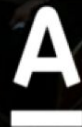


21 июня 2018, 19:00

Backend stories MeetUp



Альфа·Банк

Павел Дерендяев



dddpaul



Как жить, если вы неправильно готовите Elasticsearch



План

- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо
- 7) Выводы



- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо
- 7) Выводы



Что нужно от Выписки?

- 1) Полнотекстовый поиск по транзакциям
- 2) Отображение списка транзакций
- 3) Печать транзакций в PDF, XLS и т.д.
- 4) Доступ к транзакциям за последние 5 лет



Фильтры и поиск

A Альфа-Банк

Главная

Новый платёж

Платежи в работе

Выписка

Сервисы ▾

Письма

ВЭД ▾

Благотворительность

Приведи друга

Клуб клиентов

Предыдущая версия

Выписка

Счета **Выписка** Обороты Незавершённые операции Аналитика

Счёт
9 916 559 695,15 ₽
40702 810 5 0000 0012881

Фильтры

Сегодня Вчера Неделя Месяц 01.09.2015 — 27.09.2017

Текущая неделя Текущий месяц

Сумма ₽ — ₽

Тип операций
Все Приход Расход

Контрагент
Введите название контрагента Введите номер карты

Баланс за выбранный период

Приход	Расход	Остаток на конец периода
+ 34 044 ₽	- 3 612,36 ₽	9 916 559 695,15 ₽

Полнотекстовый
поиск и
автодополнение



Список транзакций



A Альфа-Банк

- Главная
- Новый платёж
- Платежи в работе
- Выписка
- Сервисы
- Письма
- ВЭД
- Благотворительность

- Приведи друга
- Клуб клиентов
- Предыдущая версия

20 сент. 2017

АО "Альфа-Банк" – 5 900 ₽

Комиссия за обслуживание р/с за период с 20АВГ17 по 19СЕН17 С

[Подробнее](#)

19 сент. 2017

АО "Альфа-Банк" – 3 590 ₽

Комиссия за обслуживание р/с за период с 19АВГ17 по 18СЕН17 С

[Подробнее](#)

17 сент. 2017

АО "Альфа-Банк" – 4 002 ₽

Комиссия за обслуживание р/с за период с 17АВГ17 по 16СЕН17 С

[Подробнее](#)

16 сент. 2017

АО "Альфа-Банк" – 4 002 ₽

Комиссия за обслуживание р/с за период с 16АВГ17 по 15СЕН17 С



Печать транзакции

A Альфа-Банк

- Главная
- Новый платёж
- Платежи в работе
- Выписка
- Сервисы
- Письма
- ВЭД
- Благотворительность

Приведи друга
Клуб клиентов
Предыдущая версия

20 сент. 2017

АО "Альфа-Банк"

– 5 900 ₽

Комиссия за обслуживание р/с за период с 20АВГ17 по 19СЕН17
Согласно тарифам Банка ООО "Не трогать"

№ 3

Платательщик

Получатель

Название:



Название: АО "Альфа-Банк"

ИНН: 5050058510

ИНН: 7728168971

КПП: 123123123

КПП: 770801001

Счёт: 40702 810 4 0110 0000999

Счёт: 47423 810 7 0110 0000528

Банк: АО "АЛЬФА-БАНК"

Банк: АО "АЛЬФА-БАНК"

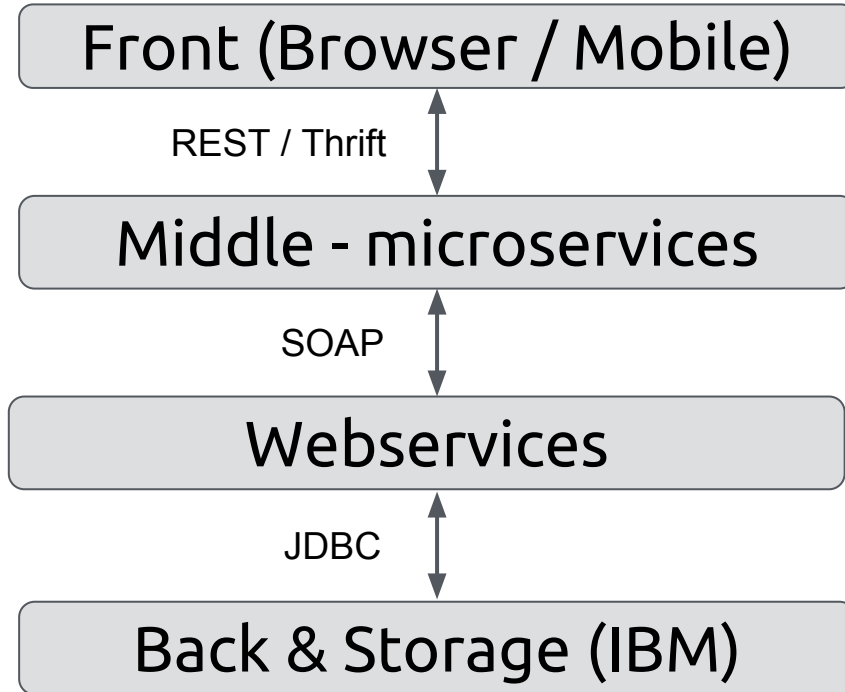
БИК: 044525593

БИК: 044525593

Скачать | PDF

Скрыть

Типовое приложение в банке





- 1) Приложение Выписка
- 2) **Зачем нужен Elasticsearch?**
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо
- 7) Выводы

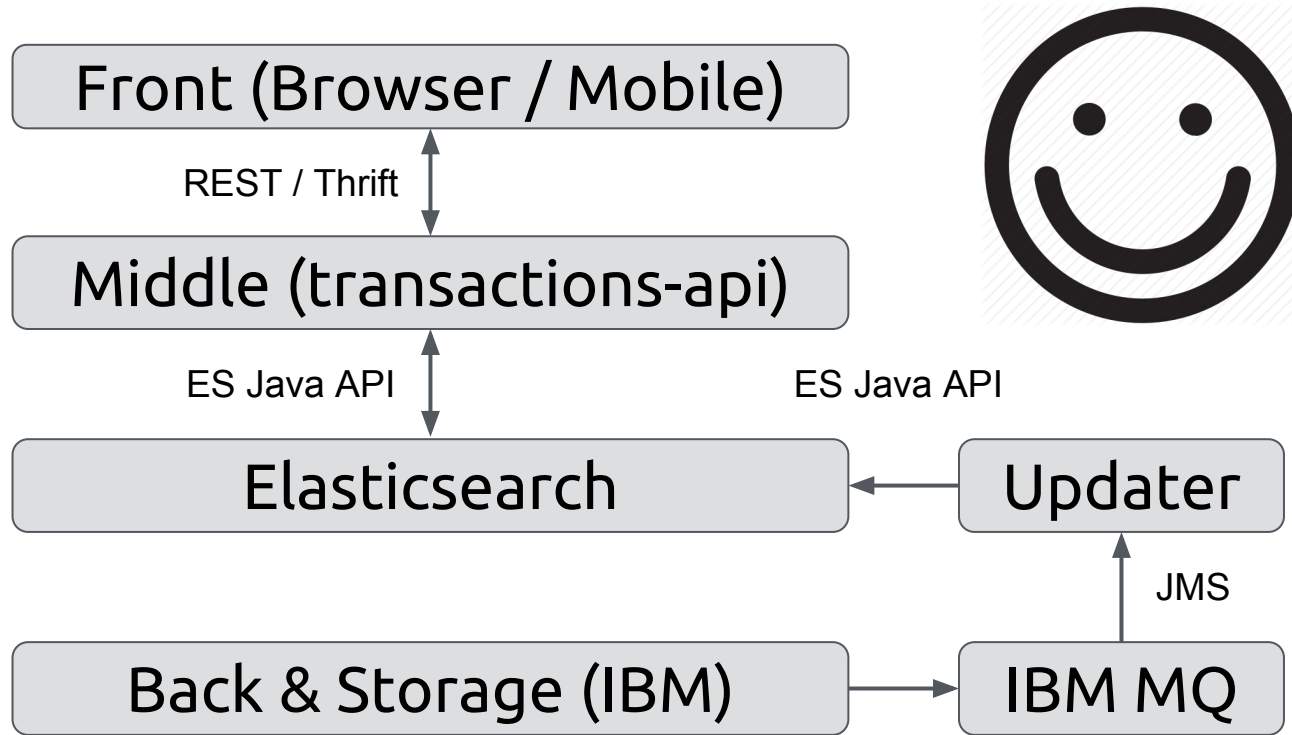


elasticsearch



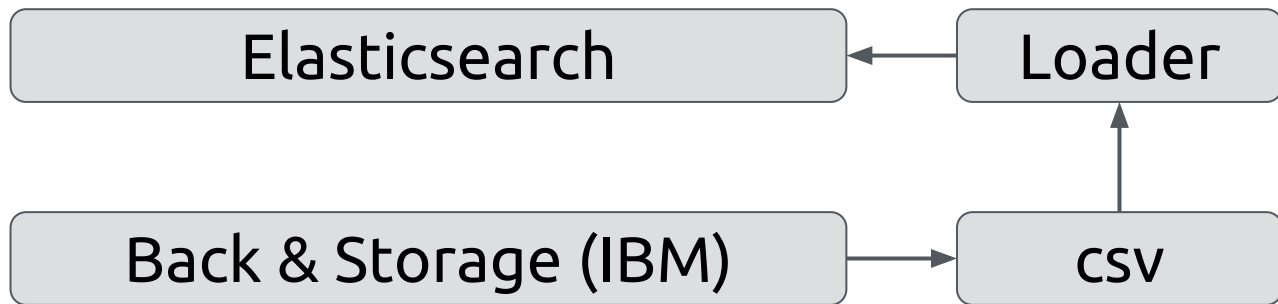
- + Есть встроенный полнотекстовый поиск
- + Хорошо масштабируется
- + Надежность (по слухам)
- + Уже использовали для хранения логов

Промежуточный storage



Первоначальная заливка

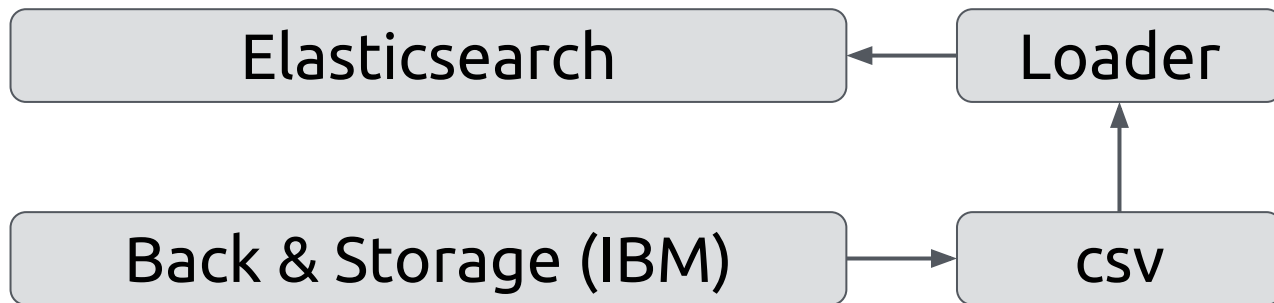
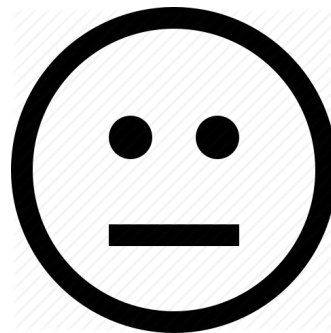
Нельзя протащить через очередь все транзакции с начала времен



Первоначальная заливка

Too slow ...

Данные за месяц
выгружаются 1 день





- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды**
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо
- 7) Выводы

*SQL и Elasticsearch



*SQL	Elasticsearch
Databases	Indices
Tables	Mapping types
Rows	Documents
Columns	Properties



Mapping type

```
"transaction" : {  
  "properties" : {  
    "direction" : { "type" : "boolean" },  
    "doc" : {  
      "type": "object",  
      "properties" : {  
        "accountNumber" : { "type" : "keyword" },  
        "currency" : { "type" : "keyword" },  
        ...  
      }  
    }  
  }  
}
```



Mapping type

```
"transaction" : {  
  "properties" : {  
    "direction" : { "type" : "boolean" },  
    "doc" : {  
      "type": "object",  
      "properties" : {  
        "accountNumber" : { "type" : "keyword" },  
        "currency" : { "type" : "keyword" },  
        ...  
      }  
    }  
  }  
}
```

Mapping type

```
"transaction" : {  
  "properties" : {  
    "direction" : { "type" : "boolean" },  
    "doc" : {  
      "type": "object",  
      "properties" : {  
        "accountNumber" : { "type" : "keyword" },  
        "currency" : { "type" : "keyword" },  
        ...  
      }  
    }  
  }  
}
```

Разбивка по индексам

- 1) Один индекс на весь период - **listing**
- 2) Разбивка по годам, месяцам или дням:
 - listing-2017-12-**01**
 - listing-2017-12-**02**
 - ...
- 3) Выбрали **yyyy-mm-dd**

Про шарды

- 1) Elasticsearch index = collection of shards
- 2) Elasticsearch shard = Lucene index

Term	Doc 1	Doc 2	Doc 3	...

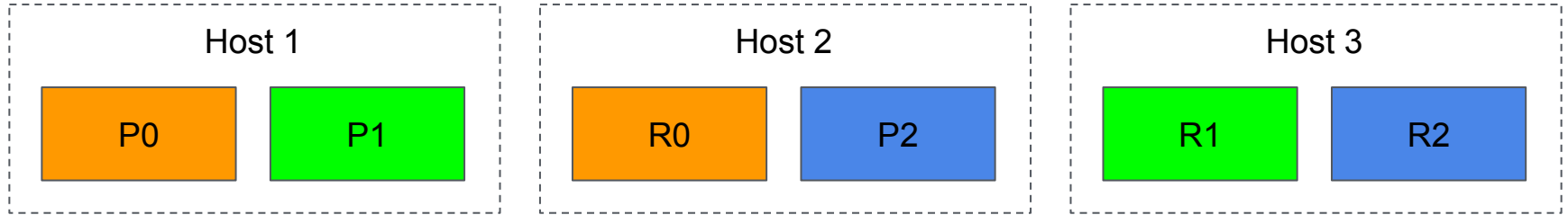
brown	X		X	...
fox	X	X	X	...
quick	X	X		...
the	X		X	...

7 шардов на индекс

- 1) 7 - по количеству нод в кластере
- 2) Две копии шарда (primary + replica)
- 3) Keep the number of shards per node below 20 to 25 per GB heap (20-25 шардов на 1 Гбайт хипа)

<https://www.elastic.co/blog/how-many-shards-should-i-have-in-my-elasticsearch-cluster>

Аллокация шардов



Пример для 3 primary + 3 replica



- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы**
- 5) Как побороли
- 6) Демо
- 7) Выводы



10к клиентов

- 1) 7 нод в кластере
- 2) Elasticsearch 1.5
- 3) 32 Гбайт на ноду (16 heap + 16 native)
- 4) 1 день выписки по дефолту

50к клиентов



- 1) 9 нод в кластере
- 2) Оптимизировали mapping type



200к клиентов

- 1) 12 нод в кластере
- 2) Elasticsearch 5.2.2
- 3) 70 Гбайт на ноду
- 4) Дефолтный период подняли до 3 месяцев
- 5) Начались звоночки

300к клиентов



- 1) Поняли, что так жить нельзя
- 2) Железо уже не помогает

Симптомы деградации



- 1) Выписка загружается слишком долго
- 2) transactions-api залипают и рестартуют
- 3) Ноды Эластика пожирают CPU
- 4) Кластер желтеет



Симптомы деградации

- 1) Выписка отдается слишком долго
- 2) transactions-api залипают и рестартуют
- 3) Ноды Эластика пожирают CPU
- 4) Кластер желтеет



Симптомы деградации

- 1) Выписка отдается слишком долго
- 2) transactions-api залипают и рестартуют
- 3) **Ноды Эластика пожирают CPU**
- 4) Кластер желтеет



Симптомы деградации

- 1) Выписка отдается слишком долго
- 2) transactions-api залипают и рестартуют
- 3) Ноды Эластика пожирают CPU
- 4) Кластер желтеет

Простейший алертинг :)



```
watch -n1 \  
'curl -sS host:9200/_cluster/health | jq .status | \  
{ read st; [ "$st" != "\"green\"" ] && echo "Alarm!!!"; }'
```

Логи приложения



```
Rejected execution of SearchDfsQueryThenFetchAsyncAction ON  
EsThreadPoolExecutor[search, queue capacity = 1000,  
EsThreadPoolExecutor[Running, pool size = 49, active  
threads = 49, queued tasks = 1000, completed tasks =  
80857358]]
```



А на ноде Эластика GC ...

```
[o.e.m.j.JvmGcMonitorService] [corpapp15]  
[gc][667721] overhead, spent [56.7s]  
collecting in the last [56.9s]
```

Длинный GC, CPU в полку, нода тупит ...

... and timeouts



```
[o.e.t.TransportService] [corpapp15] Received  
response for a request that has timed out,  
sent [34578ms] ago, timed out [12ms] ago
```

Срабатывает fault detection на дефолтном ping_timeout = 30s



Рибалансировка шардов

- 1) Нода вылетает из кластера
- 2) Кластер в YELLOW state
- 3) Начинается ребалансировка шардов
- 4) Возрастает disk IO и GC на других нодах
- 5) Эффект лавины

Эластик здорового человека



```
$ vmstat 1
```

```
procs      -----memory-----  ---swap--  -----io-----
 r  b    swpd      free    buff    cache    si    so    bi
 2  0      0 16133580  6264 64689804    0    0    7
 1  0      0 16133932  6264 64689900    0    0    0
 2  0      0 16134460  6264 64689872    0    0    0
 1  0      0 16135192  6264 64689872    0    0    0
```

Эластик курильщика



```
$ vmstat 1
```

```
procs -----memory----- ---swap-- -----io-----
r  b   swpd   free   buff   cache   si   so   bi
2 24 7263652 7068424   876 16363044   2   2  132
1 21 7263648 7066076   876 16372148   0   0 143388
6 50 7263636 7028148   876 16371732   0   0 130900
1 43 7263636 7034704   876 16371768   0   0 141360
```



- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли**
- 6) Демо
- 7) Выводы



Что делать?

- 1) Вынести master-ноды отдельно
- 2) Тюнинг ребалансировки
- 3) Роутинг (шардирование)
- 4) Помесячные индексы
- 5) Уменьшить кол-во шардов
- 6) Оптимизация поисковых запросов



Что сделали

- 1) Вынести master-ноды отдельно
- 2) Тюнинг ребалансировки
- 3) Роутинг (шардирование)
- 4) Помесячные индексы
- 5) Уменьшить кол-во шардов
- 6) Оптимизация поисковых запросов

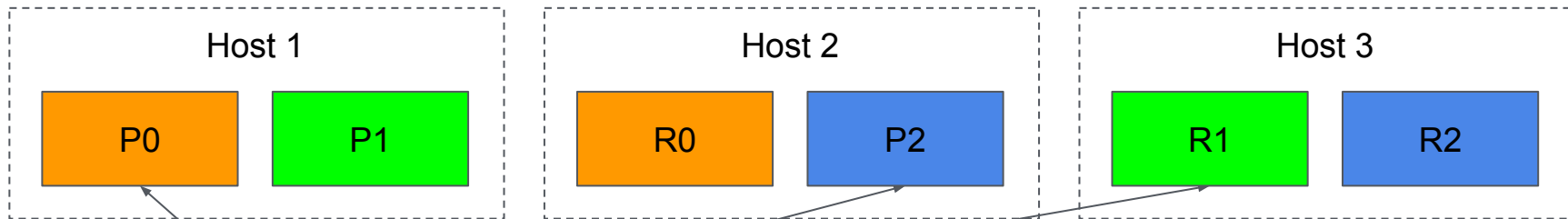


Роутинг (шардирование)

```
shard_num = hash(_routing) % num_primary_shards
```

- 1) Default routing value is “_id”
- 2) Роутинг по ID транзакции
- 3) Кажется, что это неоптимально

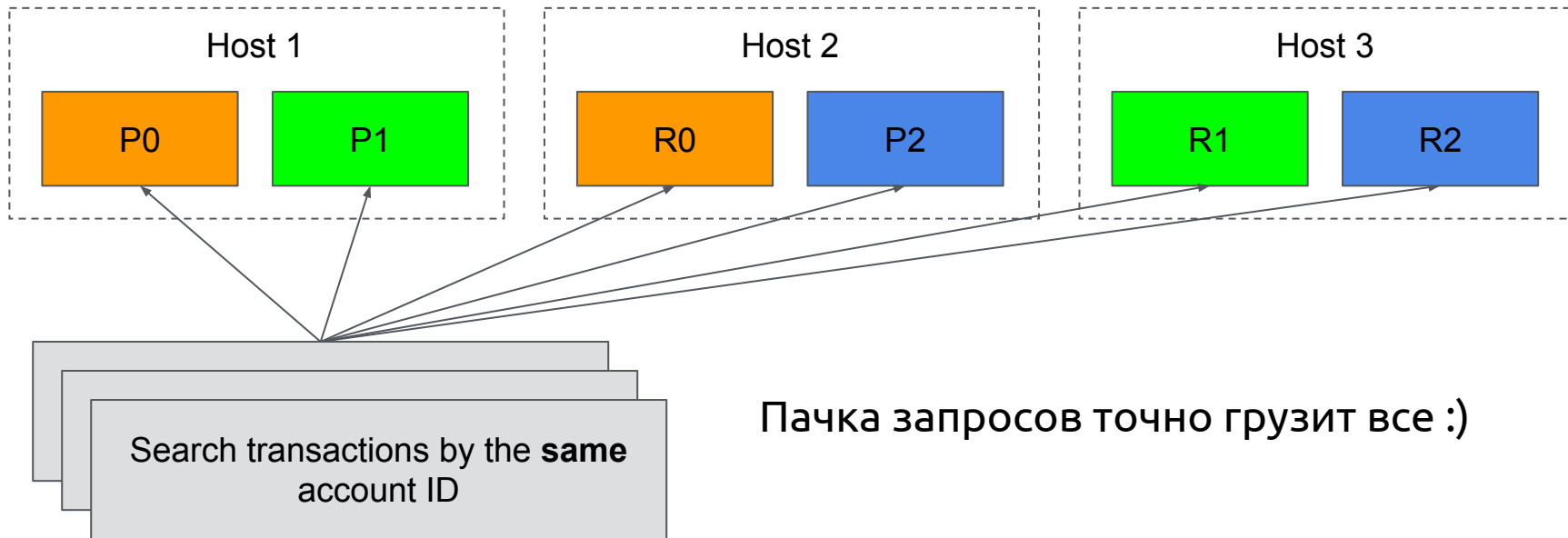
Поиск с роутингом по `_id`



Search transactions by account ID

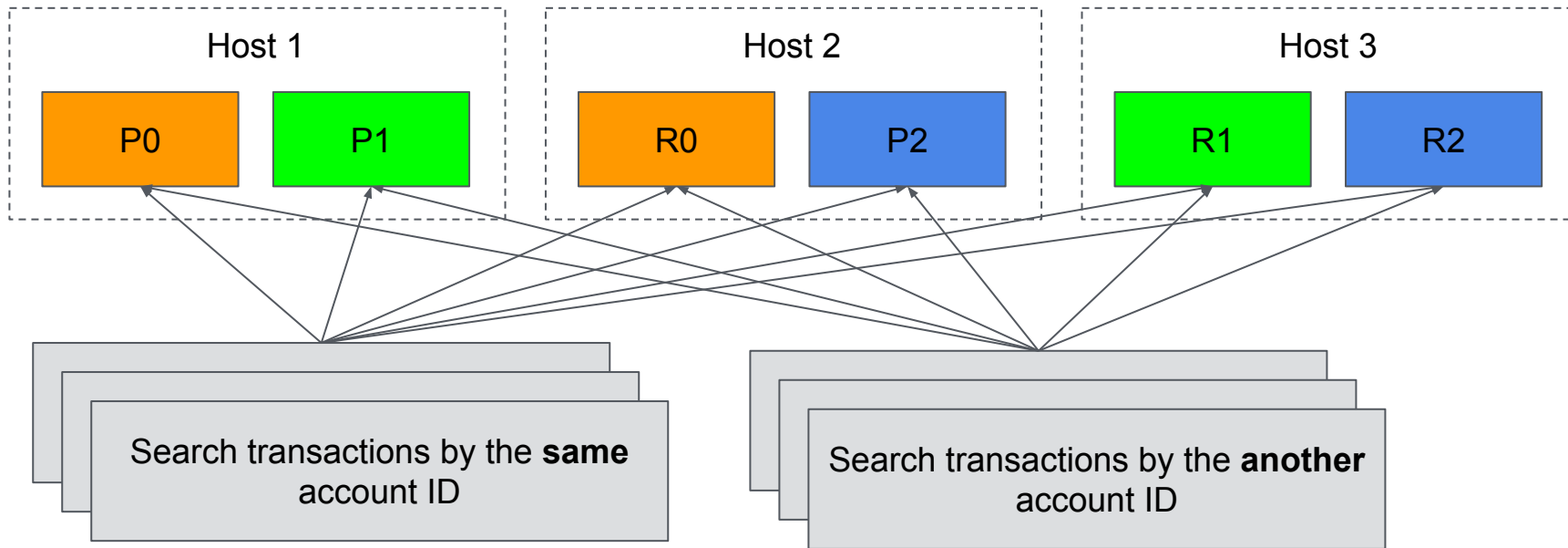
- 1) Транзакции даже одного счета лежат в разных шардах (и хостах)
- 2) Запрос выписки по одному счету нагружает все шарды индекса всех хостов

Поиск с роутингом по id

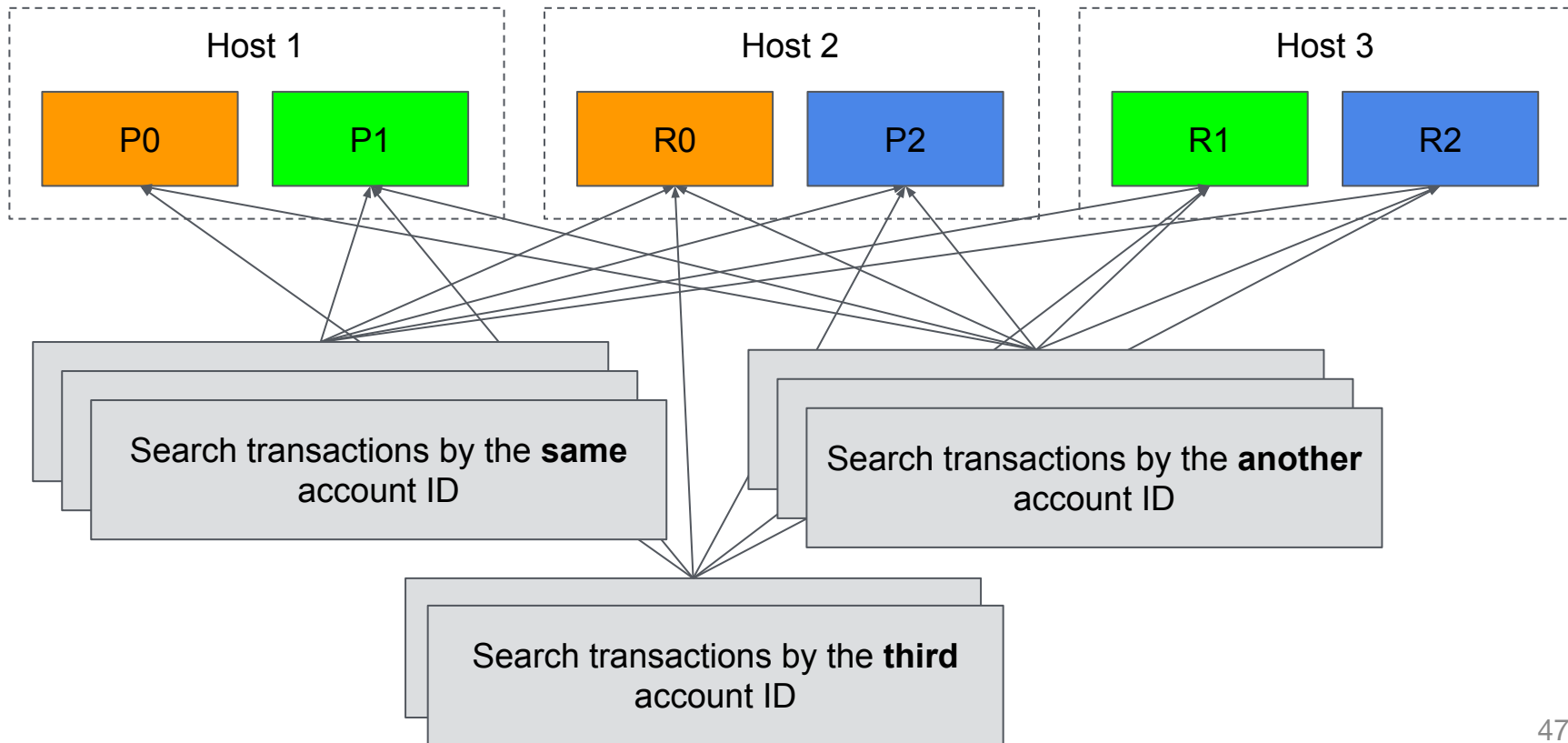


Пачка запросов точно грузит все :)

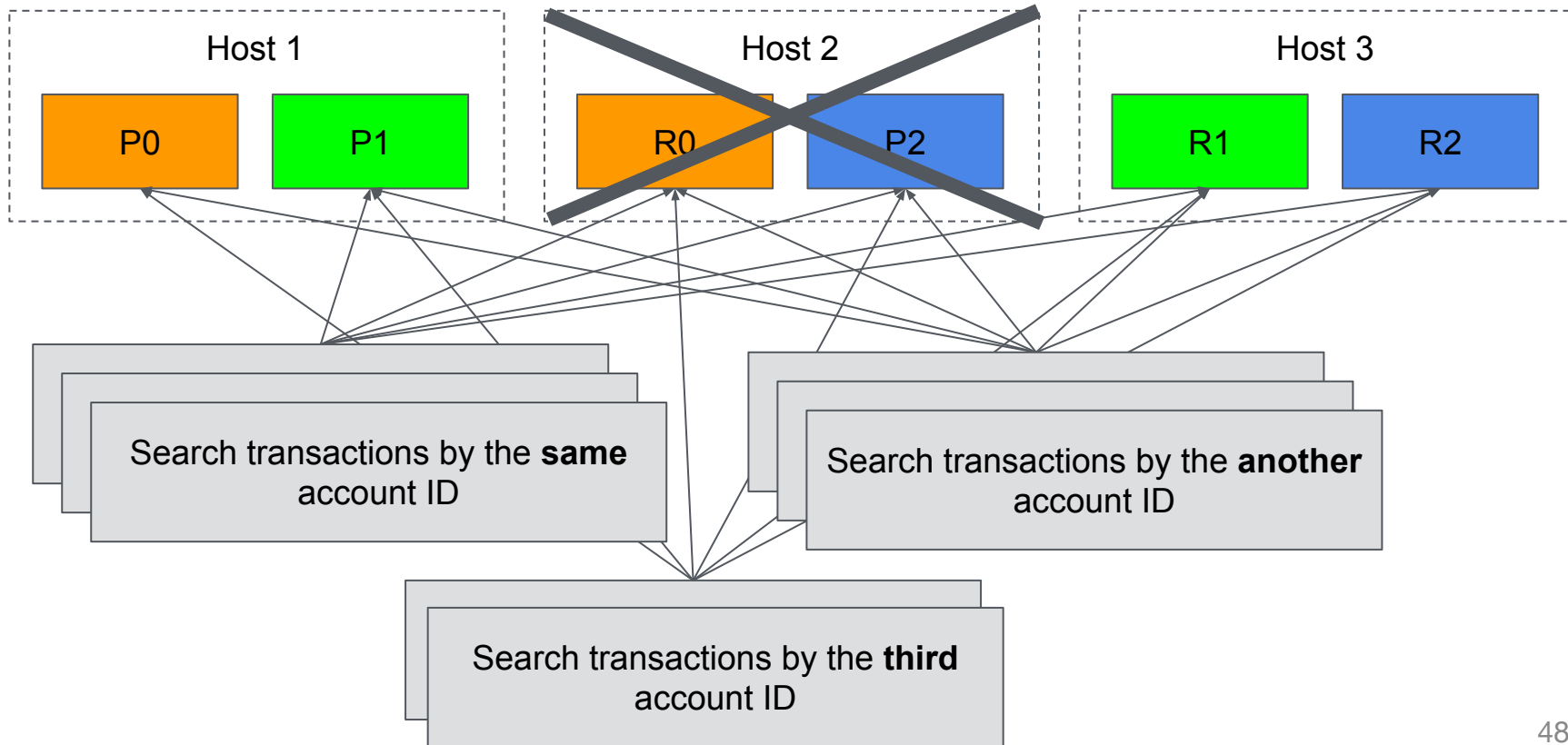
Поиск с роутингом по id



Поиск с роутингом по id

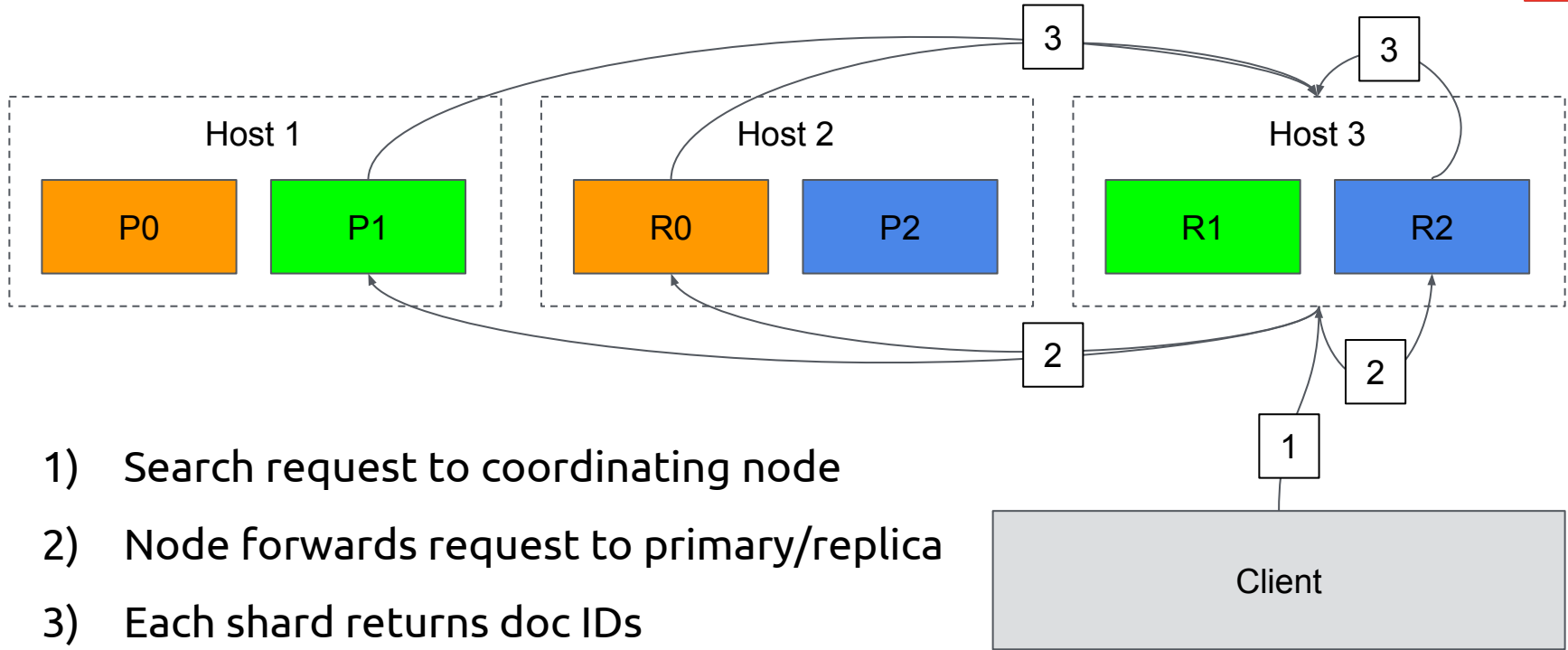


Поиск с роутингом по id



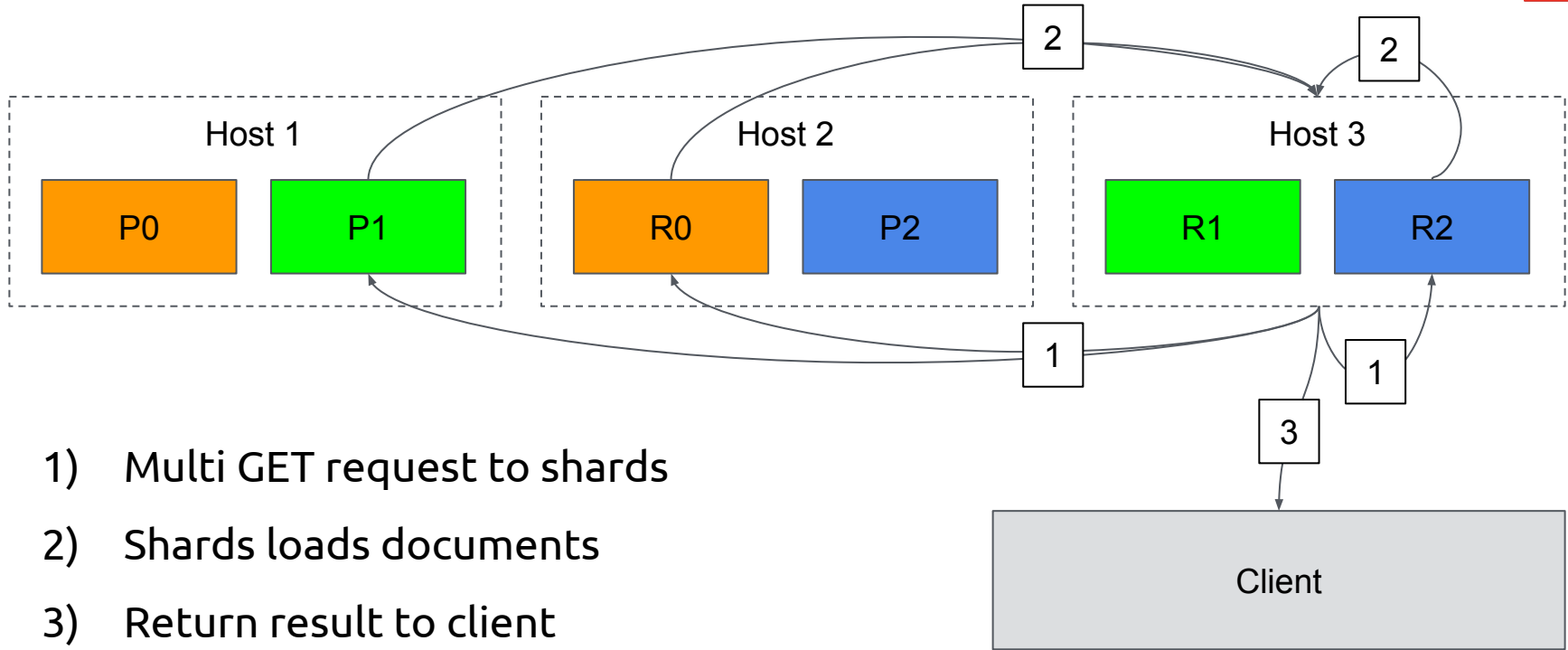
Query phase

A



- 1) Search request to coordinating node
- 2) Node forwards request to primary/replica
- 3) Each shard returns doc IDs

Fetch phase



Итого с дефолтным роутингом



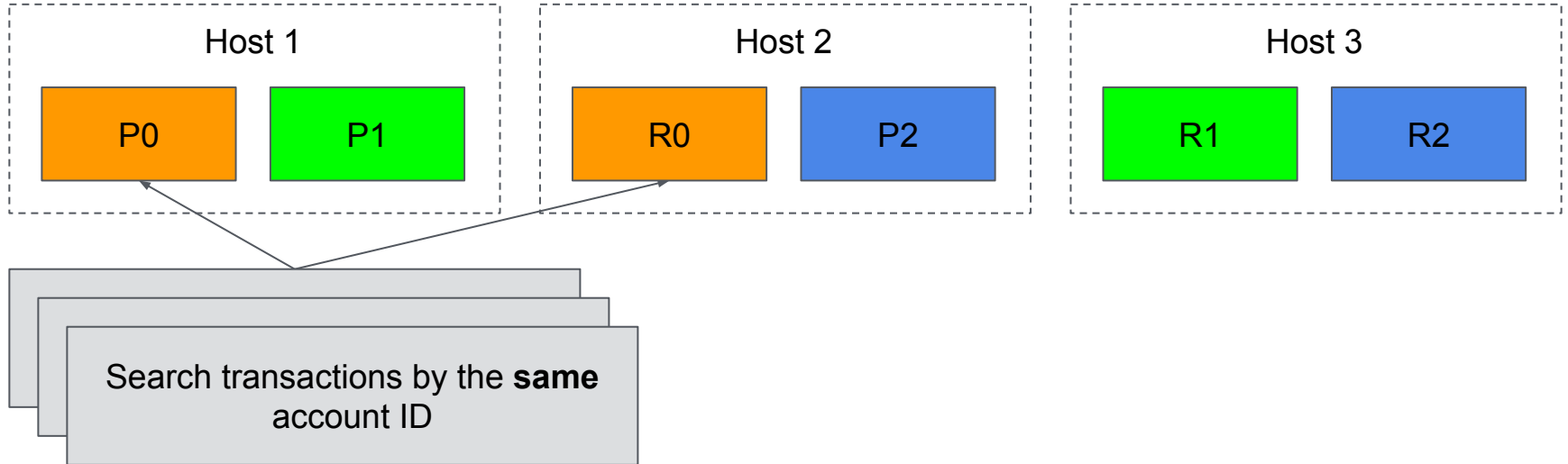
- 1) Нельзя расти по железу
- 2) Чем больше нод, тем больше вероятность выхода из строя одной из нод
- 3) Выход любой ноды из кластера тормозит поиск по любому счету

Роутинг по номеру счета

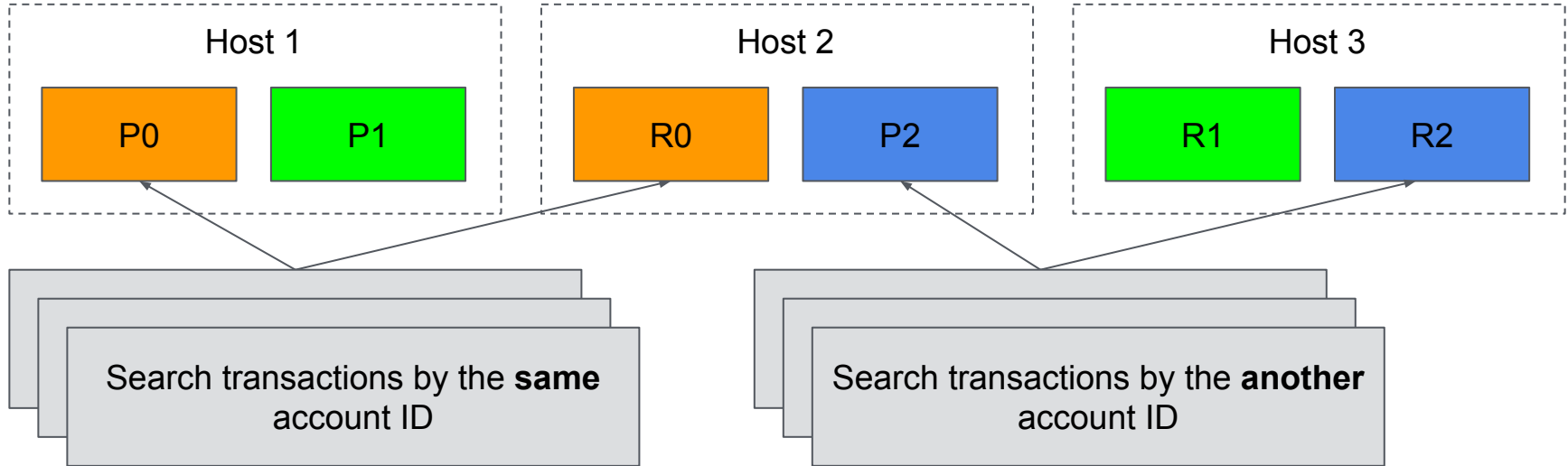


- 1) Клиент запрашивает только свои транзакции
- 2) Меньше аффектит на других клиентов

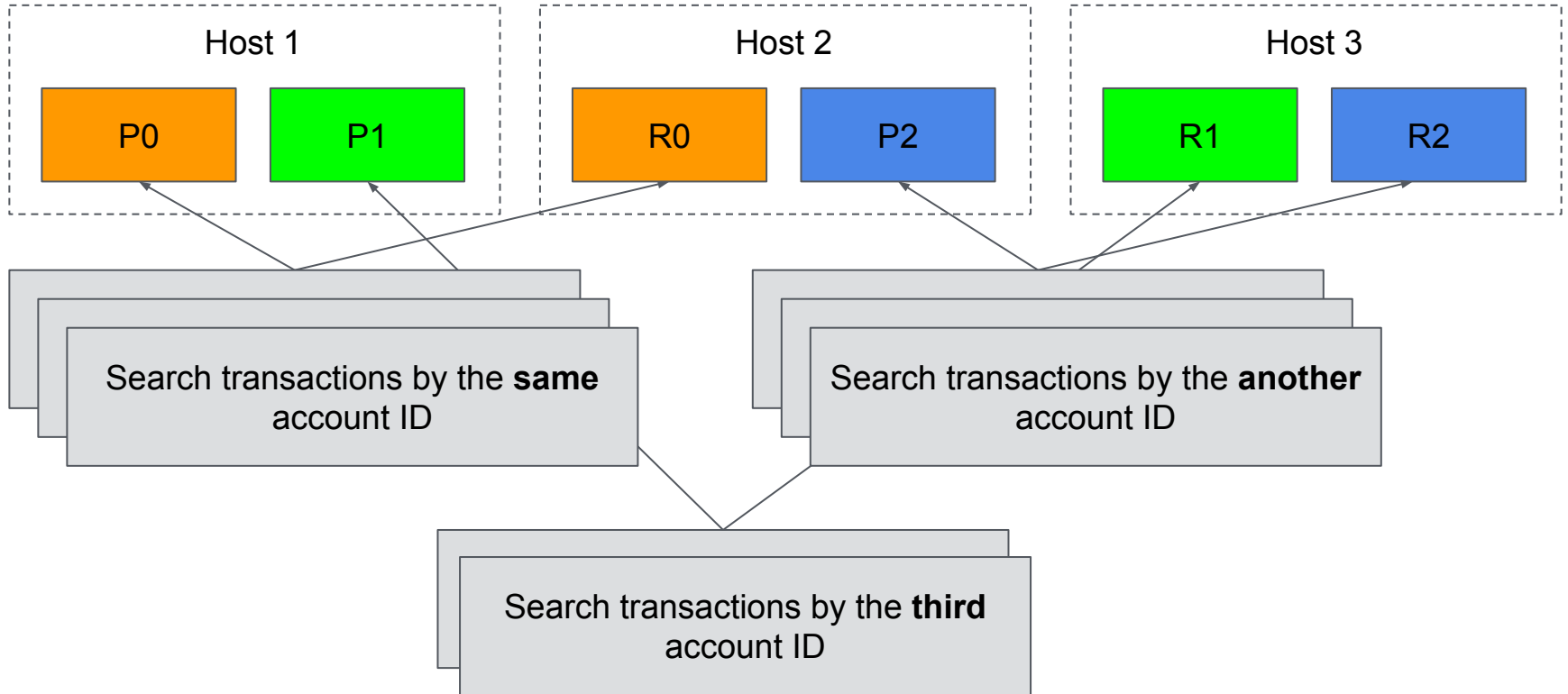
Роутинг по номеру счета



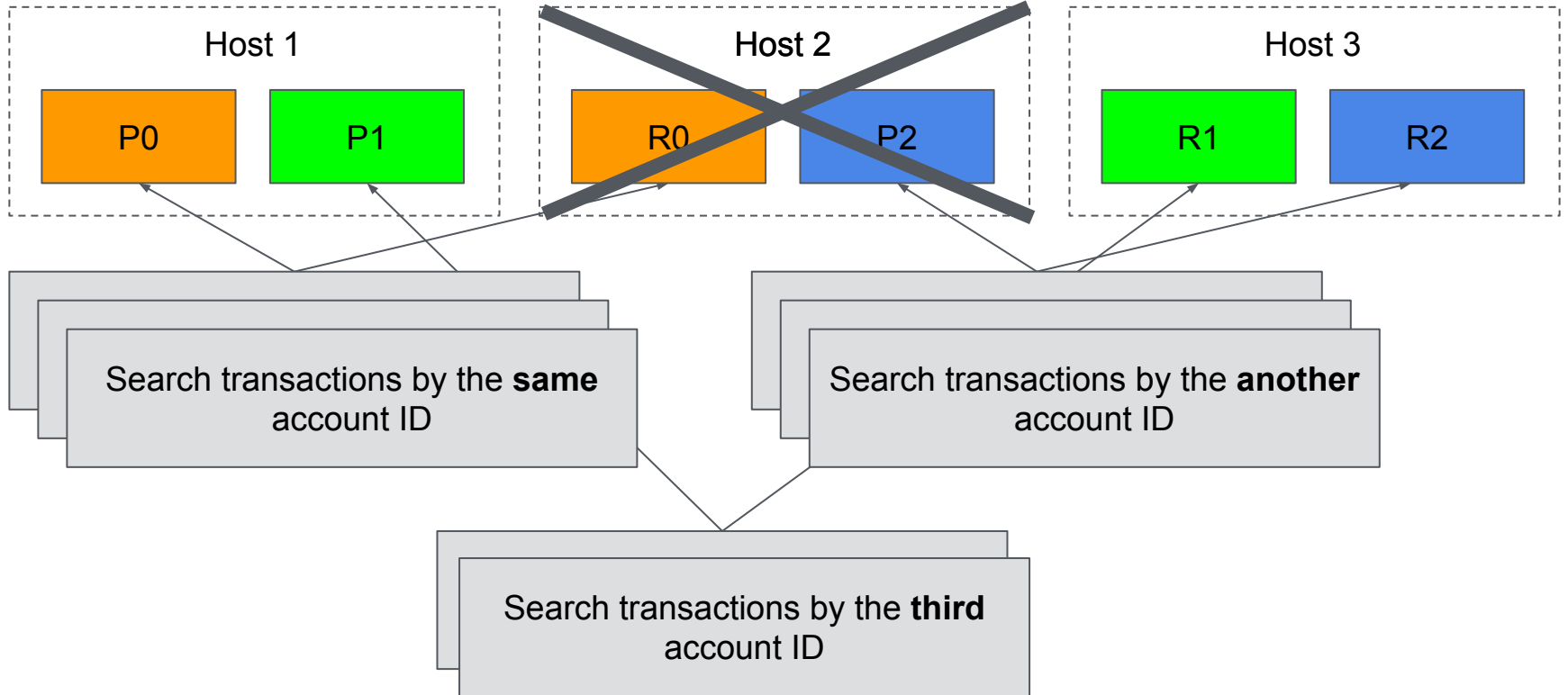
Роутинг по номеру счета



Роутинг по номеру счета



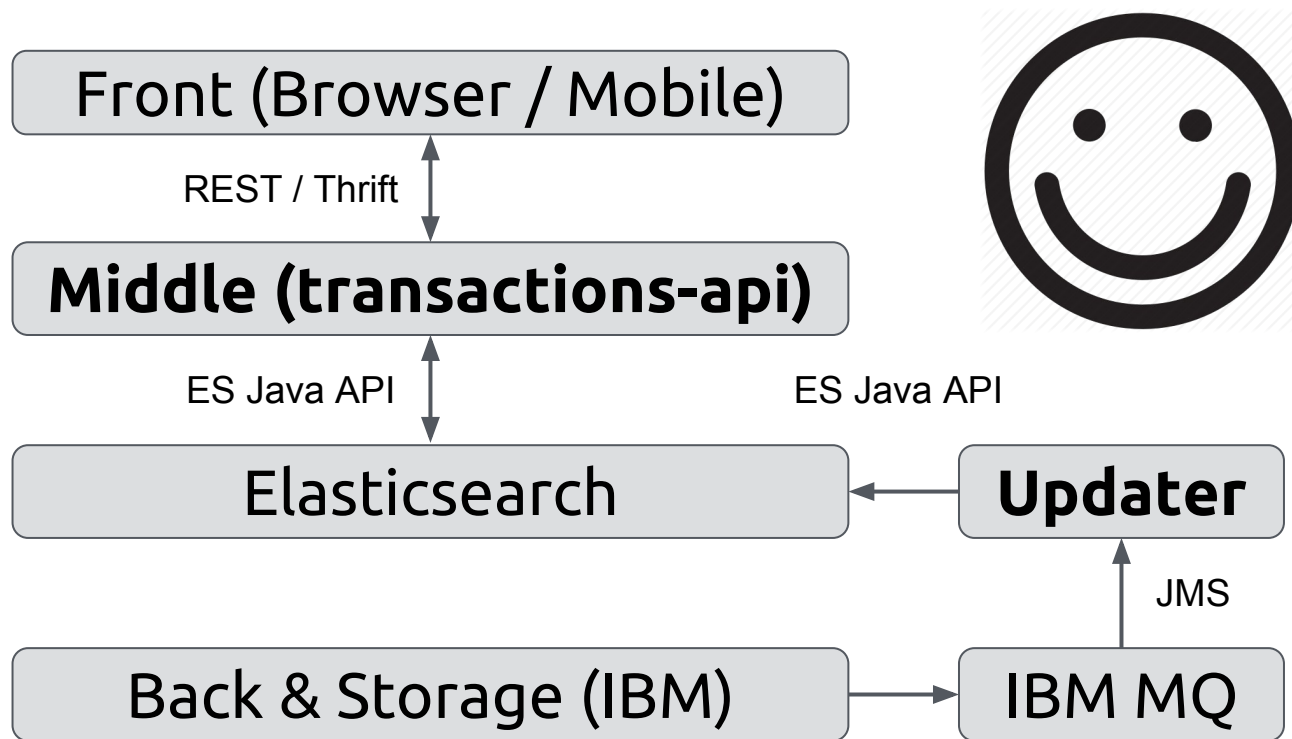
Роутинг по номеру счета



Меньше индексов и шардов

- 1) Ушли от дневных индексов к помесечным
- 2) Кол-во шардов оставили достаточно большим (150), так как на малом кол-ве есть риск раздувания шарда из-за попадания туда нескольких крупных счетов
- 3) В сумме кол-во шардов стало меньшим - 150 против 210 (7x30)

Переписали приложения





Перезалили транзакции

С включенным роутингом по номеру счета

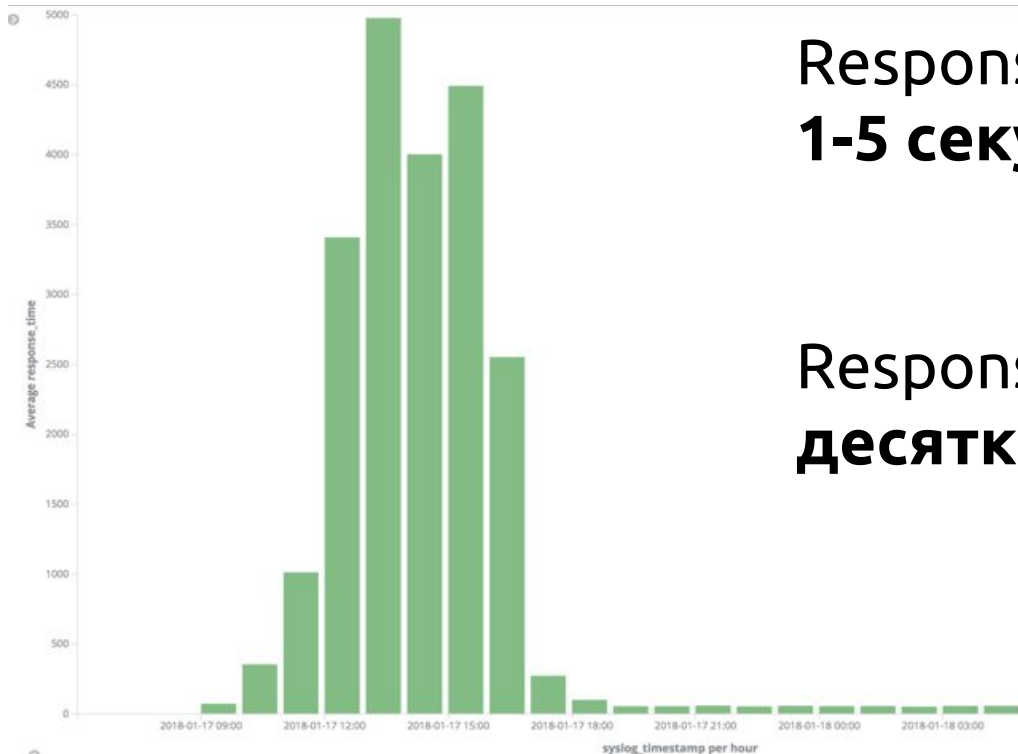
```
{
  "_index": "listing-2017-08",
  "_type": "transaction",
  "_id": "XXXXXXXXXXXXXXXXXXXXXXX",
  "_routing": "12345678901234567890",
  "_source": { ... }
}
```



Полечили тяжелые запросы

`java.lang.IllegalArgumentException: Trying to query 14525 shards, which is over the limit of 2000. This limit exists because querying many shards at the same time can make the job of the coordinating node very CPU and/or memory intensive.`

Эффект



Response time до:
1-5 секунд

Response time после:
десятки миллисекунд



- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо**
- 7) Выводы



- 1) Приложение Выписка
- 2) Зачем нужен Elasticsearch?
- 3) ES inside: индексы и шарды
- 4) Открытие на клиентов и проблемы
- 5) Как побороли
- 6) Демо
- 7) **Выводы**

Выводы

- 1) Роутинг можно и нужно использовать
- 2) Нужно оптимизировать запросы и индексы под конкретный кейс
- 3) Нужно уметь легко перезаливать данные в ES
- 4) Перезаливать из CSV - так себе
- 5) Нужно уметь использовать несколько кластеров параллельно для сравнения произв-ти

Выводы

- 1) Роутинг можно и нужно использовать
- 2) Нужно оптимизировать запросы и индексы под конкретный кейс
- 3) Нужно уметь легко перезаливать данные в ES
- 4) Перезаливать из CSV - так себе
- 5) Нужно уметь использовать несколько кластеров параллельно для сравнения произв-ти



Выводы

- 1) Роутинг можно и нужно использовать
- 2) Kafka
- 3) Kafka
- 4) Kafka
- 5) Нужно уметь использовать несколько кластеров параллельно для сравнения произв-ти