SWIFT: A High Capacity Wavelength-striped Optically Switched Network with Electronic Control

Michael Dales
University of Cambridge Computer Laboratory
15 JJ Thomson Avenue
Cambridge, CB3 0FD, UK

Madeleine Glick
Intel Research Cambridge
15 JJ Thomson Avenue
Cambridge, CB3 0FD, UK

***Abstract*— High capacity optical networks are well established in the long-haul networks arena, with new developments, such as all optical switching, being pursued to further benefit these networks. Research at Intel has shown that high capacity optical links can offer advantages for short range networks such as device interconnects and computer clusters. However, there are many challenges in taking the technologies used in the long-haul networks and applying them to short range networks.**

Here we present a semisynchronous, multi-wavelength striped optical network architecture with an electronic control system, suitable for short range networks from SANs and small LANs down to chip interconnects. We outline the challenges involved, and examine how our short range architecture overcomes them. We also present SWIFT, the Semisynchronous Wavelength Striped Interconnect Flexible Testbed, a novel optical testbed designed to enable evaluation of our network design with a combination of physical medium evaluations, low-level synthetic traffic and real world applications.

I. Introduction

Optical networking technologies are being agressively pursued in the WAN and MAN domains. The low loss characteristics, physical robustness, and ability to carry very high bandwidth links makes fibre optic cable an ideal medium for long-haul networks. Recently, the advent of Wavelength Division Multiplxing (WDM) has allowed fibre to scale much better than copper cable, allowing multiple channels of information to be sent down a single fibre: current top end Dense WDM (DWDM) hardware can carry up to 160 channels of 10 Gbps each [7]. In addition to advances in packing channels down a single fibre, researchers are pursuing all optical switching [1], signal regeneration, and buffering techniques [3].

Current optical networks still require that packets undergo optical to electrical conversion at switches, but current research aims to provide core networks with the ability to have packets remain in the photonic domain from ingress to egress, providing advantages in terms of reduced latency and greatly reduced power consumption. A major obstacle to building a general purpose LAN from all-optical switches is that there is currently no practical solution to the problem of optical memory; it is not possible to buffer packets optically for an arbitrary length of time at the switches. Without this buffering, an all-optical network requires that a light-path be constructed across all the switches on the path over the network for each packet transmission. On a large LAN, this would lead to unfeasible scheduling requirements and synchronization problems. Thus the technology required to make an all-optical LAN is not currently available.

Given the success of WDM optical networks and the potential advantages of optical switching, it is a logical step to try to apply these technologies to small scale networks: from SANs and smaller application specific networks, computer cluster interconnects within supercomputers, all the way down to device interconnects within computer systems [2], [6]. All these domains are constantly seeking faster interconnects; for example, 10 gigabit Ethernet is already in development, and people are considering optical backplanes for computer systems.

The aim of SWIFT is to investigate the design and construction of a high capacity wavelength-striped optically switched network suitable for short range applications taking advantage of modern advances in optical components and networking. By utilising WDM to stripe data from a single packet across multiple wavelengths (effectively creating a photonic bus) we can increase the link bandwidth and reduce the latency for the entire packet. The challenges include retaining the network speed seen in long-haul networks whilst scaling down the network components in terms of size, power, and cost; designing a suitable media access mechanism to allow the network to be efficiently used; and making the network self-configuring, removing the need for support engineers when adding or removing devices from the network.

Designing this type of network requires evaluation to be carried out at many levels: examining the optical properties of the switch, testing how close the network can perform to theoretical throughput estimates, and how the network performs under real world loads using conventional networking stacks (e.g., TCP/IP). To enable these evaluations we have built a flexible testbed implementation of the initial network architecture which provides scope for testing at all levels. The testbed is built in a modular fashion using reconfigurable hardware to enable fast alteration for evaluating design alternatives. No other optical network testbed in this field provides such a wide range of evaluation possibilities.

II. The SWIFT Architecture

We propose the SWIFT network architecture as a way to build a high speed, low latency network around a central optical switch. The SWIFT network uses WDM to provide parallel

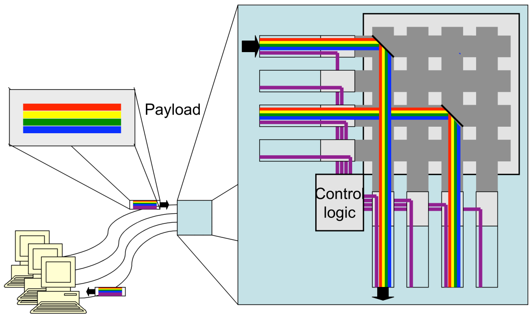

Fig. 1. Overview of the SWIFT network

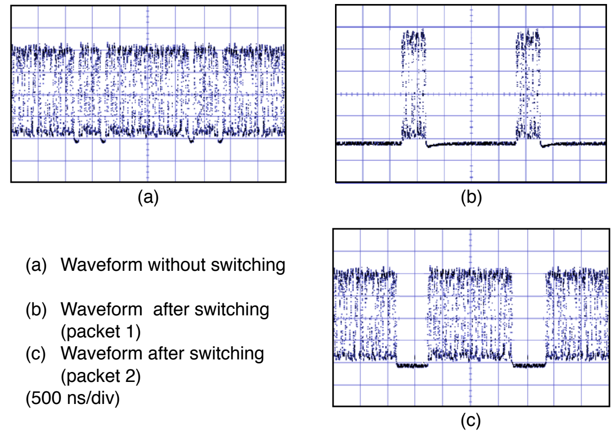

Fig. 2. Demonstration of switching packets, showing the input to the switch, and the output on one port, and the remaining signal.

data transmission over multiple wavelengths simultaneously, allowing for higher bandwidth transmission. An additional wavelength is used to carry asynchronous control communication between the switch and the nodes. Having the control information on a separate wavelength allows it to be handled electronically at the switch, whilst the data wavelengths pass through optically (without any optical-electrical conversion). Access to the network is managed by the switch, which constructs light-paths between nodes through the switching fabric. These light-paths are valid for a fixed time, called a *slot*, and the switch signals to nodes when they can transmit in a given slot. During a slot a node may send a single data packet.

An overview of the architecture can be seen in Figure 1. Nodes stripe data packets over $n - 1$ wavelengths in the network interface, with the remaining wavelength going to the control logic, effectively providing a point-to-point communication channel between the switch and each node. All the data wavelengths are simultaneously optically switched to the appropriate output port on the switch. The switch fabric is managed by the control logic, which accepts transmission requests from nodes over the respective control channel and will set up the switch fabric appropriately in the next unallocated slot.

A. The SWIFT Switch

The switch in the SWIFT network has two main components: the electronic control layer, and the optical switching fabric. The electronics control part is the "brain" of the network, which decides how to configure the network for the next time slot, based on the requests for transmission it has seen from the connected nodes. The scheduling of the network is described in more detail in the next section. The electronics are also responsible for general network configuration tasks. For instance, the switch must know the distance of the links between it and each node in order to compensate for delays when advertising slots, and the switch will gather this information each time a node is added to the network (as demonstrated in [5]).

At an abstract level, the network architecture is agnostic of the actual optical switching fabric: the fabric could be a cross-bar switch, a ring based fabric, or any other arrangement. From the point of view of the nodes they simply request permission to send packets using the control channel, and once they receive that permission from the switch they transmit their packets. We are currently building a cross-bar fabric for our initial prototype (see Section III).

Although the layout of the fabric is not important at the high level, the actual response characteristics of the components the switch is constructed from is very important. There are a number of technologies available for switching light, ranging from mechanically moved mirrors to semiconductor devices [8]. For SWIFT we have opted to use Semiconductor Optical Amplifiers (SOAs), which can turn a light beam on or off based on an electrical input on nanosecond timescales, providing a very small switching overhead. Multiple SOAs can be joined together in different arrangements to produce a variety of switching fabrics.

In a separate experiment we have demonstrated the feasibility of using SOAs by testing them with packets sent at 80 Gbps using eight wavelengths each operating at 10 Gbps [4]. Alternate large and small packets were sent back to back through an SOA based switch, and the results of separating the two packet types can be seen in Figure 2. Stable operation at a bit error rate of 10^{-12} with no error floor was observed. Note that the aggregate bit-rate in this experiment was limited only by equipment availability rather than any fundamental technology limitations.

B. Network Access Scheduling

As there is currently no practical way to buffer optical packets for an arbitrary length of time, the switch needs to ensure that a clear light-path can be constructed in the switching fabric between the sender and receiver. This requires that the control electronics in the switch schedule packet requests such that they do not clash in the switching fabric. Depending on the type of fabric used it may be that multiple paths are possible through the switch based on who the sender/receiver pairs are. To simplify access scheduling, the SWIFT network divides time into a series of fixed sized slots, with each slot representing a period of time where the switch

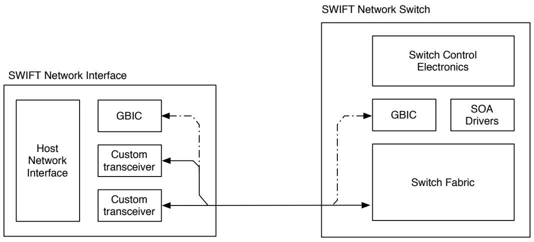

Fig. 3. Overview of the initial demonstrator

is configured a particular way to allow transmission between a given set of sender/receiver pairs. Once the switch has been configured for a slot, a node may transmit at any point within that slot, as long as it has sent all its data through the switch before the slot ends.

The length of slots is specified to allow a single packet of up to a reasonable size (compared to existing network technologies, such as Ethernet), and will be of the order of several microseconds. One challenge with high speed switched optical networks is that the end systems need time to perform clock synchronization on every packet, so each packet needs to be prefaced by a guard band which contains idle signals that can be used by the receiver to lock on to the transmitter's clocking signal. The size of the guard band, which may need to be as long as a microsecond, will effect the efficiency of the network, thus the challenge here is to find appropriate low cost technology with acceptable synchronization time.

To use the network, a node will make a request of the switch over the control channel. This request may be for a single packet, or it maybe a complex request, such as permission to send a series of packets over a certain time with a fixed interval. A request per packet adds significant to the latency of the network. We will therefore investigate techniques such as QoS provisioning and pre-allocation of slots based on historical usage in order to remove the latency overheads associated with requests over the control channel.

III. Testbed Implementation

We have constructed a testbed to demonstrate the principles of the SWIFT network architecture and to enable practical evaluation at many levels. We use custom end point interfaces, which use two data wavelengths each running at 1 Gbps, connected to a 3×3 SWIFT switch based on a cross-bar switching fabric. A block diagram of the switch and one of the interfaces can be seen in Figure 3.

The current switch fabric is built using a series of individual SOAs connected together and controlled by electronics implemented on an FPGA. The electronics in the switch is split into three parts: the slot scheduling logic, an interface to the SOAs, and interfaces to the individual nodes on the network. The node interface is connected to a GBIC laser transceiver which is multiplexed/demultiplexed with the data wavelengths at the edge of the switch. The end point interface is implemented on an FPGA with an internal microprocessor using a combination of hardware and software. For testing purposes the microprocessor can be used to generate synthetic traffic, or communicate to a PC over a gigabit Ethernet link.

Slots are currently specified as 10 μs with 1 μs of preamble for clock synchronization, which allows nodes to transmit up to 2250 bytes per slot. The theoretical maximum transmission speed for the testbed is 1.8 Gbps if all slots are used by a single node. Initially this can not be achieved as nodes need to request slots as they need them, but the aim is to use the testbed to develop the control plane to reduce this overhead.

Having constructed the testbed we plan on using it to evaluate the optical properties of the switching fabric, throughput evaluation based on synthetic traffic driven by software in the FPGAs on the end point interfaces, and real application behaviour using PCs connected to the end point interfaces. No previous optical testbed has provided such a range of evaluation possibilities.

IV. Summary

In this abstract we have described the SWIFT network architecture, which aims to provide a high-speed optical network for short ranges networks. We have described how we plan to use WDM to provide high bandwidth transmission, and use an all optical data path through the switch to reduce latency and power consumption. We have built a novel small scale prototype of this network which provides us with a range of opportunities for experimentation not found in any other single testbed, with which we plan to run experiments to allow us to analyze how the network performs and investigate possible enhancements to the network management.

References

[1] Vilson R. Almeida, Carlos A. Barrios, Roberto R. Panepucci, and Michael Lipson. All-optical control of light on a silicon chip. *Nature*, pages 1081–1084, October 2004.

[2] Dawei Huang, Theresa Sze, Anders Landin, Rick Lytel, and Howard L. Davidson. Optical Interconnects: Out of the Box Forever? *IEEE Journal of Selected Topics in Quantum Electronics*, 9(2):614–623, March/April 2003.

[3] P. Ku and C. Chang-Hasnain. Semiconductor All-Optical Buffers Using Quantum Dots in Resonator Structures. In *Optical Fiber Communications Conference*, pages 76–78, March 2003.

[4] T. Lin, W. Tang, K. A. Williams, G. F. Roberts, R. V. Penty, I. H. White, M. Glick, and D. McAuley. 80Gb/s optical packet routing for data networking using FPGA based 100 Mb/s control scheme. In *European Conference on Optical Communication*, 2004.

[5] T. Lin, K. A. Williams, P. V. Penty, I. H. White, M. Glick, and D. McAuley. Self-Configuring Intelligent Control for Short Reach 100GB/s Optical Packet Routing. In *Optical Fiber Communications Conference*, 2005.

[6] E. Mohammed, A. Alduino, T. Thomas, H. Braunisch, D. Lu, J. Heck, A Liu, I. Young, B. Barnett, G. Vandentop, and R. Mooney. Optical interconnect system integration for ultra-short reach networks. *Intel Technology Journal*, 8:115–128, 2004.

[7] A. R. Pratt, P. Harper, S. B. Alleston, P. Bontemps, B. Charbonnier, W. Forysiak, L. Gleeson, D. S. Govan, G. L. Jones, D. Nesset, J. H. B. Nijhof, I. D. Phillips, M. F. C. Stephens, A. P. Walsh, T. Widdowson, and N. J. Doran. 5,745 km DWDM Transcontinental Field Trial Using 10 Gbit/s Dispersion Managed Solitons and Dynamic Gain Equalization. In *Optical Fiber Communications Conference*, pages 1–3, March 2003.

[8] Rajiv Ramaswami and Kumar N. Sivarajan. *Optical Networks: A Practical Perspective*. Morgan Kaufmann Publishers, 2nd edition, 2002. ISBN: 1558606556.