IP Integration Challenges and Solutions in 7nm and Beyond

Raymond Nijssen, Achronix VP and Chief Technologist, DAC 2019

Achronix®

Data Acceleration

Motivation

- Technology scaling continues 16nm -> 7nm -> 5nm etc.
  - Main driver: continued higher integration of functionality onto the same die
  - 1.8x higher density going from 16nm to 7nm (*)

- Opposing Forces :
  - Higher integration :
    - More IP instances
    - IP blocks getting larger
    - More functionally complex IP blocks

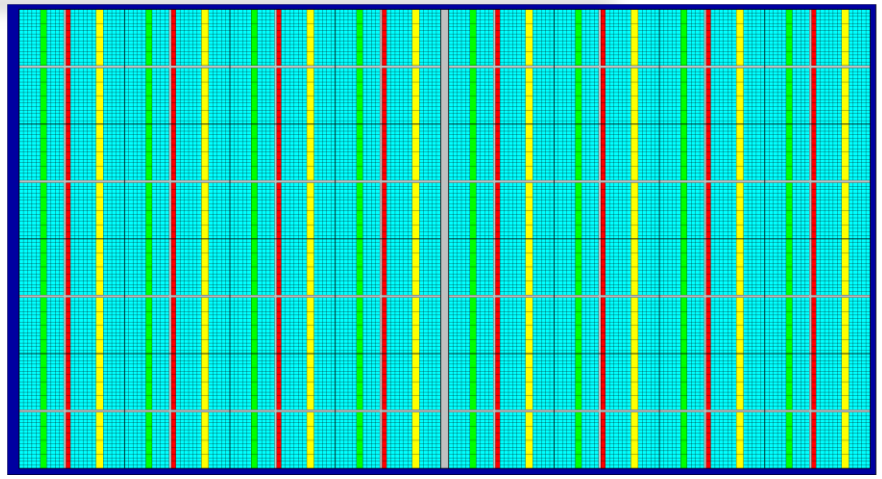
    This requires that IP integration must be made easier so IP integration does not become the limiting factor

  - Physical design requirements @7nm and beyond are getting
    - more restrictive
    - more complex
    - more pervasive
    - less abstractable

    This is making IP integration even harder

Example of a large IP Block

- Technology scaling enables IP blocks that would until recently be a separate chip
- Example: Embedded FPGA block ~250k LUTs



Characteristic	Magnitude
Internal clock insertion delay	2 .. 10ns
External non-common path typical	2 … 10ns
Boundary signal clock periods	> 1.33 ns
GOCV on x-boundary paths	< 4ns

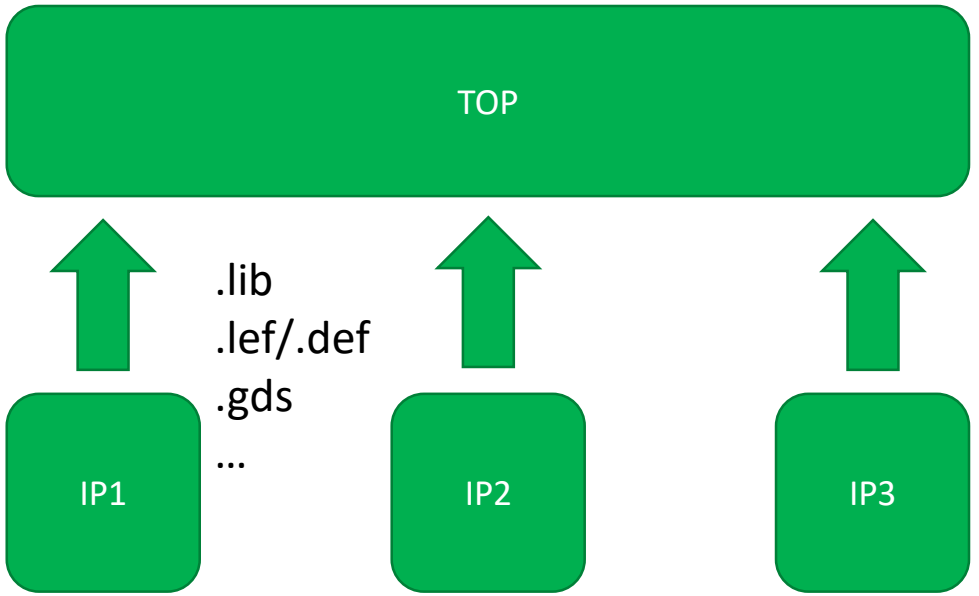
Characteristic	Magnitude
Technology	7nm
#layers	13
#LUT	246k
#BRAM Bytes	2.5M
#MLP	256
#Available pins	30,000
Area	~54mm2
Static Pwr FF@Tj=105C	<2.3W

Ideal Hard IP Integration Scenario

- Ideally, the hard IP-vendor creates the full layout and all integration collaterals once
  - And then delivers this packaged as a “kit” multiple times to multiple customers
  - The IP block is designed and implemented separately, in isolation

  - OK if the IP block does not depend on the environment into which it’s being integrated
    - The metal stacks of the IP and the ASIC are compatible
    - The configuration of the IP is the same across all applications
      - Same functional options
      - Same pin locations, block size
      - Same timing
      - Same power delivery network
    - Compatible sign-off methodologies

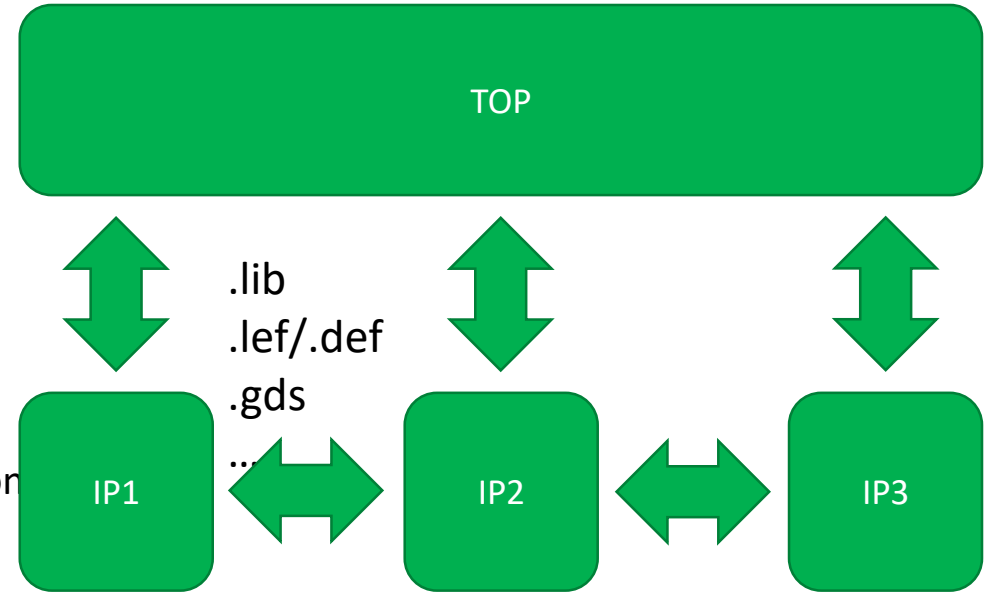
- This model is used for most smaller hard IPs
  - PLLs, IOs
  - Memory blocks (generators)
  - SerDes
  - Hardened PHYs



One-Directional IP Integration Dependency does not Scale

- Larger, more complex IP blocks do depend on the environment into which they're being integrated

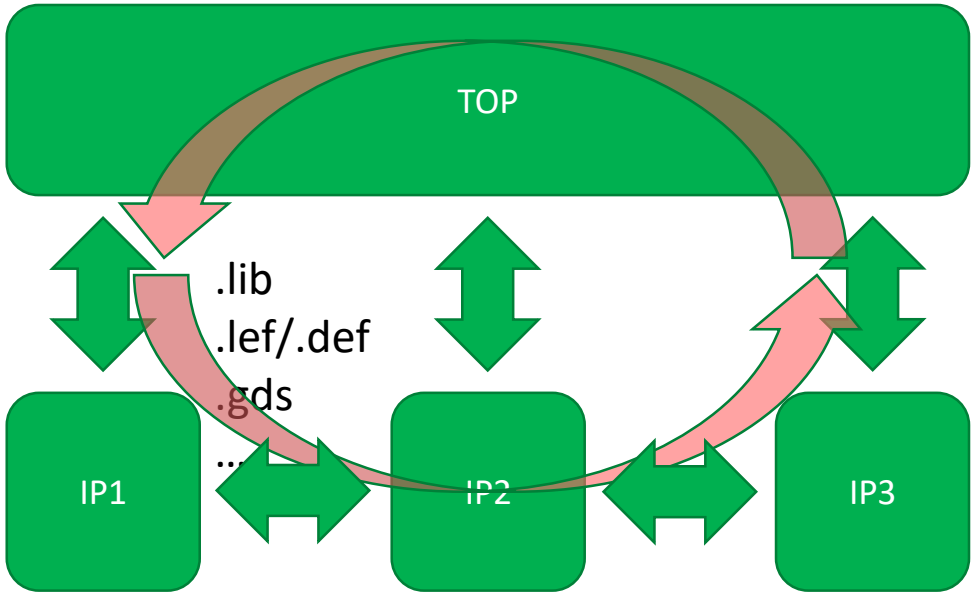
- Increasingly so with technology node scaling :
  - Blocks are getting bigger
    - Physical design dependencies
      - Custom pin locations, custom block size
      - Custom timing
      - Custom power delivery network
  - Blocks are getting more complex
    - Custom functional modes and options
  - More & stronger technology induced effects
    - Cross-boundary effects are breaking down levels of abstraction
    - Getting harder to create blocks in isolation



One-Directional IP Integration Dependency does not Scale

- Multiple iterations between integrator and IP block vendor needed per instantiation of the IP block
  - Cyclical dependencies must be broken with methodology
    - Not limited to paths between repeated instantiations
  - Iterations may converge slowly
  - Dependencies between instances of IP blocks
    - Including from different vendors

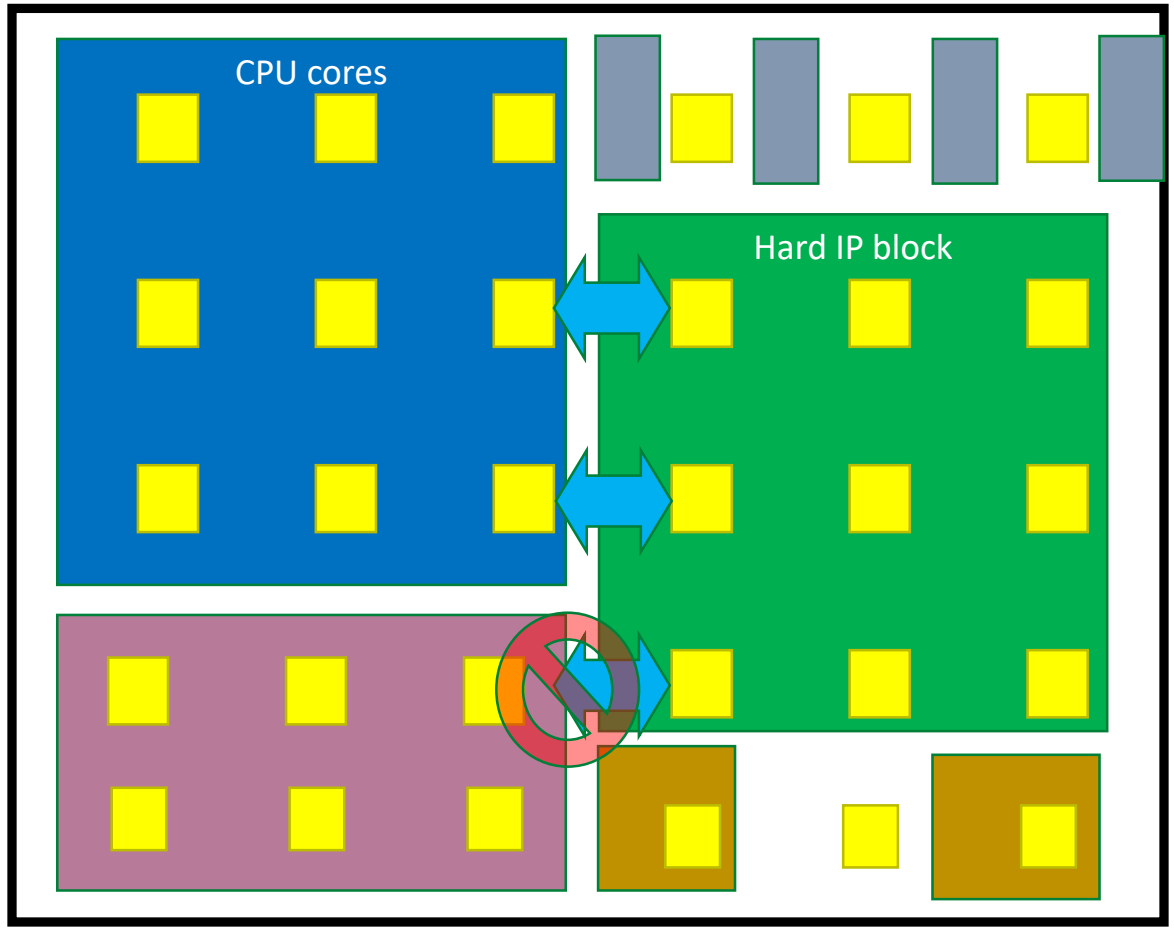
- This is fundamentally changing the engagement model of IP blocks
  - From a one-size-fits-all, off-the-shelf kit
  - To a collaborative contract-work sub-block co-design



Example: Process Monitoring Blocks

- Foundries require insertion of process monitoring and guidance blocks
- Design Rule: Must be distributed across die with max separation requirements

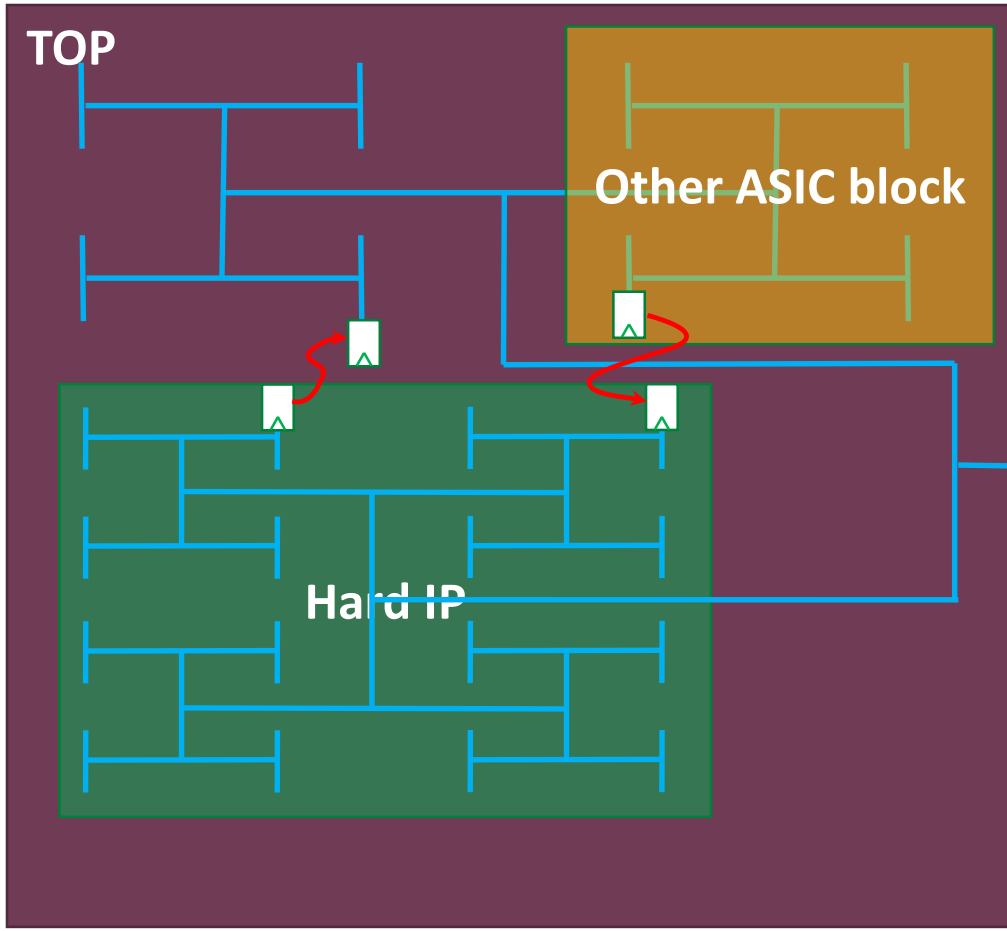
- These are large so their number must be minimized
- Large IP blocks must contain these
  - Must be placed inside IP block in a way that depends on context



Cross IP boundary Timing Closure

- Timing closure between blocks getting more challenging
  - Clock frequencies going up: 1 GHz +
  - Impact of OCV getting much worse
    - Large blocks most affected by this
    - Requiring more sophisticated modeling POCV – more complex
    - Simple global OCV derating factor no longer supported
      - GOCV would have had to be > 10%
  - The number of timing corners continues to increase
  - Clock balancing : task to match insertion delays across IP boundaries
    - Hard IP block comes with frozen internal clock distribution network
    - Larger blocks will likely cause larger insertion delay mismatches
    - Environment needs to match this

    Getting a lot harder as well
- Timing Corner Explosion

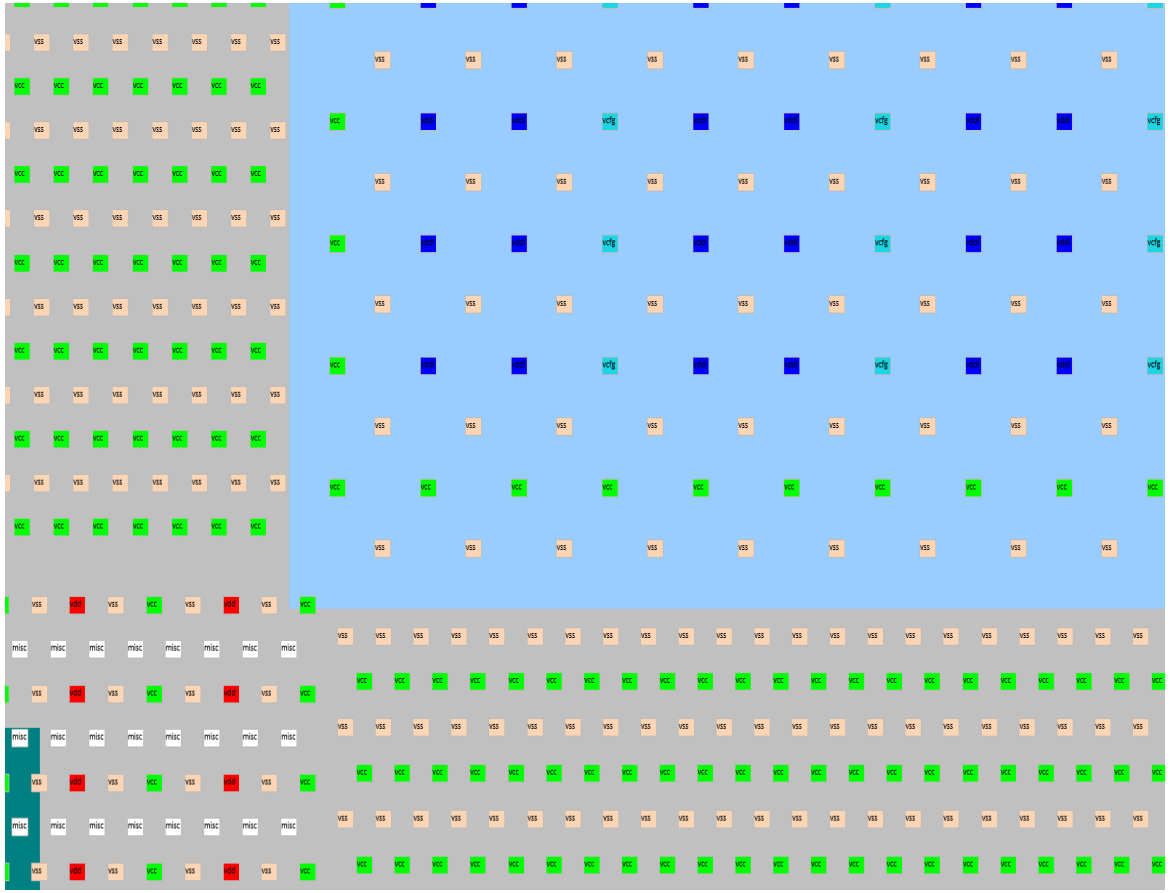


Power Integrity

- Large IP blocks have a large impact on the chips's PI

- Large IP blocks require their own power bumps
- Different customer power bump methodologies can often not be satisfied with same IP layout:
  - Some customers want IP block to own all bumps
  - Other customers want to own RDL over IP

- Power transients of IP block affect PDN resonance on-die, in package and on PCB
  - Chip PI analysis must be done with detailed models of IP block



Full-Chip Flat Sign-Off Analysis Requirements

- Sign-off analysis requires that models of the IP blocks be provided :

  STA, PEX, PI, SI, DRC, LVS, SDF back-annotated simulations

- Weakening of separation of abstraction between blocks calls for deeper cross-boundary analysis
  - Black boxes with abstract models are too inaccurate
  - Top-level sign-off analyses must look deep into IP blocks
  - IP vendors need to protect their IP and blocks
    - Need to keep models of IP opaque

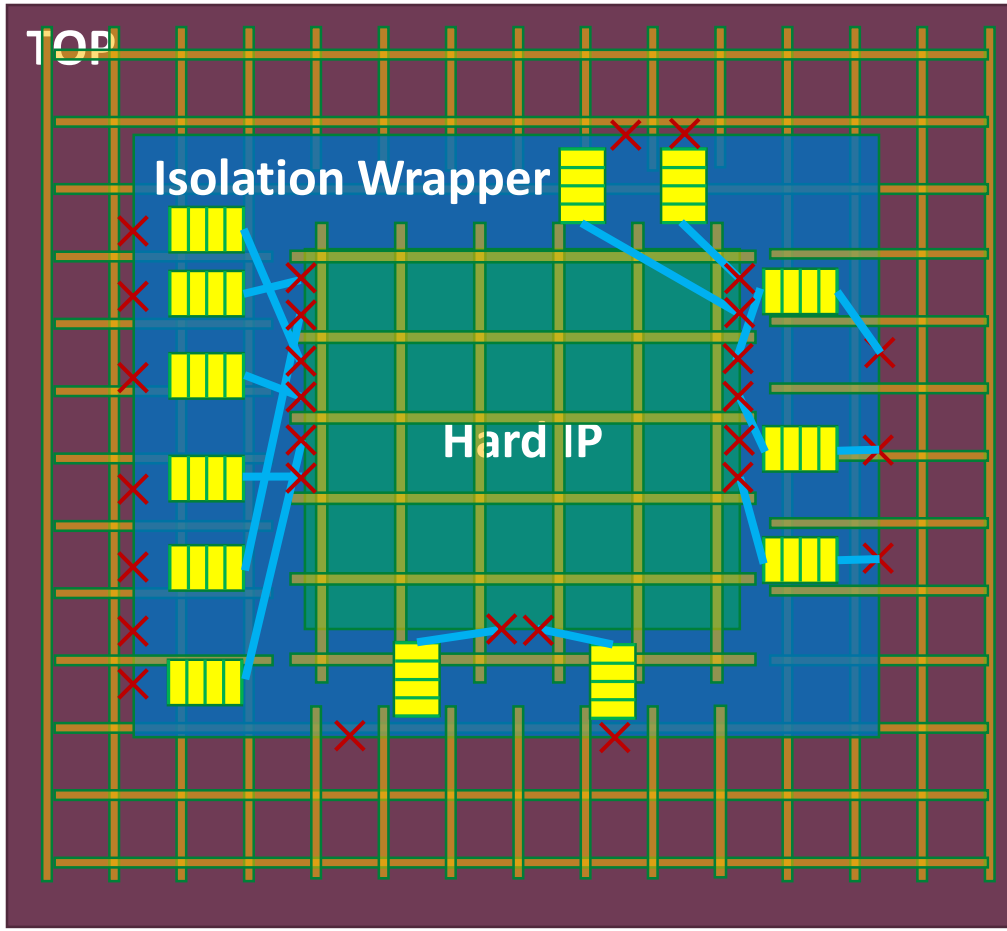
- Technology scaling enabled growth of IP blocks Full-Chip
  - Full-chip flat analysis without abstraction becoming computationally intractable

- Confidentiality of IP block difficult to maintain with customers insisting on flat sign-off analysis

- Solutions vary from encryption, different abstractions, contractual protections

Methodology Solutions

- Rigorously Eliminate / mitigate dependencies of IP block on its environment
  - Physical Design dependency Separation:
    - Physical wrappers in which the interfacing challenges are solved to insulate IP block from top level
    - Pin location adaptor
  - Electrical / Power dependency Isolation
    - Level shifters on all signal pins of IP block
    - Isolate PDN
      - Completely, or
      - Up to package, or
      - Up to RDL
      - Return of the power ring ?
  - Timing dependency Isolation:
    - Tighten existing methologies to decouple cross-block timing dependencies
      - Pre-inserted buffers with a-priori size on all boundary crossings
    - GALS
    - Mesochronous FIFOs



Methodology Solutions

- Hard IP: Isolation methodologies
  - “Chip-within-a-chip”
  - These measures come at a significant cost:
    - Area, performance (increased latency) and power
    - This may amortize to become acceptable as blocks grow in size
- Semi-Hard IP:
  - Ever closer collaboration between Integrator and IP vendor
  - Iterations
  - Similarities with PD of IP block done as design service for integrator
- Customer-owned PD of IP
  - Integrator must perform several PD tasks of the IP block
  - IP vendor provides many PD rules, relative placements, etc.

Summary

- Continued technology scaling >= 7nm is changing IP Integration
  - More, larger and more complex IP blocks
  - Separation of levels of abstraction is under increasing pressure
  - Blocks are becoming more dependent on environment
  - Increased collaboration between integrator and IP vendor required