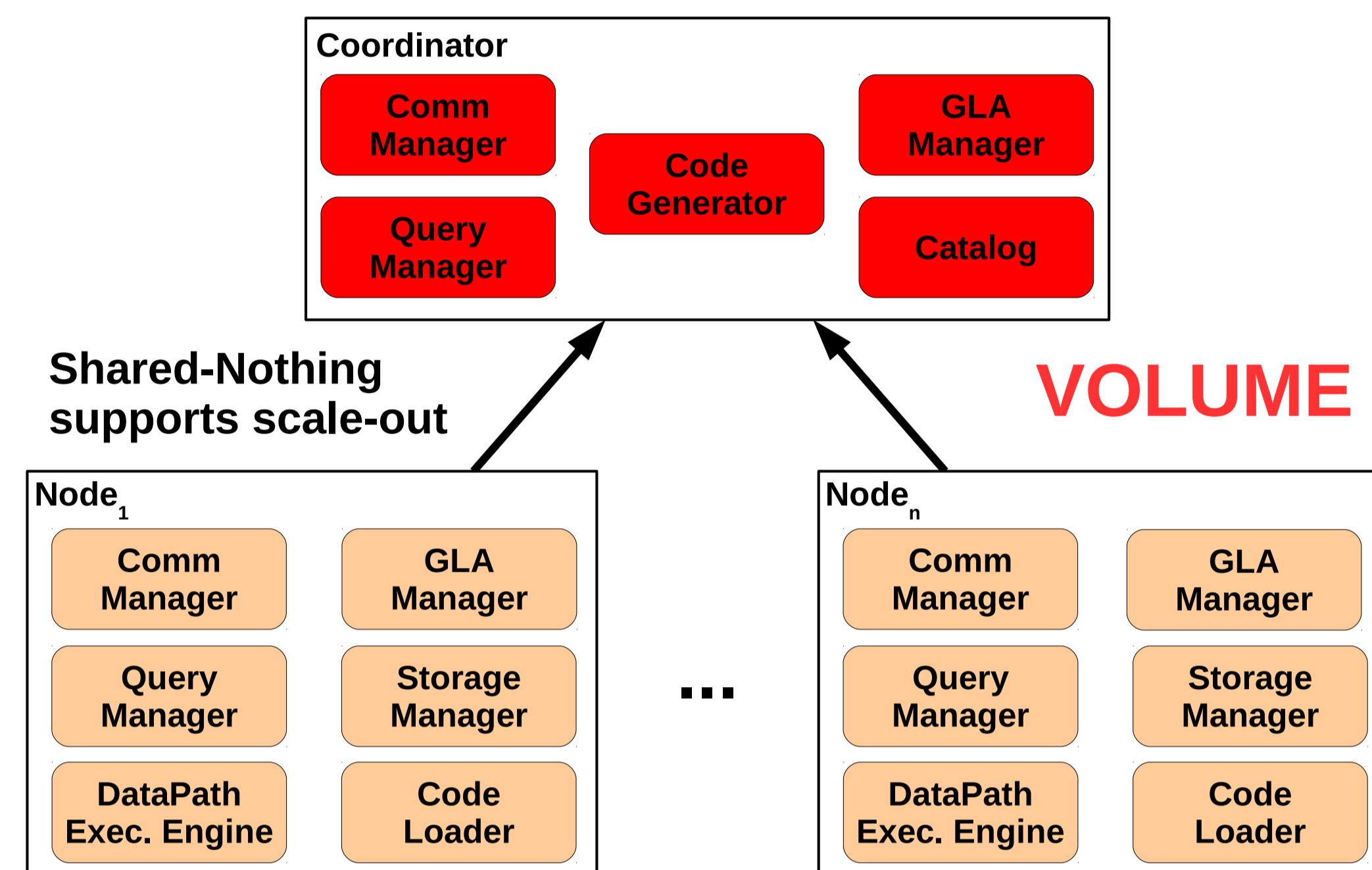


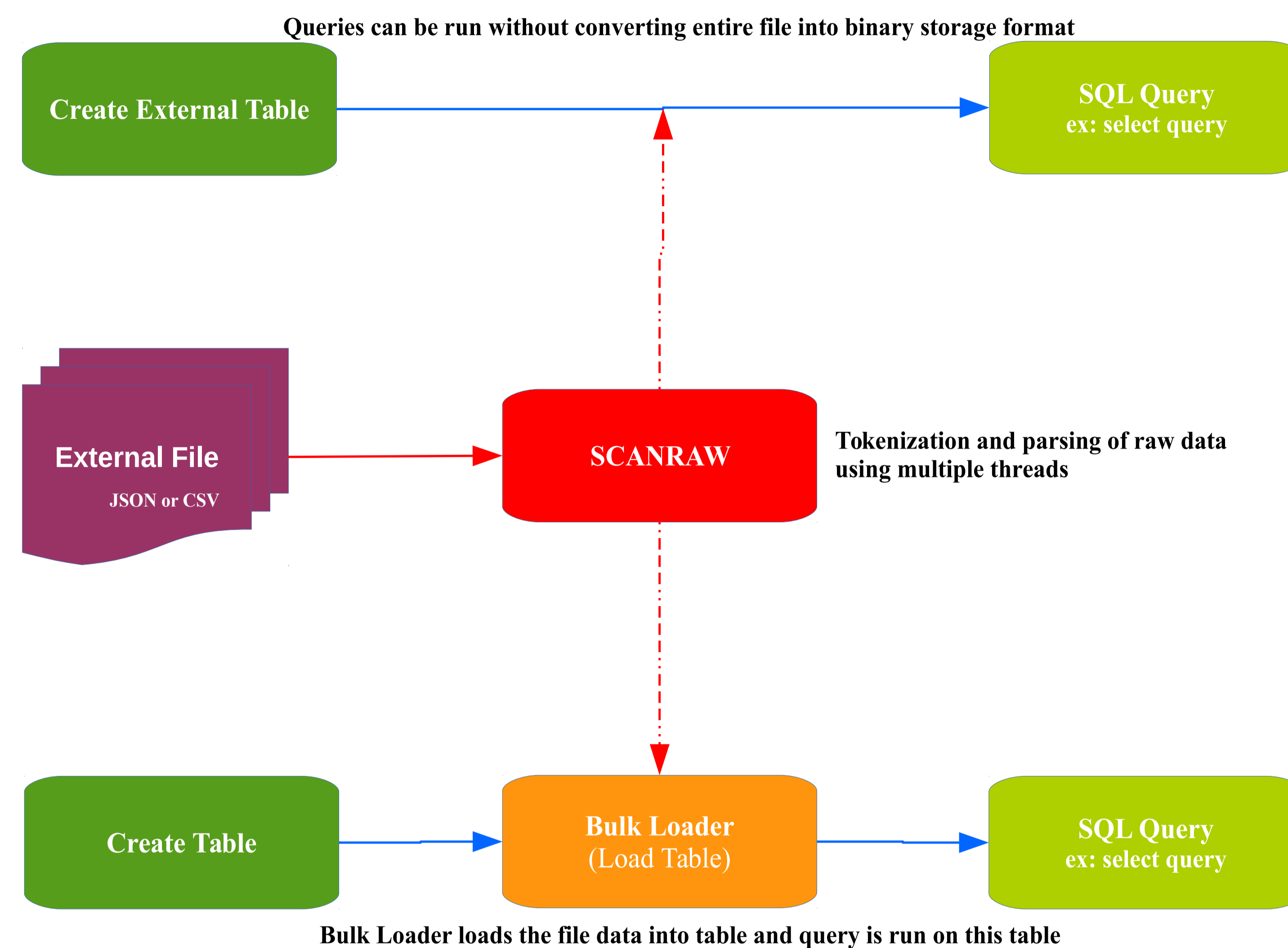
System Architecture

GLADE is a parallel Big Data processing system



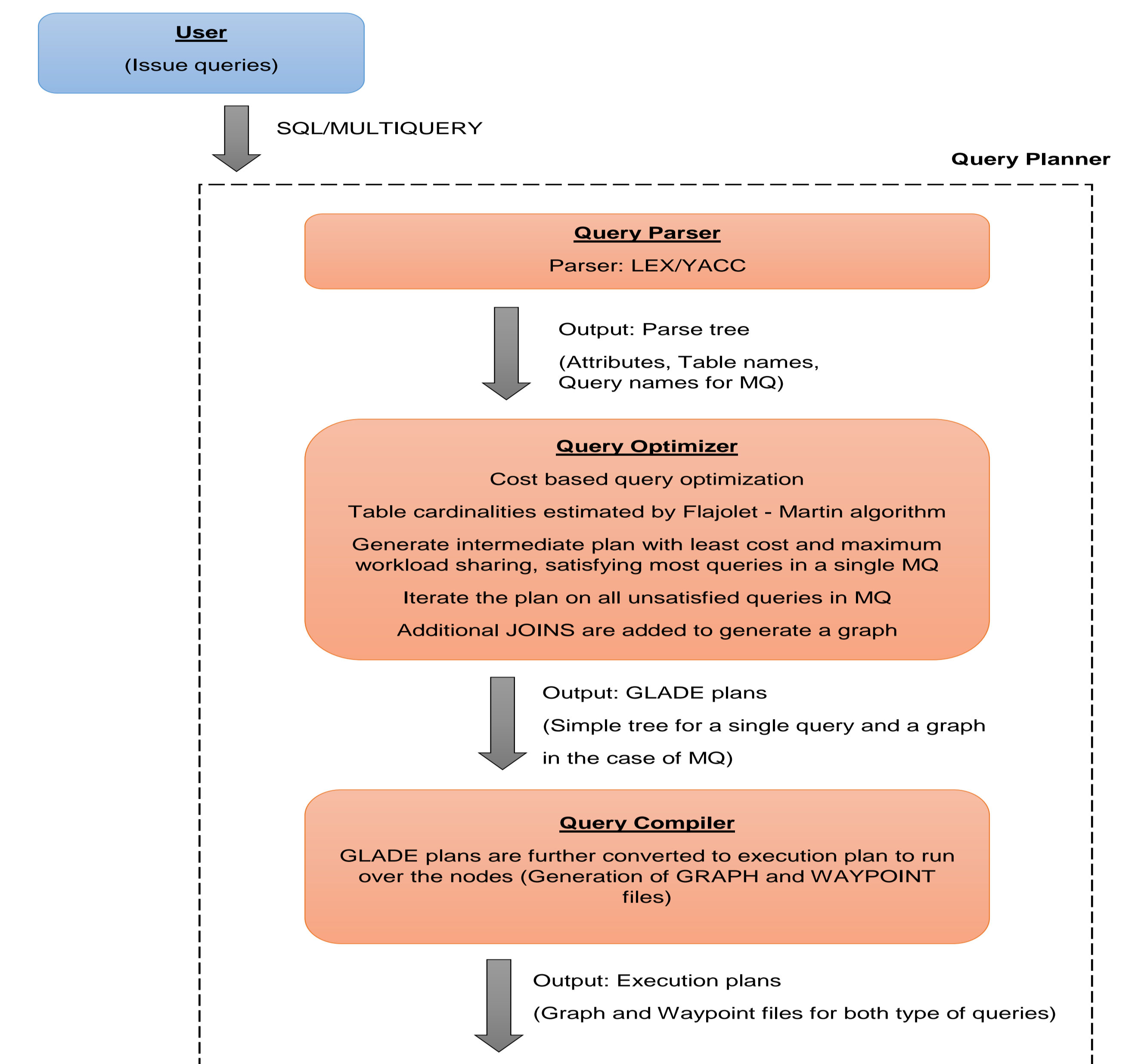
- Massive heterogeneous data: data partitioning; parallel execution; relational and array data model
- Extensible complex analytics: user code executed inside the engine; enhanced UDA interface
- Architecture independence: multi-thread (shared memory and shared disk) and inter-node (shared nothing) parallelism

External File Support



Multi-Query Optimization

Structure of Query Planner



MULTIQUERY Example

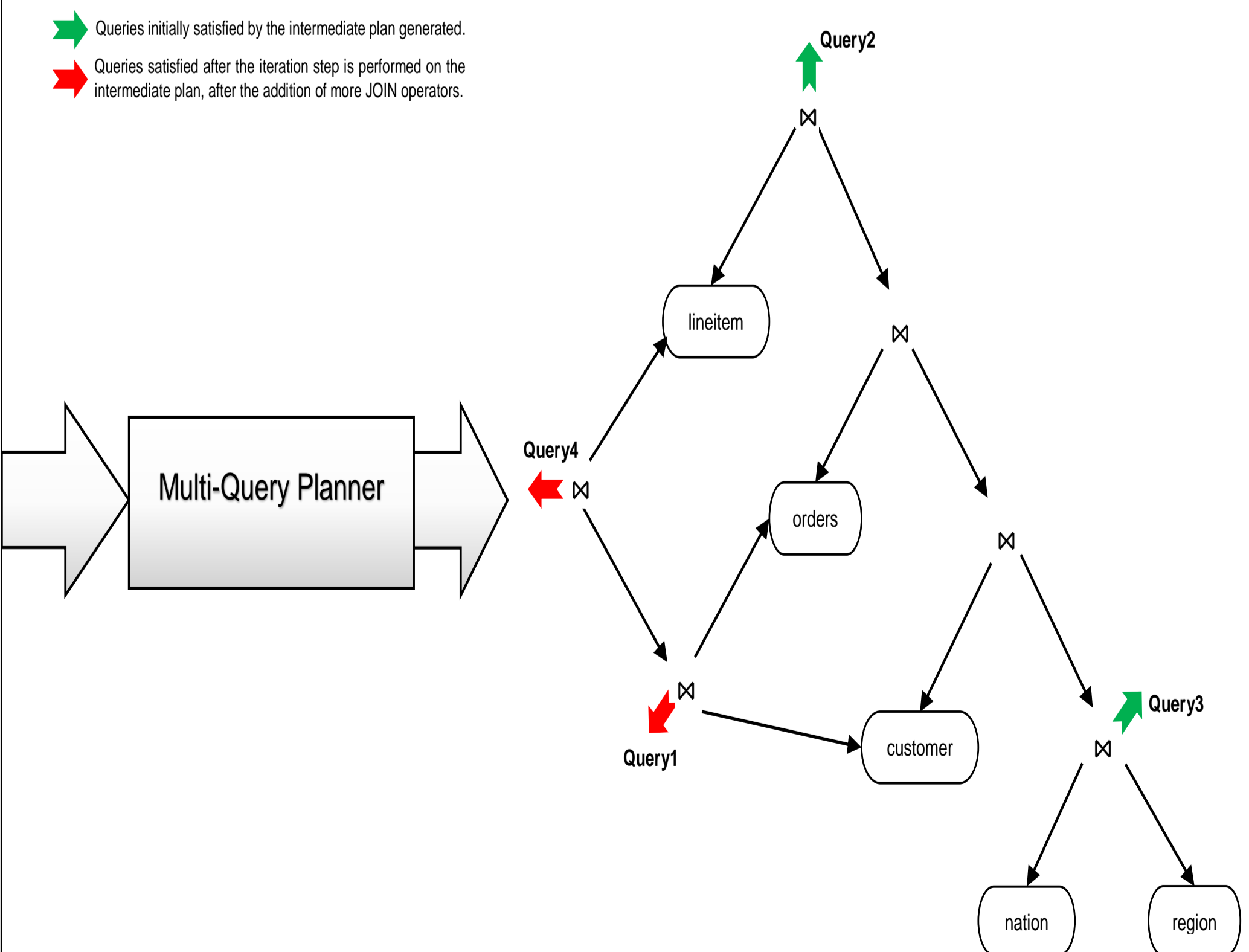
```

MULTIQUERY
Query1:
SELECT c_acctbal
FROM customer, orders
WHERE c_custkey = o_custkey AND
o_totalprice < 10000

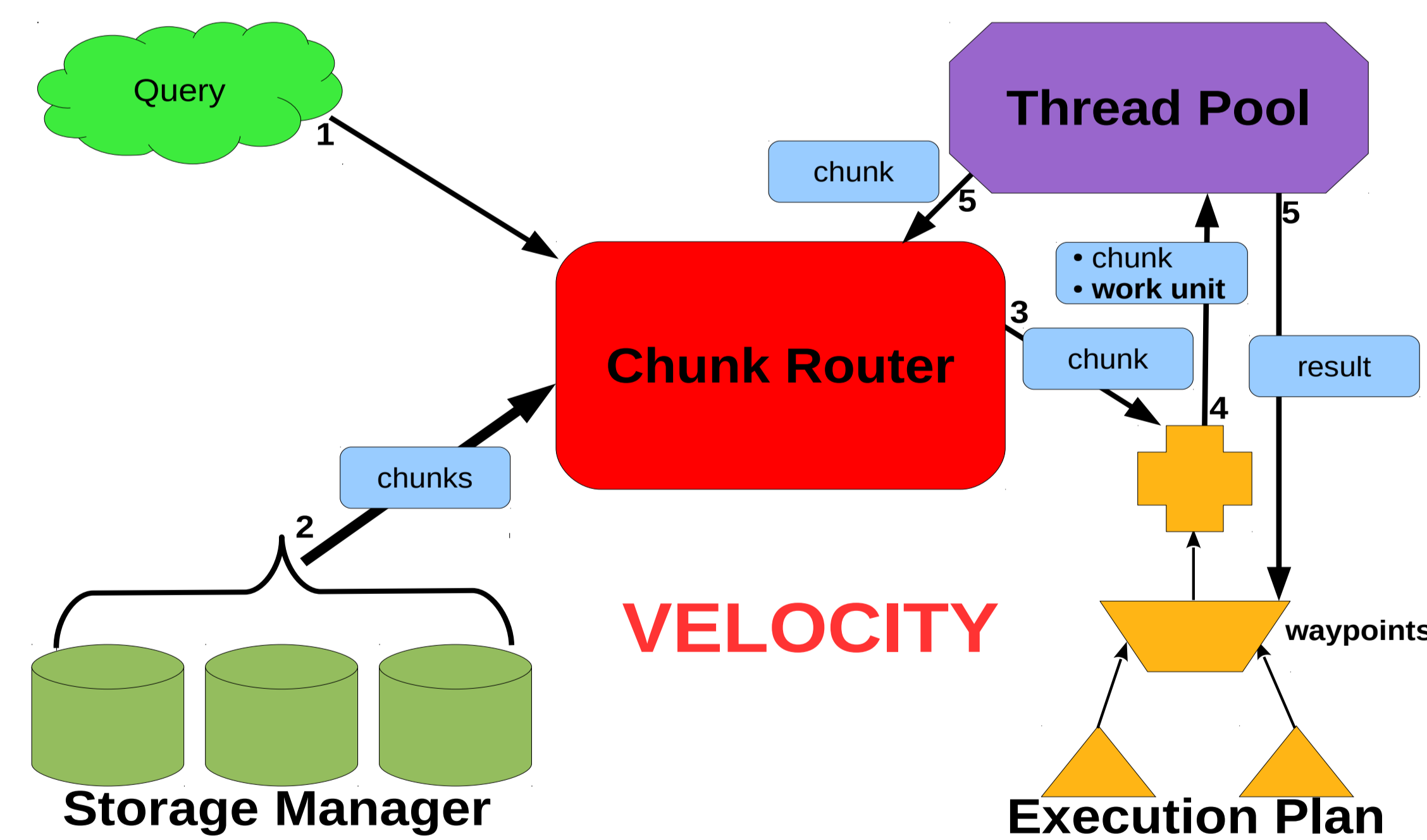
Query2:
SELECT c_name
FROM lineitem, orders, customer, nation, region
WHERE l_orderkey = o_orderkey AND
o_custkey = c_custkey AND
c_nationkey = n_nationkey AND
n_regionkey = r_regionkey AND
r_regionkey = 1 AND o_orderkey < 10000

Query3:
SELECT n_name
FROM nation, region
WHERE n_regionkey = r_regionkey

Query4:
SELECT l_discount
FROM customer, orders, lineitem
WHERE c_custkey = o_custkey AND
o_orderkey = l_orderkey AND
c_name = 'Customer#000070919' AND
l_quantity > 30 AND l_discount < 0.03
END
  
```

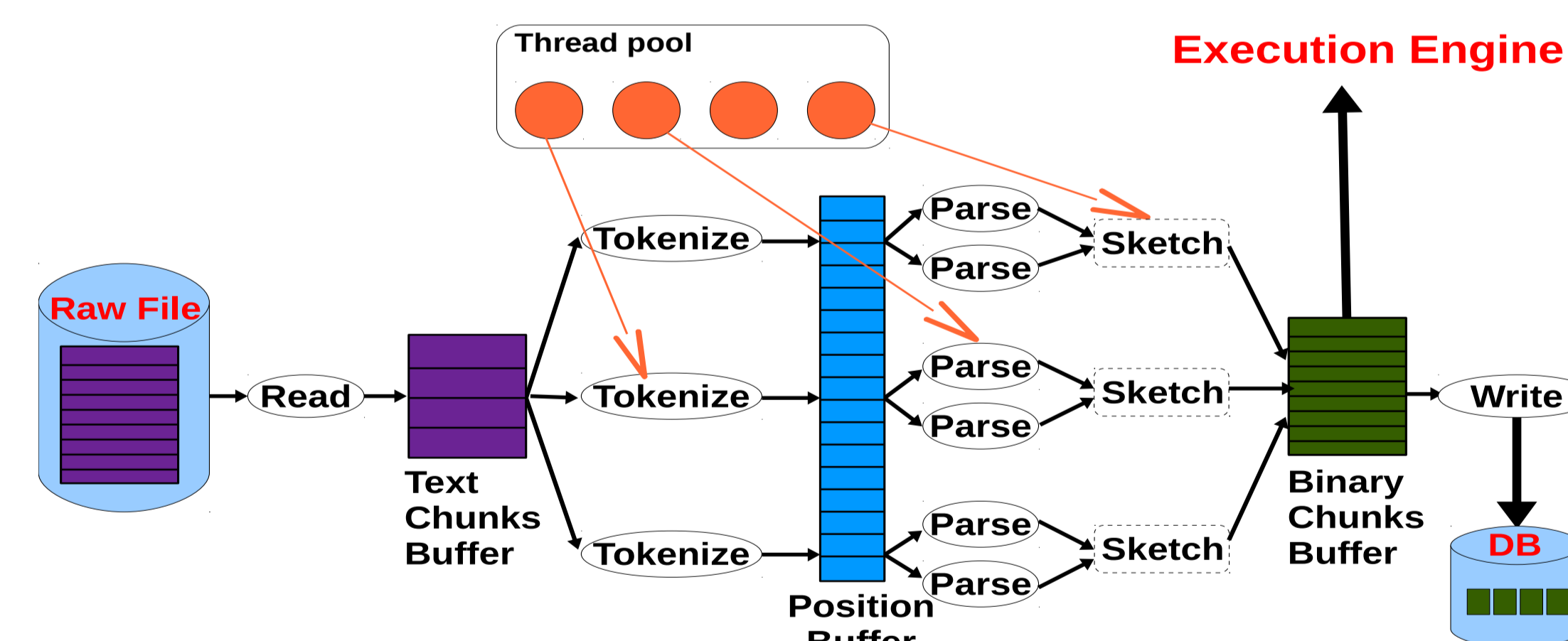


Execution Engine



- The storage manager is responsible for organizing data, reading, and delivering the data to the execution engine.
- Execution Engine implements a series of relational operators like SELECT, PROJECT, JOIN, AGGREGATE and a special operator for the execution of arbitrary user code specified using the GLA interface.

SCANRAW Architecture



Install Script

