Clay Baenziger

Contact Information	Boulder, CO 80301	*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 Linux computing environment.

Research Interests   Data analysis (automated data gathering, statistical modeling and presentation); Open source and software architecture; Apache, Eclipse and CNCF Foundation projects in Python and Java; Visual communications (UML, BPML and data presentation); Structural data modeling (XML, (non)relational database design, file-format efficiencies Apache Parquet and Apache Arrow); Group presentation technique (team education, documentation design and conference effectiveness)

Professional Experience

Bloomberg   *2013 – Current*

Engineering Culture

2016 – Current   **Distributed Systems Guild**

- Inaugural Member
- Represented Distributed Systems Guild interests to peer guild chairs
- Managed conference attendance and information dissemination with over 1,000 contacts per year across Bloomberg engineering before mentoring new Guild leaders to advance conference attendance further
- Brought in influential external speakers on technologies of interest and facilitated internal, engineering-wide training events
- Have mentored new Guild leaders running talk series achieving consistent attendance in the hundreds of engineers

2019 – Current   **Open Source (Internal) Meetup**

- Inaugural member of internal "Open Source Meetup"
- Became adjunct member of CTO Open Source Program Office for outbound open-source software reviews
- Review numerous contribution requests from co-workers (participated in mentorship achieving numerous successful HBase, HDFS, core Hadoop and Hive contributions by co-workers)
- Organize speakers presenting on licensing, community building and documentation on external best practices and etiquette
- A mentor for contributing open software software externally

Project Teams

2021 – Current   **Cloud Native Compute Services**

- Team operates and develops a Golang platform for broad-scale Kubernetes API server federation using Kine (Postgres)-backed K8s API servers for state storage, CRD's for run-time platform definitions and controllers for credential injection
- Productionized changes to Hadoop allowing secure access to Kerberized services from non-Kerberized clients

2014 – 2021   **Data and Analytics**

- Started an engineering team tasked with introducing Hadoop to all Bloomberg Engineering
- Managed seven full-time engineers growing team via one intern conversion, one junior hire, two contractors (converted), two internal transfers and one external senior hire as well as managing the work of three internal "visiting engineers" before handing team off to support remote work
- Led open development on a Chef and Ruby based Hadoop deployment and management platform
- Represented Bloomberg to various vendors in multi-year, multi-million dollar co-engineering and support engagements on Apache HBase, Apache Kafka, Python, the Linux OS
- Various Java, shell script and C contributions to Apache HBase, Apache HAWQ, Apache Hadoop HDFS, Apache Oozie, Apache Ratis, procps (pgrep(1))

External Presentations:
- "Open Source Recipes for Chef Deployments of Hadoop" – Hadoop Summit San Jose, June 4th, 2014 (slides)
- "Open Source Recipes for Chef Deployments of Hadoop" – San Francisco Hadoop User Group, July 16th, 2014 (slides)
- "Chef Patterns At Bloomberg Scale" with Amit Anand and Biju Nair – Chef Conf, April 1st, 2015 (slides)
- "Impromptu Interview at Chef Conf" with Amit Anand – Chef Conf, April 1st, 2015
- "Smooth Operators Panel" – HBaseCon, May 7th, 2015
- "Coprocessors - Uses, Abuses, Solutions" with Esther Kundin – HBaseCon 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 San Jose, June 14th, 2017 (slides)
- "Breathing New Life into Apache Oozie with Apache Ambari Workflow Manager" with Artem Ervits – DataWorks Summit Berlin, April 19th, 2018(slides)
- "Breathing New Life into Apache Oozie with Apache Ambari Workflow Manager" with Artem Ervits – DataWorks Summit San Jose, June 20th, 2018(slides)
- "Breathing New Life into Apache Oozie with Apache Ambari Workflow Manager" with Artem Ervits – DataWorks Summit San Jose, June 20th, 2018(slides)
- "Of Data Dropboxes and Data Gloveboxes" – StrangeLoop, September 28th, 2018 (slides)
- Moderator HBase Birds-of-a-Feather with Artem Ervits – DataWorks Summit Barcelona, March 20th, 2019
- "Data Gloveboxes: A Philosophy of Data Science Data Security " – DataWorks Summit Barcelona, March 21st, 2019 (slides)
- "Intermediate HBase Operations" – NoSQL Day, May 21st, 2019 (slides)

Initiatives:
- Successfully provided novel infrastructure to for key teams required by the Bloomberg Professional Service:
  - Was infrastructure liaison for teams migrating end-of-day time-series product from DB2 and large trading history system from Oracle
  - Provided database systems capable of returning 10's of billions of queries a day with 5ms median SLO

- Developed dashboards and training for use by operators for ensuring SLO compliance
- Championed Data Glovebox for secure data science workflows (with core Hadoop WebHDFS change)
- Developed cluster driven application deployment pattern (including Apache Oozie support)
- First broadly Kerberized infrastructure deployment for Bloomberg
- Leveraged "BFD" via OpenBFD for host influenced BGP routing with static routes

2013 – 2014   **Enterprise OpenStack**

- First Engineering team-member heading an OpenStack team to build out OpenStack-as-a-service for the then 4,000 person Engineering organization
- Developed a Jenkins continuous integration system for our opensource OpenStack platform from bare-metal, to OS install and Open Stack platform with monitoring up and running
- Built a production deploying continuous integration environment for deploying Bloomberg's consumer 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 firewall configuration per VM

Initiatives:
- Pioneered the first open-development (open source code, developed openly) team model at Bloomberg

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 into Apache Pig. Wrote unit-tests to show equivalence of Pig to original SAS equivalence.
- 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 commercial banking portfolio analytics application based on Hadoop, Pig, Oozie and Jython designed as a 3-tier application (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), among 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 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
- Managed 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**

- Managed client relationships
- Ensured productivity of clients in artistic, pre-press and professional service occupations
- Educated clients on optimal computing techniques
- Led clients in describing problems to provided resolution

FORMAL
EDUCATION

Colorado School of Mines, Golden, Colorado USA

Studies toward Masters of Science, Master of Science in Engineering and Technology Management, August 2007 – 2010 Courses completed:

- 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 (Optimization)
  Willingness to Buy Models

Aims Community College, Greeley, Colorado USA

Associate Studies, Photography, September 2002 – September 2004

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

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

Cited 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 meet-ups in a diversity of environments. Collaborated with businesses, hosting was negotiated between the Sun Microsystems campusand universities (CU Boulder, Colorado School of Mines) and a national lab (National Center for Atmospheric Research)
- 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)
    – April 2nd, 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
- Troubleshot 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

MISCELLANEOUS: Working understanding of electronics test equipment (oscilloscopes, signal generators, multi. meters), Amateur (Ham) Radio Licensee – Technician Class, Class-A Commercial Drivers License with tractor trailer, tanker and hazardous materials endorsements, working knowledge of German, chooses not to use non-open source operating systems (Windows, OS X) for general purpose computing

REFERENCES   Available upon request