Ujjwal Kanth

Phone: +91 9864202749
Email: ujjwal.kanth@gmail.com
Link: https://uknth.me/about

Principle Architect @ Netcore/Unbxd

Professional Summary

Engineer with 15+ YOE with expertise in high performance, scalable systems & cloud infrastructure, having worked on:

- Search & Recommendations pipeline handling 160+ million requests daily
- Catalog sync engine handling 1TB+ catalog data daily
- Inter region, inter-datacenter cloud agnostic infrastructure built on K8s
- Data pipeline tracking 250+ million events daily
- AI Data pipeline using S3+Parquet Data-lake & Kubeflow

Work Experience

Principle Architect

[Jan, 2022 - Present] [Team Size: 20 - 25]

Search Index Improvements:
Improved query performance by **20-30%** and reduced cost by **10%** overall by upgrading the search-index infrastructure
- Improved the query formation by **~10%**, by introducing new Query Engine which uses custom DSL to communicate with search-indexes
- Improved system performance & reduced failure rate due to custom replica-placement, replication and custom operators
- Infrastructure upgrades to **ARM** & **Apache Solr9**

Data Pipeline (version 2):
Next iteration of data pipeline built to solve the scaling problem faced by the existing pipeline
- Built on **Apache Flink** for streaming, **Apache Druid** & **Apache Superset** for warehousing & **CockroachDB** for reporting
- Dynamic, on-demand scaling of Flink Jobs using custom operator, improved cost of running the pipeline by **30%**
- Tiered data storage for various reporting use-cases

Product Features:
- Multiple Services developed to handle support for user-level segmentation & merchandising
  • Service to classify incoming traffic based on user-data using **Redis** & **CockroachDB**
  • Service to handle classified traffic and apply merchandising based on user-level data
- Improved features set of Schema/Lifecycle management services to reduce failures due to edge cases & support differential update to schema
- Analytics reporting service and improvements to Warehousing using BigQuery

Catalog Ingestion Pipeline (version 3):
Highly scalable, event driven data pipeline to handle massive amounts of catalog data from thousands of products to hundred millions

- Designed from ground up to handle high rate of partial updates and huge reconciliation steps of hundreds of GB data
- Built entirely on **Argo Workflows** leveraging Kubernetes's dynamic scaling along with Spot Instances from AWS, reducing cost of operation by **40-60%** depending on time of the day

Document Store:
Catalog data storage for high speed retrieval when serving the request from index
- Improved the serving latency by **20%** by reducing data overhead within the stack
- Built on top of **PostgreSQL** & **Aerospike**

Architect at Unbxd
[Jan, 2019 - Present]

Inter Region, Inter Datacenter Network Topology:
Design and development of inter-region communication for the micro-services
- Event Driven Design, using **NATS** for inter-region communication of configurations, metadata & events
- Network topology & routing based on index-affinity

AI Data Pipeline:
Designed and developed the data pipeline used for training AI & machine learning models
- Custom data-lake built on top of **S3 & Parquet** Files
- Connectors built for easy access from Kubeflow

Standardisation of Internal Libraries:
- Go (OSS) https://github.com/unbxd/go-base
- Java (Non-OSS) Starter Project & internal functions built on **Quarkus**

Senior Software Engineer at Unbxd
[Jan, 2014 - Jan, 2019]

Edge Server:
Custom API gateway designed to handle all network related operations, observability and logging

Centralised Configuration Management:
Hierarchal rule system based on tiers and groups, built on **DGraph** & **NATS**

Query Pipeline:
Migration from a single Monolith to Microservices Based Architecture
- Custom Index Solution built on Solr(6.6)'s Fork
- Custom Solr-Plugin, Query Parser, Similarity Class, Search Component etc

Other Projects @Netcore/Unbxd:
Some notable projects across different roles
- Git based CD Pipeline using GitOPS & Flux CD reducing deployment times
- Internal API Gateway handling inter-team API calls reducing error rates and improving observability
- **Protobuf** & **GRPC** adoption for inter-service communication improving overall latency of internal requests by **20%**

- Multiple service contributions across the stack (Orchestrators, Query Analysis, Promotions, Tracker, Analytics Pipeline v1)

Product Engineer at Parablu Inc.

[May, 2013 - Jan, 2014]

PHP Framework, Client Side Javascript, HTML5 APP using Angular JS & PhoneGap

Software Engineer at Manhattan Associates

[July, 2012 - May, 2013]

Warehouse Management System: Development, maintenance, performance tuning & Unified multi staged operation to single stage

Assistant Software Engineer at Tata Consultancy Services

[July, 2010 - July, 2012]

London Clearing House (Project): Maintained customisation for the client on TCS Bancs & Enhancements in Equity Management

Skills

Information Retrieval	Distributed Systems	Architecture Design
Engineering Leadership	Software As A Service (SaaS)	Event Driven Design

Education

B.Tech (2006-2010) Manipal Institute of Technology

Technologies

Solr & Zookeeper	Kubernetes	Big-Data, EMR, Parquet
Streaming Pipelines	Event Driven Design	Apache Flink
Kafka	Flux CD	CockroachDB
Redis	NATS	Aerospike
Elastic Search	Apache Druid & Superset	Go/Java

Personal Projects

Self Hosted Homelab built using Ansible: https://github.com/uknth/homelab

Dotfiles for personal setup: https://github.com/uknth/dotfiles