

# Harnessing the Power of Parallel Computation on the IBM BlueGene/L to Analyze Complex Digital and RF Systems

Raj Mittra

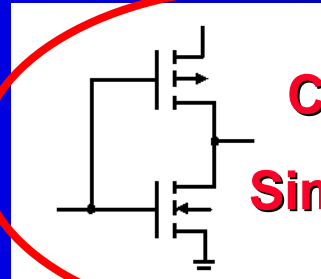
Electromagnetic  
Communication Laboratory

Penn State University

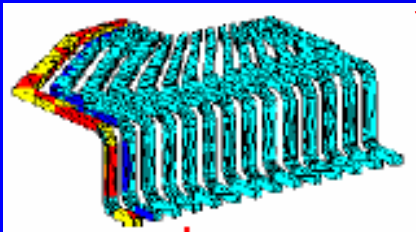
The International Technology Roadmap for Semiconductors (ITRS) projects the need for package delay accuracy of 1% of the off-chip clock frequency. With projected on-chip speeds of 15/33 GHz by 2010/2015, the off-chip interconnect is projected to reach 10/30 Gbps by 2010/2015

# Introduction

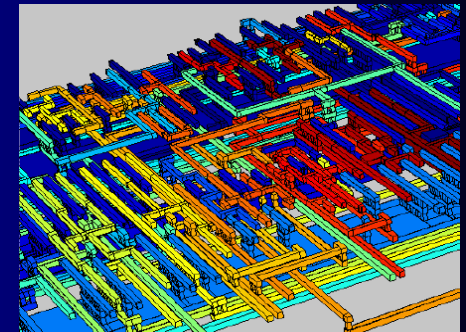
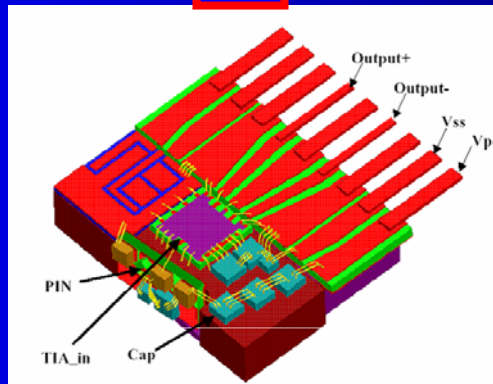
## Incorporating EM Structures into Circuit Simulation



**Circuit  
Simulator**

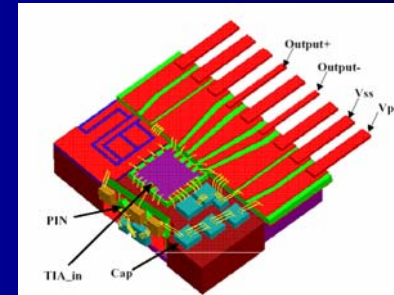
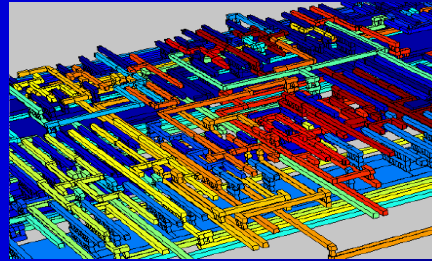
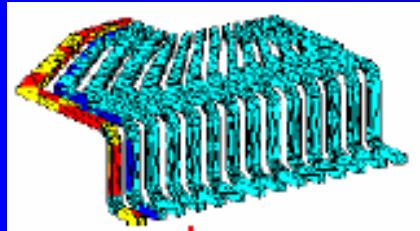


**Packages**



**Interconnects**

# Introduction (cont'd)



**EM simulation, Measurement,  
Circuit model extraction**

**A large equivalent  
circuit network**

**Frequency  
response with a  
large number of  
sampling points**

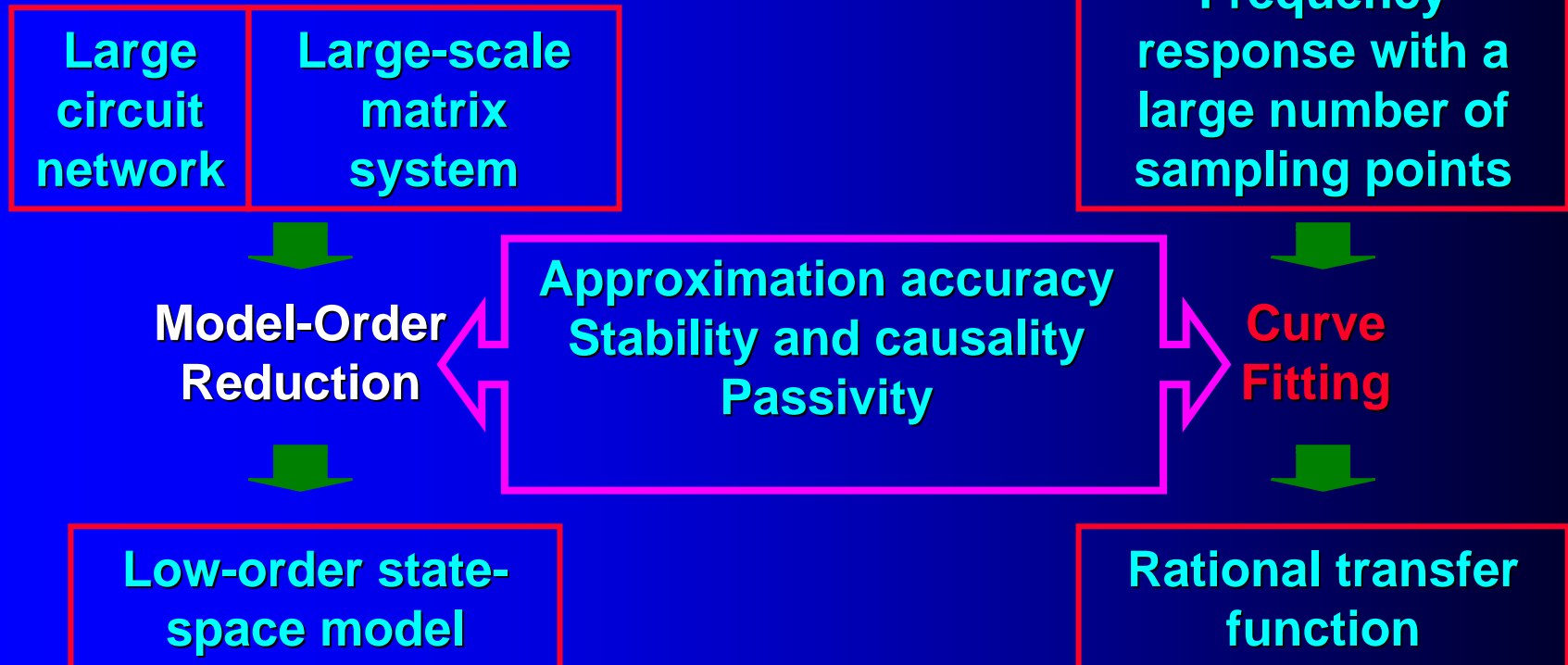
**A large-scale  
matrix system**

**Not Efficient for Circuit Simulation**

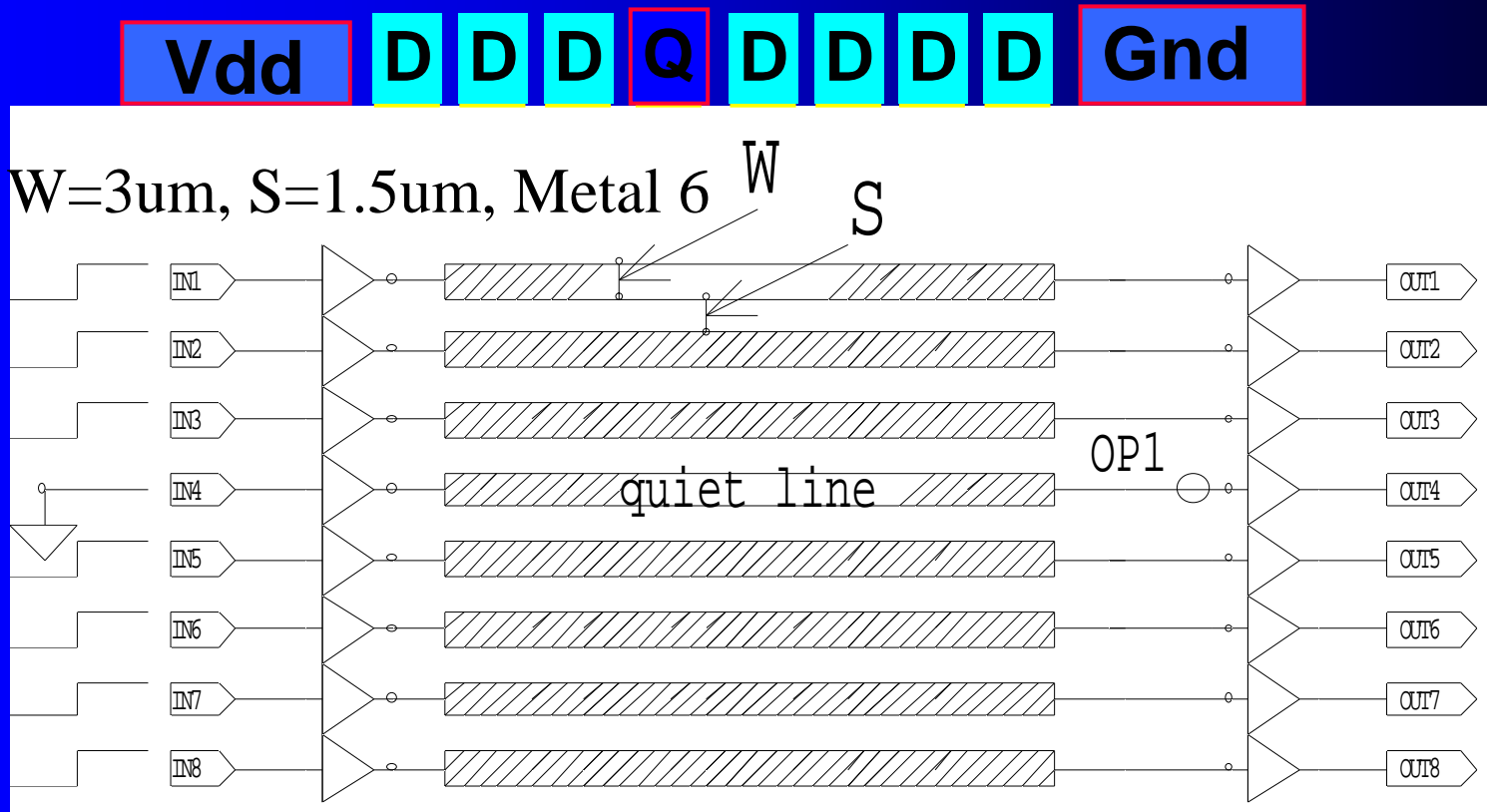
# Introduction (cont'd)

State-Space Model  
Available

Black Box Modeling



# Why is Inductance important?

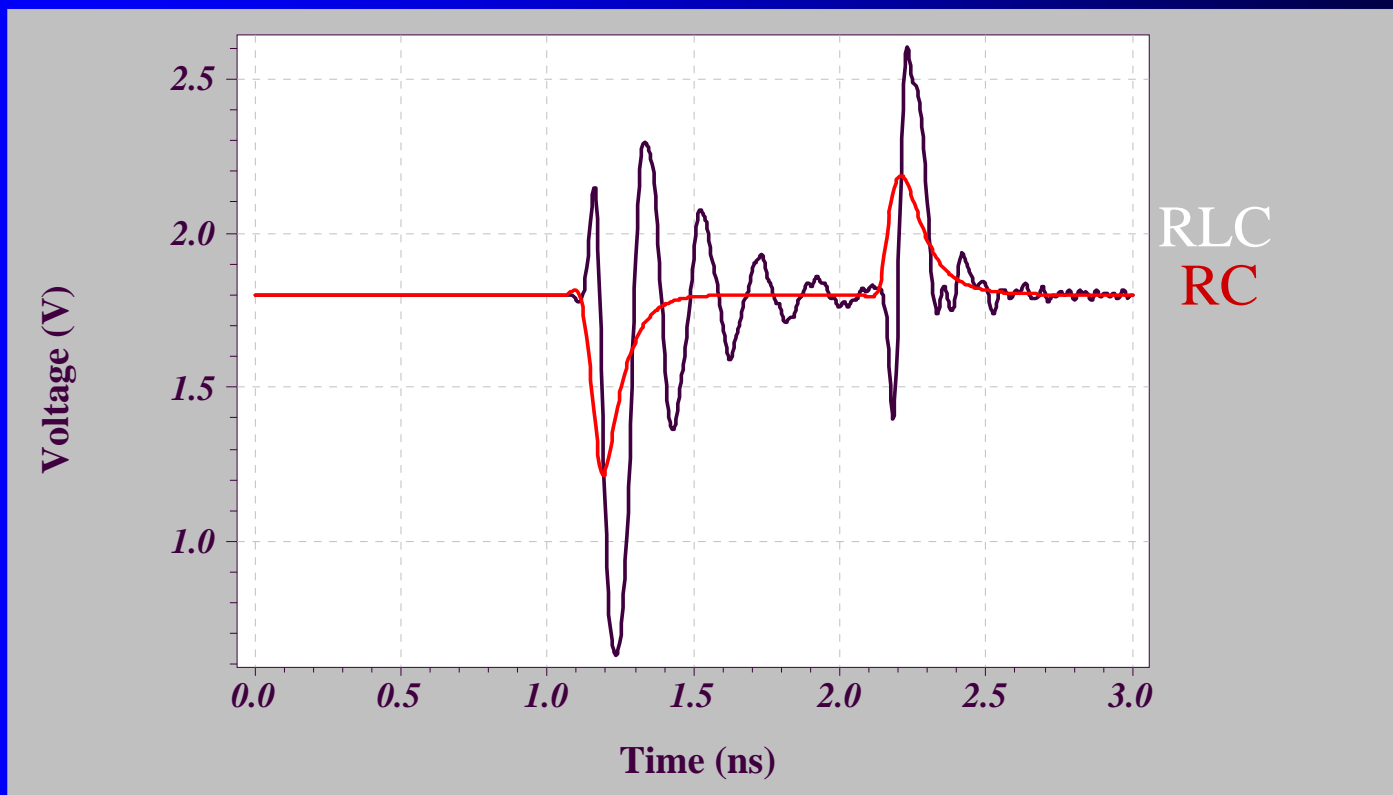


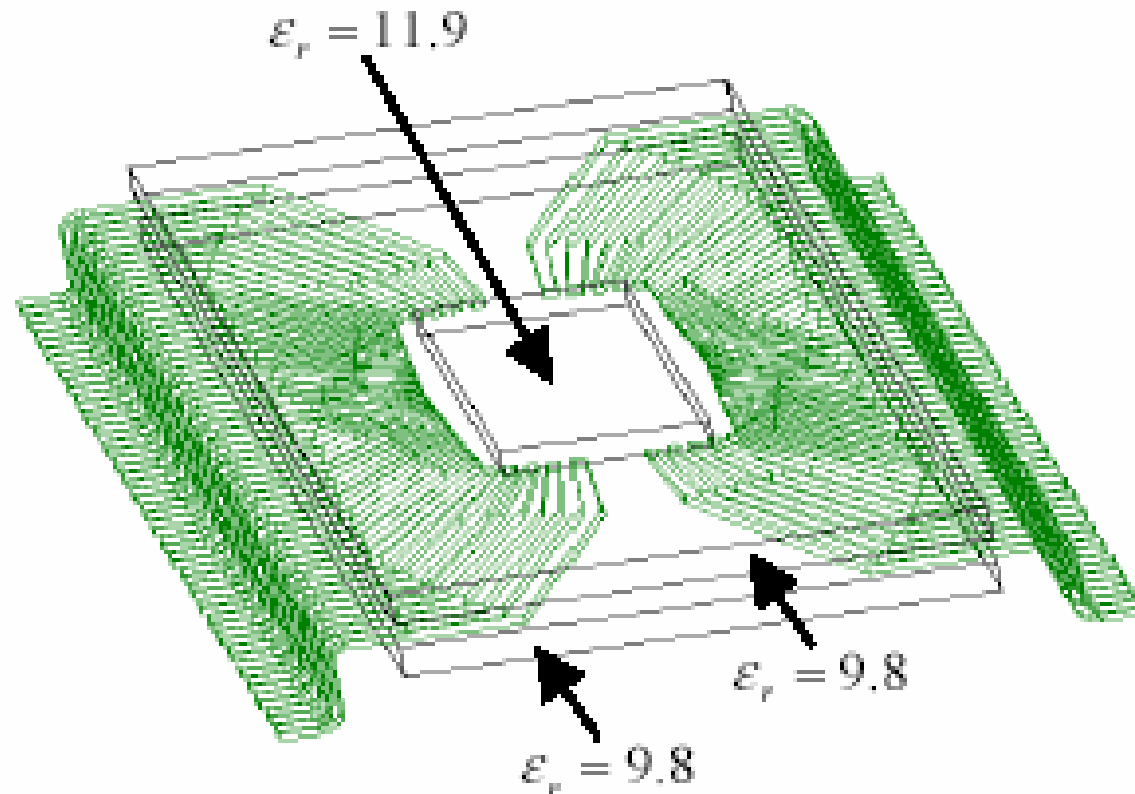
**Eight-Bit Buss Cross-Talk Example**

**Tech=0.18um, Vdd=1.8v, Wire length=3000um,**

# Noise induced on the quiet line?

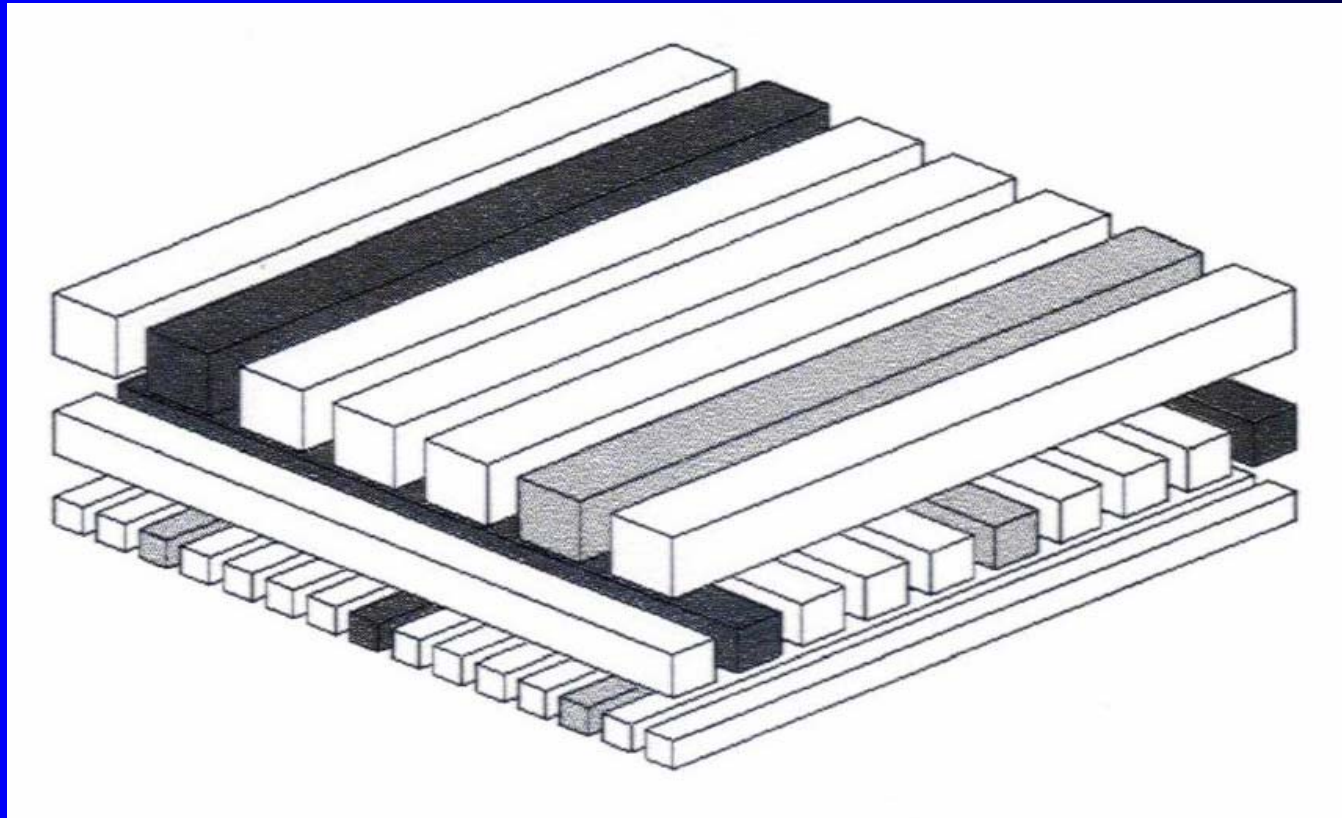
- Peak Noise including L = 1.17V (V<sub>dd</sub>=1.8V)
- Peak Noise ignoring L = 0.54V      0.18um tech
  - Do not model L, miss the glitch!



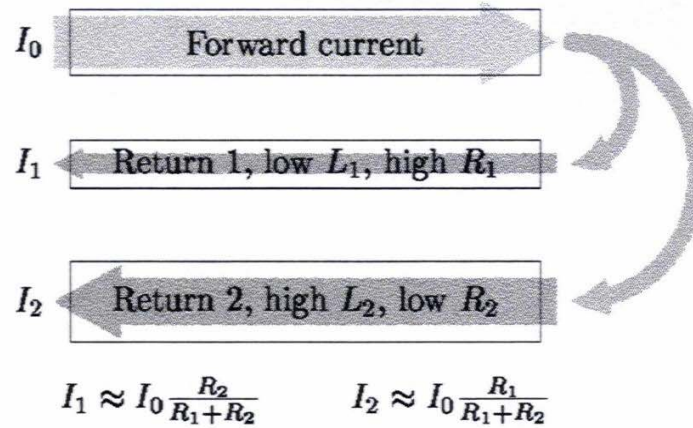


**Figure 7: Package structure  $280\mu\text{m} \times 680\mu\text{m} \times 56\mu\text{m}$  with 56 leads. The package is sandwiched by a top and a bottom  $\text{Al}_2\text{O}_3$  ceramic layer ( $\epsilon_r = 9.8$ ) of thickness  $20\mu\text{m}$ . A dielectric slab of  $\epsilon_r = 11.9$  and thickness of  $10\mu\text{m}$  is placed in the space between the leads ( $96\mu\text{m} \times 220\mu\text{m}$ ).**

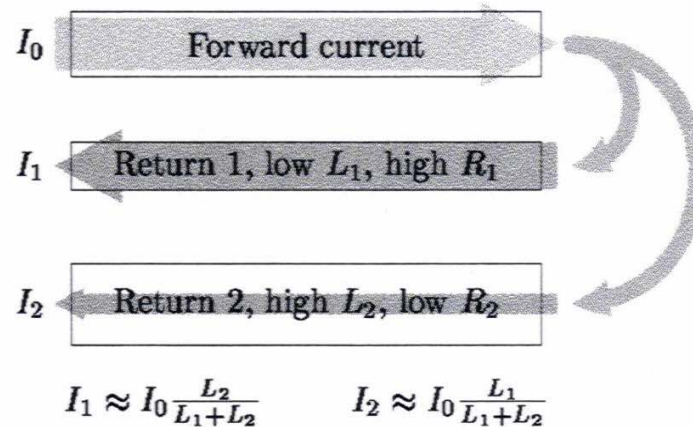




**Fig. 1. A multilayer interconnect with the power distribution grid highlighted; the ground lines are light gray, the power lines are dark gray, and the signal lines are white.**

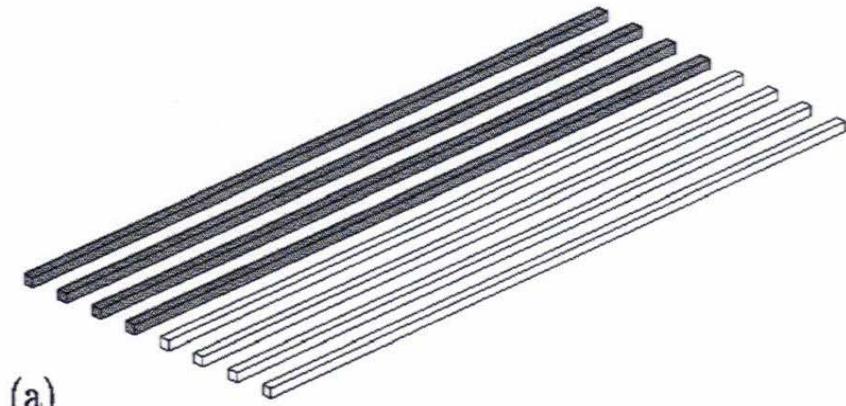


(a) Low frequency,  $R \gg j\omega L$

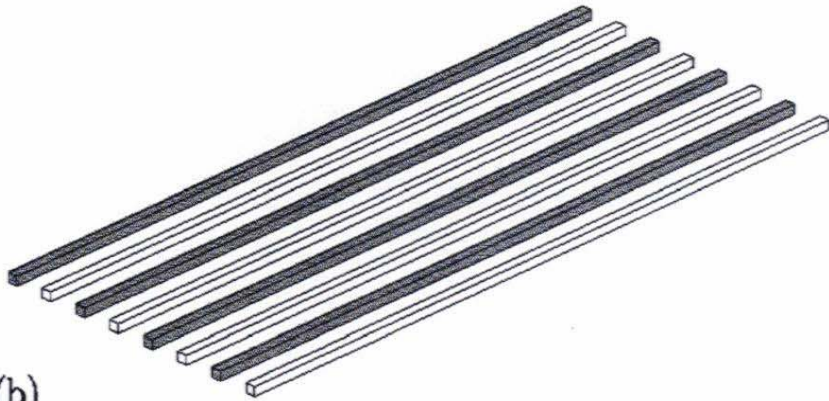


(b) High frequency,  $R \ll j\omega L$

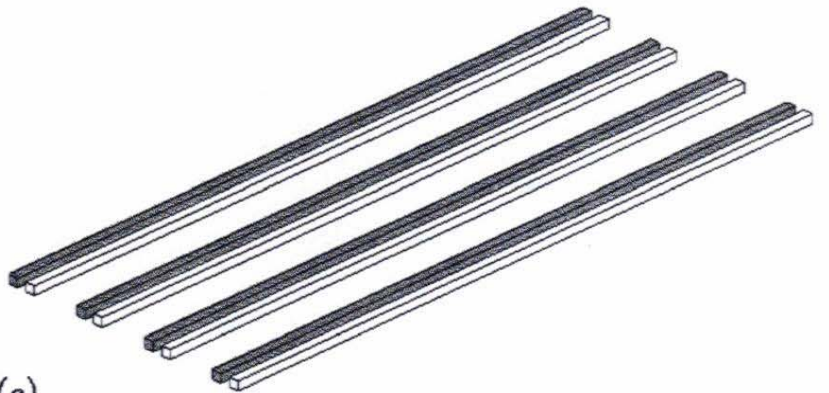
Fig. 3. Current loop with two alternative current return paths. The forward current  $I_0$  returns both through return path one with resistance  $R_1$  and inductance  $L_1$ , and return path two with resistance  $R_2$  and inductance  $L_2$ . In this structure,  $L_1 < L_2$  and  $R_1 > R_2$ . At low frequencies (a), the path impedance is dominated by the line resistance and the return current is distributed between two return paths according to the resistance of the lines. Thus, at low frequencies, most of the return current flows through the return path of lower resistance, path two. At very high frequencies (b), however, the path impedance is dominated by the line inductance and the return current is distributed between two return paths according to the inductance of the lines. Most of the return current flows through the path of lower inductance, path one, minimizing the overall inductance of the circuit.



(a)



(b)

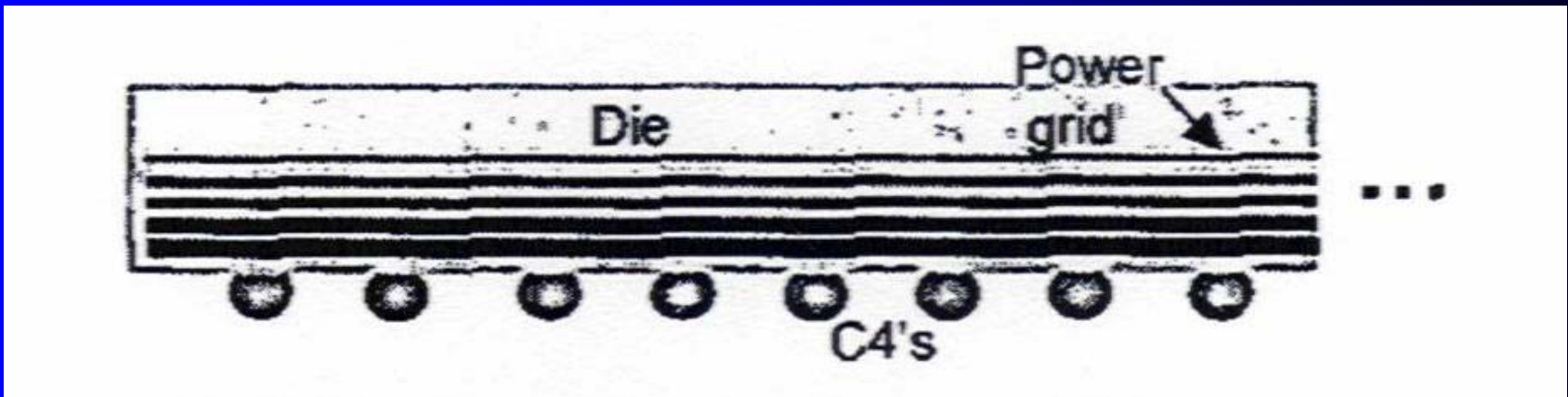


(c)

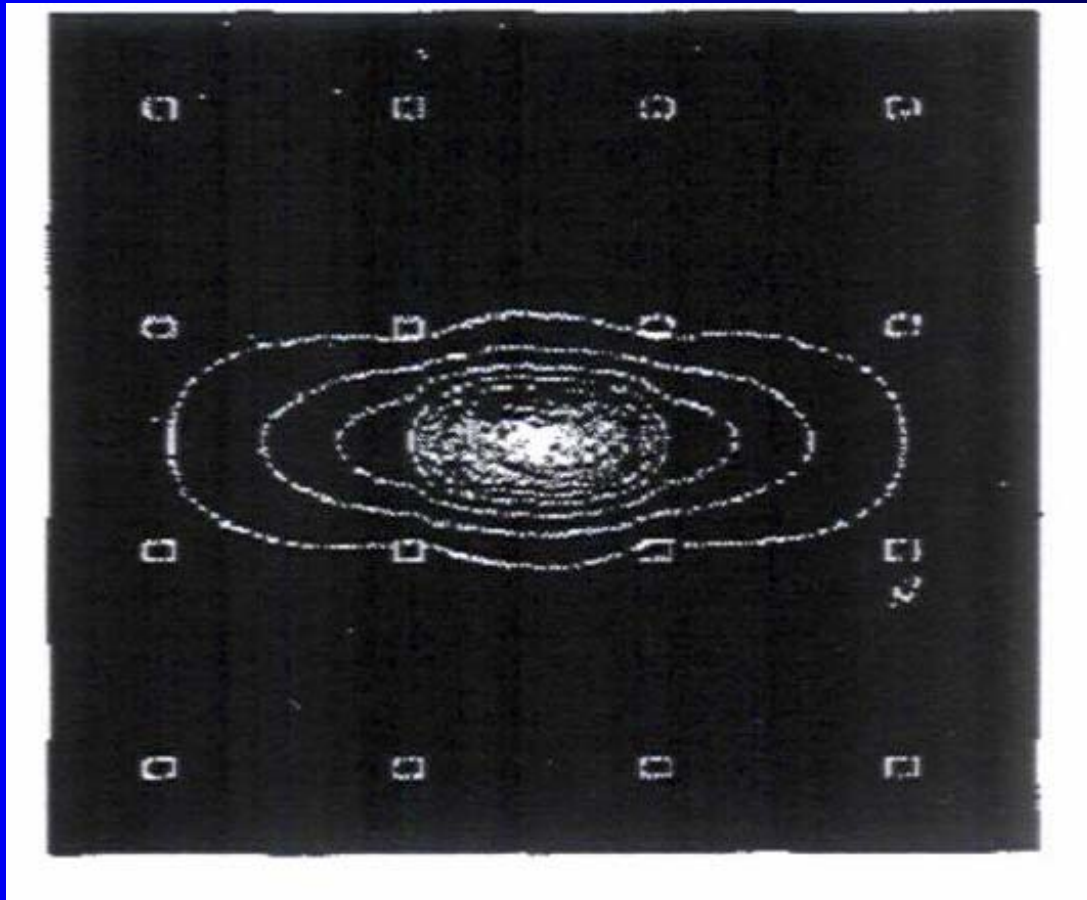
Power/ground grid structures under investigation: (a) a noninterdigitated grid, (b) a grid with the power lines interdigitated with the ground lines, and (c) paired grid, the power and ground lines are in close pairs. The power lines are gray colored, the ground lines are white colored.

# Fast Flip-chip Power Grid Analysis Via Locality and grid Shells

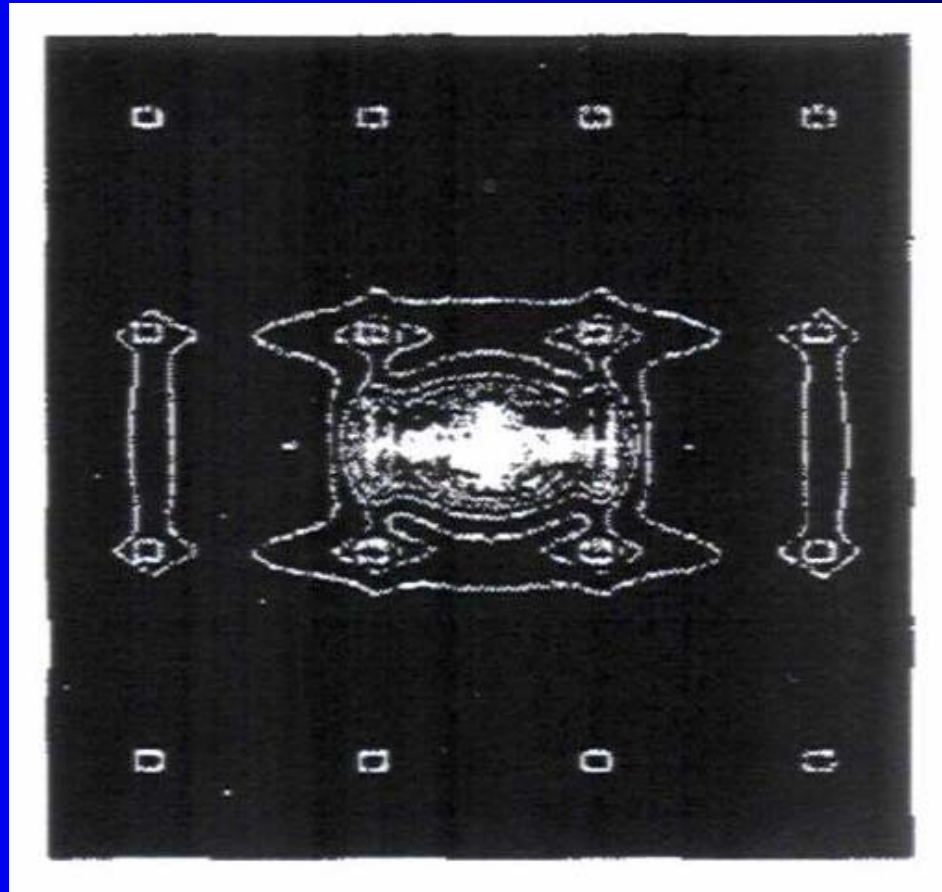




**Fig. 1. Flip-chip die showing C4 bump connections.  
(from Chiprout)**



**Fig. 2. M2 voltage contour map of a Pentium® grid within a 4x4 array of C4 bumps.**



**Fig. 3. M4-M5 via Currents.**



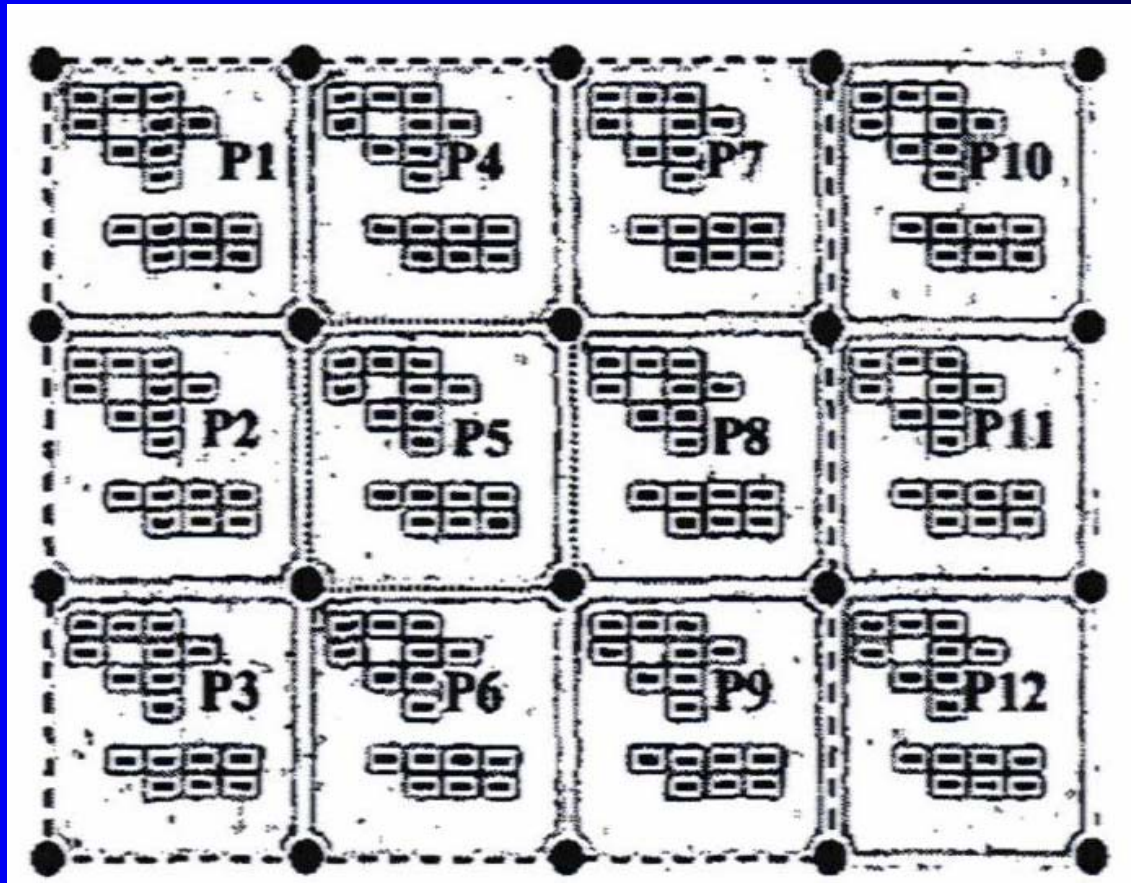
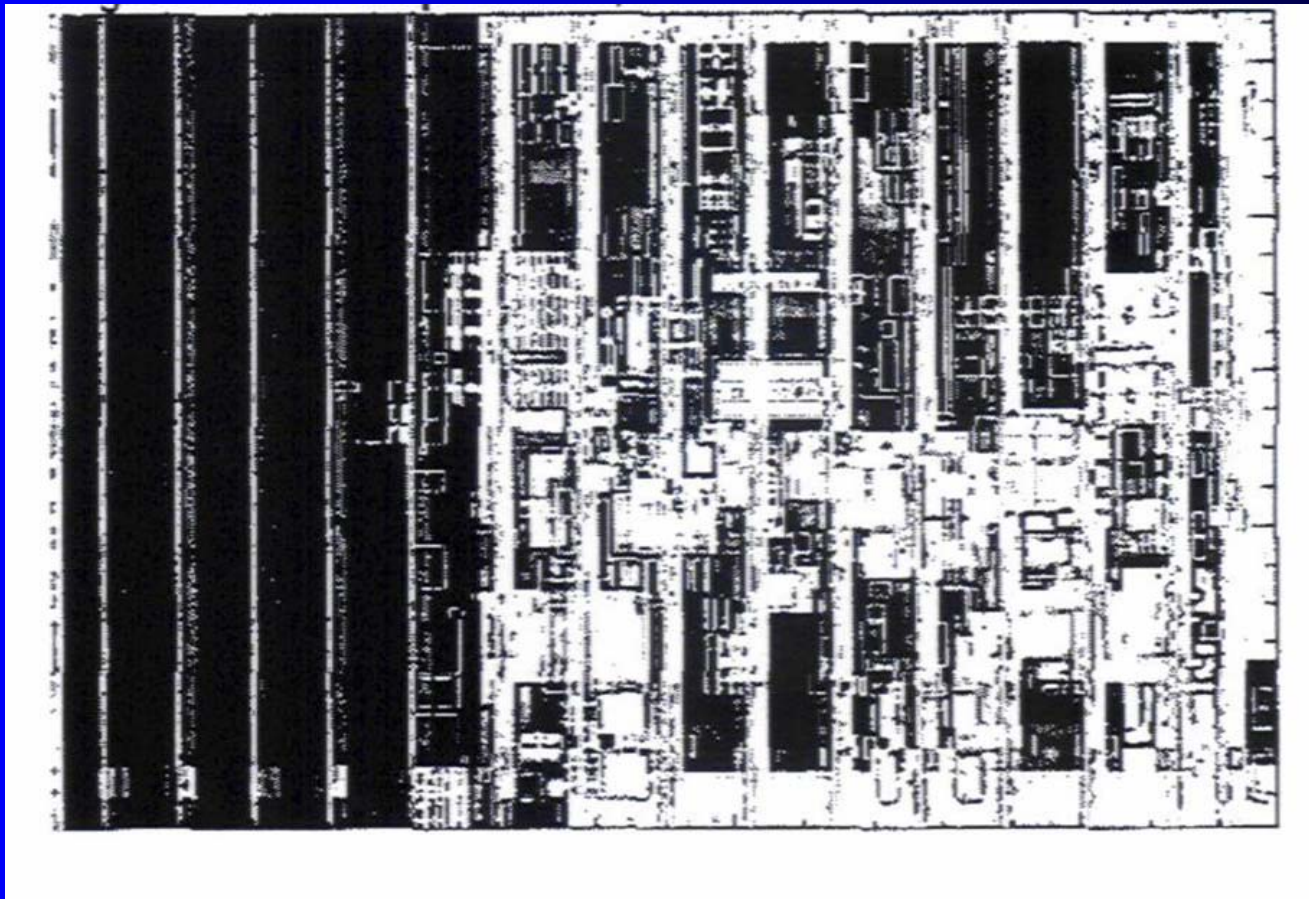


Fig. 4. Current source partitions with P5 source partition high-lighted, with its power grid shell, P1-P9.



**Fig. 5. Currents of a complete Pentium® microprocessor power grid model with top 4 metals, obtained in less than 30 minutes.**

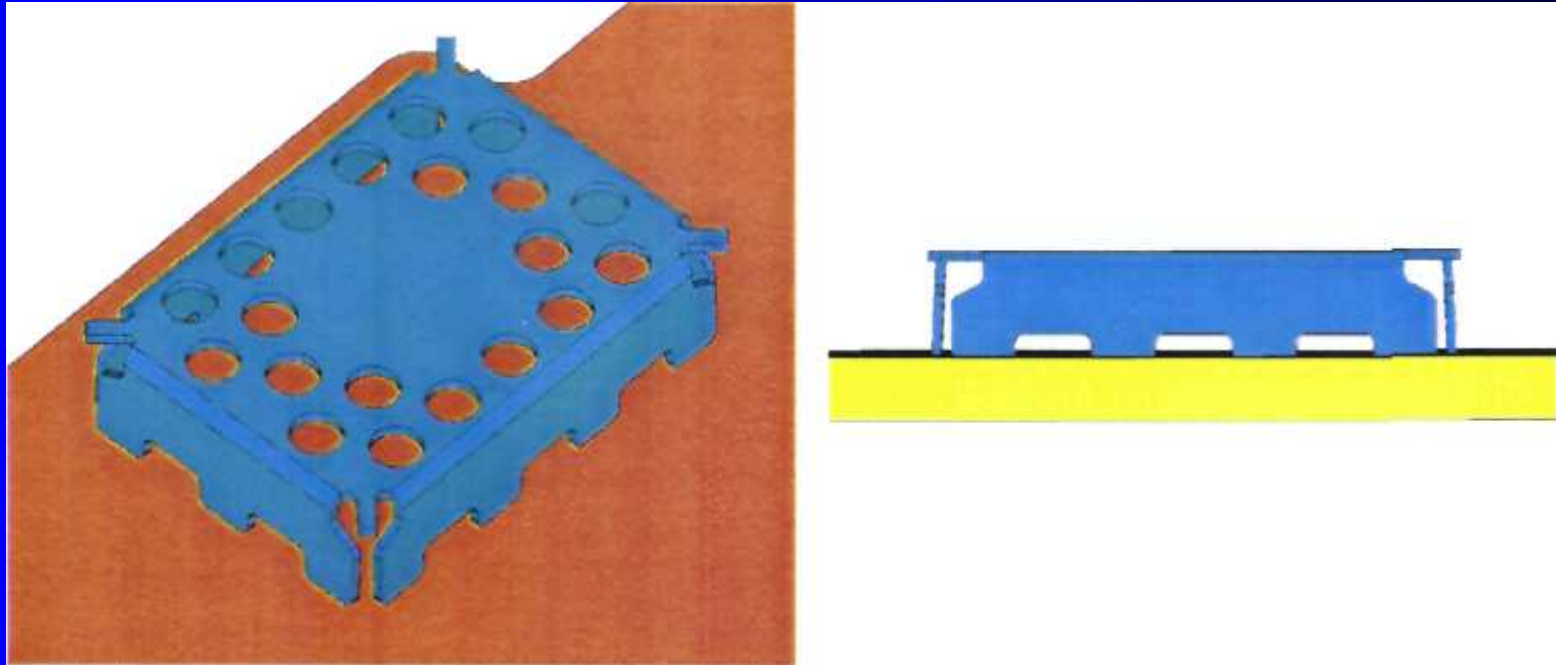
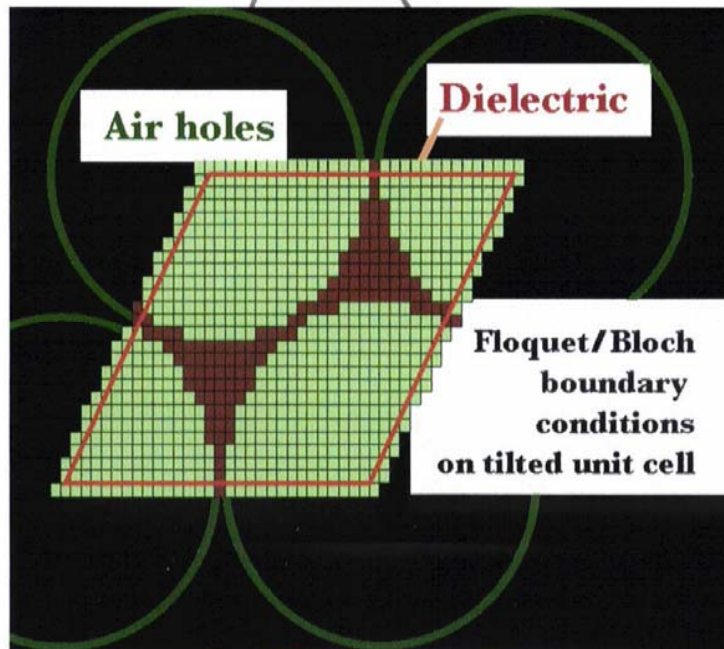
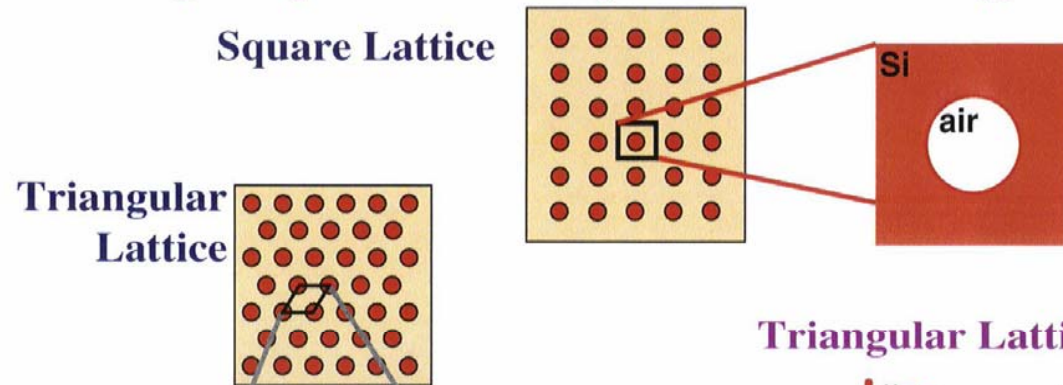


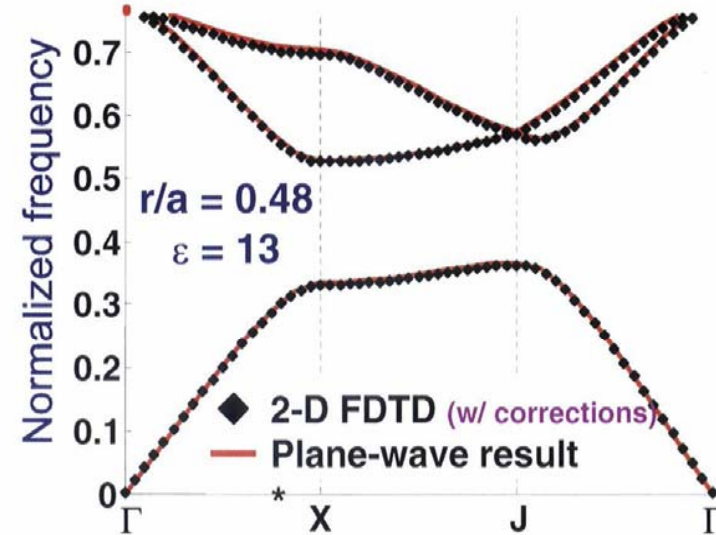
Fig. 1 Mechanical mode! of the shield and PCB



# Example: photonic crystal band-diagrams



## Triangular Lattice – TE bands



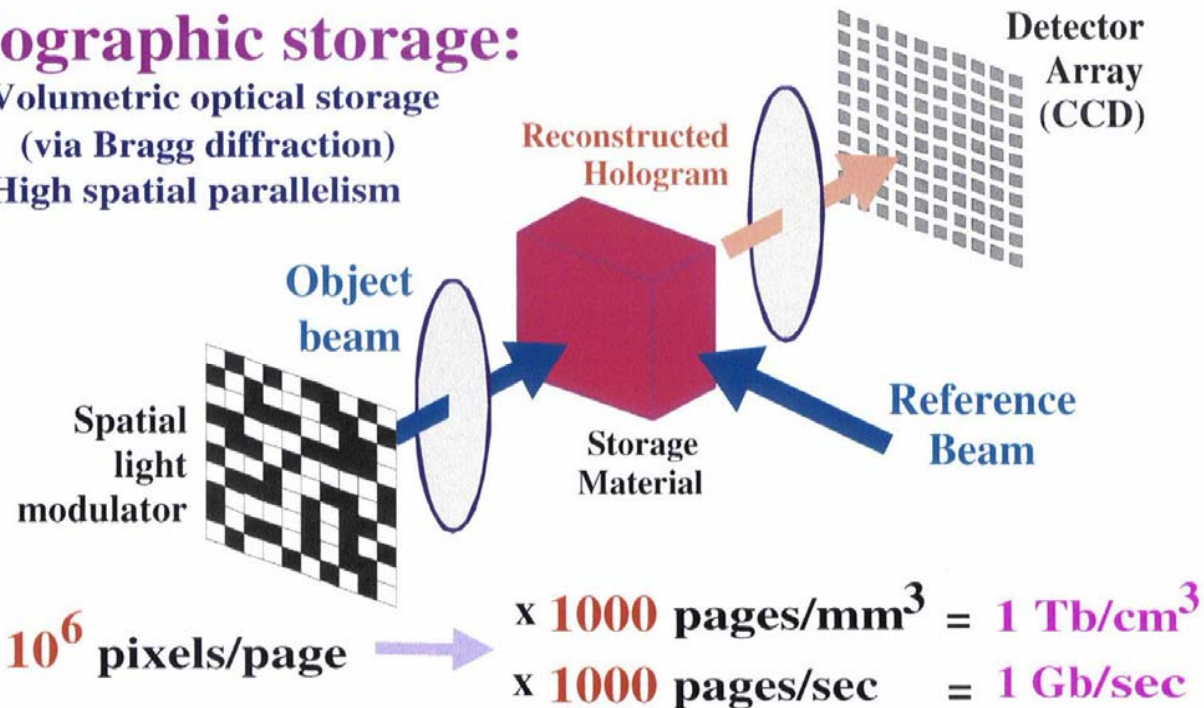
\* MIT "Photonic bands" code – thanks to Bob Shelby of IBM Almaden for his help...

**IBM** Almaden Research Center

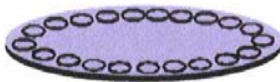
Geoffrey W. Burr (IBM Almaden Research Center, San Jose, CA 95120), "Numerical modeling for nanophotonics design."

# Holographic storage:

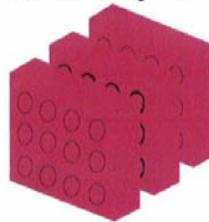
- Volumetric optical storage (via Bragg diffraction)
- High spatial parallelism



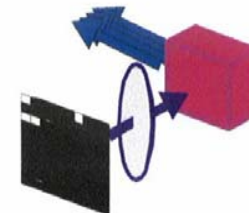
**WORM system**  
(moving media,  
for archival storage)



**Read/write system**  
(stationary media  
for low-latency storage)



**Content-addressable storage**  
(for rapidly searching  
massive databases)



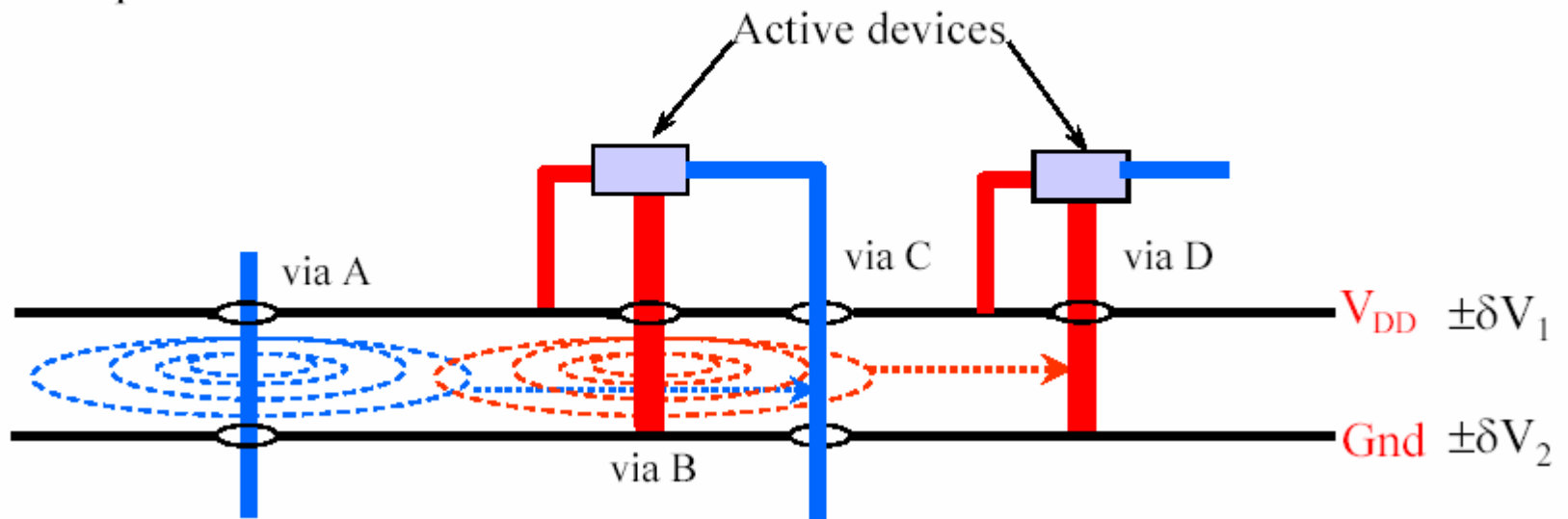
G. W. Burr

IBM Almaden Research Center

Geoffrey W. Burr (IBM Almaden Research Center, San Jose, CA 95120), "Numerical modeling for nanophotonics design."

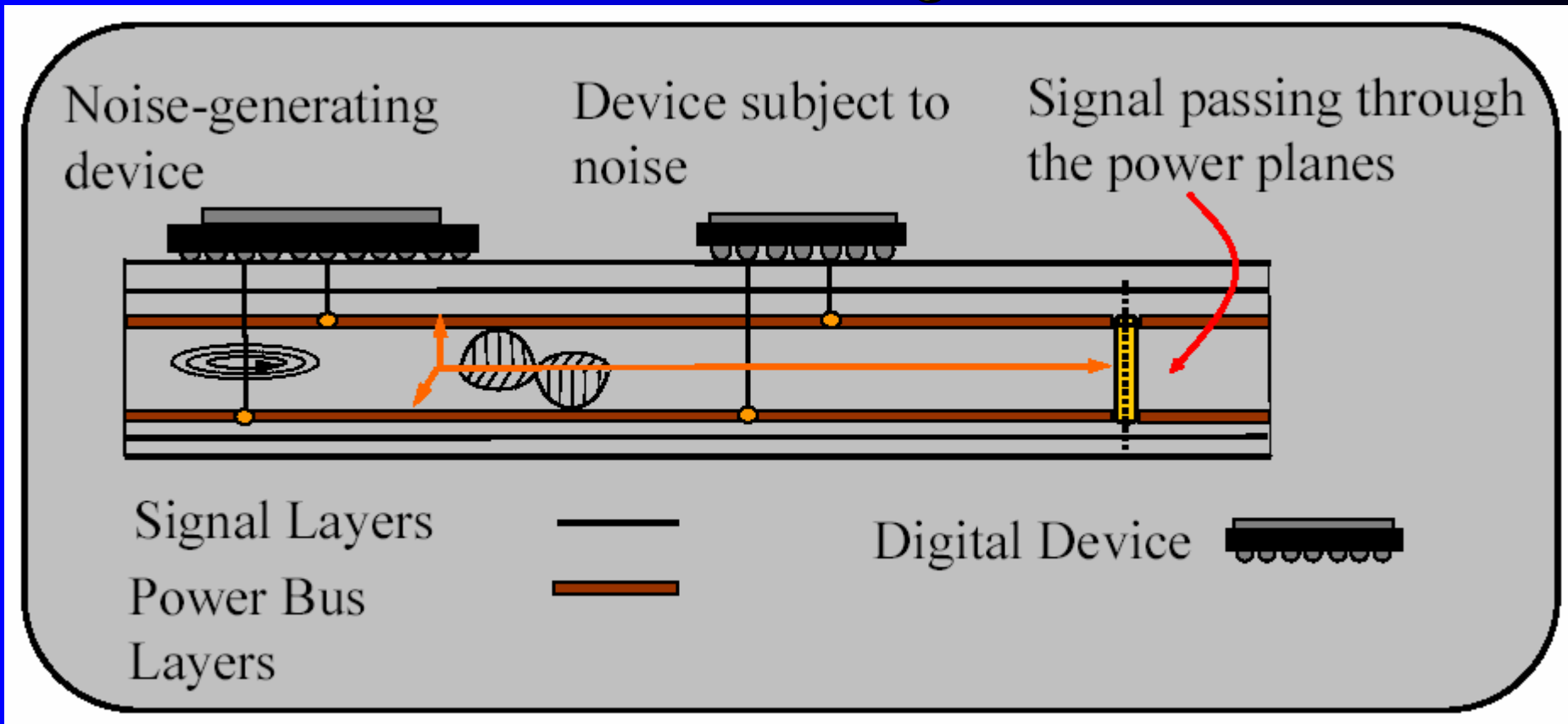
# EM Noise in PCB

— signal  
— power



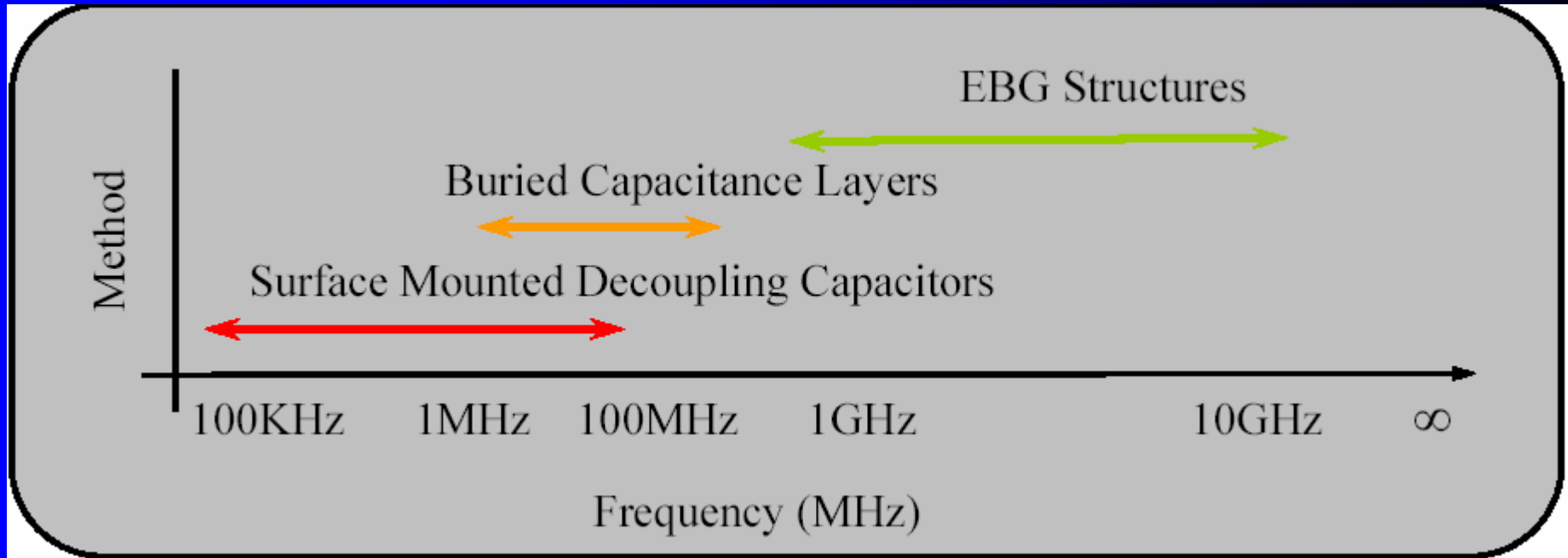
# Motivation

## Problem: Switching noise ....



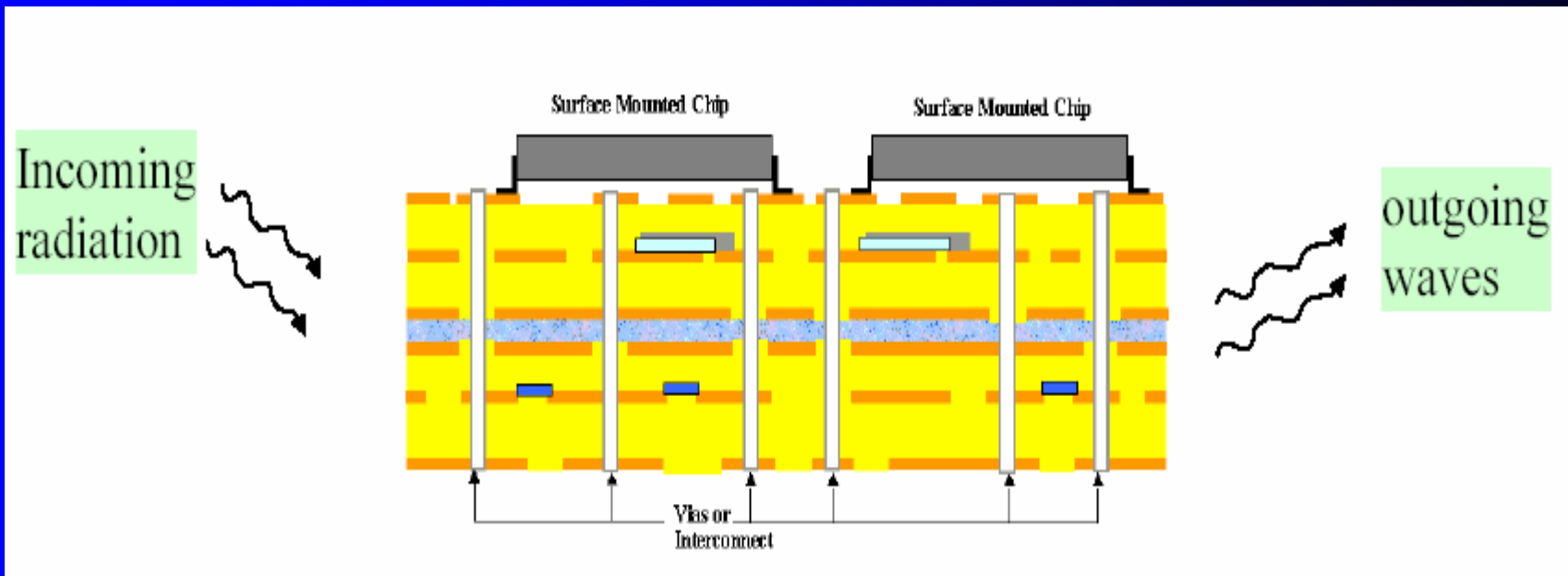
- How to mitigate the noise beyond 500MHz

# State of Current Technology



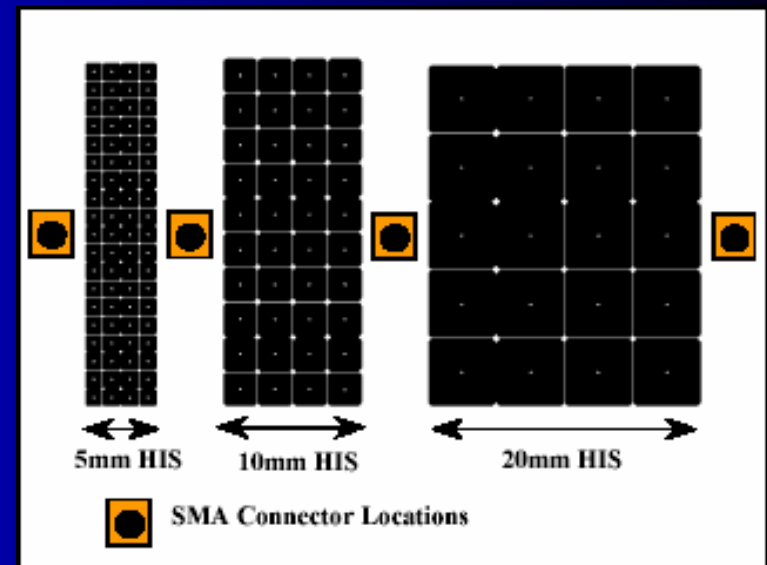


# Coupling to Sensitive Devices in a Multi-Layer Stack up

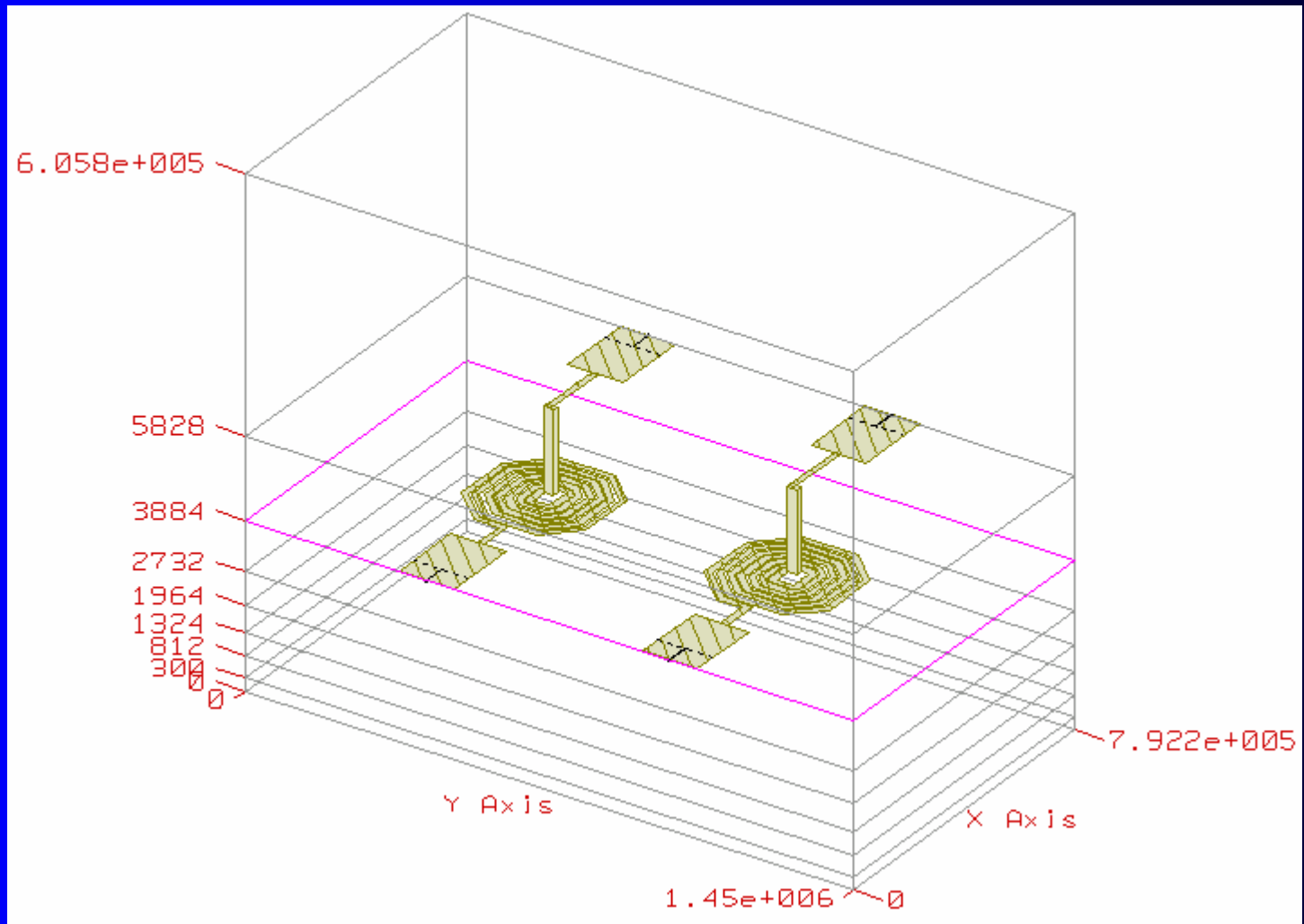


# Wideband Noise Mitigation in PCBs

*Concept: cascaded  
filter design*



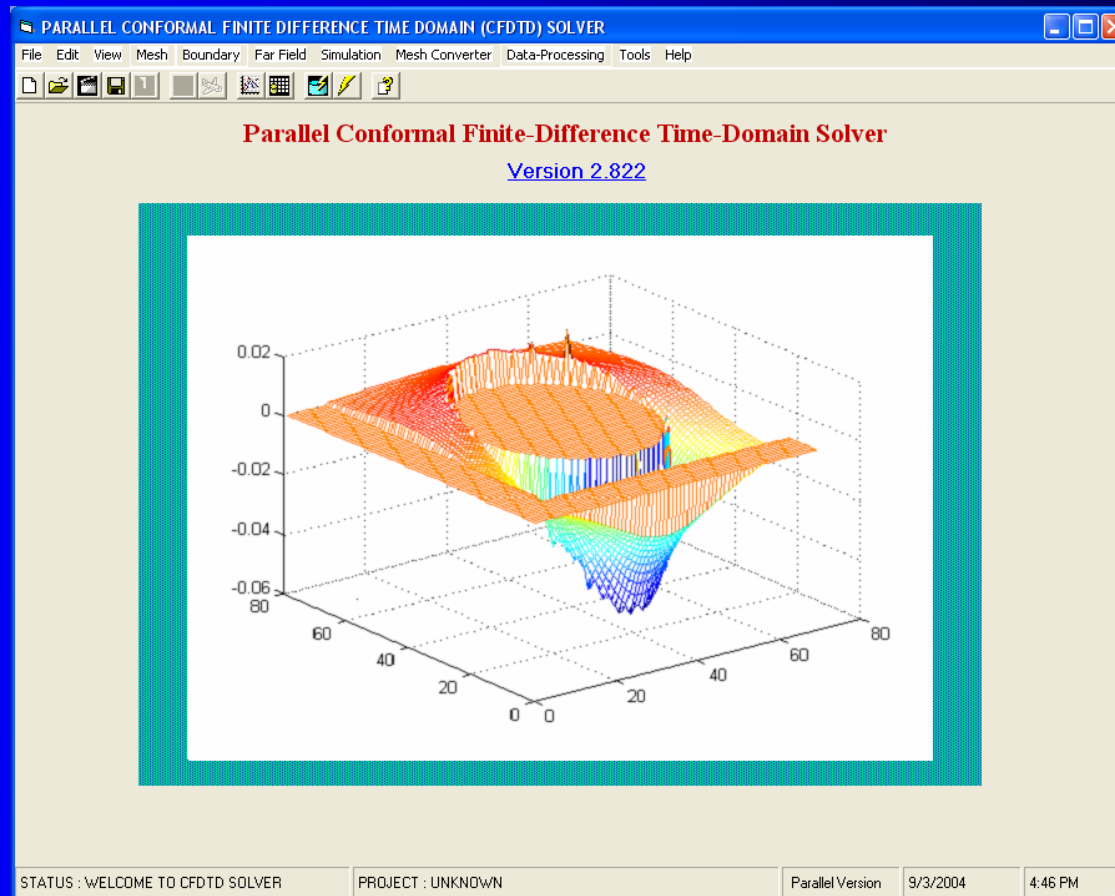
# Spiral Inductors



## Details of the Geometry

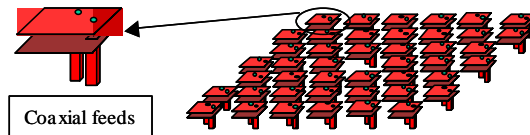
- **Box Dimension:**
- $792150nm$  along  $x$ ;  $1450210nm$  along  $y$ ; 159 divisions along  $x$ ; 500 divisions along  $y$ .
- **Layer stackup** (starting from bottom{thickness,  $\epsilon_r$ }):  
{ $300nm, 2.9$ }; { $512nm, 2.9$ }; { $512nm, 2.9$ }; { $640nm, 2.9$ }; { $768nm, 2.9$ };  
{ $1152nm, 2.9$ }; { $1944nm, 2.9$ }; { $600mm, 1.0$ }.
- **Frequency:** 1GHz.
- **Number of Unknowns:** 6940.
- **Number of Ports:** 4.
  
- CBFM REDUCES THE MATRIX SIZE TO 8
- CBMOM IS HIGHLY PARALLELIZED

# Parallel Conformal FDTD Solver

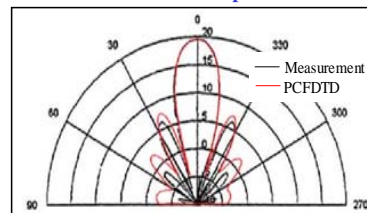


# Conformal FDTD (CFDTD) Solver

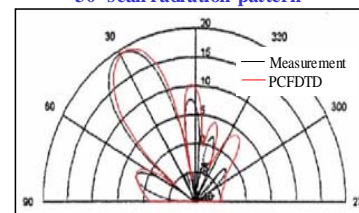
Patch antenna array: problem size =  $4\lambda \times 4\lambda \times 1\lambda$  (500x500x200), 4 processors



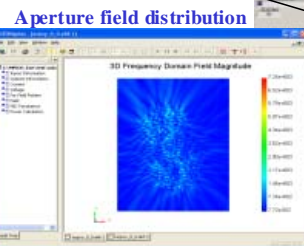
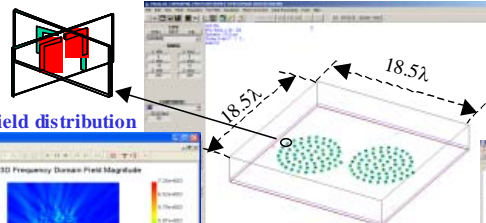
0° scan radiation pattern



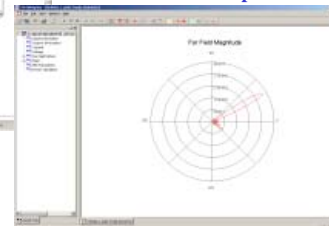
30° scan radiation pattern



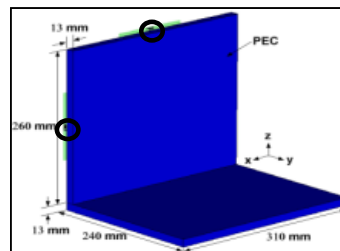
Crossed dipole array: problem size =  $18.5\lambda \times 18.5\lambda \times 3.05\lambda$  (1400x1400x254), 60 processors



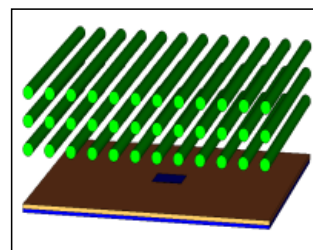
30° scan radiation pattern



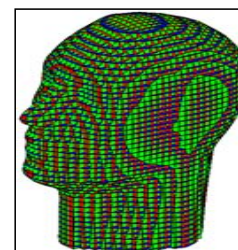
Dual-band Inverted-F WLAN Antennas on notebook computer



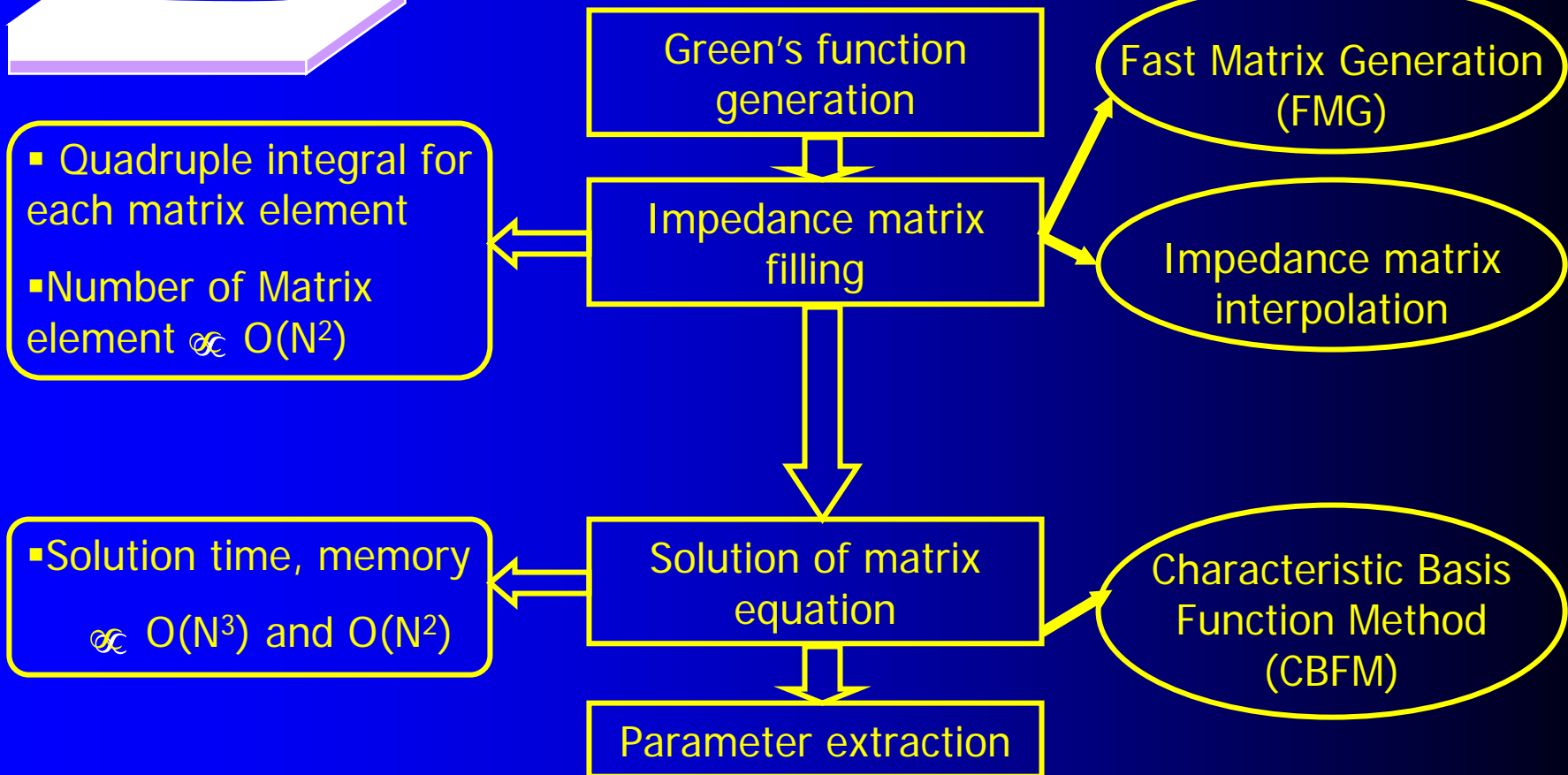
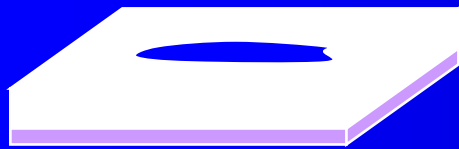
Patch antenna with EBG superstrate



Bio-electromagnetics



# Computational problems and suggestions for each step of MoM



		<b>Conventional MOM/FMM</b>	<b>CBMOM</b>
<b>1</b>	<b>Localization of the Problem</b>	<b>Yes &amp; No</b> <b>Computes + stores only the near field interactions but iterates on the entire solution vector</b>	<b>Localization is achieved by using a windowed incident field</b>
<b>2</b>	<b>Multiple Frequencies</b>	<b>Treated individually-one at a time</b>	<b>CB's can be generated up front for a few frequencies and then used over the range.</b>
<b>3</b>	<b>Other attributes</b>	<b>(a) Kernel-dependent</b> <b>(b)Can lead to Ill-conditioned Matrices specially near resonance</b>	<b>(a) Not kernel-dependent.</b> <b>(b)Handles resonant structures without difficulty because solves problems without iteration</b> <b>© Highly parallelizable</b>



# A small cluster (including 3 PCs):



**Number of processors: 3; Computer: Dell Precision Workstation 340  
Processor: Intel Pentium 4 3.06 GHz; RAM: RDRAM 2GB Dual channel  
NIC: Intel 1 GHz; Switch: Dell Power Connect 2508 1GHz  
Operating System: Redhat Linux 9.0; Fortran Compiler: Intel Fortran 7.0  
C++ compiler: GCC 3.2; MPI: MPICH 1.2.5**

# Large cluster (including 256 processors)



**Number of nodes: 128; Number of processors: 256; Computer: Dell PowerEdge 1750;**

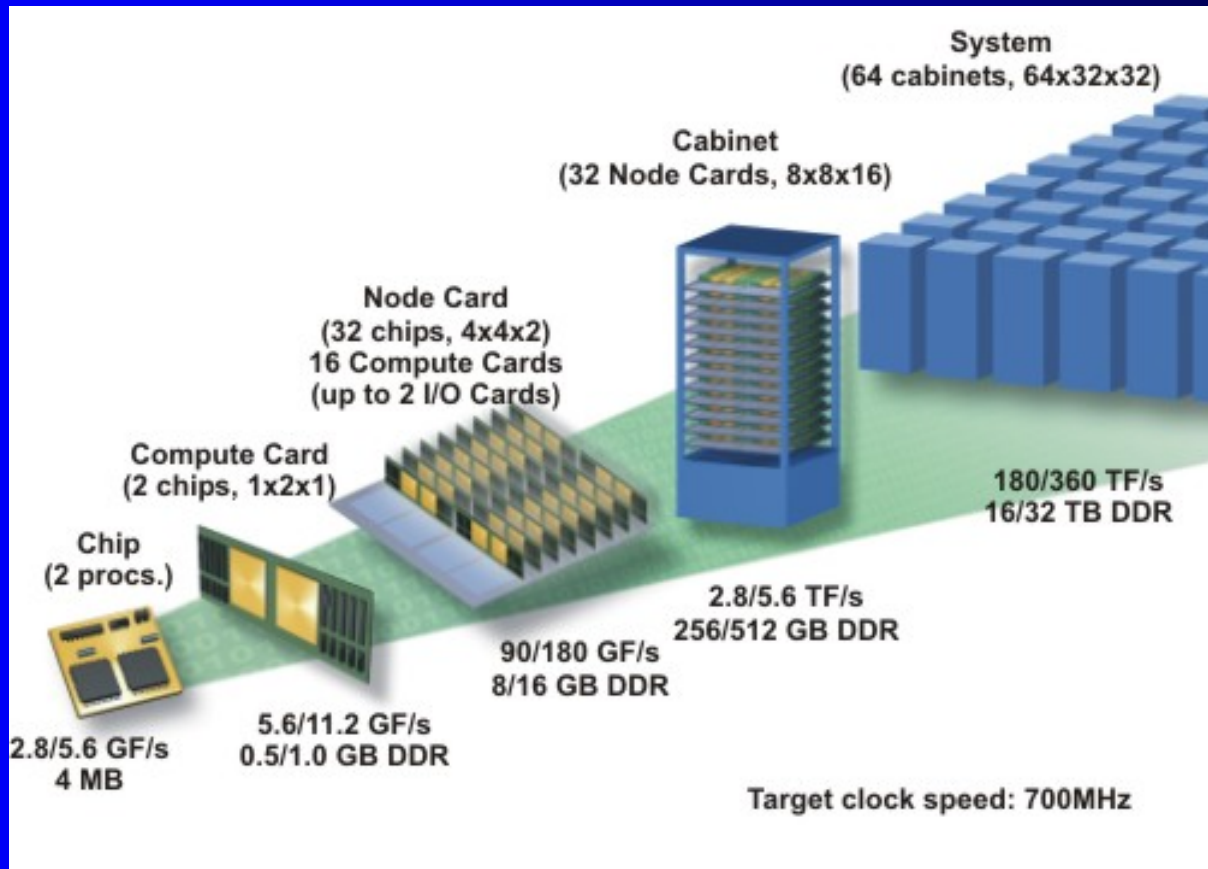
**Processor: Dual Intel Xeon 3.2 GHz, 1 MB advanced Transfer Cache**

**RAM: ECC DDR SDRAM (2x2 GB); NIC: Dual embedded Broadcom 10/100/1000 NICs**

**Switch: Myricom; Operating System: Redhat Linux AS 2.1**

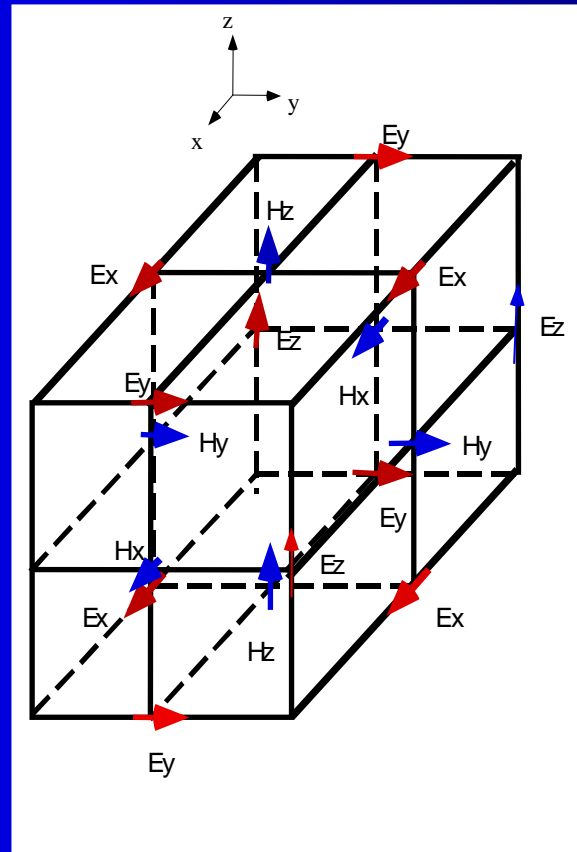
**Fortran Compiler: Intel Fortran 7.0; C++ compiler: Intel C++ ; MPI: MPIGM**

# Architecture of IBM BlueGene/L

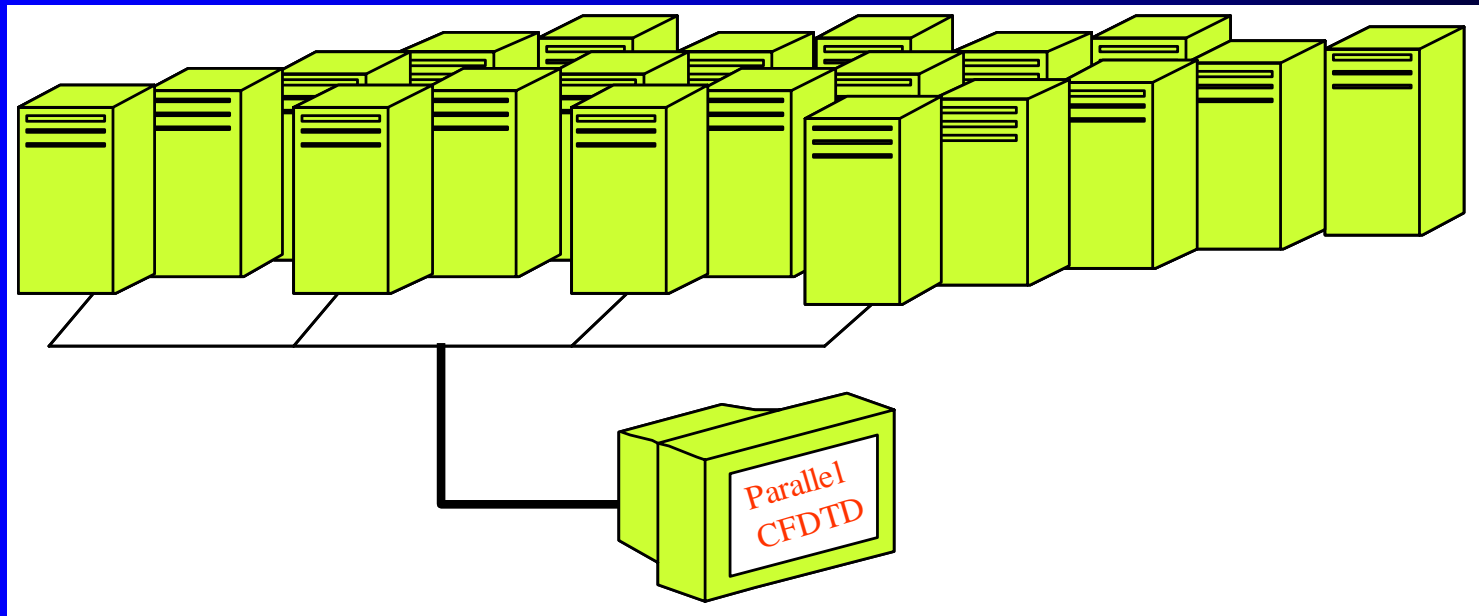


# Finite Difference Time Domain

## Yee Algorithm

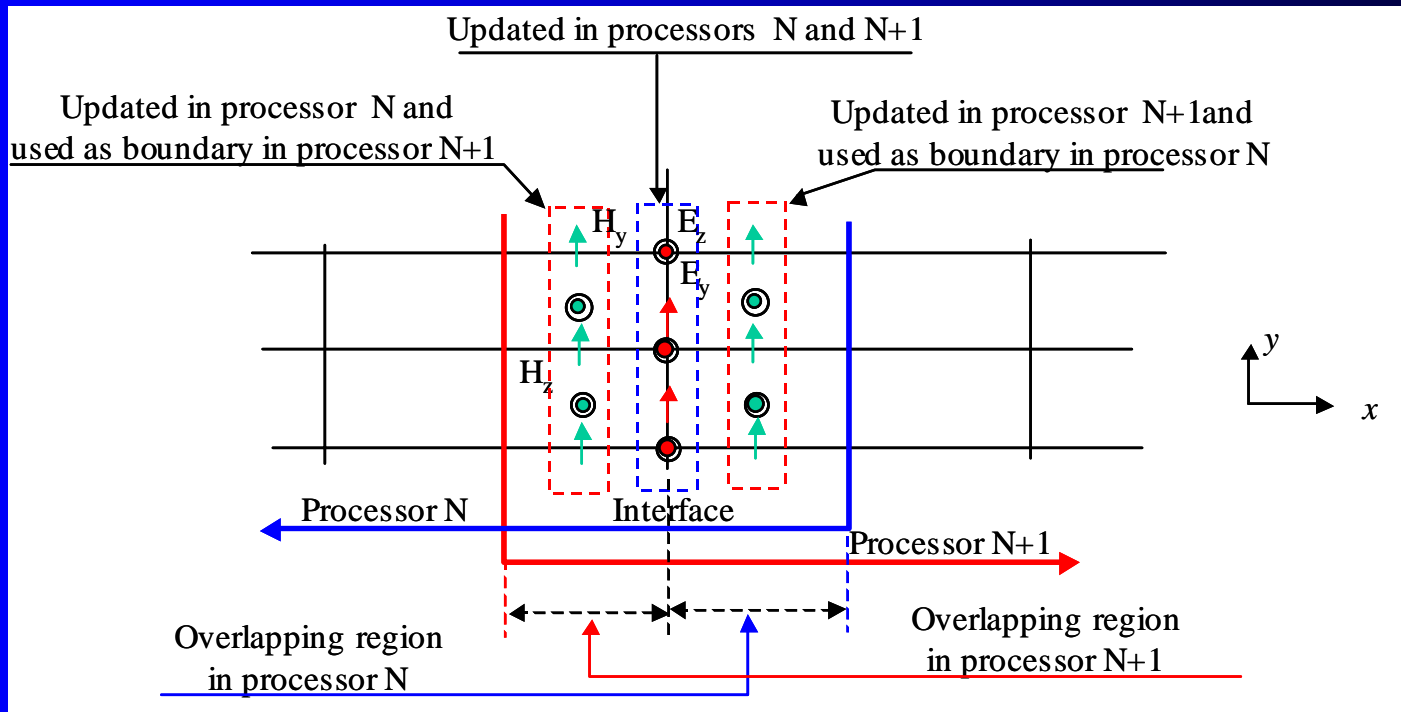


# Use Multiple Processors to simulate a Large Problem



UPWARD OF 10 BILLION UNKNOWNNS  
SOVED ROUTINELY N PSU CLUSTER  
CAN DO MUCH BIGGER PROBLEMS ON THE BLUE GENE

# Information exchange procedure



# Subdomain division

CFDTD Solver ----- Parallel processing

No. of Processes	Direction	Special Processes	Current Process	Index of Process
6	X_DIRECTION	Left Side 0 0%	45 0%	1

Processor distribution along the X direction

- + 1

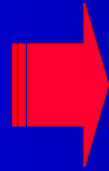
Browse Machine List Edit Machine List Help

machine.mlf

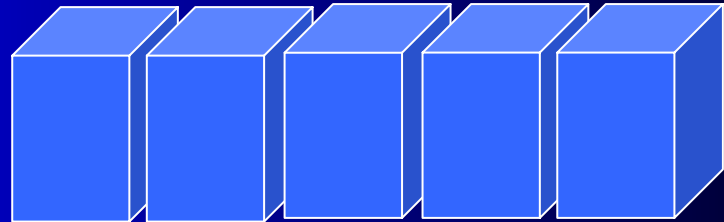
Cancel Ok

# 1-D parallel processing configuration.

Original problem



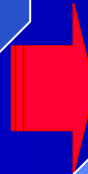
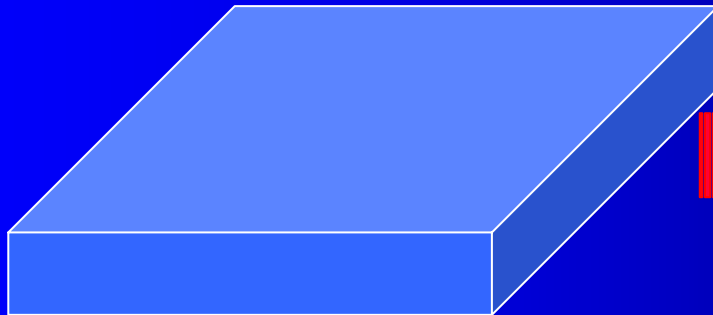
1-D domain decomposition



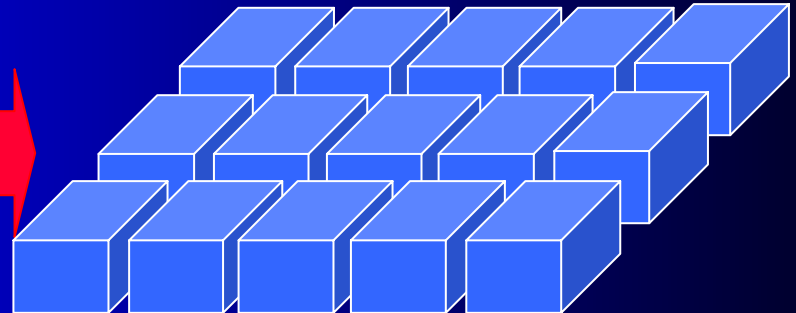


# 2-D parallel processing configuration.

Original problem

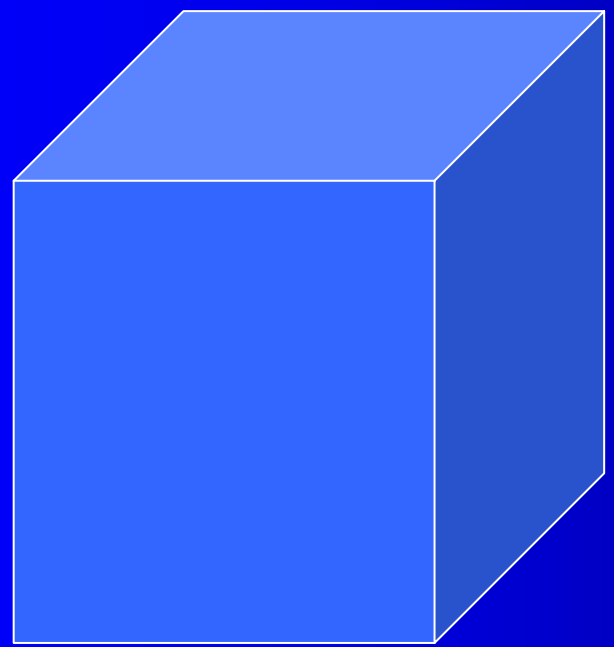


2-D domain decomposition

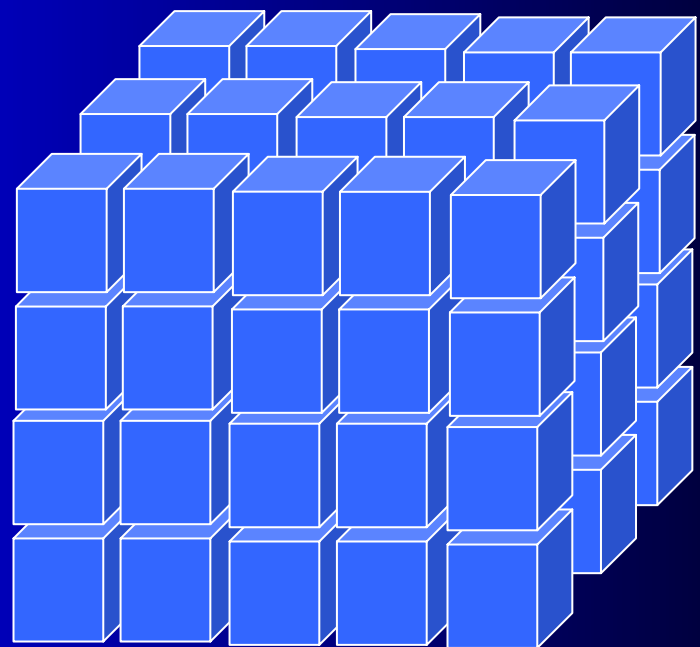


# 3-D parallel processing configuration.

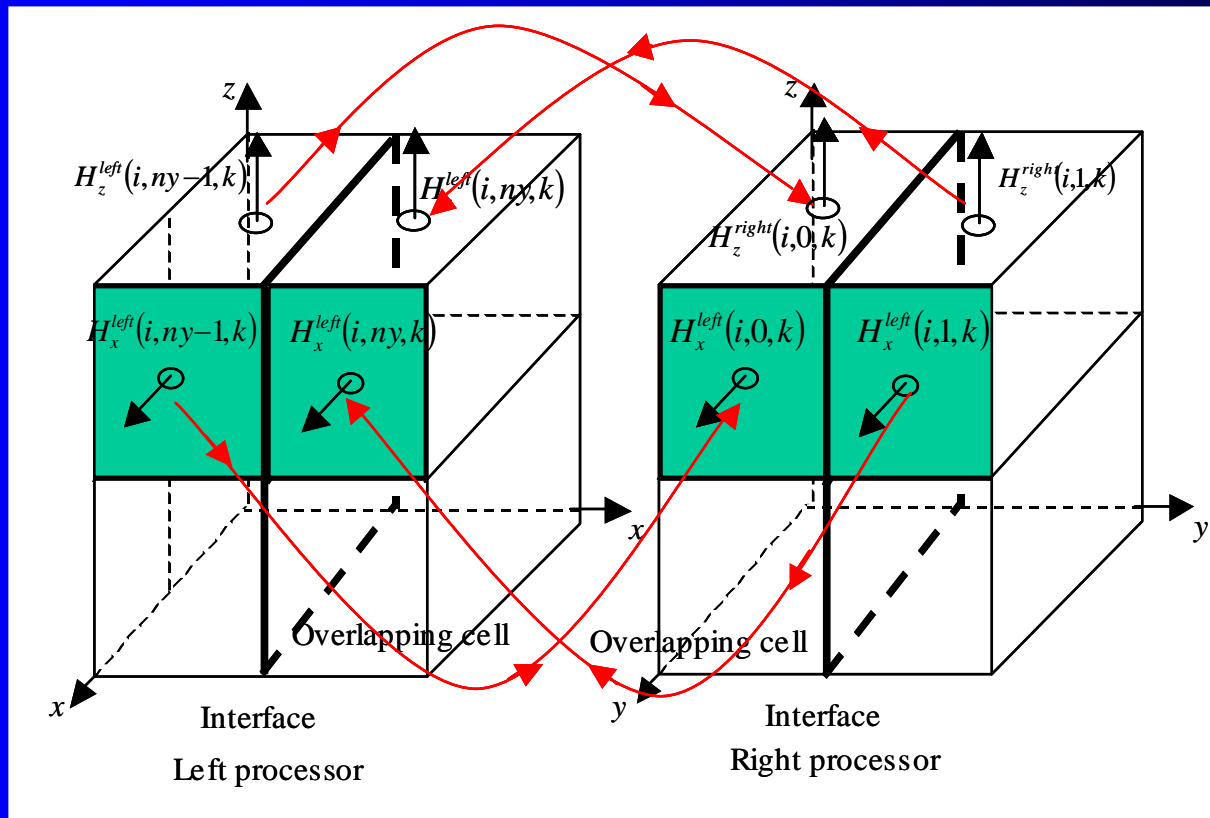
Original problem



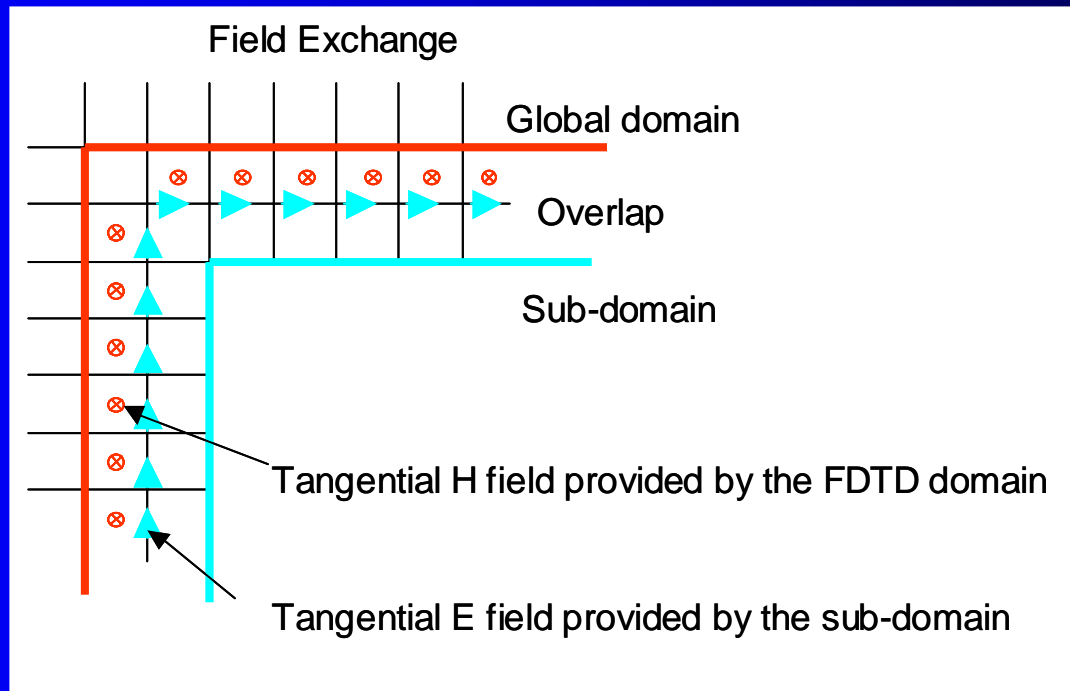
3-D domain decomposition



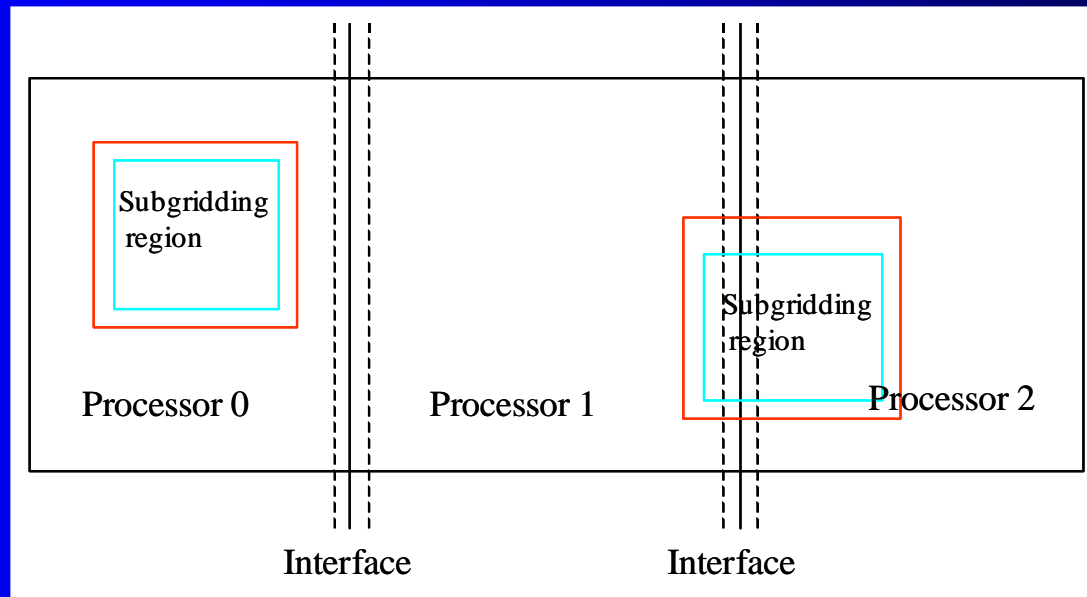
# Data exchange configuration for processes along the y-direction



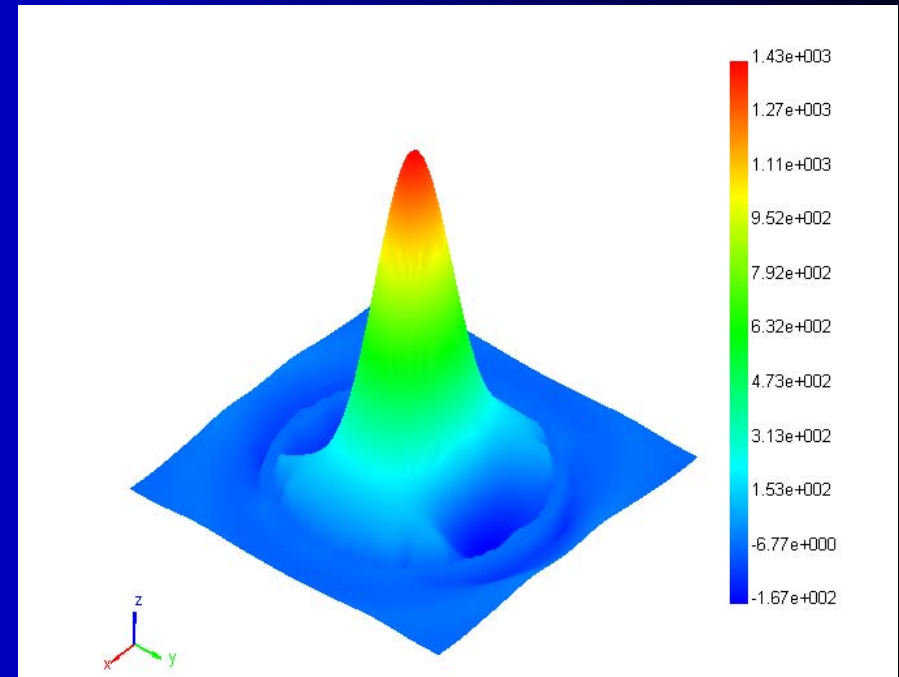
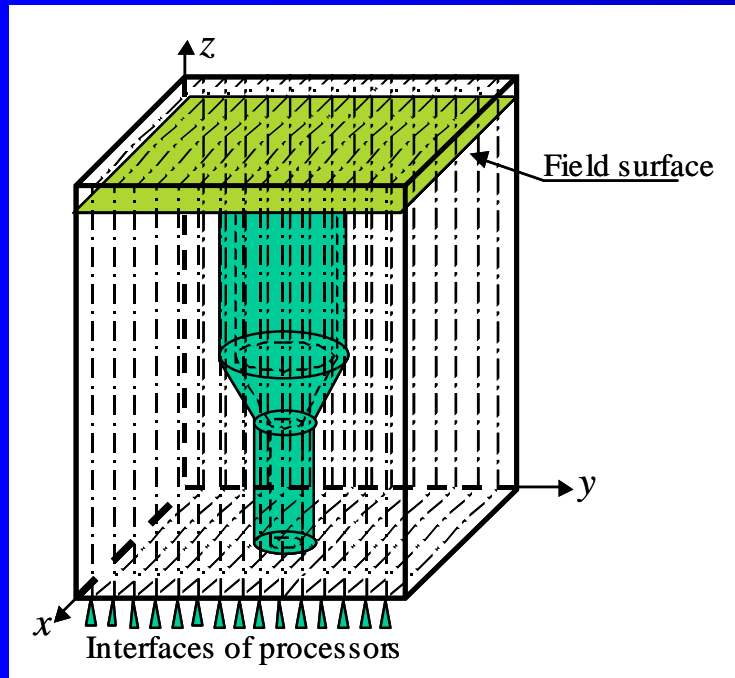
# Overlapping region between regular FDTD and subgridding regions



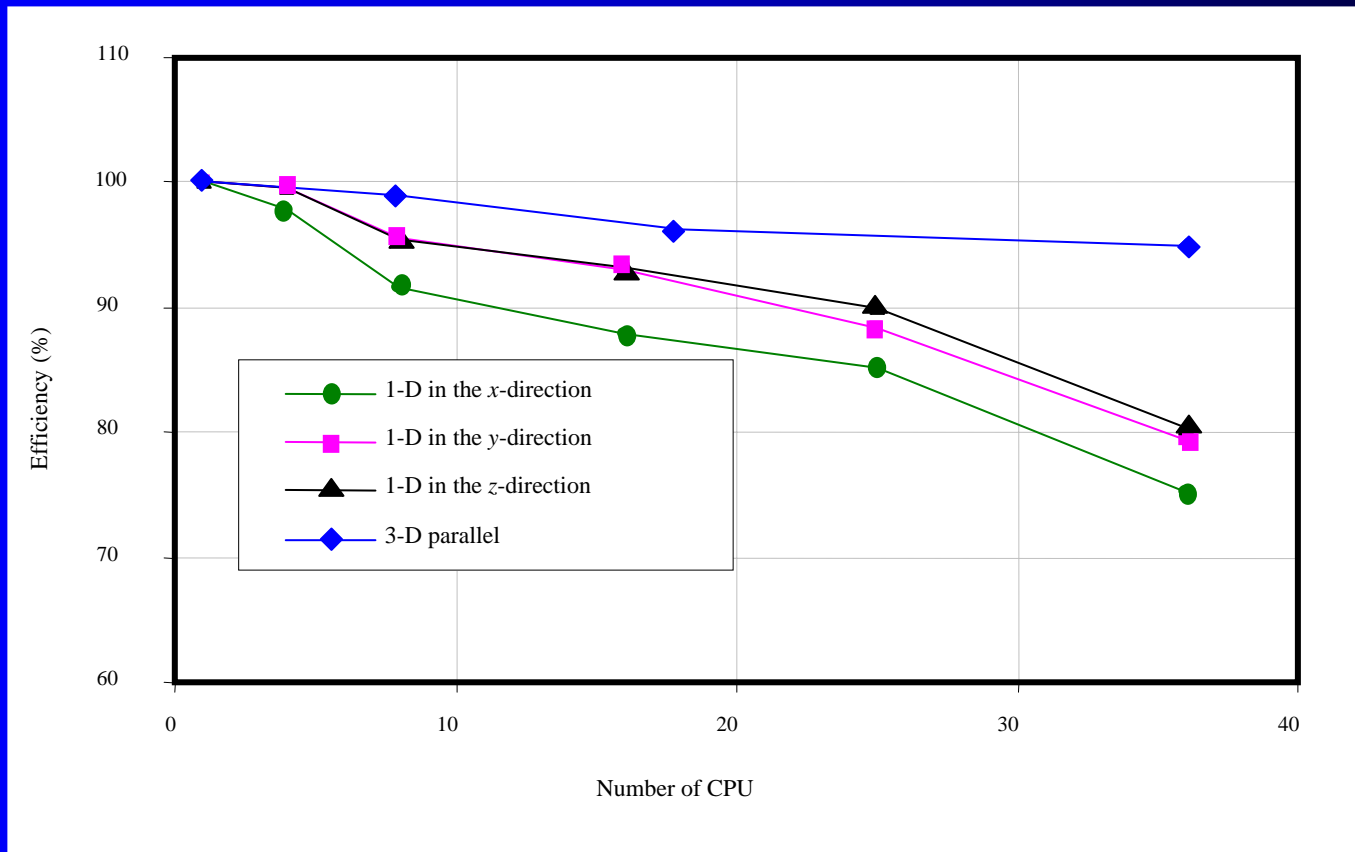
# Subgridding implementation in the parallel processing



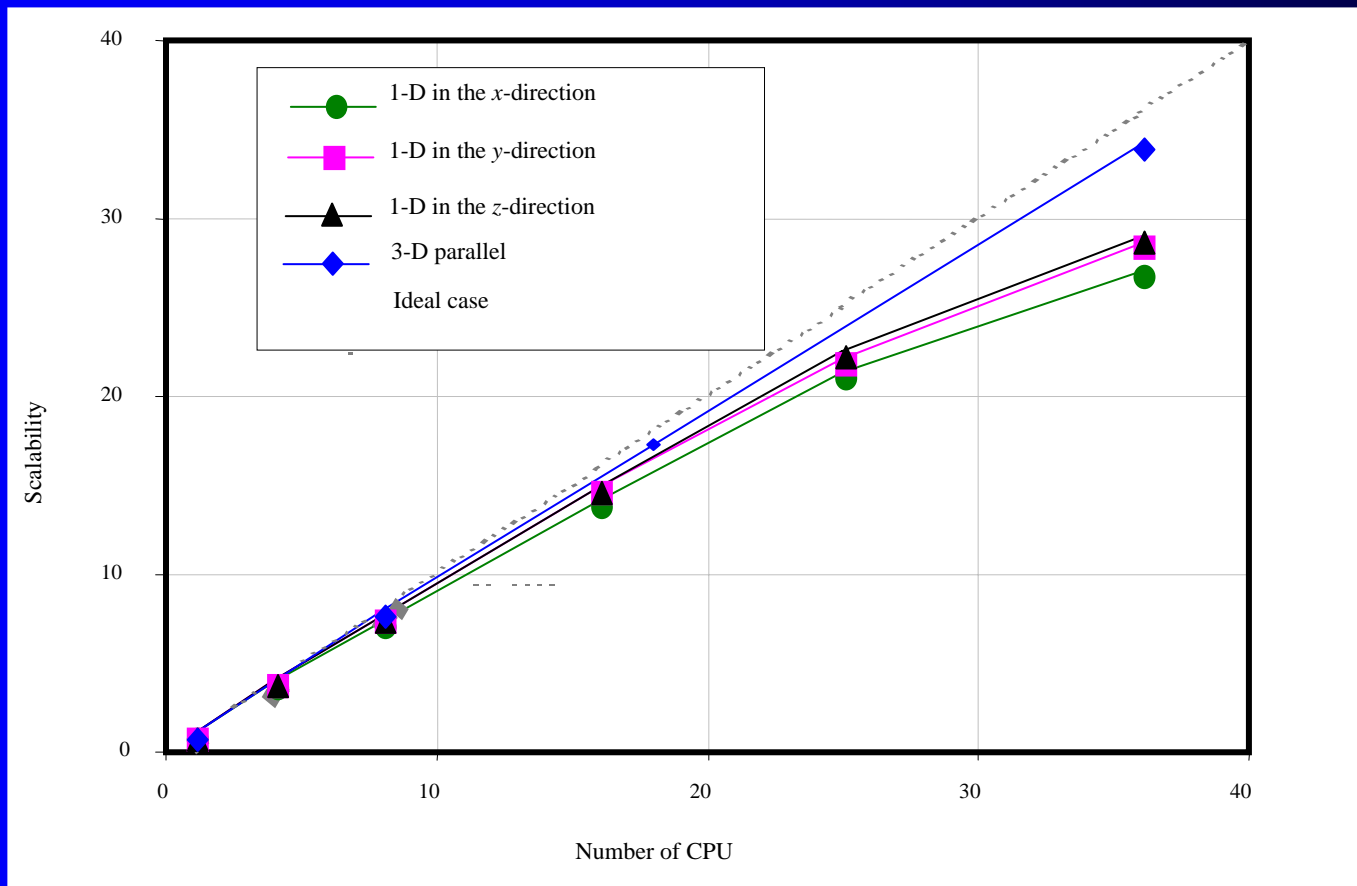
# Result collection



# Efficiency of parallel FDTD solver (Penn State Lion-xm cluster)

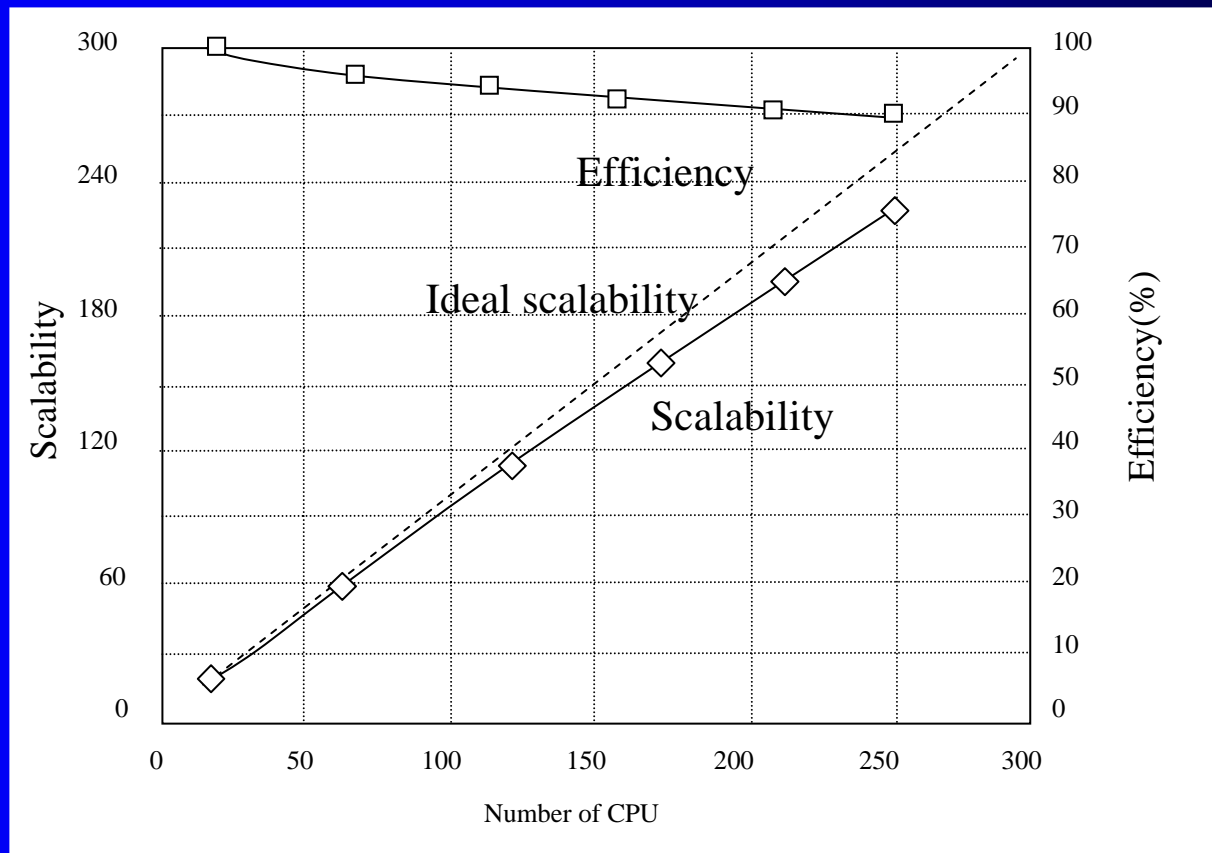


# Scalability of parallel FDTD solver (Penn State Lion-xm cluster)

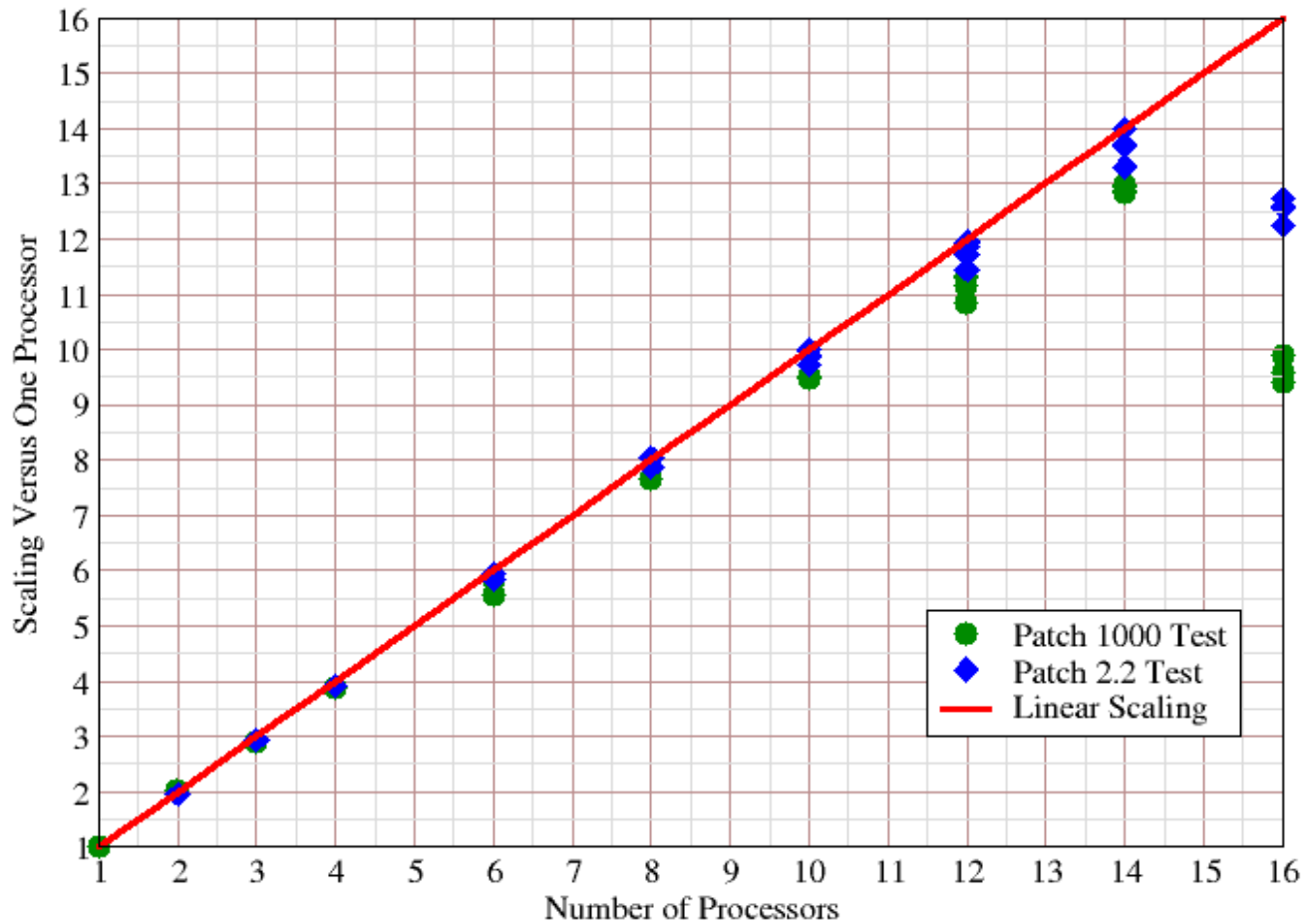




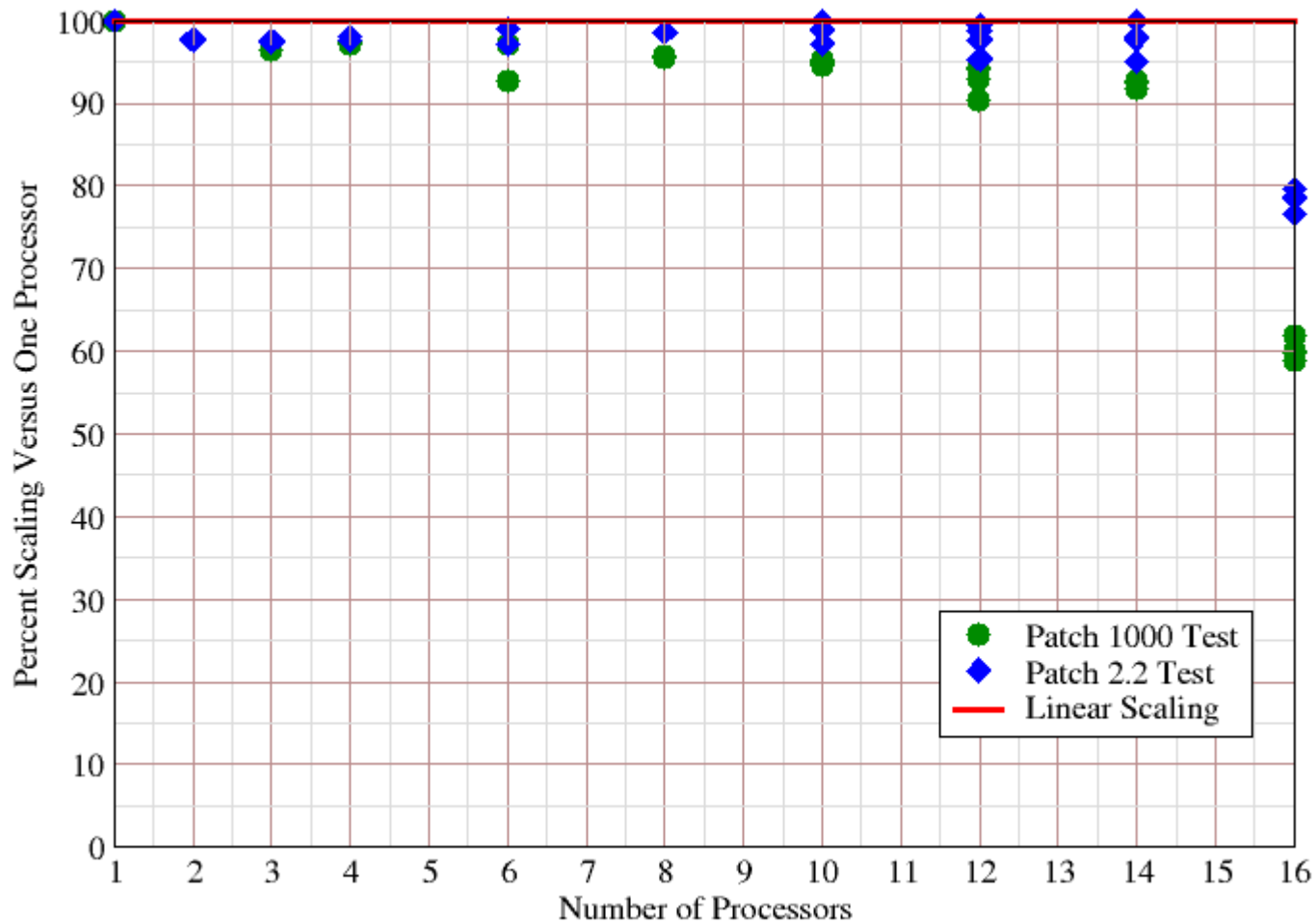
# Scalability and Efficiency of parallel FDTD solver (IBM BlueGene/L)



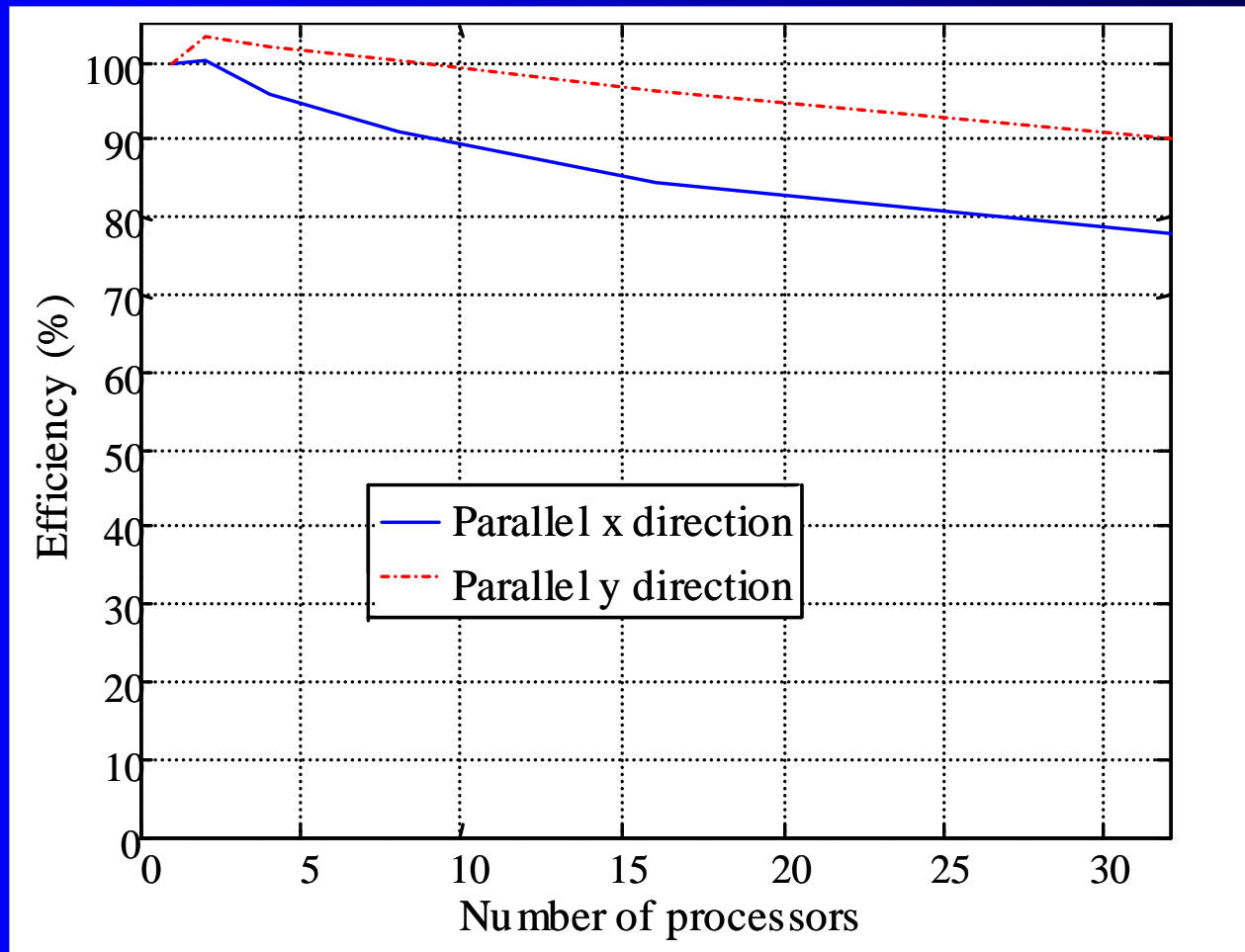
# PFDTD NP Scaling Results



# PFDTD Percent Scaling Results

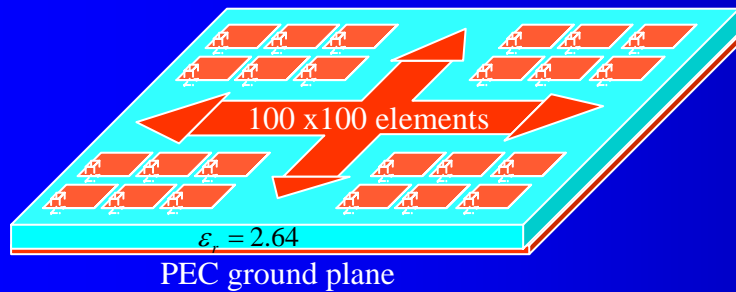


# Antenna array fed by coaxial cable (continued)

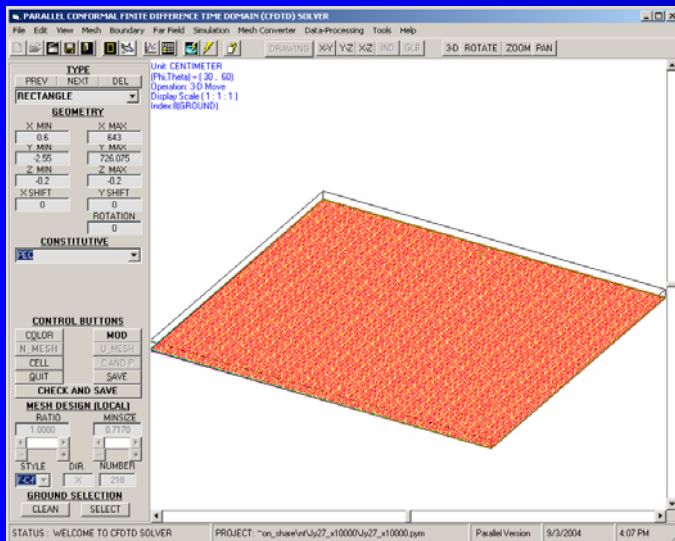


# 100x100 patch antenna array (array size: 48.25 wavelengths)

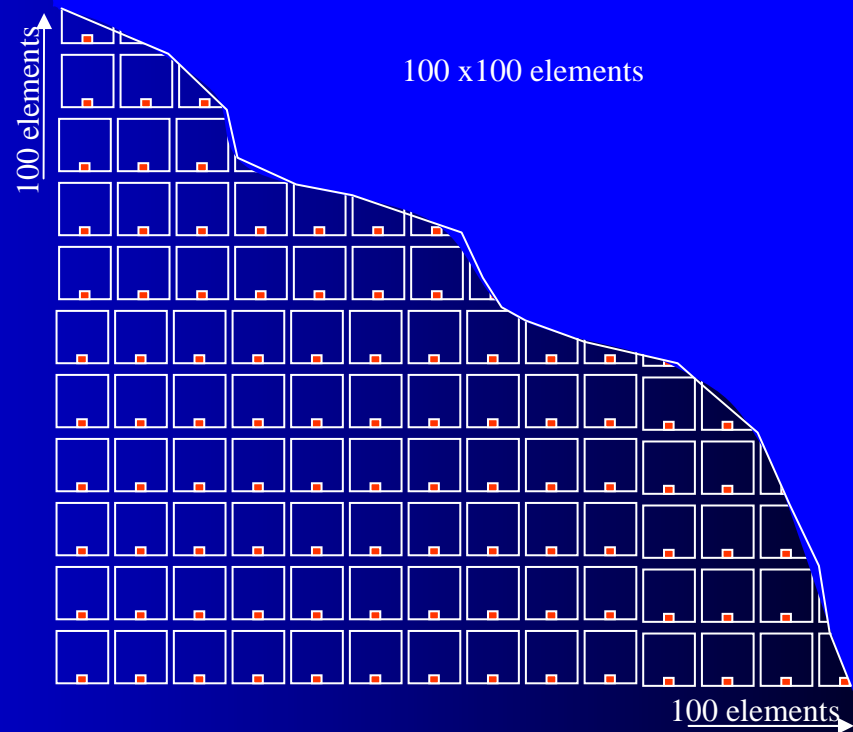
3-D view



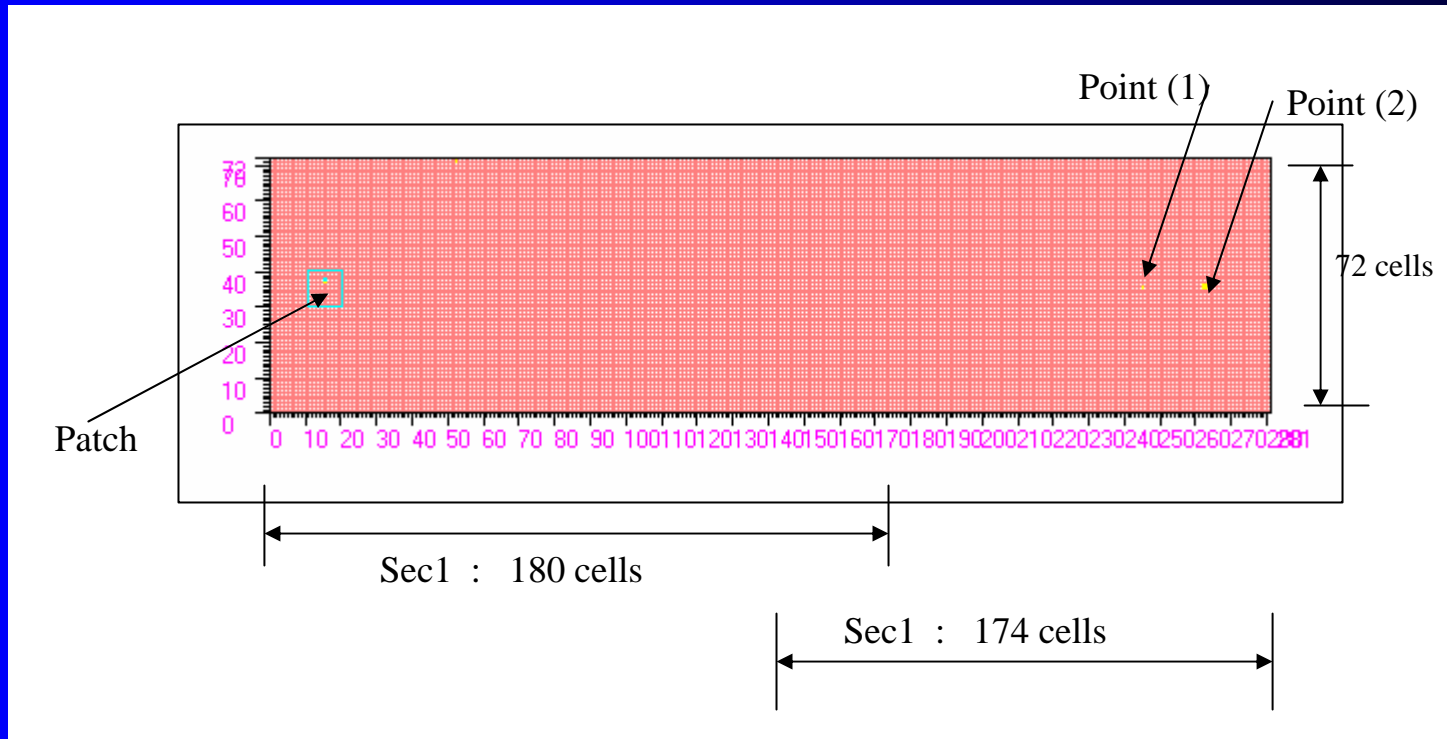
FDTD project configuration



Top view

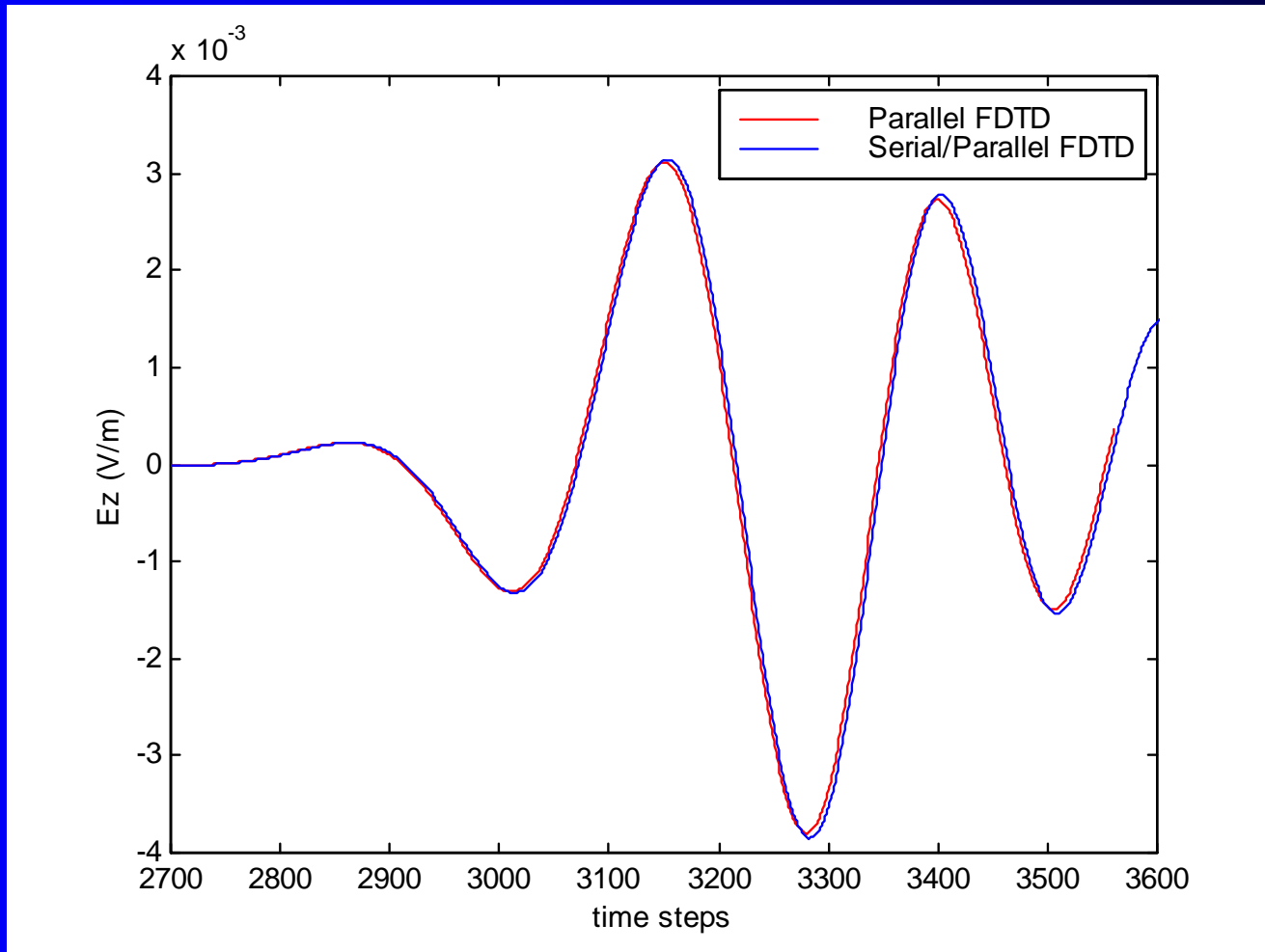


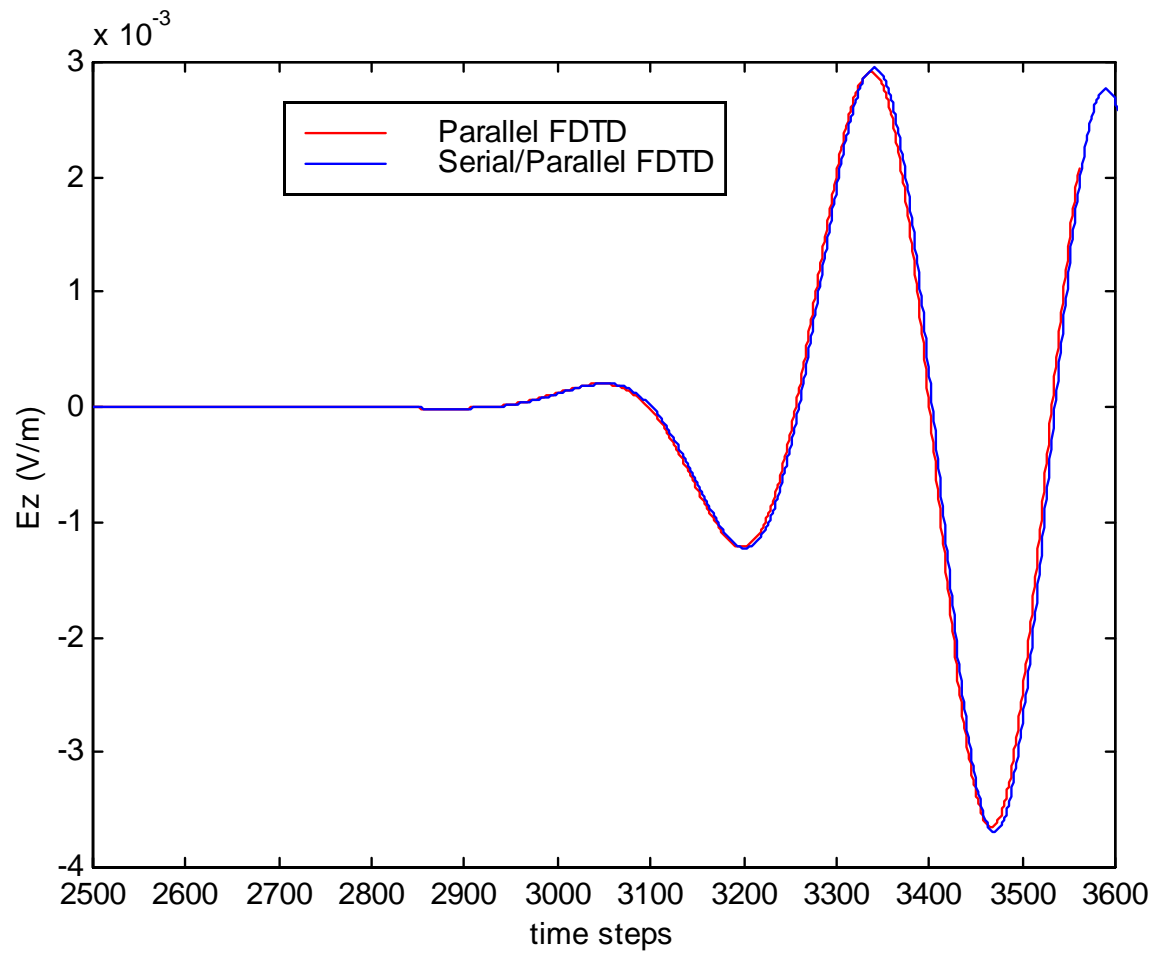
# Serial/Parallel FDTD Mimics TDR



**CAN ENHANCE PROBLEM SOLVING CAPABILITY  
BY ORDERS OF MAGNITUDE IN TERMS  
OF NUMBER OF UNKNOWNNS**

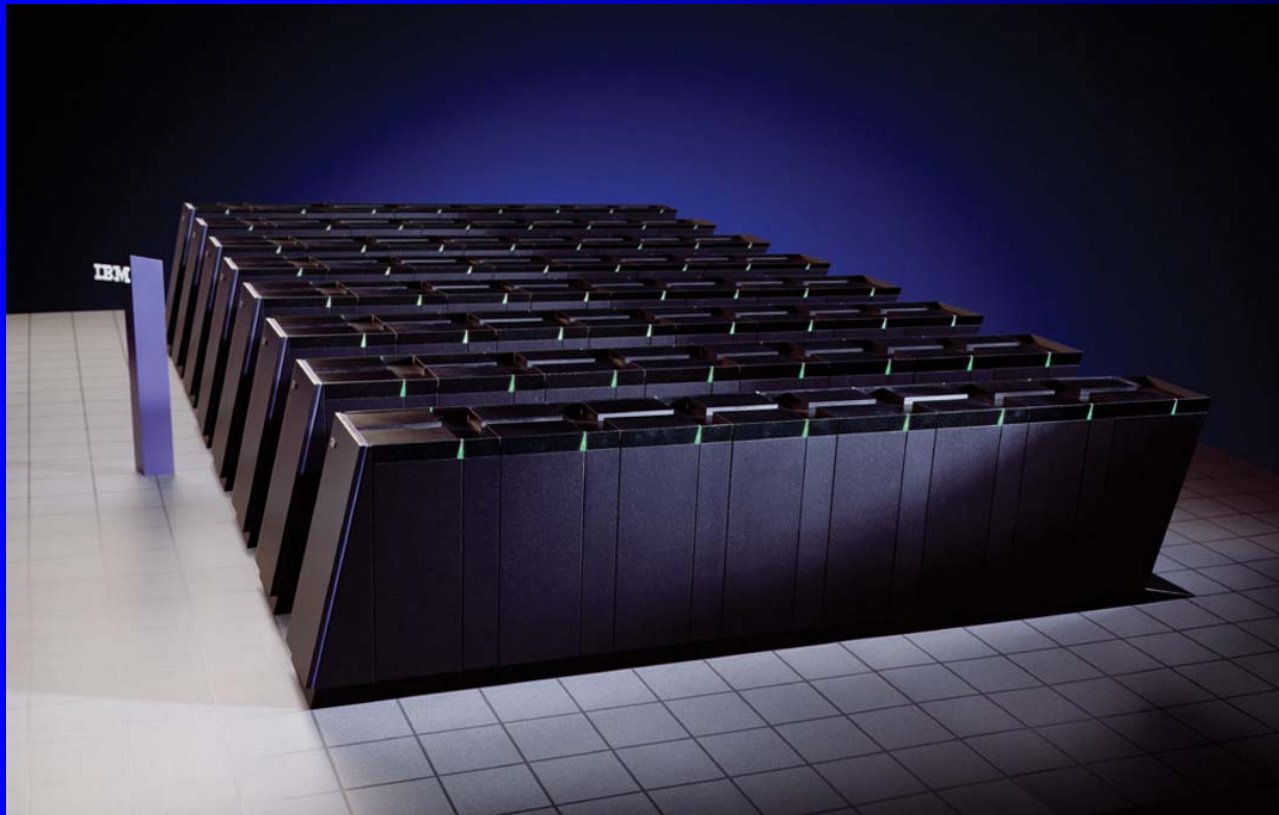
# Point (1)







# IBM BlueGene/L



# Code Enhancer—Beowulf on the Go--in a Box--Just Plug into the USB Port and Play



**POOR MAN'S BLUE GENE??**