Mark Roddy

117 South Oxford Street   Phone: (856) 371-5486
Apt 1   Email: markroddy@gmail.com
Brooklyn, NY 11217   Homepage: http://dibon.homelinux.org/

Employment

- Senior Software Engineer – Mortar Data, January 2013–Present

  - Lead team that won the Best New Feature category in the 2013 Netflix Cloud Prize competition. Under my supervision the team improved Lipstick, an Apache Pig job visualization tool open sourced by Netflix. Contributed new features as well as adding key integration points which allowed Lipstick to be integrated with the Mortar platform.

  - Developed the Mortar Local framework which transparently installs, configures, and manages updates to the Hadoop and Apache Pig runtimes allowing data scientists to author and execute Hadoop jobs in their local environments without having to perform any software installation or configuration.

  - Developed Oz, an internal support platform for managing various system inside of Mortar Data. By providing a central platform for support functionality, which had previously only been available in adhoc scripts, the visibility into system operations greatly increased. In addition, support burden was greatly reduced as all common support functions became trivial to perform.

  - Built Aristo, a tool for continuous deployment. Aristo consists of web application for performing deployment steps and a DSL for easily defining steps to build, test, and deploy code to staging and production environments. The previous deploy processes varied greatly across projects, and each step had a varying level of documentation which made them susceptible human error. Using the Aristo DSL, deploy steps were easily automated and the user facing deploy process for each project was reduced to a single button for the build, test, and deployment steps which was consistant across all projects.

- Senior Software Engineer – Etsy.com, August 2011–January 2013

  - Payments Core Engineering - Responsible for the general availability and automated recovery of internal credit card processing systems. Work includes upgrades to existing systems to provide graceful degradation in the event of downtime in external dependencies and the addition of new services which provide automated detection and recovery of recurring classes of system failures.

  - Developed Hadoop application using Apache Pig for analyzing performance of A/B tests running on the website. This application also served as a prototype for evaluating a possible transition from existing data processing tool chain.

  - Created a library for writing map reduce programs in PHP which abstracted away details of interfacing with Hadoop such as IPC and data type coercion. This allowed Hadoop applications to leverage existing in-house APIs, infrastructure, and knowledge.

  - Automated configuration and software package deployment for hadoop and payment processing servers using Chef.

- Senior Software Engineer – Opera Solutions, January 2011–August 2011

  - Developed a distributed implementation of analytical model for recommending financial news and information based on client portfolios. Using Hadoop and Apache Pig, this new implementation was able to process 24 hours of news information for each client portfolio in under

an hour, improving upon the original implementation which required over 12 hours to process a testing data set of 1/10 the size.

– Created Jython based unit testing tool kit for Apache Pig. This provided functionality for automatically verifying the correctness of analytical models which reduced the overall defect rate. By using Jython to interface with Java based Hadoop and Pig APIs, this approach allowed developers more familiar with Python to quickly iterate on models in Pig via standard unit testing techniques.

– Developed work flow system for Apache Pig which provided basic flow control and segmentation of individual models while still allowing for shared consumption of data extraction and cleansing pipeline.

– Mentored junior developers on software engineering techniques for implementing analytical models and process automation.

- Research Developer – Ackoff Center for Advancement of Systems Approaches, University of Pennsylvania 2006–2011

    – Developed framework for designing and conducting monte carlo simulations on software models created using in-house Agent Based Modeling (ABM) tool. Supports conducting both stochastic and monte carlo analysis. Also developed external API for easy integration with third party tools and models which reduced the time to completion for integration with third parties from weeks to days.

    – Implemented a data storage back end for retaining simulation state to be used in post run analysis. Previously, all data storage was done on a per model basis which resulted in duplicated work between many projects.

    – Developed a system for analysts to design models for prediction of abstract events and situations based upon the results of a completed simulation. These high level situations better express the outcome of a simulation than individual agent decisions making it easier to communicate the overall result of an experiment.

    – Leveraging the tools above, created a customized software stack for the Lockheed Martin Advanced Technology Lab as part of the DARPA sponsored ICEWS project (International Crisis Early Warning System). Due to the customizable nature components in the above tools, was able to develop an initial prototype in less than three days unlike the previous integration with the same group at Lockheed which took two months to produce the initial prototype.

Education

- B.S. Computer Science–Rutgers University, New Brunswick, NJ

Publications

- Modeling and Simulation Methodologies–A Systems Approach to the Social Sciences, In: Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains, Wiley Publishing, 227-270, Barry G. Silverman, Gnana K. Bharathy, Benjamin Nye, G. Jiyun Kim, Mark Roddy, and Mjumbe Poe

Skills

- Python
- OOP
- Unit Testing

- Hadoop
- Apache Pig
- Map/Reduce

- Bash
- Linux
- Chef