

Clay Baenziger

CONTACT

INFORMATION *Blog:* <http://clayb.net/blog>

CITIZENSHIP USA

OBJECTIVE To find an exciting position providing technical leadership for a research or cutting-edge industrial computer science organization – a position allowing hands-on interaction with mathematics, economics, open-source and the Unix computing environment.

RESEARCH INTERESTS Data analysis (automated data gathering, statistical modeling and presentation), Object oriented development (Python, C++ and Java); Visual communications (UML, data presentation), Structural data modeling (XML, relational database design); Group presentation technique (group education, business meeting and conference settings)

PROFESSIONAL EXPERIENCE

Bloomberg *2013 – Current*

2013 – Current

Enterprise Big Data

- First Engineering team-member tasked with developing operation for Hadoop-as-a-service for all Bloomberg Engineering teams
- Led development on [opensource Hadoop](#) deployment platform
- Grew team to six full-time engineers converting one intern, two contractors, three internal transfers and one external senior hire before handing team off
- Represented Bloomberg to various vendors in multi-year, multi-million dollar engineering and support engagements
- Contributor to [Apache HBase](#), [Apache HAWQ](#)

External Presentations:

- "Open Source Recipes for Chef Deployments of Hadoop" – Hadoop Summit San Jose, June 4th, 2014 ([slides](#))
- "Chef Patterns At Bloomberg Scale" with Amit Anand and Biju Nair – Chef Conf, April 1st, 2015 ([slides](#))
- "Impromptu Interview at Chef Conf" – Chef Conf, April 1st, 2015
- "Coproductors - Uses, Abuses, Solutions" with Esther Kundin – HBase Con East, September 26th, 2016 ([slides](#))
- "Cluster Continuous Delivery with Oozie" – Apache Con North America - Big Data, May 18th, 2017 ([slides](#))
- "Multitenancy At Bloomberg - HBase and Oozie" – Dataworks Summit, June 14th, 2017 ([slides](#))

2013 – 2014

Enterprise OpenStack

- First Engineering team-member heading an OpenStack team to build out OpenStack-as-a-service for the entire 4,000 person Engineering organization
- Developed a Jenkins continuous integration system for our [opensource OpenStack](#) platform
- Built a proof-of-concept continuous integration environment for deploying Bloomberg's consumer-level mobile API service on top of OpenStack
- Mentored an intern (converted to employee) developing a firewall-as-a-service plugin to OpenStack Neutron to drive physical Palo Alto firewalls per VM

Opera Solutions, Global Markets

2011 – 2013

Consulting

- Brought Apache CloudStack into the organization (filing three bugs for the upstream community, producing user documentation for end-user use and documenting install and configuration for future administrators)
- Brought Hadoop into the organization starting on Amazon EC2 moving to a 21 physical node Hadoop Map/Reduce and HBase cluster (managed with Puppet and Cacti) leveraging Amazon EMR for spill-over of map/reduce jobs
- Reverse-engineered and translated SAS code producing Apache Pig. Wrote unit-tests to show equivalence of Pig to original SAS
- Designed a multi-tier architecture for an Internet facing web application. Wrote prototype code for all layers in the stack from Pervasive Data Rush, Apache Avro, Apache Pig, Apache Oozie on Cloudera Hadoop populating a MongoDB and served by LifeRay Portal.
- Provided leadership and committed 43K lines of code per year from start to production on a Fortune 500 banking Hadoop, Pig, Oozie and Jython 3-tier project (otherwise using MS SQL and IBM Websphere)
- Communicate project status and direction, both to internal stakeholders, as well as client technical staff and project management on a daily basis
- Provide executive management direction and data-flow, runtime and network architectures on areas of strategic concern or capital expense. Examples are: Platform-as-a-Service (Hadoop) analytics, Infrastructure-as-a-Service (Cloud Stack) compute virtualization, network security isolation (particularly how to provide internal and external facing Atlassian JIRA and Git source repositories), amongst others.

Sun Microsystems/Oracle, Solaris Engineering

2004 – 2011

2007 – 2011

Software Engineer

- Communicated and worked in a distributed team of 30 engineers spread across 10 sites spanning North America, Europe and Asia
- Maintained a working understanding of the legal environment surrounding free and open-source software, analyzed 20 years of licenses used by Solaris install to go live with OpenSolaris
- Developed documentation and presentations for company and external audiences (slide decks, blog – achieving over 2,500 unique visitors a month, career fairs and academic colloquium presentations)
- Researched new technologies to improve team efficiency; further, presented and educated my team (moved team from primarily C and shell-script code to object-oriented Python code, introduced UML for better architecture communications and a continually pushed a better understanding of XML standards)
- Led Solaris Automated Installer webserver and XML installer manifest feature development; further, I was the leading filer of bugs for the OpenSolaris install team in 2010.
- Performed nightly builds of complete OpenSolaris distributions, eventually automating and performing initial Hudson server integration
- Subject matter expert on the OpenSolaris Automated Installer; worked implementing LZMA de-compression for in-kernel lofi(7) driver and performing extensive work calling the C++ compression code from user-land C code; kernel debugger capable and DTrace proficient
- Had a \$100,000/year budget for new hardware acquisition and bring up, focusing on driver compatibility and operating system feature completeness requiring extensive driver tweaking to bring-up new hardware
- Created Python Ctypes bindings between OO Python APIs and underlying C system API (Solaris SMF and ZFS); also produced code using the Python C-

API for both calling C library code in shared libraries and calling Python from C shared libraries

2004 – 2007

Systems Administration

- Developed Perl, PHP and shell-script products to save engineering time through lab automation resulting in a Metal-as-a-Service (MaaS) web application
- Provided support for lab network, serial, fibre and power cabling and routing
- Deployed and retired assets throughout lab of roughly 600 machines; further performed automated data clean-up and data visualization on group-wide MySQL/PHP based lab asset database (containing almost 4,000 pieces of equipment)
- Participated in new-lab bring up (from design to rack-and-stack) as well, participated in old-lab shutdown (from design to asset redeployment)
- Resolved interoperability issues (e.g. multi-vendor and pre-production hardware/firmware incompatibilities)
- Produced documentation for lab users, as well as best practices for affiliated labs

University of Colorado at Boulder

1997 – 2004

2002 – 2004 Discovery Learning Center

Software Engineer

- Redesigned public interface to 16 panel video wall and 108 seat conference center
- Employed: AMX, Barco, ClearOne, Clarion, Crown, Extron and Jupiter equipment
- Produced 20,000 lines of embedded systems control code, and documentation
- Reported to College of Engineering Assistant Dean of Administration
- Maintained building routing, firewall and tenant lab networks and kept conference center technology running during development work

1999 – 2002 CS Ops., Dept. of Computer Science

Systems Administration

- Worked with seven full-time coworkers managing approximately 750 machines running:
NEXTSTEP, FreeBSD, IRIX, OpenBSD, OS/MP, RedHat, Solaris, and various flavors of Linux
- Developed multi-machine room temperature monitoring system using Perl and RRDTOOL to provide web based on-demand graphs and heuristic based SMS alarms
- Setup RedHat Kickstart based automated installation environment for an instructional lab of 50 machines

1997 – 1999 Department of Applied Mathematics

Systems Administration

- Performed backups using Amanda; maintained a public Apache webserver; produced HTML markup
- Provided desktop page-layout abilities; including work on a faculty directory for the successful receipt of a \$2.3 million NSF grant

Independent Consulting, Multiple Individuals and Small Offices

1996 – 1998

Consulting

- Acquired new clients and ensured productivity of existing clients
- Educated clients on optimal computing techniques
- Led clients in describing problems and provided resolution

FORMAL
EDUCATION

Colorado School of Mines, Golden, Colorado USA

Incomplete Masters Studies, Master of Science in Engineering and Technology Management, August 2007 – 2010

- Accounting
- Ethics
- Patent Filing
- Project Management
- Technology Product Development

Bachelors of Science in Mathematics, June 2007

- Focus in statistics
 - Confidence Intervals
 - Markov Chains
 - Probability Distributions
 - Regression Analysis
- Minor in Economics and Business
 - Micro- and Macro-Economic Theory
 - Operations Research
 - Linear Optimization

Aims Community College, Greeley, Colorado USA

Associate Studies, Photography, September 2002 – September 2004

CONTINUING
EDUCATION

Presenting Data and Information, Edward Tufte **Data Visualization**

- Techniques for presenting statistical data: tables, graphics and animation
- Achieving multi-dimensional scientific visualizations
- Data presentation dealing with complexity and ensuring clarity

Reliability Engineering, David Trindade **Applied Reliability**

- Techniques for modeling failure in repairable systems
- Techniques for modeling failure in non-repairable systems
- Worked through P.A. Tobias, D. C. Trindade, *Applied Reliability*, 3rd ed., Chapman and Hall/CRC

ACTIVITIES

Reviewer Unix and Linux System Administration Handbook, Nemeth, et al., 4th ed., 2010

Front Range OpenSolaris User Group, 2005–2010

- Advisory member of open source meet-up with attendances from 10-50 people monthly; leader from mid-2008 until suspension of OpenSolaris program in mid-2010
- Managed logistics of meetings in a diversity of environments, working collaboratively with local businesses, universities and hosted at the Sun Microsystems campus
- Provided meeting material for a variety of audiences (proximate, live-webcast and archived recordings)
- Presentations:
 - February 28th, 2008 – "LZMA Compression Algorithm"
 - April 24th, 2008 – "Debugging OpenSolaris"
 - February 26th, 2009 – "Creating Custom OpenSolaris USB Sticks"
(co-presented with Jack Schwartz)
 - March 26th, 2009 – "Bugs!: Finding and Filing"
(co-presented with Patrick Korsnick)
 - May 28th, 2009 – "OpenSolaris - Automated Install"

- September 24th, 2009 – "SMF Bridge for Python"

Colloquium Presentation, University of Denver, Department of Computer Science, October 30th, 2009 – "OpenSolaris Operating System for Users and Application Developers"

Funkiphino, 2004–2008

- Ran sound for a traveling, for-hire, funk band ranging between 8-14 members
- Troubleshoot house sound systems at a variety of venues across Colorado
- Provided DMX based theatrical lighting for especially large events
- Performed mixing for TV broadcast events

"LANL Supercomputing Data Analysis", Colorado School of Mines Technical Report, 2007

- Over a six week period, investigated a 4.5 million record data set provided by Los Alamos National Labs chronicling super-computer failure and usage data for five super computer systems/clusters in a four person team
- Did initial clean-up and data quality analysis (discovery of bad data, corrupted data)
- Performed data categorization for failure (i.e. sub-system groupings) and usage data (summary statistics, user profiles, job characteristics)
- Performed rudimentary regression and factor analysis to discern patterns potentially leading to failure cases

House Manager, Sigma Nu Fraternity, Colorado School of Mines, 2007

- Resolved mechanical and electrical issues in a 40 year old fraternity house
- Acted as liaison between University and Fraternity on issues of house occupancy, maintenance and janitorial services
- Led Fraternity in house improvement projects and maintenance duties

TECHNICAL COMPETENCIES

Build Maintenance Tools: Apache Ant, Maven, SVN, CVS, Git, GNU Make, Jenkins, Mercurial, RCS, SCCS

Databases Programs: MS SQL 2005/2008 (query tuning), MySQL (used for data-analysis with GNU R), Oracle 10i (produced queries only), RRDTool (used for data recording and HTML/CGI information display), SQLite3 (interfaced with Python)

Digital Media Tools: Adobe After Effects, Illustrator, InDesign, and Photoshop, Apple FinalCut, GNU Image Manipulation Program, L^AT_EX, MediaWiki, OpenGL, Xfig

Mathematics Programs: AMPL, Frontline Solver, Microsoft Excel, GNU R, MATLAB, Mathematica, OpenOffice, SAS

Networking Protocols and Management: ARP, Cisco IOS (router rules and 802.11q VLAN configuration), DHCP (ISC server), DNS (multicast, ISC BIND), Extreme ExtremeWare (routing rules and 802.11q VLAN configuration), HTTP[S] (Apache, CherryPy), Intel PXE boot, Mail (Sendmail, SMTP, IMAP, POP), SNMP, TCP/IP, UDP

Operating Systems: Apple OS X, Linux (Debian and RedHat), BSD, Solaris, and other UNIX variants

Programming Languages: Apache Hadoop, Apache Oozie, Apache Pig, AppleScript, C, C++, Java, JavaScript, Perl, PHP, Python, UNIX shell scripting (bash, ksh93), SQL

Project Management: Microsoft Project, Oracle Primavera, Scrum & Waterfall Methodologies

Miscellaneous: Working understanding of electronics test equipment (oscilloscopes, signal generators, multi. meters), Amateur (Ham) Radio Licensee – Technician Class, USITT DMX512 lighting protocol, working knowledge of German, USENIX and LISA member, chooses not to use Microsoft Windows for general purpose computing

REFERENCES

Available upon request