- [User] [Hadoop.java] application_148004886726_693054 Wed Jan 04 2017 08:41:00 GMT+0800 (CST)
MRResUtilMkLayer
JobTracker: http://mapfml.madech.19989.com/history/job/148004886726_693054
41.777 GB Hours 0% 0:08:05 26.04%

Mapper Data Skew

Severity: None

- Group A 766 tasks @ 1021 MB avg
- Group B 807 tasks @ 1 GB avg
- Number of tasks 1573

Mapper GC

Severity: None

- Avg task CPU time (ms) 41880
- Avg task GC time (ms) 1780
- Avg task runtime (ms) 59859
- Number of tasks 1573
- Task GC/CPU ratio 0.0421707488831016

Mapper Time

Severity: Critical [Explain]

- Average task input size 1 GB
- Average task runtime 99 sec
- Max task runtime 4 min 35 sec
- Min task runtime 19 sec
- Number of tasks 1573

发现问题 — 流计算产品的场景

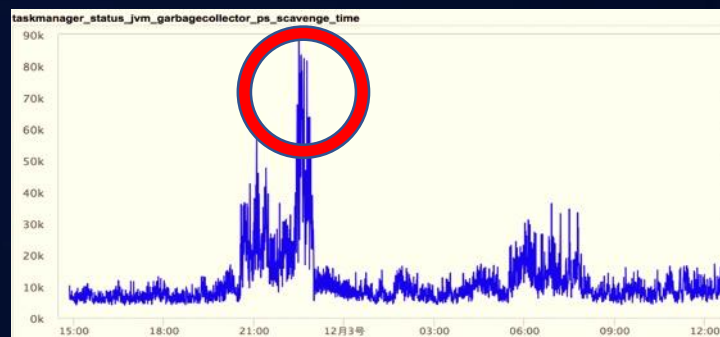
问题：

- 时延和吞吐要求高
- 同一个指标，不同业务场景的告警的要求不同
- 传统阈值告警存在误报、漏报的问题

发现问题 — 流计算产品的场景

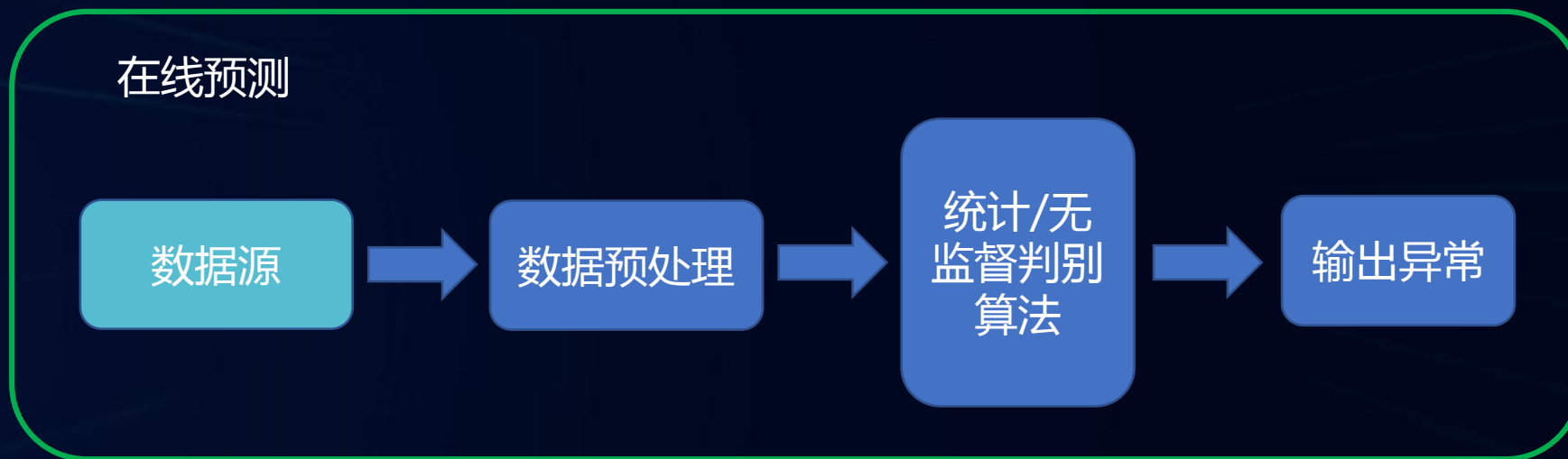
问题：

- 时延和吞吐要求高
- 同一个指标，不同业务场景的告警的要求不同
- 传统阈值告警存在误报、漏报的问题

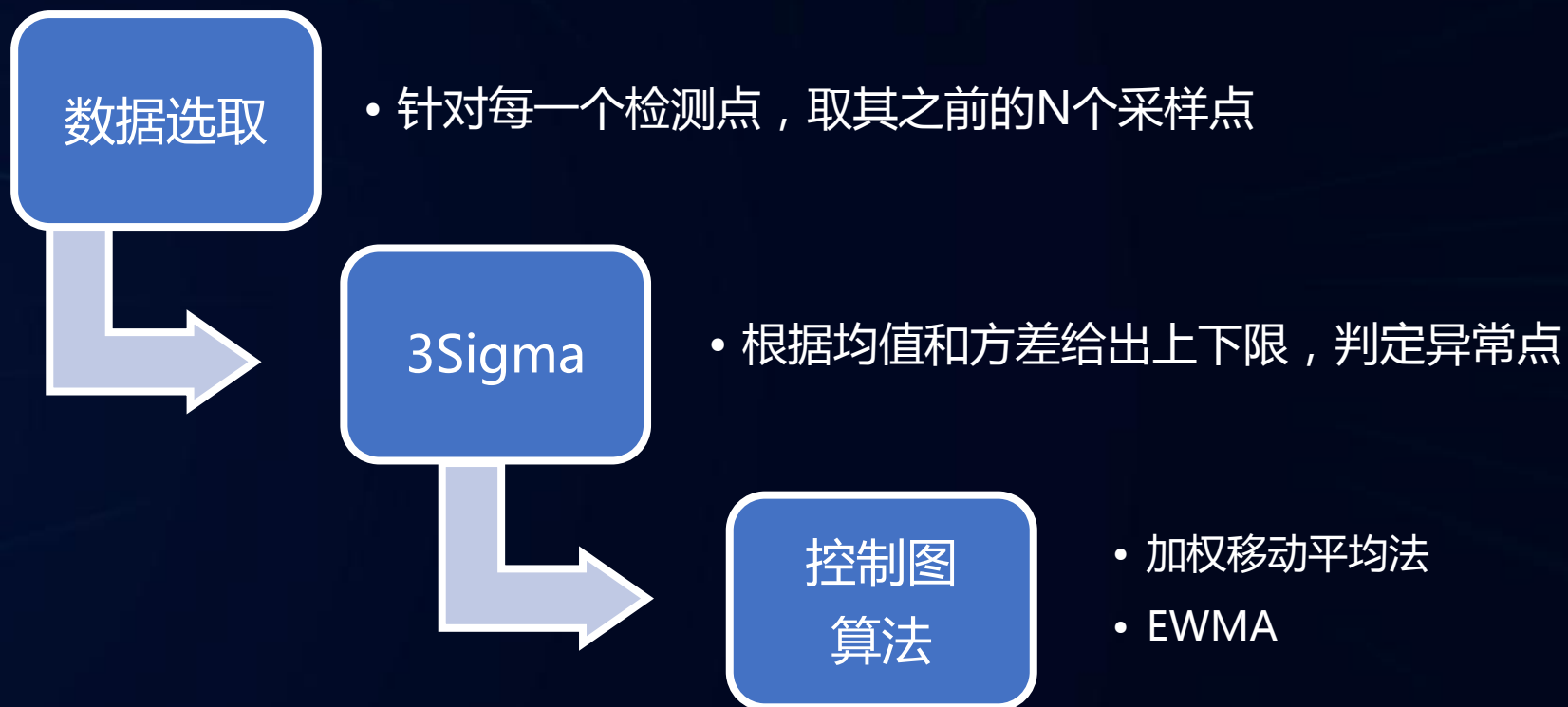


思路：

异常检测：采用基于统计、机器学习算法来自动化判断时序指标是否异常



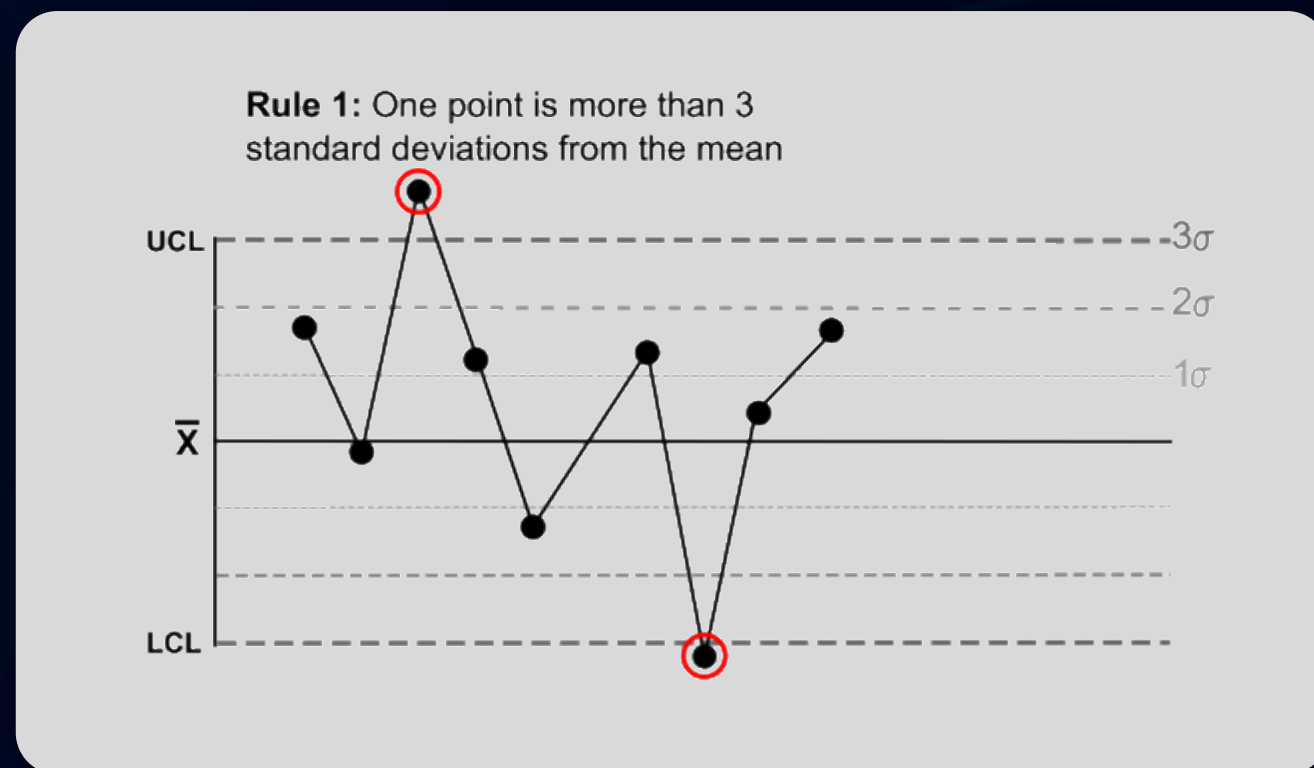
异常检测 — 基于统计的判别方法



异常检测 — 基于统计的判别方法 3Sigma

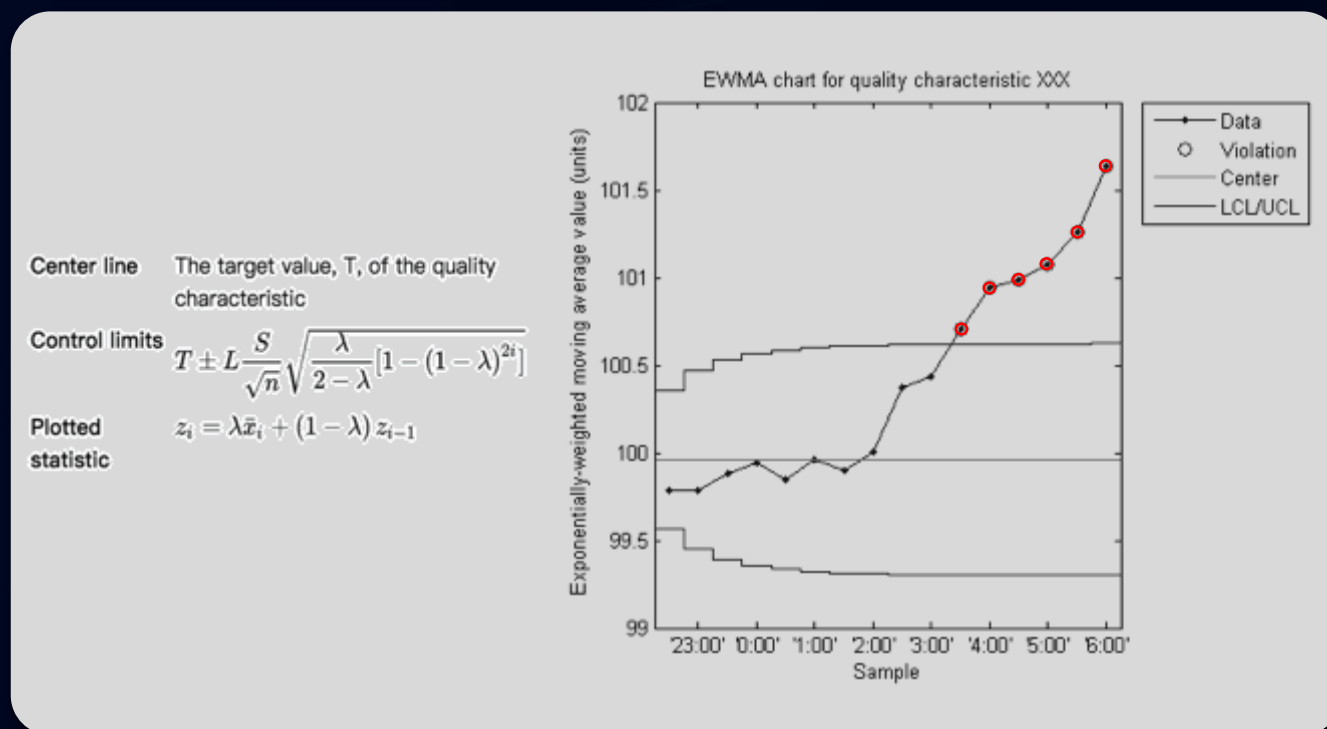
● 3Sigma

- 数据符合正态分布
- 样本落在3sigma范围外为异常点



异常检测 — 基于统计的判别方法 控制图算法

- 加权移动平均法
 - 对观察值给予不同权重，求得移动平均值来确定预测值
- 指数加权移动平均法 EWMA
 - 各数值的加权系数随时间呈指数级递减

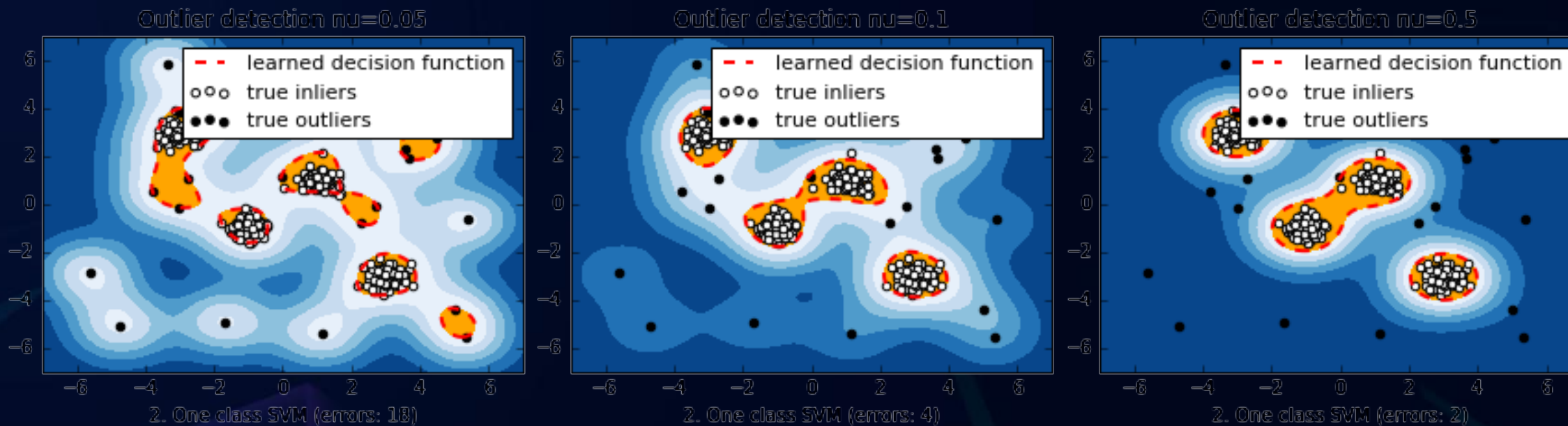


异常检测 — 基于无监督算法的判别方法

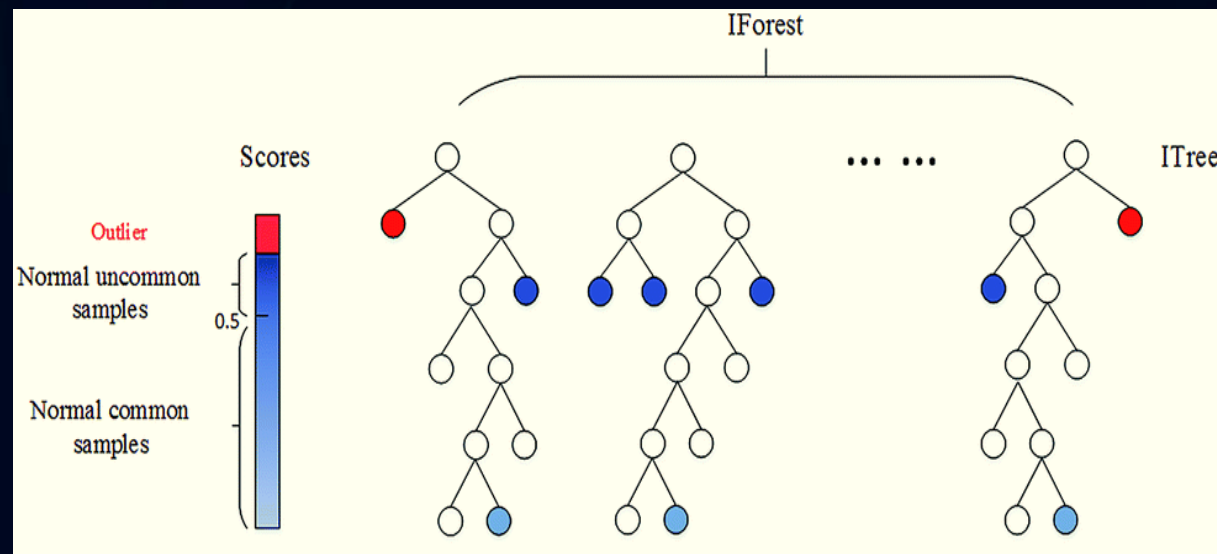


异常检测 — 基于无监督算法的判别方法 One class SVM

- 训练过程使用正常点数据
- 使用核函数训练，训练完之后得到高维平面的决策超平面
- 根据待检测数据与决策超平面的关系决定该点是否异常
- 对于非高斯数据有更加优秀的判定效果



- 基于集成学习的思想
- 基于异常点会很快被划分到叶子结点的思想
- 根据被检测点在多棵树上的与根节点的距离判断该点是否异常
- 内存要求低
- 线性的时间复杂度



Algorithm 1 $iTree(X, e, h)$

Input: X - input data; e - current height; h - height limit.

Output: an $iTree$.

- 1: if $e \geq h$ OR $|X| \leq 1$ then
- 2: return $exNode\{Size \leftarrow |X|\}$;
- 3: else
- 4: Randomly select an attribute q ;
- 5: Randomly select a split point p between min and max values of attribute q in X ;
- 6: $X_l \leftarrow filter(X, q < p)$, $X_r \leftarrow filter(X, q \geq p)$;
- 7: return $inNode\{ Left \leftarrow iTTree(X_l, e + 1, h)$,
 $Right \leftarrow iTTree(X_r, e + 1, h)$,
 $SplitAttr \leftarrow q, SplitValue \leftarrow p\}$;
- 8: end if

异常检测 — 基于统计/无监督算法的判别方法

- 优点

 - 性能较好

 - 无需标注

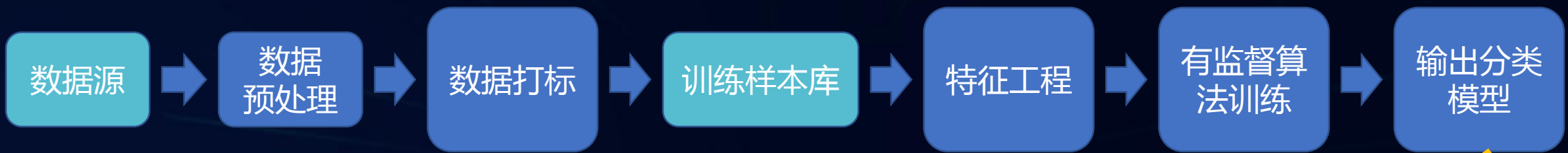
- 缺点

 - 适应面窄

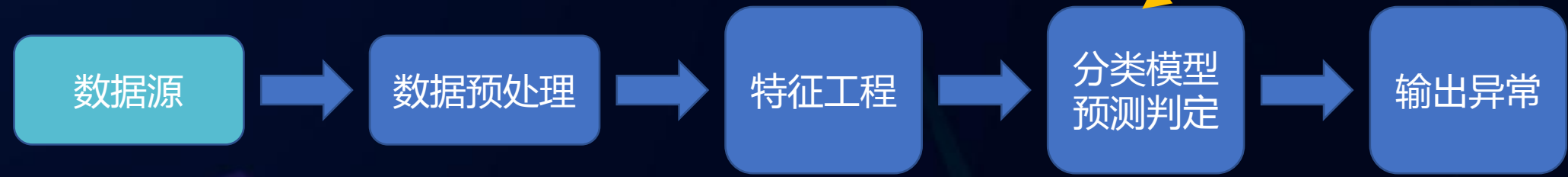
 - 分辨能力一般

异常检测 — 基于有监督算法的判别方法 主要流程

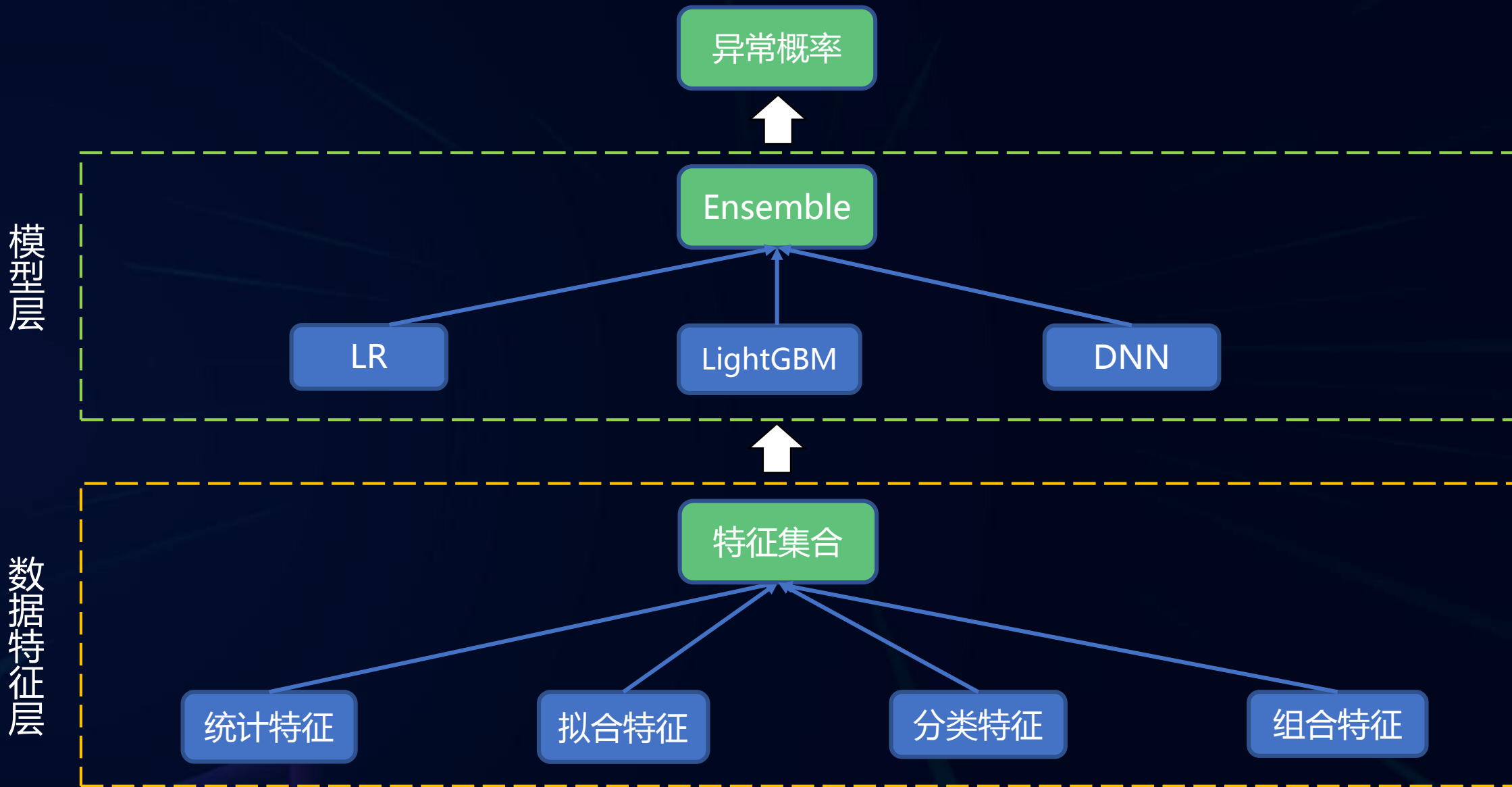
离线训练



在线预测



异常检测 — 基于有监督算法的判别方法 算法框架



- 最值、均值、方差
- 中位数、二分位数、四分位数...

统计特征

1

- 移动平均
- 带权重移动平均
- 指数移动平均
- ...

拟合特征

2

- DTW特征
- 值分布特征
- ...

分类特征

3

- 多维度特征组合

组合特征

4

异常检测 — 基于有监督算法的判别方法

- 优点

 - 适应面广

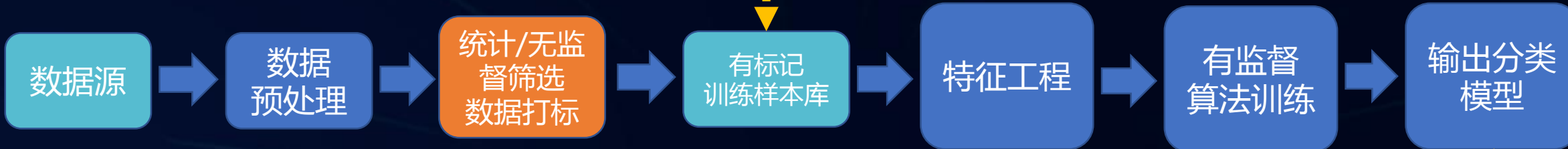
 - 分辨能力强

- 缺点

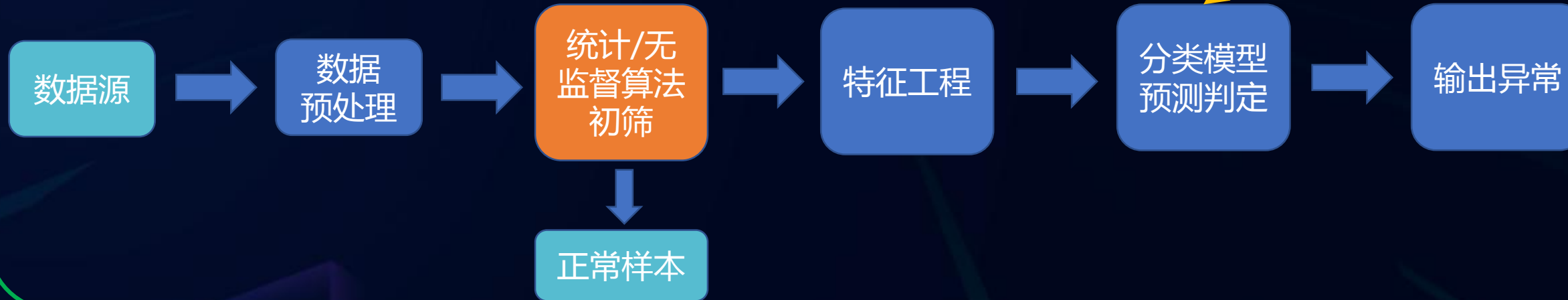
 - 标注成本高

 - 特征工程和训练预测，计算开销大

离线训练



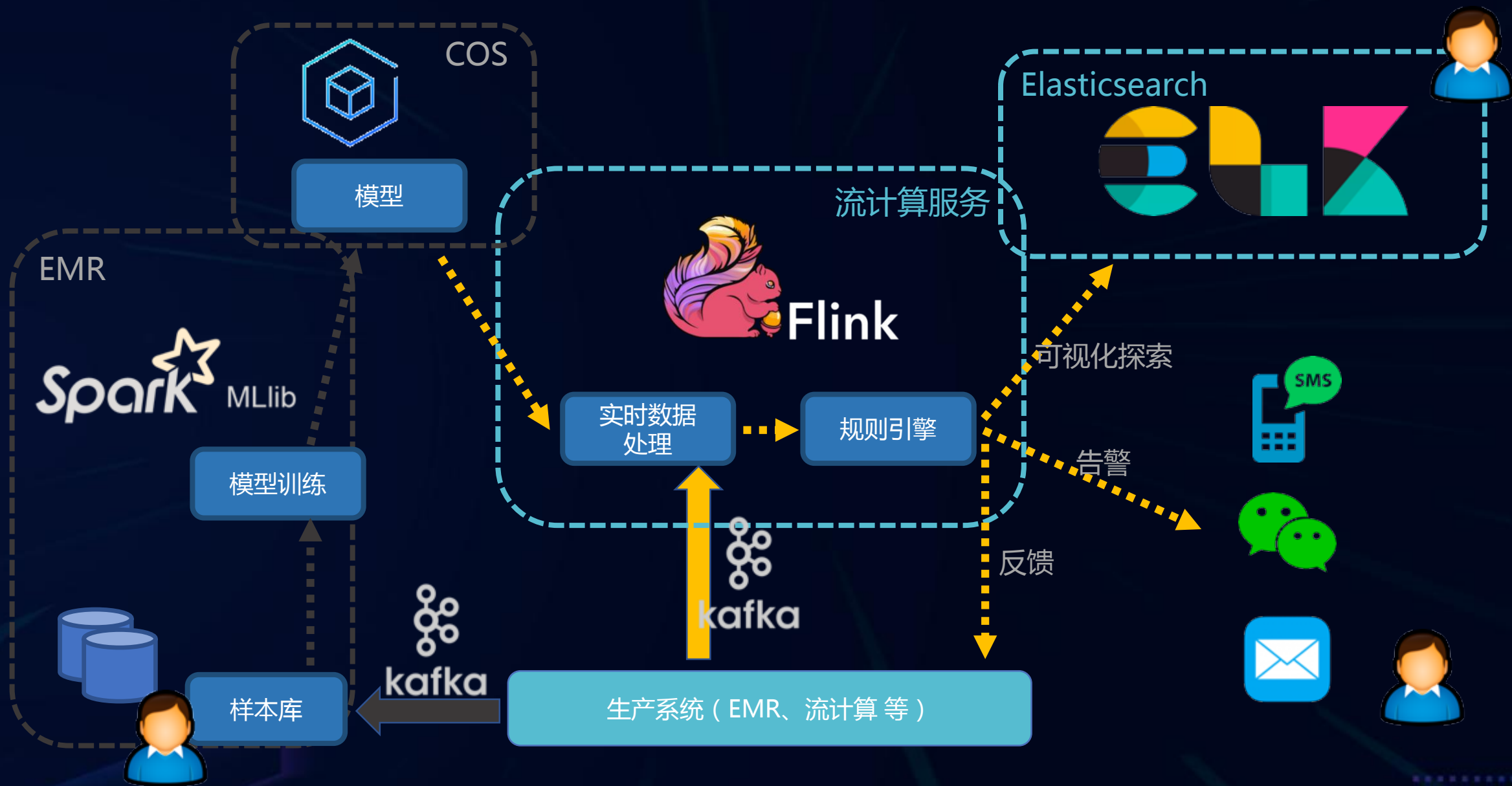
在线预测



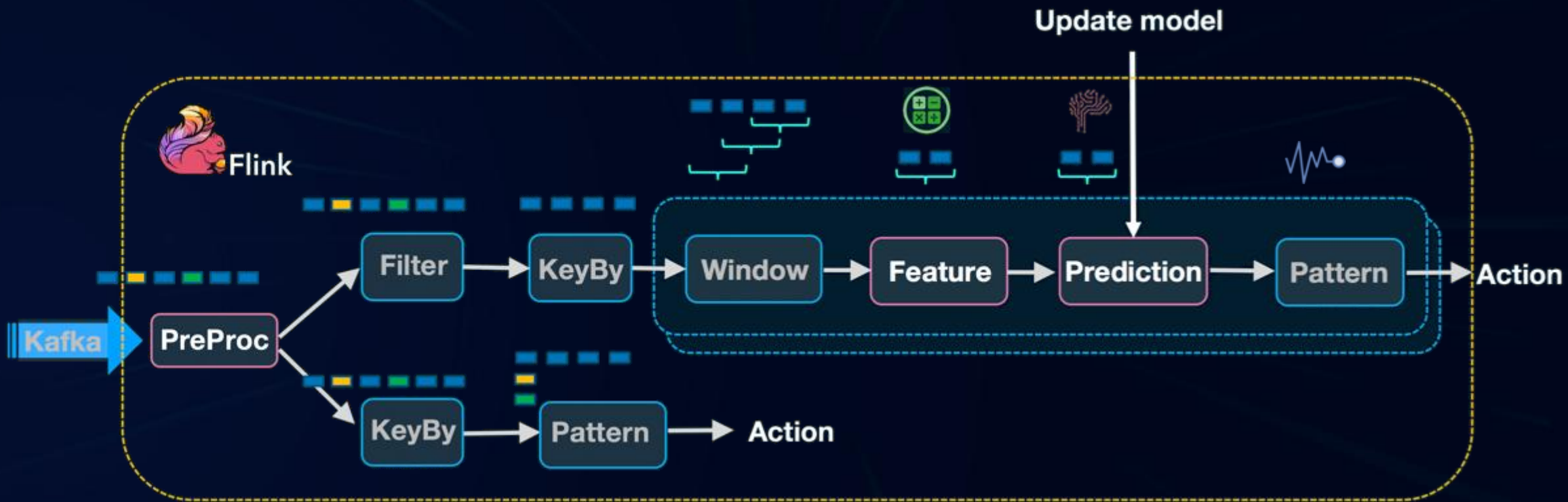
异常检测 — 整体系统概要图



异常检测 — 整体系统概要图



异常检测 — 实时计算逻辑流程



异常检测 — 验证效果

- 准确率：检查异常上报的视图，人工确认是否异常
- 召回率：收集一段时间的时序数据，并在模型里进行测试
- 检测模型给出异常概率大于0.5即判定为异常

准确率 > 88%

召回率 > 79%

优化方向：

- 使用聚类算法做分类，匹配相应的预测模型
- 数据周期性分析，参考周期数据采样提升准确率

优化问题 — EMR产品的应用场景

Summary												
	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write
Active(16)	111	6.9 MB / 45.7 GB	0.0 B	45	106	50	2186	2342	8.94 h (1.36 h)	80.6 GB	2.2 MB	1019.7 MB
Dead(18)	174	10.8 MB / 54.4 GB	0.0 B	54	0	43	3086	3129	17.00 h (7.05 h)	183.7 GB	2.1 MB	93.5 MB
Total(34)	285	17.6 MB / 100.2 GB	0.0 B	99	106	93	5272	5471	25.94 h (8.41 h)	264.3 GB	4.3 MB	1113.2 MB

doop Application application_1543602466024_0026

Kill Application

User: [redacted]
Name: [redacted]
Application Type: MAPREDUCE
Application Tags:
YarnApplicationState: FINISHED
Queue: root.default
FinalStatus Reported by AM: SUCCEEDED
Started: Sun Dec 02 17:09:46 +0800 2018
Elapsed: 26mins, 7sec
Tracking URL: History
Diagnostics:

mapreduce.task.io.sort.mb
mapreduce.job.reduces

doop Application application_1543602466024_0030

Kill Application

User: [redacted]
Name: [redacted]
Application Type: MAPREDUCE
Application Tags:
YarnApplicationState: FINISHED
Queue: root.default
FinalStatus Reported by AM: SUCCEEDED
Started: Sun Dec 02 20:51:50 +0800 2018
Elapsed: 21mins, 12sec
Tracking URL: History
Diagnostics:

优化问题 — EMR产品的应用场景

问题：

- 产品涉及组件多
- 业务配置复杂
- 人工配置参数繁琐，容易出错

思路：

智能调参：使用算法正确预测每一组参数配置下的运行时间与资源利用率，实质是一个回归问题

Summary												
	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write
Active(16)	111	6.9 MB / 45.7 GB	0.0 B	45	106	50	2186	2342	8.94 h (1.36 h)	80.6 GB	2.2 MB	1019.7 MB
Dead(18)	174	10.8 MB / 54.4 GB	0.0 B	54	0	43	3086	3129	17.00 h (7.05 h)	183.7 GB	2.1 MB	93.5 MB
Total(34)	285	17.6 MB / 100.2 GB	0.0 B	99	106	93	5272	5471	25.94 h (8.41 h)	264.3 GB	4.3 MB	1113.2 MB

Application application_1543602466024_0026

Kill Application

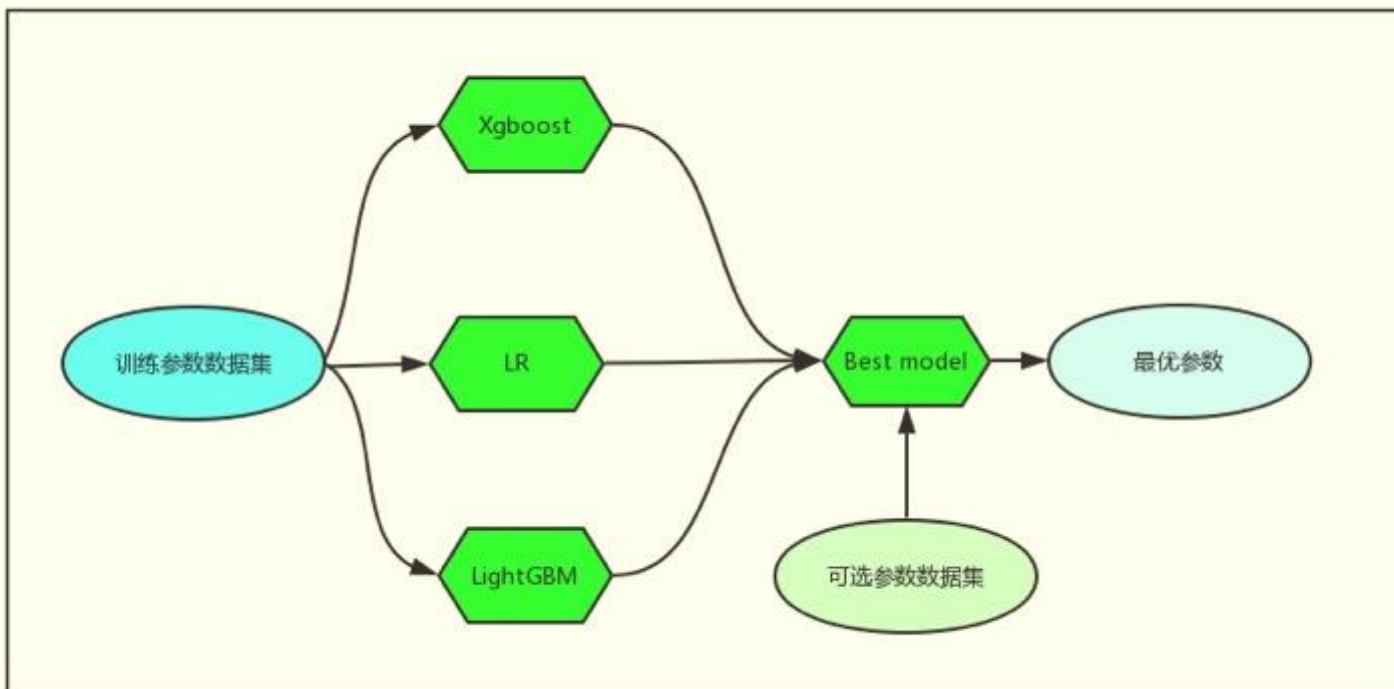
User: [redacted]
 Name: [redacted]
 Application Type: MAPREDUCE
 Application Tags:
 YarnApplicationState: FINISHED
 Queue: root.default
 FinalStatus Reported by AM: SUCCEEDED
 Started: Sun Dec 02 17:09:46 +0800 2018
 Elapsed: 26mins, 7sec
 Tracking URL: History
 Diagnostics:

mapreduce.task.io.sort.mb
 mapreduce.job.reduces

Application application_1543602466024_0030

Kill Application

User: [redacted]
 Name: [redacted]
 Application Type: MAPREDUCE
 Application Tags:
 YarnApplicationState: FINISHED
 Queue: root.default
 FinalStatus Reported by AM: SUCCEEDED
 Started: Sun Dec 02 20:51:50 +0800 2018
 Elapsed: 21mins, 12sec
 Tracking URL: History
 Diagnostics:



- 参数数据集作为训练数据
- 训练回归模型
- 通过模型寻找任务参数与目标性能指标之间的关系

问题：过程较为复杂，依赖数据数量和质量

智能调参 — 强化学习是否可以解决调参问题？

- 模拟人类与环境的交互学习方式
- AlphaGO/机器人

学习一个最优策略，可以让Agent在Environment中，根据当前的State，做出Action，从而获得最大回报

State

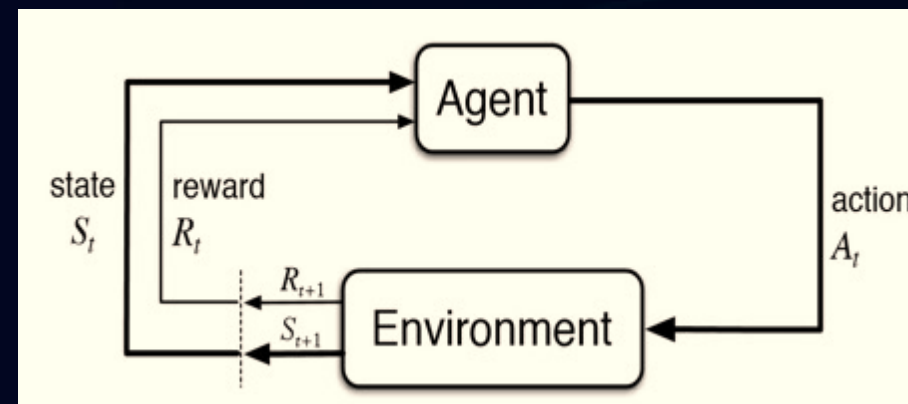
- 一个表示Environment的数据

Action

- Agent可以做的动作

Reward

- 每个动作得到的奖励



智能调参 — Qlearning

初始化 $Q(s, a), \forall s \in S, a \in A(s)$, 任意的数值, 并且 $Q(\text{terminal} - \text{state}, \cdot) = 0$

重复 (对每一节 episode) :

初始化 状态 S

重复 (对 episode 中的每一步) :

使用某一个 policy 比如 ($\epsilon - greedy$) 根据状态 S 选取一个动作执行

执行完动作后, 观察 reward 和新的状态 S'

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha (R_{t+1} + \lambda \max_a Q(S_{t+1}, a) - Q(S_t, A_t))$$

$$S \leftarrow S'$$

循环直到 S 终止

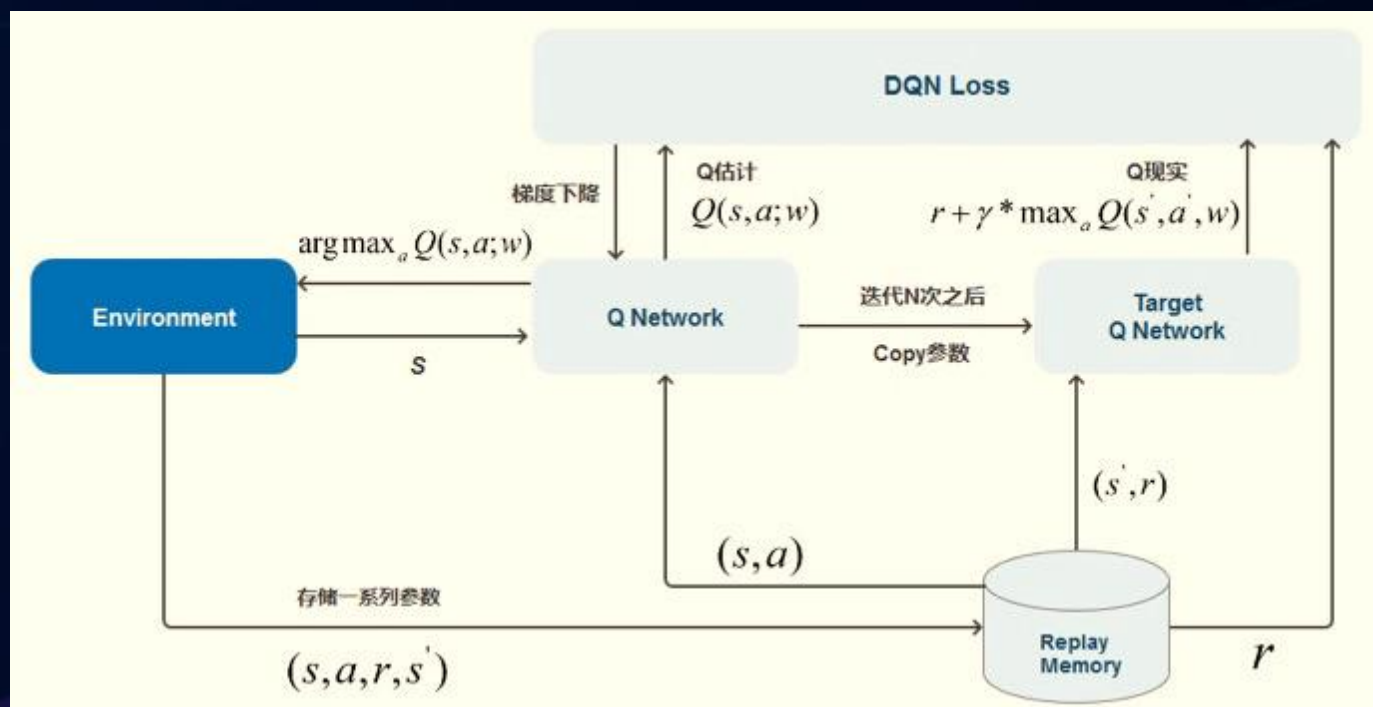
Q learning : Q值函数更新算法



问题 : 无法解决状态或动作空间维度爆炸的场景

智能调参 — 深度强化学习算法 (DQN)

- 采用DNN用来逼近Q函数
- 使用两个神经网络，一个产生当前Q值，另一个产生目标Q值
- 使用经验池Replay Memory来回放随机小批量训练数据



智能调参 — DQN在EMR场景的落地

● State、Action、Reward

- State : EMR参数数值列表 $\langle \text{para0}, \text{para1}, \dots, \text{paraN} \rangle$
- Action: 对于每一个参数采用三维向量描述对State的改变 $\langle 0:\text{变大}, 1:\text{不变}, 2:\text{变小} \rangle$
- 在某个State，选择Action之后，执行action，根据性能变化情况获取此次的Reward

● Exploration & Exploitation

- 以10%的概率：随机选择一个Action（探索）；否则，选择当前Q值最大的Action（开发）

● Replay memory

- 每次从大数据平台（环境）得到 $\langle s, a, r, s' \rangle$ 之后将其放入到记忆池中，训练时从中随机拿出小批量测试样本来训练

● Double DQN

- 动作的选择和评估用不同的值函数，避免Q值偏高估计引起的收敛速度偏慢

智能调参 — 效果对比

优点：

- 对样本数据的依赖较小
- 能够较快达到较优解

改进方向：

- 如何将在一个业务上的经验积累可以直接用到其他业务

算法	初始依赖样本量	目标达成时间(h)	任务执行时间(s)	优化效果
默认配置	N/A		42	
LR、TreeModel	800+	9	33	21%
DQN	240+	4	32	23%
Double DQN	240+	3	32	23%

- 测试使用intel-hadoop的Hibench—<https://github.com/intel-hadoop/HiBench>
- hibench.scale.profile = bigdata(32亿条数据)
- 在腾讯云大数据集群上执行Terasort任务，统计性能参数

腾讯云大数据产品现状

01

智能化运维的概念和价值

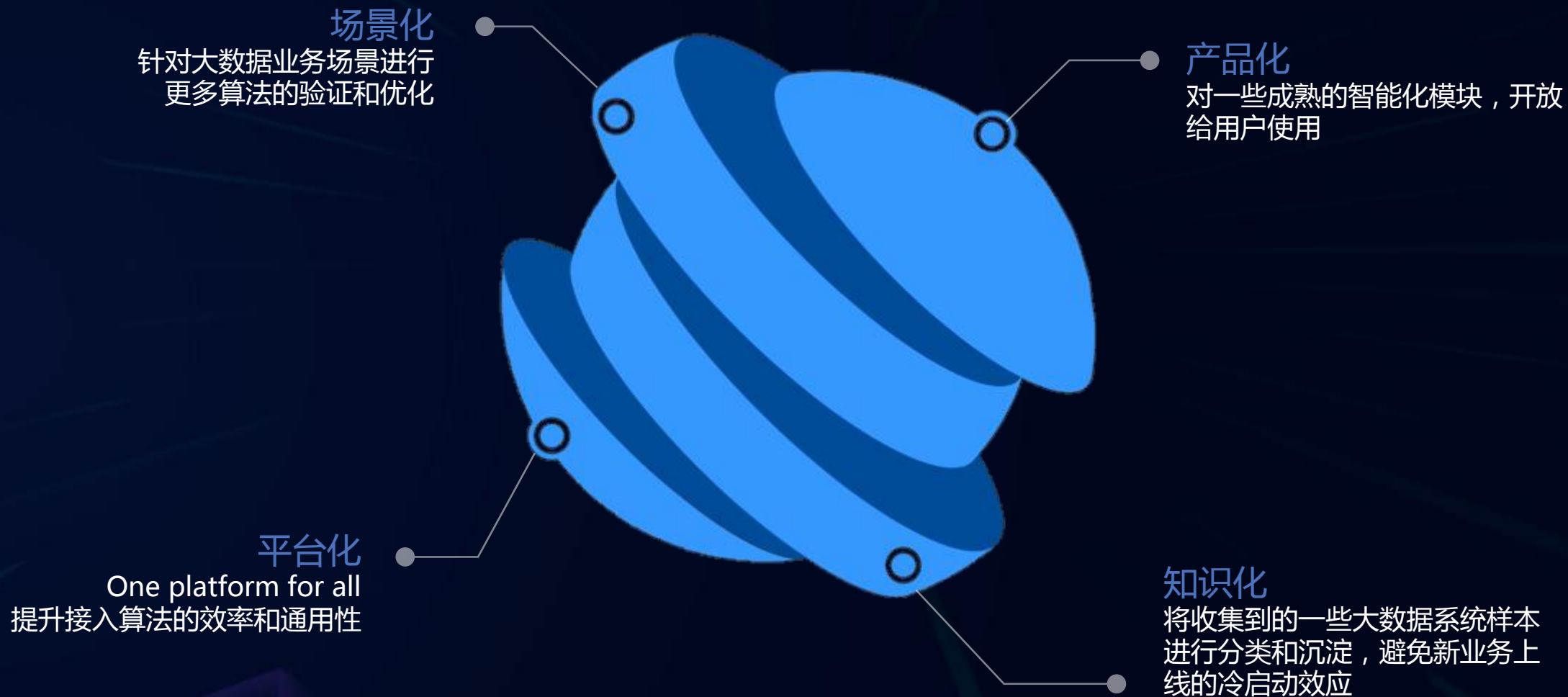
02

智能化大数据平台最佳实践

03

未来展望

04



腾讯云大数据产品现状

01

智能化运维的概念和价值

02

智能化大数据平台最佳实践

03

未来展望

04



THANKS

new trend
new technology
new application

Cloud + community
Developer conference