

Izhar ul Hassaan

SYSTEMS ENGINEER · BIG DATA ARCHITECT · DEVOPS ENGINEER

Stockholm, Sweden

☎ (+46) 72-537-5777 | ✉ ezhaar@gmail.com | 🏠 ezhaar.github.io | 📱 ezhaar | 🌐 ezhaar

"We are bound by our choices, but we are more than our mistakes."

Summary

- Systems Engineer with over 5 years experience working with Big data projects.
- Experience working with large scale datasets (100s of TBs) and distributed clusters (100s of nodes)
- Experience working with Docker, Ansible
- Hands on with Spark programming and optimization (python, scala)
- Passionate about code optimization
- Worked with AWS, GCE and OpenStack cloud environments
- Worked with Cloudera, Hortonworks distributions
- Experience with designing, architecting and implementation Big Data infrastructures from scratch

Education

KTH, Royal Institute of Technology

Stockholm, Sweden

MASTERS IN SCIENTIFIC COMPUTING (COURSEWORK)

2008

- Specialization in High Performance Computing

National University of Science and Technology

Islamabad, Pakistan

BACHELORS IN INFORMATION TECHNOLOGY

2002 - 2006

- Focus on Algorithms and Data Structures

Skills

Data Analytics Spark, Hadoop, Kafka, Yarn, Hive, Cassandra, Oozie, Zeppelin, Jupyter

Programming Python, Scala, C, MPI, OpenMP

DevOps Docker, Ansible, Fabric, Bash

Cloud Google Compute, AWS, Openstack

Experience

Telia Company

Stockholm, Sweden

SOLUTION ARCHITECT & ANALYTICS DEVELOPER

Mar. 2016 - Present

- Working with **Data Pipelines** for **Crowd Analytics**, Network Data (probes, CDRs), **IPTV**
- Architecting and implementing **data ingestion**, storage, compression and archival strategies in a multi-tenant kerberized Data Lake
- Leading the move from **Map-Reduce, Hive jobs to Spark**
- Optimized Spark code for data anonymization for Crowd Analytics (**10x Speedup**)
- Optimization of Cloudera Hadoop Cluster (**small files issues, block size issues, slow Hive queries, slow MR jobs**)
- Spark application development
- Data discovery using Jupyter/Zeppelin notebooks

Tradedoubler AB

Stockholm, Sweden

SYSTEMS ENGINEER

Sep. 2015 - Feb. 2016

- Setup/administration of **Hortonworks cluster** in **Google Compute Cloud**
- Design and architecture of data pipelines (DropWizard, Kafka, Storm, HDFS, Parquet, Spark)
- Spark application development
- Data discovery using ipython/zeppelin notebooks

PDC Center for HPC, KTH

Stockholm, Sweden

SYSTEMS ENGINEER

Jan. 2013 - Aug. 2015

- Design and deploy highly scalable cloud-based SaaS infrastructure (OpenStack).
- Administration of Ceph Cluster and integration with Openstack cloud
- Writing software and scripts to support creation of **virtual clusters** for data intensive computing on our private cloud.
- Virtualization (IAAS, PAAS) on **OpenStack, AWS**
- **Docker** for **Hadoop** and **Spark** dev environments
- Advance **OpenVPN** deployment

Nosiva AB

Stockholm, Sweden

SYSTEMS ENGINEER, FREELANCE CONSULTANT (SIDE PROJECT)

Feb. 2015 - Jul. 2015

- Setup **BigData infrastructure** in **AWS**
- Administration of **Hadoop and Spark** Clusters in AWS
- Setup development environment (data exploration) for data scientists **jupyter notebooks, spark, cassandra**
- Spark application development (python)
- Development of Services to handle **streaming data** from **Amazon Kinesis**
- Service deployment using **docker**

PDC Center for HPC, KTH

Stockholm, Sweden

RESEARCH ENGINEER

Jun. 2009 - Dec. 2012

- Provide support for **MPI, OpenMP** and **CUDA** Programming applications on PDC clusters.
- Worked with **Cray XE6** system (35000+ cores)
- Management and installation of software on HPC Clusters.
- Developed dashboard for monitoring, accounting of jobs on PDC clusters (**Django, MySQL, JQuery, HighCharts**).
- Collaboration with researchers to develop and **optimize parallel applications**.
- Developed and updated documentation for end users on how to use HPC resources.
- Provided technical support and training to new users and employees.