Hive là một kho dữ liệu (data warehouse)  xử lý các dữ liệu dạng cấu trúc trên nền tảng hadoop. Sử dụng hive để tổng hợp, tạo truy vấn và phân tích dữ liệu một cách dễ dàng mà không cần phải hiểu nhiều về MapReduce. Hive cung cấp một giao diện web để tương tác với hdfs. Tương tác command line.

Điểm hay của HIVE là hỗ trợ truy xuất giống SQL đến dữ liệu có cấu trúc, được biết với tên HiveSQL (hoặc HQL) cũng như phân tích big data với MapReduce. Hive không được xây dựng để hồi đáp nhanh các câu truy vấn nhưng nó được xây dựng cho các ứng dụng khai thác dữ liệu (data mining). Các ứng dụng khai thác dữ liệu có thể mất nhiều phút đến nhiều giờ để phân tích dữ liệu và HIVE được dùng chủ yếu.