NanoFabrics: Spatial Computing Using Molecular Electronics

Seth Copen Goldstein and Mihai Budiu

Computer Architecture, 2001. Proceedings. 28th Annual International Symposium on
30 June-4 July 2001 Page(s):178 - 189

Introduction

- **CAEN** : Chemically Assembled Electronic Nanotechnology
- A promising alternative to CMOS-based computing under intense investigation
- A form of electronic nanotechnology (EN) which uses **self-alignment** to construct electronic circuits out of nanometer-scale devices that take advantage of quantum-mechanical effects

Introduction

- Claim: CAEN can be harnessed to create useful computational devices with more than **10^{10} gate-equivalents per cm^2**
- The fundamental strategy is to **substitute compile time** (which is inexpensive) **for manufacturing precision** (which is expensive)
- Through a combination of reconfigurable computing, defect tolerance, architectural abstractions and compiler technology

Introduction

- We introduce an architecture based on fabricating dense regular structures, which we call ***nanoBlocks***
- Nanoblocks can be programmed after fabrication to implement complex functions
- We call an array of connected nanoBlocks a ***nanoFabric***

Introduction

- Compared to CMOS, CAEN-based devices have a higher defect density
- Such circuits will thus require built-in defect tolerance
- A natural method of handling defects is to first configure the nanoFabric for self-diagnosis and then to implement the desired functionality by configuring around the defects
- Reconfigurabilty is thus integral to the operation of the nanoFabric

Introduction

- One advantage of nanoFabrics over CMOS-based reconfigurable fabrics (like FPGAs) is that the **area overhead for supporting reconfiguration** is virtually eliminated

Electronic Nanotechnology

- CAEN devices are very small: A single RAM cell will require 100 nm^2 as opposed to 100,000 nm^2 for a single laid out CMOS transistor
- For the CAEN device we assume that the nanowires are on 10nm centers
- A CMOS transistor with a 4:1 ratio in a 70nm process, with no wires attached measures 210nm x 280nm
- Attaching minimally-sized wires to the terminals increases the size to 350nm x 350nm

Electronic Nanotechnology

- A simple logic gate or an static memory cell requires several transistors, separate p- and nwells, etc., resulting in a factor of 10^5 difference in density between CAEN and CMOS (these numbers are not very accurate in my view --Reza)
- CAEN devices use much less power, since very few electrons are required for switching

Fabrication and Architectural Implications

- In the first step, wires of different types are constructed through chemical self-assembly
- The next step aligns groups of wires:
- Also through self-assembly, two planes of aligned wires will be combined to form a two-dimensional grid with configurable molecular switches at the crosspoints

Fabrication and Architectural Implications

- The resulting grids will be on the order of a few microns
- A separate process will create a silicon-based die using standard lithography
- The circuits on this die will provide power, clock lines, an I/O interface, and support logic for the grids of switches
- The die will contain “holes” in which the grids are placed, aligned, and connected with the wires on the die

Fabrication and Architectural Implications

- The precise alignment required to co-locate three wires at the device makes them unsuitable for producing real circuits with inexpensive chemical assembly
- We thus assume that **CAEN devices will be limited to performing logic using two terminal devices; i.e. diode-resistor logic**
- As the active components will be diodes and configurable switches, there will be no inverters
- Because we cannot build inverters, all logic functions will generally compute both the desired output and its complement

Fabrication and Architectural Implications

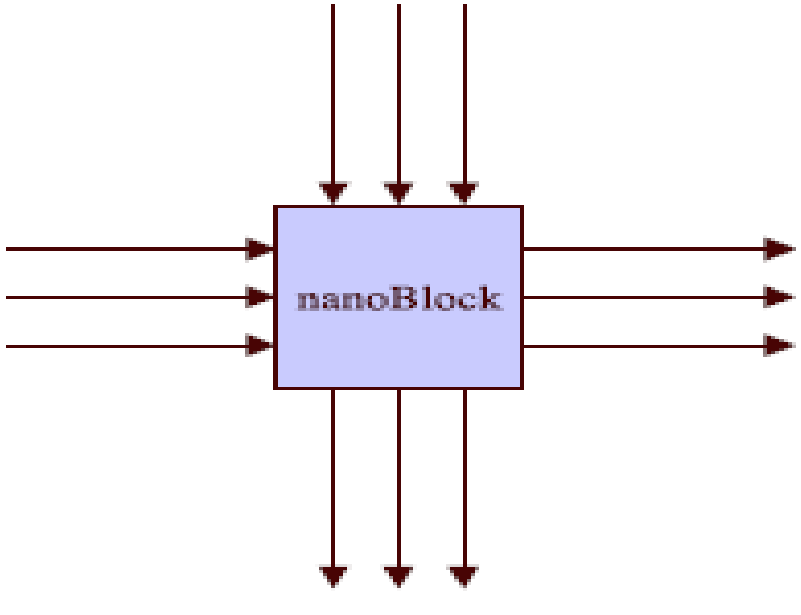
- The lack of a transistor means that **special mechanisms will be required for signal restoration** and for building registers
- Using CMOS to buffer the signals is unattractive for two reasons:
  - First, CMOS transistors are significantly larger and would decrease the density of the fabric
  - Second, the large size of CMOS transistors would slow down the nanoFabric
- We have successfully designed and simulated a **molecular latch** motivated by work in tunnel diodes
- The latch is composed of a wire with two inline NDR molecules at either end
- The latch combined with a clocking methodology, provides signal restoration, latching, and I/O isolation

Fabrication and Architectural Implications

- The fabrication process also disallows the precise alignment required to make end-to-end connections between nanoscale wires
- Our architecture ensures that **all connections between nanoscale wires occur by crossing the wires**

NanoFabric

- The nanoBlocks are logic blocks that can be programmed to implement a three-bit input to three-bit output Boolean function and its complement (see Figure 1a).



(a)

Single nanoBlock showing I/O lines

NanoFabric

- The nanoBlocks are organized into **clusters** (See Figure 2)
- Within a cluster the nanoBlocks are connected to their nearest four neighbors
- **Long wires**, which may span many clusters (long-lines), are used to route signals between clusters.

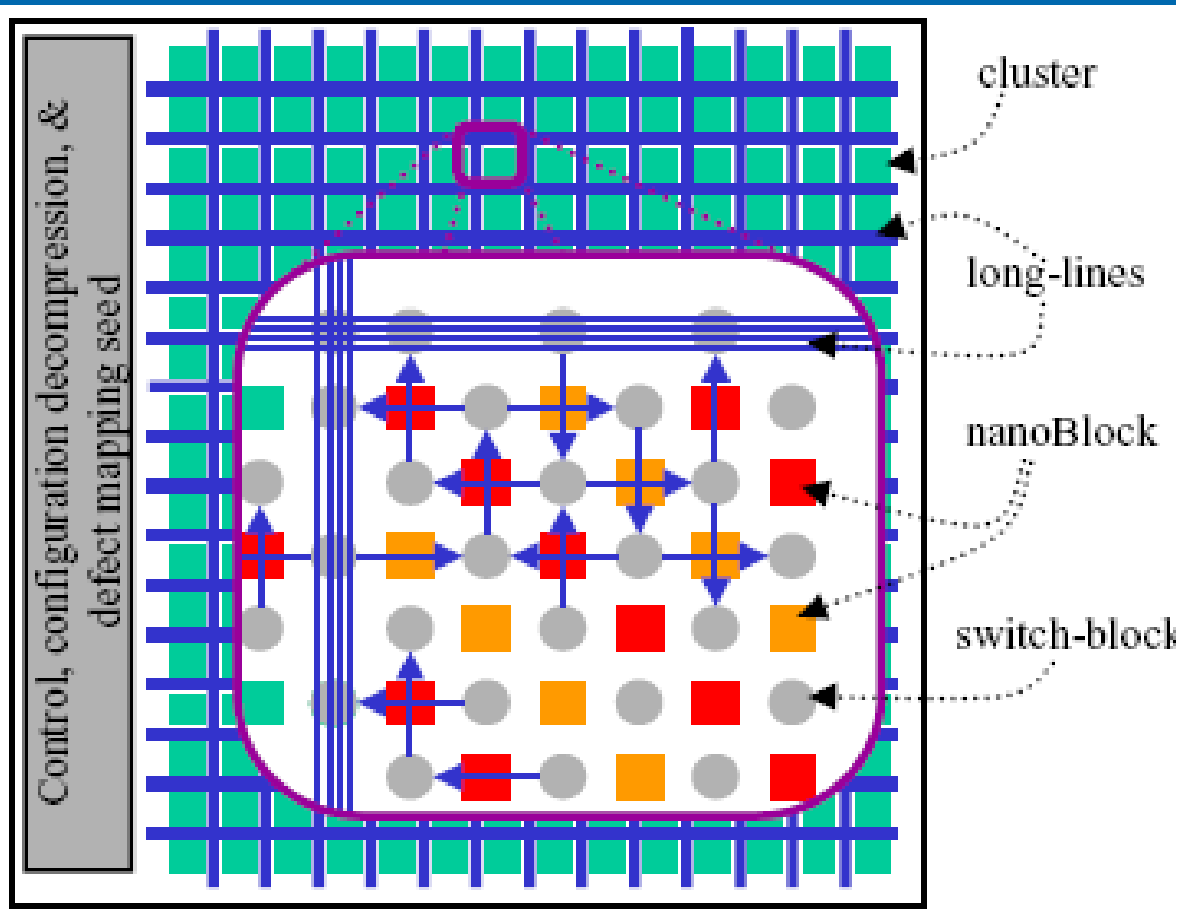
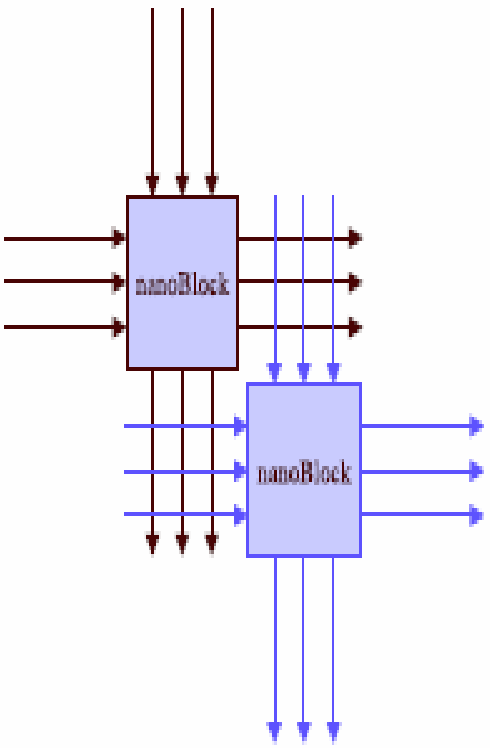


Figure 2. *The layout of the nanoFabric with a partial blowup of a single cluster and some of the adjacent long-lines.*

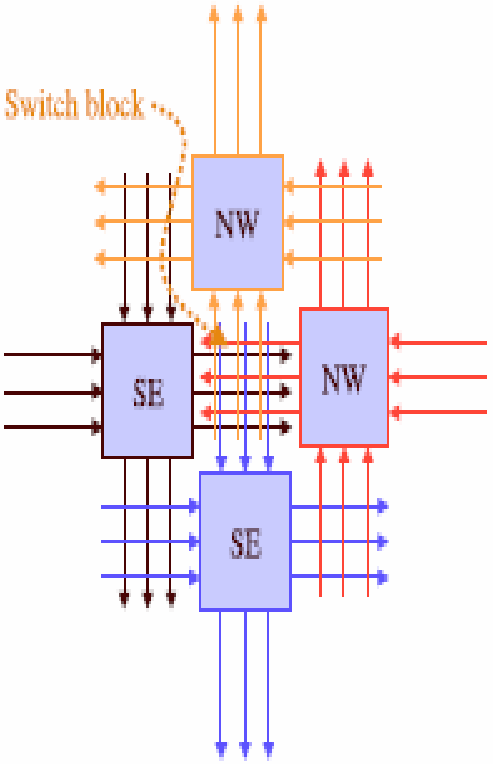
NanoFabric

- Figures 1 (b,c) show how the outputs of one nanoBlock connect to the inputs of another
- We call the area where the input and output wires overlap a *switch block*



(b)

Half of a switchblock



(c)

Switchblock with four surrounding nanoBlocks

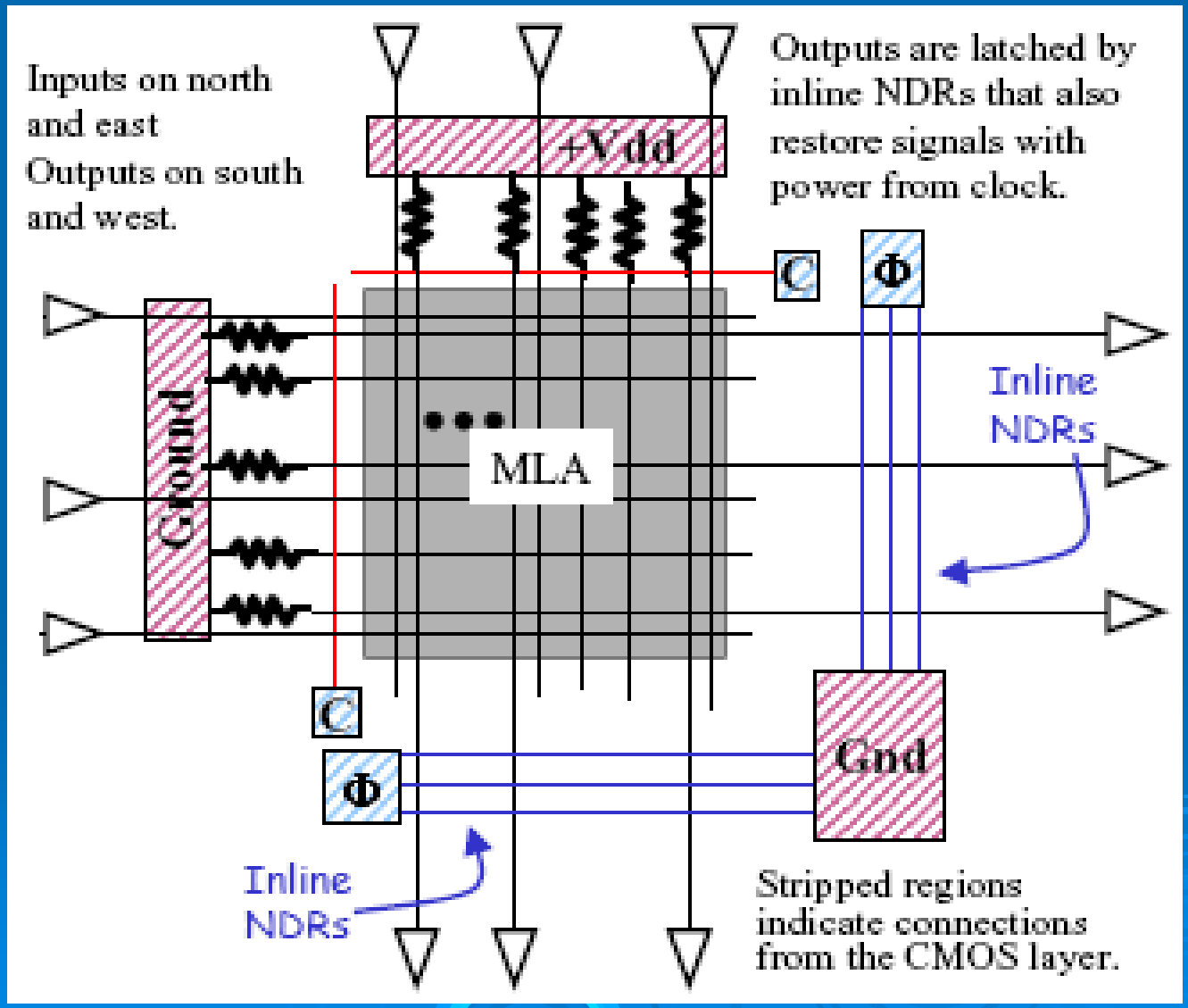
NanoFabric

- As the number of components increases we can increase the number of long lines that run between the clusters. This supports **routability of netlists**
- Each **cluster is designed to be configured in parallel**, allowing configuration times to remain reasonable even for very large fabrics
- The power requirements remain low because we use molecular devices for all aspects of circuit operation
- Finally, because we assemble the nanoFabric hierarchically **we can exploit the parallel nature of chemical assembly**

NanoBlock

- Nanoblock is composed of three sections (see Figure 3):
  (1) the molecular logic array, where the functionality of the block is located
  (2) the latches, used for signal restoration and signal latching for sequential circuit implementation
  (3) the I/O area, used to connect the nanoBlock to its neighbors through the switch block

NanoBlock

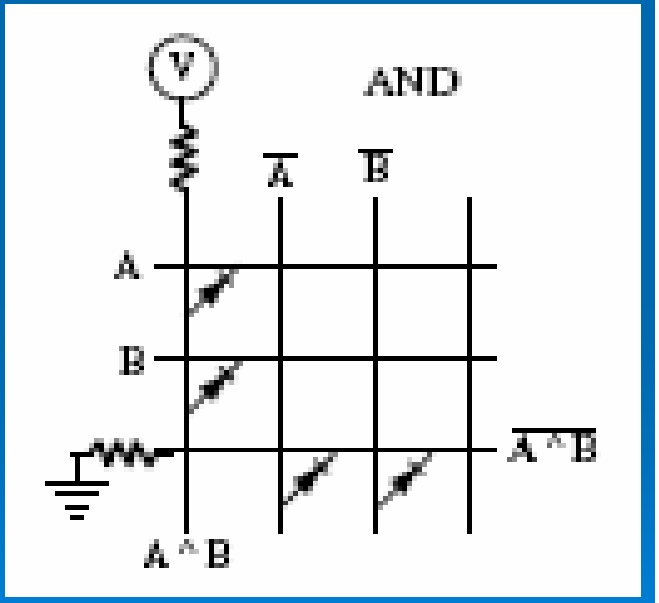

Inputs on north
and east
Outputs on south
and west.
+Vdd
Outputs are latched by
inline NDRs that also
restore signals with
power from clock.
C
Φ
Ground
MLA
Inline
NDRs
C
Φ
Gnd
Inline
NDRs
Stripped regions
indicate connections
from the CMOS layer.

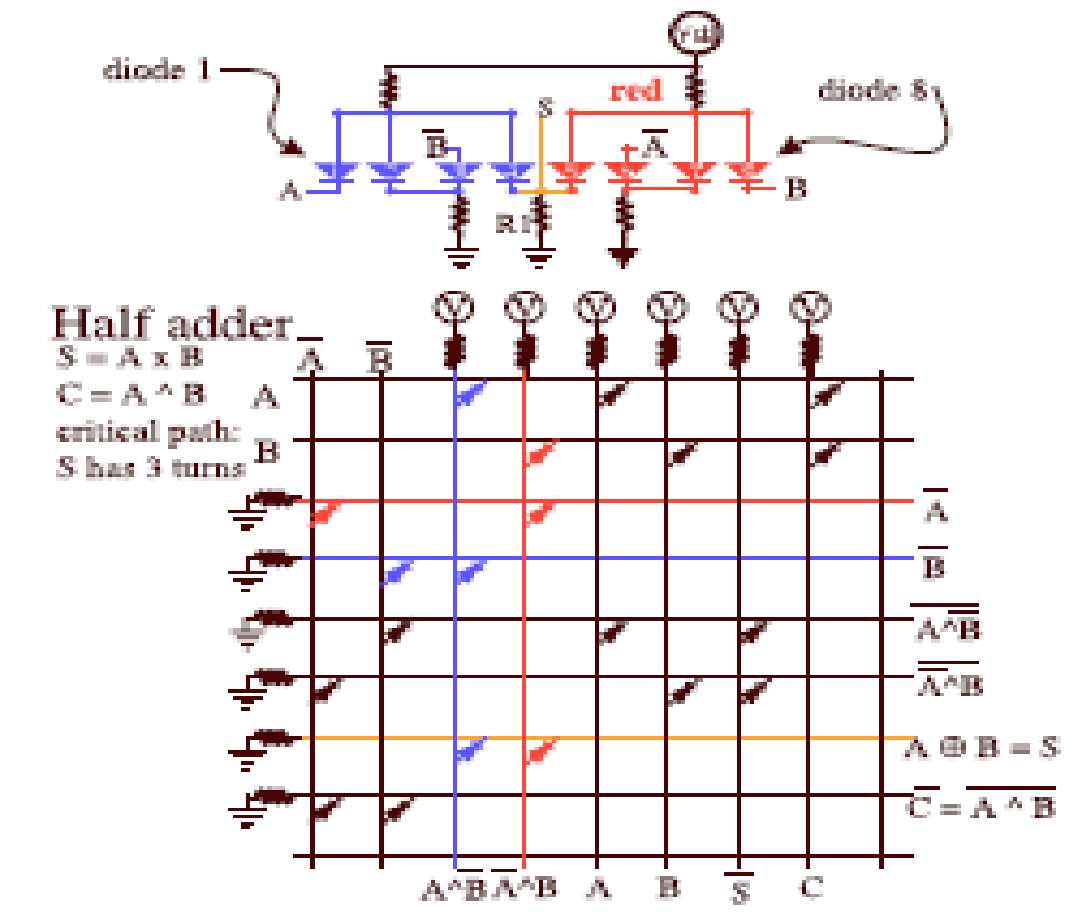
NanoBlock

- The molecular logic array (MLA) portion of a nanoBlock is composed of two orthogonal sets of wires
- At each intersection of two wires lies a configurable molecular switch
- The switches, when configured to be “on”, act as diodes.

NanoBlock

- Figure 4 shows the implementation of an AND gate
- Figure 5 shows the implementation for a half-adder





NanoBlock

- The drawback is that the signal is degraded every time it goes through a configurable switch
- In order to restore signals to proper logic values without using CMOS gates, we will use a molecular latch
- The layout of the MLA and of the switch block makes rerouting easy in the presence of faults
- By examining Figure 5, one can see that a bad switch is easily avoided by swapping wires that only carry internal values

Defect Tolerance

- The nanoFabric is defect tolerant because:
  - **It is regular:** The regularity allows us to choose where a particular function is implemented
  - **It is highly configurable:** The configurability allows us to pick which nanowires, nanoBlocks, or parts of a nanoBlock will implement a particular circuit
  - **It is fine-grained:** The fine-grained nature of the device combined with the local nature of the interconnect reduces the impact of a defect to only a small portion of the fabric
  - **It has a rich interconnect:** Finally, the rich interconnect allows us to choose among many paths in implementing a circuit

Defect Tolerance

- Thus, with a defect map we can create working circuits on a defective fabric
- Researchers on the Teramac project faced similar issues
- Because the number of tests required to isolate any specific defect does not grow as the total size of the device grows, the computational work needed to test a device is at worst linear in the size of the device (….. ?!!)

Defect Tolerance

- Once a defect map has been generated the fabric can be used to implement arbitrary circuits
- While the molecules are expected to be robust over time, inevitably new defects will occur over time
- Finding these defects, however, will be significantly easier than doing the original defect mapping because the unknown defect density will be very low

Configuration

- A molecular switch is configured when the voltage across the device is increased outside the normal operating range
- There are two factors that contribute to the configuration time:
  - The first factor is the time that it takes to download a configuration to the nanoFabric
  - The second factor is the time that it takes to distribute the configuration bits to the different regions of the nanoFabric

Configuration

- The fabric has been designed so that the clusters can be programmed in parallel
- A very conservative estimate is that we can simultaneously configure one nanoBlock in each of 1000 clusters in parallel
- Our preliminary calculations indicate that we can load the full nanoFabric, which is comprised of 10^9 configuration bits at a density of 10^{10} configuration bits/cm^2, in less than one second (.. ??)

Putting It All Together

- SPICE simulations show that a nanoBlock configured to act as a half-adder can operate at between 100MHz and 1GHz
- Preliminary calculations show that the fabric as a whole will have a static power dissipation of 1.2 watts and dynamic power consumption of 4watts at 100Mhz