IDC

SPOTLIGHT

Sponsored by: Riverbed

As AI adoption and cloud complexity increase, organizations face growing pressure to move large data sets faster, more reliably, and at lower cost across distributed infrastructure.

Enterprise Data Movement in the AI Era: Turning a Critical Bottleneck into a Strategic Advantage

June 2026

Written by:
Paul Nicholson, Research Vice President, Cloud and Datacenter Networks, and Mark Leary, Research Director, Network Analytics and Automation

Introduction

The rapid expansion of AI workloads is reshaping the demands placed on enterprise data infrastructure. As organizations deploy AI models at scale, they face a structural challenge with data gravity: the compute resources required to train and run those models (GPU-dense AI factories, often concentrated where power costs are lowest) are geographically separated from the enterprise data that feeds them. Data stored in ERP systems, CRM platforms, supply chain applications, point-of-sale environments, and edge locations must be moved to where it can be processed. The data must move to the model; the model cannot move to the data.

This creates a new class of infrastructure pressure, a challenge not solved just with bandwidth. IDC research confirms that 91% of enterprises using on-premises datacenters expect inter-datacenter bandwidth to grow by 11% or more within the next year due to AI workload deployments, with 36% anticipating growth of over 51% (source: IDC's 2025 Worldwide AI in Networking Special Report). Yet simply adding bandwidth does not solve the issue. It adds complexity: more to manage, more routing decisions, more failure points, and more demand on already-stretched IT teams. The opportunity lies in rethinking how existing and new infrastructure is used, extracting maximum efficiency for data in transit, before defaulting to raw capacity additions.

AT A GLANCE

KEY STATS

» AI workloads are driving sharp increases in bandwidth demand between datacenters. IDC research shows that 91% of organizations will add 11% or more inter-datacenter bandwidth within the next year due to AI, with 36% planning increases of more than 51% (source: IDC's 2025 *Worldwide AI in Networking Special Report)*.

» Data movement is increasingly a strategic capability, not a back-office function. Organizations that can relocate large data sets quickly gain direct advantages in AI model training timelines, cloud migration execution, and business continuity response.

» A managed, SaaS-delivered approach to enterprise data movement eliminates hardware procurement, multiteam coordination, and extended planning cycles, enabling organizations to initiate large-scale transfers quickly and redirect IT capacity toward higher-value priorities.

The challenge is not confined to AI. Cloud migrations with hard business deadlines, mergers and acquisitions requiring rapid data consolidation, and ongoing disaster recovery scenarios demanding urgent data relocation (illustrated by the experience of organizations in the Middle East that needed to rapidly move petabyte-scale data sets from one cloud region to another following infrastructure disruptions due to the Middle East War) all place similar demands on enterprise data movement capabilities. Across all these scenarios, organizations identify pain points: speed, operational simplicity, and cost.

At the same time, AI investment is under increasing pressure to demonstrate value. IDC finds that while more than 54% of enterprises report that over half of their AI projects have delivered measurable business outcomes, scaling those outcomes and demonstrating ongoing value to the business remain the two top challenges to realizing the full potential of AI investments (source: IDC's February 2026 Future Enterprise Resiliency and Spending Survey [FERS Survey] Wave 1). A repeatable, reliable data supply chain, one that consistently delivers fresh, accurate training data to AI infrastructure, is a foundational requirement for moving AI from pilot to production at scale.

Benefits: Strategic value of modern enterprise data movement

Leading organizations now recognize that data movement is no longer a back-end operational task; it is a strategic capability. The ability to move data quickly, reliably, and cost-effectively between clouds, cloud regions, and datacenters directly determines the success of an organization's AI models, enabling cloud migration benefits, responses to business continuity events, and realizing value from distributed multicloud investments.

IDC identifies the following as strategic dimensions organizations should consider when evaluating enterprise data movement capabilities:

- **Enabling AI at scale with repeatable data pipelines:** AI model accuracy depends on the currency of training data. Organizations that rely on ad hoc, manual data movement approaches will consistently struggle to keep models current. IDC research shows that scaling AI outcomes and demonstrating ongoing business value are the two top challenges organizations face after initial deployment (source: IDC's February 2026 *Future Enterprise Resiliency and Spending Survey* Wave 1). A reliable, automated, and repeatable data movement capability is a prerequisite for closing the gap between AI experimentation and production-grade AI delivery.
- **Reducing the cost and complexity of cloud migration:** Cloud migration projects can quickly exceed planned timelines, in part because data movement tooling is underspecified relative to the volume being transferred. Extended migration windows increase costs because organizations must pay for both source and destination environments simultaneously during the overlap period, a period of dual expenditure known as the "double bubble" that grows longer when transfer speeds are inadequate. Accelerating migration completion shortens this window and brings forward efficiency and capability gains that motivated migration.
- **Supporting resilience with data mobility:** The ability to rapidly relocate petabyte-scale data sets across cloud providers and geographic regions is a genuine resilience capability, not only a technical convenience. Organizations that can execute urgent data relocations in response to infrastructure disruptions, geopolitical events, or regulatory requirements have a materially stronger business continuity posture than those that depend on pre-provisioned, static replicas.

- **Reducing cloud egress cost exposure:** Cloud egress fees represent one of the most controllable yet frequently underestimated cost variables in hybrid and multicloud operations. At petabyte scale, egress charges accumulate rapidly and are proportional to the volume of data transmitted. Data optimization before transmission, reducing the effective volume through compression and deduplication, results in a direct proportional reduction in egress charges, representing material ongoing cost savings for workloads that require recurring large-scale transfers.
- **Simplifying operations and reallocating IT capacity:** Managing data movement with DIY approaches requires coordinating multiple teams across source and destination environments, integrating disparate tools, and maintaining that integration over time. This operational burden consumes skilled engineering capacity that organizations are simultaneously being pressured to redirect toward AI development, application modernization, and other higher-value initiatives. IDC finds that infrastructure optimization is among the most protected spending priorities in current enterprise budgets, reflecting broad recognition that operational efficiency enables the delivery of strategic objectives source: IDC's February 2026 *Future Enterprise Resiliency and Spending Survey* Wave 1).
- **Optimizing before expanding:** The instinctive response to slow data transfers, procuring additional bandwidth, addresses only part of the problem while introducing its own additional complexity. Applying compression and deduplication before transmission as a first step can dramatically reduce the effective data volume in transit, allowing existing infrastructure to achieve higher throughput without proportional capacity, cost, and complexity additions. Organizations that optimize transfer efficiency first will find that bandwidth requirements are often lower than initial estimates, and that the infrastructure complexity of adding capacity can be deferred or avoided entirely.

Vendor profile: Considering Riverbed Data Express

Riverbed has positioned its Data Express product as a SaaS-delivered, high-speed data movement service designed specifically for the enterprise data transfer requirements described in the preceding sections. The service targets the four primary use cases that can drive enterprise demand for accelerated data movement: cloud migration, mergers and acquisitions, disaster recovery and business resilience, and AI data pipelines.

Riverbed Data Express is positioned as a managed service alternative to DIY data movement workflows. Where conventional approaches require organizations to assemble and integrate multiple tools, including transfer utilities, compression software, monitoring systems, verification tooling, and compliance tracking, Data Express seeks to consolidate these capabilities into a single SaaS offering. According to Riverbed, if an organization has credentials and permission to access the data, Data Express can be set up and operational in minutes, with integration available via REST APIs and agentic controls.

Product highlights

- **Riverbed Data Express Service:** The core offering is a fully managed SaaS service designed to accelerate the movement of massive data sets between public clouds, cloud regions, and on-premises datacenters. According to Riverbed, Data Express transfers data at rates of 5–40TB/hour and can move up to 1PB/day, which is described as up to 10 times faster than conventional data movement methods. The service applies compression and deduplication before transmission, which Riverbed states can reduce data volume by up to 90% for typical enterprise data sets, thereby reducing both transfer time and cloud egress charges. According to Riverbed, what

took months can now be completed in days. The service supports AWS and Oracle Cloud Infrastructure (OCI) and their respective regions, with additional hyperscaler and datacenter capabilities planned for the near future. Riverbed describes the data transfer process as requiring only permissions, source identification, destination selection, job creation, and execution.

Key features and capabilities

- **High-speed, petabyte-scale transfers:** Riverbed states that Data Express is designed to handle enterprise-scale workloads, with transfer rates substantially higher than native cloud transfer tools or open source utilities. The service is positioned to address scenarios where migration timelines, AI training cadence, or DR response time requirements cannot be met with conventional approaches.
- **Compression and deduplication before transmission:** Data is compressed and deduplicated at the source before it crosses the network. The service builds what Riverbed describes as a dynamic Data Fabric that scales to hundreds of gigabits per second (Gbps) of network capacity, with egress fee billing applied only to the net data volume that crosses the network rather than the logical source volume. This approach can substantially reduce cloud egress charges by a percentage equivalent to the data reduction ratio achieved.
- **SaaS delivery with minimal operational overhead:** Data Express requires no hardware procurement, no virtual machine deployment, and no software installation. Riverbed positions the service simple enough to be operable by one or two staff members for projects that conventional DIY approaches would require multiple specialist teams spread across source cloud, destination cloud, networking, and compliance functions to manage over months.
- **Post-quantum cryptography (PQC) security:** For advanced encryption, the service protects data in transit through PQC-secure tunnels, addressing enterprise requirements for end-to-end data security and supporting compliance with data sovereignty and privacy regulations.
- **Automated verification and chain-of-custody logging:** The service includes automated verification of transfers and chain-of-custody logging that provides auditability without the need for additional compliance tooling. A unified portal provides real-time visibility into throughput, job status, and transfer trends.
- **Multicloud and multiregion support: Data Express** supports transfers between public cloud providers and across cloud regions, as well as between cloud and on-premises environments. Riverbed states that the transfer flexibility supports organizations in navigating data residency requirements, geopolitical risk, and multicloud optimization strategies without locking data into a specific provider or region.
- **Consumption-based pricing:** The pricing model is consumption based and aligned with cloud economics, with charges applied to the actual data volume moved and only the net egress fees; with volume discounts available depending on commitment.

Taken together, Data Express is positioned to address the full spectrum of enterprise data movement requirements, from AI data pipelines that require consistent throughput to onetime migrations that demand speed and auditability, within a single managed service. The combination of high transfer rates, source-side optimization, SaaS delivery, and consumption-based pricing reflects a design approach oriented toward reducing both the time and operational overhead of large-scale data movement.

Challenges

Riverbed Data Express is a relatively recent standalone SaaS offering, and organizations should consider its stated capabilities alongside a few practical considerations.

Performance results will vary with data type, network conditions, and environment configuration. Compression and deduplication ratios depend heavily on data content: structured data and logs typically compress well, while already-compressed formats such as video, images, or encrypted files achieve lower reduction ratios. Organizations should validate expected outcomes against their specific workloads before building egress cost projections around a particular reduction ratio.

Organizations with existing data movement toolchains or established Riverbed deployments should assess the integration path. The SaaS model removes hardware overhead but introduces a dependency on Riverbed's service infrastructure, and IT teams will need to configure and maintain the access credentials and cloud permissions that enable Data Express to reach the source and destination environments.

As with all vendors in rapidly evolving technology spaces, Riverbed will need to continue developing Data Express capabilities as enterprise data movement requirements, including sovereign datacenter destinations, emerging cloud providers, and expanding data governance requirements continue to evolve. If Riverbed continues to invest in the service's breadth and depth of capabilities, it will be well placed to address a growing range of enterprise data movement requirements.

Conclusion

The convergence of AI workload growth, cloud migration complexity, geopolitical risk, and multicloud strategy execution has elevated enterprise data movement from a back-end operational concern to a strategic infrastructure priority. Organizations that can move large data sets quickly, reliably, and cost-effectively across clouds, regions, and datacenters are better positioned to operationalize AI at scale, complete cloud transitions on schedule, and respond to business continuity events without unacceptable disruption.

The bandwidth growth that IDC research identifies, with 91% of enterprise datacenters expecting to add at least 11% inter-datacenter bandwidth within the year and 36% expecting growth above 51%, reflects both the scale of the challenge and the risk of treating it as a capacity-only problem. Adding bandwidth without optimizing its use adds cost and complexity without proportionally improving outcomes. The organizations that address both dimensions, optimizing transfer efficiency and scaling capacity where genuinely needed, will achieve better results than those that rely exclusively on either approach.

Riverbed Data Express addresses the efficiency dimension of this challenge through compression, deduplication, and managed transfer acceleration delivered as a SaaS service. The positioning as a consumption-based managed service with minimal setup overhead is well-suited to organizations that need to initiate large-scale data movement quickly, reduce the operational burden on IT teams, and control egress costs without adding infrastructure complexity. Organizations evaluating data movement solutions should assess Riverbed Data Express alongside the specific data types, cloud environments, and transfer requirements relevant to their use cases.

About the analysts

Paul Nicholson, *Research Vice President , Cloud and Datacenter Networks*

Paul Nicholson is IDC's Research Vice President, Cloud and Datacenter Networks. He provides thought leadership and actionable insights on Cloud and Datacenter Networking markets and technologies. Paul has a deep understanding of the networking market along with its business and application requirements, technologies, product roadmaps, competitive differentiation, and go-to-market strategies, enabling him to provide informed guidance for vendors, cloud providers, enterprise IT buyers and practitioners.

Mark Leary, *Research Director, Network Analytics and Automation*

Mark Leary is IDC's Research Director within enterprise infrastructure global research domain. He covers Network Observability, Automation, and Artificial Intelligence as part of the Network Infrastructure and Services subdomain. Mark's research covers the advancement and adoption of network observability, the development of network automation capabilities, and AI-powered network solutions, engineering, and operations.

MESSAGE FROM THE SPONSOR

Data Movement as Powerful as AI Itself

Riverbed Data Express enables organizations to move massive volumes of data across clouds and regions, turning data mobility into a strategic advantage for migration, resilience, and AI. It removes the friction from large-scale data movement, delivering high-speed, secure, and predictable transfer of massive data sets across major cloud providers, their regions, and customer datacenters. Organizations can migrate data faster, build resilient multicloud architectures, and fuel AI with the data that matters, wherever it lives.

With Data Express, data is no longer something you relocate once and optimize around forever. It becomes portable, strategic, and continuously optimized. By replacing slow, complex data transfers with a secure, cloud-delivered data movement service, Data Express cuts project timelines from months to days, improving time to value while significantly lowering data transfer costs. The result is fast, reliable, and secure data delivery at the scale modern AI and multi-cloud environments demand.

Learn more at https://www.riverbed.com/products/data-express/

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
One Beacon Street
Suite 33100
Boston, MA 02108, USA
T 508.872.8200
F 508.935.4015
blogs.idc.com
www.idc.com