

14.1 Финальный тест

Шаг 14 из 25



Когда мы дизайним ETL-процесс, имеет ли значение выбор формата хранения/пересылки данных? И если да, то как он влияет на производительность?

Выберите ответ из списка

- Различия между форматами не настолько серьезные, чтобы думать об этом при написании ETL
- Выбор формата может сильно повлиять на производительность ETL, особенно с применением компрессии on wire/at rest и колоночного хранения. Лучше выбирать что-то вроде Parquet или ORC для записи на диск
- Формат хранения и пересылки в первую очередь зависит от источника и приемника, при этом производительность больше зависит от их скорости и задержек, чем от сериализации
- Производительность ETL в первую очередь зависит от железной инфраструктуры, всегда гораздо важнее брать сеть и диски побыстрее

Отправить

Далее >

