Logistics

- E-mail:
  - pgvillar@us.ibm.com
  - rodman@reshape.com
  - abk@cs.ucsd.edu
- Website: http://vlsicad.ucsd.edu/
  - (.pdf's but probably not all .ppt's)
- "Tutorial end" with Section V, ~4:30pm
- Continue with questions until ???

Section I: Introduction

Schedule

- 9:00 – 9:45 I: Introduction
  - Technology roadmap implications, baseline flat methodology, problem motivations (SI, timing closure)
- 9:45 – 10:45 II: Basic Issues
  - Hierarchy, data prep, packaging, tool selection, tapeout issues, implications of test / pad layout / verification / library / clocking choices
- 10:45 – 11:00 Coffee Break
- 11:00 – noon III: Partitioning and Floorplanning
  - Partitioning into P&R blocks, block-level floorplanning, area I/O vs. peripheral I/O, clock distribution

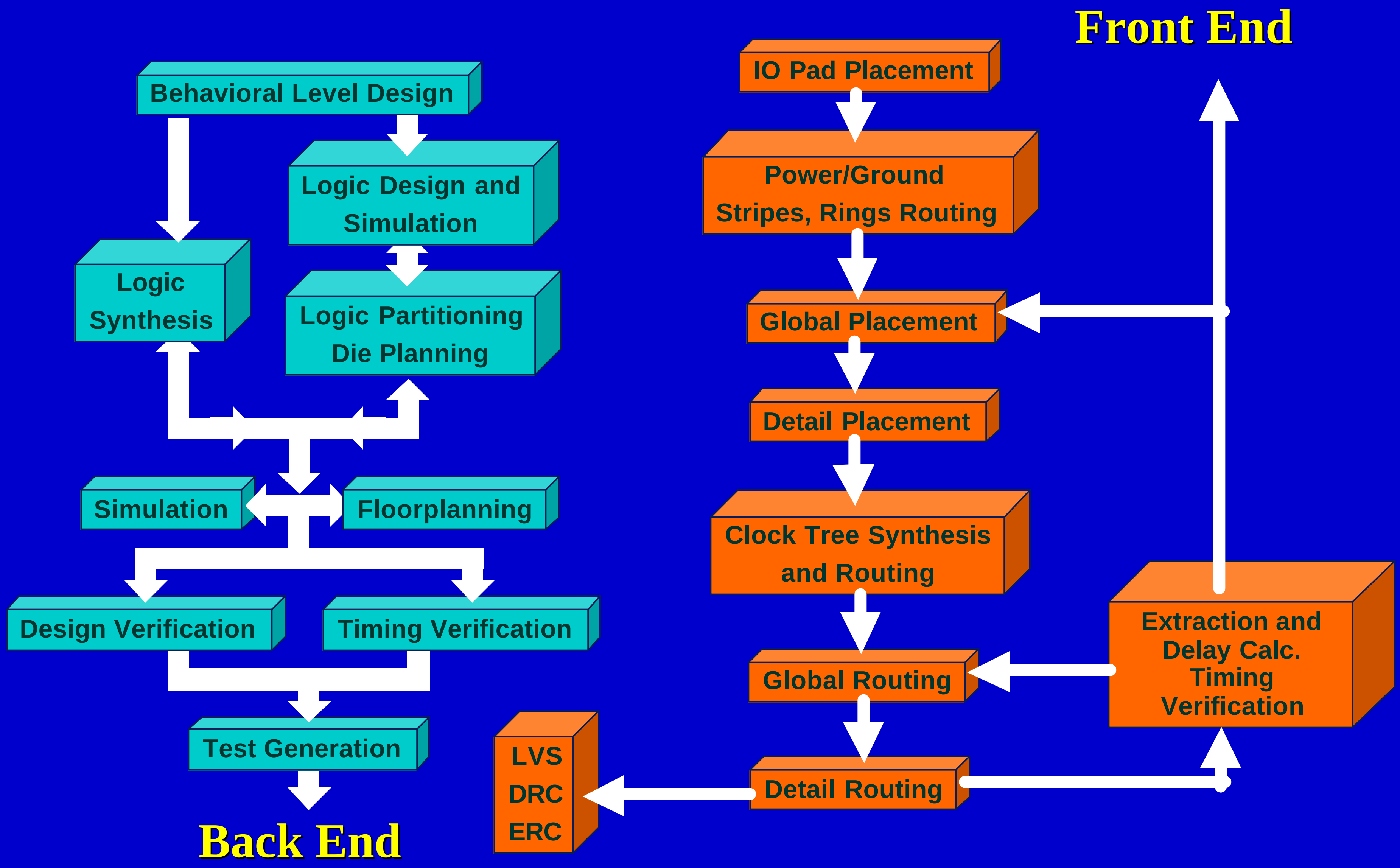
Schedule

- 1:00 – 2:30 IV: Timing Closure Techniques
  - Integrated timing/synthesis/placement/wiring for ASIC design, placement algorithms, congestion management, use of timing-driven features, timing and routability convergence
- 2:30 – 2:45 Coffee Break
- 2:45 – 4:00 V: Analysis and Verification
  - Manufacturability, inductance modeling, IR drop and ground bounce, power analysis and decoupling, signoff timing verification, special LVS/DRC issues
- 4:00 – 4:45 VI: Other Topics
  - Test, formal verification, vendor / tool gossip, your call (based on questionnaire feedback during lunch)…

Overview

- Introduction
- Technology roadmap
- Design convergence approaches

Traditional Flow



Front End

IO Pad Placement

Power/Ground Stripes, Rings Routing

Global Placement

Detail Placement

Clock Tree Synthesis and Routing

Global Routing

Detail Routing

Extraction and Delay Calc. Timing Verification

LVS
DRC
ERC

Behavioral Level Design

Logic Design and Simulation

Logic Synthesis

Logic Partitioning Die Planning

Simulation

Floorplanning

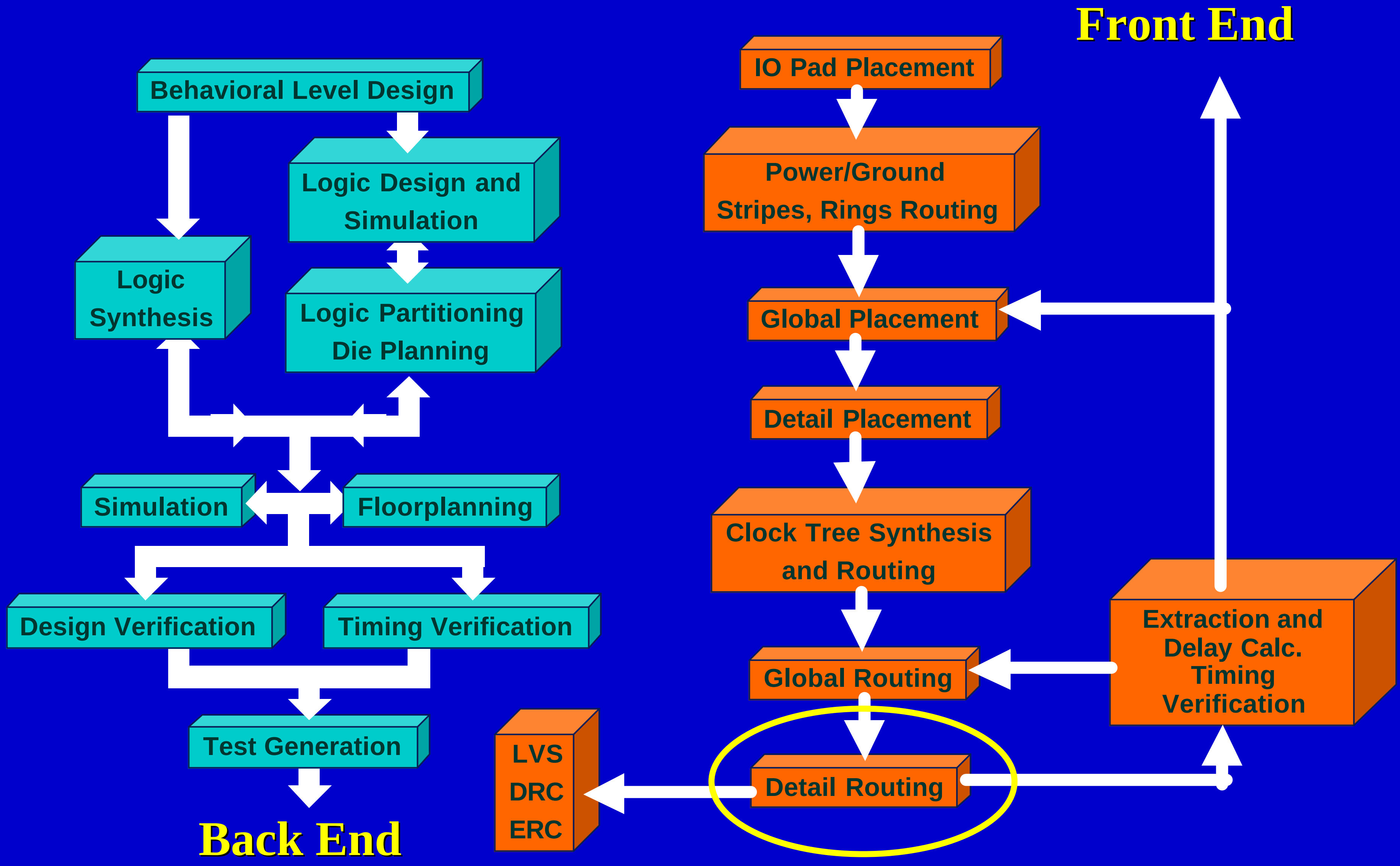
Design Verification

Timing Verification

Test Generation

Back End

It’s Tough All Over (ex: Detail Routing)



Front End

IO Pad Placement

Power/Ground Stripes, Rings Routing

Global Placement

Detail Placement

Clock Tree Synthesis and Routing

Global Routing

Detail Routing

Extraction and Delay Calc. Timing Verification

LVS DRC ERC

Behavioral Level Design

Logic Design and Simulation

Logic Synthesis

Logic Partitioning Die Planning

Simulation

Floorplanning

Design Verification

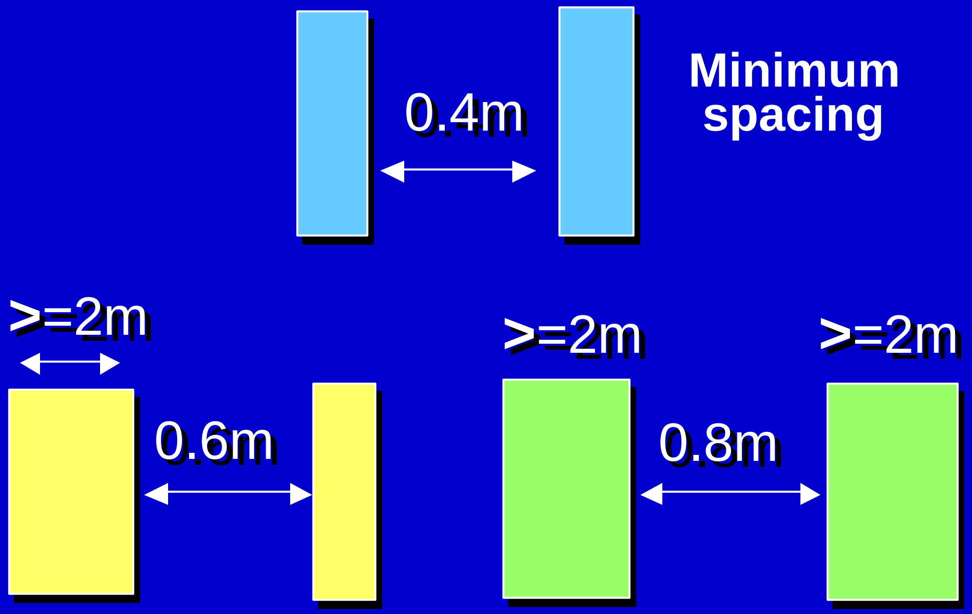
Timing Verification

Test Generation

Back End

Width / Spacing Rules

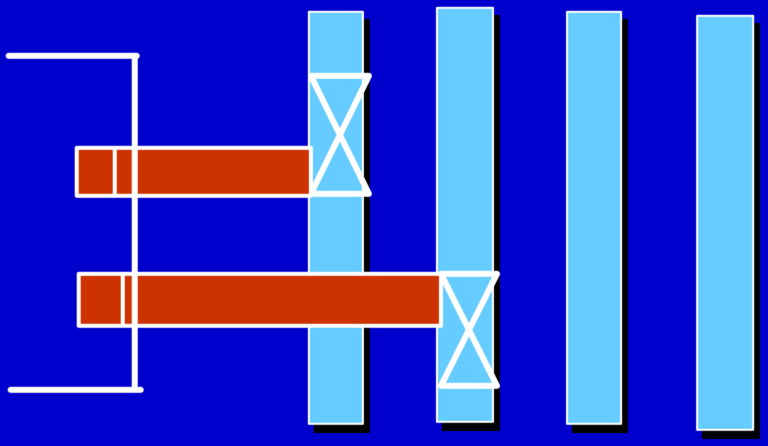
- Per connection, per net, per class, …



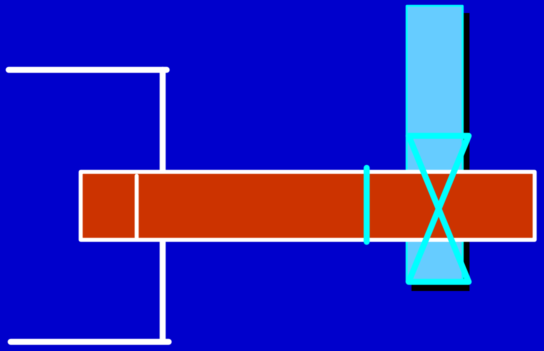
Width-based Spacing

Via Selection

- Via array based on wire size or resistance
- Rectangular via rotation and offset

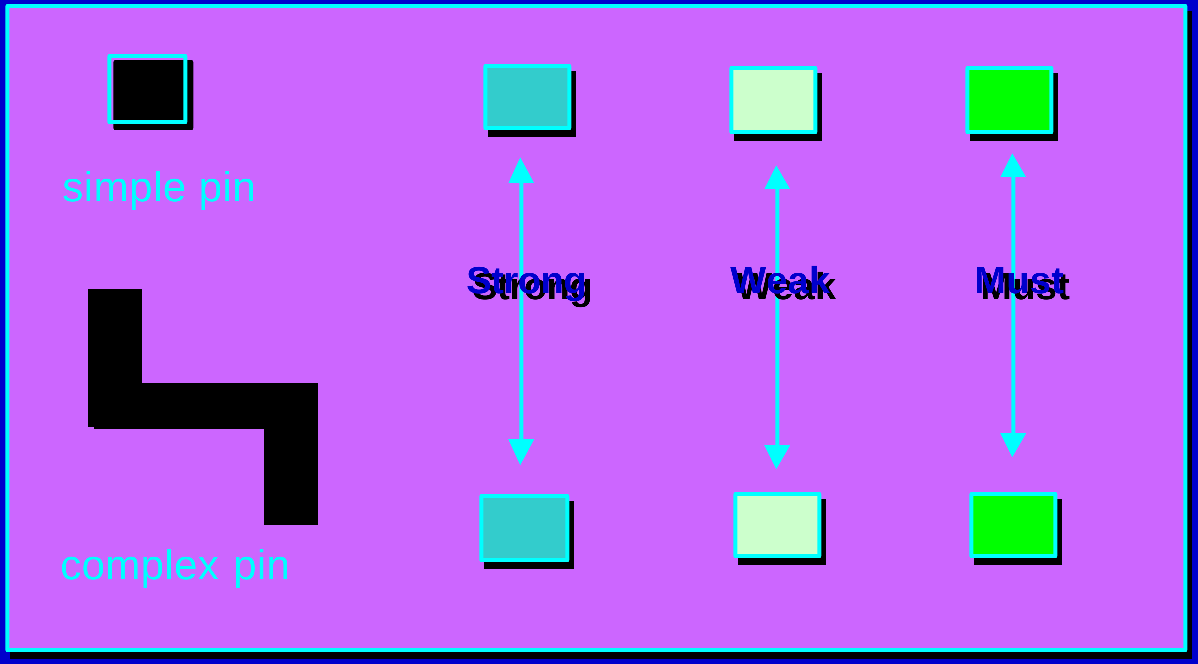


Rotate and offset horizontal vias

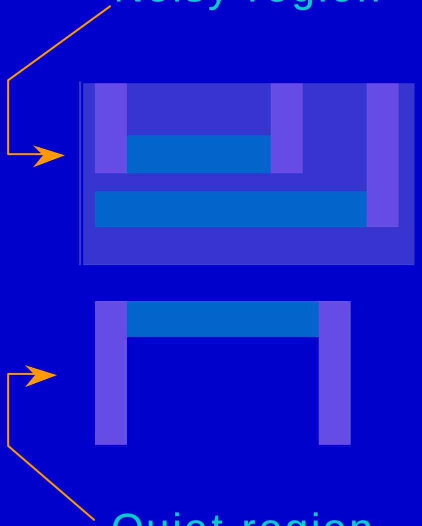


No rotation for a “cross” via

Complex Pins, Equivalent Pin Modeling



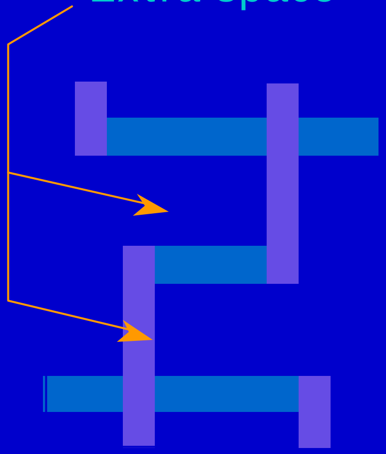
Noise-Driven



Noisy region

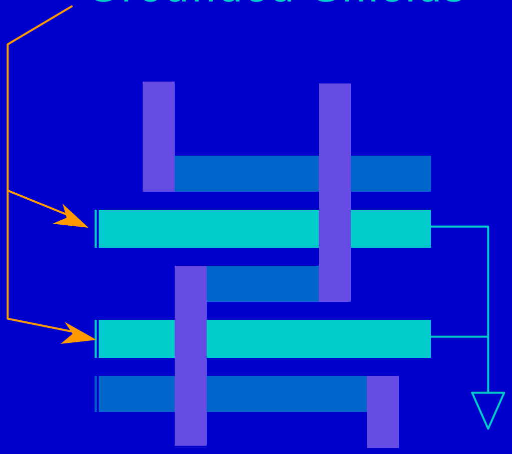
Quiet region

Segregation



Extra space

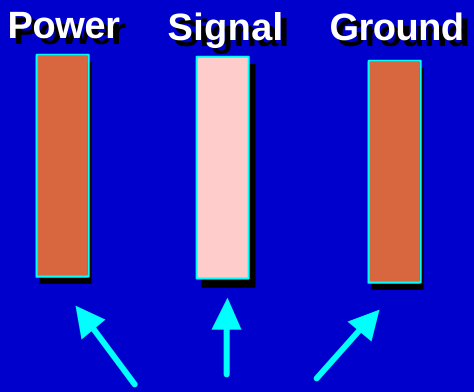
Spacing



Grounded Shields

Shielding

Same-Layer, Adj-Layer Shielding



Power

Signal

Ground

Same-Layer Shielding

M2

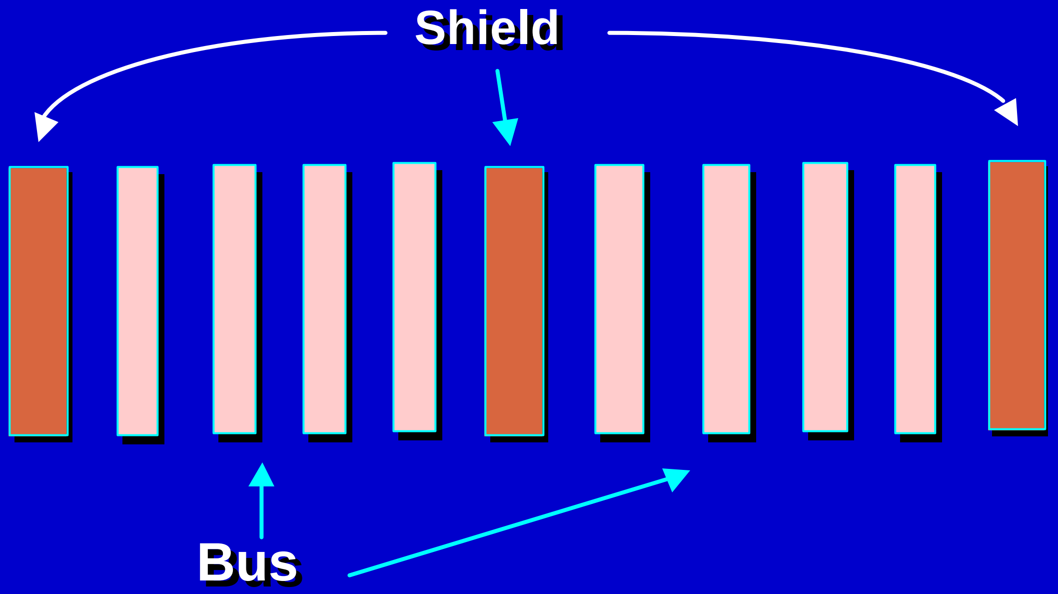
Signal

M1

Poly

Adjacent-Layer Shielding

Bus Shielding and Interleaving

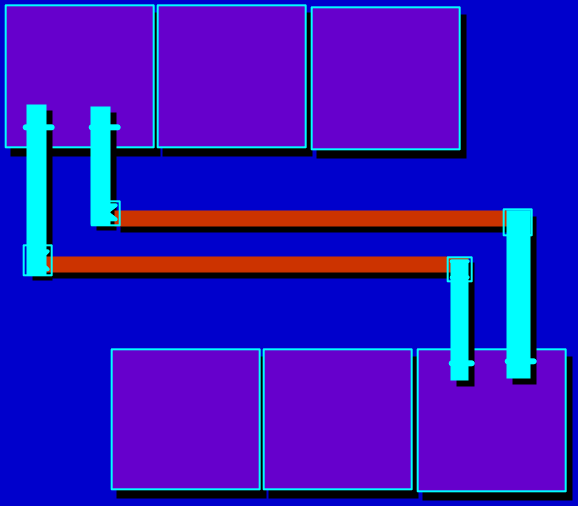


Shield

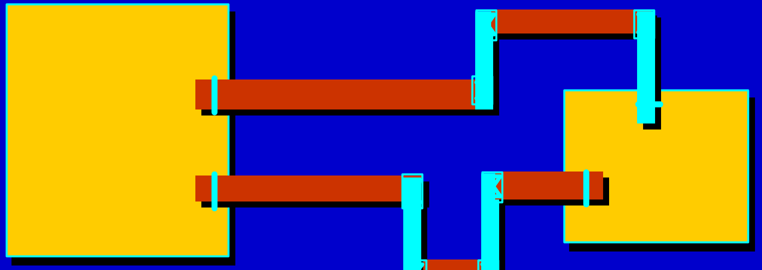
Bus

Bus Shielding

Differential-Pair, Balanced (length, cap) Routes

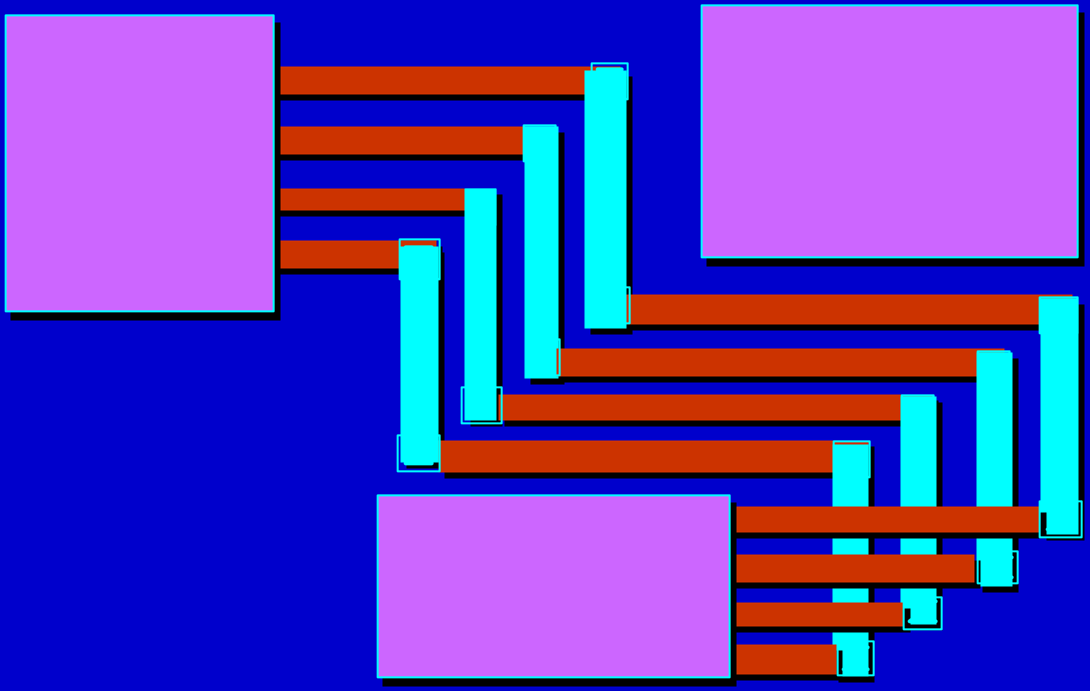


Differential

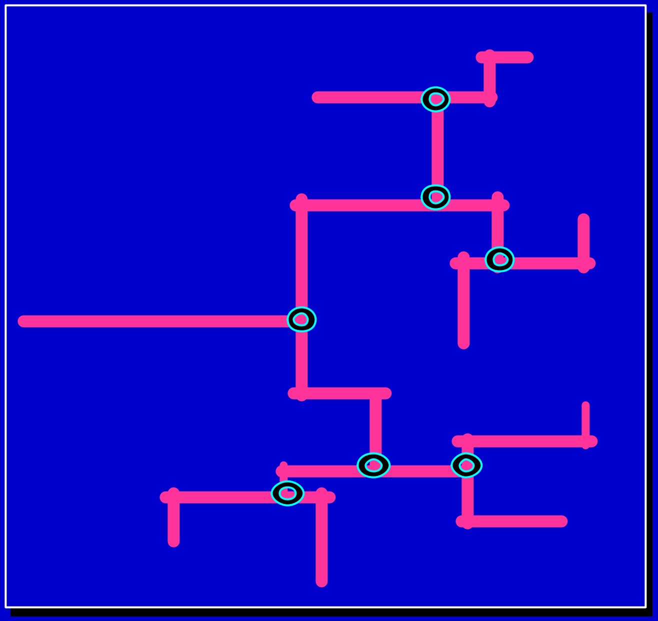


Balanced length

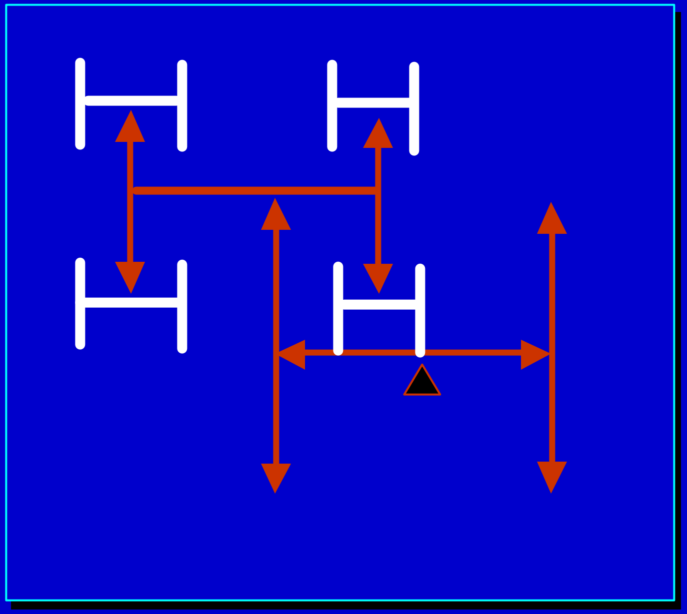
Bus Routers



Clock Routers



Balanced Tree

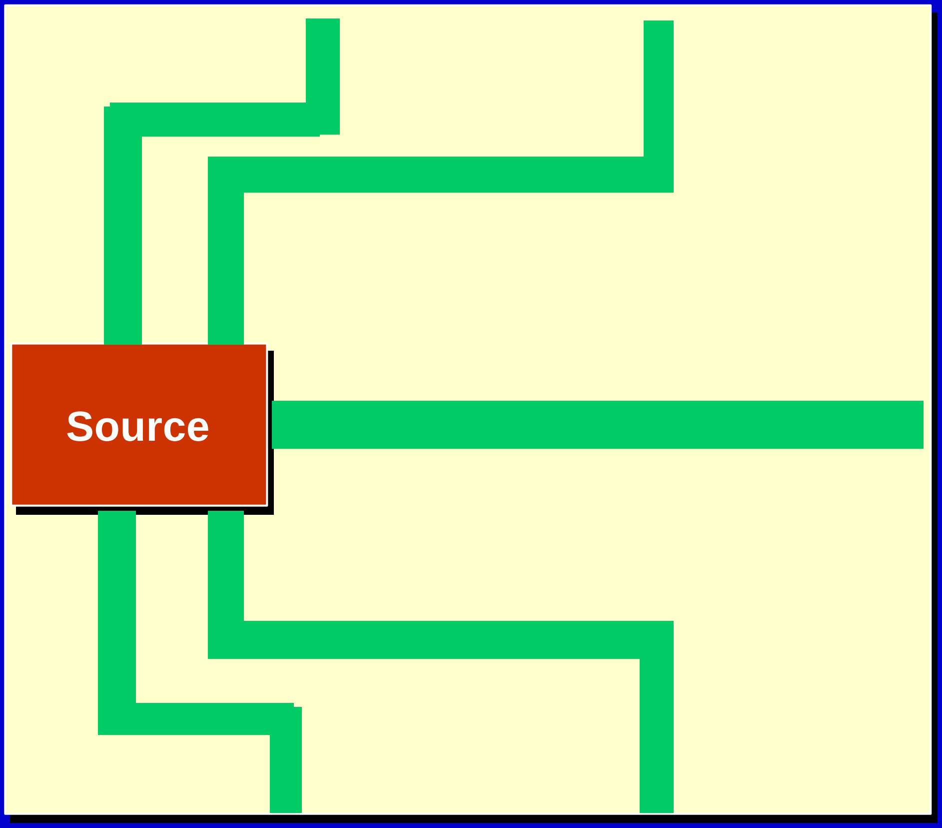


H-Tree

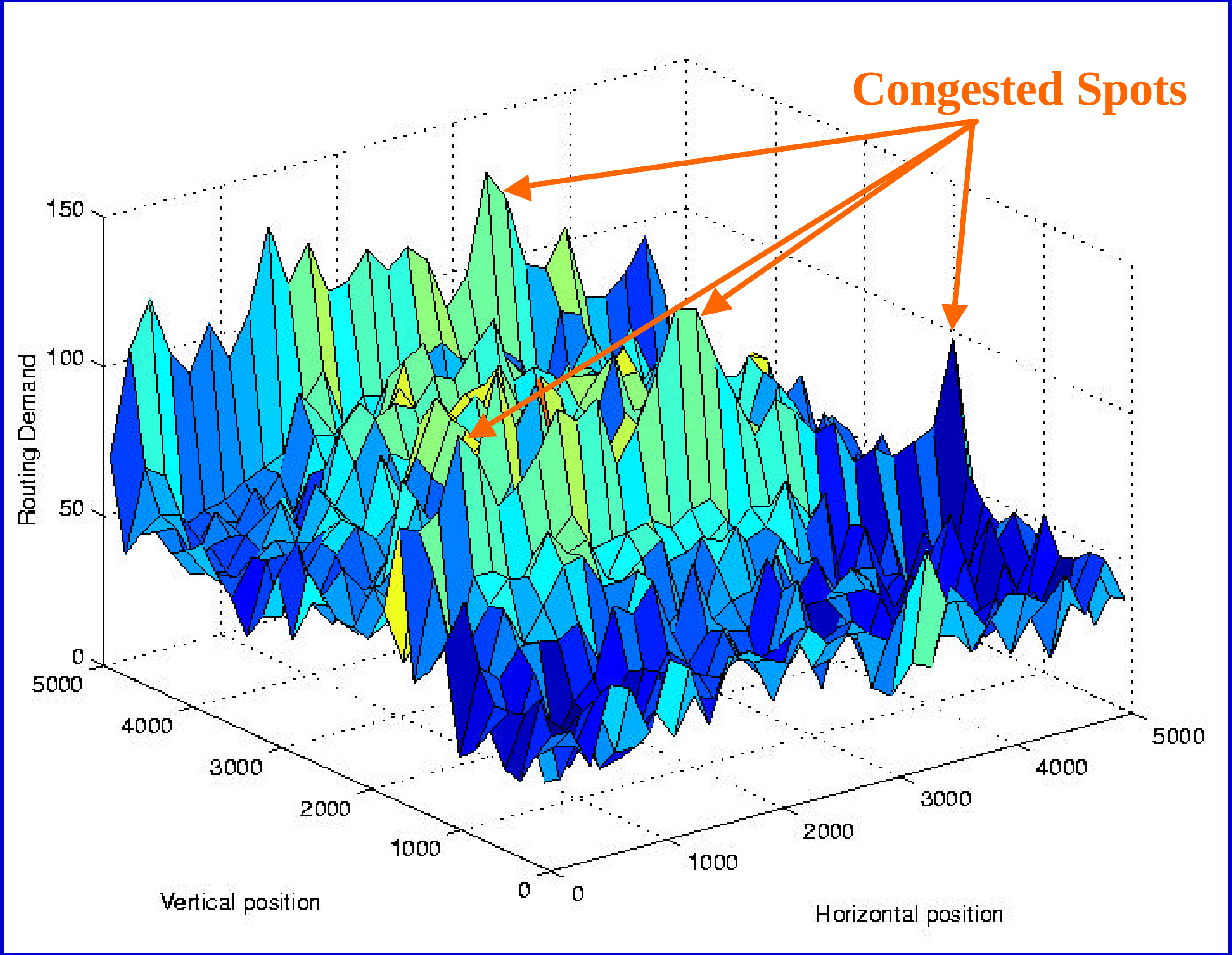
Power Routers

- Power Mesh
- Power Ring
- Star Routing
- Grounded Fill

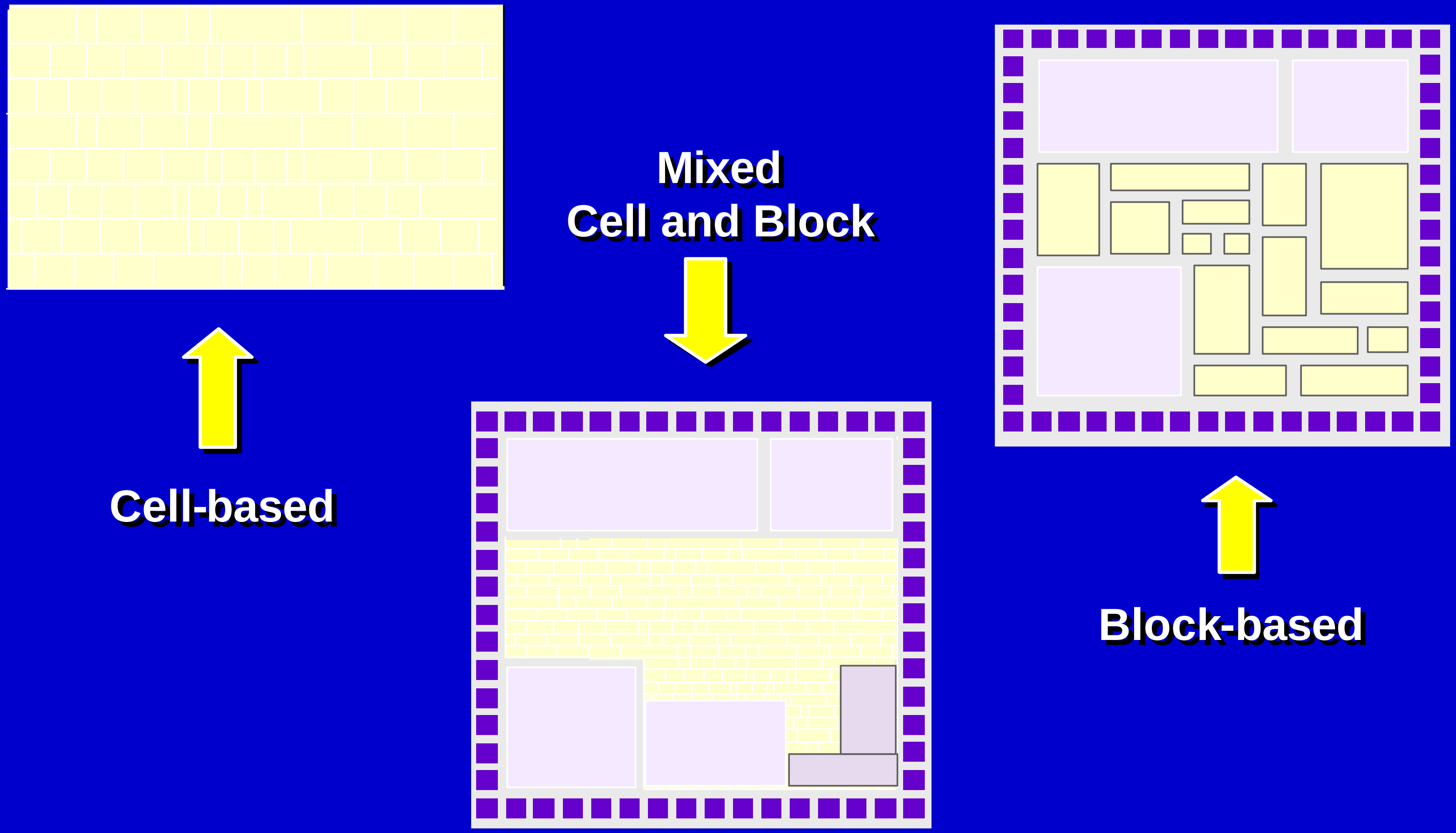
Star Routing



Congestion


Congested Spots
Routing Demand
150
100
50
0
Vertical position
5000
4000
3000
2000
1000
0
Horizontal position
0
1000
2000
3000
4000
5000

Different Applications

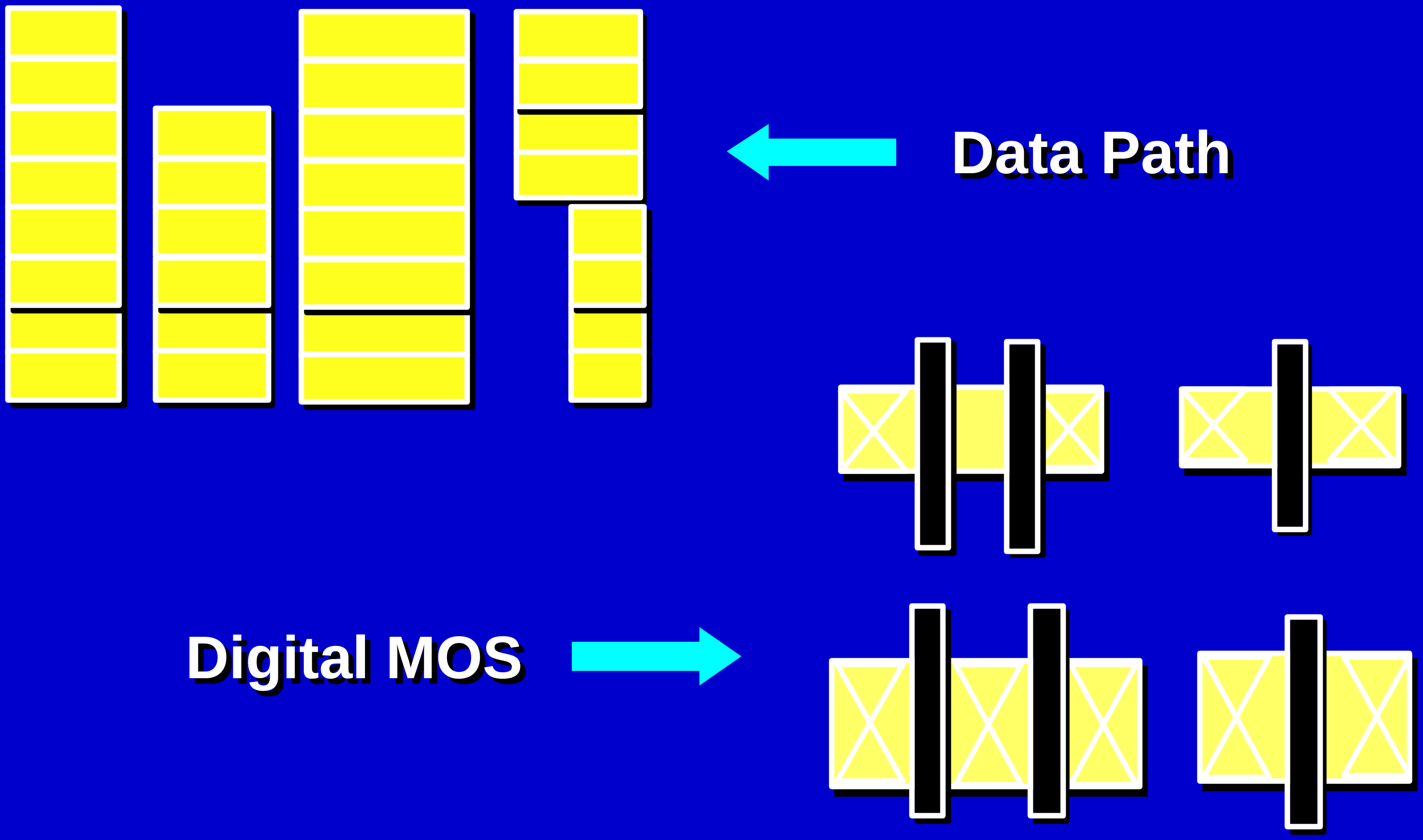


Mixed
Cell and Block

Cell-based

Block-based

Different Applications



Data Path

Digital MOS

And More…

- Other rules
  - Process antenna, phase shift mask, OPC rules
- Frameworks
  - Channel routing
  - Switch box routing
  - Maze routing
  - Line probe routing
  - Shape-based routing
  - Fixed die vs. variable die
  - Gridded vs. gridless
- This tutorial: Navigating seas of uncertainties, changes, choices (which define methodologies)…

Technology Roadmap

Roadmap Changes Since 2000

- Next “node” = 0.7x half-pitch or minimum feature size
  - → 2x transistors on the same size die
- 90nm node in 2004 (100nm in 2003)
  - 90nm node → physical gate length = 45nm
- MPU/ASIC half-pitch = DRAM half-pitch in 2004
  - Previous ITRS (2000): convergence in 2015
- Psychology: everyone must beat the Roadmap
  - Reasons: density, cost reduction, competitive position
  - TSMC CL010G logic/mixed-signal SOC process: risk production in 4Q02 with multi-Vt, multi-oxide, embedded DRAM and flash, low standby power derivatives, …

Red Brick Wall - 2001 ITRS vs 1999

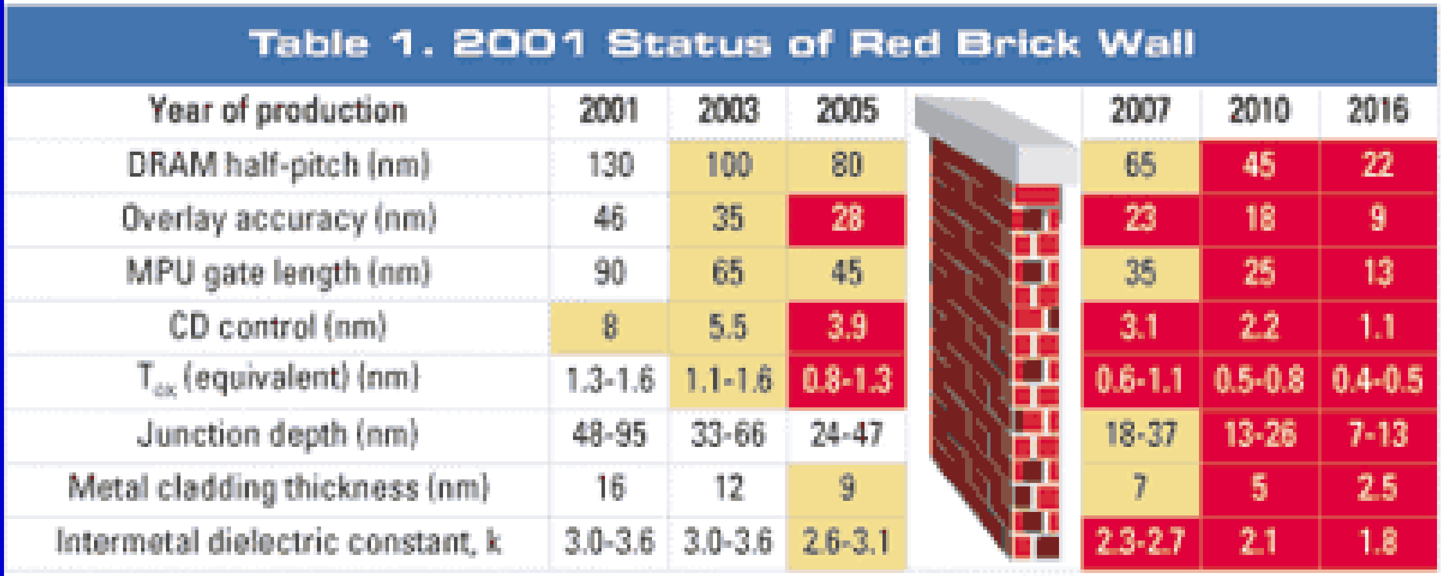


Table 1. 2001 Status of Red Brick Wall

Year of production	2001	2003	2005	2007	2010	2016
DRAM half-pitch (nm)	130	100	80	65	45	22
Overlay accuracy (nm)	46	35	28	23	18	9
MPU gate length (nm)	90	65	45	35	25	13
CD control (nm)	8	5.5	3.9	3.1	2.2	1.1
T_{ox} (equivalent) (nm)	1.3-1.6	1.1-1.6	0.8-1.3	0.6-1.1	0.5-0.8	0.4-0.5
Junction depth (nm)	48-95	33-66	24-47	18-37	13-26	7-13
Metal cladding thickness (nm)	16	12	9	7	5	2.5
Intermetal dielectric constant, k	3.0-3.6	3.0-3.6	2.6-3.1	2.3-2.7	2.1	1.8

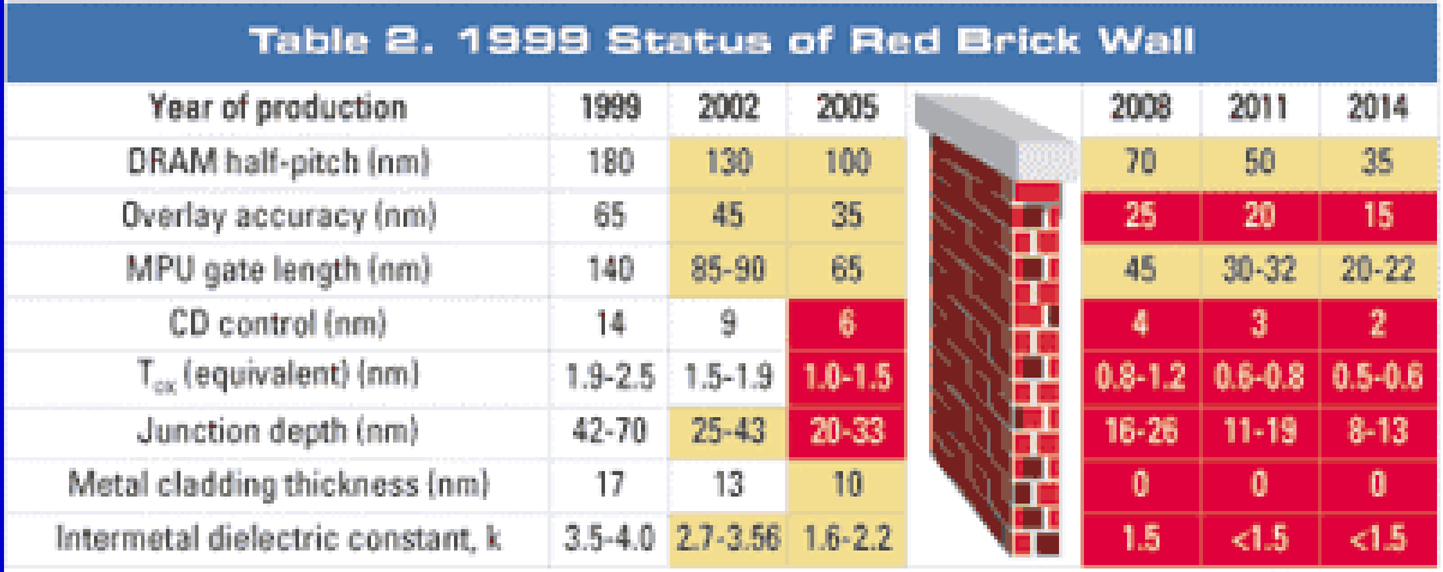


Table 2. 1999 Status of Red Brick Wall

Year of production	1999	2002	2005	2008	2011	2014
DRAM half-pitch (nm)	180	130	100	70	50	35
Overlay accuracy (nm)	65	45	35	25	20	15
MPU gate length (nm)	140	85-90	65	45	30-32	20-22
CD control (nm)	14	9	6	4	3	2
T_{ox} (equivalent) (nm)	1.9-2.5	1.5-1.9	1.0-1.5	0.8-1.2	0.6-0.8	0.5-0.6
Junction depth (nm)	42-70	25-43	20-33	16-26	11-19	8-13
Metal cladding thickness (nm)	17	13	10	0	0	0
Intermetal dielectric constant, k	3.5-4.0	2.7-3.56	1.6-2.2	1.5	<1.5	<1.5

Source: Semiconductor International - http://www.e-insite.net/semiconductor/index.asp?layout=article&articleId=CA187876

Roadmap Acceleration and Deceleration

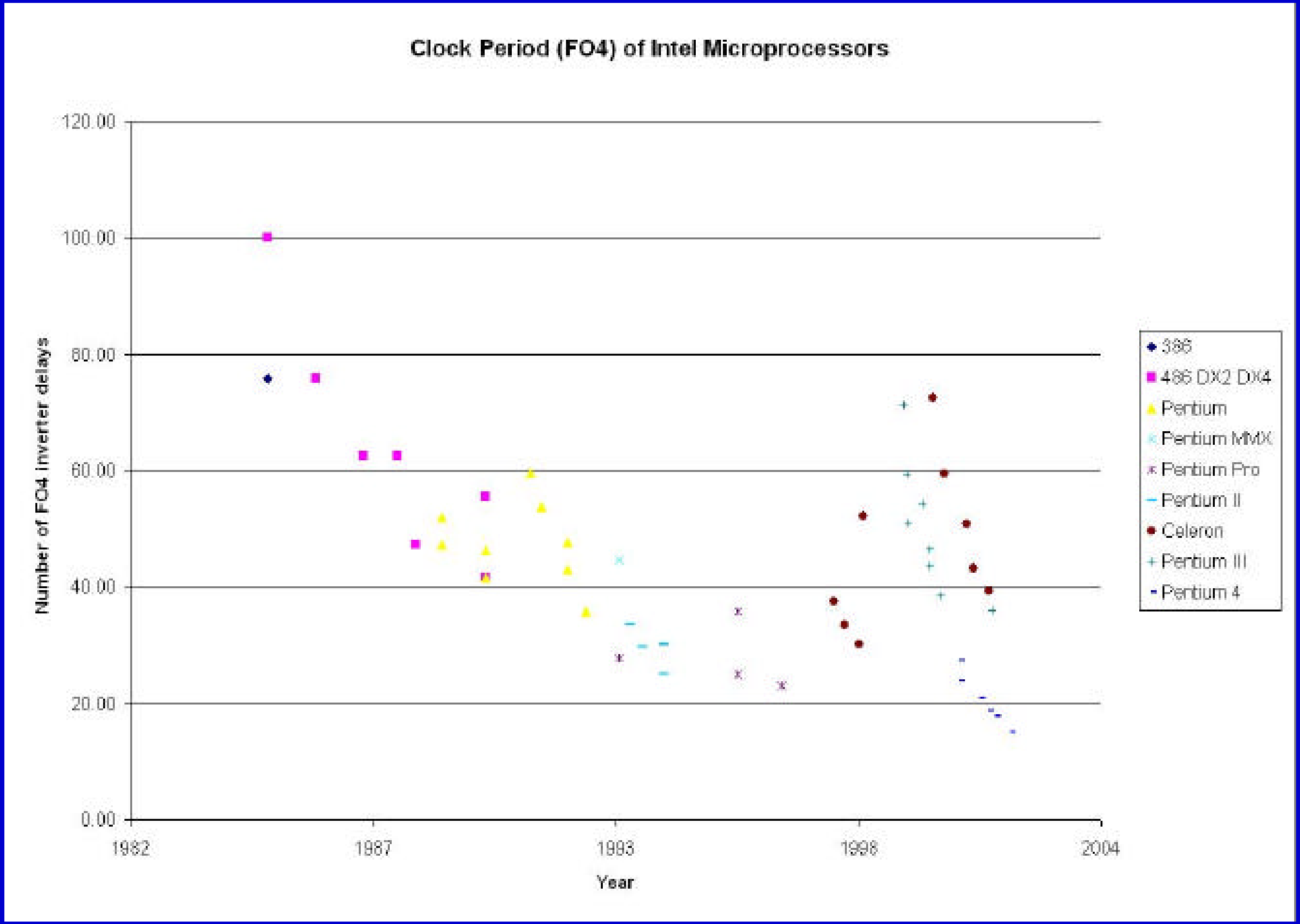
2001 versus 1999 Results

Year of Production:	1999	2002	2005	2008	2011	2014
DRAM Half-Pitch [nm]:	180	130	100	70	50	35
Overlay Accuracy [nm]:	65	45	35	25	20	15
MPU Gate Length [nm]:	140	85-90	65	45	30-32	20-22
CD Control [nm]:	14	9	6	4	3	2
T_{OX} (equivalent) [nm]:	1.9-2.5	1.5-1.9	1.0-1.5	0.8-1.2	0.6-0.8	0.5-0.6
Junction Depth [nm]:	42-70	25-43	20-33	16-26	11-19	8-13
Metal Cladding [nm]:	17	13	10			0.0
Inter-Metal Dielectric K:	3.5-4.0		2.7-3.5		1.6-2.2	1.5

HP / LOP / LSTP Device Roadmaps

Parameter	Type	**99**	**01**	**03**	**05**	**07**	**10**	**13**	**16**
Vdd	MPU	1.5	1.2	1.0	0.9	0.7	0.6	0.5	**0.4**
	LOP	1.3	1.2	1.1	1.0	0.9	0.8	0.7	**0.6**
	LSTP	1.3	1.2	1.2	1.2	1.1	1.0	0.9	**0.9**
Vth (V)	MPU	0.21	0.19	0.13	0.09	0.05	0.021	0.003	**0.003**
	LOP	0.34	0.34	0.36	0.33	0.29	0.29	0.25	**0.22**
	LSTP	0.51	0.51	0.53	0.54	0.52	0.49	0.45	**0.45**
Ion (uA/um)	MPU	1041	926	967	924	1091	1250	1492	**1507**
	LOP	636	600	600	600	700	700	800	**900**
	LSTP	300	300	400	400	500	500	600	**800**
CV/I (ps)	MPU	2.00	1.63	1.16	0.86	0.66	0.39	0.23	**0.16**
	LOP	3.50	2.55	2.02	1.58	1.14	0.85	0.56	**0.35**
	LSTP	4.21	4.61	2.96	2.51	1.81	1.43	0.91	**0.57**
Ioff (uA/um)	MPU	0.00	0.01	0.07	0.30	1.00	3	7	**10**
	LOP	1e-4	1e-4	1e-4	3e-4	7e-4	1e-3	3e-3	**1e-2**
	LSTP	1e-6	1e-6	1e-6	1e-6	1e-6	3e-6	7e-6	**1e-5**

FO4 INV Delays Per Clock Period



- FO4 INV = inverter driving 4 identical inverters (no interconnect)
- Half of freq improvement has been from reduced logic stages

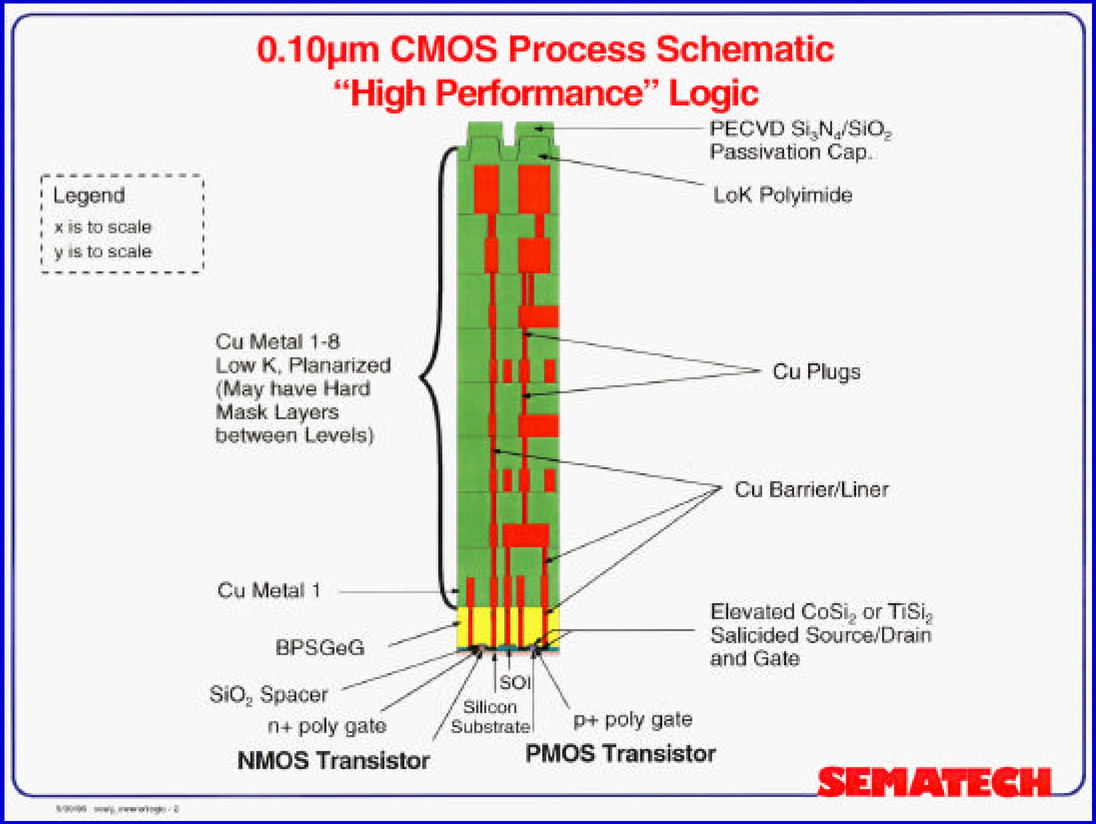
Silicon Complexity Challenges

- Silicon Complexity = impact of process scaling, new materials, new device/interconnect architectures
- Non-ideal scaling (leakage, power management, circuit/device innovation, current delivery)
- Coupled high-frequency devices and interconnects (signal integrity analysis and management)
- Manufacturing variability (library characterization, analog and digital circuit performance, error-tolerant design, layout reusability, static performance verification methodology/tools)
- Scaling of global interconnect performance (communication, synchronization)
- Decreased reliability (SEU, gate insulator tunneling and breakdown, joule heating and electromigration)
- Complexity of manufacturing handoff (reticle enhancement and mask writing/inspection flow, manufacturing NRE cost)

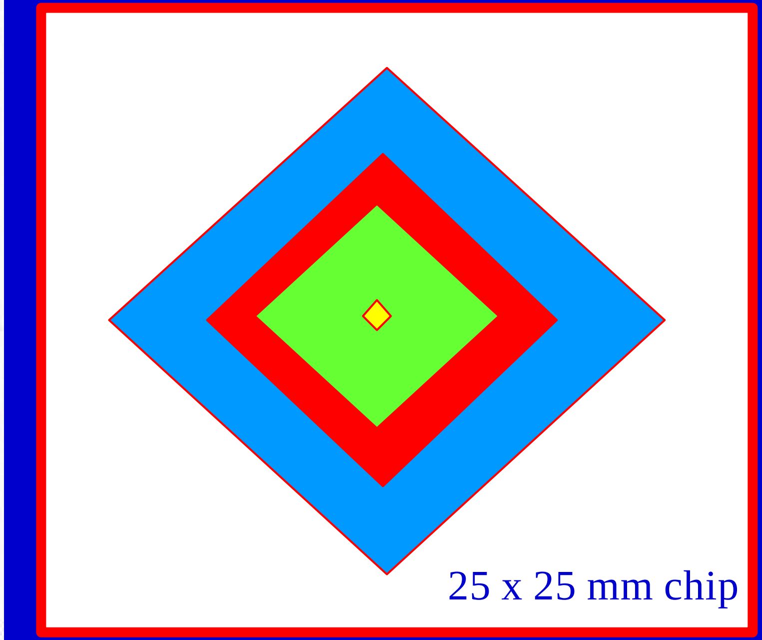
System Complexity Challenges

- System Complexity = exponentially increasing transistor counts, with increased diversity (mixed-signal SOC, …)
- Reuse (hierarchical design support, heterogeneous SOC integration, reuse of verification/test/IP)
- Verification and test (specification capture, design for verifiability, verification reuse, system-level and software verification, AMS self-test, noise-delay fault tests, test reuse)
- Cost-driven design optimization (manufacturing cost modeling and analysis, quality metrics, die-package co-optimization, …)
- Embedded software design (platform-based system design methodologies, software verification/analysis, codesign w/HW)
- Reliable implementation platforms (predictable chip implementation onto multiple fabrics, higher-level handoff)
- Design process management (team size / geog distribution, data mgmt, collaborative design, process improvement)

Reachability



Reachability in t_{crit} = 80 ps



Summary of Technology Scaling

- Scaling of 0.7x every three (two?) years

.25u	.18u	.13u	.10u	.07u	.05u
1997	1999	2002	2005	2008	2011
5LM	6LM	7LM	7LM	8LM	9LM

- Interconnect delay dominates system performance
  - consumes 70% of clock cycle
- Cross coupling capacitance is dominating
  - cross capacitance → 100%, ground capacitance → 0%
  - 90% in .18u
  - huge signal integrity implications (e.g., guardbands in static analysis approaches)
- Multiple clock cycles required to cross chip
  - whether 3 or 15 not as important as fact of “multiple” > 1

New Materials Implications

- Lower dielectric permittivity
  - reduces total capacitance
  - doesn’t change cross-coupled / grounded capacitance proportions
- Copper metallization
  - reduces RC delay
  - avoids electromigration (factor of 4-5 ?)
  - thinner deposition reduces cross cap
- Multiple layers of routing
  - enabled by planarization; 10% extra cost per layer
  - reverse-scaled top-level interconnects
  - relative routing pitch may increase
  - room for shielding

Technical Issues

- Manufacturability (chip can't be built)
  - antenna rules
  - minimum area rules for stacked vias
  - CMP (chemical mechanical polishing) area fill rules
  - layout corrections for optical proximity effects in subwavelength lithography; associated verification issues
- Signal integrity (failure to meet timing targets)
  - crosstalk induced errors
  - timing dependence on crosstalk
  - IR drop on power supplies
- Reliability (design failures in the field)
  - electromigration on power supplies
  - hot electron effects on devices
  - wire self heat effects on clocks and signals

Noise

- Analog design concerns are due to physical noise sources
  - because of discreteness of electronic charge and stochastic nature of electronic transport processes
  - example: thermal noise, flicker noise, shot noise
- Digital circuits due to large, abrupt voltage swings, create deterministic noise which is several orders of magnitude higher than stochastic physical noise
  - still digital circuits are prevalent because they are inherently immune to noise
- Technology scaling and performance demands make noisiness of digital circuits a big problem

Courtesy Hormoz/Muddu, ASIC99

Design Convergence Approaches

Goal: Design Convergence

- What must converge ?
  - logic, timing, and spatial embedding
  - support front-end signoff with a <u>predictable</u> back-end
- Achieve Convergence through <u>Predictability</u>
  - correct by construction ("assume, then enforce")
    - constraints and assumptions passed downstream; not much goes upstream
    - ignores concerns via guardbanding
    - separates concerns as able (e.g., FE logic/timing vs. BE spatial embedding)
  - construct by correction ("tight loops")
    - logic-layout unification; synthesis-analysis unification, concurrent optimization
  - elimination of concerns
    - reduced degrees of freedom, pre-emptive design techniques
    - e.g., power distribution, layer assignment / repeater rules, GALS/LIS

“Design Convergence Tool”

- “Silicon Virtual Prototype”, “IC Implementation Suite”
- Input
  - RT-level HDL + technology + constraints
- Output
  - “go”: recipe for invoking SP&R, composing results
  - “no go”: diagnosis of RTL code problems
- Logical and physical hierarchies co-evolve
  - spatial: top-down coarse placement → physical hierarchy
  - logic/timing: implementable RTL → logical hierarchy
  - Evolution: no-FP, phys-FP, RTL-FP, …
- Big details (construct, predict, ignore, eliminate, ...)
  - pin optimizations, interconnect planning, hierarchy reconciliations, budgeting mechanisms, compatibility with downstream SP&R, ...

Planning Technology Elements

- RTL partitioning
  - understand interaction b/w block definition and placement quality
  - recognize and cure a physically challenged logic hierarchy
- Global interconnect planning and optimization
  - symbolic route representations to support block plan ECOs
- Controllable SP&R back end (including power/clock/scan)
- Incremental / ECO optimizations, and optimizations that are "robust" under partial or imperfect design knowledge
- Better estimators ("initial WLMs")
  - to account for resource, topological heterogeneity
  - to account for optimizations (placement, ripup/reroute, timing)
- → "earliest RTL signoff with detailed P&R knowledge"

Taxonomy of Traditional Planning / Implementation Methodologies

- Centered on logic design
  - wire-planning methodology with block/cell global placement
  - global routing directives passed forward to chip finishing
  - constant-delay methodology may be used to guide sizing
- Centered on physical design
  - placement-driven or placement-knowledgeable logic synthesis
- Buffer between logic and layout synthesis
  - placement, timing, sizing optimization tools
- Centered on SOC, chip-level planning
  - interface synthesis between blocks
  - communications protocol, protocol implementation decisions guide logic and physical implementation

Issue: Perf Optimization DOF’s

- Design Optimization
  - global restructuring optimization -- logic optimization on layout using actual RC, noise peak values etc.
  - localized optimization -- with no structural changes and least layout impact
  - repeater/buffer insertion for global wires
- Physical optimization
  - high fanout net synthesis (eg. for clock nets); buffer trees to meet delay/skew and fanout requirements
  - automatically determine network topology (# levels, #buffers, and type of buffers)
  - wire sizing, spacing, shielding etc.
- Fixing timing violations automatically
  - fix setup/hold time violations
  - fix maximum slew and fanout violations

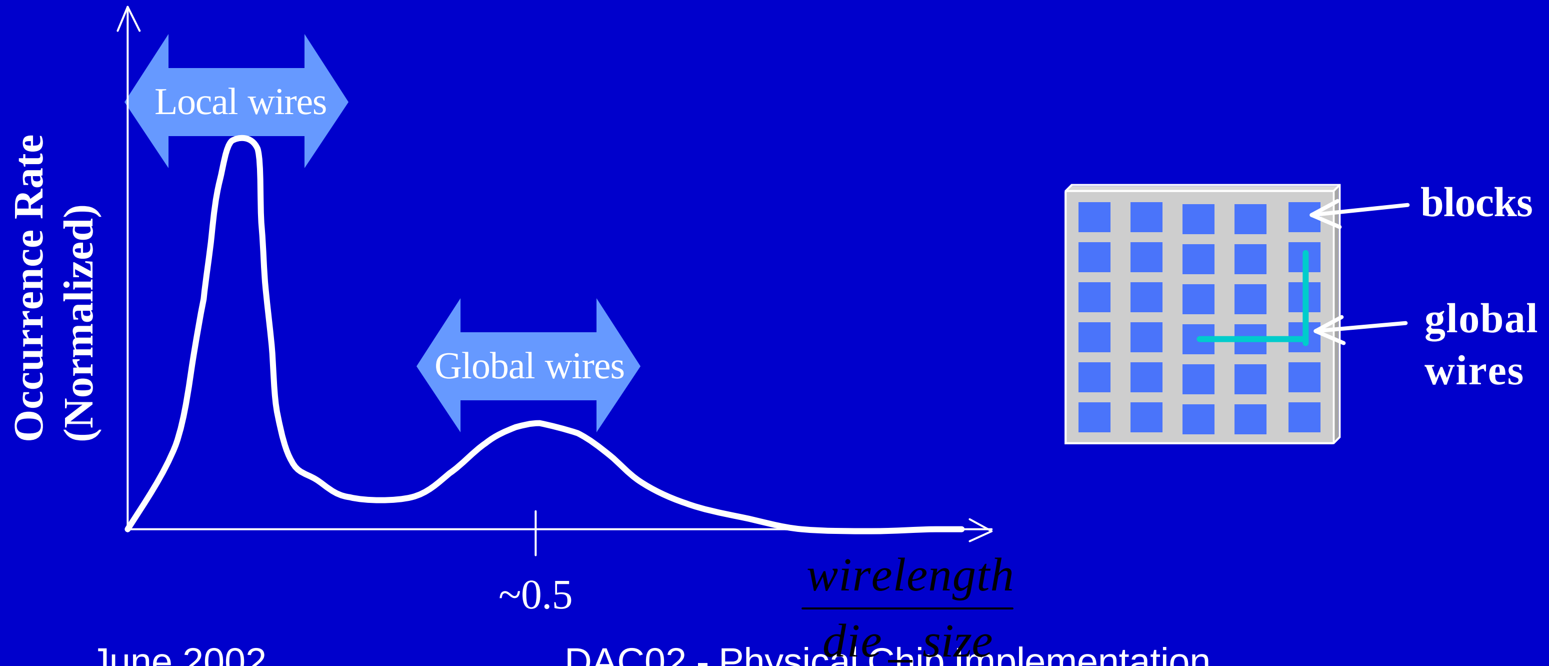
Courtesy Hormoz/Muddu, ASIC99

Issue: Hierarchy

- Two hierarchies: logical/functional, and physical
  - schematic hierarchy also typical in structured-custom
- RTL design = logical/functional hierarchy
  - provides valuable clues for physical embedding: datapath structure, timing structure, etc.
  - can be incredibly misleading (e.g., all clock buffers in a single hierarchy block)
- Main issues:
  - how to leverage logical/functional hierarchy during embedding
  - when to deviate from designer’s hierarchy
  - methodology for hierarchy reconciliation (buffers, repartitioning / reclustering, etc.)

Issue: Interconnect Complexity

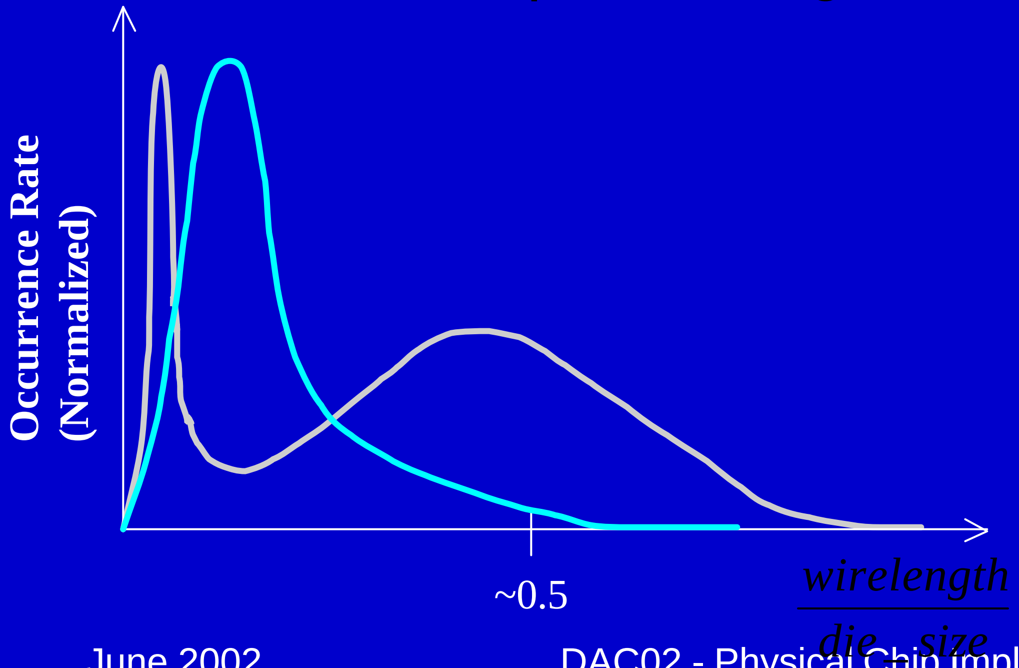
- Interconnect effects play a major role in the increasing costs for large hard-block or rectilinear-outline based design styles
- Probabilistic wireload models fail
- Need “soft block” design and assembly

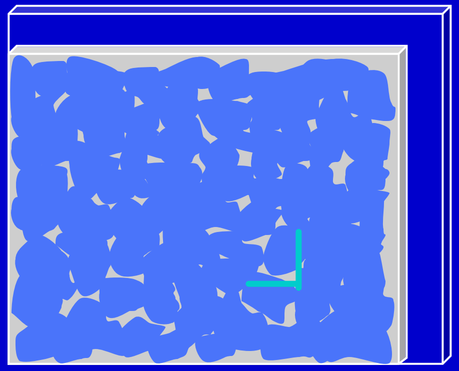


Courtesy Pileggi, MARCO GSRC

Soft Blocks

- Flexible blocks allow system assembly to more thoroughly exploit the available technology
- Interconnect problem is controlled via: soft boundaries for area re-shaping; re-synthesis and re-mapping for timing; smart wires; and top-down specified block synthesis
- Cf. “Amoeba” placement, coloring analysis of “good” placements with respect to original logic hierarchy, etc.





Superior timing, power and cost

Courtesy Pileggi, MARCO GSRC

Placement Directions

- Global placement
  - engines (analytic, top-down partitioning based, (iterative annealing based) remain the same; all support “anytime” convergent solution
  - becomes more hierarchical
    - block placement, latch placement before “cell placement”
  - support placement of partially/probabilistically specified design
- Detailed placement
  - LEQ/EEQ substitution
  - shifting, spacing and alignment for routability
  - ECOs for timing, signal integrity, reliability
  - closely tied to performance analysis backplane (STA/PV)
  - support incremental “construct by correction” use model

Routing Directions

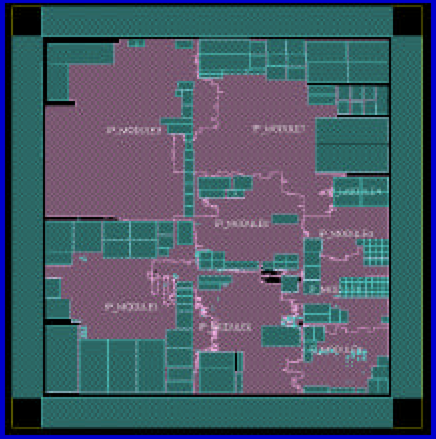
- Router ultimately responsible for meeting specs/assumptions
  - slew, noise, delay, critical-area, antenna ratio, PSM-amenable …
- Checks performability throughout top-down physical impl.
  - actively understands, invokes analysis engines and macromodels
- Many functions
  - circuit-level IP generation: clock, power, test, package substrate routing
  - pin assignment and track ordering engines
  - monolithic topology optimization engines
  - <u>owns</u> key DOFs: small re-mapping, incremental placement, device-level layout resynthesis
  - is hierarchical, scalable, incremental, controllable, well-characterized (well-modeled), detunable (e.g., coarse/quick routing), ...

Methodology Criteria/Directions

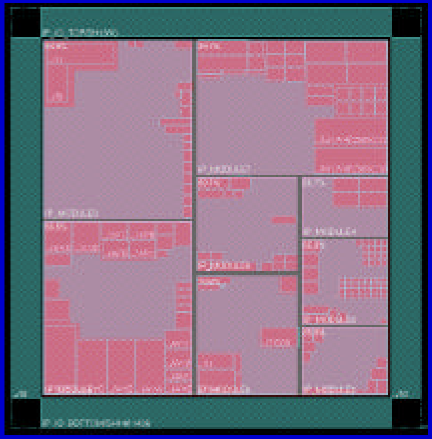
- Mixed-signal capability
  - Fully-hierarchical block-based SoC design
  - Timing, electrical design & verification
  - Power (current) distribution, electromigration
  - Concurrent engineering
- Functional design flexibility
  - Late-stage ECOs handled near tape-out
- Analog circuit design
  - System, chip, package, I/O optimization
  - Robust clock architecture
  - Verified crosstalk, Signal Integrity
- Technology readiness, electrical design experience
  - Validated, robust design margins → stable (high) yield
- Program management

Notes on “Silicon Virtual Prototyping” and Convergence Methodologies

Coarse Placement Drives Partitioning, Coarse Routing Drives Pin Assignment / Timing Opt

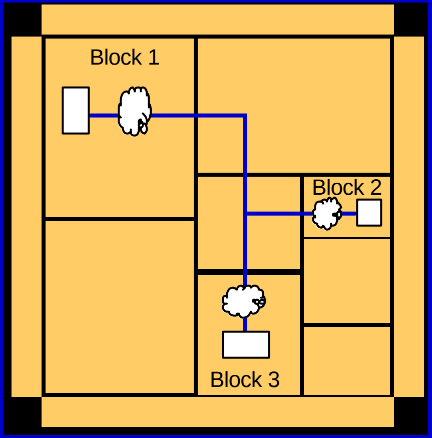


Physical Prototype



Partitioning

Block-Level Pin Assignments



Block-Level Timing Budgets

- Full-chip prototype results in optimal pin placement
  - Results in narrower channels and reduced die size
  - Reduces the routing congestion
  - Improves the chip timing
- Accurate timing budgets result in predictable timing convergence

“Physical Prototyping Philosophy”

RTL

↓

Gates

← *Functionality known*

↓

Physical Prototype

↓

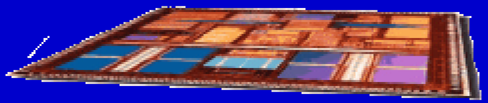
← *Timing / routability known*

Floorplan / Placement

↓

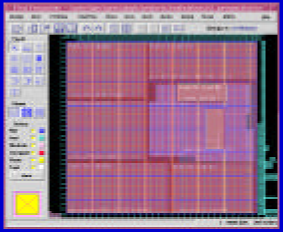
Routing

↓



- Prototype delivers accurate physical data
- Based on tape-out quality placement and ‘detail’ route
- Includes timing, clock tree and power analysis engine
- Hierarchical:
  - Chip-level CTS, top-level route and IPO, power analysis and grid design
  - Block-level synthesis, placement, IPO, routing
- “Handoff with enough physical information to ensure correct results”

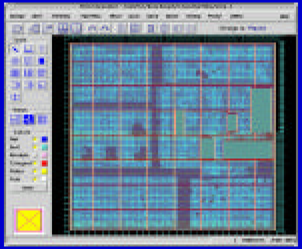
Cool Pictures of the Pieces…



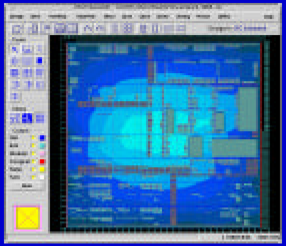
Full Chip Power Planning

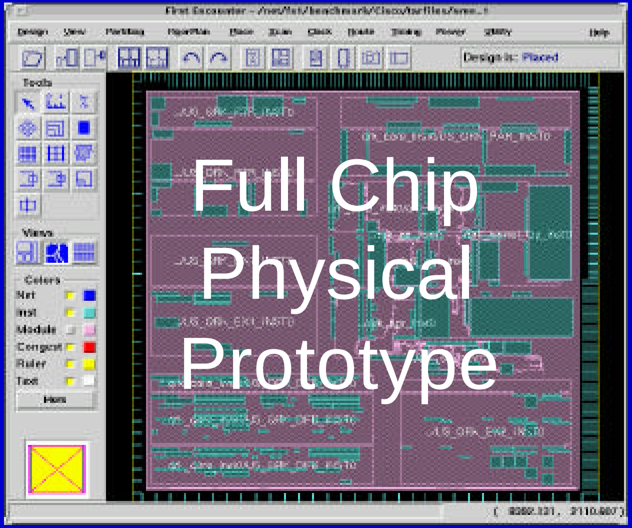
Place
Detailed Trial Route
RC Extraction
Delay Calc / STA
IPO

Timing Closure



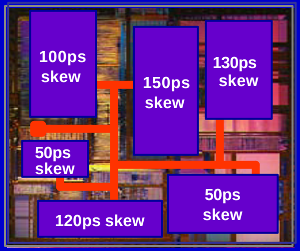
Power IR Drop Analysis



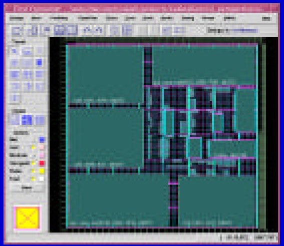


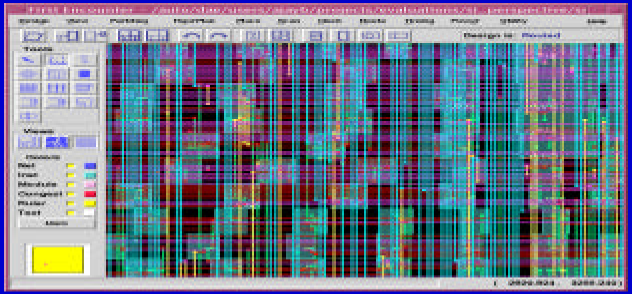
Full Chip Physical Prototype

Hierarchical Clock Tree Synthesis

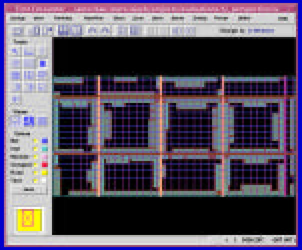


Partition



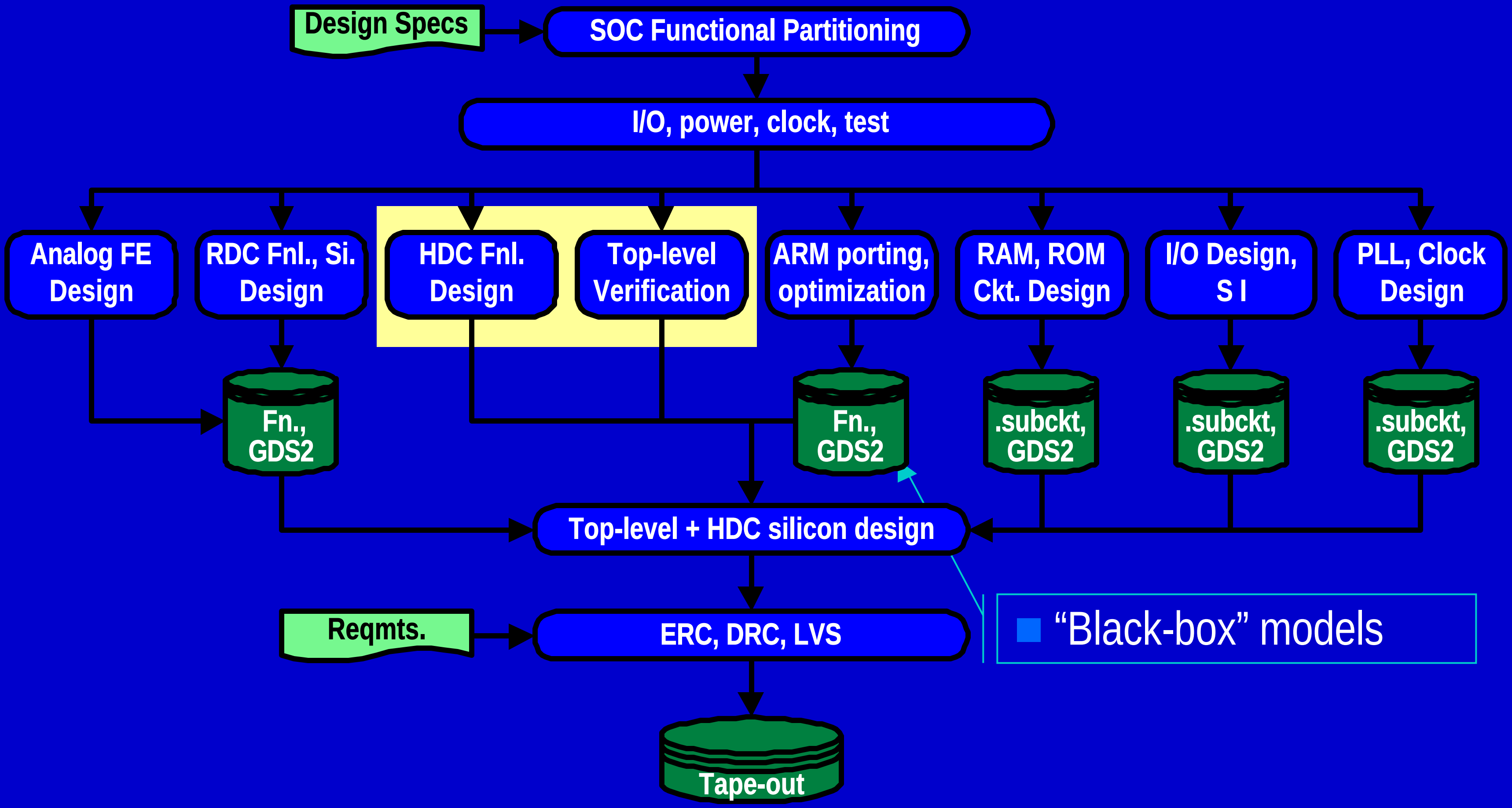


Block-Level Optimization



"Tape Out Every Day"

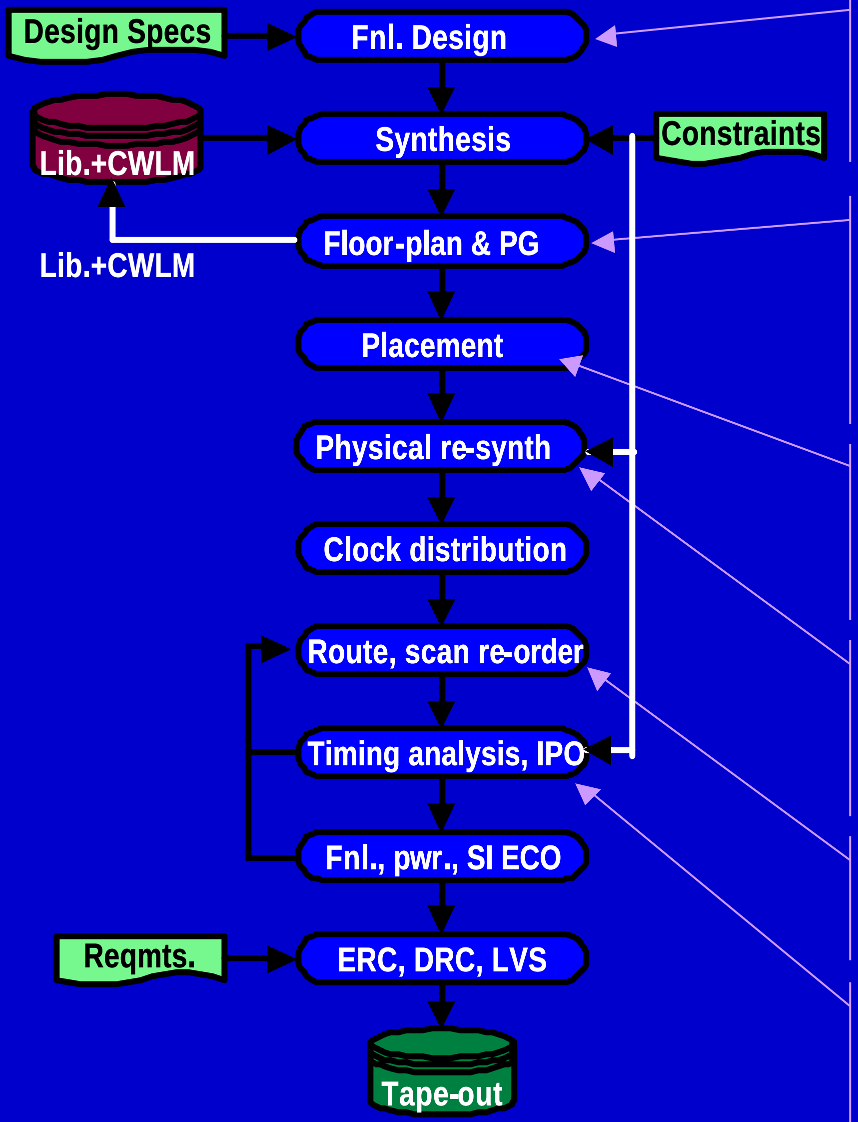
Top-Level Methodology



- "Black-box" models

- Functional design → hierarchical
- Electrical / physical design → hierarchical
- IP leverage; Customer-specific design

Block-Level Methodology



- Architectural optimization (timing)
- Inter-group buses, bandwidth
- Clock, SI, test; validation

- Floorplanning and custom WLM
- Power distribution (Internal, I/O)
- I/O driver, padring design
- Board-level timing, SI

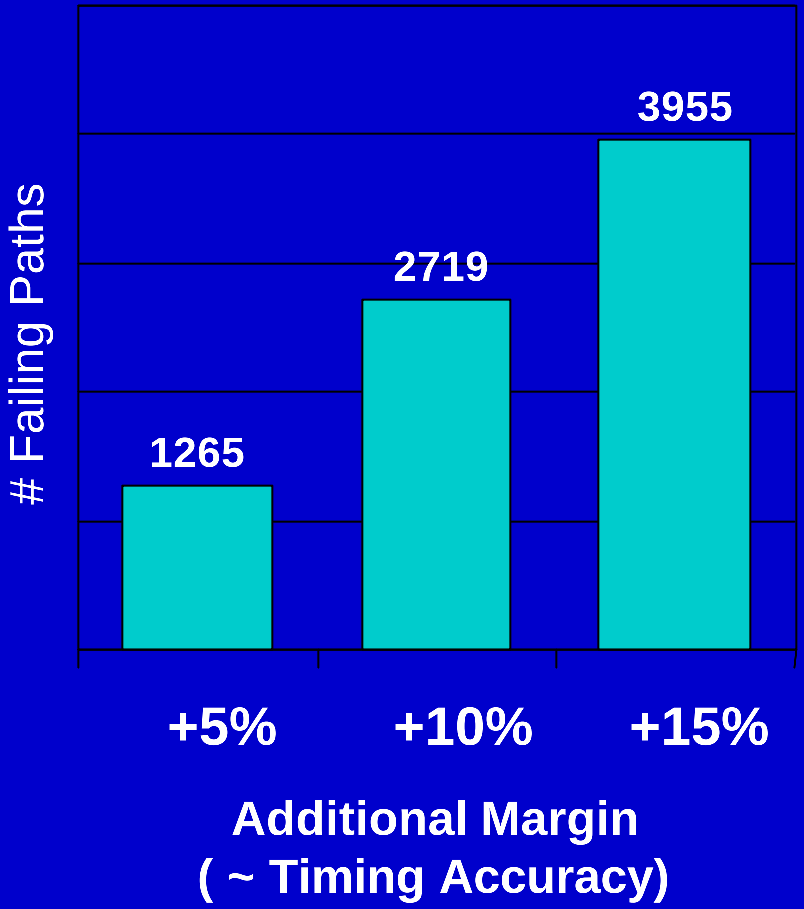
- Row definitions
- Placement of cells
- Congestion analysis

- Placement-based re-synthesis
- Noise minimization, isolation
- Clock distribution

- Full routing
- Scan stitching, re-ordering

- Full RC back-annotation
- Hierarchical timing, electrical and SI analysis and IPO/ECO

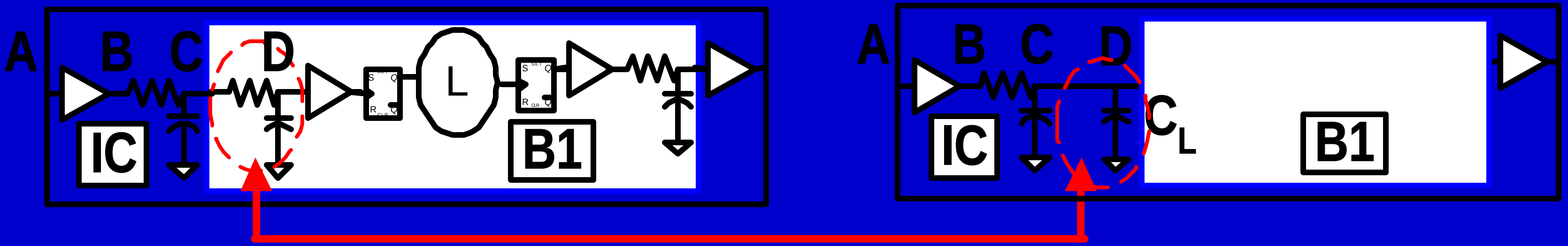
Accuracy → Less Margin Needed



- Margin has costs
  - More failed paths
  - Additional engineering resources
  - Longer to tape-out
  - Missed market opportunity
- Lost revenues, business opportunity
- .18µ? .13µ? .10µ = more cross-coupling, lower Vdd (more IR risk)

Hierarchical Delay Calculation

- Fully-hierarchical block-based timing analysis
  - Analyze large designs (scalable capacity)
  - Enable concurrent design
  - Faster timing convergence, verification (STA)
- Signal paths traverse hierarchy
  - Block inputs with ~0 – 2 mm metal → RC delay



- Model block boundary pin input RC as CL
- CL → timing inaccuracies when RC significant

Power-Timing Interactions

- Buffers get different VDD voltage
- This and IR drops cause timing closure problems if not accounted for
  - Additional failed paths
  - Race conditions