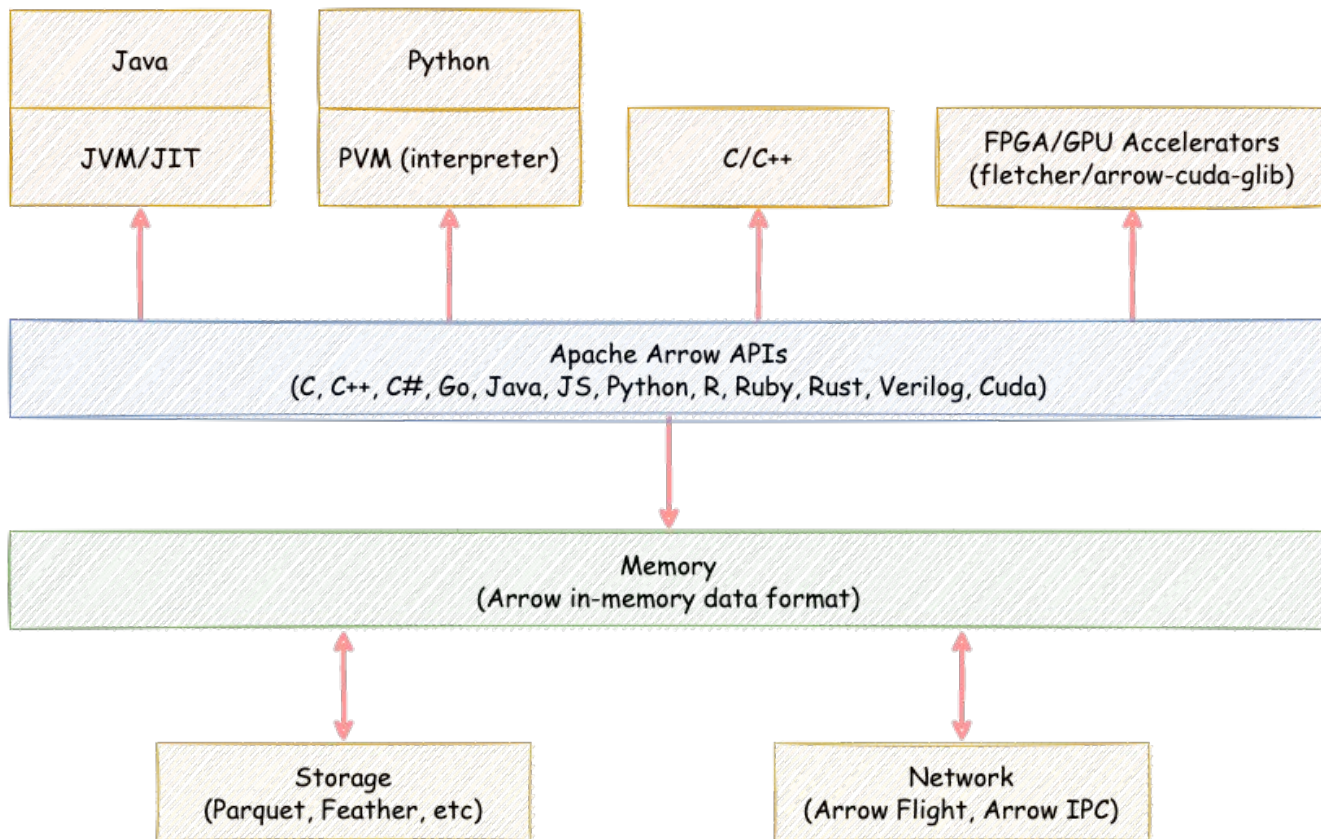


Apache Arrow. В погоне за скоростью

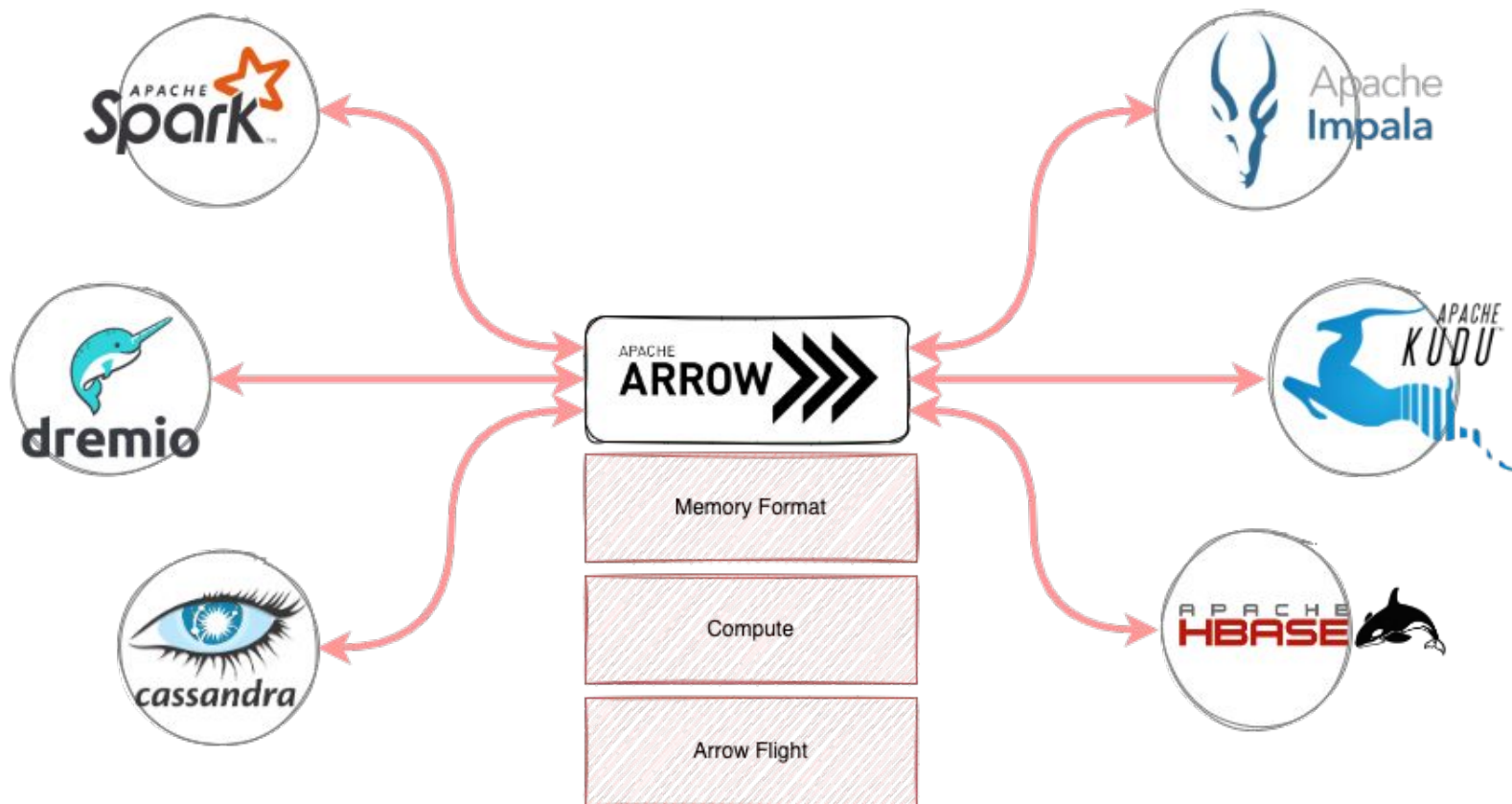
О нас

- Компания Querify Labs [\[1\]](#)
 - Разработка новых СУБД и data management систем: движки, оптимизаторы, storage, протоколы.
- Работали много лет над распределенными in-memory движками:
 - Apache Ignite
 - Hazelcast
- Контрибьюторы в Apache Calcite [\[2\]](#) и Apache Ignite

Apache Arrow ecosystem



Arrow powered

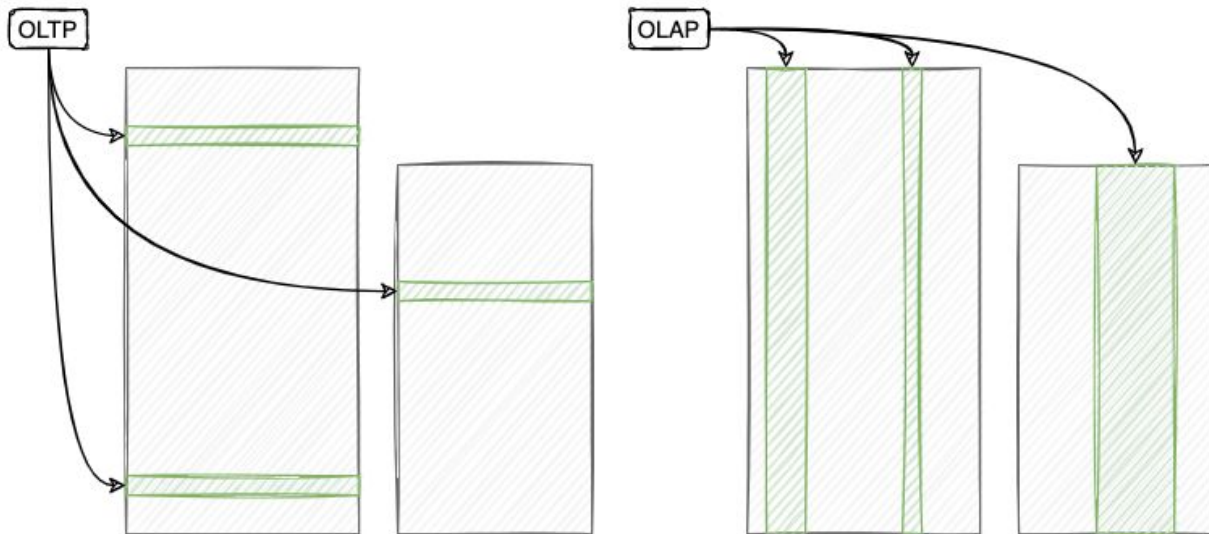


Columnar data. What for?

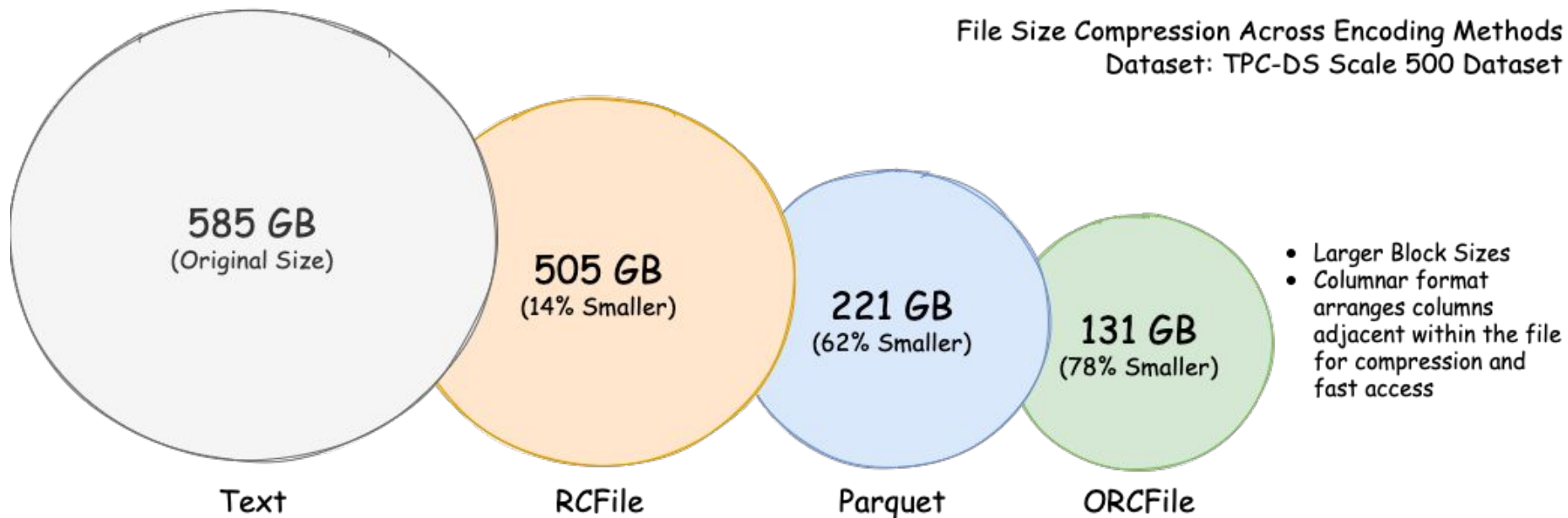
OLTP vs OLAP

OLTP vs OLAP

- OLTP: точечные запросы, широкий набор колонок
 - Относительно небольшое количество строк и трансформаций
 - Критичен выбор оптимального метода доступа (скан vs Index, выбор оптимального индекса)
- OLAP: “длинные” запросы, узкий набор колонок



Optimizes data footprint

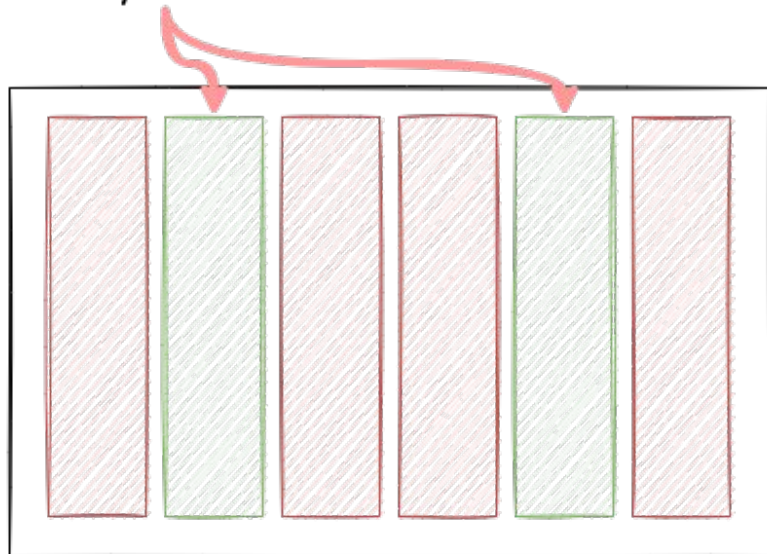


Optimizes reads

```
SELECT AVG(col_2), SUM(col_5) FROM table_1
```

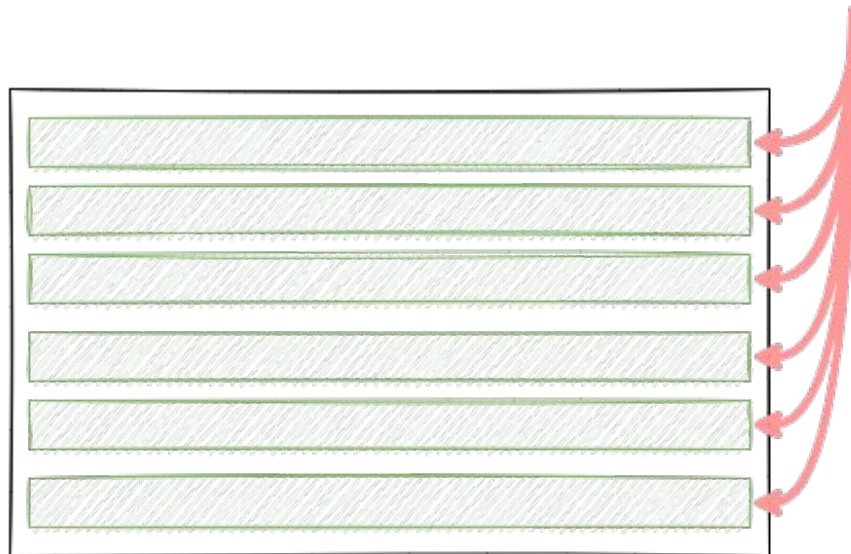

Optimizes reads

Read only needed data



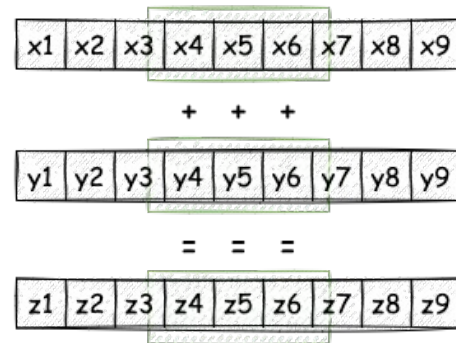
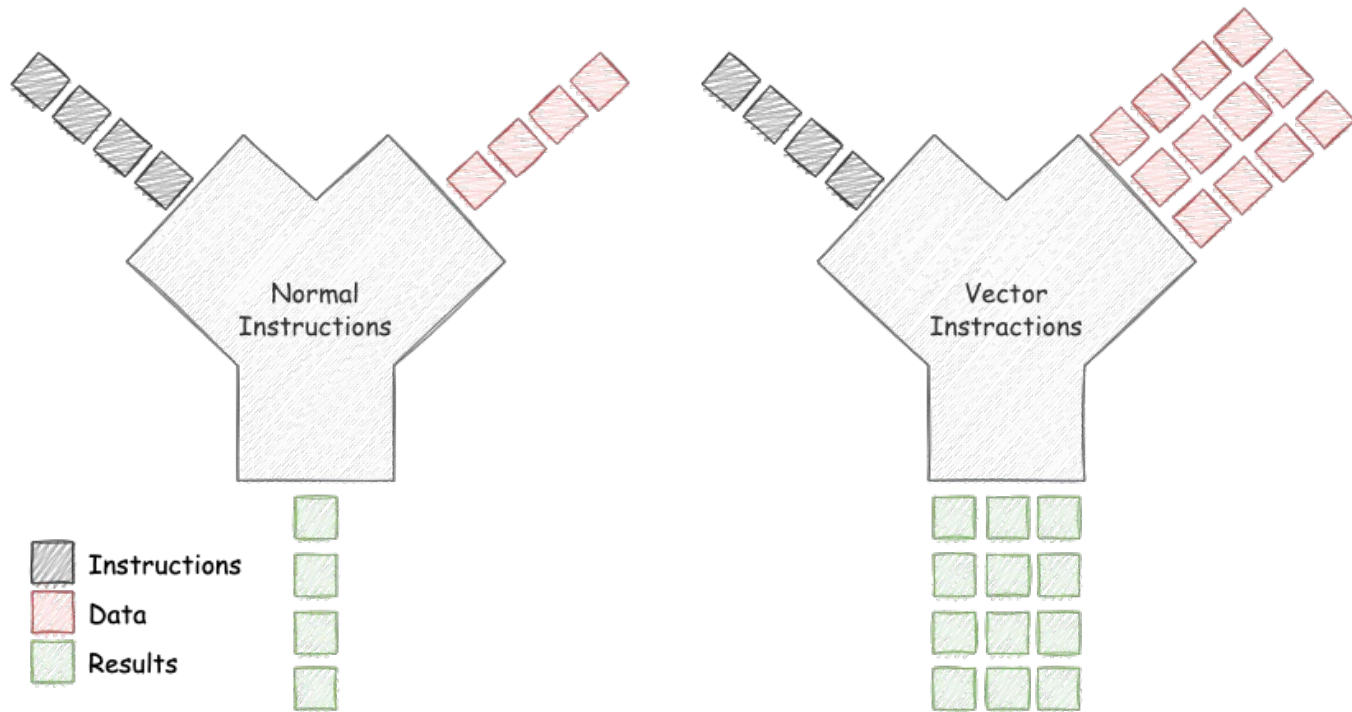
File in columnar format

Read all the data

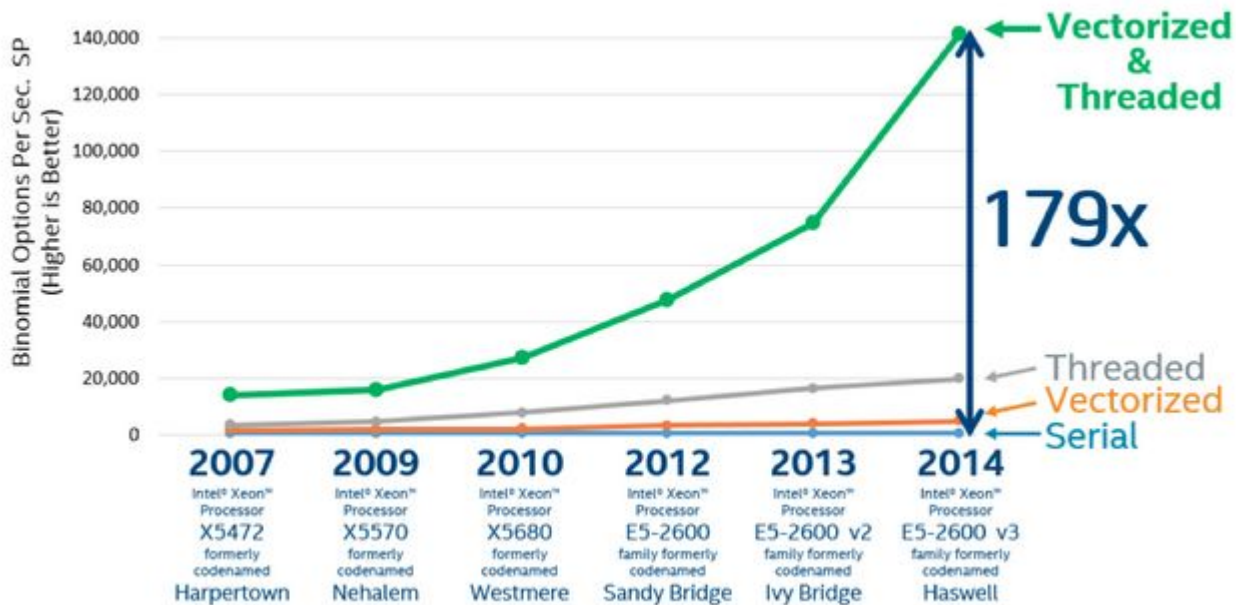


File in row format

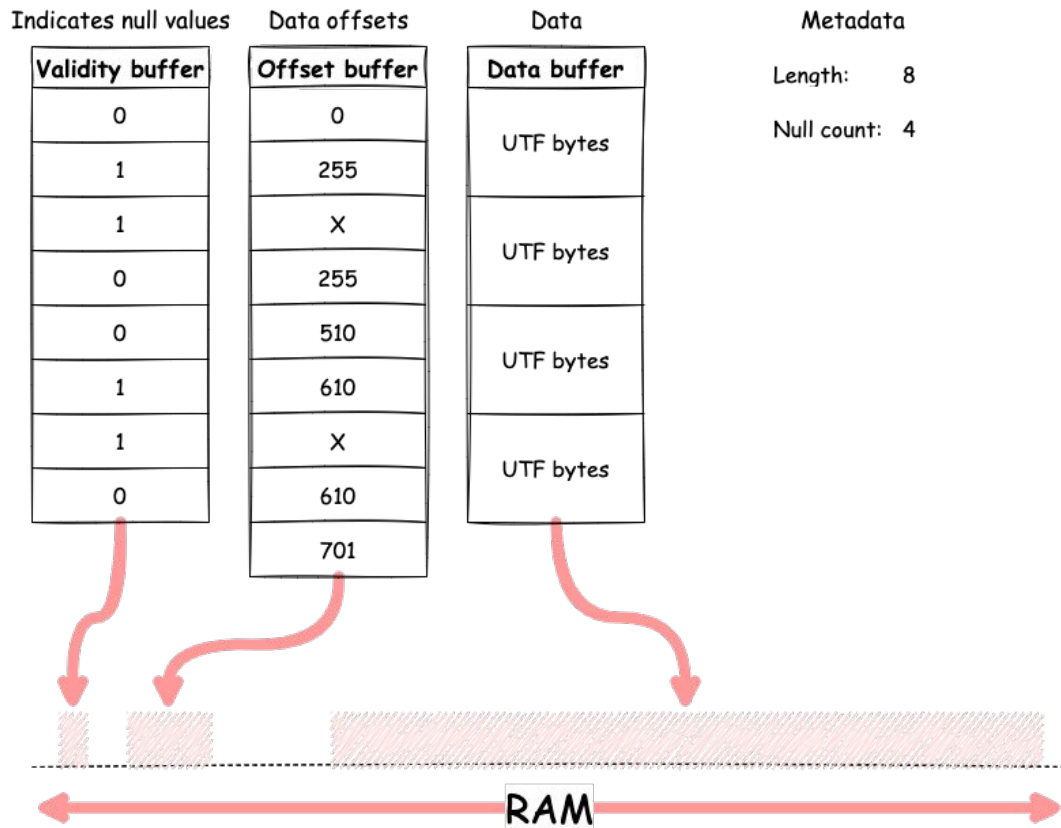
Optimizes compute operations



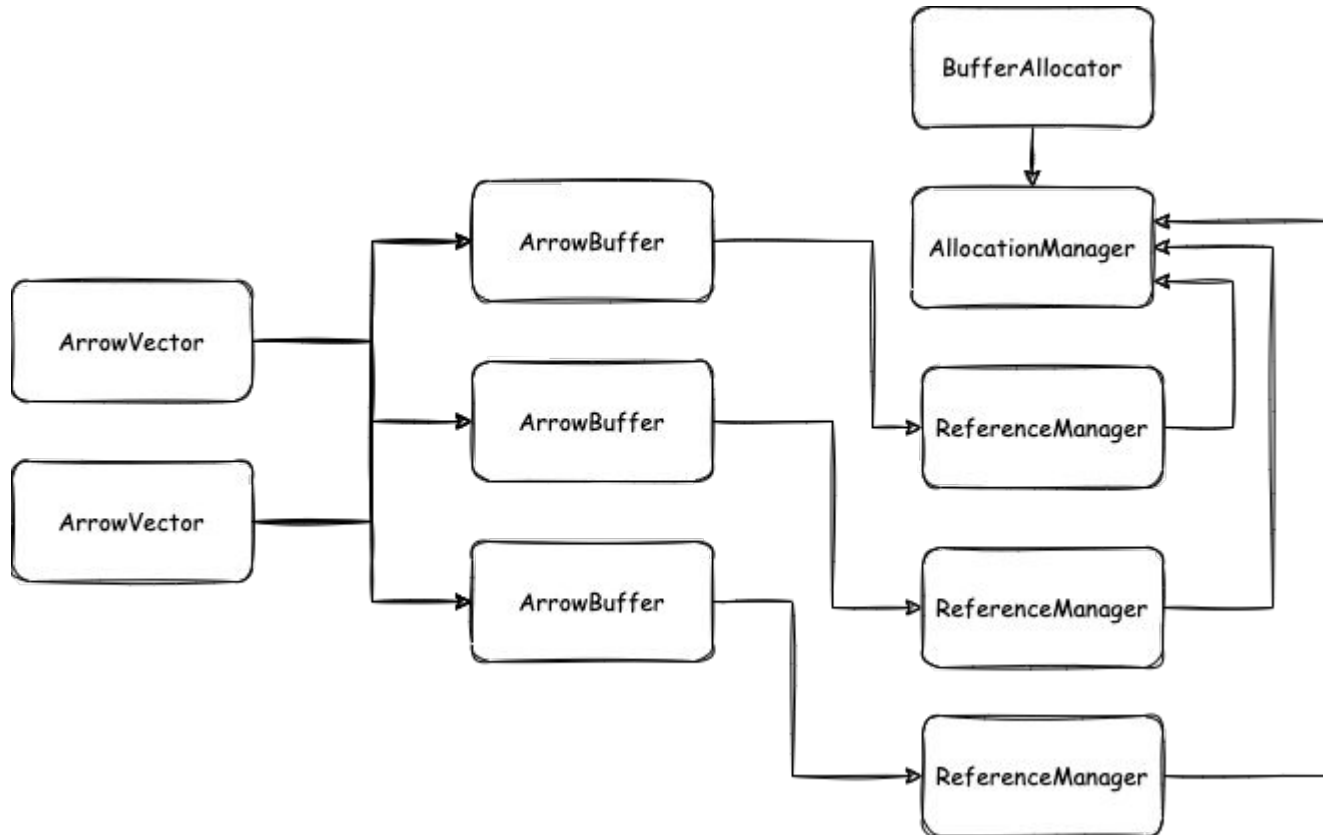
Optimizes compute operations



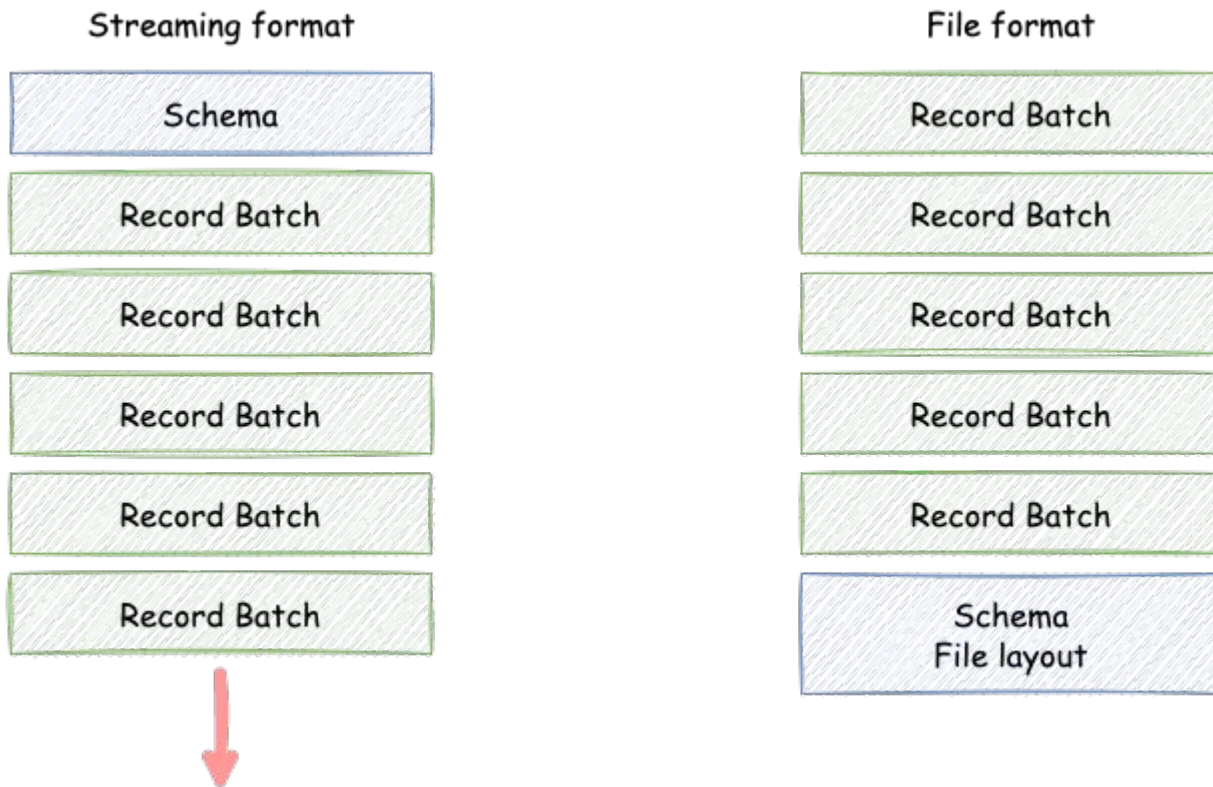
Arrow: memory format



Arrow: memory format (Java)

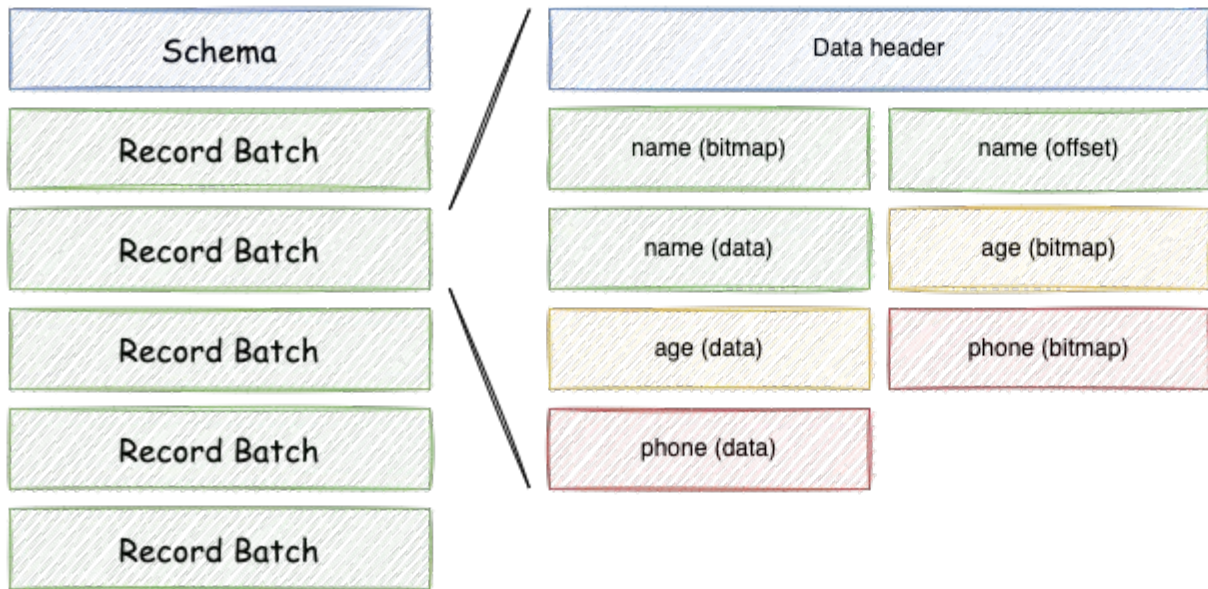


No-marshalling



No-marshalling

Streaming format



Row record:

```
{  
  name: 'Joe',  
  age: 18,  
  phone: 555-444-666  
}
```

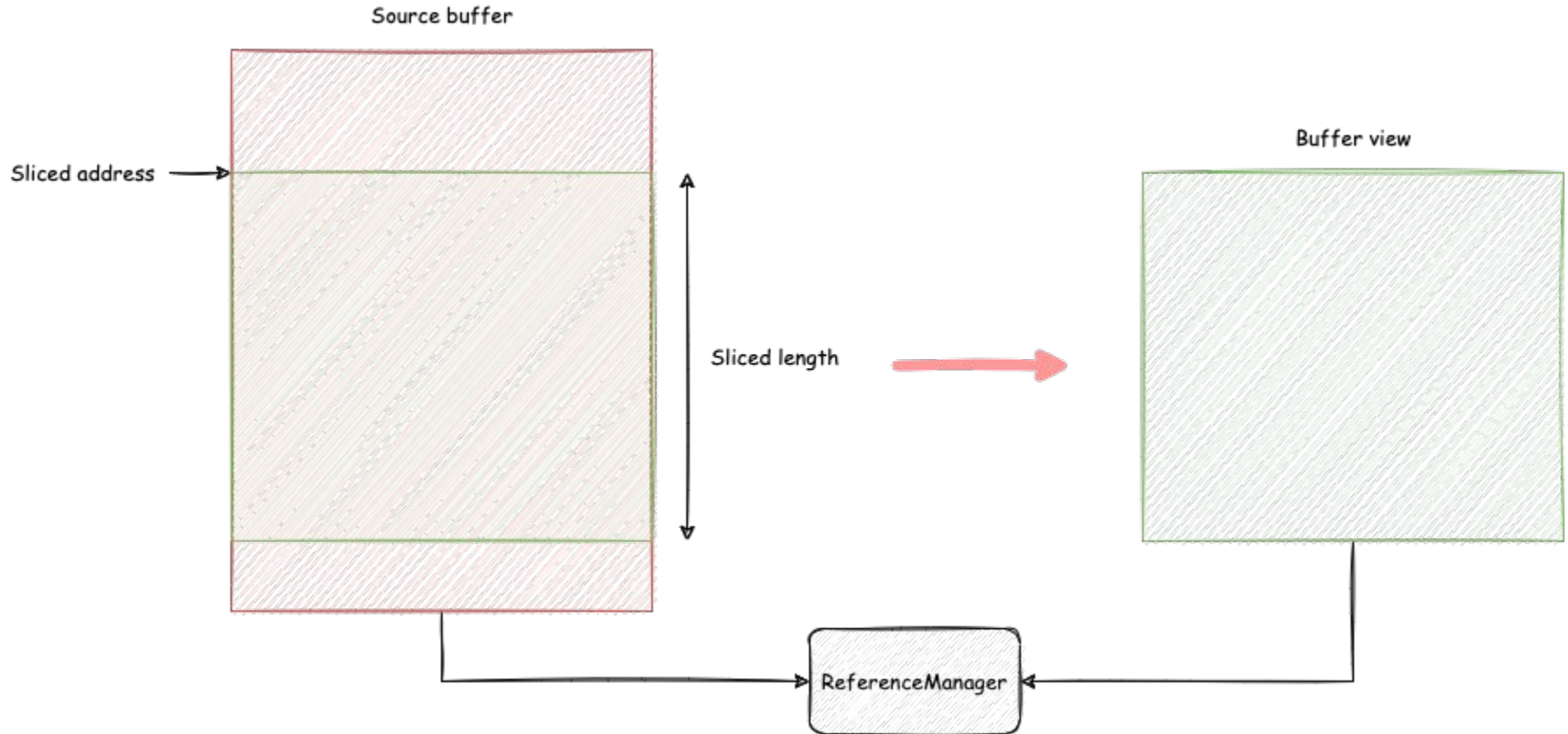
Zero-copy

- Slicing/transferring
- Zero-copy reads
- Zero-copy IPC

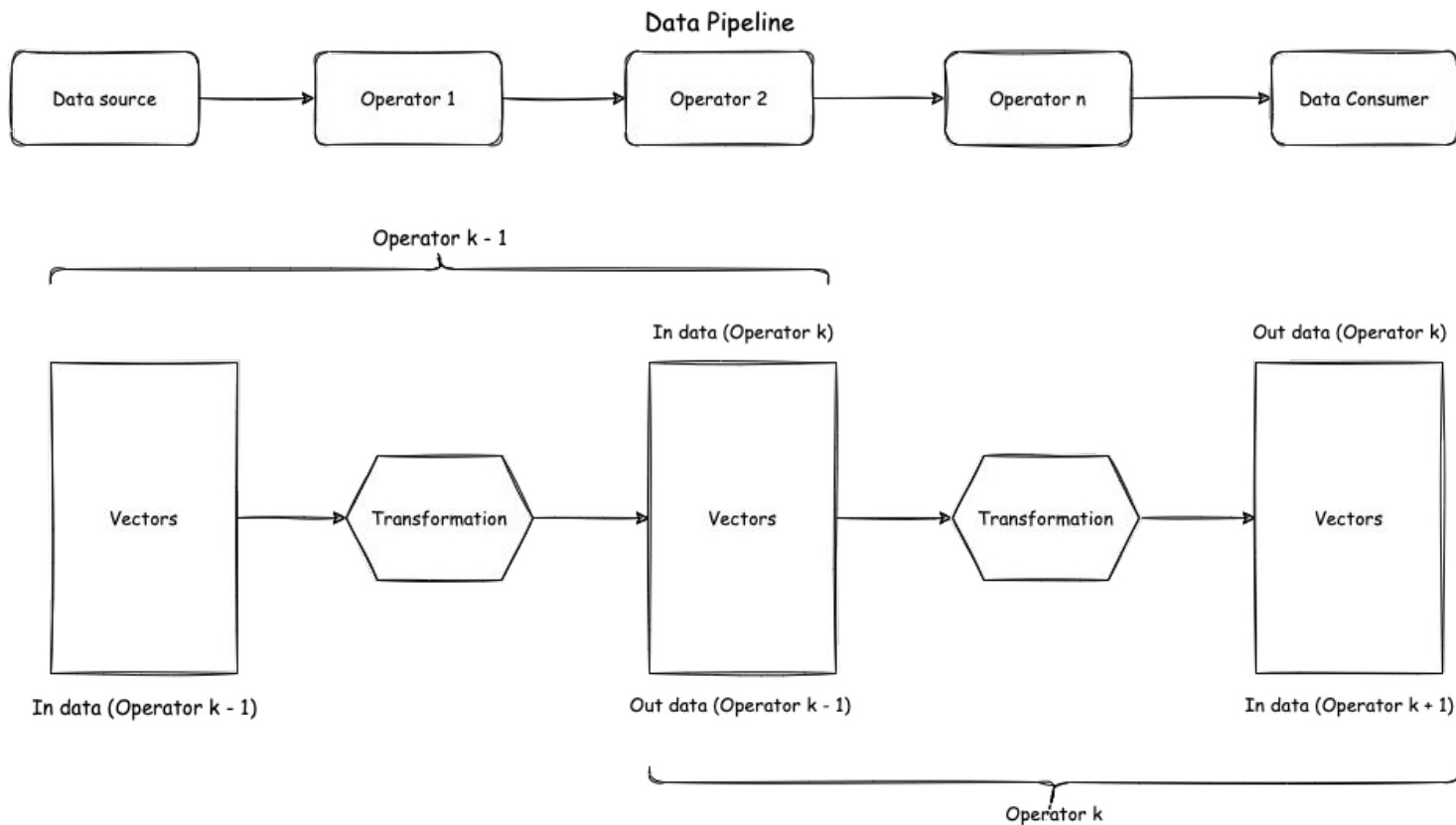
Zero-copy: slicing

```
SELECT col_1 FROM table_1 LIMIT 10 OFFSET 15
```

Zero-copy: slicing



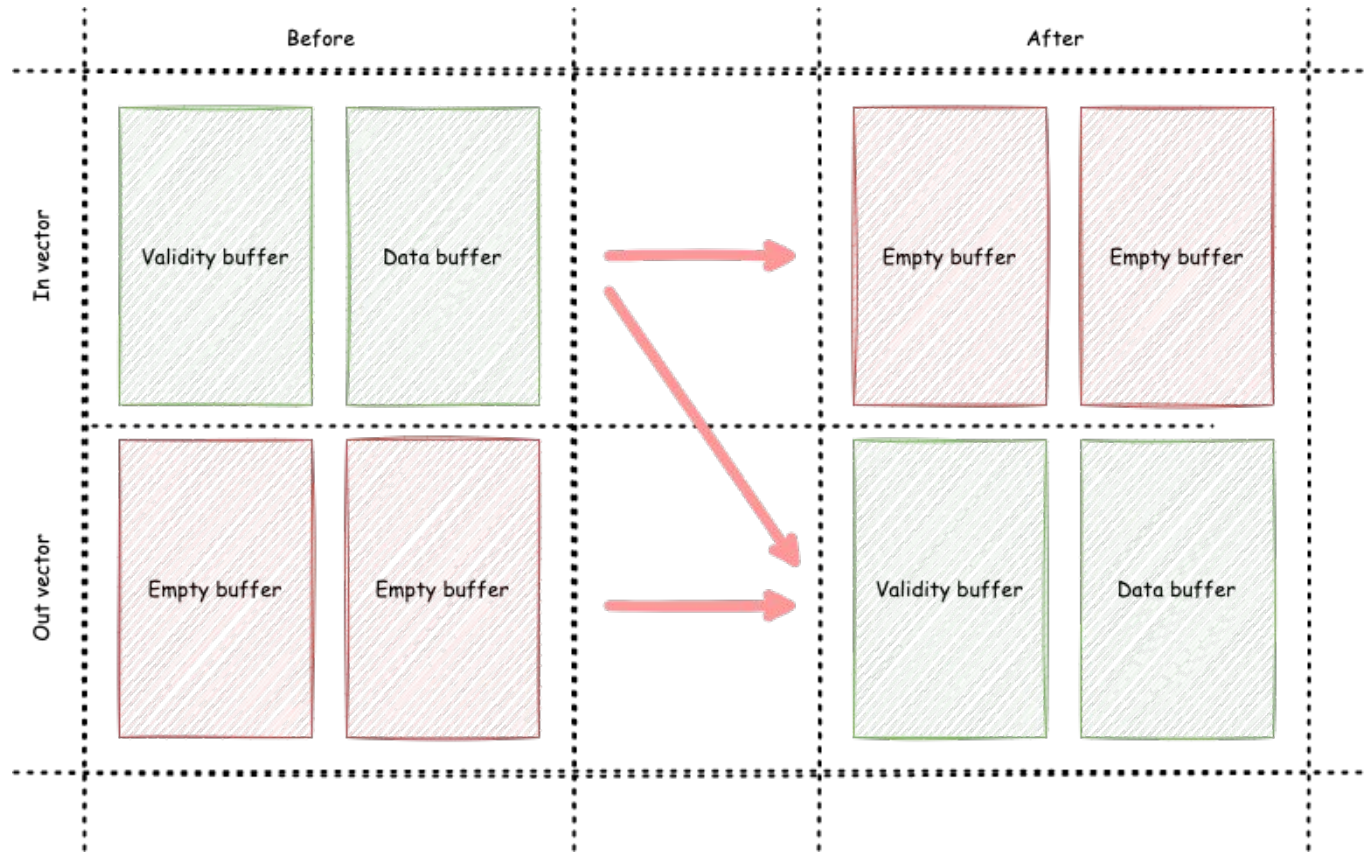
Zero-copy: transferring



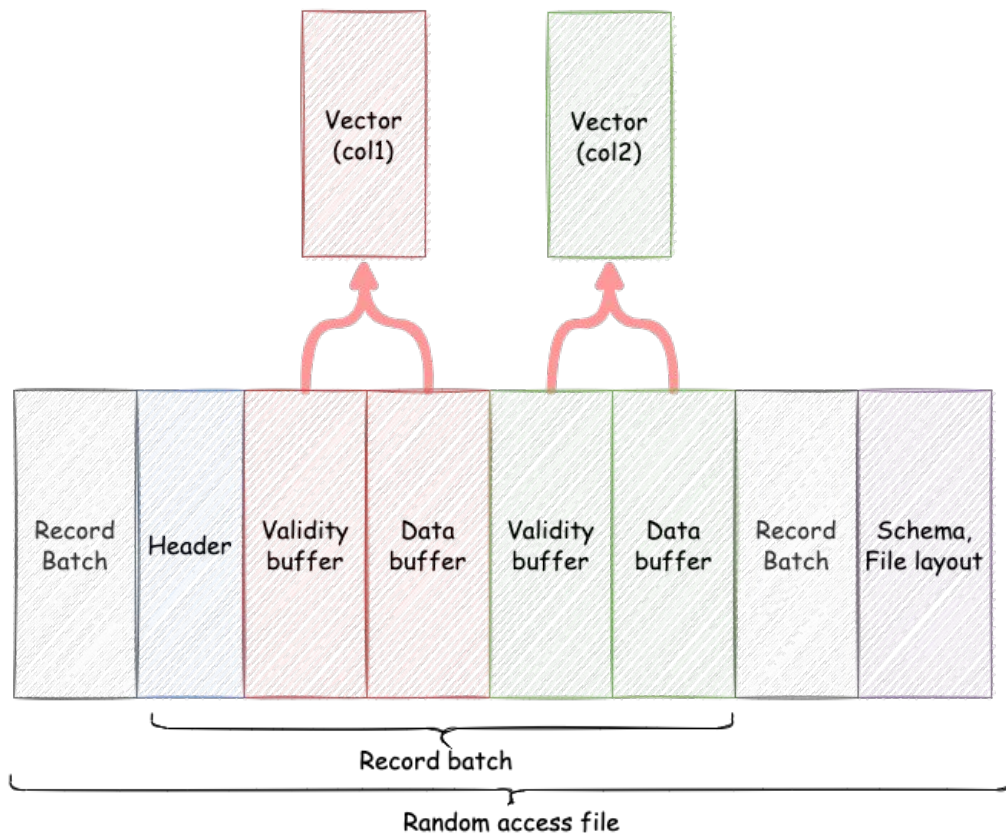
Zero-copy: transferring

```
SELECT col_1, col_2 + col_3 FROM table_1
```

Zero-copy: transferring



Zero-copy: reads



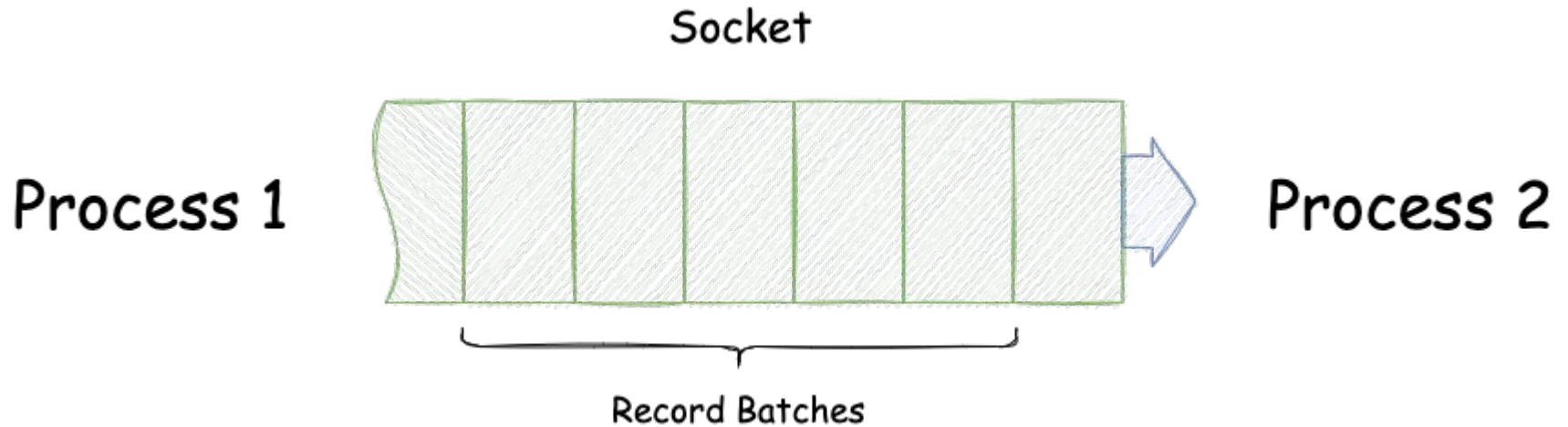
Header - содержит офсеты на буферы внутри Record Batch

Layout буфера внутри Record Batch такой же как и у вектора в памяти (No-marshalling)

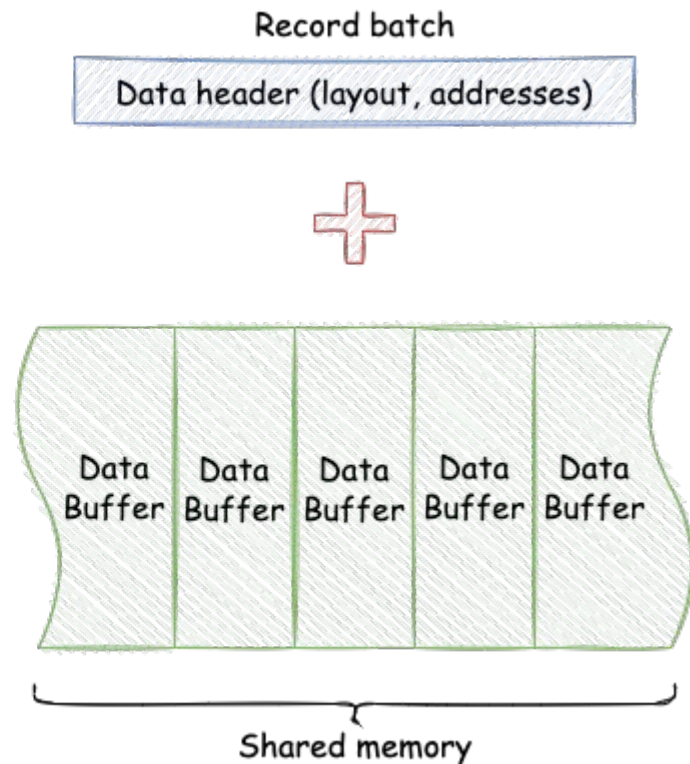
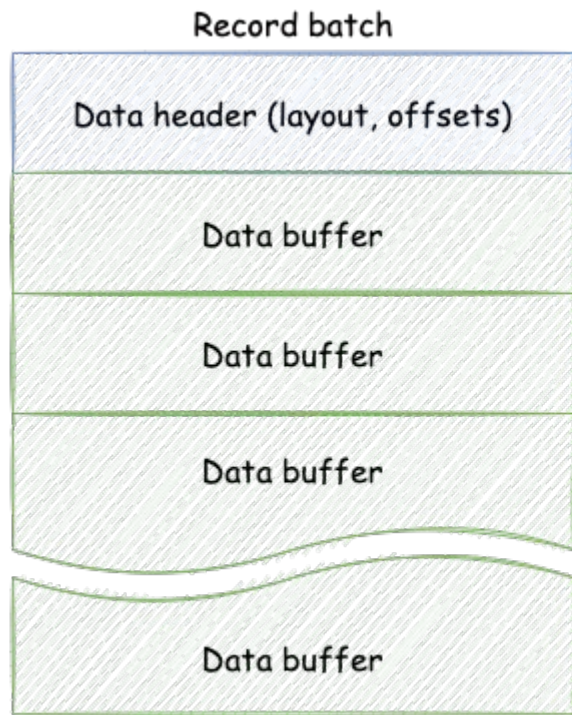
Выполнив mmap файла офсеты можно транслировать в адрес буфера в памяти

Используя адреса буферов из Record Batch можно создать вектор который "смотрит" на данные из файла

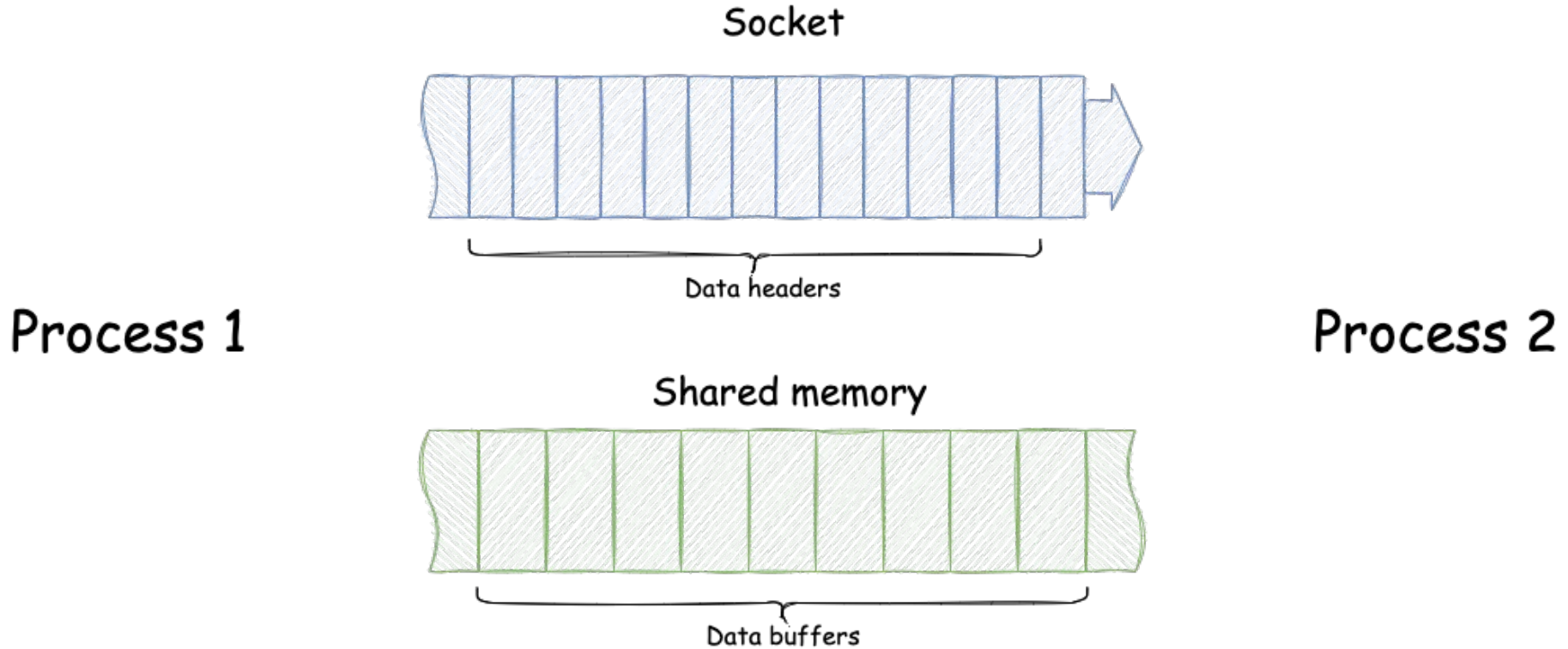
Zero-copy: IPC



Zero-copy: IPC



Zero-copy: IPC

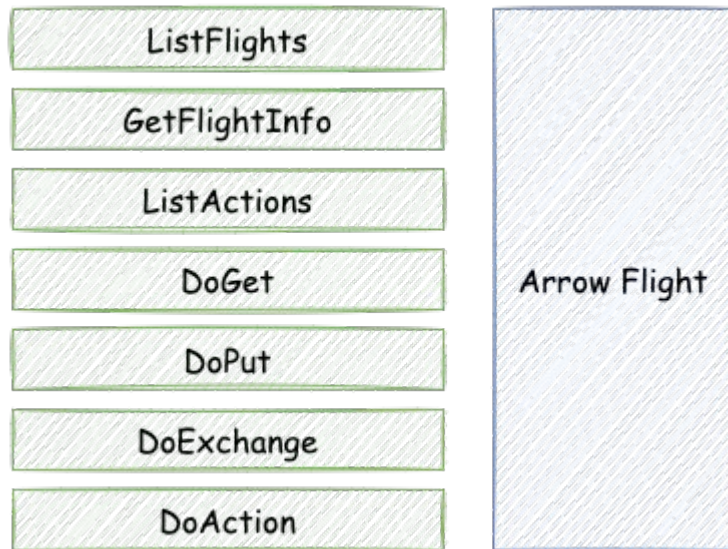


Arrow: What else?

- Arrow Flight
- Arrow Gandiva
- Arrow Plazma
- Arrow Execution Engine

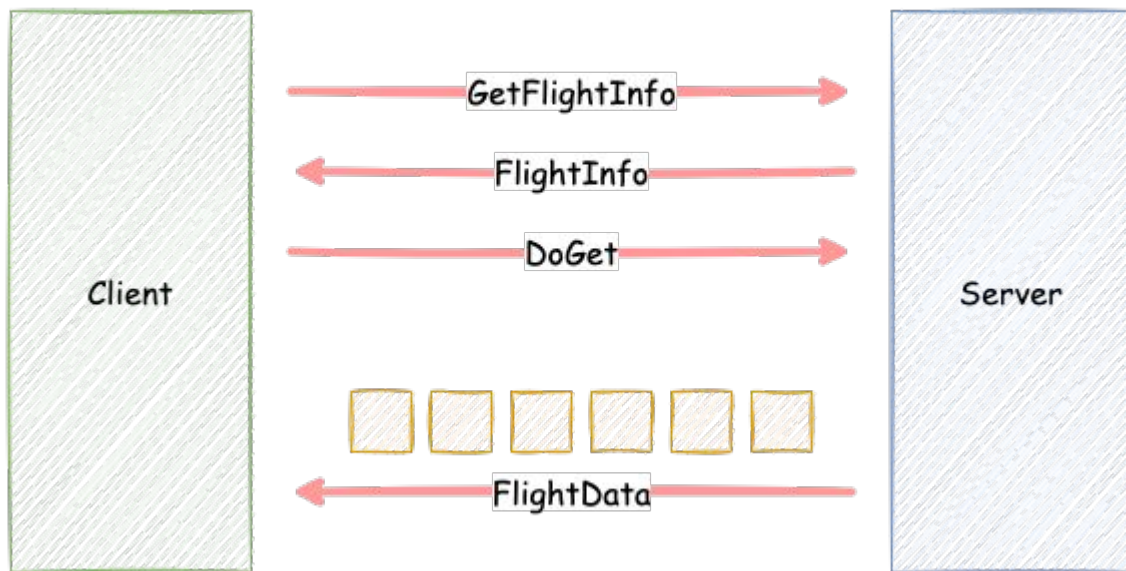
Arrow Flight

General-purpose client-server framework to simplify high performance transport of large datasets over network interfaces.



Arrow Flight

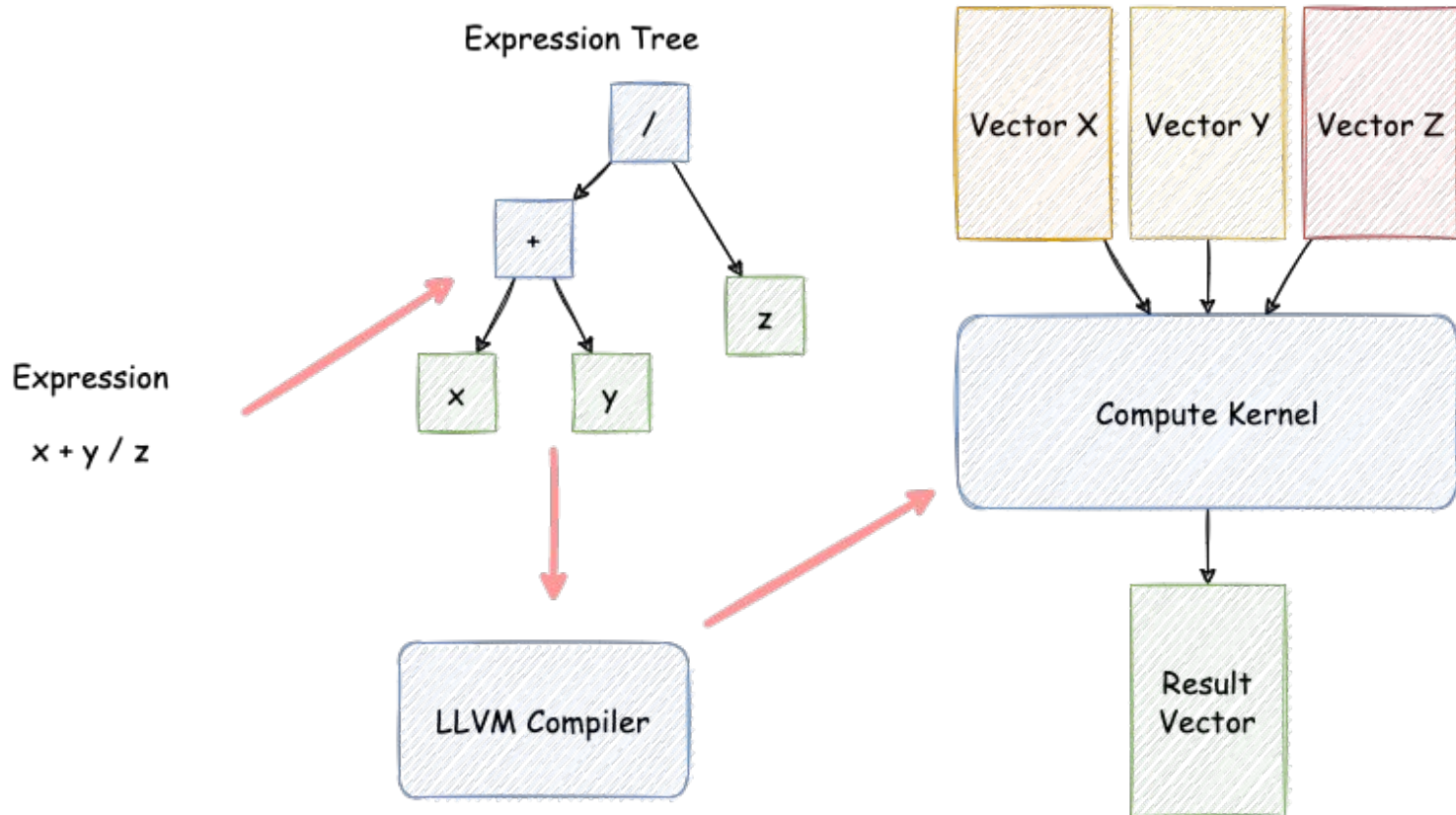
Simple Client-Server Execution Flow



Arrow Gandiva

LLVM-based Analytical Expression Compiler
for Apache Arrow

Arrow Gandiva



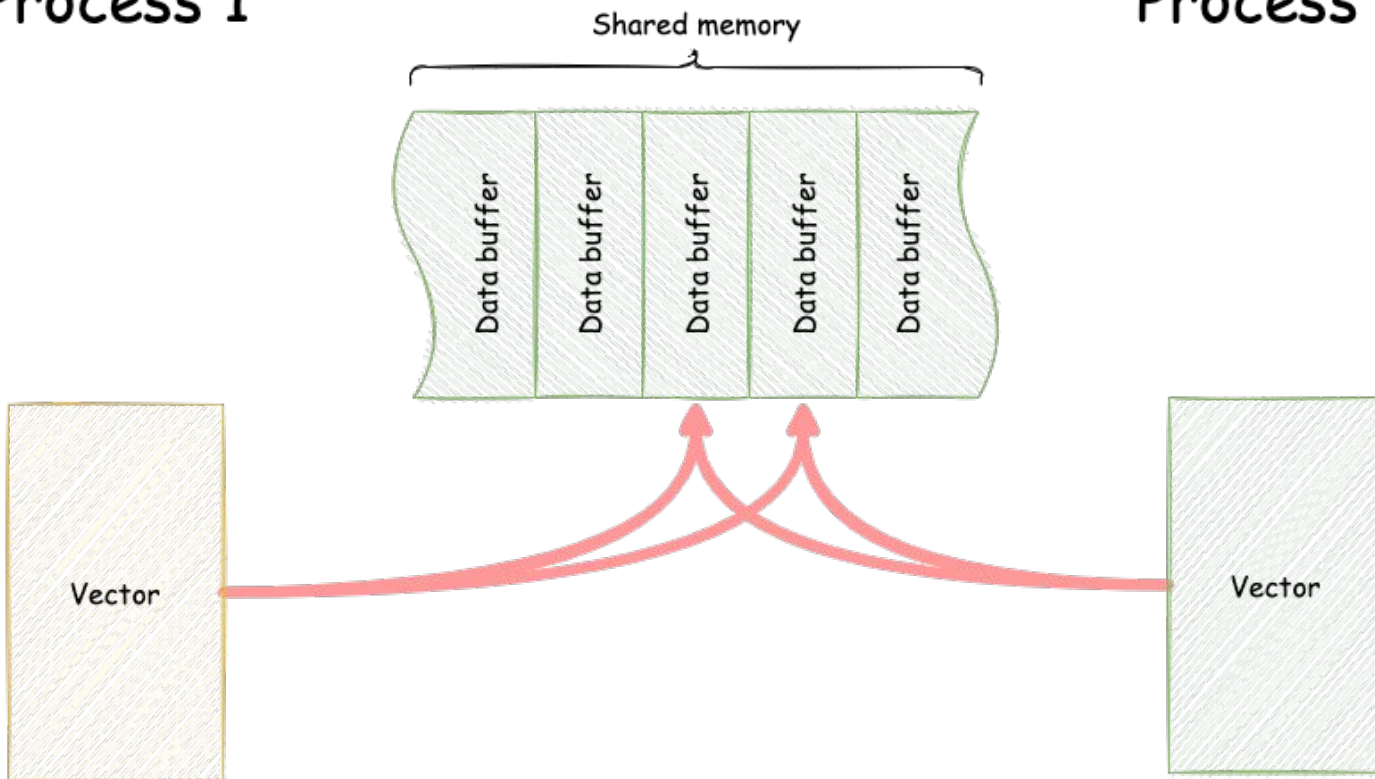
Arrow Plazma

High-Performance Shared-Memory Object Store

Arrow Plazma

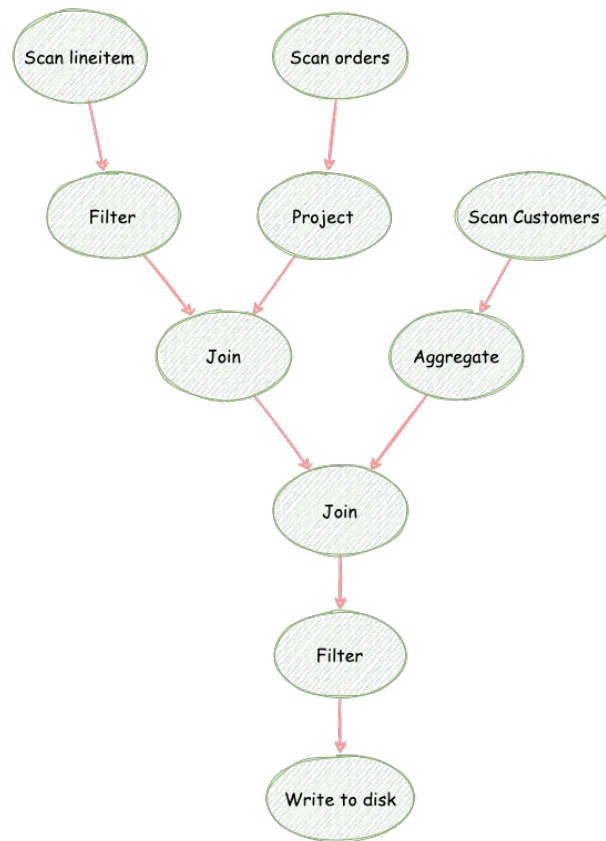
Process 1

Process 2



Arrow: Execution Engine

Arrow-native columnar streaming execution engine



Conclusion



- Идеально подходит для построения аналитических движков исполнения
- Инструменты для организации эффективной пересылки данных
- Инструменты для построения data transformation пайплайнов (ETL сценарии)
- Поддерживается во многих языках программирования (Java, C++, Python, Ruby, etc)
- Используется во многих популярных продуктах (Spark, Dremio, Cassandra, etc)
- Отличный потенциал для кастомизации под конкретные сценарии

Apache Arrow. В погоне за скоростью

Questions