# Proposal: Backup và Archive Migration với Tiered Storage

## 1. 📄 Executive Summary

### Problem Statement

In the era of data explosion, organizations face increasing costs for storing large volumes of unstructured data, especially backups and archives. Traditional backup systems are often static, expensive, and lack automation in data retrieval and compliance enforcement.

### Solution Overview

This proposal introduces a cloud-native backup system leveraging AWS S3 Intelligent-Tiering to automatically transition backup data based on access frequency. The system integrates Object Lock for compliance, Glacier tiers for low-cost archival, and monitoring tools for retrieval automation and cost control.

### Business Benefits & ROI

Up to 70% storage cost savings via automatic tiering.

Regulatory compliance with Object Lock and retention policies.

Improved operational efficiency with automation and monitoring.

Scalable and secure data storage without infrastructure overhead.

### Investment & Timeline

Estimated cloud infrastructure budget: $150/month.

Development time: 6 weeks.

Break-even point expected within 5 months.

### Success Metrics

60%+ data moved to cold storage tiers within 90 days.

99.99% availability and durability.

Zero incidents of unauthorized deletions.

30% reduction in backup-related operational effort.

### Evaluation Criteria

**Clarity and conciseness**: Clearly summarizes the proposal's core value without unnecessary complexity.

**Compelling business case**: Demonstrates a clear return on investment and operational impact.

**Accurate summary of main points**: Covers all critical aspects from problem to expected outcomes.

**Executive-level language**: Uses professional and concise language suitable for leadership audiences.

## 2. Problem Statement

### Current Situation

Most on-premise or legacy cloud backup systems store all data at the same tier regardless of access patterns, leading to high costs. Additionally, ensuring data immutability and compliance involves complex manual policies.

### Pain Points

Rising storage costs (especially for infrequently accessed data).

Manual management of backup lifecycles.

Compliance risks from accidental deletions.

Lack of centralized cost monitoring and alerts.

### Stakeholders Affected

IT Operations Team: burdened by manual lifecycle management.

Compliance Officers: need proof of data integrity.

Finance: affected by unpredictable storage costs.

### Business Consequences

Without optimization, backup costs could rise 35-50% YoY. A breach in compliance or loss of critical backups can lead to financial penalties or data recovery failure.

### Market Opportunity

With AWS offering built-in lifecycle automation, this system can serve SMEs and regulated industries (e.g., fintech, healthcare) seeking secure, cost-optimized backups.

## 3. Solution Architecture

### High-Level Architecture Diagram

(Included in Appendix A)

### AWS Services Selected

**Amazon S3 with Intelligent-Tiering**: Automated storage class transitions.

**S3 Object Lock**: Enforces retention for compliance.

**S3 Glacier & Glacier Deep Archive**: Cost-effective cold storage.

**AWS CloudWatch**: Logs and monitors access patterns.

**AWS Cost Explorer**: Tracks and forecasts storage spending.

### Component Interactions

Uploaded data goes to S3 > Monitored by CloudWatch > Access frequency triggers class transitions > Object Lock protects data > Glacier retrieval initiated if cold files are accessed.

### Security & Compliance

IAM Policies: Restrict delete/modify actions.

Object Lock: Enables WORM (Write Once Read Many).

Encryption: S3 Default SSE enabled (AES-256).

Bucket Policies: Allow access only from specific roles/accounts.

### Scalability & Performance

Automatically scales with storage demand.

Retrieval from Glacier optimized for batch requests.

### Integration Points

Existing on-premise systems can use AWS CLI or SDKs for integration.

Third-party backup tools (e.g., Veeam, Commvault) can push backups to S3.

## 4. Technical Implementation

### Implementation Phases

**Setup AWS Environment**: IAM roles, billing alerts, bucket policies.

**Create S3 Buckets**: Enable Intelligent-Tiering, Object Lock.

**Upload Backups**: Import historical backups or schedule uploads.

**Access Simulation**: Generate realistic traffic to validate tiering.

**Monitoring Setup**: CloudWatch, S3 Analytics, Cost Explorer.

**Testing**: Perform retrieval from Glacier, delete protection check.

**Optimization**: Adjust retention periods, enable Deep Archive.

### Technical Requirements

Compute: AWS Lambda (for automation scripts, optional).

Storage: S3 Intelligent-Tiering with ~1 TB monthly usage.

Network: Secure endpoints via VPC or Gateway Endpoints.

### Development Methodologies

Infrastructure-as-Code: Use CloudFormation or Terraform.

CI/CD (Optional): GitHub Actions for automation deployments.

### Testing Strategy

Unit: Validate script logic.

Integration: Ensure tier transition works.

Performance: Simulate access spikes.

Compliance: Try unauthorized deletions (should fail).

### Deployment & Rollback

Phased rollout per bucket/project.

Restore previous backup via S3 Versioning or Glacier Restore.

### Configuration Management

Use AWS Config for compliance rules.

Maintain version-controlled config files in repository.

## 5. Timeline & Milestones

| **Phase** | **Duration** | **Milestone** |
| --- | --- | --- |
| Setup & Planning | Week 1 | AWS accounts, IAM, bucket strategy defined |
| Development | Week 2-3 | S3 buckets, Object Lock, scripts ready |
| Testing | Week 4 | Tier transition validation, data restore tests |
| Deployment | Week 5 | Migration of real backup data |
| Optimization | Week 6 | Final tuning, reporting, documentation |

Dependencies: Account provisioning, billing permissions.  
Buffer Time: +1 week for AWS approval delays or script bugs.

## 6. Budget Estimation

### Monthly AWS Costs

S3 Intelligent-Tiering (1 TB): ~$20

Glacier & Deep Archive storage: ~$4 (est. 200 GB cold)

PUT/GET requests: ~$10

CloudWatch Metrics: ~$8

Cost Explorer: Free

### One-time Development Costs

Engineer Time: 40 hours @ $30/hr = $1,200

### Total 1st Year Cost

Monthly Infra: ~$42 × 12 = $504

One-time Dev: $1,200

**Total: $1,704**

### ROI & Optimization

Previous static S3 storage: ~$100/month × 12 = $1,200

Estimated savings = $696/year (58%)

Break-even in 4-5 months

Use lifecycle policies & monitoring to further reduce costs.

## 7. Risk Assessment

| **Risk** | **Impact** | **Likelihood** | **Mitigation** |
| --- | --- | --- | --- |
| Glacier retrieval delays | Medium | High | Use Instant Retrieval tier where needed |
| Misconfiguration of Object Lock | High | Medium | Automate via IaC templates |
| Cost spike due to misclassified access | Medium | Medium | Use S3 Analytics to review tiering accuracy |
| Permission mismanagement | High | Low | Strict IAM policies + audit logging |

Monitoring: Weekly checks via CloudWatch, AWS Config.  
Escalation: Alert stakeholders via SNS + email for anomalies.

## 8. Expected Outcomes

### Success Metrics

60% data moved to cold tiers within 3 months

30% less time spent by Ops team managing backups

100% compliance for retention policy

### Benefits

**Short-term (0–6 mo)**:

Cost savings visible in billing dashboard

Proof of compliance for audit

**Mid-term (6–18 mo)**:

Integrated into other teams’ backup flows

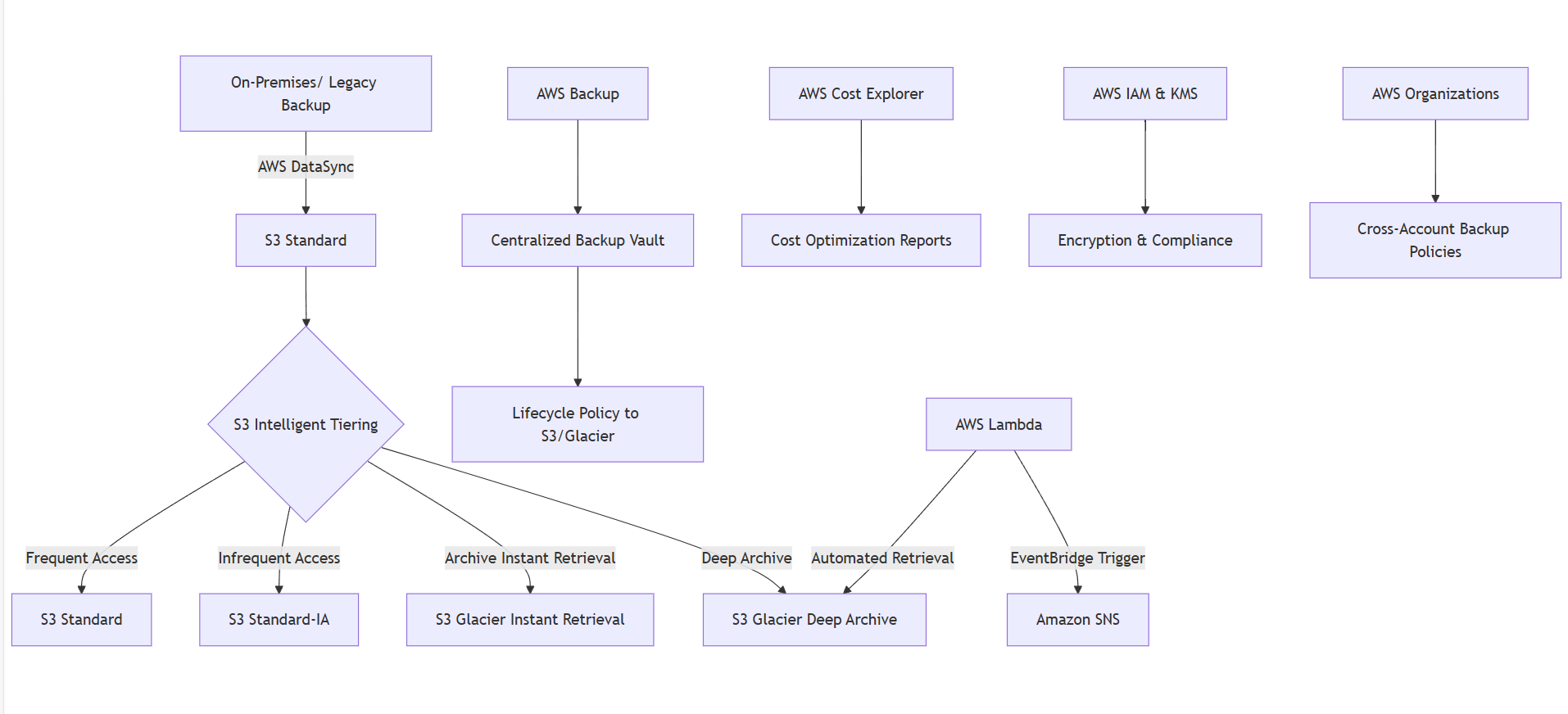
Redundancy added (e.g., replication)

**Long-term (>18 mo)**:

Scalable foundation for full backup platform

Strategic compliance asset for audits/investors

Long-term cost predictability and savings



### ****Amazon S3 Intelligent-Tiering****

[https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-class-intelligent-tiering.html](https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-class-intelligent-tiering.html" \t "_new)

### ****Object Lock và Compliance****

[https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lock-overview.html](https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lock-overview.html" \t "_new)

### ****AWS S3 Storage Lens (Monitoring)****

[https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-lens.html](https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-lens.html" \t "_new)

### ****Cost Optimization Best Practices****

AWS Well-Architected – Storage Lens: [https://docs.aws.amazon.com/wellarchitected/latest/storage-lens/storage-cost-optimization.html](https://docs.aws.amazon.com/wellarchitected/latest/storage-lens/storage-cost-optimization.html" \t "_new)

### ****Khôi phục dữ liệu từ Glacier****

[https://docs.aws.amazon.com/AmazonS3/latest/userguide/restoring-objects.html](https://docs.aws.amazon.com/AmazonS3/latest/userguide/restoring-objects.html" \t "_new)

AWS CLI S3 Commands: [https://docs.aws.amazon.com/cli/latest/reference/s3/index.html](https://docs.aws.amazon.com/cli/latest/reference/s3/index.html" \t "_new)

AWS Blog – Optimizing Storage with S3 Intelligent-Tiering:  
[https://aws.amazon.com/blogs/storage/optimize-costs-with-s3-intelligent-tiering/](https://aws.amazon.com/blogs/storage/optimize-costs-with-s3-intelligent-tiering/" \t "_new)