

Evgenii Seliavka - Sr. Hadoop Administrator/DevOps/SRE Engineer

- mobile: 404-496-8247 • evgenii@seliavka.site
 - linkedin: [linkedin.com/in/eseliavka](https://www.linkedin.com/in/eseliavka) • github: eselyavka
-

Work experience

“Twitter” (twitter.com) Staff Site Reliability Engineer SF, CA / ATL, GA

Nov 2018-current

Domain: Internet, online microblogging and social networking service

Responsibilities:

- Administration and management of the hadoop fleet in the company, tens of thousands of hosts
- Implementation and integration Kerberos security on top of the cloud / on-prem clusters, without any downtime to enable GDPR
- Troubleshooting and debugging technical issues for my team and across the company
- Continuous automation of all processes within the team and company
- Bring test driven development methodology to the team
- Monitoring, SLA/SLO/SLI
- Working with the open source community
- Development of the different applications and scripts
- Refactoring and continuous improving existing code base
- Cloud migration process, implementation, documentation, review and feedback
- Mentoring other engineers in the team and interviewing new candidates
- Capacity planning and cost reduction
- Incident management
- Docs creation and review, such as Technical Design Docs (TDD), Production Readiness Docs (PRD), etc.
- Creating / Improving existing educational/training documentation (wiki/run books)
- Technical leadership, knowledge transfer, team roadmap planning, project management

“IAS” (integralads.com) Sr. Big Data Engineer ETL NYC, NY

Jan 2018-Nov 2018

Domain: Internet, technology company that analyzes the value of digital advertising placements

Responsibilities:

- Development and enhancing various data pipelines, integrating those pipelines with existing CI/CD pipelines using tools such as Fabric, Ansible and Jenkins, troubleshooting
- Data impact analysis in case of any code changes
- Unit test coverage of the existing pig scripts with Pig Unit as an enhancement of the existing CI/CD pipelines.

“Cigna” (cigna.com) Sr. Hadoop Administrator contractor Lafayette, LA

Aug 2017 - Jan 2018

Domain: Health care/Insurance, one of the biggest insurance players in US/Europe

Responsibilities:

- Automation of the hadoop master/worker nodes provisioning with Ansible
- Performance bottleneck illumination for CDH and ecosystem components especially Apache Hbase
- Kerberized and secure CDH cluster

“Grid Dynamics” (griddynamics.com) Sr. DevOps Engineer Russia, Sankt-Petersburg

Feb 2015-Aug 2017

Domain: IT Service/Staffing

Responsibilities:

- Supporting various ETL pipelines
- Building CI/CD pipelines
- Performance tuning and bottleneck analysis of the MR jobs and overall Hadoop cluster(s)

- Various automation activities based on the AWS stack
- Creator of schema migration process for Impala

“Okko” (okko.tv) Hadoop Administrator / DBA *Russia, Sankt-Petersburg*

Jan 2012- Feb 2015

Domain: Internet, one of the major VOD(video on demand) provider in Russia

Responsibilities:

- Administration of the CDH cluster, Postgresql cluster
- Set up Puppet for provisioning virtual and on premise infrastructure
- Building home grown HA solution for the CDH, Postgresql and Redis clusters
- Set up performance auditing of the PostgreSQL and Hadoop clusters on everyday basic which includes profiling of the SQL queries, JMX metric collection and log analysis
- Vanilla Hadoop to CDH migration
- Building event driven pipeline based on Flume

“Megafon” (megafon.ru) Engineer *Russia, Sankt-Petersburg*

Jul 2008 - Jan 2012

Domain: Telecommunication, one of the major 4G/LTE provider in Russia

Responsibilities:

- Administration of the Oracle and Postgresql RDBMS
- Set up monitoring for various telecom services(IVR, VoiceMail, VideoMail, USSD, SMS) based on Nagios
- Develop bunch of monitoring/housekeeping scripts

Software projects:

Impala/Hive Liquibase plugin: <https://github.com/eselyavka/liquibase-impala>

- Author of an open source Liquibase plugin for Impala and Hive
- Created support documentation
- Engaged with Hadoop / Liquibase community on support of the plugin

Contributed to **dnsmasq** open source project: <https://bit.ly/3QkZziM>

- Enhance --conf-dir parameter to load files in a deterministic order in version **2.81**

Publications:

- Kerberizing Hadoop Clusters at Twitter: <https://bit.ly/3YX8Oco>
- How we scaled Reads On the Twitter Users Database: <https://bit.ly/3YX3mG9>

Education

Saint Petersburg ITMO State University, faculty of science, department of physics. Bachelor's Degree in Information Systems and Technology with final graduation paper “Design of Computer Learning-Aid Model ‘Interactive Abbe Diagram’” - 2008

Certifications: CCA-410 Cloudera Certified Administrator for Apache Hadoop CDH4; CCB-400 Cloudera Certified Specialist in Apache HBase; RHCE Red Hat Certified Engineer; RHCSA Red Hat Certified System Administrator; OSS-DB Silver Certified Engineer; AWS Certified Developer - Associate

Languages and Technologies

Languages: Python, Go, Bash, awk, sed, Java, C

Development: Git, Github, Subversion, MVN, JDK

Virtualization: KVM, VirtualBox, Docker, Kubernetes

Cloud: AWS, GCP

Provisioning: Ansible, Chef, Puppet, Fabric, Vagrant, Terraform

CI/CD: Jenkins, Nexus, Artifactory, Sonar, Liquibase, Github actions

RDBMS: Postgresql, Vertica, Oracle

Big Data: Hadoop, HDFS, YARN, MapReduce, Spark, Flume, Kafka, Hive, Impala
OS: Linux, Unix