IBM STORAGE SCALE

Accelerating WAN data transfers with Bridgeworks PORTrockIT

BRIDGEWORKS

OCF

BRILLIANCE. TOGETHER.

EXECUTIVE SUMMARY

IBM Storage Scale, previously known as IBM Spectrum Scale, GPFS, MMFS and TigerShark, is a clustered file system that provides performance and scalability for data storage.

Whether your use case is aggregation and analysis, training AI models, or backup and recovery, data is only useful if you can access it. Many companies use IBM Storage Scale to access their data remotely, often across long distances. Distance introduces latency into the connection and increases the chance of packet loss. Both latency and packet loss significantly slow transfer rates, and can make day-to-day data access a painful process.

This paper is the result of joint testing by Bridgeworks and IBM Gold Business Partner OCF. It demonstrates how Bridgeworks PORTrockIT can solve the remote data access problem for IBM Storage Scale customers by significantly mitigating the effects of both latency and packet loss on WAN transfer jobs. Even in the most challenging scenario we tested, with 300ms of latency and 0.1% packet loss, PORTrockIT was able to complete backup jobs 50 times faster than a traditional network architecture.

PORTrockIT can help your business accelerate data transfers with IBM Storage Scale – keeping your data properly protected, reducing the risk of disruption to your business, and helping to ensure that you get the full value from your investment in high-speed network infrastructure.

> *"Even in the most challenging scenario we tested... **PORTrockIT was able to complete backup jobs 50 times faster** than a traditional network architecture."*

WHY SPEED MATTERS

Modern organisations depend on fast, predictable access to data. Whether for analytics, model training, replication, or backup, the right data must be in the right place at the right time. IBM Storage Scale provides multiple pathways to access your data, through Remote Cluster sharing, NFS, S3, SMB, or HDFS via Cluster Export Services (CES). Yet getting data to applications and users in a timely manner remains a challenge.

The rate at which data can be moved has a significant effect on the business. When data movement is slow, we must structure our routines around it, setting transfers to run overnight or during downtime. This does more than reduce productivity, it limits our ability to understand and respond to the insights

that data provides. Customers feel the impact too, because user-facing systems slow down when the data they depend on isn't available quickly enough.

IBM Storage Scale is also commonly used for data protection and replication, especially when the contents of a filesystem need to be backed up and replicated across a wide area network (WAN) to a remote server or cluster. When dealing with multiple sites, IBM Storage Scale clusters may sync data with each other, so each has a full copy in case of disaster. Additionally, in many cases, it's important to guarantee that the cluster remains available even during a failure. IBM Storage Scale offers Active File Management Disaster Recovery (AFM DR), which allows you to keep a hot standby constantly synchronised with your critical systems, ready to take over immediately in the event of a failure.

Whenever data is being backed up or replicated – and especially if the data needs to move across a WAN – transfer speed is a critical consideration. Slow backups are more than a technical issue for the IT department: they affect business users too. If overnight backup jobs overrun into working hours, users may find that transactional systems become less responsive and reporting jobs take much longer. This is because the backup job is still consuming network bandwidth, and the congestion reduces performance for all the other systems.

Moreover, if a backup job runs so slowly that it cannot be completed within a reasonable window, the only options are to back the data up less frequently, or to back up a smaller amount of data – both of which reduce the company's ability to protect its systems and data effectively, and increase the business impact of cyber-attacks.

THE PROBLEMS: LATENCY AND PACKET LOSS

In general, there are two main issues that cause the majority of performance problems when transferring data across a WAN.

The first is latency – the round-trip time delay between sending a packet and receiving its acknowledgement. The main causes of latency are the physical distance that the packet must travel, and the time taken to receive, queue and process packets at either end of the connection and at any intermediate gateways. The further the data must travel, and the more gateways it must pass through, the greater the latency.

IBM Storage Scale offers several technologies for connecting between clusters across a WAN. NFS mount points can be created separately or integrated with IBM Storage Scale's Cluster Export Services (CES) system, which also supports SMB, S3 and HDFS. Clusters can also share file systems natively using

Remote Cluster. However, all of these technologies fundamentally rely on TCP/IP, which has two serious issues.

The first issue is that high latency can cripple TCP/IP transfer rates, even over a theoretically high-bandwidth WAN infrastructure. TCP/IP works by sending a group of packets, then waiting for an acknowledgement that the packets have been received before it sends the next group. If the latency of the connection is high, then the sender spends most of its time waiting for acknowledgements, rather than actually sending data. During these periods, the network is effectively idle, with no new data being transferred.

The second issue is packet loss, which occurs when a packet sent from one system never reaches the intended recipient, or when the acknowledgement is lost before it returns to the sender. When this happens, TCP/IP automatically reduces the number of packets it sends in the next group to compensate for the unreliability of the connection. As a result, network utilisation is greatly reduced, because the sender is sending fewer packets in the same amount of time.

Organisations often try to solve TCP/IP performance issues by investing in more expensive network infrastructure that offers a larger maximum bandwidth. However, this does not fix the problem. As we have seen, latency and packet loss prevent TCP/IP connections from fully utilising the available bandwidth – so any extra investment in bandwidth will simply be wasted unless the latency and packet loss issues can be addressed.

THE SOLUTION: PORTROCKIT

PORTrockIT offers a solution to the twin challenges of network latency and packet loss. Instead of sending a group of packets down a single physical connection and waiting for a response, the solution creates multiple parallel virtual TCP/IP connections that send a constant stream of data across the physical connection.

As soon as a virtual connection has sent its packets and starts waiting for an acknowledgement from the recipient, PORTrockIT immediately opens another virtual connection and sends the next set of packets. Further connections are opened until the first connection receives its acknowledgement; that first connection is then reused to send another set of packets, and the process repeats continuously.

This parallelisation effectively eliminates the effects of latency by ensuring that the physical connection is constantly transferring new packets from the sender to the recipient: there is no longer any idle time, and the network's bandwidth can be fully utilised.

The solution also significantly reduces the impact of packet loss. If one of the virtual connections loses a packet, TCP/IP will only reduce the number of packets in the next group sent by that specific virtual connection. All the other virtual connections continue to operate at full speed, maintaining aggregate throughput even under adverse network conditions.

Unlike WAN optimisation products that rely on compression or deduplication techniques, PORTrockIT transfers data exactly "as-is". This means it can accelerate pre-compressed, deduplicated or encrypted data just as effectively.

DEPLOYMENT

Bridgeworks PORTrockIT integrates well with existing IBM Storage Scale installations because of its variety of deployment options:

- Dedicated PORTrockIT hardware appliances deliver the highest performance, particularly at scale.
- Virtual PORTrockIT appliances on platforms such as VMware ESXi, Proxmox, and Windows Hyper-V, provide compatibility with existing VM hosts.
- PORTrockIT Docker containers offer the greatest flexibility, deploying on any modern Linux server with minimal setup.

Each approach provides identical acceleration technology; the choice depends on scale, throughput requirements and operational preference. Hardware suits high-throughput, long-distance WANs, while virtual and containerised deployments simplify roll-out and automation.

TURNING THEORY INTO PRACTICE

To demonstrate the kind of results that PORTrockIT can deliver for IBM Storage Scale customers, Bridgeworks conducted a set of performance tests at a UK-based testing facility. The concept was initially explored on open-source virtualisation platform Proxmox, running PORTrockIT in Docker containers for rapid deployment. When the results showed promise, further tests were performed, using VMware ESXi for virtualisation and hardware PORTrockITs to achieve maximum performance.

The test infrastructures replicated a real-world WAN transfer, using a WAN emulator (WANulator) to simulate different levels of latency and packet loss between the source and the target systems.

The tests transferred data between two clusters, each with a gateway and a pair of NSD servers. Each NSD server was equipped with two high-performance solid state drives (SSDs) to ensure that the results were not bottlenecked by

the write speed of the storage media. Instead, the only limitations on the transfer speed were the 10Gb/s Ethernet link and the simulated conditions of the WAN.

Each cluster had a local IBM Storage Scale filesystem and was configured to remotely mount the other cluster's filesystem, providing read and write access. Transfers were performed using the IO500 storage benchmark suite, specifically the IOR-easy-write test. The IO500 suite is an industry standard tool for assessing the performance of filesystems, so it was ideal for this comparison.

The first set of tests were performed on an unaccelerated architecture, where the clusters were connected directly to the WANulator (see Figure 1). The same tests were then repeated on an architecture that was accelerated by introducing two hardware PORTrockIT appliances, connected to either side of the WANulator, between the cluster and the WAN (see Figure 2).

TEST EQUIPMENT

SOFTWARE:

- IBM Storage Scale version 5.2.3.2
- IOR 4.0.0
- IO500 SC20 benchmark suite
- VMware ESXi 8.0 Update 3

HARDWARE:

- 2 x Dell PowerEdge R730 as ESXi hosts
- 2 x PORTrockIT 400 Nodes (hardware test only)
- 2 x Dell PowerEdge R260 running Red Hat Enterprise Linux 9 as Docker hosts (Docker test only)
- WANulator host

For a final comparison, tests were also performed with two types of virtual PORTrockIT: one running as a virtual machine hosted on VMware ESXi (Figure 3), and another running under Docker on Red Hat Enterprise Linux 9 (Figure 4).

Figure 1: Unaccelerated test rig

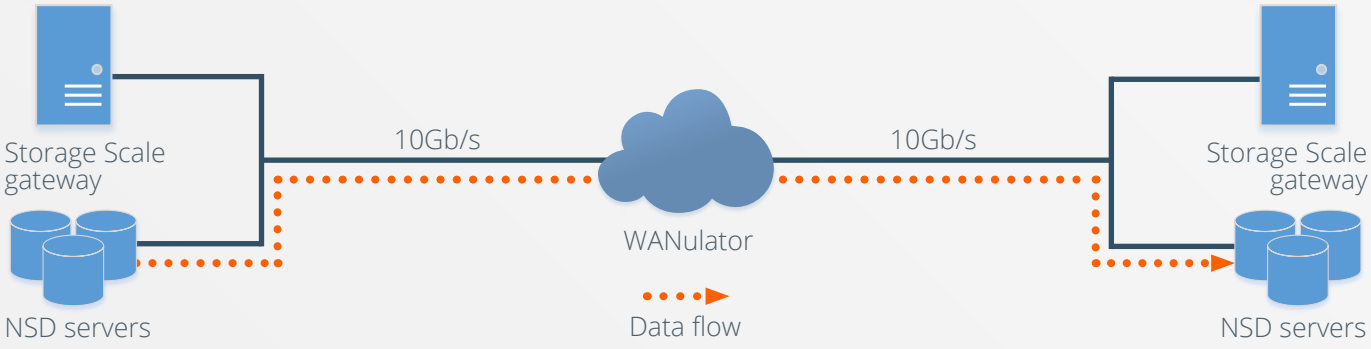


Figure 2: Accelerated test rig for the hardware test

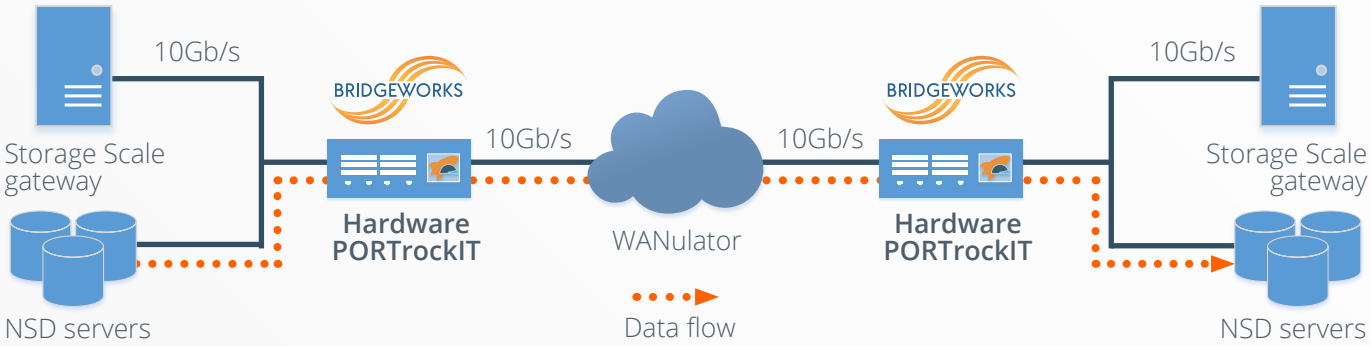


Figure 3: Accelerated test rig for the virtual machine test

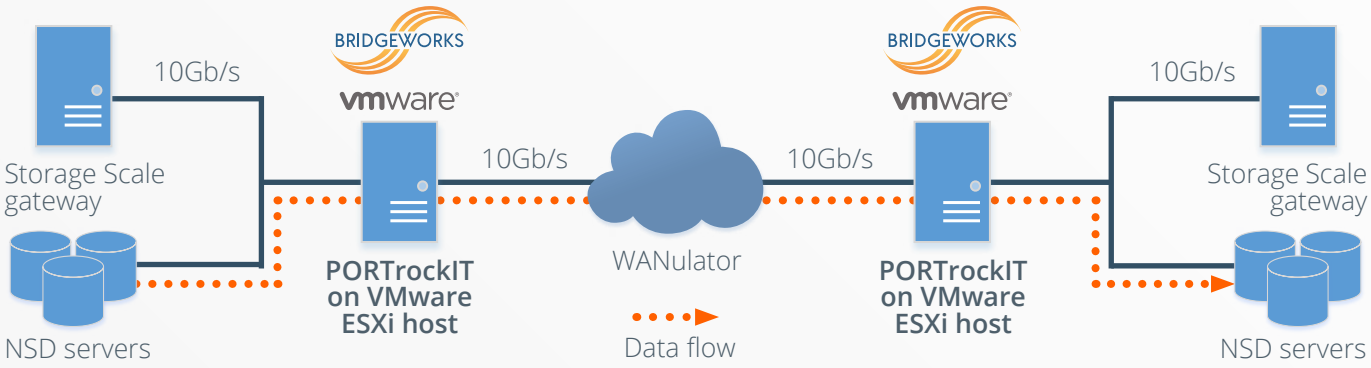
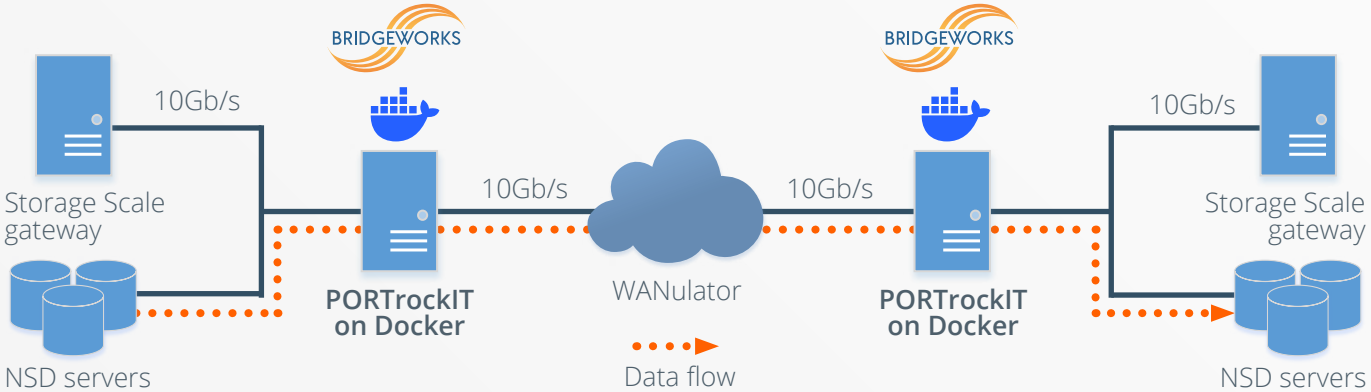


Figure 4: Accelerated test rig for the Docker test



WHAT THE DATA TELLS US

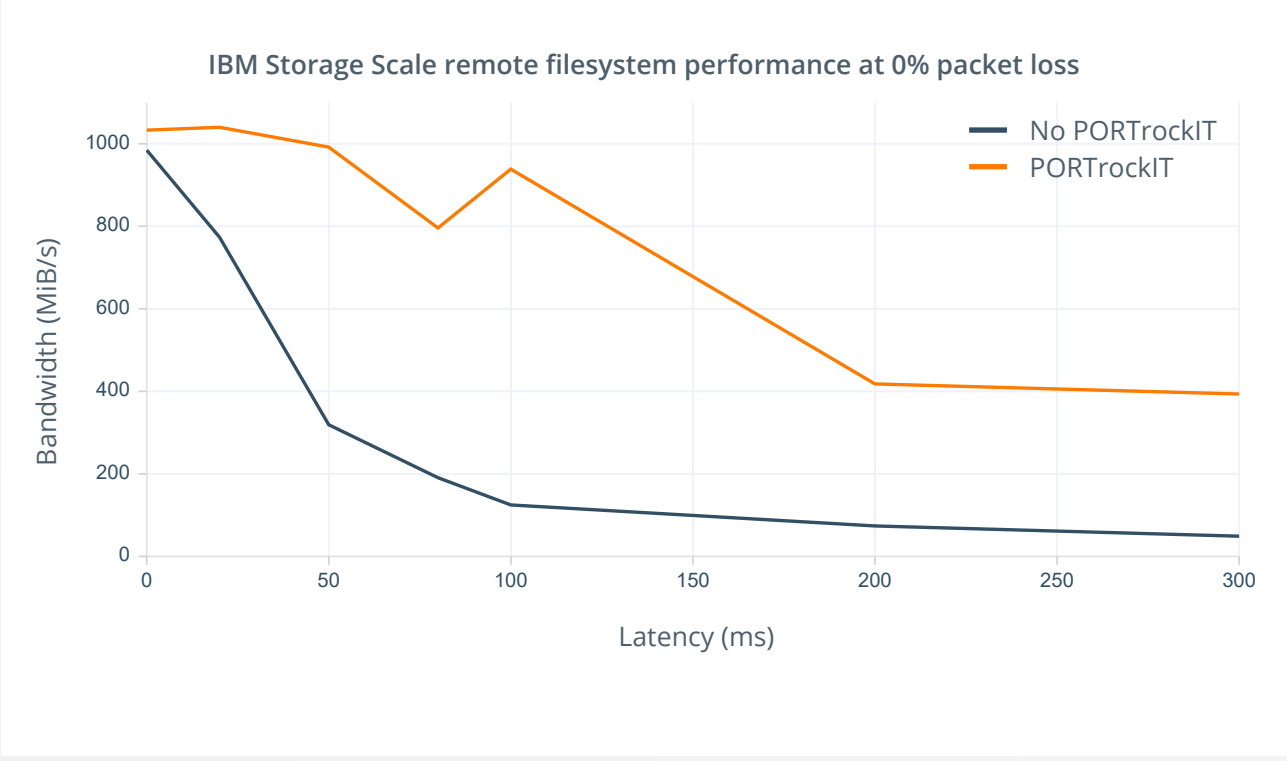
Note: All throughput figures are reported in MiB/s to match IO500 output. A 10Gb/s link has a practical ceiling of approximately 1,200MiB/s.

The first test simulated a scenario with no packet loss, at latencies ranging from 0ms to 300ms round trip time (RTT). The IOR test was allowed to run for 5 minutes in total, during which time it wrote between 100GB and 400GB of data to the remote filesystem. The test was performed twice, first via the unaccelerated architecture, and then again via the accelerated architecture with PORTrockIT, while keeping the WAN link limited to 10Gb/s.

Note: Unlike WAN optimisation products which use compression or deduplication techniques to improve throughput, PORTrockIT transfers data exactly "as-is", without modifications. This means that PORTrockIT can accelerate deduped, compressed or encrypted data transfers to exactly the same extent as it accelerates any other data type.

Looking at Figure 5, the results show that performance on the unaccelerated architecture degraded significantly as soon as even a small amount of latency was introduced. Just 50ms of latency reduced performance from 1,000MiB/s to 319MiB/s, and when latency rose above 100ms, transfer rates dropped below 100MiB/s.

Figure 5: Accelerated and unaccelerated performance at various latencies with 0% packet loss



This effect was even more extreme when a small amount of packet loss was introduced. As Figure 6 shows, with 0.1% packet loss, it takes just 20ms of latency to bring the transfer speed down from 1,000MiB/s to 34MiB/s.

With PORTrockIT it is a different story. At 0% packet loss, PORTrockIT maintains 938MiB/s even at 100ms of latency, while with 0.1% packet loss it reaches 450MiB/s when the unaccelerated transfer is struggling to reach 5MiB/s.

Meanwhile, the 0.1% packet loss and 20ms of latency that previously crippled the transfer now barely has an effect, as PORTrockIT increases the speed by 25x.

A similar graph can also be created to compare the effect of packet loss, when latency is not a factor. Figure 7 demonstrates that, for typical packet losses of <1%, latency is by far the greatest determiner of bandwidth loss. However, even with no latency, packet loss can significantly harm the effectiveness of a connection. IBM Storage Scale is able to cope with up to 0.1% packet loss in a latency-free environment, but performance decreases rapidly thereafter. But by introducing PORTrockIT, we can significantly mitigate the impact of higher levels of packet loss.

Figure 6: Accelerated and unaccelerated performance at various latencies at 0.1% packet loss

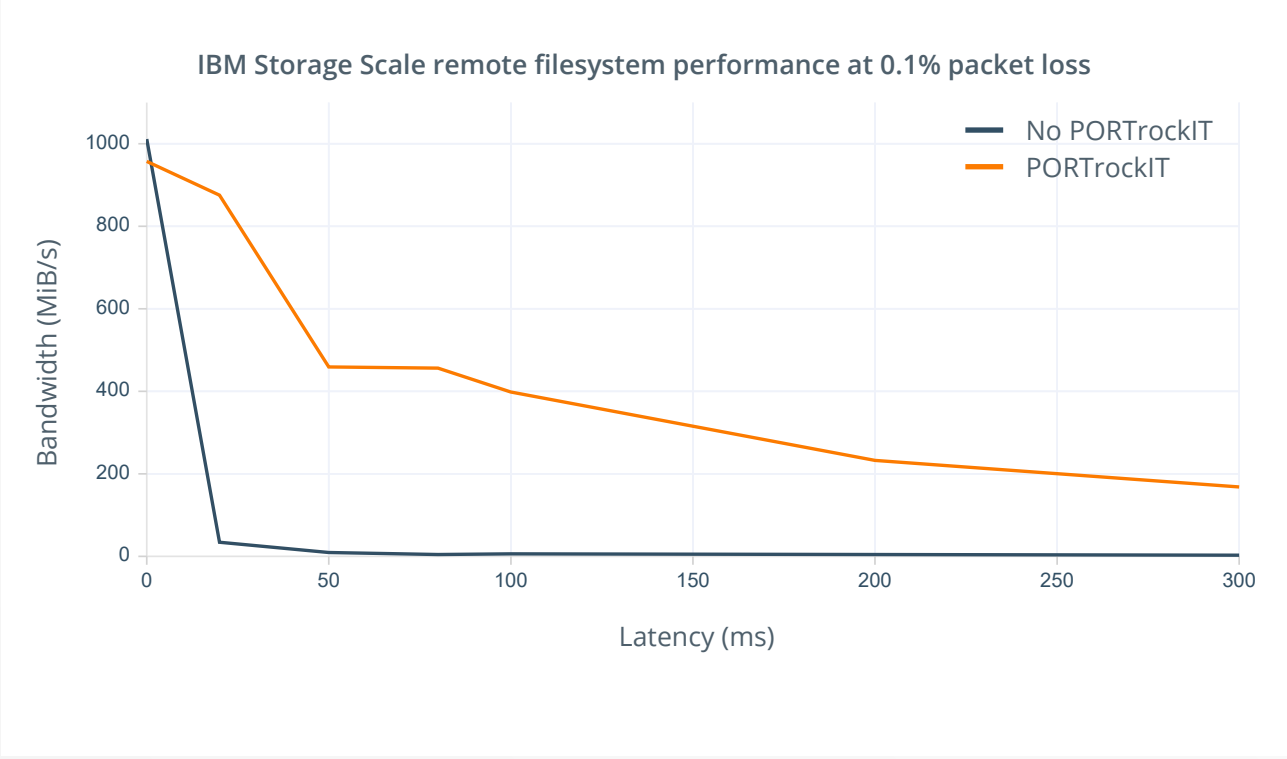


Figure 7: Accelerated and unaccelerated performance with various rates of packet loss at 0ms latency

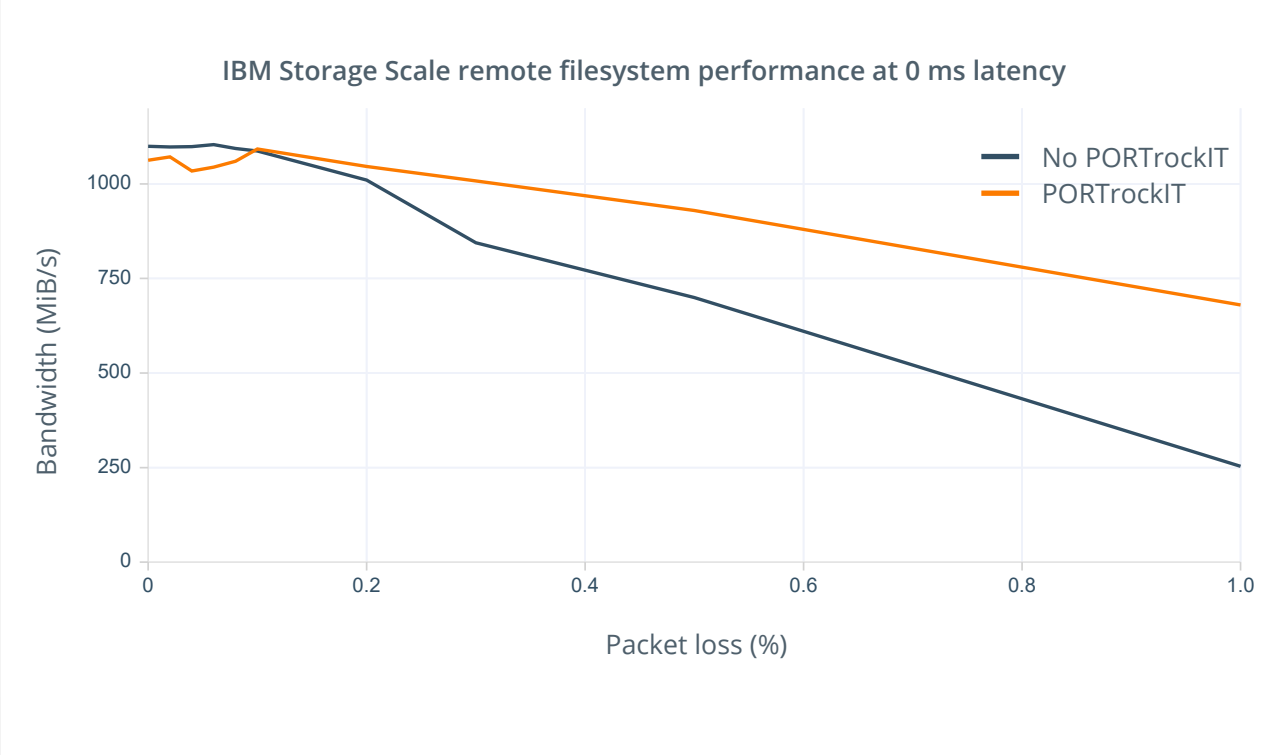
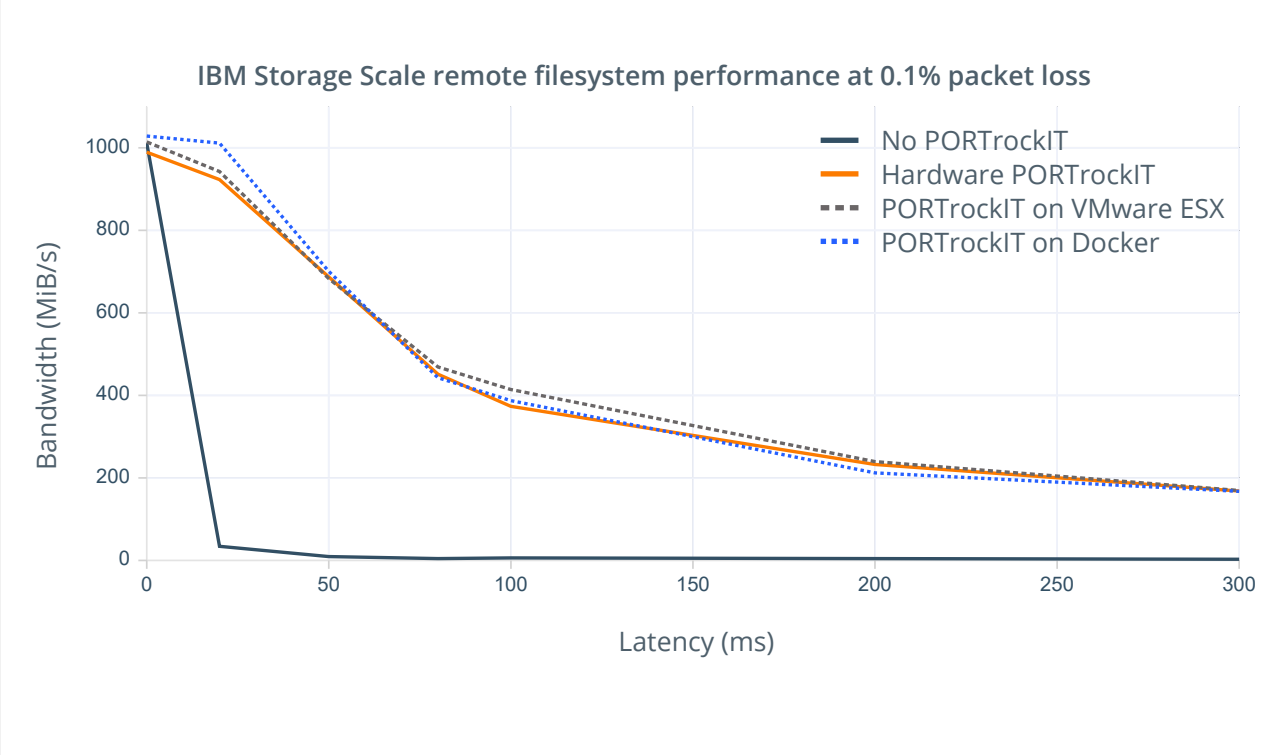


Figure 8: Performance comparison of various PORTrockIT deployment models versus the unaccelerated baseline



PLATFORM COMPARISON

In our platform comparison tests, all three PORTrockIT deployment models yielded substantial gains over the unaccelerated baseline. Figure 8 shows that under adverse WAN conditions, PORTrockIT consistently achieved superior performance regardless of platform.

With typical TCP traffic, PORTrockIT hardware appliances will outperform virtual formats in the highest latency and packet loss scenarios. With their dedicated NIC queues and offload consistency, hardware appliances can sustain higher aggregate throughput and lower tail latencies. Meanwhile, ESXi and Docker will deliver near-parity at moderate latency values (≤ 100ms), while providing their own benefits. Docker offers the fastest deployment and easiest automation, while ESXi integrates cleanly in environments where virtual networking and vDS policy are standardised.

However, Figure 8 shows our results for Storage Scale are similar on all platforms, which suggests that the lower processing power of virtual appliances is not limiting performance. Since the underlying SSDs are easily capable of 10Gb/s, the bottleneck must be in the speed at which IBM Storage Scale is providing data.

*"Under adverse WAN conditions, **PORTrockIT consistently achieved superior performance** regardless of platform."*

At higher latencies, acknowledgements and control signals take longer to arrive, which can cause applications to limit their transfer rates despite the increased bandwidth offered by PORTrockIT. Nevertheless, these results prove that PORTrockIT can make a significant difference, and highlight avenues for further exploration.

In practice, our recommendation is to use hardware where links are long-distance, lossy, or require bandwidth above 10Gb/s. Meanwhile, use Docker or ESXi for rapid adoption and environments with well-tuned host resources and NIC features (RSS, NUMA pinning and MTU configuration).

In all cases, the performance uplift remained consistent, demonstrating that PORTrockIT's acceleration operates independently of the underlying platform and scales linearly with link capacity.

REALISING THE BUSINESS BENEFITS

For companies that manage their data storage with IBM Storage Scale, PORTrockIT can dramatically improve operational performance over wide area links. PORTrockIT can extend the geographical range of your IBM Storage Scale clusters without compromising performance – creating new opportunities for growth while maintaining high availability of critical data.

Furthermore, if a backup process is threatening to overrun the available window, or if it is desirable to reduce transfer times to free up server and network resources for other important jobs, PORTrockIT provides an elegant, low-risk solution.

PORTrockIT offers plug-and-go technology that can be implemented quickly with minimal impact on existing infrastructure, keeping deployment cost and risk to a minimum. By maximising the performance of existing infrastructure, PORTrockIT also reduces the need to invest in expensive high-bandwidth connections or more powerful servers – delivering measurable cost-avoidance and faster return on investment.

Most important of all, PORTrockIT removes the risk of relying on an infrastructure that is not fast, robust or reliable enough to keep your business-critical data flowing smoothly. By accelerating WAN data transfers, the solution ensures that data can move wherever and whenever the business needs it – not just when network conditions allow it.

As a result, you can access critical data more quickly, optimise your recovery point and recovery time objectives (RPO/RTO), and deliver higher service levels to business users and customers.

TAKE THE NEXT STEPS

To learn more about PORTrockIT and other smart networking solutions from Bridgeworks, please visit www.4bridgeworks.com, or call us on +44 (0) 1590 615 444.

For expert guidance, design, and deployment of IBM Storage Scale environments, contact OCF, an IBM Gold Business Partner specialising in high-performance computing and parallel storage. Visit www.ocf.co.uk or call +44 (0)114 257 2200 to discuss how OCF's technical specialists can help optimise, scale, and support your IBM Storage Scale infrastructure.

Bridgeworks' UK-based technical team, together with partners such as OCF, can assist with proof-of-concept deployments, WAN performance assessments, and integration of PORTrockIT with IBM Storage Scale environments.

CONTRIBUTORS

This proof of concept and whitepaper were made possible through the collaboration of technical specialists from Bridgeworks and OCF. Their combined expertise in WAN acceleration, network performance optimisation, and IBM Storage Scale environments ensured the success and accuracy of this evaluation.

AUTHORS

- Matthew Dyson, BSc (Hons), Software Engineer, Bridgeworks
- Mário Fernandes, BSc (Hons), Software Engineer, Bridgeworks
- Sasha Paterson, MEng (Hons), Lead Software Engineer, Bridgeworks

SPECIAL THANKS

To Laurence Horrocks-Barlow, Technical Director at OCF, whose guidance, support and years of experience with IBM Storage Scale have been invaluable throughout.

To Peter McNamara, Business Development Manager EMEA at Bridgeworks, for project coordination and resource management.

To David Trossell, CEO of Bridgeworks, for his advice and expertise in WAN acceleration technology.