

△ > Reference > Release notes > 0.7.1

Release Version 0.7.1: November 2023

Significant Apache Pinot updates since last StarTree release

For complete details on Pinot changes, see [Releases](#).

- Skip unparseable records in the CSV reader. To enable, set the `skipUnParseableLines` flag to true ([pull request](#)).
- Protocol buffer ingestion supports null values with Proto 3 ([pull request](#))
- Upgrade Confluent libraries from 5.5.3 to 7.2.6 ([pull request](#))
- Faster real-time table ingestion with updates to the segment builder. To enable, edit the table configuration to set `realtime.segment.flush.enable_column_major` to `true` ([pull request](#))
- Improve alias handling in single-stage engine with multiple fixes to column aliases ([pull request](#))
- Enhance handling of new partitions when using `StrictReplicaRouting` to prevent "instance unavailable" exceptions ([pull request](#))
- Optimize performance in the multi-stage engine:
 - For a single join key and group key scenario, operate directly on the key values without wrappers ([pull request](#))
 - Operate on column indexes in multi-stage aggregations to prevent extra conversion steps
 - Avoid converting unnecessary rows in aggregations ([pull request](#))
- Enhance segment assignments for upsert tables with more checks to ensure that the conditions required for upsert functionality to work are not violated ([pull request](#))
- Fix handling of literals used in aggregation for v2 engine ([pull request](#))

Breaking changes

- You must now specify the data type of literals in Pinot queries. Before this change, for example, `2022-02-02 22:22:22.123` was automatically treated as a timestamp data type. Now, following standard SQL behavior, use `CAST('2022-02-02 22:22:22.123' AS TIMESTAMP)` instead ([pull request](#)).
- Change the "forbidden" error to "unauthorized" ([pull request](#))
- Table configurations that point to a different schema name no longer work ([pull request](#)).
- You can no longer change the table state using the `GET` call ([pull request](#)).
- You can no longer create a schema with `NaN` as the default value ([pull request](#)).
- `BigDecimal` responses are now stored as a string with double quotes instead of a number ([pull request](#)).

Dependencies

- Upgrade to Parquet 1.12.3 ([pull request](#))
- Upgrade to Hadoop 3.2.4 ([pull request](#))
- Upgrade to AVRO 1.10.2 ([pull request](#))

StarTree extensions for Apache Pinot

The following updates are available only in StarTree Cloud.

- Improvements to [file ingestion task](#):
 - Enhancements to batch ingestion using minion to improve atomic ingestion and backfill operations
 - Control size-based segment creation with `desiredSegmentSize` to improve performance
- Automatically tune segment size for segment refresh task without configuring `maxNumRecordsPerTask` and `maxNumRecordsPerSegment`. Size-based tuning helps make predictable segment sizes and avoid memory- or size- related exceptions
- Validation is stricter for using sync mode in conjunction with other tasks. You can no longer schedule the segment refresh task at the same time as sync mode.
- Separate RocksDB log from server logs to improve debugging experience and allow you to set different retention and rollover policies

- Improve Kafka logs by changing the following classes to error-level:
 - `KafkaConsumer`
 - `AppInfoParser`
 - `ConsumerConfig`
- Enhancements to upsert tables:
 - Correctly track primary key count and add corresponding metrics
 - Improve stability during deletion
- Improve performance and navigation in broker and server Grafana dashboards
- Move to Google Trust Services Certificate Authority to improve certification management

Data Manager

- Improve data sampling from Kafka topics with large numbers of partitions by preventing "no data" error in preview
- Automate Google Cloud Platform (GCP) credentials in Data Manager so you can ingest instead of having to contact StarTree support
- Improve error messages to aid troubleshooting

ThirdEye

- Improve loading time for multi-dimension alerts and dashboard statistics
- Simplified alert creation with advanced anomaly detection and tuning options, reducing complexity of data patterns and seasonality

Last updated on May 21, 2024