***Литература:***

[*https://oracle-patches.com/db/mysql/3814-tranzaktsii-v-baze-dannyh-mysql*](https://oracle-patches.com/db/mysql/3814-tranzaktsii-v-baze-dannyh-mysql) *(или любая другая статья по теме «уровни изоляции транзакций» и «ACID»)*

[*https://zen.yandex.ru/media/id/5aef49c279885e47d5eb6199/o-bazah-dannyh-prosto-razlichiia-olap-i-oltp-5bf5557e5184cc00a99028ff*](https://zen.yandex.ru/media/id/5aef49c279885e47d5eb6199/o-bazah-dannyh-prosto-razlichiia-olap-i-oltp-5bf5557e5184cc00a99028ff) *(или любая другая статья по теме «OLAP и OLTP»)*

**Задания:**

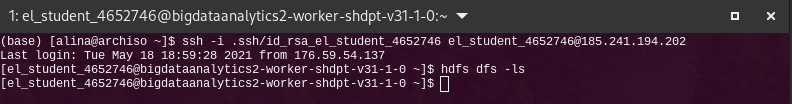
***Порассуждать после прочтения литературы: чем грозит отсутствие ACID?***

Если БД не предоставляет гарантий ACID, то их стоило бы реализовать в самом приложении, ведь если транзакция не атомарна, то высок риск выпадения в race condition вследствие попыток одновременного доступа. Я так поняла, примерно тем же опасны и несогласованность, и неизолированность процессов, если система будет использовать промежуточные результаты для параллельных операций. Если база не гарантирует долговечность/надёжность, возможно, это значило бы отсутствие резервных копий данных/метаданных, и при сбое системы/оборудования выполнение транзакции могло быть прервано, вследствие чего очень вероятно нарушение консистентности базы. Однако, в статье <https://habr.com/ru/post/535616/> приводятся и NoSQL БД, которые могут существовать и приносить пользу и без принципов ACID, предлагая «согласованность в конечном счете» или eventual consistency (из BASE). Надеюсь услышать по этому поводу какие-то пояснения человека из среды, насколько это нужная технология, используют ли её и сильно ли на деле страдает надёжность транзакций.

***Учитывая информацию о OLAP и OLTP, для каких бизнес задач подходит Hadoop, для каких нет и почему?***

Hadoop, как инструмент анализа больших данных, разумно было бы использовать для маркетинговых типов бизнес-задач, когда нужно проанализировать поведение пользователя и сделать какие-либо материально-значимые/конвертируемые выводы. Я в бизнесе полный ноль, но гугл подсказывает, что моделирование (прогнозирование) отлично подходит для эффективного планирования и учёта, контроля и управления бизнес-процессами, что звучит логично. Сюда же можно отнести управление производством/закупочным сектором, где построение модели на уже имеющихся данных предоставит возможность более грамотно распределять ресурсы в будущем. Одёрните, если что не так.  
Очевидно, hadoop никак не подойдет для тех задач, которые подразумевают insert’ы и update’ы в БД, онлайн-обработку транзакций.

***Получить у меня доступы и зайти на кластер. Выполнить hdfs dfs -ls***



*Инструкция по доступу в Ambari на учебном кластере.*

*1. Пройти по ссылке http://37.139.32.56:8080/#/login*

*2. Вести имя/пароль student/v3rVyK8AVgAgWfy*

*3. Наслаждаться непонятными графиками :)*

* наслаждаюсь изо всех сил :)

