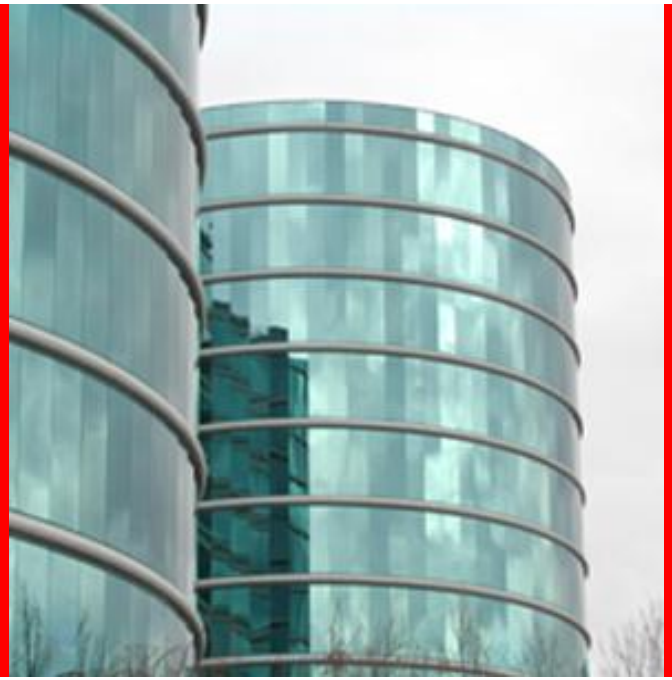


ORACLE®

Digitally-assisted analog circuits for a 10 Gbps, 395 fJ/b optical receiver in 40 nm CMOS

Philip Amberg, F. Liu, M. Dayringer, J. Lexau, D. Patil, J. Gainsley, H. Moghadam, E. Alon, X. Zheng, J. Cunningham, A. Krishnamoorthy, and R. Ho

Trends in computer system design

Exploiting Moore's Law

Improved performance from increased integration

- 100s (to 1000s) of cores per chip in the near future

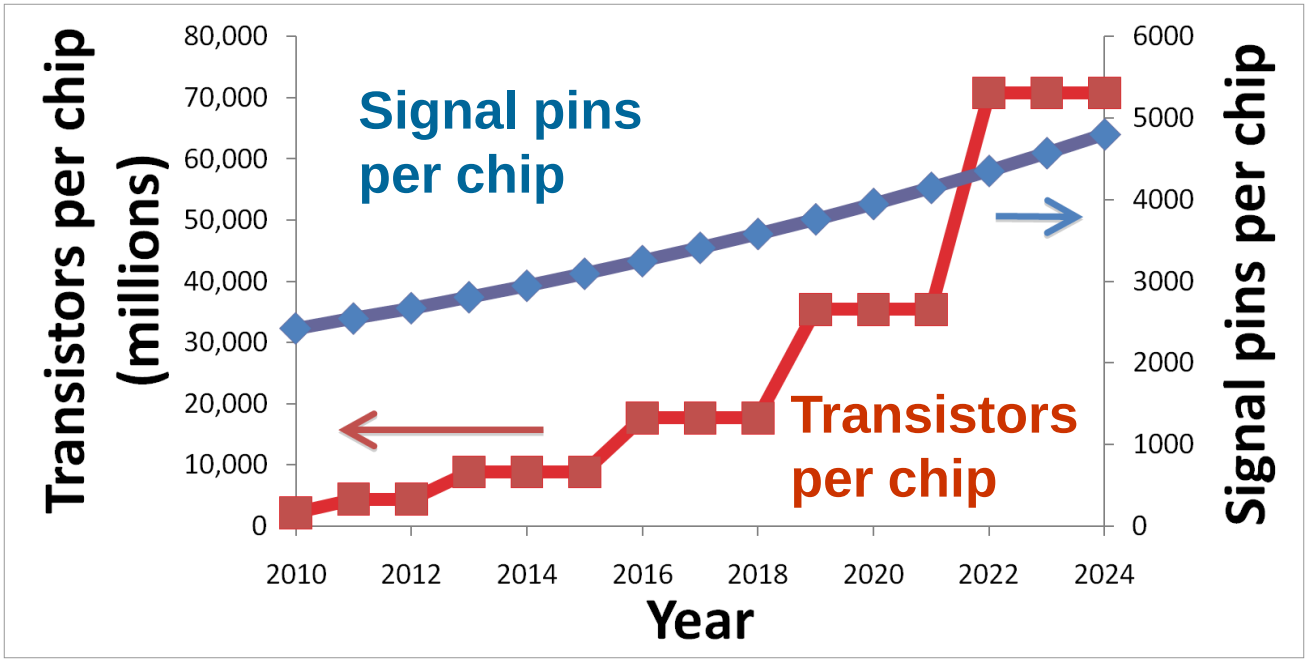
Several benefits

- Pack together lots of functional units
  - A big win if we can exploit task/thread parallelism
- Cut global communication latency
  - "Bandwidth is about money. Latency is about God."
- Cut total system power dissipation
  - Need less power to communicate if things are close together

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

There is no free lunch

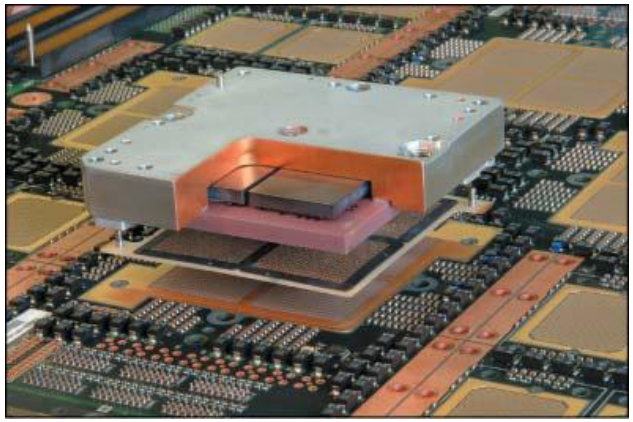
- Aggregate off-chip bandwidth is scaling slowly
  - We don't have enough chip input/output signals
  - This causes an IO bottleneck at the chip boundary
  - Forces us to *further increase* chip integration (caches)



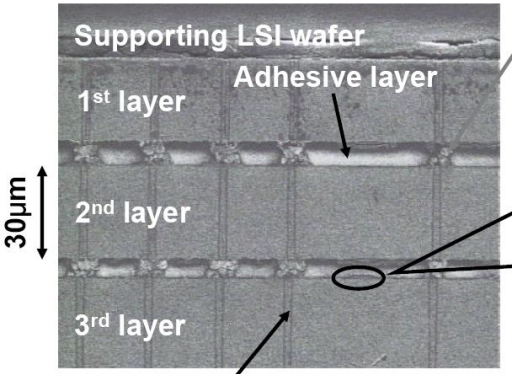
This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Lots of work trying to solve this problem…

- Integrate vertically with TSVs (build up)
  - Complex packaging and assembly, but it will be here soon
  - Thermal issues limit the architectural use cases (memory)
- Integrate horizontally with fine pitch IO (build out)
  - Complex package, assembly, and test with MCMs
- Known-good die issues



Source: IBM



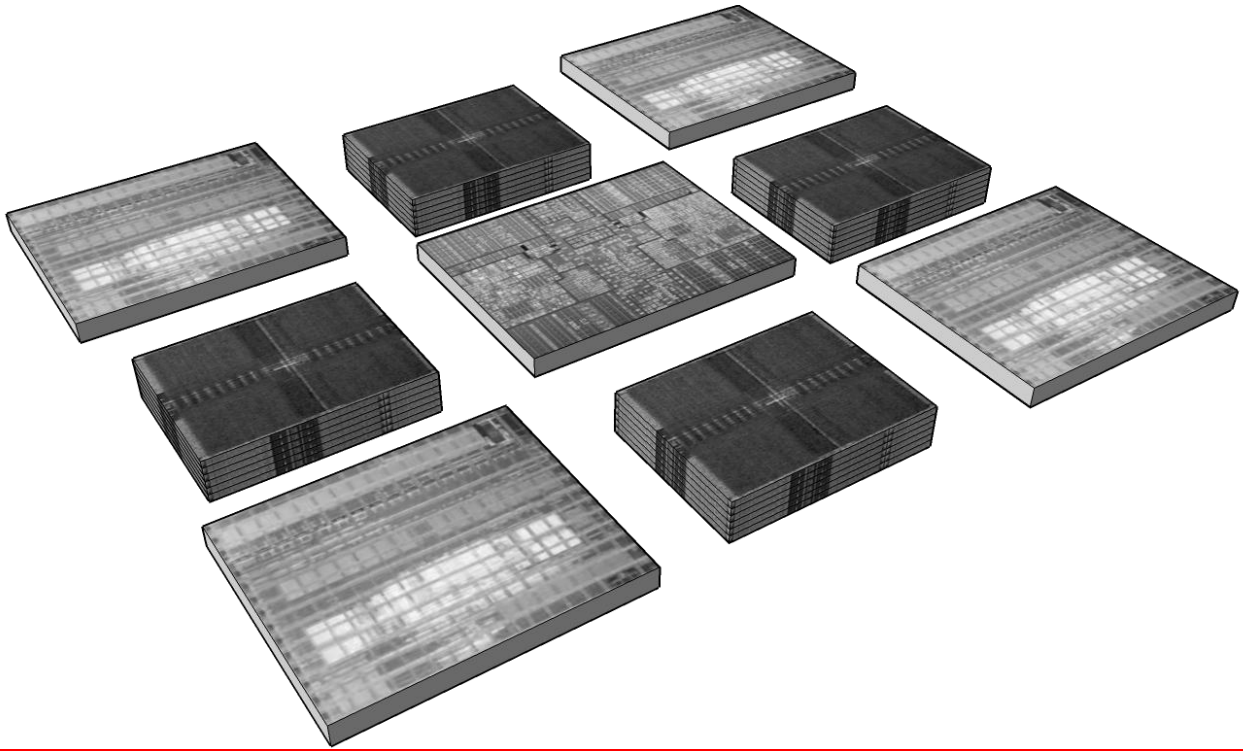
Source: Samsung

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

An ideal (?) solution

Enable systems with chips that build both up and out

- Aggregate TSV'ed chip stacks with a fine pitch IO
  - But without conductive (soldered) attachments, to help yield
  - Needs a chip-to-chip IO with enormous BW and low energy

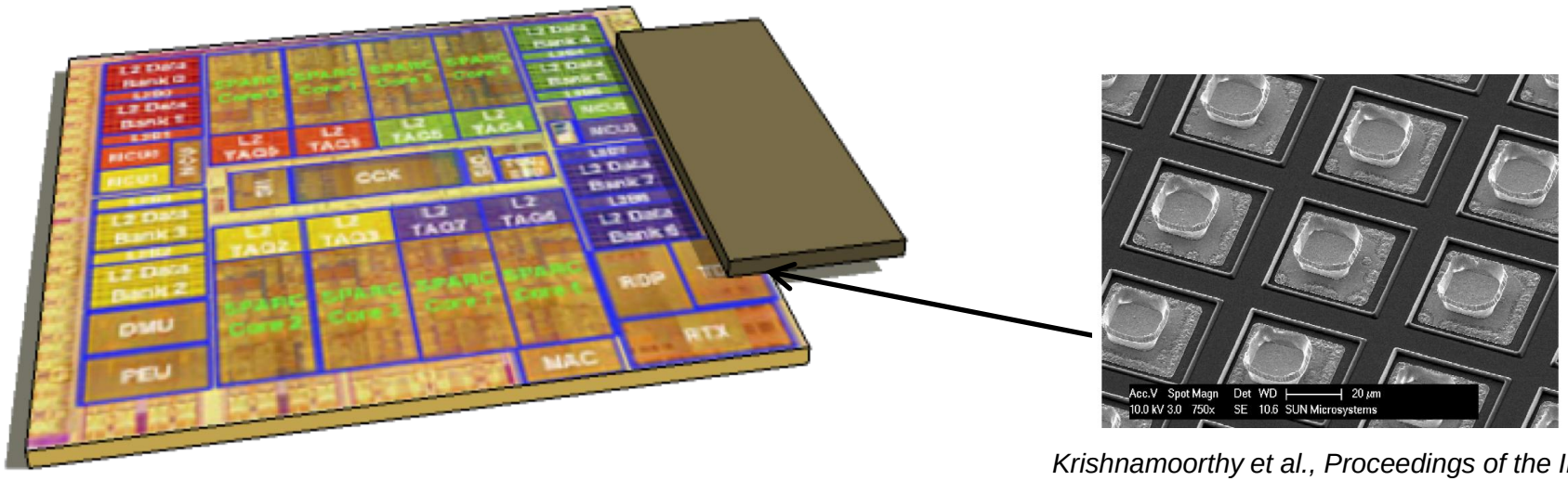


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

One possibility: optically-enabled silicon

Create enormously dense IO channels over fast optical media

- Optical channels are an interesting solution
  - Wavelength division mux'ing: lots of λs in a single waveguide
  - Signals travel at the speed of light on a silicon chip
- We need an electrical-to-optical interface
  - Use a "bridge chip" face-bonded to a CPU or mem controller
  - Bond with fine pitch solder tailored to IO (not for power)



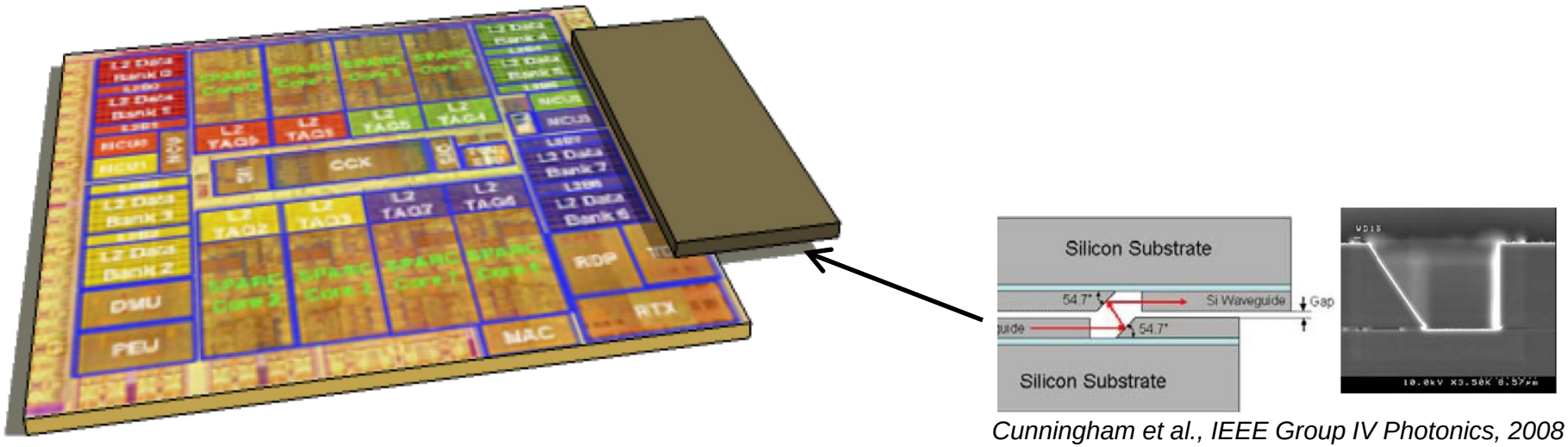
Krishnamoorthy et al., Proceedings of the IEEE, 2009

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Why this segregation of optics and electronics?

Why not put the optics directly on the silicon chips?

- "Hybrid" integration allows for optimal designs
  - CPUs and DRAMs want inexpensive bulk CMOS processes
  - Optical receivers need III-V materials
    - An unnecessary cost for processors and memories
- Optical waveguides on the bridge can point down
  - "Optical proximity communication" – etched mirrors in Si



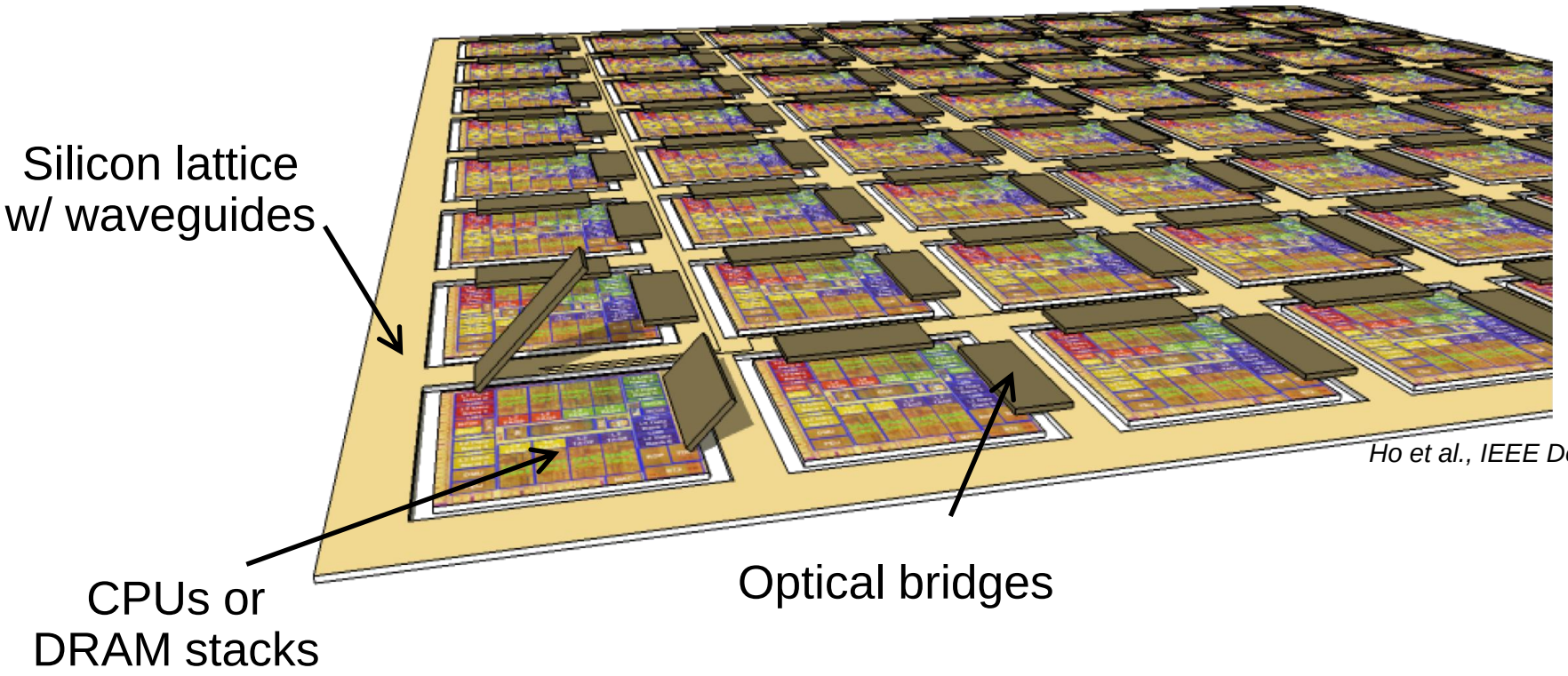
Cunningham et al., IEEE Group IV Photonics, 2008

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

The “macrochip” – a server-on-a-chip vision

Aggregate the hybrid chips onto a substrate with embedded waveguides

- Silicon lattice carrying CPUs and DRAMs
  - With a very rich waveguide network connecting the sites
  - Bridges perform OE/EO and couple optically to lattice

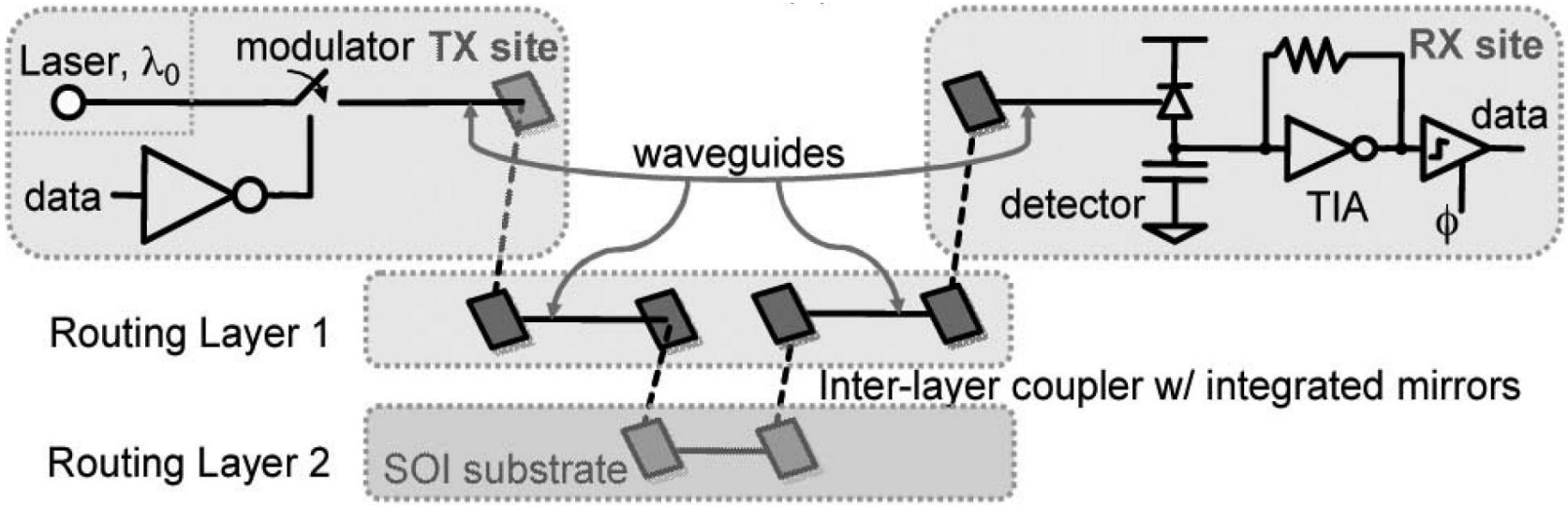


Ho et al., IEEE Design and Test, 2010

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

And of course the most important part: the link

- The goal is to hit 0.3 pJ/b energy consumption
  - Commercial electronic links today around 20 pJ/b
- We have demonstrated < 0.5 pJ/b thus far (2011)

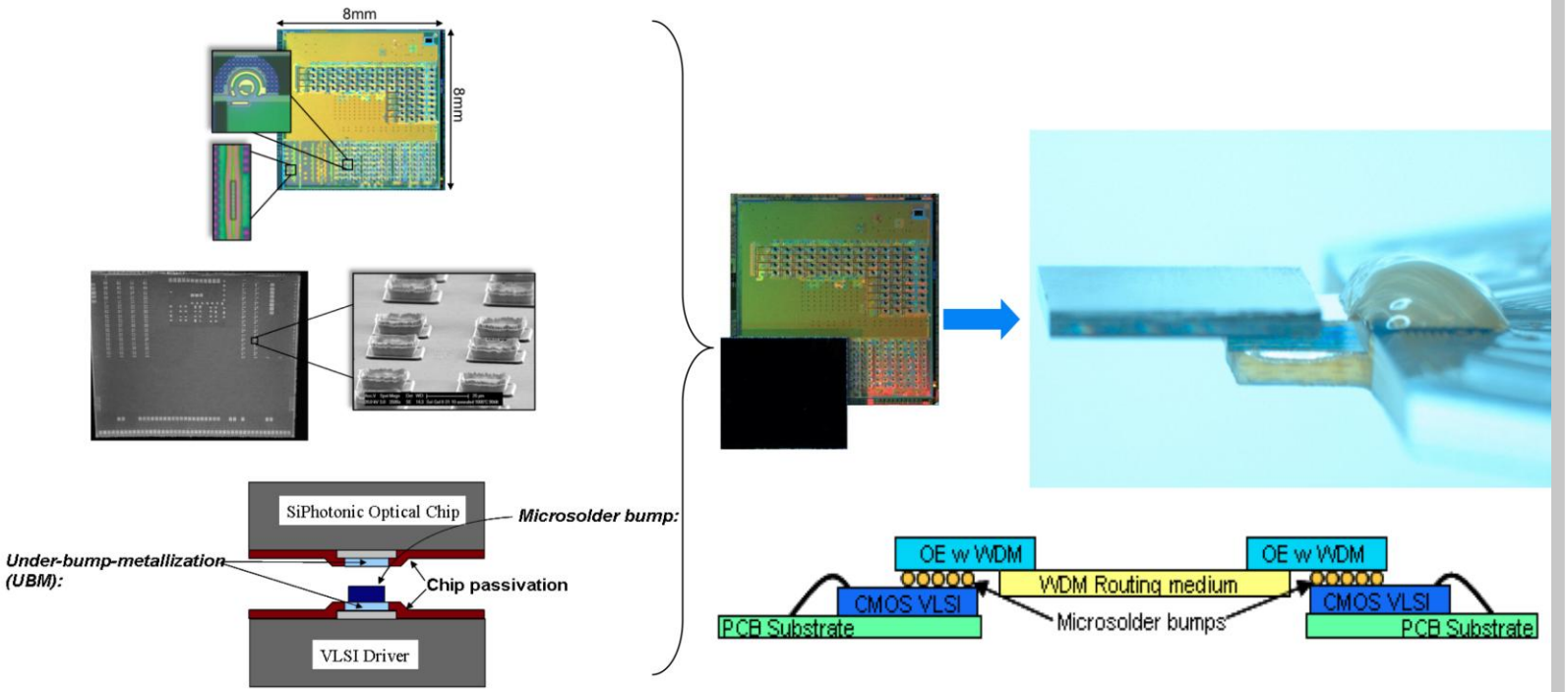


Krishnamoorthy et al., JSTQE 2011

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Some recent results

Low parasitic hybrid integration



- Hybrid integration using microsolder bumps
  - 25μm pads (R<1ohm, C<20fF) scalable to 10μm with a few fF.
  - Integration of best-of-breed photonic and electronic circuits with little performance compromise

ORACLE

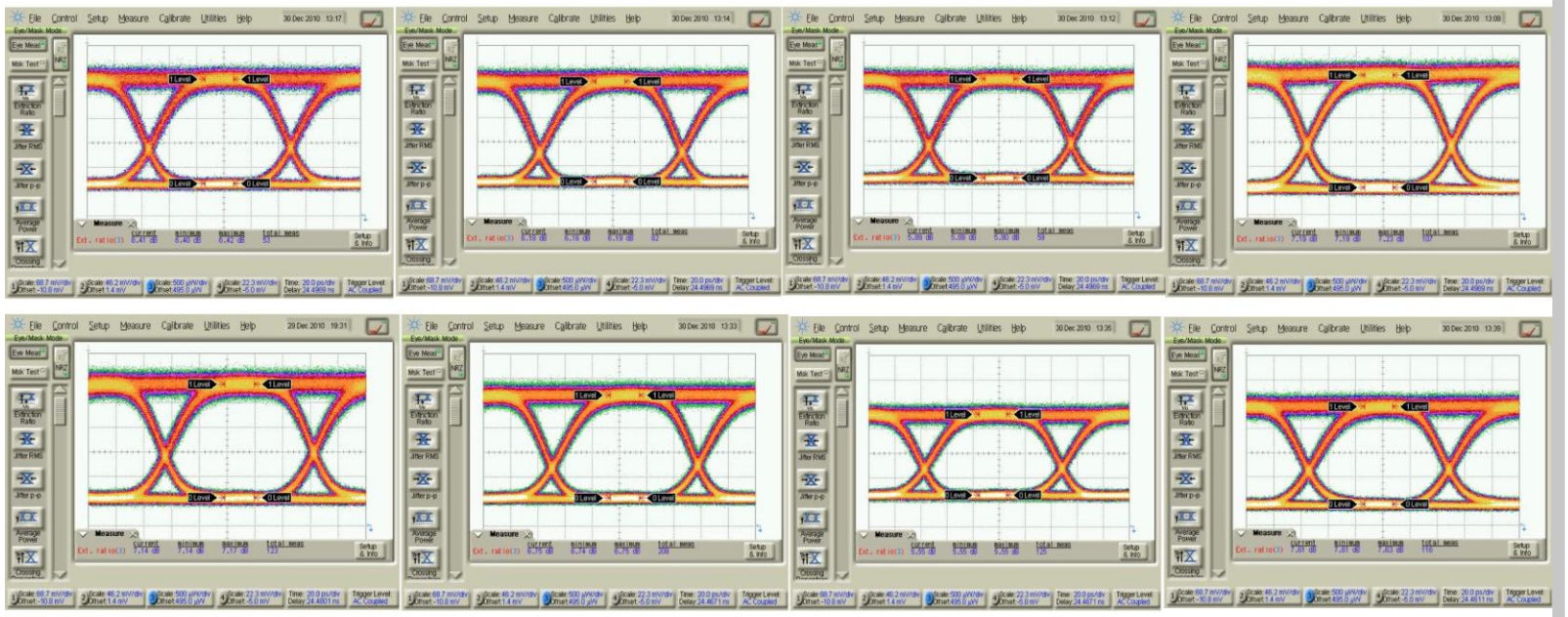
Approved for public release. Distribution unlimited. 9

Zheng, OFC post-deadline, 2011

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Some recent results, cont.

10G Optical "Eye" Diagrams



Chanel	1	2	3	4	5	6	7	8
Power(mW)	0.94	0.71	0.70	0.70	0.92	0.68	0.66	0.91

- 1x8 10G silicon photonic WDM transmitter array
  - Dynamic extinction ratio 5.6~8dB
  - Measured power less than 1mW (100fJ/bit) per channel

ORACLE

Approved for public release. Distribution unlimited.

12

Zheng, OFC post-deadline, 2011

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Why 10 Gbps?

- Practical, low-power design point
- Simple, power-efficient circuits can be used
  - $f_{op} << f_t$
- Avoid power hungry Ser-Des circuits
  - Data rate low multiple of CPU clock rate
  - DDR clocking used instead
- Inexpensive optical channels can be used
  - Low-loss at high frequency less important

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

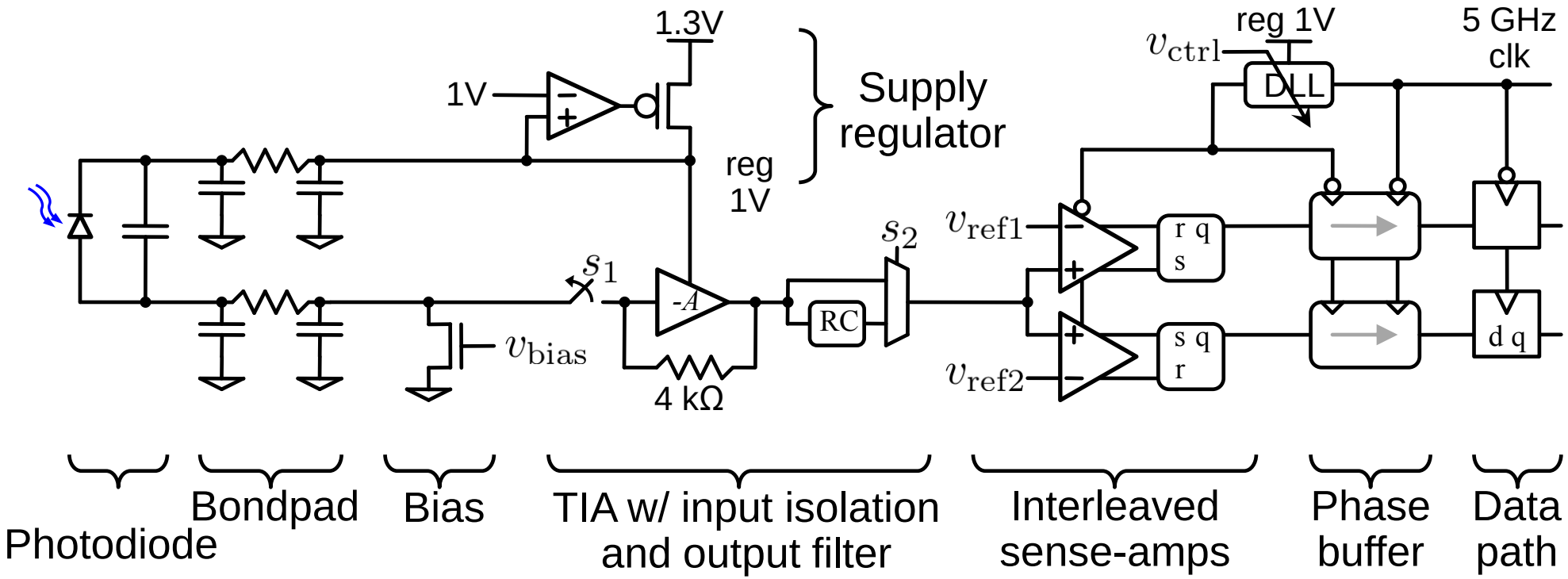
The big challenge is energy efficiency

- Traditional analog techniques are well-understood
  - 10Gbps is a very achievable target for a 40nm technology
  - However, doing it at under 0.5 pJ/b is not so simple
- Focus on energy leads to interesting issues
  - No limiting amplifiers, but rather sense-amplifiers
    - Need an efficient way to compensate offsets
  - No DC-balanced data
    - Need a way to DC-bias the receiver input
  - No continuous CDR
    - Need a way to center the clock in the data eye

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver topology

Overall receiver model

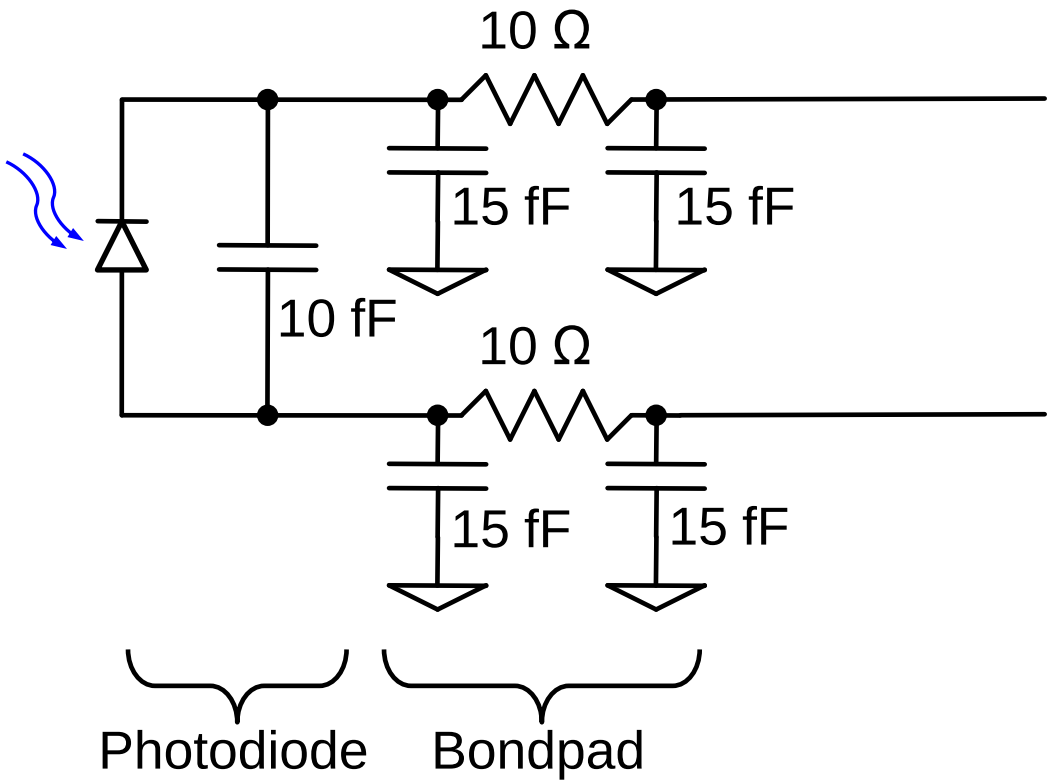


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver topology

Photodiode and Bondpad model

- Hybrid integration with microsolder enables low capacitance connection between VLSI and Optical chips

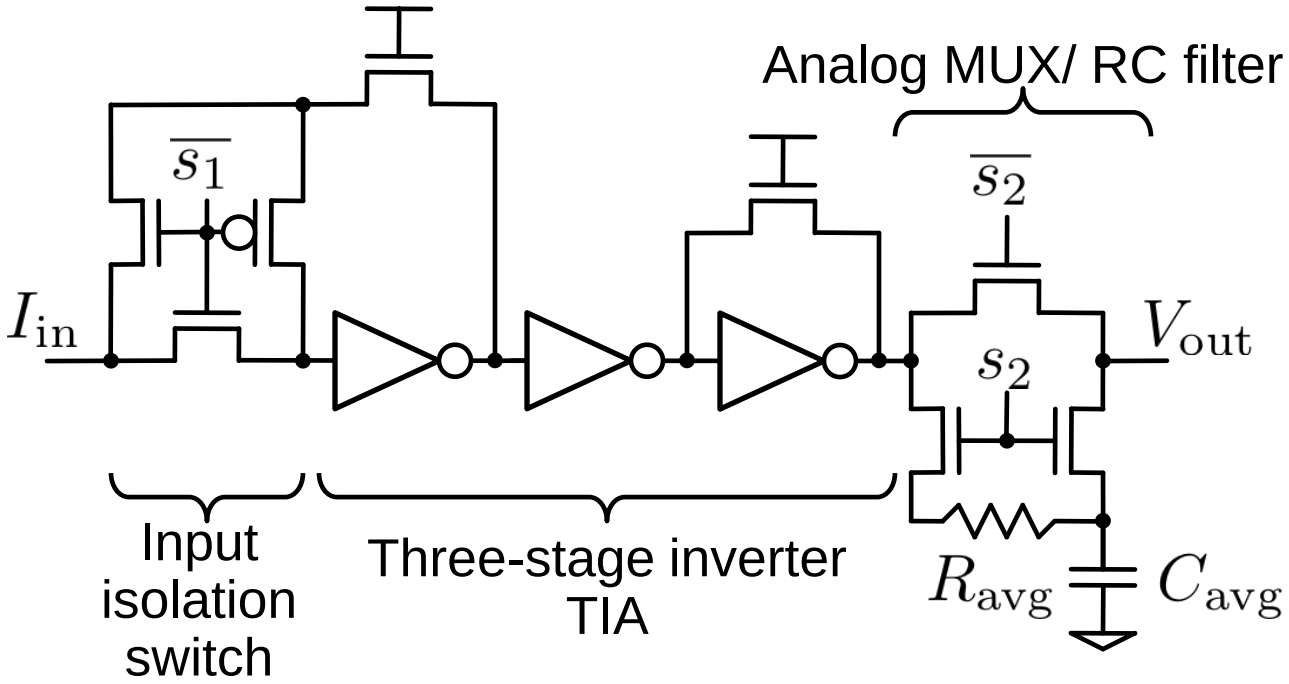


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver topology

TIA design

- Isolation switch disconnects TIA from photodiode during sense amp calibration
- Analog MUX chooses the RC filtered output during TIA input bias calibration

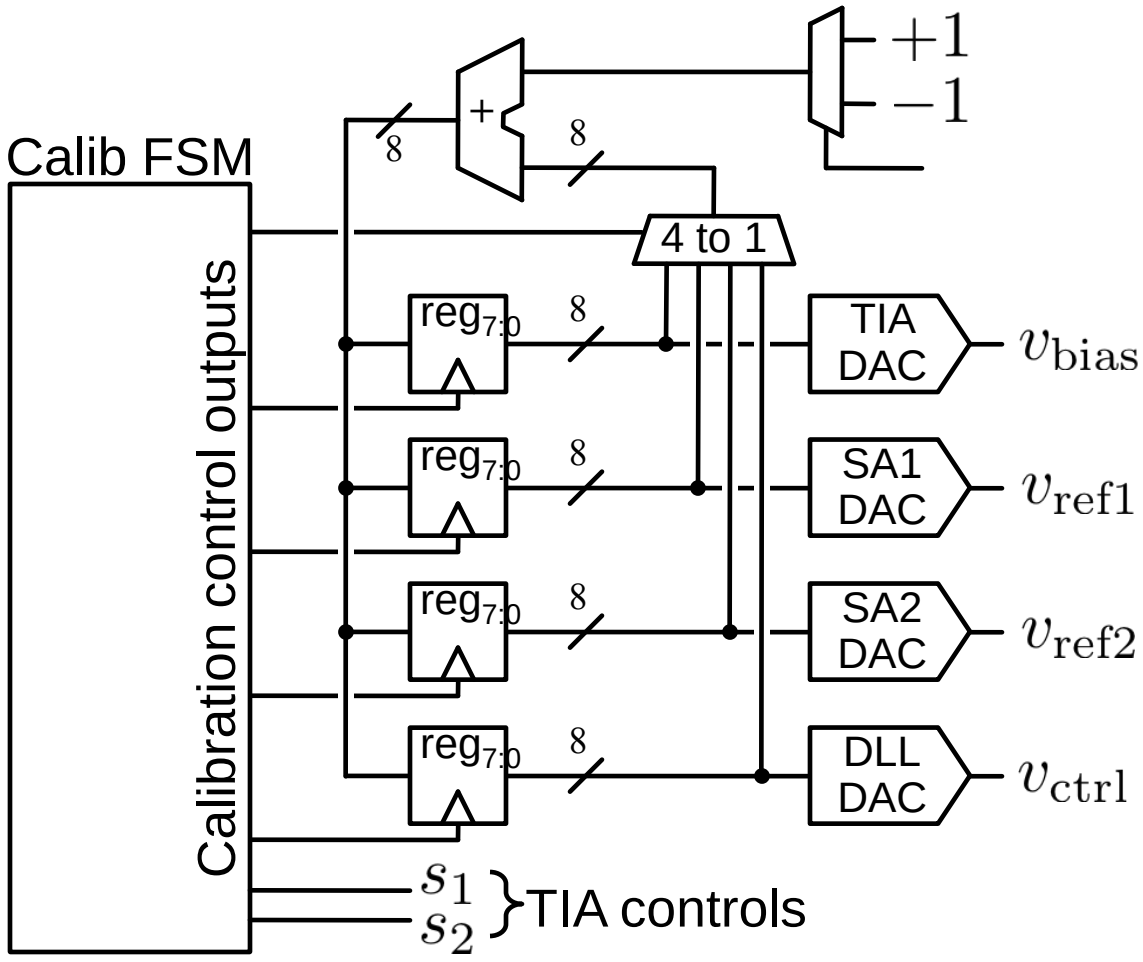


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver topology

Calibration FSM

- Calibration process is controlled with a FSM and 4 DAC's
- FSM controls the TIA operating mode and active counter
- Sense amp output increments or decrements the desired counter
- Clocked at 156 MHz



This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Overview

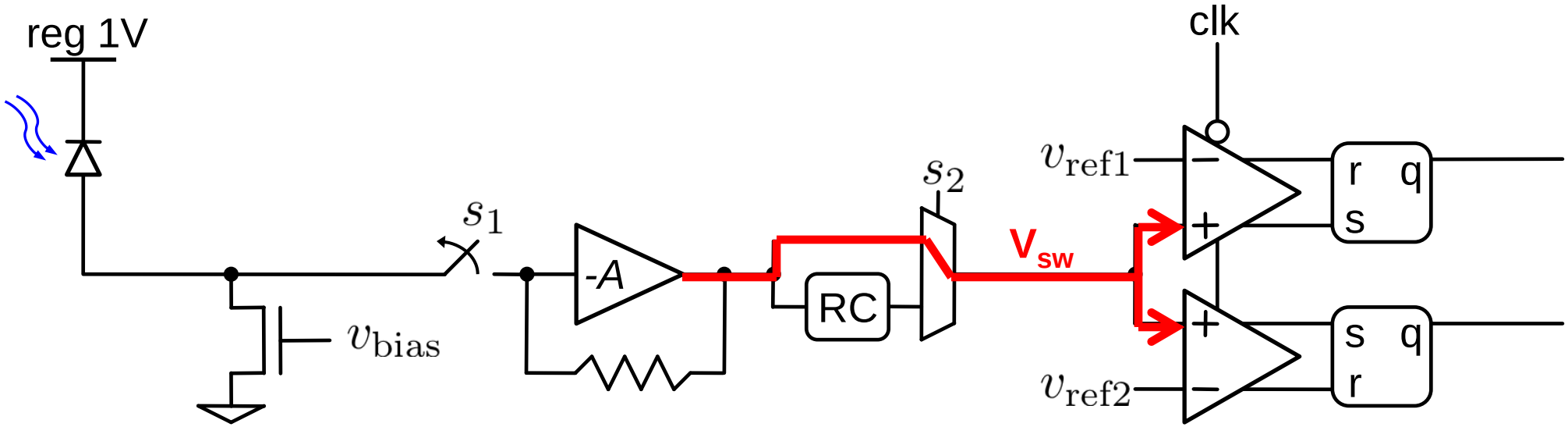
- Calibration needs to
  - Remove each sense amplifier's offset
  - Set DC bias level on the TIA input
  - Center the DLL output in the data eye
- Calibration initiated by service processor that broadcasts calibrate signal to all sites on the macrochip over a low-bandwidth electrical path
- Calibration performed periodically (~1 ms)
  - Frequent enough to correct for circuit drift
  - Infrequent enough to amortize power/performance overhead
- Transmitter sends 0101… pattern during calibration

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 1: Calibrate sense amplifiers

- Sense amplifiers are window into the receiver
  - Remove their offset first
- Disconnect photodiode from TIA (s_1 low)
  - Self biases TIA to high-gain region
  - TIA output held at switching threshold

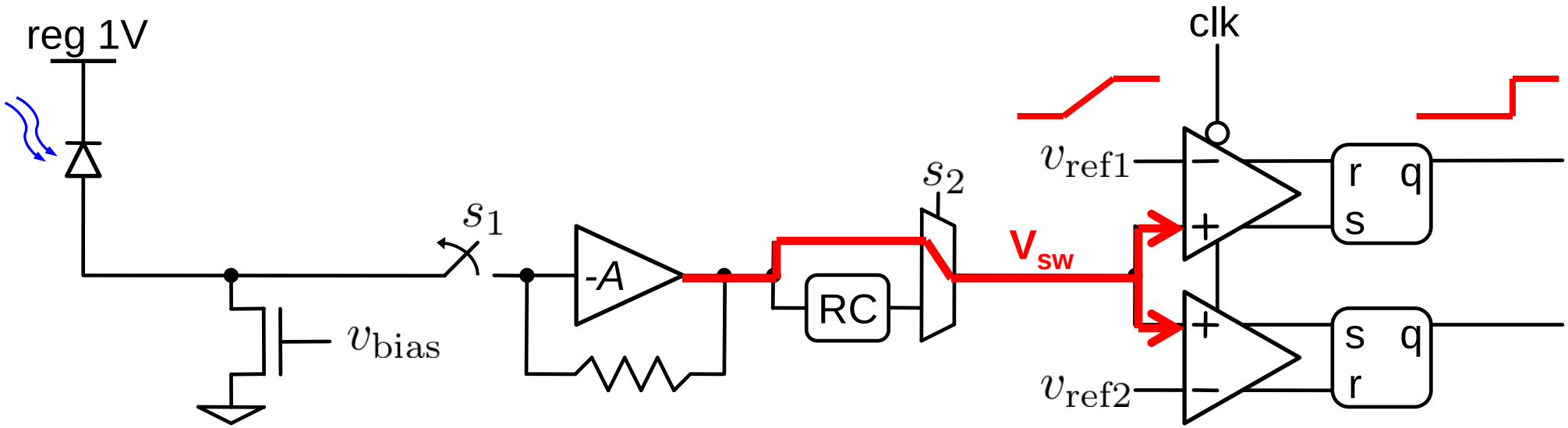


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 1: Calibrate sense amplifiers

- Ramp V_{ref1} from 0 V until sense amp output switches
  - $V_{ref1} = V_{sw} + V_{os1}$
- Repeat process for V_{ref2}
  - $V_{ref2} = V_{sw} + V_{os2}$

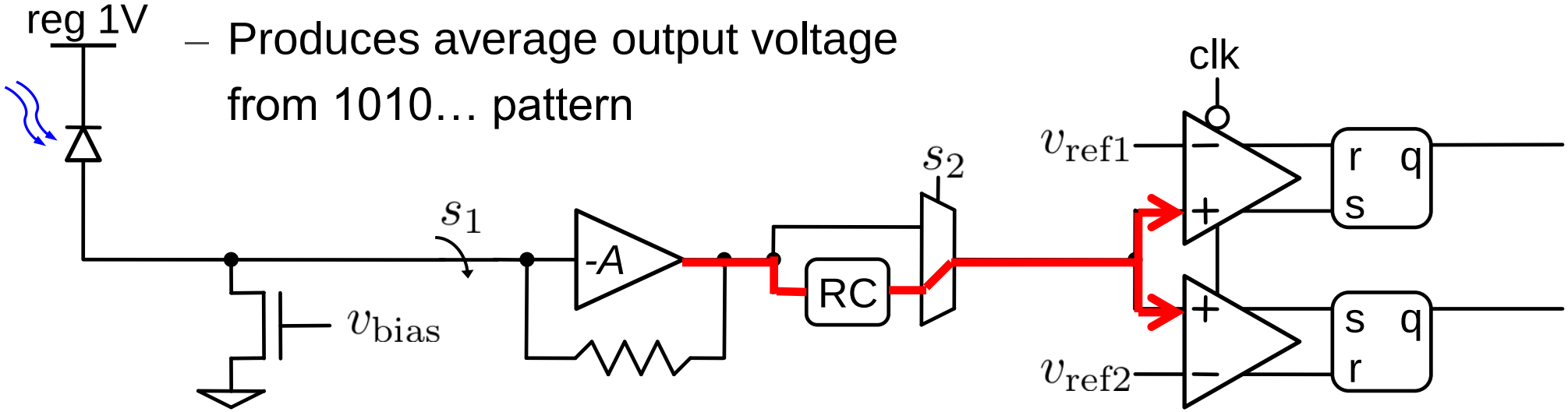


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 2: Calibrate TIA input level

- Photodiode has DC output current due to finite extinction ratio and dark current
  - Must be removed to bias TIA in high-gain region
- Connect photodiode to TIA (s_1 high)
- Filter TIA output through large RC pole (s_2 high)
  - Produces average output voltage from 1010… pattern

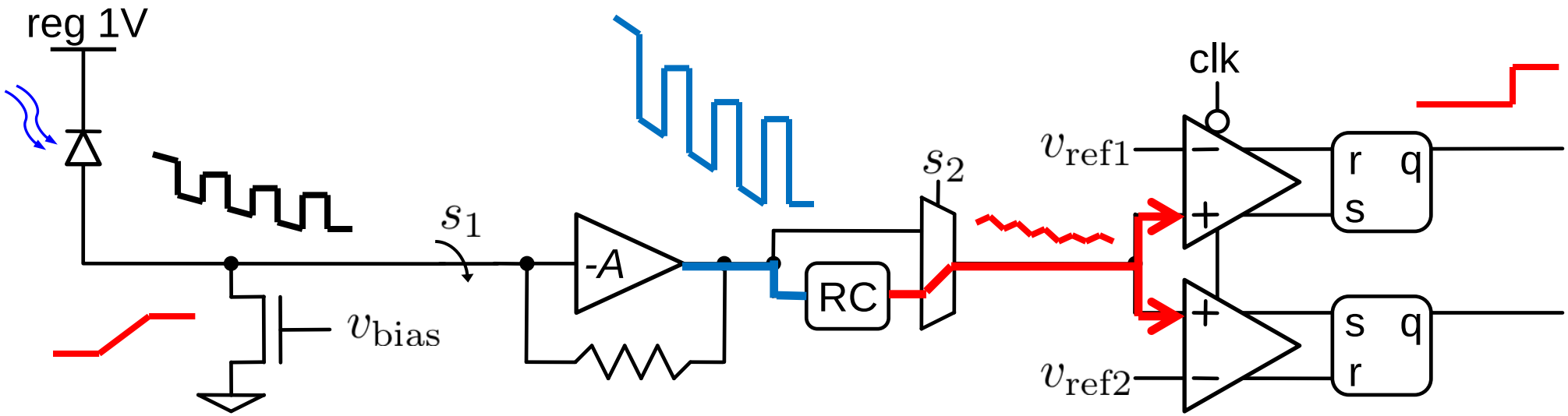


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 2: Calibrate TIA input level

- Ramp Vbias from 0V until sense amp output switches
- Sense amp output will switch when I_{DC} is pulled through bias device

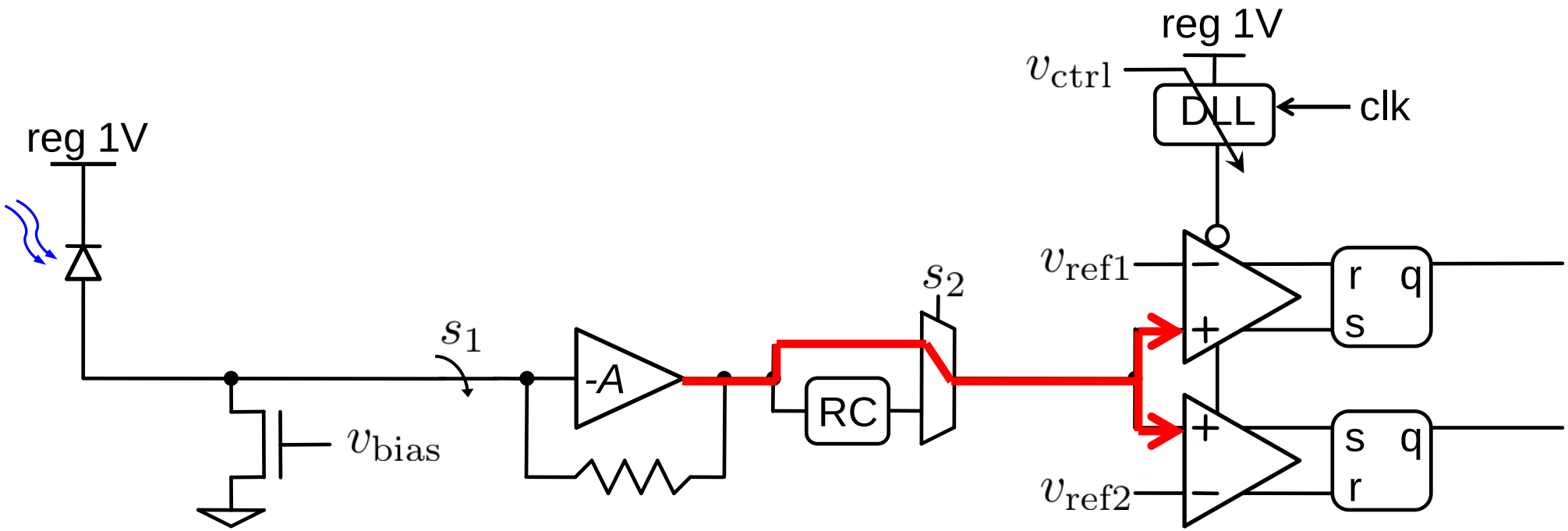


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 3: Calibrate sense amplifier timing

- Sense amplifier must sample at the middle of the data eye
- Select non-filtered TIA output (s2 low)

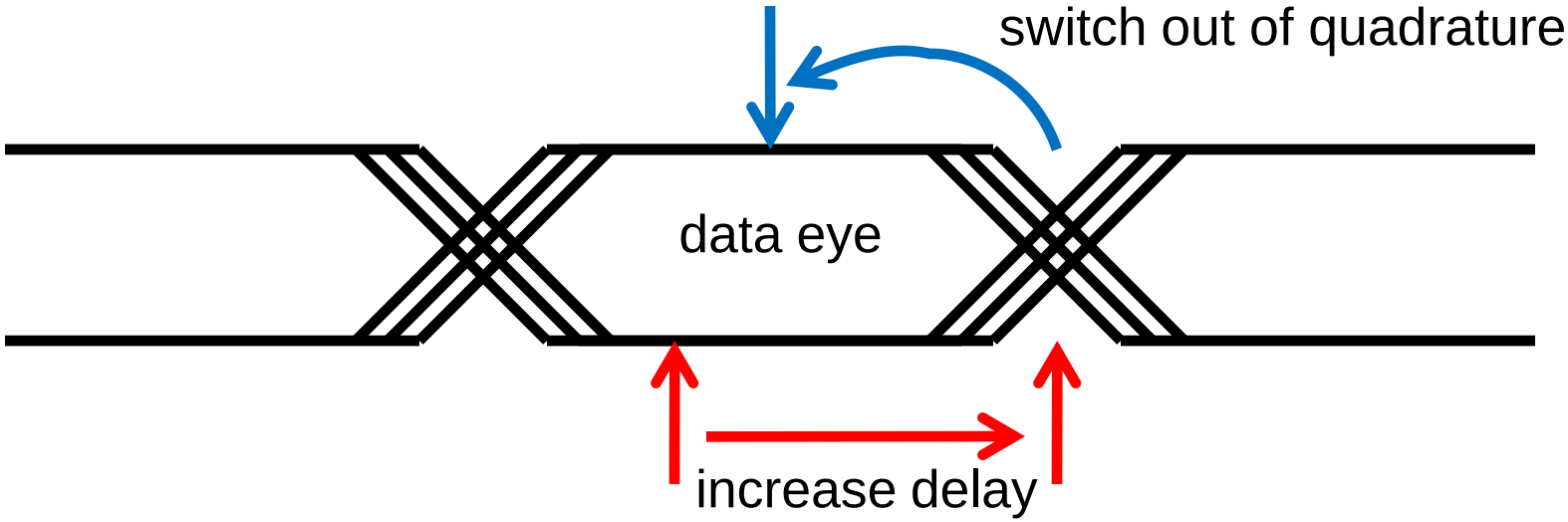


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Receiver Calibration Process

Step 3: Calibrate sense amplifier timing

- Receiver clock is switched into quadrature
- DLL delay is increased from minimum delay until sense amp output switches
  - Should occur at edge of data eye
- Receiver clock is switched out of quadrature
  - Centers clock edge in data eye

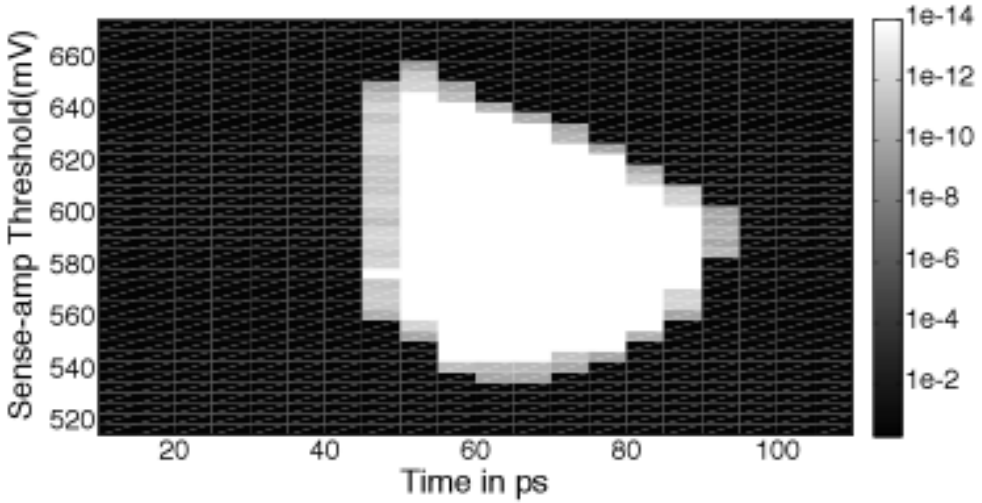


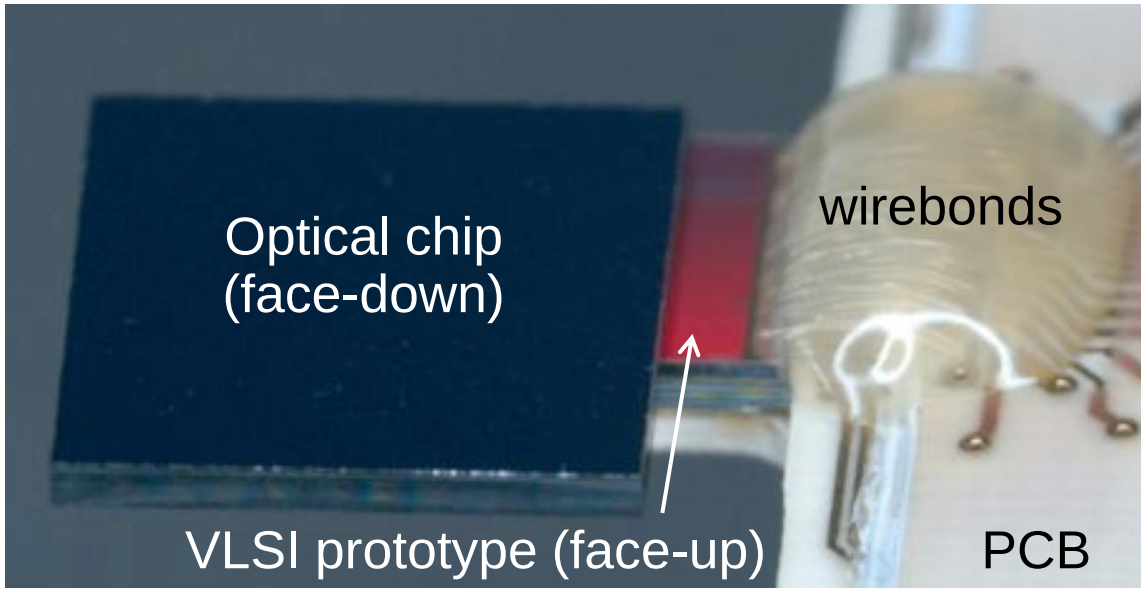
This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Results

Chip photo and measured data eye

- Data eye is shown at -15 dBm sensitivity (24 µA average current)
- Face-down optical chip hybrid-bonded to face-up VLSI chip
- VLSI chip wirebonded to PCB for test



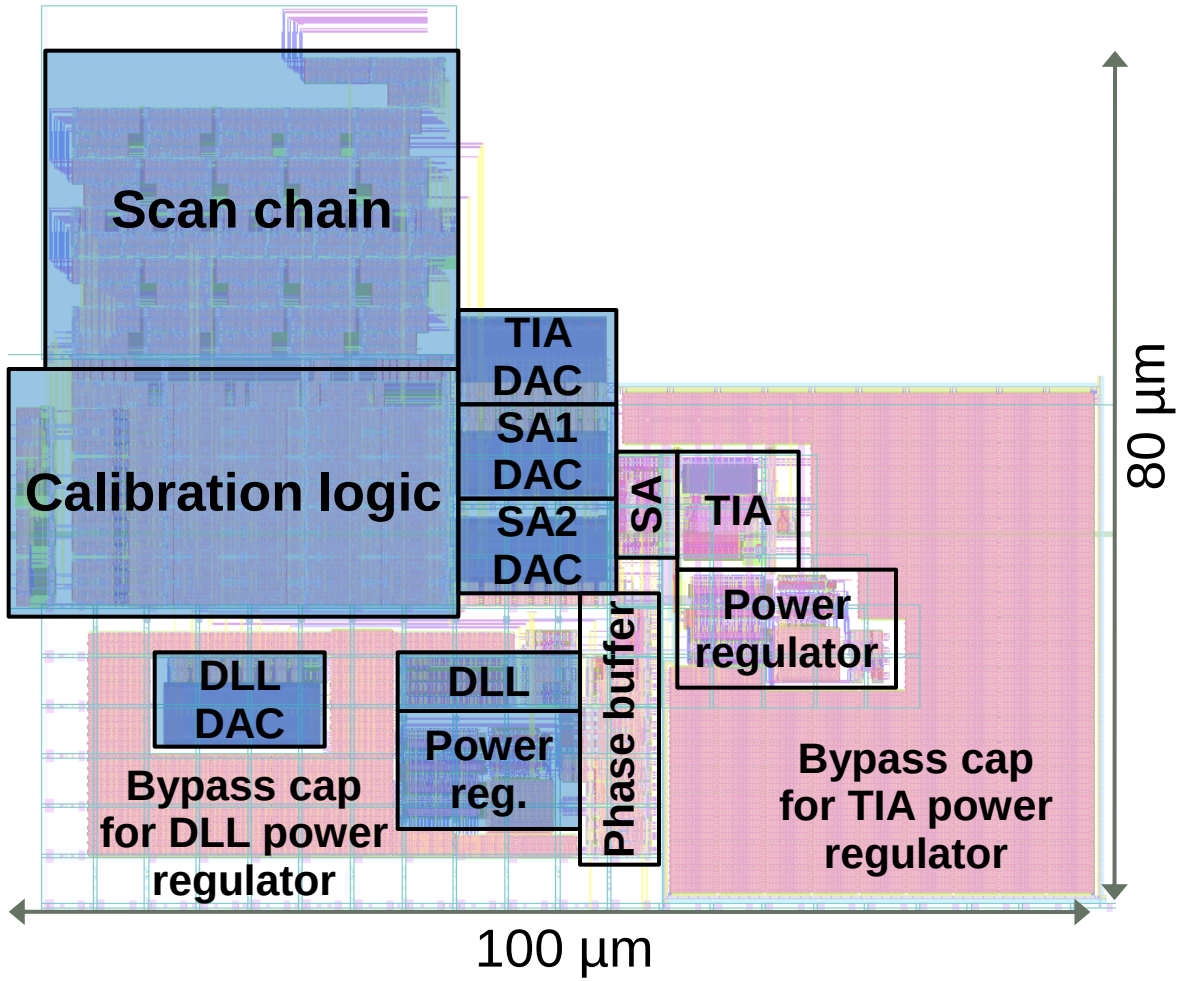


This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Results

Receiver layout

- Digital-assist circuits take 2800 µm²
  - Highlighted in blue
  - 40% of receiver area
  - Dominated by registers in calibration logic and scan chain



This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Results

Power overhead

- Calibration run every 1 ms, takes 8.2 µs
- Effective power from scalable blocks 1.12 µW
  - Static leakage current dominates
- The delay line consumes the vast majority of the power
  - 5 GHz clock with full activity factor
  - Its 1.78 mW already counted in 3.95 mW total rx power

Circuit	Power (scalable)	Power (static)
Digital calibration loops	136 µW	100 µW
DAC (x4)	n/a	376 µW
Delay Line	n/a	1.78 mW

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

Summary

- Demonstrated high-performance receiver with simple, low-power analog parts
- Digital calibration circuits compensate for non-ideal analog behavior
- Periodic tuning with digital circuits is an energy-efficient way to increase performance and will likely scale with technology

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.
The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

This work was supported in part by DARPA under Agreement HR0011- 08-09-0001. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.


ORACLE®