Defect Tolerance for Nanocomputer Architecture

Arvind Kumar and Sandip Tiwari

Cornell University

February 15, 2004

Outline

- Defects and Reconfigurability
  - TeraMAC (large defect tolerance)
- A priori wire length estimation
  - Area
  - Power
  - Delay
  - Congestion
- Defect Analysis
  - Case 1: Defective Logic Blocks
  - Case 2: Defective Interconnects
  - Case 3: Defective Logic Blocks & Interconnects
- Summary and Caveats

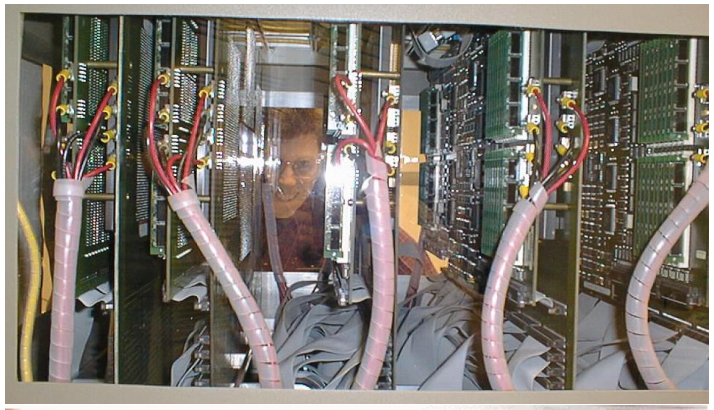
Defect Tolerance

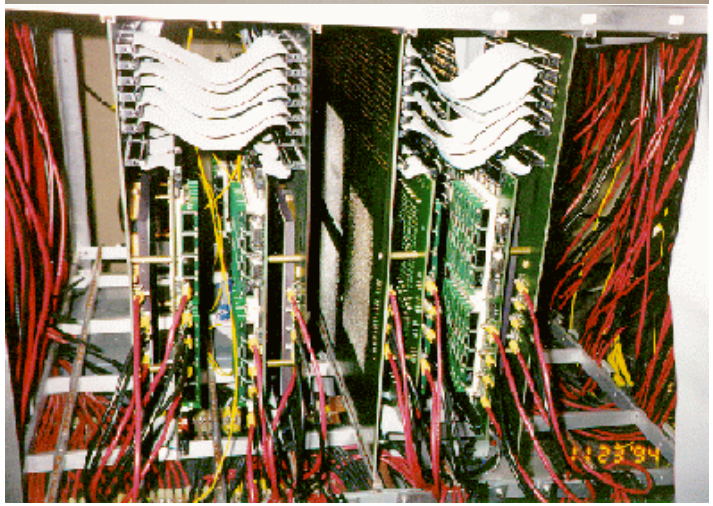
- Diagnosis: Granularity and Scalability
- Reconfiguration: Scalability
- Performance: ?

CV of TeraMAC

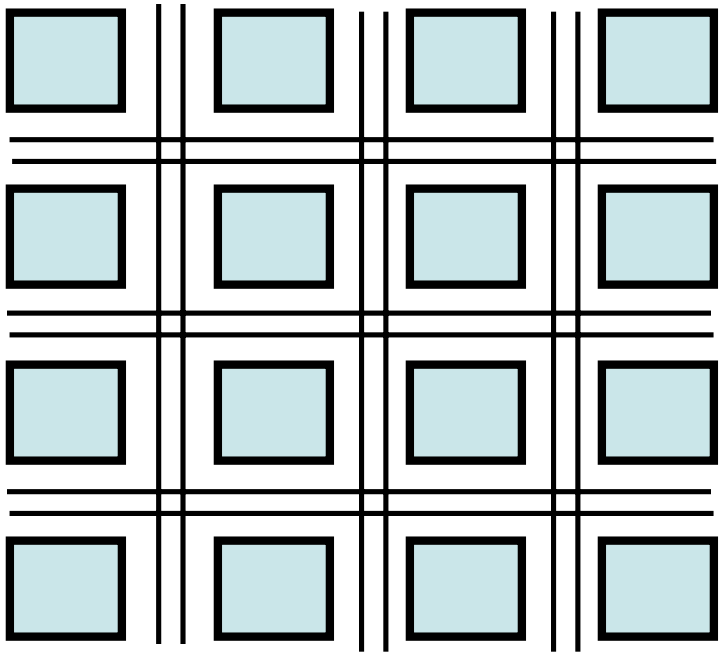
- Name: TeraMAC=Tera Multiple Architecture Computer
- Birth Place: HP labs
- Current Address: BYU
- 65,536 LUT
- 864 FPGAs
- Defective components: 75% FPGAs, 10% Logic Cells, 10% Interconnect
- 1 Million Logic Blocks @ 1 MHz
- Graph Partitioning, DNA String Match, Long Integer Multiply

TeraMAC

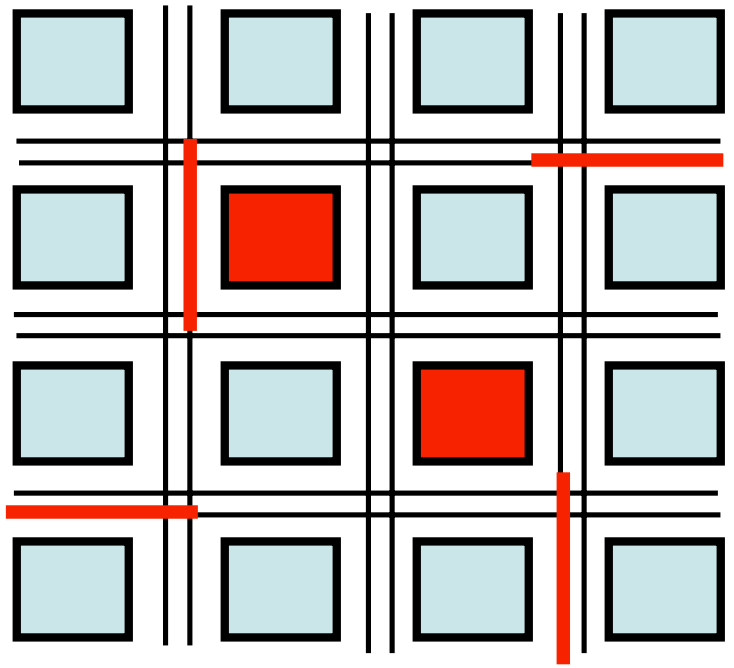




Reconfiguration basics

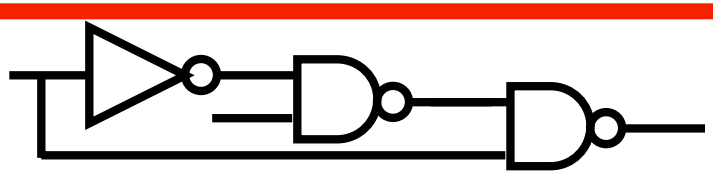


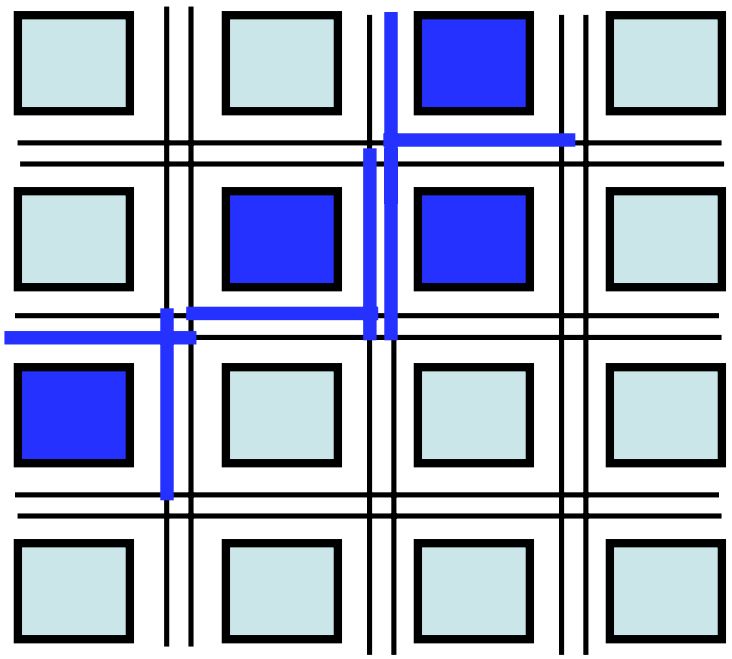
Ideal FPGA with no defects



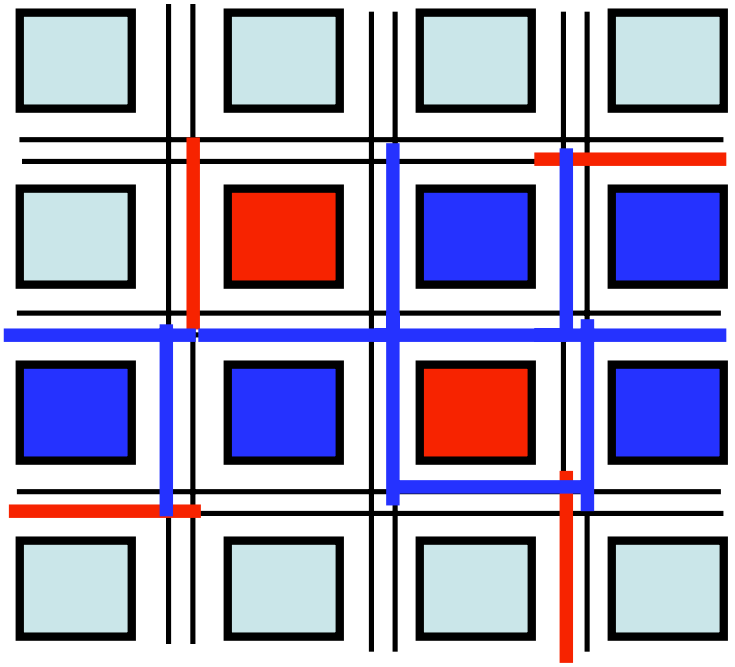
After diagnosis, defects are located

Reconfiguration Basics





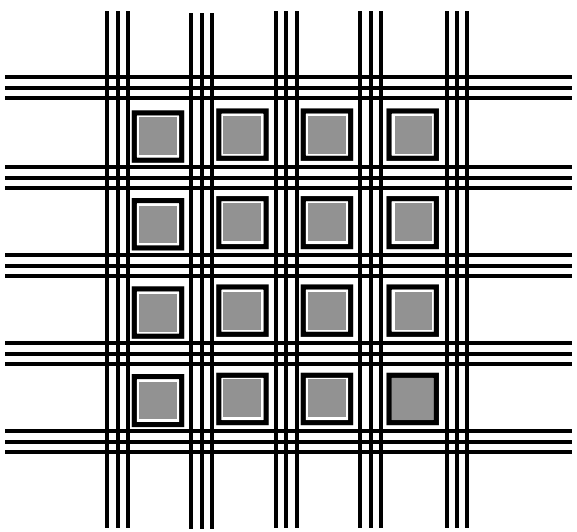
Placement/Routing on ideal FPGA

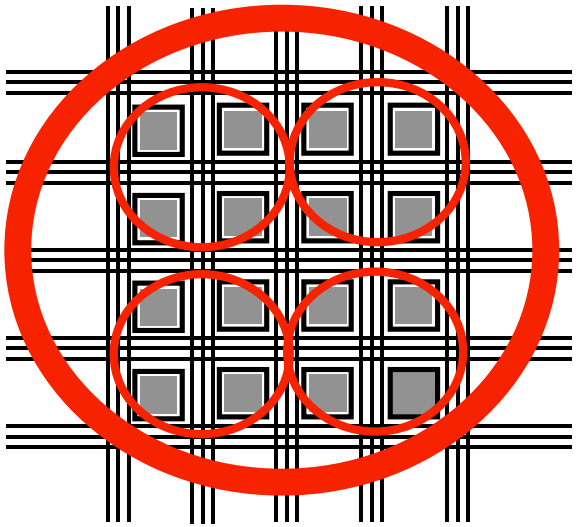


Placement/Routing on defective FPGA by avoiding the defects

MODEL

- Sea-of-gates FPGA architecture
- *N* Logic blocks, K Levels (4^K=N)
- Programmable network with Rent's exponent *p*





Rent's Rule

- E. F. Rent (1960, IBM) : log plot of # of pins vs. # of circuits in a logic design
- Landman and Russo (1971)
  - Empirical *Law* $T = tN^p$

  T = # of I/O terminals, *N* = # of gates

  t and *p* are empirical constants
- Rent's exponent *p* ⇔Placement Optimization

A priori wire length estimation

- Donath (1979)
- Davis (1998), Stroobandt (1999)
- Average wire length

$$l_{st}^{2D} = R(p)\frac{1-\gamma}{\gamma}\frac{H(K,p,1)}{H(K,p,2)}$$

$$H(K,p,x) = \frac{2^{K(2p-x)}-1}{2^{2p-x}-1}$$

$$R(p) = 4\frac{2p-3}{2p+1}\frac{4^{2p-1}-(p+2)2^{2p-1}+(p+1)}{4^{2p-1}-(2p+3)2^{2p-1}+(4p+2)}$$

K **= # of hierarchical levels, 0<γ<0.5**

Delay

- Longest source-sink pair at the highest hierarchical level

$$l_K^{2D} = 2^K R(p)$$

- Global and local distributed capacitance is similar
- RC delay is proportional to square of wire length (without repeaters)

Power Dissipation

- Wiring capacitance
- Assume constant activity factor
- Power ~ Capacitive load x Frequency

$$L_{TOTAL} = R(p)(1-\gamma)t4^{K}(1-4^{p-1})\frac{2^{(2p-1)K}-1}{2^{2p-1}-1}$$

$$P_{DYN} \sim \frac{L_{TOTAL}}{(l_K^{2D})^2} = (1-\gamma)t(1-4^{p-1})\frac{1}{R(p)}\frac{2^{(2p-1)K}-1}{2^{2p-1}-1}$$

Layout Area

- Area dominated by interconnects
- Donath and many others used minimum number of wire-segments for estimating wiring space.
- Davis: Area ~ Wire length x Gate Pitch

$$Area \sim L_{TOTAL}$$

Congestion and Routability

- Bakoglu, Lou, and Parthsarthy.
- Yang: More appropriate for gate arrays and is based on cut ratio in recursive bipartitioning

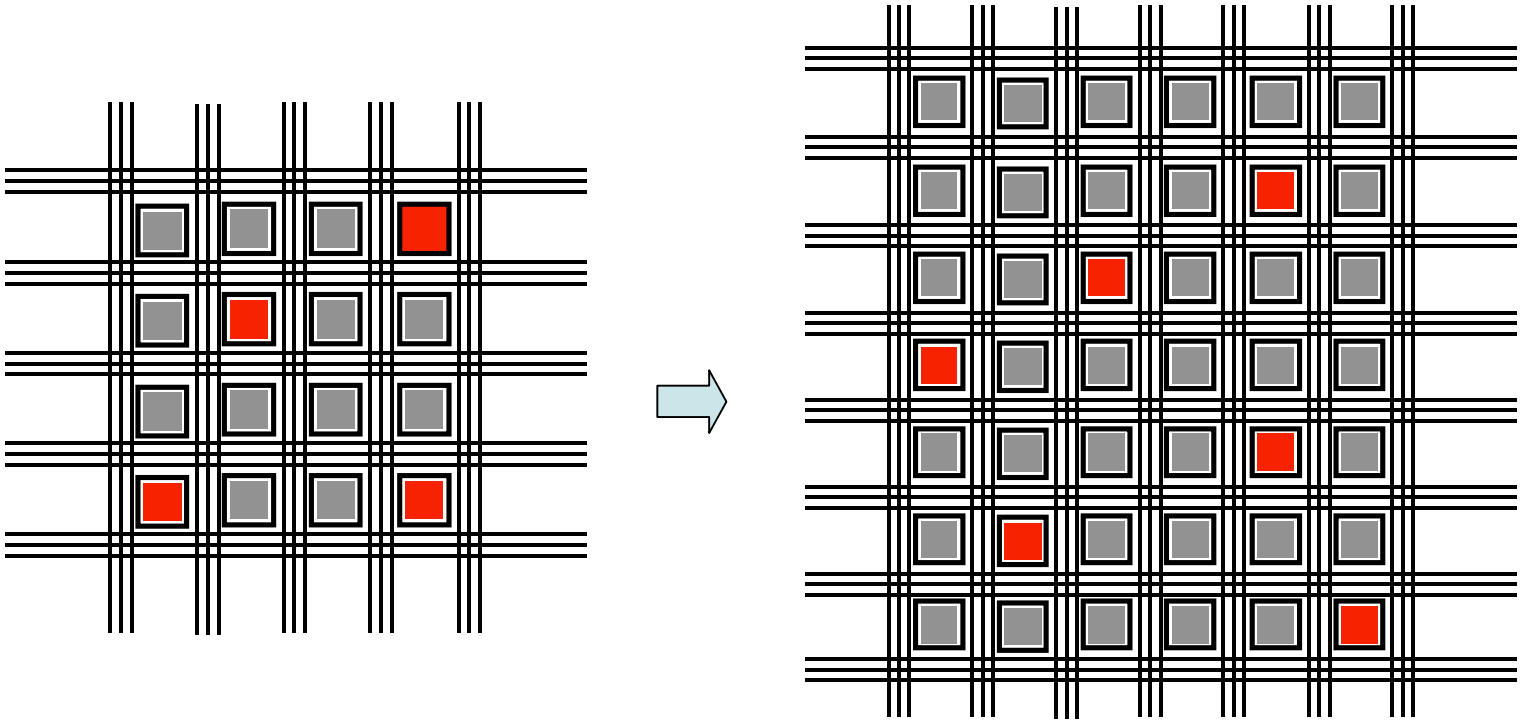
$$C_{max} < \frac{C_1(1-\alpha^{2\log_4 N})}{(1-\alpha)} \quad C_1 = \frac{tN}{2} \quad \alpha = \frac{C_{i+1}}{c_i} = 2^{-p}$$

- Gamal: Channel width α Wire length

$$W = \frac{tL}{2}$$

Case 1

- Logic Blocks are defective
- Interconnects are perfect



Case 1

- Uniform defect density d_{LB}
- Logic Block Scaling

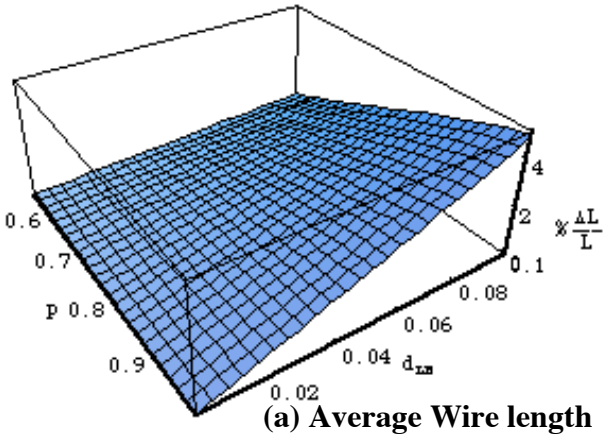
$$C_{LB} = \frac{1}{1 - d_{LB}}$$

- Increase in terminal count

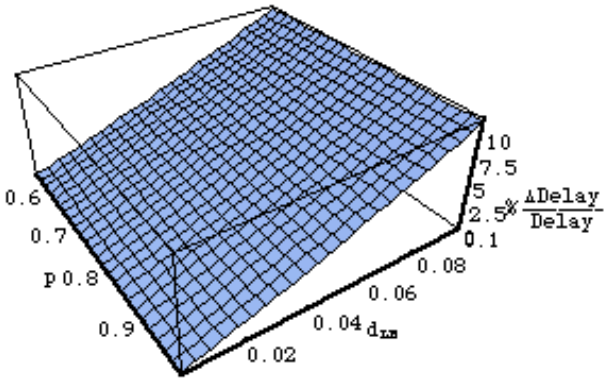
$$\Delta T_{EXT} = tN^p \left[\left(\frac{1}{1 - d_{LB}} \right)^p - 1 \right]$$

$$\Delta T_{INT} = tN \left[\left(\frac{d_{LB}}{1 - d_{LB}} \right) - \left(\left(\frac{1}{1 - d_{LB}} \right)^p - 1 \right) \right]$$

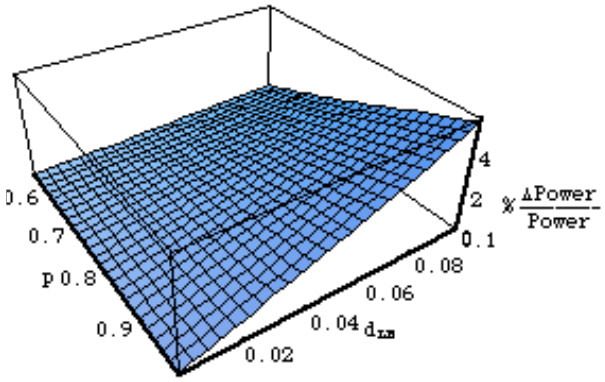
Case 1



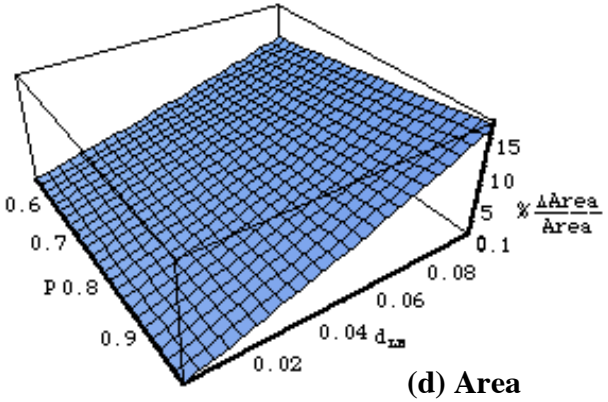
(a) Average Wire length



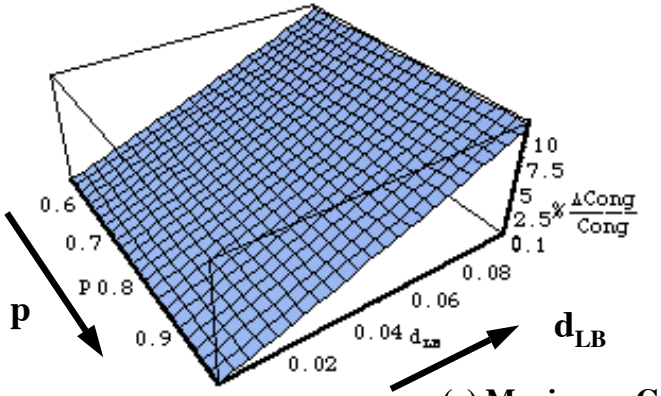
(b) Largest Delay



(c) Power Dissipation



(d) Area

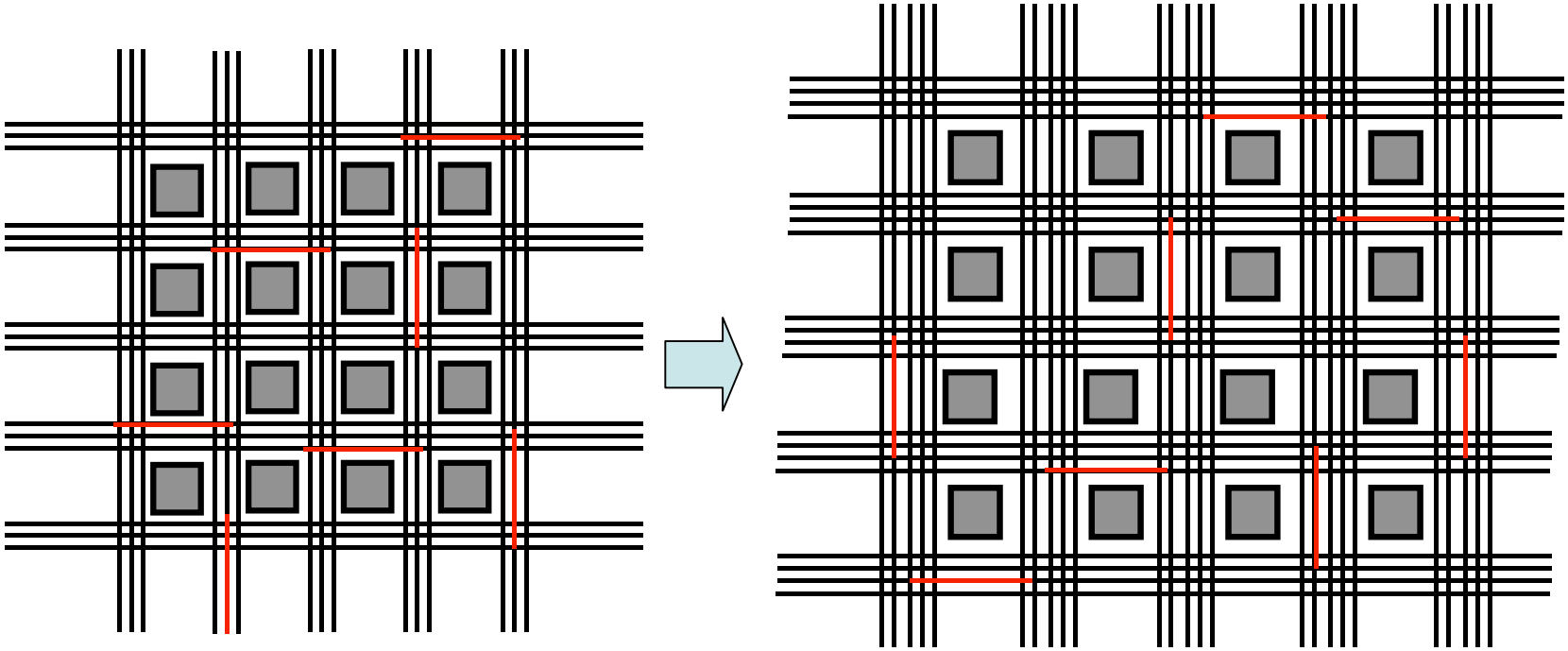


(e) Maximum Congestion

- K=20

Case 2

- Logic blocks are perfect, Interconnect is defective
- Assume same reduction in communication at all hierarchical levels.



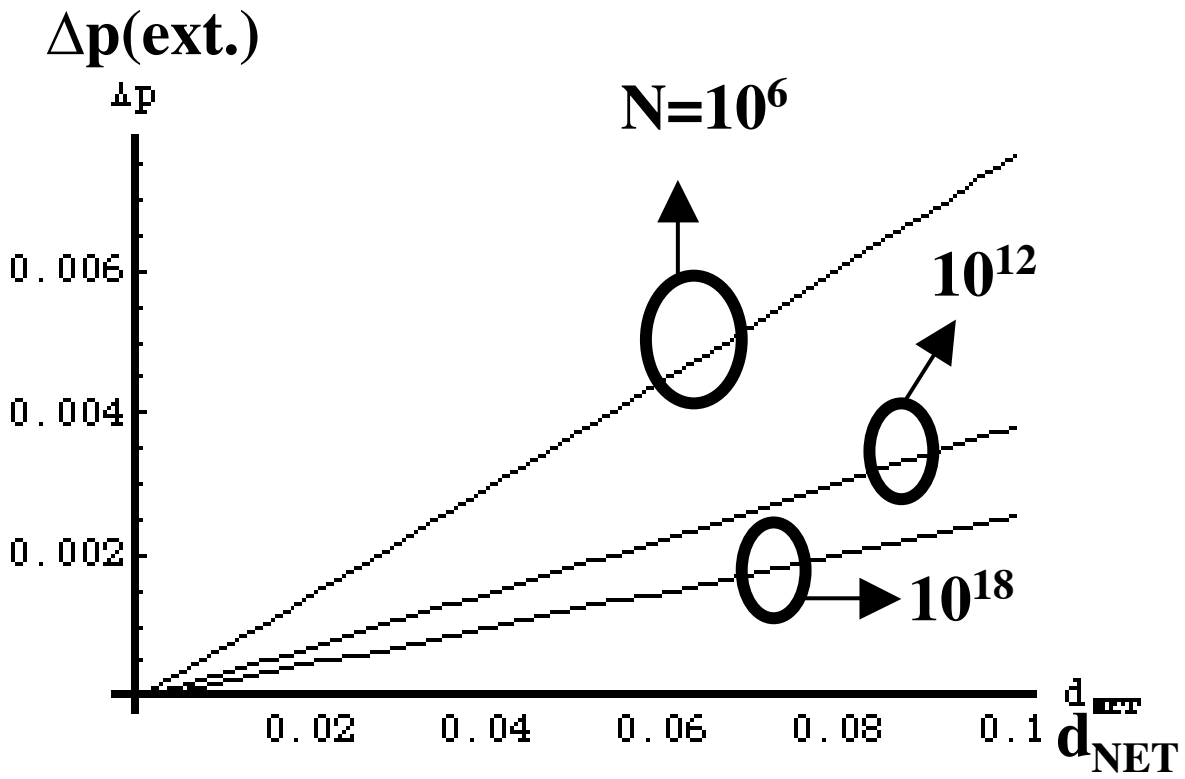
Case 2

- **Decrease in Rent's exponent**

$$t\,(N)^{p'} = (1 - d_{NET})t\,(N)^{p}$$

$$\Delta p_{EXT} = p - p' = \left(\frac{\log \frac{1}{1-d_{NET}}}{\log N} \right)$$

Case 2



- Decrease in Rent's exponent

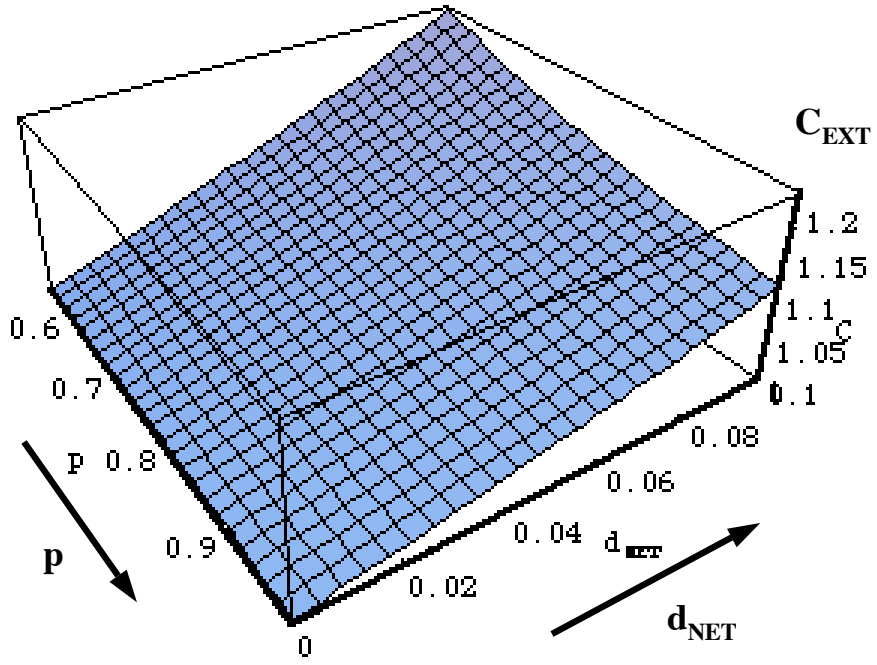
Case 2

- **Defect tolerance by additional logic blocks**

$$(1 - d_{NET})t(C_{EXT} \times N)^p = tN^p$$

$$C_{EXT} = \left(\frac{1}{1 - d_{NET}}\right)^{\frac{1}{p}}$$

Case 2



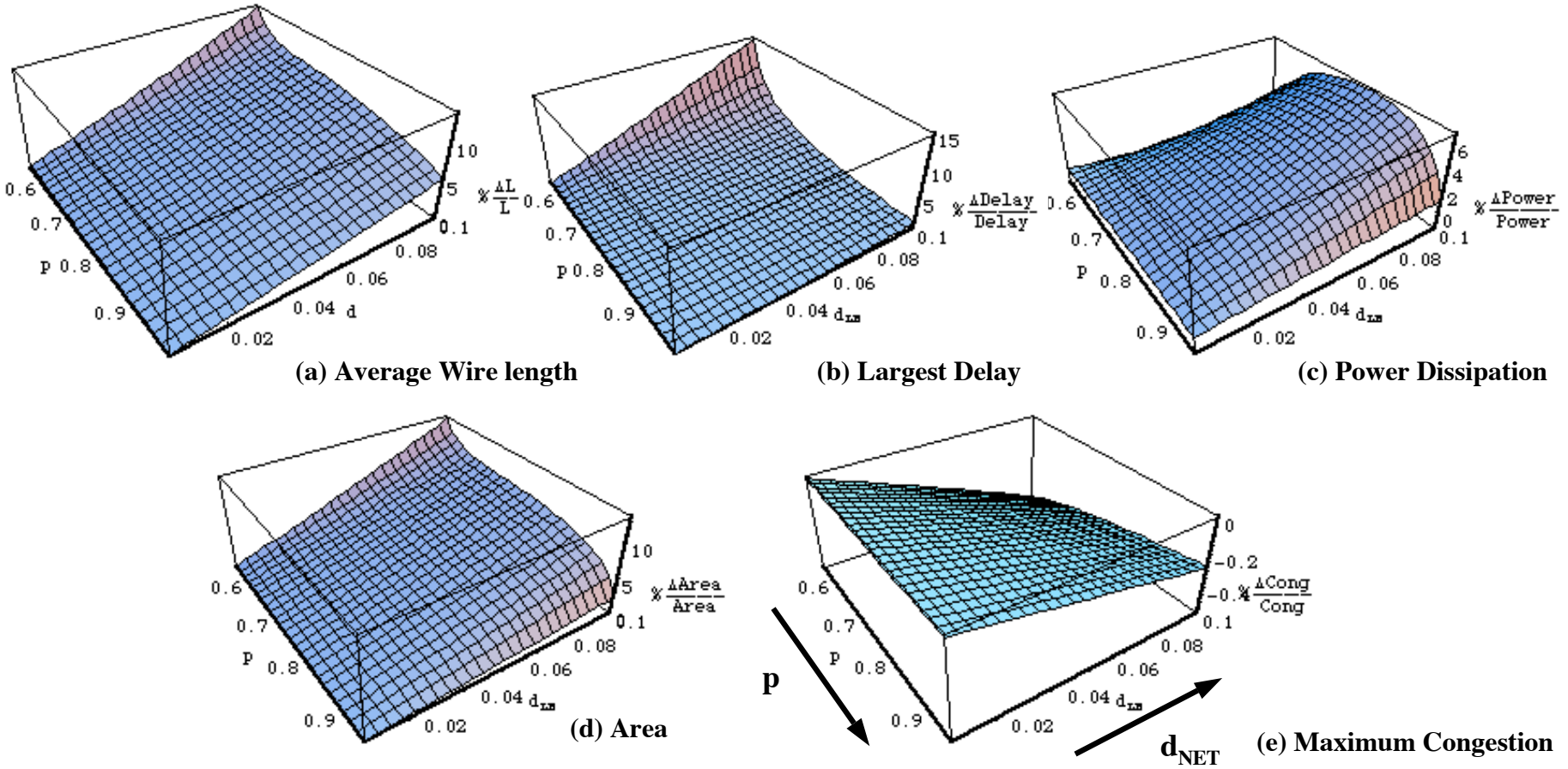
- Defect tolerance by additional logic blocks
- More scaling required for well optimized circuits.

Case 2

- **Defect tolerance through richer interconnect**

$$p' = p + \left(\frac{\log \frac{1}{1-d_{NET}}}{\log N} \right)$$

Case 2



(a) Average Wire length (b) Largest Delay (c) Power Dissipation

(d) Area (e) Maximum Congestion

- **Defect tolerance through richer interconnect**

Case 2

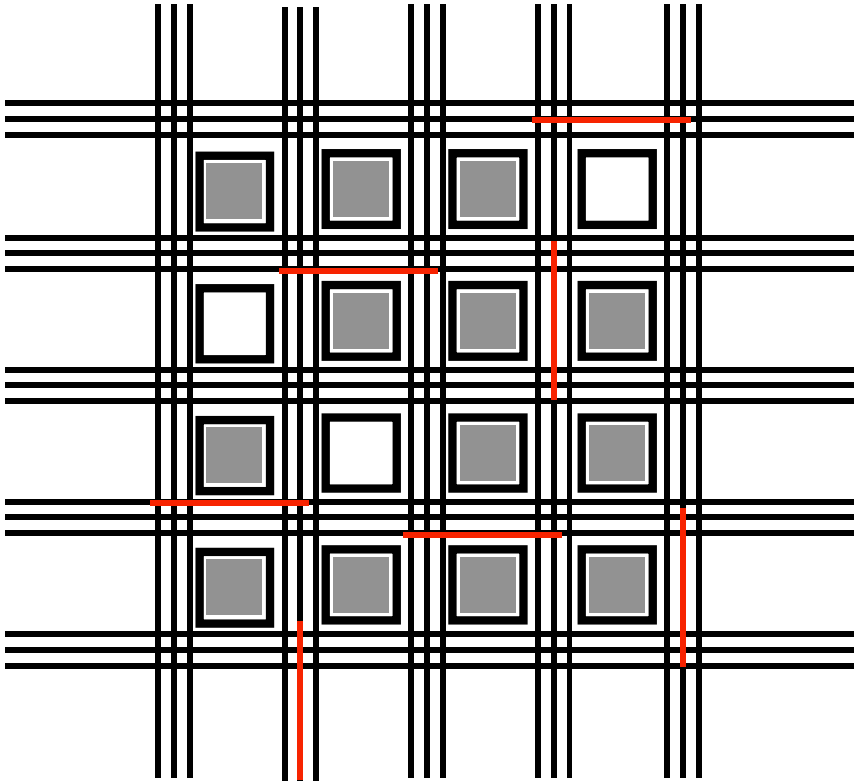
- **Interconnect vs. logic block trade off**

$$(1 - d_{NET})K(C \times N)^{p'} \geq K N^{p}$$

$$p' \geq p\left(\frac{\log N}{\log(C \times N)}\right) + \left(\frac{\log \frac{1}{1-d_{NET}}}{\log(C \times N)}\right)$$

Case 3

- Both logic blocks and interconnects are defective



Case 3

- **Decrease in Rent's exponent**

$$(1-d_{NET})t\left((1-d_{LB})\times N\right)^{p}=tN^{p'}$$

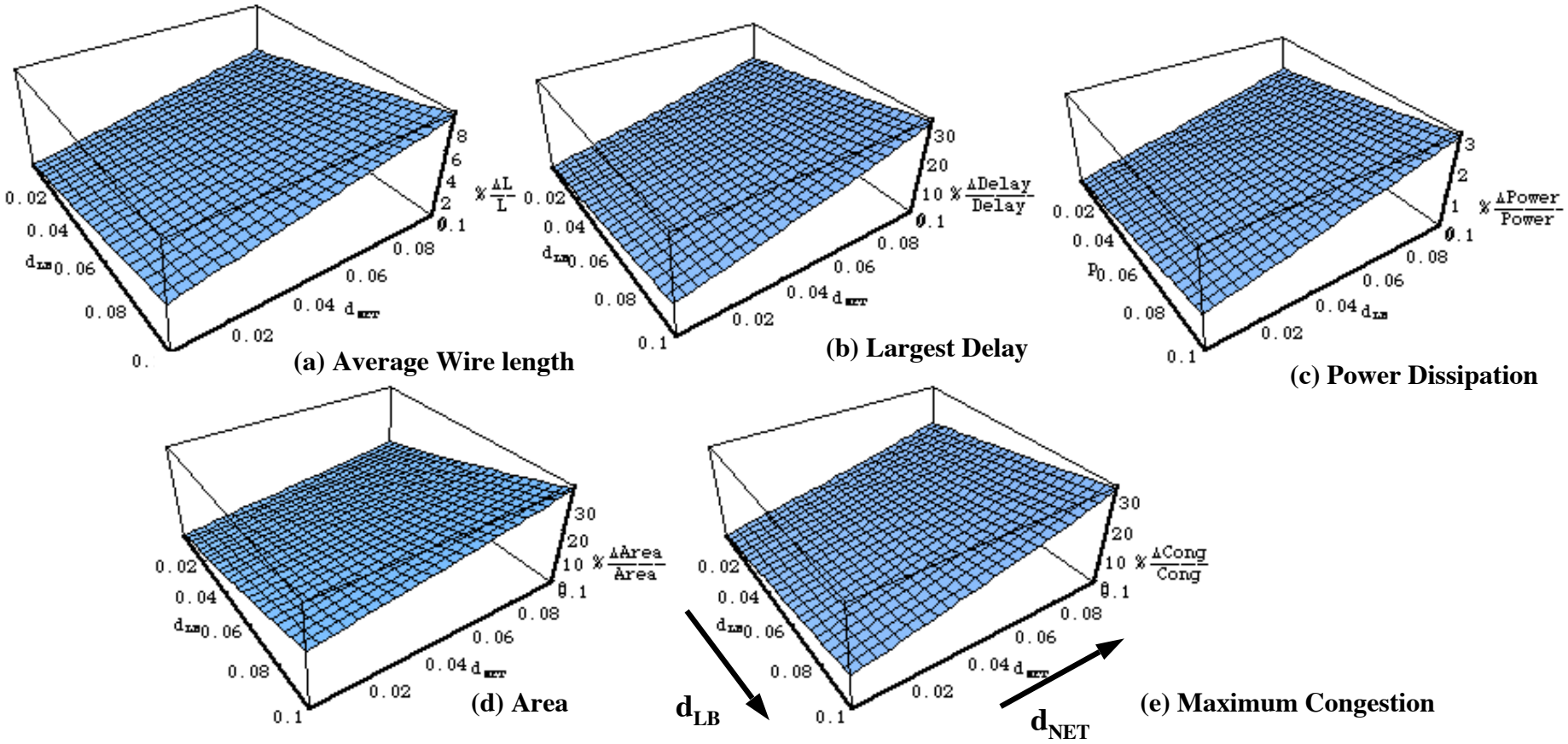
$$\Delta p=p-p'=p-\left(\frac{\log(1-d_{NET})}{\log N}\right)-\left(p\frac{\log\left(N(1-d_{LB})\right)}{\log N}\right.$$

Case 3

- **Defect tolerance through logic blocks only.**

$$C_{LB} = \left(\frac{1}{1 - d_{NET}}\right)^{\frac{1}{p}} \left(\frac{1}{1 - d_{LB}}\right)$$

Case 3



- **Defect tolerance through logic blocks only.**

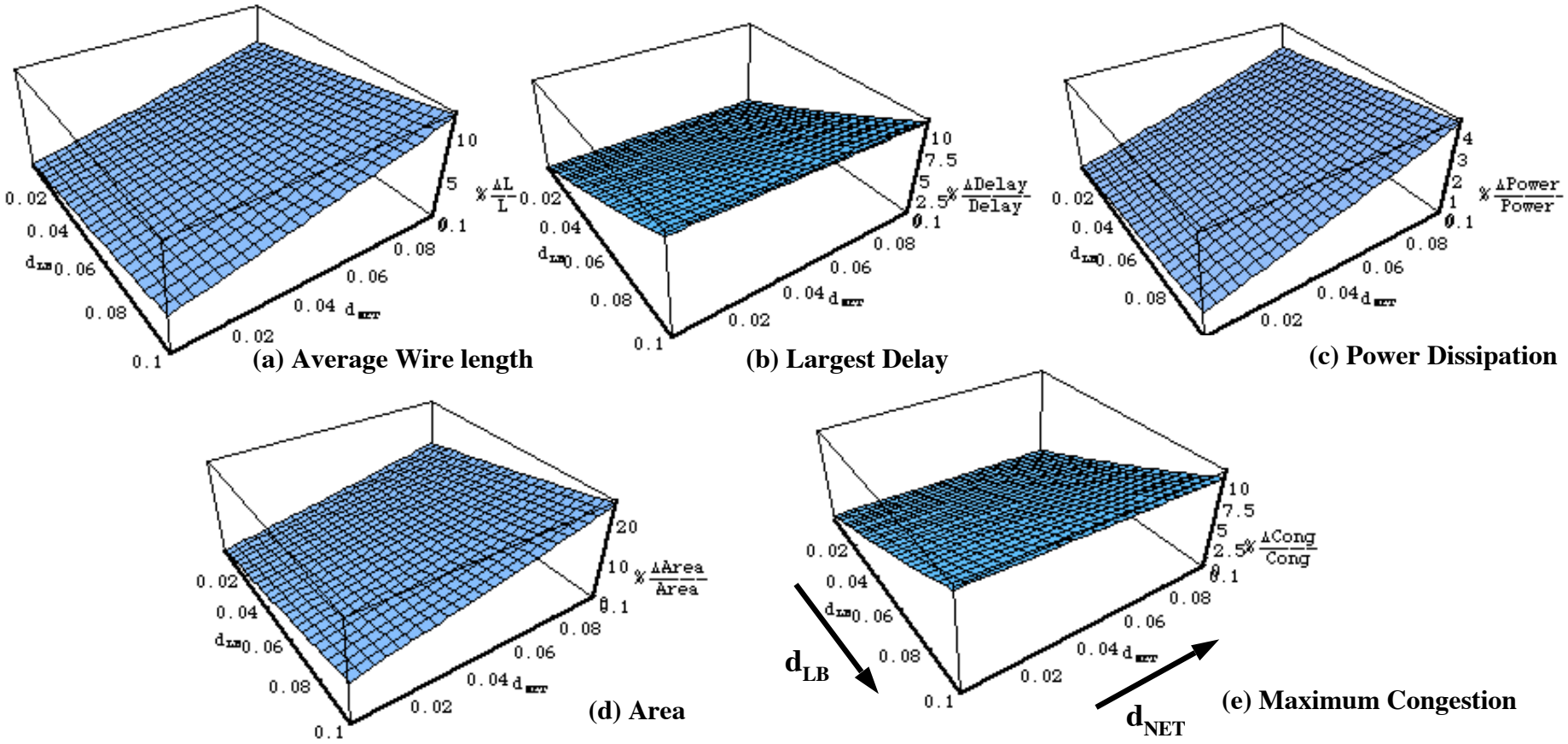
Case 3

- **Defect tolerance through logic blocks and interconnect.**

$$p' = p + \left(\frac{\log \frac{1}{1-d_{NET}}}{\log N} \right)$$

$$C_{LB} = \frac{1}{1 - d_{LB}}$$

Case 3



- Detect tolerance through logic blocks and interconnect.

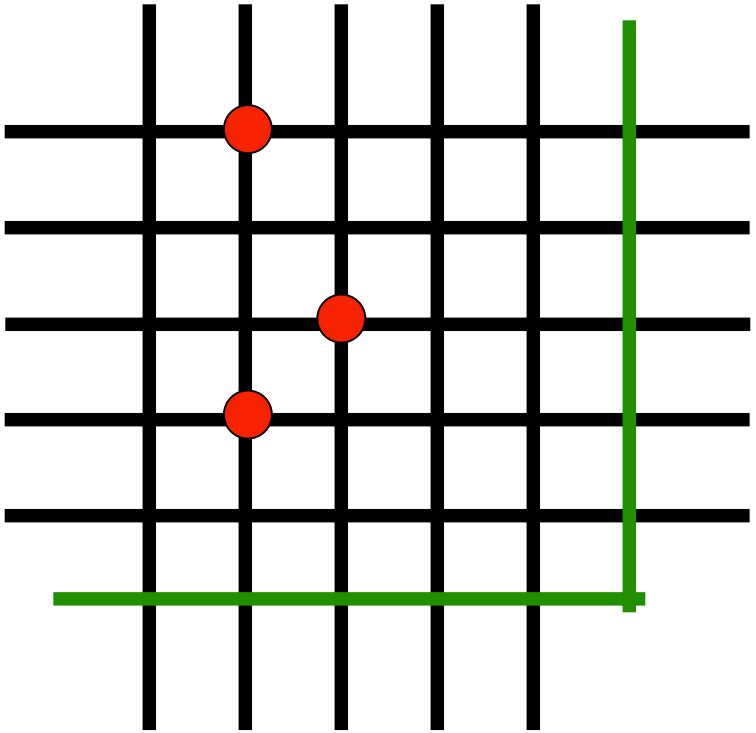
Case 3

- **Logic block vs. interconnect trade off**

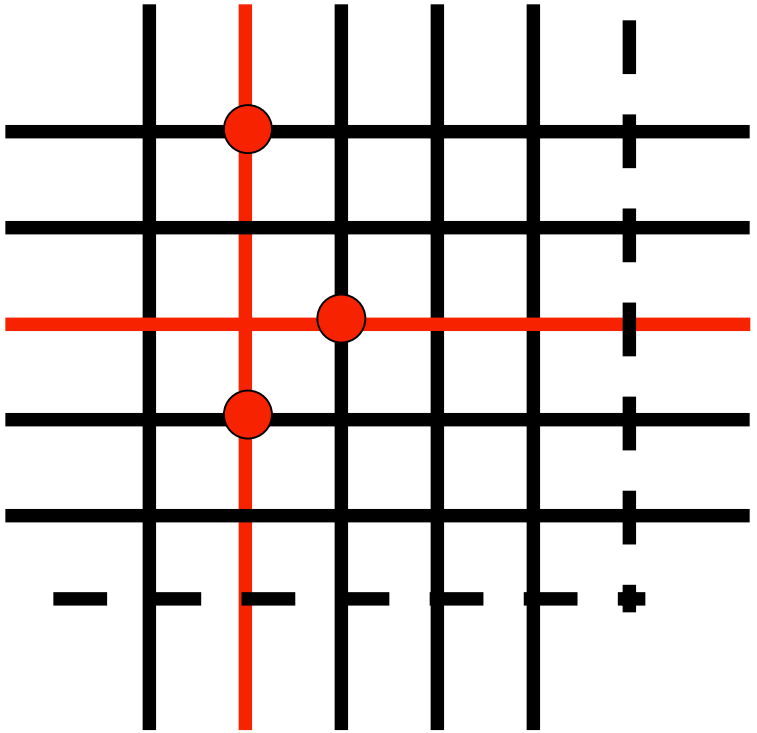
$$(1-d_{NET})t\,(C \times (1-d_{LB}) \times N)^{p'} \geq tN^{p'}$$

February 15, 2004

Testing Overhead ?



Defective elements are identified



Spare row/columns are used instead of the defective ones

- **Use of spare row/column for yield enhancement in memory arrays**

Complexity

- Shi and Fuchs (1989)
- Fault coverage
- *Result* : An NxN array with rN and cN spare rows and columns respectively (0<r,c<1) is irreparable if failure probability

$$f(n) > \frac{Constant(r,c)}{N}$$

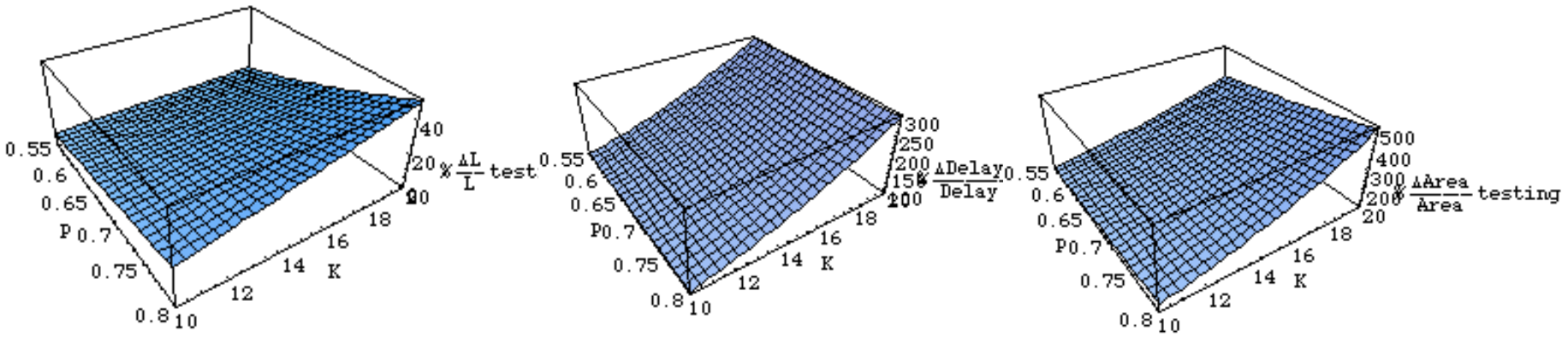
- For the memory array p=0.5

$$f(n) > \frac{Constant(r,c)}{(N \times N)^p}$$

An exponent rule ?

- **The testing cost**

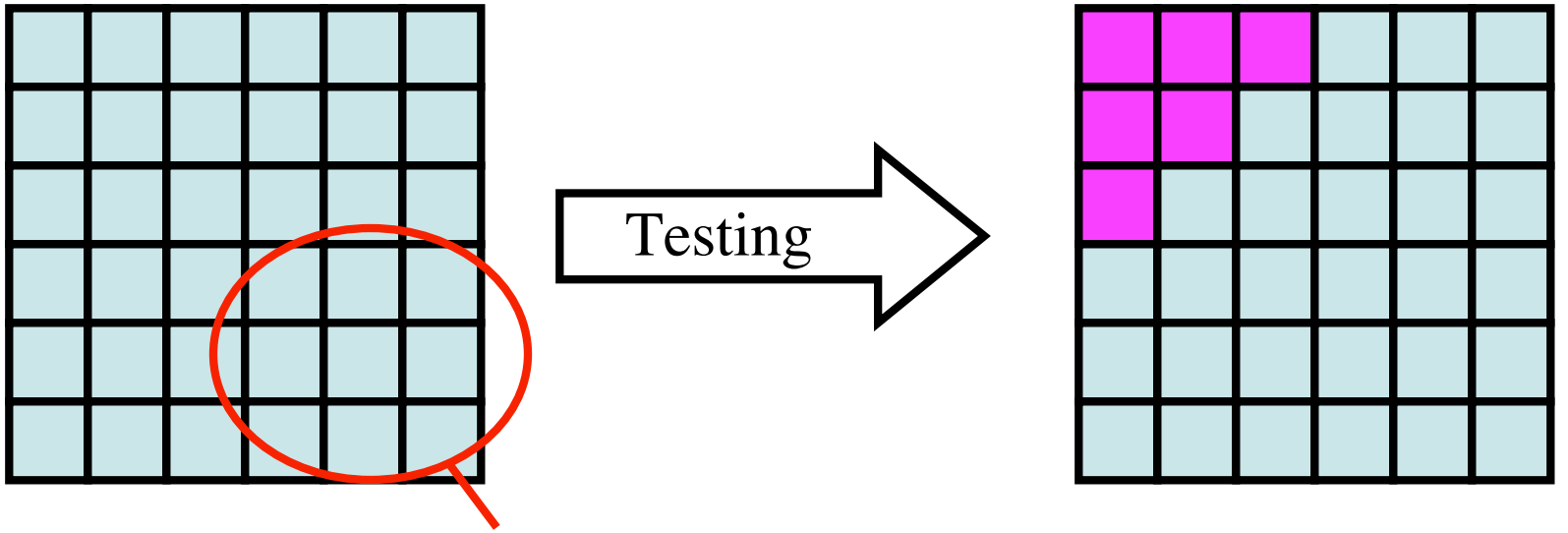
$$N_{testing} \sim N^{\frac{Constant}{p}}$$

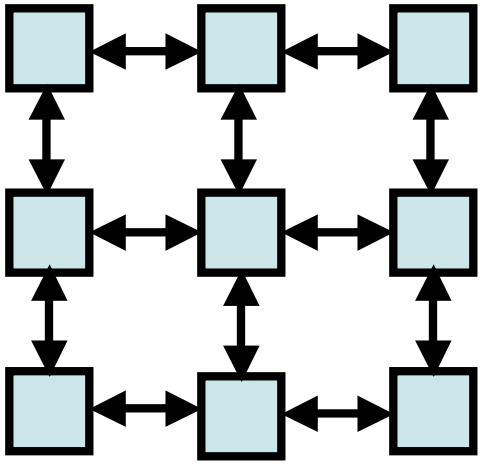


(a) Average wirelength **(b) Largest Delay** **(c) Area**

$$N_{testing} = N^{1.05}$$

CellMatrix™





- Sqrt(N) for 2-D
- Simple logic blocks (2 gates & LUT)

Summary

- Defect Tolerance places measurable overhead on interconnect resources.
- Reconfigurability provides a cheap alternative for yield enhancement.
- Return of local communication based architecture.
- Have to make a judicious decision based on power-delay-area-routability trade offs.