

Managing Data Distribution

This section describes some of the advanced controls that are available for managing data distribution within Xpand.

Key Terms Used in this Section:

Relation - Each table in Xpand is referred to as a "relation".

Representation - Each index is called a "Representation" in Xpand. Table data is stored in the "Base Representation", the relational representation indexed by an internal key that covers all of the columns of the base table. For tables keyed by a primary key, the data for the "Base Representation" is stored with the primary key.

Distribution Key - Each representation has all or a portion of its index hashed using a consistent hashing algorithm. A "distribution key" defines which columns of an index are used to construct that hash. The default distribution for indexes is 1, meaning the first column of an representation (index) will be hashed and become the distribution key for that representation.

Slices - Xpand breaks each representation into smaller, more manageable segments called "slices". Slices are then distributed throughout the cluster to facilitate evenly distributed query processing.

Replicas - Xpand maintains multiple copies of each slice of data to provide fault tolerance and high-availability. Replicas are distributed throughout the cluster to optimize performance and to ensure all data is protected in the event of a node failure.

The default behavior for distribution, slices, and replicas is optimal for the majority of workloads. For more information please refer to the following:

[DISTRIBUTE](#)

[SLICES](#)

[REPLICAS](#)

On Xpand, the SHOW CREATE TABLE command will display information for key distribution, slices, and replicas defined for a table.

[Xpand Support](#) is available for recommendations regarding fine-tuning data distribution strategies.