Mandelbrot Computation Using Spark

Data Management 2019

Outline

➔ Mandelbrot
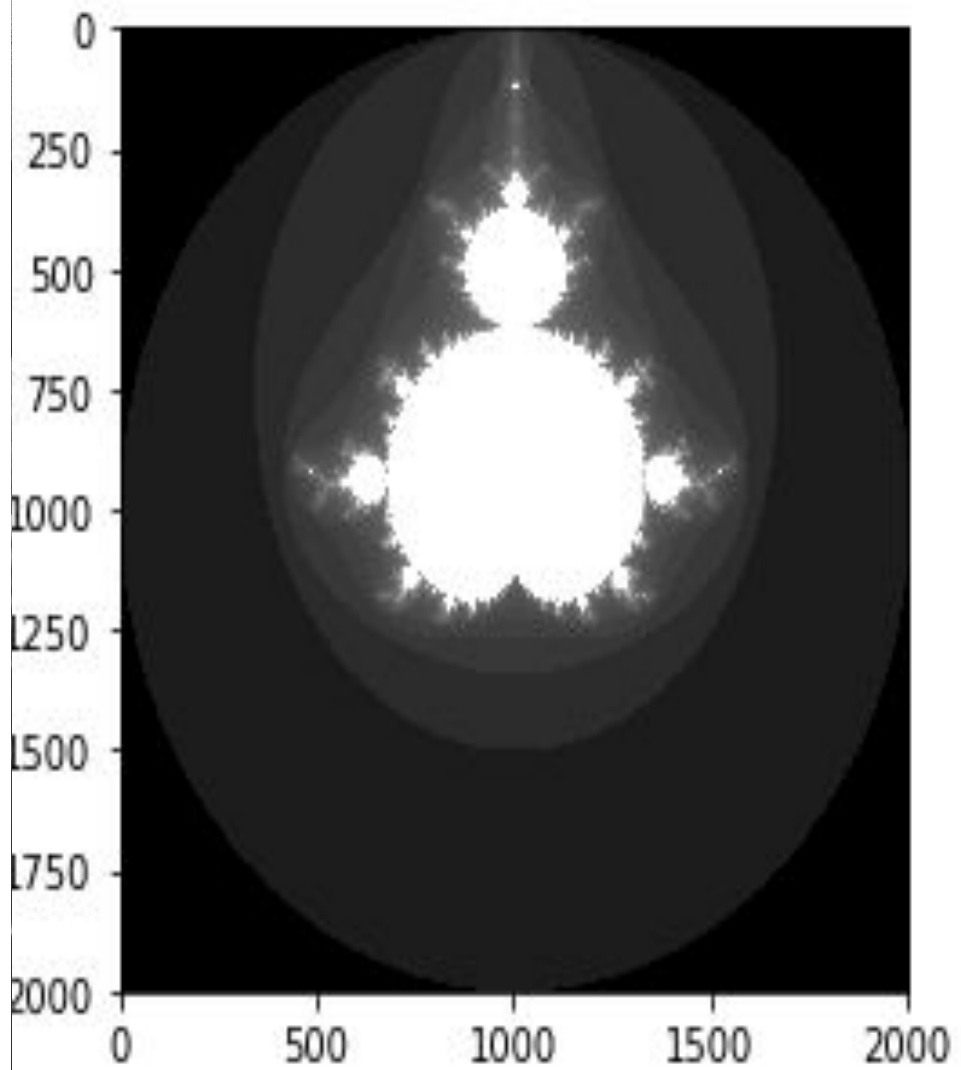 ◆ What it is

➔ Mandelbrot Computation On Spark
  ◆ Default Partitioning
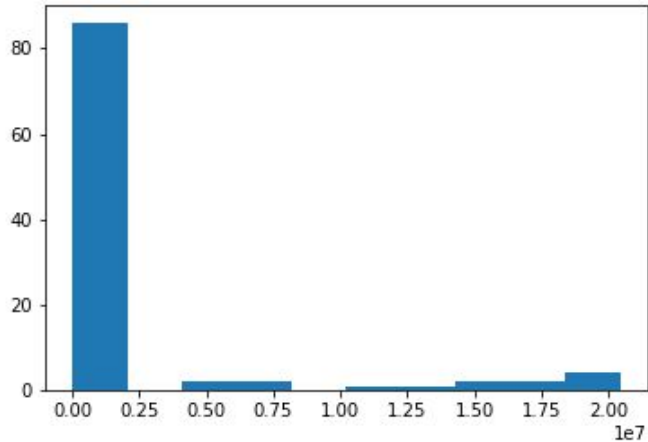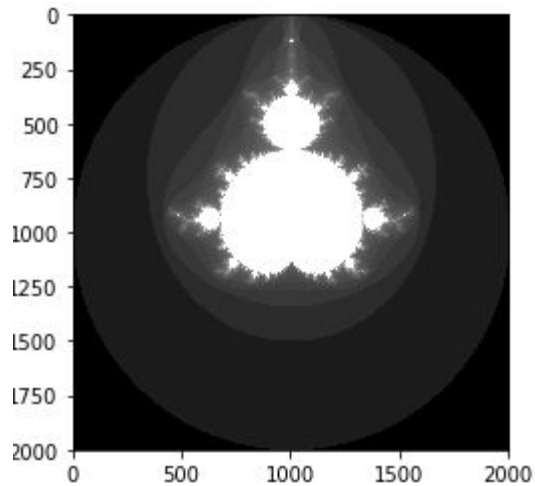 ◆ Load Balancing

Mandelbrot

What It Is

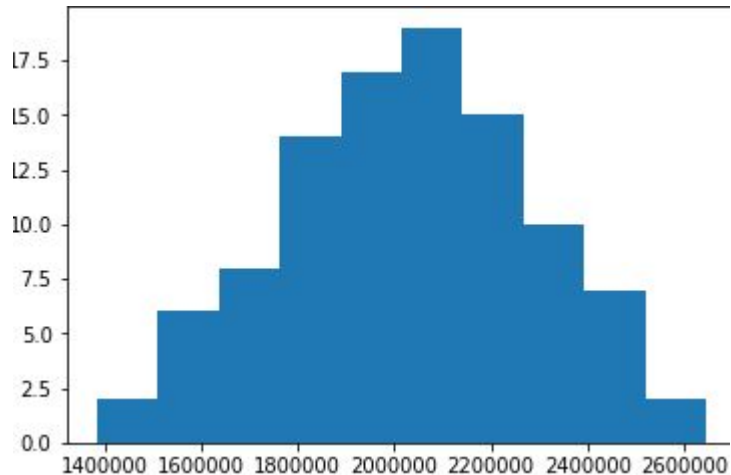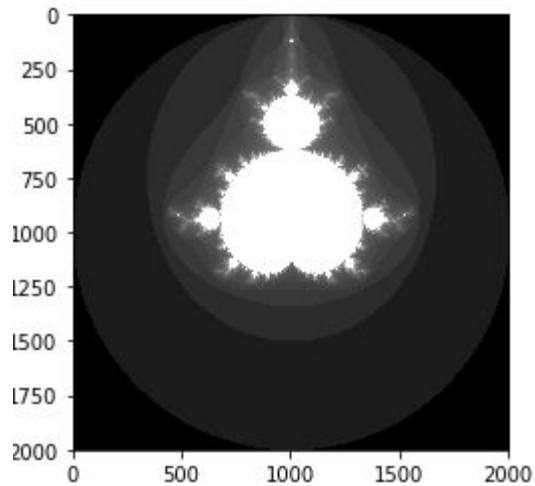Set Of Complex Numbers

Computation On Spark

Default Partitioning

Strategy:-
Partition the data into 100 partitions, within a fixed order.

Not Balanced:-
As some pixels are fully devoted to some certain partitions,
Which actually leads the processing time to be slow

Load Balancing

Strategy:-
Partition the data, was done using randomised partitioning, using random.shuffle()

Balanced:-
As, the pixels have been reordered into different partitions, which is actually leading the computation to be fast and balanced.

Default Partitioning

- Time:- 14 Min.

Load Balancing

- Time:- 2 Min.

Thank You

By:- Sujoy Roy

Guided By:- Dr. Dip Sankar Banerjee

Indian Institute Of Information Technology, Guwahati