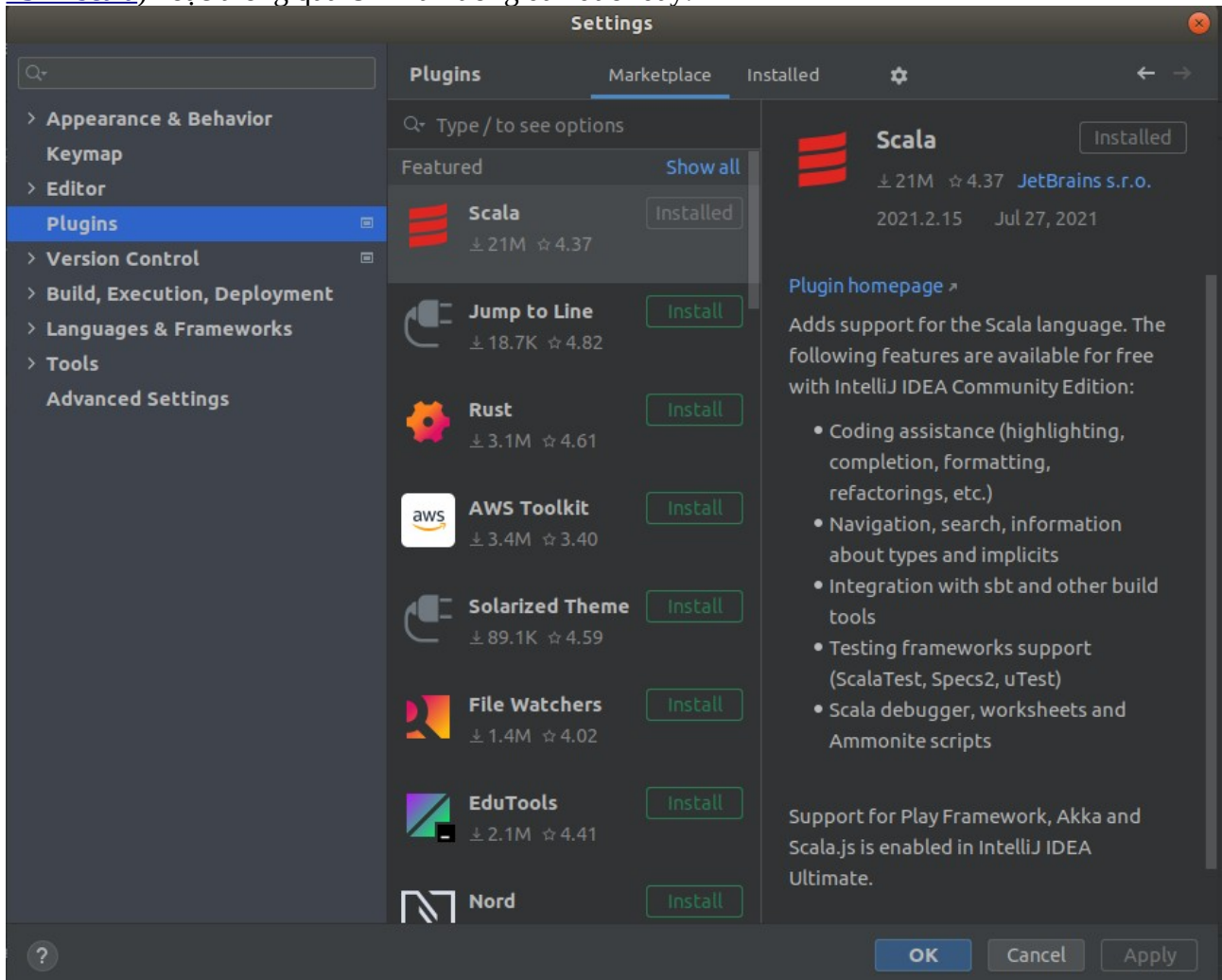


HƯỚNG DẪN CÀI ĐẶT SPARK SCALA

PHẦN 1: Hướng dẫn cài đặt trên IntelliJ

- Yêu cầu: Đã tải IntelliJ (chỉ cần bản Community)
- Đã cài Spark, Java JDK Oracle (đề xuất sử dụng Java JDK 11)

Bước 1: cài Plugin Scala cho IntelliJ thông qua marketplace (<https://plugins.jetbrains.com/plugin/1347-scala>) hoặc thông qua UI như hướng dẫn dưới đây:

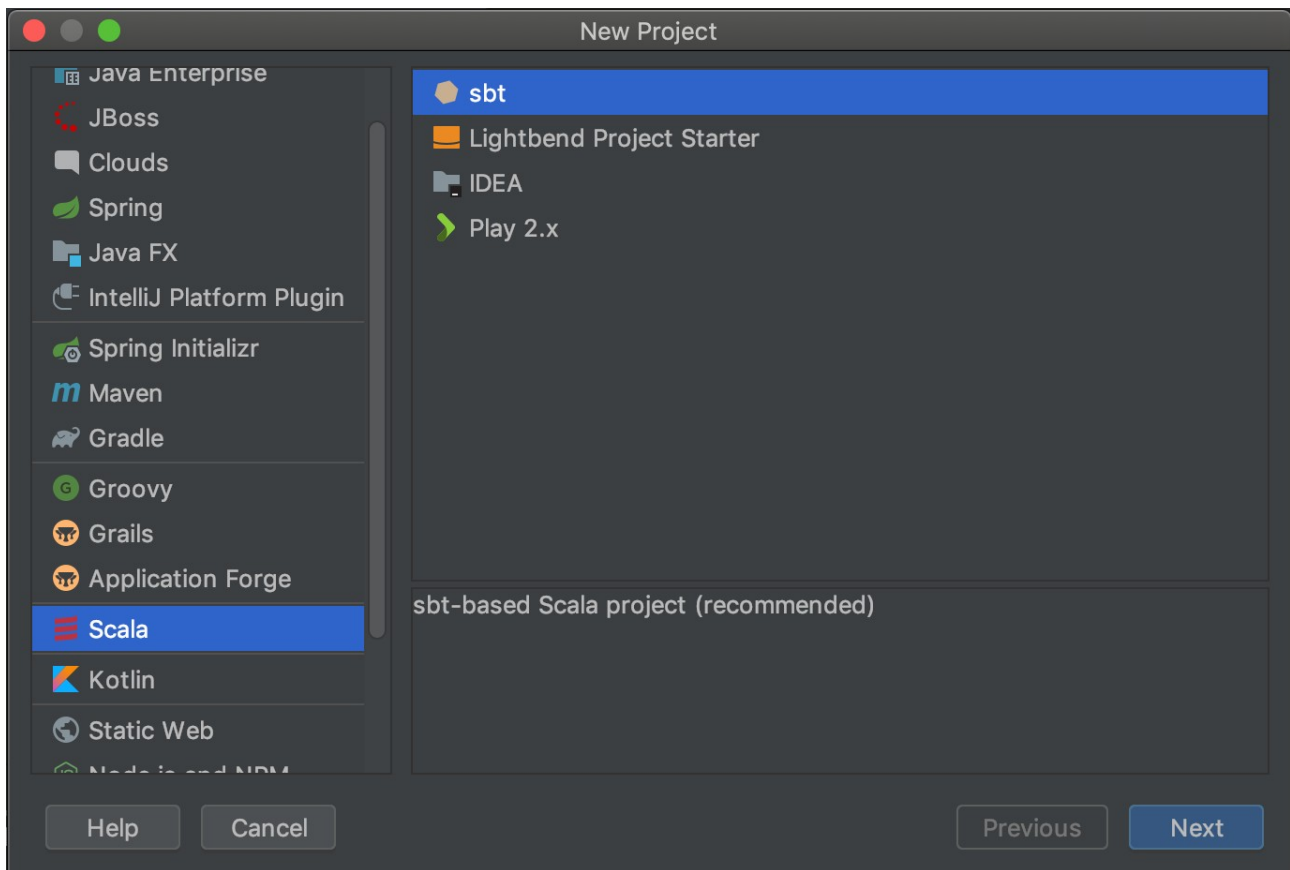


Lựa chọn scala và chọn Plugin như mô tả, ấn “Apply” >> “OK”

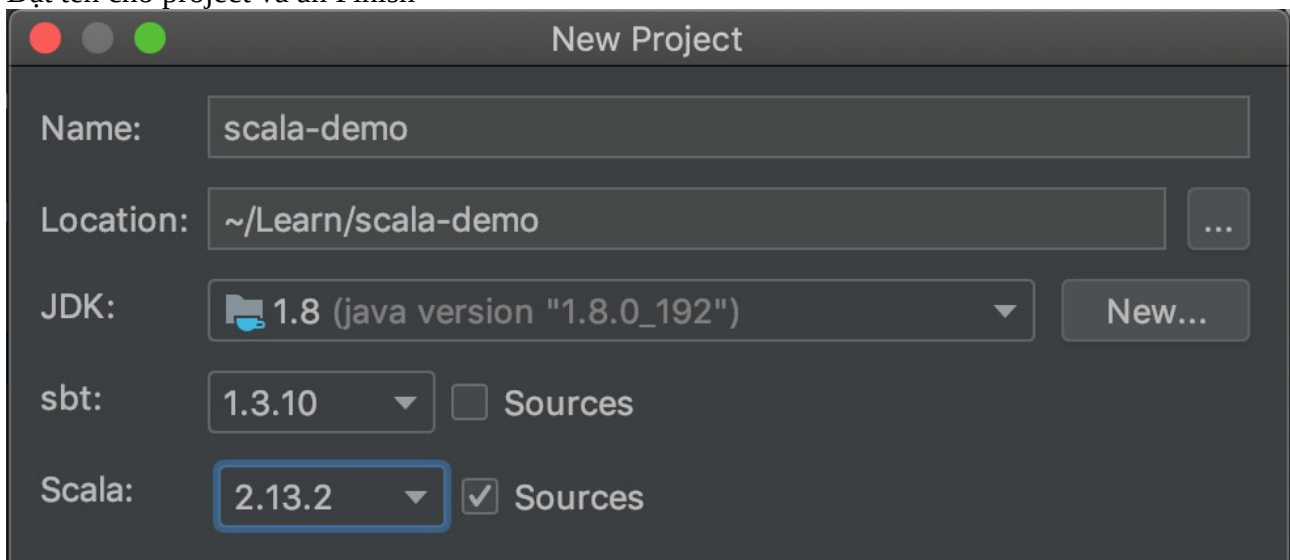
Bước 2:

Tạo project mới File → New → Project...

Lựa chọn “Scala” >> “sbt”, chọn “Next”



Đặt tên cho project và ấn Finish



Bước 3: Viết chương trình Hello World đầu tiên

Đầu tiên chuột phải vào folder “src” và chọn “Mark Directory as” >> “Resource Root” để đặt chế độ thư mục gốc tại “src”

Tạo 1 file tên “HelloWorld.scala” trong thư mục “src” với nội dung

```
object HelloWorld extends App {  
  println("Hello, World!")  
}
```

Có hai cách để chạy chương trình này:

Cách 1: chạy trên terminal:

```
> scalac HelloWorld.scala
```

câu lệnh sẽ sinh ra file thực thi dạng “HelloWorld.class” và “HelloWorld\$.class” trong đó file “HelloWorld.class” là file giúp nhận diện các class có trong project và file “HelloWorld\$.class” là class chứa chương trình chính.

Để tiếp tục chạy, thực hiện câu lệnh sau:

```
> scala HelloWorld
```

Cách 2: Chuột phải vào file “HelloWorld.scala” và chọn run “Hello World”

Cách 3: Chạy lệnh

```
> sbt run
```

Bước 4: Thêm các dependency

Mở file build.sbt và thấy có nội dung như sau:

```
name := "scala-demo"
```

```
version := "0.1"
```

```
scalaVersion := "2.12.8"
```

Sửa đổi lại và thêm các dòng sau để được kết quả cuối là

```
name := "scala-demo"
```

```
version := "0.1"
```

```
scalaVersion := "2.12.8"
```

```
val sparkVersion = "2.4.0"
```

```
libraryDependencies += "org.apache.spark" %% "spark-core" %  
sparkVersion
```

```
libraryDependencies += "org.apache.spark" %% "spark-sql" %  
sparkVersion
```

Chú ý: Phiên bản của scalaVersion và sparkVersion phải hỗ trợ cho nhau. Chi tiết của các phiên bản hỗ trợ này xem tại server của Maven theo link sau:

Link xem các phiên bản Scala được sbt hỗ trợ:

<https://repo1.maven.org/maven2/org/scala-lang/scala-reflect/>

Link xem các phiên bản spark ứng với scala qua link:

<https://repo1.maven.org/maven2/org/apache/spark/>

Khi đó các thư viện của spark được tự động add vào thư viện support của sbt

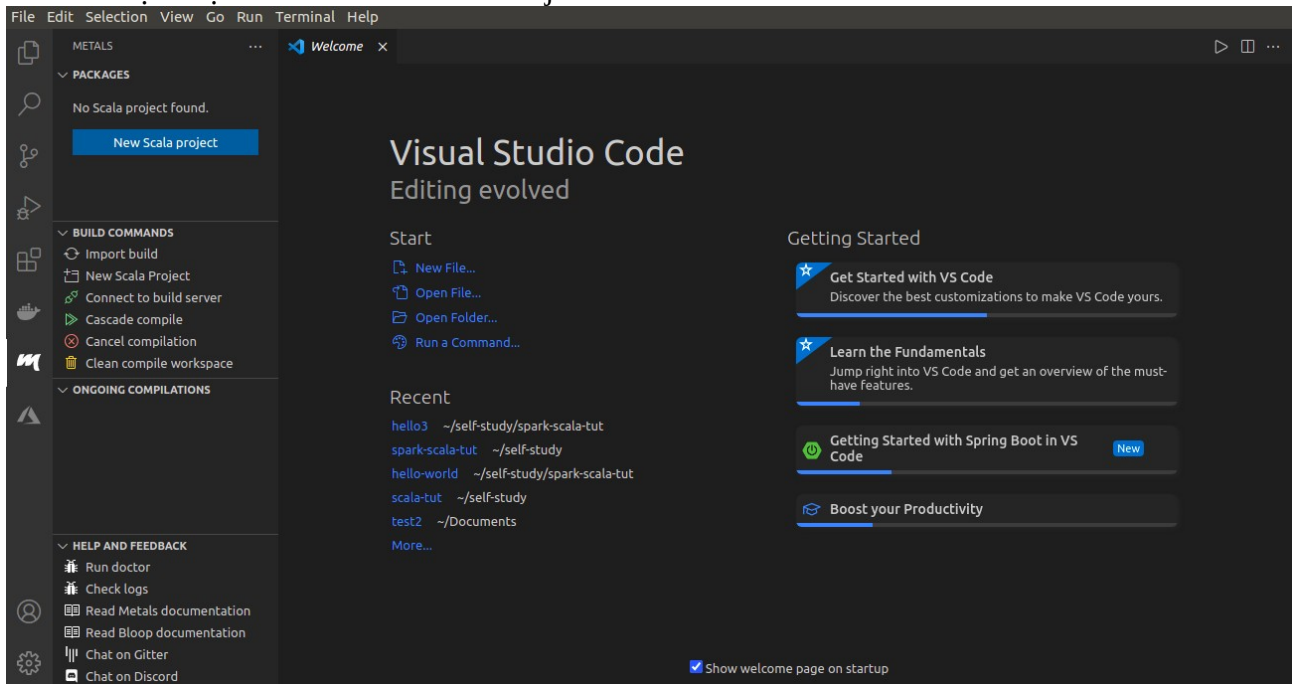
Nguồn tham khảo chính: <https://www.linkedin.com/pulse/running-spark-application-using-intellij-murali-poola>

PHẦN 2: Cài đặt trên VS Code:

Việc cài đặt trên VS Code cho phép triển khai nhanh hơn so với IntelliJ

Bắt đầu bằng việc tải các Extension Scala Language Scala(sbt), Scala(metal), Scala Language Server

Bước 1: Lựa chọn Metal > New Scala Project



Các bước còn lại tương tự như đối với IntelliJ. Chỉ khác ở chỗ khi cập nhật lại file build.sbt thì thực hiện các lệnh sau trong màn hình terminal;

>bst

khi màn hình bst shell hiện lên, thực hiện câu lệnh

>reload

>update