

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: 23 April 2026

D. King  
Lancaster University  
T. Chown  
Jisc  
C. Rapier  
Pittsburgh Supercomputing Center  
D. Huang  
ZTE Corporation  
K. Yao  
China Mobile  
20 October 2025

Current State of the Art for High Performance Wide Area Networks  
draft-kcrh-hpwan-state-of-art-03

Abstract

High Performance Wide Area Networks (HP-WANs) represent a critical infrastructure for the modern global Research and Education (R&E) community, facilitating collaboration across national and international boundaries. These networks include global education and research networks, such as G<sub>テ</sub>陰NT, Internet2, Janet, ESnet, CANARIE, CERNET, and others, and also refer to large scale commercial dedicated networks built by hyperscalers and operators. They are designed to support the ever-growing transmission of vast amounts of data generated by scientific research, high-performance computing, distributed AI-training and large-scale simulations.

This document provides an overview of the terminology and techniques used for existing HP-WANs. It also explores the technological advancements, operational tools, and future directions for HP-WANs, emphasising their role in enabling cutting-edge scientific research, AI training and massive R&E data analysis.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 23 April 2026.

## Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	3
1.1. Background . . . . .	4
2. Terminology . . . . .	4
3. Example Use Cases for HP-WANs . . . . .	6
4. Current Technologies Used in HP-WANs: Key Components . . . . .	7
4.1. Architectural Elements . . . . .	7
4.2. Topology . . . . .	8
4.3. Bandwidth and Latency . . . . .	9
4.4. Localised Data Movement . . . . .	9
4.5. Forwarding Optimisation . . . . .	10
4.6. Reliability and High Availability . . . . .	11
4.7. Quality of Service . . . . .	12
4.8. Congestion Control . . . . .	12
4.9. Performance Monitoring . . . . .	12
4.10. Scalability . . . . .	13
4.11. Sustainability and Energy Efficiency . . . . .	13
4.12. Resource Scheduling . . . . .	13
5. Examples of HP-WANs . . . . .	13
5.1. Gテ陰NT . . . . .	14
5.2. Janet . . . . .	14
5.3. Google Effingo . . . . .	15
5.4. Energy Sciences Network . . . . .	16
5.4.1. Practical Examples of Dynamic Network Management . . . . .	16
5.5. Internet2 . . . . .	17
5.6. CANARIE . . . . .	17

5.7. Asia-Pacific Advanced Network . . . . .	17
5.7.1. CERNET . . . . .	18
5.7.2. China Mobile Cloud Dedicated Network . . . . .	18
6. Emerging Trends and Future Directions . . . . .	19
6.1. Integrated Resource and Network Control . . . . .	19
6.2. Intent-Based Networking and Automation . . . . .	19
6.3. Network Signalling . . . . .	20
7. IANA Considerations . . . . .	20
8. Security Considerations . . . . .	20
9. Acknowledgements . . . . .	20
Contributors . . . . .	20
Normative References . . . . .	21
Informative References . . . . .	21
Authors' Addresses . . . . .	22

## 1. Introduction

High Performance Wide Area Networks (HP-WANs) are the backbone of global Research and Education (R&E) infrastructure, enabling the seamless transfer of vast amounts of data and supporting advanced scientific collaborations worldwide. These networks are designed to meet the demanding requirements of data-intensive research fields, including high-energy physics, climate modeling, genomics, and artificial intelligence.

The evolution of HP-WANs is deeply intertwined with the growing need for advanced scientific research and the increasing globalisation of collaboration. Traditional WANs, which were sufficient for general business and communication needs, quickly became inadequate for the specialised requirements of research institutions. As scientific endeavours began to generate larger datasets, ranging from terabytes to petabytes, there arose a need for networks capable of transferring these massive volumes of data reliably and securely across long distances.

Specialised (inter)National R&E Networks (NRENs), such as ESnet in the United States, GÉANT in Europe, Janet in the UK, and CERNET in China, have evolved to support the unique needs of the scientific community, while also carrying more generalised and less demanding research and education traffic. These networks are designed by their operators to provide high bandwidth and ensure low latency, high reliability, and robust security, critical for applications like real-time data analysis, distributed computing, and remote instrumentation.

This evolution has been made possible both by the design decisions made by NREN operators and also the campus operators who build the localised network and systems infrastructures. HP-WANs require an

end-to-end engineering approach, where the NREN operators provide capacity while the local campus operators engineer their network architectures and their data transfer and compute systems ensure optimal use of that capacity, e.g., by adopting the Science DMZ approach described in this document.

Today, HP-WANs are foundational to the research community and are leading the way in demonstrating how advanced networking technologies can be applied to other sectors. They serve as testbeds for innovations in networking that eventually trickle down to broader commercial applications. As we look towards the future, HP-WANs will continue to play a critical role in enabling scientific discoveries and fostering international collaboration, particularly as emerging technologies such as quantum computing and the Internet of Agents (IoA) push the boundaries of what these networks must support.

This document explores the current state of the art in HP-WANs, examining the technological advancements, operational challenges, and emerging trends shaping the future of networks built for research, education, massive data analysis and collaborative AI training at scale and speed. Through this exploration, we aim to provide a better understanding of the current state of the art in high performance computing across wide area networking.

## 1.1. Background

High Performance Wide Area Networks (HPWANs) evolved as specialised networks initially designed to facilitate scientific research requiring high-speed data transfer, high reliability, and minimal latency. ESnet, Janet, G<sub>テ</sub>陰NT and CERNET emerged in response to the increasing data volumes generated by scientific and educational institutions, transforming traditional WAN capabilities.

HP-WAN is an end-to-end capability, architected by collaborative efforts by the NREN operators, campus IT operators, and the science communities, and have since grown integral to research and educational communities, supporting distributed scientific collaborations, large-scale simulations, and intensive data analysis. Their capabilities have been continually enhanced to meet rising demands, laying foundations for future networking technologies.

## 2. Terminology

This document provides a lexicon terminology that relates to high performance WANs.

CERN: The European Organization for Nuclear Research, housing the Large Hadron Collider (LHC).

High Performance Computing (HPC): Is a general term for computing with a high level of performance. Often high performance computing specifically refers to running jobs which are very parallel, often running on hundreds or even thousands of cores.

High Performance Wide Area Network (HP-WAN): A type of Wide Area Network (WAN) designed specifically to meet the high-speed, low-latency, and high-capacity needs of scientific research, education, and data-intensive applications. These networks connect research institutions, universities, and data centers across large geographical areas.

Infiniband: Traditionally, a localised data interconnect used by many high performance computing (HPC) systems providing high bandwidth and low latency.

National Research and Education Network (NREN): Provide a specialised network by operators supporting the research and education community within a specific country or region. NRENs provide high-speed connectivity and other services tailored to the needs of academic and research institutions.

Remote direct memory access (RDMA): Enables one networked node to access another networked nodes memory without involving either computer's operating system or interrupting either nodes processing. This helps minimise latency and maximise throughput, reducing memory bandwidth bottlenecks.

RDMA over Converged Ethernet (RoCE): Traditionally, a network protocol which allows remote direct memory access (RDMA) over a local Ethernet network. There are multiple RoCE versions. RoCE v1 [UEC] is an Ethernet link layer protocol and hence allows communication between any two hosts in the same Ethernet broadcast domain. RoCE v2 is an internet layer protocol which means that RoCE v2 packets can be routed.

Worldwide LHC Computing Grid (WLCG): Is a global network of over 170 computing centres across more than 40 countries, designed to process, store, and analyse the vast amounts of data generated by the Large Hadron Collider (LHC) at CERN.

Performance Service Oriented Network monitoring Architecture(PerfSONAR): Is a network performance monitoring toolkit designed to provide end-to-end performance measurement and monitoring across multi-domain network infrastructures.

Science DMZ: A model for deployment of infrastructure at a site

(campus) to optimise the performance of data transfers in and out of data transfer nodes (DTNs) at the site 参见 <https://fasterdata.es.net/science-dmz/>. Elements of the model include the local network architecture, tuning of DTNs, selection of data transfer software, efficient implementation of security policies, and provide persistent monitoring.

### 3. Example Use Cases for HP-WANs

HP-WAN development has become synonymous with large-scale research and experimentation, big data, and AI. HPC and therefore HP-WAN, is driven by the continuous need to move large data between HPC facilities, facilitating the following industries:

- \* High-Energy Physics Research, e.g., the Large Hadron Collider (LHC)
- \* Climate Modeling
- \* Radioastronomy, e.g., the Square Kilometre Array (SKA) project
- \* Healthcare, Genomics and Life Sciences
- \* AI training
- \* Data Backup

The data rates required by HPC applications vary significantly based on the application type and data scale.

Scientific simulations, such as climate modeling and molecular dynamics, typically demand data rates from 10 Gbps to over 100 Gbps due to the large volumes of data processed and moved between nodes and storage systems.

In high-energy physics, such as experiments at CERN, data rates can reach hundreds of gigabits per second, with aggregate peaks between site exceeding 1 Tbps currently, and predicted to rise to 10 Tbps, during intensive data processing.

Healthcare, Genomics, and Life Sciences might typically operate at rates between 1 Gbps and 40 Gbps. These applications require high throughput to handle large datasets efficiently, often through parallel data streams.

AI learning and tasks, particularly those involving deep learning, require data rates ranging from 10 Gbps to 100 Gbps to ensure efficient data movement, keeping GPUs and other accelerators fully utilised.

These varying data rates underscore the high demands of HPC applications, which are expected to grow as the field evolves and datasets become larger, and the growing need for movement of large data sets between sites (including data centers).

#### 4. Current Technologies Used in HP-WANs: Key Components

High Performance Computing (HPC) networks are specialised networks designed to connect supercomputers and other high-performance computing resources, enabling them to collaborate on computational tasks that require significant processing power, memory, and data storage. These networks facilitate large-scale scientific research, complex simulations, and data-intensive tasks that exceed the capabilities of standard computing systems.

The following sub-sections outline typical characteristics and requirements for HP-WANs. These technical requirements ensure that wide-area interconnects can meet the demanding needs of distributed HPC environments, enabling researchers and scientists to collaborate effectively globally.

##### 4.1. Architectural Elements

Not all HP-WAN deployments rely on backbone controllers. In many NRENs, capacity is provisioned ahead of demand on private dedicated, with no QoS or bandwidth-on-demand systems in the core.

Some HP-WAN network providers may choose or need to use specific resource and/or network controllers in support of delivering that functionality

Resource Controllers provide detailed control over individual network resources, such as routers and switches, ensuring efficient usage and reliable network performance through comprehensive monitoring and configuration.

Network Controllers maintain global visibility of network topology, resource availability, and status, essential for path computation, resource reservation, and dynamic reconfiguration to meet stringent performance demands.

End-to-End Orchestration translates user and application requirements into actionable network operations, enabling automated, policy-driven management and significantly improving resource responsiveness and optimisation.

#### 4.2. Topology

HPC networks can be broadly categorised into intra-site networks, which connect components within a single HPC site, such as a data centre, and inter-site networks, which link multiple HPC sites across different geographical locations. The intra-site (HP-WAN) networks typically use high-speed, low-latency non-Internet interconnects like InfiniBand or high-speed Ethernet. In contrast, inter-site networks rely on dedicated high-capacity wide area networks (WANs) to facilitate distributed computing and data sharing on a regional and global scale.

Each NREN operator, e.g., Jisc in the case of Janet in the UK, will build and operate the NREN infrastructure for its research and education users. This may typically take the form of a well-provisioned backbone, with regional access networks extending to the end sites (campuses, research organisations, etc). The NREN demarcation is typically at the campus edge. In some countries the regional networks are operated separately.

The NRENS then typically have interconnects to other NRENS, forming a worldwide RE network infrastructure. In Europe, GANT provides connectivity between the European NRENS and then wider connectivity to the rest of the world. And NRENS will have other interconnects to non-RE networks, e.g., via one or more national IXs, direct peerings to content providers (including the big cloud providers) and then "catch-all" commodity connectivity via one or more Tier 1 ISPs.

Dedicated infrastructure is commonly used in HPC environments where performance, security, and reliability are paramount. In these cases, the network infrastructure is built exclusively for HPC applications, including dedicated fibre-optic connections, private data centres, and specialised network transport like RDMA over Converged Ethernet (RoCE) and InfiniBand nodes. The primary benefits of dedicated infrastructure are its ability to provide optimised performance for HPC tasks, ensure high levels of security by preventing unauthorised access, and maintain consistent reliability by avoiding congestion or performance issues caused by other network traffic.

Usually, the responsibility for networking within an end site or campus lies with that organisation, e.g., a university IT department, while the operation of an HPC facility may have dedicated (separate)



staff. With the additional administrative domains of the NRENs and inter-NREN backbones like GANT, end-to-end traffic may pass through many networks operated by different organisations. To achieve optimal e2e performance, everyone on the end-to-end path needs to implement best practices.

#### 4.3. Bandwidth and Latency

The technical requirements for wide area interconnects between HPC sites are stringent, given the unique demands of distributed high-performance computing. High bandwidth is a primary requirement, as these interconnects must support the rapid transfer of large datasets between sites, ensuring that data movement does not become a bottleneck in computational workflows. HPC data flows might typically consume 1Gbit to beyond 400Gbit/s.

Low latency is equally critical, for many HPC applications. Latency requirements for inter-DC locations will be in the low-millisecond range. This low latency is essential for applications that require real-time or near-real-time data processing.

#### 4.4. Localised Data Movement

Network-intensive applications like networked storage or cluster computing need a network infrastructure with high bandwidth and low latency.

These interconnects may need to support specialised communication protocols designed for HPC environments, such as Remote Direct Memory Access (RDMA) [RFC5040] and [RFC7306], which optimises the performance of distributed HPC applications by reducing overhead and improving data transfer efficiency.

InfiniBand (IB) is another computer networking communications standard used in high-performance computing that features very high throughput and very low latency. InfiniBand is also used as either a direct or switched interconnect between servers and storage systems, as well as an interconnect between storage systems.

The advantages of RDMA and IB over other network application programming interfaces, are lower latency, CPU load, and bandwidth. The downside with these specialised protocols is the need for all interfaces and nodes to support the technique on the end-to-end path.

iWARP is a computer networking protocol that implements remote direct memory access (RDMA) for efficient data transfer over Internet Protocol networks. Several IETF techniques are used for iWARP:

- \* [RFC5040] A Remote Direct Memory Access Protocol Specification is layered over Direct Data Placement Protocol (DDP). It defines how RDMA Send, Read, and Write operations are encoded using DDP into headers on the network.
- \* [RFC5041] Direct Data Placement over Reliable Transports is layered over MPA/TCP or SCTP. It defines how received data can be directly placed into upper layer protocols receive buffer without intermediate buffers.
- \* [RFC5042] Direct Data Placement Protocol (DDP) / Remote Direct Memory Access Protocol (RDMAP) Security analyzes security issues related to iWARP DDP and RDMAP protocol layers.
- \* [RFC5043] Stream Control Transmission Protocol (SCTP) Direct Data Placement (DDP) Adaptation defines an adaptation layer that enables DDP over SCTP. Elephant flows: For each burst, the intensity of each flow could reach up to the line rate of NICs.
- \* [RFC5044] Marker PDU Aligned Framing for TCP Specification defines an adaptation layer that enables preservation of DDP-level protocol record boundaries layered over the TCP reliable connected byte stream.
- \* [RFC6580] IANA Registries for the Remote Direct Data Placement (RDDP) Protocol defines IANA registries for Remote Direct Data Placement (RDDP) error codes, operation codes, and function codes.
- \* [RFC6581] Enhanced Remote Direct Memory Access (RDMA) Connection Establishment fixes shortcomings with iWARP connection setup.
- \* [RFC7306] Remote Direct Memory Access (RDMA) Protocol Extensions extends [RFC5040] with atomic operations and RDMA Write with Immediate Data.

#### 4.5. Forwarding Optimisation

The scaling of HPC applications, especially across a WAN between multiple sites, requires the ability to route the massive traffic. Specifically, this requires network infrastructure to provide several routing and forwarding characteristics, which are detailed below.

- \* Low entropy: Compared to traditional data center workloads, the number and the diversity of flows for workloads and flow patterns are usually repetitive and predictable.
- \* Burstiness: Flows usually exhibit the "on and off" nature in the time granularity of milliseconds.

- \* Jumbo frames: Ethernet frames larger than the standard maximum transmission unit (MTU) size of 1,500 bytes, typically carrying payloads of up to 9,000 bytes. Using jumbo frames can significantly enhance network efficiency and reduce CPU overhead.
- \* Elephant flows: For each burst, the intensity of each flow could reach up to the line rate of NICs.

It should be noted that efficiently handling these elephant flows is crucial in HPC as they can otherwise saturate network links, leading to congestion and reduced performance for other network traffic. Strategies to manage elephant flows effectively, such as prioritising these flows or segmenting network traffic, help maintain overall network performance and ensure that large data transfers do not hinder the execution of other critical tasks within the HPC environment.

HPC transport options include IP (both UDP and TCP), and emerging mechanisms such as QUIC. However, each transport technology provides strengths and weaknesses. In all cases, the primary goal is to ensure the effective high-throughput, low latency and jitter, low-packet loss ratio, transmission of massive data sets.

#### 4.6. Reliability and High Availability

In HPC networks, the resilience of the data stream is important due to the critical need for precise, high-speed data transfer. These networks must maintain continuous data flow to support large-scale computations, where even minor interruptions or packet loss can severely impact performance, causing delays or incorrect results. Therefore, resilience must be implemented to ensure the network can recover from disruptions without compromising speed or integrity.

For retransmission and lossless data transfer, HPC networks must have mechanisms to handle data loss efficiently. They must quickly retransmit lost or corrupted packets while maintaining a seamless data flow to avoid performance degradation. The requirement for lossless communication is essential to meet the needs of scientific computations, simulations, and data-intensive tasks.

High availability and redundancy are also essential to prevent data loss and ensure continuous operation, especially given that HPC tasks often run for extended periods and involve critical research. These networks must also incorporate advanced security measures, including encryption and secure access controls, to protect the often sensitive or classified data being transmitted.

#### 4.7. Quality of Service

Depending on the provisioning and contention for resource, the network may need to support Quality of Service (QoS) mechanisms to prioritise traffic, ensuring that critical HPC tasks receive the necessary bandwidth and low-latency performance.

An approach may be needed to enable applications to request specific bandwidth or latency guarantees, ensuring that high-priority tasks receive required resources, hard to provide over shared multi-domain infrastructure.

Differentiated Services (Diffserv) offers a flexible method to manage traffic prioritization without the need for an explicit request-and-grant process. Diffserv operates by marking packets with different priority levels, allowing the network to prioritize and protect access to capacity for critical tasks. This approach may be useful in HPC environments where dynamic traffic patterns require adaptive resource management.

#### 4.8. Congestion Control

Congestion control mechanisms ensure that data transfers between nodes and across networks are efficient and do not overwhelm the HPC network infrastructure. By managing and regulating the flow of data, congestion control mechanisms help prevent bottlenecks, reduce latency, and maintain high throughput, which are essential for the performance and reliability of HPC applications that require the rapid movement of large volumes of data across distributed systems.

Depending on the transport technology used in the HPC environment, several congestion control schemes may be use:

- \* InfiniBand Congestion Control
- \* RDMA-based Data Center Quantized Congestion Notification (DCQCN)
- \* TCP-based Bottleneck Bandwidth and Round-Trip Time (BBRv3)
- \* Explicit Congestion Protocol (XCP)

#### 4.9. Performance Monitoring

End-to-end performance measurement and monitoring across multi-domains and network infrastructures are important in HPC environments. They provide a method to diagnose and troubleshoot network performance issues that can affect data-intensive applications and distributed computing tasks commonly found in HPC.

perfSONAR is a network measurement toolkit commonly used. It is designed to provide federated coverage of network paths. It provides an interface that allows for the scheduling of measurements, storage of data, and generate visualisations.

Data transfer applications should log their transfers such that monitoring tools can assess their end-to-end, disk-to-disk performance. e.g., FTS does this for WLCG data transfers.

#### 4.10. Scalability

Scalability is another crucial aspect, allowing the network to expand efficiently as computational needs grow, accommodating additional sites or increased capacity without significant reconfiguration. Interoperability is also necessary, ensuring that the network can communicate seamlessly across different types of hardware, software, and protocols used at various HPC sites.

#### 4.11. Sustainability and Energy Efficiency

As HPWANs continue to expand, sustainability and energy efficiency are becoming critical considerations. The operational scale of these networks—spanning global infrastructures and data-intensive applications—poses significant environmental and economic challenges. Future HP-WAN deployments will increasingly prioritise energy-efficient network components, smart power management systems, and sustainable operational practices.

Emerging approaches include adaptive network management strategies designed to reduce energy consumption during periods of lower utilisation and leveraging advanced technologies such as optical networking and energy-aware routing protocols. Furthermore, industry-wide initiatives are focusing on measuring and reducing the carbon footprint of data transfers and network operations, contributing to broader climate goals.

#### 4.12. Resource Scheduling

[Editor's Note - Do we need to discuss service and resource scheduling?]

### 5. Examples of HP-WANs

The following sub-sections highlight examples of HP-WANS, and their technical specifications.

### 5.1. GANT

The GANT network is a pan-European data network dedicated to research and education, providing high-speed, high-capacity connectivity across Europe, between European NRENs and to other worldwide NRENs. It is an essential infrastructure for HPC applications, enabling collaboration and data sharing among research institutions, universities, and HPC centers across the continent and beyond.

The core of GANT operates at speeds of up to 600 Gbps, using Dense Wavelength Division Multiplexing (DWDM) technology. This provides connectivity suitable for HPC applications, particularly those involving large-scale simulations, scientific research, and real-time data processing. Reliability is provided by using multiple optical underlay paths for data to travel between GANT nodes. This design ensures high availability and reliability, which is crucial for the continuous operation of HPC environment.

The GANT network integrates perfSONAR for real-time network performance monitoring and reporting of IP performance metrics [RFC6703], allowing HPC users to detect and troubleshoot potential issues that could impact data transfer and overall performance. This ensures that the high-performance requirements of HPC applications are met consistently across the network.

GANT provides specialised services for specific HPC projects, such as the LHC Optical Private Network (LHCOPN) and LHC Open Network Environment (LHCONE), which are critical for supporting the data-intensive needs of the Large Hadron Collider (LHC) at CERN. These services offer dedicated, high-bandwidth connections that are optimised for the massive data flows generated by LHC experiments.

The GANT network connects over 50 million users across more than 10,000 institutions in 40 countries. This extensive reach supports a wide range of HPC applications by enabling seamless collaboration between geographically dispersed research facilities. Beyond Europe, GANT connects to other major research and education networks, including Internet2 in the United States and CANARIE in Canada, allowing for global HPC collaborations and data exchanges.

### 5.2. Janet

The Janet network is the UK NREN, operated by Jisc. First established in 1984, backbone links now run at up to 800Gbps, with a growing number of sites connected at 100Gbps, in some cases with multiple 100G links. A typical university site will have multiple 10G links.

Janet connects to other RE networks via a 400G resilient link to GANT. It has a presence in multiple IXes, predominantly LINX, connects/peers directly to many content and cloud providers, and has commodity connectivity via Tier1 ISPs. The total aggregate external capacity is around 4-5 Tbit/s.

Some private, dedicated links are used by Janet sites, e.g., the CERN to RAL (UK Tier 1 site) LHCOPN link, which is a dedicated 200G path. While LHCONE is a L3VPN that may be on its own circuit or as an overlay on general RE paths. Otherwise, Jisc seeks to provide sufficient capacity to its science communities ahead of demand, and work with them (via its network performance team) to provide advice and guidance on how sites can optimise their use of the Janet network, e.g., by following Science DMZ principles.

### 5.3. Google Effingo

Google Effingo is a state-of-the-art, high-performance infrastructure designed to meet the demanding data processing and storage needs of large-scale machine learning (ML), artificial intelligence (AI), and computational workloads. As part of Google's cloud offering, Effingo is an example of how WAN infrastructure supports high-performance computing applications across diverse industries and research areas.

Effingo leverages a global network of data centers interconnected with high-capacity, low-latency WAN links. These links facilitate rapid data exchange and provide the performance required to handle real-time AI model training, complex simulations, and large-scale data analytics. The network is optimised for high-throughput workloads, where low latency and reliability are critical for processing large datasets across vast geographical areas, and more than 100 data center sites.

Effingo utilises a private global network of high-capacity fiber links, combined with packet-layer protocols to deliver low-latency, high-speed data transfer across continents. This connectivity enables global collaboration between research centers, universities, and data-driven enterprises, allowing them to share large datasets and results.

Currently, Effingo daily data transfers exceeds 1 exabytes.

#### 5.4. Energy Sciences Network

The Energy Sciences Network (ESnet) is a high-performance network dedicated to supporting scientific research within the United States, operated by the U.S. Department of Energy (DOE). Established in 1986, ESnet interconnects national laboratories, supercomputing centres, universities, and research institutions, enabling collaborative scientific projects, data-intensive applications, and high-performance computing (HPC) tasks across multiple geographical locations.

ESnet delivers high-capacity, low-latency connectivity through its robust fibre-optic backbone, employing advanced optical networking technologies and dynamic circuit provisioning services. It supports data transfer rates ranging from tens of gigabits per second up to multi-hundred gigabit per second capacities, essential for demanding scientific workflows such as high-energy physics experiments, climate modelling, and large-scale genomic research.

A key feature of ESnet is its use of specialised services such as the On-Demand Secure Circuits and Advance Reservation System (OSCARS), providing dynamic, guaranteed-bandwidth paths that allow researchers to reserve network capacity tailored specifically to their project's needs. Additionally, the network incorporates advanced orchestration platforms like SENSE, offering intent-driven, automated management to ensure optimal network resource utilisation and agile response to evolving scientific requirements.

ESnet's infrastructure integrates comprehensive monitoring and diagnostic tools such as PerfSONAR, ensuring end-to-end network visibility and performance analysis across institutional boundaries. This facilitates proactive identification and resolution of performance bottlenecks, maintaining the reliability and efficiency necessary for HPC operations.

With interconnections to international research networks, including GANT, Janet, Internet2, and CANARIE, ESnet provides global reach, facilitating extensive international collaboration and enabling the seamless exchange of data among scientific communities worldwide.

##### 5.4.1. Practical Examples of Dynamic Network Management

ESnet's OSCARS system exemplifies dynamic, advanced reservation, and circuit provisioning, demonstrating the practical application of HPWAN capabilities in operational scientific networks.



The SENSE platform further illustrates how intent-based networking and automation can simplify complex resource allocation processes, significantly improving network agility and scalability.

#### 5.5. Internet2

Internet2 is a high-performance networking consortium serving the United States research and education community. Established in 1996, Internet2 provides advanced networking infrastructure specifically designed to support collaborative research, scientific discovery, and innovation among educational institutions, government laboratories, and industry partners.

Internet2 operates an advanced optical backbone network capable of multi-terabit speeds, also delivering exceptionally high-capacity and low-latency connections. As with aforementioned networks it supports dynamic bandwidth allocation, advanced monitoring through tools, and federated identity management.

#### 5.6. CANARIE

CANARIE is Canada's national research and education network, established in 1993, dedicated to providing robust, high-performance connectivity for research, education, and innovation. It interconnects universities, research centres, healthcare institutions, and government laboratories across Canada, as well as facilitating international collaboration through global interconnections with networks such as GANT, Internet2, and ESnet.

As with other regions the CANARIE network operates using a high-capacity fibre-optic backbone, delivering advanced networking services tailored specifically for demanding scientific and research applications. The network provides dynamic, software-driven capabilities, including dedicated high-speed links, automated resource allocation, and integrated identity and access management solutions. Additionally, CANARIE supports advanced services like the Digital Accelerator for Innovation and Research (DAIR), enabling cloud-based research and development.

#### 5.7. Asia-Pacific Advanced Network

#### 5.7.1. CERNET

The CERNET (China Education and Research Network) is a national high-performance network serving China's education, research and innovation sectors. It delivers high-speed, secure connectivity across the country—linking universities, research institutions and HPC centers—and enables connections to global research networks, playing a key role in supporting collaborative research, large-scale data sharing and advanced HPC applications.

CERNET's backbone operates at up to 100 Gbps (with ongoing upgrades) using technologies like DWDM and SDN, handling HPC demands such as scientific simulations, climate modeling and high-throughput data processing. A redundant topology with alternate paths ensures ultra-high reliability, minimizing downtime for critical HPC operations.

CERNET offers customized services for major national projects, including dedicated high-bandwidth connections for the China Spallation Neutron Source (CSNS) and national supercomputing centers (e.g., Sunway TaihuLight). These support data-intensive workloads and seamless data exchange.

CERNET covers 31 Chinese provinces, and connects over 200 million users across 2,000+ universities, institutes and schools, fostering cross-regional research collaboration. Internationally, it peers with global networks like GANT (Europe) and Internet2 (U.S.), enabling global HPC cooperation.

#### 5.7.2. China Mobile Cloud Dedicated Network

China Mobile Cloud Dedicated Network is a high-performance dedicated service for enterprises, governments and industrial users. It has built a large-scale SRv6 Policy-based cloud dedicated network, offering secure, low-latency connectivity between on-premises data centers, branches and its public cloud—supporting mission-critical tasks like real-time industrial control and large-scale data analytics.

It covers 31 provinces and over 300 cities, it has nearly 50 Tbps inter-provincial bandwidth and over 800 SRv6 Policy-enabled devices. Its backbone uses Segment Routing (SR); some backbone links support 400 Gbps OTN to handle high-bandwidth demands. A multi-path redundant architecture ensures 99.99% availability, reducing business disruptions.

It integrates an intelligent management platform for real-time monitoring of latency, jitter and bandwidth, and connects tens of thousands of enterprise users domestically to eases inter-office

collaboration and cloud migration. Internationally, it links to global cloud providers and overseas data centers for secure cross-border data transmission.

## 6. Emerging Trends and Future Directions

As HP-WANs continue to evolve, driven by emerging requirements from scientific research, high-performance computing, distributed artificial intelligence, and industrial data analytics, several key trends and future directions are shaping the next generation of HP-WANs.

While these capabilities may not yet be widely deployed on RE infrastructures, there is clear interest in the NREN operator community in tracking their development and readiness for production deployment.

### 6.1. Integrated Resource and Network Control

Enhanced integration between resource controllers and network controllers for scheduled services to maximise network efficiency. This tighter integration aims to deliver more granular and efficient control over network resources, enabling dynamic, on-demand bandwidth allocation and optimised resource allocation decisions. Such integration facilitates more effective orchestration of network resources, aligning network performance closely with application requirements

For example, A typical trend for HP-WANs is to satisfy the need of distributed or decentralized AI, which is so-called "scale-across" network between multiple large data centers over long distance. The integration of network resources and applications should touch the underlying resources, such as optical transport layer [HIC-OTN], and routing policies configured by network controllers, as well as service requests. This integration needs the coordination of optical resource management platform, the network controller, and the service orchestrator.

### 6.2. Intent-Based Networking and Automation

Intent-based networking (IBN) and automation technologies are increasingly used in the role in the management and orchestration of HP-WANs. IBN allows network administrators to define desired network states or outcomes, with automated systems translating these intents into actionable network configurations. As discussed earlier, platforms such as ESnet's SENSE provide valuable practical demonstrations of how intent-driven orchestration can significantly enhance agility, scalability, and operational efficiency.

### 6.3. Network Signalling

As the scale and complexity of HP-WAN deployments grow, efficient signalling mechanisms become increasingly critical, especially when running HPWAN services over shared public infrastructure.

Applications may want to signal their desired bandwidth to the network, enabling more precise rate negotiation and collaborative congestion control, to achieve a targeted competition time for the data transfer.

Therefore, efficient and scalable signalling approaches are vital for dynamic resource allocation in HPWAN environments. Effective protocols must support rapid dissemination of resource states and swift propagation of requests between network components, minimising latency and overhead.

Desirable signalling mechanisms in HPWAN include extensibility, low overhead, real-time responsiveness, and robustness, supporting diverse technologies and ensuring reliable, high-performance communication.

### 7. IANA Considerations

This document makes no requests for action by IANA.

### 8. Security Considerations

The security requirements for HPC networks, particularly in inter-data center scenarios, are crucial to ensuring the integrity, confidentiality, and availability of sensitive data and computational resources. These requirements are stringent due to the high-value and often sensitive nature of the data processed within HPC systems, such as research data in fields like national defense, pharmaceuticals, and climate science.

### 9. Acknowledgements

This document was partly motivated by the discussion occurring on the IETF hp-wan@ietf.org mailing list.

The authors would like to thank Gorrry Fairhurst and Zahed Sarker for their reviews and suggestions.

### Contributors

The following authors contributed significantly to this document:

Nicholas Race  
Lancaster University  
United Kingdom  
Email: n.race@lancaster.ac.uk

## Normative References

## Informative References

- [HIC-OTN] Sun, J., "Decentralized Training over 100km Based on Optical Transport Network for Artificial Intelligence", 2024, <<https://ieeexplore.ieee.org/abstract/document/10926085>>.
- [RFC5040] Recio, R., Metzler, B., Culley, P., Hilland, J., and D. Garcia, "A Remote Direct Memory Access Protocol Specification", RFC 5040, DOI 10.17487/RFC5040, October 2007, <<https://www.rfc-editor.org/info/rfc5040>>.
- [RFC5041] Shah, H., Pinkerton, J., Recio, R., and P. Culley, "Direct Data Placement over Reliable Transports", RFC 5041, DOI 10.17487/RFC5041, October 2007, <<https://www.rfc-editor.org/info/rfc5041>>.
- [RFC5042] Pinkerton, J. and E. Deleganes, "Direct Data Placement Protocol (DDP) / Remote Direct Memory Access Protocol (RDMAP) Security", RFC 5042, DOI 10.17487/RFC5042, October 2007, <<https://www.rfc-editor.org/info/rfc5042>>.
- [RFC5043] Bestler, C., Ed. and R. Stewart, Ed., "Stream Control Transmission Protocol (SCTP) Direct Data Placement (DDP) Adaptation", RFC 5043, DOI 10.17487/RFC5043, October 2007, <<https://www.rfc-editor.org/info/rfc5043>>.
- [RFC5044] Culley, P., Elzur, U., Recio, R., Bailey, S., and J. Carrier, "Marker PDU Aligned Framing for TCP Specification", RFC 5044, DOI 10.17487/RFC5044, October 2007, <<https://www.rfc-editor.org/info/rfc5044>>.
- [RFC6580] Ko, M. and D. Black, "IANA Registries for the Remote Direct Data Placement (RDDP) Protocols", RFC 6580, DOI 10.17487/RFC6580, April 2012, <<https://www.rfc-editor.org/info/rfc6580>>.

- [RFC6581] Kanevsky, A., Ed., Bestler, C., Ed., Sharp, R., and S. Wise, "Enhanced Remote Direct Memory Access (RDMA) Connection Establishment", RFC 6581, DOI 10.17487/RFC6581, April 2012, <<https://www.rfc-editor.org/info/rfc6581>>.
- [RFC6703] Morton, A., Ramachandran, G., and G. Maguluri, "Reporting IP Network Performance Metrics: Different Points of View", RFC 6703, DOI 10.17487/RFC6703, August 2012, <<https://www.rfc-editor.org/info/rfc6703>>.
- [RFC7306] Shah, H., Marti, F., Nouredine, W., Eiriksson, A., and R. Sharp, "Remote Direct Memory Access (RDMA) Protocol Extensions", RFC 7306, DOI 10.17487/RFC7306, June 2014, <<https://www.rfc-editor.org/info/rfc7306>>.
- [UEC] Ultra, "Ultra Ethernet Specification v1.0.1 ", 2025, <<https://ultraethernet.org/uec-1-0-spec>>.

## Authors' Addresses

Daniel King  
Lancaster University  
Email: [d.king@lancaster.ac.uk](mailto:d.king@lancaster.ac.uk)

Tim Chown  
Jisc  
Email: [tim.chown@jisc.ac.uk](mailto:tim.chown@jisc.ac.uk)

Chris Rapier  
Pittsburgh Supercomputing Center  
Email: [rapier@psc.edu](mailto:rapier@psc.edu)

Daniel Huang  
ZTE Corporation  
Email: [huang.guangping@zte.com.cn](mailto:huang.guangping@zte.com.cn)

Kehan Yao  
China Mobile  
Email: [yaokehan@chinamobile.com](mailto:yaokehan@chinamobile.com)