Hudhaifa Ahmed

Senior Big Data Architect

Berlin, Germany

About Me

Experienced Senior Big Data Architect with over 12 years in the IT industry, specializing in designing and implementing scalable, high-performance data solutions. Proficient in Apache Hadoop ecosystem, cloud technologies (AWS, Azure), and data processing frameworks like Apache Spark. Skilled in developing ETL pipelines, real-time streaming solutions, and optimizing data workflows. Databricks Certified Data Engineer Associate with extensive experience in Databricks CLI, API, Unity Data Catalog, and Databricks workspaces. Certified Cloudera Administrator and Scrum Master with a strong background in agile methodologies. Recently expanded expertise to include Large Language Models (LLMs) and Generative AI, integrating cutting-edge AI technologies into big data architectures to drive innovation and business value.

Overview

- Senior Big Data Architect
- Ultra Tendency Academy Trainer (Kafka, NiFi, ML, Databricks)
- Senior Big Data Architect with expertise in Large Language Models (LLMs) and Artificial Intelligence (AI)
- Ultra Tendency Academy Trainer (Kafka, NiFi, ML, Databricks, LLMs, GenAI)
- Cloudera Certified Administrator (CCA131)
- Certified Professional Scrum Product Owner & Scrum Master
- +12 years of relevant IT experience
- Expert for Azure Cloud technologies like Network and Security
- Experience with Databricks/Cloudera & Hadoop Stack (Spark and Data technologies)
- 10 years of hands-on experience in Software Development, Architecture, System integration, OOP programming, Designing and Prototyping, Database designing and Specification creating
- In-depth knowledge designing highly scalable Spark-based IoT streaming applications based on Kafka, Spark Streaming, HBase
- Deep knowledge of Agile Approaches (SCRUM, DevOps, Clean Code)
- Technical coaching for project members supporting their future development & career goals

Contact:

- E-mail: info@huthaifa-abd.com
- Website: http://huthaifa-abd.com
- GitHub: https://github.com/huthaifa-abd

Main Technologies

- Apache Spark, Apache Hive, Apache Hadoop, Apache HBase, Apache Kafka, Apache Oozie, Cloudera Impala, Cloudera Hue, Cloudera Manager, Apache Ambari
- Active Directory & Kerberos Hadoop integration
- Databricks CLI, API, Unity Data Catalog, Workspaces, Clusters, Notebooks

- Hadoop High Availability (NameNode HA, Resource Manager HA, HiveServer2 MetaStore Server HA, Impala Load Balancing, Oozie Load Balancing)
- Hadoop security & hardening (Kerberos, TLS, Redaction, Cluster LAN)
- Hadoop operations (replication, backup, maintenance, hardware replacement, version upgrades)
- Cluster sizing, cluster layout, cluster hardware specification
- RESTful Microservices Development
- Containers (Docker and Kubernetes)
- Cloud Computing / Cloud Automation (AWS, Azure)
- MySQL, HBase, Oracle, DB2, MongoDB
- Maven, SBT
- RHEL, SLES 11 & 12, CentOS, Debian, Ubuntu
- JIRA, Redmine, Jenkins, Git, SVN

Other Technologies

OpenStack, Foreman, VSphere, KVM, Windows VDI, CEPH, Chef, Puppet, Asp.net [C#- VB], PHP5, Javascript, JAVA, Ruby, CSS3, HTML5, C++, Java, C#, VB, Ruby, .Net, WPF, Winforms, ASP.net, MVC, Entity Framework, CakePhp, CodeIgniter, Drupal, Wordpress, Android, IOS, PhoneGap, Magento, Opencart, Windows Server 2008 - 2012, TFS, DC, Microsoft Exchange Server, Ubuntu Server 12.4, Cloud Computing, Centos 6.5 – 7, IIS 6-7, Apache, Nginx, MySql, MSSQL, Photoshop, 3ds Max, Zbrush, After Effect, Adobe Flash, Adobe Dreamweaver, Gimp, Blender, Topogun, Keyshot, Substance Designer, Substance Painter UDK, Unity, GameMaker, UE4

Work Experience

Ultra Tendency, Germany (Aug 2017 - Present)

Role: Berlin Territory Manager & Senior Lead Big Data Developer

Responsibilities:

- Development, implementation, design solutions, architecture review, product management, pre-sales, technical leadership
- Leveraging new and emerging practices for Enterprise Data Management utilizing Hadoop and Software-based Platforms
- Ensuring excellent practices are utilized in delivering Big Data Management and Integration Solutions
- Develop and Implement different strategies that drastically improve the performance of existing and new Big Data Solutions
- Formulating and implementing risk mitigation strategies
- Middleware Programming utilizing Java, Scala, Python
- Responsible for building and supporting a Hadoop-based ecosystem designed for enterprise-wide analysis of structured, semi-structured, and unstructured data
- Writing scalable and quality application code

Achievements:

- Maintained a high level customer satisfaction across all clients engaged with
- Implemented multiple large migration and new data systems deployments end to end
- Deployed new streamlined event based near-realtime Engineering stack

- Delivered multiple revenue impacting projects fixing numerous data quality issues
- Designed, Implemented, and Deployed patent-able ML technologies into production at scale
- Establish and maintain an engineering culture with Agile foundations
- Enhance and improved database access and caching performance
- Achieved internal goals of restructuring matrix organization and launched operations in a new geo from scratch

Promotions:

- Berlin Territory Manager & Senior Lead Big Data Developer - Jul 2021
- Regional Manager & Senior Lead Big Data Developer - Jun 2018
- Lead Big Data Developer - Jan 2018
- Big Data Developer - Aug 2017

Earthlink Telecommunications (July 2011 - Aug 2017)

Role: Research and Development Manager

Responsibilities:

- Software Development, solution design, architecture review, market analysis, risk assessment & stakeholder management, recruiting, coaching, technical leadership, reporting, documentation, presales
- Improve and renovate the research and development programs to meet organizational needs and to capitalize on potential new products
- Helping senior management and internal clients to understand the results and implications of research and development initiatives
- Selecting, supervising, training and directing the work of Research and Development staff
- Introduce and embrace the open source culture by encouraging the use open source software and migrating critical solution successfully to open source alternatives openstack / ceph

Achievements:

- Increase employee engagement by over 150% in a period of 1 year by providing a competitive environment
- Improve project management process by integrating CI / CD tools to achieve efficient well written solutions (Rundeck, Redmine, Gitlab)
- Renovated most of the company VAS services to be more in compliant with current needs that lead to 50% more engagement
- Lead the third party solution development by taking initiative to advertise the company products and to coordinate with clients to provide the most satisfaction possible
- Improved the hiring process by 80% using dedicated recruitment tools and interviewing skills to elect the best talents possible
- Decreased software deployment costs by applying automated integration / automated deployment tools saved over 16+ Hours for each deployment
- Reduce hardware resource usage significantly by planning / developing / deploying cloud technologies to save on resource and improve the solution performance
- Improve Solution up-time by implementing software monitoring tools and applying feasible threshold for the services, providing maintenance documents to quickly resolve issues

Promotions:

- Jan 2013 - Promoted to Senior Developer
- May 2014 - Promoted to Lead Developer
- Feb 2015 - Promoted to Research And Development Manager

Projects

Ultra Tendency, Germany

Austrian Insurance Company

Role: Development Lead

Tasks:

- Development Lead of Databricks migration workstream
- Planning and design cloud architecture for new Databricks implementation
- Develop Security concept and authentication methods
- Hands on migration of an application from on prem to Databricks on Azure

Banking Sector

Role: Full Stack Lead Developer (Product Owner / Architect)

Tasks:

- Requirement Engineering & Solution Documentation
- Scalable RESTful Micro service design and integration
- Planning & Supervising solutions deployments
- Planning and management of the test resources
- Assessing the progress and effectiveness of the test effort
- SCRUM & Product Life Cycle Management
- Define user task flows and interaction models
- Showcase design ideas through sitemaps, storyboards, and process flows
- Implement the latest design innovations to ensure that the product is in accordance with the cutting-edge technology
- Develop standards and procedures to determine product quality and release readiness

German Insurance Company

Role: Senior DevOps

Tasks:

- Maintenance of an existing Cloudera cluster landscape (HDFS, Spark, Kafka, Hive)
- Planning and execution of performance enhancements

German Telecommunications Company

Role: Lead developer in a Cloud-based Big Data project on Microsoft Azure

Tasks:

- Micro service development with Spring Boot and Kubernetes
- Cloudera on Azure, Databricks on Azure
- Introduction and maintenance of a KPI system to monitor and improve the code quality based on SonarQube

German Automobile Company

Role: Cloudera Expert

Tasks:

- Construction, installation and automation of a Cloudera cluster
- Kerberizing of cluster services and TLS for Web Frontends
- Automatic cluster deployment using Puppet
- Cluster Administration
- HBase, Spark, Kafka, Impala

Service Provider

Role: Big Data Developer

Tasks:

- Construction, installation, and automation of a Cloudera cluster
- ETL Workflow under deployment of Sqoop und Oozie
- Cluster Administration

German Financial Institution

Role: Big Data Developer

Tasks:

- Construction and hardening of a Hortonworks HDP Cluster
- Connection of HDP Cluster to SAP HANA
- Automatic cluster deployment using Ansible
- Cluster administration

Service Provider

Role: Big Data Developer

Tasks:

- Construction and installation of a Cloudera cluster
- ETL from MS SQL Server using Talend

Telecommunication Provider

Role: Big Data Developer

Tasks:

- Software Development (Enterprise Web / Mobile Applications)
  - Spring Boot Framework, Java, Maven, SBT
- Solution design (Distributed Microservices, Scalable SQL Databases)
  - HBase, Oracle, DB2, MongoDB, MySql, MSSQL
- Software Architecture review
  - Microservices, MVC, Entity, Openstack, Foreman, VSphere, KVM
  - Windows VDI, CEPH, Ansible, Chef, Puppet
- Market analysis, Risk assessment & stakeholder management
- Recruiting, Coaching, Technical leadership
- Reporting, Documentation

Earthlink Telecommunications

On-Premise Cloud Solutions [Openstack] (2014)

Role: Cloud Architect

Tasks:

- Implemented a private cloud infrastructure using Openstack, Qemu, and KVM
- Integrated CEPH for distributed storage and Puppet for configuration management
- Utilized Foreman and Fuel for automated deployment and Chef for infrastructure as code

LMS Manifest System (2013)

Role: Lead Developer

Tasks:

- Developed a multi-node web application for Commodity Manifests Management
- Utilized ASP.NET MVC framework with C#, integrated with MSSQL using LINQ
- Implemented responsive UI using HTML, AJAX, jQuery, and Telerik Tools

Zain DOS System (2012)

Role: Senior Developer

Tasks:

- Created an advanced ticketing and billing system for a mobile telecommunication company
- Built with ASP.NET MVC, C#, and MSSQL for robust data management
- Designed responsive UI with HTML, AJAX, jQuery, and CSS

Multiplayer Gaming System (2011)

Role: Senior Developer

Tasks:

- Designed an ISP-grade game server manager and advanced gaming portal
- Implemented using ASP.NET MVC, C#, and MSSQL with UDP networking protocols
- Integrated various game server APIs and utilized PowerShell for automation

Human Resource Management System - HR (2013)

Role: Full Stack Developer

Tasks:

- Developed a comprehensive HR solution for the Ministry of Education
- Built on ASP.NET MVC framework with C#, MSSQL, and Telerik Tools
- Implemented advanced reporting and employee management features

TV Streaming Service Web Application (2012)

Role: Full Stack Developer

Tasks:

- Built a streaming service platform using ASP.NET MVC and C#
- Integrated VLC streaming services and IMDB third-party APIs
- Developed mobile API web-services for cross-platform accessibility

Content Management System (2013)

Role: Full Stack Developer

Tasks:

- Developed a CMS with customized features using ASP.NET, PHP, and C#
- Implemented multi-database support with MSSQL and MySQL
- Utilized Apache and IIS for flexible deployment options

EGovernment Centralized Archiving Solution (2013)

Role: Systems Architect

Tasks:

- Implemented a centralized archiving system using Windows Server and VDI
- Developed with ASP.NET MVC, C#, and MSSQL, utilizing Entity Framework
- Integrated PowerShell for automation and WinForms for admin interfaces

Certifications

- Cloudera Certified Administrator (CCA131)
- Microsoft Certified: Azure Solutions Architect Expert
- Certified Professional Scrum Product Owner & Scrum Master
- Databricks Certified Data Engineer Associate

Education

- Bachelor of Science in Computer Science, University of Tikrit (2007 - 2011)
    - 3D Now [Augmented Reality Project] (2011) - BC Graduation Project

Languages

- Arabic (native)
- English (C2)
- German (B2)

Open Source Projects

- Mesh Optimization System: An Enterprise solution to optimize complex architecture models for use on desktop and mobile devices. URL: https://github.com/huthaifa-abd/Cloud-Mesh-Optimization-System
- Talent Recommendation Engine: A Map Reduce Java application to crunch job posting data into meaningful information and providing matching rates. URL: https://github.com/huthaifa-abd/Talent-Recommendation-Engine-Using-Hadoop
- Laggards: A card-based Battle Arena game set in a fierce battle for resources on a foreign planet. URL: https://github.com/Thanduriel/laggards
- Project Nightowl: An intelligent surveillance tracking system utilizing IoT / Big Data Technologies. URL: https://github.com/huthaifa-abd/Project-Nightowl

Interests

- Travel to share knowledge on conferences, meetups, podcasts, workshops, mentorship, etc.
- Solving complex IoT / Big Data problems
- Figuring out best & worst practices (by asking "Why?"), and publishing them at Professional Blog and Game Development Blog
- Cloud Computing, Hadoop, Big Data Analytics, Internet of Things, Data Warehouse, Business Intelligence, Cross Platform Enterprise Software Development, Distributed Storage Solutions, Process Automation, CMS/Ecommerce Systems, System Administration, Database Management
- Exploring the world of AI and LLM technologies and utlizing them into real life usecases.