Parallel Computing and Simulations with Revolution R Enterprise

Techniques for parallel computing with R in a distributed computing environment.

Overview

We are at the beginning of the multicore era. Computers will have increasingly many cores (processors). To take advantage of this industry wide shift in architecture, we need to use parallel computing. This course presents an overview of techniques for parallel computing with R on computer clusters, multi-core systems or in grid computing. The class uses a combination of lecture and labs to instruct students on how to effectively use and script parallel programming packages in R.

After completing Parallel Computing with Revolution R Enterprise, you will be able to:

1. Understand the motivation for and limits of parallel computing
2. Recognize when to leverage the power of parallelism
3. Use some of the key open source tools in R for parallelization
4. Leverage Revolution R Enterprise to easily run any R function in parallel.

Audience

Experienced R programmers who would like to enhance the performance of their applications by parallelizing across multicore processors and compute clusters.

Prerequisites

T-RRE-211
Basic understanding of parallel programming

Duration

One day

“One must learn by doing the thing, for though you think you know it — you have no certainty until you try.”

-Sophocles, 450 BC.
Introduction
A platform-agnostic overview of parallelism lays the foundation for a focus on parallelism implementation within R. After completing the introductory module, you will be able to:

1. Understand the motivation for and the limitations of parallelism
2. Identify common parallelism implementations in R
3. Describe how Revolution R Enterprise makes it easy to parallelize code in R

An Overview of parallel tools in R
R provides a plethora of parallel back ends. After completing this module, you will be able to:

1. Understand how and when to leverage various parallel backends: doRSR, doParallel, doSNOW, doMC
2. Use the iterators package along with parallel backends to speed up matrix operations
3. Parallelize nested loops

Parallel tools in Revolution R
Revolution R Enterprise meets the challenges of scalable analytics by addressing HPC programming’s parallelism and data management concerns. After completing this module, you will be able to:

1. Run parallel Linear Models, Logistic Regressions and Generalized Linear Models using RevolutionR Enterprise
2. Execute essentially any R function in parallel on a vectorized data set using HPC and rxExec.

Hands on Applications of Parallel Computing in R
Parallel computing addresses the massive increase in data size and increase in simulation demands that a multitude of industries have recently experienced. In this module you’ll execute a few examples of parallelization including the following

1. Implement the Black-Scholes-Merton Option pricing model
2. K-Means Clustering

Interested in This Course?

To register visit: www.academyr.com

Contact us: academyr@revolutionanalytics.com