100 Gb/s DWDM InfiniBand Transport over up to 40 km

Ulrich Schlegel and Klaus Grobe
ADVA AG Optical Networking, Fraunhoferstr. 9a, 82152 Martinsried, Germany
David Southwell
Obsidian Strategics, 2966 Parsons Road NW, Edmonton, Alberta, Canada T6N 1B1
KGrobe@ADVAoptical.com

Abstract: This paper describes InfiniBand-over-WDM transport with aggregated capacity of 100 Gb/s and distances of up to 100 km. Total capacity can be extended to 80 x 10 Gb/s which covers upcoming InfiniBand Quad Data Rate 12X interfaces and extends them to regional applications at full throughput. This allows regionally dispersed clusters of HPCs to be connected natively with high-speed InfiniBand links. These clusters can then be connected to national NRENs and the related global infrastructure either via ultra long-haul InfiniBand range extension, or via (multiple) 10GbE links.

***Keywords*—InfiniBand range extension, InfiniBand-over-WDM distance transport.**

1. INTRODUCTION

Grid computing is one of the relevant applications within National Research and Educational Networks (NRENs). An increasing number of applications, for example in the health care and scientific area, require ultra high-capacity, low-latency connections between several locations which are regionally dispersed, i.e., with distances below 200 km. This is true for both, storage and CPU clusters.

InfiniBand can be used for high-capacity, low-latency connections. So far, it lacked distance capabilities in order to extend it natively into regional distances. Here, we report on InfiniBand-over-WDM range extension and distance transport which allows any InfiniBand capacity to be transported up to 100 km without protocol conversion. Capacity-wise, this covers up to 80 x 10G InfiniBand (i.e., more than 6 InfiniBand QDR 12X links on a single fiber pair).

2. HIGH-CAPACITY INFINIBAND-OVER-WDM RANGE EXTENSION AND TRANSPORT

The first-ever 100 Gb/s DWDM (Dense Wavelength Division Multiplexing) InfiniBand-over-distance transport has been demonstrated at the 2008 SuperComputing Conference in Austin, Tx. The joint solution combines new Obsidian Strategics™ Longbow C Series™ InfiniBand range extenders and high-capacity ADVA Optical Networking optical transport technologies, uniquely enabling a powerful new class of storage-area network, cluster-clustering, remote visualization and data center-expansion applications.

The resulting metro and national network architecture is shown in the figure below. It is based on massive intra-metro InfiniBand-over-WDM coupling, together with fewer high-speed connections into the NRENs. These connections are based on 10GbE today (or STM-64), they will migrate to 40GbE and/or 100GbE in the (near) future.

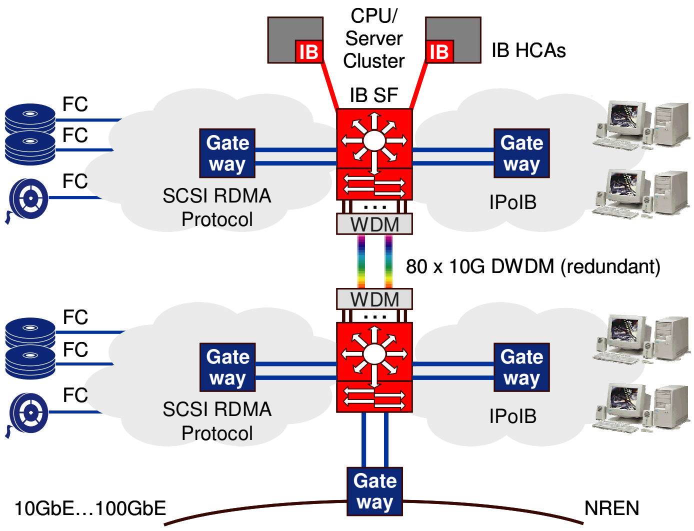

Fig. 1. Metro and national research and educational network. IB SF: InfiniBand Switch Fabric. HCA: Host Channel Adaptor.

InfiniBand is primarily used as a fast and scalable fabric interconnect for linking compute and storage nodes within a high-performance data center. The protocol is usually limited to very short-reach cables unsuitable for applications between buildings or campuses. The technology combination offers native InfiniBand performance over tens of kilometers – scaling to up to 80 InfiniBand 4X SDR links through a single dedicated fiber pair.

Many-core processors, high-capacity disks and solid-state storage technologies are resulting in exponentially growing volumes of data being accessed ever more rapidly. However, these data sets are increasingly trapped by the walls of their data centers – traditional long-haul networking is not keeping up and the time required to routinely move Petabytes of data is often prohibitive. Range extension allows the use of InfiniBand as an alternative to TCP/IP, leveraging faster InfiniBand hardware and bypassing the latency and bandwidth efficiency problems that result from lossy transport protocols carried over high bandwidth-delay product links. Striping parallel InfiniBand links over suitable DWDM equipment provides the means to scale network capacity alongside compute and storage.

InfiniBand has been limited in the past due to the inability to transport the protocol over distance. WDM can deliver protocol-agnostic transport for latency-sensitive protocols like InfiniBand, providing an excellent solution for demanding, high-capacity applications.

The FSP 3000 xWDM platform supports a multitude of data center applications for enterprise and R&E customers. It was developed with the most stringent requirements in mind regarding latency and power consumption typical for these environments. The platform supports the transparent transmission of up to 80 wavelengths carrying up to 10 Gb/s each over unregenerated distances of up to 2,000 km. It also supports 80 x 40 Gb/s (e.g., for upcoming InfiniBand QDR 4X transport) over somewhat shorter distances, and is ready for the upcoming 100 Gb/s per wavelength transport.

Longbow C Series products transparently extend InfiniBand fabrics over up to 100 km of lightpath – fully retaining InfiniBand performance and connection semantics. Longbow C products operate in pairs, transparently range-extending InfiniBand SDR 4X at full bandwidth and sub-microsecond device latencies over dedicated fibers or DWDM channels. The C Series comprises five different models that vary in optical interfaces and range-extension capability. The range extension is based on providing sufficient buffer credits.

The 100 Gb/s InfiniBand-over-distance demonstration was an extension of earlier work conducted in Germany at the University of Stuttgart's High Performance Computing Center (HLRS), which provides systems, tools and expertise for public and private supercomputing development projects. This work was done in June 2008, when Obsidian and ADVA Optical Networking demonstrated an integrated capability to extend InfiniBand connections to 40 km without performance degradation.

3. CONCLUSION

InfiniBand-over-WDM range extension and distance transport has been demonstrated at aggregated bit rates of 100 Gb/s and over distances of up to 40 km for the first time. This capability allows to natively extend high-capacity InfiniBand links into regional networks. Possible applications include high-capacity, low-latency storage and cluster connections between campuses in large metropolitan areas.

AUTHOR BIOGRAPHIES

Klaus Grobe received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Leibniz University, Hannover, Germany, in 1990 and 1998, respectively. From 1990 to 1993, he worked for an offshore engineering company and developed fiber-optic telemetry systems for deep sea exploration. From 1998, he worked for national and pan-European network operators where he planned and designed WDM transport networks. Since 2000, he is with ADVA AG Optical Networking where he is now responsible for advanced technologies.

Dr. Grobe authored and co-authored more than 50 scientific publications as well as three book chapters on WDM and PON technologies. He is member of the IEEE LEOS (Laser & Electro-Optics Society), the German VDE ITG (German Association for Electrical, Electronic & Information Technologies), and the ITG Study Group 5.3.3 on Photonic Networks.

Uli Schlegel received the Dipl.-Ing. Physikalische Technik (engineering degree) from University of Applied Sciences (TFH) Berlin, Germany. From 1999 to 2000 he worked with Lawrence Berkeley National Laboratory (LBL) at the Center of X-Ray Optics (CXRO) on the determination of optically related fundamental physical constants of Ruthenium.

In 2000, he joined ADVA AG Optical Networking and held various positions within the company, including optical engineering, technical support management and systems consultant for WDM and SAN networks.

Currently Mr. Schlegel is working with the application and solution management division focusing on Storage Area Networking (SAN).

David Southwell received M.Eng (1990) and Ph.D degrees (1994) in Electronics Engineering from the University of York, UK, working with British Telecom. Subsequently he lectured at the University of Alberta (Edmonton, Canada) and co-founded several technology startups including YottaYotta (long-haul storage area networking) and Obsidian Research Corporation (InfiniBand technologies including long-haul optical connectivity). David currently serves as President and Co-CEO of Obsidian Strategics.