
Lecture 12

Silicon Test and Validation

Intel Corporation
jstinson@stanford.edu

Reading:

Introduction:

With design complexity and raw transistor counts growing at a 2X rate per generation, issues surrounding validation of silicon and test/manufacturing have become hot topics in the industry. Unlike software, hardware cannot be “patched” and must meet a much higher level of quality before being shipped to the customer. This lecture will go thru some of basic issues in both validation and manufacturing of digital designs.

Manufacturing vs. Validation

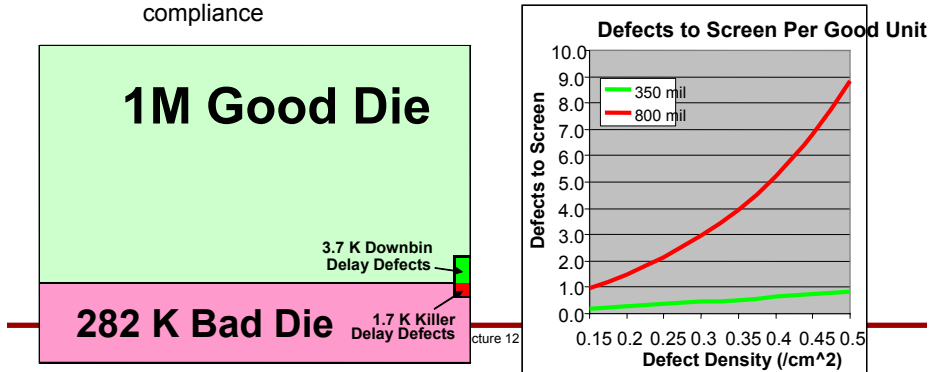
- Manufacturing Test
 - Reliable test of individual parts for volume shipment
 - Concerned with:
 - Detecting defects (yield) – identifying parts that don't work due to defects in fabrication
 - Binning parts – identifying correct “bin” for parts (e.g. frequency binning)
 - Reducing costs – test time and tester costs are key components
- Validation (Debug)
 - Verifying correct operation of the design
 - Concerned with:
 - Logical functionality – the design produces correct logical output
 - Electrical functionality – the design works at speed across entire spectrum of process, voltage, temperature, reliability spec's

Manufacturing Test

- Goal is highest possible quality at lowest possible cost
 - Quality is #1 concern
 - Unlike software, silicon cannot be “patched” (usually)
 - Design cycles are 6 months to 6 years
 - EXTREMELY expensive to recall
 - Reducing cost is still important
 - Test time directly impacts capacity and throughput
 - Tester costs are going up at an alarming rate
 - Typical automated test equipment is \$1-10M per tester
 - Multiple “sockets” directly impact both test time and tester costs

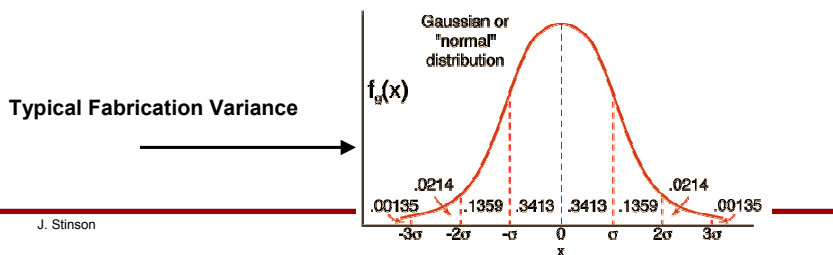
Manufacturing Test (Defects)

- All manufacturing process introduce defects into wafers
- Need to run minimum set of diagnostics on each part to ensure no defects
 - Typical CPU goal is 500-1000 DPM (defects per million)
 - Heavily relies on statistics and sampling techniques to ensure compliance



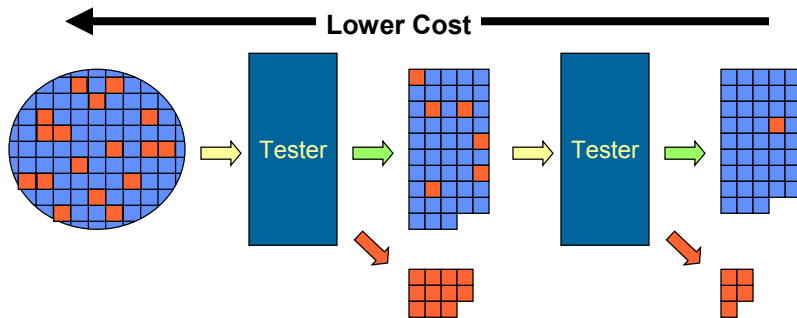
Manufacturing Test (Binning)

- Natural variation in VLSI fabrication
 - Marketing takes advantage by selling parts at multiple “bins”
 - Frequency most common example: same design sold at 1.5GHZ, 1.4GHZ and 1.3GHZ
 - Need to accurately determine to which bin each part belongs
 - Highest margin bins are usually the lowest occurrence
 - Competing Goals
 - Need adequate “guardbands” to ensure correct operation
 - Maximize highest margin bins



Manufacturing Test (Sockets)

- Process of elimination
 - Typical manufacturing test involves multiple “socketing” of parts
 - Each socket geared towards filtering defects and binning the parts
 - Lower cost the “earlier” problems can be identified

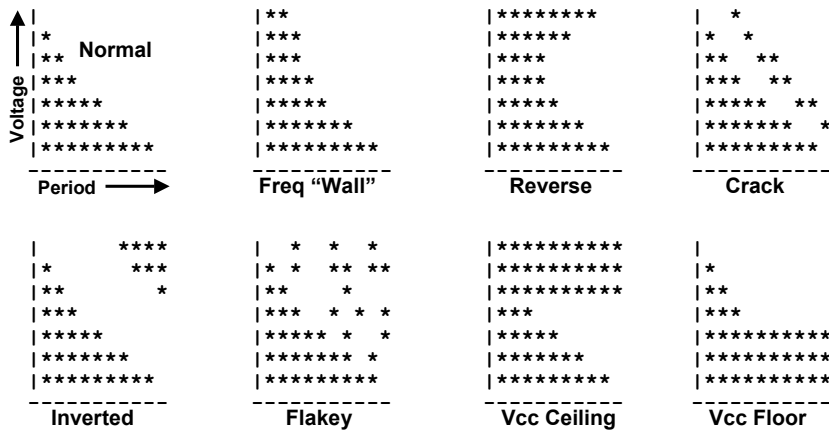


Validation (Debug)

- Goal is to ensure design meets all specifications
 - First step before going into manufacturing
 - Typical CPU debug times: 6 months to 2 years
 - Determines binning and electrical stress content for manufacturing
 - Defect content determined statistically
 - Must anticipate all usage conditions w/in spec window
 - Process Variation
 - Depending on volume, range can vary from 1-5 sigma
 - Voltage spec range
 - Power supply and di/dt variances must be accounted (2-10% nominal)
 - Temperature spec range
 - Typical from 0C to 110C junction temperature
 - “Feature” range
 - Validate across all features (diff’t core frequencies, bus fractions, thermal throttling, etc.)
 - Lifetime range
 - Typical CPU lifetime: 7-10 years

Shmoo

- Voltage vs. Frequency graph (shmoo)
 - Shape is critical in determining potential issues



BurnIn

- Need to check reliability across lifetime of design (7-10 years)
 - Obviously can't wait that long.....
 - Most wearout mechanisms are accelerated by elevated temp and voltage (burnin)
 - EM, SH, oxide wearout, PMOS degradation
- Many defect mechanisms can be accelerated with brief stress
 - "Infant Mortality" manufacturing screen to identify marginal parts
- Stress sample of parts at burnin conditions
 - Increased voltage/temperature dramatically increase power
 - Both dynamic and leakage power
 - Creates an infrastructure cost (test time, capacity, etc.)
 - Reduce frequency to help manage the power
 - Burnin functionality significantly increases design constraints
 - Design must work at extreme voltage, temp and low frequency

Test Platforms

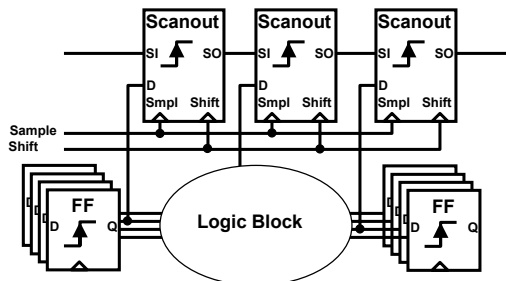
- System
 - “Real world” system (cheap)
 - Ex: Quake running on Linux on a PC
 - Easy to write code
 - Content can run for hours
 - Difficult to control electrical environment
 - Difficult to make deterministic
- Automated Test Equipment (ATE)
 - Specialized tester to “replay” waveforms at the pin level (\$\$\$\$)
 - Ex: breadboard with a logic analyzer attached to input/outputs
 - Difficult to write code (“stored response”)
 - Must simulate input stimulus to determine correct behavior
 - Simulation limits diagnostic content length
 - Excellent at controlling electrical environment
 - Should always be deterministic

Test Content

- Goals
 - Stress design to check for functionality, electrical, defects
 - “Coverage” – term used to describe quality of test content to cover specific issues or potential problems
- Focused Test
 - Diagnostics specifically written to test some aspect of the design
 - Can be functional, electrical or defect based
 - Best method of ensuring specific coverage
 - Very high cost in engineering resources
- Random Test
 - Random or pseudo-random test vectors
 - Use software to “throw the kitchen sink” at the design
 - Need method of verifying output
 - Simulation or comparison to “known good die”
 - Low engineering cost but can only achieve 60-80% coverage

Scanout

- Observability registers (non-destructive)
- Able to capture “snapshot” of state machine
 - Does not destroy machine state (great for system debug)
- Typically only done on 1-10% of state nodes
 - Adds clock loading and cap to critical paths
 - Creates it's own min/max delay issues

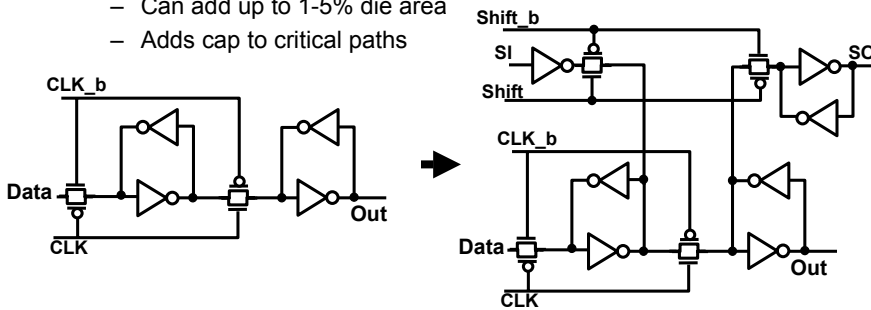


Signature Mode

- Normal scanout can only “snapshot” data as fast as the longest scanout chain
 - Must wait to shift the entire chain out before capturing new data
 - If a normal scan chain is 3000 nodes, this means only one out of every 3000 clock cycles can be “observed”
- Signature mode
 - Keep “sample” and “shift” signals asserted simultaneously
 - XOR the “Shift-In” data with the “logic” data every cycle
 - Creates a unique “signature” for the device running a particular application
 - Excellent method of adding new observability during test/validation
 - Issue: need to get this working “at speed”
 - Creates add'l max delay work

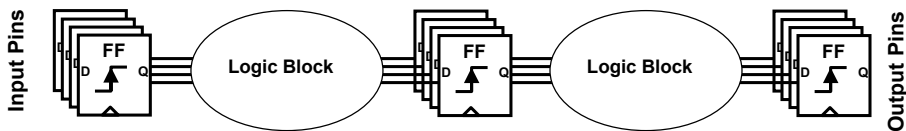
Scan

- Observability and controllability registers (destructive)
- Useful for gaining access to internal states
- Strong movement towards full scan in industry (all sequential elements are “scan-able”)
 - Difficult to control domino/arrays/clocked logic
 - Can add up to 1-5% die area
 - Adds cap to critical paths



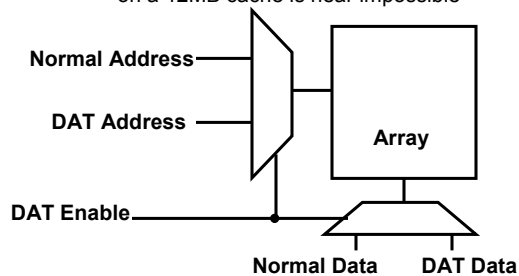
ATPG

- Automated Test Pattern Generation
 - Method of content generation with high coverage, low effort
 - Use controllability of scan to generate automated test vectors
 - Can use software to ensure that every node in the design “toggles”
- Issues
 - Accurate modeling of design captured by ATPG software
 - Full vs. partial scan
 - Capturing delay defects and events



DAT

- Direct Array Testing
 - Similar to Scan
 - Allows test manipulation of an array (or register file)
 - Uses either existing port access or adds new ports to the array
 - Powerful method of directly checking memory cells
 - Relying on functional patterns or system to “touch” every memory cell on a 12MB cache is near impossible



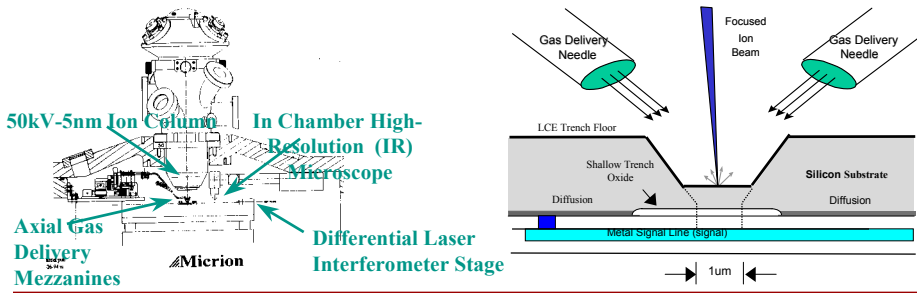
BIST

- Built-In Self Test
 - Let the chip test itself
 - Special test state machine that can control normal state machine
 - Can use either SCAN or DAT machinery to control
 - Can also be built into hardware itself (e.g. IA32 microcode)
 - Supports algorithmic test generation
 - Random “seeds” to generate internal vectors
 - Capture “signature” output and compare to expected
 - Much faster than externally manipulating internal state
 - Often only way to generate “back-to-back” cycle testing of portions of the design
 - Can be part of power-up sequence of each part
 - Good customer feature to ensure that the part is good at boot-up every time (quality)

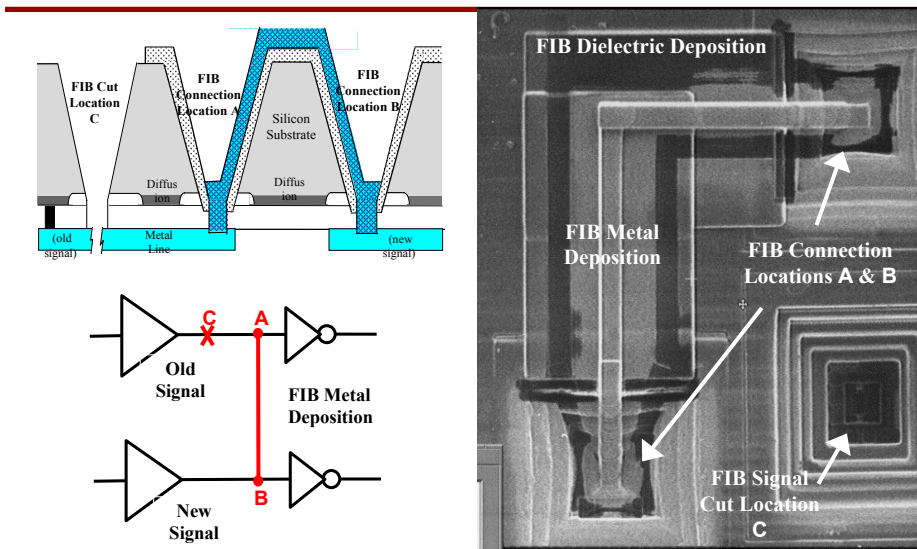
Debug Tools

Focused Ion Beam (FIB)

- Uses large particles (ions) to physically edit silicon after fabrication
 - Enables both removal and deposition
 - Can “fix” problems without having to wait for full stepping
- Limited in # of parts (5-50 hours per part)
- Limited in accessibility and scope
 - Some edits are just too complicated



Focused Ion Beam (FIB)



Device Probing

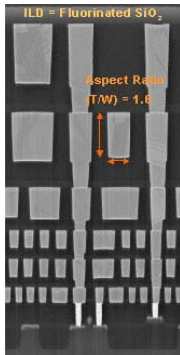
- Important to be able to collect waveforms from silicon
 - Best method of understanding “what’s going on?”
- Evolved substantially over last 15 years
 - Pico-probing – mechanical probing of metal pads on die
 - Scanning Electron Beam (SEM) – measure deflection of electrons from metal lines
 - InfraRed Emission Microscopy (IREM) – detect thermally induced IR emissions from silicon
 - Laser Voltage Probe (LVP) – measure light reflection changes in laser beam reflected off xtors
 - Time Resolved Emission (TRE)/Picosecond Imaging Circuit Analysis (PICA) – detect photonic emissions from switching xtors

Device Probe Looping

- Need deterministic loop
 - Repeat the same diagnostic over and over again
 - Must be 100% deterministic (same behavior every time)
 - Similar to oscilloscope operation
 - Most probing technologies are detecting “rare” events
 - Need MANY of these rare events to swamp out noise
 - Need excellent timing accuracy
 - Each time thru the “loop”, step forward a bit in time
 - Rely on averaging of many cycles to build a waveform
 - Short loop times necessary to prevent trigger “drift”
- Extremely difficult to probe w/in a system environment
 - Cannot easily make a system deterministic
 - Cannot easily create short loops

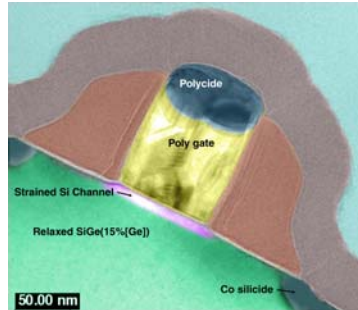
Scanning Electron Microscope (SEM)

- Measures backscatter'd electrons induced from electron beam
 - High intensity electron beam "shot" at device w/in a vacuum
 - Detector measures secondary electrons emitted by material
 - Different material will have varying levels of secondary emission
 - Used in industry for MANY years (30+)
 - Cool Big Brother: Transmission Electron Microscope (TEM)



Source: Fly's Head, Museum of Science

SEM Images



EE 371 Lecture 12

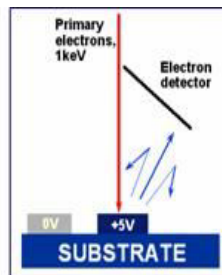
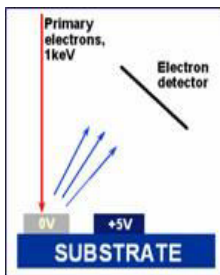
TEM Image (colorized 😊)

25

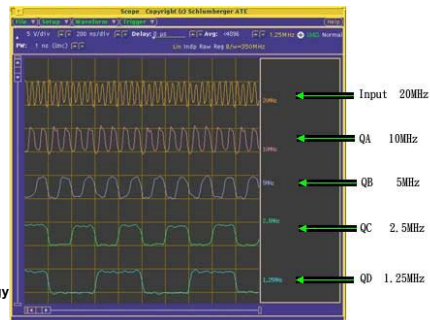
Source: 130nm technology, Intel

Ebeam Probing

- Measure secondary electron emission in time domain
 - Keep e-beam pointed at single point (invasive)
 - Pulse the beam (or pulse the detector) in time domain to build waveform
 - Technology works VERY well on metal interconnect
 - Primary probing technique with "wirebond" packaging

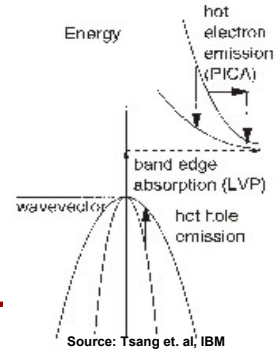


Source: Integrated Service Technology

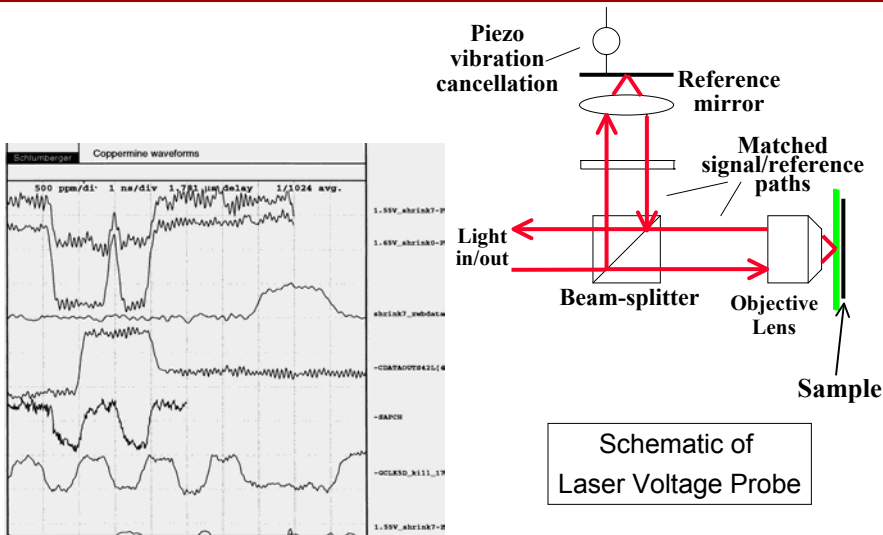


Laser Voltage Probe (LVP)

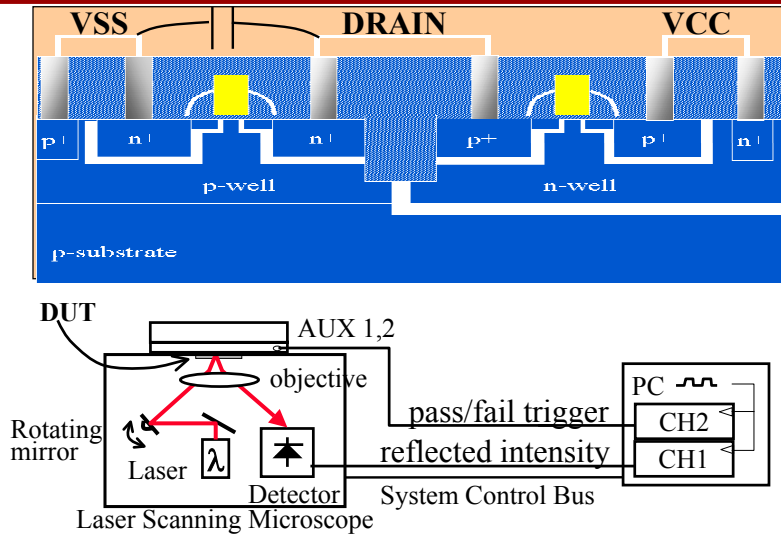
- Energy (light) absorbed by carriers in conduction band
 - Laser beam pointed at “backside” of transistors
 - Requires “flip-chip” packaging
 - Laser photon energy close to silicon band edge
 - Wavelength kept in IR or NIR band (transparent thru silicon)
 - Laser can induce carriers in conduction band
 - Need to keep intensity low enough to prevent inducing current
 - SOI tends to absorb laser light
 - Difficult to use LVP w/SOI
 - Laser must be mode-locked to test
 - Must be sync'd to test loop length



Laser Voltage Probe (LVP)

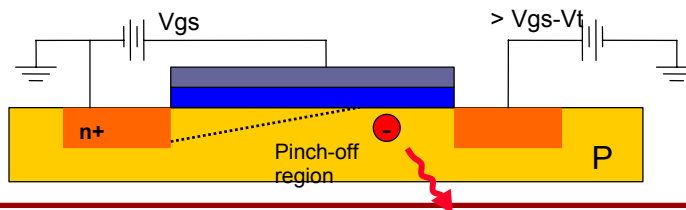


Laser Assisted Device Alteration (LADA)



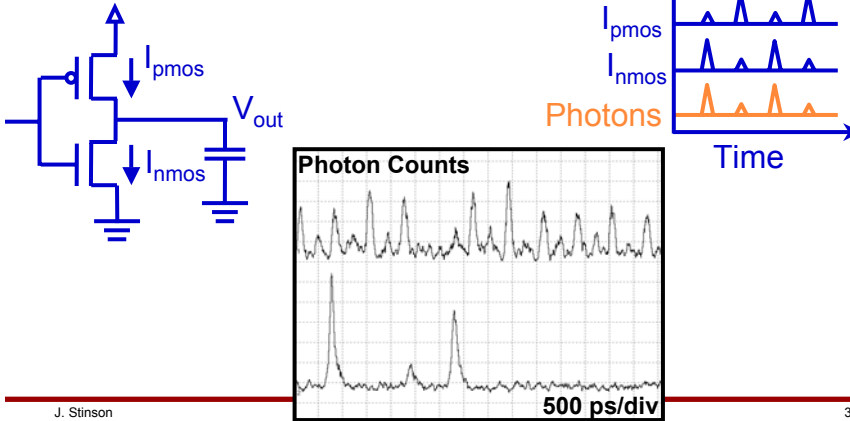
Time Resolved Emission (TRE)

- Detects photons emitted by switching xtors (similar to PICA)
 - Carriers in the channel “thermalize”, emitting NIR light
 - Silicon is transparent to IR
 - Need a REALLY good detector
 - Single photon per 10K switching events
 - Photons go in all directions; detector only at one angle
 - Need great timing resolution
 - Completely non-invasive
 - Collection times are significant
 - Longer time = better signal-to-noise ratio (SNR)



Time Resolved Emission (TRE)

- Photon emission strongest from NMOS devices
 - Falling edge vs. rising edge have diff't amplitudes
 - Linear dependence on device width
 - Exponential dependence on voltage

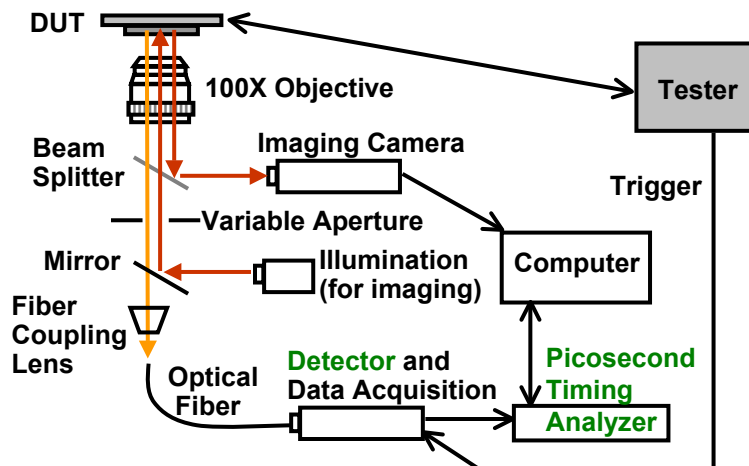


J. Stinson

Source: "Single Element Time Resolved Emission Probing for Practical Microprocessor Diagnostic Applications" E. Varner et. al, ISTFA 2002

31

TRE Basic Schematic



J. Stinson

EE 371 Lecture 12

32