

Resume

Basic Information

Name: Xinyu Pan

Gender: Male

Phone: 18500852935

Email: pxy0592@msn.com

Technical Skills

Technical Stack Microservice, API Key Management, Graph query engine, Elasticsearch, RabbitMQ, Hadoop/Spark, DevOps, Kubernetes/OpenShift, CI/CD, IBM Cloud, GitHub/Git, SonarQube

Technical Role Technical team leader, Backend Developer, Architecture on feature level

Program Language Java, JUnit, JanusGraph, Knowledge Graph, SPARQL, Neo4j/Cypher, Python, Operator, Ansible, Helm Chart, Shell, Jenkins, GitHub/Git

Soft Skills English communication, Time management, Risk control, multiple task management, Agile/Scrum, DevOps, SDLC

Project List

IBM Data Product Hub release

Data Product Hub is a new flagship product aiming to digital asset management in the data mesh, it is built based on open cloud and private cloud, make the data assets can be automatically, intelligently used by any category application, such as model training, data analysis, etc. Currently the product team distributed in North America, China, and Indian. Since 2023 the team spent one half year from the initial prototype design, first MVP until the final 1st release on July, 2024.

Team Role: Team leader, Software Development Engineer

Release cycle 2023.01 - present

Development cycle Agile + DevOps, two weeks per sprint, 3 or 4 sprints in total

Summary Led a 7 person Indian team, completed the 1st release using one half year and continue the development for several major/minor release

Project List

- User ACL management including user, role, permission management. Support for million grade enterprise user currently

- The data assets import, allow customers load the data from various data sources. Support for 10+ data source, mainly include major database, Snowflake, Log, CSV, Plain Text, and URL
- The data asset preview, simply view the structure, non-structure desensitize data
- The data asset version control policy, contains pre-release and release version control
- The data asset definition, includes metadata definition, relationship definition, etc.
- The data asset lineage, construct the knowledge graph, lineage info based on the defined metadata information
- The semantic search combines with the knowledge graph, nlp model, bio-inspired model, etc. for the data assets

IBM Global Search Service

The search service is a common core service built in IBM Knowledge Catalog product to provide the search experience for raw data and governance data for customers using the global unique search bar in browser interface.

Team Role: Software Development Engineer

Release cycle 2019.06 - 2022.01

Development cycle Agile, two weeks per sprint, 4 sprints for minor release, 8 sprints for major release

Summary An unique search bar receives the query from customers directly. As a developer in China, collaborate with the team located in NA and Indian working on feature design, coding, unit test, deployment in agile manner. In three half years, delivered multiple release on IBM public cloud and IBM private cloud solution successfully.

Project List

- Key-value search, customers can use key-value pair to search the data assets
- A customizable ES query clause to expand the query capability and improve the relevant and precise in search result
- Semantic search. based on the knowledge graph, provide a light query engine to add more relationship in the query context from the customers. In technology using JanusGraph API implemented the query engine including SPARKQL query construct, transformation, query execution optimization. In the meantime, proactive evaluate the Neo4J capability instead
- RabbitMQ events optimization in consumer side

IBM High-Performance Distributed Computing Cluster

Team Role: Team leader, Software Development Engineer in Test

Summary

Spectrum Symphony, Spectrum Conductor, Watson Machine Learning-Accelerating are global famous distributed computing cluster product. They have been used by global Top 10 Financial Industry, such as JP Morgan, Morgan Stanley, CitiGroup, Credit Suisse Group, etc. The cluster scale size is from 100+ nodes to 100000+ nodes.

Release cycle Traditional waterfall model, 3~4 months per major release

Development cycle Led 10 person team to conducted and delivered each major release focusing on feature test, system test, performance test, quality analysis, etc.

The related-to test areas have following items, but not limited

- Test goal and strategy
- Feature requirement analysis and test plan
- Test environment installation
- Test case design
- Test automation development and execution
- CI/CD
- Test execution and monitor
- Performance test
- User scenarios validation
- Regression test
- Test report and documentation
- Retrospection
- Communication and collaboration
- Training and knowledge sharing

Companies

- IBM [2012.04 - present]
- Platform Computing Inc. [2004.07 - 2012.03]
- Electric science research Beijing Inc. [2001.07 - 2004.06]

Education

Northeastern University Automation Control B.S