SPARK PROPERTIES

1. Spark Properties

Spark properties điều khiển hầu hết các cài đặt ứng dụng và được cấu hình riêng biệt cho mỗi ứng dụng. Các thuộc tính này có thể được cài đặt một cách trực tiếp trên SparkConf và được chuyển tới SparkContext.SparkConf cho phép bạn cấu hình một vài thuộc tính phổ biến (chẳng hạn như master URL và application name), cũng như các cặp key – value khác thông qua phương thức set().

Lưu ý rằng chúng ta có thể có nhiều hơn một luồng trong local mode, và trong những trường hợp như Spark Streaming, chúng ta có thể yêu cầu nhiều hơn 1 luồng để tránh bất kì vấn đề thiếu sót nào.

|  |  |
| --- | --- |
| Properties Name | Meaning |
| Spark.app.name | Tên ứng dụng của bạn. Nó sẽ xuất hiện trên trên giao diện người dùng và log data. |
| Spark.driver.cores | Số core sử dụng trong quá trình xử lý. Chỉ dùng trong cluster mode. |
| Spark.driver.maxResultSize | Giới hạn tổng kích thước của các kết quả được tuần tự hóa của tất cả các phân vùng cho mỗi Spark action tính bằng byte. Tối thiểu là 1M hoặc 0 cho không giới hạn. Việc có một giới hạn cao có thể gây ra lỗi hết bộ nhớ trong trình điều khiển (phụ thuộc vào spark.driver.memory và chi phí bộ nhớ của các đối tượng trong JVM). Đặt giới hạn thích hợp có thể bảo vệ trình điều khiển khỏi hết lỗi bộ nhớ. |
| Spark.driver.memory | Giá trị của bộ nhớ để sử dụng trong quá driver process. Nơi SparkContext được khởi tạo, có cùng định dạng với chuỗi bộ nhớ JVM với hậu tố đơn vị kích thước(“k”, “m”, “g” hoặc “t”) |
| Spark.executor.memory | Dung lượng bộ nhớ sử dụng cho mỗi quá trình thực thi, có cùng định dạng với chuỗi bộ nhớ JVM với hậu tố là đơn vị kích thước (“k”, “m”, “g”, hoặc “t”). |
| Spark.extraListeners | Một danh sách các lớp phân tách bởi dấu phẩy hiện thực SparkListener. Khi khởi tạo SparkContext, các instance của các class này sẽ được tạo và đăng kí với Spark’s listener bus. Nếu một lớp có một phương thức khởi tạo một đối số chấp nhận một SparkConf, thì phương thức khởi tạo đó sẽ được gọi; nếu không một phương thức khởi tạo không đối số sẽ được gọi. Nếu không tìm thấy hàm khởi tạo nào, quá trình tạo SparkContext sẽ không thành công với một ngoại lệ. |
| Spark.local.dir | Thư mục để sử dụng cho không gian bản thảo trong Spark, bao gồm các tệp đầu ra của map và RDDs được lưu trên đĩa và nên nằm trên một đĩa cục bộ và nhanh trong hệ thống. |
| Spark.master | Một cụm quản lý để kết nối tới. |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |