

CLAIMS

1. A system, comprising:
 - a memory to store program instructions which, if performed by at least one processor, cause the at least one processor to perform a method to at least:
 - receive a schema that describes one or more data sets in one or more remote data stores;
 - store the schema separately from a pool of computing resources configured to execute queries;
 - receive a first query directed to the one or more data sets in the one or more remote data stores; wherein the one or more data sets are stored separately from the pool of computing resources configured to execute queries;
 - in response to the receipt of the first query:
 - provision one or more computing resources to execute the first query from the pool of computing resources configured to execute queries;
 - route the first query to the provisioned computing resources to execute the first query with respect to the separately stored one or more data sets in the remote data stores, wherein the provisioned computing resources obtain the separately stored schema that describes the one or more data sets as part of executing the first query; and
 - provide a result for the first query generated at the provisioned computing resources.
2. The system of claim 1, wherein the schema that describes the one or more data sets is stored in a metadata store, and wherein the provisioned computing resources access the schema in the metadata store to obtain the schema.
3. The system of claim 2, wherein the first query and the schema are received via a same user interface.
4. The system of claim 3, wherein the at least one processor is implemented as part of a network-based managed query service, wherein at least one of the data sets is stored in a

network-based data storage service, wherein the schema is stored in a network-based data catalog service, wherein the provisioned computing resources are implemented as part of a virtual compute service, and wherein the managed query service, data storage service, data catalog service and virtual compute service are offered as part of a same provider network.

5. A method, comprising:

receiving a schema that describes one or more data sets in one or more remote data stores;

storing the schema separately from a pool of computing resources configured to execute queries;

receiving a first query directed to one or more data sets in the one or more remote data stores; wherein the one or more data sets are stored separately from the pool of computing resources configured to execute queries;

in response to receiving the first query:

provisioning one or more computing resources to execute the first query from the pool of computing resources configured to execute queries; and

routing the first query to the provisioned computing resources to execute the first query with respect to the separately stored one or more data sets in the remote data stores, wherein the provisioned computing resources obtain the separately stored schema that describes the one or more data sets as part of executing the first query.

6. The method of claim 5, wherein the provisioning the one or more computing resources comprises selecting one or more computing resources that implement a particular query engine to execute the first query.

7. The method of claim 6, wherein the particular query engine is indicated in the first query.

8. The method of claim 5, wherein at least one of the data sets is stored according to a different data format than another one of the data sets, wherein the at least one data set is stored in a different remote data store than the other one of the data sets.

9. The method of claim 5, further comprising:
monitoring one or more query execution metrics for the first query;
based on the monitoring, determining that the first query execution metrics exceed a query execution limitation; and
halting execution of the first query at the computing resources.

10. The method of claim 5,
wherein
the schema that describes the one or more data sets is stored in a metadata store, and
wherein the provisioned computing resources access the schema in the metadata store to obtain the schema.

11. The method of claim 5, wherein the first query is received via a programmatic interface and wherein the method further comprises returning a result for the first query generated at the provisioned computing resources via the programmatic interface.

12. The method of claim 5, wherein the first query indicates a result destination for the first query, and wherein the method further comprises sending a result for the first query generated at the provisioned computing resources to the result destination.

13. The method of claim 5, wherein the first query is received at a service endpoint for a network-based service that performs the receiving, the provisioning, and the routing, wherein the first query is received from a first client, wherein a plurality of other queries are received at the service endpoint from a plurality of other clients that access one or more data sets different than the data sets, and wherein the network-based service performs the receiving, the provisioning, and the routing for the other received queries.

14. A non-transitory, computer-readable storage medium, storing program instructions that when executed by one or more computing devices cause the one or more computing devices to implement:

receiving a schema that describes one or more data sets in one or more remote data stores;

storing the schema separately from a pool of computing resources configured to execute queries;

receiving a first query directed to the one or more data sets in the one or more remote data stores wherein the one or more data sets are stored separately from the pool of computing resources configured to execute queries;

in response to receiving the first query:

selecting one or more computing resources to execute the first query from a pool of computing resources configured to execute queries;

removing the selected computing resources from the pool of computing resources;

associating the removed computing resources with the first query; and

routing the first query to the selected one or more computing resources to execute the first query with respect to the separately stored one or more data sets in the remote data stores, wherein the selected one or more computing resources obtain the separately stored schema that describes the one or more data sets as part of executing the first query.

15. The non-transitory, computer-readable storage medium of claim 14, wherein the program instructions cause the one or more computing devices to implement:

monitoring one or more query execution metrics for the first query;

based on the monitoring, determining that the first query execution metrics exceed a query execution limitation; and

halting execution of the first query at the computing resources.

16. The non-transitory, computer-readable storage medium of claim 14, wherein the first query is formatted according to Structured Query Language (SQL).

17. The non-transitory, computer-readable storage medium of claim 14, wherein the first query is received via a programmatic interface and wherein the program instructions cause the one or more computing devices to further implement returning a result for the first query generated at the selected computing resources via the programmatic interface.

18. The non-transitory, computer-readable storage medium of claim 14, wherein the schema that describes the one or more data sets is stored in a metadata store, and wherein the selected computing resources access the schema in the metadata store to obtain the schema.

19. The non-transitory, computer-readable storage medium of claim 18, wherein the first query and the schema are received via a same user interface.

20. The non-transitory, computer-readable storage medium of claim 19, wherein the one or more computing devices are implemented as part of a network-based managed query service, wherein at least one of the data sets is stored in a network-based data storage service, wherein the schema is stored in a network-based data catalog service, wherein the selected computing resources are implemented as part of a virtual compute service, and wherein the managed query service, data storage service, data catalog service and virtual compute service are offered as part of a same provider network.

21. A system, comprising:

a memory to store program instructions which, if performed by at least one processor, cause the at least one processor to perform a method to at least:

monitor metrics for one or more pools of computing resources for a network-based service, wherein the pools of computing resources execute jobs selectively routed by the network-based service to different ones of the computing resources;

based on the monitoring, determine if one or more computing resources for the one or more pools is to be adjusted, and adjust the number of one or more computing resources for the one pool in response to the determination.

22. A method, comprising:

for one or more computing resources obtained from pool of computing resources configured to execute queries for a network-based service, detecting completion of one or more previously received queries executing at the one or more computing resources;

updating resource management data for the network-based service to identify the one or more computing resources as available for servicing queries for the network-based service and return the one or more computing resources to the pool;

in response to receiving a query at the network-based service:

selecting a computing resource to execute the query from the one or more computing resources identified available to execute queries in the resource management data; and

routing the query to the selected computing resource for execution.

23. A method, comprising:

receiving a first query directed to one or more data sets;

in response to receiving the first query:

selecting one or more computing resources to execute the first query based, at least in part, on an historical data of computing resources used in performing prior received queries; and

performing the first query at the selected computing resources with respect to the one or more data sets.