*ACID (Atomicity, Consistency, Isolation, Durability) — описание требований к СУБД, системе управления базами данных:****атомарность, согласованность, изолированность и прочность.***

В реляционных базах данных все связанные данные хранятся в одном узле или на одной машине.

Большинство инструментов NoSQL ослабляют критерии согласованности операций для обеспечения отказоустойчивости и масштабирования.

**Теорема CAP** (известная также как **теорема Брюера)** — [эвристическое](https://ru.wikipedia.org/wiki/%D0%AD%D0%B2%D1%80%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0) утверждение о том, что в любой реализации [распределённых вычислений](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D1%91%D0%BD%D0%BD%D1%8B%D0%B5_%D0%B2%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B5%D0%BD%D0%B8%D1%8F) возможно обеспечить не более двух из трёх следующих свойств:

* [*согласованность данных*](https://ru.wikipedia.org/wiki/%D0%A1%D0%BE%D0%B3%D0%BB%D0%B0%D1%81%D0%BE%D0%B2%D0%B0%D0%BD%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85) ([англ.](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) *consistency*) — во всех вычислительных узлах в один момент времени данные не противоречат друг другу;
* [*доступность*](https://ru.wikipedia.org/wiki/%D0%94%D0%BE%D1%81%D1%82%D1%83%D0%BF%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%B8) ([англ.](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) *availability*) — любой запрос к распределённой системе завершается корректным откликом, однако без гарантии, что ответы всех узлов системы совпадают;
* *устойчивость к разделению* ([англ.](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) *partition tolerance*) — расщепление распределённой системы на несколько изолированных секций не приводит к некорректности отклика от каждой из секций.

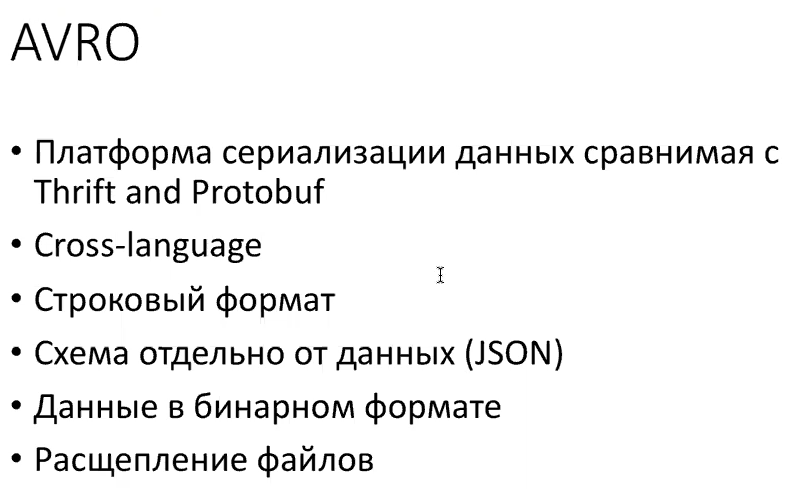
**Сериализация** (в программировании) — процесс перевода структуры **данных** в последовательность байтов. Обратной к операции **сериализации** является операция **десериализации** (структуризации) — создание структуры **данных** из битовой последовательности. **Сериализация** используется для передачи объектов по сети и для сохранения их в файлы. Например, нужно создать распределённое приложение, разные части которого должны обмениваться **данными** со сложной структурой.

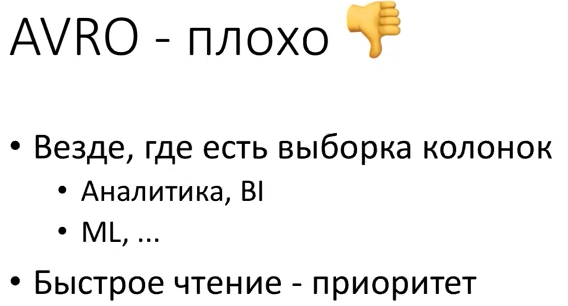
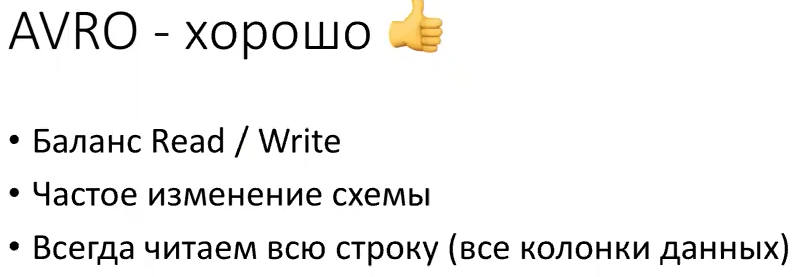
Команды hive, python, spark-submit – запускают исполняемые файлы

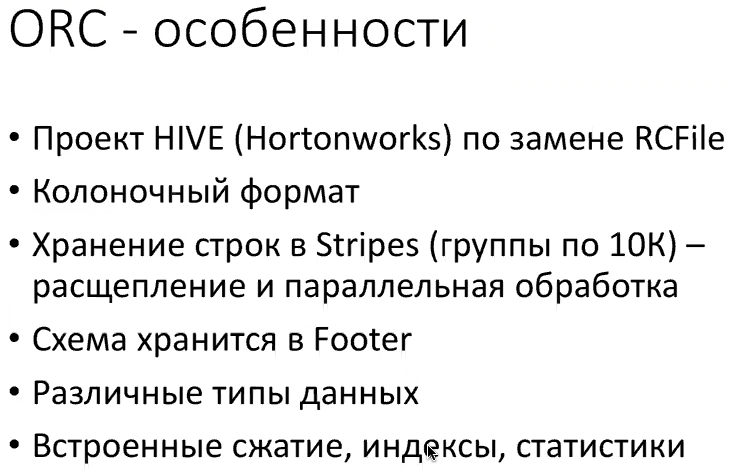
В linux 3 вида переменных окружения:

1. Локальные
2. Пользовательские
3. Системные

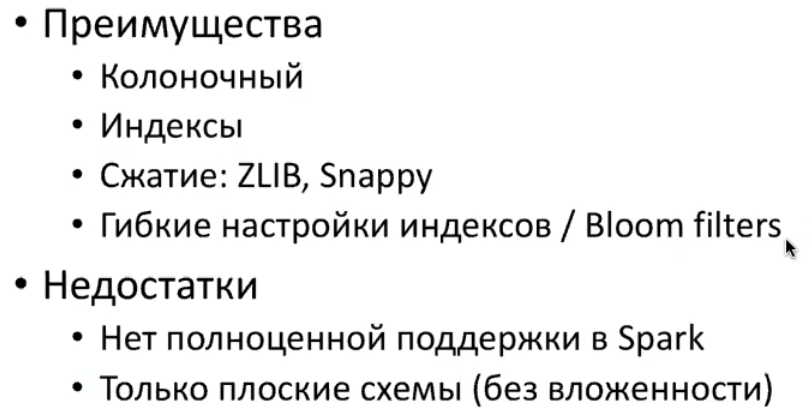
Расщепление – возможность распаралелить чтение, передачу данных

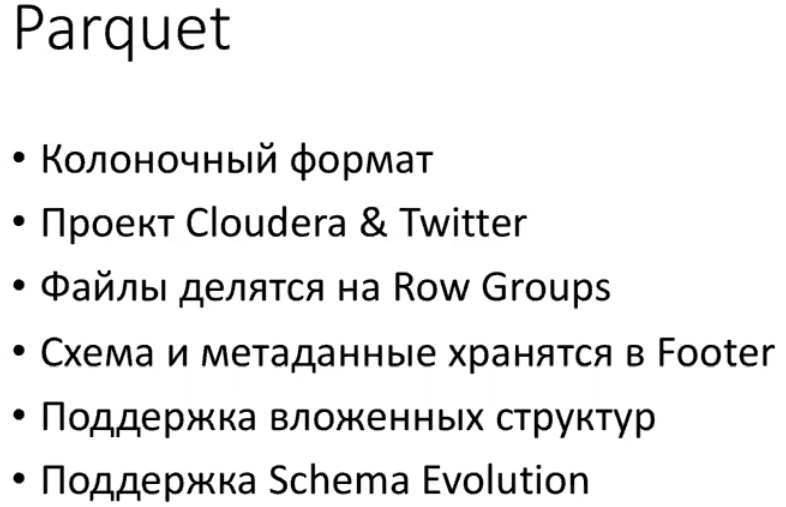












//Row groups похоже на stripes(ORC)



