

HCatalog Storage Formats

Storage Formats

- [Storage Formats](#)
 - [SerDes and Storage Formats](#)
 - [Usage from Hive](#)
 - [CTAS Issue with JSON SerDe](#)

SerDes and Storage Formats

HCatalog uses Hive's SerDe class to serialize and deserialize data. SerDes are provided for RCFile, CSV text, JSON text, and SequenceFile formats. Check the [SerDe documentation](#) for additional SerDes that might be included in new versions. For example, the [Avro SerDe](#) was added in Hive 0.9.1, the [O RC](#) file format was added in Hive 0.11.0, and [Parquet](#) was added in Hive 0.10.0 (plug-in) and Hive 0.13.0 (native).

Users can write SerDes for custom formats using these instructions:

- [How to Write Your Own SerDe](#) in the Developer Guide
- [Hive User Group Meeting August 2009](#) pages 64-70
- also see [SerDe](#) for details about input and output processing

For information about how to create a table with a custom or native SerDe, see [Row Format](#), [Storage Format](#), and [SerDe](#).

Usage from Hive

Hive and HCatalog (version 0.4 and later) share the same storage abstractions, and thus, you can read from and write to HCatalog tables from within Hive, and vice versa.

However, for HCatalog versions 0.4 and 0.5 Hive does not know where to find the HCatalog jar by default, so if you use any features that have been introduced by HCatalog, such as a table using the JSON SerDe, you might get a "class not found" exception. In this situation, before you run Hive, set environment variable `HIVE_AUX_JARS_PATH` to the directory with your HCatalog jar. (If the examples in the [Installation](#) document were followed, that should be `/usr/local/hcat/share/hcatalog/.`)

After version 0.5, HCatalog is part of the Hive distribution and you do not have to add the HCatalog jar to `HIVE_AUX_JARS_PATH`.

CTAS Issue with JSON SerDe

Using the Hive `CREATE TABLE ... AS SELECT` command with a JSON SerDe results in a table that has column headers such as `"_col0"`, which can be read by HCatalog or Hive but cannot be easily read by external users. To avoid this issue, create the table in two steps instead of using CTAS:

1. `CREATE TABLE ...`
2. `INSERT OVERWRITE TABLE ... SELECT ...`

See [HCATALOG-436](#) for details.

Navigation Links

Previous: [Command Line Interface](#)

Next: [Dynamic Partitioning](#)

SerDe general information: [Hive SerDe](#)

SerDe details: [SerDe](#)

SerDe DDL: [Row Format](#), [Storage Format](#), and [SerDe](#)

General: [HCatalog Manual](#) – [WebHCat Manual](#) – [Hive Wiki Home](#) – [Hive Project Site](#)