

About Apache Pig

Table of contents

1 What is Apache Pig?.....	2
----------------------------	---

1. What is Apache Pig?

Apache Pig is a platform for analyzing large data sets. Pig's language, Pig Latin, is a simple query algebra that lets you express data transformations such as merging data sets, filtering them, and applying functions to records or groups of records. Users can create their own functions to do special-purpose processing.

Pig Latin queries execute in a distributed fashion on a cluster. Our current implementation compiles Pig Latin programs into Map-Reduce jobs, and executes them using Hadoop cluster.

For more information about Pig, please see the [Pig wiki](#).