



## Computação de Alto Desempenho em R

*Kamila da Silva Alvarenga e Leonardo Siqueira Glória*

O *Cluster Beowulf* é um conjunto de computadores interconectados que trabalham em conjunto com o objetivo de fornecer máximo poder computacional. O *processamento e análise de grande quantidade de dados, bem como análises de cálculos complexos* estão atualmente presentes em diversas áreas, fato este que pode tornar a tarefa computacionalmente complexa e exaustiva. A utilização do *Cluster Beowulf* em conjunto com a linguagem R nos permite realizar tais tarefas. Este projeto visa disponibilizar para a comunidade acadêmica local um *Cluster Beowulf (HPC-High Performance Computing)* para realização de cálculos estatísticos com os softwares R e SAS através do acesso online. Para elaboração do projeto envolve um estudo sobre as arquiteturas de *cluster* para alta performance. O levantamento bibliográfico dos modelos de infraestrutura adotados na construção do cluster através de artigos científicos permitiu o conhecimento das tecnologias que estão sendo utilizadas. O protótipo é composto por computadores com processadores i5 de 2ª geração com 4GB de memória RAM, um Switch 8 Portas TP Link 10100 Mbps e a distribuição *Rocks Cluster* (originalmente chamado *NPACI Rocks*). Constata-se que o processamento e análise dos dados atualmente é algo imprescindível para as organizações que manipulam grande quantidade de dados. Por este motivo os softwares R e SAS tornam-se uma ferramenta aliada para realização dessas tarefas de forma eficiente. Entretanto, diversas vezes, com apenas as funções básicas destas ferramentas não é o satisfatório para que o processamento dos dados seja feito em tempo computacional hábil. Logo, surge a inevitabilidade de buscar novos meios para se alcançar o máximo poder computacional. Observa-se que a paralelização através do *Cluster Beowulf* é uma opção que nos permite analisar dados estatísticos de forma mais rápida sem elevados custos de implantação para a instituição e seus pesquisadores.

Palavras-chave: Cluster Beowulf, Processamento Paralelo e Linguagem em R.

Instituição de fomento: FAPERJ.