FUTURE OF INTELLIGENT CLOUD COMPUTING

- Real-time streaming of object access
- ML-powered tier classification (HOT / WARM / COLD)
- Dynamic storage optimization for cloud cost-saving

Team Leader- SHAURYA SHARMA
EMAIL- shaurya22006sharma@gmail.com

Impact

- High cloud bills
- Poor storage efficiency
- No real-time intelligence
  Need: Automated, intelligent, real-time tiering

Need
Automated, intelligent, real-time tiering

Storage Problem

- Enterprises store millions of cloud objects
- 70–80% become rarely accessed but stay on expensive HOT storage
- Manual tiering is slow, error-prone, and not scalable

Smart Cloud Storage Optimizer

- Streams live access logs from objects
- Predicts best storage tier (HOT / WARM / COLD) using ML
- Shows real-time dashboard + cost and migration logs

Streaming Logs → FastAPI Backend → ML Model (joblib)

↓

SQLite storage

↓

Streamlit Dashboard

How It Works

Data Pipeline

- Generator simulates access events
- SQLite stores object history
- Backend aggregates:
  - `size_gb`
  - `reads_last_7d`
  - `recency_days`
  - `latency_requirement_ms`
  - `cost_per_gb`

Machine Learning

- Random Forest Classifier
- Predicts optimal tier per object
- Activity logs show movement (HOT → COLD, etc.)

Demo Highlights

✅ Real-time streaming (data refreshes as generator runs)
✅ Tier distribution updates live
✅ Activity Log mimics actual cloud migrations
✅ Search any object by ID
✅ Works now, no cloud dependency
(Can plug into AWS S3, Azure Blob, or NetApp StorageGrid later)

Why only ~1 HOT object?

- HOT requires high reads + low recency + high performance sensitivity
- Most objects quickly become WARM , realistic enterprise pattern

Results & Roadmap

Current Output (example)

- HOT: 1
- WARM: majority
- COLD: archival data
- Total objects update as logs stream

Value

Removes manual storage decisions
Reduces monthly storage cost
Makes tiering intelligent & autonomous

Future Upgrade

- Add real AWS/Azure migration APIs
- Kafka streaming
- Online ML retraining

CONCLUSION:
A lightweight, scalable intelligent storage tiering system that saves cost and automates cloud data management.