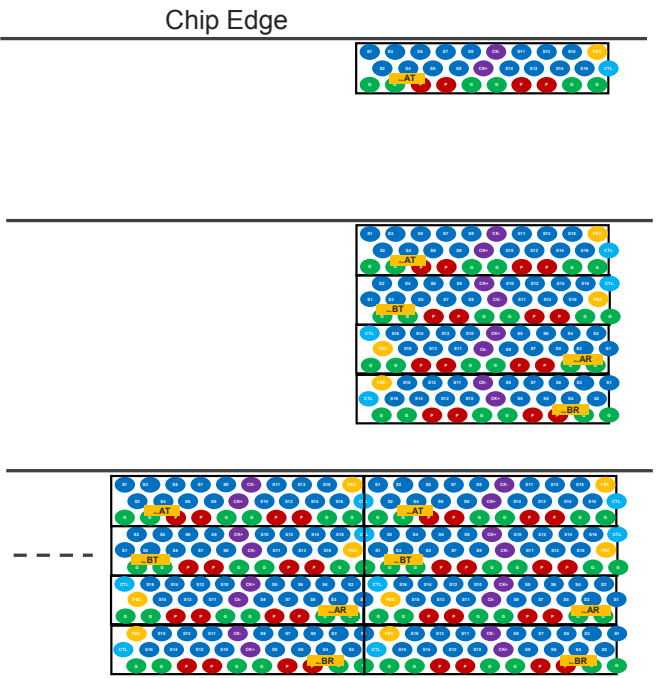
BoW Bump Maps, Routing, FEC, Redundancy

Ken Poulton	FEB 2020
Keysight Labs	

KEYSIGHT TECHNOLOGIES

BoW Hierarchy Naming Proposal

It would clarify discussions to standardize the names for BoW elements. Only ‘stack’ seems to be named consistently in the documents so far.



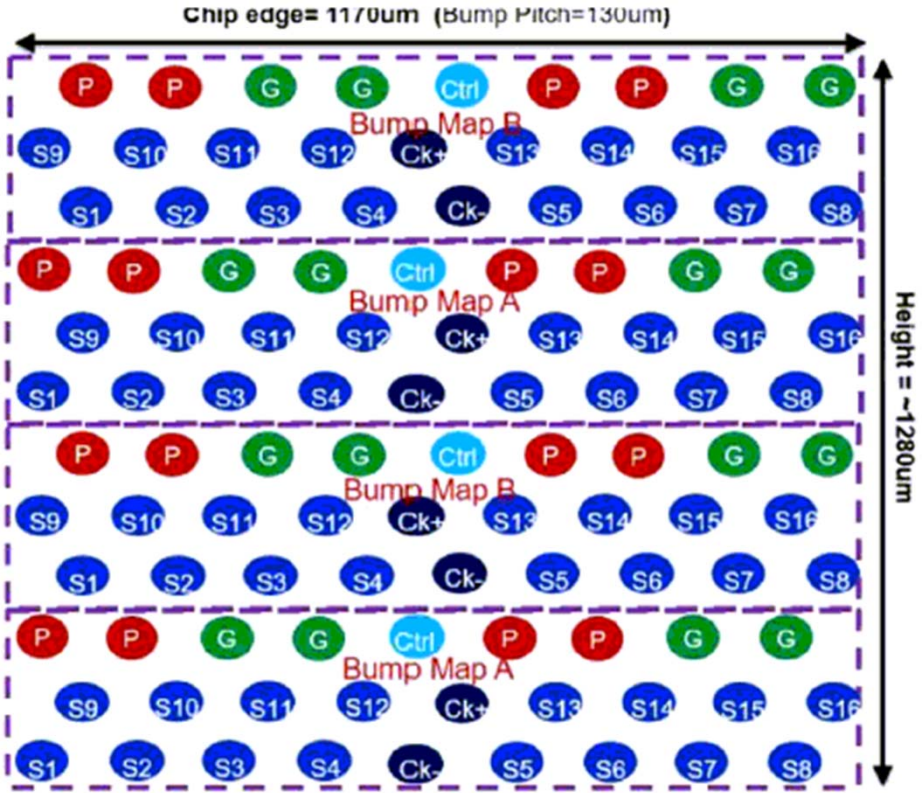
- Slice? Bunch? Group?
  bumps: 16 data + 2 Clk + FEC? + Ctl? + VDD/GND
  Plus the circuits associated with those bumps
- Stack
  up to 4 slices in a stack
- Link
  One or more stacks, operating as a single logical connection to one other chip, along with any control overhead (TBD)

Redundancy

- AIB and HBM include extra wires, with muxing on chip to utilize them
  - They can route around one defective wire or bump per group
  - An HBM2 memory stack and interface has ~1000 wires and 13,600 ~50-um-pitch bumps (mostly inside the memory chip stack)
- Yield data
  - Can anyone find yield data on existing laminate (single-die) packages with high wire count ?
    - Just use the number of signal balls, exclude supplies.
  - Can anyone provide data on how often HBM or AIB memory redundancy is actually used?
- **Do we need redundancy?**
  - What is our target complexity? 100 wires, 1000 wires, 10,000 wires?
  - If we add redundancy, it grows the BoW width by another 65 um (5%).

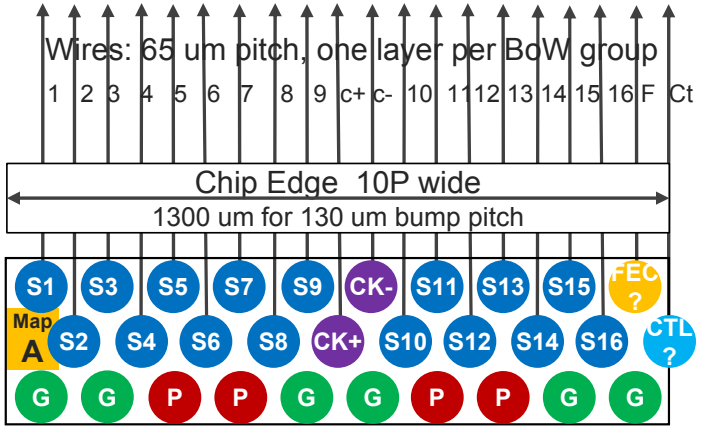
Sep 2019 BoW bump map

- Issues
  - Needs FEC bump
  - Should Ctrl be per-group or per-stack or per-link?
  - Shouldn't Ctrl bump should be in signal row to keep routing regular?
  - Why do S1 and S9 not stay in same alignment in A and B maps?
  - Do we even need A and B maps?
  - How does this align on N,S,E,W chip edges?
  - How does this align for 180-deg chiplet rotation?
  - Doubling power rows regularizes signal routing and does not impede interoperability - might be mentioned as a routing choice
  - How do chiplets with different stacking numbers interoperate?



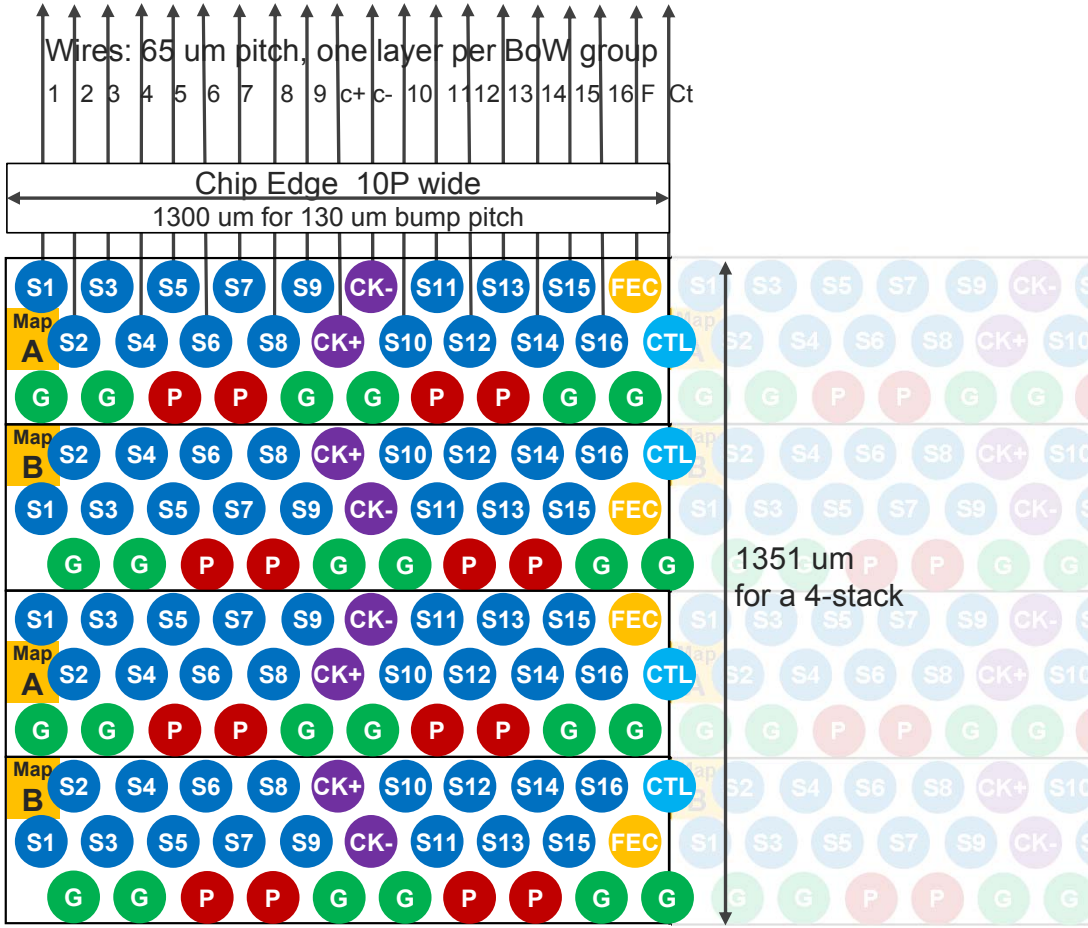
Wire Numbering

- The proposal 0.7 numbers wires starting with 1
- Most wire busses seem to start with 0
- What does PIPE use?
- Should we renumber starting with 0?



Base Bump Maps A & B

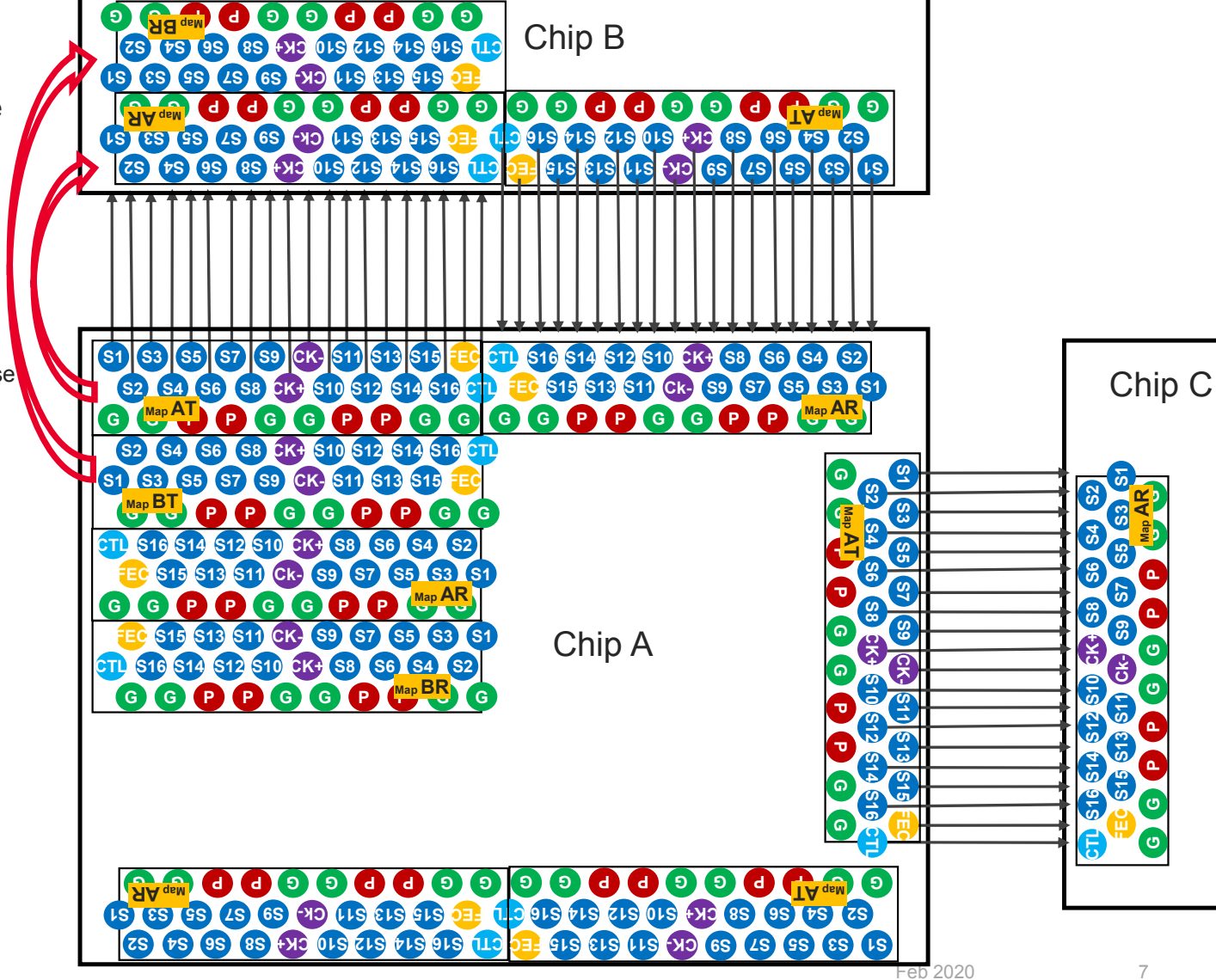
- Alternating two bump maps A & B allow hexagonal-packed bumps
- Change 1: all bumps that are wired to the other chip reside in the two signal bump rows (moved FEC + CTL)
- Change 2: Clocks are moved one bump to re-center them in the row (future bidirectional compat)
- Change 3: bumps are renumbered so that wire number is the same for Map A and Map B
- Change 4: Power/Ground row is away from chip edge
- Chip designer can choose to use only Map A (to save layout time) at a 5% cost in extra bump area
- Chip designer can choose number and order of power/ground bumps and how many rows they occupy (to ease routing in package)
- Unresolved: Function and existence of FEC, CTL, and REDundant bumps and wires



RED ?

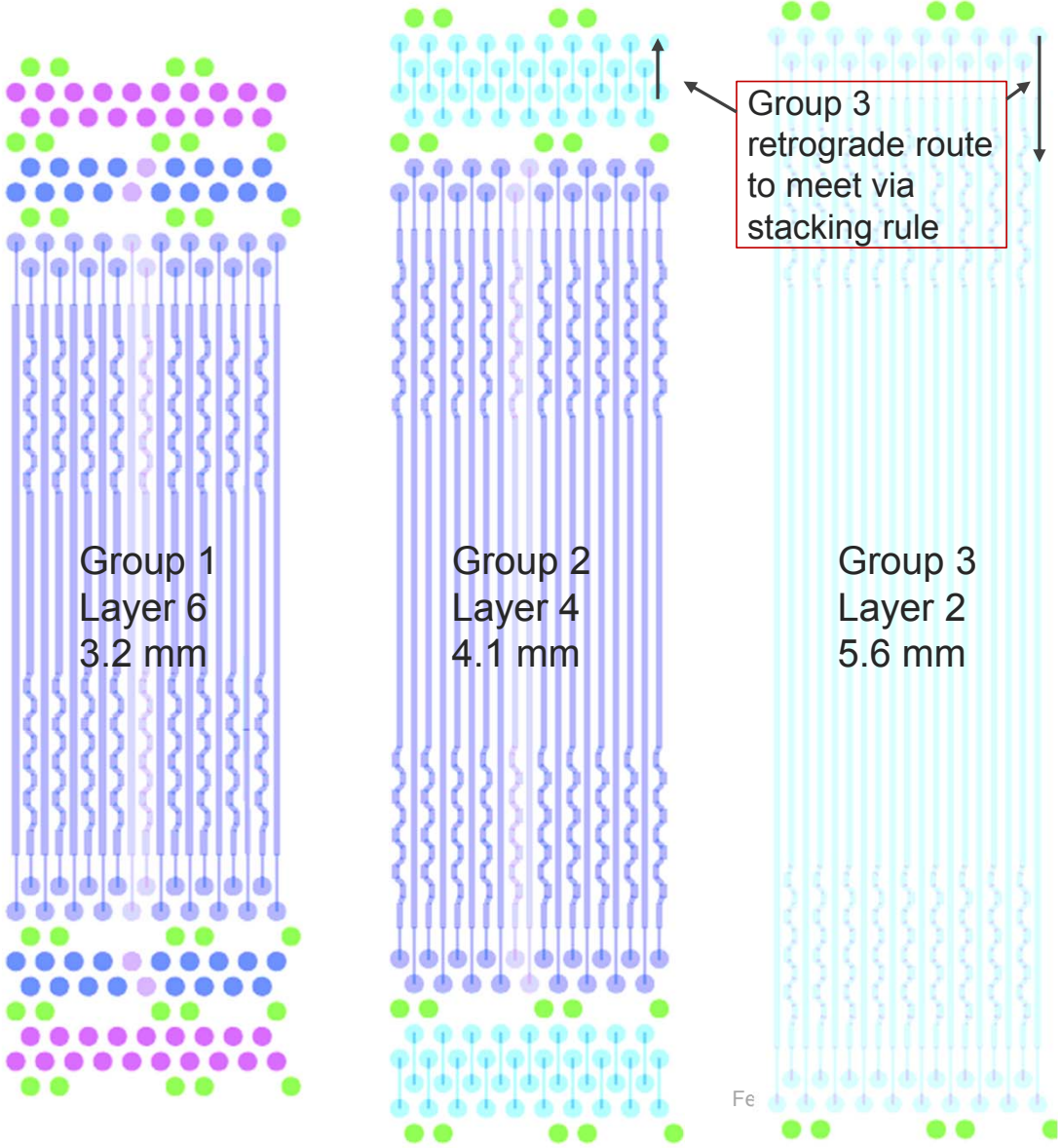
Orientation Proposal

- Each Rx map is mirrored about the chip edge from Tx map
- Now we have 4 maps: AT, AR, BT, BR
- Rules:
  - Supply row away from chip edge
  - Use Map A at chip edge
  - then alternate Map A and B for best packing, or use all Map A for simplicity
- + pin numbers line up from chip to chip
- + hexagonal closest packing when using A and B maps
- -- Tx group has reverse wire numbering from Rx group
- + CK pins centered to support bidirectional groups
- Unresolved: preferred method for T/R alternation – alternate within each stack, or alternate stack to stack?



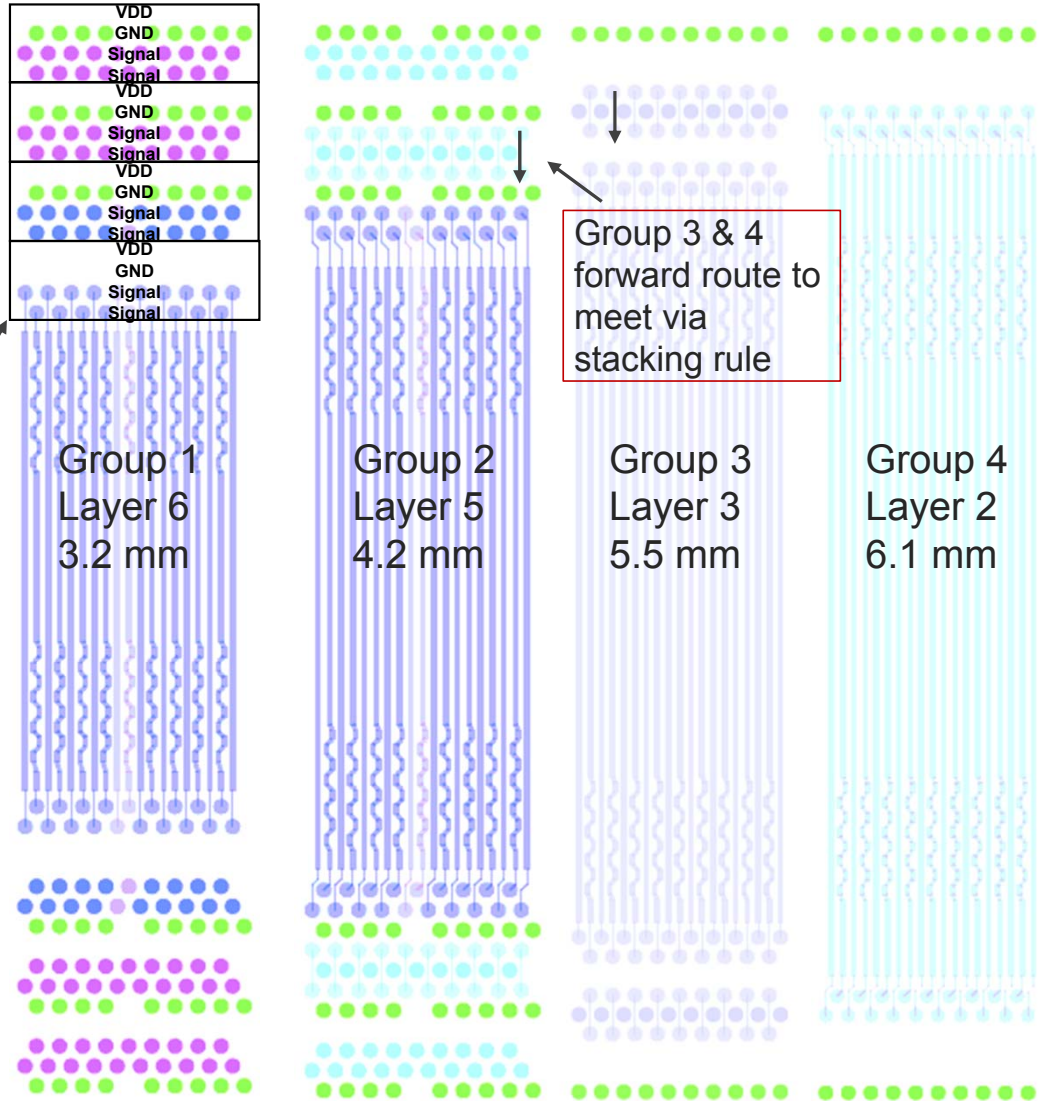
Routing Study – 3 Stack

- Package Stackup: 7 buildup – 2 core – 7 buildup
  - At this pitch, 130 µm, cannot fit large core vias and route on the core layer, so only the top 7 layers are usable for signals
- 2.5 mm die edge to die edge
  - Die edge is about 150 µm to first bump
- G-s-G layer assignment minimizes crosstalk:
  - 8 (top) PWR
  - 7 GND
  - 6 Signals for BoW group 1 (pins closest to die edge)
  - 5 GND
  - 4 Signals for BoW group 2
  - 3 GND
  - 2 Signals for BoW group 3 (pins farthest from die edge)
  - 1 GND
- Wire lengths were equalized within a group (wiggles)
  - Wiggles are not needed. Difference ≈ 225 um ≈ 55 fF ≈ 2 ps
- Group 2 wires are 0.7 mm longer than Group 1
  - Extra length from bump placement and vias to lower layers
- Group 3 wires are 1.5 mm longer than Group 2
  - Package substrate design rules are limited to 4 stacked micro-vias, therefore group 3 traces are routed retrograde on layer 4
- Multi-signal crosstalk is ~27% of the total single-trace capacitance to ground
  - Total mutual capacitance/trace = 1.8 pF for 10 mm
  - Including miller multiplication → C_m = 2.3 pF for 10 mm



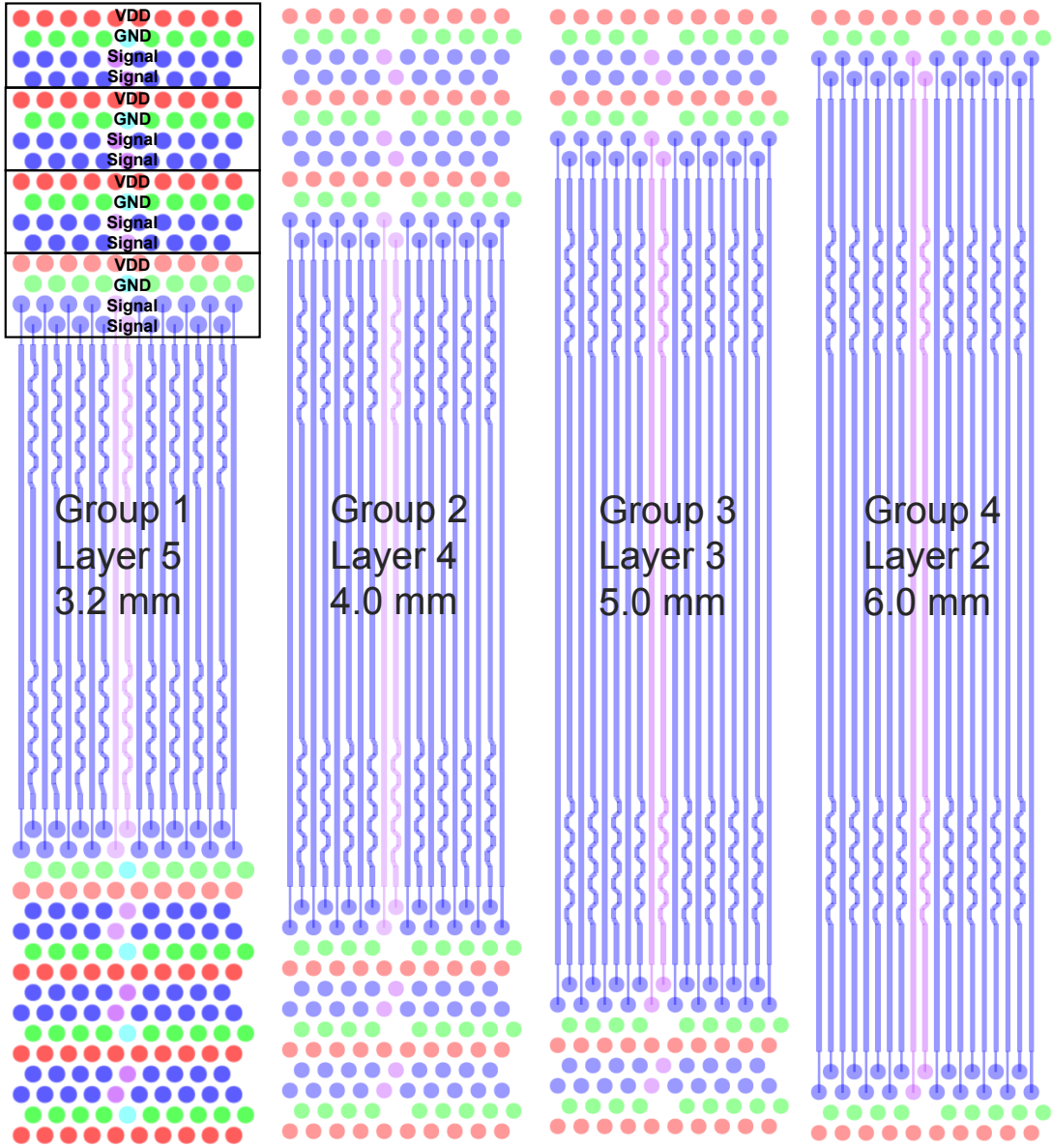
Routing Study – 4 Stack

- Package Stackup: 7 buildup – 2 core – 7 buildup
  - At this pitch, 130 µm, cannot fit large core vias and route on the core layer, so only the top 7 layers are usable for signals
- 2.5 mm die edge to die edge
  - Die edge is about 150 µm to first bump
- G-s-s-G layer assignment has more crosstalk:
  - 8 (top) PWR
  - 7 GND
  - 6 Signals for BoW group 1 (pins closest to die edge)
  - 5 Signals for BoW group 2
  - 4 GND
  - 3 Signals for BoW group 3
  - 2 Signals for BoW group 4 (pins farthest from die edge)
  - 1 GND
- Placed bumps in 4 rows: 2 signal, 1 GND, and 1 power
  - Allows via un-stacking in forward steps rather than retrograde steps
  - Result: Longest traces are only 10% longer than in the 3-stack plan on previous page
  - More bump area leads to shorter traces
- Multi-signal crosstalk is ~~65% of the total single-trace capacitance to ground
- Requires 30% more drive strength than G-s-G (3 Stack)



Routing Study – Min Layers

- All-signal layer assignment for highest density: one buildup layer per BoW group in the stack
- Package Stackup: 4 buildup – 2 core – 4 buildup
  - At this pitch, 130 µm, cannot fit large core vias and route on the core layer
  - Could go down to a 1 Stack (2-2-2) and up to a 5 stack (6-2-6), based on the package via stacking rules
- 2.5 mm die edge to die edge
  - Die edge is about 150 µm to first bump
- Layer stack:
  - 5 (top) Signals for BoW group 1 (pins closest to die edge)
  - 4 Signals for BoW group 2
  - 3 Signals for BoW group 3
  - 2 Signals for BoW group 4 (pins farthest from die edge)
  - 1 GND
- Multi-signal crosstalk is ~~100% of the total single-trace capacitance to ground
  - Total mutual capacitance/trace = 1.8 pF for 10 mm
  - Including miller multiplication → C_m = 3.6 pF for 10 mm
- Requires 55% more drive strength than G-s-G (3 Stack)



Routing Study Conclusions

- Routing works well when one signal layer is allocated to each level of a BoW stack
  - 130 um bump pitch, 65 um wire pitch is close to design rules
- The BoW group nearest the chip edge routes on the top-most signal layer
  - The farthest-farthest connection is 2-3 mm longer than the nearest-nearest
- All bumps connected to chip-to-chip wires should be in the two signal-bump rows to keep an even wire pitch
- It pays to give the chip designer freedom to choose extra power/ground row(s)

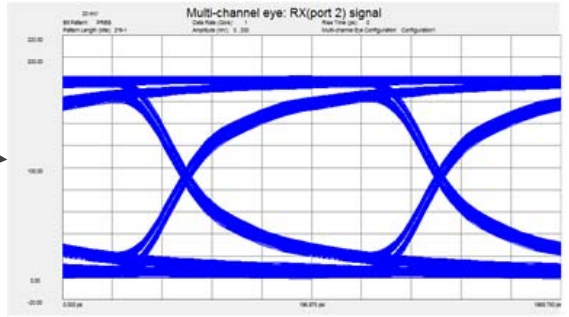
Comparison of Layer Stackups

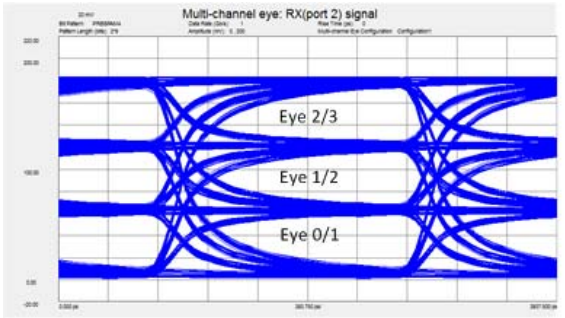
Layer assignment	Layer count for 4-BoW stack	Ctot (fF/mm)	Crosstalk/ Ctot	Cworst-transition (fF/mm)
g-S-g-S-g…	8-2-8	180	27%	229
g-S-S-g…	6-2-6	180	65%	297
S-S-S-S-g	4-2-4	180	65-100%	360

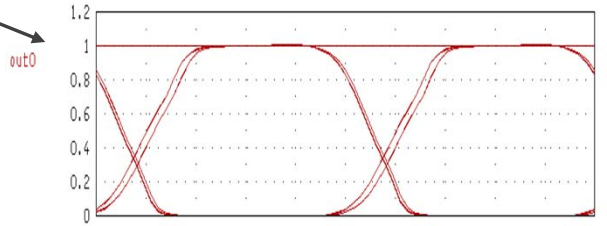
- More ground layers reduces crosstalk, but takes more package layers
  - or limits the BoW stack depth and increases the chip perimeter needed
- Crosstalk causes bounded jitter - with full-swing signals this does not affect the next symbol
  - A 4-BoW stack can be routed with longest wires <6 mm with these design rules
    - => 10 mm is a reasonable limit (allows for routing choices and limited fanout to different bump pitches)
- Power is proportional to Ctot, but the Tx driver should be sized for Cworst-transition
- Proposal for the unterminated case: spec the C, R, slew rate values (not loss)
  - Spec the wires at 2 pF max (max ~ 10 mm) and 45 ps RC product (limits interposer to ~3 mm)
  - Spec the Tx drivers for 45 ps ? risetime with 4 pF load = 200/100 um CMOS driver
- Possible length issue with terminated lines: a 50 mm line will have too much C for unterm driver

FEC Needs

- A typical spec for NRZ serdes BER is 1e-12 to 1e-15.
  - This is considered sufficient for most real-time applications
  - Many data protocols (e.g., TCP) add packet-level retry layered on top
  - Some real-time applications (e.g., electronic measurements) need lower BER (e.g., 1e-20) and will include HW FEC
- Newer PAM4 serdes have uncorrected BER in the 1e-4 to 1e-10 range
  - Most PAM4 serdes include mandatory FEC just to reach 1e-15
- Decision Feedback Equalization (DFE) used in serdes tends to convert a single error into a burst of errors
  - This drives the popularity of Reed-Solomon and other complex error-correction codes
- BoW, with modest bit rates and short wires, can be designed for full-swing signals
  - We should be able to get 1e-15 without FEC
  - Some applications will require FEC improve the BER or to prove that the link is working well



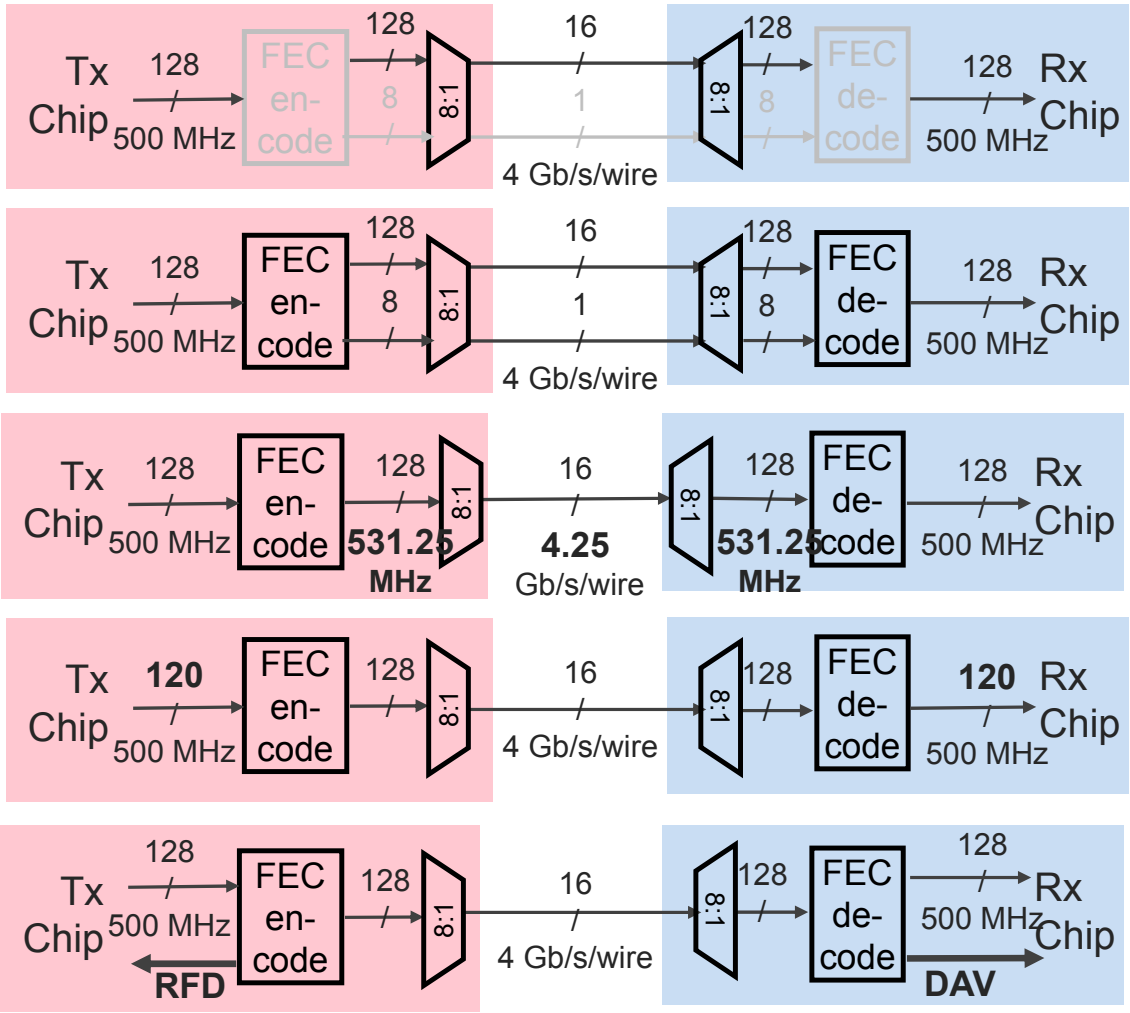




FEC Topology Choice

1. FEC Off
   - Simple 2^N values for clock ratios and bit widths
2. Add FEC wire in the interface (current plan)
   - Cost: extra wire
3. Change data rate (common in serdes)
   - Carry FEC data with an increased BoW data rate
   - Cost: generating and using odd clock frequencies, gearboxes
4. Reduce payload width
   - Cost: Odd numbers at the core interface
5. Reserve cycles for FEC data
   - Delivers payload data only 16 cycles out of 17
   - Cost: adds flow control (ReadyForData and DataValid signals)

Using 4 Gb/s and 500 MHz as round-number example values:



BoW “FEC Lighter” Proposal

- FEC Code
  - Short -reach NRZ links inherently have very low bit error rate, typically BER < 1e-15
    - An FEC that can correct one error per 256-bit frame improves BER to < 2.5e-28 if the errors are independent
  - Reed-Solomon is preferred for serdes because of its burst-correction capability
    - DFE is the primary cause of error bursts in serdes; we have no DFE
    - Reed-Solomon is complicated
  - If we are not concerned with bursts, then a simple correct-one, detect-two ECC is sufficient
    - Available as DesignWare ECC module in common libraries (no extra IP charge)
- Latency and block size
  - For either ECC or RS(34,32,8) we need a block size of >170 data bits to fit in 6% overhead (one FEC wire)
    - A block of 256 bits from one BoW group takes 16 cycles at the line rate, which is 3.2 ns at 5 Gb/s
    - A lower-latency approach: when 2-4 BoW groups are stacked, compute the FEC over the data from those 2-4 groups.
      - + The min of 256 is captured in just one logic cycle, reducing block latency to ~2 ns
      - -- Both ends have to do this the same way
    - ECC computation takes just one synthesized clock cycle (~2 ns). Reed Solomon takes more.