CWDM transceiver for mid-board optics

Paul Rosenberg*[a], Sagi Mathai[a], Wayne Sorin[a], George Panotopoulos[a], Glenn Rankin[a], George Megason[b], David Moore[b], Gregg Combs[c], Kent Devenport[c], Darrell Childers[d], DJ Hastings[d], Mike Tan[a]

[a]Hewlett Packard Labs, 1501 Page Mill Rd, Palo Alto, CA, USA 94304; [b]Hewlett Packard Enterprise, 11445 Compaq Center Dr W, Houston, Tx 77070; [c]Hewlett Packard Enterprise, 3404 E Harmony Rd, Ft Collins, Co 80528; [d]USConec Inc, 1138 25th St. SE, Hickory, NC, USA 28602

ABSTRACT

The need for additional IO bandwidth for data center device interconnection is well established. Optical interconnects can deliver required bandwidth along with energy and space efficiency at a cost that encourages adoption.

To this end, we are developing an optical transceiver incorporating multimode VCSEL emitters in a coarse wavelength division multiplex (CWDM) system capable of transmission at 25Gbps per channel, 100Gbps/fiber, and a maximum aggregate bidirectional data rate of 1.2Tbps. Electrical connection to the transceiver can be made by solder reflow or LGA connector, and optical connection is made by means of a custom optical connector supporting CWDM transmission.

Keywords: parallel optics, wave division multiplex, optical interconnect, CWDM, optical transceiver, opto-electronics, VCSEL, mid-board optics

1. INTRODUCTION

Data centers are increasingly challenged by the need to manage volumes of digital data which are doubling every two years and are predicted to reach 44 zettabytes by the year 2020[1]. There is also big revenue connected to big data. By 2019, worldwide spending for data archiving will exceed $50 billion[2]. In response to the challenge/opportunity presented by big data, Hewlett Packard Enterprise is developing a computer architecture, nicknamed 'the machine', that incorporates an optically interconnected fabric to provide compute nodes with fast access to local and remote shared memory.

Optical links within the machine fabric, known as Gen-Z[3], are implemented as mid-board optic (MBO) transceivers. As the name implies, the MBO configuration moves transceivers from the outer faces of a switch or server enclosure onto the interior system printed circuit board (PCB), typically adjacent to a connected IC. The MBO approach offers several benefits; 1) improved signal integrity; 2) simplified assembly by eliminating panel mount transceiver cages; 3) improved thermal performance; 4) improved chassis EMI/EMC performance by eliminating panel cuts; 5) cost savings from integrated circuit (IC) integration; 6) cost saving due to direct transceiver solder attach and elimination of an electrical connector. Potential drawbacks of the MBO architecture include; 1) reduced flexibility in system configuration since optical transceivers cannot easily be changed after the enclosure is buttoned up; 2) need for an additional optical jumper to bridge between transceiver and the front panel; 3) reduced serviceability. MBO architectural concepts are still relatively new, and it will take some time to determine the balance between these pros and cons.

2. PACKAGE PLATFORM OVERVIEW

2.1 General Requirements

Five key requirements for the transceiver were established during the project definition phase; 1) 100Gbps bandwidth per fiber; 2) the electro-optical subassembly must be capable of withstanding exposure to a standard surface mount solder reflow process; 3) optical alignment processes must be non-active; 4) no use of active temperature control methods; 5) maximum density of bandwidth per PCB area in close adjacency to a flip chipped IC. The need to control transceiver cost influenced decision making at every stage of the design process. The common industry cost metric has been $1/Gbps, and our team was constantly reminded of the need to beat that metric significantly

The requirements led directly to the final form of the transceiver. In order to satisfy the 100Gbps per fiber requirement, multimode CWDM was chosen. The need for solder reflow compatibility, ruled out pig-tailing and led to the development of a custom optical connector system. In order to facilitate vision-aided optical alignment, the optical actives were designed with integrated collimating lenses. An early schematic of the transceiver concept is shown in figure 1.

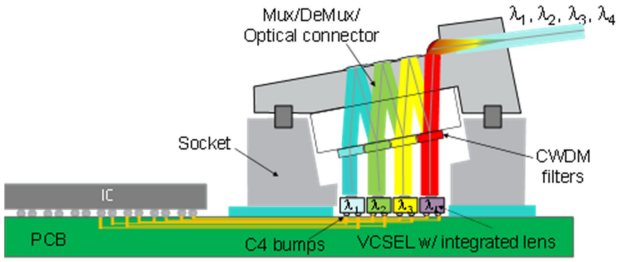


Figure 1. Schematic of prototype co-packaged CWDM optical transceiver module showing MUX/DEMUX, opto-electronic arrays, filters and integrated optical fiber.

2.2 Opto-electronic subassembly

At the heart of our CWDM transceiver is a build-up organic substrate. Electrical connection to a system PCB is made through a 576 pad BGA solder ball arrangement on 1mm pitch. Though designed for surface mount connection to the system PCB, the BGA contacts are also compatible with LGA electrical connectors. The opposite side of the build-up substrate is patterned with several flip chip solder pad arrays for assembling the IC and opto-electronic (OE) arrays. Emitter and detector optical arrays are attached using standard pick and place and solder reflow processes that achieve passive alignment accuracy between the arrays of better than 5um and 0.5 degree tilt. SAC-305 solder is used to avoid the cost, higher temperature, and reflow atmosphere control associated with AuSn eutectic solder. The use of SAC-305 also simplifies assembly, allowing all active and passive components to be reflow attached during the same standard surface mount solder process.

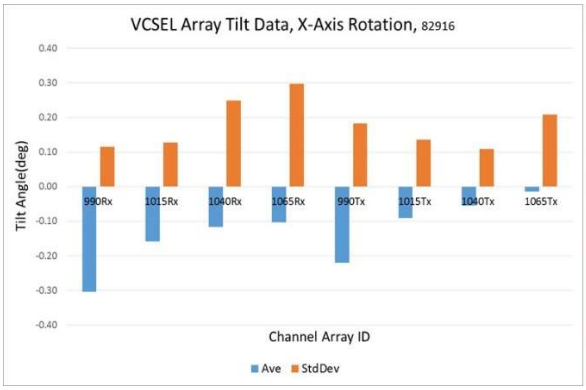
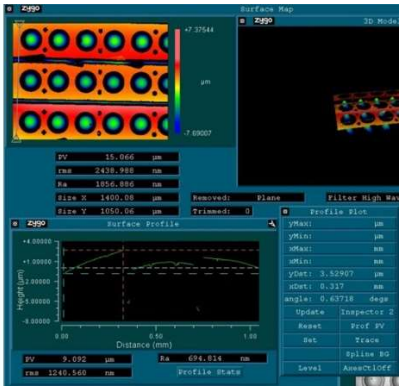


Figure 2a. VCSEL array flip chip tilt angle about X axis.



2b. Array flatness measurement data.

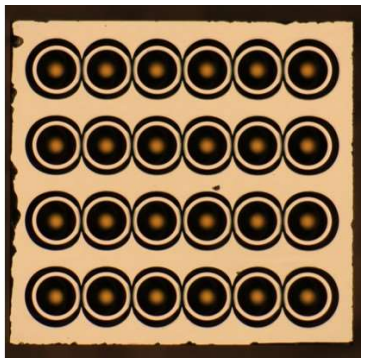
2.3 Active electro-optical devices

Bottom emitting VCSELs (990, 1015, 1040 & 1065nm) were chosen for flip-chip solder reflow compatibility. The integration of a collimating lens on each VCSEL eliminates the need for a separate collimating lens array component, as well as an additional precision alignment process. The large collimated VCSEL output beam reduces x-y alignment sensitivity to the next optical element, while requiring more precise angular control. The VCSELs are nominally of the same design with modifications to the distributed Bragg reflector (DBR), oxide layer, and quantum well layer to adjust for the four emission wavelengths over a range 75nm.

Strain compensation allows the use of highly strained InGaAs quantum wells resulting in higher differential gain and better over-temperature performance[4]. A 25nm channel spacing accounts for manufacturing variation among the VCSELs and filters, and eliminates the need for VCSEL wavelength temperature control, thus boosting yield and

reducing cost. The photodetectors are implemented as a 2D array of lensed, bottom-entry InGaAs PIN photodiodes (PD). The PDs have a 3dB bandwidth of 27 GHz and a responsivity of 0.6A/W.

Figure 3a. 1X6 lensed VCSEL array.



3b. 4x6 lensed photodetector array.

2.4 Optical interface

An assembly consisting of an injection molded plastic ferrule and discrete wavelength filters, performs the optical multiplex (MUX) and de-multiplex (DEMUX) function. This CWDM light turn (CLT) ferrule/filter assembly functions as an optical zig-zag or 'bounce cavity', as shown in figure 4, and can be thought of as a folded confocal cavity. When operating in the MUX mode, the ferrule combines light output from 4 VCSELs, and focuses the combined light signals into a 50um core multimode optical fiber. The ferrule/filter assembly performs the reciprocal function in the DEMUX mode, separating the four optical signals by wavelength and directing them to the appropriate PD. Because optical signals entering and exiting the MUX/DEMUX are collimated, X-Y alignment sensitivity between the ferrule and the optic arrays is reduced substantially. This benefit is somewhat offset by increased sensitivity to tilt misalignment. The ferrule can accommodate up to 16 fibers on a 250um pitch, and with NRZ signaling at 25Gbps delivers bandwidth of 100Gbps per fiber and aggregate bandwidth of 1.2Tbps.

An injection molded optical 'socket' provides precise alignment between the OE arrays and the ferrule. The socket is assembled to the OE substrate using pattern recognition to align features on the socket with respect to features on the OE arrays. Light cure adhesive is used to tack the socket in place on the electrical substrate. A secondary adhesive applied around the base of the socket provides additional bond strength and environmental sealing. In addition to aligning the ferrule and OE devices, the socket encases and protects the OE arrays and the optical connector assembly.

Photographs of the ferrule/filter assembly is shown in figures 10 and 11.

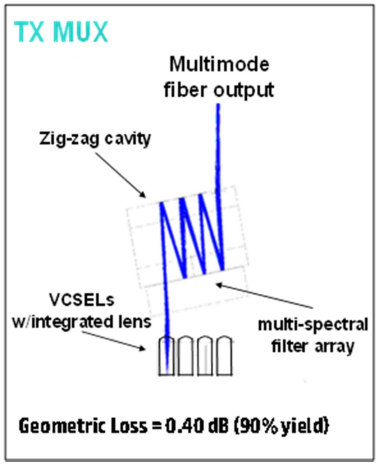
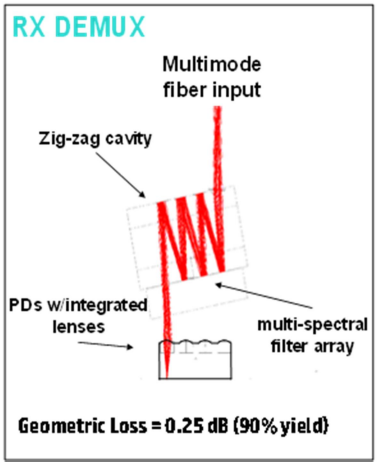


Figure 4a. Ray trace schematic for longest TX optical path.



4b. Ray trace schematic for longest RX optical path.

3. SYSTEM DESIGN

3.1 Thermal analysis

Thermal performance is a key aspect of all optical transceiver design. MBO designs can face additional challenges because of the close proximity between emitter chips and one or more power dissipating ICs. Typically, optical transceiver temperature limits are set by the requirements of the lasers in order to insure good device lifetime and efficient optical power output. Our thermal design goal was to insure a maximum substrate temperature under the VCSELs of 70C.

A wide range of computational fluid dynamics (CFD) analyses was conducted using Flotherm in order to identify the critical factors contributing to VCSEL substrate temperature. These factors included 1) heat sink geometry and material; 2) spacing between IC and OE arrays; 3) use of heat pipe; 4) rate and direction of airflow; 5) use of dedicated heat spreader under the OE arrays.

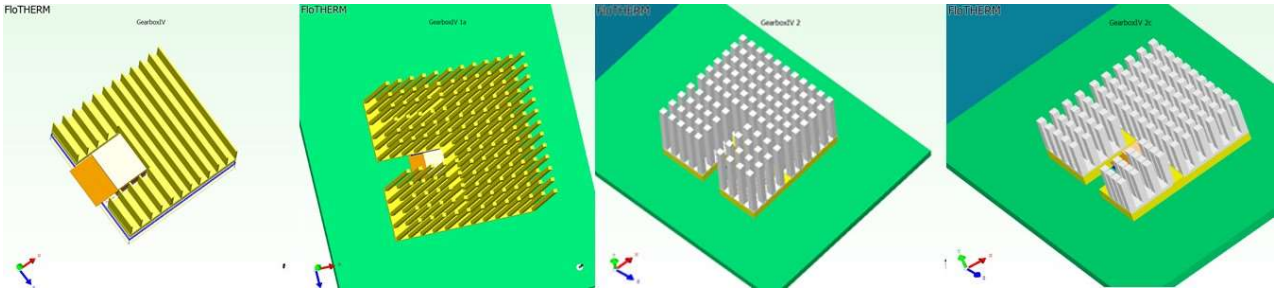


Figure 5a. Design1	5b. Design4	5c. Design6	5d. Design8
Heatsink area: 20x20mm	Heatsink area: 40x40mm	Heatsink area: 40x40mm	Heatsink area: 40x40mm
Base thickness: 1mm	Base thickness: 1mm	Base thickness: 3mm Cu*	Base thickness: 3mm Cu*
Fin height: 6mm	Fin height: 38mm	Fin height: 20mm	Fin height: 10mm
Material: Cu	Material: Cu base/Al fin	Material: Cu base/Al fin	Material: Cu base/Al fin

*Heat pipe incorporated in Cu base.

Thermal geometry is shown in figure 5 for four of the analyses. Figure 6 below shows cross sectional heat maps that correspond to the geometries in figure 5. The temperature of critical VCSELs, PDs and IC are shown in figure 7.

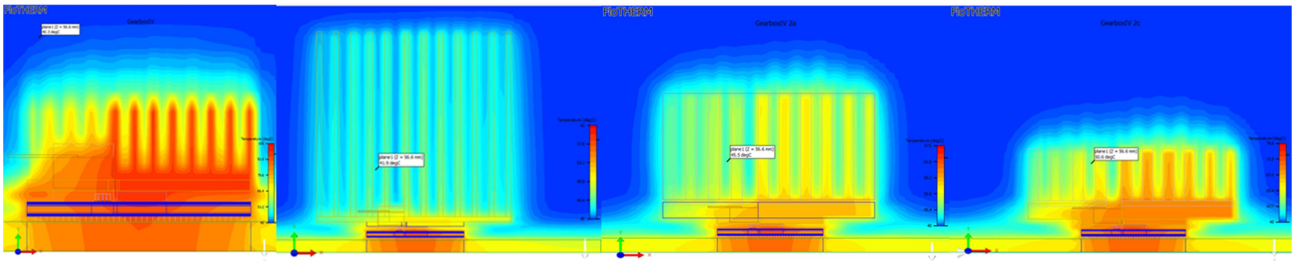


Figure 6a. Design1 6b. Design4 6c. Design6 6d. Design8

Multiple analyses showed that the first order factors limiting VCSEL substrate temperature are IC Heatsink geometry, including base area and fin height, air flow rate, and inlet air temperature. Other design factors that had less impact on the substrate temperature include:

- Increasing IC to VCSEL gap. This effect was approximately 1C/mm for gaps between 1 and 5mm.
- Minimizing thermal transfer through the heatsink from IC to OE arrays by use of thermal reliefs.
- Mounting OE arrays onto a Si interposer and utilizing an isolated heatsink to extract heat from the interposer.

Using the silicon interposer + OE arrays as a separate subassembly enables the optical actives to be assembled into the transceiver as ‘known-good’ components. Nonetheless, we chose not to use the interposer in order to reduce cost and avoid potential reliability issues associated with adding another layer of solder contacts.

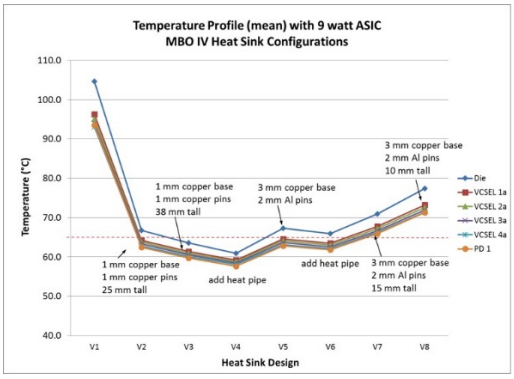
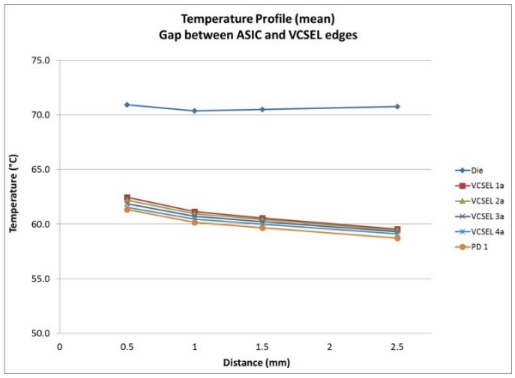


Figure 7a. Temperature vs IC to OE array gap.



7b. Active device temperature for 8 heatsink designs.

3.2 Optical design and analysis

Modeling and simulation was utilized in every phase of optical design and system validation. Zemax design software was selected as the primary simulation tool, used for obvious tasks such as optimizing lens profiles for the OE arrays and ferrule. At the system level, optical modeling was used extensively to evaluate the effect on coupling of a variety of geometric variables such as OE array tilt, fiber location within the ferrule, and assembly offsets. Ultimately, 40+ optical system variables were identified, along with a mathematical description of their expected geometric variation. Optical coupling sensitivity was evaluated for each variable while holding other variables constant. Typical results for two of these analyses, VCSEL array tilt along the long axis of the VCSEL array, and PD lens clear aperture, are shown in figure 8.

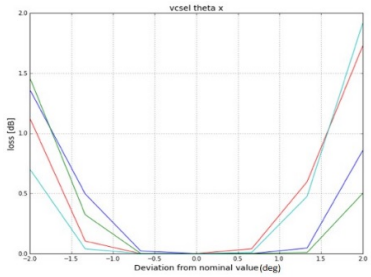
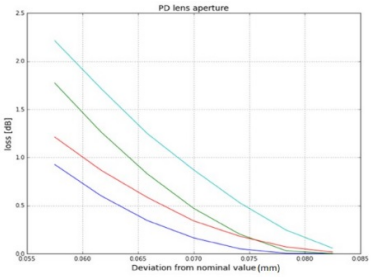


Figure 8a. Coupling loss vs VCSEL array tilt about X axis.



8b. Coupling loss vs photodiode lens clear aperture.

Monte Carlo analysis was also used to evaluate expected coupling loss due to a combination of geometric sources of variation. These non-idealities included piece part imperfections and assembly position errors. Simulation helped determine part and assembly tolerances, and proved useful as a debug tool to evaluate the performance impact of parts that were out of specification. Once the link budget was finalized, the Monte Carlo analyses provided insight into yield expectations for different system designs. Figure 9 shows the cumulative loss distribution function for the link operating in both MUX and DEMUX modes.

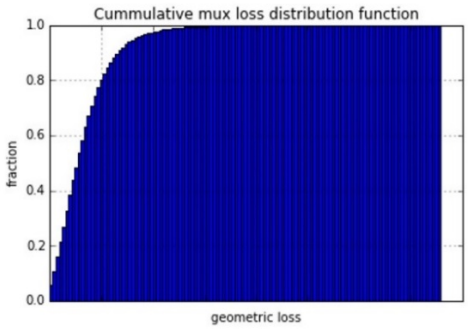
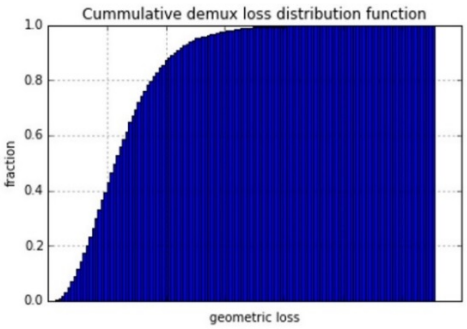


Figure 9a. Monte Carlo simulation for MUX loss.



9b. Monte Carlo simulation for DEMUX loss.

3.3 Optical connector system

The CLT ferrule/filter assembly provides the optical interface to the VCSELs and PDs. Within the CLT, multiple fibers are precisely located with respect to an array of lenses that act to turn and focus both incoming and outgoing light beams. The top side of the CLT ferrule also incorporates 3 rows of 'relay' lenses that work together with spectral filters attached to the back side of the ferrule to form a zig-zag optical cavity. A high reflectivity dielectric coating is deposited onto the relay and turning lenses to minimize optical losses. Figure 10 shows the CLT ferrule in more detail. All critical optical and mechanical ferrule features are formed during the injection molding process. Passive assembly processes are used to attach the optical fibers and the spectral filter assembly.

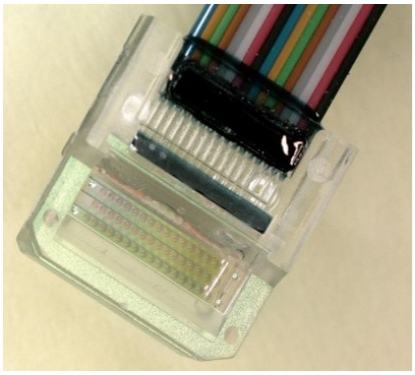
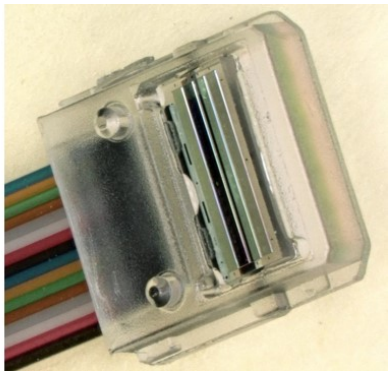


Figure 10a. CLT ferrule lens side.



10b. CLT ferrule filter side.

The decision to build a solderable transceiver led naturally to the development of a connectorized optical cable. Pigtailing is suitable in certain applications, but pigtailed optics (especially those containing optical thermoplastics) typically are not compatible with solder reflow processing. Connectorization adds design complexity, and cost, but also provides benefits in the form of greater test flexibility, simplified troubleshooting, simplified cable management, and rework flexibility.

Two areas of focus during connector development were; 1) opto-mechanical alignment and; 2) control of environmental contamination. A double guide pin geometry was chosen to control opto-mechanical alignment between the CLT lensed ferrule and the mating optical socket. If space is available, the ferrule alignment pins should be positioned on the axis connecting the centers of the integrated ferrule lenses, or on an axis perpendicular to the lens center axis. This approach is used on the MT ferrule® and the PRIZM® LightTurn® ferrule, respectively. In order to limit ferrule size, we chose to position the alignment pin axis parallel to and offset from the ferrule lens axis. This geometry can increase X-Y alignment error by 2-3um in the least material condition for ferrule pin and socket guide pin hole. This potential alignment error was viewed as an acceptable tradeoff in order to minimize ferrule size and maximize optical channel density.

MBO transceivers typically operate inside a server or switch enclosure utilizing fast moving air to cool electronic components. This air typically carries with it dust and other particulate contamination. Figure 11 shows an extreme case of dust accumulation onto a pair of heatsinks inside an enclosure.

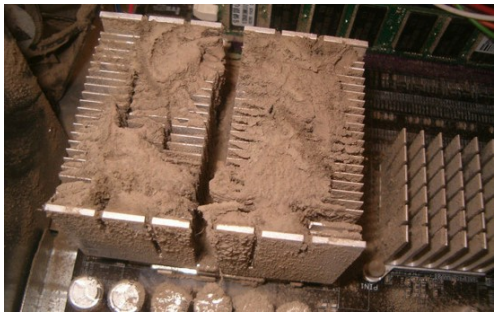


Figure 11. Heatsink inside electronic enclosure exhibiting extreme dust accumulation.

The CLT assembly incorporates perimeter seals around the base and top side of the optical socket to prevent contamination from entering the socket. The base seal is provided by the adhesive that bonds the socket to the electronic substrate. A flexible liquid silicone rubber (LSR) component seals the gap between the top of the optical socket and the optical

connector. The seal, optical connector, and substrate assembly are shown in figure 12. In order for the seal to function as required, it is necessary to use the proper combination of seal material and part geometry. The base material needs to be low outgassing and resistant to compression set (permanent deformation under load and temperature).

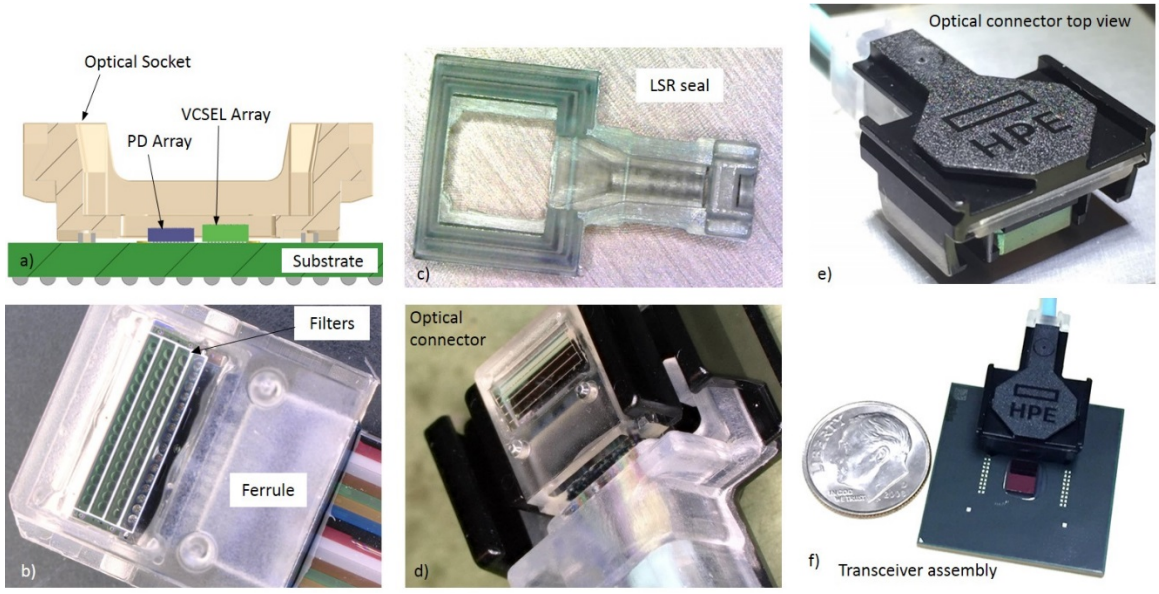
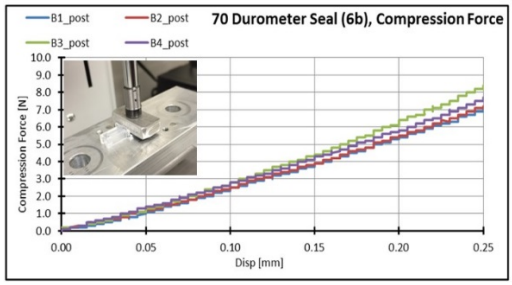


Figure 12. CWDM Tx/Rx module components a) electro-optical actives on build-up substrate with optical socket; b) CWDM MUX/DEMUX assembly; c) seal; d) optical connector, filter-side view; e) optical connector cover-side view; f) transceiver assembly.

The seal is compressed as the connector is mated to the socket, and effort was made to minimize the effective seal spring force. Figure 13a shows the force deflection curve for the portion of the seal that undergoes compression during connector latching. Figure 13b shows force/deflection data for the entire CLT connector assembly.



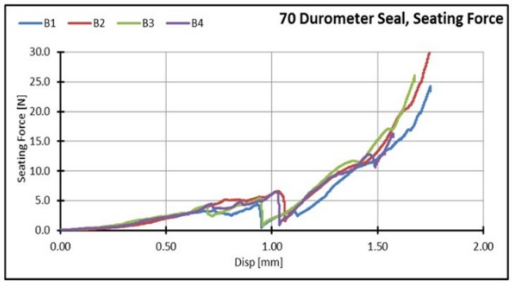


Figure 13a. 70 durometer seal force-displacement curve

13b. Optical connector latching force-displacement curve.

4. SYSTEM PERFORMANCE

4.1 Insertion Loss

A fiber optic connector system needs to satisfy multiple performance criteria; foremost is minimizing optical insertion loss. Insertion loss in a single fiber connector, like the LC or SC, depends mainly on the dimensional fidelity of the simple cylindrical ferrule[5]. The lensed CWDM ferrule assembly is affected by additional factors such as the quality of the high reflectivity (HR) coating on the relay lenses, lens geometry, overall ferrule dimensional accuracy, and fiber and filter assembly accuracy.

The CLT ferrule is designed to accept up to 16 optical fibers. Early device assembly has focused on 4 and 12 fiber versions. However, the optical ferrule can accept 16 fibers, enabling a 33% increase in bandwidth. Figures 14a and 14b show insertion loss histograms for MUX and DEMUX testing of four fiber CLT parts. This data includes results for all four transmission wavelengths. 1065nm signals travel the shortest path through the bounce cavity and experience only a single reflection. Not surprisingly, these signals account for the lowest loss portion of the loss distribution. The 990nm signals experience the longest optical path and most bounces through the zig-zag; as such they account for the high loss tail of the distribution.

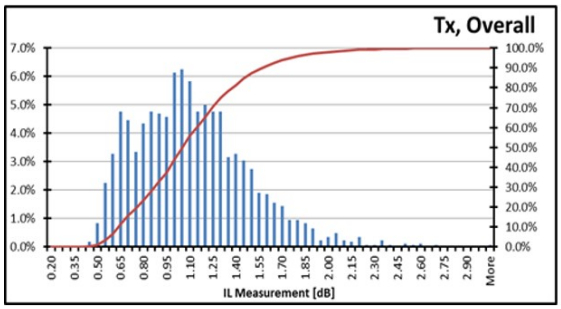


Figure 14a. Histogram of MUX insertion loss.

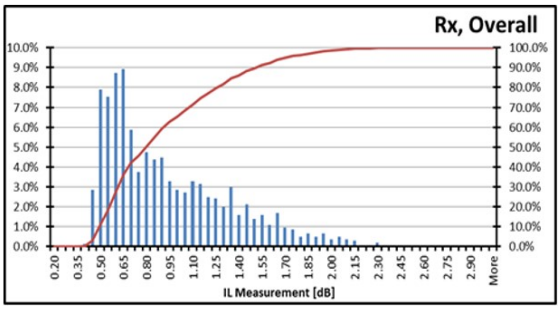


Figure 14b. Histogram of DEMUX insertion loss.

Table 1. Insertion loss for 24 VCSEL, 6 fiber CLT MUX assembly.

990nm		1015nm		1040nm		1065nm	
VCSEL ID	Cable Loss(dB)	VCSEL ID	Cable Loss(dB)	VCSEL ID	Cable Loss(dB)	VCSEL ID	Cable Loss(dB)
0	-1.08	6	-1.04	12	-0.75	18	-0.78
1	-0.95	7	-0.91	13	-0.83	19	-0.59
2	-0.924	8	-0.93	14	-1.08	20	-0.55
3	-1.012	9	-0.92	15	-0.96	21	-0.62
4	-1.086	10	-1.03	16	-0.99	22	-0.95
5	-1.140	11	-1.00	17	-1.03	23	-0.46
Average	**-1.03**	**Average**	**-0.97**	**Average**	**-0.94**	**Average**	**-0.66**
StdDev	**0.08**	**StdDev**	**0.05**	**StdDev**	**0.11**	**StdDev**	**0.16**

In addition to measuring insertion loss, it is useful to characterize the alignment sensitivity between components and also to visualize the optical spot position and shape at various locations relative to ferrule datums. Figure 15a shows an image of optical power from 24 VCSEL sources at the CLT relay lens surface. Figure 15b shows coupled power ($1/e^2$) as a function of X axis translation for each combination of laser and associated relay lens. Ideally each curve will exhibit a flat top section, indicating constant coupled power over the region within which a properly shaped optical beam falls entirely within the relay lens clear aperture. Deviation from the flat top profile indicates room for improvement in beam shape, ferrule performance, alignment, or some combination of these factors.

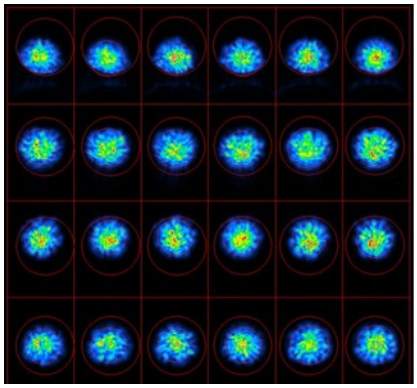
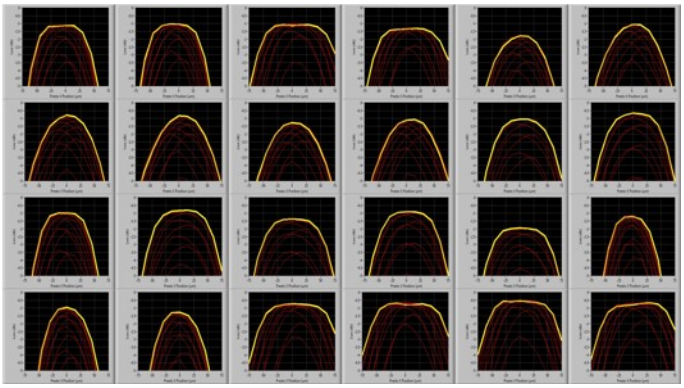


Figure 15a. Optical spots at relay lens surface.



15b. Coupled MUX power as a function of ferrule X position.

4.2 Link Performance

The CWDM optical transceiver described here has demonstrated the ability to support link lengths of up to 75m using standard OM3 fiber. With all 4 wavelength VCSELs operating, we show in figure 16a the optical spectrum from the

output of a single fiber overlaid with the measured transmission spectra from the four corresponding bandpass filters. There is good spectral alignment between the VCSELs and bandpass filters, resulting in optical crosstalk of less than 35 dB between spatially adjacent channels. Figure 16b shows 25Gbps optical eye diagrams from each CWDM channel at the output of the DEMUX ferrule.

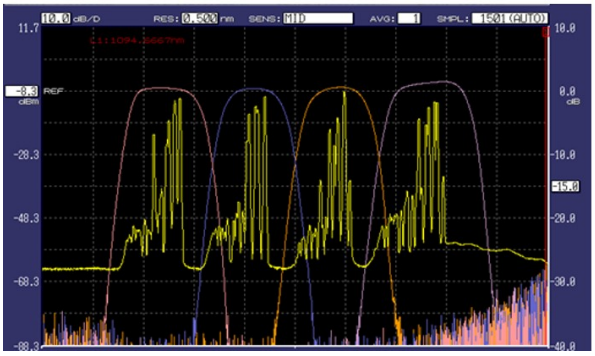
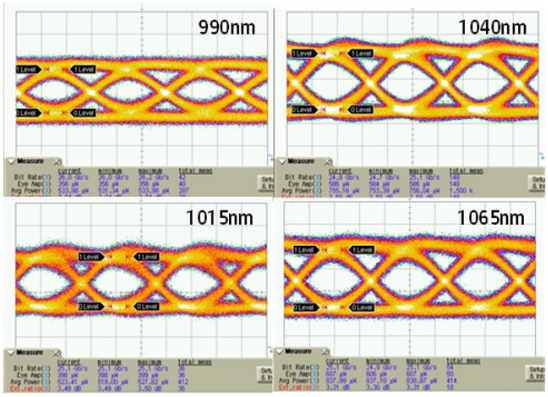


Figure 16a. Overlay of VCSEL and filter spectra.



16b. 25Gbps CWDM optical eyes at DEMUX output.

4.3 Side-load testing

Fiber optic connector systems are required to exhibit low additional insertion loss when a specified force is applied to the optical cable. This reflects real world cable usage, wherein cables may undergo loading due to being bundled together, tie-wrapped, and routed through the data center. In order to assess connector performance under load, side-load testing is performed[6]. The CLT connector design isolates the ferrule holder part from the outer connector housing, and applies a compressive spring-load to seat the ferrule against the socket mating surface. The interface seal also provides strain relief for the cable, and prevents excessive cable loads from being transferred directly to the ferrule. Side-load results for a typical CLT connector subject to a 2.2N cable side load are summarized in figure 17.

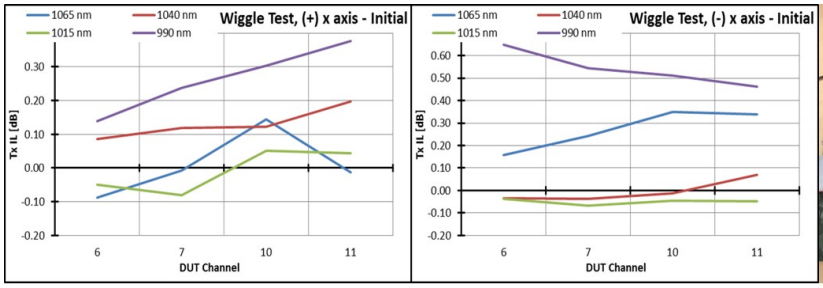
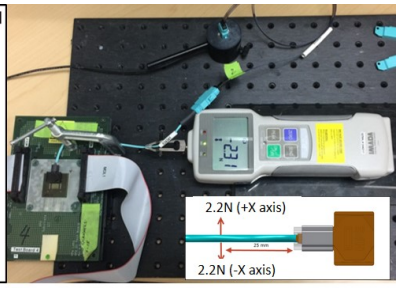


Figure 17a. +/-X axis side load wiggle test data at 2.2N load.



17b. Side-load test hardware setup.

4.4 Insertion Durability Testing

MBO optical connectors are typically mounted inside a server or switch system enclosure where they are not readily accessible. Therefore the typical CLT connector is expected to be connected and removed fewer than 5 times. However, some connectors will be used for testing and other applications that require higher numbers of connect/disconnect cycles. A sample of CLT connectors was subjected to 60 cycles of connect/disconnect- see figure 18. Insertion loss was tested after every 10cycles, and photos taken of the guide pins and lenses to evaluate contamination generation.

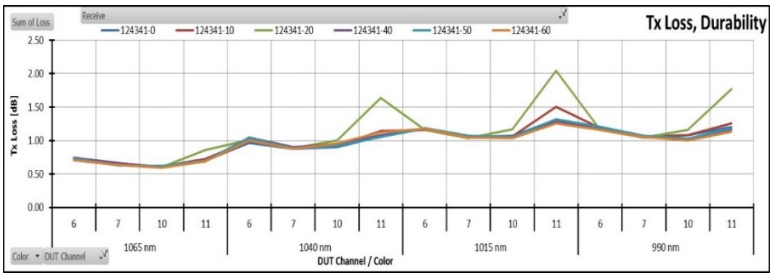


Figure 18. CLT connector durability test data over 60 connect/disconnect cycles; testing after each 10 mate set.

4.5 Dust Testing/Environmental Contamination Protection

Designing the optical connector assembly to protect active and passive components from environmental contamination proved to be a challenge. There is an irregular geometry at the interface between the optical connector and the mating socket, due mainly to the presence of the optical cable exit. In order to test the integrity of the sealed interfaces at the top and bottom of the socket, we subjected a sample of transceivers to an Arizona road dust exposure test[7]. An engineering team at HP Enterprise has recently developed a dedicated test system for subjecting optical components to a controlled dust stream. Key elements of the system, shown in figure 19a, are the solid aerosol generator (SAG), mixer-buffer chamber, and a dust concentrator nozzle. The HPE developed system allows close control over the dust distribution in the air stream, and also reduces dust clumping which can limit test repeatability. Figure 19b shows a typical OE array lens surface following the dust exposure test. Visual inspection of the samples before and after testing demonstrated that essentially no road dust was able to penetrate into the interior of the socket.

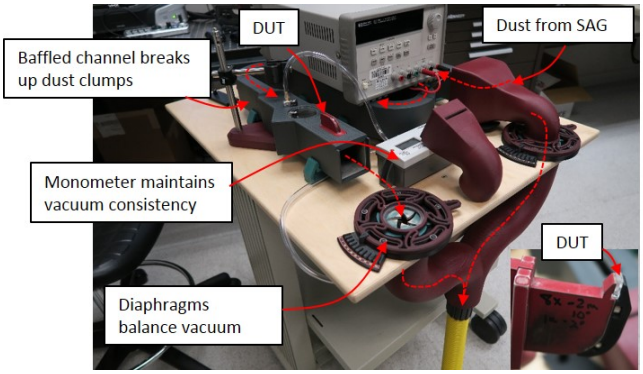


Figure 19a. Road dust test system.

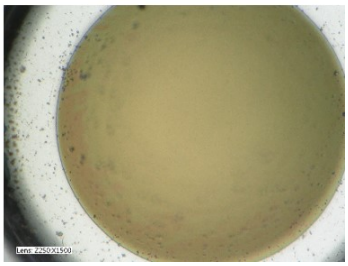


Figure 14b. OE array lens after dust exposure.

5. SUMMARY

We present a compact, low cost, solder-reflowable, optically connectorized CWDM optical engine with 1.2Tbps (100Gbps/fiber) bandwidth capacity. Standard flip chip processes are employed to precisely assemble the active electronic and optical components. The PD and laser arrays are fabricated with integrated collimating lenses. This significantly reduces optical alignment sensitivity and enables the use of vision alignment, rather than active alignment, for positioning the optical socket used to locate our custom optical connector.

The transceiver is expected to achieve a cost target < $1/Gbps, and to be capable of scaling to high volume manufacturing. Because the CWDM architecture employed achieves data capacity of 100Gbps per fiber, fewer connectors and fiber cables are needed to transfer data within the data center. This provides further cost savings, especially as cable run lengths exceed distances of a few meters.

As currently configured, this multimode CWDM platform offers significant bandwidth per unit area of installation. Through the implementation of forward error correction (FEC) and PAM-4, module bandwidth can be increased by a factor of 2 or more times without replacing the fiber infrastructure. Using single mode VCSELs at 1um wavelength, together with SMF28 fiber, offers the potential to extend link lengths to > 2km.

ACKNOWLEDGMENT

The authors would like to acknowledge significant technical contributions from Dan Berkram and the Ft Collins ASIC team, and Jason Culler. We also would like to thank Cullen Bash, Jon Smela, Michael McBride and John Sontag for their support.

REFERENCES

[1] http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm
[2] https://www.statista.com/topics/3150/data-storage/
[3] http://genzconsortium.org
[4] Tan, Michael, et al, "Universal photonic interconnect for data centers" OFC 2017, DSN6, Optical devices, subsystems, and networks for Datacenter and Computercom
[5] Kyocera corporation; 'Standard ferrule product list, document R0185A
[6] IEC International Standard; 62150-3, 'Fibre optic active devices and components- Test and measurement procedures'
[7] Moore, David A., "Dust Exposure Trial," HPE System Reliability Report, February 2, 2016