Benjamin Stokes, Ph.D.

+1 (801) 839-8993 • stokes@cosmic.utah.edu • github.com/benjamin-stokes
www.linkedin.com/in/benjamintstokes • plus.google.com/+BenjaminStokesPhD

Summary

I am a problem solver with extensive experience collecting, processing, and analyzing large amounts of data in distributed settings. Through a high degree of creative persistence, I have successfully tackled extremely difficult problems. Specific areas of expertise include:

- Big data management
- Data reduction and analysis
- Statistical methods and machine learning
- Algorithm development and implementation
- Distributed systems
- Advanced troubleshooting
- Software architecture and design
- Monte Carlo modeling and simulations
- HPC design and administration
- Strategic problem solving
- International collaboration
- Public presentations

Skills

Computer Languages: C, Bash, R, Go, Python, LaTeX, C++, FORTRAN, MySQL, Tcsh, Visual Basic, Wiki, XML

Data Management, Statistical Analysis, and Machine Learning: C, Bash, NoSQL, R, Python, MySQL, Excel

Distributed Systems and Virtualization: PBS, Rocks, VirtualBox, VMware, Hadoop, MapReduce, Spark, AWS

Frameworks, Environments, and Operating Systems: Linux, Emacs, Windows, SVN, Git, Unix, Xilinx

Natural languages: English *(native proficiency)*, Russian *(elementary proficiency)*, French *(beginner)*

Experience

University of Utah — Salt Lake City, Utah

Postdoctoral Research Associate — *May '10 – present*

I am a research collaborator with the Telescope Array (TA) cosmic ray observatory. As a one-of-a-kind facility, the 150 scientists and technologists in the TA collaboration are responsible for design and maintenance of remote facilities (in Millard County, Utah), data acquisition, data storage, data analysis and public dissemination of resulting scientific knowledge. This mission has required the TA collaboration to develop extensive computational resources and has resulted in groundbreaking discoveries about the origin and composition of cosmic rays. My principal achievements include:

- ***Developing a technique for mapping and reducing parallel computations.*** I pursued this effort with the aim of open-ended scalability and robust fault protection. The resulting software could be described as a highly specialized reinvention of MapReduce implemented on Linux clusters with bash scripting employing a high degree of concurrency.
- ***Designing the primary framework for the Monte Carlo simulation of the TA observatory.*** This effort entailed integrating 40 years worth of legacy code (FORTRAN and C), writing 20,000 lines of new code (C, C++ and NoSQL), and innovating an entirely novel algorithmic approach. The resulting simulations are unprecedented in detail and accuracy.
- ***Engaging in big data management.*** In addition to the two Linux clusters described below, I built and manage a 50 TB Linux data server and manage a 10 TB remote data space. The simulation framework described above is capable of producing 500 GB per day. In order to effectively utilize this data, I have developed a keen sense of foresight towards data logistics.
- ***Becoming a master troubleshooter.*** The TA observatory is operational 24 hours per day, 365 days per year. With operational costs and capital outlays exceeding $100,000 per month, it is imperative that any problems be addressed promptly. As a mid-level collaboration member, I have developed the skills to rapidly troubleshoot a broad array of hardware and software problems.
- ***Engaging in international collaboration.*** The TA collaboration is 70% Japanese and, in addition to the University of Utah, also has member institutions in South Korea, Belgium, and Russia. Working in this diverse setting has taught me to value, above all else, clear communication while respecting the cultural differences and sensitivities of those around me.

Benjamin Stokes, Ph.D.

+1 (801) 839-8993 • stokes@cosmic.utah.edu • github.com/benjamin-stokes
www.linkedin.com/in/benjamintstokes • plus.google.com/+BenjaminStokesPhD

Experience (cont.)

Rutgers, The State University of New Jersey — Piscataway, New Jersey

Postdoctoral Associate — *Jul '08 – Apr '10*

I first joined the TA collaboration as a researcher with the affiliated group at Rutgers University. While my position was officially in New Jersey, I spent most of my time on remote deployment in Utah. In addition to commencing the projects described in the previous section, relevant experience included:

- ***Building and administering two 20-node Rocks Linux clusters.*** This proved to be an important test bed for later distributed computing initiatives.
- ***Learning to work remotely.*** Spending the majority of my time 3000 km from the rest of my research group was an excellent opportunity to develop independence, clear communication, and self-motivation.

University of Hawai'i at Mānoa — Honolulu, Hawai'i

Research Fellow — *Jan '05 – Aug '06*

I spent 18 months in a research fellowship studying radio detection of cosmic rays and neutrinos as a member of both the ANITA and AMBER collaborations. Major activities included:

- ***Designing and debugging embedded systems.*** I developed both embedded software (Linux and C) and firmware (Xilinx).
- ***Designing and running Monte Carlo simulations.*** I designed simulations both by writing new code (C) and by adapting legacy code (C and FORTRAN) with the objective of studying different observation scenarios.

Junior Researcher — *Sep '06 – Jun '08*

After my astrophysics research fellowship ended, I decided to remain in Honolulu and secured a position with the Neuroscience and MRI Research Program. My primary responsibilities included analyzing fMRI data, administering a 25-node Linux cluster, and performing clinical duties. Major milestones included:

- ***Successfully executing a major domain change.*** In making the transition to neuroscience, I quickly learned a whole new nomenclature and worldview. In doing so, I developed extensive strategies for efficiently tackling unfamiliar problems.
- ***Contributing to enterprise software development.*** After attending a two-week training course in North Carolina, I participated in development of the operational source code (C++) of Siemens MRI scanners.

Education

University of Utah — Salt Lake City, Utah

Doctor of Philosophy in Physics

Dissertation: *A Search for Anisotropy in the Arrival Directions of Ultra High Energy Cosmic Rays Observed by the High Resolution Fly's Eye Detector*

Bachelor of Arts in Physics

Timpview High School — Provo, Utah

High School Diploma

Valedictorian

Awards

- Outstanding Postdoctoral Researcher, *University of Utah Department of Physics and Astronomy*
- U.S. Presidential Scholar, *White House*
- National Science Scholar, *U.S. Department of Education*
- Eagle Scout, *Boy Scouts of America*

Other Interests

Among many passions, I am a four-season mountaineer, a classically-trained double bassist, and an internationally published amateur photographer.