rows are stored across hosts

Node - maps to a host

vnode

A node may define many vnodes.

VNodes provide a layer of indirection when assigning responsibility for token ranges to a node.

Mechanism for breaking up the token ring.

Alternative for manually assigning token ranges to nodes

The hosts in a cluster can be considered a ring

Each row is stored on a single host (but replicated on multiple hosts)

Tables are stored in order

Makes data access efficient when queries span rows

Supports sequential reads

Each node in a ring is assigned a token on startup

Rows are distributed (partitioned)

Rowkey - another name for primary key

Tables are a collection of columns.

Partition - A segment of the token ring range

Token ring - set of all tokens to which the primary key of a row may be hashed to. The ring is broken up into partitions, with each partition assigned to a node of the cluster

Cassandra hashes the rowkey into a token

Each token gets compared to the partition ranges that make up the ring in order to determine where a row is stored.

Partition key

The row key / primary key

Main purpose is to evenly distribute data across a cluster

Primary key -

Another name for the partition key

Or the row key

Non primary key columns are stored along with the primary key columns in a particular partition.

Non-primary key columns can be grouped together within a partition by specifying a clustering key.

Within a cluster, rows are ordered based on clustering column. Clustering order can be manually specified using CLUSTERING ORDER clause

Storage mechanisms force cassandra to require that queries restrict on no columns, partitioning columns or clustering columns if ALLOW FILTERING is specified.

ALLOW FILTERING is a hint that the user understands that the query could be inefficient

Secondary indexes allow restricting based on non partition / clustering columns.

Querying by only secondary index requires all nodes to be contacted

Including partition key restrictions will result in more efficient access, only contacting nodes matching the partition-key restrictions

Compaction