

Graphical Analysis of I/O in High Performance Computing

Denver Coker

University of North Georgia

University of Tennessee & Oak Ridge National Laboratory

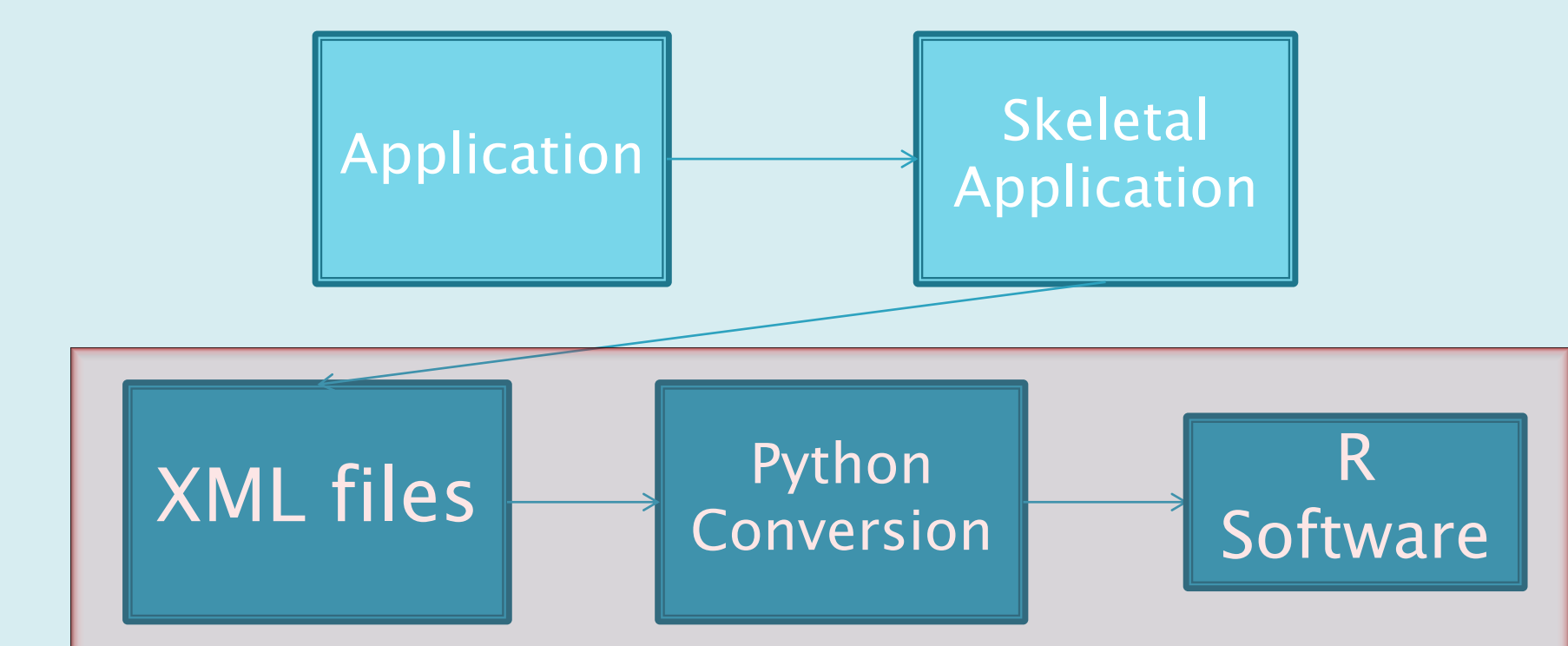
Motivation

- On large machines, such as **Kraken** or **Titan**, machine time is extremely valuable
- I/O is often a **bottleneck**
- I/O optimization can greatly reduce I/O time
- Analysis and understanding of I/O performance is time consuming

Solution

- A skeletal application, built by Skel, observes only the **I/O performance**
- The Skeletal application runs much faster
- The information needed to analyze the I/O remains
- **Saving machine time** and computing hours, thus making parameter studies possible
- The output is imported to the statistical software R

Project Flow



Automated Plot Presentation

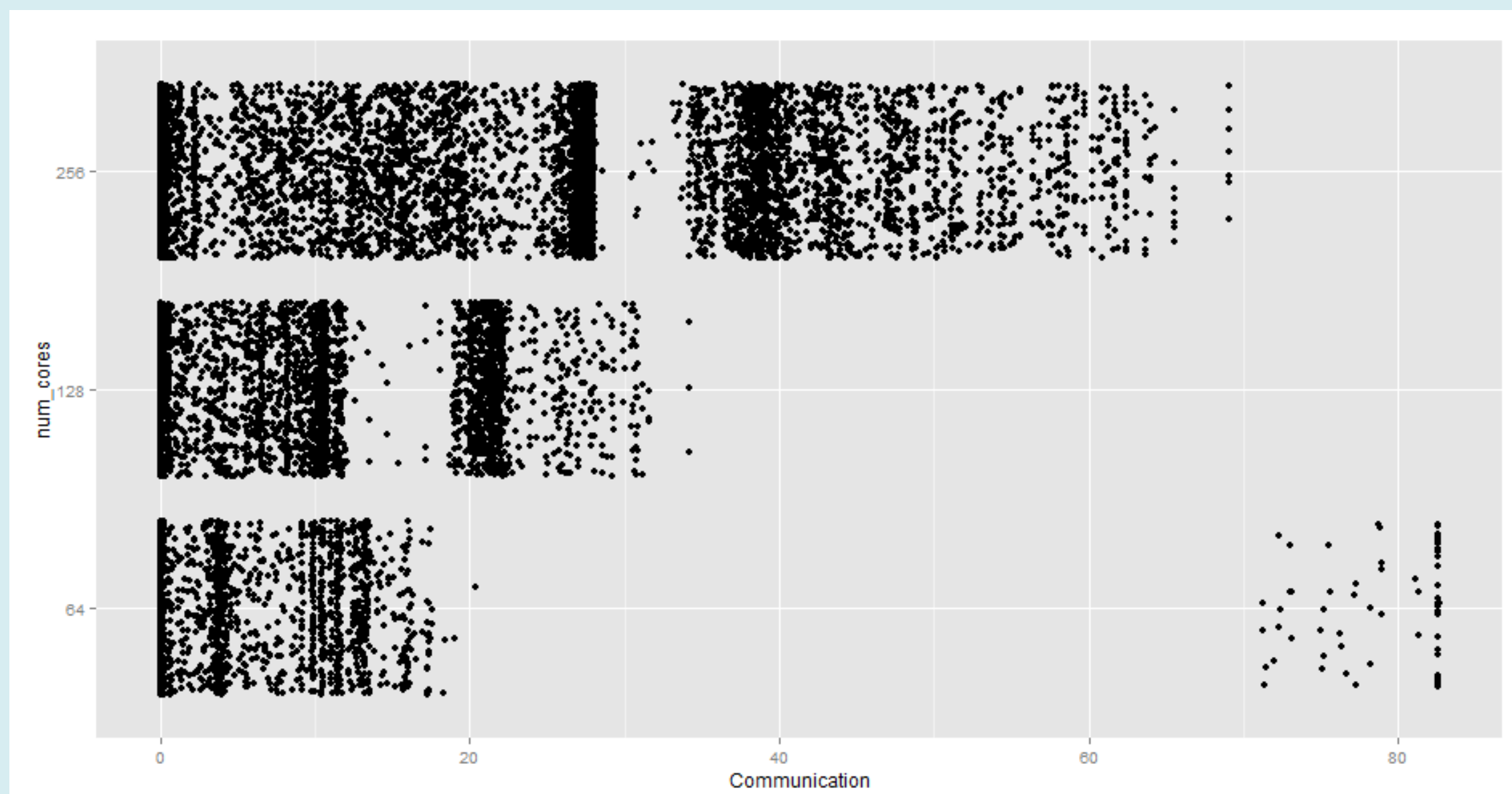
A parameter study has *factors* and *responses*. For each response, graphs of *significant terms* are presented in decreasing order of significance since thousands of combinations are possible. Significance is based on an ANOVA model with all two-factor interaction terms

A significant main effect term

The R code automatically generates the graphs of varying (significant) terms that may exhibit interesting behavior such as outliers.

A significant two-way interaction term

The same graph but colored by the interacting variable shows the outlying group is all the same I/O method



Conclusion

This new method of analyzing I/O performance can

- lead to better understanding, refinement and
- ultimately automated optimization of I/O performance in HPC

Acknowledgements

Oak Ridge National Laboratory
RDAV
Scientific Data Group
University of Tennessee
University of North Georgia
Dr. Scott Klasky
Dr. Jeremy Logan
Dr. George Ostrouchov