Previously

- Saw
  - need to exploit locality/structure in interconnect
  - a mesh might be useful
  - Rent’s Rule as a way to characterize structure

Today

- Mesh:
  - Channel width bounds
  - Linear population
  - Switch requirements
  - Routability
  - Segmentation
  - Clusters

Mesh

Street Analogy

Manhattan

Mesh

switchbox
**Mesh**

- Strengths?

**Mesh Channels**

- Lower Bound on \( w \)?
  - Bisection Bandwidth
    - \( BW \propto N^p \)
    - channels in bisection \( \propto N^{0.5} \)
    - \( W \propto \frac{N^p}{\sqrt{N}} = N^{(p-0.5)} \)
    - Channel width grows with \( N \).

**Straight-forward Switching Requirements**

- Total Switches?
- Switching Delay?

**Switch Delay**

- Switching Delay:
  - Manhattan distance
    - \(|X_i - X_j| + |Y_i - Y_j|\)
    - \(2 \sqrt{(N_{\text{subarray}})}\)
    - worst case:
      - \( N_{\text{subarray}} = N \)

**Total Switches**

- Switches per switchbox:
  - \( 4 \times (3w \times w)/2 = 6w^2 \)
  - Bidirectional switches
    - \( (N \rightarrow W \text{ same as } W \rightarrow N) \)
    - double count

- Switches into network:
  - \( (K+1)w \)

- Switches per PE:
  - \( 6w^2 + (K+1)w \)
  - \( w = cN^{p-0.5} \)
  - Total \( \propto w^2 \times N^{2p-1} \)
- Total Switches: \( N \times (\text{Sw}/\text{PE}) \propto N^{2p} \)
Routability?

- Asking if you can route in a given channel width is:
  - NP-complete
- Contrast with Beneš, Beneš-crossover tree….

Linear Population Switchbox

Traditional Mesh Population: Linear

- Switchbox contains only a linear number of switches in channel width

Linear Mesh Switchbox

- Each entering channel connect to:
  - One channel on each remaining side (3)
  - 4 sides
  - W wires
  - Bidirectional switches
    - (N→W same as W→N)
    - double count
  - 3×4×W/2=6W switches
  - vs. 6w² for full population

Total Switches

- Switches per switchbox:
  - 6w
- Switches into network:
  - (K+1) w
- Switches per PE:
  - 6w +(K+1) w
  - w = cN^p-0.5
  - Total = N^p-0.5
- Total Switches: N×(Sw/PE) = N^p+0.5 > N

Total Switches (linear population)

- Total Switches
  - ∝ N^p+0.5
  - N < N^p+0.5 < N^2p
- Switches grow faster than nodes
- Wires grow faster than switches
Checking Constants (Preclass 3)
When do linear population designs become wire dominated?
• Wire pitch = 4 F
• switch area = 625 F^2
• wire area: (4w)^2
• switch area: 6x625 w
• Crossover?

Checking Constants: Full Population
Does full population really use all the wire physical tracks?
• Wire pitch = 4F
• switch area = 625 F^2
• wire area: (4w)^2
• switch area: 6x625 w^2
• effective wire pitch: 60F
  ~15 times pitch

Practical
• Full population is always switch dominated
  – doesn’t really use all the potential physical tracks
  – …even with only two metal layers
• Just showed:
  – would take 15x Mapping Ratio for linear population to take same area as full population (once crossover to wire dominated)
• Can afford to not use some wires perfectly
  – to reduce switches (area)

Diamond Switch
• Typical linear switchbox pattern:
  – Used by Xilinx

Mapping Ratio?
• How bad is it?
• How much wider do channels have to be?

Mapping Ratio
• Empirical:
  – Seems plausibly, constant in practice
• Theory/provable:
  – There is no Constant Mapping Ratio
    • At least detail/global
  – can be arbitrarily large!
Domain Structure

- Once enter network (choose color) can only switch within domain

Detail Routing as Coloring

- Global Route channel width = 2
- Detail Route channel width = N
  - Can make arbitrarily large difference

Routing

- Lack of detail/global mapping ratio
  - Says detail can be arbitrarily worse than global
  - Doesn’t necessarily say domain routing is bad
    - Maybe can avoid this effect by changing global route path?
  - Says global not necessarily predict detail
  - Argument against decomposing mesh routing into global phase and detail phase
    - Modern FPGA routers do not
    - VLSI routers and earliest FPGA routers did

Buffering and Segmentation

Buffered Bidirectional Wires
Segmentation

- To improve speed (decrease delay)
- Allow wires to bypass switchboxes
- Maybe save switches?
- Certainly cost more wire tracks

Segmentation

- Segment of Length $L_{seg}$
  - 6 switches per switchbox visited
  - Only enters a switchbox every $L_{seg}$
  - SW/sbox/track of length $L_{seg} = 6/L_{seg}$

Segmentation

- Reduces switches on path $\sqrt{N/L_{seg}}$
- May get fragmentation
- Another cause of unusable wires

Segmentation: Corner Turn Option

- Can you corner turn in the middle of a segment?
- If can, need one more switch
- SW/sbox/track = $5/L_{seg} + 1$
Delay of Segment

\[ T_{\text{seg}} = T_{\text{sw}} + \left( L_{\text{seg}} \right)^2 \times R_{\text{seg}} \times C_{\text{seg}} \]

Segment R and C

\[ T_{\text{seg}} = T_{\text{sw}} + \left( L_{\text{seg}} \right)^2 \times R_{\text{seg}} \times C_{\text{seg}} \]

Preclass 4

• Fillin Tseg table together.

Preclass 4

• What \( L_{\text{seg}} \) minimizes delay for:
  • Distance=1?
  • Distance=2?
  • Distance=6?
  • Distance=10?
  • Distance=20?

VPR \( L_{\text{seg}} = 4 \) Pix

VPR \( L_{\text{seg}} = 4 \) Route
Effect of Segment Length?

- Experiment with on HW9

C-Box Depopulation

- Not necessary for every input to connect to every channel
- Saw last time:
  - $K \times (N-K+1)$ switches
- Maybe use fewer?

IO Population

- Toronto Model
  - $F_c$ fraction of tracks which an input connects to
- IOs spread over 4 sides
- Maybe show up on multiple
  - Shown here: 2

Clustering
Leaves Not LUTs

- Recall cascaded LUTs
- Often group collection of LUTs into a Logic Block

Logic Block

What does clustering do for delay?

Delay versus Cluster Size

[Lu et al., FPGA 2009]

Area versus Cluster Size

[Lu et al., FPGA 2009]

Review: Mesh Design Parameters

- Cluster Size
  - Internal organization
- LB IO (Fc, sides)
- Switchbox Population and Topology
- Segment length distribution
  - and staggering
- Switch rebuffering

Big Ideas

[MSB Ideas]

- Mesh natural 2D topology
  - Channels grow as $\Omega(N^{2.5})$
  - Wiring grows as $\Omega(N^{2.5})$
  - Linear Population:
    - Switches grow as $\Omega(N^{2.5})$
      - Worse than shown for hierarchical
    - Unbounded global-detail mapping ratio
    - Detail routing NP-complete
    - But, seems to work well in practice…
Admin

- HW8 due Wednesday
- HW9 out
- Reading for Wednesday on web