

DataPath – Specification of components and messages sent between them

June 2, 2009

For each component the name, purpose and interaction with other components.

1 Coordinator

This component provides the global coordination of the system in terms of injecting and removing queries.

Single threaded component since it does not have too much work to do. Since most of the actions are serial anyway (to keep things simple) there is no reason to have a multi-threaded coordinator.

The biggest decision to make is when to remove the queries from the system Should the removal be delayed until new queries are inserted or a separate request for removal is better.

1.1 Messages Received

CCNewQuery Message sent form the UI with an new query specification. The specification is in SQL or another language the Quey Planner understands. The request has to be placed in a queue until the query finishes (should contain the place where the answer is sent).

NewQueryPlan Message from the Query Planner with information on how to insert the new query. The message contains three parts.

1. new connectivity graph (waypoint pairs)
2. reconfiguration information for each waypoint. Specifies also new waypoints created and old waypoints to be deleted.

3. reconfiguration information for each file scanner. Specifies what columns need to be read for each new query and where they have to be placed into the chunk.
4. set of new queries (container of QueryIDs) together with the query masks (map from QueryID to mask, i.e. BITSET)

NewQueryCode Message from the Code Generator containing the new code that the waypoints will execute.

QueryTerminationInfo Message sent by each File Scanner with an object that can track query termination. This object will be passed to the query execution engine together with the other information needed.

CCStartNewQuery Message from the execution engine that it is ready for Chunks from the new query.

QueryFinished Message from the execution engine that a query has finished.

2 Messages Sent

UIQueryCompleted Message sent to the UI when the query is completely executed. Will contain information indicating where the result is.

QPNewQuery Message sent to the Query Planner to indicate that action is needed. The message will contain a new query that has to be inserted and possible old queries to be removed.

CGNewQuery Message sent to the Code Generator to produce the code the waypoints will execute in the future. Only the information about Waypoints from the message received from the Query Planner will be sent (not the graph or the file scanner config).

FSNewQuery Tell the File Scanners about the new query. This will get them ready for a latter message specifying they can produce Chunks for the new query. The message will specify what waypoints need query termination for that queries.

EENewQuery Tell the execution engine to inject a new query. This message contains information about:

1. Code for the waypoints together with list of deleted waypoints (Map from WaypointID to configObject, and collection of waypiontID for deleted waypoints).
2. New graph (Graph Object)
3. Query termination objects for the waypoints that need them (map from WaypointID to object)
4. new queries inserted (Collection of QueryID) together with their masks (map from QueryID to mask).

FSStartNewQuery Message to the File Scanners that they can produce chunks for the new query. At this moment a new query can be considered. All queries arriving until this point should be buffered.

3 Query Planner

Does the planning (insertion and removal of queries).

4 Messages Received

QPNewQuery from the Coordinator with the new query.

5 Messages Sent