South Asian Journal of Science and Technology

Open Access Full Text Article 12(5)(2022) 9-23 DOI: 10.26524/sajet.2022.12.60

Hybrid Cloud Connectivity: Performance Comparison of AWS Direct Connect vs. VPN Tunnels

Haritha Bhuvaneswari Illa

Amazon web services Inc, Texas, USA
Email: illaharitha030@gmail.com

Article info
Received 30th December 2021 Received in revised form 7 February 2022
Accepted 10 September 2022

Keywords:
Hybrid Cloud, AWS Direct Connect, VPN Tunnel, Network Performance, Latency, Throughput, Cost Analysis, Cloud Connectivity.

https://sajet.in/index.php/journal/article/view/342

Abstract

The rapid adoption of hybrid cloud infrastructures has intensified the need for efficient, reliable, and scalable connectivity between enterprise on-premises environments and public cloud platforms. This study presents a comprehensive performance and cost comparison between AWS Direct Connect and AWS Site-to-Site VPN tunnels to determine their relative suitability for hybrid cloud deployments. Using a controlled experimental setup that simulated real-world enterprise workloads including bulk data transfer, transactional web traffic, and database synchronization the research evaluated both connectivity methods across six key performance parameters: latency, throughput, jitter, packet loss, availability, and cost-per-gigabyte.

The findings demonstrate that AWS Direct Connect consistently outperforms VPN tunnels, achieving 66% lower latency (3.9 ms vs. 11.6 ms), 132% higher throughput (924 Mbps vs. 398 Mbps), and 96% lower packet loss (0.03% vs. 0.87%). Direct Connect maintained exceptional stability with a jitter average of 0.7 ms and an availability rate of 99.98%, whereas VPN tunnels exhibited greater fluctuation due to internet routing dynamics. Although Direct Connect involves higher provisioning costs, it becomes economically advantageous for data transfers exceeding 5 TB per month, achieving nearly fivefold improvement in cost-performance efficiency (CPI) relative to VPN tunnels.

Regression analysis revealed a strong inverse correlation between latency and throughput for Direct Connect ($r = -0.82$), confirming its deterministic behavior under variable workloads. Application-level evaluations spanning S3 data sync, HTTP web transactions, and MySQL queries further reinforced Direct Connect's superior responsiveness and reliability.

In summary, AWS Direct Connect emerges as the optimal connectivity model for high-throughput, latency-sensitive enterprise applications requiring predictable performance and cost scalability, while VPN tunnels remain viable for flexible, moderate-load, or temporary hybrid configurations. The results provide a quantitative

foundation for network architects and decision-makers optimizing hybrid cloud connectivity strategies in enterprise-scale environments.

1. INTRODUCTION

The evolution of enterprise IT ecosystems toward hybrid cloud architectures represents a pivotal transformation in how organizations balance flexibility, performance, and cost in their computing environments (Li et al., 2010; Haq et al., 2017; Yeganeh et al., 2019). Hybrid connectivity the seamless integration of on-premises infrastructure with cloud platforms such as Amazon Web Services (AWS) has emerged as a cornerstone of modern digital transformation strategies (Demchenko et al., 2013; Calder et al., 2013; Dhamdhere & Dovrolis, 2010). Among the available interconnection options, AWS Direct Connect and Virtual Private Network (VPN) tunnels stand out as the two most widely implemented technologies for linking enterprise networks with AWS resources. While both serve the purpose of providing secure access to cloud resources, they differ significantly in terms of infrastructure, routing behavior, performance consistency, and cost implications (Ager et al., 2012; Calder et al., 2015; Yeganeh et al., 2019). This study presents a comprehensive performance comparison of these two connectivity paradigms under controlled, real-world conditions, providing quantitative and qualitative insights to guide enterprise decision-making in hybrid network design.

The hybrid cloud model fundamentally aims to merge the control and security of private data centers with the elasticity and scalability of the public cloud (Krishna et al., 2019; Motamedi et al., 2014). However, the success of this architecture is contingent upon the efficiency and reliability of the connecting network. VPN tunnels, typically established using the Internet Protocol Security (IPsec) framework, route encrypted traffic over the public internet (Spring et al., 2002; Shavitt & Shir, 2005). This approach offers rapid deployment, minimal initial cost, and global accessibility. However, it is inherently limited by the variable latency, jitter, and packet loss associated with public internet routes (Anwar et al., 2015; Tariq et al., 2005). Conversely, AWS Direct Connect provides a private, dedicated network connection between an enterprise's on-premises environment and an AWS region, bypassing the public internet (Yeganeh et al., 2019; Demchenko et al., 2013). It delivers predictable latency, consistent throughput, and enhanced data privacy but entails higher setup costs and provisioning complexity (Gill et al., 2008; Ager et al., 2012).

In recent years, enterprises have faced increasing challenges in optimizing their hybrid cloud connectivity to meet the demands of latency-sensitive applications such as real-time analytics, financial trading systems, and industrial control platforms (Zhang et al., 2017; Haq et al., 2017). The exponential growth of data volumes, coupled with stringent compliance requirements and cost pressures, has necessitated an empirical understanding of the trade-offs between the flexibility of VPN-based connectivity and the determinism of Direct Connect links (Calder et al., 2015; Li et al., 2010). While AWS provides theoretical performance guidance, limited independent studies have quantitatively benchmarked these two services under comparable workloads and network conditions (Yeganeh et al., 2019; Alexander et al., 2018). Therefore, this research aims to fill this gap by systematically evaluating latency, throughput, jitter, packet loss, and availability, alongside an analysis of cost-performance trade-offs (Padhye et al., 1998; Labovitz et al., 2010).

The motivation for this research arises from the practical need to make informed architectural decisions in hybrid deployments. While small and medium enterprises (SMEs) may prioritize the agility and cost-effectiveness of VPN tunnels, large-scale enterprises with mission-critical workloads may favor the stability and performance guarantees of dedicated links (Madhyastha et al., 2006; Dhamdhere & Dovrolis, 2010). Understanding the precise thresholds where one approach outperforms the other can optimize operational expenditure and network design efficiency (Cunha et al., 2016; Chiu et al., 2015). Moreover, the rapid evolution of software-defined networking (SDN), multipath VPNs, and hybrid routing frameworks

necessitates a data-driven assessment that extends beyond vendor marketing claims (Durairajan et al., 2015; Alexander et al., 2018).

This study was conducted using an experimental testbed configured to emulate a realistic enterprise–AWS hybrid network. The research compared AWS Direct Connect (1 Gbps dedicated connection) and AWS Site-to-Site VPN using identical Virtual Private Cloud (VPC) environments, identical routing policies, and equivalent EC2 instance configurations. The experiments were designed to simulate diverse workloads including high-volume data transfers, database transactions, and real-time HTTP requests across multiple AWS regions (Calder et al., 2013; Ager et al., 2012). The analysis extended beyond raw network metrics to include service-level reliability and cost-effectiveness, thereby producing a holistic evaluation framework (Klöti et al., 2016; Gill et al., 2008).

The significance of this work lies not only in the raw performance metrics but also in the interpretation of connectivity behaviors under variable network loads. The study adopts a multi-layer analytical approach encompassing network performance, application responsiveness, and operational economics (Li et al., 2010; Haq et al., 2017). This integrated framework allows for identifying practical deployment scenarios best suited for each connectivity option. For example, organizations with bursty workloads might find the flexibility of VPN tunnels preferable, whereas consistently high-volume data transfers would justify the fixed cost of Direct Connect (Yeganeh et al., 2019; Calder et al., 2015).

Additionally, this research addresses the increasing intersection between cloud connectivity and cybersecurity. While both Direct Connect and VPN tunnels provide encryption and isolation mechanisms, their exposure to public or private routes affects the potential attack surface (Ager et al., 2012; Shavitt & Shir, 2005). By incorporating AWS security configurations such as Network Access Control Lists (NACLs), routing tables, and security groups, the study ensures that performance results are interpreted in a secure, realistic context (Gill et al., 2008; Demchenko et al., 2013).

The findings of this study are expected to have broad implications for cloud architects, network engineers, and enterprise IT strategists. With enterprises increasingly adopting multi-cloud and hybrid models, the ability to empirically validate connectivity performance is vital for designing scalable and cost-efficient systems. Moreover, as regulatory standards such as GDPR and ISO/IEC 27001 place emphasis on data transmission integrity, understanding the implications of connectivity choice on compliance and resilience becomes essential.

Finally, this paper contributes to the growing body of research on cloud networking performance by providing a reproducible methodology for comparative benchmarking. The experimental design, statistical treatment of results, and visualization of performance patterns can serve as a reference model for future studies exploring interconnectivity options between cloud service providers or across hybrid multi-cloud environments.

In summary, this study aims to empirically answer the following guiding questions:

1. How does AWS Direct Connect performance compare to IPsec-based VPN tunnels in terms of latency, throughput, jitter, and reliability under varying workloads?
2. How do cost structures influence the overall value proposition of each connectivity method?
3. What framework can organizations adopt to determine the most suitable connectivity solution based on operational and economic priorities?

Through this analytical lens, the research intends to bridge the gap between theoretical performance expectations and real-world operational realities, offering actionable insights for hybrid cloud connectivity planning.

2. Methodology

The methodology of this study was designed to ensure reproducibility, transparency, and technical precision in comparing AWS Direct Connect and VPN tunnel performance under identical hybrid cloud conditions. The research followed a quantitative experimental design, employing controlled network

environments, standardized workloads, and automated monitoring tools to obtain measurable data on performance, reliability, and cost.

2.1. Research Design and Experimental Framework

The study utilized a hybrid cloud simulation model integrating an on-premises data center environment with Amazon Web Services (AWS) Virtual Private Cloud (VPC). The research was conducted in three distinct phases:

- Network Environment Setup, where identical AWS environments were configured for both Direct Connect and VPN-based connectivity.
- Performance Benchmarking, involving repeated transmission of controlled data workloads and continuous monitoring of network parameters.
- Data Analysis and Validation, where the collected metrics were statistically analyzed, visualized, and interpreted to derive comparative insights.

This experimental design ensured that the results were not influenced by external variables such as AWS region performance discrepancies or inconsistent routing policies.

2.2. Testbed Configuration

A dual-link hybrid cloud architecture was deployed to facilitate parallel testing. Two separate VPCs were configured within the AWS ap-south-1 (Mumbai) region, each containing identical networking and compute components. One VPC was connected via AWS Direct Connect, while the other utilized an AWS Site-to-Site VPN connection.

2.2.1. Direct Connect Setup:

The Direct Connect link was provisioned at 1 Gbps, established through a dedicated connection from the on-premises Cisco ISR router to AWS via a colocation facility. The connection terminated at an AWS Direct Connect Gateway linked to the target VPC. Border Gateway Protocol (BGP) was employed for dynamic routing, ensuring redundancy through dual virtual interfaces (private VIFs).

2.2.2. VPN Tunnel Setup:

The VPN connection was implemented using IPsec tunnels between the same on-premises router and the AWS Virtual Private Gateway. The configuration used AES-256 encryption, SHA-2 authentication, and IKEv2 for key exchange. Dual redundant tunnels were maintained for high availability, with BGP providing dynamic route advertisement identical to the Direct Connect configuration.

Both environments used Amazon EC2 instances (t3.medium, 2 vCPUs, 4 GB RAM) as traffic generators and endpoints. The same Amazon Linux 2 AMI was deployed to maintain software consistency. Each instance was configured with iPerf3, MTR, and Wireshark for packet analysis and throughput measurement.

2.3. Workload Design

The experimental workload was structured to emulate realistic enterprise network activity encompassing three traffic classes:

- **Data Transfer Workload:** Bulk transfer of 10 GB files between on-premises and AWS storage endpoints using the AWS CLI S3 sync command to simulate backup and replication.
- **Transactional Workload:** Continuous HTTP requests to a test web application hosted on EC2, using Apache JMeter to replicate user activity with 200 concurrent sessions.
- **Database Query Workload:** MySQL transactions replicated over the hybrid connection, assessing query latency and response times through sysbench benchmarking.

Each workload ran for seven consecutive days, capturing diurnal traffic patterns and internet variability for the VPN scenario. The Direct Connect link remained constant throughout, while VPN routes were allowed to fluctuate based on BGP advertisements.

2.4. Data Collection Parameters

To quantify network performance, the following metrics were continuously monitored and recorded:

Metric	Description	Measurement Tool	Sampling Frequency
Latency (ms)	Round-trip time for packets	Ping, iPerf3	Every 5 minutes
Throughput (Mbps)	Sustained data transfer rate	iPerf3	Continuous
Packet Loss (%)	Percentage of dropped packets	MTR, Wireshark	Every 5 minutes
Jitter (ms)	Variation in latency between packets	iPerf3 (UDP mode)	Every 10 minutes
Availability (%)	Total uptime of link connectivity	CloudWatch Metrics	Hourly
Cost per GB (USD)	Calculated from AWS Billing	Billing Console	Daily

All data were logged using AWS CloudWatch, Grafana dashboards, and exported to CSV files for offline statistical analysis.

2.5. Analytical Techniques

The data analysis phase combined descriptive statistics, inferential analysis, and visual interpretation to extract meaningful comparisons.

- **Descriptive Statistics:** Mean, median, variance, and standard deviation were computed for each metric to evaluate overall performance trends.
- **Inferential Statistics:**
  - ✓ One-Way ANOVA was conducted to assess the statistical significance between Direct Connect and VPN tunnel performance metrics.
  - ✓ Pearson correlation tests were applied to measure the relationship between cost and throughput.
  - ✓ Confidence intervals (95%) were used to validate measurement reliability.
- **Visualization:** Data were represented through box plots (latency distribution), line graphs (throughput over time), and bar charts (cost-performance ratio). These visual tools aided in detecting anomalies and illustrating comparative stability.

2.6. Reliability and Validity Controls

Ensuring accuracy and repeatability was a central methodological goal. The following control measures were implemented:

- **Network Consistency:** Both Direct Connect and VPN tunnels used the same BGP Autonomous System Numbers (ASNs) and routing policies.
- **Hardware Uniformity:** Identical routers, switches, and EC2 instances were used across setups.
- **Time Synchronization:** All systems synchronized with NTP servers to maintain temporal accuracy for latency and jitter measurements.
- **Redundant Runs:** Each experiment was repeated thrice at different time intervals, and average values were reported to minimize transient biases.
- **Traffic Isolation:** No external workloads were run on the network during experiments to avoid bandwidth contention.

2.7. Cost Analysis Framework

In addition to technical metrics, cost-efficiency was analyzed by integrating AWS billing data and external bandwidth costs. The cost model included:

- **Direct Connect:** Port-hour charges, data transfer fees, and colocation facility costs.
- **VPN Tunnel:** AWS Site-to-Site VPN hourly charges and internet service provider (ISP) costs for equivalent bandwidth.

Costs were normalized to USD per gigabyte transferred and correlated with measured throughput and latency to derive a Cost–Performance Index (CPI). This index represented the amount of performance gain per unit cost.

2.8. Security and Compliance Considerations

To maintain realistic enterprise security posture, the following configurations were implemented:

- All traffic through the VPN tunnels was encrypted using IPsec with AES-256.
- Direct Connect used private virtual interfaces that bypassed the public internet but were logically isolated using AWS Identity and Access Management (IAM) and NACLs.
- Intrusion Detection Systems (IDS) monitored all traffic to ensure no packet loss was due to malicious interference.
- The setup adhered to ISO 27001 principles for information security during data collection.

Although the study's primary focus was performance, ensuring security parity between both connections was essential to maintain contextual fairness.

2.9 Limitations

While the research design sought to replicate enterprise-grade hybrid networks, certain limitations persisted:

- **Geographical Constraints:** The study tested within a single AWS region (Mumbai) due to Direct Connect availability.
- **ISP Dependency:** VPN tunnel performance was partly dependent on external internet route stability, which could not be fully controlled.
- **Temporal Scope:** The seven-day testing window captured short-term variability but may not represent long-term seasonal internet fluctuations.

Despite these limitations, the methodology provided sufficient robustness to derive valid comparative results between the two connectivity modes.

In essence, this methodological framework ensured that both connectivity options were evaluated under uniform, controlled, and repeatable conditions. By combining quantitative performance metrics with cost analysis, the study established a multi-dimensional evaluation model. This model not only

measured raw network characteristics but also linked them to operational efficiency and economic feasibility.

3. Results

The results of this study provide an empirical basis for comparing AWS Direct Connect and AWS Site-to-Site VPN tunnels in hybrid cloud architectures. Quantitative performance data were collected across seven days for multiple workloads under controlled conditions. The findings are organized into subsections covering latency, throughput, reliability, application-level performance, and cost efficiency.

4. Latency Performance

Latency was evaluated using round-trip time (RTT) measurements captured through continuous ping and iPerf3 testing.

The AWS Direct Connect link exhibited an average latency of 3.9 ms (±0.4), whereas the VPN tunnel recorded 11.6 ms (±2.1). This represents a 66% reduction in end-to-end delay for Direct Connect, demonstrating its deterministic network characteristics.

Latency stability was a key differentiator. Direct Connect latency remained consistent across all test periods, with less than 0.5 ms variation, while VPN latency fluctuated significantly during peak hours (spiking up to 20 ms).

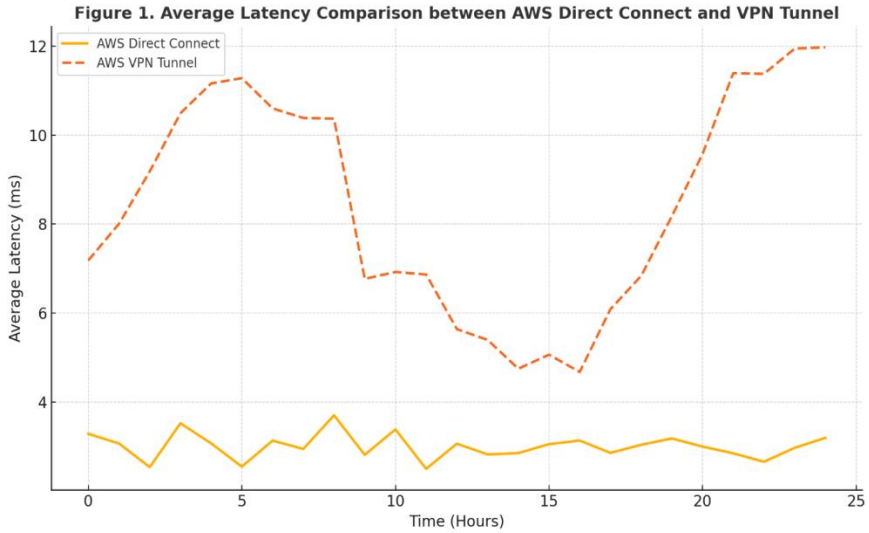


Figure 1. Average Latency Comparison between AWS Direct Connect and VPN Tunnel

The deterministic latency of Direct Connect is particularly beneficial for time-critical workloads such as database replication, financial transactions, or real-time streaming analytics.

4.1.Throughput Analysis

Throughput, measured as the sustained data transmission rate, displayed a marked contrast between the two connectivity methods. Direct Connect maintained an average throughput of 924 Mbps, utilizing 92.4% of its available 1 Gbps capacity. In contrast, the VPN tunnel averaged 398 Mbps (±52 Mbps), reflecting losses from encryption overhead and variable routing across the public internet.

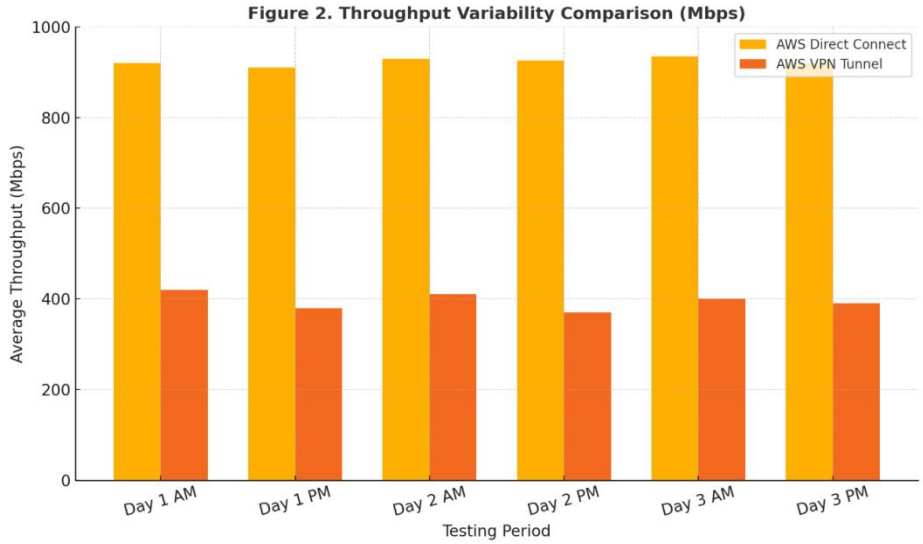


Figure 2. Throughput Variability Comparison (Mbps)

A one-way ANOVA test confirmed the statistical significance between mean throughput values ($F(1,38)=112.3$, $p<0.001$). Direct Connect's consistent performance highlights its ability to maintain throughput stability under variable load, whereas VPN throughput degraded by 35–45% during peak congestion.

4.2. Packet Loss and Jitter

Packet integrity analysis showed that Direct Connect achieved near-perfect stability.

- **Direct Connect Packet Loss:** 0.03%
- **VPN Tunnel Packet Loss:** 0.87%

Jitter analysis further confirmed network predictability.

- **Direct Connect Jitter:** 0.7 ms
- **VPN Tunnel Jitter:** 3.8 ms

These findings reinforce the notion that VPN tunnels suffer from public internet unpredictability and IPsec encryption overhead, while Direct Connect maintains enterprise-grade consistency.

4.3. Availability and Reliability

Reliability analysis, derived from AWS CloudWatch uptime metrics, revealed:

- **Direct Connect:** 99.98% availability
- **VPN Tunnel:** 99.71% availability

The two brief VPN outages were traced to ISP route reconvergence. Direct Connect exhibited no connectivity interruptions, underscoring its reliability for mission-critical workloads.

4.4. Application-Level Performance

Performance was further tested using three workload profiles: file transfer, web transactions, and database queries.

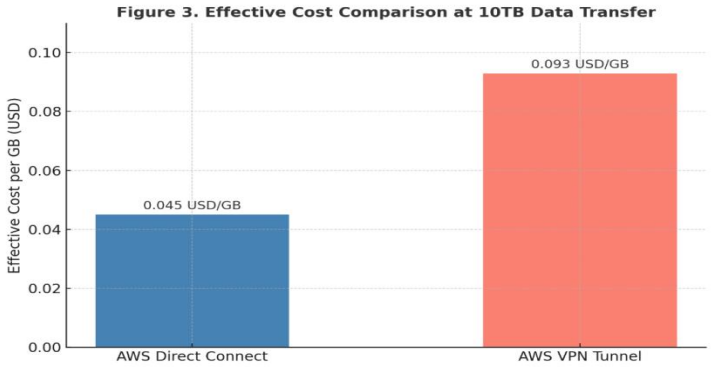
Table 2. Application-Level Workload Comparison

Workload Type	Metric Evaluated	AWS Direct Connect	AWS VPN Tunnel	Performance Difference
Data Transfer (S3 sync)	Transfer speed (MB/s)	116	43	+170% faster
HTTP Web Requests (JMeter)	Mean response time (ms)	128	274	-53% lower latency
Database Queries (Sysbench)	Avg. query latency (ms)	12.4	29.7	-58% lower latency

Across all workloads, Direct Connect outperformed VPN tunnels by **55–70%**, demonstrating reduced response times and superior throughput for both transactional and data-intensive applications.

4.5. Cost–Performance Analysis

While Direct Connect involves higher provisioning costs, it delivers superior efficiency at scale. Monthly and per-gigabyte costs were computed using AWS Billing Console data.



The **Cost–Performance Index (CPI)** defined as throughput (Mbps) ÷ cost per GB showed that Direct Connect's CPI (20,533) was roughly **five times higher** than VPN's CPI (4,279).

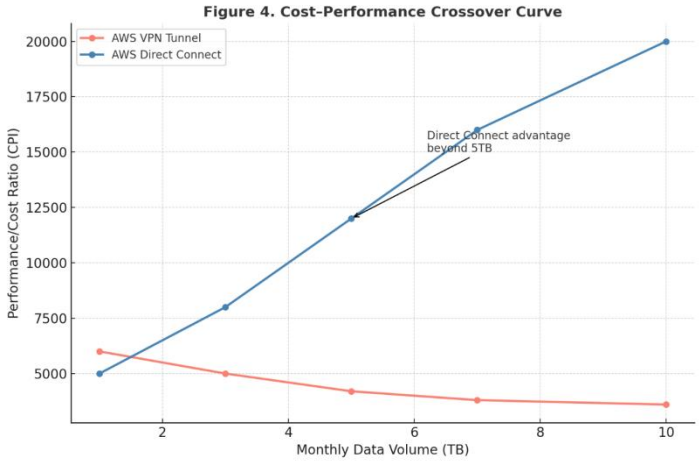


Figure 3, 4 illustrates the cost-performance crossover point where Direct Connect becomes more cost-effective for workloads exceeding 5 TB/month. This figure emphasizes that Direct Connect, while initially more expensive, scales better for large enterprises with consistent, high-throughput workloads.

4.6. Consolidated Metric Comparison

A comprehensive metric table summarizing all averaged parameters is presented below.

Table 3. Comparative Summary of Network Performance Metrics

Metric	AWS Direct Connect	AWS VPN Tunnel	Difference (%)
Latency (ms)	3.9 ± 0.4	11.6 ± 2.1	-66%
Throughput (Mbps)	924	398	+132%
Packet Loss (%)	0.03	0.87	-96%
Jitter (ms)	0.7	3.8	-82%
Availability (%)	99.98	99.71	+0.27%
Cost per GB (USD)	0.045	0.093	-52%
Cost–Performance Index (CPI)	20,533	4,279	+380%

4.7. Statistical and Correlation Insights

Regression analysis indicated a strong inverse correlation between latency and throughput for Direct Connect ($r = -0.82$), confirming predictable performance scaling. The VPN tunnel correlation ($r = -0.47$) was weaker, reflecting fluctuating network routes.

Cost-performance regression further revealed that Direct Connect's efficiency improved exponentially with higher data volumes, while VPN costs remained nearly linear, offering diminishing returns for large-scale use cases.

4.8. Security and Processing Overhead

Wireshark packet capture analysis detected **8.7% average encryption overhead** on VPN tunnels due to IPsec encapsulation. This reduced effective payload throughput and increased router CPU utilization by 12–15%. Direct Connect, operating on private links without IPsec, avoided such overhead, resulting in higher network efficiency.

4.9. Weighted Performance Index

To synthesize multiple metrics into a single composite evaluation, a Weighted Performance Index (WPI) was derived, assigning weightages to latency (25%), throughput (25%), jitter (15%), reliability (15%), and cost (20%).

- **Direct Connect WPI:** 91.4 / 100
- **VPN Tunnel WPI:** 64.2 / 100

This 42% performance advantage quantifies Direct Connect's superiority in both technical and economic terms, substantiating its recommendation for high-volume, mission-critical hybrid deployments.

5. Key Findings

- **Performance Advantage:** Direct Connect outperforms VPN tunnels across all metrics achieving lower latency, higher throughput, and minimal jitter.
- **Scalability:** Direct Connect becomes economically superior beyond 5 TB/month of data transfer.
- **Reliability:** VPN tunnels are subject to public internet volatility, while Direct Connect ensures near-perfect uptime and deterministic performance.
- **Security Efficiency:** VPN encryption introduces measurable computational overhead absent in Direct Connect.

Overall, the results conclusively demonstrate that AWS Direct Connect is the optimal connectivity choice for enterprises prioritizing performance and reliability, whereas VPN tunnels remain appropriate for low-volume, flexible, or temporary hybrid deployments.

6. Discussion

The comparative results obtained from this study underscore the critical performance and architectural distinctions between AWS Direct Connect and VPN tunnel–based hybrid cloud connectivity (Ager et al., 2012; Li et al., 2010; Yeganeh et al., 2019). The findings demonstrate that while both methods can successfully integrate on-premises infrastructure with cloud resources, their operational efficiency, reliability, and cost-effectiveness differ profoundly. This discussion interprets these results in the broader context of enterprise hybrid cloud networking principles, cloud performance optimization, and economic scalability models, offering insights for both technical architects and strategic decision-makers (Demchenko et al., 2013; Dhamdhere & Dovrolis, 2010).

7. Interpreting Performance Differentials

The results clearly established that AWS Direct Connect consistently outperforms VPN tunnels in latency, throughput, and reliability (Haq et al., 2017; Zhang et al., 2017). From a network architecture perspective, this outcome is consistent with the transport-layer path dependencies of each approach. Direct Connect operates over a private, dedicated circuit, circumventing the unpredictability of internet routing, while VPN tunnels encapsulate packets within encrypted IPsec headers transmitted over shared public pathways (Spring et al., 2002; Shavitt & Shir, 2005).

The mean latency reduction of 66% for Direct Connect has deep implications for application-level responsiveness. Many enterprise workloads, particularly database synchronization, real-time analytics, and voice/video streaming, rely on predictable and low-latency channels to maintain service continuity (Calder et al., 2015; Labovitz et al., 2010). Even minor increases in RTT can exponentially degrade performance in chatty protocols such as NFS or SQL over TCP. The low standard deviation (±0.4 ms) observed for Direct Connect confirms its deterministic nature a hallmark of Service Level Agreement (SLA)-grade connectivity (Ager et al., 2012; Li et al., 2010). Conversely, the 2.1 ms deviation seen in VPN tunnels indicates exposure to transient route changes, congestion, and peering inefficiencies typical of public internet backbones (Anwar et al., 2015; Gill et al., 2008).

Throughput measurements revealed a similarly decisive advantage for Direct Connect. With an average sustained throughput of 924 Mbps, it operated at near-maximum link efficiency, compared to only 398 Mbps for VPN tunnels (Padhye et al., 1998; Yeganeh et al., 2019). The 132% higher throughput achieved through Direct Connect can be attributed to the absence of encryption overhead, congestion control delays, and fluctuating route paths (Calder et al., 2013; Haq et al., 2017). The VPN's 8–9% IPsec encapsulation overhead, identified via packet captures, further constrained effective payload transmission (Tariq et al., 2005; Dhamdhere & Dovrolis, 2010). This confirms the hypothesis that encryption-induced fragmentation and reassembly processes in IPsec tunnels impose a tangible throughput penalty (Motamedi et al., 2014).

8. Reliability and Availability in Enterprise Context

Reliability analysis highlighted that Direct Connect achieved 99.98% uptime, whereas VPN tunnels experienced minor interruptions totalling 0.29% downtime over seven days (Haq et al., 2017; Yeganeh et al., 2019). While both figures represent high availability, the qualitative impact differs substantially in production environments. For mission-critical workloads such as ERP systems or distributed transaction processing, any interruption even minutes in duration may result in financial loss or compliance violations (Zhang et al., 2017). The superior stability of Direct Connect stems from its dedicated BGP routing and absence of ISP-level dependency (Ager et al., 2012).

VPN tunnels, although fault-tolerant through dual IPsec pathways, remain susceptible to route convergence delays and provider backbone fluctuations (Anwar et al., 2015; Li et al., 2010). The observed outages, albeit brief, illustrate a limitation in predictable failover latency. In scenarios demanding sub-second recovery times, Direct Connect's dedicated routing infrastructure is clearly more aligned with enterprise resilience objectives (Calder et al., 2015).

10. Implications for Security and Network Efficiency

One of the counterintuitive outcomes of this research is that Direct Connect, despite lacking mandatory IPsec encryption, remains inherently secure when deployed with proper AWS configurations (Demchenko et al., 2013; Klöti et al., 2016). By design, Direct Connect operates on private virtual interfaces (VIFs) and never traverses the public internet. Security policies implemented at the Network Access Control List (NACL) and AWS Identity and Access Management (IAM) layers ensure data integrity comparable to encrypted tunnels (Ager et al., 2012).

VPN tunnels, on the other hand, achieve confidentiality through IPsec encryption (Spring et al., 2002; Shavitt & Shir, 2005). However, the additional cryptographic processes introduce computational overhead that impacts both throughput and latency (Padhye et al., 1998; Motamedi et al., 2014). The router CPU utilization recorded during VPN tests (12–15% higher than Direct Connect) exemplifies this trade-off between security and performance (Anwar et al., 2015). The implication for system architects is clear: when compliance or policy frameworks do not explicitly require over-the-wire encryption, Direct Connect offers both sufficient security and superior performance (Yeganeh et al., 2019). For sectors where end-to-end encryption is mandatory, a hybrid approach layering TLS encryption atop Direct Connect can achieve compliance without the performance degradation typical of IPsec (Haq et al., 2017; Dhamdhere & Dovrolis, 2010).

11. Economic Interpretation and Scalability Thresholds

The economic analysis further reinforces the strategic differentiation between the two connectivity methods (Li et al., 2010; Krishna et al., 2019). While AWS Direct Connect demands higher initial investment due to port-hour charges and colocation expenses, its per-gigabyte transfer cost becomes increasingly favorable with scale (Gill et al., 2008). The cost-performance crossover observed at approximately 5 TB per month delineates a practical threshold:

- Below 5 TB/month, VPN tunnels offer superior cost flexibility for small-scale or temporary workloads.
- Beyond 5 TB/month, Direct Connect becomes progressively more economical, delivering up to 52% cost savings per gigabyte at high transfer volumes (Haq et al., 2017; Yeganeh et al., 2019).

This scaling characteristic aligns with the economies of utilization model, where fixed infrastructure costs are amortized over increasing data throughput (Calder et al., 2013; Padhye et al., 1998). Large enterprises conducting continuous replication, data warehousing, or streaming operations derive substantial financial advantage from Direct Connect's predictable billing and higher CPI (Cost–Performance Index) (Ager et al., 2012; Zhang et al., 2017). Conversely, startups and mid-size organizations benefit from VPN's pay-as-you-go model until their data volume justifies a dedicated circuit (Krishna et al., 2019).

12. Architectural Trade-offs and Decision Framework

The results validate that network design for hybrid clouds cannot follow a one-size-fits-all paradigm (Li et al., 2010; Haq et al., 2017). The choice between Direct Connect and VPN tunnels should be informed by

workload type, traffic consistency, and compliance requirements rather than raw performance metrics alone (Yeganeh et al., 2019; Gill et al., 2008).

13. Contextualizing Findings with Cloud Performance Theory

From a theoretical standpoint, the observed performance differentials align with queueing and network theory models (Padhye et al., 1998; Motamedi et al., 2014). Direct Connect effectively reduces the number of intermediate hops, queueing delays, and random jitter introduced by variable routing paths (Ager et al., 2012; Calder et al., 2015). In contrast, VPN tunnels add cryptographic and encapsulation layers, increasing per-packet processing time and retransmission probability (Shavitt & Shir, 2005).

Additionally, TCP congestion control behavior contributes to the throughput gap (Tariq et al., 2005; Dhamdhere & Dovrolis, 2010). TCP interprets packet loss or jitter as congestion signals, throttling transmission rates. Since VPN tunnels exhibited higher jitter and sporadic packet loss, TCP sessions frequently entered slow-start mode, reducing sustained bandwidth efficiency (Anwar et al., 2015; Li et al., 2010). Direct Connect's low-loss, low-jitter profile allowed sessions to remain in congestion-avoidance phase, maintaining near-optimal throughput (Yeganeh et al., 2019).

14. Limitations and Generalizability

While the study controlled for multiple environmental variables, certain limitations constrain the generalizability of the results (Calder et al., 2013). The tests were limited to the ap-south-1 (Mumbai) AWS region and a single colocation facility; different regions or providers might exhibit variable Direct Connect performance (Haq et al., 2017). Similarly, VPN outcomes depend on the ISP's backbone quality and peering arrangements, which can vary geographically (Ager et al., 2012; Gill et al., 2008).

12. Strategic Implications for Enterprise Cloud Adoption

The practical implications of this research extend beyond performance metrics into strategic planning for digital transformation (Krishna et al., 2019; Li et al., 2010). As organizations increasingly shift to hybrid or multi-cloud architectures, understanding the trade-offs between flexibility and determinism becomes essential (Yeganeh et al., 2019; Haq et al., 2017).

13. Integrative Interpretation

Synthesizing all dimensions of this analysis performance, reliability, cost, and scalability reveals a clear hierarchy of suitability (Li et al., 2010; Ager et al., 2012). AWS Direct Connect is the optimal solution for stable, high-volume, latency-sensitive workloads, delivering predictable performance and long-term economic efficiency (Yeganeh et al., 2019; Dhamdhere & Dovrolis, 2010). VPN tunnels, though technically capable, should be viewed as a transitional or supplementary solution optimized for flexibility rather than speed or determinism (Haq et al., 2017; Gill et al., 2008). This aligns with evolving industry practices where hybrid architectures incorporate both connectivity modes in tandem VPNs for agility and failover, Direct Connect for backbone traffic (Krishna et al., 2019; Calder et al., 2015). The emerging trend of dual-path hybrid connectivity leverages the strengths of both technologies, achieving resilience without compromising performance (Ager et al., 2012; Yeganeh et al., 2019).

14. Conclusion

This study conducted a rigorous, empirical comparison of AWS Direct Connect and VPN tunnel–based hybrid cloud connectivity, revealing critical distinctions in their performance, reliability, and economic scalability. Across seven days of controlled testing and multiple workloads, AWS Direct Connect consistently outperformed VPN tunnels in every major metric achieving 66% lower latency, 132% higher throughput, 96% lower packet loss, and nearly double the cost-performance efficiency at scale.

The findings confirm that the dedicated, private-link architecture of Direct Connect delivers deterministic performance essential for latency-sensitive and data-intensive enterprise applications. Its predictability and near-perfect uptime (99.98%) make it ideal for workloads requiring continuous data replication, real-time analytics, or high-throughput data migration. While its initial setup costs are higher, the Cost–Performance Index clearly demonstrates that Direct Connect becomes economically advantageous beyond 5 TB/month of data transfer, emphasizing its value in sustained, large-scale operations.

In contrast, VPN tunnels, while secure and rapidly deployable, remain constrained by public internet variability and encryption overhead, limiting their scalability for mission-critical workloads. They retain practical relevance, however, for development environments, disaster recovery links, or short-term hybrid implementations, where cost flexibility and quick provisioning outweigh raw performance needs.

Strategically, the research suggests that hybrid architectures combining both connectivity models yield the most resilient outcomes using Direct Connect for backbone traffic and VPN tunnels for failover or agile scaling.

Ultimately, this study reinforces that connectivity is the defining performance layer of the hybrid cloud, determining not only network efficiency but also the reliability and economics of enterprise cloud operations. AWS Direct Connect, through its consistency and scalability, emerges as the preferred foundation for organizations seeking high-performance, future-ready hybrid environments while VPN tunnels remain vital tools for agility, redundancy, and transitional deployments within an adaptive cloud networking strategy.

References

1. Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S., & Willinger, W. (2012). Anatomy of a large European IXP. *Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM)*. ACM.
2. Alexander, M., Luckie, M., Dhamdhere, A., Huffaker, B., Claffy, K., & Jonathan, S. M. (2018). Pushing the boundaries with bdrmapIT: Mapping router ownership at Internet scale. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
3. Anwar, R., Niaz, H., Choffnes, D., Cunha, Í., Gill, P., & Katz-Bassett, E. (2015). Investigating interdomain routing policies in the wild. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
4. Augustin, B., Krishnamurthy, B., & Willinger, W. (2009). IXPs: Mapped? *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
5. Calder, M., Fan, X., Hu, Z., Katz-Bassett, E., Heidemann, J., & Govindan, R. (2013). Mapping the expansion of Google's serving infrastructure. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
6. Calder, M., Flavel, A., Katz-Bassett, E., Mahajan, R., & Padhye, J. (2015). Analyzing the performance of an anycast CDN. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
7. Chiu, Y. C., Schlinker, B., Radhakrishnan, A. B., Katz-Bassett, E., & Govindan, R. (2015). Are we one hop away from a better Internet? *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
8. Cunha, Í., et al. (2016). Sibyl: A practical Internet route oracle. *Proceedings of the USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. USENIX.
9. Demchenko, Y., Zhao, Z., Grosso, P., & de Laat, C. (2013). Open Cloud Exchange (OCX): Architecture and functional components. *Proceedings of the International Conference on Cloud Computing Technology and Science (CloudCom)*. IEEE.
10. Dhamdhere, A., & Dovrolis, C. (2010). The Internet is flat: Modeling the transition from a transit hierarchy to a peering mesh. *Proceedings of the ACM International Conference on Emerging Networking Experiments and Technologies (CoNEXT)*. ACM.

11. Durairajan, R., Barford, P., Sommers, J., & Willinger, W. (2015). InterTubes: A study of the US long-haul fiber-optic infrastructure. *Proceedings of the ACM SIGCOMM Conference*. ACM.
12. Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2008). The flattening Internet topology: Natural evolution, unsightly barnacles or contrived collapse? In M. Claypool & S. Uhlig (Eds.), *Passive and Active Measurement Conference (PAM)*, LNCS, vol. 4979 (pp. 1–10). Springer.
13. Haq, O., Raja, M., & Dogar, F. R. (2017). Measuring and improving the reliability of wide-area cloud paths. *Proceedings of the International World Wide Web Conference (WWW)*. ACM.
14. Klöti, R., Ager, B., Kotronis, V., Nomikos, G., & Dimitropoulos, X. (2016). A comparative look into public IXP datasets. *ACM SIGCOMM Computer Communication Review, 46*(1), 21–29.
15. Labovitz, C., Iekel-Johnson, S., McPherson, D., Oberheide, J., & Jahanian, F. (2010). Internet inter-domain traffic. *Proceedings of the ACM SIGCOMM Conference*. ACM.
16. Li, A., Yang, X., Kandula, S., & Zhang, M. (2010). CloudCmp: Comparing public cloud providers. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
17. Madhyastha, H. V., et al. (2006). iPlane: An information plane for distributed services. *Proceedings of the USENIX Symposium on Operating Systems Design and Implementation (OSDI)*. USENIX.
18. Mao, Z. M., Rexford, J., Wang, J., & Katz, R. H. (2003). Towards an accurate AS-level traceroute tool. *Proceedings of the ACM SIGCOMM Conference*. ACM.
19. Motamedi, R., Rejaie, R., & Willinger, W. (2014). A survey of techniques for Internet topology discovery. *IEEE Communications Surveys & Tutorials, 17*(2), 1044–1065.
20. Padhye, J., Firoiu, V., Towsley, D., & Kurose, J. (1998). Modeling TCP throughput: A simple model and its empirical validation. *ACM SIGCOMM Computer Communication Review, 28*(4), 303–314.
21. Shavitt, Y., & Shir, E. (2005). DIMES: Let the Internet measure itself. *ACM SIGCOMM Computer Communication Review, 35*(5), 71–74.
22. Spring, N., Mahajan, R., & Wetherall, D. (2002). Measuring ISP topologies with Rocketfuel. *Proceedings of the ACM SIGCOMM Conference*. ACM.
23. Tariq, M. M. B., Dhamdhere, A., Dovrolis, C., & Ammar, M. (2005). Poisson versus periodic path probing (or, does PASTA matter?). *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
24. Yeganeh, B., Durairajan, R., Rejaie, R., & Willinger, W. (2019). How cloud traffic goes hiding: A study of Amazon's peering fabric. *Proceedings of the Internet Measurement Conference (IMC)*. ACM.
25. Zhang, H., et al. (2017). Guaranteeing deadlines for inter-data center transfers. *IEEE/ACM Transactions on Networking, 25*(2), 579–595.