

VOLTDDB Do's & Don'ts

The following are a few tips to make your work with VoltDB successful.

Do's

1. **Partition your tables.** Partitioning allows you to dynamically scale the size of your database. It also enables simultaneous processing of transactions (see tip #2).
2. **Partition your stored procedures.** To maximize throughput, maximize the frequency of single-partition transactions and minimize multi-partition transactions.
3. **Connect to all nodes of the cluster.** Creating multiple client connections avoids possible bottlenecks and allows VoltDB to take advantage of client affinity (in the Java client).
4. **Use K-Safety.** VoltDB protects against data loss due to network or node failures through internal replication (called K-Safety). With K-Safety, if a node goes down, the database continues to run because copies of the data still exist. Note, however, that when using K-Safety, we recommend using three or more nodes. Because with only two servers, if a node fails the cluster cannot distinguish between a node failure and a network partition, and the remaining node may stop itself to avoid a partition.

Don'ts

1. **Don't return excessive amounts of data from stored procedures.** Too large result sets clog the network and impact latency. VoltDB limits the procedure results to 50MB. Keep your results small to maximize throughput and minimize latency.
2. **Don't create huge tables.** VoltDB works best with small to medium-sized tables. Creating tables with hundreds of columns or many large columns can impact query performance. Create multiple smaller tables instead. VoltDB limits columns to 1MB each and all of the columns in a single table to 2MB.
3. **Don't do unnecessary processing in your stored procedures or callback routines.** Keep your procedures neat and clean to avoid long-running procedures blocking other transactions. Also, callback routines are processed one at a time within the client, so make the callback procedures efficient to avoid stalling the client after the transaction completes.
4. **Don't perform non-deterministic operations in stored procedures.** VoltDB requires deterministic results to guarantee consistency across partitions and during recovery. Avoid any operations in stored procedures that give variable results such as system time, random numbers, file I/O, etc. Instead, generate the variable data in the client and use it as input to the procedure.