

scala for spark in one slide

Spark is written in scala - no extra language API libraries or wrappers needed.

Runs on the Java Virtual Machine. Interoperates with JAVA and is converted to JAVA.

Integrated Object Oriented and Functional Programming

- Data is generally immutable (val). Great for spark programming model!

- Anonymous functions.

Statically Typed

- Types checked at compile-time.

- Convenient Type Inference for shorter code.

Implicit Type Conversion

- If a method is not found for your object, it will be converted to an object which does have that method. Example: RDD methods and PairRDD methods.

Case Class

- convenient Serializable object with override methods such as equals(), hash() and toString().

Can use in REPL/shell and Notebook (ie. Jupyter, Zeppelin, DBCE).