# Executive Summary

DevOps Engineer with 28 years of experience managing Linux, UNIX, [Windows](https://www.microsoft.com/en-us/windows/), and OSX/macOS systems. Programming languages have included Python, PHP, Perl, and Java. Work environments have been heterogeneous (several flavors of Linux, [Windows](https://www.microsoft.com/en-us/windows/), and OSX/macOS), small to medium sized (from 10 to 1200 servers, 20 to 300 workstations), and mixed locations (all local to all remote teams).

# Relevant Technical Skills

* **Big Data:** [Google Big Query](https://cloud.google.com/bigquery), [HDFS](https://hadoop.apache.org/), [Hive](https://hive.apache.org/), [YARN](https://hadoop.apache.org/), [Alluxio](https://www.alluxio.io/), [Impala](https://impala.apache.org/), [Trino](https://trino.io/), [Kafka](https://kafka.apache.org/), [Kubernetes](https://kubernetes.io/), [Dagster](https://dagster.io/)
* **Database Servers:** [MySQL](https://www.mysql.com/), [PostgreSQL](https://www.postgresql.org/), [Microsoft SQL Server](https://www.microsoft.com/en-us/sql-server)
* **Database Skills:** [PostgreSQL](https://www.postgresql.org/) Database Administration, Relational Schema Design, Structured Query Language (SQL)
* **Programming and Scripting Languages:** [Bash](http://www.gnu.org/savannah-checkouts/gnu/bash/manual/bash.html), C/C++, [Java](https://www.java.com/en/), Javascript, [Perl](https://www.perl.org/), [PHP](https://www.php.net/), [Python](https://www.python.org/)
* **Programming Skills:** [Docker](https://www.docker.com/), [Jenkins](https://www.jenkins.io/), [Jira](https://www.atlassian.com/software/jira), [Intellij IDEA](https://www.jetbrains.com/idea/), Object-Oriented Design, Object-Oriented Programming, Refactoring
* **Software Configuration Management Tools:** [Git](https://www.git-scm.com/), [GitHub](https://github.com/), [GitHub Actions](https://github.com/features/actions), [Mercurial](https://www.mercurial-scm.org/), [Subversion](https://subversion.apache.org/)

# Relevant Job History

## EvolutionIQ - Senior Software Engineer

New York City, NY - Feb 2024 - Current

* Orchestrated [Google Cloud](https://cloud.google.com/) infrastructure with [Terraform](https://www.terraform.io/), managing [GKE](https://cloud.google.com/kubernetes-engine), [BigQuery](https://cloud.google.com/bigquery), and developer virtual machines for scalability and efficiency.
* Redesigned EvolutionIQ’s CI/CD pipeline, enhancing efficiency in development workflows.
* Fixed critical bugs in locally developed [GitHub Actions](https://github.com/features/actions), ensuring smooth and reliable environment promotion.
* Built a template repository to guide EvolutionIQ teams in using [GitHub Actions](https://github.com/features/actions) for deployment, as well as setting up data pipelines with [Dagster](https://dagster.io/).
* Developed a Python-based CI/CD library to manage complex actions beyond [Bash](http://www.gnu.org/savannah-checkouts/gnu/bash/manual/bash.html)’s capabilities.
* Implemented [nektos/act](https://github.com/nektos/act) to enable developers to test build and deploy actions locally before committing to [GitHub](https://github.com/).
* Refined [data retention policies](https://cloud.google.com/bigquery/docs/managing-partitioned-tables" \l "partition-expiration) to guarantee compliance and ensure critical data availability for legal requirements.
* Implemented SLA monitoring with [Grafana](https://grafana.com/), [SendGrid](https://sendgrid.com/), and [PagerDuty](https://www.pagerduty.com/) to proactively detect and alert customers and engineers about overnight job failures, ensuring swift issue resolution.
* On-call support on a rotating basis.

## Pulsepoint - Data Engineer and Director of Infrastructure for Data

New York City, NY & Newark, NJ (Telecommute) - Mar 2015 - Nov 2023

### Director of Infrastructure for Data, May 2018 - Nov 2023

* Architected data streaming that manages 40T of data/day.
* Established new data centers in Europe and in Virginia.
* Migrated data center, moving processing of data pipelines to new data center.
* Replaced [Vertica](https://www.vertica.com/) with [Trino](https://trino.io/).
* Deployed and configured [Alluxio](https://www.alluxio.io/) for caching and data orchestration.
* Performance tuned [Kafka](https://kafka.apache.org/).
* Tested new tools for suitability, including [MariaDB](https://mariadb.org/), [Clickhouse](https://clickhouse.com/), and [Kudu](https://kudu.apache.org/).
* Switched build server from [TeamCity](https://www.jetbrains.com/teamcity/) to [Jenkins](https://www.jenkins.io/), recreating all build jobs.

### Data Engineer, Mar 2015 - May 2018

* Upgraded [Kafka](https://kafka.apache.org/) with zero downtime for producers and consumers.
* Enabled integration with [Active Directory](https://en.wikipedia.org/wiki/Active_Directory) for [Hadoop](https://hadoop.apache.org/) systems.
* Built tool to graphically show the ETL pipelines.
* Transitioned ETL pipeline from crontabs to [Mesos](https://mesos.apache.org/) and then into [Kubernetes](https://kubernetes.io/).
* Troubleshooting of issues with [Hadoop](https://hadoop.apache.org/), [Kafka](https://kafka.apache.org/), [SQL Server](https://www.microsoft.com/en-us/sql-server), and [Kubernetes](https://kubernetes.io/).
* Production maintenance of data pipelines, including after hours support.
* Implemented data duplication between two [Hadoop](https://hadoop.apache.org/) clusters.
* Upgraded [Hadoop](https://hadoop.apache.org/) clusters with minimal downtime.
* Optimized the performance and reliability of [Hadoop](https://hadoop.apache.org/), ensuring high availability.
* Added over a hundred automated tests to our ETL pipeline.
* Performed root cause analysis on ETL and cluster level failures.

## Weight Watchers - Systems Engineering Lead

New York City, NY - Nov 2014 - Feb 2015

* Developed lightweight monitoring tool for use within my group.
* Configured [Vormetric](https://cpl.thalesgroup.com/encryption/vormetric-data-security-manager) products to ensure [HIPAA](https://en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act) compliance for customer data.
* Worked to transfer from [Rackspace Cloud](https://www.rackspace.com/cloud) to [Openstack](https://www.openstack.org/) based private cloud.

## OrcaTec, LLC - Developer

Atlanta, GA (Telecommute) - Jun 2012 - Oct 2014

* Found major security hole (remote code execution) and closed it.
* Debugged and resolved memory issues that were causing systems to shut down.
* Incorporated memcached into our stack to handle sessions and cached data.
* Switched web server from [Paster](https://pypi.org/project/Paste/) to [Apache](https://httpd.apache.org/) with [mod\_wsgi](https://code.google.com/archive/p/modwsgi).
* Corrected Unicode handling errors in the code.
* Added holds and matters framework, allowing customers to state that documents belong to specific cases and should not be deleted while the cases are ongoing.
* Identified weaknesses in the database model, and added code to prevent those weaknesses from being hit.

## Choopa.com - Developer

Sayreville, NJ - Jan 2012 - May 2012

* Developed library to manage [OpenStack](https://www.openstack.org/) nodes, and gather billing information.
* Built [Nagios](https://www.nagios.org/) configuration file generator for in-house web interface for [Nagios](https://www.nagios.org/).
* Configured [Bacula](https://www.bacula.org/) backup system as replacement for custom backup scripts.
* Reconfigured [Nagios](https://www.nagios.org/) monitoring, reducing full check from 8 hours to 2 minutes.
* Refactored in-house [Nagios](https://www.nagios.org/) web interface. This reduced the workload from six files down to one when adding new checks.
* Several smaller bug fixes and features throughout the internal code base.

## Education

Bachelor of Science in Computer Science, 2000  
East Stroudsburg University, East Stroudsburg, Pennsylvania

## Professional Certificates

[Google Cloud Certified Professional Data Engineer](https://www.credly.com/badges/6c91d726-a924-4b8d-994b-80e49e002428/public_url) (Jan 2024)  
Online Course - Google