Dr. Geena Louise Rollins, BS,MS,Ph.D. Computer Science • geena@leapfire.com • in/rollins • (408) 823-6040 • Sunnyvale, CA

Strengths

- Distributed Systems, Distributed Data, Query & Analytics
- Innovative, Influential, Reinventing the Future, Business Acumen, Cross Functional
- Recognized by outside peers and companies
- Architectural Design, Coding, Wide horizontal, Deep vertical experience
- Mentoring/Coaching, Diversity & Inclusion
- Relational, NoSQL, Hadoop & Spark, Agile leadership, web applications, Microservices

Major Influential Accomplishments

- 1997: Developed one of the first comparison shopping search engines. Founders sold company in one year to Inktomi.
- 2004: Founded and sold the first Contextual-Advertising company. Beat Google AdSense by 2 full years. Sold company a year later.
- 2007: Lead architect/engineer for CBS.com, serving 40M unique visitors a month, satisfying 20,000 SQL queries per second.
- 2014: Distributed-Data Architect for Symantec Norton Engineering, a $2B/year division. Consulted for over a dozen use cases.
- 2017: Query & Analytics, Time-Series Event Systems. Developed innovative technique for Elasticsearch to solve Star Joins delayed until query time.
- Track record of reinventing the future several times: comparison shopping, screen scraping, contextual ad-targeting, delayed-join, …

On the job Experience

- Java since 1997**,** Scala, Linux, SQL, Agile
- XML, Avro, JSON, REST APIs, Microservices, AWS, Azure, OpenStack, multi data-centers.
- Elasticsearch, Lucene, SOLR, MySQL, Clustrix, Spring, Hibernate, Kafka, Hadoop, Spark
- NoSQL since 2009: Redis, memcached, MongoDB, HBase, Cassandra, DynamoDB, Couchbase
- Distributed. high-performance, highly available, multi-threaded, loosely-coupled, service oriented, scalable architectures
- Tradeoffs between Consistency, Availability, Partition Tolerance
- Analytics: real-time, near real-time, batch
- Technical Presentations
- Performance improvements for high-traffic web site (10 million page views per day).

Positions

- June 2017 – Present: **Adobe**, San Jose, CA, *Senior Computer Scientist II*

  Software developer working in the Machine Learning Framework team. Batch Analytics with Spark/MLLib in Scala. Still learning the vast area of ML on the job.

  Worked on the Adobe Pipeline project that defined Microservices over raw Kafka. For all Adobe.io Microservices that define a _query endpoint, I designed *AQL: Adaptable Query Language*. Implemented compilers AQL -> SQL and AQL -> EQL (Elasticsearch) to enable applications to query relational, Elasticsearch, or both. Gave talk on the *Join-Compare Solution* described below.

- Jan 2017 – June 2017: **Leapfire Solutions, LLC**, Sunnyvale, CA, *Elasticsearch Consulting Architect.*

  Same work as at Symantec, below, but for external customers. Worked for two customers that each were building Time-Series Event Analytics systems that consumed 10,000 events per second. Designed an architecture based on Elastic Stack and included Kafka to enable disaster recovery and to level event-traffic spikes. This was critical for a company that builds network switches and did not want that hardware to absorb back pressure from the ETL program (Logstash). Other components pass the buck upstream; with Kafka, *"The buck stops here"*.

  For scatter-gather distributed queries, common advice is to denormalize, denormalize, and then denormalize. Premature denormalization of event documents eventually leads to stale information and missing query results in future investigations. I designed the *Join-Compare Solution*, an innovative technique that solves this problem for many use cases. Implemented Join-Compare in Java; deployed to AWS.

  I developed a technical presentation on the *Join-Compare Solution* in practice with performance results; gave talks.

• Feb 2014 – Dec 2016: **Symantec**, Mountain View, CA, *Senior Principal Software Engineer*

Distributed Data Architect for Norton Engineering, the $2B/year consumer-products business-unit.

Collaborated with the Norton Chief Software Architect. Together we designed the *Norton Distributed Data Stack* (NDD Stack) as a cloud-ready reference architecture for cost-effective scaling of big data. NDD Stack, comprising open source components that complement each other, is an excellent choice for many Use Cases. Cassandra for cross-data-center Key-Value Store, Elasticsearch as an Indexing Engine with unparalleled ad-hoc query performance, Spark as a Compute Engine for sorting, joining, and map/reduce, and Kafka as the shock absorber. Gave talks on NDD Stack and Elastic Stack training for several Norton Engineering product development groups.

Collaborated with over a dozen product software architects to develop their application architecture adapting the NDD Stack to their Use Cases. I did data modeling, query design, NDD training presentations, performance testing, cluster topology/operations planning, and more. I was taking whole Java-Oracle teams and transitioning them to become good distributed-database programmers.

These projects were either new applications or re-architecting of existing applications with disparate data sources and several application levels. Many, but not all, indexed and analyzed security events such as device-pings and detected-threats.

Developed first design and implementation of Join Compare; later at Leapfire, re-implemented from scratch.

I was Symantec's leading Near-Real-Time Analytics expert and was a member of Elastic's Customer Advisory Counsel.
Elastic asks *"How did Geena become so knowledgeable of Elasticsearch?"*
https://www.elastic.co/blog/elasticsearch-support-an-investment-that-keeps-paying-off-at-symantec

• May 2013 – Jan 1014: **Disney Interactive**, Palo Alto, CA, *Lead Software Engineer*

Developed Java back-end REST APIs for use across many mobile apps in Java with MongoDB.

• Jan 2012 – Nov 2012: **Yahoo!, Inc.**, Sunnyvale, CA, *Software Engineer*

Developed Hadoop software in Java for behavioral ad targeting. Introduced key-value store HBase to the targeting platform to reduce expensive shuffling of data. Developed Hadoop Map-Reduce jobs processing up to 2TB/hour across 1,000 nodes. Developed performance-test harness enabling the use of subclasses defining HBase schemas to compare different object modeling proposals.

• Oct 2010 – Nov 2011: **Software Consultant, Company-Founding Investigation**

Designed and implemented a web-site back-end using Java, Redis, MySQL, JSON, Spring MVC and Jetty.

• Dec 2007 – Sept 2010*:* **CBS Interactive**, San Francisco, CA, *Software Architect*

Software Architect and Lead Engineer for team that replaced the old static web site to make CBS.com a dynamic site with social activities. *"During her three years with the network's digital group, revenue increased by 700% and the network's portal, CBS.com, increased its audience tenfold."* Designed the CBS.com back-end as a Java Tomcat application with Spring MVC, Hibernate, mySQL, Linux, Perl, JSON, XML, memcached, Redis and Lucene.

Architected and personally implemented several mission-critical feature sets, doing all coding myself:

- Queryable Video Feed (QVF) meta-data integrator and query engine for all videos on CBS.com.
- Configurable Fantasy Games for TV Shows similar to Fantasy Football (*e.g.,* Fantasy Survivor, Fantasy Amazing Race, …)
- Contests (*e.g.,* to predict Grammy Award winners).
- Integrated CBS.com with Facebook Connect doing all front-end Javascript and back-end Java work.
- CBS.com's first-ever Site Search, using Lucene and exposed as a JSON REST API (Search as a Service).
- Extreme Voting API for high-traffic high-stakes events (*e.g.,* Grammy and Academy of Country Music Awards Shows).

Collaborated on feature development and interaction with marketing product managers and software engineers within our team, outside CBS groups and partner companies such as Grammy, ACM, Victoria's Secret, Sprint, and others.

• July 2006 – Nov. 2007: **MetroFi**, Mountain View, CA, *Software Architect*
The municipal WiFi broadband ISP with free ad-supported networks nationwide.
Ad Server in Java, JSPs, JavaScript and AJAX.

• Nov. 1999 – Jan. 2006
Tribal Fusion, Emeryville, CA*, Chief Contextual Architect*
Leapfire Technologies, Inc, Sunnyvale, CA, *Founder, Software Architect*
Tribal Fusion acquired (Startup) Leapfire in April 2004.

Principal Technical Lead of the Context Project. Joined when Tribal Fusion had 15 employees; by 2006 grew to over 100 employees and served 11 billion monthly impressions. The *Context Software* implements linguistic text analysis to target advertising to the content of web pages. It is a highly-available distributed multi-threaded Java system using AJAX techniques for ad displays, Java Tomcat application server, and JDBC to Oracle database (first version used mySQL).

At Leapfire, designed and implemented software that links products, text ads, and relevant information to content of web pages. It is a proprietary linguistics-based text-filtering system built with Java using scalable distributed multi-threaded implementation. Text-ads campaign on Associated Press had an astounding 1.68% Click-Through Rate over six months across a variety of news areas including health, entertainment, business, sports, technology and more.

Leapfire placed text ads on AP Wire stories on over 400 newspaper web sites nationwide. Designed and implemented proprietary map-reduce application to implement distributed text-analysis algorithms (*e.g.,* to compute document frequencies of two-word phrases over a corpus of Associated Press news articles over a three-year period).

• Aug. 1997 – Sept. 1999
Inktomi Corporation, San Mateo, CA*, Software Engineering Manager*
C2B Technologies, Inc, San Mateo, CA, *Director of Software Technology*
Inktomi acquired (Startup) C2B in Sept 1998

Designed the first version of the *Shopping Engine*, Inktomi's third product line. Lead software engineer for the *Integrated Order Process* (IOP) designed to integrate portals' shopping applications with merchants' web stores. All projects were in Java.

At C2B, lead software engineer for the *Shopping Engine*, C2B's flagship product, a pioneering *comparison-shopping* query engine. Designed the overall software architecture for this system, detailed design, and much of the software development. Did presentations and technical due diligence for prospective customers, investors, and acquisition partners.

The C2B Shopping Engine is implemented as multi-threaded Java Servlets and connects to a relational database system through JDBC, and a geographic database through sockets. Invented (independently at a similar time with Junglee) the robot technology known as *Screen Scraping*; developed a spidering and scraping technology used to build the multi-merchant product database and to query merchant web sites in real-time.

US Patents Issued: ** Primary Inventor*

7,415,429 * Providing navigation objects for communications over a network
7,363,248 * Pre-filling order forms for transactions over a communications network
7,349,867 * Tracking transactions by using addresses in a communications network
6,385,602 Presentation of search results using dynamic categorization

Refereed Publications

- D. Jackson and G. L. Rollins, "A New Model of Program Dependences for Reverse Engineering", 12/1994.
- D. Jackson and G. L. Rollins, "Abstraction Mechanisms for Pictorial Slicing", 11/1994.
- R. Harper, P. Lee, F. Pfenning, and G. L. Rollins, "A Compilation Manager for SML/NJ", 6/1994.
- J. M. Wing, G. L. Rollins and A. M. Zaremski, "Thoughts on Larch/ML and a new application for LP", 7/1992.
- G. L. Rollins, "SourceGroup: A Selective Recompilation System for SML", 9/1991.
- G. L. Rollins and J. M. Wing, "Specifications as Search Keys for Software Libraries", 6/1991.
- G. L. Rollins, "A Simple System for Object Storage in Common Lisp", 6/1991.
- P. Lee, F. Pfenning, G. L. Rollins and W. S. Scherlis, "The Ergo Support System: An integrated set of tools for prototyping integrated environments", 11/1988.