Application of the GPU to a Two-Part Computational Electromagnetic Algorithm

NVIDIA GPU Technology Conference 2012

Eric Dunn, Electromagnetic Research Scientist
Jeremy Seuring, Electromagnetic Research Scientist

NATIONAL SECURITY • ENERGY & ENVIRONMENT • HEALTH • CYBERSECURITY

SAIC®

Application of the GPU to a Two-Part Computational Electromagnetic Algorithm

- **The Team**
  - Who we are …
  - What tools we use …
- **The Problem**
  - What is our research trying to address …
- **The Solution**
  - How have we used GPU technology to help us …

GPU = graphics processing unit

Our Company

SAIC is a FORTUNE 500® scientific, engineering, and technology applications company that uses its deep domain knowledge to solve problems of vital importance to the nation and the world in national security, energy & environment, health and cybersecurity.

Our Successes

- 40 years of continuous growth
  - Approximately $11 billion in annual revenues for fiscal year 2011
  - FORTUNE 500 company – No. 219
- Superb staff of qualified professionals
  - Approximately 41,000 personnel worldwide
  - Approximately 10,000 employees with advanced degrees
  - Over 20,000 employees with security clearances
- Leading provider of contracted R&D services

FORTUNE 500 is a registered trademark of Time, Inc. in the U.S. and/or other countries.

This slide for use until January 1, 2013.

SAIC®

Market Leadership

SAIC uses its deep domain knowledge to solve problems of vital importance to the nation and the world.

No. 6	**World's Most Admired Companies: Information Technology Services** *FORTUNE* (March 2012)
No. 6	**Top 100 Government Contractors** *Washington Technology* (June 2011)
No. 4	**Top 100 DHS Contractors** *Government Security News* (April 2011)
No. 219	**FORTUNE 500** *FORTUNE* (May 2011)

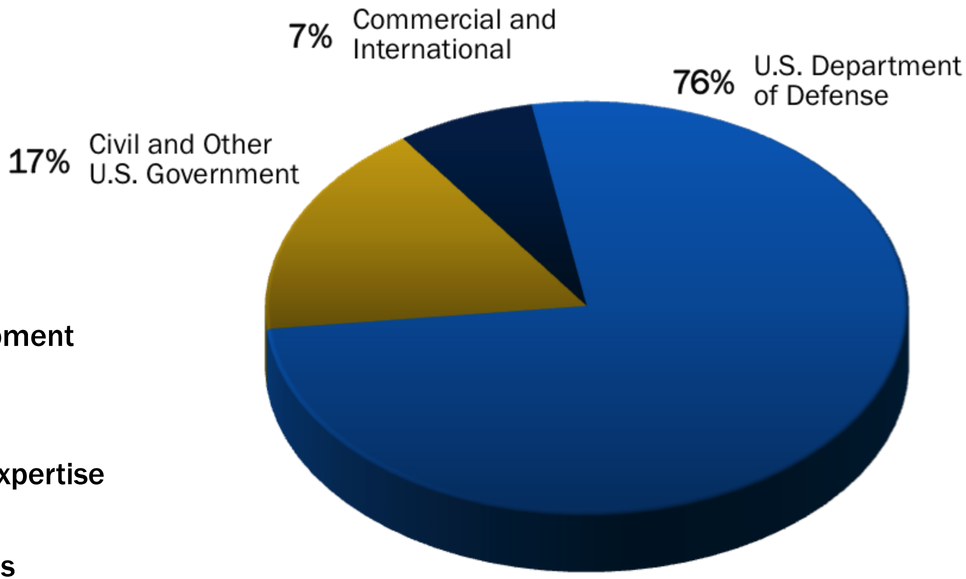
This slide for use until January 1, 2013.

SAIC®

SAIC Business Overview

- **Business Areas**
  - National Security
  - Energy & Environment
  - Health
  - Cybersecurity

- **Competitive Strengths**
  - Innovative applications of research and development
  - Platform independence
  - Reputation for succeeding on the tough jobs
  - Breadth and depth of technology and domain expertise
  - Proven management track record
  - Proven best practices, technologies and systems
  - Customer focus that leads to in-depth understanding of customer missions



Approximately $11 billion

(Fiscal Year 2011)

This slide for use until January 1, 2013.

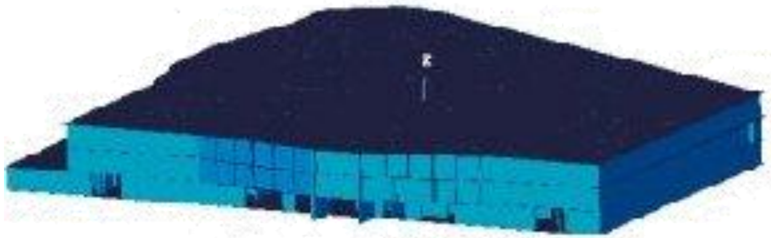
Our Organization

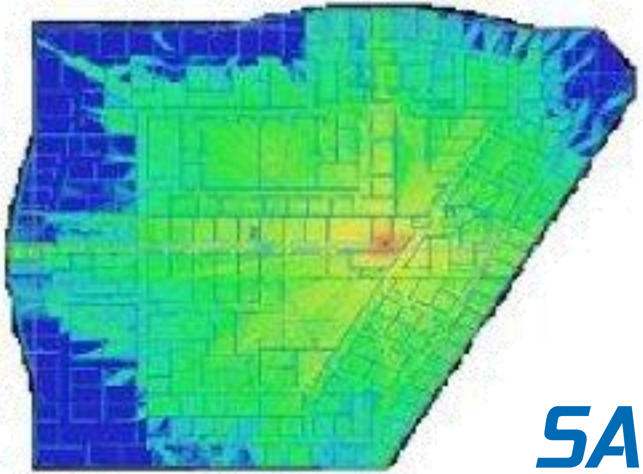
- **Our team is part of the Sensors and Phenomenology Operation at SAIC**

Our Team

We apply modeling, simulation, and analysis to solve problems related to electromagnetic radiation, scattering, and coupling.

– Radiation problems involve calculating
  - how an antenna will perform when installed on a platform
  - how the energy radiated by an antenna will propagate in a complex environment

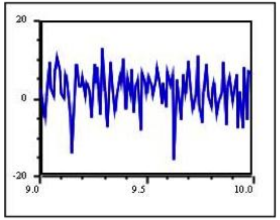




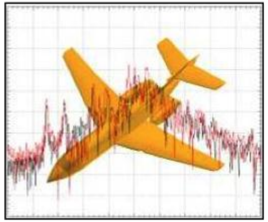
Our Team

We apply modeling, simulation, and analysis to solve problems related to electromagnetic radiation, scattering, and coupling.

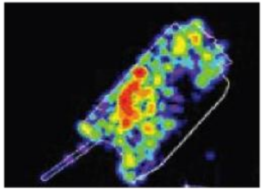
- Radiation problems involve calculating
  - how an antenna will perform when installed on a platform
  - how the energy radiated by an antenna will propagate in a complex environment
- Scattering problems involve calculating
  - what a target signature looks like (RCS, range profiles, SAR images, scattering centers)
  - where the sources of signature components come from



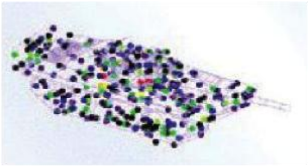
RCS



Range profile



SAR image



Scattering centers

Our Team

We apply modeling, simulation, and analysis to solve problems related to electromagnetic radiation, scattering, and coupling.

- Radiation problems involve calculating
  - how an antenna will perform when installed on a platform
  - how the energy radiated by an antenna will propagate in a complex environment
- Scattering problems involve calculating
  - what a target signature looks like (RCS, range profiles, SAR images, scattering centers)
  - where the sources of signature components come from
- Coupling problems involve calculating
  - how much energy is received and coupled into a system

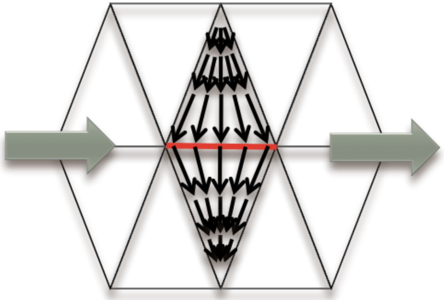
SAIC

Our Tools

We develop, distribute, and employ many tools to solve these problems. The tools fit into three categories of methods called full-wave, asymptotic, and hybrid.

– Full-wave methods are
- the most accurate solution available,
- expensive in terms of memory-usage and run-time,
- executable only on geometric models that have satisfied strict requirements,
- typically used for electrically small problems

$$\nabla \times \vec{E} = -j\omega\mu\vec{H}$$
$$\nabla \times \vec{H} = \vec{J} + j\omega\varepsilon\vec{E}$$

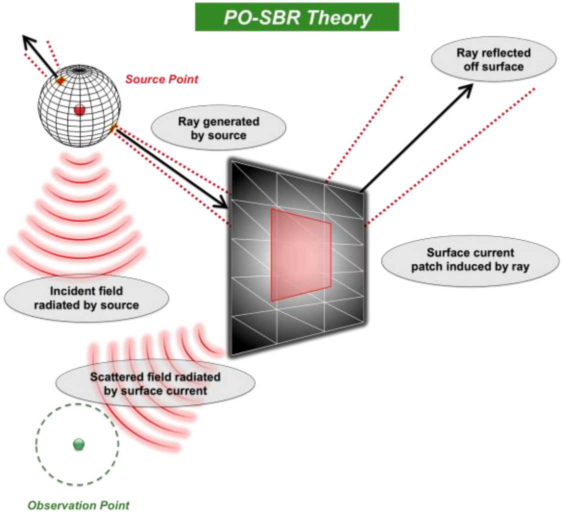


$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \\ a_{N1} & a_{N2} & & a_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

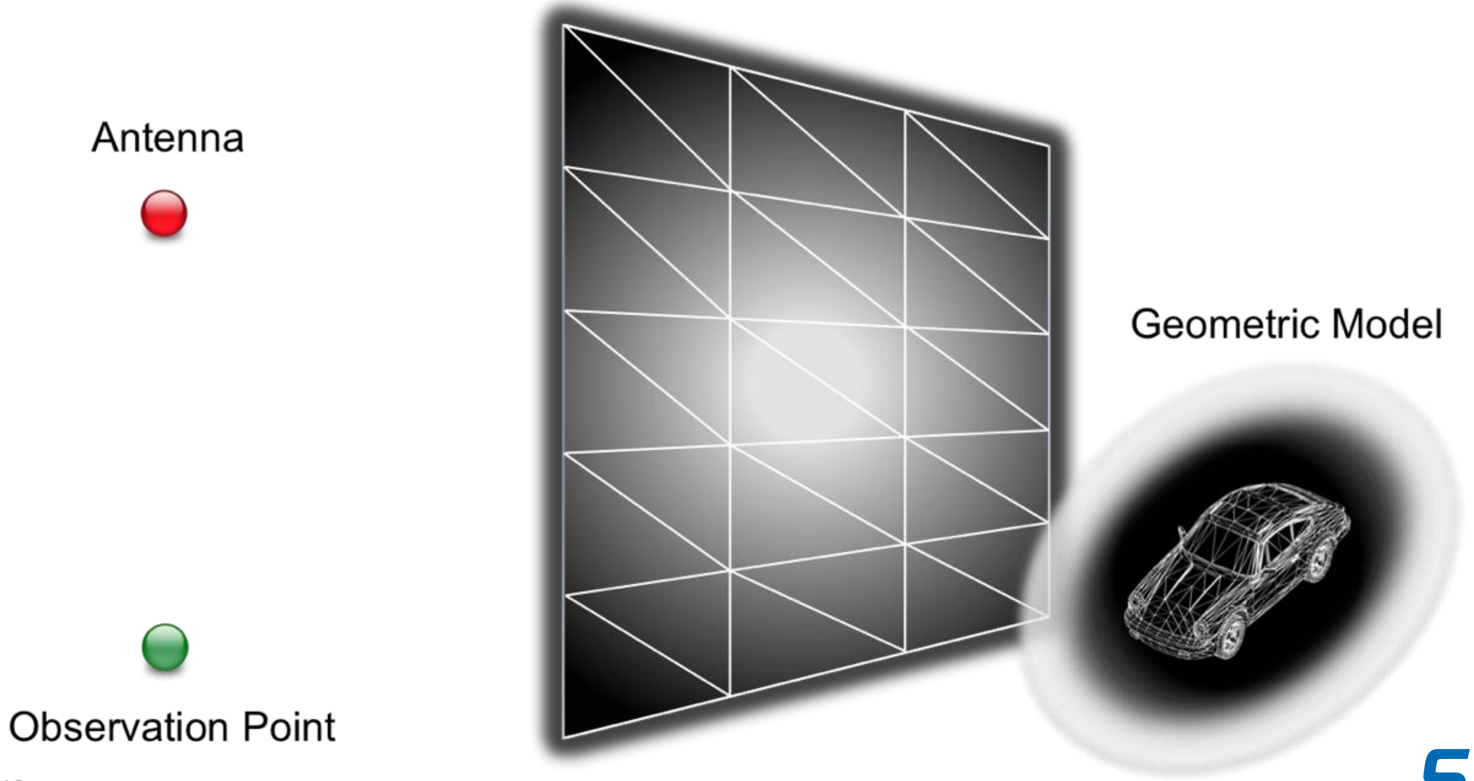
Our Tools

We develop, distribute, and employ many tools to solve these problems. The tools fit into three categories of methods called full-wave, asymptotic, and hybrid.

– Asymptotic methods are
  - much faster than full-wave solutions,
  - less expensive in terms of memory-usage and run-time,
  - ready to execute with less strict requirements for the geometric model,
  - not all inclusive and do not model all physical phenomenon,
  - typically used for electrically large problems

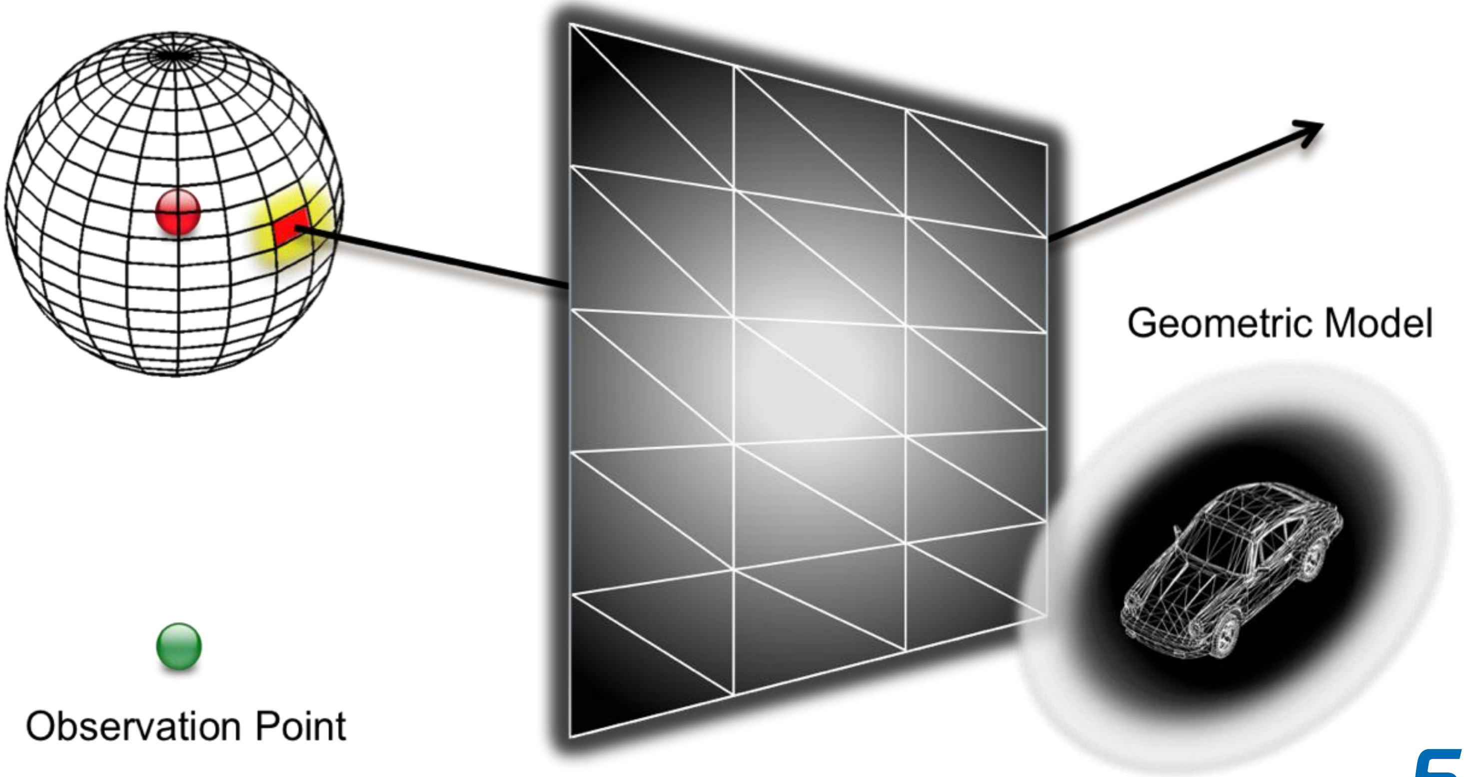


Simulation Setup

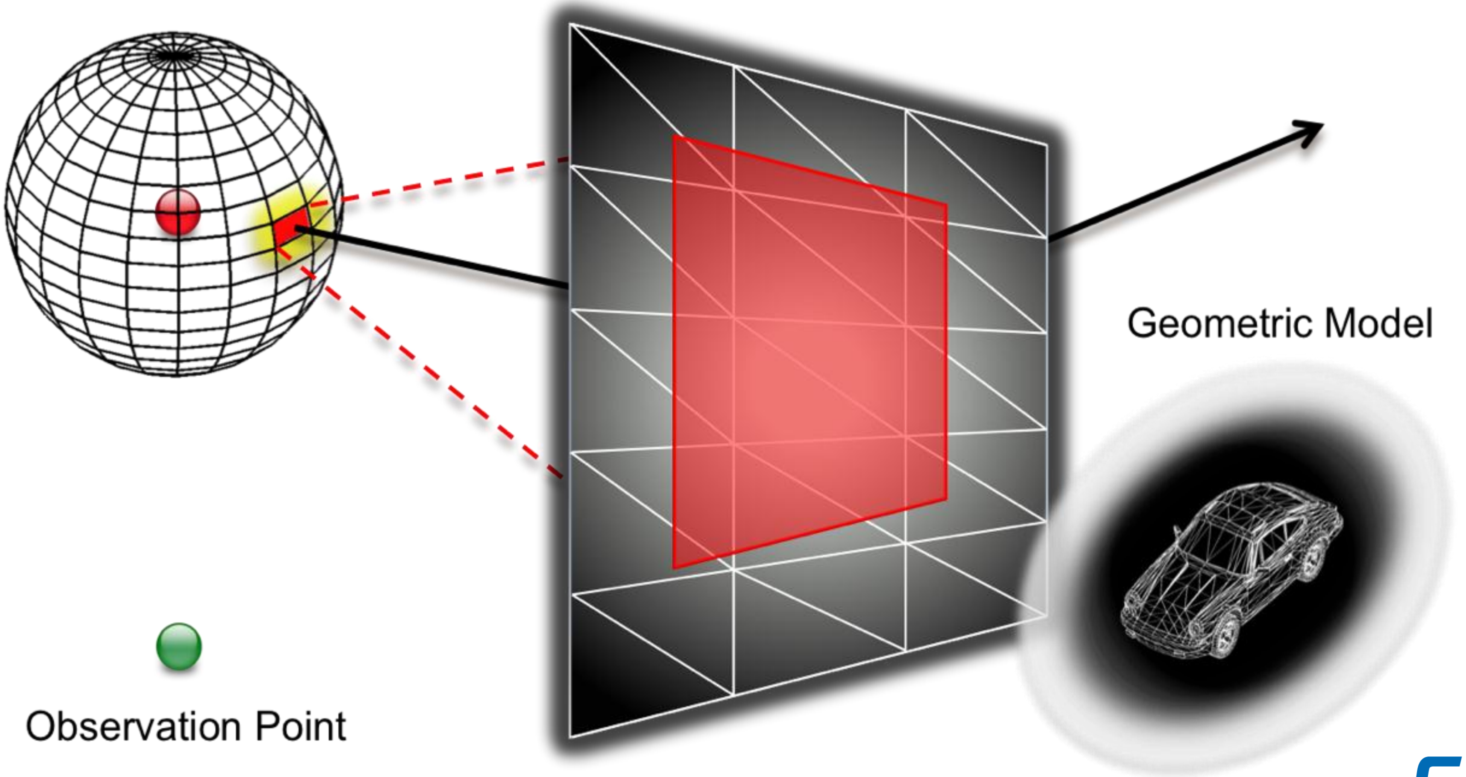


SAIC

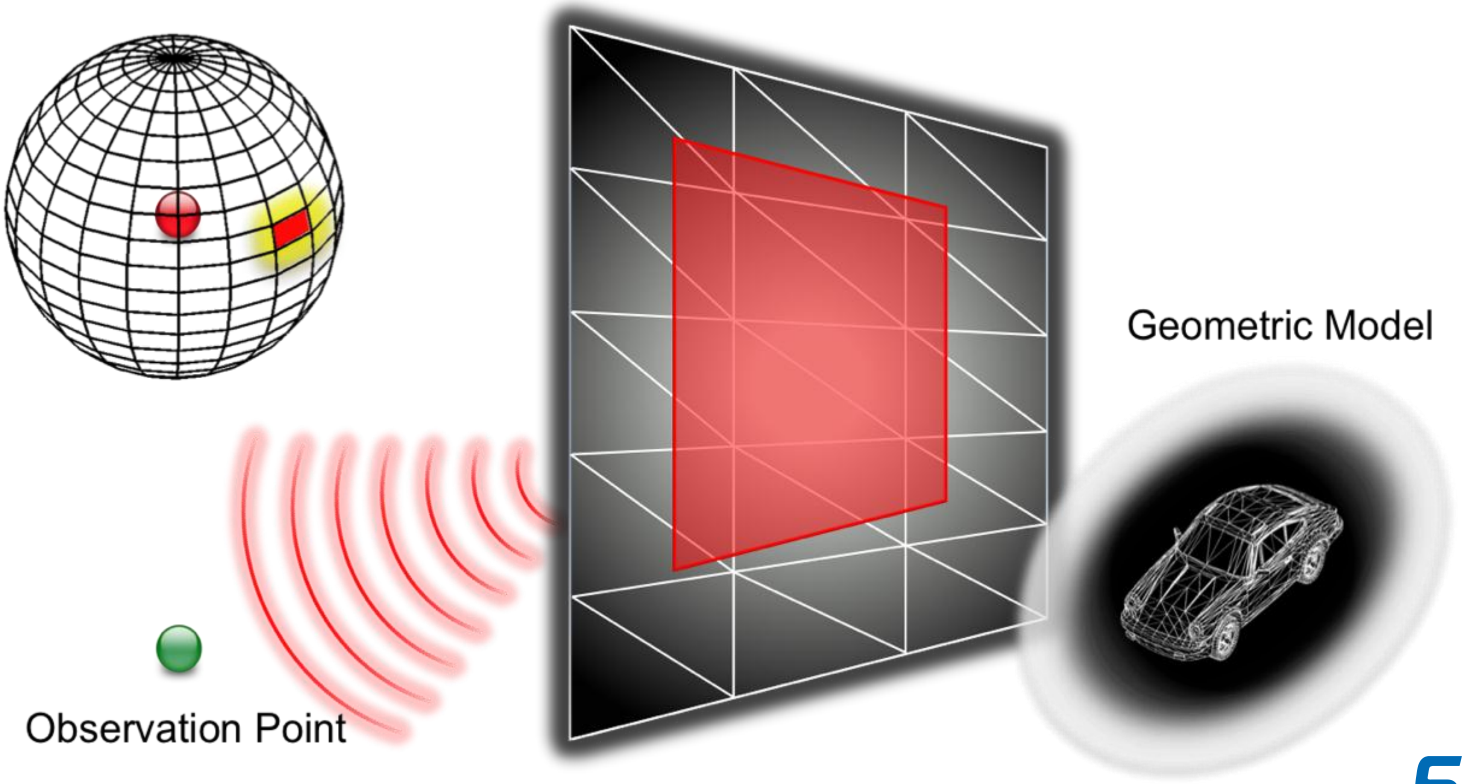
Shooting and Bouncing Ray (SBR)



Shooting and Bouncing Ray (SBR)



Shooting and Bouncing Ray (SBR)

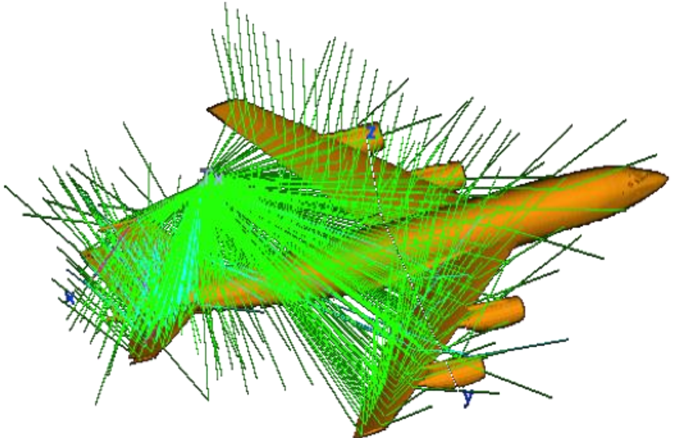


SAIC®

Our Tools

We develop, distribute, and employ many tools to solve these problems. The tools fit into three categories of methods called full-wave, asymptotic, and hybrid.

- Hybrid methods are
  - a combination of a full-wave solution and an asymptotic solution,
  - ways to capture the best of both approaches,
  - typically used for electrically large problems that contain important smaller features



Application of the GPU to a Two-Part Computational Electromagnetic Algorithm

- **The Team**
  - **Who we are ...**
    We solve electromagnetic radiation, scattering, and coupling problems.
  - **What tools we use ...**
    We develop, distribute, and employ tools that use full-wave, asymptotic, and hybrid methods.
- **The Problem**
  - **What is our research trying to address ...**
- **The Solution**
  - **How have we used GPU technology to help us ...**

GPU = graphics processing unit

SAIC®

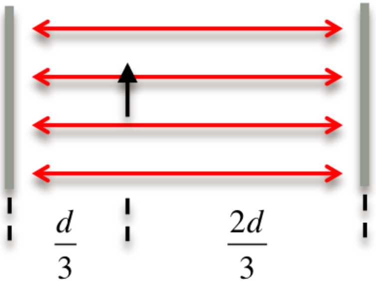
Resonance

- Energy in the system
  - Leaves through friction in the bearings
  - Enters through kicking your legs

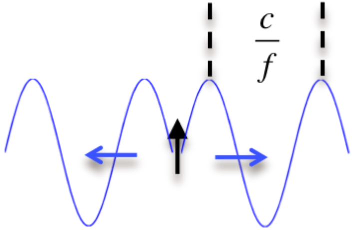
- Resonance amplifies the energy you put into the system
  - Achieved by kicking your legs at a rate equal to the natural frequency of oscillation

Resonance

- Analogous to an electromagnetic wave trapped between two metal plates



- Energy in the system
  - Leaves though escaped electromagnetic waves since the plates only have a finite size
  - Enters through radiation from a Hertzian dipole antenna



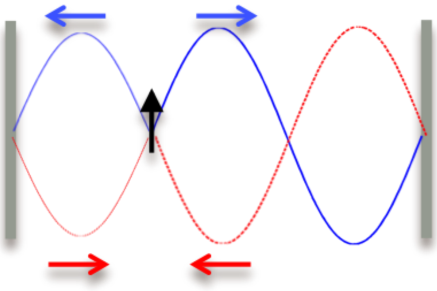
f = frequency c = speed of light d = plate separation

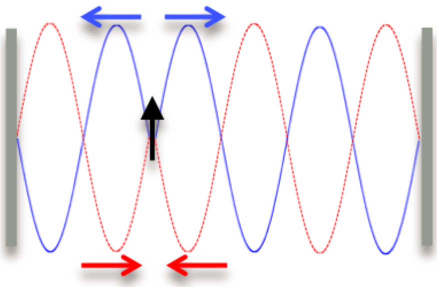
Resonance

- **Waves resonating between the plates constructively and destructively combine**
  - **A metal plate has a negative reflection coefficient**

frequency f_1

$$\frac{c}{f_1} = 2\left(\frac{d}{3}\right)$$





frequency $f_2 = 2f_1$

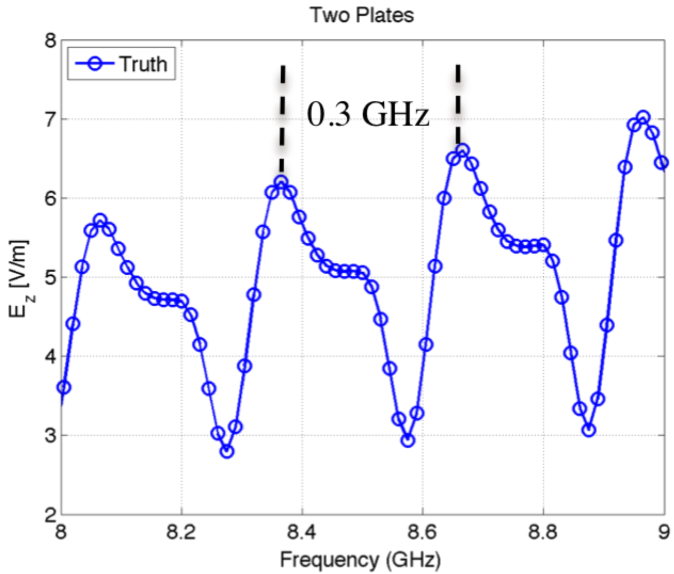
$$\Delta f = f_2 - f_1 = \frac{3c}{2d}$$

- **With plate separation of 1.5 meters, this causes some periodicity as a function of frequency**
  - **Scattered electric field intensity at the antenna will repeat every 0.3 GHz**

SAIC®

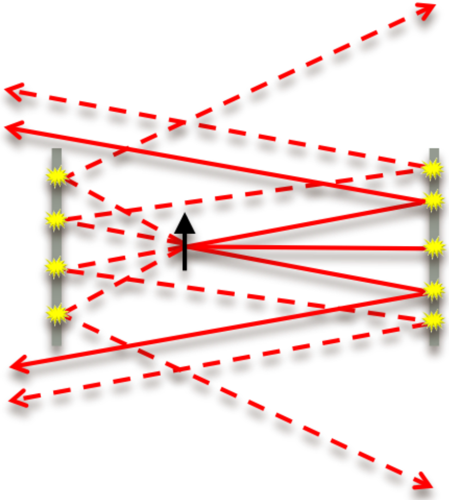
Full-wave Simulation

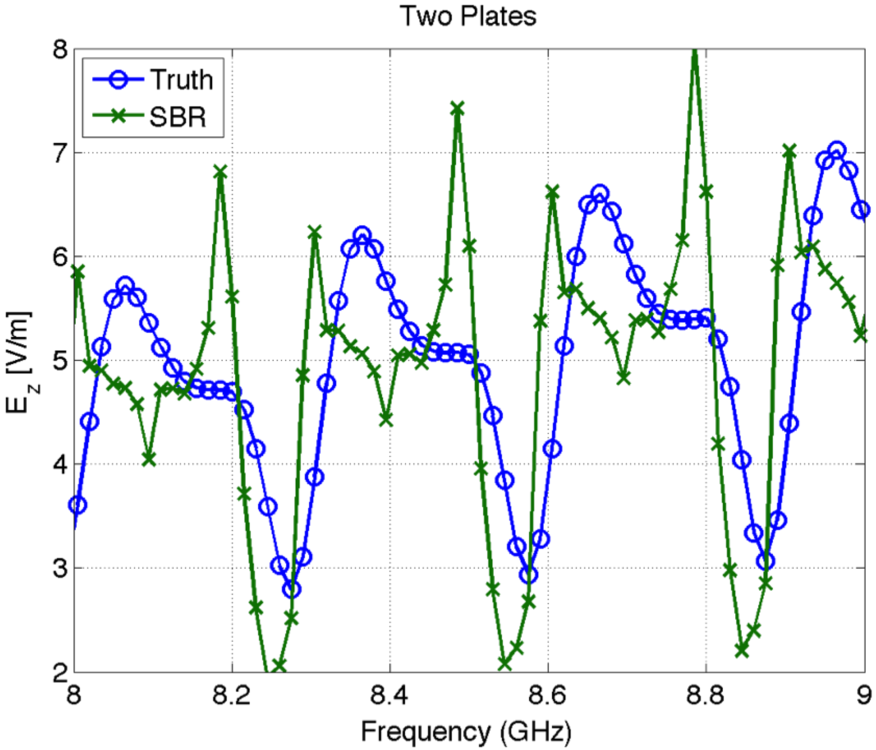
- **Using a full-wave solver we can calculate exact scattered electric field at the antenna**



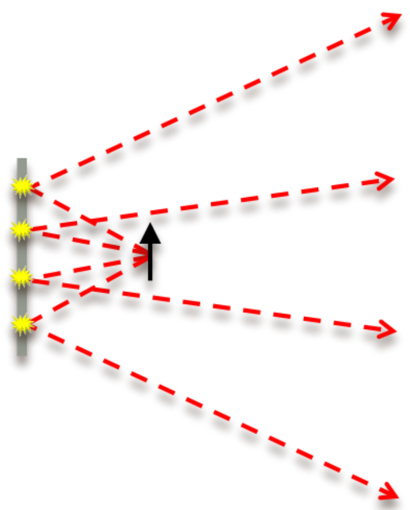
- **Increase in magnitude is because field radiated by the antenna is proportional to frequency**

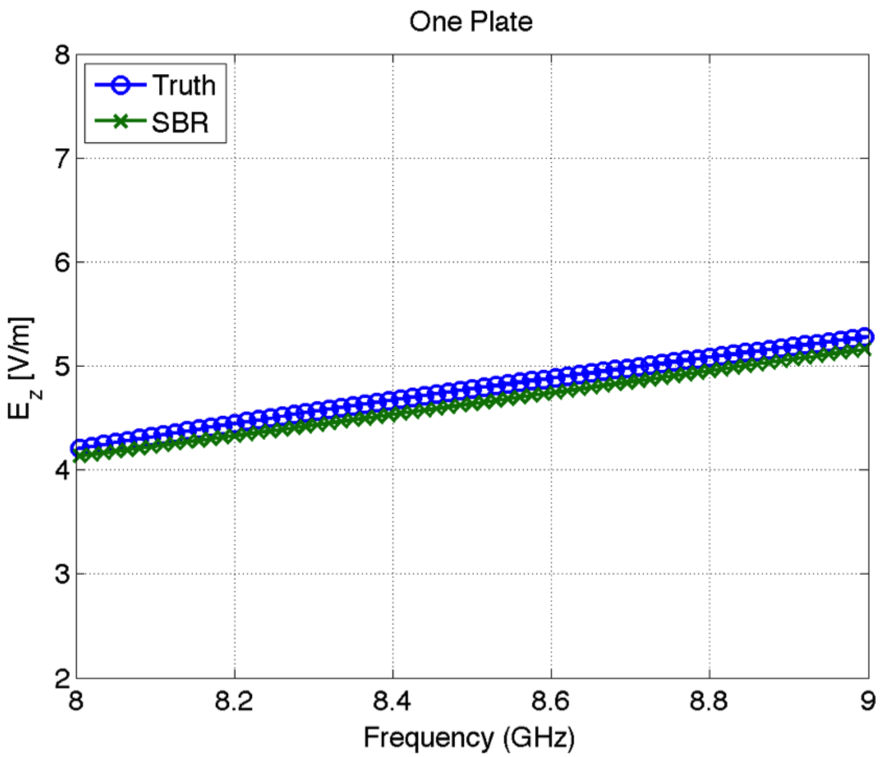
Asymptotic Simulation





Asymptotic Solution





Application of the GPU to a Two-Part Computational Electromagnetic Algorithm

- **The Team**
  - **Who we are ...**

    We solve electromagnetic radiation, scattering, and coupling problems.
  - **What tools we use ...**

    We develop, distribute, and employ tools that use full-wave, asymptotic, and hybrid methods.
- **The Problem**
  - **What is our research trying to address ...**

    The basic SBR asymptotic method does not accurately solve problems involving resonance.
- **The Solution**
  - **How have we used GPU technology to help us ...**

GPU = graphics processing unit   SBR = shooting and bouncing ray

GPU Technology

Hardware for executing many simultaneous threads

- **GPU hardware parallelism**
  - Organizes many **threads** into **blocks** and **grids**
    - Each thread executes the same device code (**kernels**)
    - Each block contains a group of threads (maximum of 512)
    - Each grid contains a group of blocks (maximum of 65,535 x 65,535)
  - Executes **thread processors** with a collection of **multi-processors**
    - Each multiprocessor has eight thread processors
    - Each multiprocessor handles a single block of threads and executes 32 instructions simultaneously (**warps**)
- **Design goal is for threads to be as homogenous as possible**
  - Allows all thread processors in each multi-processor to execute the same instruction in parallel
- **NVIDIA *Quadro® FX 5800***
  - **4GB** of global memory with **240** thread processors

GPU = graphics processing unit

NVIDIA Quadro and the NIVIDIA logo are registered trademarks of NVIDIA Corporation in the U.S. and/or other countries.

GPU Technology

This work was done on a *Quadro® FX 5800*

- **Trade-offs**
  - Thread processor speed is slower than CPU core processor speed, but have many more of them
    - 240 threads at 650 MHz (vs. eight threads at 2.4 GHz *Quad Core Intel® Xeon® E5530*)
  - Can require significant algorithm changes to formulations
    - Languages like CUDA® are needed to manage threads, blocks, and grids
    - Code changes might be needed to optimize performance on different hardware
    - Libraries like MAGMA (Matrix Algebra on GPU and Multicore Architectures) make it easier to minimize re-coding
  - Can be bulky and consume a lot of power
    - 189 W (vs. 80 W *Quad Core Intel® Xeon® E5530*)
    - Purchased separate BOOSTERX 5 power supply

CPU = central processing unit, GPU = graphics processing unit, CUDA = Compute Unified Device Architecture
Intel and Xeon are registered trademarks of Intel Corporation in the U.S. and/or other countries.
Quadro and CUDA are registered trademarks of NVIDIA Corporation in the U.S. and/or other countries.

GPU Technology

This work was done on a *Quadro® FX 5800*

- **Other GPU hardware**
  - ***GeForce® GTX 470*** **(vs.** ***Quadro® FX 5800*****)**
    - **$350 approximate cost (vs. $3000)**
    - **448 cores (vs. 240)**
    - **1215 MHz thread processor speed (vs. 650 MHz)**
    - **1.2 GB global memory (vs. 4 GB)**
    - **133.9 GiB/sec memory bandwidth (vs. 102 GiB/sec)**
    - **215 W power consumption (vs. 189 W)**
    - **GF-100 Fermi architecture (vs. GT-200 Tesla)**

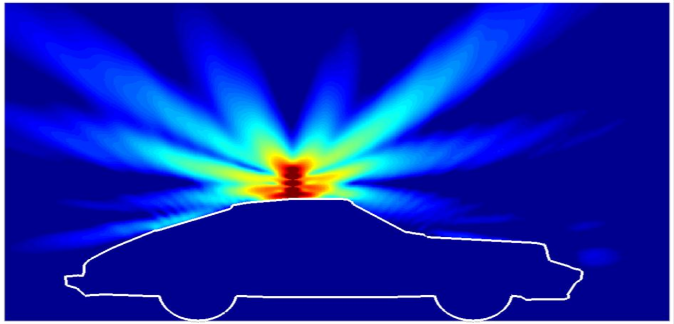
- **Still using Quadro® FX 5800 because of its larger global memory size**

GPU = graphics processing unit
Quadro and GeForce are registered trademarks of NVIDIA Corporation in the U.S. and/or other countries.

SAIC

GPU Applications

- First problem we looked at with GPUs was accelerating field profile computation for SBR
  - High degree of parallelism
  - Each footprint radiates to many observation points
- Details presented at HPEC 2010 (presentation) and GTC 2010 (poster session)
  - OptiX™ for accelerating the ray tracing
  - CUDA™ for accelerating the rest of the computation

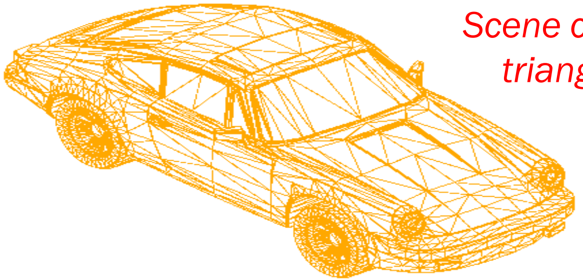


GPU = graphics processing unit   HPEC = High Performance Extreme Computing Conference   GTC = GPU Technology Conference

CUDA and OptiX are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries.

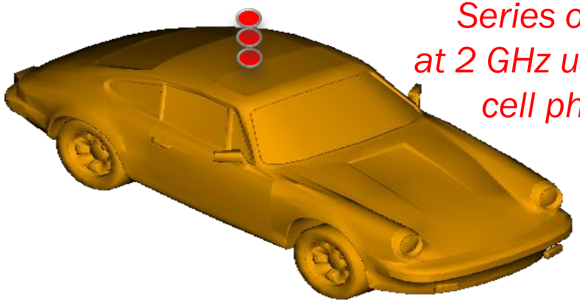
Example: Fun Car

- **Single frequency and multiple observation points**
  - *Maximum number of bounces = 1*
  - *Burst interval size = 0.25 degrees*
  - *Ray shoot sections = 1*
  - *Ray surface intersections = 485,406 per source*



Scene contains ***10,461*** *triangular surfaces*

- **Compare simulation times for field profiles**
  - 15 cm resolution (1 point per wavelength)
  - 6 cm (2.5 points per wavelength)
  - 3 cm (5 points per wavelength)
  - 15 mm (10 points per wavelength)
  - 7.5 mm (20 points per wavelength)

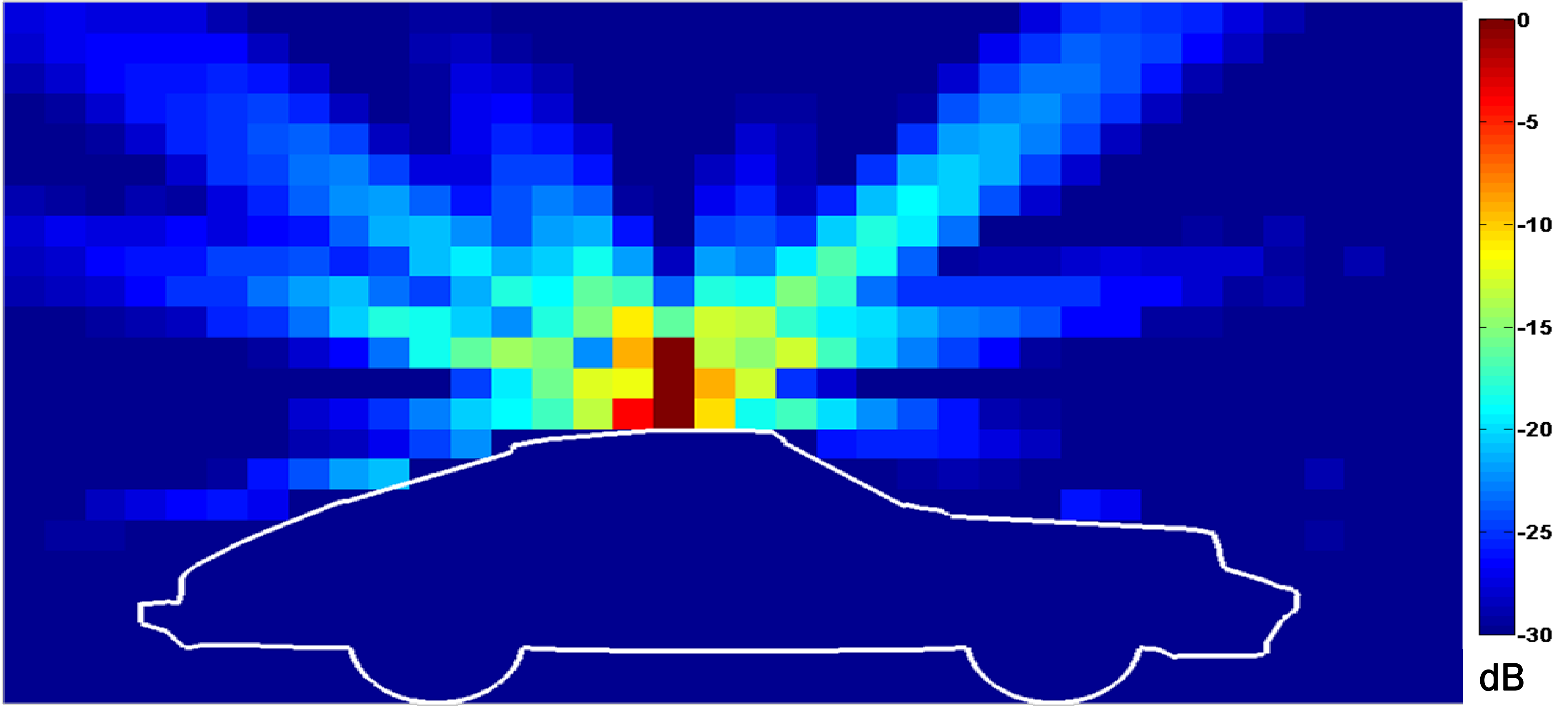


Series of 41 point sources at 2 GHz used to model 12-inch cell phone car antenna

SAIC®

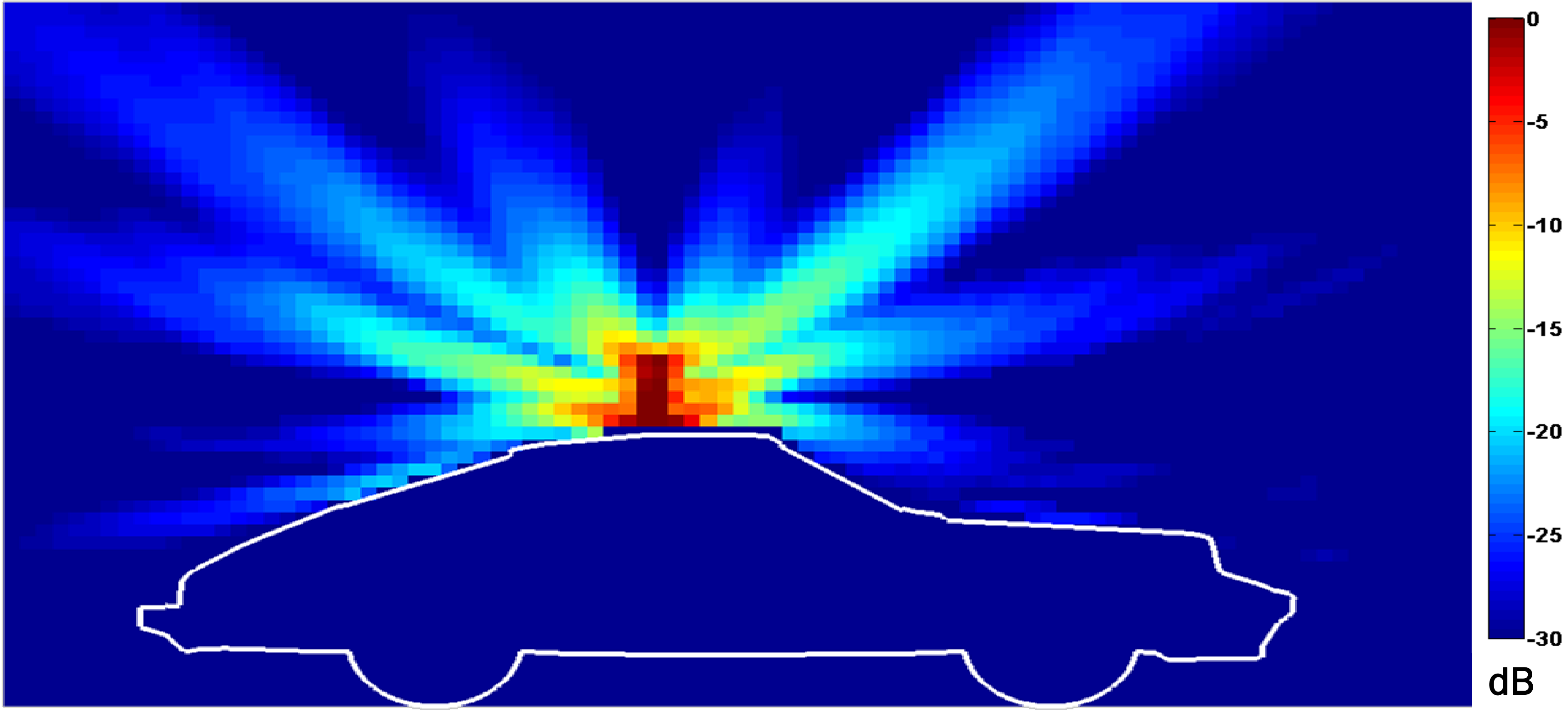
Example: Fun Car

15 cm resolution (737 observation points), 3.1 min (GPU version)



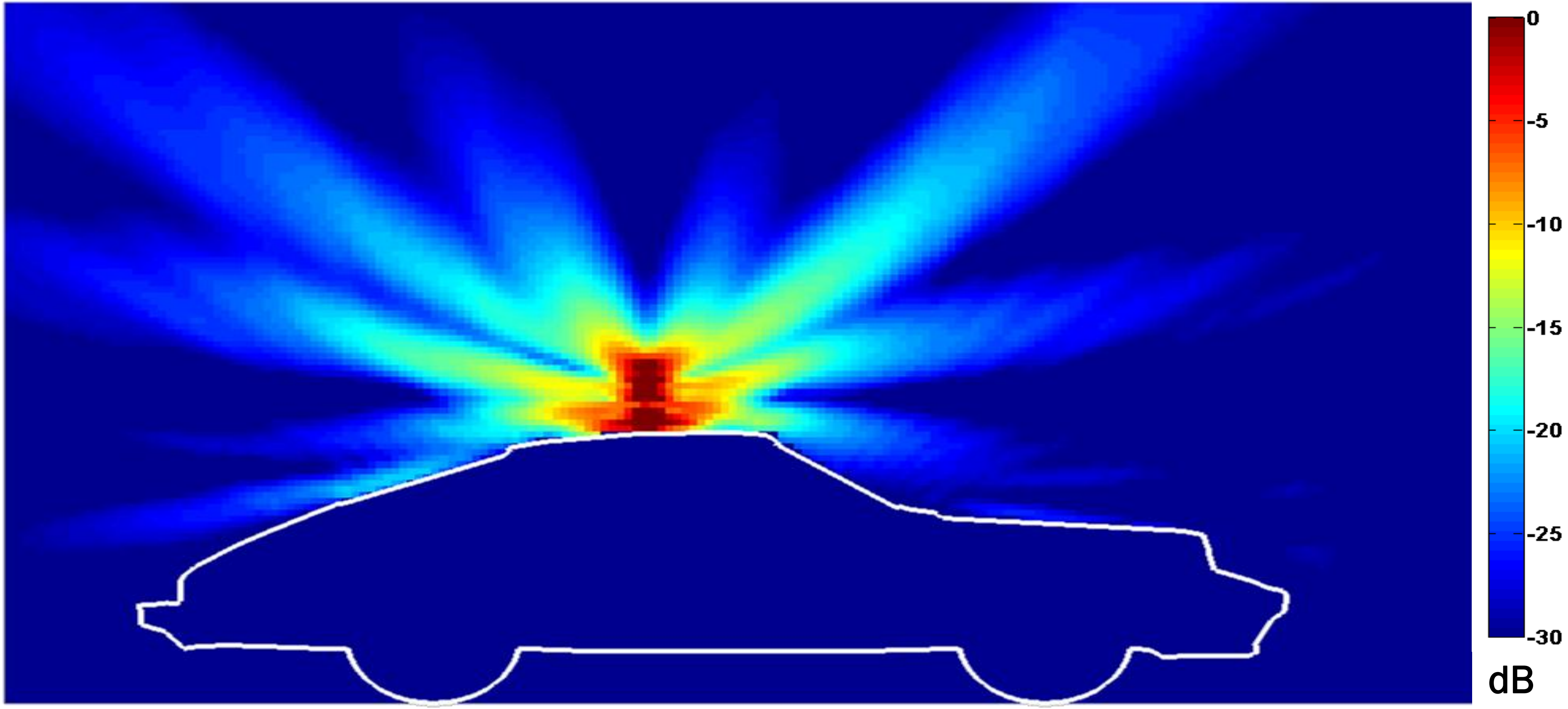
Example: Fun Car

6 cm resolution (4,483 observation points), 4.4 min (GPU version)



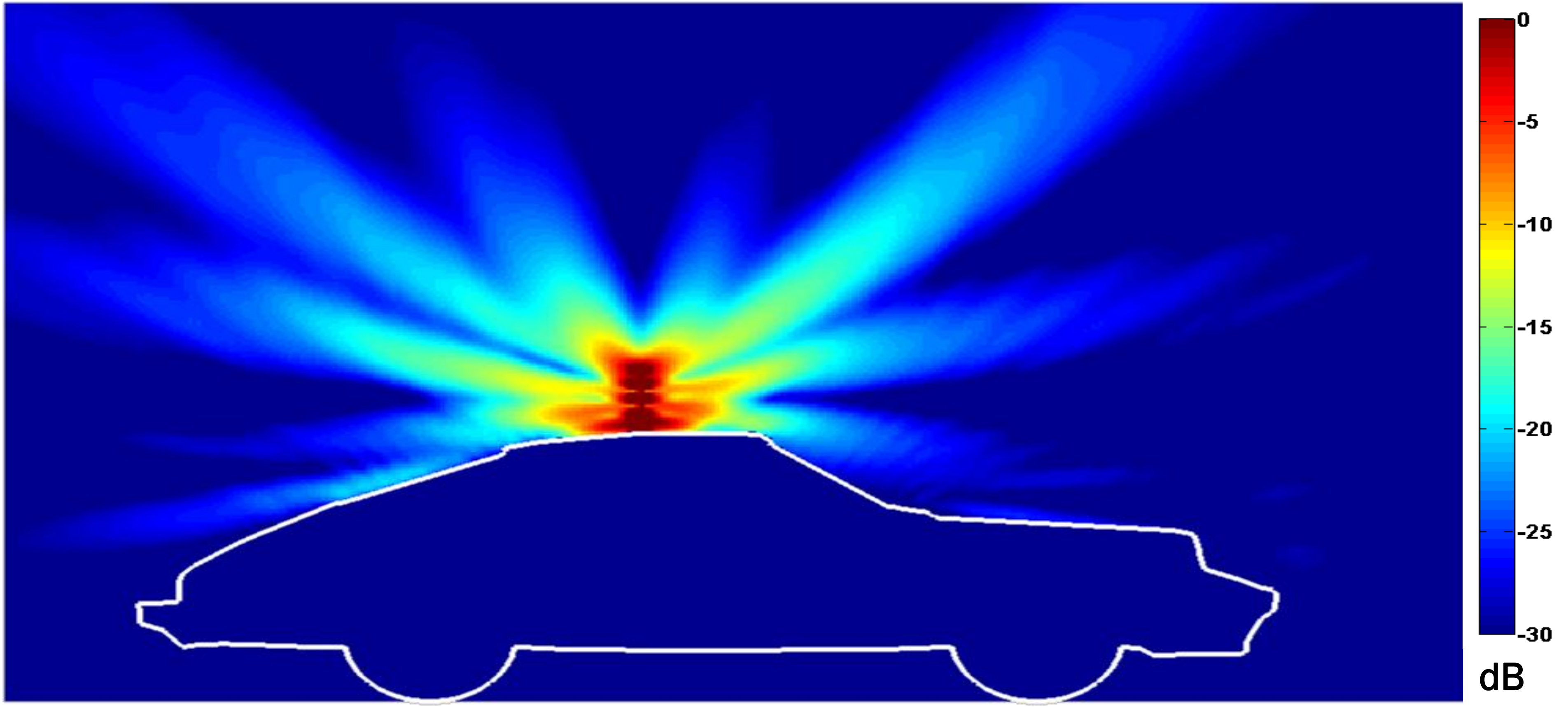
Example: Fun Car

3 cm resolution (17,741 observation points), 15.3 min (GPU version)



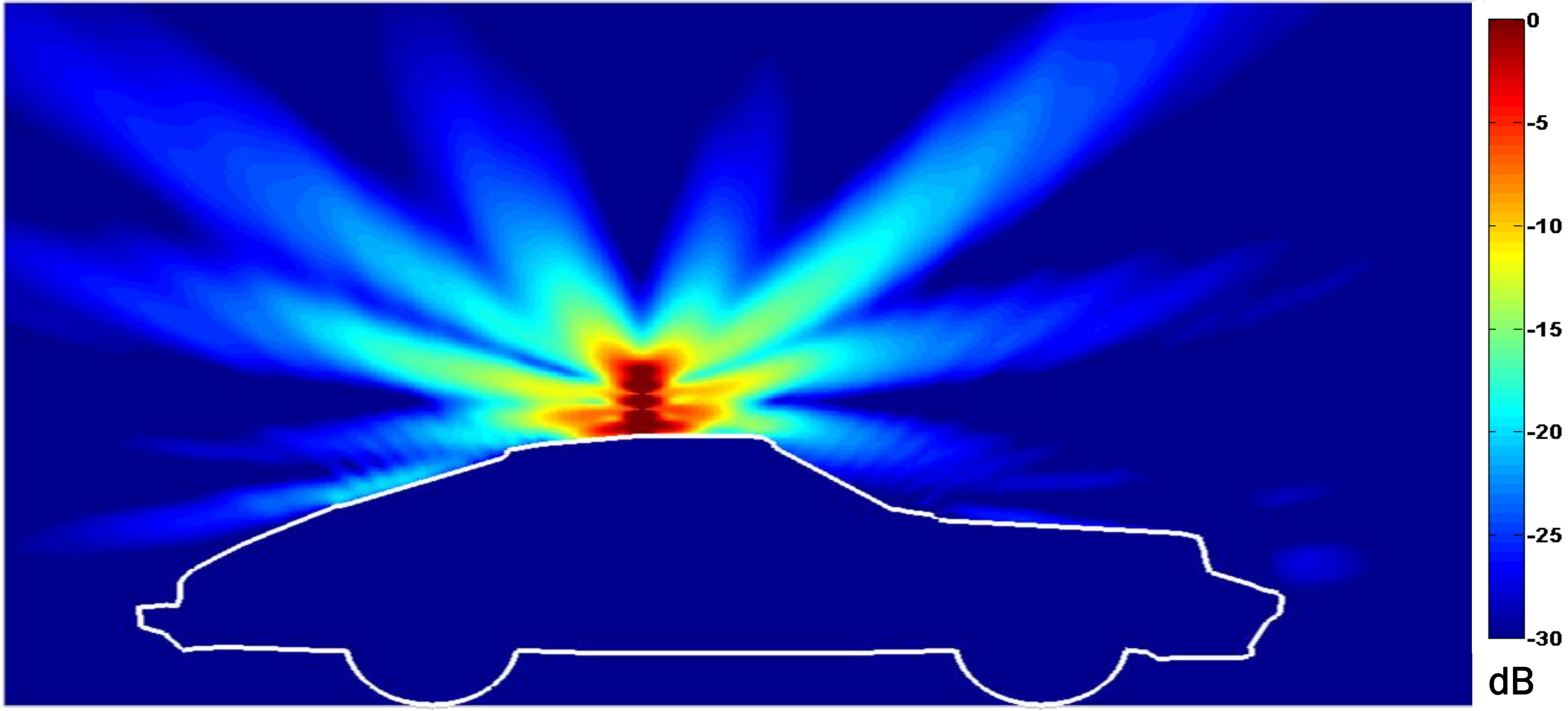
Example: Fun Car

15 mm resolution (70,754 observation points), 0.95 hrs (GPU version)

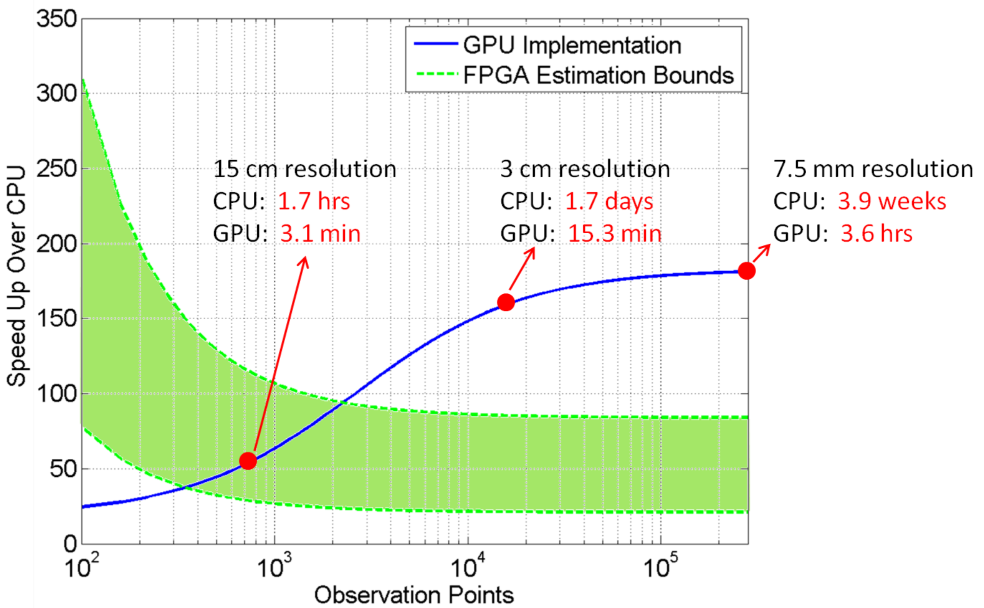


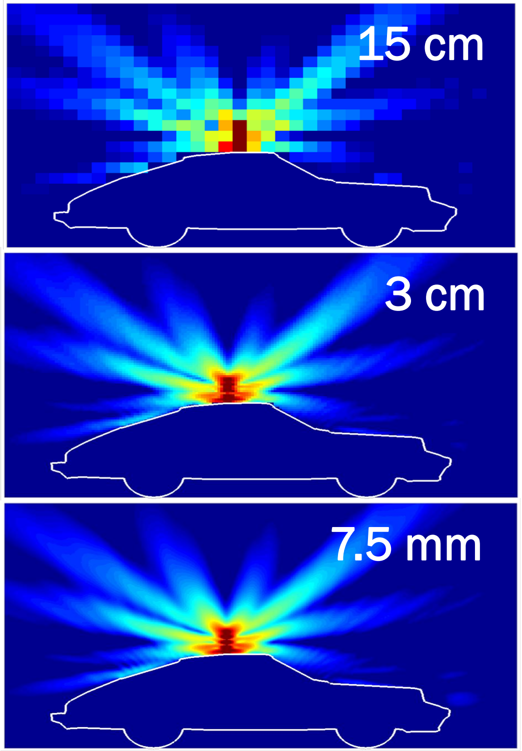
Example: Fun Car

7.5 mm resolution (282,242 observation points), 3.6 hrs (GPU version)



Example: Fun Car



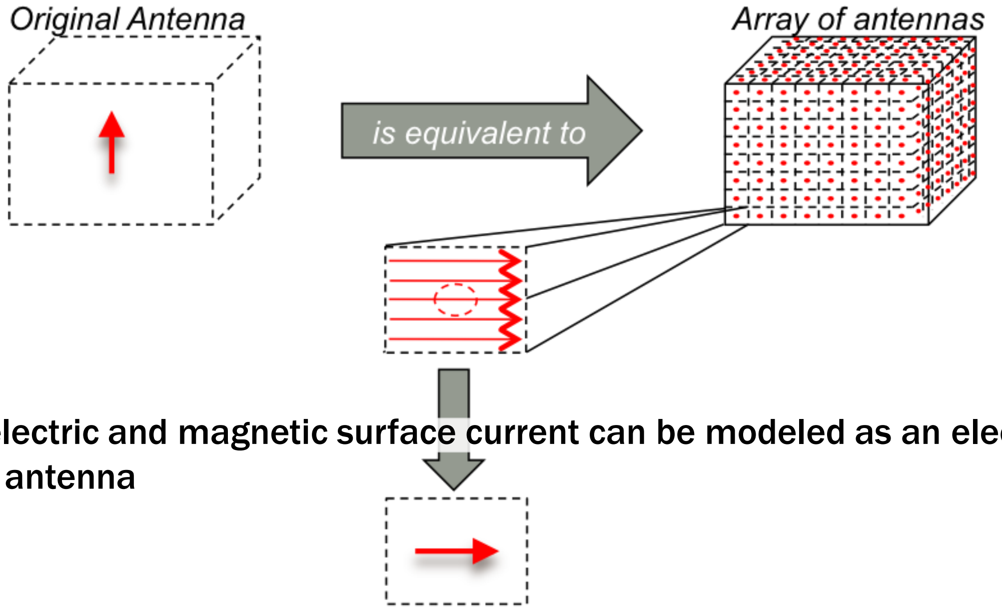


CPU / FPGA times are PO time (does not include SBR time)
GPU times are total execution time (PO+SBR)
CPU time measured for single source and scaled by 41

CPU = central processing unit GPU = graphics processing unit PO = physical optics
SBR = shooting and bouncing ray   FPGA = field programmable gate array

Could this be applied to many antennas?

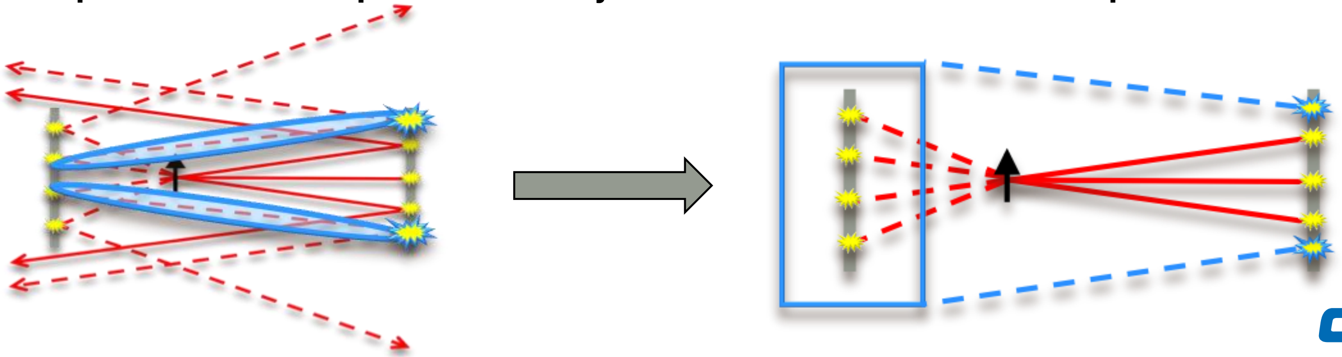
- **Surface equivalence theorem**
  - **A single antenna can be replaced with many antennas distributed over a surface**



  - **Each patch of electric and magnetic surface current can be modeled as an electric and magnetic Hertzian dipole antenna**

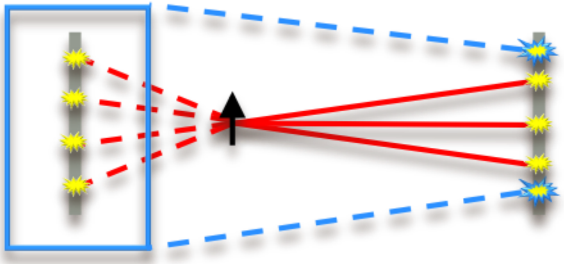
Could this be applied to many antennas?

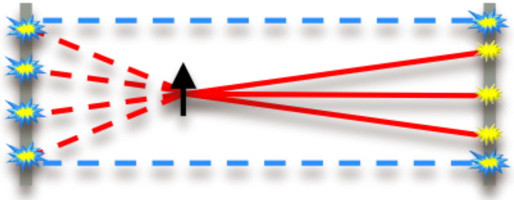
- **Surface equivalence theorem**
  - **A single antenna can be replaced with many antennas distributed over a surface**
- **Back to our resonance problem …**
  - **Single plate simulation fine, double plate simulation had errors**
  - **Error must come from using rays to propagate fields between the plates**
- **Replace each plate with an equivalent array to find the fields incident upon the other plate**



Alternative Formulation

- **Why not just use the original footprints to radiate between plates instead of first radiate to an equivalent surface?**





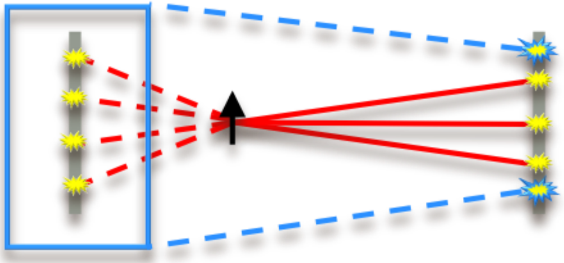
- **This approach allows for hybridization where instead of using SBR alone, could use a full-wave solver to simulate one plate**

- **Details of implementation are identical whether work with footprints or equivalent surface**

SBR = shooting and bouncing ray

Implementation

- **This is well suited for a GPU …**



Many sources get radiated to many observation points

- **Based on device memory, break combinations of sources and observations into chunks**

- **Two-stage process**
  - Evaluate the fields for each chunk (`computeField` kernel)
  - Parallel reduction scheme to combine the fields from all the sources (`reduceField` kernel)

GPU = graphics processing unit

Implementation

First step is to group the combinations into chunks

- Sources and observations are grouped into chunks
  - Sources into chunks of 65,535 (or less)
  - Observations into chunks based on remaining device memory

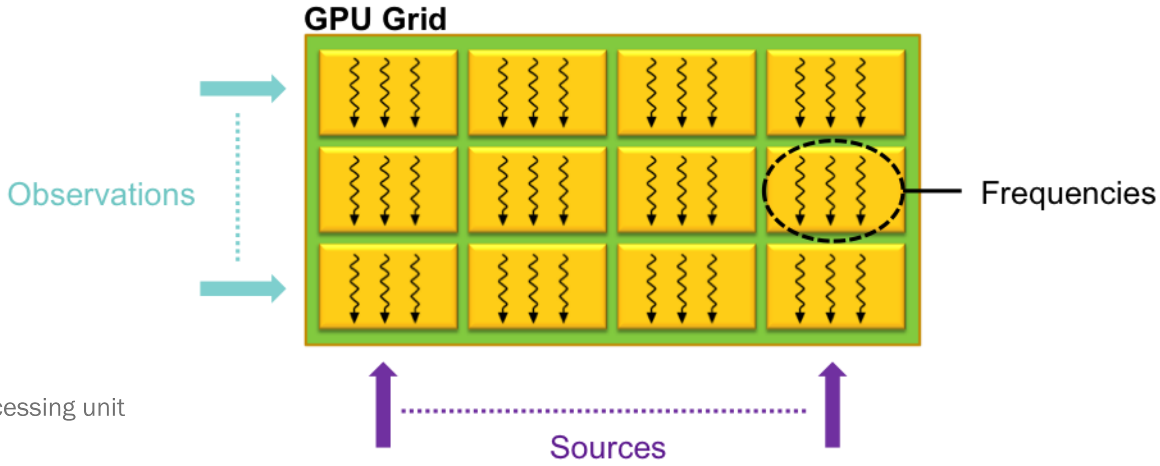
- Higher source densities will require more chunks
  - 1 point per wavelength
    - 384 sources grouped into one chunk of 384 sources
    - 10,000 observations grouped into one chunk of 10,000 observations
    - `computeField` kernel only called once
  - 20 points per wavelength
    - 153,600 sources grouped into three chunks
    - 10,000 observations grouped into eight chunks of 1,343 observations
    - `computeField` kernel called 24 times

Implementation

First step is to group the combinations into chunks

```
for each observation chunk ...
   for each source chunk ...
            computeField<<<dimGrid,dimBlock>>>
   end
   for each tree level ...
            reduceField<<dimGrid,dimBlock>>>
   end
end
```

majority of simulation time



GPU = graphics processing unit

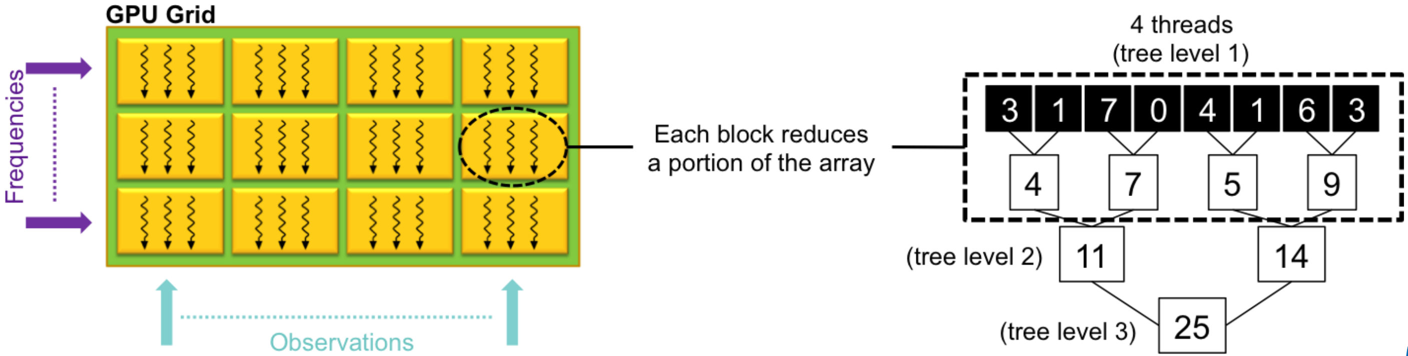
SAIC®

Implementation

Second step is to simulate and reduce each chunk

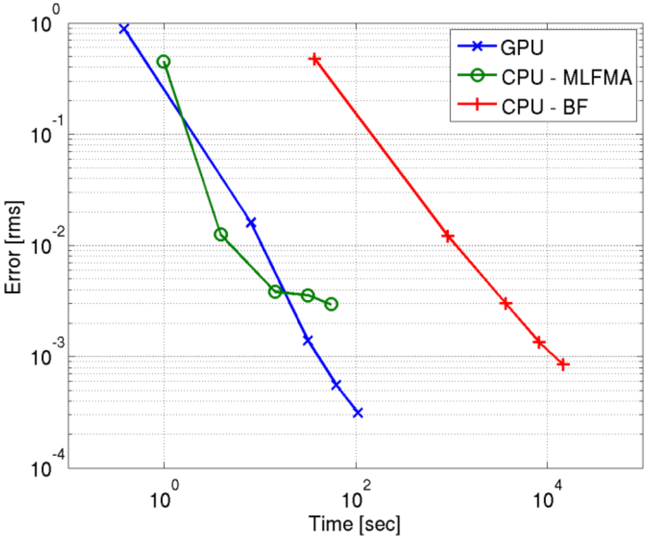
```
for each observation chunk ...
   for each source chunk ...
            computeField<<<dimGrid,dimBlock>>>
   end
   for each tree level ...
            reduceField<<dimGrid,dimBlock>>>
   end
end
```

negligible simulation time



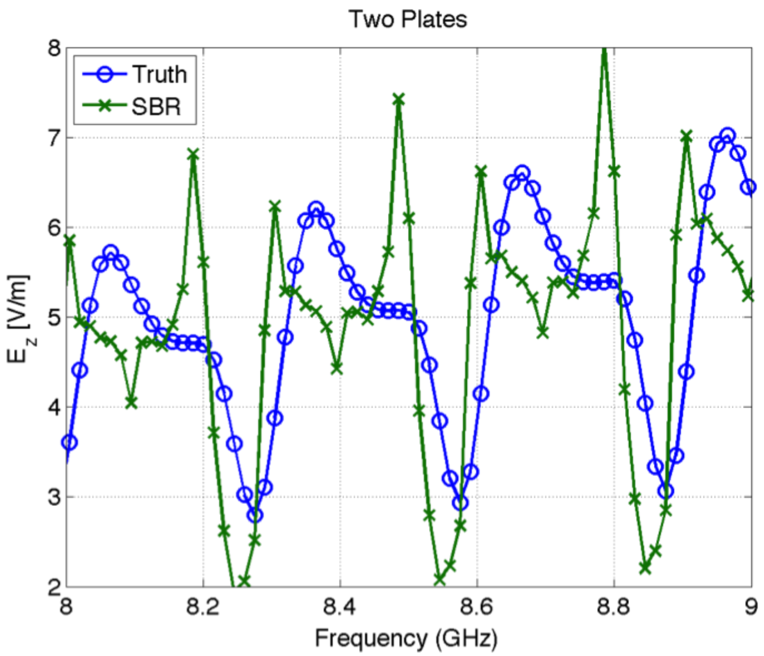
Results

- **Recreated a uniform plane plane (at a single frequency)**
  - **Increasing the number of antennas improves accuracy, but increases computation**



GPU = graphics processing unit CPU = central processing unit MLFMA = multilevel fast multipole algorithm BF = brute force

Results



Two Plates
Truth
SBR
E_z [V/m]
Frequency (GHz)

Two Plates
Truth
Two-Part SBR/SBR
E_z [V/m]
Frequency (GHz)

Application of the GPU to a Two-Part Computational Electromagnetic Algorithm

- **The Team**
  - **Who we are ...**
    We solve electromagnetic radiation, scattering, and coupling problems.
  - **What tools we use ...**
    We develop, distribute, and employ tools that use full-wave, asymptotic, and hybrid methods.
- **The Problem**
  - **What is our research trying to address ...**
    The basic SBR asymptotic method does not accurately solve problems involving resonance.
- **The Solution**
  - **How have we used GPU technology to help us ...**
    Replacing each resonant part with an equivalent antenna array improves the accuracy.
    Since each antenna's radiation can be computed individually, the GPU computation is very efficient.

GPU = graphics processing unit SBR = shooting and bouncing ray

SAIC®

Thank You

Eric A. Dunn, Electromagnetic Research Scientist

4001 N Fairfax Dr, Fourth Floor | Arlington, VA 22203

Tel: 703.312.6507 | Email: eric.a.dunn@saic.com

Visit us at saic.com