

Preferred Aggregate Storage

AtScale Administrators now have the ability at the project-level to set a data warehouse as the Preferred Aggregate Storage location. If an aggregate is eligible, it will be promoted to the data warehouse set as the preferred aggregate storage location. The Preferred Aggregate Storage feature enables you to connect to a read-optimized database that can act as an aggregate table cache. Using an aggregate table storage that is optimized for fast-reads improves query performance.

About This Task

When a Preferred Aggregate Storage location is set, the AtScale engine assesses aggregates for **promotion** to the target data warehouse. Administrators can configure a maximum per-table row limit for a promoted aggregate candidate, and a cumulative row count limit for the data warehouse chosen as the preferred aggregate store. Additionally, Aggressive Aggregate promotion can be enabled on a per-project basis.

Aggregate Promotion

- ▲ **Aggregate Promotion** is defined as the assignment of an aggregate to a preferred storage location. Promoted aggregates may replicate an existing aggregate from a non-Preferred Aggregate Storage location. Aggregates may also be chosen to be directly created in the Preferred Aggregate Storage location, for example, if the source dataset has [local](#) aggregate creation disabled. When promoting a copy of an existing aggregate, aggregate promotion does not cause the existing aggregate instance to be deleted or deactivated.
 - ▲ An aggregate is eligible for promotion if:
 1. AtScale receives a query that is entirely resolved with aggregates
 2. The dataset [Aggregate Locality](#) options of all contributing datasets would permit the selection of that data warehouse.
 3. That aggregate instance is not already in preferred storage.
 4. That aggregate is not a User-Defined Aggregate (UDA).
 5. The row-count of that aggregate is less than or equal to the row-count limit for the preferred storage location (if a limit has been set).
 6. The cumulative row-count limit of the aggregate, along with other aggregates in the store, is below the configured cumulative row count limit configured for the preferred storage location (if a limit has been set).



Note: Both the source and target data warehouses must have **portable** aggregates. Please refer to the [Data Warehouse support matrix](#) for more information.

- ▲ **Aggregate Promotion Status:** To discern the promotion status of an aggregate, navigate to the Aggregates > Aggregate Definitions tab and observe the Promoted column. If promoted, the column value will list YES.

Figure 1. The Promoted Column on the Aggregate Definitions tab.

CUBE NAME	PROJECT NAME	TYPE	PROMOTED
Test Sales ...	agg_promotio...	System-Defined	No
Test Sales ...	PromotedAgg...	User-Defined	No
Test Sales ...	agg_promotio...	System-Defined	No
Test Sales ...	agg_promotio...	User-Defined	No
Test Sales ...	agg_promotio...	System-Defined	No
▶ Test Sales ...	agg_promotio...	System-Defined	Yes

Aggressive Aggregate Promotion

AtScale 2020.3.0 and onward introduces the concept of **Aggressive Aggregate Promotion**. Aggressive Aggregate Promotion is enabled at the project level. When enabled, all aggregates referenced by a query part will be considered for promotion, regardless of whether a join to other non-preferred or non-aggregate datasets was required. Aggressive promotion of an aggregate will be considered if all of the conditions in **Aggregate Promotion** topic above are met, except if the query is resolved entirely with aggregates.

In order for aggregates to be built, it must also be a valid local aggregate. Joins that will not produce an aggregate in the local data warehouse will not be created in the promoted warehouse. One example of this is when the compression is too low to generate an aggregate that improves performance adequately.

Row Count Limits

To prevent the AtScale engine from writing a large aggregate to a database that cannot adequately handle tables above a certain size, Administrators can configure a maximum per-table row limit for a promoted aggregate candidate.

Additionally, a cumulative row count limit across all aggregates from the current project for the data warehouse selected as the Preferred Aggregate store is configurable from the **Edit a Project** dialog. Setting a cumulative row count limits the amount of disk space that AtScale uses for promoted aggregates (on behalf of a single project) in that Data Warehouse, preventing performance issues.



If you have set the `query.result.max_rows` engine setting on your AtScale system, the per-table row count limit for Preferred Aggregate Storage must be set lower than or equal to that engine setting.

Procedure

To configure a data warehouse as the Preferred Aggregate Storage location:

1. From the Project Overview page, select the draft version of the project from the left hand drawer.
2. Select the **Edit** tab.
3. In the Edit a Project dialog, scroll down to the **Preferred Aggregate Storage** section.
4. Select the data warehouse to be used as the **Preferred Aggregate Storage** location.
5. Set the desired row count limit. If a row-count limit is not set, no limit is enforced.
6. Set the desired cumulative row count limit. If a cumulative row-count limit is not set, no limit is enforced.
7. Select **Save**.
8. **Publish** the project. If you want to set a different data warehouse as the Preferred Aggregate Storage location, a re-publish of the project is required to initiate the change.



Note: It is not necessary for any of your AtScale project's datasets to reside in the data warehouse that you plan to use as the preferred aggregate storage location.

Figure 2. The Preferred Aggregate Storage setting on the Edit a Project Dialog.

Preferred Aggregate Storage

Using Preferred Aggregate Storage will cause query failures unless the virtualization engine is running.

DATA WAREHOUSE

BigQuery Data Warehouse

ROW LIMIT ⓘ

✓

CUMULATIVE ROW COUNT ⓘ

✓

To configure a data warehouse for **Aggressive Aggregate** Promotion:

1. From the Project Overview page, select the draft version of the project from the left hand drawer.
2. Select the **Edit** tab.
3. Move the slider to the right to enable **Aggressive Aggregate Promotion**.
4. Select **Save**.
5. **Publish** the project. If you want to disable Aggressive Aggregate promotion, a re-publish of the project is

required to initiate the change.

Figure 3. The Aggressive Aggregate promotion setting on the Edit a Project Dialog.

