22AIE498 Project Phase- 1

Migrion : AI-powered platform for pre-migration assessment and dynamic planning

Empowering enterprises to **accelerate data migrations**, **optimize costs**, and eliminate vendor lock-in

Domain: Cloud Computing

Group - 13 Batch - A

Nandini Kuppala - CB.EN.U4AIE22030

Monish. G. V - CB.EN.U4AIE22033

Nivedha. V - CB.EN.U4AIE22037

Balu Pinnenti - CB.EN.U4AIE22060

GUIDE

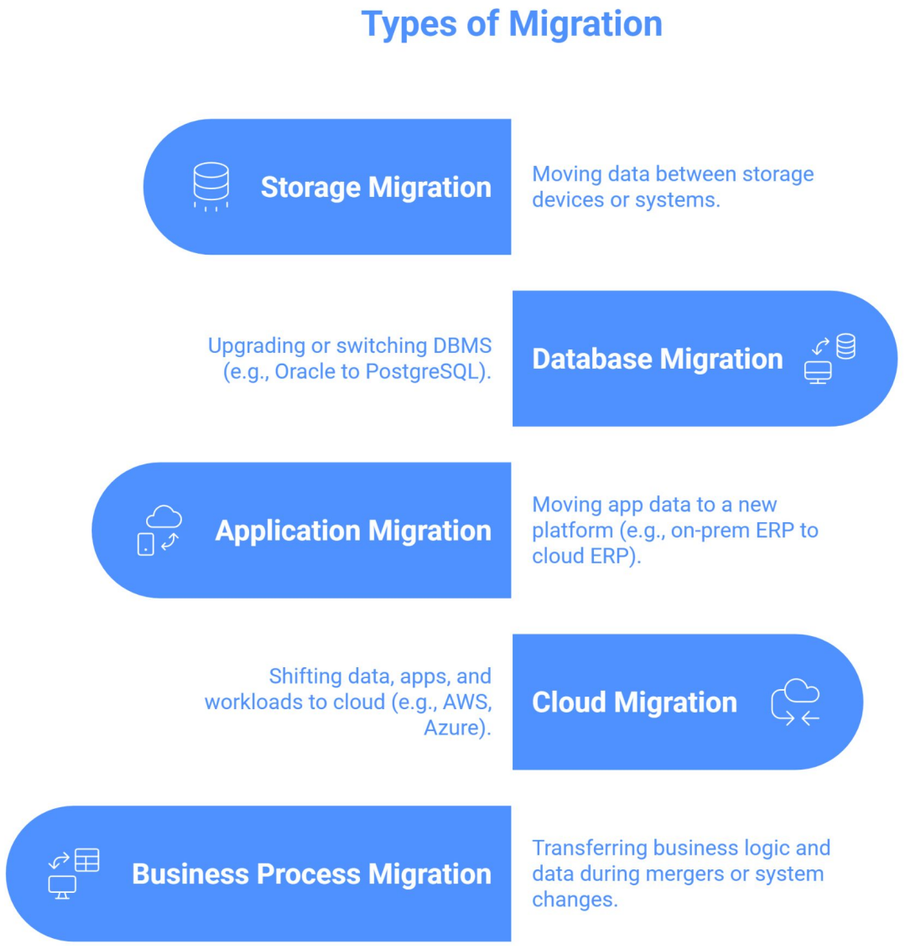
Dr. Jaisooraj. J
Assistant Professor,
Department of AI,
Amrita School of Artificial Intelligence,
Coimbatore

1. Introduction

📦 What is Data Migration?

Data migration is the process of moving data from one system, format, or storage to another typically during upgrades, cloud adoption, data center consolidation, or application changes.

- ➢ **Why is Data Migration Needed?**
- ➢ **Problems in Data Migration:**
  - Downtime and business disruption
  - Data loss or corruption, Mismatch in schemas or formats
  - Security and compliance risks, High cost and resource demand
- ➢ **How Much Do Enterprises Depend on It?** (90% of organizations undergo data migration)



1. Introduction

Which data migration strategy should be chosen?

Lift and Shift

Fast but risky, suitable for quick moves without changes.

Replatforming

Minor changes for optimization, balanced approach.

Refactoring

Rewriting for efficiency, longer but more effective.

Data Migration Testing Strategies

Pre-migration testing

Test workloads on test instances

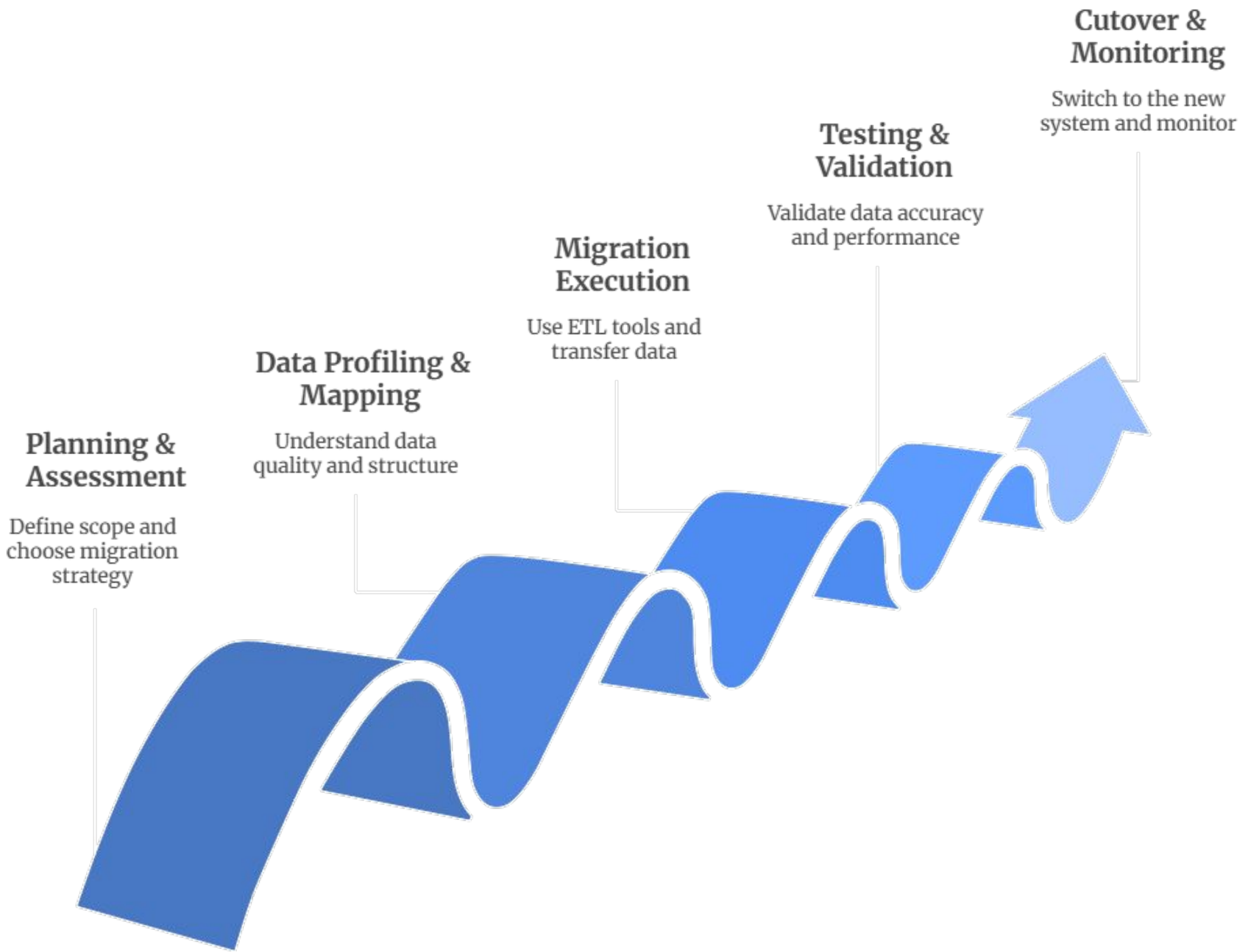
Parallel run

Run old and new systems side-by-side

Data validation scripts

Ensure nothing is missed or misaligned

Data Migration Process



2. Market Research (Existing Tools)

2.1 Cloud Provider Native Tools

Tool	Provider Capabilities	Limitations
AWS Database Migration Service (DMS)	AWS Data transfer, schema conversion	Vendor-specific, limited planning intelligence
Azure Database Migration Service	Microsoft Schema detection, compatibility issues	Static assessment, no adaptive planning
Google Cloud Database Migration Service	Google Heterogeneous migrations, AI-assisted conversion	Limited to Google Cloud ecosystem
Oracle Zero Downtime Migration (ZDM)	Oracle Automated Oracle migrations	Oracle-specific, limited multi-cloud support

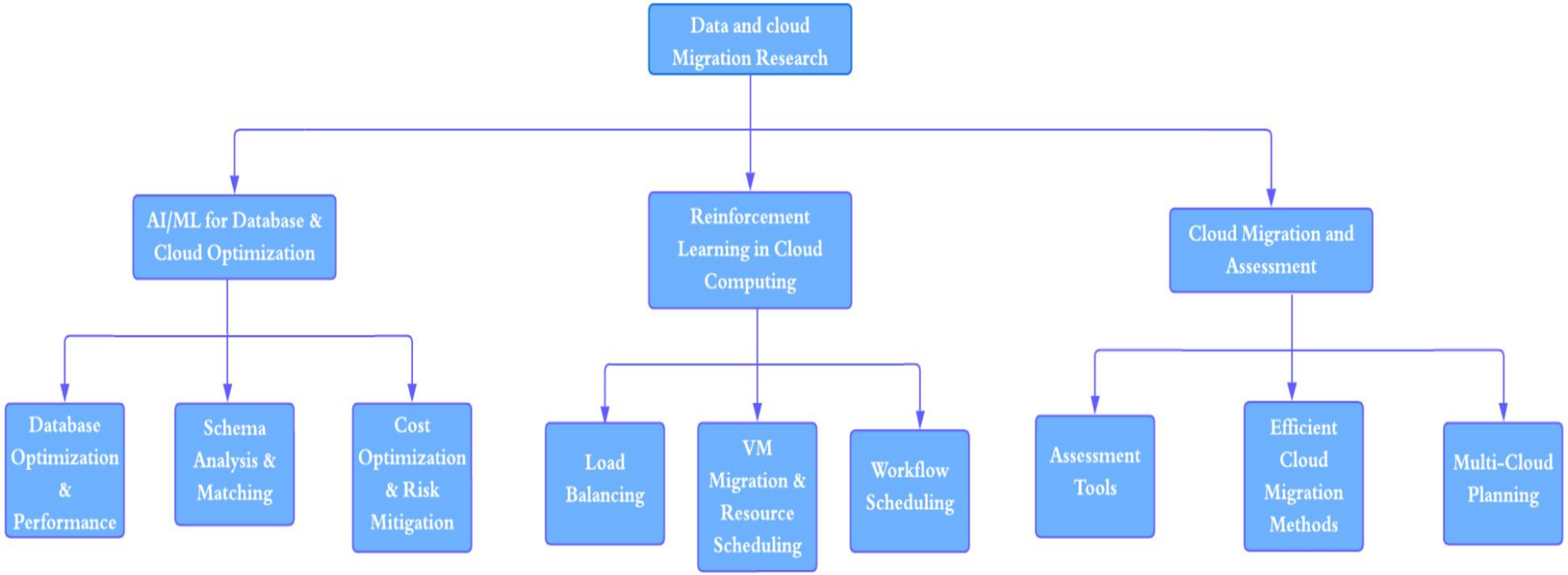
2.2 Third-Party Tools

Tool	Type	Focus Limitations
Matillion	ETL/ELT	Data pipeline automation Limited assessment capabilities
Fivetran	Data Integration	Automated data pipelines No migration planning
Airbyte	Open-source	Data integration Lacks dynamic planning
Flyway/Liquibase	Database Migration	Schema version control Limited to schema changes

3. Core Research Areas & Literature review

DOMAIN	TITLE	YEAR	DESCRIPTION
AI/ML for Database Migration and Optimization	Reinforcement Learning for Database Configuration Tuning	2021	This paper explores RL applications in database optimization, relevant for dynamic resource allocation.
	Graph Neural Networks for Database Schema Matching	2022	It uses GNNs for schema analysis and mapping, applicable to dependency analysis.
	Intelligent Cloud Migration: A Survey	2023	This comprehensive survey of AI-driven migration approaches identifies gaps in current automation.
Cloud Migration and Assessment	Automated Assessment of Cloud Migration Readiness	2023	This work focuses on pre-migration assessment automation, though it is limited to rule-based approaches.
	Multi-Cloud Database Migration Strategies	2023	It addresses multi-cloud scenarios but lacks intelligent planning mechanisms.
Reinforcement Learning in Cloud Computing	Deep Reinforcement Learning for Cloud Resource Management	2023	This paper explores RL applications in cloud resource allocation, relevant for dynamic planning.

3. Core Research Areas



Literature review

Paper Title (Link)	Year	Authors	Source	Key Contribution
A Survey on Deep Reinforcement Learning for Cloud Resource Scheduling	2024	Zhou et al.	Springer (AI Review)	Comprehensive survey covering DRL applications in cloud environments; highlights suitability of RL for complex scheduling and migration tasks.
Security-aware and Cost-effective Database Migration	2023	Acikalin et al.	Springer (Constraints Journal)	Proposes an NP-hard problem formulation for risk-aware migration and uses memetic algorithms to optimize security shifts and cost.
Reinforcement Learning-Based Resource Management for VM Migration	2021	Dai et al.	MDPI (Applied Sciences)	Presents a multi-agent DRL model for VM migration in edge-cloud scenarios using centralized training and decentralized execution.
Load Balancing with TGNN-RL in Cloud Environments	2025	Rajammal & Chinnadurai	Springer Nature (Scientific Reports)	Combines Temporal Graph Neural Networks with RL for dynamic load balancing, modeling inter-resource dependencies.
GrapheonRL: Graph-Based RL for Workflow Scheduling	2025	Sharma & Kunkel	IEEE (COMPSAC 2025)	Introduces a hybrid GNN-RL approach for heterogeneous cloud workflow scheduling, achieving near-optimal makespan.
SMig-RL: An Evolutionary DRL Framework for Cloud Migration	2020	Ren et al.	ACM (TOIT)	Evolutionary Deep RL system to learn cloud migration strategies balancing performance and cost across services.
DMML: Predicting Data Migration Performance with ML	2025	Ghaneshirazi et al.	ACM (ICPE Companion)	ML model trained on migration benchmarks to predict migration time and bottlenecks, enabling planning.
Heterogeneous Load Balancing in Cloud Environments	2023	Nutipalli & Das	Springer (SN Computer Science)	Proposes a custom non-ML heuristic for load balancing in heterogeneous clouds, optimizing energy and time.
A Systematic Review on Dynamic Load Balancing Techniques	2022	Tawfeeg et al.	IEEE Access	SLR covering cloud-based fault tolerance and dynamic resource allocation methods; identifies gaps for RL use.
Intelligent Cloud Migration: A Survey	2023	Patel et al.	IEEE Cloud Computing	Survey of intelligent migration approaches, highlighting ML/AI-based tools and automation gaps.

4. Research Gap

4.1 Technical Gaps

- **Absence of End-to-End RL-Driven Planning**: Current tools use static, rule-based approaches
- **Insufficient Risk Prediction**: Lack of AI-powered risk assessment and mitigation
- **No dynamic adaptation**: Static planning approaches dominate

4.2 Market Gaps

- **Cloud-Agnostic Solutions**: Market dominated by vendor-specific tools
- **SME Accessibility**: High-cost enterprise solutions with limited SME options
- **No end-to-end automation**: Most of the existing tools require significant manual intervention

5. Problem Statement

Database migration to cloud environments has become a critical need for enterprises, but current approaches suffer from high failure rates, cost overruns, and operational disruptions.

1 **High Failure Rates**

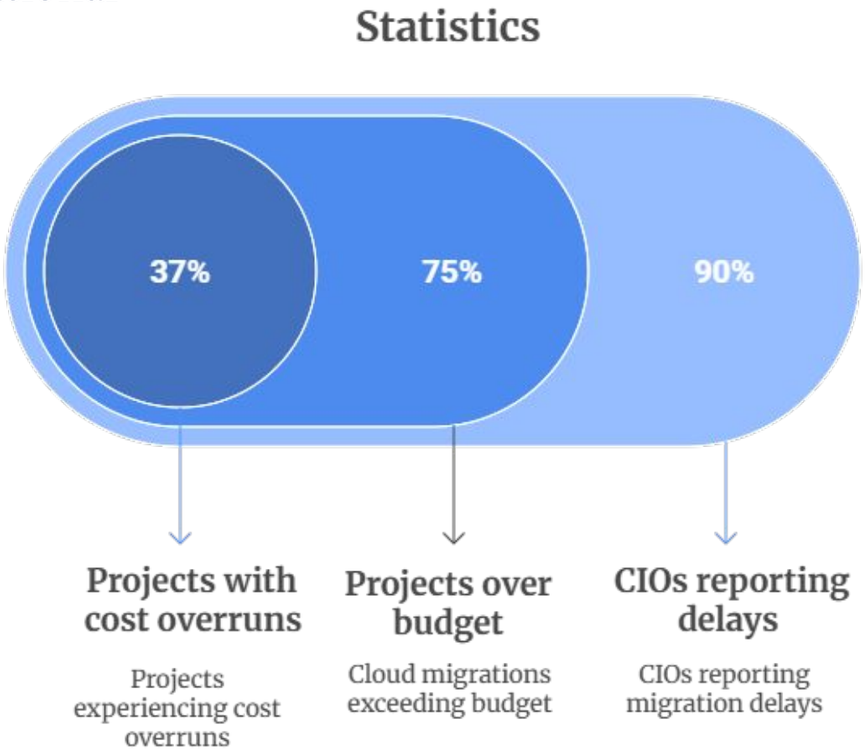
Over 80% of data migration projects fail to meet deadlines or budgets.

2 **Vendor lock-in**

Limited cloud-agnostic planning capabilities

3 **Lack of automated assessment**

Manual discovery and assessment of complex database environments



6. Proposed Solution

6.1 Proposed System Architecture

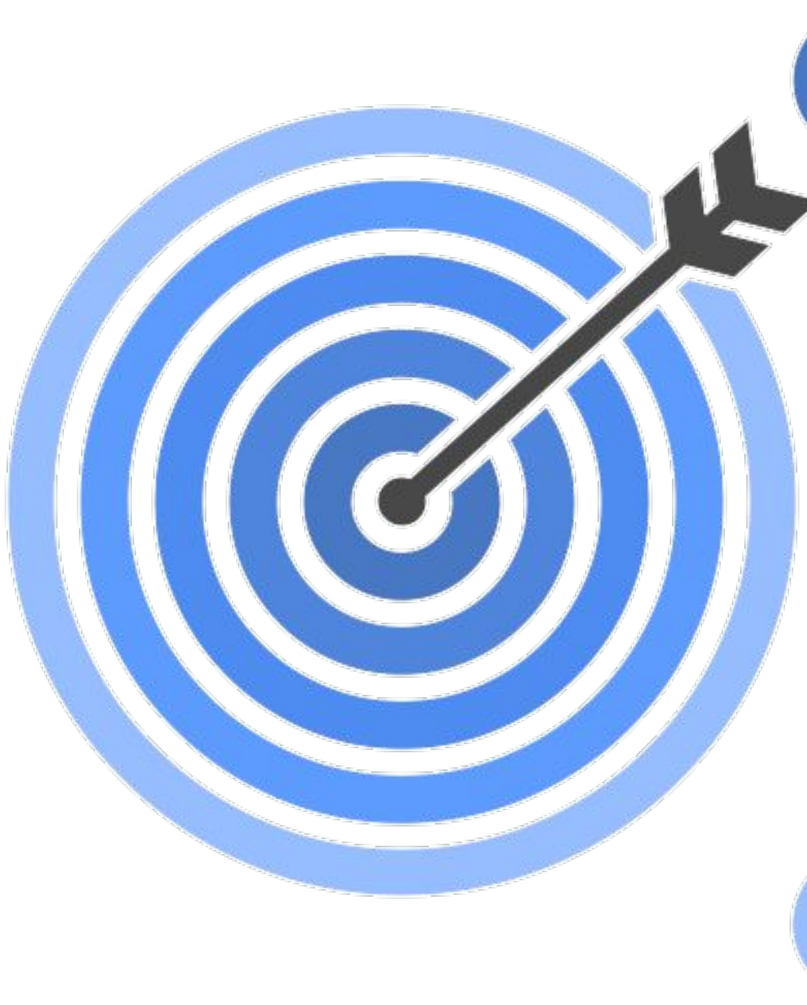
Pre-Migration Assessment and Dynamic Planning Tool using:

- **Reinforcement Learning** for dynamic planning optimization
- **Multi-Agent Architecture** for specialized assessment tasks
- **Graph Neural Networks** for complex dependency analysis
- **Predictive Analytics** for risk assessment and cost estimation (ML)

6.2 Key Innovation Points

- **RL-Driven Planning**: Dynamic adaptation to changing conditions and continuous learning
- **Multi-Cloud Intelligence**: Cloud-agnostic recommendations and planning
- **Predictive Risk Assessment**: AI-powered risk prediction and mitigation strategies
- **End-to-End Automation**: Minimal manual intervention required

7. Project Objectives



Autonomous Assessment System

Core system for database evaluation

RL-based Migration Strategies

Dynamic planning using reinforcement learning

Predictive Models

Forecasting risks and costs

Cloud-agnostic Planning

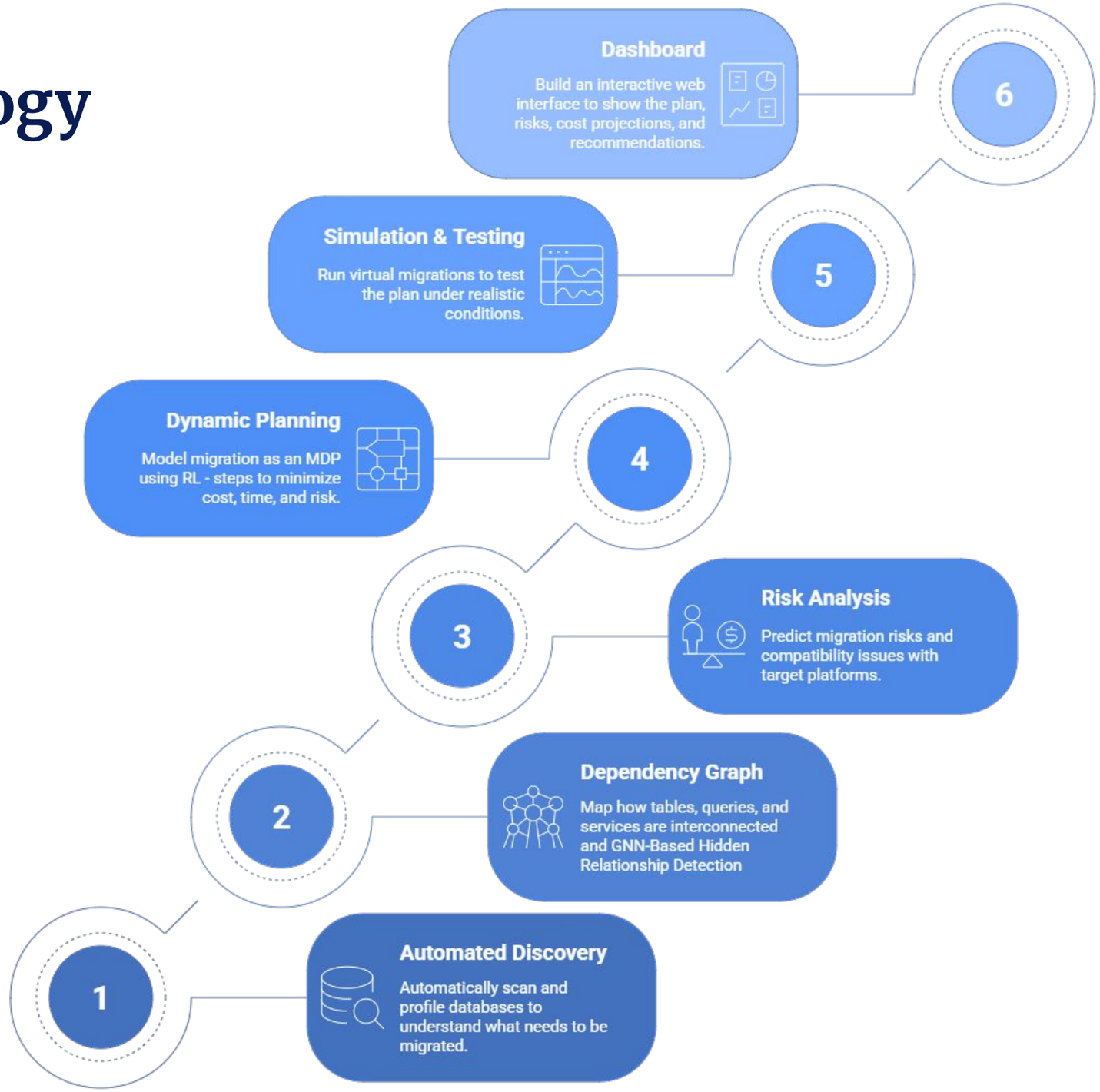
System adaptable to various cloud platforms

Explainable AI Recommendations

Transparent AI insights for decision-making

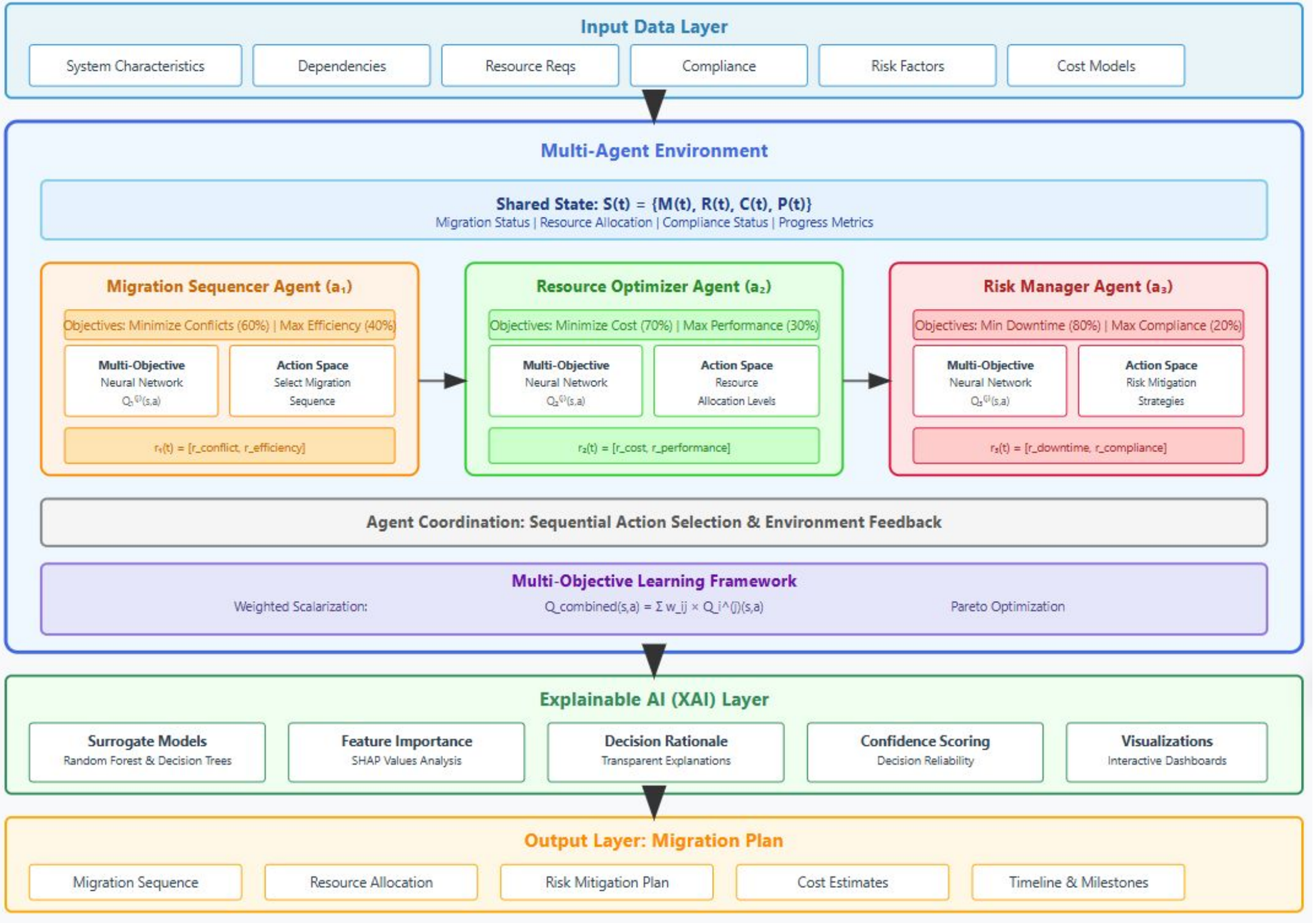
Web app features

8. Methodology



Multi-Agent Multi-Objective Reinforcement Learning Framework

for Intelligent Cloud Migration Planning



8. RL

Reward Function:

$$\mathbf{r}_1(t) = [r_{conflict}, r_{efficiency}]$$

where:

$$r_{conflict} = -\alpha_1 \sum_{s_i \in A(t)} |unmet\_deps_i|$$

$$r_{efficiency} = \beta_1 \cdot |A(t)| + \gamma_1 \cdot parallel\_score$$

$$\mathbf{r}_2(t) = [r_{cost}, r_{performance}]$$

where:

$$r_{cost} = -\alpha_2 \cdot \frac{current\_cost}{baseline\_cost}$$

$$r_{performance} = \beta_2 \cdot \frac{achieved\_perf}{target\_perf}$$

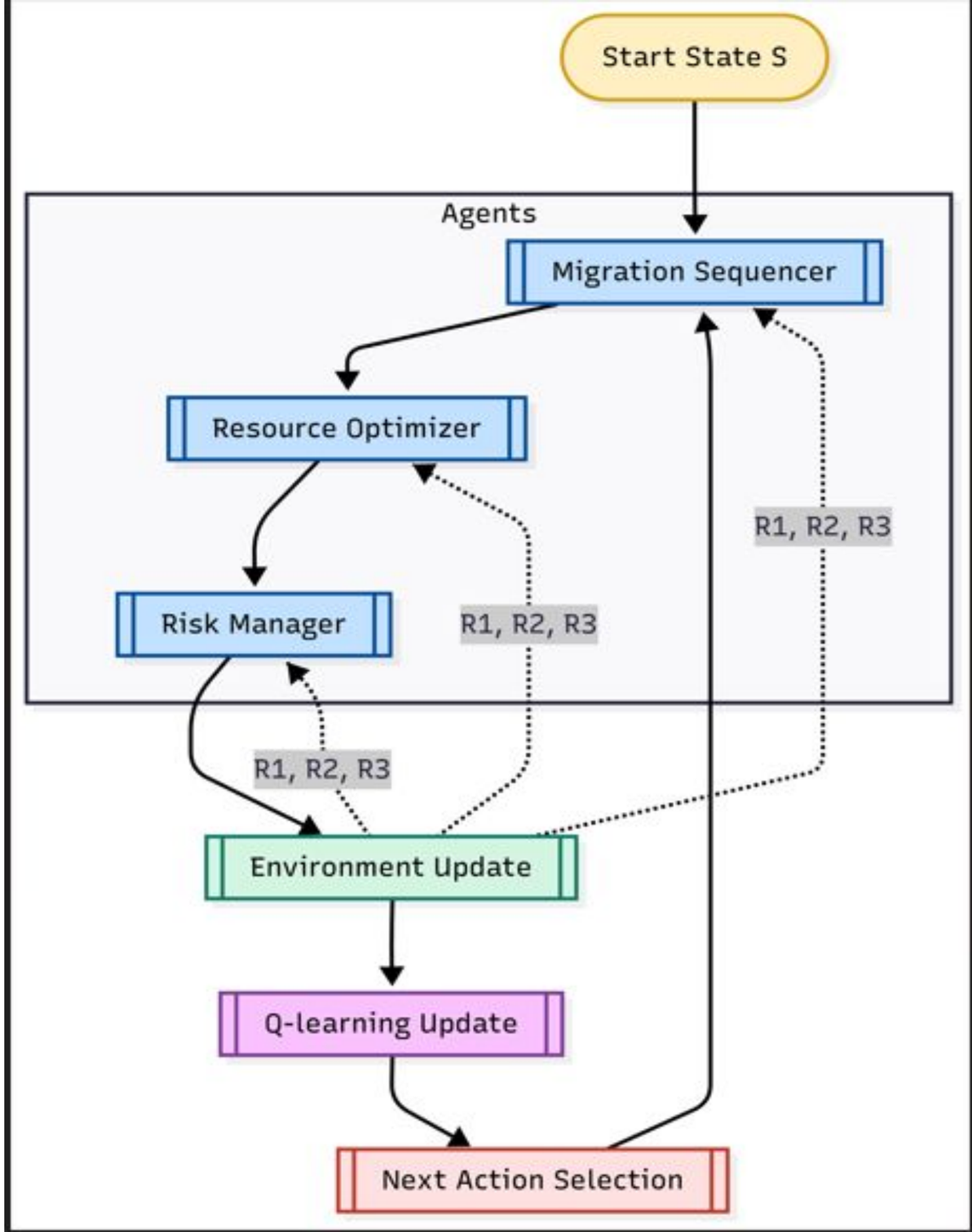
$$\mathbf{r}_3(t) = [r_{downtime}, r_{compliance}]$$

where:

$$r_{downtime} = -\alpha_3 \cdot estimated\_downtime$$

$$r_{compliance} = \beta_3 \cdot compliance\_score$$

$$Q_i^{(j)}(s_t, a_t) \leftarrow Q_i^{(j)}(s_t, a_t) + \alpha[r_{i,j} + \gamma \max_{a'} Q_i^{(j)}(s_{t+1}, a') - Q_i^{(j)}(s_t, a_t)] \quad (16)$$



Tech stack components

Programming Language

Python is the primary language used for development. Java for few components

Data Processing

Pandas, NumPy, Matplotlib/Plotly are used. These technologies are used for data manipulation, analysis, and visualization.

AI/ML Frameworks

Reinforcement Learning(Gymnasium), Deep Learning(PyTorch), Traditional ML and Graph Processing(NetworkX) are used. These frameworks are essential for AI and ML tasks.

Web Framework

Streamlit (intially), Next.js

Database Technologies

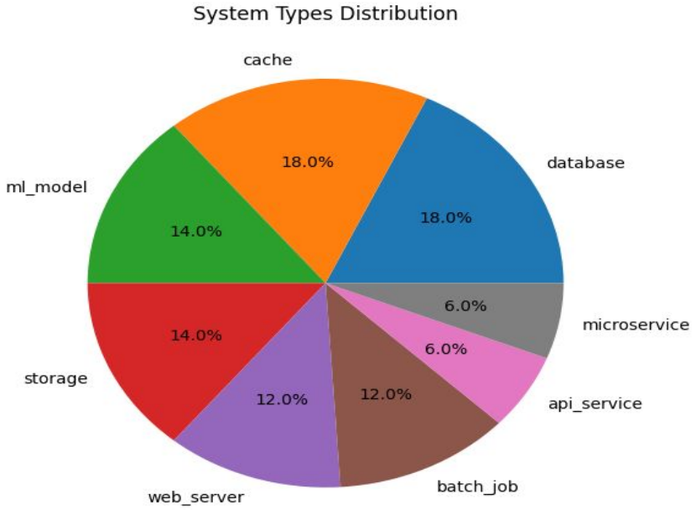
SQLAlchemy, psycopg2, mysql-connector-python, PostgreSQL, and MySQL are used. These technologies are used for database interaction and management.

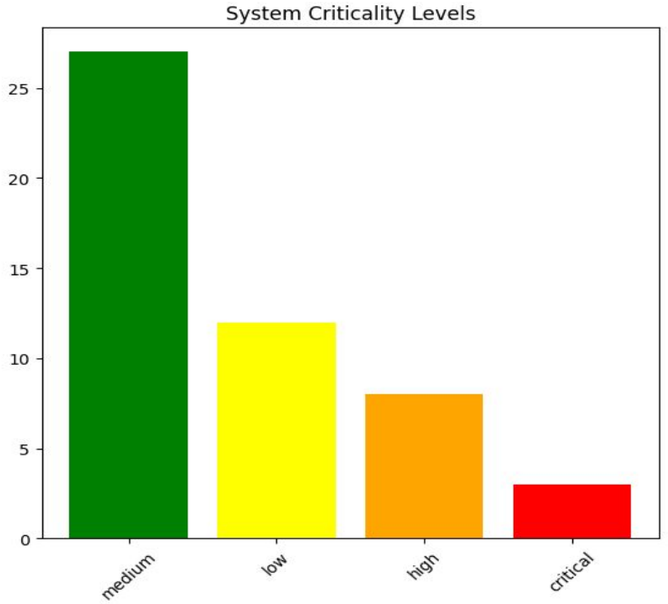
Cloud Technologies

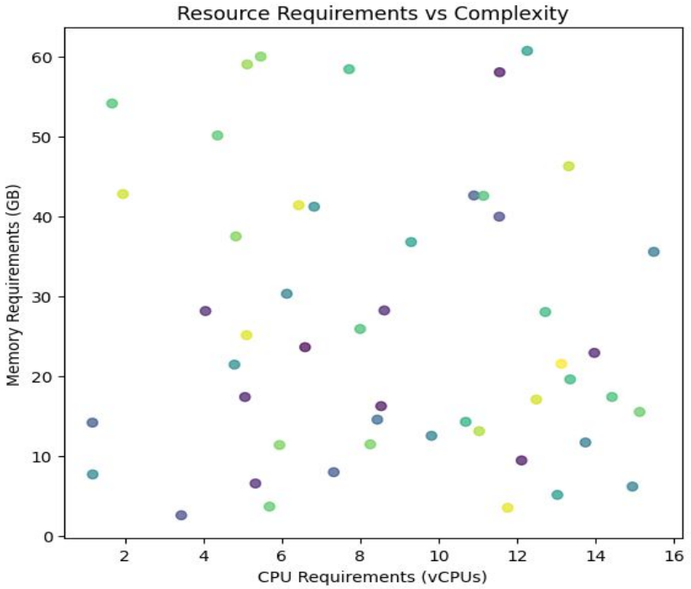
AWS SDK, Azure SDK, and Google Cloud are used. These technologies are used for cloud service interaction.

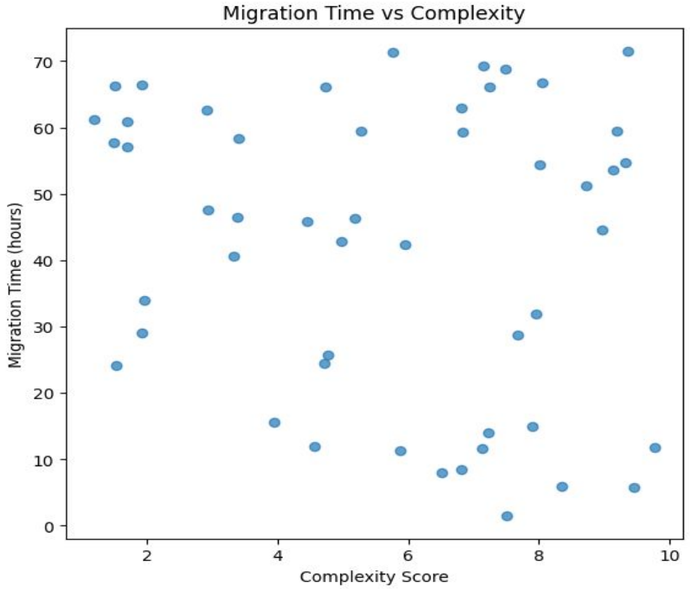
RESULTS

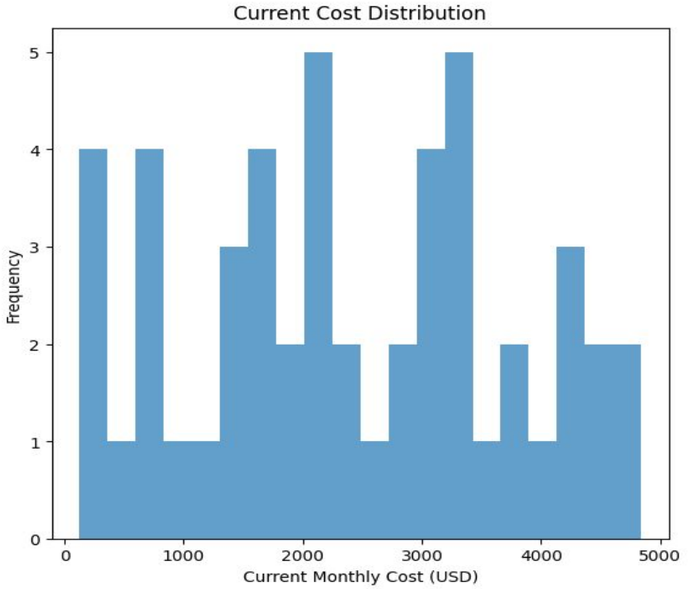
Cloud Migration Dataset : System Characteristics

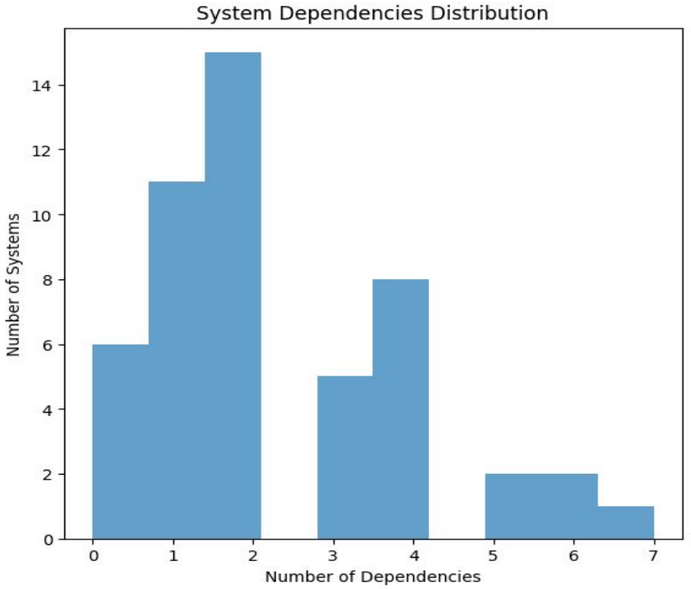




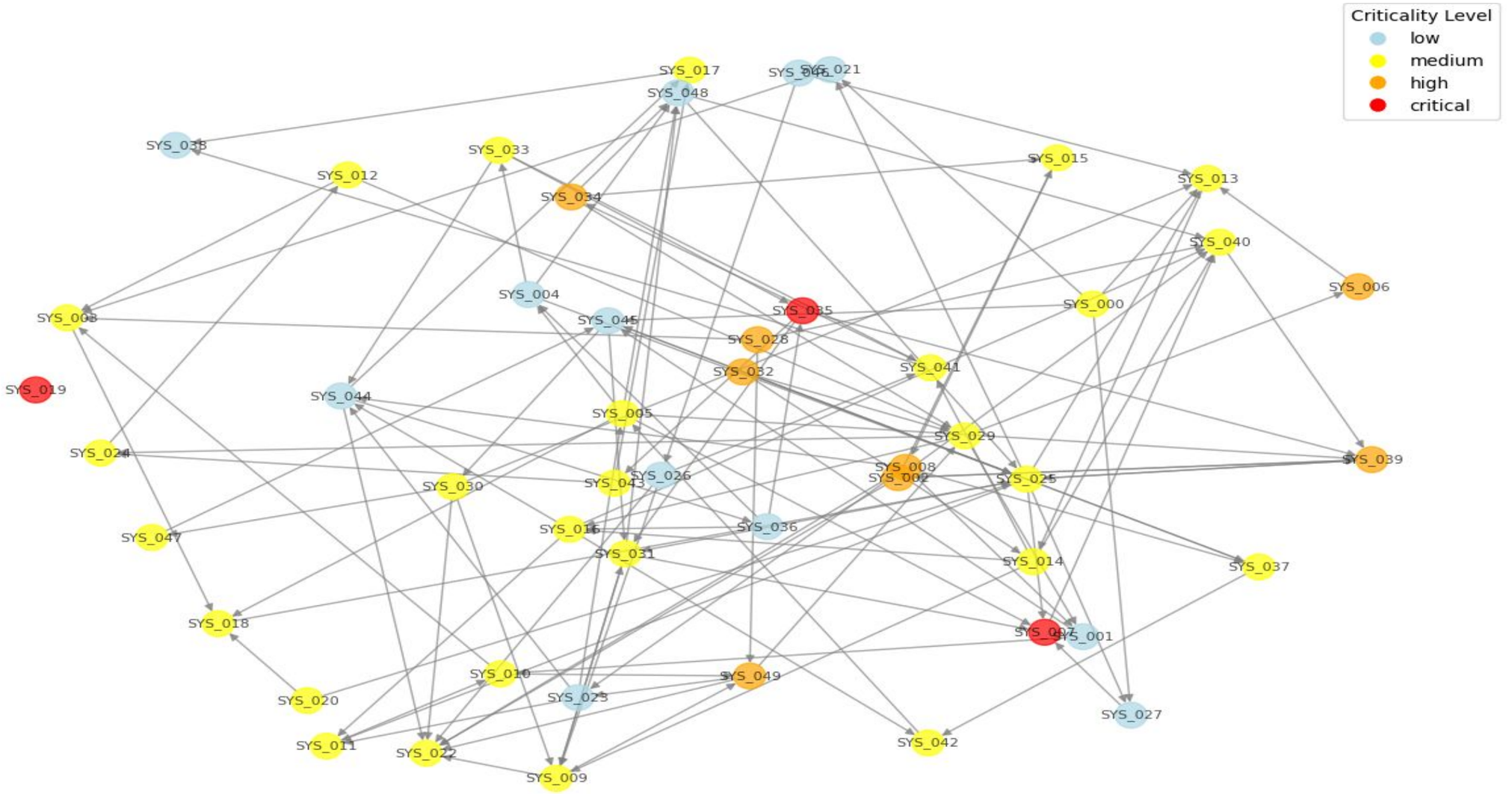




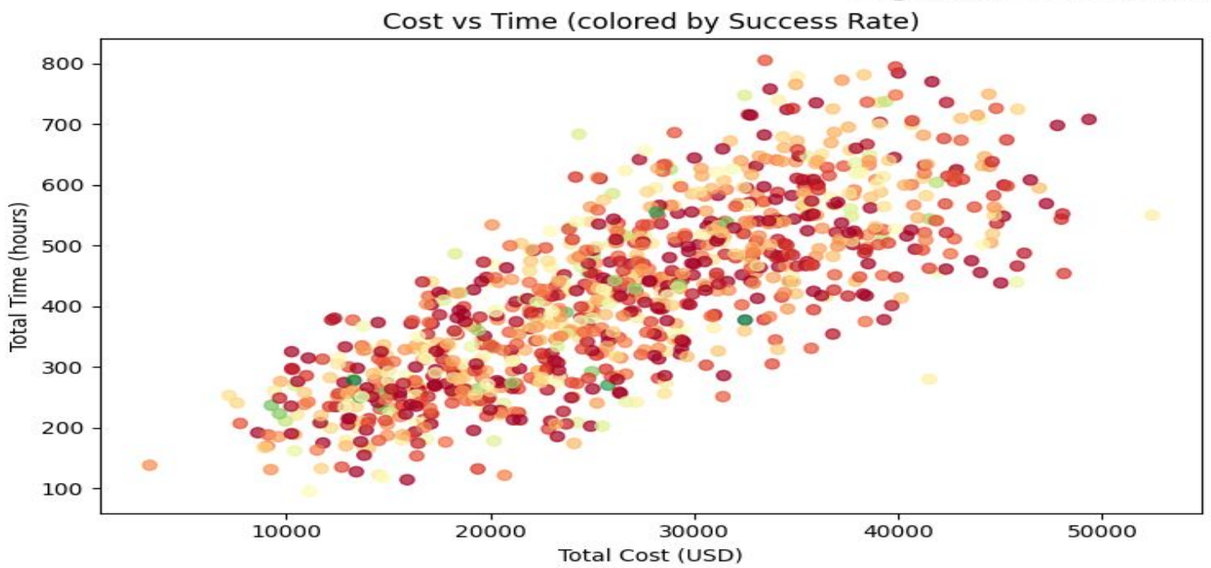


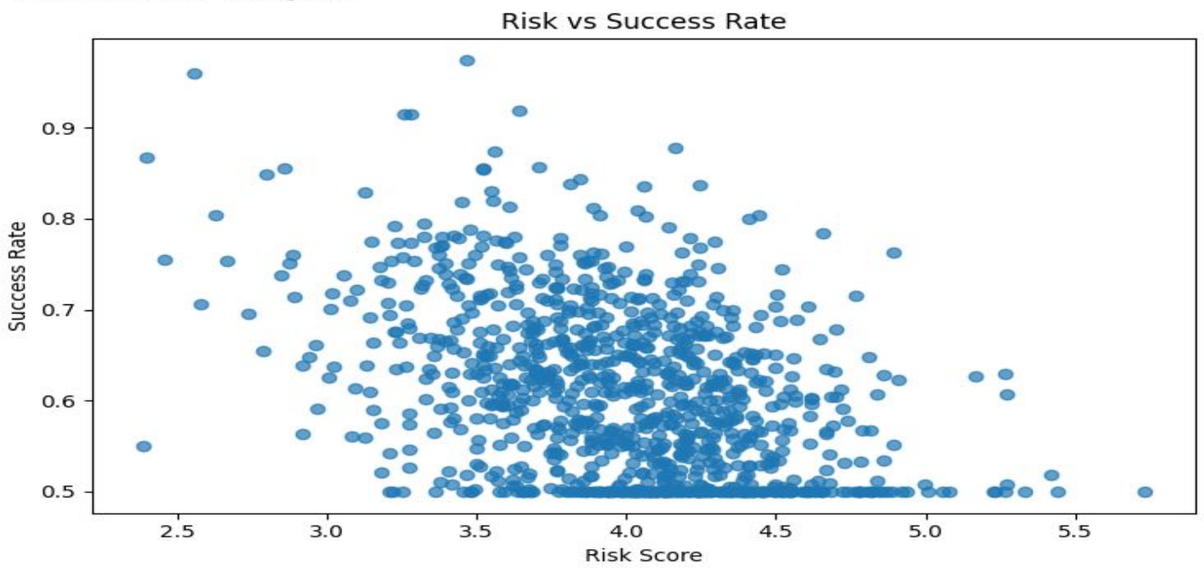


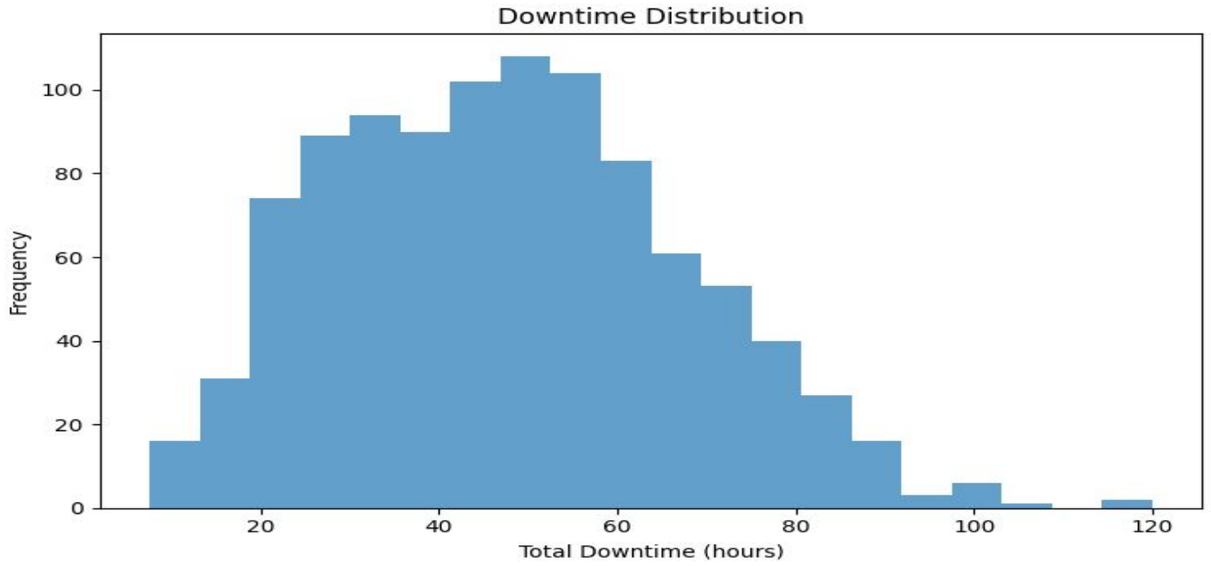
System Dependencies Network

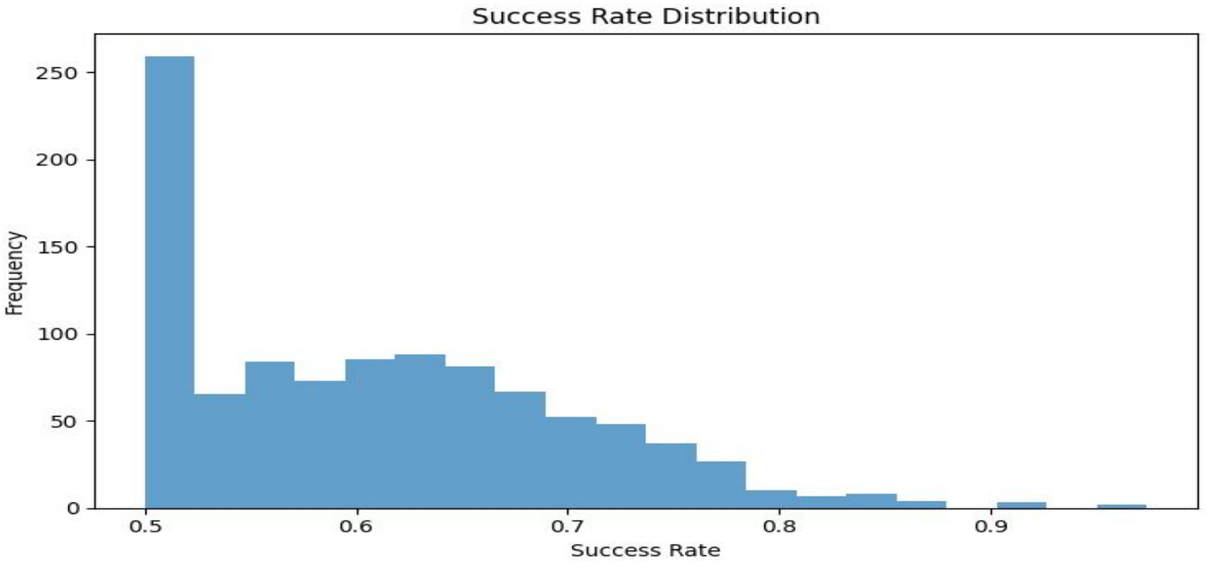


Migration Scenarios: Outcome Analysis

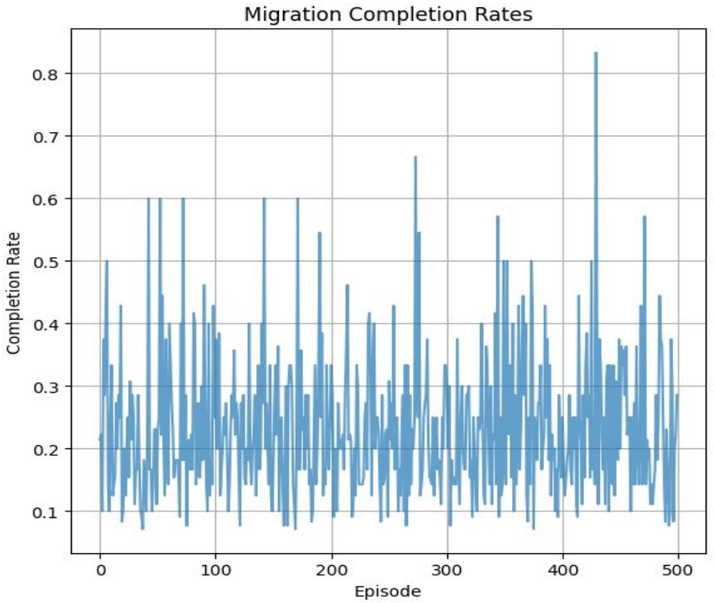


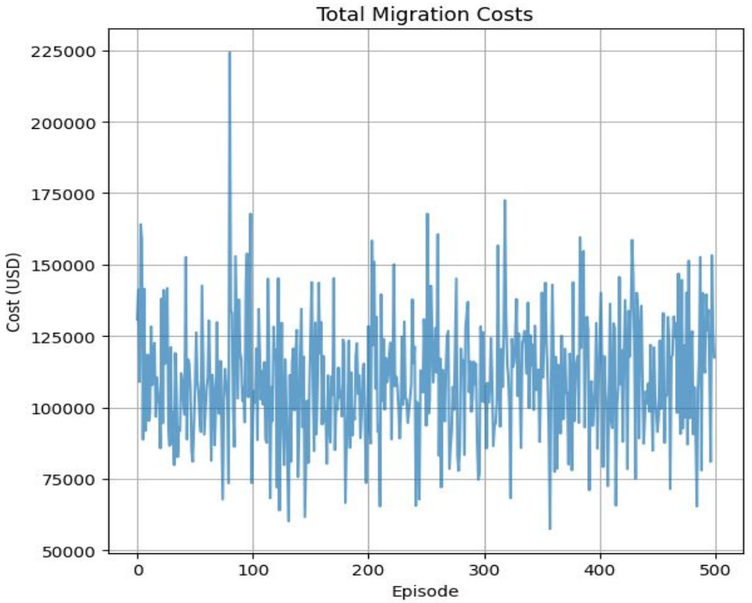


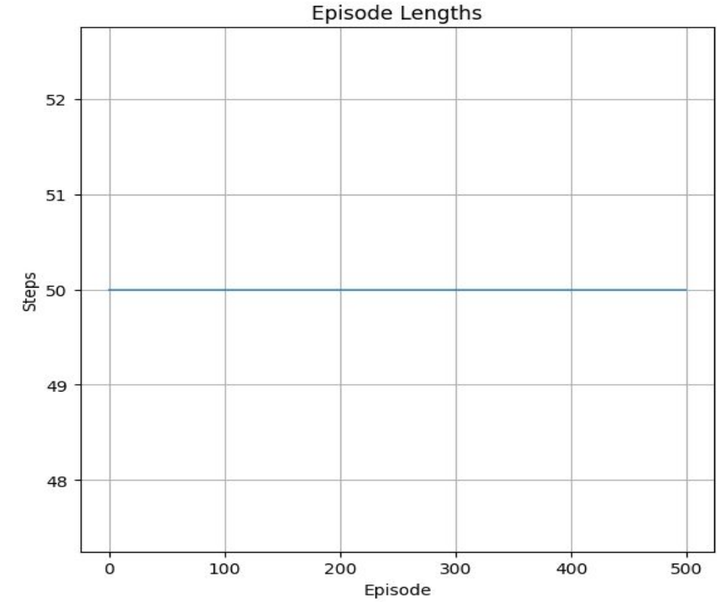


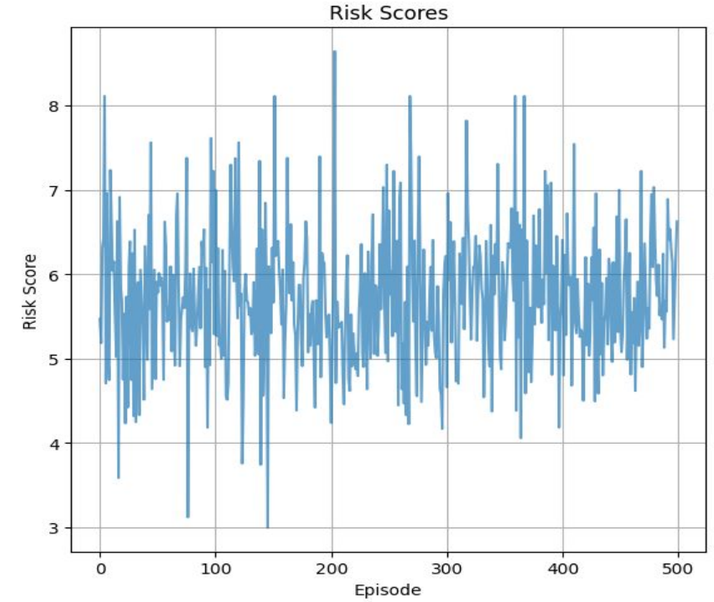


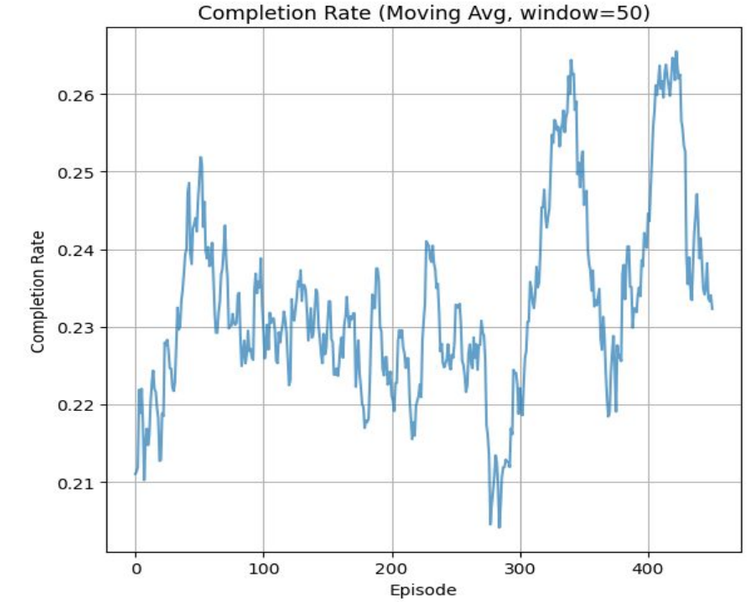
Multi-Agent Training Process



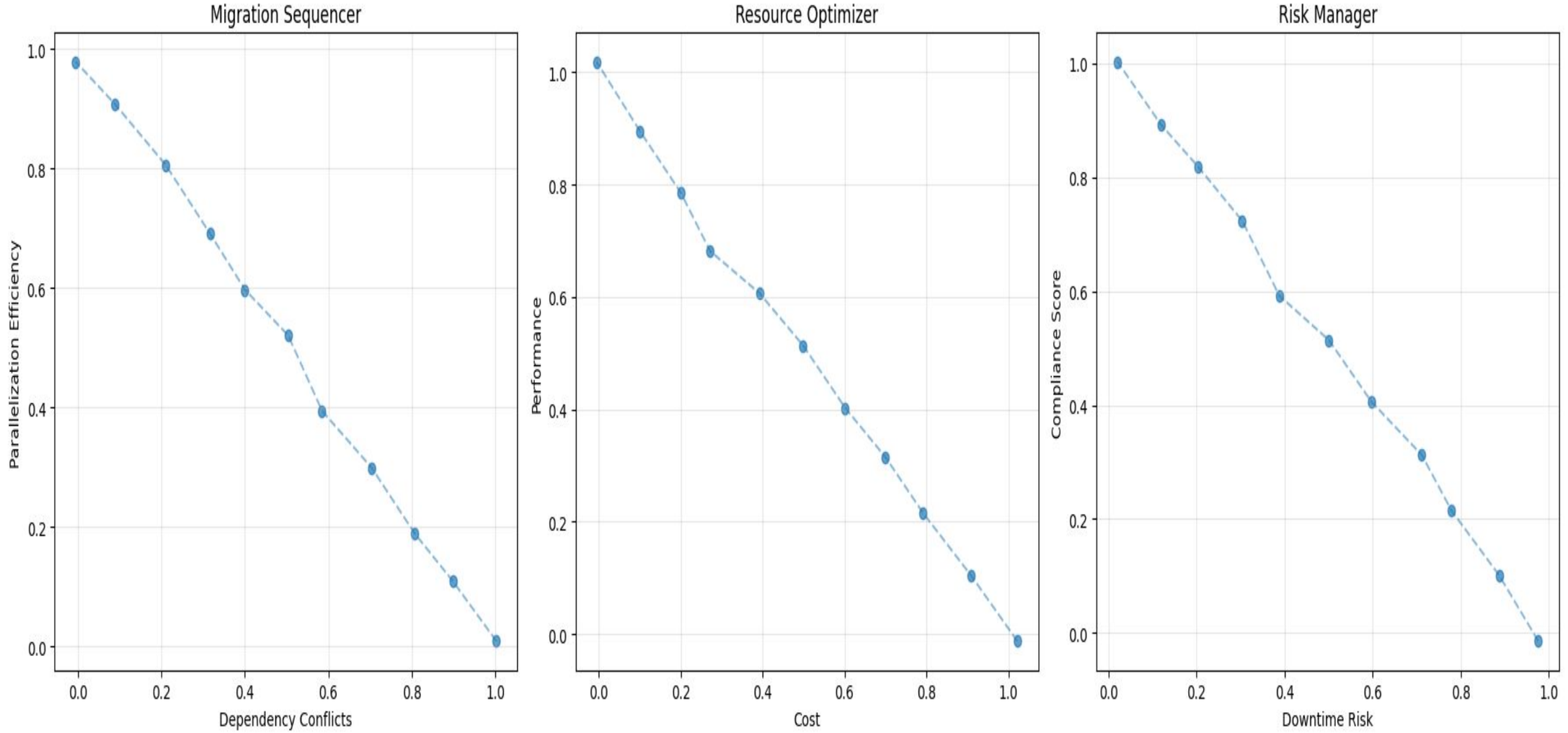




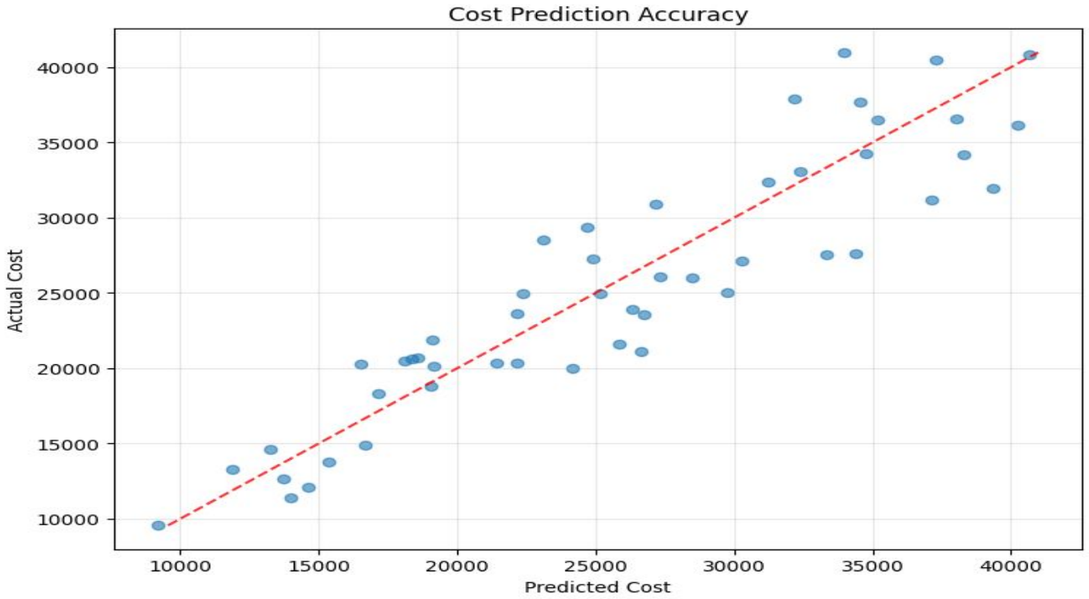


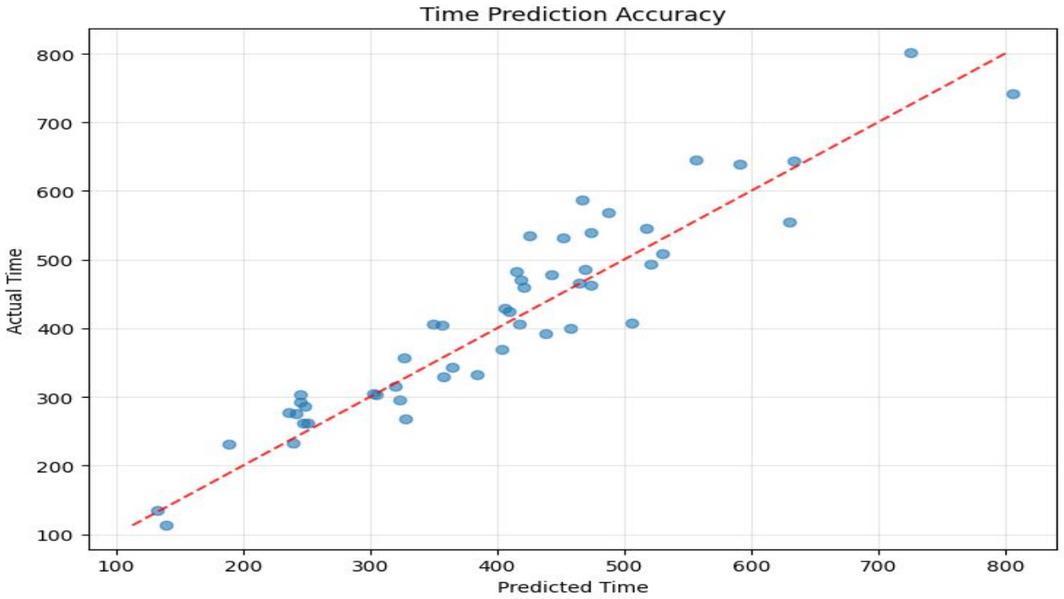


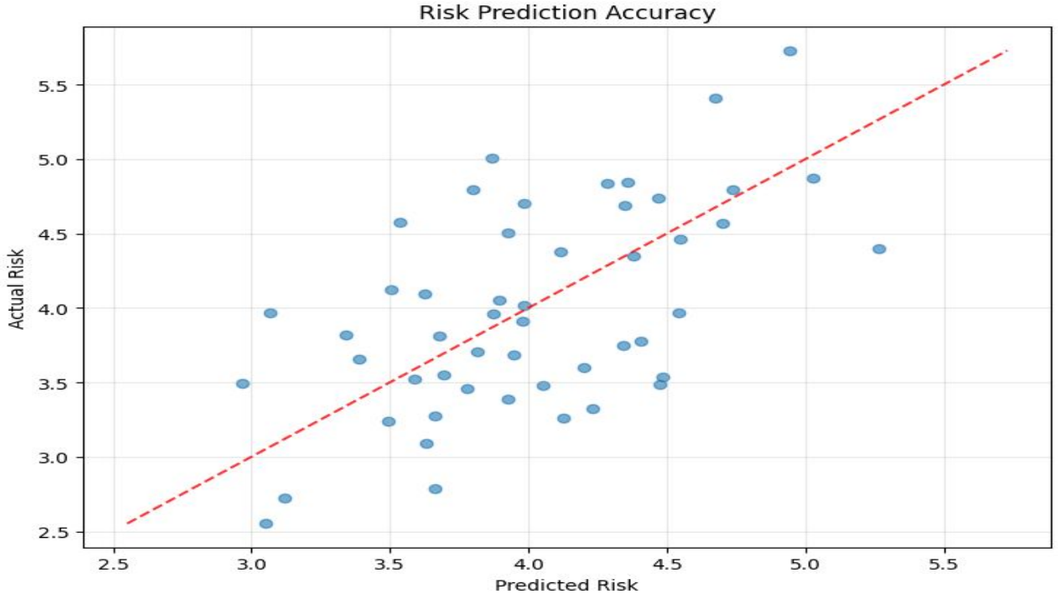
Parento Fonts for Multi-Objective Optimisation

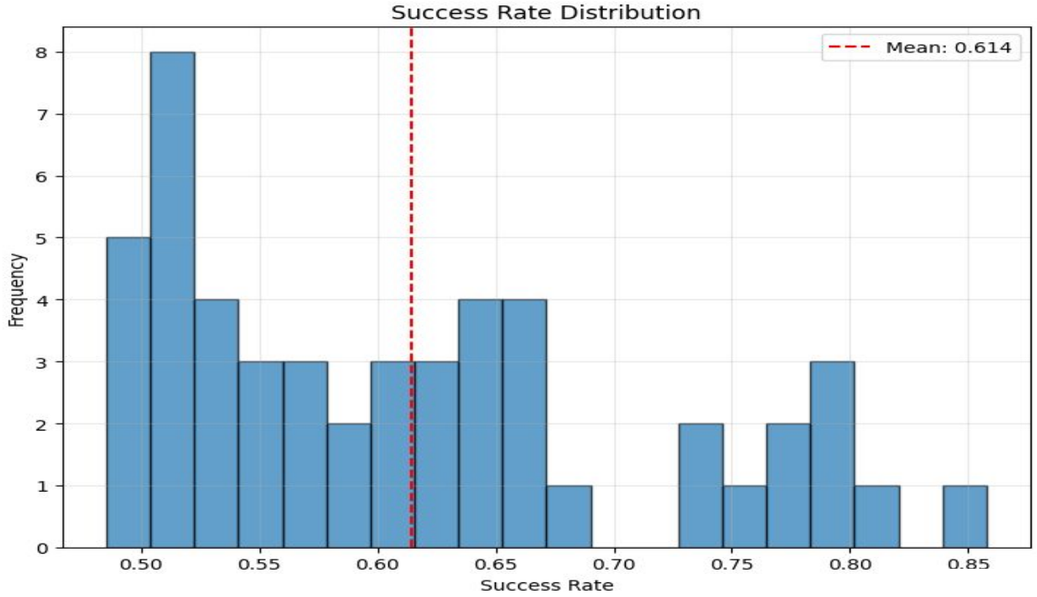


Test Scenario Result Analysis

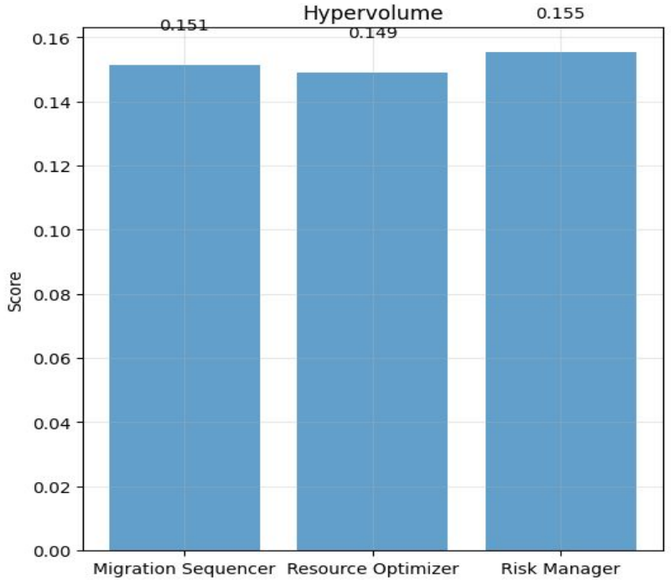


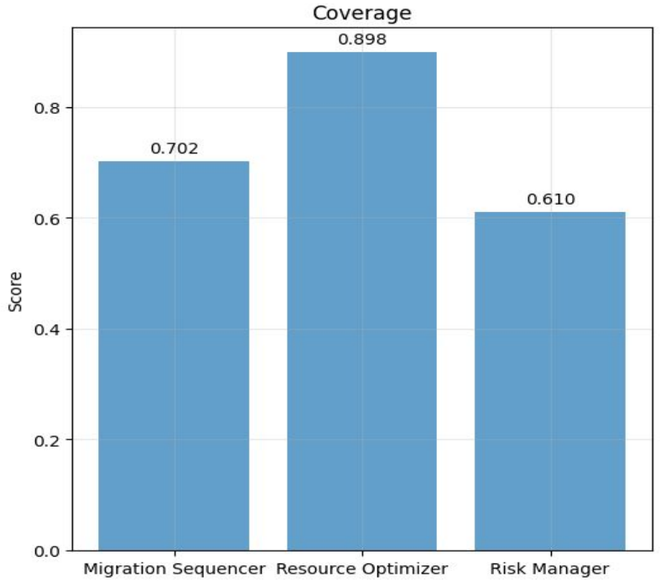


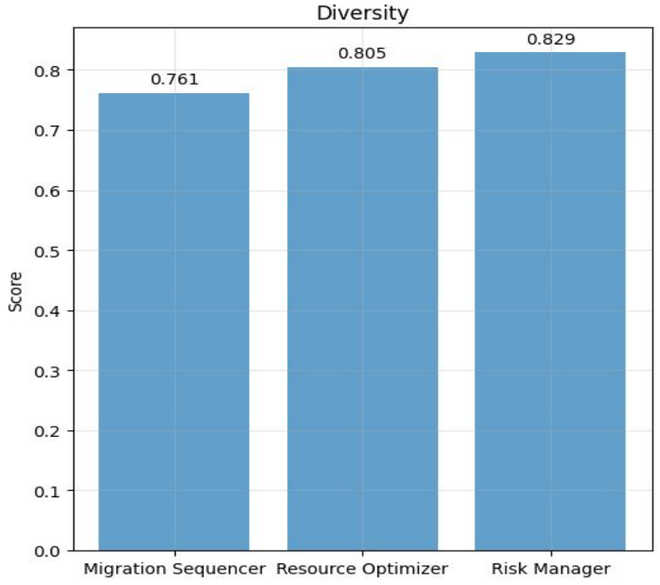


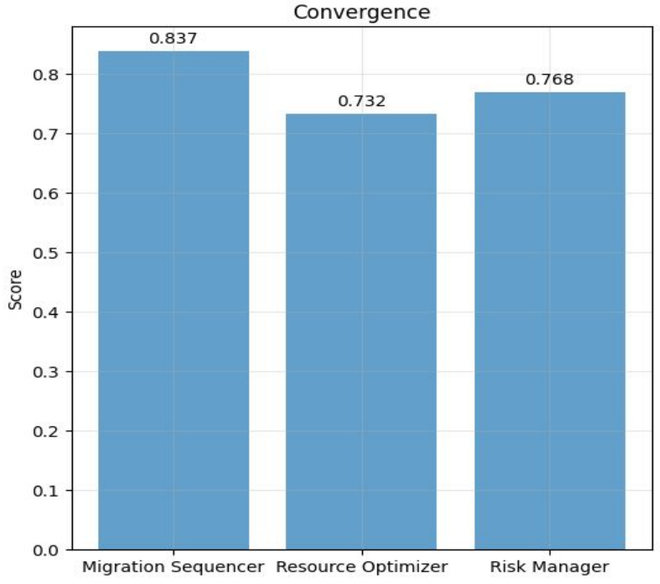


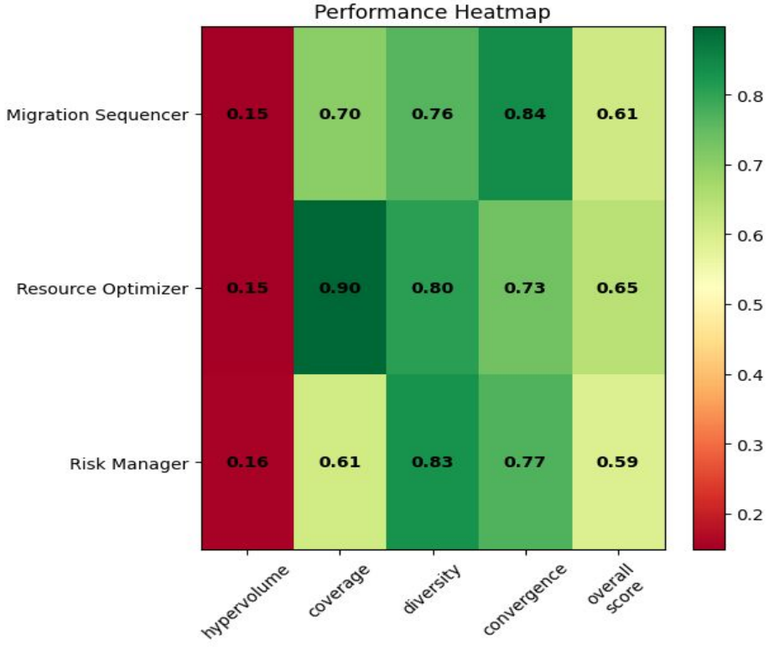
Multi-Objective Performance Comparison





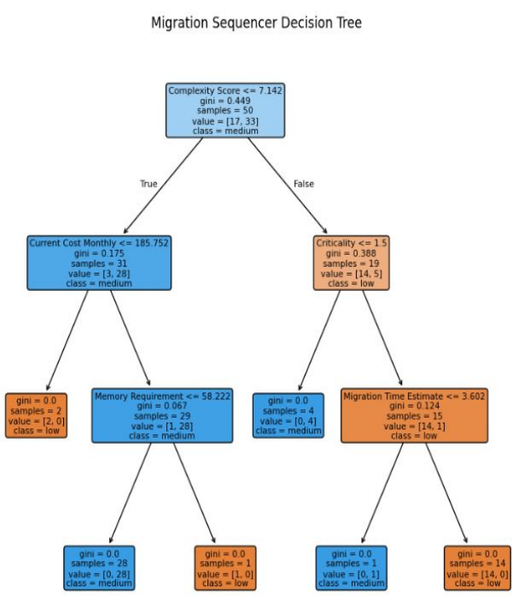


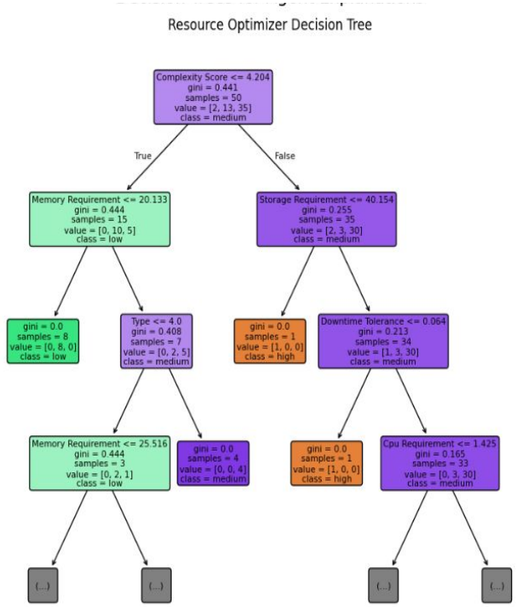


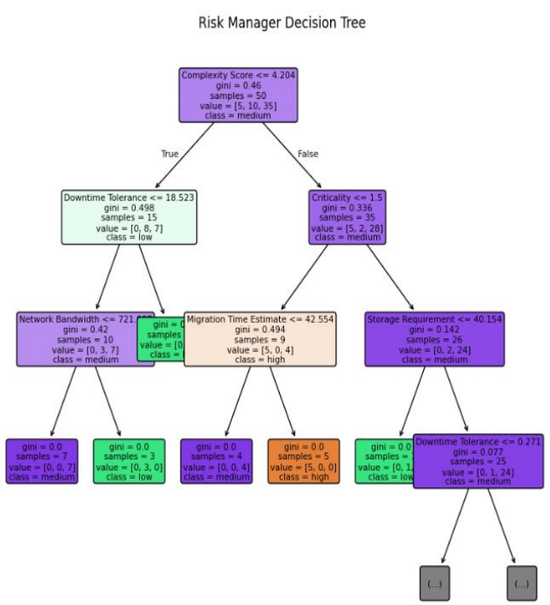


Evaluation Report Summary

```
Generating comprehensive evaluation report...
Evaluating individual agent performance...
Testing agents on 50 scenarios...
Calculating multi-objective performance scores...

================================================================================
EVALUATION REPORT SUMMARY
================================================================================

Scenarios Tested: 50
Average Success Rate: 0.618
Cost Prediction Accuracy: 0.905
Time Prediction Accuracy: 0.920
Risk Prediction Accuracy: 0.908

Agent Rankings:
  1. Resource Optimizer: 0.651
  2. Risk Manager: 0.644
  3. Migration Sequencer: 0.602

Recommendations:
  • Migration Sequencer: Focus training on improving convergence_rate through additional scenarios
  • Resource Optimizer: Focus training on improving convergence_rate through additional scenarios
  • Risk Manager: Focus training on improving convergence_rate through additional scenarios

Evaluation results saved to 'evaluation_results.json'
Evaluation completed successfully!
```

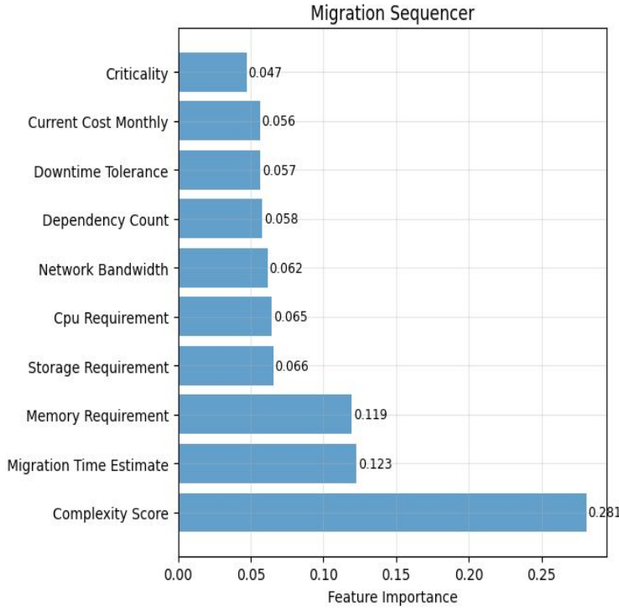


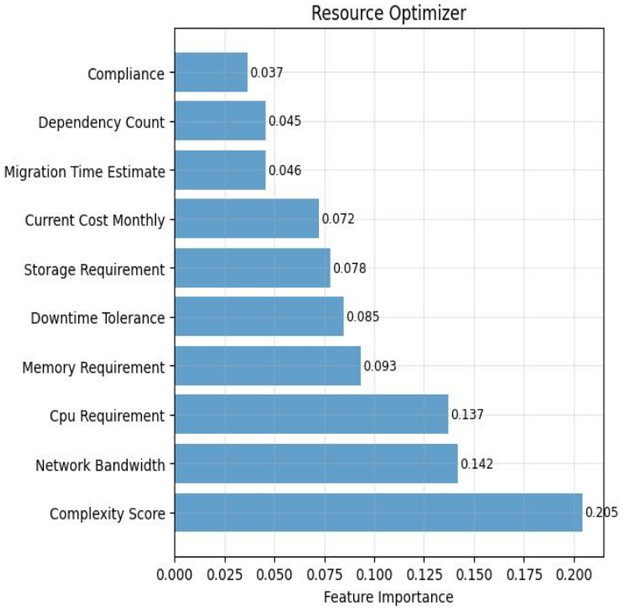


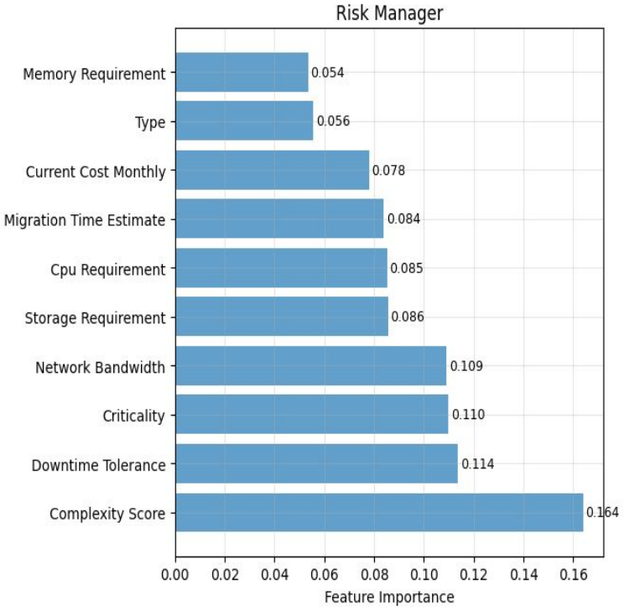


Decision Tree for Agent Explanations

Feature Importance Comparison across Agents



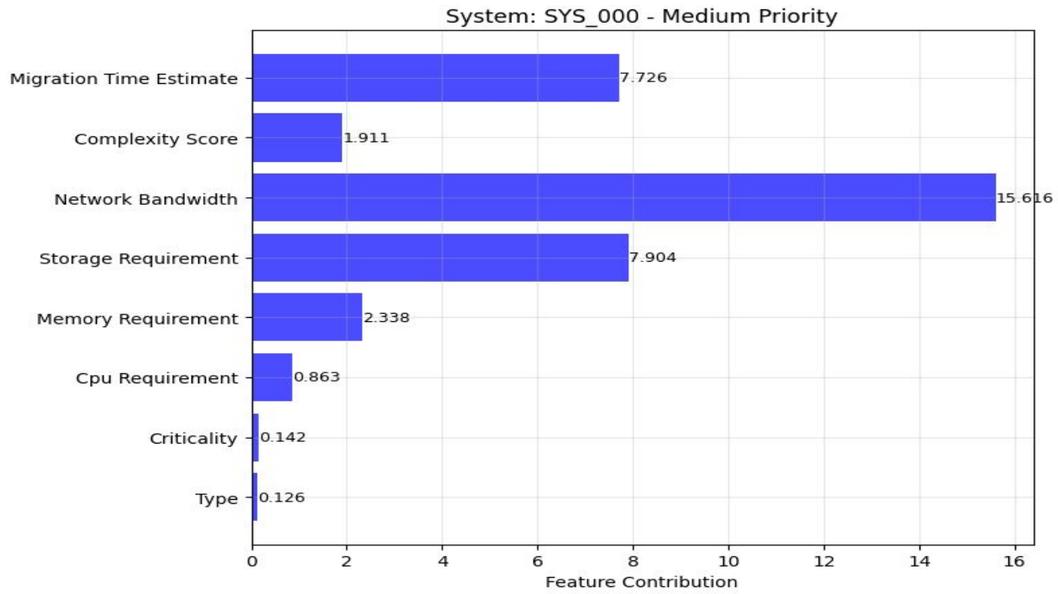


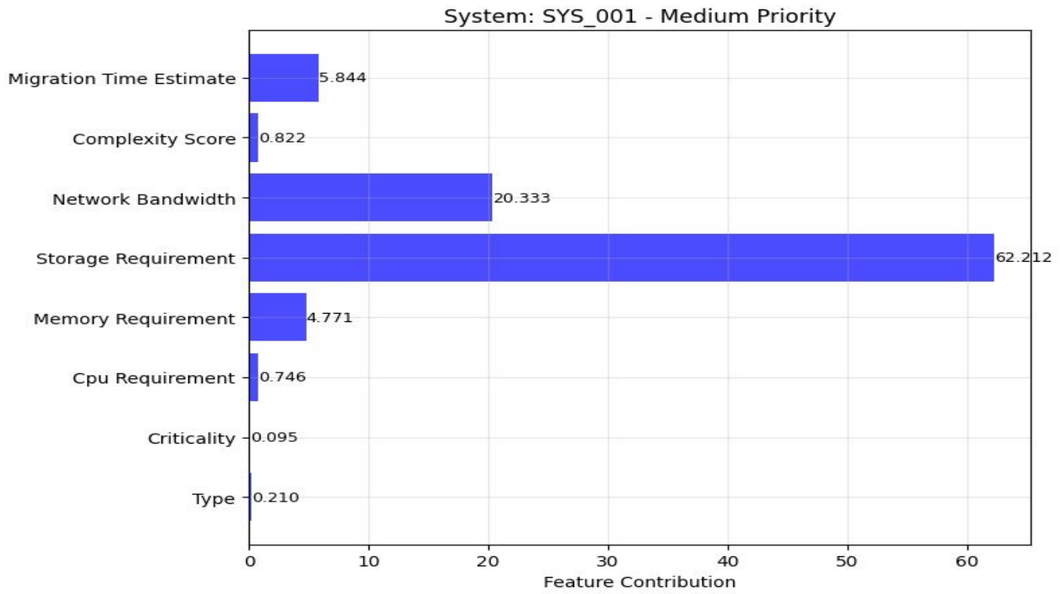


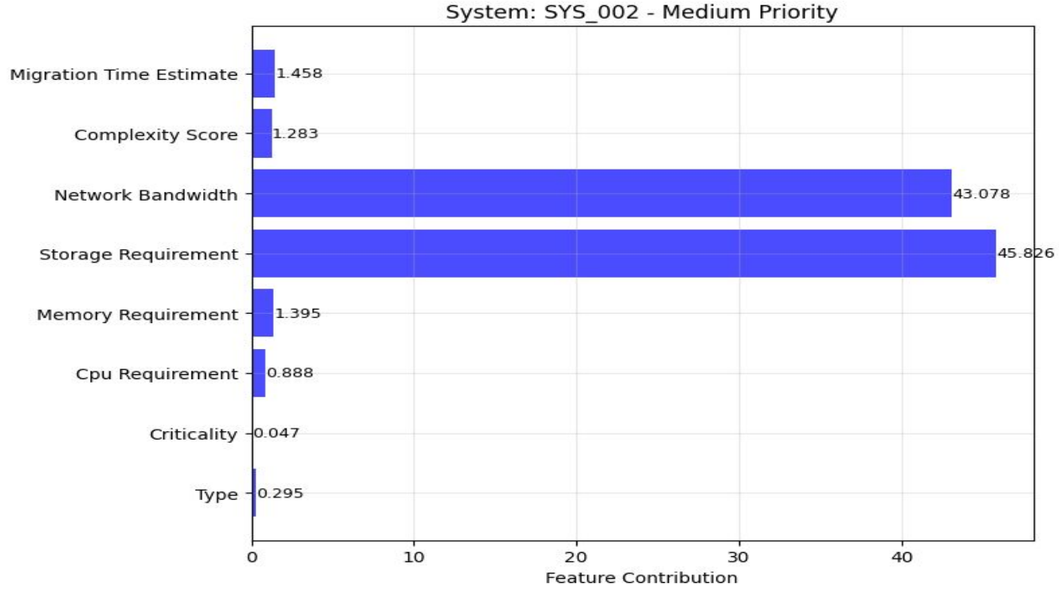
Feature Importance Analysis

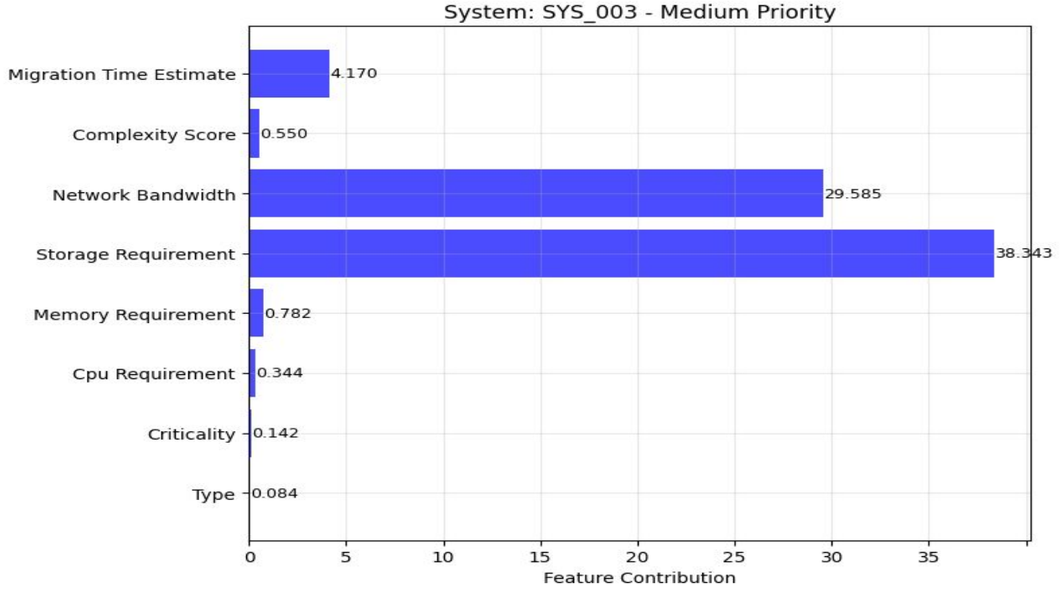
```
============================================================
FEATURE IMPORTANCE ANALYSIS
============================================================

Migration Sequencer:
--------------------
  Complexity Score: 0.281
  Migration Time Estimate: 0.123
  Memory Requirement: 0.119
  Storage Requirement: 0.066
  Cpu Requirement: 0.065

Resource Optimizer:
-------------------
  Complexity Score: 0.205
  Network Bandwidth: 0.142
  Cpu Requirement: 0.137
  Memory Requirement: 0.093
  Downtime Tolerance: 0.085
...
    - Storage Requirement: 10.304
```

SHAP – Style Features Contribution Explanations

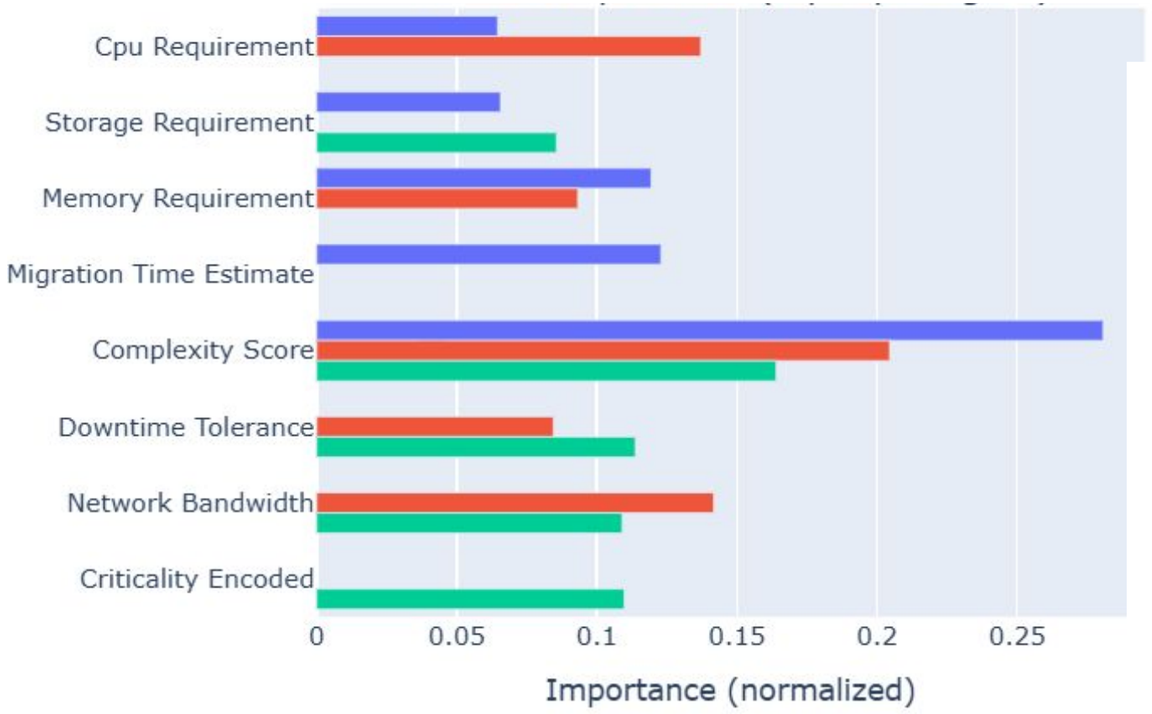




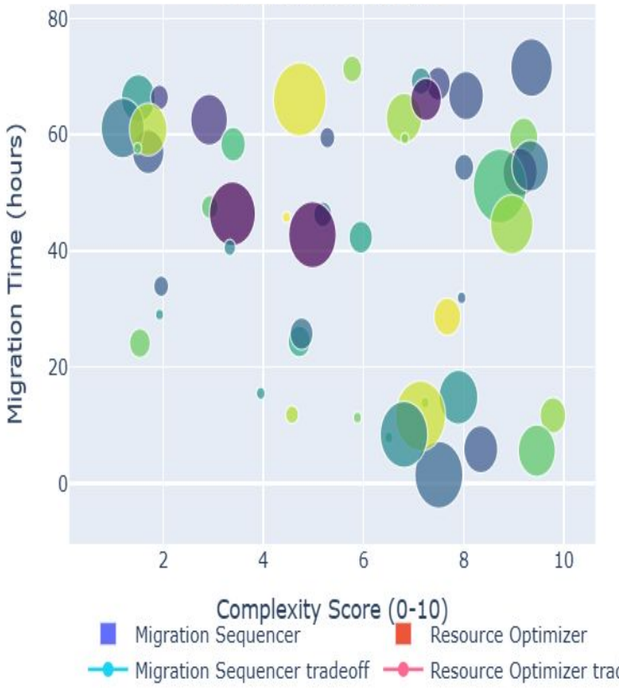




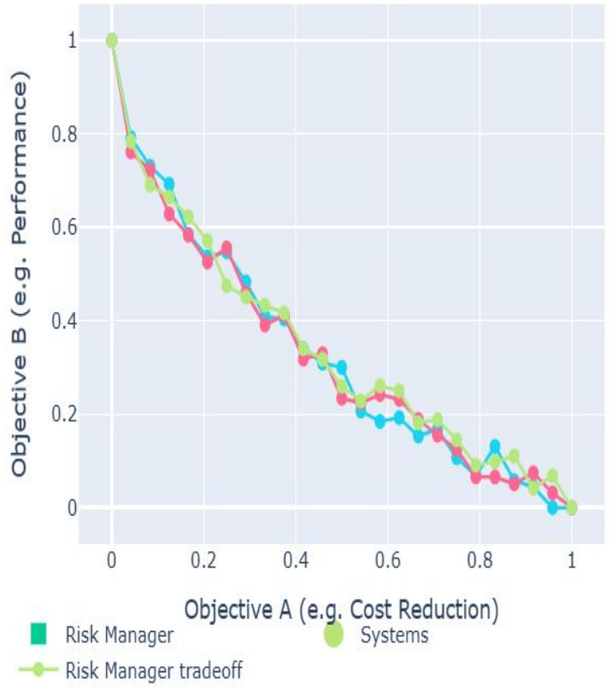
Feature Importance



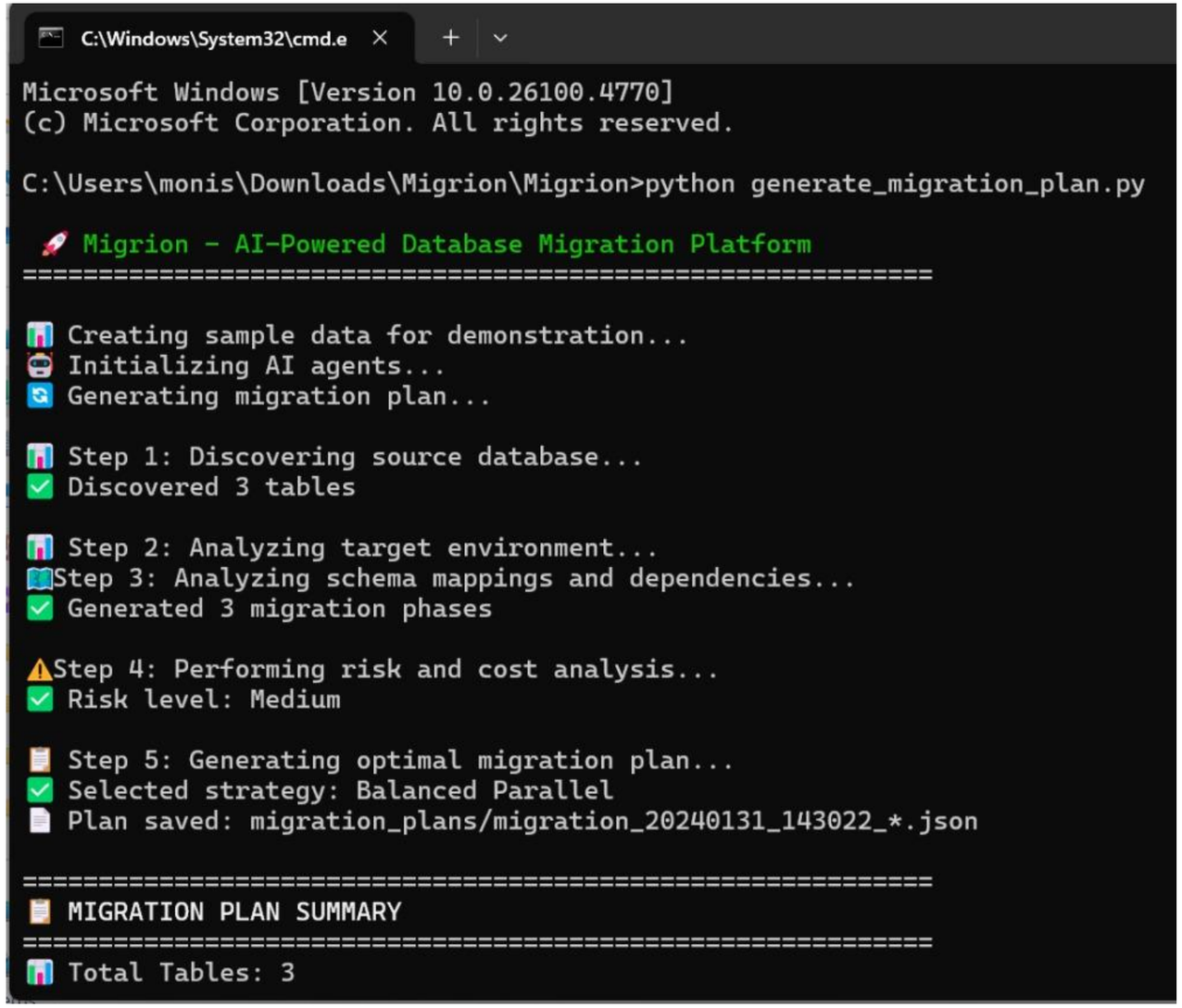
System Risk Profile

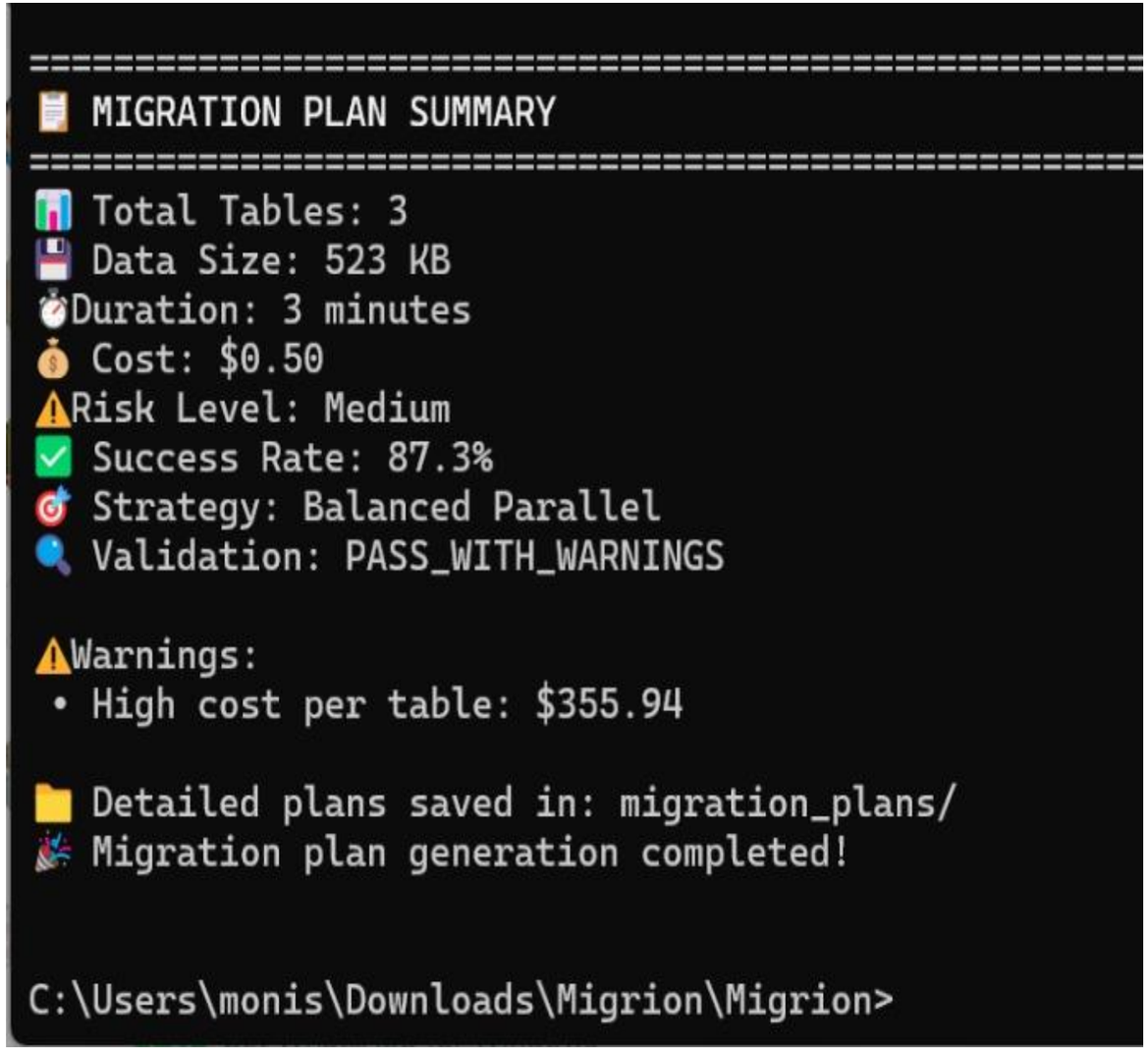


Multi Objective Trade off



Results - terminal output





ER diagrams generated after analyzing data

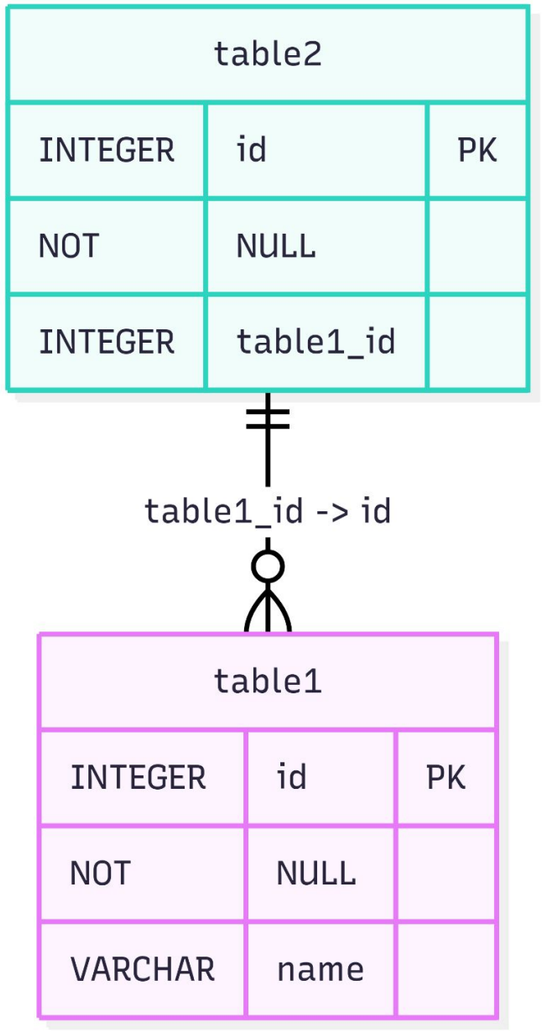


table3		
INTEGER	id	PK
NOT	NULL	
TEXT	description	

Plan generated for Orange League Ventures Technologies Pvt. Limited

Enterprise Database Migration Assessment Report

Prepared for: Orange League Ventures Technologies Private Limited
Prepared by: Migrion Technologies Pvt. Ltd.
Date: December 31, 2024
Project Code: TCS-MIG-2024-Q4
Classification: Confidential - Internal Use Only

Executive Summary

This document presents a comprehensive pre-migration assessment and strategic migration plan for Orange League Ventures Technologies Private Limited's database infrastructure. The assessment covers the migration of the organization's core Oracle Database 19c Enterprise Edition system to Amazon RDS PostgreSQL 15.4, encompassing 2.8 TB of critical business data across 47 tables and 8 database schemas.

Key Findings

Assessment Area	Current State	Recommended Target	Business Impact
Database Platform	Oracle Database 19c Enterprise	Amazon RDS PostgreSQL 15.4	Cost reduction of ₹39.84L annually
Data Volume	2.8 TB across 47 tables	Same volume, optimized structure	Improved query performance
Infrastructure Cost	₹66L annually (licensing + maintenance)	₹32.16L annually (AWS RDS)	51.2% cost reduction
Migration Duration	N/A	28.5 hours over 4 weekends	Minimal business disruption
Risk Assessment	High complexity system	Mitigated through phased approach	91.2% success probability

2. Migration Strategy & Approach

2.1 Strategic Objectives

Primary Goals:

1. **Cost Optimization:** Achieve 51% reduction in database operational costs
2. **Scalability Enhancement:** Leverage cloud-native scaling capabilities
3. **Performance Improvement:** Reduce average query response time by 20%
4. **Business Continuity:** Maintain >99.9% system availability during migration
5. **Compliance Adherence:** Ensure continued regulatory compliance

2.2 Migration Methodology

Approach: Phased Migration with Zero-Downtime Strategy

This methodology minimizes business risk by:

- Migrating non-critical systems first to validate processes
- Implementing real-time data synchronization
- Maintaining rollback capabilities at each phase
- Conducting comprehensive testing before each cutover

2.3 Target Architecture

Amazon RDS PostgreSQL Configuration:

- **Instance Type:** db.r6g.8xlarge (32 vCPU, 256 GB RAM)
- **Storage:** 3 TB General Purpose SSD (GP3)
- **Multi-AZ Deployment:** Yes (Mumbai and Hyderabad)
- **Backup Retention:** 35 days automated, 7 years archived
- **Monitoring:** Enhanced monitoring with 1-second granularity

Security Configuration:

- Encryption at rest using AWS KMS
- SSL/TLS for all connections

Plan generated for Orange League Ventures Technologies Pvt. Limited

3. Risk Assessment & Mitigation

3.1 Risk Analysis Matrix

Risk Category	Risk Level	Probability	Impact	Mitigation Strategy
Data Volume Complexity	High	80%	High	Phased approach with parallel processing
Business Continuity	Critical	30%	Critical	Zero-downtime migration with CDC
Schema Conversion	Medium	60%	Medium	Automated tools + manual validation
Performance Degradation	Medium	40%	High	Comprehensive performance testing
Regulatory Compliance	High	20%	Critical	Detailed compliance validation
Team Readiness	Medium	70%	Medium	Structured training program
Network Connectivity	Low	10%	High	Redundant connectivity options

3.2 Critical Risk Mitigation Plans

3.2.1 Data Volume Management

Challenge: 2.8 TB data migration within acceptable timeframes **Mitigation:**

- AWS Database Migration Service (DMS) for bulk data transfer
- Parallel loading across 8 concurrent streams
- Data compression during transfer (40% size reduction expected)
- Pre-staging in Amazon S3 for optimized RDS import

5.1 Cost-Benefit Analysis

5.1.1 Migration Investment Breakdown

Cost Component	Amount (₹)	Amount ($)	Percentage
Professional Services	18,75,000	22,500	100%
- Migration Planning	2,50,000	3,000	13.3%
- Schema Conversion	4,75,000	5,700	25.3%
- Data Migration	6,25,000	7,500	33.3%
- Testing & Validation	2,50,000	3,000	13.3%
- Go-Live Support	1,25,000	1,500	6.7%
- Training & Knowledge Transfer	1,50,000	1,800	8.0%

5.1.2 Operational Cost Comparison

Current Oracle Environment (Annual):

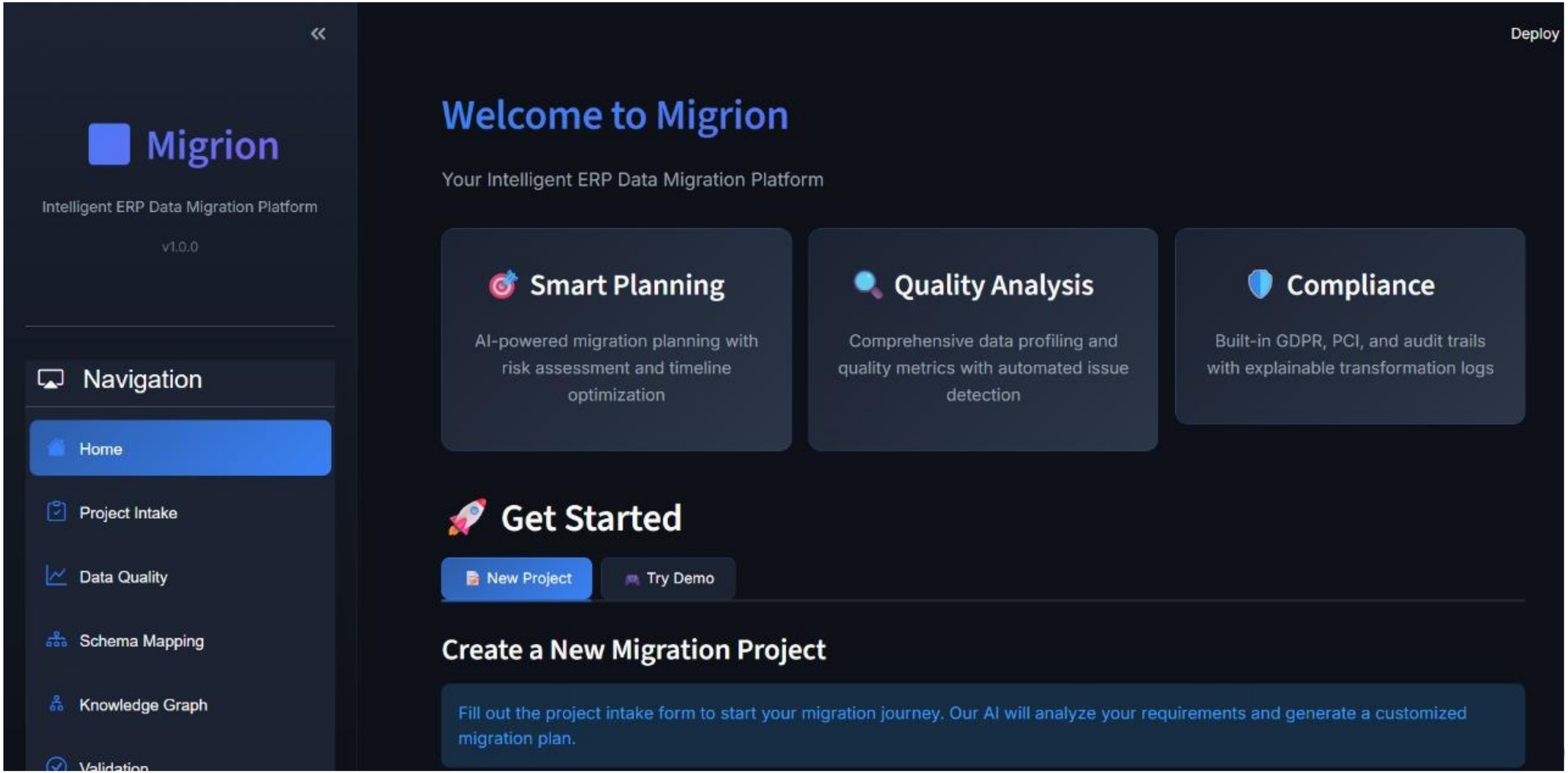
- Database licensing: ₹54,00,000 ($64,800)
- Hardware maintenance: ₹9,00,000 ($10,800)
- Storage and backup: ₹5,40,000 ($6,480)
- Support contracts: ₹3,60,000 ($4,320)
- **Total Annual Cost:** ₹72,00,000 ($86,400)

Target AWS Environment (Annual):

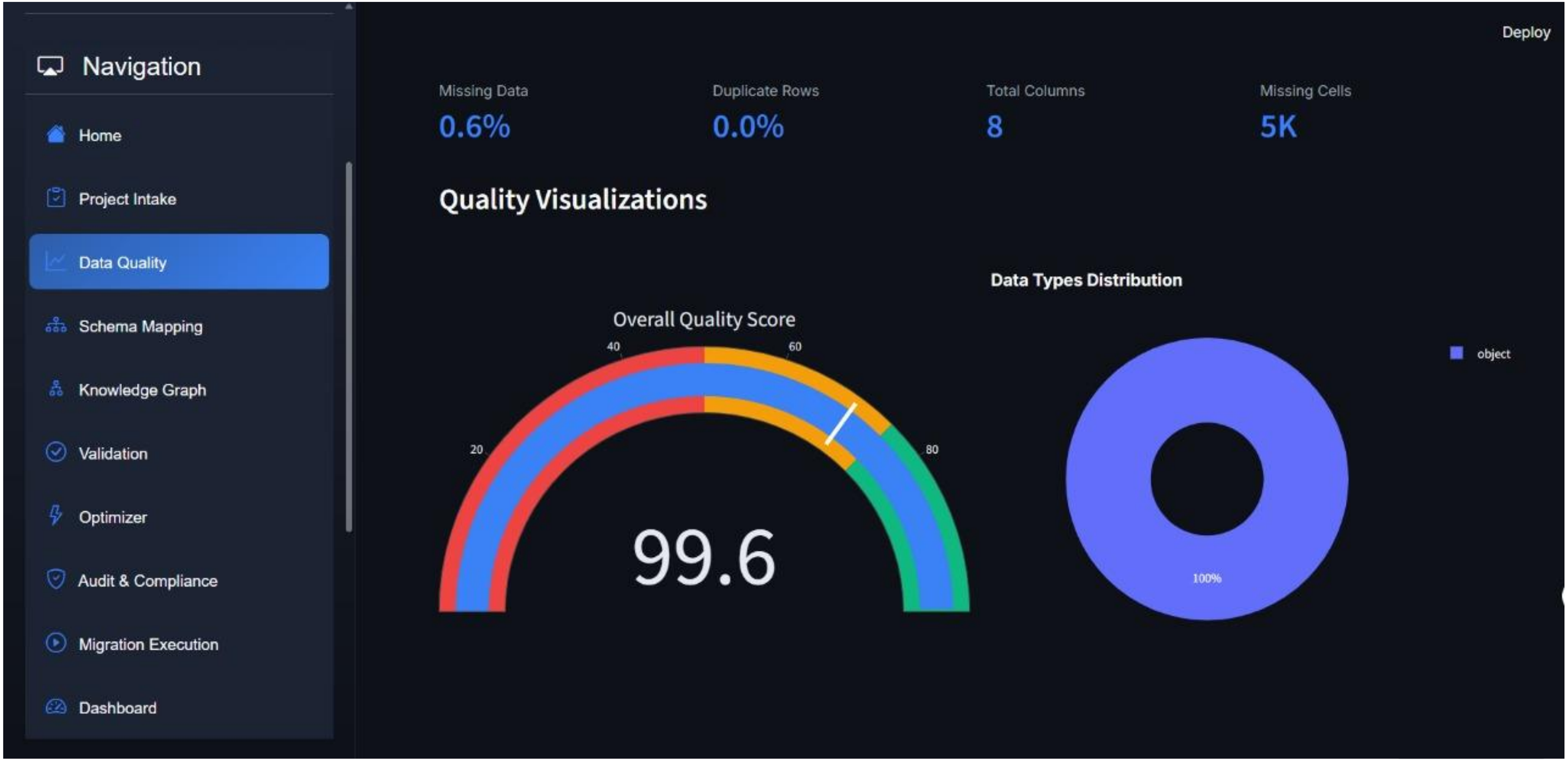
- RDS PostgreSQL: ₹22,20,000 ($26,640)
- Storage (3TB GP3): ₹4,20,000 ($5,040)
- Backup storage: ₹3,00,000 ($3,600)
- Data transfer: ₹1,80,000 ($2,160)
- Monitoring: ₹96,000 ($1,152)
- **Total Annual Cost:** ₹32,16,000 ($38,592)

Annual Savings: ₹39,84,000 ($47,808)
Cost Reduction: 55.3%

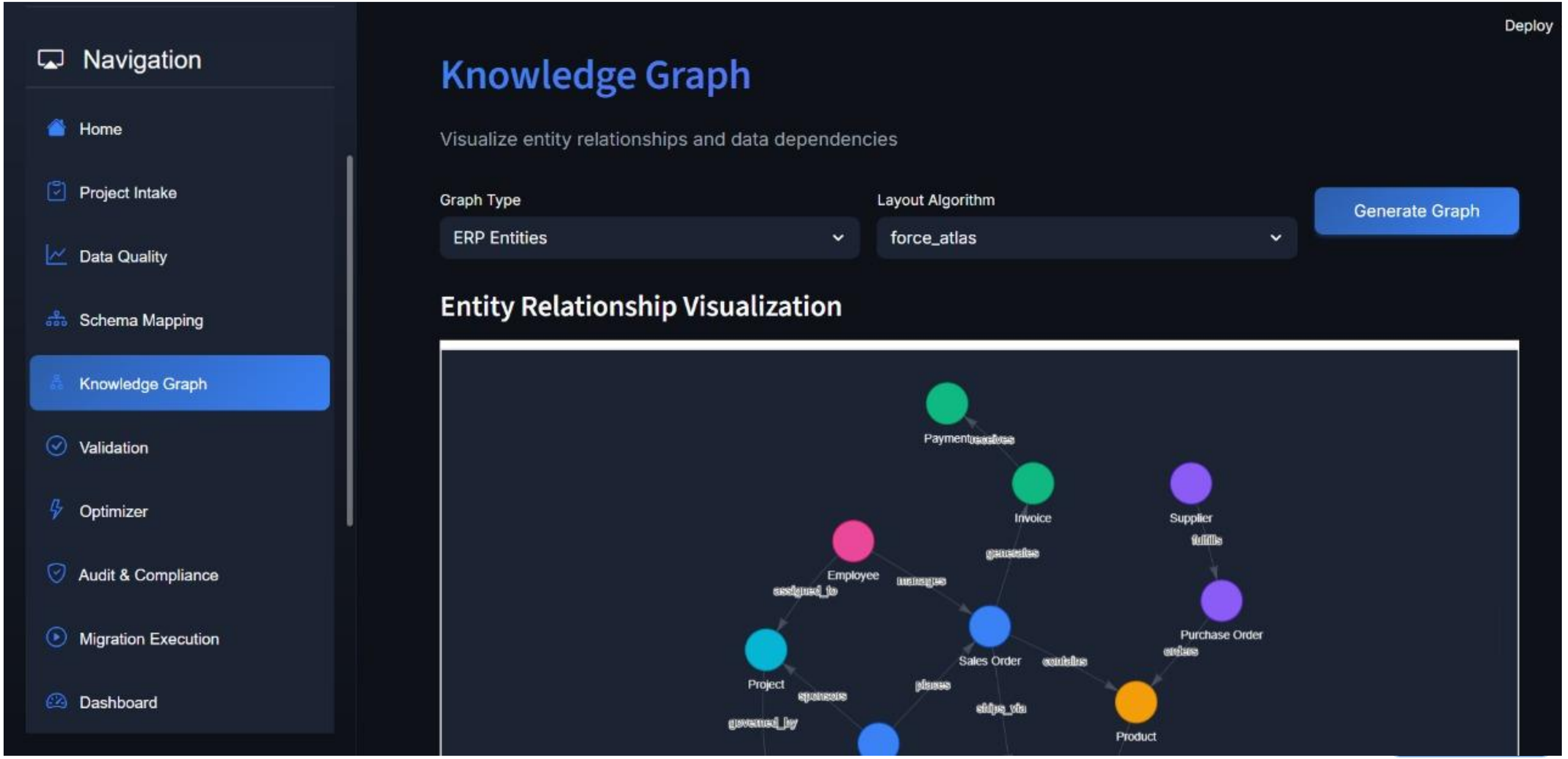
Web Application Snapshots



Web Application Snapshots



Web Application Snapshots



NOTIFICATION of Acceptance

G

EDAS Conference Manager<help@edas.info> on behalf of gstomar=ieee.org@edas.info

To: KUPPALA NANDINI - [CB.EN.U4AIE22030]; **+4 others**

Tue 10/14/2025 9:42 PM

Dear Ms. Nandini Kuppala: Congratulations - On behalf of the organizing committee; I am happy to inform you that your paper #1571209932 ('Multi-Agent Multi-Objective Reinforcement Learning for Intelligent Cloud Migration Planning: A Comprehensive Framework with Explainable Decision Support') submitted by you has been **accepted** with some minor changes for oral Presentation in the conference "17th IEEE International Conference on Computational Intelligence and Communication Networks 2025" on the basis of the recommendations given by the peer review group of experts. The Conference proceedings of previous years are published by IEEE Xplore and are online, the proceedings of this conference will be submitted to IEEE Xplore for publication. The proceedings will be indexed by EI Compendex, SCOPUS, ISI Proceeding etc One of the author (s) must register by paying required registration fee till 18 Oct 2025. Unregistered paper will not be considered as accepted and will not be included in the proceedings. You are requested to visit conference website http://www.cicn.in registration page for further instructions for registration and use payment link for payment of fee. You are advised to incorporate reviewers' comments in your final manuscript to be submitted on conference email cicnmir@gmail.com in word format following ieee format with proof of payment. The conference will be held on Dec 20-21, 2025 at National Institute of Technology, Goa. Should you have any questions, please feel free to inform us through email.

Regards TPC chairs The reviews are below or can be found at 1571209932.

Thank You