Demonstration of High-Speed MIMO OFDM Flexible Bandwidth Data Center Network

Philip N. Ji, Ting Wang, Dayou Qian, Lei Xu, Yoshiaki Aono* and Tsutomu Tajima*
NEC Laboratories America, Inc., 4 Independence Way, Princeton, New Jersey 08540, USA
*Optical Network Division, NEC Corporation, 1131 Hinode, Abiko, Chiba 270-1198, Japan

Christoforos Kachris, Konstantinos Kanonakis and Ioannis Tomkos
Athens Information Technology (AIT), 0.8 km Markopoulo Av., 19002, Athens, Greece
pji@nec-labs.com

Abstract: We propose a novel datacenter network architecture utilizing OFDM and parallel signal detection technologies and efficient subcarrier allocation algorithms. Fast, low latency, fine granularity, bandwidth flexible, and low power consumption MIMO switching is demonstrated experimentally.

© 2012 Optical Society of America

OCIS codes: (200.4650) Optical interconnects; (060.4265) Networks, wavelength routing

1. Introduction

As the global Internet traffic growing exponentially, the data centers, which host many Internet application servers, are facing rapid increase in bandwidth demands. Furthermore, next generation data center networks (DCN) need to achieve low latency, high throughput, high flexibility, high resource utilization, low power consumption, and low cost. Optical technology has been adopted in DCN due to its high bandwidth capacity. However, it is mainly used for point-to-point links, while the intra DCN interconnects are still based on electrical switching fabric, which has high power consumption and limited bandwidth capacity [1]. In recent years, several hybrid optical/electrical or all-optical interconnect schemes have been proposed [2-9]. Many of them rely on large scale fiber cross-connect (FXC) [2-4] or multiple wavelength-selective switches (WSS) [4, 5], which are costly and have slow switching speed (at ns level). Having a large scale FXC also present an undesirable potential single source-of-failure. A recent work in [6] uses silicon microring WSS and semiconductor optical amplifier-based switch to achieve ns scale switching, making all-optical packet level routing possible. However the key components are not commercially available and have poor scalability. Several other architectures use cyclic arrayed waveguide grating (CAWG) as the core optical router [7-9], the cyclic wavelength arrangement characteristics of the CAWG avoids the wavelength contention and eliminates core FXC. However these architectures require costly tunable wavelength converters and do not allow bandwidth resource sharing among the connections. Some of them also require electrical or optical buffer.

In this paper we propose and experimentally demonstrate a novel all-optical inter-rack and inter-server DCN architecture that combines the passive CAWG core router with orthogonal frequency division multiplexing (OFDM) modulation and parallel signal detection (PSD) technologies. The architecture achieves fast (ns speed), low latency, low power consumption, multiple-input multiple-output (MIMO) switching while allowing fine granularity bandwidth sharing without requiring FXC or WSS.

2. MIMO OFDM flexible bandwidth DCN architecture

The schematic of the MIMO OFDM DCN architecture is illustrated in Fig. 1. It contains \( N \) racks, each accommodating multiple servers connected through a top-of-the-rack switch (ToR). Inter-rack communications are performed by interconnecting these ToRs through the DCN. The inter-rack signals at each rack are aggregated and sent to a transmitter, which contains an OFDM modulator that modulates the aggregated signals into up to \( N \) OFDM data streams with appropriate subcarrier assignments, one for each destination rack. These OFDM data streams are converted to WDM optical signals through an array of \( N \) directly modulation lasers or \( N \) sets of laser/modulator with different wavelengths. These signals are combined through a WDM combiner to form an OFDM-modulated WDM signal and sent to an \( N \times N \) CAWG. Through the cyclic non-blocking wavelength arrangement of the CAWG, the WDM channels are routed to the respective output ports for the destination racks. Each optical receiver receives one WDM channel from each input port. Through a centralized subcarrier allocation scheme, the WDM channels at each receiver do not have subcarrier contention, so that a single photo-detector (PD) can receive all WDM channels simultaneously through the PSD technology [10], which has been demonstrated in OFDM WDM-based optical networks [11]. The received OFDM signal is then demodulated back to the original data format and send to the appropriate servers through the destination ToR.
When a new switching state is required, the OFDM modulators execute the new subcarrier assignments determined by the centralized controller, and the respective lasers turns on and off to generate new OFDM WDM signals. This can be achieved at sub-ns level, making this system feasible for packet operation. All switched signals take exactly one hop (i.e. passing through the switch only once), therefore the latency is very low and uniform. This architecture does not require switching at the core optical router, and does not require WSS or FXC, therefore the optical components have lower power consumption and lower cost compared with other optical DCN architectures.

In this architecture, the signals from each rack can be switched to multiple racks simultaneously, and each rack can receive signals arriving from multiple racks simultaneously, delivering the MIMO switching capability. Since electrical OFDM modulation is used, the switching granularity can be as fine as tens of Mb/s, achieving flexible bandwidth allocation and efficient spectrum utilization. The PSD-based OFDM receiver does not require guard band between subcarrier bands from different sources. OFDM modulation also provides the capability to change the bandwidth allocation and efficient spectrum utilization. The PSD-based OFDM receiver does not require guard band

If the number of the racks increases, it is not cost efficient to install N lasers at each transmitter, this is also not necessary because it is not likely that each rack needs to communicate with all other racks simultaneously. Therefore the N fixed wavelength lasers in the transmitter can be replaced with fewer tunable lasers. The switching speed will then be determined by both the laser on/off speed and tuning speed, which can also be realized at ns level [12].

This architecture can be further extended to provide direct communication between “super servers” that require constant large bandwidth. Dedicated ports can be reserved for these servers to bypass the ToR (Fig. 1(a)).

3. Experimental demonstration of the MIMO OFDM flexible bandwidth DCN

A testbed is constructed to demonstrate the flexible MIMO interconnect capability of the proposed architecture. The optical core router is an 8×8 CAWG with 100 GHz spacing. Each transmitter contains two tunable external cavity lasers with 10 GHz intensity modulators. The OFDM signals are generated by a Tektronix 7122C arbitrary electrical OFDM modulation is used, the switching granularity can be as fine as tens of Mb/s, achieving flexible bandwidth allocation and efficient spectrum utilization. The PSD-based OFDM receiver does not require guard band between subcarrier bands from different sources. OFDM modulation also provides the capability to change the bandwidth allocation and efficient spectrum utilization. The PSD-based OFDM receiver does not require guard band

If the number of the racks increases, it is not cost efficient to install N lasers at each transmitter, this is also not necessary because it is not likely that each rack needs to communicate with all other racks simultaneously. Therefore the N fixed wavelength lasers in the transmitter can be replaced with fewer tunable lasers. The switching speed will then be determined by both the laser on/off speed and tuning speed, which can also be realized at ns level [12].

This architecture can be further extended to provide direct communication between “super servers” that require constant large bandwidth. Dedicated ports can be reserved for these servers to bypass the ToR (Fig. 1(a)).

### Table 1. Wavelength and subcarrier assignments in the experiment (SC: subcarrier).

<table>
<thead>
<tr>
<th>Test</th>
<th>From ToR 1</th>
<th>From ToR 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SC 100-450</td>
<td>SC 300-450</td>
</tr>
<tr>
<td></td>
<td>1549.32 nm</td>
<td>1551.72 nm</td>
</tr>
<tr>
<td></td>
<td>(To ToR 3)</td>
<td>(To ToR 3)</td>
</tr>
<tr>
<td>SC</td>
<td>SC 451-800</td>
<td>SC 850-950</td>
</tr>
<tr>
<td>QPSK</td>
<td>193.5 THz</td>
<td>193.1 THz</td>
</tr>
<tr>
<td>2</td>
<td>SC 200-600</td>
<td>SC 30-150</td>
</tr>
<tr>
<td></td>
<td>16QAM</td>
<td>700-900</td>
</tr>
<tr>
<td></td>
<td>1552.52 nm</td>
<td>1550.12 nm</td>
</tr>
<tr>
<td></td>
<td>(To ToR 6)</td>
<td>(To ToR 3)</td>
</tr>
<tr>
<td>SC</td>
<td>SC 150-150</td>
<td></td>
</tr>
<tr>
<td>QPSK</td>
<td>193.4 THz</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SC 50-150</td>
<td></td>
</tr>
<tr>
<td></td>
<td>193.7 THz</td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>SC 250-250</td>
<td></td>
</tr>
<tr>
<td>QPSK</td>
<td>193.2 THz</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SC 700-700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>193.3 THz</td>
<td></td>
</tr>
</tbody>
</table>

Test1: TO ToR3

Test2: TO ToR6

Test1 & 2: TO ToR6

Fig. 2. Measured RF spectra before and after the PSD receiver.
The measured RF spectra (Fig. 2) show that different OFDM signals from the same transmitter can possess overlapping or non-contiguous subcarriers, and through the PSD technology each receiver can successfully detect and receive OFDM signals from multiple sources, provided that these OFDM subcarriers do not overlap. No guard band is required when assigning the OFDM subcarriers. Besides allocating subcarriers, the centralized controller also balances the subcarrier powers among signals from different inputs. The optical spectra of the WDM signals at the transmitter outputs and the receiver inputs are shown on Fig. 3, which confirms the non-blocking cyclic routing operation of the CAWG. The PSD receiver performance is measured and compared with single channel receiver (Fig. 4), no significant degradation is observed in all tests. This shows that PSD technology is feasible for receiving multiple OFDM WDM signals simultaneously using a single PD, and thus realizing MIMO switching.

4. Subcarrier allocation algorithms

The proposed architecture has to be complemented by an appropriate subcarrier assignment scheme to avoid contention at each receiver and to maximize bandwidth resource utilization across the entire DCN. Here we evaluate three algorithms: The first is the simplest scheme that examines the demands at a certain receiver sequentially (SEQ). The second one, a heuristic called Most Subcarriers First (MSF), arranges the connection demands per receiver in decreasing order of requested subcarrier numbers, then allocates subcarriers in order. The rationale is to avoid the case of a large demand appearing at the end of the algorithm execution, thus causing a large gap of unused subcarriers. The third algorithm provides the optimal solution (OPT), where subcarrier assignment is formulated as a knapsack problem with the demands considered as items (with weights equal to their requested number of subcarriers) that must fit in a knapsack of capacity equal to the number of subcarriers per receiver. Since we want to maximize the number of subcarriers assigned versus the ones requested, the profit for each item equals their weight.

Computer simulations are performed for a 32×32 DCN with a maximum of 1024 subcarriers per link under various connectivity degrees (i.e. the percentage among the total number of racks that each rack has active connections with) and traffic loads. The results (Fig. 5) reveal that when connectivity is high, subcarrier utilization for all algorithms is always > 90%, even for high traffic loads. This is due to the fact that each demand consists of only a small number of subcarriers, thus they can be efficiently packed in the receiver subcarrier pool. On the other hand, smaller connectivity degrees imply coarser demands, hence reduced utilization (albeit still at quite high levels, due to the fine granularity offered by the use of OFDM for bandwidth sharing). As expected, OPT and SEQ offer the best and the worst performance respectively. However, it is worth noting that the performance of MSF closely approximates OPT. Therefore MSF, due to its much lower complexity (involving only a sort operation), could be employed for a hardware implementation of the subcarrier assignment mechanism without sacrificing performance.

5. Conclusion

We propose and experimentally demonstrate a novel DCN architecture utilizing OFDM and PSD technologies. This architecture delivers fast, low and uniform latency, and low power consumption MIMO switching. Along with the developed subcarrier allocation algorithms, it achieves flexible bandwidth sharing at fine granularity, making this architecture suitable for all-optical inter-rack and inter-server communication in next generation DCN application.

References