Scala for Apache Spark

Markus Dale, medale@asymmetrik.com

Jan 2019

Intro, Slides And Code

- Slides: https://github.com/medale/scala-spark/blob/master/presentation/ScalaSpark.pdf
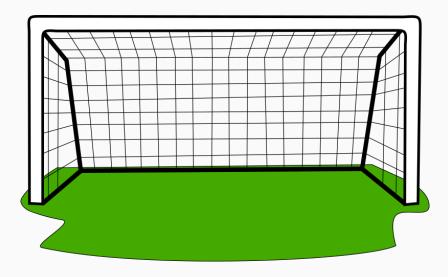- Scala Spark Code Examples: https://github.com/medale/scala-spark

Figure 1: Intro to Scala for Spark

Why Scala for Spark?

- full interoperability with Java
    - strong type system
    - elegant multi-paradigm (functional & OO)
    - less keyboard typing
- JVM

Java to Scala - Java Main

```java
package com.uebercomputing.scalaspark.common;
public class JavaMain {
  private int answer = 0;
  public JavaMain(int answer) {
    this.answer = answer;
  }
  public int getAnswer() {
     return answer;
  }
  public static void main(String[] args) {
    System.out.println("Starting a Java program...");
    JavaMain jaMain = new JavaMain(42);
    int answer = jaMain.getAnswer();
    System.out.println("The answer was " + answer);
  }
}
```

Scala Main

```scala
package com.uebercomputing.scalaspark.common

object ScalaMainOne {

  def main(args: Array[String]): Unit = {
    println("Starting a Scala program...")
    val scMain = new ScalaMainOne(42)
    val answer = scMain.answer
    println(s"The answer was ${answer}")
  }
}

class ScalaMainOne(val answer: Int)
```