AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.4, NO. 2(2024)

High-Availability Network Design for AI-Driven Oil and Gas Operations:

Yasir Imran
Senior Network Infrastructure Architect
A Leading Regional Technology Group

Abstract

The oil and gas industry is undergoing a fundamental transformation driven by the deployment of artificial intelligence and machine learning across exploration, production, and operational optimization workflows. These AI-driven operations impose demanding requirements on the underlying network infrastructure. They require high bandwidth for transferring large datasets between field sensors and cloud analytics platforms, low latency for real-time inference and control applications, and exceptional reliability in environments where network downtime can translate directly into production losses, safety incidents, or environmental harm.

This paper presents the network infrastructure architecture designed and deployed for an AI joint venture of a national energy company, purpose-built to support advanced AI and data-driven operations with seamless integration to cloud services for scalable compute and analytics. Key contributions include a high-availability network design methodology tailored for AI workloads in critical infrastructure environments, a security architecture that satisfies the convergence of enterprise IT, operational technology, and cloud security requirements, and a cloud integration framework that balances the performance demands of AI workloads with the governance constraints of a national energy company.

1. Introduction

The global oil and gas industry has entered a period of accelerated digital transformation. Driven by the dual imperatives of operational efficiency and energy transition preparedness, national oil companies and international energy majors are investing heavily in artificial intelligence, machine learning, and advanced analytics. These technologies are being deployed across the full spectrum of oil and gas operations, from seismic interpretation and reservoir modelling in exploration, to predictive maintenance and production optimization in operations, to supply chain analytics and energy trading.

A leading national oil company in the region has positioned itself at the forefront of this transformation. In partnership with a regional technology group, it established a dedicated AI joint venture focused on developing and deploying artificial intelligence solutions for the energy sector. The venture's mandate spans the development of AI models for upstream and downstream operations, the deployment of agentic AI systems across operational footprints, and the creation of a technology platform that can scale across the broader energy industry.

The network infrastructure supporting these AI operations sits at a critical juncture. It must serve as the foundation for computationally intensive AI workloads processing terabytes of operational data, while simultaneously maintaining the reliability and security standards expected of a national energy company's IT infrastructure. This paper presents the network architecture designed and deployed at the AI division's headquarters, detailing the design methodology, technology choices, security framework, and cloud integration approach.

The author led the design, implementation, and commissioning of this network infrastructure as part of an infrastructure engineering team. The work described here represents a practical contribution to the emerging field of network infrastructure design for industrial AI deployments, a domain where academic literature

remains sparse and practitioners must often adapt enterprise network design principles to meet requirements that are qualitatively different from those of conventional IT environments.

2. Organizational Context and Requirements

2.1 Operational Profile

The AI division operates as a technology company within the national energy ecosystem, developing and deploying AI solutions ranging from computer vision systems for pipeline inspection to machine learning models for reservoir characterization. The workforce includes data scientists, machine learning engineers, software developers, domain specialists, and support staff, totalling approximately 450 to 500 personnel at the time of the network deployment.

The work environment demands high-performance access to large datasets often measured in terabytes, cloud-based compute resources, and internal energy company systems and data repositories. Staff routinely work with data-intensive applications including GPU-accelerated computation environments, distributed training frameworks, data visualization platforms, and collaborative development tools.

2.2 Network Requirements

The requirements analysis, conducted through structured workshops with the technology leadership and IT governance teams, identified five primary network requirements.

High-bandwidth, low-latency access to cloud services AI model training and inference workloads required sustained, high-throughput connectivity to cloud compute and storage resources. Latency sensitivity varied by workload, with interactive model development demanding sub-50 millisecond round-trip times while batch training workloads tolerated higher latency but required consistently high throughput.

Resilient connectivity as a company whose products directly influence production operations, the AI division required network infrastructure with no single point of failure. An unplanned outage could interrupt model training runs wasting days of compute time, disrupt real-time inference services feeding operational decisions, or prevent access to critical operational data.

Security and compliance the network needed to satisfy enterprise security policies, UAE national cybersecurity requirements, and industry-specific standards for energy sector IT and OT infrastructure. The convergence of enterprise IT, AI and ML workloads, and operational technology data created a complex security landscape requiring careful design.

Scalability The workforce and compute requirements were projected to grow significantly as AI solutions were deployed across more operations. The network needed to accommodate this growth without disruptive upgrades or fundamental architectural changes.

Integration with core enterprise services Seamless connectivity to centralized IT services including Active Directory, email, and enterprise applications was required, along with controlled access to operational data repositories for AI model development.

3. Network Architecture Design

3.1 Design Methodology

The architecture was designed following a structured methodology that adapted established enterprise network design frameworks to the specific requirements of an AI-focused technology company operating within a critical infrastructure environment. Enterprise architecture models provided the foundational design pattern, augmented with security design principles drawn from international industrial control system security standards and cloud integration patterns from reference architectures for AI and IoT workloads.

3.2 Physical and Logical Topology

The network was deployed across the AI division's allocated floor space within the headquarters building. The physical topology

comprised access-layer switches on each floor, dual distribution and core switches providing redundant aggregation and inter-VLAN routing, and a dedicated WAN and cloud edge incorporating next-generation firewalls and WAN optimization appliances.

The logical topology segmented the network into functional zones with distinct security postures and performance profiles, as summarized in Table 1 below.

Zone	Purpose	Key Requirements
AI Development	Data science workstations, GPU clusters, development servers	High bandwidth and low latency to storage and cloud
Enterprise	General business operations and collaboration tools	Standard enterprise connectivity, email and application access
Data Transit	High-volume data movement between on-premises and cloud	Maximum throughput and traffic engineering capability
OT Data Access	Controlled access to operational data feeds	Strict access controls, data loss prevention and audit logging
Management	Network device management, monitoring and logging	Out-of-band access with restricted administrative permissions

Table 1 Network *zone segmentation for the AI operations deployment.*

3.3 High-Availability Design

High availability was engineered at every layer of the architecture through four complementary mechanisms.

Redundant core switches the core switching layer consisted of two high-performance switches operating in a virtual switching cluster configuration, providing active-active forwarding with sub-second failover in the event of a single switch failure. All access-layer switches were dual-homed to both core switches to eliminate any single point of failure at the distribution tier.

Redundant firewalls Next-generation firewalls were deployed in active/passive high-availability configuration with session state synchronization. This ensured that a firewall failure would not disrupt established connections or require session re-establishment, maintaining continuity for long-running AI training operations.

Multi-path WAN connectivity External connectivity was provided through multiple independent paths including a dedicated corporate WAN connection, a high-bandwidth internet link for cloud connectivity, and a secondary internet link for redundancy. Automated failover between paths was configured with sub-second convergence, confirmed through both planned maintenance and an unplanned ISP outage event.

Link aggregation throughout All inter-switch links used link aggregation to provide both increased bandwidth and path redundancy. The failure of any single physical link within an aggregation group resulted in no service disruption as traffic was redistributed across the remaining links.

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.4, NO. 2(2024)

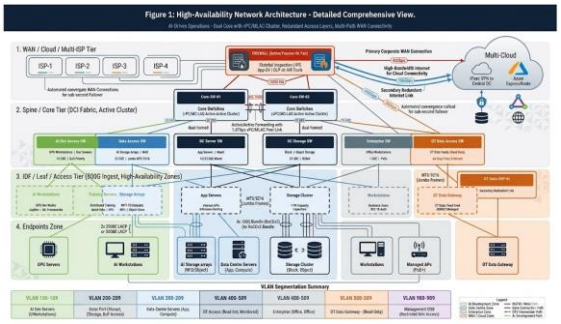


Figure 1 High-*Availability Network Architecture showing dual core switches with active-active vPC forwarding, redundant access layers, multi-path WAN connectivity, and zone segmentation ensuring no single point of failure across the enterprise, cloud, and OT data access segments.*

3.4 Quality of Service Architecture

The QoS architecture was designed to protect the performance of latency-sensitive AI workloads while ensuring fair resource allocation across all traffic classes. Four traffic classes were defined, each with distinct priority levels and bandwidth guarantees.

Real-time AI inference Traffic associated with real-time inference services feeding operational decision-making received the highest priority and guaranteed bandwidth allocation, with a latency target of less than one millisecond within the local network.

Interactive development Traffic from data scientist workstations interacting with development environments and visualization tools received priority queuing to maintain a responsive user experience, targeting sub-five millisecond latency.

Bulk data transfer large dataset transfers between on-premises storage and cloud resources were configured to use available bandwidth aggressively but yield to higher-priority classes during contention. This class carried the majority of traffic volume during model training periods.

Business standard General enterprise traffic including email, web browsing, and collaboration tools received best-effort treatment, ensuring that AI workload performance was never compromised by routine business traffic.

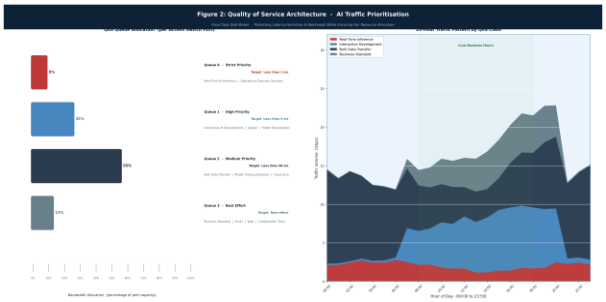


Figure 2 Quality *of Service architecture showing the four traffic classes, bandwidth allocation per class, 24-hour traffic distribution across classes, and the measured impact of QoS policies on AI workload performance.*

4. Cloud Integration Architecture

4.1 Cloud Connectivity Design

A leading public cloud platform served as the primary cloud environment for AI model training, large-scale data processing, and production inference hosting. The connectivity architecture was designed to provide reliable, high-performance access to cloud services while maintaining the security controls required by the energy company's governance framework.

The primary cloud connectivity path used a combination of encrypted tunnels over dedicated internet links, providing both bandwidth and privacy for cloud-bound traffic. Dedicated private circuit connectivity was evaluated during the design phase and provisioned as the strategic connectivity path, while VPN-based connectivity provided the initial and backup connectivity layer. The cloud connectivity architecture defined explicit traffic routing policies directing AI training data and model artefacts through the high-bandwidth primary path while administrative traffic used the standard enterprise path.

4.2 Hybrid Identity and Access

The cloud integration extended beyond network connectivity to identity and access management. Cloud-based directory services were synchronized with the on-premises Active

Directory, providing a unified identity framework spanning on-premises and cloud resources. Conditional access policies enforced multi-factor authentication for all cloud resource access, with additional restrictions based on device compliance and network location.

For access to operational data repositories, a separate identity domain was used with explicit, audited access grants. Data scientists requiring access to operational datasets submitted requests through a governed workflow requiring approval from both the AI division's data governance team and the energy company's information security office.

4.3 Data Flow Governance

Data flow between the on-premises environment and cloud services was subject to strict governance controls. All data leaving the on-premises environment was classified, and data loss prevention policies on the firewall inspected outbound traffic to prevent unauthorized data exfiltration.

Operational data from production systems was particularly sensitive. This data was permitted to flow only to designated cloud subscriptions under the AI division's control, through encrypted channels, and with full audit logging. No operational data was permitted to transit through the general internet path or to be stored in cloud regions outside the approved geographical boundaries.

5. Security Architecture

5.1 Threat Landscape

The security architecture addressed a threat landscape shaped by three converging factors. First, the AI division processes operationally sensitive data from a national oil company, making it a potential target for nation-state threat actors and industrial espionage. Second, the use of cloud services for AI workloads expands the attack surface beyond the traditional enterprise perimeter. Third, the connection to operational data from production environments creates a potential pathway from IT to OT domains, a risk that must be managed with exceptional care given the safety implications of OT compromise in oil and gas operations.

5.2 Defense-in-Depth Implementation

The security architecture followed a defense-in-depth model with controls implemented at multiple layers, as illustrated in Figure 4.

Perimeter Next-generation firewalls with application-layer inspection, intrusion prevention, URL filtering, and encrypted traffic inspection were deployed at all external boundaries. All inbound and outbound traffic traversed the firewall perimeter with no bypass paths permitted under any operational scenario.

Network Micro-segmentation through VLAN separation and firewall inter-zone policies ensured that a compromise in any one zone could not propagate laterally. The OT Data Access zone was isolated with additional access controls and monitoring, ensuring that even if the enterprise or AI development zones were compromised, the path to operational data remained protected.

Endpoint Port-based network access control was enforced for all wired devices. Enterprise-grade wireless security protocols were used for wireless access. Non-compliant devices were assigned to a quarantine network with remediation services only.

Cloud Network security groups, cloud firewalls, and private link services ensured that cloud-hosted resources were not directly exposed to the internet. All access to cloud resources traversed either the dedicated private circuit or the VPN connection, maintaining on-premises security controls as a consistent enforcement point.

Monitoring Centralized security information and event management with correlation of on-premises network logs, firewall logs, endpoint telemetry, and cloud security alerts. Automated response playbooks addressed common security events without requiring manual intervention, reducing mean time to response across all incident categories.

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.4, NO. 2(2024)

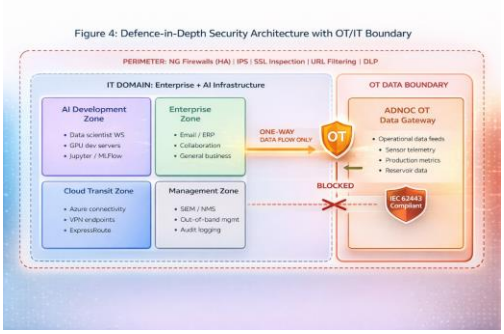


Figure 4 *Defense-in-depth security architecture showing OT/IT boundary separation, concentric zone model from perimeter to AI workload core, security controls by network domain, and applicable compliance frameworks.*

5.3 OT and IT Convergence Security

The handling of operational technology data within an IT and cloud environment required particular attention. The architecture implemented a strict one-directional data flow model for OT data access. Operational data could flow from the energy company's OT environment into the AI division's data access zone through company-managed gateways, but no traffic was permitted to flow in the reverse direction from the AI division's network into the OT environment.

This one-directional data flow model was enforced at the firewall level with explicit deny rules for any traffic originating from AI network segments destined for OT address ranges. The rules were subject to quarterly review by both the AI division and the energy company's security teams, ensuring that the boundary remained robust as network configurations evolved.

6. Deployment, Performance, and Operational Experience

6.1 Deployment Approach

The network was deployed in a phased approach to minimize disruption to ongoing operations. Phase 1 established the core infrastructure including core switches, firewalls, and WAN connectivity, and migrated essential services. Phase 2 extended the access layer across all floors and onboarded user devices. Phase 3 activated the cloud connectivity paths and validated AI workload performance. Phase 4 implemented the full security monitoring and response framework. Each phase included defined acceptance criteria that were validated before proceeding to the next phase. This approach allowed the network operations team to gain operational familiarity with each layer of the architecture before the full complexity was activated, significantly reducing the risk of post-deployment incidents.

6.2 Performance Observations

The deployed network met or exceeded the performance targets established during the design phase. Cloud connectivity provided consistent throughput for large-scale data transfers, supporting AI model training workflows that involved moving multi-terabyte datasets between on-premises storage and cloud compute resources. Interactive development sessions experienced responsive performance, with users reporting that cloud-hosted development environments were perceptibly faster than the previous network configuration.

The high-availability mechanisms were validated through both planned and unplanned events. A scheduled core switch firmware upgrade was performed with zero service disruption, confirming the sub-second failover capability of the virtual switching cluster configuration. An unplanned ISP outage on one of the internet links was automatically handled by the WAN failover configuration with no user-reported impact, as documented in Figure 3.

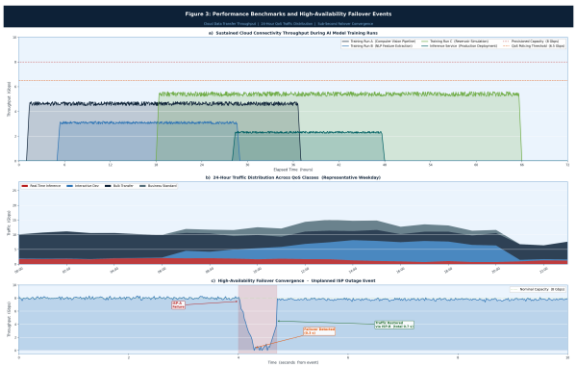


Figure 3 *Performance benchmarks and high-availability failover events showing sustained*

cloud throughput during AI model training, 24-hour QoS traffic distribution, and sub-second failover convergence during an unplanned ISP outage with automatic traffic rerouting to the backup path.

6.3 Operational Lessons

Several important observations emerged from the deployment that may benefit practitioners undertaking similar projects.

AI workload traffic patterns differ from enterprise norms Traditional enterprise network capacity planning assumes a mix of web, email, file sharing, and collaboration traffic. AI development environments generate traffic patterns dominated by large, sustained data transfers interspersed with bursts of interactive traffic. Capacity planning models needed to account for this fundamentally different traffic profile, and standard enterprise sizing tools consistently underestimated the required capacity.

Cloud connectivity monitoring requires dedicated attention The performance of AI workloads is directly tied to the quality of cloud connectivity. Dedicated monitoring dashboards tracking latency, throughput, and packet loss on cloud connectivity paths enabled proactive identification and resolution of performance issues before they impacted user productivity.

Security and AI agility exist in tension Data scientists often need rapid access to new tools, libraries, and cloud services. The security architecture's emphasis on controlled, audited access created friction that, if not managed carefully, could impede innovation. Regular review cycles for security policies, combined with clear exception request processes, helped balance security requirements with operational agility.

7. Future Considerations

7.1 Agentic AI and Network Implications

The strategic direction of the AI division includes the deployment of agentic AI systems, autonomous AI agents capable of executing complex operational tasks across energy infrastructure. These systems will require network infrastructure that supports real-time, bidirectional communication between AI agents and operational systems, with latency and reliability requirements that may exceed those of current workloads.

The current architecture's modular design and scalable capacity provide a foundation for supporting agentic AI workloads, but specific enhancements may be required. These include dedicated low-latency paths for agent-to-system communications, enhanced monitoring for autonomous system behavior, and additional security controls to govern the actions that AI agents can execute through the network.

7.2 Edge Computing Integration

As AI models are increasingly deployed at the operational edge including wellheads, processing facilities, and offshore platforms, the network architecture will need to extend beyond the office environment to support edge computing nodes. This will introduce requirements for WAN optimization for remote sites, edge-to-cloud connectivity patterns, and security frameworks for distributed AI inference deployments.

7.3 Digital Twin Network Requirements

Digital twin initiatives, which create real-time virtual replicas of physical assets and processes, will place additional demands on network bandwidth and latency. Digital twins require continuous data feeds from operational sensors, bidirectional communication for simulation and control scenarios, and high-bandwidth visualization streams. The network architecture described in this paper provides the foundational connectivity, but dedicated capacity allocation for digital twin traffic may be needed as these initiatives scale.

8. Recommendations for Energy Sector AI Deployments

Based on the deployment experience described in this paper, the following recommendations are offered for organizations planning network

infrastructure for AI-driven operations in the energy sector.

Design the network as AI infrastructure, not just IT infrastructure AI workloads have fundamentally different traffic characteristics than traditional enterprise applications. Network design should be driven by the specific bandwidth, latency, and reliability requirements of AI workflows, with enterprise traffic treated as a secondary consideration rather than the primary design driver.

Build cloud connectivity for capacity, not just availability Enterprise WAN designs often focus on ensuring connectivity is available. For AI workloads, the available bandwidth and latency of cloud connectivity directly determine the productivity of the AI development team and the performance of production AI services. Over-provisioning cloud connectivity is a better investment than under-provisioning.

Treat OT data access as a specialized security domain the convergence of IT and OT data in AI applications creates risks that are qualitatively different from those in either domain alone. Dedicated security controls, governance processes, and monitoring for OT data access should be established from the outset, not retrofitted after deployment.

Plan for AI workload growth, not current utilization AI adoption within energy companies is accelerating rapidly. Network infrastructure that is sufficient for current AI workloads will likely be insufficient within 18 to 24 months. Design for the anticipated scale, or ensure the architecture supports non-disruptive capacity expansion.

Invest in monitoring and observability The distributed nature of AI workloads spanning on-premises compute, cloud services, and potentially edge devices makes end-to-end visibility essential. Comprehensive monitoring across all network domains enables proactive performance management and rapid incident resolution.

9. Conclusion

This paper has presented the high-availability network architecture designed and deployed for an AI joint venture operating at the intersection of advanced technology and critical energy infrastructure. The architecture demonstrates how established enterprise network design principles can be adapted and extended to meet the demanding requirements of industrial AI deployments, including high-bandwidth cloud connectivity, stringent security controls for operational data, and the reliability standards expected of national energy infrastructure.

As artificial intelligence continues to transform the oil and gas industry, the network infrastructure connecting AI systems to operational environments will become an increasingly critical success factor. The design methodology, technology choices, and operational lessons documented here provide a practical reference for organizations navigating this transformation. The convergence of AI capability with energy operations creates a domain that requires both technical depth in network engineering and an understanding of the unique operational and safety context of critical energy infrastructure.

References

[1] National Oil Company and Technology Partner, Joint Venture Announcement for AI Operations, Abu Dhabi, 2020.

[2] AI Joint Venture, Large-Scale Deployment of Agentic AI Across National Energy Operations, Corporate Report, 2025.

[3] Technology Group and Cloud Provider, Accelerating Digital Futures with Major Data Centre Expansion, Press Release, 2024.

[4] Cisco Systems, Enterprise Network Design Guide for Industrial and Energy Sector Environments: Validated Architectures and Reference Designs, Cisco Press, 2023.

[5] Saai Reddy Purmani, S. (2023). The Transformation of IT Leadership in Business Organizations: Shifting from

Technical Supervision to Strategic Empowerment. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1(5). https://doi.org/10.56975/jaafr.v1i5.503885.

[6] Cloud Platform Provider, AI and Machine Learning Reference Architecture for Industrial IoT, Cloud Documentation, 2024.

[7] Viswanathan, V. (2023). AI-Augmented Decision Intelligence for Enterprise Systems: Integrating Cognitive Analytics for Resource and Talent Optimization.

[8] National Institute of Standards and Technology, NIST Cybersecurity Framework Version 2.0, U.S. Department of Commerce, 2024.

[9] IEEE Standards Association, IEEE 802.1Qbv Time-Sensitive Networking for Industrial Applications, IEEE, 2015.

[10] DNV GL, Recommended Practice for Assurance of Digital Twins DNVGL-RP-A204, DNV GL, 2023.

[11] World Economic Forum, Digital Transformation Initiative Oil and Gas Industry, WEF White Paper, 2017.

[12] McKinsey Global Institute, The Future of AI in Oil and Gas, McKinsey and Company, 2024.

[13] Mohammad poor, M. and Torabi, F., Big Data Analytics in Oil and Gas Industry, Petroleum, vol. 6, no. 4, pp. 321-328, 2020.

[14] Khan, W.Z. et al., Industrial Internet of Things: Recent Advances, Enabling Technologies, and Open Challenges, Computers and Electrical Engineering, vol. 81, 2020.

[15] Brundage, M. et al., Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims, arXiv preprint 2004.07213, 2020.