FPGA Interconnect Planning

Amit Singh

Xilinx Inc.
San Jose, CA
amit.singh@xilinx.com

Malgorzata Marek-Sadowska

University of California, Santa Barbara
Santa Barbara, CA
mms@ece.ucsb.edu

Outline

- Introduction
- Previous Work
- Architecture Model
  - Architecture & Design Rent's exponent
- Clustering
  - Spatial regularity
- Fanout distribution
  - Area minimization
  - Area-delay minimization
  - Performance
- Conclusions

Introduction

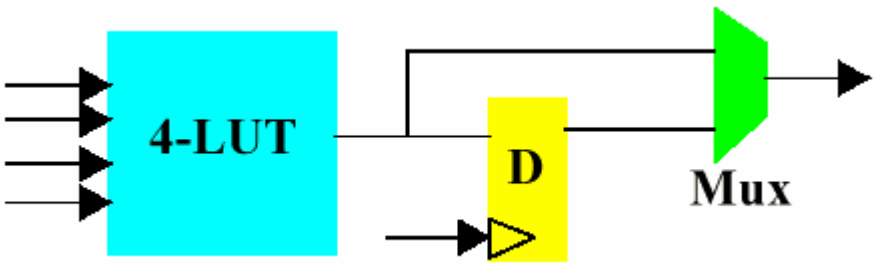
- Clustered FPGAs
  - System-on-Chip (SOC)
- Performance, Power, Area
  - Matching design and architecture complexities
  - Circuit packing/clustering
- Fanout distribution (Zarkesh-Ha et al. 2000)
  - Rent's rule
- Segment length planning in FPGAs
  - Area reduction
  - Area-Delay product minimization

Previous Work

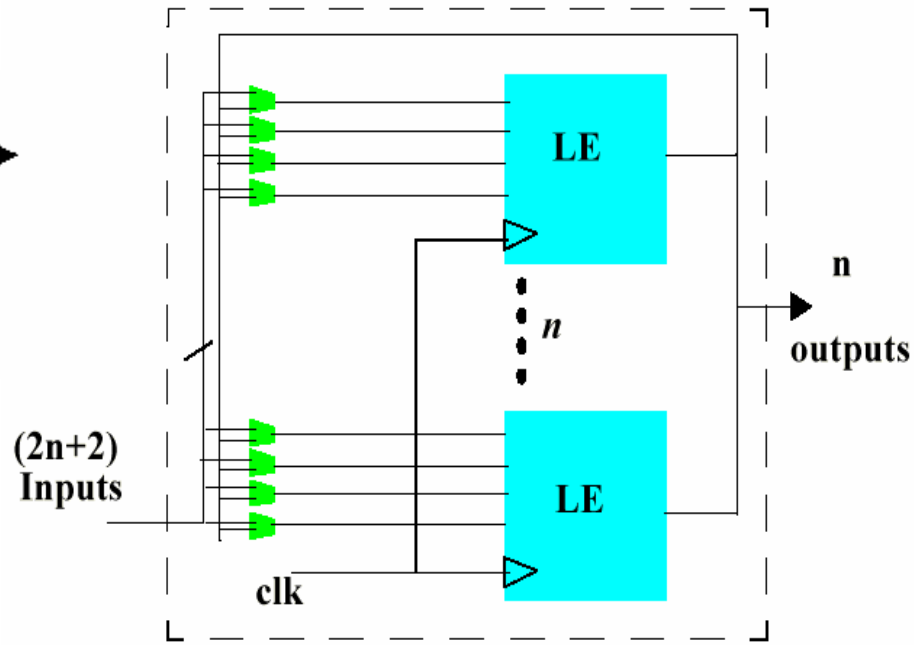
- Rent's rule
  - Landman, Russo (1974), Donath (1979)
- Interconnect distribution model
  - Davis et al. (1998), Zarkesh-Ha et al. (2000)
  - Local, semi-global, global wiring requirement
    - Homogeneous, heterogeneous systems
- FPGA segment length distribution
  - Betz et al. (1999)
  - Type of routing switches
  - Impact – segment length distribution on area-delay product
    - Best area-delay product: Mix of length 4 and 8 segments

Clustered FPGAs

- Clusters of LEs, connection boxes and switch-boxes
- Regular 2-D mesh array
- Example: Xilinx Virtex series, Altera APEX & Stratix

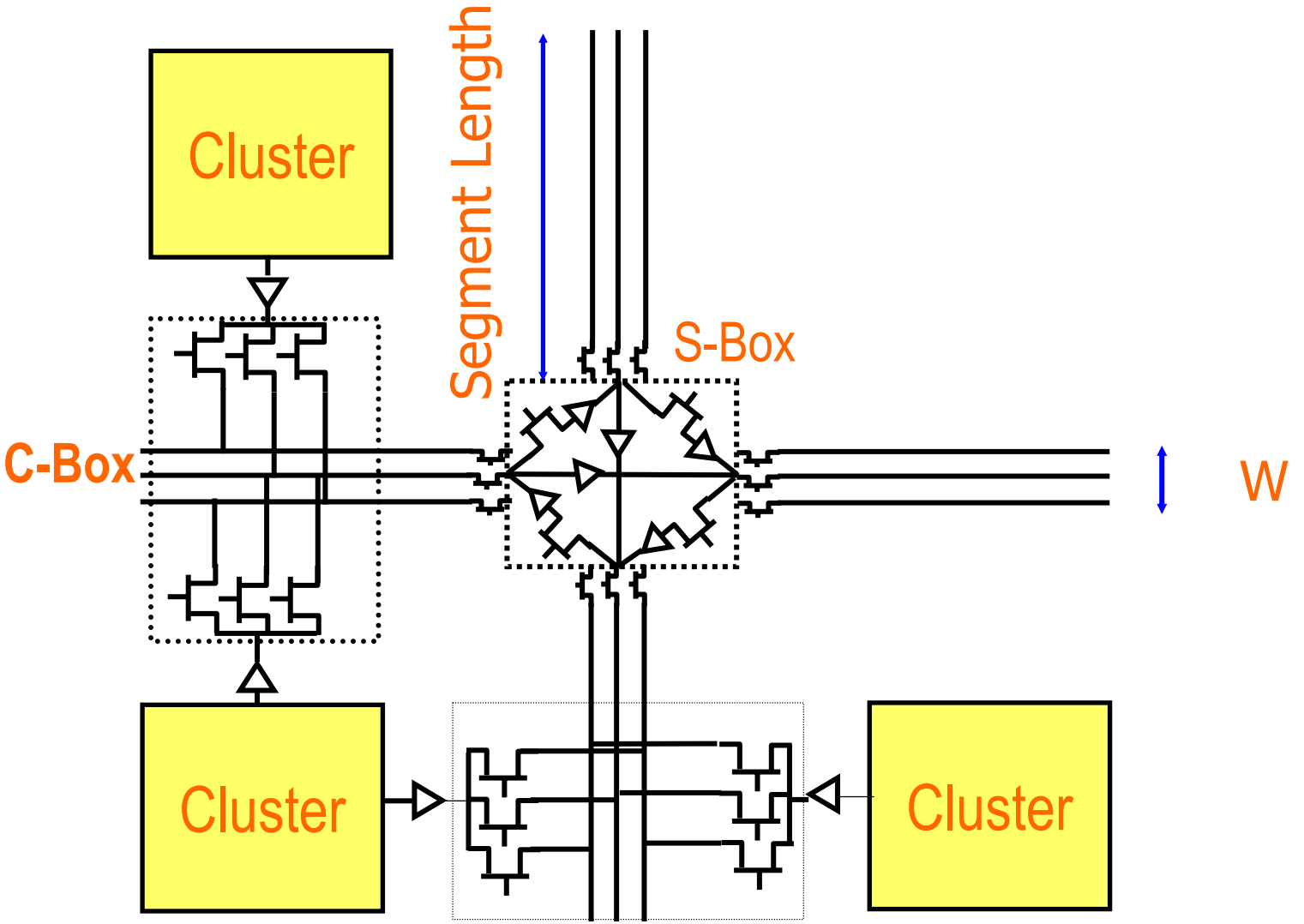


A Logic Element

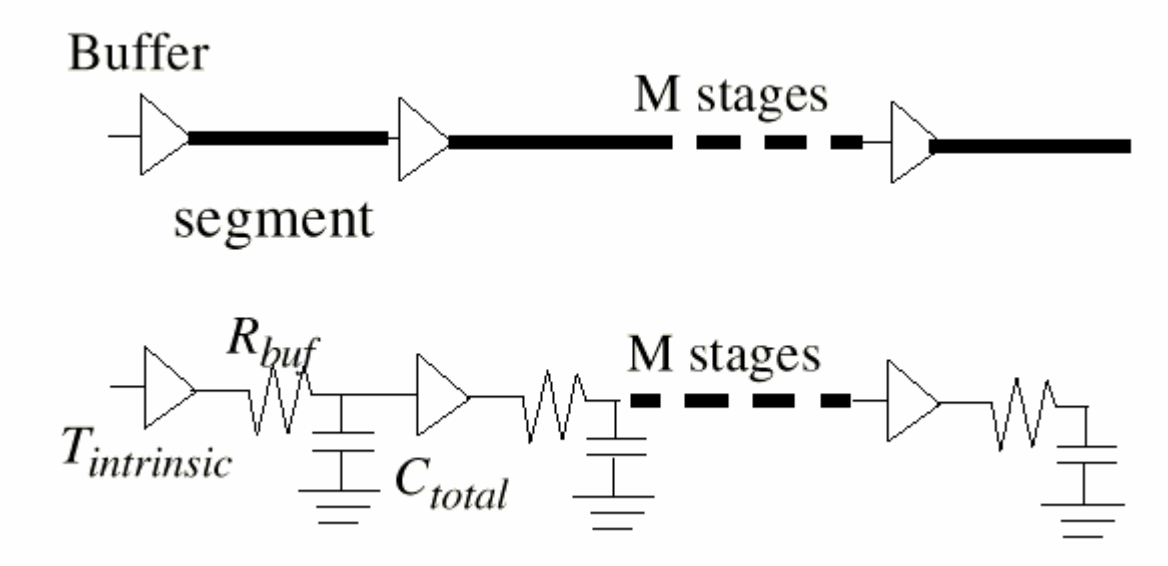


Cluster of LEs

Routing Model



Routing Model



Buffered routing switches

$$T_{buffer} = M \bullet [T_{intrinsic} + R_{buf}C_{total}]$$

Buffer chain delay

$$T = \sum_{i=1}^{M} iR_{pass}C_{total} = \frac{M}{2}(M+1)R_{pass}C_{total}$$

Pass-transistor chain delay

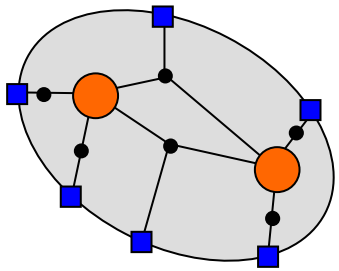
How much interconnect?

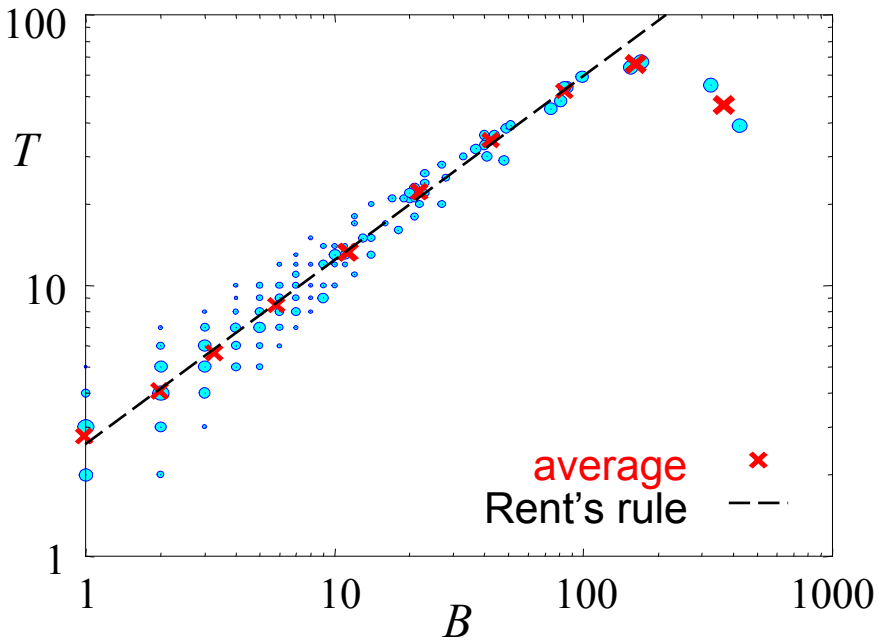
- ~80% of FPGA area = interconnects
- Routing resource utilization (RRU) is low
  - 100% logic utilization
- Depopulating logic clusters
  - Regularity
- Interconnect complexity guided fanout distribution-Rent's rule
  - Average fanout
    - Segment lengths (shorter segments or longer segments?)
    - Switch type (tri-state buffers or pass-transistor?)

Rent's Rule

Rent's rule: Landman and Russo in 1971.

Average number of terminals and blocks per module in a partitioned design:





$T = t\,B^{p}$

p = Rent exponent

$t \cong$ average # term./block

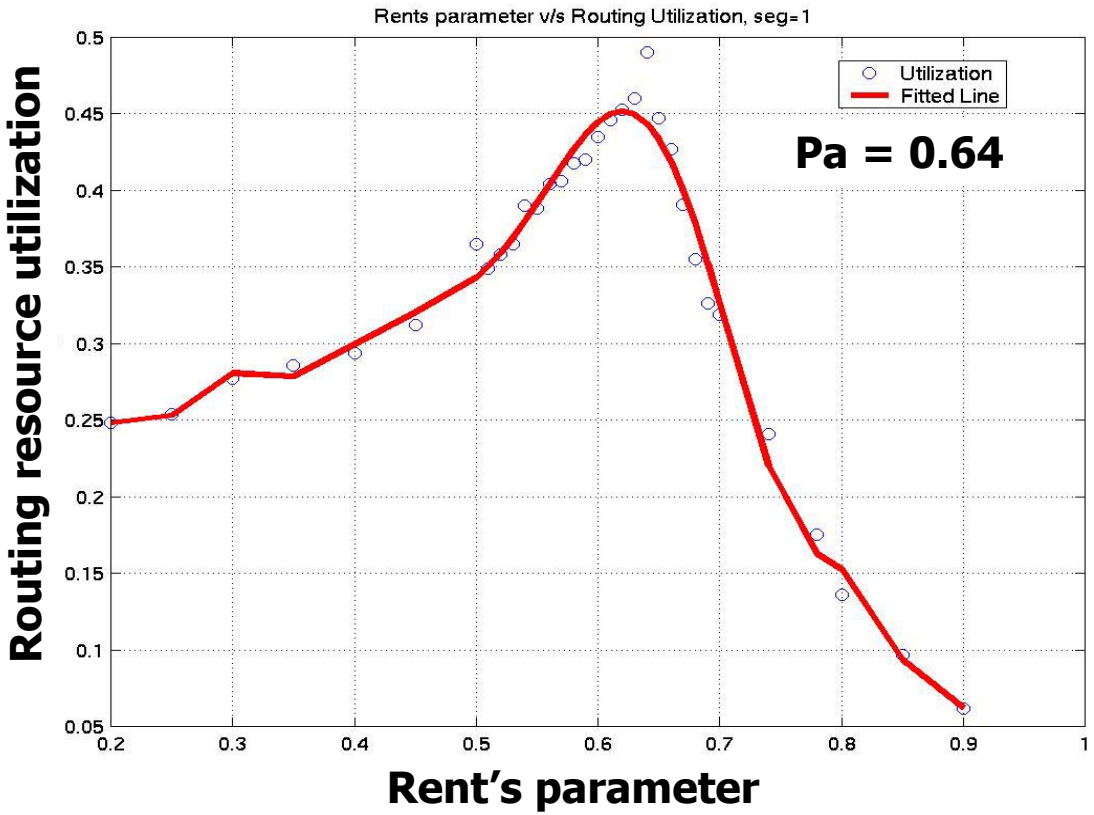
Measure for the complexity of the interconnection topology.

(simple) $0 \le p \le 1$ (complex)

Typical values: $0.5 \le p \le 0.75$

Rent's Rule

- Definitions
  - P_d – Rent's parameter for Design
  - P_a – Rent's parameter for Architecture



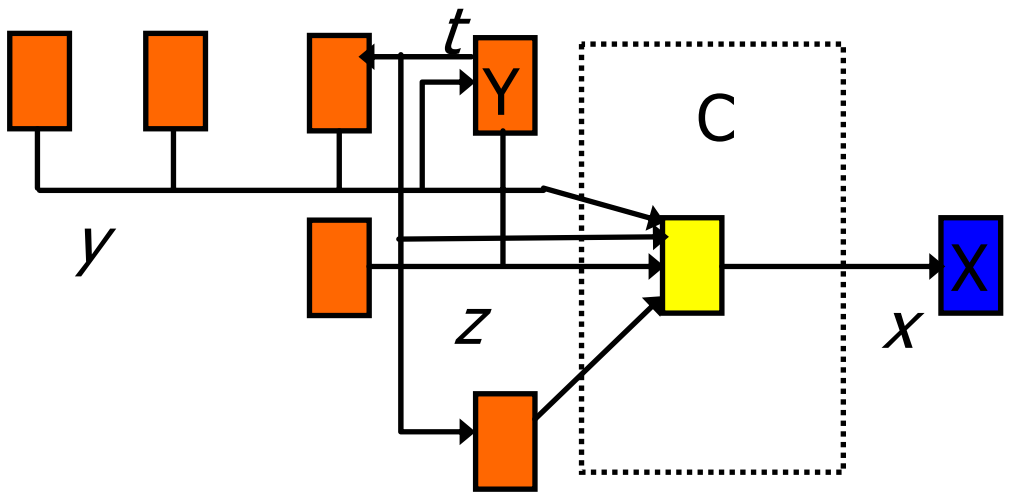
Logic Clustering

- *Separation* : Sum of all terminals of nets incident to LE
- *Degree* : Number of nets incident to LE

$$c = \frac{separation}{d^2}$$

Net weight: $w(x) = \frac{2}{r}$

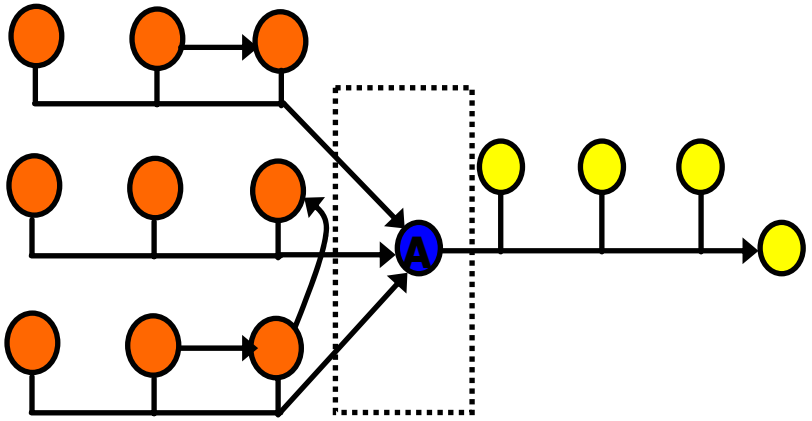
Example



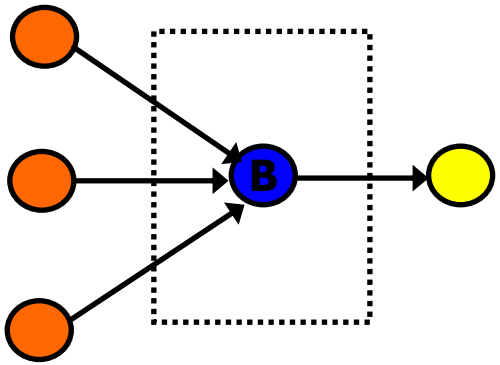
$$G(X) = [2nw(x) \bullet (1+\alpha)]\ k$$

$$IO \le (k+1)n^{P_a}$$

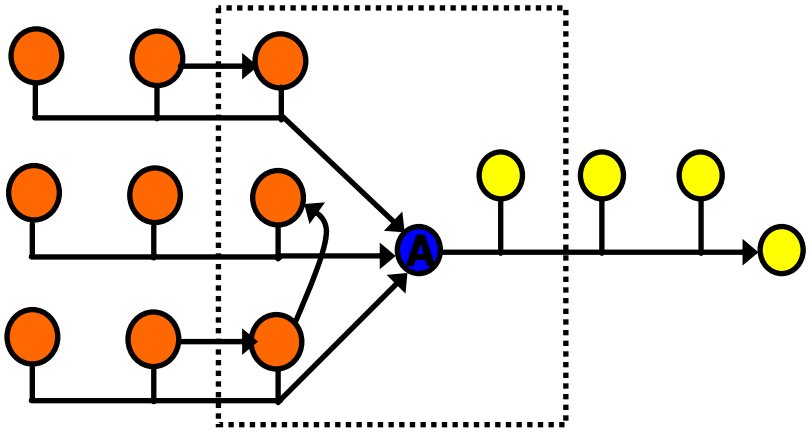
Clustering: Seed selection



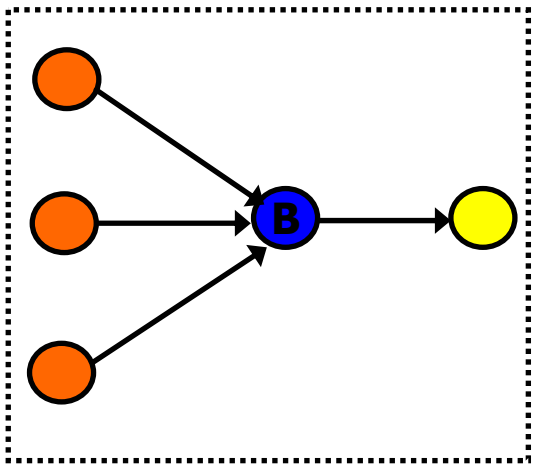
degree = 4; separation = 18, c = 1.25



degree = 4; separation = 8, c = 0.5



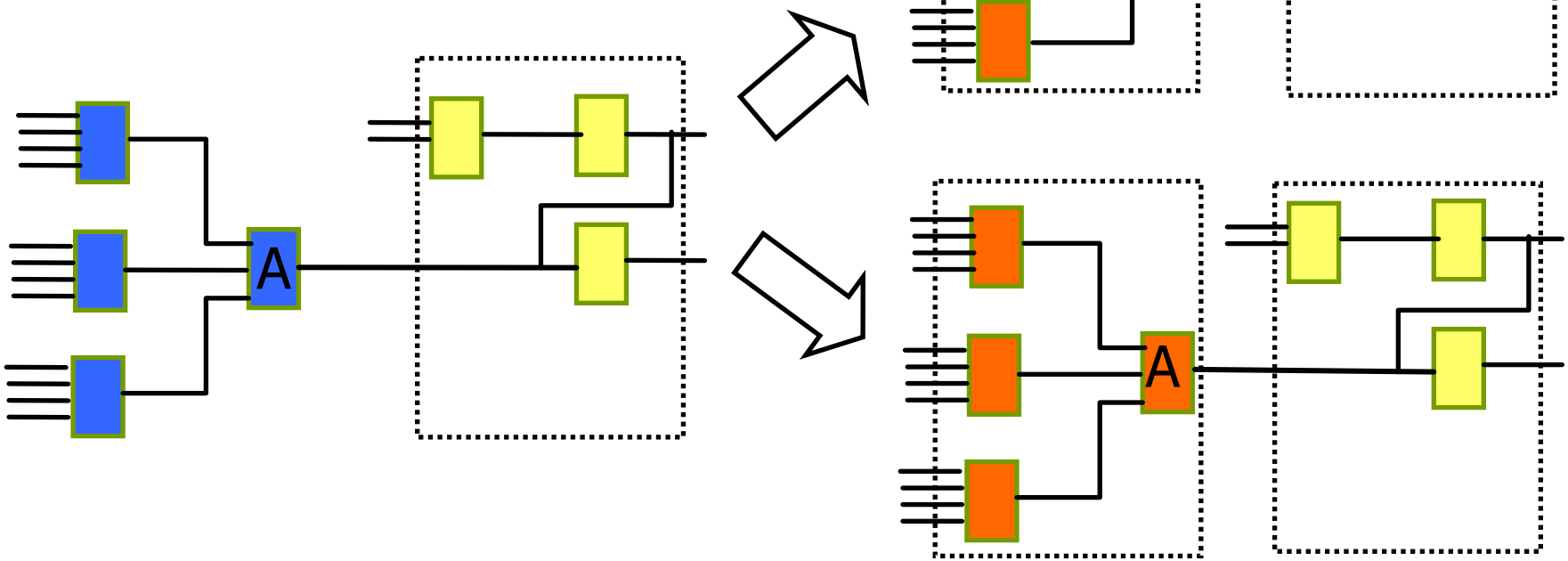
Nets absorbed = 1



Nets absorbed = 4

Rent's Rule: Depopulation

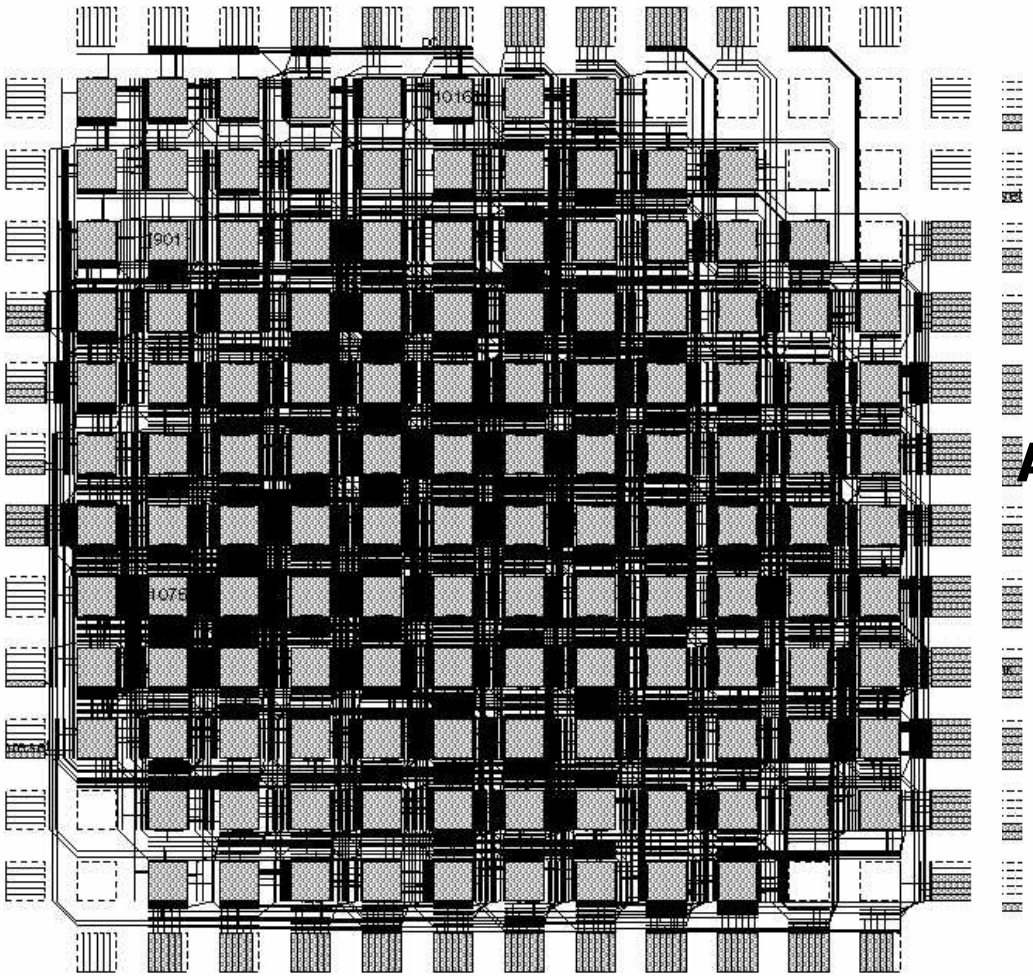
- Case 1: $P_d <= P_a$
  - Achieve spatial uniformity.
- Case 2: $P_d > P_a$
  - Need more routing resources.
  - Solution – Depopulate clusters



Regularity

- Better clustering for increased routability

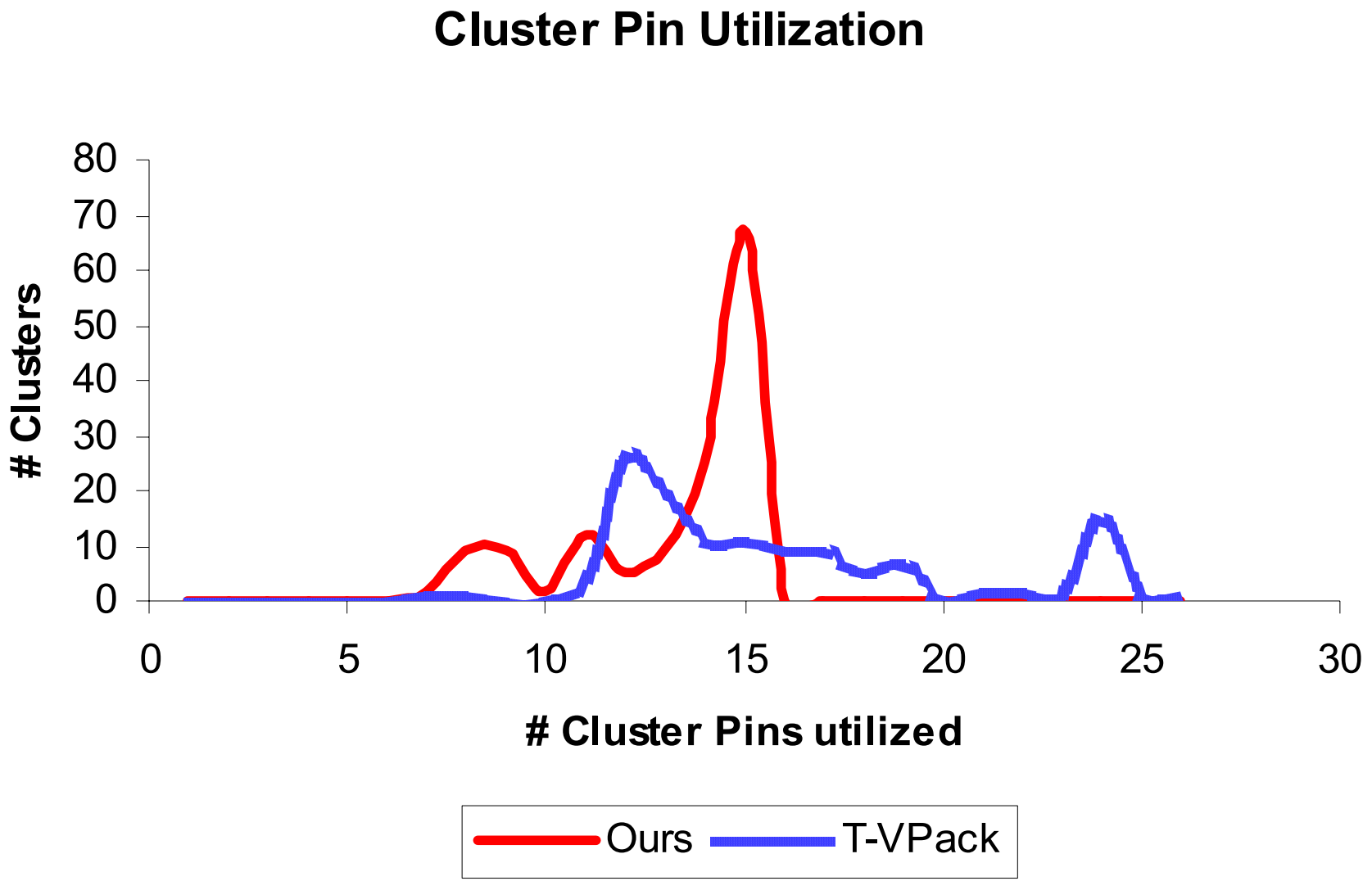
Avg. fan-out 2.7



Ours: Avg. fan-out 3.7

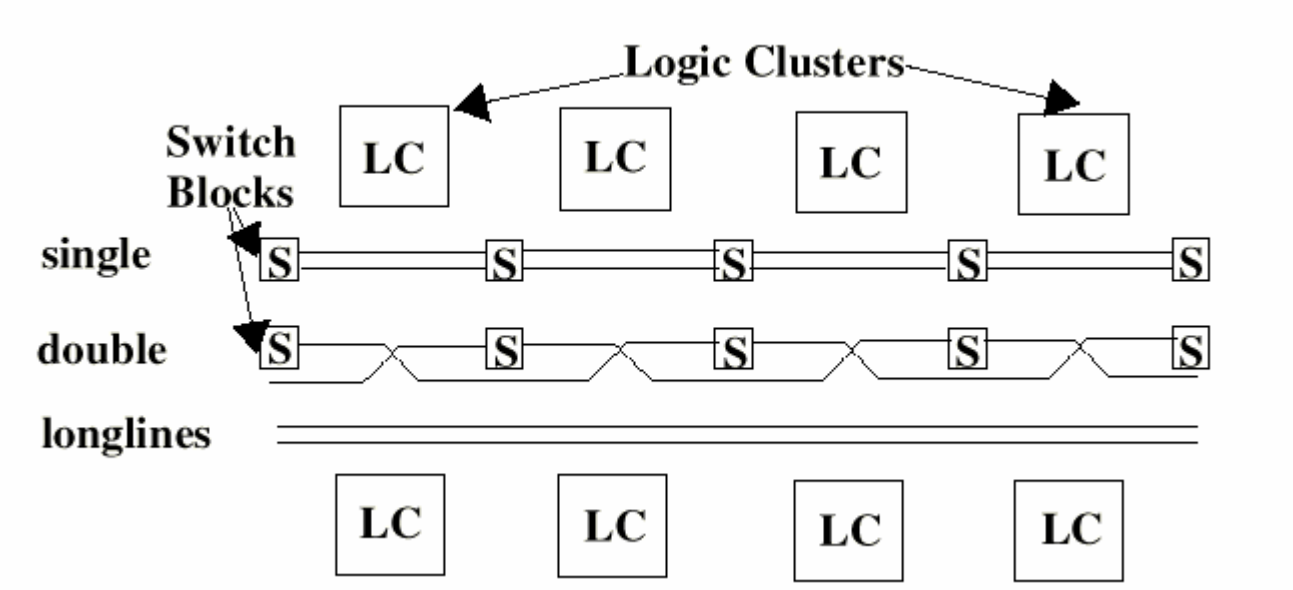
Routing succeeded with a channel width factor of 21 of 12

Rent's Rule: Depopulation



Segment length Planning

- Typical segment distribution



What is the best mix of segments for:
i) Area minimization ii) Area-delay minimization?

Fanout Distribution

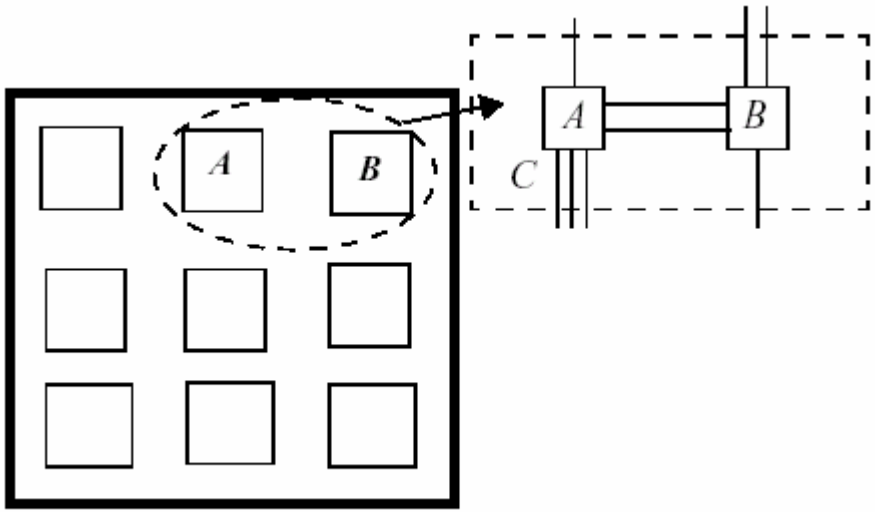
- Inter-cluster wire requirement
- Equivalent Rent's parameter of system

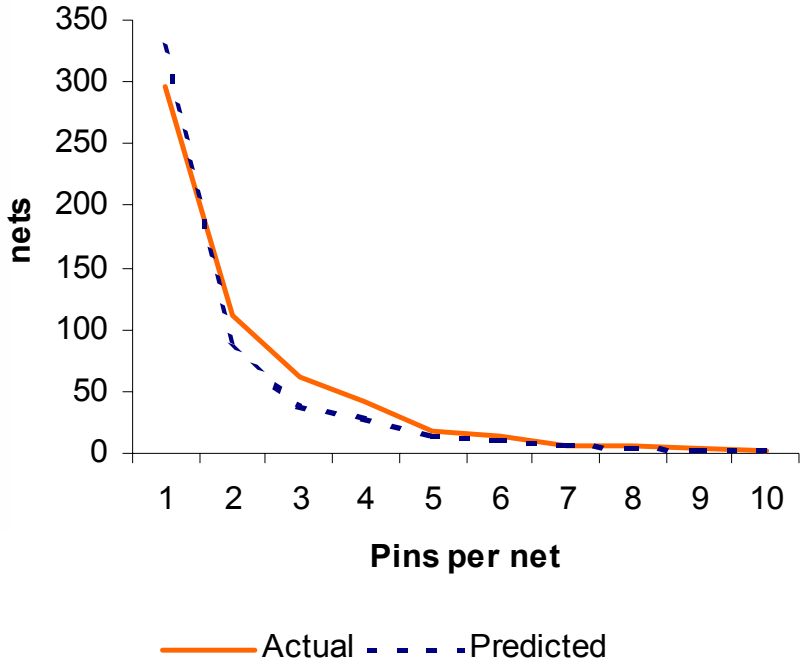
$$k_{eq} = \sqrt[N]{\left(\prod_{i=1}^{n} k_i^{N_i}\right)} \qquad p_{eq} = \frac{\sum_{i=1}^{n} p_i N_i}{N}$$

- Netlist profile
  - Low fanout nets – smaller length segments
  - Global nets – Longer segments (long lines)
- Global segments – buffered
- Shorter segments – not buffered(pass transistor switches)

Fanout distribution

- Array of clusters





$$Net(m) = \frac{kN((m-1)^{p-1} - m^{p-1})}{m}$$

$$Fo_{avg} = \frac{1-(Fo_{Max}+1)^{p-1}}{1-(Fo_{Max}+1)^{p-2}-\Omega(p,Fo_{Max})} - 1$$

$$\Omega(p,Fo_{Max}) = \sum_{n=1}^{Fo_{Max}} \frac{n^p}{n^2(n+1)}$$

Fanout distribution: Area-Minimization

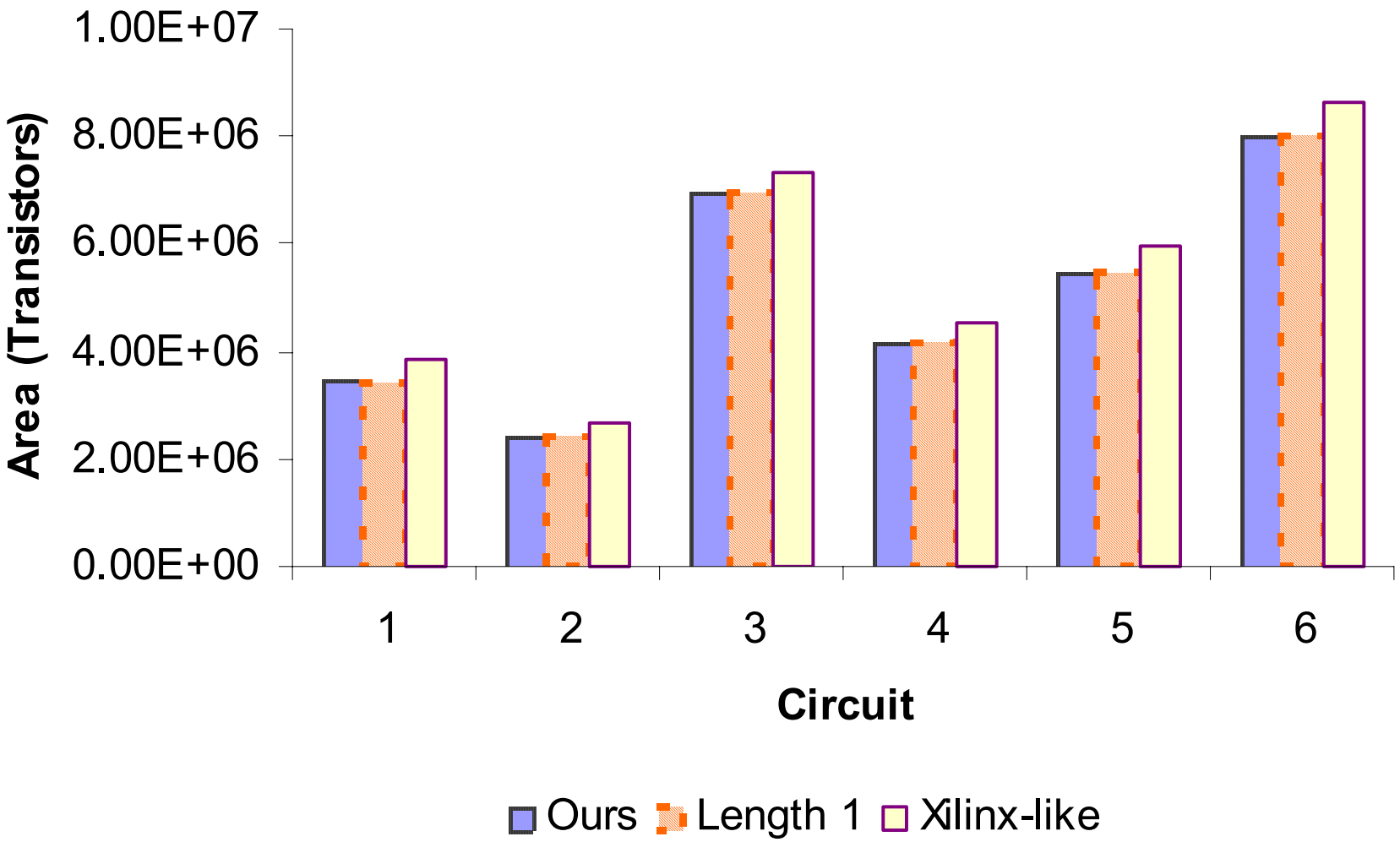
- Good Placement

$$f_l \propto l^{2p-3}$$

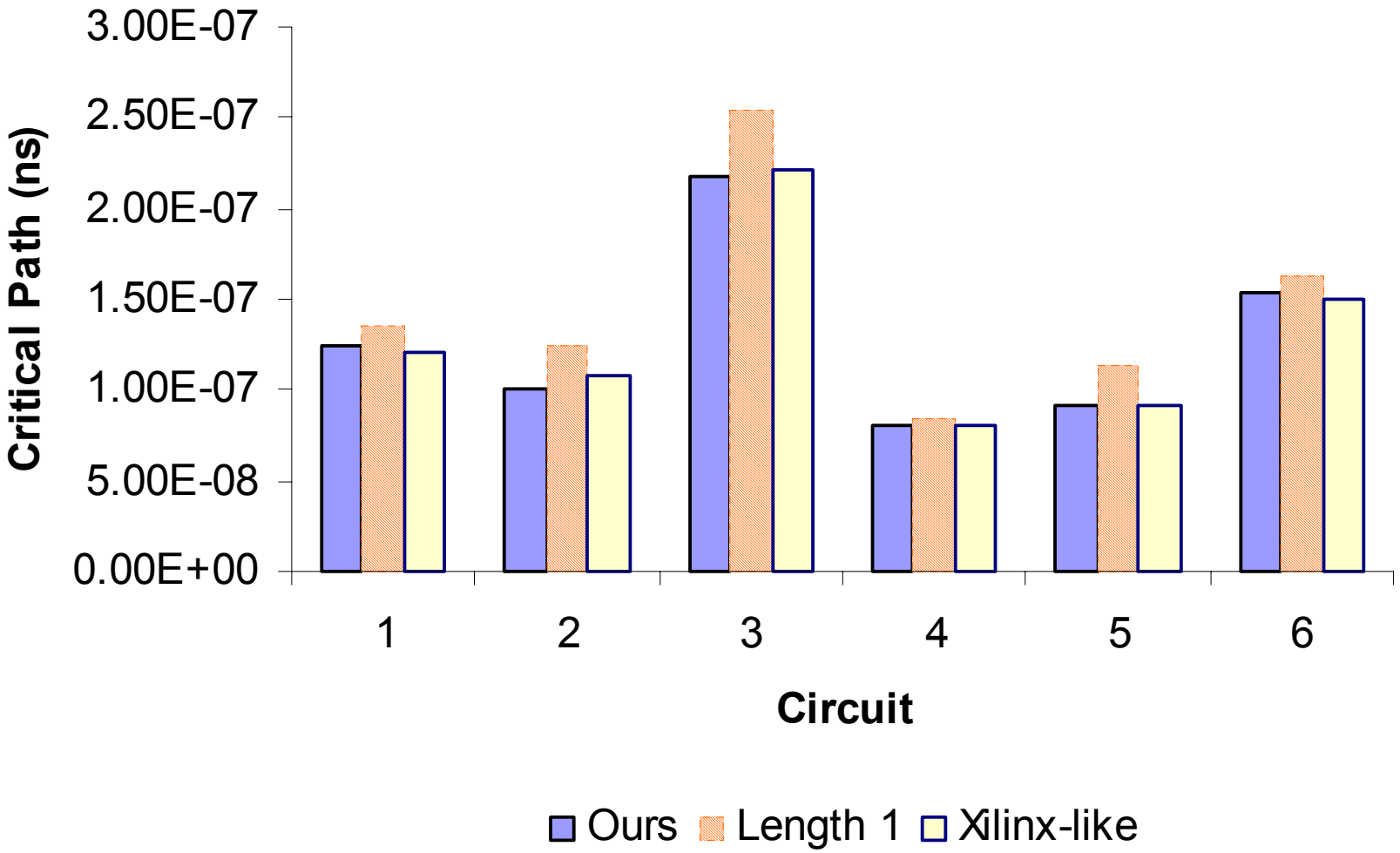
Example

Length	Ours	Length 1	Xilinx-like
1	0.47	1.00	.25
2	0.18	-	0.125
4	0.35	-	0.375
quarter-long	-	-	0.25

Fanout distribution: Area-Minimization



Fanout distribution: Area-Minimization



Fanout distribution:Area-Delay Product

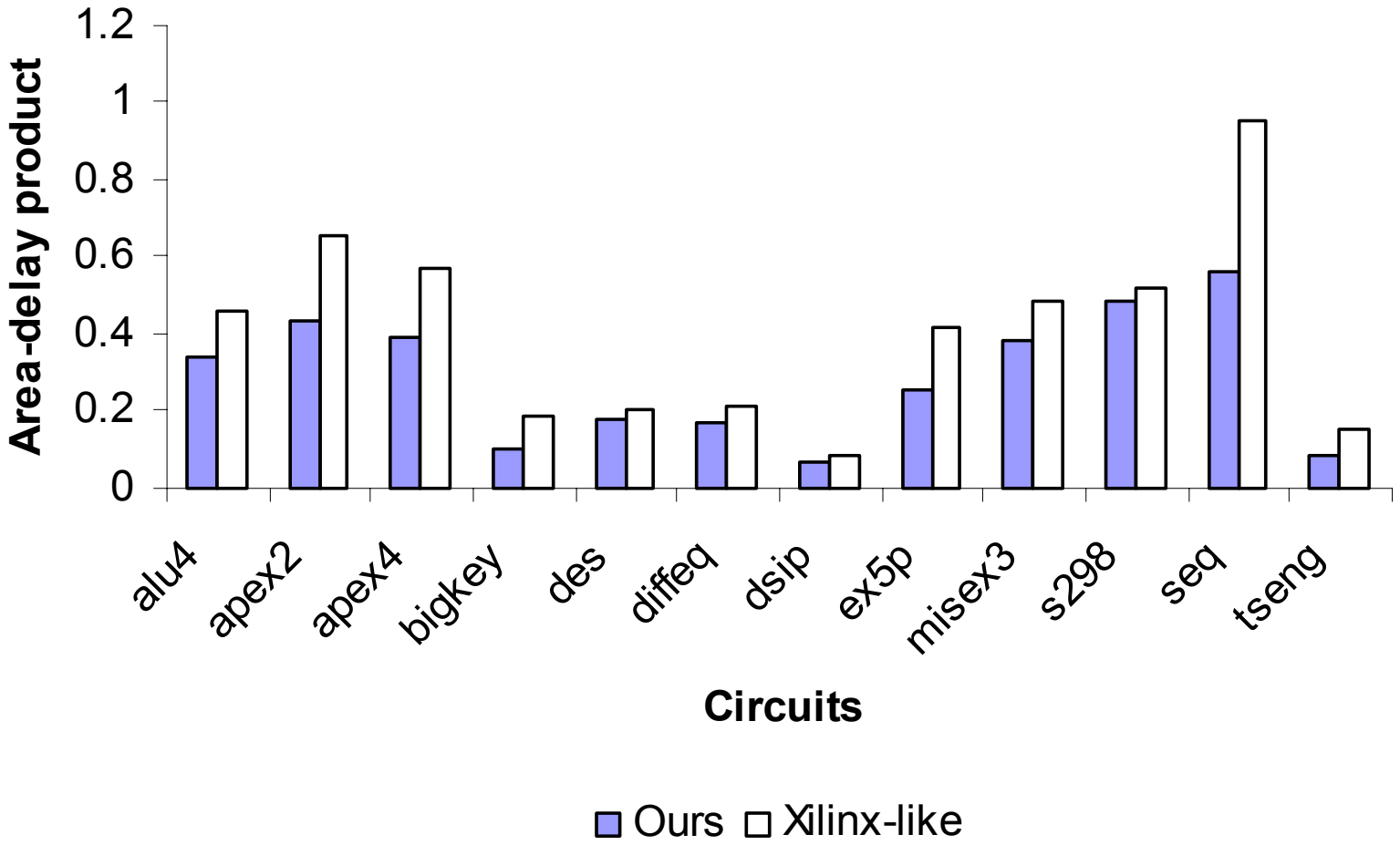
- Performance!
- Average fanout: Critical path model

$$FO_{Avg} = \frac{\sum_{FO=1}^{FO_{MAX}} FO \bullet Net(FO)}{\sum_{FO=1}^{FO_{MAX}} Net(FO)} \qquad n(f_m) = \frac{\sum_{FO=1}^{FO_m} FO \bullet Net(FO)}{\sum_{FO=1}^{FO_{MAX}} FO \bullet Net(FO)}, m = 4$$

$$n(f_m) = \frac{\sum_{FO=5}^{FO_m} FO \bullet Net(FO)}{\sum_{FO=1}^{FO_{MAX}} FO \bullet Net(FO)}, m = 8$$

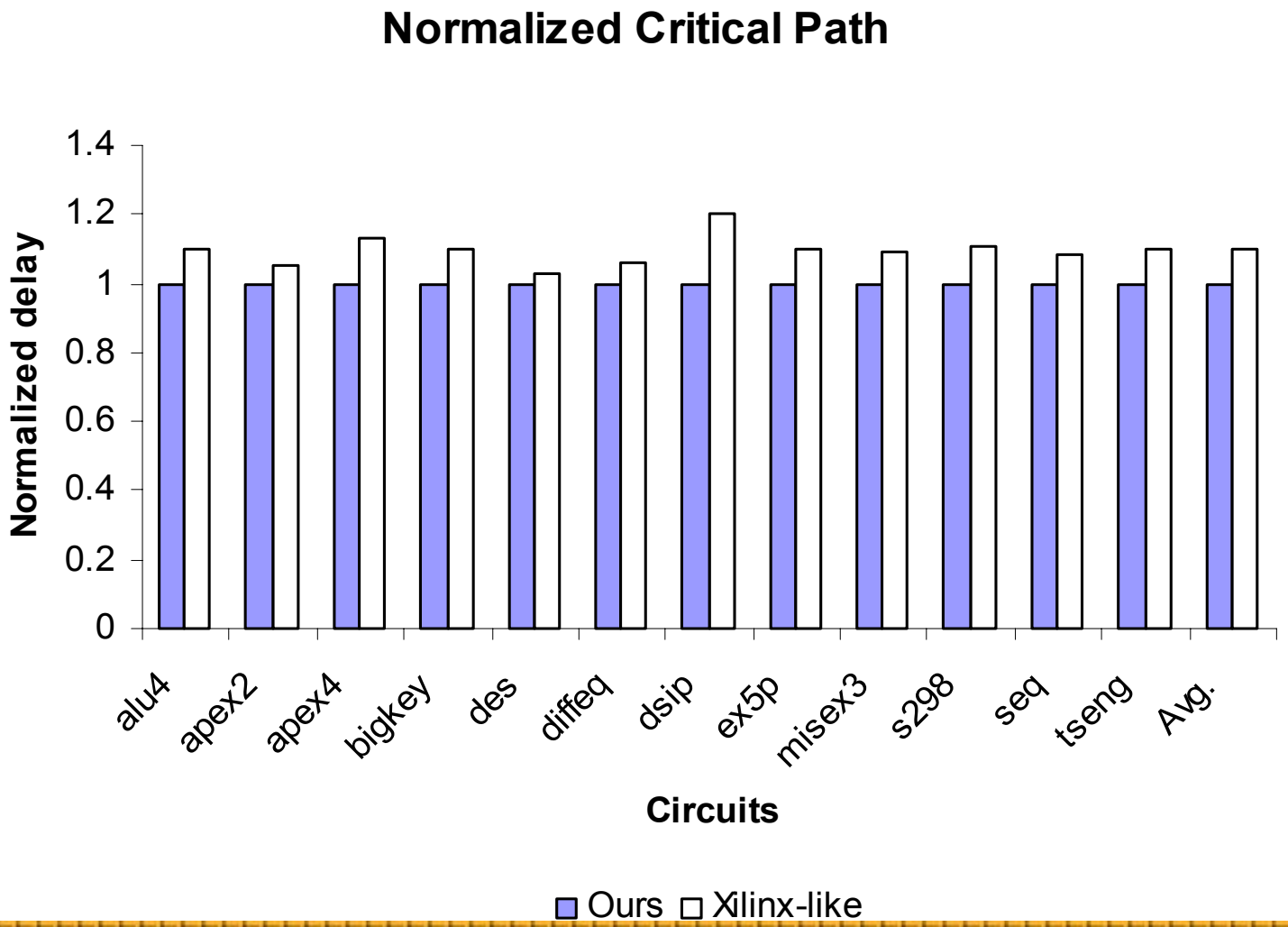
Fanout distribution:Area-Delay Product

- Timing-driven Placer and Router



Fanout distribution:Performance

- Using a Timing-driven Placer and Router



Conclusions & Future Work

- FPGA Clustering for Regularity
  - Rent's rule
- Fanout distribution based segment length planning
  - Area minimization
    - Reduced wire-length
  - Area-Delay minimization
    - Reduction of 29% over state-of-art
- Fixed FPGA Architecture
  - 20% better area-delay product
- Future work: Metal layer assignment
  - Applying technique to a pipelined FPGA

Thank You for Your Attention