**Load-Balanced EC2 Linux SFTP Cluster - Migration Proposal**

**Executive Summary**

This proposal presents a comprehensive solution for implementing a highly available SFTP infrastructure using load-balanced EC2 Linux instances with shared EFS storage, integrated with your existing Windows Active Directory through LDAP authentication. This approach provides complete control over the SFTP infrastructure while addressing high availability and scalability requirements.

**Key Benefits:**

* **Complete infrastructure control** with customizable SFTP server configuration
* **High availability design** with multi-AZ deployment and automated failover
* **Flexible group processing** through custom LDAP integration logic
* **Scalable architecture** with Auto Scaling Groups and load balancing
* **Full operational visibility** with custom monitoring and alerting

**Consideration:** This solution requires significant operational expertise and ongoing maintenance overhead compared to managed service alternatives, but provides maximum flexibility and control.

**Table of Contents**

1. [Current State Analysis](#current-state-analysis)
2. [Proposed Solution Architecture](#proposed-solution-architecture)
3. [Technical Requirements](#technical-requirements)
4. [Implementation Plan](#implementation-plan)
5. [Resource Requirements](#resource-requirements)
6. [Cost Analysis](#cost-analysis)
7. [Risk Assessment](#risk-assessment)
8. [Success Metrics](#success-metrics)
9. [Next Steps](#next-steps)

**Current State Analysis**

**Current Environment**

* **SFTP Service**: AWS Transfer Family with native Active Directory integration
* **Identity Provider**: Windows EC2-based Microsoft Active Directory
* **Storage Backend**: Amazon S3 (single backend)
* **User Base**: 200+ users with anticipated growth to 500+ users
* **Network**: VPC-based infrastructure with connectivity to existing AD

**Limitations Addressed by EC2 Approach**

**Single AD Group Restriction**

* **Current Impact**: Users can only belong to one AD group in Transfer Family
* **EC2 Solution**: Custom LDAP integration can process multiple group memberships
* **Implementation**: PAM (Pluggable Authentication Modules) with custom group processing logic
* **Benefit**: Flexible group membership handling through custom authentication scripts

**Infrastructure Control Requirements**

* **Current Limitation**: Limited configuration options with AWS Transfer Family
* **Business Need**: Custom SFTP configurations, specific security requirements, specialized logging
* **EC2 Solution**: Complete control over OpenSSH configuration and system-level customizations
* **Flexibility**: Custom authentication hooks, specialized file processing, integration with existing tools

**High Availability Design**

* **Current State**: Single managed endpoint with AWS-controlled availability
* **Business Requirement**: Visible high availability architecture with controlled failover
* **EC2 Solution**: Multi-AZ deployment with load balancing and automated instance replacement
* **Control**: Full visibility into failover procedures and infrastructure health

**Solution Requirements**

* **High Availability**: Multi-AZ deployment with automatic failover and recovery
* **Scalability**: Automatic scaling based on load and connection requirements
* **LDAP Integration**: Direct integration with Windows AD supporting multiple group memberships
* **Shared Storage**: Consistent user experience across multiple SFTP server instances
* **Custom Configuration**: Flexible SFTP server configuration for specialized requirements

**Proposed Solution Architecture**

**Architecture Overview**

The Load-Balanced EC2 Linux SFTP Cluster provides a highly available SFTP service using multiple Linux instances behind a Network Load Balancer, with shared EFS storage for user data and direct LDAP integration with your existing Windows Active Directory.

**Core Components**

**Network Load Balancer**

* **TCP Load Balancing**: Distributes SFTP connections (port 22) across healthy instances
* **Health Monitoring**: Application-level health checks for SFTP service availability
* **Cross-Zone Balancing**: Even distribution of traffic across multiple Availability Zones
* **Session Affinity**: Optional sticky sessions for consistent user experience during file transfers

**EC2 Auto Scaling Group**

* **Multi-AZ Deployment**: Instances distributed across 3 Availability Zones for fault tolerance
* **Automatic Scaling**: CPU and connection-based scaling policies for demand management
* **Instance Replacement**: Automatic replacement of unhealthy instances with health checks
* **Launch Template**: Standardized instance configuration with automated setup scripts

**OpenSSH SFTP Configuration**

* **Ubuntu/CentOS**: Hardened Linux instances with latest security patches
* **OpenSSH Server**: Configured for SFTP-only access with security best practices
* **PAM Integration**: Pluggable Authentication Modules for Windows AD authentication
* **Custom Authentication**: Scripts for processing multiple AD group memberships

**Shared Storage Infrastructure**

* **Amazon EFS**: POSIX-compliant file system mounted across all SFTP instances
* **User Home Directories**: Consistent user experience regardless of instance connection
* **Performance Mode**: General Purpose or Max I/O based on concurrent user requirements
* **Backup Integration**: Automatic backups with point-in-time recovery capabilities

**Integration Architecture**

**Windows AD LDAP Integration**

* **Direct LDAP**: Native LDAP protocol connectivity to existing domain controllers
* **Service Account**: Dedicated AD account for authentication queries and group lookups
* **PAM Configuration**: System-level integration for seamless user authentication
* **Group Processing**: Custom scripts to handle multiple group memberships and permissions

**High Availability Design**

* **Active-Active Configuration**: Multiple active instances handling concurrent connections
* **Health Monitoring**: Comprehensive health checks at network, system, and application levels
* **Automatic Failover**: Load balancer removes unhealthy instances from rotation automatically
* **Data Consistency**: Shared EFS ensures consistent user data across all instances

**Network Architecture**

**Load Balancer Configuration**

* **Network Load Balancer**: Layer 4 TCP load balancing for optimal SFTP performance
* **Target Groups**: Health check configuration with custom SFTP availability validation
* **Listener Configuration**: Port 22 listener with SSL passthrough for SSH encryption
* **DNS Integration**: Route 53 alias record for consistent SFTP endpoint

**Security Group Configuration**

* **SFTP Access**: Inbound port 22 from authorized IP ranges and security groups
* **LDAP Connectivity**: Outbound ports 389/636 to Windows AD domain controllers
* **EFS Mounting**: NFS traffic (port 2049) between instances and EFS mount targets
* **Management Access**: SSH access from administrative networks for maintenance

**Data Flow Diagram**

[SFTP Users] → [Network Load Balancer] → [EC2 SFTP Instances (Multi-AZ)]  
 ↓  
[EFS Shared Storage] ← [Auto Scaling Group] ← [Windows AD via LDAP]  
 ↓  
[CloudWatch Monitoring] ← [Health Checks] ← [Application Logs]

**Technical Requirements**

**Prerequisites**

**Linux Infrastructure Expertise**

* **System Administration**: Ubuntu/CentOS administration with OpenSSH configuration
* **LDAP Integration**: Experience with PAM configuration and Windows AD connectivity
* **Load Balancing**: Network Load Balancer configuration and troubleshooting
* **Storage Management**: EFS mounting, permissions, and performance optimization

**AWS Infrastructure Knowledge**

* **Auto Scaling**: Auto Scaling Groups with launch templates and scaling policies
* **Networking**: VPC configuration, security groups, and multi-AZ deployment
* **Monitoring**: CloudWatch configuration for infrastructure and application monitoring
* **Load Balancers**: Network Load Balancer health checks and target group management

**Windows AD Environment**

* **Service Account**: Dedicated account with read permissions for user and group queries
* **LDAP Access**: Network connectivity on ports 389/636 from EC2 instances to domain controllers
* **Group Structure**: Understanding of AD group hierarchy for custom processing logic
* **Security Permissions**: Appropriate permissions for LDAP bind and user authentication

**Infrastructure Specifications**

**EC2 Instance Configuration**

* **Instance Type**: t3.medium (2 vCPU, 4GB RAM) minimum, scalable to t3.large based on load
* **Operating System**: Ubuntu 22.04 LTS or CentOS Stream 8 with security hardening
* **Storage**: 20GB GP3 root volume with optimized IOPS for system operations
* **Network**: Enhanced networking enabled for improved performance and lower latency

**Auto Scaling Group Design**

* **Minimum Capacity**: 3 instances across 3 Availability Zones for high availability
* **Maximum Capacity**: 6 instances to handle peak load and traffic spikes
* **Scaling Policies**: CPU utilization (>70%) and network connections (>80% capacity)
* **Health Checks**: EC2 and ELB health checks with 300-second grace period

**Load Balancer Specifications**

* **Type**: Network Load Balancer with cross-zone load balancing enabled
* **Scheme**: Internet-facing with public IP addresses for external SFTP access
* **Listeners**: TCP port 22 with target group health check configuration
* **Health Check**: Custom TCP health check on port 22 with 3 healthy/unhealthy thresholds

**Storage Configuration**

* **EFS File System**: General Purpose performance mode with encryption in transit
* **Mount Targets**: One per Availability Zone with security group configuration
* **Access Points**: User home directory structure with POSIX permissions
* **Backup**: Daily backups with 30-day retention and point-in-time recovery

**Software Components**

**OpenSSH Configuration**

* **SFTP-Only Access**: Restricted shell configuration preventing shell access
* **Authentication Methods**: Password and public key authentication support
* **Logging**: Comprehensive audit logging for all file transfer activities
* **Security Hardening**: Disabled root login, key-based authentication preferences

**PAM Integration**

* **LDAP Module**: pam\_ldap configuration for Windows AD authentication
* **Group Processing**: Custom PAM modules for multiple group membership handling
* **Session Management**: pam\_mkhomedir for automatic home directory creation
* **Security Policies**: Account lockout and password policies synchronized with AD

**Monitoring and Logging**

* **CloudWatch Agent**: System and application metrics collection
* **Log Aggregation**: Centralized logging for authentication and file transfer activities
* **Custom Metrics**: SFTP connection counts, authentication success rates, file transfer volumes
* **Alerting**: Automated alerts for system health, performance, and security events

**Implementation Plan**

**Phase 1: Infrastructure Setup (4 weeks)**

**Objectives**

Deploy core AWS infrastructure components for highly available SFTP cluster with proper networking, security, and monitoring foundations.

**Key Activities**

**Weeks 1-2: VPC and Networking**

* Design and deploy VPC with private and public subnets across 3 Availability Zones
* Configure security groups for SFTP, LDAP, EFS, and management access
* Set up Network Load Balancer with target groups and health check configuration
* Implement Route 53 DNS configuration for SFTP endpoint resolution

**Weeks 3-4: Storage and Scaling Infrastructure**

* Deploy EFS file system with mount targets and access points configuration
* Create Auto Scaling Group with launch template and scaling policies
* Configure CloudWatch monitoring and alerting for infrastructure components
* Set up AWS Systems Manager for configuration management and patching

**Deliverables**

* **VPC Infrastructure**: Complete networking setup with security group configuration
* **Network Load Balancer**: Operational load balancer with health check validation
* **EFS Configuration**: Shared file system accessible from all Availability Zones
* **Auto Scaling Setup**: Launch template and scaling policies configured and tested
* **Monitoring Framework**: Basic infrastructure monitoring and alerting established

**Resources Required**

* **Solutions Architect**: Full-time for infrastructure design and deployment
* **Network Engineer**: Full-time for load balancer and VPC configuration
* **Linux Administrator**: Part-time for EFS and system configuration planning
* **DevOps Engineer**: Part-time for automation and monitoring setup

**Success Criteria**

* VPC and networking components deployed and tested across all Availability Zones
* Network Load Balancer operational with successful health checks
* EFS file system mounted and accessible with proper permissions
* Auto Scaling Group creating and terminating instances successfully
* Basic monitoring and alerting operational with baseline metrics

**Phase 2: SFTP Configuration (3 weeks)**

**Objectives**

Configure OpenSSH SFTP services with Windows AD integration, user authentication, and shared storage access.

**Key Activities**

**Week 1: OpenSSH and System Configuration**

* Install and configure OpenSSH server on EC2 instances with SFTP-only access
* Implement security hardening and configuration best practices
* Configure EFS mounting with proper permissions and performance optimization
* Set up automated instance configuration with launch template user data scripts

**Weeks 2-3: LDAP Integration and Authentication**

* Configure PAM for LDAP authentication against Windows Active Directory
* Implement custom group processing logic for multiple AD group memberships
* Set up user home directory creation and permission management
* Configure comprehensive logging and audit trail for all SFTP activities

**Deliverables**

* **OpenSSH Configuration**: SFTP-only access configured on all instances
* **LDAP Authentication**: Working authentication against Windows AD with group processing
* **User Management**: Automated home directory creation and permission assignment
* **Security Configuration**: Hardened system configuration with comprehensive audit logging
* **Documentation**: Configuration guides and operational procedures

**Resources Required**

* **Linux Administrator (2)**: Full-time for OpenSSH configuration and LDAP integration
* **Solutions Architect**: Part-time for technical guidance and architecture validation
* **Windows Administrator**: Support for AD service account and LDAP configuration testing
* **Security Specialist**: Part-time for security hardening and configuration review

**Success Criteria**

* OpenSSH SFTP service operational on all instances with proper security configuration
* Users successfully authenticating against Windows AD with multiple group processing
* Shared EFS storage accessible with consistent user experience across instances
* Comprehensive audit logging operational with centralized log aggregation
* Security hardening completed and validated through penetration testing

**Phase 3: High Availability Setup (3 weeks)**

**Objectives**

Implement and test comprehensive high availability, monitoring, and automated scaling capabilities.

**Key Activities**

**Week 1: Health Monitoring and Failover**

* Configure advanced health checks for SFTP service availability and performance
* Implement automated failover procedures and instance replacement testing
* Set up CloudWatch custom metrics for SFTP connections and performance monitoring
* Configure automated alerting for health status and performance degradation

**Weeks 2-3: Scaling and Performance Optimization**

* Test Auto Scaling policies under various load conditions and connection scenarios
* Optimize EFS performance settings and connection handling for concurrent users
* Implement connection pooling and resource management for optimal performance
* Configure automated scaling based on custom metrics and business requirements

**Deliverables**

* **High Availability Validation**: Completed failover testing with documented procedures
* **Monitoring Dashboards**: Comprehensive CloudWatch dashboards for all components
* **Auto Scaling Configuration**: Verified scaling policies responding to load conditions
* **Performance Optimization**: Tuned system configuration for optimal concurrent user support
* **Operational Procedures**: Documented procedures for monitoring and maintenance

**Resources Required**

* **Network Engineer**: Full-time for load balancer optimization and failover testing
* **Linux Administrator**: Full-time for performance tuning and scaling configuration
* **DevOps Engineer**: Part-time for monitoring automation and alerting setup
* **QA Engineer**: Part-time for load testing and failover validation

**Success Criteria**

* Automated failover successfully tested with <5 minute recovery time
* Auto Scaling policies responding appropriately to load conditions
* Performance optimization supporting target concurrent user capacity
* Comprehensive monitoring operational with appropriate alerting thresholds
* High availability architecture validated through disaster recovery testing

**Phase 4: Testing & Validation (2 weeks)**

**Objectives**

Conduct comprehensive end-to-end testing including load testing, security validation, and disaster recovery procedures.

**Key Activities**

**Week 1: Load and Performance Testing**

* Execute load testing with concurrent SFTP connections matching expected peak usage
* Perform security testing including penetration testing and vulnerability assessment
* Validate disaster recovery procedures including instance failure and AZ outage scenarios
* Test backup and recovery procedures for EFS and configuration data

**Week 2: User Acceptance and Documentation**

* Conduct user acceptance testing with representative users from different organizational groups
* Validate multi-group authentication and permission inheritance functionality
* Complete operational documentation including troubleshooting guides and procedures
* Provide training to operations and support teams on new infrastructure

**Deliverables**

* **Load Testing Results**: Comprehensive performance validation with capacity recommendations
* **Security Assessment**: Penetration testing results with vulnerability remediation
* **Disaster Recovery Validation**: Tested recovery procedures with documented timelines
* **User Acceptance**: Validated functionality with user sign-off and feedback
* **Operational Documentation**: Complete guides for monitoring, maintenance, and troubleshooting

**Resources Required**

* **QA Engineer**: Full-time for testing execution and results documentation
* **Security Specialist**: Part-time for security validation and penetration testing
* **Linux Administrator**: Part-time for issue resolution and optimization
* **Business Users**: Pilot group participation for user acceptance testing

**Success Criteria**

* Load testing demonstrates capacity to handle target concurrent user load
* Security assessment shows no critical vulnerabilities with remediation plan
* Disaster recovery procedures validated with acceptable recovery time objectives
* User acceptance criteria met with positive feedback from test groups
* Operational documentation completed and training provided to support teams

**Phase 5: Go-Live & Support (2 weeks)**

**Objectives**

Execute production deployment with comprehensive operational support and monitoring establishment.

**Key Activities**

**Week 1: Production Cutover**

* Update DNS configurations to point SFTP traffic to new load balancer endpoint
* Execute production migration with user validation and rollback procedures ready
* Implement 24/7 monitoring and alerting with escalation procedures
* Establish operational support procedures and team assignments

**Week 2: Optimization and Support**

* Monitor production performance and optimize configuration based on actual usage patterns
* Provide intensive support during initial production period with rapid issue resolution
* Conduct post-implementation review with lessons learned and optimization recommendations
* Establish ongoing maintenance procedures and capacity planning processes

**Deliverables**

* **Production Environment**: Fully operational SFTP cluster serving all users
* **24/7 Monitoring**: Comprehensive monitoring and alerting with support procedures
* **Performance Baseline**: Established production performance metrics and trends
* **Post-Implementation Review**: Lessons learned and optimization roadmap
* **Operational Procedures**: Ongoing maintenance and support documentation

**Resources Required**

* **Operations Team**: Full-time for production cutover and initial support
* **Linux Administrator**: Full-time for performance optimization and issue resolution
* **Network Engineer**: Part-time for traffic monitoring and load balancer optimization
* **Support Team**: Part-time for user assistance and incident management

**Success Criteria**

* Zero-downtime production cutover with successful user validation
* 24/7 monitoring operational with appropriate response procedures
* Performance metrics meeting or exceeding target specifications
* Support team trained and ready for ongoing operational responsibilities
* Post-implementation review completed with stakeholder approval

**Resource Requirements**

**Human Resources**

**Solutions Architect**

* **AWS Certifications**: Solutions Architect Professional with EC2, EFS, and Load Balancer expertise
* **Experience**: 5+ years designing highly available AWS infrastructure
* **Skills**: Multi-AZ deployment, Auto Scaling, Network Load Balancer, VPC design
* **Commitment**: 25 hours/week average across infrastructure phases
* **Role**: Technical leadership, architecture decisions, and infrastructure design

**Linux Administrator (2 Required)**

* **Technical Skills**: Ubuntu/CentOS administration, OpenSSH configuration, PAM/LDAP integration
* **AWS Experience**: EC2 management, EFS configuration, CloudWatch monitoring
* **Security Knowledge**: System hardening, audit configuration, security compliance
* **Commitment**: 40 hours/week each during configuration and deployment phases
* **Role**: SFTP server configuration, AD integration, and operational procedures

**Network Engineer**

* **Networking Expertise**: Load balancer configuration, VPC design, DNS management
* **AWS Experience**: Network Load Balancer, Auto Scaling, CloudWatch monitoring
* **High Availability**: Failover design, health check configuration, disaster recovery
* **Commitment**: 35 hours/week during infrastructure and HA configuration phases
* **Role**: Load balancer setup, network configuration, and failover testing

**QA Engineer**

* **Testing Expertise**: Load testing, security testing, infrastructure validation
* **Tools Experience**: Performance testing tools, security scanners, monitoring validation
* **Documentation Skills**: Test plan development and results analysis
* **Commitment**: 40 hours/week during testing and validation phases
* **Role**: Comprehensive testing execution and validation documentation

**Technical Resources**

**AWS Services**

* **Amazon EC2**: Linux instances with Auto Scaling Groups and launch templates
* **Network Load Balancer**: TCP load balancing with health check configuration
* **Amazon EFS**: Shared POSIX file system with backup and encryption
* **Amazon CloudWatch**: Comprehensive monitoring, logging, and alerting
* **AWS Systems Manager**: Configuration management and automated patching

**Infrastructure Management Tools**

* **Ansible/Chef**: Configuration management and automation for consistent deployment
* **Terraform**: Infrastructure as Code for repeatable deployments
* **Git Repository**: Version control for configuration scripts and documentation
* **Monitoring Tools**: Custom monitoring scripts and alerting integration

**Testing and Validation Tools**

* **Load Testing**: Apache JMeter or similar for concurrent connection testing
* **Security Testing**: Nessus, OpenVAS for vulnerability assessment
* **Network Analysis**: Wireshark, tcpdump for network troubleshooting
* **Performance Monitoring**: iostat, htop, sar for system performance analysis

**Cost Analysis**

**Implementation Costs (One-time)**

**Professional Services**

* **Solutions Architect**: $30,000 (200 hours @ $150/hour)
* **Linux Administrator (2)**: $28,800 (360 hours @ $80/hour combined)
* **Network Engineer**: $21,000 (175 hours @ $120/hour)
* **QA Engineer**: $9,000 (120 hours @ $75/hour)
* **Project Management**: $8,000 (80 hours @ $100/hour)
* **Total Professional Services**: $96,800

**AWS Development and Testing Environment**

* **Development Infrastructure**: $3,000 (Load balancer, instances, storage for 3.5 months)
* **Load Testing Environment**: $2,000 (Dedicated testing infrastructure and tools)
* **Security Assessment Tools**: $1,500 (Penetration testing tools and third-party assessment)
* **Total Development Environment**: $6,500

**Training and Documentation**

* **Specialized Training**: $4,000 (Linux administration and AWS load balancer training)
* **Technical Documentation**: $3,500 (Architecture guides and operational procedures)
* **Knowledge Transfer**: $4,000 (Structured training sessions and handoff procedures)
* **Total Training and Documentation**: $11,500

**Contingency and Risk Management**

* **Implementation Contingency (20%)**: $22,960 (Complex infrastructure deployment risks)
* **Total Contingency**: $22,960

**Total Implementation Cost: $137,760**

**Operational Costs (Monthly)**

**AWS Infrastructure Services**

* **EC2 Instances**: $500 (3 t3.medium instances, 24/7 operation)
* **Network Load Balancer**: $250 (Data processing and hourly charges)
* **EFS Storage**: $400 (1TB General Purpose with backup)
* **Auto Scaling**: $150 (Additional instances during peak usage)
* **CloudWatch**: $125 (Detailed monitoring and custom metrics)
* **Data Transfer**: $200 (Inter-AZ and internet data transfer)
* **Total AWS Services**: $1,625

**Operational Support and Maintenance**

* **24/7 Monitoring**: $600 (Dedicated monitoring and alerting services)
* **System Administration**: $400 (Ongoing maintenance and configuration management)
* **Security Updates**: $200 (Regular patching and security maintenance)
* **Backup Management**: $150 (Backup monitoring and restoration testing)
* **Incident Response**: $300 (On-call support and issue resolution)
* **Total Support**: $1,650

**Total Monthly Operational Cost: $3,275**

**Annual Cost Projections**

**Year 1 (Implementation + Operations)**

* **Implementation Costs**: $137,760
* **Operational Costs (12 months)**: $39,300
* **Total Year 1**: $177,060

**Year 2 and Beyond (Operations Only)**

* **Annual Operational Costs**: $39,300
* **Annual Optimization**: -$3,000 (Reserved instances and efficiency improvements)
* **Net Annual Cost**: $36,300

**5-Year Total Cost of Ownership**

* **Implementation**: $137,760
* **Operations (5 years)**: $183,600
* **Total 5-Year TCO**: $321,360
* **Average Annual Cost**: $64,272

**Cost Optimization Opportunities**

**Short-term Optimizations (Year 1)**

* **Reserved Instances**: 1-year reserved instances for baseline capacity (30% savings)
* **EFS Performance Optimization**: Right-size throughput provisioning based on actual usage
* **Load Balancer Optimization**: Optimize data processing charges through traffic analysis
* **Estimated Savings**: $300-400/month

**Long-term Optimizations (Years 2-5)**

* **3-Year Reserved Instances**: Maximum discount for stable baseline capacity (50% savings)
* **Spot Instance Integration**: Use spot instances for non-critical scaling capacity
* **Storage Lifecycle Management**: Implement EFS Intelligent Tiering for cost optimization
* **Estimated Savings**: $600-800/month

**Risk Assessment**

**Risk Classification Matrix**

**High Risk Items**

**Complex Infrastructure Management**

* **Risk**: Multi-component architecture increases failure points and operational complexity
* **Probability**: High (Custom infrastructure requires specialized expertise)
* **Impact**: High (Service disruption affects all users, complex troubleshooting)
* **Mitigation**: Comprehensive documentation, 24/7 support, automated monitoring and alerting

**Linux Expertise Requirements**

* **Risk**: Specialized Linux and OpenSSH skills may not be available in-house
* **Probability**: Medium (Depends on current team capabilities)
* **Impact**: High (Implementation delays, ongoing operational challenges)
* **Mitigation**: External consulting, comprehensive training, knowledge transfer procedures

**Ongoing Maintenance Overhead**

* **Risk**: Significant operational burden for security updates, configuration management
* **Probability**: High (Inherent in self-managed infrastructure)
* **Impact**: Medium (Ongoing costs and resource allocation requirements)
* **Mitigation**: Automation tools, scheduled maintenance procedures, dedicated operations team

**Medium Risk Items**

**LDAP Integration Complexity**

* **Risk**: Windows AD integration may encounter connectivity or authentication issues
* **Probability**: Medium (Network dependencies and configuration complexity)
* **Impact**: Medium (Authentication failures, delayed implementation)
* **Mitigation**: Comprehensive testing, Windows AD expertise on team, fallback procedures

**Load Balancer Configuration**

* **Risk**: Network Load Balancer setup for SFTP may require specialized tuning
* **Probability**: Medium (TCP load balancing specifics for SSH/SFTP)
* **Impact**: Medium (Performance issues, connection distribution problems)
* **Mitigation**: AWS Professional Services consultation, extensive testing

**Performance Scaling Challenges**

* **Risk**: Manual tuning required for optimal performance under varying loads
* **Probability**: Medium (Depends on usage patterns and scaling requirements)
* **Impact**: Medium (User experience degradation during peak usage)
* **Mitigation**: Comprehensive load testing, automated scaling policies, performance monitoring

**Low Risk Items**

**AWS Service Reliability**

* **Risk**: EC2, EFS, and Load Balancer service outages
* **Probability**: Low (Mature AWS services with high availability SLAs)
* **Impact**: Medium (Service disruption during outages)
* **Mitigation**: Multi-AZ deployment, AWS Enterprise Support

**Risk Mitigation Strategies**

**Technical Risk Mitigation**

* **Infrastructure as Code**: Terraform deployment for consistent and repeatable infrastructure
* **Comprehensive Testing**: Multi-phase testing including failover and disaster recovery scenarios
* **Automated Configuration**: Ansible playbooks for consistent server configuration
* **Performance Monitoring**: Real-time monitoring with automated scaling and alerting

**Operational Risk Mitigation**

* **24/7 Support Coverage**: Dedicated operations team with escalation procedures
* **Documentation Excellence**: Comprehensive operational guides and troubleshooting procedures
* **Knowledge Transfer**: Structured training and cross-training for operational continuity
* **Vendor Support**: AWS Enterprise Support for critical issue resolution

**Business Risk Mitigation**

* **Phased Implementation**: Gradual rollout with rollback procedures at each phase
* **Parallel Operations**: Maintain existing system during transition with quick fallback
* **Communication Management**: Regular stakeholder updates and change management procedures
* **Service Level Management**: Clear SLAs and performance metrics with regular reporting

**Contingency Planning**

**Technical Contingencies**

* **Rollback Procedures**: Documented steps to revert to current Transfer Family system
* **Alternative Deployment**: Single-AZ deployment option if multi-AZ encounters issues
* **Performance Alternatives**: Instance type scaling options for performance requirements

**Resource Contingencies**

* **Backup Expertise**: Identified external consultants for specialized Linux and AWS skills
* **Extended Timeline**: Additional phases and resources for complex implementation scenarios
* **Budget Reserves**: 20% contingency budget for unforeseen infrastructure and operational challenges

**Success Metrics**

**Technical Performance Metrics**

**System Availability**

* **Target**: >99.5% uptime (acknowledging custom infrastructure complexity)
* **Measurement**: CloudWatch uptime monitoring and health check success rates
* **Monitoring**: Real-time availability dashboards with automated alerting
* **Recovery**: Mean Time to Recovery (MTTR) <5 minutes for automatic failover

**Performance and Scalability**

* **Target**: Support 500+ concurrent SFTP connections with <3 second response time
* **Measurement**: Connection metrics, file transfer throughput, system resource utilization
* **Monitoring**: Custom CloudWatch metrics and performance dashboards
* **Scaling**: Automated scaling response within 5 minutes of threshold breach

**Failover and Recovery**

* **Target**: Automatic failover within 5 minutes, manual recovery within 15 minutes
* **Measurement**: Health check failure detection and instance replacement timing
* **Monitoring**: Failover event tracking and recovery time analysis
* **Testing**: Monthly disaster recovery testing and documentation updates

**Authentication Success Rate**

* **Target**: >99.5% successful authentication attempts
* **Measurement**: LDAP authentication logs and failure analysis
* **Monitoring**: Authentication metrics with failure pattern analysis
* **Improvement**: Continuous optimization of AD connectivity and authentication logic

**Operational Excellence Metrics**

**Incident Response Performance**

* **Target**: <15 minutes initial response time for critical issues
* **Measurement**: Incident ticket timestamps and escalation tracking
* **Monitoring**: Automated alerting with escalation procedures
* **Improvement**: Regular review and optimization of response procedures

**Maintenance and Updates**

* **Target**: <4 hours monthly maintenance window, 99% successful patch deployment
* **Measurement**: Maintenance window duration and success rate tracking
* **Monitoring**: Automated patching status and rollback procedures
* **Planning**: Proactive maintenance scheduling and change management

**Cost Management**

* **Target**: Maintain operational costs within $3,275/month budget
* **Measurement**: Monthly AWS billing analysis and resource utilization
* **Monitoring**: Real-time cost dashboards with budget alerts
* **Optimization**: Quarterly cost review and optimization initiatives

**Business Impact Metrics**

**User Satisfaction**

* **Target**: >90% positive feedback from users (lower than managed service due to complexity)
* **Measurement**: User surveys and support ticket sentiment analysis
* **Monitoring**: Regular user feedback collection and trend analysis
* **Improvement**: Continuous enhancement based on user feedback and usage patterns

**Administrative Efficiency**

* **Target**: 25% reduction in user management overhead (less than Lambda approach due to complexity)
* **Measurement**: Time tracking for user provisioning and support activities
* **Monitoring**: Monthly efficiency reports and process analysis
* **Enhancement**: Automation initiatives to reduce manual administrative tasks

**Security and Compliance**

* **Target**: Zero security incidents, successful quarterly security audits
* **Measurement**: Security event logs and audit compliance reports
* **Monitoring**: Continuous security monitoring and vulnerability assessment
* **Maintenance**: Regular security updates and configuration reviews

**Next Steps**

**Immediate Actions (Week 1)**

**Stakeholder Decision and Resource Commitment**

* **Executive Review**: Present EC2 cluster proposal with comprehensive cost-benefit analysis
* **Resource Assessment**: Evaluate internal Linux expertise and determine training/hiring needs
* **Budget Approval**: Secure implementation budget of $137,760 and ongoing operational budget
* **Timeline Commitment**: Confirm 3.5-month implementation timeline and resource availability

**Technical Prerequisites Validation**

* **Infrastructure Capacity**: Assess current AWS account limits and request increases if needed
* **Network Connectivity**: Validate VPC connectivity requirements for LDAP integration
* **Security Requirements**: Review security policies and compliance requirements for self-managed infrastructure
* **Operational Readiness**: Assess current operational capabilities for 24/7 infrastructure management

**Team Assembly and Expertise Gap Analysis**

* **Linux Administration**: Identify or recruit two qualified Linux administrators with OpenSSH expertise
* **Network Engineering**: Assign network engineer with load balancer and high availability experience
* **Operational Support**: Plan for 24/7 operations team and support procedures
* **Training Planning**: Develop training plan for existing team members on new technologies

**Alternative Evaluation**

* **Managed Service Comparison**: Final comparison with Lambda Custom IdP approach for executive decision
* **Risk Tolerance Assessment**: Evaluate organizational appetite for operational complexity vs. control
* **Long-term Strategy**: Consider EC2 approach within broader infrastructure management strategy
* **Hybrid Considerations**: Evaluate potential for hybrid approach or future migration path

**Short Term Actions (Weeks 2-6)**

**Detailed Architecture and Planning**

* **Infrastructure Design**: Complete detailed VPC, load balancer, and scaling architecture
* **Security Architecture**: Develop comprehensive security hardening and monitoring plan
* **Capacity Planning**: Detailed analysis of instance types, scaling parameters, and performance requirements
* **Integration Planning**: Finalize Windows AD integration approach and authentication flow design

**Proof of Concept Development**

* **Small-scale Deployment**: Deploy proof of concept with single instance and basic LDAP integration
* **Load Balancer Testing**: Validate Network Load Balancer configuration with SFTP traffic
* **EFS Performance Testing**: Test shared storage performance and concurrent access patterns
* **Authentication Validation**: Prove Windows AD integration and multi-group processing concept

**Operational Preparation**

* **Monitoring Strategy**: Design comprehensive monitoring and alerting framework
* **Support Procedures**: Develop initial operational procedures and escalation processes
* **Documentation Framework**: Establish documentation standards and knowledge management approach
* **Training Curriculum**: Develop training materials for team members and ongoing support

**Vendor and Tool Selection**

* **Configuration Management**: Select and implement Ansible or Chef for automated configuration
* **Monitoring Tools**: Implement additional monitoring and alerting tools beyond CloudWatch
* **Security Tools**: Select and implement security scanning and compliance tools
* **Backup Solutions**: Design and implement backup and disaster recovery procedures

**Medium Term Actions (Month 1)**

**Infrastructure Foundation Deployment**

* **VPC and Networking**: Deploy complete networking infrastructure across multiple AZs
* **Load Balancer Setup**: Configure Network Load Balancer with proper health checks
* **Auto Scaling Configuration**: Implement Auto Scaling Groups with launch templates
* **EFS Deployment**: Deploy shared file system with proper access points and security

**SFTP Service Development**

* **OpenSSH Configuration**: Develop and test SFTP-only configuration with security hardening
* **LDAP Integration**: Implement and test Windows AD authentication with group processing
* **User Management**: Develop automated user provisioning and home directory management
* **Security Hardening**: Implement comprehensive security measures and audit logging

**Monitoring and Operations Setup**

* **CloudWatch Configuration**: Implement comprehensive monitoring and custom metrics
* **Alerting Framework**: Configure automated alerting with appropriate thresholds
* **Operational Dashboards**: Create monitoring dashboards for operations team
* **Support Procedures**: Finalize operational procedures and team training

**Long Term Considerations**

**Continuous Improvement and Optimization**

* **Performance Optimization**: Ongoing tuning based on production usage patterns
* **Cost Optimization**: Regular review and optimization of AWS resource utilization
* **Security Enhancement**: Continuous security assessment and improvement initiatives
* **Operational Excellence**: Regular review and enhancement of operational procedures

**Strategic Alignment and Evolution**

* **Technology Roadmap**: Plan for future technology evolution and upgrade paths
* **Scalability Planning**: Ensure architecture can accommodate continued organizational growth
* **Integration Opportunities**: Identify opportunities for integration with other systems
* **Migration Considerations**: Plan for potential future migration to managed services if appropriate

This comprehensive implementation approach ensures successful deployment of the Load-Balanced EC2 Linux SFTP Cluster while managing the inherent complexity and operational requirements. The detailed planning and risk management approach provides clear accountability and measurable progress toward achieving high availability SFTP services with complete infrastructure control.