ESE680-002 (ESE534): Computer Organization

Day 18: March 21, 2007

Interconnect 6: MoT

Previously

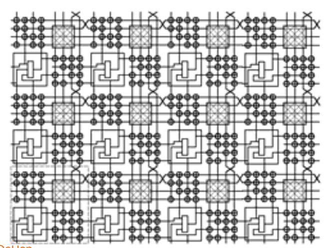
- HSRA/BFT – natural hierarchical network
  - Switches scale O(N)
- Mesh – natural 2D network
  - Switches scale $\Omega(N^{p+0.5})$

Today

- Good Mesh properties
- HSRA vs. Mesh
- MoT
- Grand unified network theory ☺
  - MoT vs. HSRA
  - MoT vs. Mesh

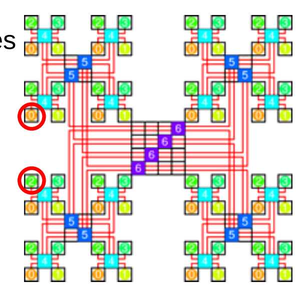
Mesh

1. Wire delay can be Manhattan Distance
2. Network provides Manhattan Distance route from source to sink



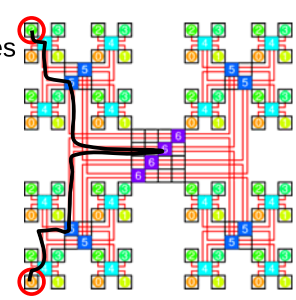
HSRA/BFT

- Physical locality does not imply logical closeness



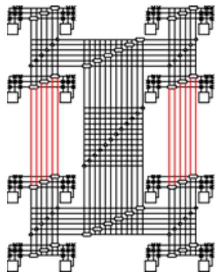
HSRA/BFT

- Physical locality does not imply logical closeness
- May have to route twice the Manhattan distance



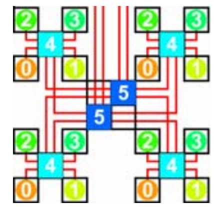
Tree Shortcuts

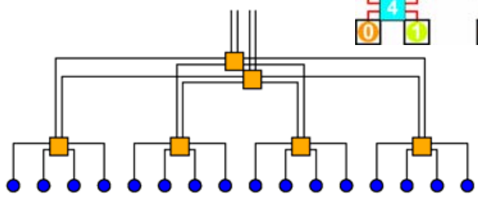
- Add to make physically local things also logically local
- Now wire delay always proportional to Manhattan distance
- May still be 2× longer wires



BFT/HSRA ~ 1D

- Essentially one-dimensional tree
  - Laid out well in 2D

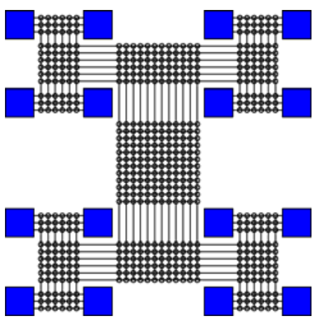




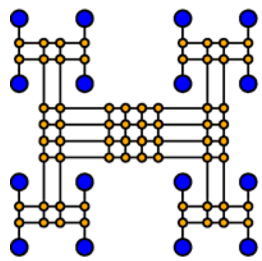
Consider Full Population Tree

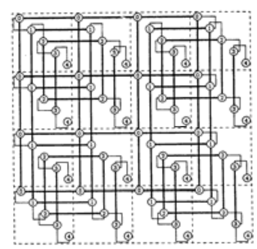
ToM

Tree of Meshes

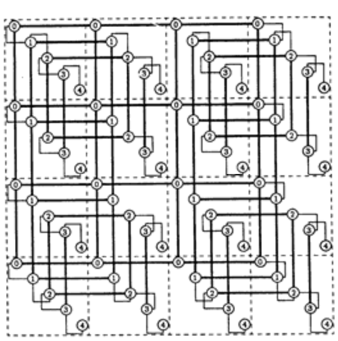


Can Fold Up





Gives Uniform Channels

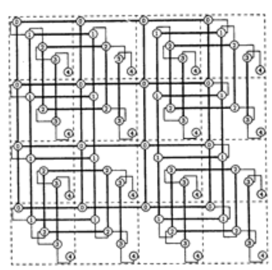


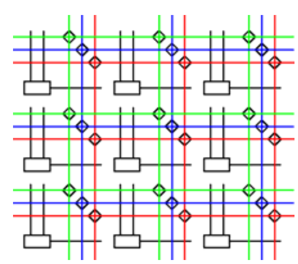
Works nicely p=0.5

Channels log(N)

[Greenberg and Leiserson, *Appl. Math Lett.* v1n2p171, 1988]

Gives Uniform Channels





(and add shortcuts)

How wide are channels?

$$W = \left[\frac{w(N)+w(N/2)}{\sqrt{N}}\right] + \left[\frac{w(N/4)+w(N/8)}{\sqrt{\frac{N}{4}}}\right] + \ldots$$

$$w(N) = c\,N^{p}$$

$$W = \left(\frac{c\,N^{p}}{\sqrt{N}}\right)\times\left(1+2^{-p}+\left(1+2^{-p}\right)\times 2\times 2^{-2p}+\ldots\right)$$

How wide are channels?

$$W = \left(\frac{c\,N^{p}}{\sqrt{N}}\right)\times\left(1+2^{-p}+\left(1+2^{-p}\right)\times 2\times 2^{-2p}+\ldots\right)$$

$$W = \left(c\,N^{p-0.5}\right)\times\left(1+2^{-p}\right)\times\left(1+2^{1-2p}+2^{2\times(1-2p)}+\ldots\right)$$

$$W = \left(c\,N^{p-0.5}\right)\times\left(1+2^{-p}\right)\times\left(\frac{1}{1-2^{1-2p}}\right)$$

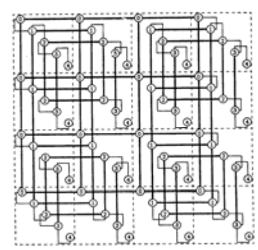
How wide are channels?

$$W = \left(c\,N^{p-0.5}\right)\times\left(\frac{1+2^{-p}}{1-2^{1-2p}}\right)$$

- A constant factor wider than lower bound!
- P=2/3 → ~8
- P=3/4 → ~5.5

Implications

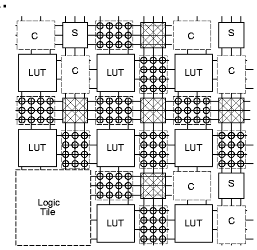
- Tree never requires more than constant factor more wires than mesh
  - Even w/ the non-minimal length routes
  - Even w/out shortcuts
- Mesh global route upper bound channel width is $O(N^{p-0.5})$
  - Can always use fold-squash tree as the route
  - Matches lower bound!



MoT

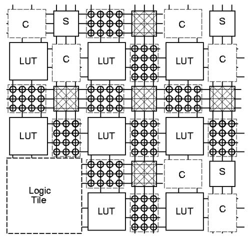
Recall: Mesh Switches

- Switches per switchbox:
  - $6w/L_{seg}$
- Switches into network:
  - (K+1) w
- Switches per PE:
  - $6w/L_{seg} + Fc\times(K+1)\,w$
  - $w = cN^{p-0.5}$
  - Total $\propto N^{p-0.5}$
- Total Switches: N*(Sw/PE) $\propto N^{p+0.5} > N$



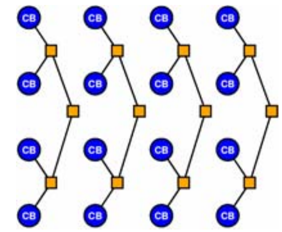
Recall: Mesh Switches

- Switches per PE:
  - $6w/L_{seg} + Fc\times(K+1)\ w$
  - $w = cN^{p-0.5}$
  - Total $\propto N^{p-0.5}$
- Not change for
  - Any constant Fc
  - Any constant L_{seg}



Mesh of Trees

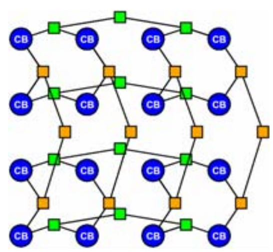
- Hierarchical Mesh
- Build Tree in each column



[Leighton/FOCS 1981]

Mesh of Trees

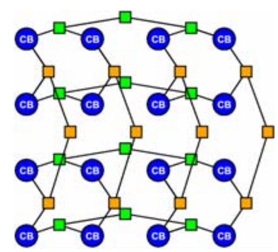
- Hierarchical Mesh
- Build Tree in each column
- …and each row



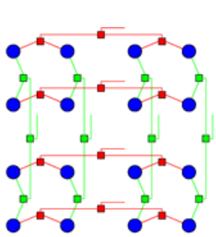
[Leighton/FOCS 1981]

Mesh of Trees

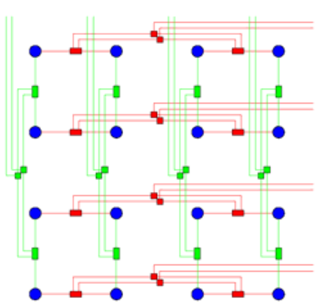
- More natural 2D structure
- Maybe match 2D structure better?
  - Don't have to route out of way



MoT Parameterization: P



P=0.5

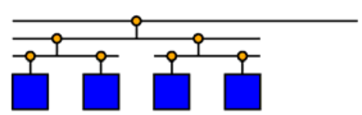


P=0.75

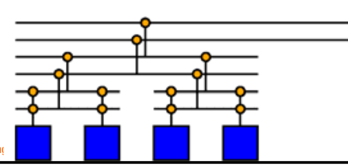
MoT Parameterization

- Support C with additional trees
  - (like BFT)

C=1

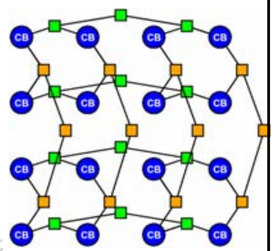


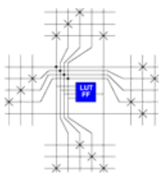
C=2



Mesh of Trees

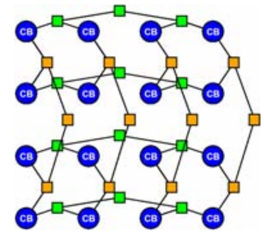
- Logic Blocks
  - Only connect at leaves of tree
- Connect to the C trees
  - Per side
  - 4C total





Switches

- Total Tree switches
  - 2 C√N×(switches/tree)
  - 2={X,Y}
  - C per Row and Col.

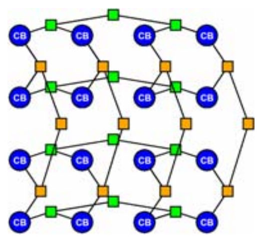


Switches

- Total Tree switches
  - 2 C√N× (switches/tree)
- Sw/Tree:

$$\left(\frac{\sqrt{N}}{2}\right)\times\left(1+\frac{2^{p-0.5}}{2}+\left(\frac{2^{p-0.5}}{2}\right)^2+\ldots\right)$$

$$\left(\frac{\sqrt{N}}{2}\right)\times\left(\frac{1}{1-2^{p-1.5}}\right)$$

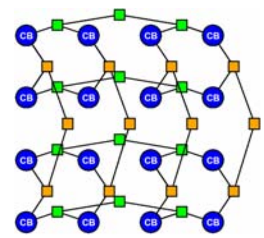


Switches

- Total Tree switches
  - 2 C √N× (switches/tree)
- Sw/Tree:

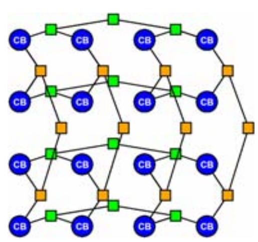
$$\left(\frac{\sqrt{N}}{2}\right)\times\left(\frac{1}{1-2^{p-1.5}}\right)$$

$$TreeSwitches=\left(\frac{C\times N}{1-2^{p-1.5}}\right)=O(N)$$



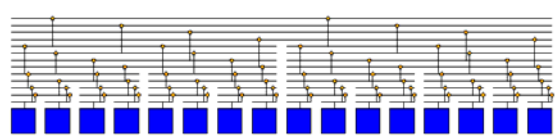
Switches

- Only connect to leaves of tree
- C×(K+1) switches per leaf
- Total switches
  - Leaf + Tree
  - O(N)



Wires

- **Design:** O(N^p) in top level
- Total wire width of channels: O(N^p)
  - Another geometric sum
- No detail route guarantee (at present)
  - Likely amenable to expander design (Day16)



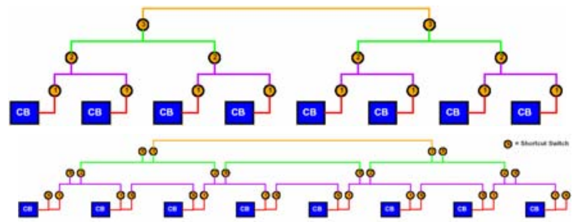
Empirical Results

- **Benchmark:** Toronto 20
- Compare to L_{seg}=1, L_{seg}=4
  - CLMA ~ 8K LUTs
    - Mesh(L_{seg}=4): w=14 → 122 switches/LB
    - MoT(p=0.67): C=4 → 89 switches/LB
  - Benchmark wide: 10% less
    - CLMA largest
    - Asymptotic advantage

[Rubin,DeHon/FPGA2003]

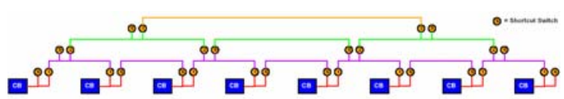
Shortcuts

- Strict Tree
  - Same problem with physically far, logically close



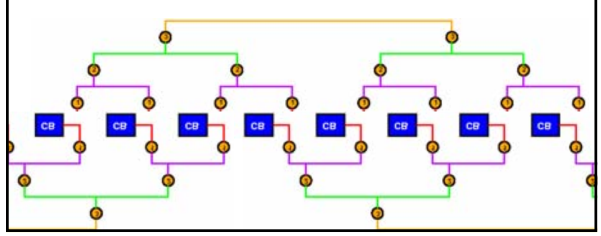
Shortcuts

- Empirical
  - Shortcuts reduce C
  - But net increase in total switches



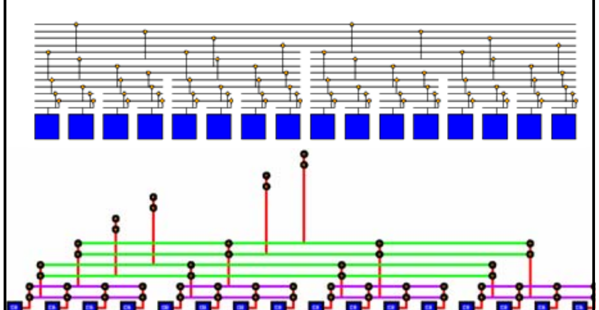
Staggering

- With multiple Trees
  - Offset relative to each other
  - Avoids worst-case discrete breaks
  - One reason don't benefit from shortcuts



Flattening

- Can use arity other than two



Overall 26% fewer than mesh

TABLE V
TOTAL SWITCHES VERSUS ARITY AND RENT EXPONENT (p)

arity	2		3		4			5	8
p	0.67	0.75	0.65	0.81	0.625	0.67	0.75	0.60	0.67
alu4	86	101	88	94	95	88	74	91	90
apex2	106	98	88	95	109	87	91	105	103
apex4	110	129	108	99	113	89	95	123	104
bigkey	62	72	52	71	62	51	54	60	58
clma	103	96	86	99	106	85	93	100	103
des	65	79	69	69	63	71	60	58	61
diffeq	88	77	71	72	64	71	75	61	76
dsip	62	72	70	71	62	68	54	60	58
elliptic	82	93	85	91	91	84	71	88	86
ex1010	107	102	84	90	93	88	98	101	106
ex5p	113	106	108	99	114	90	96	123	107
frisc	103	94	85	91	91	85	89	103	101
misex3	108	100	89	97	95	87	93	106	88
pdc	128	128	118	112	124	124	117	144	136
s298	84	73	70	71	62	69	72	74	73
s38417	84	75	70	77	61	69	76	72	74
s38584.1	84	100	70	77	77	69	76	72	74
seq	107	98	104	93	93	86	91	105	104
spla	123	117	101	91	106	101	106	117	114
tseng	90	80	72	74	65	72	77	62	76
max	128	129	118	112	124	124	117	144	136
sum	1895	1890	1688	1733	1746	1634	1658	1825	1792

[Rubin&DeHon/TRVLSI2004]

TABLE VI

TREE DOMAINS (C) AND WIRES PER CHANNEL (W_x) VERSUS ARITY AND RENT EXPONENT (p)

arity	2				3				4						5		8	
p	0.67		0.75		0.65		0.81		0.625		0.67		0.75		0.60		0.67	
property	C	W_x	C	W_x	C	W_x	C	W_x	C	W_x	C	W_x	C	W_x	C	W_x	C	W_x
alu4	4	76	4	116	5	65	4	124	6	54	5	55	4	60	6	54	6	42
apex2	5	95	4	116	5	65	4	124	7	63	5	55	5	75	7	63	7	49
apex4	5	95	5	145	6	78	4	124	7	63	5	55	5	75	8	72	7	49
bigkey	3	57	3	87	3	39	3	93	4	36	3	33	3	45	4	36	4	28
clma	5	135	4	180	5	105	4	252	7	91	5	95	5	155	7	63	7	105
des	3	81	3	135	4	52	3	93	4	52	4	76	3	93	4	36	4	60
diffeq	4	76	3	87	4	52	3	93	4	36	4	44	4	60	4	36	5	35
dsip	3	57	3	87	4	52	3	93	4	36	4	44	3	45	4	36	4	28
elliptic	4	76	4	116	5	65	4	124	6	54	5	55	4	60	6	54	6	42
ex1010	5	135	4	180	5	65	4	124	6	78	5	95	5	155	7	63	7	105
ex5p	5	95	4	116	6	78	4	124	7	63	5	55	5	75	8	72	7	49
frisc	5	95	4	116	5	65	4	124	6	54	5	55	5	75	7	63	7	49
misex3	5	95	4	116	5	65	4	124	6	54	5	55	5	75	7	63	6	42
pdc	6	162	5	225	7	91	5	155	8	104	7	133	6	186	10	90	9	135
s298	4	76	3	87	4	52	3	93	4	36	4	44	4	60	5	45	5	35
s38417	4	108	3	135	4	84	3	189	4	52	4	76	4	124	5	45	5	75
s38584.1	4	108	4	180	4	84	3	189	5	65	4	76	4	124	5	45	5	75
seq	5	95	4	116	6	78	4	124	6	54	5	55	5	75	7	63	7	49
spla	6	114	5	145	6	78	4	124	7	63	6	66	6	90	8	72	8	56
tseng	4	76	3	87	4	52	3	93	4	36	4	44	4	60	4	36	5	35
max	6	162	5	225	7	105	5	252	8	104	7	133	6	186	10	90	9	135
sum	89	1907	76	2572	97	1365	73	2583	112	1144	94	1266	89	1767	123	1107	121	1143

Arity 5 → 42% fewer wires than arity 2

[Rubin&DeHon/TRVLSI2004]

MoT Parameters

- Shortcuts
- Staggering
- Corner Turns – to come
- Arity

Day 6

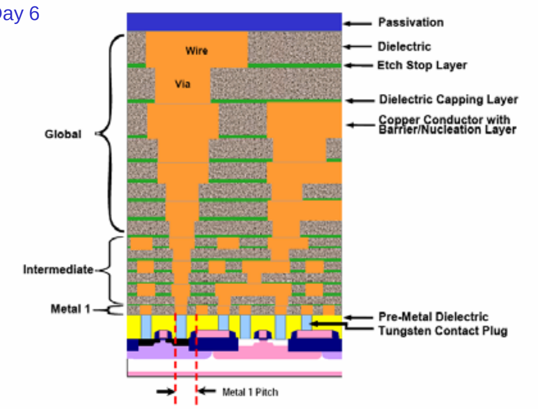
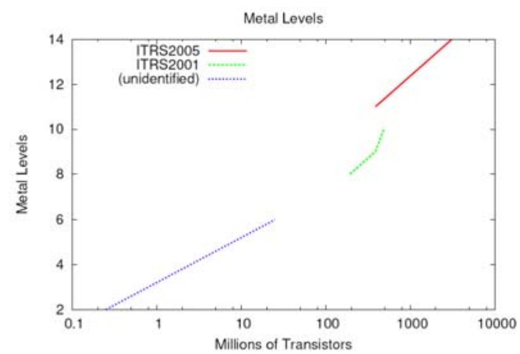


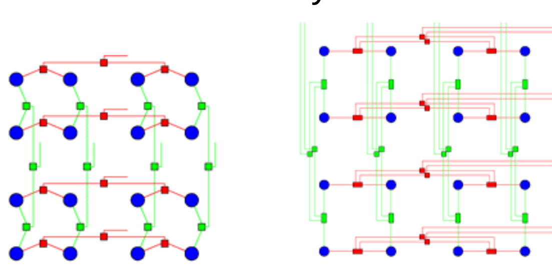
Figure 70 Cross-section of Hierarchical Scaling—MPU Device

Day 6

Wire Layers = More Wiring

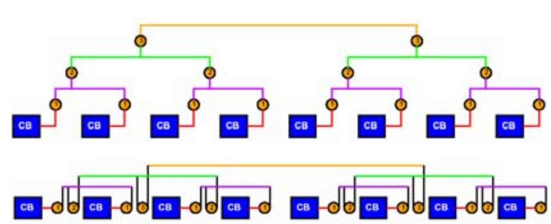


MoT Layout



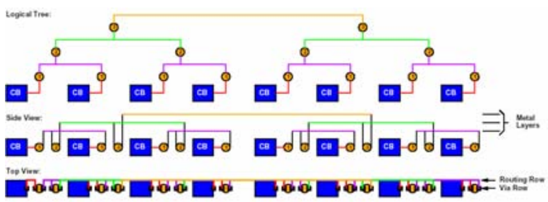
Main issue is layout 1D trees in multilayer metal

Row/Column Layout

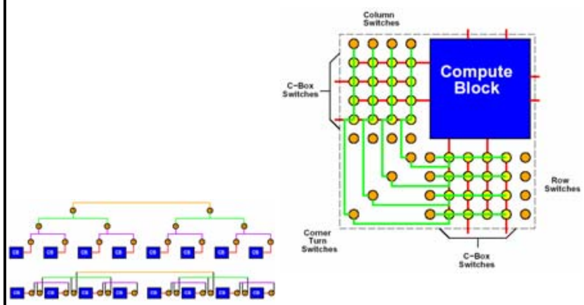


Geometric Progression → does not saturate via space!

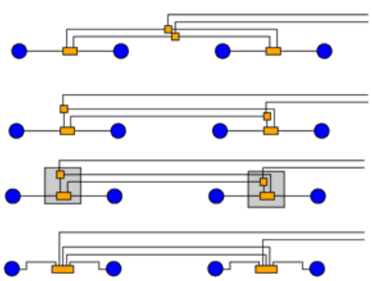
Row/Column Layout



Composite Logic Block Tile



P=0.75 Row/Column Layout



P=0.75 Row/Column Layout

MoT Layout

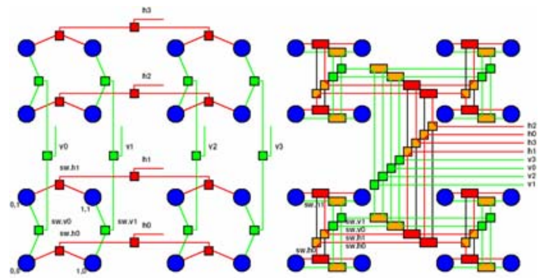
- Easily laid out in Multiple metal layers
  - Minimal $O(N^{p-0.5})$ layers
- Contain constant switching area per LB
  - Even with p>0.5

Relation?

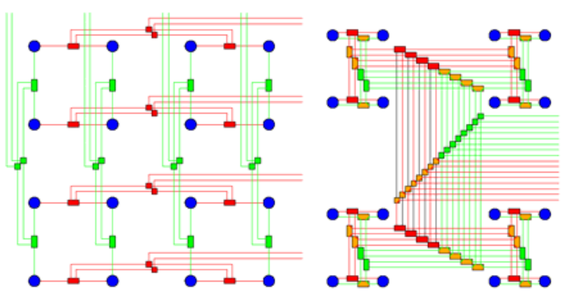
How Related?

- What lessons translate amongst networks?
- Once understand design space
  - Get closer together
- Ideally
  - One big network design we can parameterize

MoT → HSRA (P=0.5)

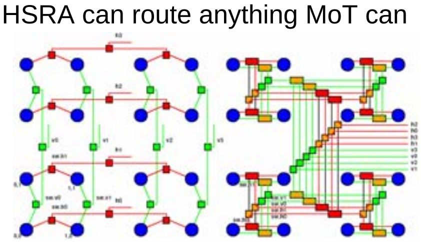


MoT→HSRA (p=0.75)

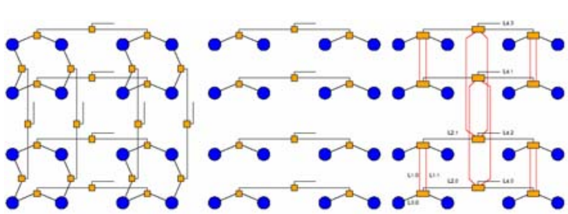


MoT → HSRA

- A C MoT maps directly onto a 2C HSRA
  - Same p's
- HSRA can route anything MoT can

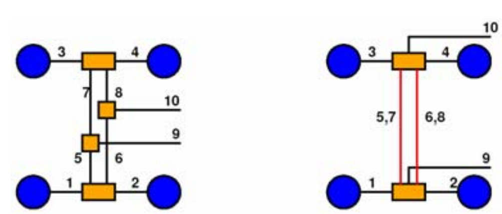


HSRA → MoT

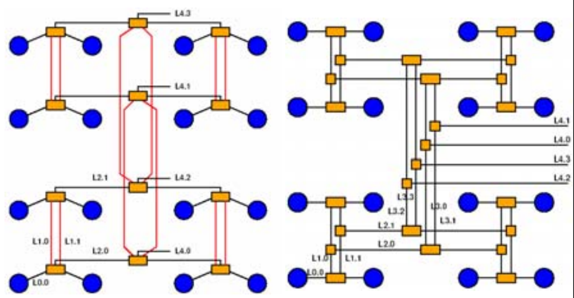


- Decompose and look at rows
- Add homogeneous, upper-level corner turns

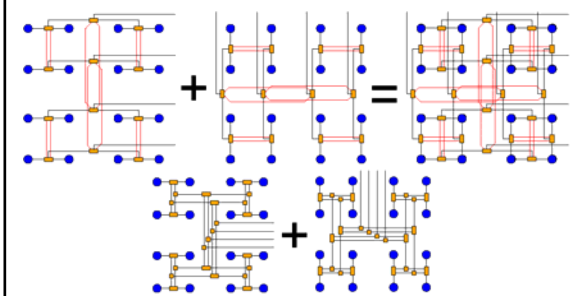
HSRA→MoT



HSRA→MoT

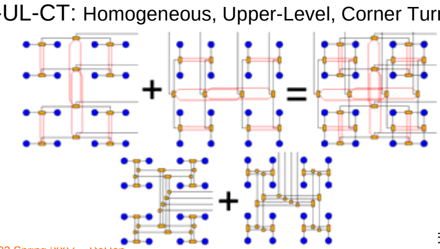


HSRA→MoT

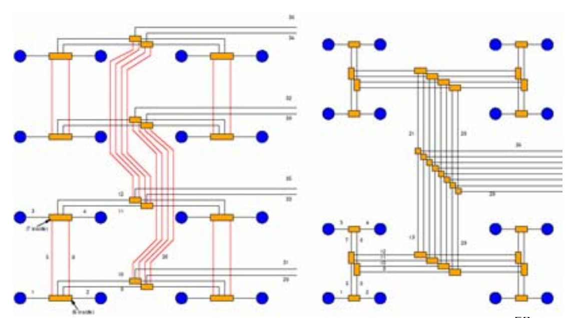


HSRA→MoT

- HSRA + HSRA^T = MoT w/ H-UL-CT
  - Same C, P
  - H-UL-CT: Homogeneous, Upper-Level, Corner Turns

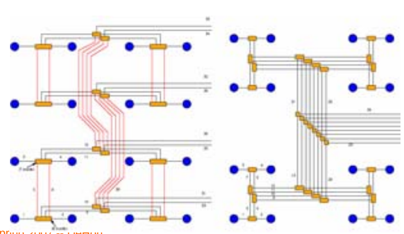


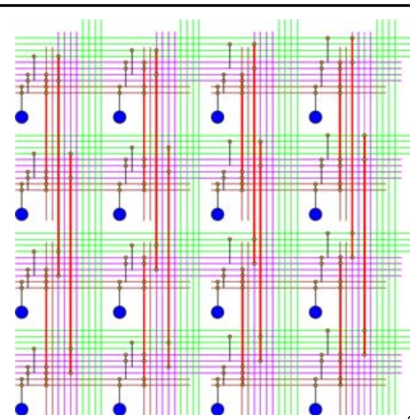
HSRA→MoT (p=0.75)



HSRA→MoT (p=0.75)

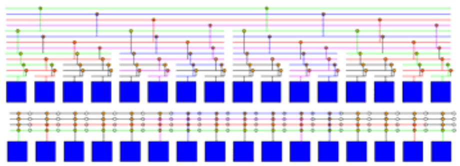
- Can organize HSRA as MoT
- P>0.5 MoT layout
  - Tells us how to layout p>0.5 HSRA





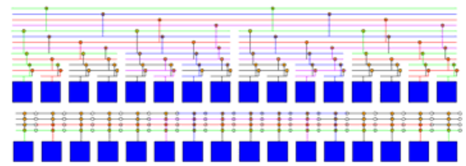
MoT vs. Mesh

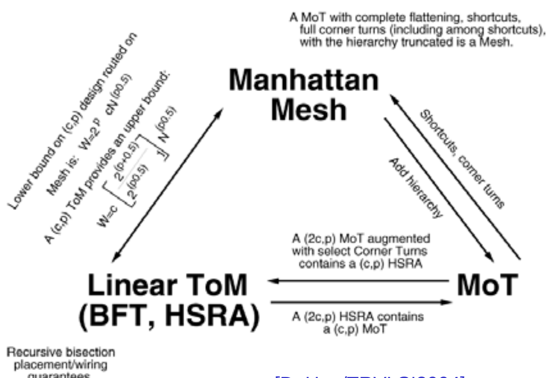
- MoT has Geometric Segment Lengths
- Mesh has flat connections
- MoT must climb tree
  – Parameterize w/ flattening
- MoT has $O(N^{p-0.5})$ fewer switches



MoT vs. Mesh

- Wires
  – Asymptotically the same (p>0.5)
  – Cases where Mesh requires constant less
  – Cases where require same number





[DeHon/TRVLSI2004]

Admin

- Interconnect assignment due today
- Retiming assignment out today
  – Monday lecture is key
  – Reading handed out last time for Monday

Big Ideas

- Networks driven by same wiring requirements
  – Have similar wiring asymptotes
- Can bound
  – Network differences
  – Worst-case mesh global routing
- Hierarchy structure allows to save switches
  – $O(N)$ vs. $\Omega(N^{p+0.5})$