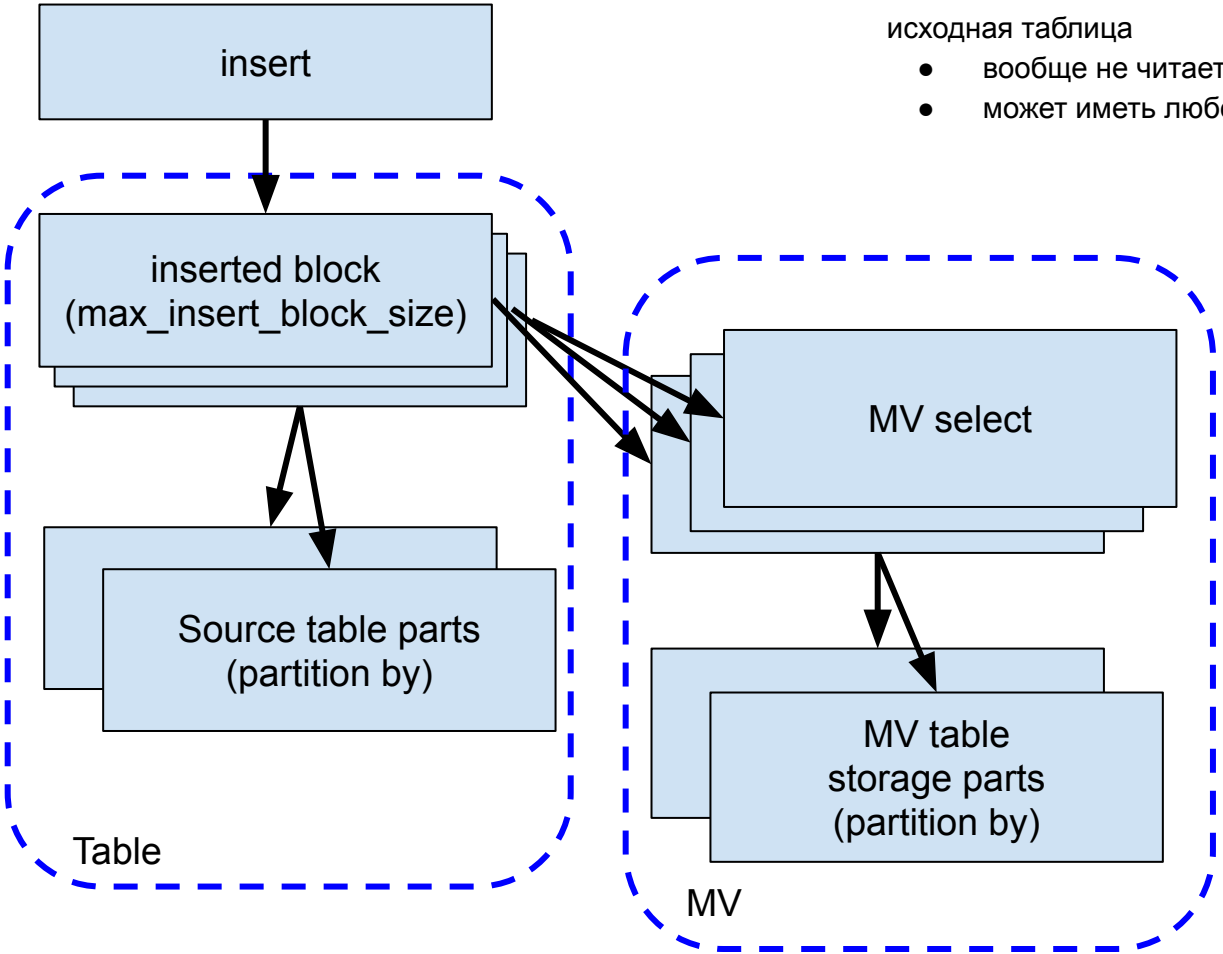
ClickHouse

MV, на самом деле

Журавлев Денис, DBA, Revjet

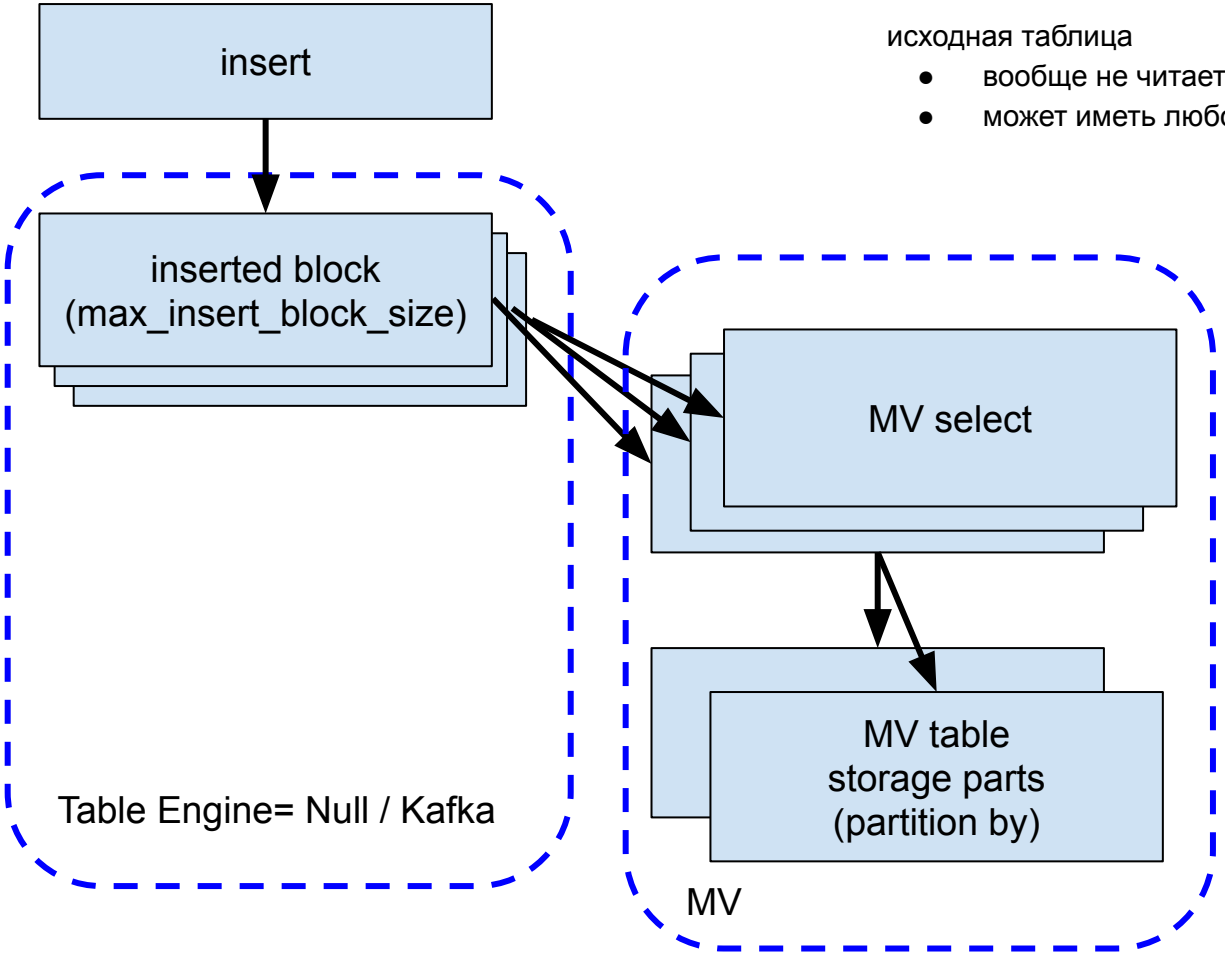
MV изнутри



исходная таблица

- вообще не читается
- может иметь любой движок, даже Null

MV изнутри



исходная таблица

- вообще не читается
- может иметь любой движок, даже Null

Зачем MV нужны

- Пред-агрегации (Summing/Aggregating)
- Единая точка входа / Трансформации (ExtractTransform)
- Дублирование таблицы в таблицу с другим ключом (+эмуляция инверсного индекса)
- Kafka

MV c TO vs MV .inner

Неявная таблица `.inner.mv1`

```
CREATE MATERIALIZED VIEW mv1
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, b, d)
AS
SELECT a, b, d, count() AS cnt
FROM source
GROUP BY a, b, d;
```

```
DESCRIBE TABLE `.inner.mv1`
```

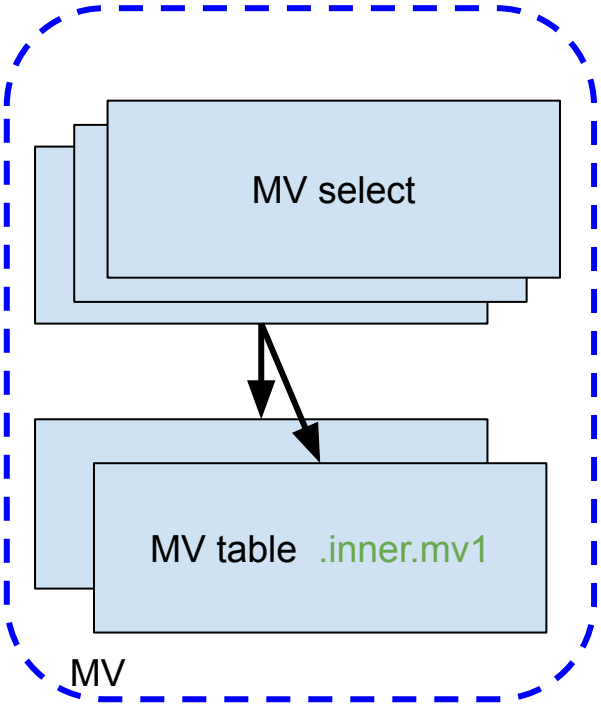
name	type	default_type
a	Int64	
b	Int64	
d	Date	
cnt	UInt64	

Явная таблица dest

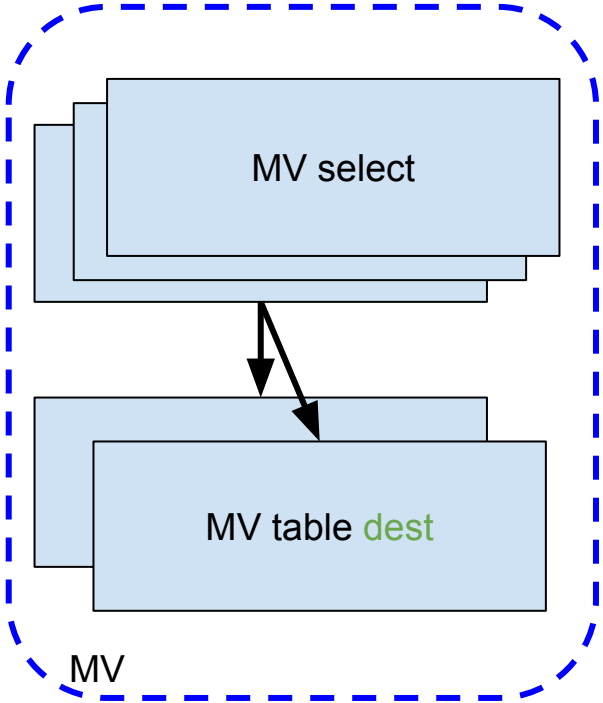
```
CREATE TABLE dest
(a Int64, b Int64, d Date, cnt UInt64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, b, d);

CREATE MATERIALIZED VIEW mv1
TO dest
AS
SELECT a, b, d, count() AS cnt
FROM source
GROUP BY a, b, d;
```

MV c TO vs MV .inner

Неявная таблица `.inner.mv1`



Явная таблица dest



MV c TO vs MV .inner

Неявная таблица `.inner.mv1`

- не работает optimize_move_to_prewhere
  при запросах к mv приходится писать from `.inner.mv1`
- работает populate
- drop table mv1 неявный сторадж удаляется

Явная таблица dest

- не работает populate (можно сделать insert)
- drop table mv1 явный сторадж НЕ удаляется

MV c TO vs MV .inner

- insert into mv1 / insert into `.inner.mv1` / insert into dest

Populate

- Ни разу не пригодился

Основные ошибки

- Таблица-сторадж не знает про MV и select в MV !!!

Основные ошибки

Неправильный ключ (order by) таблицы

Неправильно

```
CREATE MATERIALIZED VIEW mv1
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, b)
AS
SELECT a, b, d, count() AS cnt
FROM source
GROUP BY a, b, d;
```

Правильно

```
CREATE MATERIALIZED VIEW mv1
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, b, d)
AS
SELECT a, b, d, count() AS cnt
FROM source
GROUP BY a, b, d;
```

Основные ошибки

Двойные группировки (неправильно)

```
CREATE MATERIALIZED VIEW mv1
ENGINE = AggregatingMergeTree
PARTITION BY toYYYYMM(hour)
ORDER BY hour
POPULATE
AS
SELECT toStartOfHour(time) hour,
       maxState(cnt_by_minute) max_by_hour,
       sumState(cnt_by_minute) cnt_by_hour
FROM
(
     SELECT minute, count() AS cnt_by_minute
     FROM source
     GROUP BY minute
)
GROUP BY hour
```

```
insert into source
values (now()), (now())

результат max_by_hour = 2
```

```
insert into source
values (now());
insert into source
values (now());

результат max_by_hour = 1
```

Основные ошибки

MV никак не связан с исходной таблицей !!!

- MV не следит за Replacing/CollapsingMT в исходной таблице
- MV ничего не знает про truncate / alter delete / alter update / drop partition / drop table / rename исходной таблицы
- MV может хранить данные за другой период

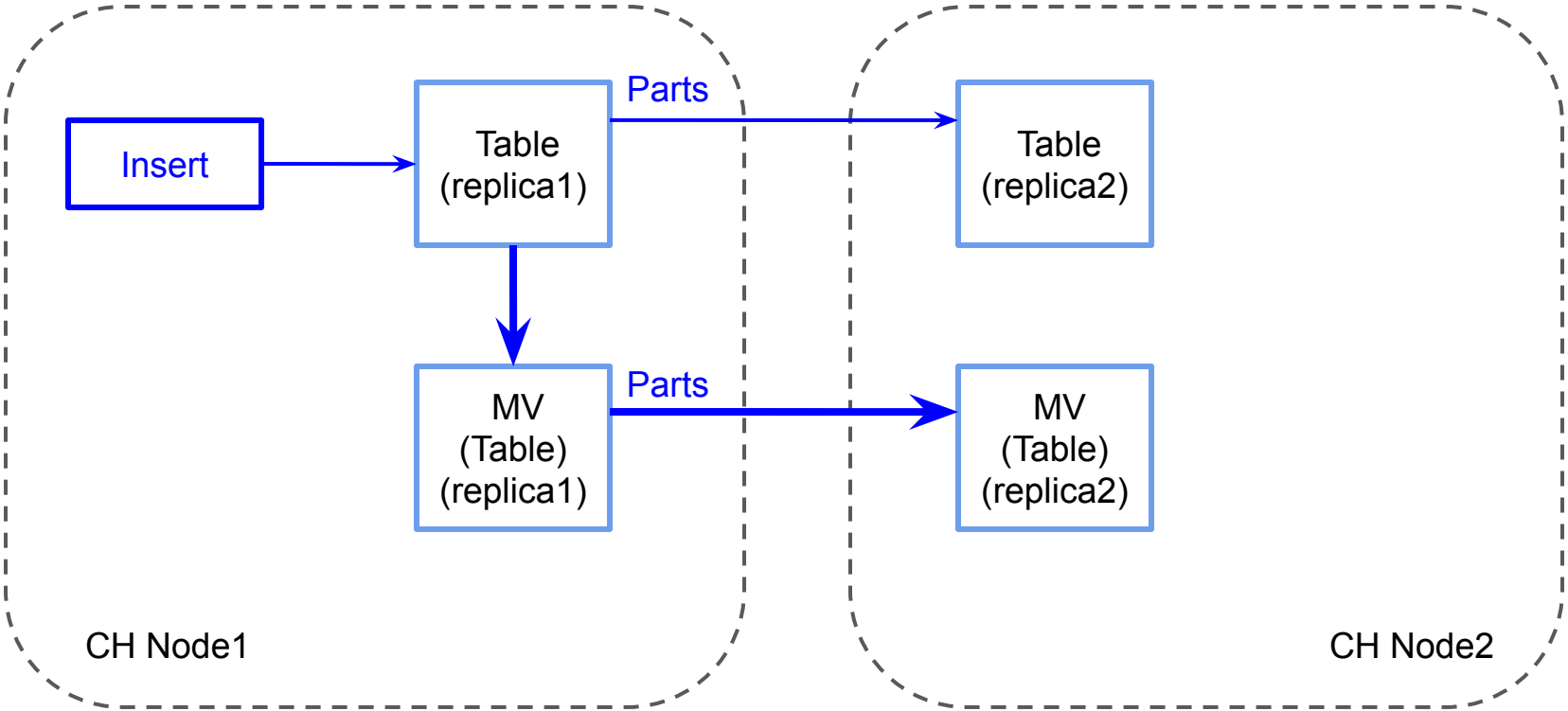
Основные ошибки

MV использует имена полей !!!

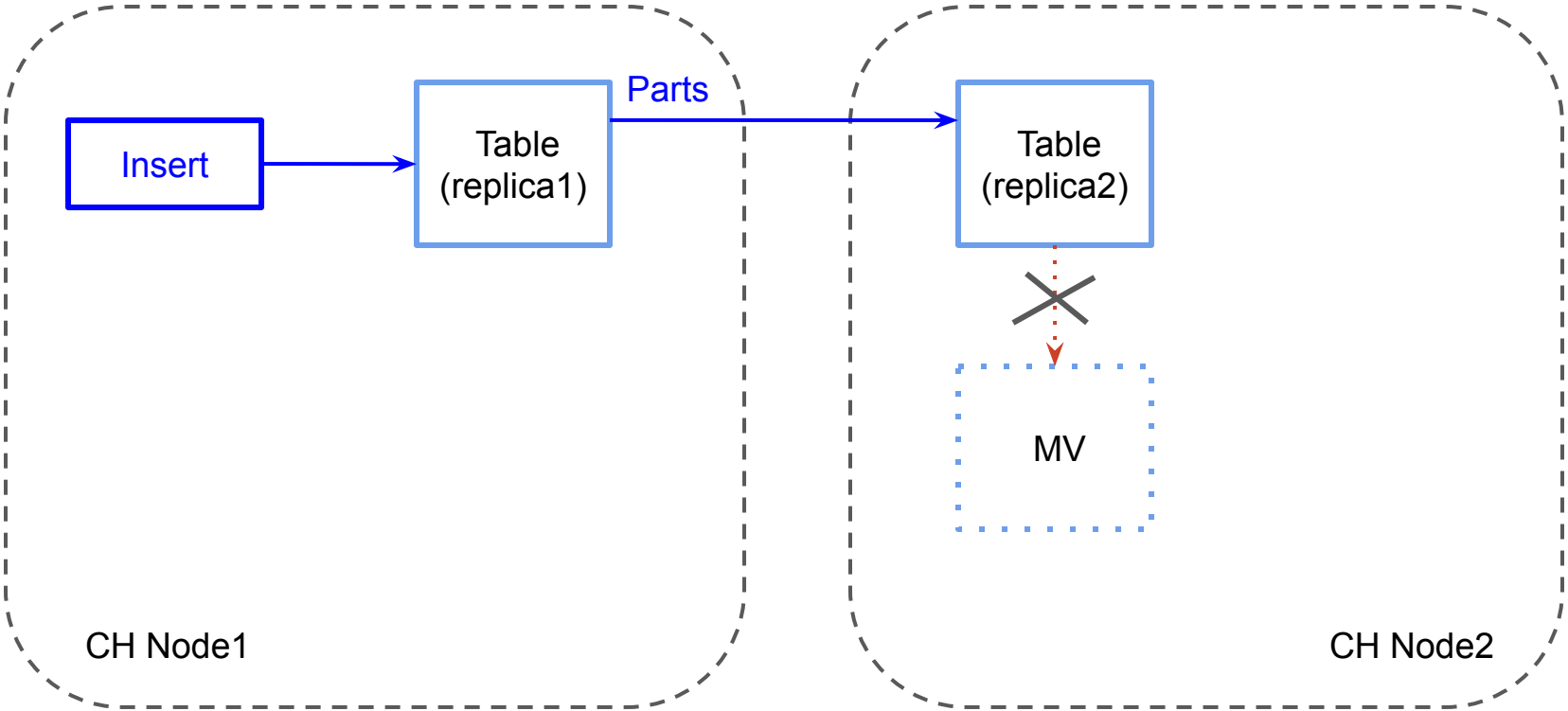
```
CREATE MATERIALIZED VIEW
mv1 (a Int64, d Date, cnt Int64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, d)
POPULATE
AS
SELECT a, d, count()
FROM source
GROUP BY a, d;
```

```
CREATE MATERIALIZED VIEW
mv1 (a Int64, d Date, cnt Int64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d)
ORDER BY (a, d)
POPULATE
AS
SELECT a, d, count() as cnt
FROM source
GROUP BY a, d;
```

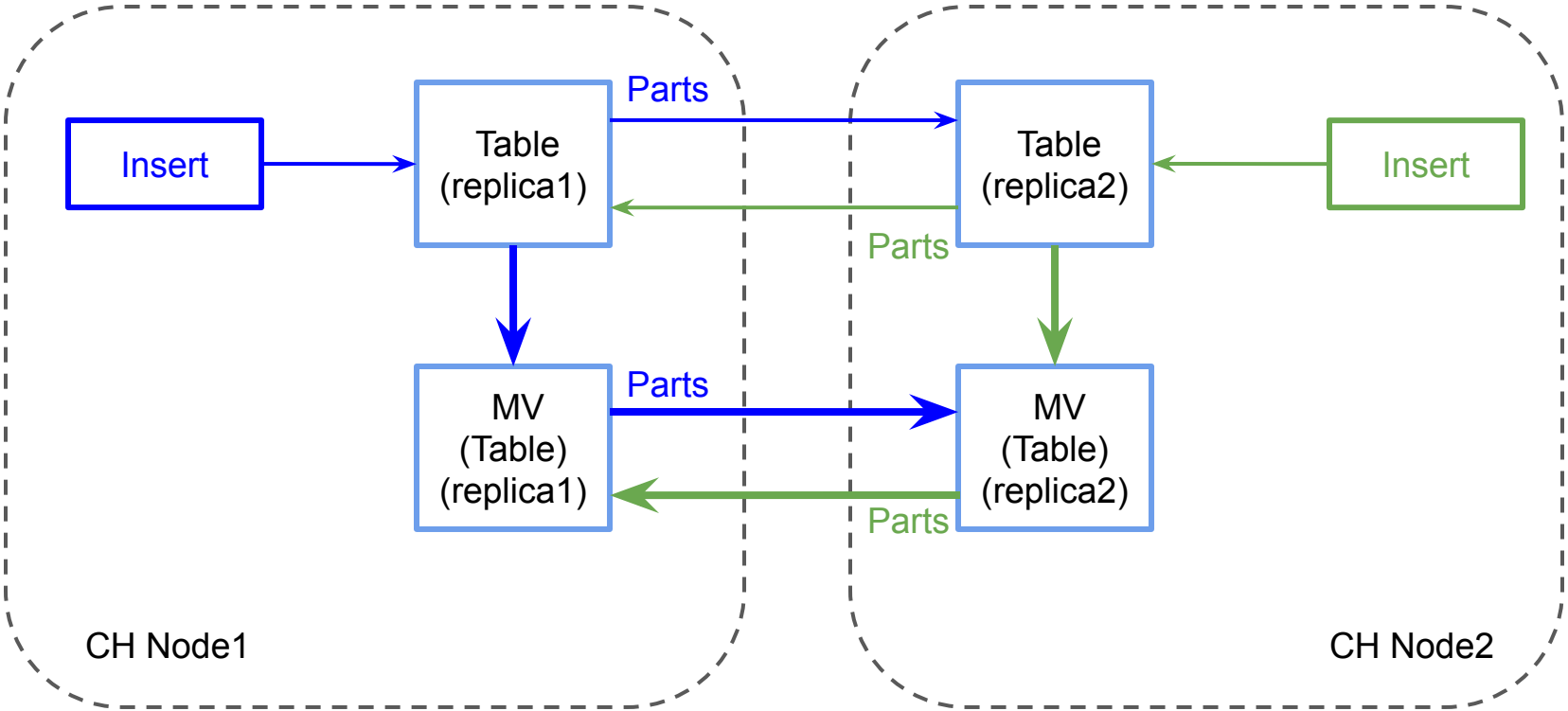
ReplicatedMT и MV



ReplicatedMT и MV

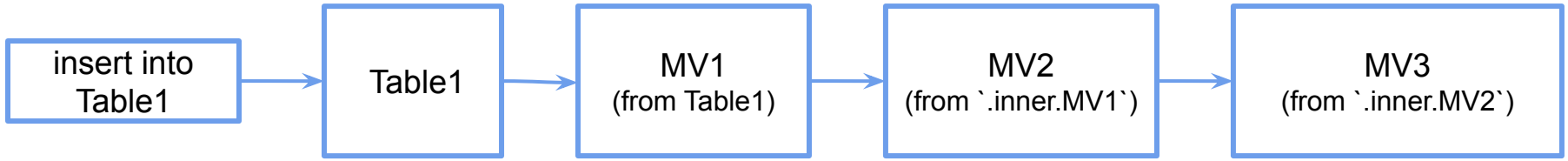


ReplicatedMT и MV

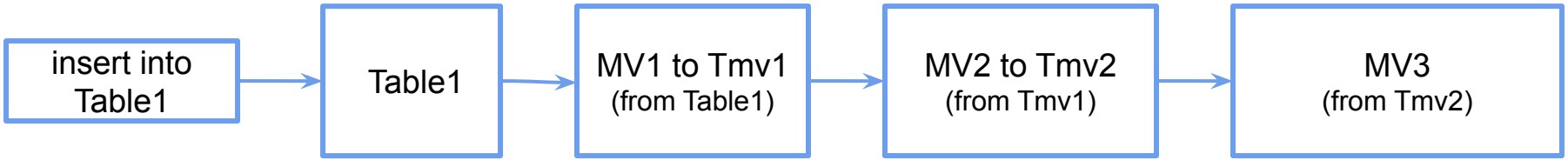


Каскадирование MV

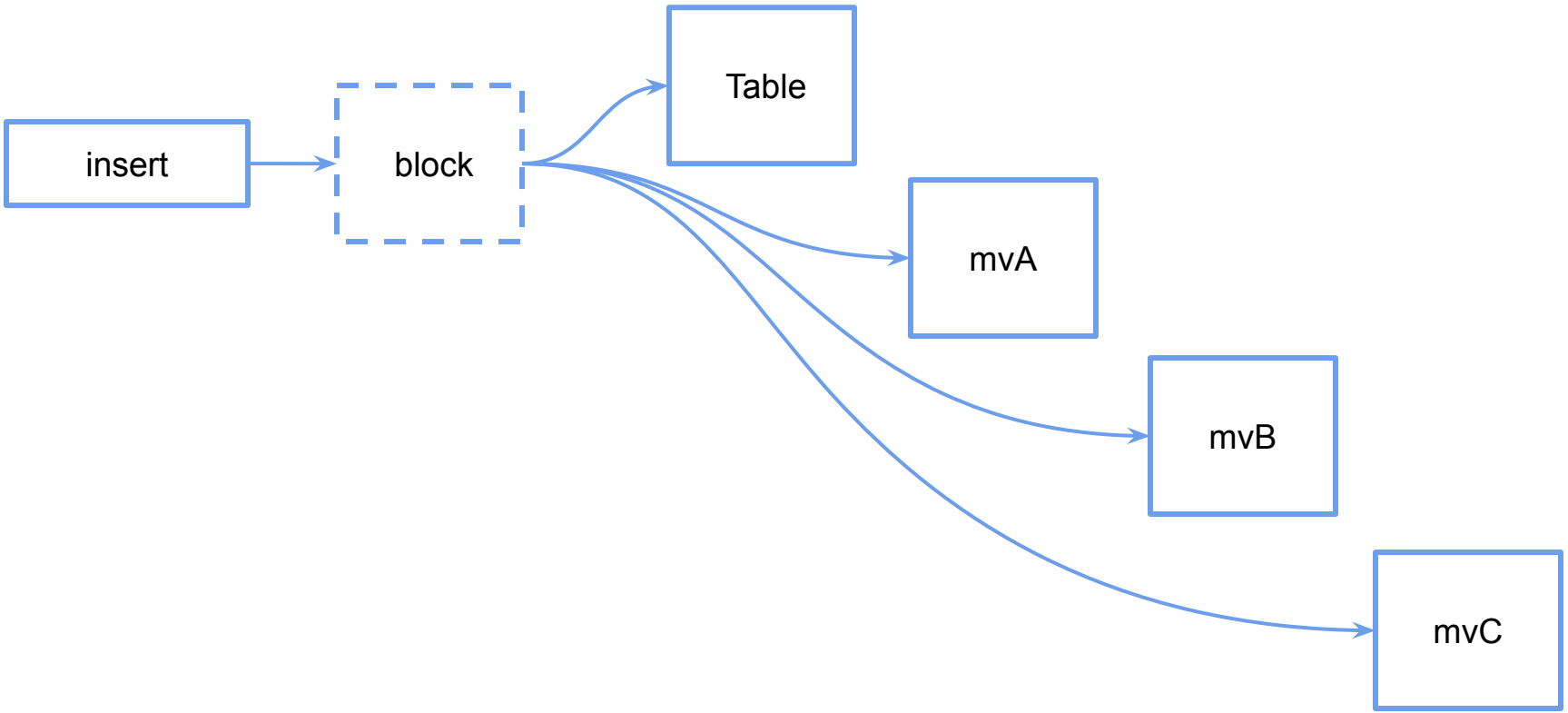
.inner



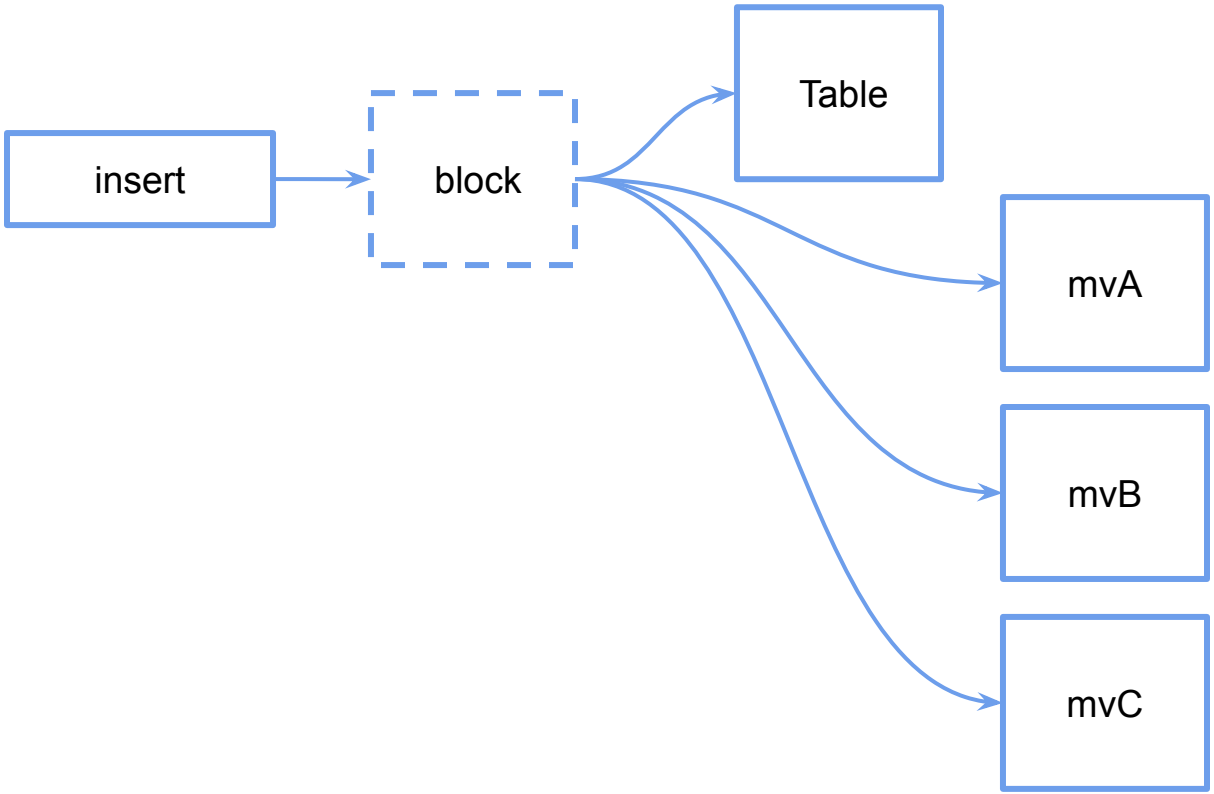
to



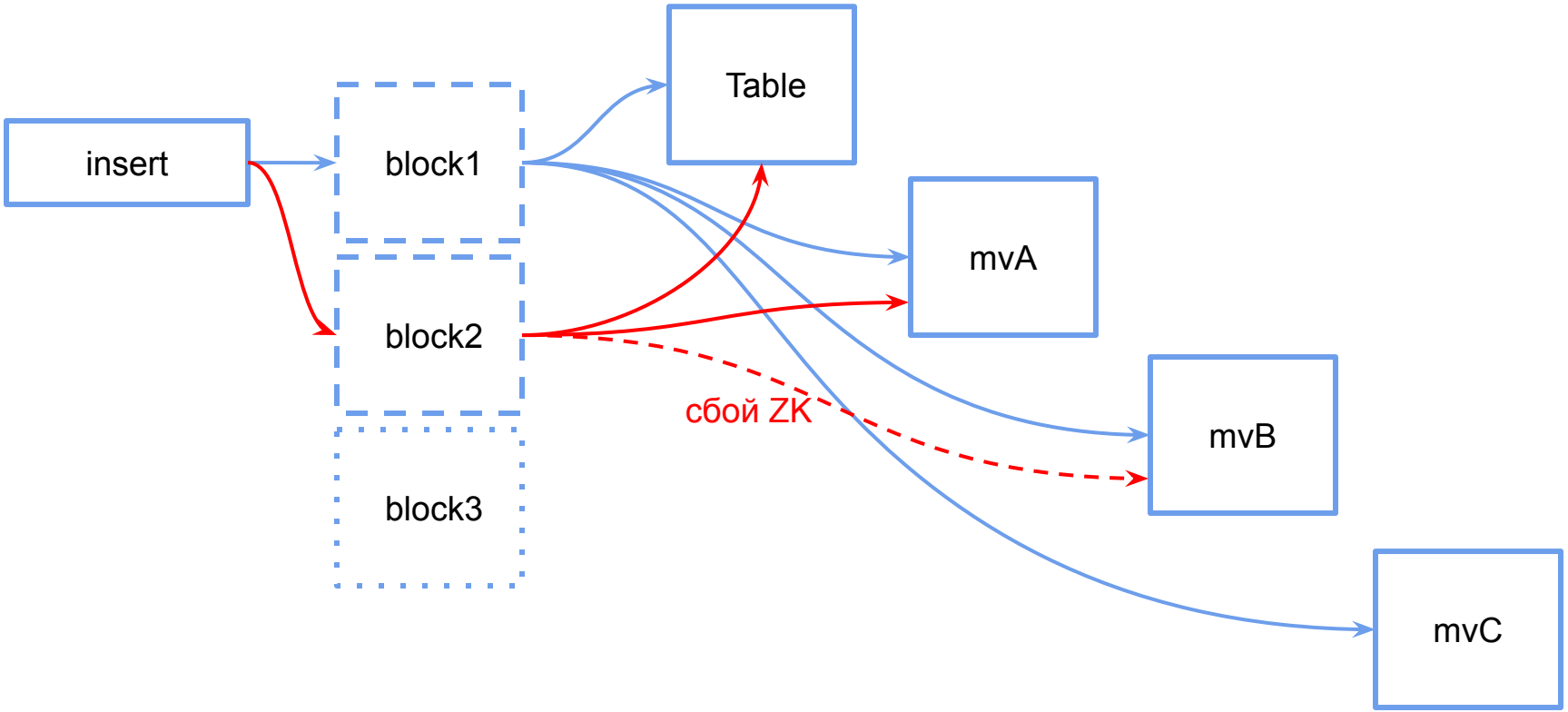
Порядок вставки в MV



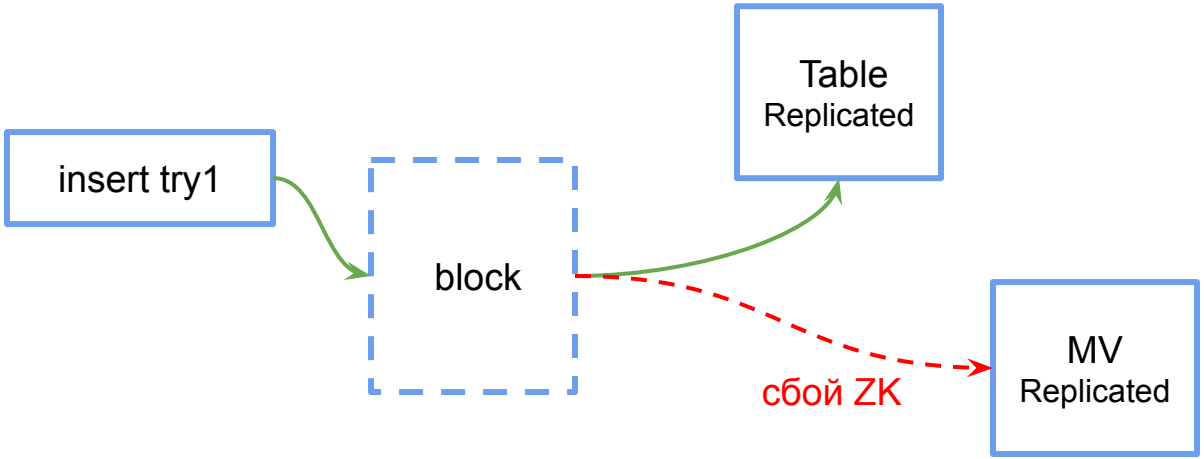
Порядок вставки в MV и parallel_view_processing



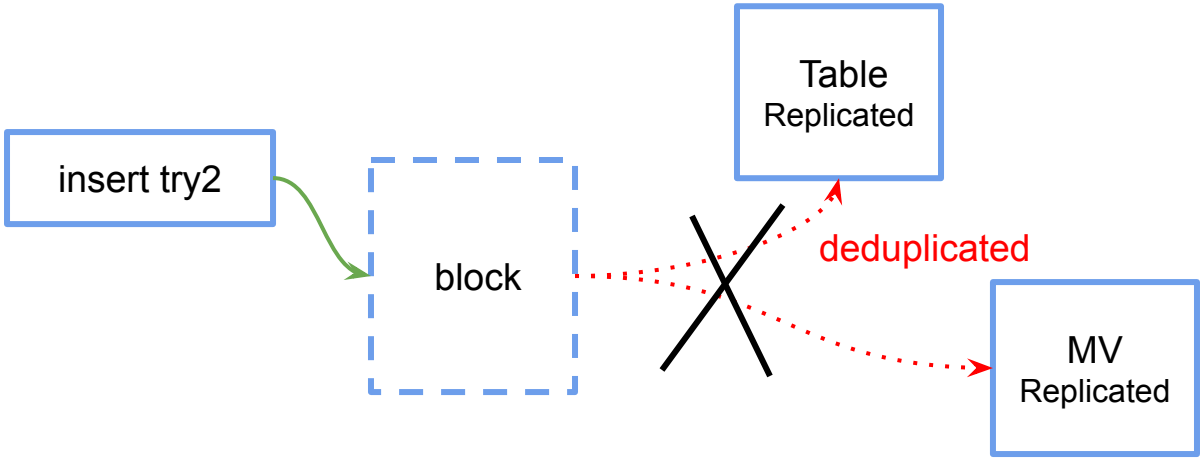
Порядок вставки в MV



MV insert deduplication



MV insert deduplication



- параметр deduplicate_blocks_in_dependent_materialized_views = 0 (default)

JOIN в MV

- Несколько блоков -> несколько JOIN
- Внешние словари и dictGet или engine=Join и joinGet
- MV на обе таблицы в JOIN (2 MV to dest + RIGHT JOIN)

MV from distributed / To distributed

- From Local* to Local
- From Distributed to Distributed
- From Local* to Distributed
- From Distributed to Local

* (то что Local за Distributed -- в этом случае несущественно)

Как изменить MV

Неявная таблица (.inner.mv1)

1. останавливаем insert-ы
2. detach table mv1
3. alter table `.inner.mv1`
4. attach materialized view mv1

4a. поправить .sql + attach table

Явная таблица (TO dest)

1. останавливаем insert-ы
2. alter table dest
3. drop table mv1
4. create materialized view mv1

Как изменить MV

Неявная таблица (.inner.mv1)

```
DETACH TABLE mv1

ALTER TABLE `.inner.mv1`
    ADD COLUMN b Int64 AFTER a,
    MODIFY ORDER BY (a, b)

ATTACH MATERIALIZED VIEW mv1
ENGINE = SummingMergeTree
ORDER BY (a, b) AS
SELECT a, b, sum(amount) AS s
FROM source
GROUP BY a, b
```

Явная таблица (TO dest)

```
ALTER TABLE dest
    ADD COLUMN b Int64 AFTER a,
    MODIFY ORDER BY (a, b)

DROP TABLE MV1

CREATE MATERIALIZED VIEW mv1
TO dest
SELECT a, b, sum(amount) AS s
FROM source
GROUP BY a, b
```

Как изменить MV

```
SET allow_experimental_alter_materialized_view_structure = 1;

ALTER TABLE mv1 MODIFY QUERY
    SELECT v * 2 as v, 1 as v2
    FROM source;
```

Как наполнить MV существующей таблицы (populate)

- создаем MV с where date >= в_будущем
  ждем в_будущем, вставляем все за до в_будущем

- останавливаем инсерты
  вставляем во вью все за сегодня
  запускаем инсерты
  вставляем все за до сегодня

- snapshot / alter table freeze

- создаем MV
  вставляем все за до вчера , (до начала партиции)
  ждем новый месяц (новую партицию)
  дропаем неполную партицию, пересчитываем

Как наполнить MV существующей таблицы (populate)

```
CREATE TABLE dest(a Int64, d Date, cnt UInt64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d) ORDER BY (a, d);

-- создаем MV с where date >= в_будущем
CREATE MATERIALIZED VIEW mv1 TO dest AS
SELECT a, d, count() AS cnt
FROM source
WHERE d >= '2020-10-02'
GROUP BY a, d;

-- наступает 2020-10-02
INSERT INTO dest -- вставляем все за до в_будущем
SELECT a, d, count() AS cnt
FROM source
WHERE d < '2020-10-02' -- кусками по 1-му месяцу (или .. дню)
GROUP BY a, d;
```

Как наполнить MV существующей таблицы (populate)

```
CREATE TABLE dest(a Int64, d Date, cnt UInt64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d) ORDER BY (a, d);

-- останавливаем инсерты

INSERT INTO dest -- вставляем во вью все за сегодня
SELECT a, d, count() AS cnt
FROM source
WHERE d = today()
GROUP BY a, d;

CREATE MATERIALIZED VIEW mv1 TO dest AS
SELECT a, d, count() AS cnt
FROM source
GROUP BY a, d;

-- запускаем инсерты

INSERT INTO dest -- вставляем все за до сегодня
SELECT a, d, count() AS cnt
FROM source
WHERE d < today() -- кусками по 1-му месяцу (или .. дню)
GROUP BY a, d;
```

Как наполнить MV существующей таблицы (populate)

```
CREATE TABLE dest(a Int64, d Date, cnt UInt64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d) ORDER BY (a, d);

-- останавливаем инсерты

ALTER TABLE dest FREEZE;

CREATE MATERIALIZED VIEW mv1 TO dest AS
SELECT a, d, count() AS cnt
FROM source
GROUP BY a, d;

-- запускаем инсерты

create table temp_snapshot as source;

-- mv .../shadow/1/data/default/source/* .../data/default/temp_snapshot/detached/
-- cd .../data/default/temp_snapshot/detached/
-- let i=1;for f in `ls -1` ; do echo $i $f;((i++)); echo "alter table temp_snapshot attach
part '$f';"|clickhouse-client ; done

INSERT INTO dest -- вставляем во вью из temp_snapshot
SELECT a, d, count() AS cnt
FROM temp_snapshot -- кусками по 1-му месяцу (или .. дню)
GROUP BY a, d;
```

Как наполнить MV существующей таблицы (populate)

```
CREATE TABLE dest(a Int64, d Date, cnt UInt64)
ENGINE = SummingMergeTree
PARTITION BY toYYYYMM(d) ORDER BY (a, d);

CREATE MATERIALIZED VIEW mv1 TO dest AS
SELECT a, d, count() AS cnt
FROM source
GROUP BY a, d;

INSERT INTO dest -- вставляем все за предыдущий период
SELECT a, d, count() AS cnt
FROM source
WHERE d < '2020-10-01' -- кусками по 1-му месяцу (или .. дню)
GROUP BY a, d;

-- наступает 2020-11-01
ALTER TABLE dest DROP PARTITION 202010;

INSERT INTO dest
SELECT a, d, count() AS cnt
FROM source
WHERE d >= '2020-10-01' and d < '2020-11-01'
GROUP BY a, d;
```

полезняшки

Alter MV and SummingMergeTree
https://gist.github.com/den-crane/3a8d57253711e277b2a60a75b5dfeef6

How to convert not replicated MV to replicated
https://gist.github.com/den-crane/80cb95e74f046be2d8ffae58d9b04e8f

How to alter MV
how to convert MV with implicit storage .inner to explicit storage (with TO)
https://gist.github.com/den-crane/431010ca08b9e51b960e55344b1dbbe3

MV_populating_with_freeze
https://gist.github.com/den-crane/64c05be67ef8a926242011188ce96f44

Several_MV_one_internal_store
https://gist.github.com/den-crane/005633470c70877dd28c00211cd9fcfb

Populate AggregatingMergeTree through null table
https://gist.github.com/den-crane/f7382cd4f1f859ff6ac46afe7dc9925a

Rename MV in CH before 19.8.3.8
https://gist.github.com/den-crane/d9a0bceb58a7d5314ea66e8edbebc9f7