CS184a: Computer Architecture (Structure and Organization)

Day 14: February 10, 2003

Interconnect 4: Switching

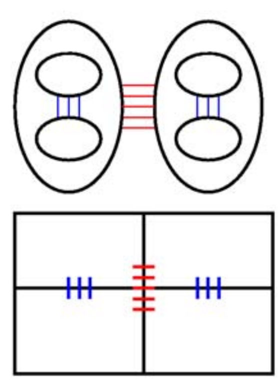
Previously

- Used Rent's Rule characterization to understand wire growth

$$IO = c\, N^p$$

- Top bisections will be $\Omega(N^p)$
- 2D wiring area

$$\Omega(N^p) \times \Omega(N^p) = \Omega(N^{2p})$$



We Know

- How we avoid $O(N^2)$ wire growth for "typical" designs
- How to characterize locality
- How we might exploit that locality to reduce wire growth
- Wire growth implied by a characterized design

Today

- Switching
  - Implications
  - Options
- Multilayer metalization

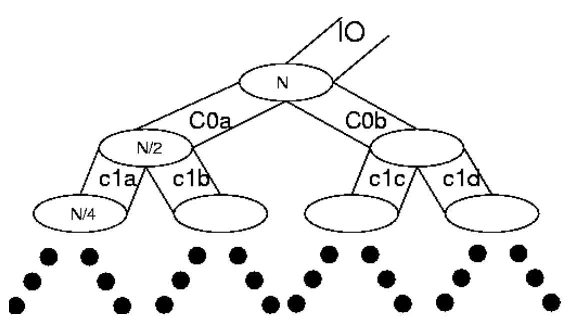
Switching:

How can we use the locality captured by Rent's Rule to reduce switching requirements? (How much?)

Observation

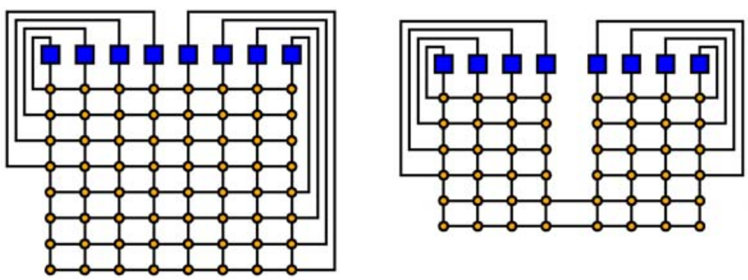
- Locality that saved us wiring,
  also saves us switching

$$IO = c\, N^{p}$$



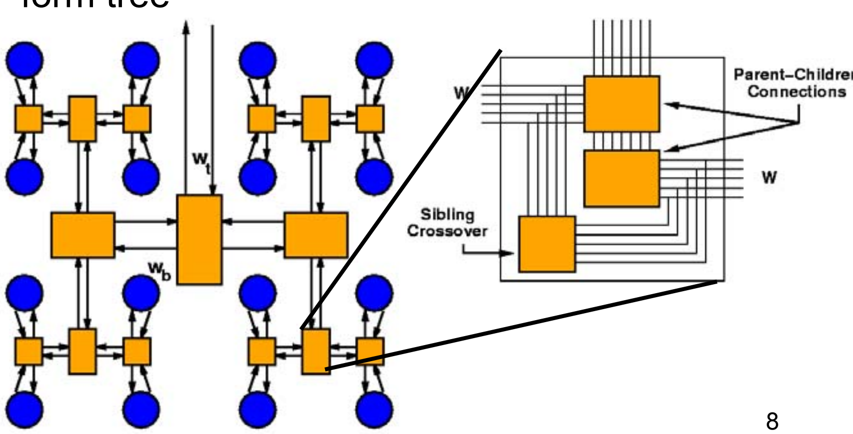
Consider

- Crossbar case to exploit wiring:
  - split into two halves, connect with limited wires
  - N/2 x N/2 crossbar each half
  - $N/2 \times (N/2)^p$ connect to bisection wires
  - $2(N^2/4) + 2(N/2)^{p+1} < N^2$



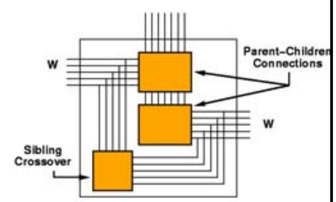
Recurse

- Repeat at each level
  – form tree



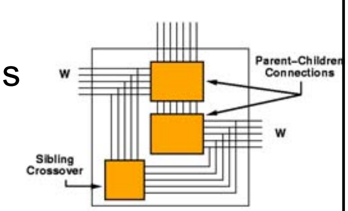
Result

- If use crossbar at each tree node
  - $O(N^{2p})$ wiring area
    - $p>0.5$, direct from bisection
  - $O(N^{2p})$ switches
    - top switch box is $O(N^{2p})$
    - switches at one level down is
      - $2\times(1/2^p)^2 \times$ previous level
      - $(2/2^{2p})=2^{(1-2p)}$
      - coefficient < 1 for $p>0.5$



Result

- If use crossbar at each tree node
  - $O(N^{2p})$ switches
    - top switch box is $O(N^{2p})$
    - switches at one level down is
      - $2^{(1-2p)} \times$ previous level
    - Total switches:
      - $N^{2p} \times (1+2^{(1-2p)}+2^{2(1-2p)}+2^{3(1-2p)}+\ldots)$
      - get geometric series; sums to $O(1)$
      - $N^{2p} \times (1/1-2^{(1-2p)})$
      - $= 2^{(2p-1)}/(2^{(2p-1)}-1) \times N^{2p}$



Good News

- Good news
  - asymptotically optimal
  - Even without switches area $O(N^{2p})$
    - so adding $O(N^{2p})$ switches not change

Bad News

- Switches area >> wire crossing area
  - Consider 6λ wire pitch $\Rightarrow$ crossing $36\ \lambda^2$
  - Typical (passive) switch $\Rightarrow$ $2500\ \lambda^2$
  - Passive only: 70x area difference
    - worse once rebuffer or latch signals.
- Switches limited to substrate
  - whereas can use additional metal layers for wiring area

Additional Structure

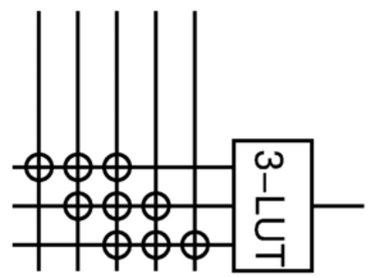
- This motivates us to look beyond crossbars
  - can depopulate crossbars on up-down connection without loss of functionality?

Can we do better?

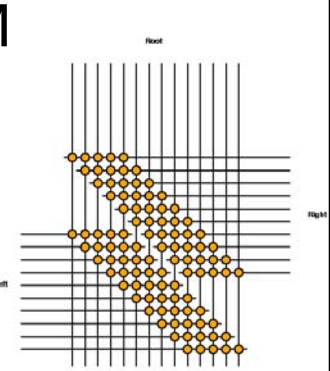
- Crossbar too powerful?
  - Does the specific down channel matter?
- What do we want to do?
  - Connect to *any* channel on lower level
  - Choose a subset of wires from upper level
    - order not important

N choose K

- Exploit freedom to depopulate switchbox
- Can do with:
  - $K\times(N-K+1)$ swtiches
  - Vs. $K \times N$
  - Save $\sim K^2$



N-choose-M

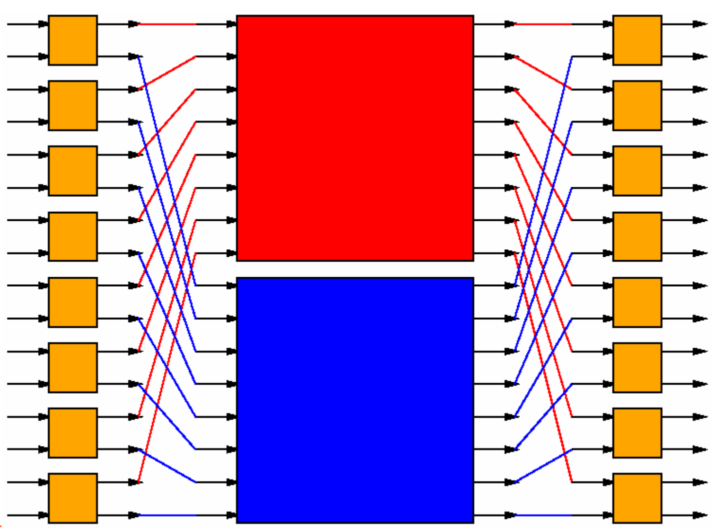


- Up-down connections
  - only require concentration
    - choose M things out of N
  - *i.e.* order of subset irrelevant
- Consequent:
  - can save a constant factor ~ $2^p/(2^p-1)$
    - $(N/2)^p \times N^p$ vs $(N^p - (N/2)^p+1)(N/2)^p$
    - $P=2/3 \rightarrow 2^p/(2^p-1) \approx 2.7$
- Similary, Left-Right
  - order not important $\Rightarrow$ reduces switches

Multistage Switching

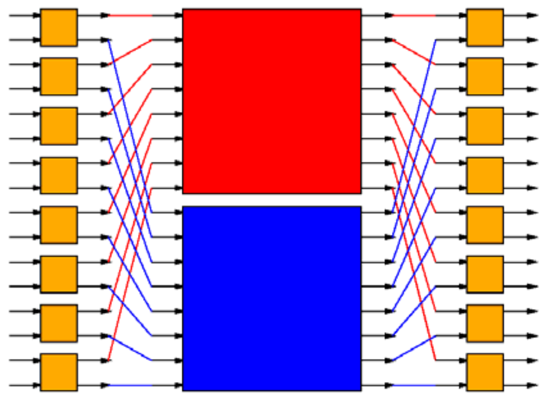
- Can route any **permutation** w/ less switches than a crossbar
- If we allow switching in stages
  - Trade increase in switches in path
  - For decrease in total switches

Decomposition



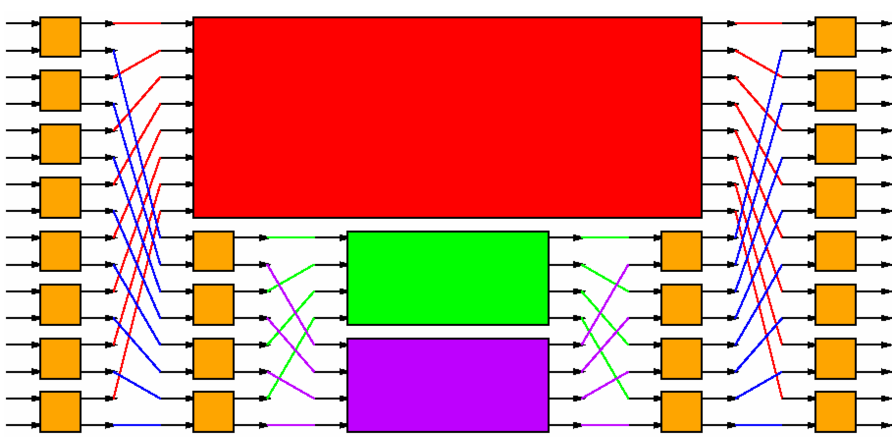
Decomposition

- Switches: $N/2 \times 2 \times 4 + (N/2)^2 < N^2$



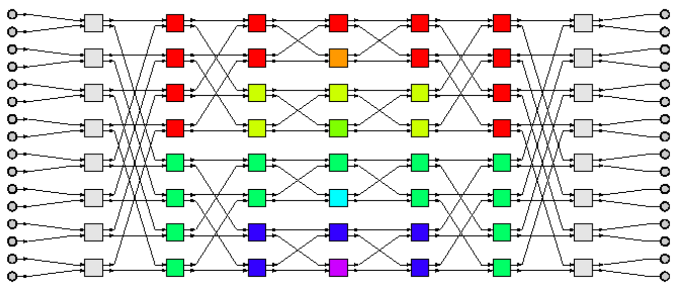
Recurse

If it works once, try it again…



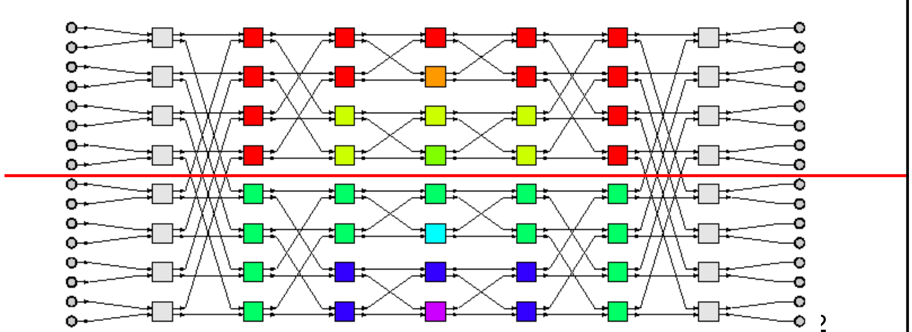
Result: Beneš Network

- $2\log_2(N)-1$ stages (switches in path)
- Of N/2 2×2 switchpoints (4 switches)
- $4N\times\log_2(N)$ total switches
- Compute route in O(N log(N)) time



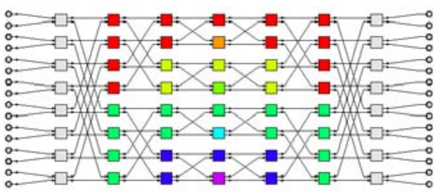
Beneš Network Wiring

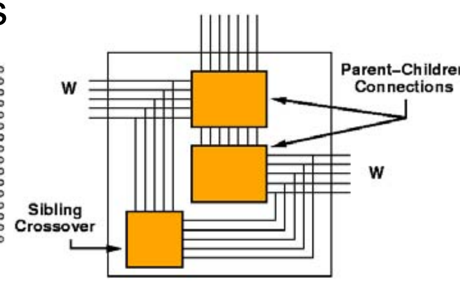
- Bisection: N
- Wiring → $O(N^2)$ area (fixed wire layers)



Beneš Switching

- Beneš reduced switches
  - N^2 to N(log(N))
  - using multistage network
- Replace crossbars in tree with Beneš switching networks





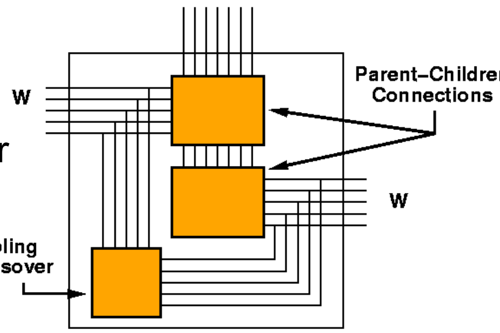
Beneš Switching

- Implication of Beneš Switching
  - still require $O(W^2)$ wiring per tree node
    - or a total of $O(N^{2p})$ wiring
  - now O(W log(W)) switches per tree node
    - converges to O(N) total switches!
  - $O(\log^2(N))$ switches in path across network
    - strictly speaking, dominated by wire delay $\sim O(N^p)$
    - but constants make of little practical interest except for very large networks ☹

Better yet…

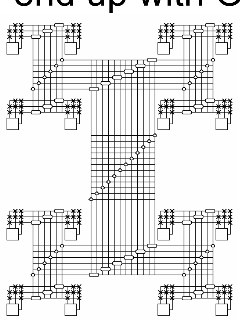
- Believe do not need Beneš on the up paths
- Single switch on up path
- Beneš for crossover
- Switches in path:

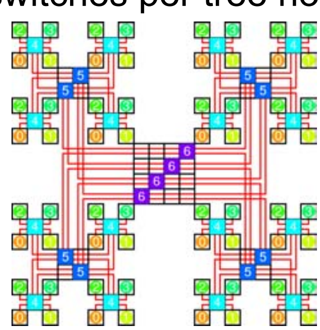
  log(N) up
  + log(N) down
  + 2log(N) crossover
  = 4 log(N)
  = O(log (N))



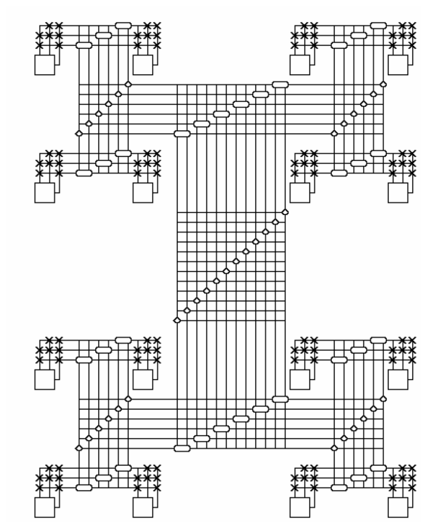
Linear Switch Population

- Can further reduce switches
  - connect each lower channel to O(1) channels in each tree node
  - end up with O(W) switches per tree node





Linear Switch (p=0.5)



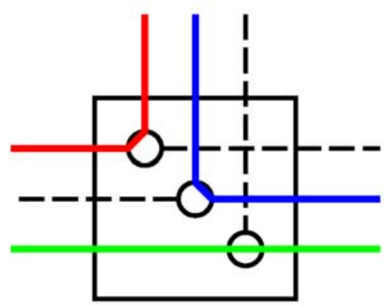
Linear Consequences: Good News

- Linear Switches
  - O(log(N)) switches in path
  - $O(N^{2p})$ wire area
  - O(N) switches
  - More practical than Beneš case

Linear Consequences: Bad News

- Lacks guarantee can use all wires
  - as shown, at least mapping ratio > 1
  - likely cases where even **constant** not suffice
    - expect no worse than logarithmic
    - **open** to establish tight lower bound for **any** linear arrangement
- Finding Routes is harder
  - no longer linear time, deterministic
  - **open** as to exactly how hard

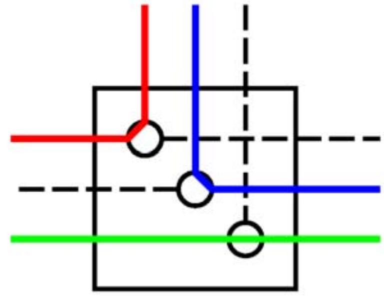
Mapping Ratio



- Mapping ratio says
  - if I have W channels
    - may only be able to use W/MR wires
      - for a particular design's connection pattern
    - to accommodate any design
      - forall channels

        physical wires $\geq$ MR $\times$ logical

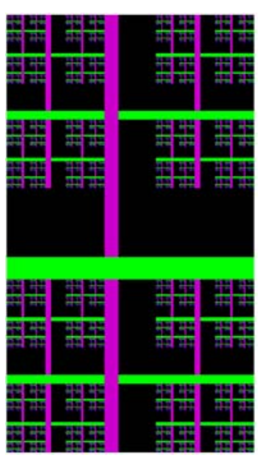
Mapping Ratio



- Example:
  - Shows MR=3/2
  - For Linear Population, 1:1 switchbox

Area Comparison

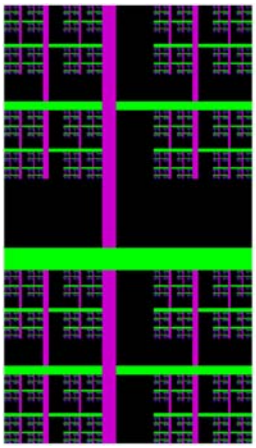
Both:
p=0.67
N=1024



M-choose-N perfect map

Linear MR=2

Area Comparison

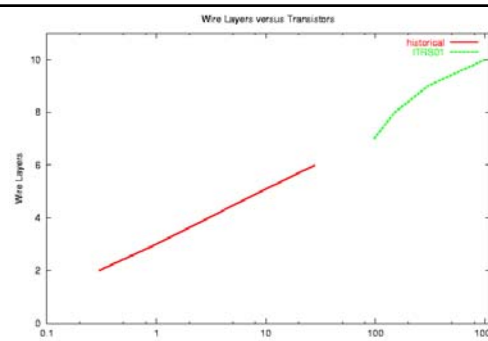


M-choose-N perfect map

Linear MR=2

- Since
  - switch >> wire
- may be able to tolerate MR>1
- reduces switches
  - net area savings
- Open:
  - Prove any constant mapping ratio
- Empirical:
  - Never seen greater than 1.5

Multi-layer metal?



- Preceding assumed
  - fixed wire layers
- In practice,
  - increasing wire layers with shrinking tech.
  - Increasing wire layers with chip capacity
    - wire layer growth ~ O(log(N))

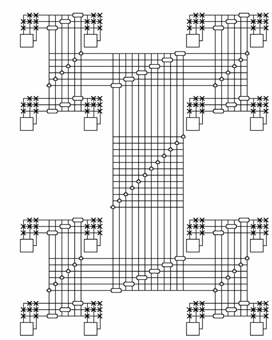
Multi-Layer

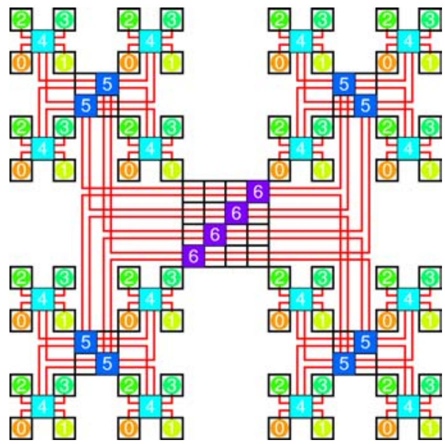
- Natural response to $\Omega(N^{2p})$ wire layers
  - Given N^p wires in bisection
    - rather than accept N^p width
      - use $N^{(p-0.5)}$ layers
      - accommodate in $N^{0.5}$ width
    - now wiring takes $\Omega(N)$ 2D area
      - with $N^{(p-0.5)}$ wire layers
    - for p=0.5,
      - log(N) layers to accommodate wiring

Linear + Multilayer

- Multilayer says can do in $\Omega(N)$ 2D-area
- Switches require 2D-area
  - more than O(N) switches would make switches dominate
  - Linear and Beneš have O(N) switches
- There's a possibility can achieve O(N) area
  - with multilayer metal and linear population

Fold and Squash Layout

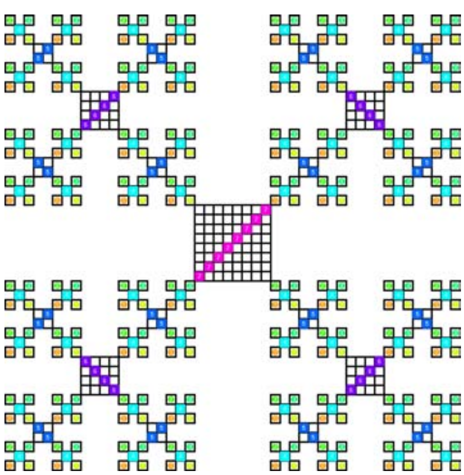


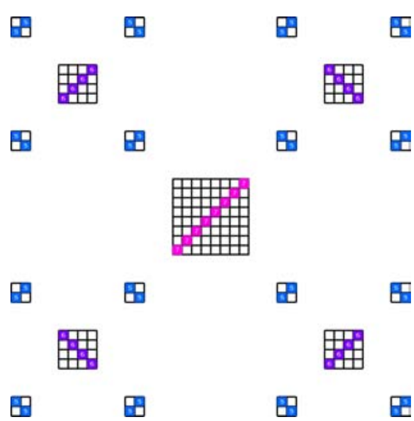


Caveat: this formulation for p=0.5 …

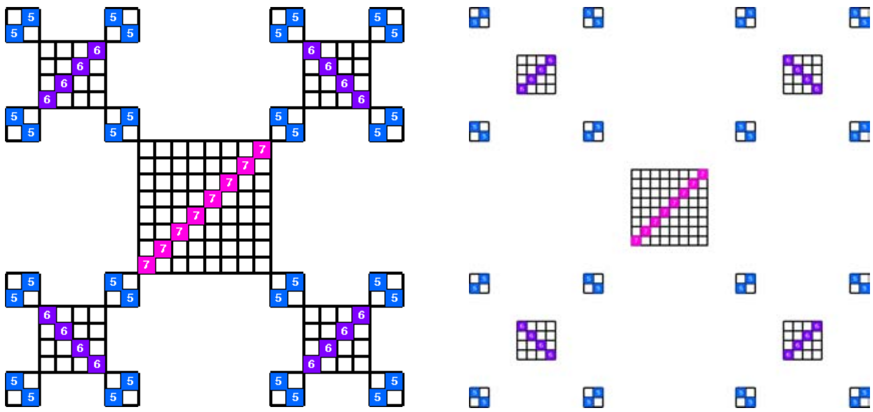
Maybe p>0.5 on Day16

Fold and Squash Layout



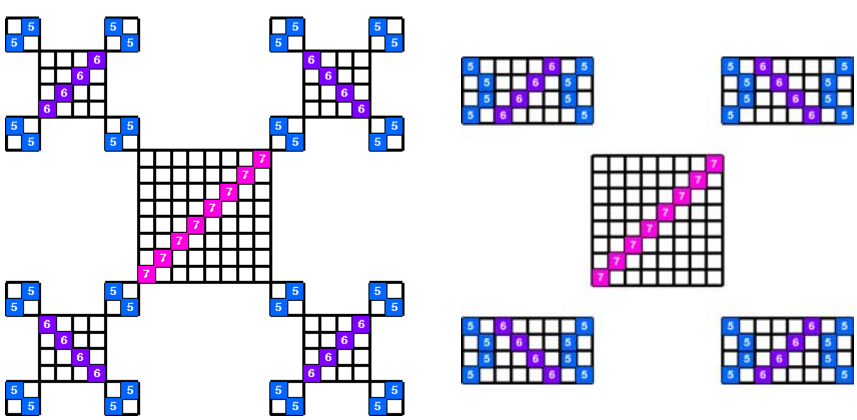


Folding



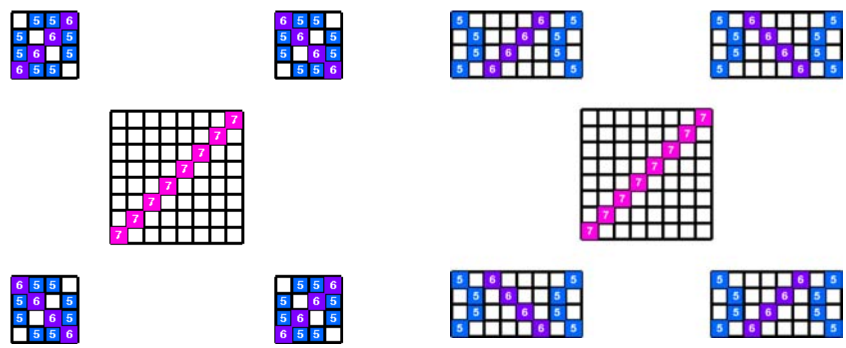
39

Folding

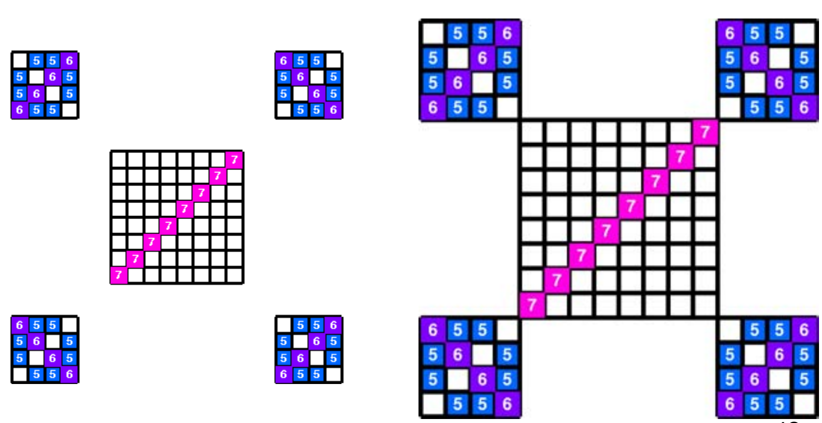


40

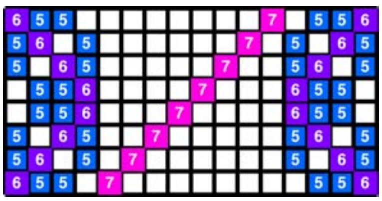
Folding

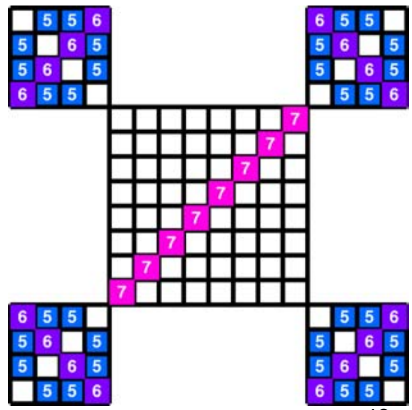


Folding

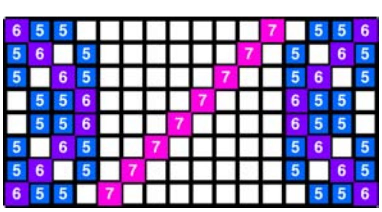


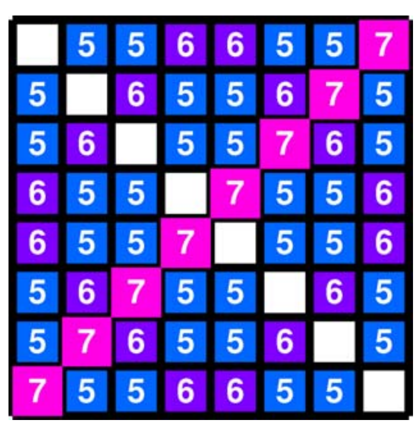
Folding





Folding





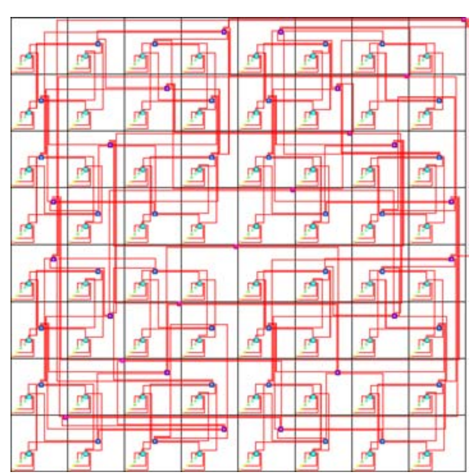
Folding Invariants

- Lower folds leave both diagonals free
- Current level consumes one, leaving other free

	5	5	6	6	5	5	7
5		6	5	5	6	7	5
5	6		5	5	7	6	5
6	5	5		7	5	5	6
6	5	5	7		5	5	6
5	6	7	5	5		6	5
5	7	6	5	5	6		5
7	5	5	6	6	5	5	

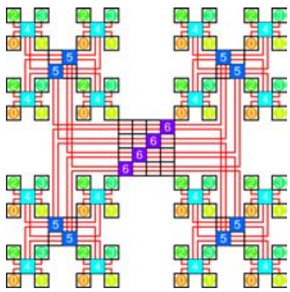
Compact Folded Layout

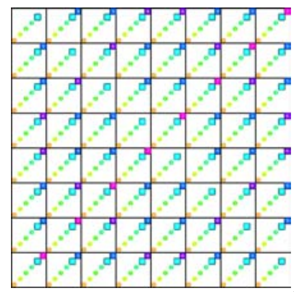
- Can contain switches to constant area
- Wires still grow faster than linear
- Can use extra wire layers to accommodate wire growth
- (whereas switches not helped by additional wire layers)



Fold and Squash Result

- Can layout BFT
  - in O(N) 2D area
  - with O(log(N)) wiring layers





Big Ideas [MSB Ideas]

- In addition to wires, must have switches
  - Have significant area and delay
- Rent's Rule locality reduces
  - both wiring and switching requirements
- Naïve switches match wires at $O(N^{2p})$
  - switch area >> wire area
  - prevent benefit from multiple layers of metal

Big Ideas [MSB Ideas]

- Can achieve O(N) switches
  - plausibly O(N) area with sufficient metal layers
- Switchbox depopulation
  - save considerably on area (delay)
  - will waste wires
  - May still come out ahead (evidence to date)
  - So far:
    - routing no longer guaranteed
    - routing becomes NP-complete?