Efficient Sources for Chip-to-Chip → Box-to-Box Communication within Data Centers

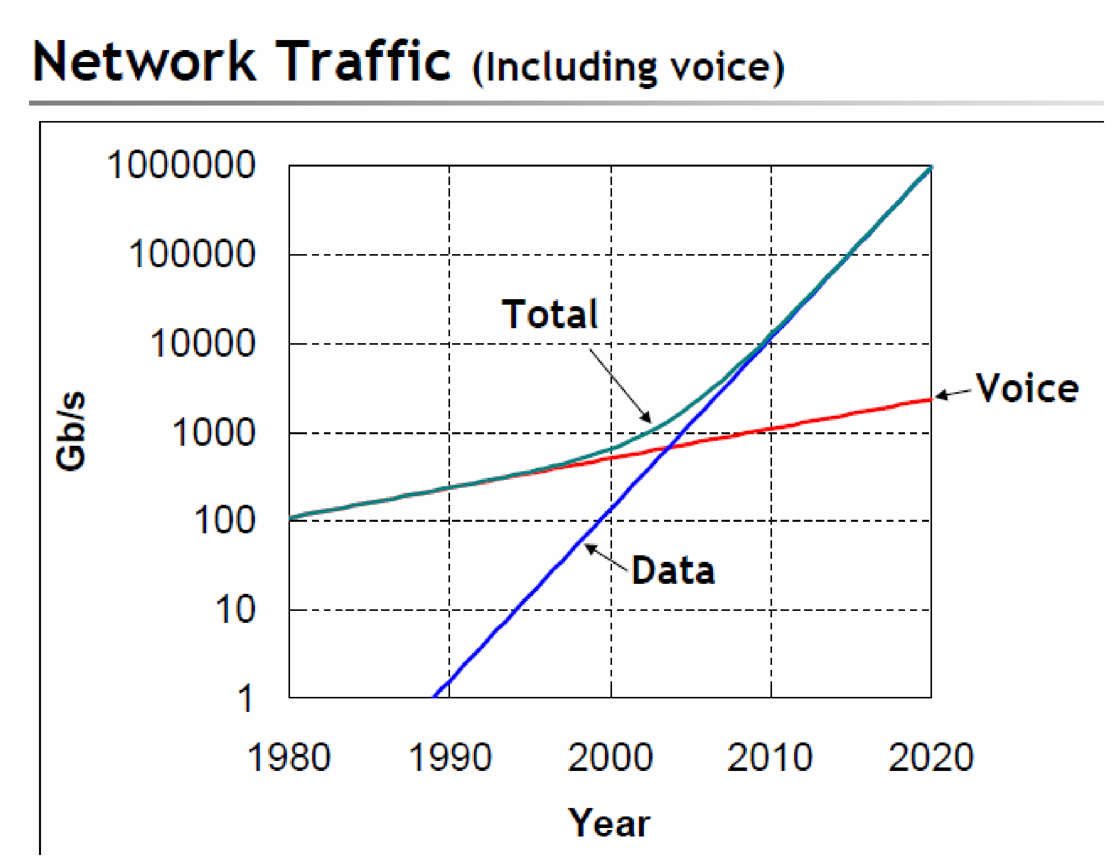
IEEE Photonics Society Summer Topicals
Optical Networks and Devices for Data Centers

PaperTuD2.1
(Invited)
20 July 2010

Larry A. Coldren
ECE and Materials Departments
University of California, Santa Barbara, CA 93106
coldren@ece.ucsb.edu

Data is King

- **Today traffic on the core network is nearly all data**



Drivers:
- Goggle
- Microsoft
- Yahoo
- Facebook
- Ebay
- You Tube
- Programmed Stock trading
- Amazon
- AOL
- Super computer com

New super-computer intraconnection also major driver for data-links

A Typical Data Center

- > 30 MW power requirements
- Requires many Gb/s of bandwidth—justifies 100Gb-Ethernet

Terabit Links Required

UCSB

- In addition to high data rates via serialization, some degree of parallelism is necessary
  - WDM
  - √ Space Division (Fiber arrays; multicore)
  - Higher Spectral Efficiency (as in long-haul)
    - External modulation/advanced formats
    - Photonic ICs for SWaP and $
    - Coherent receivers
    - Questions of power and cost for Datacom

➢ What about photonic switching?

Both Communication and Switching Power Dissipation a Concern

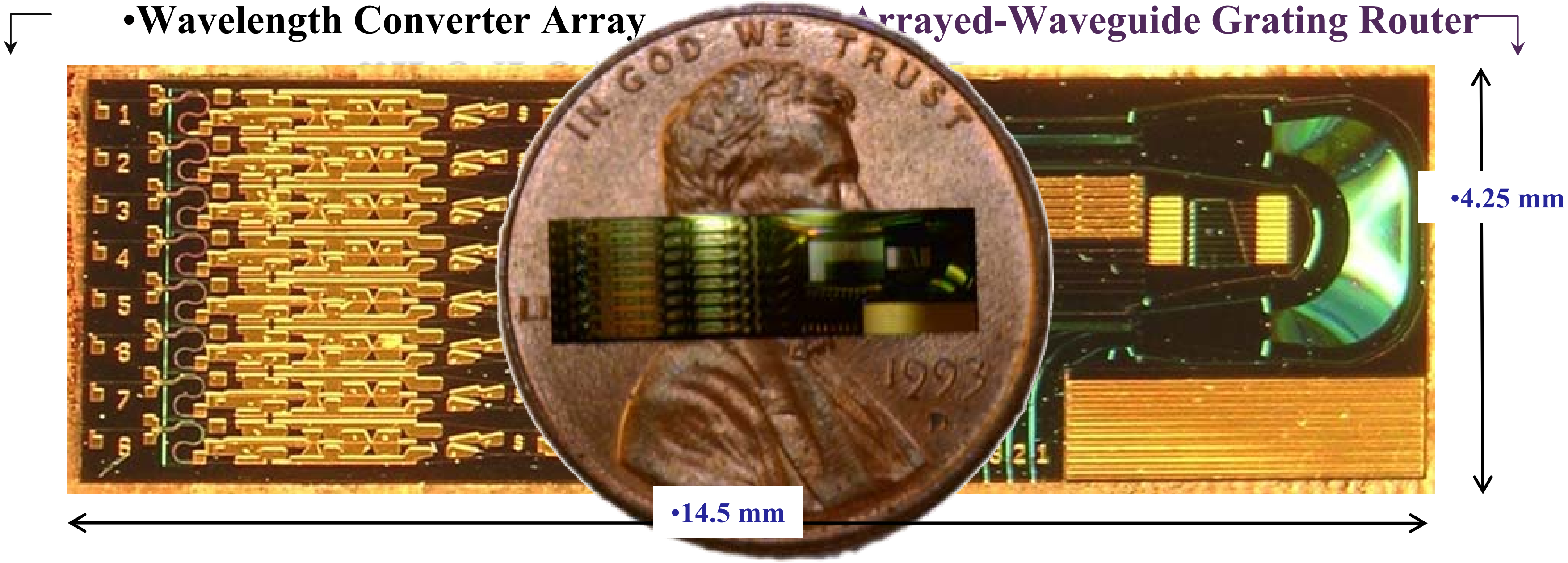
- **<u>Problem</u>:** Bandwidth demands scaling faster than both silicon and cooling technologies: Communication power = 40%; Processing/sw power = 60%

•Maximum configuration for CRS-1: 92 Tbps
➔ 72 line card shelves + 8 fabric shelves

Cisco CRS-1 Router

Photonic Switching: the MOTOR Chip— an 8 x 8 Space Switch

UCSB

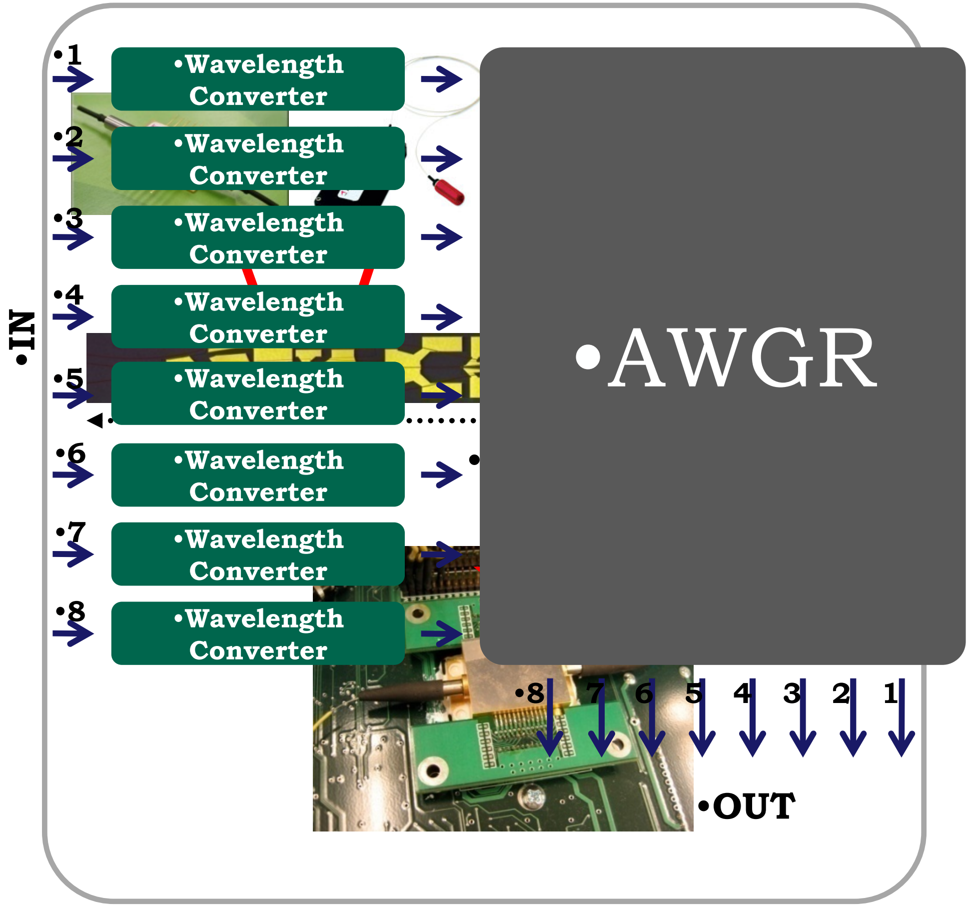


- **A monolithic tunable optical router (MOTOR) chip to function as the switch fabric of an all-optical router**
  - **Line rate: 40 Gbps / channel**
  - **Total capacity: 640 Gbps**
  - **Error-free operation**
- **Photonic integration technologies designed for high-yield, large-scale applications**
- **World's largest and most complex Photonic IC (2009)**

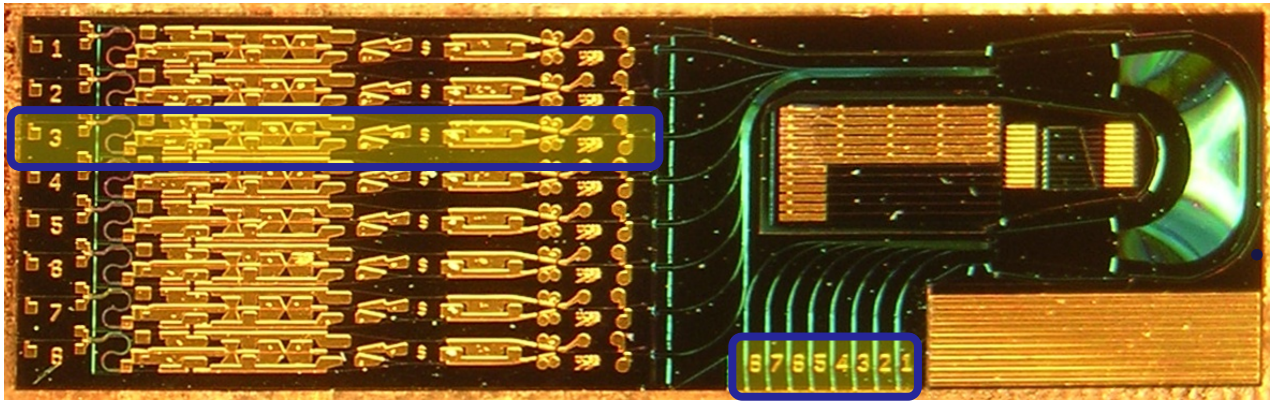
•Steven C. Nicholes, M. L. Mašanović, E. Lively, L. A. Coldren, and D. J. Blumenthal, *IPNRA '09*, Paper IMB1 (July, 2009); also *JLT*, (Jan. 2010).

Large-Scale Photonic Integration

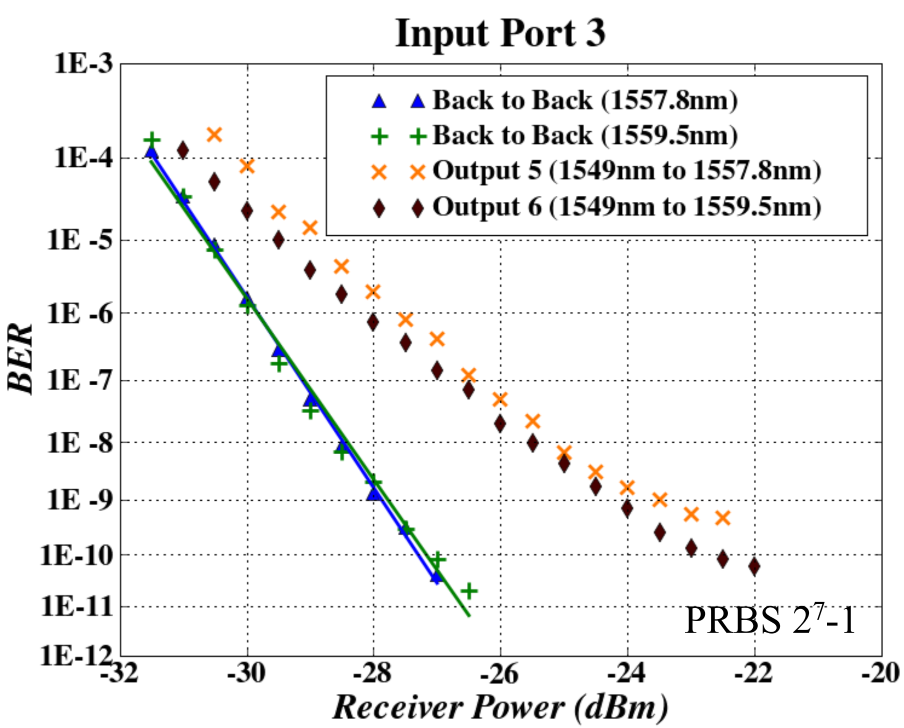
- Many components replaced by a single chip that integrates the functionality of all together
- Saves the cost and complexity of manufacturing as well as interconnecting many parts
- First we illustrate an integrated wavelength converter
- Second, we illustrate eight wavelength converters integrated with an 'Arrayed-Wavelength-Grating-Router' (AWGR) –which acts like a prism



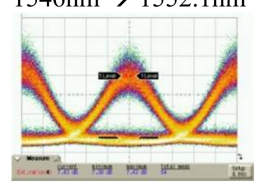
MOTOR Results : Constant Input Port— 40 Gbps RZ



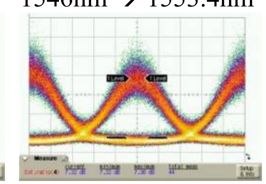
- Power penalty at BER = 1E-9 for PRBS 2^7-1 data at 40 Gbps < 3.5 dB(no AR coating)
- Power Diss: < 2 W/channel



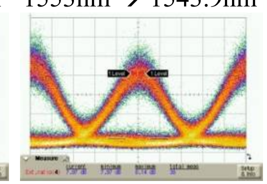
Output #1
1546nm → 1552.1nm



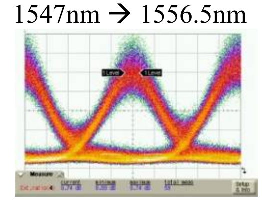
Output #2
1546nm → 1553.4nm



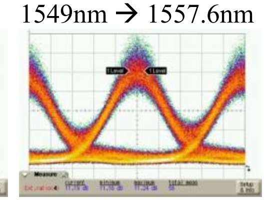
Output #3
1553nm → 1543.9nm



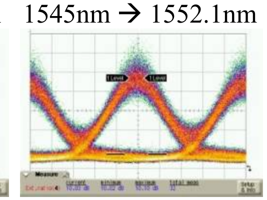
Output #4
1547nm → 1556.5nm



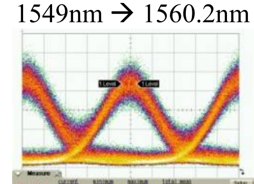
Output #5
1549nm → 1557.6nm



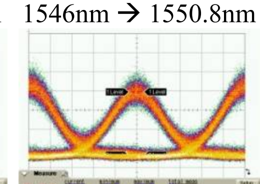
Output #6
1545nm → 1552.1nm



Output #7
1549nm → 1560.2nm



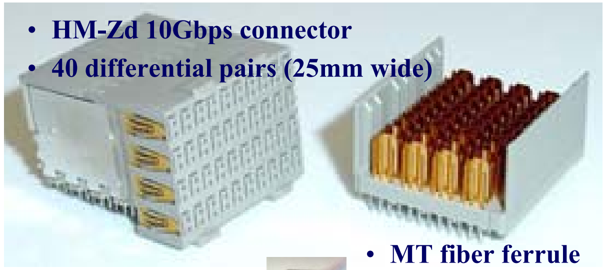
Output #8
1546nm → 1550.8nm



Optical Interconnects: High Data Rate Density

- **High Speed Cables**

- **High Speed Connectors**



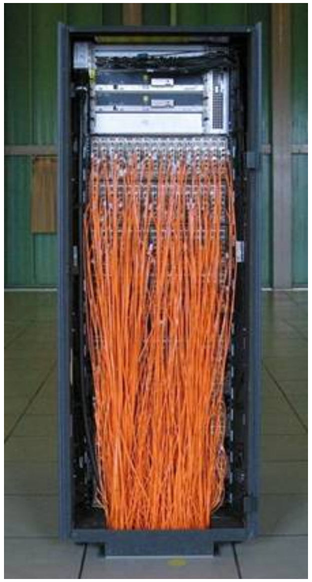
- **MT fiber ferrule**
- **48 fibers (7mm wide)**

•IBM Federation Switch Rack

•Electrical

•Mixed

•Optical



➢ **Optics enables high-density integration and better cooling efficiency**

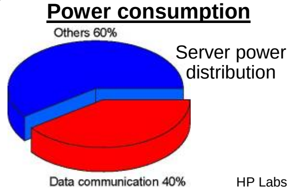
➢ **More that just Power → SWaP**

Courtesy S. Nakagawa

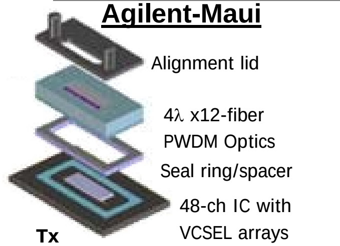
Optical Datacom Directions

Box - Box → Board - Board → Chip - Chip → On-chip Interconnects

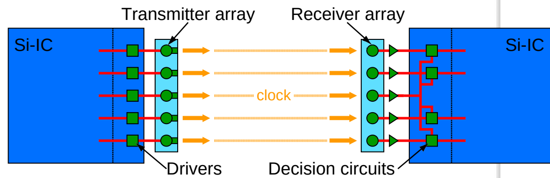
Power consumption



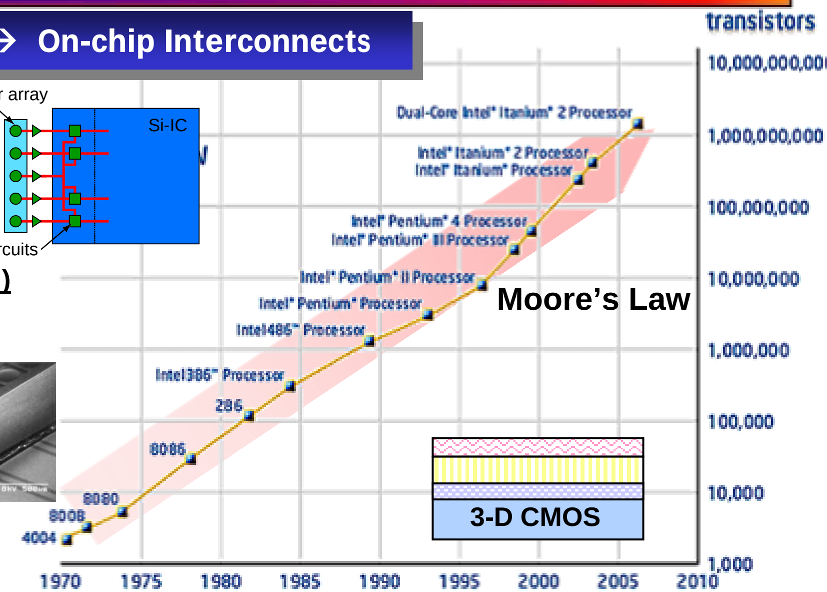
Agilent-Maui

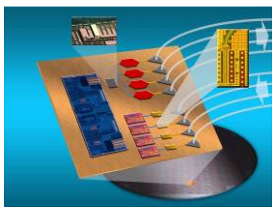


Fiber-array Transceivers



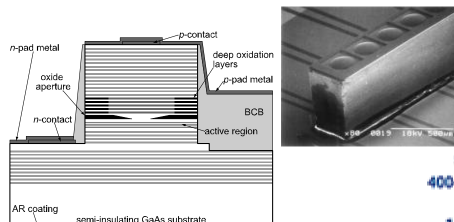
Chip-to-chip (C2OI)





Luxtera-Si-photonics

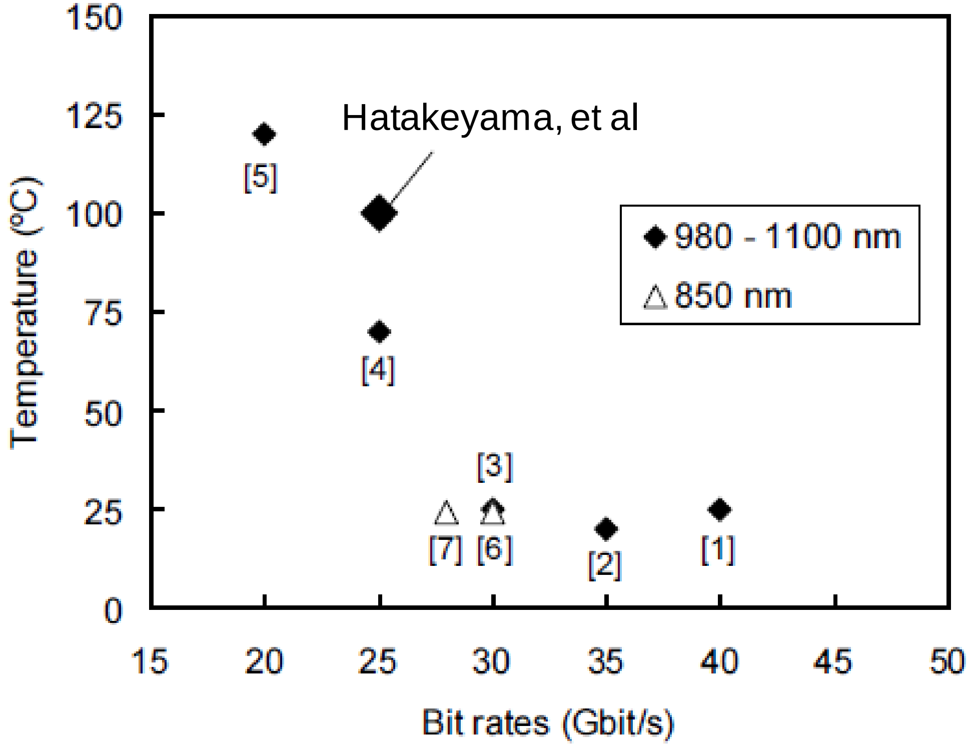
Active Cables



35 Gps VCSEL

Component	Today (mW/Gps)	Future
<u>VCSEL</u> @ 35 Gps (UCSB) VCSEL-Det Link @ 35 Gps (UCSB + UVA)	0.3 mW/Gps (efficient VCSEL) (~0.6 mW/Gps --"receiverless")	0.2 mW/Gps (2010) @ 50 Gps 0.4 mW/Gps (2010) @ 50 Gps
<u>PARALLEL FIBER LINK</u>: Maui: 48 x 10.4 Gps = 0.5 Tb/s fiber link	6.6 mW/Gps (VCSEL - detector arrys with driver and reciever electronics)	~1 mW/Gps Active Cables/VCSELs (2015) ~10 mW/Gps /Si-photonics with sources
<u>PARALLEL BOARD LINK</u> (chip-to-chip): Terabus: 24+24 x 12.5 Gps = 0.3 Tb/s Full Duplex	10 mW/Gps (VCSEL –detector arrays with polymer waveguides and electronics)	~1 mW/Gps /VCSEL arrays or Si-photonics not counting off-chip sources. ~10 mW/Gps/Si-photonics inc. sources
<u>ON-CHIP NETWORK</u> (future dream): 3-D CMOS: Logic, memory & photonic planes Parallel + WDM + Serialization in Photonics?		~300 cores (2018) >70 Tb/s on-chip (<1pJ/bit → 1 mw/Gps) Need serialization (latency)

Progress in High-Speed VCSELs



Return to 980+ nm strained-InGaAs Actives

[1] NEC ('08)
[2] UCSB ('07)
[3] NEC ('07)
[4] Agilent ('07)
[5] TU Berlin ('08)
[6] Finisar ('08)
[7] TU Chalmers ('08)

H. Hatakeyama, T. Anan, et al, "Highly reliable high speed 1.1μm-InGaAs/GaAsP-VCSELs," *Proc. SPIE-VCSEL XIII*, **7229**, 02-1 (2009)

Major VCSEL Advance circa 1989: High-Q cavity/QWs at antinode near 980 nm

Y. Lee, J. Jewell, et al, *Electron. Letts*., **25** (20) 1989

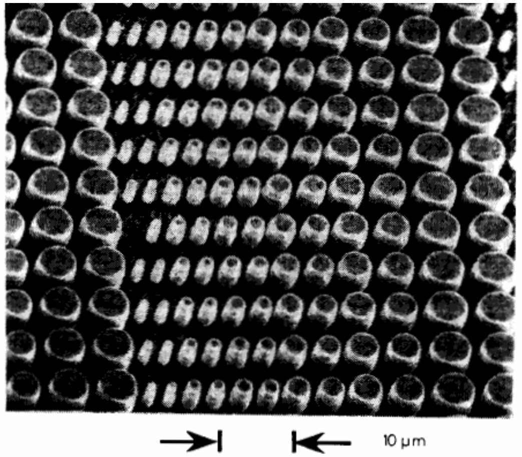


Fig. 1 *Very small portion of vertical cavity microlasers of diameter 1–5 µm*

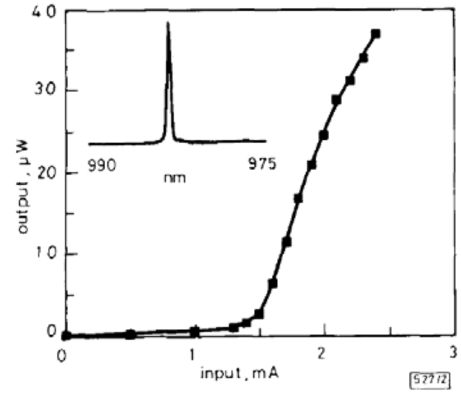


Fig. 2 *Output against current for 5 µm-square microlaser operating CW at room temperature without heat-sinking*

R. Geels, S. Corzine, et al, *PTL*., **2** (4) 1990

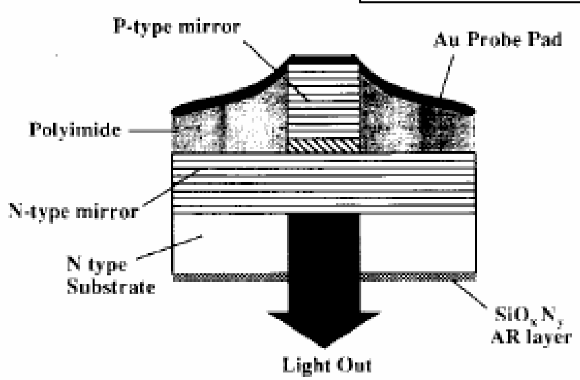


Fig. 1. Schematic cross-sectional view of a finished SEL device fabricated using the self-aligned planarized process. Cross hatching is used to indicate the active region. The substrate thickness is not drawn to scale.

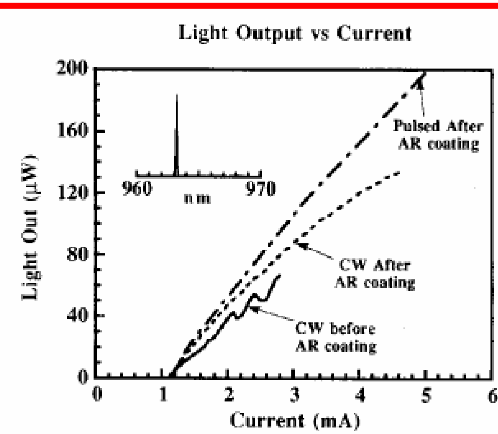


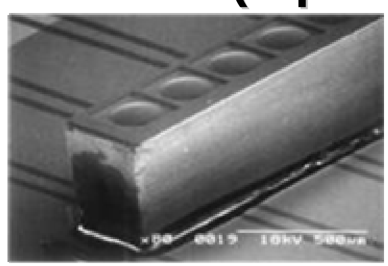
Fig. 2. Light output versus current under room temperature continuous wave and pulsed operation before and after AR coating for a 12 µm square device. Applied voltage at threshold was 4.0 V. Pulsed testing was done with 500 ns pulses resulting in a maximum output power of 400 µW and an external differential efficiency of 4%. Inset shows the output spectrum which showed a full width at half maximum of less than 1.5 Å limited by the resolution of the spectrometer.

The case for 980-1100nm (strain/GaAs*)

- **Higher-intrinsic modulation bandwidth @ lower current density**

$$f_R = \left[\frac{v_g a}{q V_p} \eta_i (I - I_{th})\right]^{1/2}$$

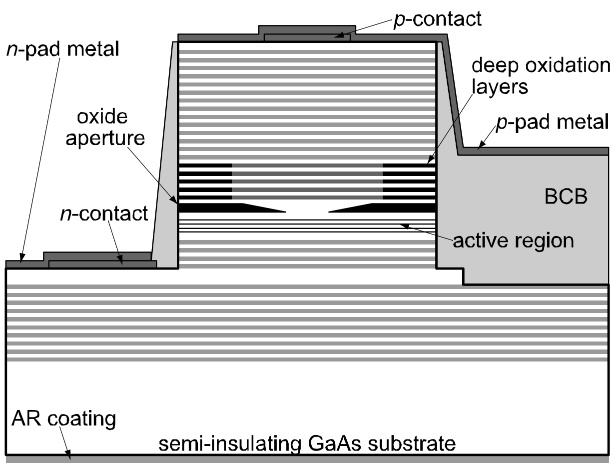
- **Lower threshold & better efficiency**
- **Improved reliability**
- **Lower fiber loss**
- **Lower fiber material dispersion**
- **Transparent substrate (flip-chip/backside optics/simple contacts, etc.)**

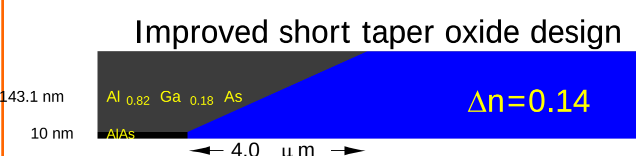


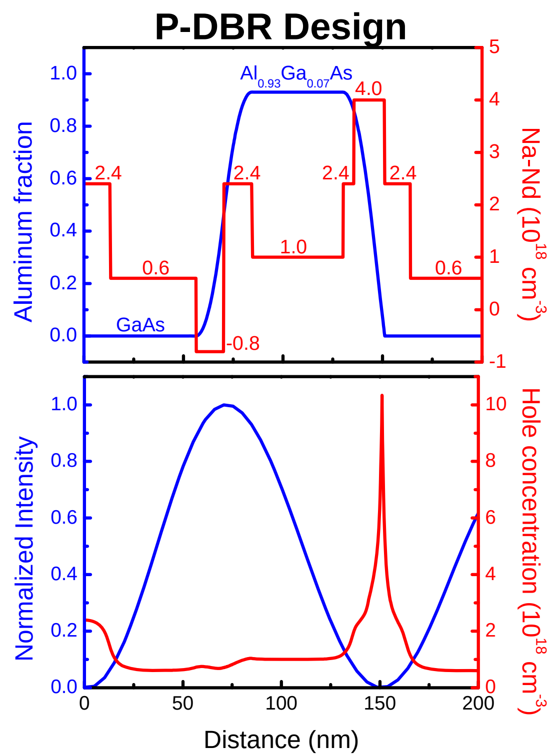
- **Standards not an issue in Active Cables**

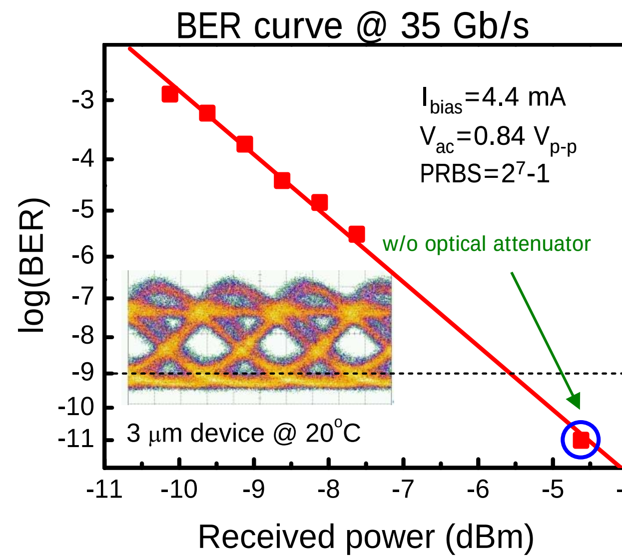
*vs 850 nm / GaAs

High-Speed, Efficient VCSELs (2007)





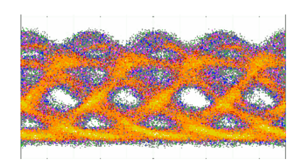




IMPROVEMENTS

- **Novel tapered oxide aperture**
  - **- Small mode**
  - **- Low loss**
- **Parasitics reduced**
  - **- Deep oxidation layers**
  - **- DBR BG-engr /low loss**

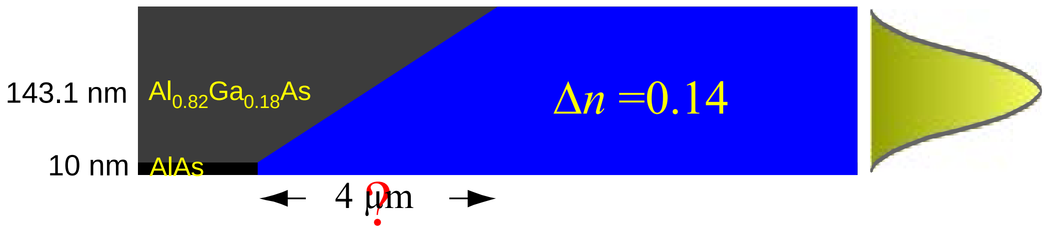
- **Limited by multimoding and detector bandwidth**

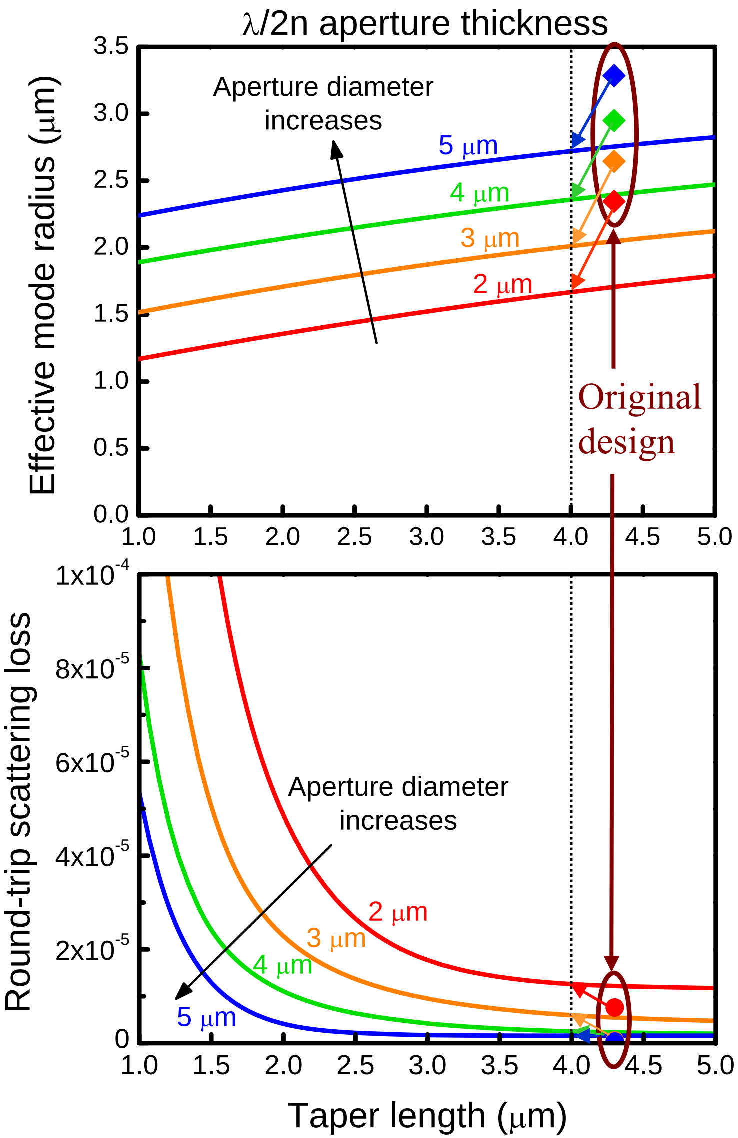


40Gb/s shows errors

- Highest bandwidth/power dissipation: (35 Gb/s)/10 mW total (2007-2009)
- Highest bandwidth 980 nm VCSELs

Blunter Tapered Oxide Reduces Mode Volume





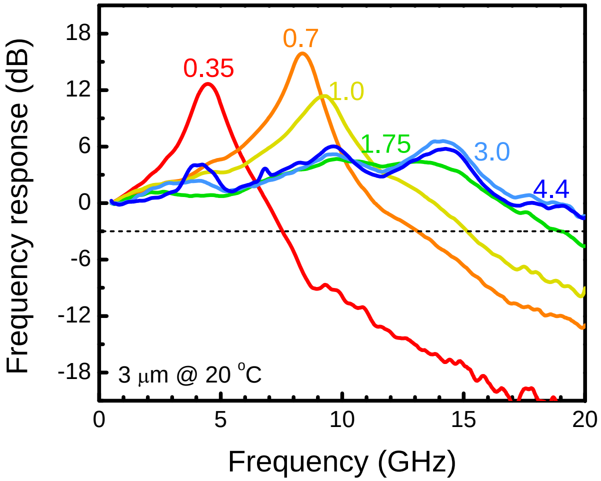
Trade-off between optical scattering loss and mode confinement

Advantages over the original design:

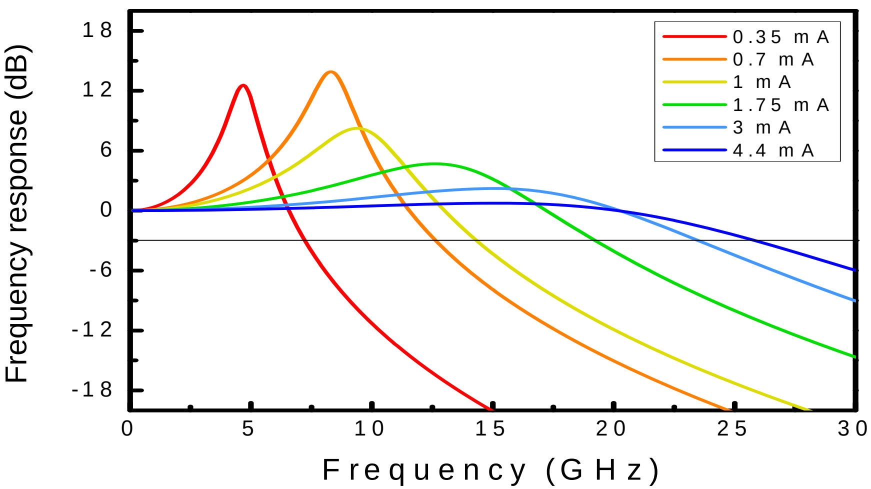
- **Better mode confinement**
  - **For 3 μm devices, mode volume reduces 1.73 times, corresponding to a 31% increase in *D*-factor**
- **Does not introduce significant optical loss**
- **Lower parasitic capacitance due to thicker oxide**

* E. R. Hegblom *et al.*, *IEEE J. Sel. Top Quantum Electron.*, vol. 3, pp. 379–389, 1997

Extrapolations for Single-mode VCSELs



- Experimental small-signal response
  - 15 GHz achieved at 1 mA (P_{diss} = *1.3 mW*)
  - BW > 20 GHz @ *I* > 2 mA → highest for 980 nm VCSELs
  - Limited by multimoding (*I* >1.5 mA)
  - Data rate limited to 35 Gb/s @ 10 mW (0.29 pJ/bit)



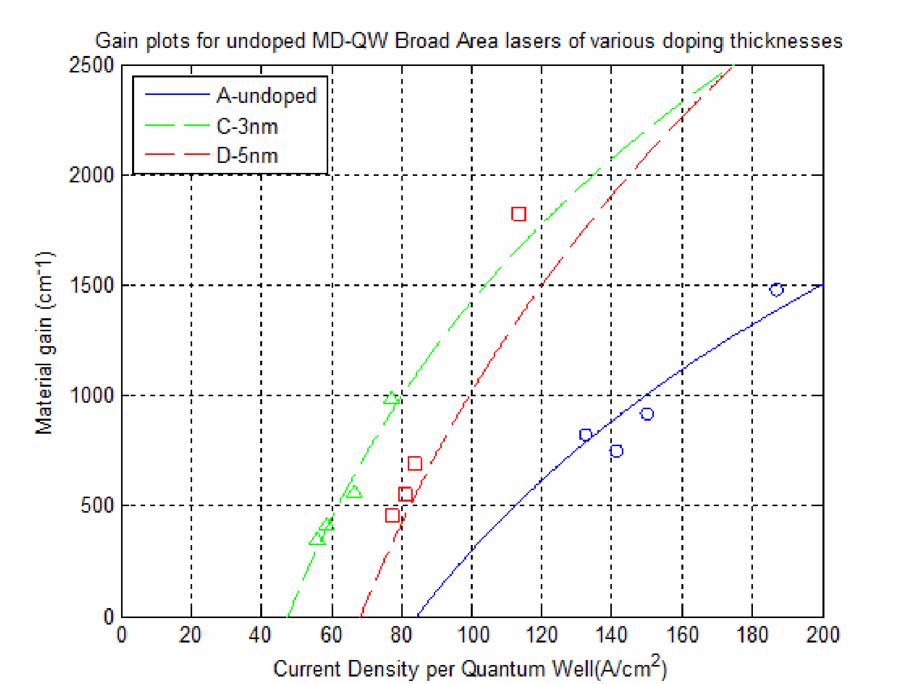
- Theoretical curve fits to single mode data (*I* < 1.5 mA)
  - Extrapolated to higher currents assuming single mode is maintained
  - > 40 Gb/s predicted

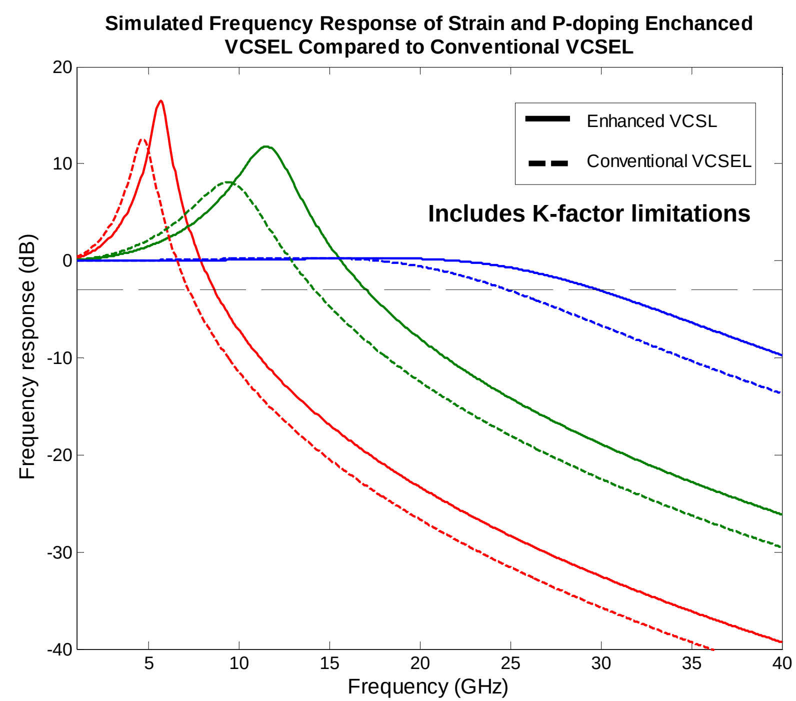
- Further improvements anticipated
  - Add p-modulation doping
  - Add additional strain
  - 50% increase in differential gain realized
  - > 50 Gb/s predicted

P-Modulation Doping Improvements

- **Strained InGaAs QWs with p-type modulation doping**
  - **increase differential gain**
  - **Reduction in transparency carrier density**

- **Broad area lasers demonstrated >50% reduction in threshold current**
- **Can lead to marked increase in frequency response over conventional VCSELs**

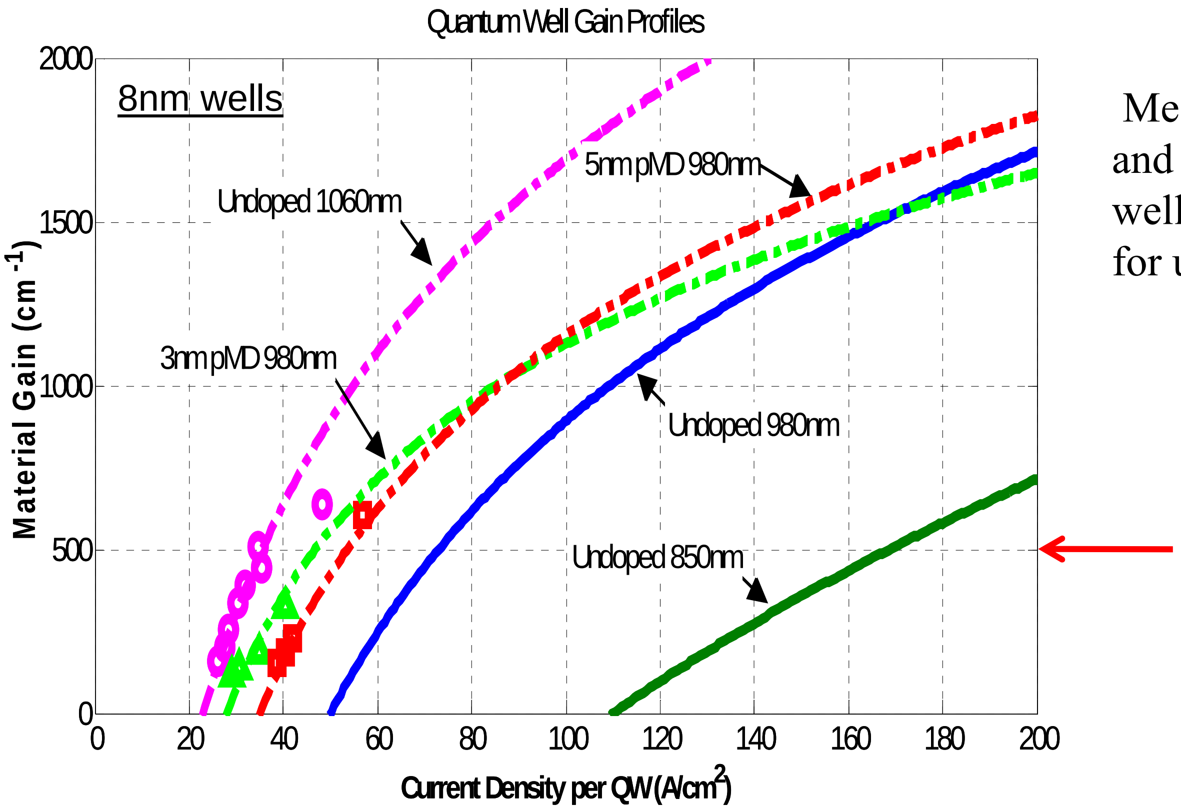




Gain Characteristic Comparison: 850, 980, 980 MD, 1060 nm

UCSB

- **Data from edge-emitters cleaved to various lengths**



Measured data used for 1060 nm and modulation-doped 980nm; well-established theoretical plots for undoped 980 nm and 850nm

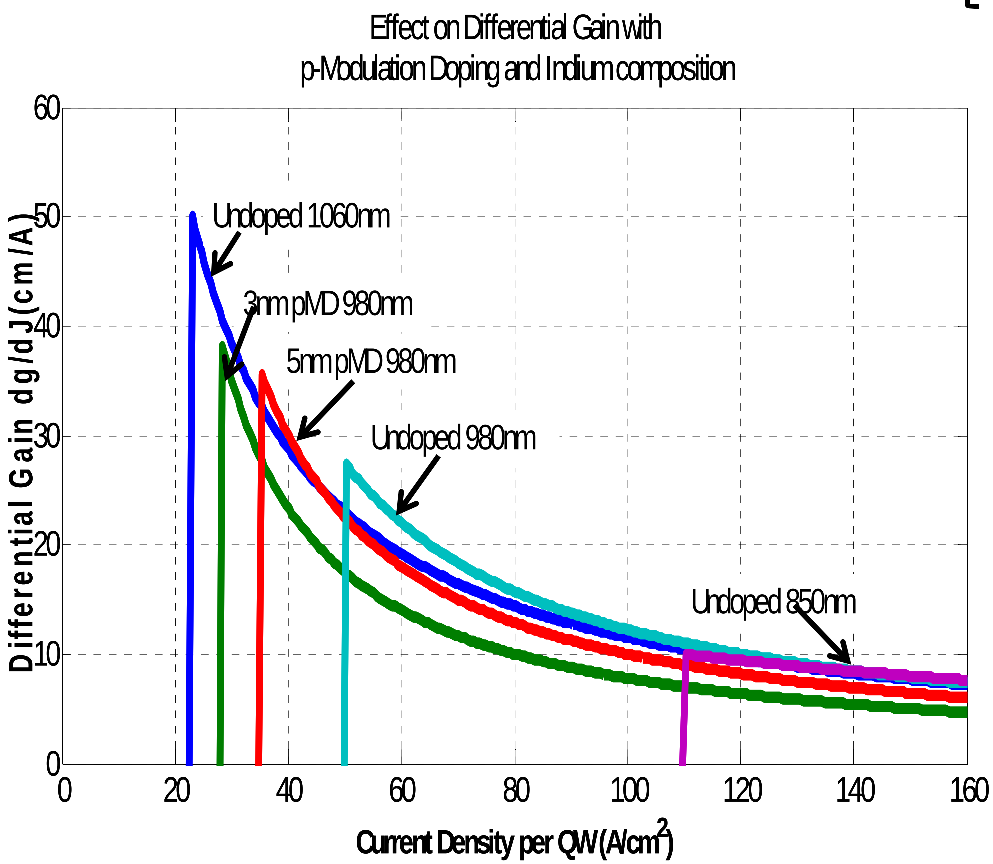
- **For 500 cm^{-1}:**
  - **InGaAs QWs @ 1060 nm require ~ 21% of the current vs GaAs**
  - **InGaAs QWs @ 980 nm require ~ 43% of the current vs GaAs**

Differential gain – the "*a*-factor"

Differential Gain, $a = dg/dN$

$N \;\alpha\; \sqrt{J}$

$$f_R = \left[\frac{v_g a}{qV_p}\eta_i(I - I_{th})\right]^{1/2}$$



"*a*" is ~ 5x larger for 1060 nm InGaAs compared to GaAs QWs

→ Implies ~ 2.2x inherently larger modulation bandwidth (neglecting K-factor limitations)

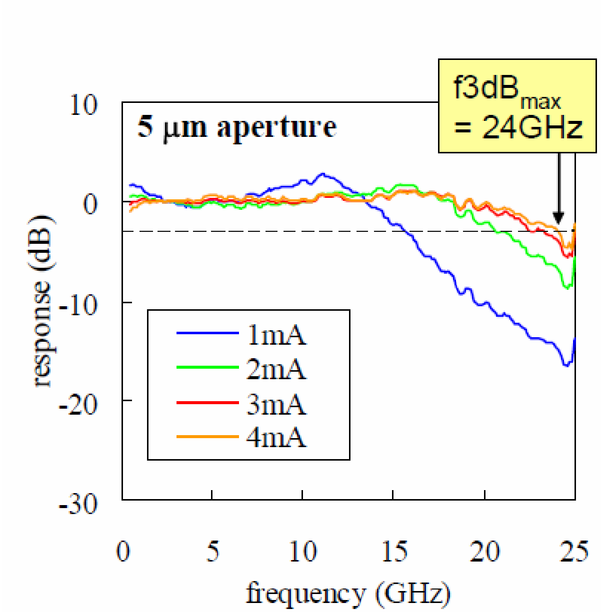
NEC Results: 1100 nm High-Speed

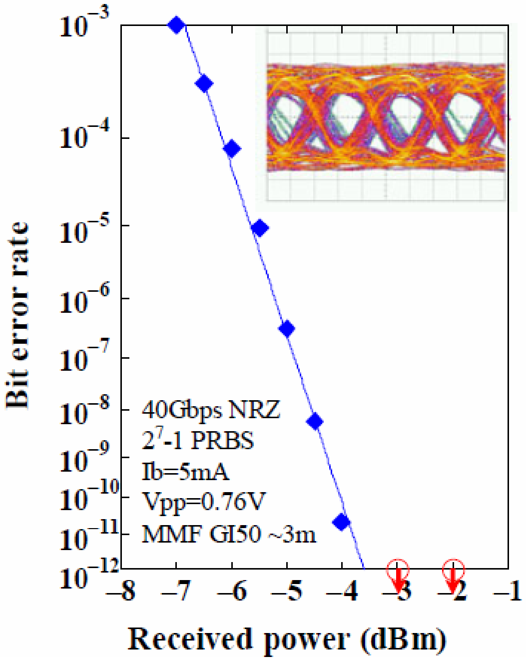
High-speed 1.1-μm-range InGaAs VCSELs

T. Anan, N. Suzuki, K. Yashiki, K. Fukatsu, H. Hatakeyama, T.Akagawa, K. Tokutome and M. Tsuji

Nanoelectronics Research Laboratories, NEC Corporation, Shiga

Paper OThS5, OFC 2008





• Developed high-speed 1.1 μm InGaAs VCSELs for optical interconnections. A wide bandwidth of 20 GHz and error-free 30 Gbps 100 m transmission have been achieved with oxide confined VCSELs.

• Developed BTJ-VCSELs with high modulation bandwidth up to 24 GHz.

NEC Results: 1100 nm & Reliability

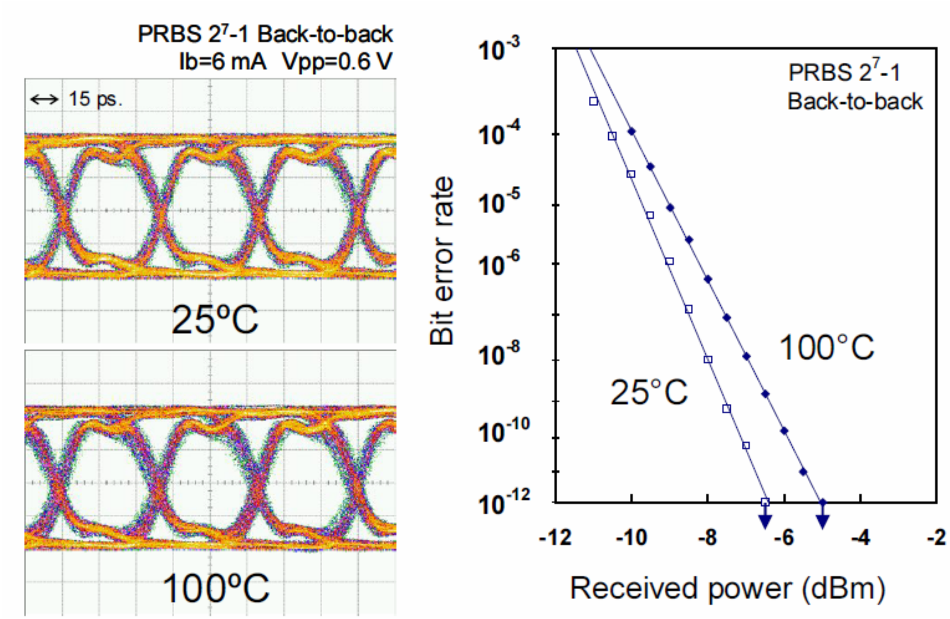
UCSB

Highly reliable high speed 1.1µm-InGaAs/GaAsP-VCSELs

H. Hatakeyama*, T. Anan, T. Akagawa, K. Fukatsu, N. Suzuki, K.Tokutome, M. Tsuji,
Nanoelectronics Research Laboratories, NEC Corporation, 2-9-1 Seiran, Otsu, Shiga, Japan.

Proc. of SPIE Vol. 7229 722902-2 (2009)

- 25 Gb/s
- BER < 10^{-12}



- Developed 1.1-µm-range oxide-confined VCSELs with InGaAs/GaAsP MQWs, and demonstrated 25 Gbit/s-100°C error-free operation.
- Investigated reliability of the VCSELs, and results of accelerated life tests showed extremely long wear-out MTTF life times of <u>10 thousand hours </u>under a junction temperature of <u>208°C.</u>
- Revealed that a major failure mode of the device was caused by <110> DLDs, which generated in the n-DBR layers.

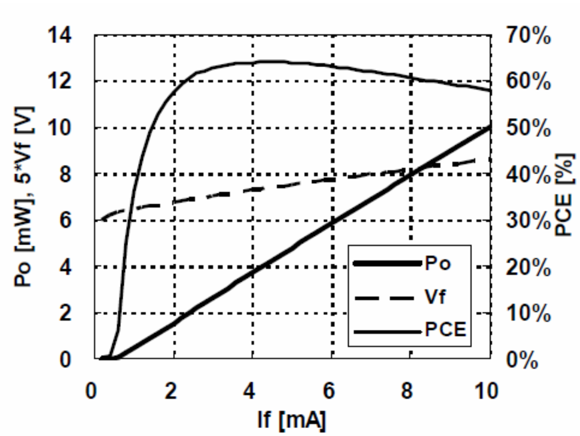
Furukawa Results: 10 x 12Gb/s, High Efficiency & Reliability

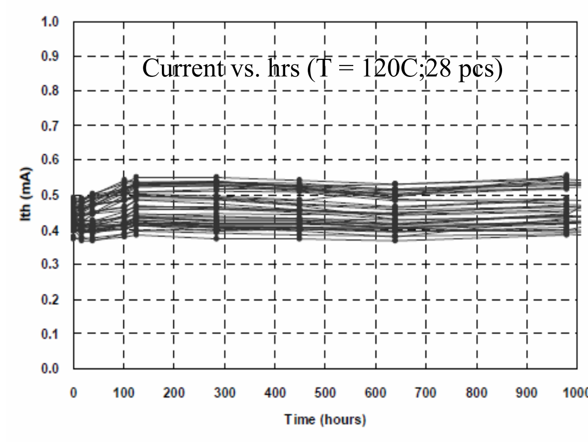
Experimental demonstration of low jitter performance and high reliable 1060nm VCSEL arrays for 10Gbpsx12ch optical interconnection

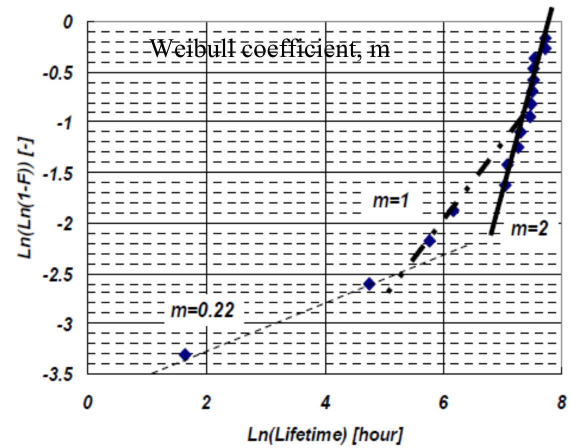
Keishi Takakia, et al

Photonics Device Research Center, Furukawa Electric co., Chiba, Japan, 290-8555

Proc. of SPIE Vol. 7615 761502-2 (2009)

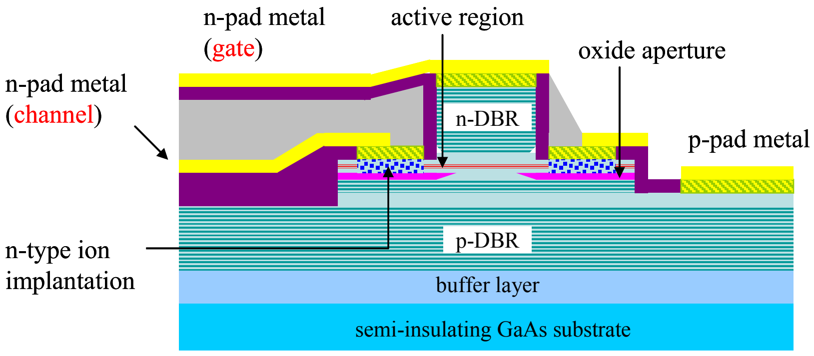


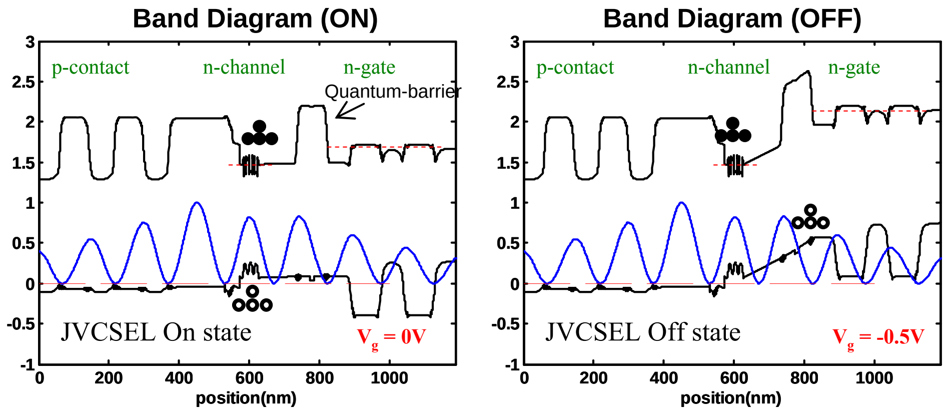




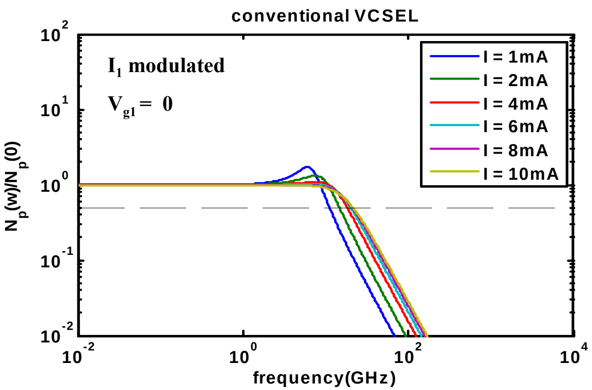
- Demonstrated high reliability and low power consumption operation with Furukawa's 1060nm VCSEL arrays.
- Power dissipation per speed of <u>5.5Gbps/mW</u> would be the most power saving VCSEL operation to our knowledge.
- Also, highly reliable performance was verified, especially no degradation of threshold current and eye diagram was demonstrated experimentally. Estimated FIT number was 81 FIT/ch and if one solved failure mode would be removed, 48 FIT/ch and 576 FIT/array would be potentially expected.
- A wide operating time in random failure regime was shown through the extra high stress test as <u>120^{o}C and >40kA/cm^2</u>.
- More than 800 years of operation in normal operating condition as 40^{o}C and 6mA was obtained under E_a=0.7and n=3.

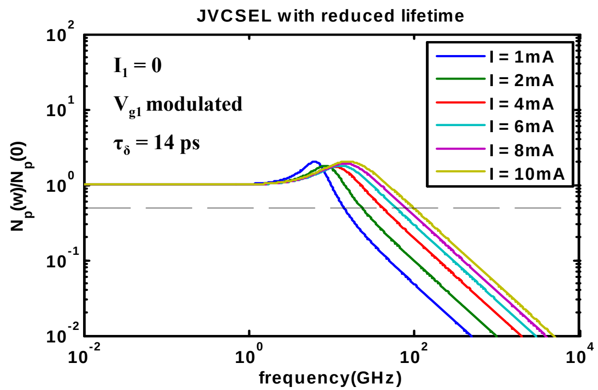
Summary of Charge Depletion VCSEL (JVCSEL)

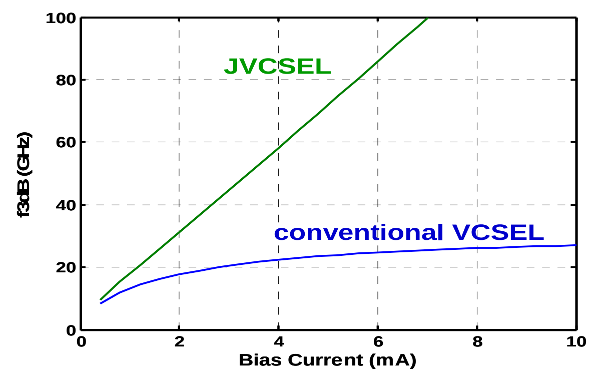




• **Novel quantum-barrier design to separate n regions**







- Parameters from experimental diode VCSELs (f_{3dB} ~ 25 GHz @ 6 mA)
- JVCSEL (f_{3dB} ~ 85 GHz @ 6 mA)→ BW continues to increase with bias (photon rate equation is directly modulated by modulating charge separation—gain—not current)
- Gate parasitic capacitance can be made low by increasing set-back to trade off RC cut-off with drive voltage (for V_g = 0.5V, f_{RC} ~ 150 GHz)

Coldren IEEE Summer Topicals 072010

Conclusions/Summary

- Photonic devices can address power dissipation as well as SWaP issues in both the switching and optical interconnect fabrics of data centers
- Practical, highly-reliable, high-efficiency, high-bandwidth VCSEL sources having a number of advantages can be made, but are not being widely developed commercially—1060 vs. 850 nm issue. Should this change?
- Direct modulation rates ~ 50Gb/s seem viable; rates of 100Gb/s are conceivable in the future without the need for high current densities
- Maximum data rates determined by electronics and multiplexing circuits, not O/E devices
- Techniques as WDM or advanced modulation formats may add additional parallelism desired to keep data rates finite, once serialization & space-division approaches saturate.