

Paras Malik

Github|paras_malik_mca_iit@yahoo.com|+917204537001|Bangalore,IN

TECHNICAL SKILLS

Programming/Scripting Languages: Java, Golang, Python, C/C++, Bash Scripting, R, Javascript

Databases/Caching Technologies: Redis, HBase, Aerospike, Mysql, Mssql, Memcache, Cassandra

Big Data Solutions: Hive, Hadoop, Spark/PySpark/Spark Structured Streaming

Machine Learning Techniques: Statistical Modeling, UCB, XGBoost, LightGBM, Random Forest, FM, FFM, LDA, NaiveBayesClassifier

Messaging Systems: Kafka

Search Technologies: ES, Lucene

CICD: Docker, Messos, Kubernetes, Jenkins

Other Technologies: Tomcat, SOA(Service Discovery, Client side load balancing), AWS(Auto-scaling/S3/ELB/EBS/Machine Selection/Placement Groups)

EXPERIENCE

Media.Net [Sr. EM|Sr. Tech Lead|Sr. SDE]

Jan 2014 - Present|Bangalore

- Lead a team of 15 SE/Dev-ops/DS/DB engineers to build end to end solution for bidding in RTB auctions

Amazon.com [Software Development Engineer]

Oct 2010 - Dec 2013|Bangalore

- Developed Java services for Kindle registration and book lending
- Developed Amazon Instant Video Living room client applications using Java

Directi [Software Development Engineer]

June 2010 - Oct 2010|Mumbai

- Developed Implemented caching for java application using JNI

EDUCATION

Masters in Computer Application [Indian Institute of Technology|2010]

Bachelors in Computer Application [Panjab University, Chandigarh|2007]

PROJECTS

RTB Bidder Service [Java|Redis|Aerospike]

- Developed & Maintained a complex API to compute and deliver 1/2 Million bids per second for RTB auctions.

Realtime Configuration Manager [Mysql|Kafka|Java]

- Developed system which update 100s of 1000s of frequently changing configurations to production in real-time

ML Model Computation Service [Golang|S3|Redis|LightGBM|Catboost|Consul|Java]

- Developed low latency Golang service to compute various ML model.
- Used Consul for service discovery in autoscaled environment using heterogeneous hosts.
- Built intelligent Java client with capability of WRR

Event Stream Synchronization & Processing [Spark Structured Streaming|Kafka]

- Developed structured streaming application to time synchronize multiple streams of different events in different Kafka topics.

Data Pipelines [Spark|Hadoop|oozi|Mysql]

- Developed complex data pipelines to process large amount of data

SRE Projects [ELK|Autoscaling|AWS|GCP]

- Developed Autoscaling algorithm to add/remove hosts
- Migration from AWS to GCP and Messos to GKE

ML Projects

- Developed System to classify webpages as commercial non-commercial using LDA/NaiveBayes/XGBoost
- Developed RPM/CTR Prediction models using LightGBM/Catboost/XGboost/Statistical Modeling
- Developed Slideshow detection system which automatically identify if given page is part of slideshow.
- Worked on UCB technique to explore new web pages

Redis to RedisCluster Migration [Redis|RedisCluster|Jedis]

- Implemented functionality to handle pipeline requests in Jedis
- Used different functionality to optimize RedisCluster
- Benchmarked and controlled migration

HBase & RedisCluster to Aerospike Migration [HBase|Aerospike|RedisCluster]

- Defined Data model to enable migration in a cost efficient way.