Minor E. Gordon, Ph.D.
Curriculum Vitae

cv256@minorgordon.net
http://mg4.consulting   Tulsa, OK

Education

2009	**Ph.D., Computer Science**, University of Cambridge Dissertation: Stage scheduling for CPU-intensive servers
2005	**Diplom Informatiker**, Technische Universität Berlin Diplomarbeit: Staged design for highly concurrent web servers
2003	**B.Sc., Computer Science**, Oklahoma State University

Skills

Languages	English (native); German (fluent)
Programming	Java/J2EE; Python; C++; SQL; C#; TypeScript; PHP; JavaScript; HTML+CSS
Frameworks	Thrift; Protobufs; Guice; Guava; Vaadin; GWT; Shiro; jQuery; Bootstrap
Formats	XML; JSON; CSV; YAML; XDR
Testing	JUnit 4; Hamcrest; jMock; Jenkins; Pyunit; nose; gcov, gprof
Documentation	Doxygen; epydoc; Javadoc; Jsdoc
Databases	ElasticSearch; MongoDB; Redis; SQLite; H2; SQL Server
Platforms	Windows; Debian-, RedHat, and SuSE-based Linux; OS X
IDEs and tools	Visual C++; Eclipse CDT/JDT/PyDev; PyCharm; Maven; CMake; gcc, gmake, gdb
Revision control	Git; Mercurial; Subversion
Administration	Shell scripting; Nagios; Ganglia; Apache httpd and Tomcat; nginx

Commercial Projects

9/2014 – ongoing Part-time	**Software architecture and backend development for the Notablist email newsletter search engine** New York, NY USA and remote
	Notablist indexes millions of email newsletters from hundreds of thousands of publishers and provides both a search interface and real-time change alerts to users.
Activities	Prototyping in Python: • Classified signup responses and mapped form inputs using numpy/scipy/scikit-learn Production rewrite in Java: • Generated Python JSON-RPC clients, Python command line tools, Java service interfaces and abstract implementations, and Java JSON-RPC servlets from Thrift interface and data structure definitions • Implemented Guice-injected Java micro-services that encapsulated MongoDB collections; ElasticSearch indices; Redis databases; S3 buckets; screenshotting and signups with Selenium; SpamAssassin checking; logstash queries; Stripe integration; Drip (CRM) integration; MailChimp, SparkPost, and SendGrid transactional email posting; etcd locking; DNS and whois querying and parsing; DMOZ category and Alexa and Quantcast rank lookups • Designed and implemented a scalable distributed system for processing Common Crawl WARC records and submitting newsletter signups • Created administrative user interface in Vaadin • Docker+Compose deployment on real hardware

7/2012 – ongoing Full-time then part-time	**Full-stack development of the <u>TeraScript</u> web application server** Tronics Software LLC / Stirling, NJ USA and remote
	TeraScribe and TeraScript Server are a visual development environment and server-based runtime for TeraScript Action Files (TAFs), TeraScript Class Files (TCFs), and the TeraScript Markup Language (TML). TML is a markup-based web application language, similar in spirit to ColdFusion. TAF and TCF are ways of organizing TML.
Activities	<u>TeraScript Server 8</u>: Rewrote TeraScript 7 in Java 7 and 8 while maintaining strict backwards compatibility: • ANTLRv4-based compilers for TML and associated little languages • Tree interpreters for TML, TAF, and TCF • Versioned OSGi bundles with Maven and Apache Felix • Library of standard functions, primitive types, and collections with extensive unit tests
	Taught the product's owner the fundamentals and best practices of software design and implementation in Java so that he could maintain and improve the code base himself
	<u>TeraScript Server 7</u> (formerly Witango Server): Reorganized, cleaned up, and modernized a 17-year-old C++ code base and addressed numerous bugs and feature requests for the server's first major release in over two years; now in maintenance releases
	Eliminated diverging platform-specific build systems (on Win32, Linux, and OS X) in favor of CMake and ported the code base to Win64 using Visual C++ 2012
	Designed and implemented a SQL generator that visits dialect-specific trees of SQL-99 constructs
	<u>TeraScribe 8</u>: Reorganized and cleaned up a 10-year-old Java Swing code base in order to add a number of features to the code editor, including autocomplete and syntax highlighting, and refit the data source management interface to use JDBC metadata
	<u>Subcontracting</u>: Cross-platform (Android, iOS, Windows) mobile app development with Xamarin Forms and C#
4/2017 – 5/2017 Part-time	**Drug name comparison project** PSW Applied Research Inc. / Toronto, ON and remote
Activities	Implemented Kondrak's phonetic alignment and similarity algorithm (ALINE) and a bigram-based orthographic algorithm (BI-SIM) as a PostgreSQL C extension
	Created a JavaScript demo application using jQuery and Bootstrap
4/2016 – 6/2016 Part-time	**Full-stack development for Polygon Analytics** Polygon Analytics Ltd. / London, UK and remote
Activities	Implemented a geocoding and geotagging library for text documents, using libpostal and proprietary natural language processing algorithms in C++ (Visual Studio, CMake, Linux port)
	Wrapped the library in Python with Boost.Python and as a REST API
	Implemented a signup, payment (Stripe), and administration interface in Python with Flask and Redis
4/2014 – 7/2014 Part-time	**Application development for the <u>Muninn Project</u> under contract to the Library and Archives Canada** PSW Applied Research Inc. / Toronto, ON and remote
Activities	Created a command-line Java application for transcribing handwritten World War I medical records from the Canadian Expeditionary Force (http://blog.muninn-project.org/node/79) using Amazon's Mechanical Turk crowdsourcing service: precut images for groups of lines, redundancy between workers, artificially-introduced mistakes and edit distances to detect cheating, and feedback to workers

3/2013 – 7/2013 Part-time	**Back end development for <u>Txtmrkt</u>, a marketplace for digital publications** Txtmrkt LLC / New York, NY USA and remote
Activities	Designed and implemented a Model-View-Controller-based Java web application with Guice-injected Thryft services backed by Amazon S3
	Authenticated and authorized users with Shiro
	Integrated co-branded order placement and payment notification with Amazon's Flexible Payments Service
12/2011 – 6/2012 Full-time	**Haystack: performance monitoring platform for online businesses** GrokIO LLC / New York, NY USA
Activities	Designed an agent (Python) and manager (Java with Spring Security+MVC and Guice) architecture for gathering, storing, and querying time series data and metadata
	Implemented time series databases from scratch using relational tables and memory-mapped files
	Wrote RFC 3986 grammar-based URL and URN parsers using Ragel (Java) and pyparsing (Python)
	Generated service interfaces in Java and Python with custom Protocol Buffers compiler plugins
	Coded agent plugins in Java and Python for discovering and fetching metric metadata and data: • Open source monitoring systems: Nagios, Ganglia, collectl, Graphite • Open source server software: Apache httpd, nginx, MySQL, memcached, MongoDB, JMX • Platform interfaces: procfs, sysctl, WMI, SNMP • Third party APIs: Google Analytics, MailChimp, Spring Social, Amazon CloudWatch
	Supervised user experience and graphic designers in Ukraine
	Implemented a client-side user interface with the Google Web Toolkit, starting from wireframes: • Architectural best practices: Model-View-Presenter; event bus; activities and places • Custom widgets: date time range picker, cell table pager, search box, selection tree • .war deployment to Tomcat with Maven
8/2011 – 12/2011 Full-time	**Back end development for <u>Birchbox</u>, a personalized subscription service for women's cosmetics samples** Birchbox Inc. / New York, NY, USA
Activities	Designed, implemented, documented, and successfully deployed warehouse integration code in Java (back end) and PHP (administrative front end)
	Wrote a Thrift compiler and targets in Python+pyparsing to generate Java and PHP from Thrift definitions
	Produced ad hoc business reports from MySQL using Java and Jython
5/2011 – 8/2011 Full-time	**Benchmarking server stacks in order to accurately provision EC2 instances at <u>Chartbeat</u>, a real-time web analytics software-as-a-service provider** Chartbeat Inc. / New York, NY, USA
Activities	Synthesized a realistic benchmark for front-line web servers from statistical analyses of nginx access logs
	Significantly reduced network bandwidth and latency on production servers with targeted optimizations
	Wrote Python scripts for analyzing Ganglia RRDs, replaying HTTP request streams, harnessing httperfs
10/2010 – 4/2011 Full-time	**File I/O library for <u>X10</u>, a type-safe, parallel object-oriented language for high-productivity computing** IBM Research / Hawthorne, NY, USA
Activities	Designed, implemented, tested, and documented a new low-overhead, buffer-based file library for X10: • inspired by POSIX, FUSE, Boost.Filesystem, and Java NIO.2 • X10 and native code for the Java and C++ source-to-source compilation backends • Scatter/gather I/O, aligned buffers, memory-mapped files, advisory locking
	Microbenchmarked the X10 runtime

| 1/2009 – 7/2010 | **LXFS: fast and reliable data storage for High-Performance Computing clusters** |
| Full-time | NEC High Performance Computing Europe / Stuttgart, Germany |

Activities — Wrote Python and bash scripts for configuring, deploying, and administrating LXFS installations:
- Lustre and lnet configuration (MGS, MDS, OSS, clients)
- Redundant NFS exports
- Nagios, Ganglia, and collectl monitoring
- Heartbeat/Linux-HA services for failover
- Promise RAID devices
- Network interfaces (Infiniband, bonded Ethernet)

Finished two major LXFS releases, used in numerous deployments

Debugged and resolved issues in production parallel file systems

| 1/2009 – 7/2010 | **XtreemFS: a distributed and replicated file system for wide-area networks** |
| Full-time | NEC High Performance Computing Europe / Stuttgart, Germany |

Activities — Designed, implemented, and tested the XtreemFS userspace client:
- FUSE and Dokan (Windows FUSE-like library) interfaces
- Multiple pipelined ONC-RPC streams to a a single server
- Staged, event-driven concurrency for robust performance under load
- Close-to-open file caching with per-file page sizes
- Automatic failover between file replicas on timeout
- Heavily benchmarked under different I/O loads (iobench, dbench, metadata benchmarks)

Finished two major XtreemFS releases, both used in production

Collaborated with academic and industry partners in Europe, Israel, and China

Teaching Experience

2/2014 – 4/2014	Two sections of „SQL Fundamentals" and „Python fundamentals" workshops for 47 (total) Ph.D. students and post-doctoral researchers in the natural sciences Iowa State University / Ames, IA USA
2002 – 2004	Technical supervision of 10 Masters' theses in the U.S. and Poland Oklahoma State University / Tulsa, OK USA
1/2002 – 3/2002	Mathematics tutoring for 6th and 7th graders Tulsa Public Schools / Tulsa, OK USA