

Школа Java Middle Developer

Kafka

Мониторинг Kafka

Содержание

1. Показатели брокеров Kafka
2. Показатели топиков и партиций
3. Мониторинг клиентов

Основы показателей

Как получить доступ к показателям

- ✓ Ко всем показателям Kafka можно обращаться через JMX
- ✓ Чтобы упростить настройку приложений, подключающихся к JMX брокера Kafka напрямую, следует указать настроенный JMX-порт в настройках брокера
- ✓ Чтобы получить информацию о настройках, можно воспользоваться утилитой ZooKeeper CLI

Внешние и внутренние показатели

- Предоставляемые по интерфейсу JMX показатели — внутренние, они формируются и выдаются самим контролируемым приложением
- К внешним показателям можно отнести общее время выполнения запроса или доступность конкретного типа запроса, т.е. те, которые можно оценить извне приложения

Контроль состояния приложения

Необходимо также контролировать общее состояние процесса приложения с помощью простой проверки рабочего состояния. Сделать это можно двумя способами:

- с помощью внешнего процесса, который сообщает, работает брокер или отключен;
- посредством того, что брокер Kafka оповещает об отсутствии показателей.

Охват показателей

- Важно тщательно выбирать, на какие показатели обращать внимание (особенно при настройке уведомлений)
- Например, вы можете получать одно уведомление о масштабной проблеме, с тем чтобы сразу начать собирать дополнительные данные для выяснения конкретики

Показатели брокеров Kafka

Множество компаний используют Kafka для сбора показателей приложений, системных показателей и журналов логов для дальнейшей отправки в централизованную систему мониторинга. Чтобы не попасть в ловушку, необходимо:

- ❖ воспользоваться для Kafka отдельной, независимой от нее системой мониторинга
- ❖ при наличии нескольких ЦОД сделать так, чтобы показатели кластера Kafka в ЦОД А отправлялись в ЦОД Б и наоборот

Недореплицированные партиции

Показатель	Недореплицированные разделы
Управляемый компонент (MBean) JMX	kafka.server.type=ReplicaManager, name=UnderReplicatedPartitions
Диапазон значений	Целое число, равное нулю или больше него

- Если многие брокеры кластера сообщают о постоянном (не меняющемся) числе недореплицированных партиций, то обычно это значит, что один из брокеров кластера отключен
- Прежде чем пытаться выяснить причины проблемы, следует проверить, была ли выбрана предпочтительная реплика

Недореплицированные партиции

- Если число недореплицированных партиций меняется или оно постоянно, но отключенных брокеров нет, то дело обычно в проблеме с производительностью кластера: связана проблема с отдельным брокером или со всем кластером?
- Если недореплицированные партиции есть на нескольких брокерах, дело может быть в проблеме с кластером или с отдельным брокером
- Если мы увидим, что все недореплицированные партиции относятся к одному брокеру, то это указывает на то, что у него есть проблемы с репликацией сообщений, поэтому имеет смысл сосредоточиться на нем

Проблемы уровня кластера

Проблемы с кластером обычно относятся к одной из двух категорий:

- ✓ дисбаланс нагрузки;
- ✓ истощение ресурсов.

Для диагностики от брокеров кластера понадобятся данные по следующим показателям:

- количество партиций;
- количество ведущих партиций;
- суммарная по всем топикам входящая скорость передачи данных [байт/с];
- суммарная по всем топикам исходящая скорость передачи данных [байт/с];
- суммарная частота входящих сообщений по всем топикам.

Проблемы уровня кластера

В идеально сбалансированном кластере показатели будут примерно одинаковыми для всех брокеров кластера, как в таблице:

Брокер	Партиция	Ведущая реплика	Входящих байтов, Мбайт/с	Исходящих байтов, Мбайт/с
1	100	50	3,56	9,45
2	101	49	3,66	9,25
3	100	50	3,23	9,82

Проблемы уровня кластера

Распространенная проблема с производительностью кластера – превышение пределов возможностей брокеров по обслуживанию запросов. Для диагностики подобных проблем существует множество показателей, которые можно отслеживать на уровне операционной системы, в том числе:

- использование CPU;
- пропускная способность сети на вход;
- пропускная способность сети на выход;
- среднее время ожидания диска;
- процент использования диска.

Проблемы уровня хоста

Если проблемы возникают на одном или двух брокерах, то имеет смысл взглянуть на соответствующий сервер и разобраться, чем он отличается от остального кластера.

Подобные проблемы делятся на следующие общие категории:

- ✓ отказы аппаратного обеспечения;
- ✓ конфликты между процессами;
- ✓ различия локальных настроек.

Проблемы уровня хоста

- Сбои аппаратного обеспечения - вещь очевидная, при этом сервер просто перестает работать, а снижение производительности бывает вызвано менее очевидными проблемами
- Распространенный тип аппаратного сбоя, приводящий к снижению производительности Kafka - отказ диска. Отказ одного диска на одном-единственном брокере может свести на нет производительность всего кластера
- Передача данных по сети - еще одна сфера, в которой частичные сбои могут вызвать проблемы
- Если аппаратных проблем нет, то часто имеет смысл поискать работающее в той же системе другое приложение, которое потребляет ресурсы и затрудняет работу брокера Kafka

Показатели брокеров

Признак текущего контроллера

- Показатель «признак текущего контроллера» (active controller count) указывает, является ли данный брокер текущим контроллером кластера
- Если два брокера утверждают, что являются текущим контроллером кластера, то имеется проблема: поток выполнения контроллера не завершил работу как полагается, а завис

Показатель	Признак текущего контроллера
Управляемый компонент (MBean) JMX	kafka.controller:type=KafkaController, name=ActiveControllerCount
Диапазон значений	0 или 1

Коэффициент простоя обработчиков запросов

- Kafka использует два пула потоков выполнения для обработки всех запросов клиентов: сетевые потоки и обработчики запросов
- Нет необходимости задавать в настройках больше потоков, чем процессоров на брокере
- Показатель «коэффициент простоя обработчиков запросов» отражает долю времени (в процентах), в течение которого обработчики запросов

Показатель	Признак текущего контроллера
Управляемый компонент (MBean) JMX	kafka.server:type=KafkaRequestHandlerPool, name=RequestHandlerAvgIdlePercent
Диапазон значений	Число с плавающей запятой между 0 и 1 включительно

Суммарная входящая скорость передачи данных

- Суммарная по всем топикам входящая скорость передачи данных, выраженная в байтах в секунду, полезна в качестве показателя количества сообщений, получаемых брокерами от клиентов-производителей
- Не мешает вычислить этот показатель в случае, когда один из брокеров кластера получает больше трафика, чем другие

Показатель	Входящая скорость, байт/с
Управляемый компонент (MBean) JMX	kafka.server:type=BrokerTopicMetrics, name=BytesInPerSec
Диапазон значений	Скорость - число с двойной точностью, количество - целочисленное значение

Суммарная входящая скорость передачи данных

Атрибуты подобных показателей. Первые два атрибута не показатели, но они полезны для понимания:

- EventType — единица измерения для всех атрибутов, в данном случае байты;
- RateUnit — для атрибутов скорости/частоты представляет собой период времени, за которое рассчитывается скорость, в данном случае секунды.

Атрибуты скорости с различным шагом детализации:

- OneMinuteRate — среднее значение за предыдущую минуту;
- FiveMinuteRate — среднее значение за предыдущие 5 минут;
- FifteenMinuteRate — среднее значение за предыдущие 15 минут;
- MeanRate — среднее значение за всё время, прошедшее с момента запуска брокера.

Суммарная исходящая скорость передачи данных

- Суммарная исходящая скорость передачи данных - обобщенный показатель масштабирования, отражает исходящую скорость чтения потребителями данных
- Исходящая скорость включает и трафик реплик. Это значит, что если коэффициент репликации всех топиков в настройках равен 2, то при отсутствии клиентов-потребителей исходящая скорость передачи данных будет равна входящей. При чтении всех сообщений кластера одним клиентом-потребителем исходящая скорость будет вдвое превышать входящую

Показатель	Исходящая скорость, байт/с
Управляемый компонент (MBean) JMX	kafka.server:type=BrokerTopicMetrics, name=BytesOutPerSec
Диапазон значений	Скорость - число с двойной точностью, количество - целочисленное значение

Суммарное по всем топикам число входящих сообщений

Скорости передачи данных отражают трафик брокера в абсолютных показателях - в байтах, в то время как показатель входящих сообщений отражает количество отдельных сгенерированных в секунду входящих сообщений вне зависимости от их размера

Показатель	Сообщений в секунду
Управляемый компонент (MBean) JMX	kafka.server:type=BrokerTopicMetrics, name=MessagesInPerSec
Диапазон значений	Скорость - число с двойной точностью, количество - целочисленное значение

Количество партиций

- Показатель количества партиций для брокера обычно незначительно меняется с течением времени, ведь он представляет собой общее количество партиций, назначенных данному брокеру

Показатель	Количество партиций
Управляемый компонент (MBean) JMX	kafka.server:type=ReplicaManager, name=PartitionCount
Диапазон значений	Целое число, равное нулю или больше него

Количество ведущих реплик

- Количество ведущих реплик отражает количество партиций, для которых данный брокер в настоящий момент является ведущей репликой
- Отражает несбалансированность кластера даже в случае идеального баланса реплик по числу и размеру

Показатель	Количество ведущих реплик
Управляемый компонент (MBean) JMX	kafka.server:type=ReplicaManager, name=LeaderCount
Диапазон значений	Целое число, равное нулю или больше него

Отключенные партиции

- Критически важен мониторинг количества отключенных партиций
- Без ведущей реплики партиция может остаться по двум причинам.
 - Останов всех брокеров, на которых находятся реплики данной партиции;
 - Ни одна согласованная реплика не может стать ведущей из-за расхождения числа сообщений (в случае, когда отключен «нечистый» выбор ведущей реплики).
- При промышленной эксплуатации кластера Kafka отключенные партиции могут влиять на клиенты-производители, приводя к потере сообщений или отрицательному обратному влиянию в приложении

Показатель	Число отключенных партиций
Управляемый компонент (MBean) JMX	kafka.controller:type=KafkaController, name=OfflinePartitionCount
Диапазон значений	Целое число, равное нулю или больше него

Показатели запросов

С помощью определенных показателей можно контролировать функционирование следующих типов запросов:

- | | | |
|-----------------------|-----------------|-------------------|
| ❖ ApiVersions; | ❖ Heartbeat; | ❖ Offsets; |
| ❖ ControlledShutdown; | ❖ JoinGroup; | ❖ Produce; |
| ❖ CreateTopics; | ❖ LeaderAndIsr; | ❖ SasHandshake; |
| ❖ DeleteTopics; | ❖ LeaveGroup; | ❖ StopReplica; |
| ❖ Fetch; | ❖ ListGroups; | ❖ SyncGroup; |
| ❖ Fetchconsumer | ❖ Metadata; | ❖ UpdateMetadata. |
| ❖ FetchFollower; | ❖ OffsetCommit; | |
| ❖ GroupCoordinator; | ❖ OffsetFetch; | |

Показатели запросов

Название	Управляемый компонент (MBean) JMX
Общее время	kafka.network:type=RequestMetrics,name=TotalTimeMs, request=Fetch
Время нахождения запроса в очереди	kafka.network:type=RequestMetrics,name=RequestQueueTimeMs, request=Fetch
Локальное время	kafka.network:type=RequestMetrics,name=LocalTimeMs, request=Fetch
Удаленное время	kafka.network:type=RequestMetrics,name=RemoteTimeMs, request=Fetch
Длительность притормаживания	kafka.network:type=RequestMetrics,name=ThrottleTimeMs, request=Fetch
Время нахождения ответа в очереди	kafka.network:type=RequestMetrics,name=ResponseQueueTimeMs, request=Fetch
Длительность отправки запроса	kafka.network:type=RequestMetrics,name=ResponseSendTimeMs, request=Fetch
Запросов в секунду	kafka.network:type=RequestMetrics,name=RequestsPerSec, request=Fetch

Показатели запросов

Показатели подсчитываются с момента запуска брокера и отражают следующие этапы обработки запроса:

- Общее время
- Время нахождения запроса в очереди
- Локальное время
- Удаленное
- Длительность притормаживания
- Время нахождения ответа в очереди
- Длительность отправки запроса

Показатели запросов

У всех показателей имеются следующие атрибуты:

- процентиля - 50thPercentile, 75thPercentile, 95thPercentile, 98thPercentile, 99thPercentile, 999thPercentile;
- Count – фактическое количество запросов с момента запуска процесса;
- Min – минимальное значение по всем запросам;
- Max – максимальное значение по всем запросам;
- Mean – среднее значение по всем запросам;
- StdDev – стандартное отклонение показателей хронометража запросов в совокупности.

Показатели топиков и партиций

- ❖ Удобны при отладке конкретных проблем с клиентами
- ❖ Показатели уровня топиков можно использовать для поиска конкретного топика, который вызывает большой прирост объема трафика кластера

Название	Управляемый компонент (MBean) JMX
Скорость входящей передачи данных, байт/с	kafka.server:type=BrokerTopicMetrics,name=BytesInPerSec,topic=TOPICNAME
Скорость исходящей передачи данных, байт/с	kafka.server:type=BrokerTopicMetrics,name=BytesOutPerSec,topic=TOPICNAME
Частота неудачного извлечения данных	kafka.server:type=BrokerTopicMetrics,name=FailedFetchRequestsPerSec,topic=TOPICNAME
Частота неудачной генерации сообщений	kafka.server:type=BrokerTopicMetrics,name=FailedProduceRequestsPerSec,topic=TOPICNAME
Частота входящих сообщений	kafka.server:type=BrokerTopicMetrics,name=MessagesInPerSec,topic=TOPICNAME
Частота запросов на извлечение данных	kafka.server:type=BrokerTopicMetrics,name=TotalFetchRequestsPerSec,topic=TOPICNAME
Частота запросов от производителей	kafka.server:type=BrokerTopicMetrics,name=TotalProduceRequestsPerSec,topic=TOPICNAME

Показатели уровня топика

Показатели уровня партиции

- ❖ Отражают объем данных (в байтах), хранящийся в настоящий момент на диске для определенной партиции
- ❖ Если их суммировать, можно узнать объем хранимых данных отдельного топика

Название	Управляемый компонент (MBean) JMX
Размер раздела	kafka.log:type=Log,name=Size,topic=TOPICNAME,partition=0
Количество сегментов журнала	kafka.log:type=Log,name=NumLogSegments,topic=TOPICNAME,partition=0
Начальное смещение журнала	kafka.log:type=Log,name=LogEndOffset,topic=TOPICNAME,partition=0
Конечное смещение журнала	kafka.log:type=Log,name=LogStartOffset,topic=TOPICNAME,partition=0

Мониторинг JVM

Помимо показателей для брокера Kafka следует контролировать стандартный набор показателей для всех серверов и самой виртуальной машины Java JVM. Благодаря им вы также сможете понять причину изменения показателей далее по конвейеру, в брокере.

- Сборка мусора
- Мониторинг операционной системы из Java
- Мониторинг ОС

Мониторинг JVM

Следите по крайней мере за статистикой дисков, на которых хранятся данные Kafka:

- число операций записи и чтения в секунду
- средние размеры очередей на чтение и запись
- среднее время ожидания
- эффективность использования диска в процентах

Журналирование

Существует два механизма журналирования, записывающих информацию в отдельные файлы на диске:

- `kafka.controller` на уровне INFO, он служит для получения информации конкретно о контроллере кластера
- `kafka.server.ClientQuotaManager`, тоже уровня INFO. Он используется для отображения сообщений, связанных с квотами на операции генерации и потребления

Журналирование

Не помешает также занести в журнал информацию о состоянии потоков сжатия журналов. Не существует отдельного показателя, отражающего состояние этих потоков, и сбой сжатия одной партии может полностью застопорить потоки сжатия журналов, причем пользователь не получит никакого оповещения об этом. Для вывода информации о состоянии этих потоков необходимо включить механизмы журналирования `kafka.log.LogCleaner`, `kafka.log.Cleaner` и `kafka.log.LogCleanerManager` на уровне `DEBUG`

Журналирование

Будет полезно включить еще некоторые виды журналирования при отладке проблем с Kafka. Например, `kafka.request.logger`, на уровне DEBUG или TRACE. Он заносит в журнал информацию обо всех отправленных брокеру запросах. На уровне DEBUG данный журнал включает конечные точки соединений, хронометраж запросов и сводную информацию. На уровне TRACE - также информацию о топике и партиции - практически всю информацию запроса, за исключением содержимого самого сообщения

Мониторинг клиентов

Показатели производителя

Новый клиент-производитель Kafka существенно повысил компактность имеющихся показателей, сделав их доступными в виде атрибутов небольшого числа управляемых компонентов (mbeans)

Название	Управляемый компонент (MBean) JMX
В целом по производителю	kafka.producer:type=producer-metrics, client-id=CLIENTID
Для отдельного брокера	kafka.producer:type=producer-node-metrics, client-id=CLIENTID, node-id=node-BROKERID
Для отдельного топика	kafka.producer:type=producer-topic-metrics, client-id=CLIENTID, topic=TOPICNAME

Общие показатели производителя

- ❖ Имеет смысл настроить оповещение для атрибута `record-error-rate`. Этот показатель всегда должен быть равен 0, и если он > 0 , значит, производитель отменяет сообщения, которые пытается отправить брокерам Kafka
- ❖ Можно отслеживать `record-retry-rate`, но он не так важен, как частота ошибок, поскольку повторы отправки свидетельствуют о нормальном функционировании
- ❖ Атрибут `outgoing-byte-rate` говорит о трафике сообщений в байтах в секунду
- ❖ `record-send-rate` описывает трафик в терминах числа сгенерированных сообщений в секунду
- ❖ `request-rate` позволяет узнать число отправленных от производителей брокерам запросов в секунду

Общие показатели производителя

- ❖ С помощью `request-size-avg` можно получить средний размер запросов, отправляемых брокерам производителями, в байтах
- ❖ С помощью `batch-size-avg` - средний размер отдельного пакета сообщений, состоящего по умолчанию из сообщений, предназначенных для отдельной партии топика, в байтах
- ❖ `record-size-avg` показывает средний размер отдельной записи в байтах
- ❖ Существует `records-per-request-avg`, описывающий среднее число сообщений в отдельном запросе от производителя
- ❖ `record-queue-time-avg` – представляет собой среднее время, которое отдельному сообщению приходится ожидать в производителе после отправки его приложением и до фактической генерации его для Kafka

Показатели уровня брокера и топика

- ❖ Наиболее полезный из показателей производителей, относящихся к отдельным брокерам, - `request-latency-avg`. Значение этого показателя практически всегда постоянно (при стабильной работе пакетной отправки сообщений), параметр может отражать проблемы с подключением к конкретным брокерам
- ❖ `outgoing-byte-rate` и `request-latency-avg`, меняются в зависимости от партиций, для которых данный брокер является ведущим
- ❖ Атрибуты `record-send-rate` и `record-error-rate` можно использовать для выяснения того, к какому топика относятся отмененные сообщения

Показатели потребителей

- ❖ Клиент-потребитель в Kafka объединяет множество показателей в атрибуты всего лишь нескольких компонентов показателей. Из этих показателей исключены процентные показатели для задержки и скользящие средние скорости/частоты

Название	Управляемый компонент (MBean) JMX
В целом по потребителю	kafka.consumer:type=consumer-metrics, client-id=CLIENTID
Диспетчер извлечения	kafka.consumer.type=consumer-fetch-manager-metrics, client-id=CLIENTID
Для отдельного топика	kafka.consumer:type=consumer-fetch-manager-metrics, client-id=CLIENTID
Для отдельного брокера	kafka.consumer:type=consumer-node-metrics, client-id=CLIENTID, node-id=node-BROKERID
Координатор	kafka.consumer:type=consumer-coordinator-metrics, client-id=CLIENTID

Показатели диспетчера извлечения

- ❖ `fetch-latency-avg` – с его помощью можно выяснить, сколько времени занимает выполнение запросов на извлечение к брокерам
- ❖ Параметрами `fetch.min.bytes` и `fetch.max.wait.ms` конфигурации потребителя определяется длительность задержки
- ❖ Показатели `bytes-consumed-rate` или `records-consumed-rate` помогут узнать объемы обрабатываемого клиентом-потребителем трафика сообщений
- ❖ `fetch-rate` сообщает число выполняемых потребителем запросов на извлечение в секунду
- ❖ `fetch-size-avg` - средний размер этих запросов на извлечение в байтах
- ❖ Показатель `records-per-request-avg` дает среднее число сообщений в каждом запросе на извлечение

Спасибо за внимание