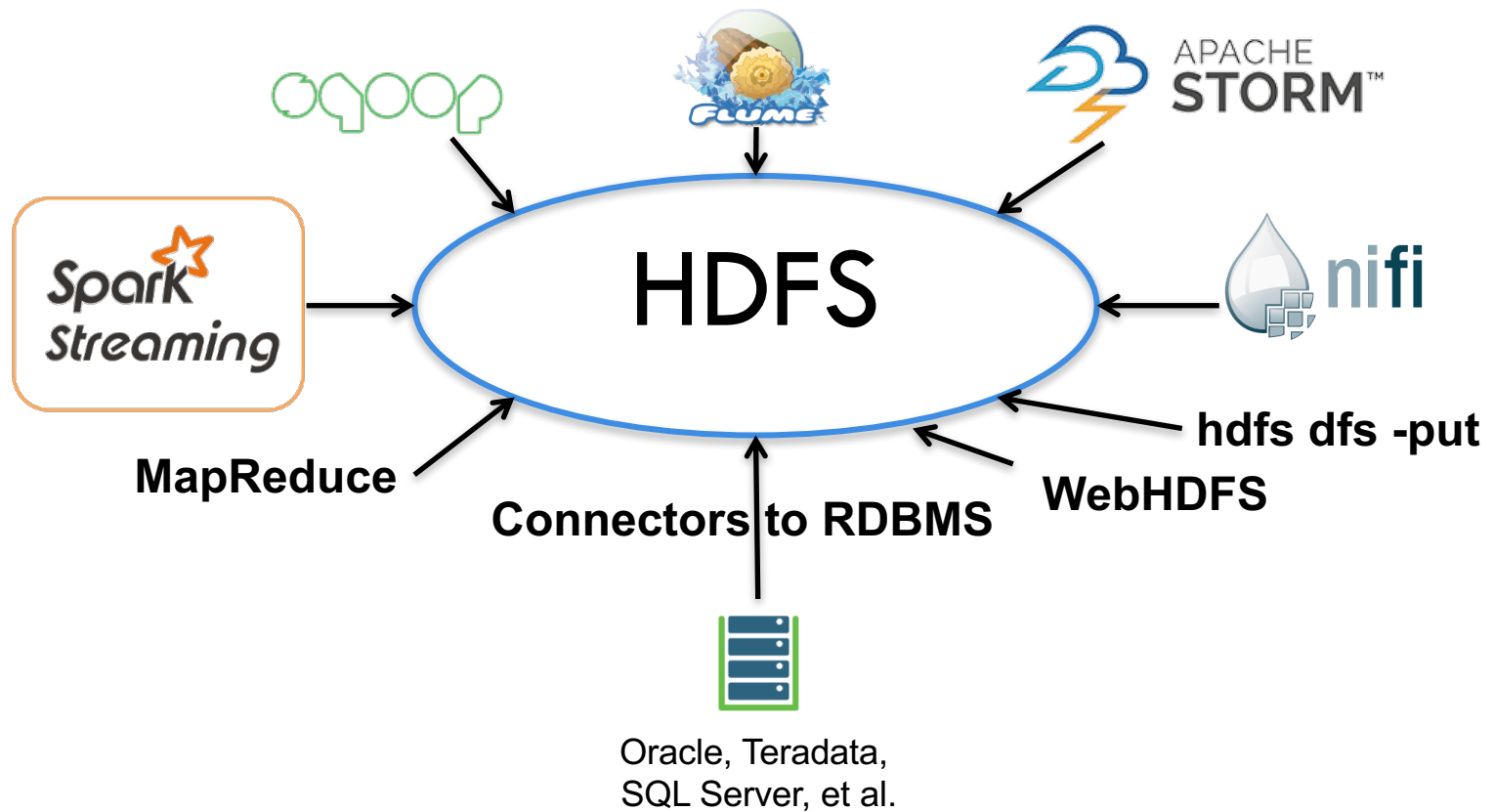




Ingesting Data into HDFS

School of Information Studies
Syracuse University

Options for Data Ingestion



When Integrating...

Remember the “Hadoop way”

- Store data “as is” (the way it was extracted from its source).
- Store common datasets/entities in a folder, not a file.
- Apply a schema to the data when they are read from HDFS.