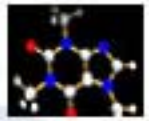
SciDAC

Scientific Discovery through Advanced Computing

Enabling Petascale Science through Combinatorial Algorithms

Alex Pothen

CSCAPES Institute

www.cscapes.org

Old Dominion University

Sandia National Laboratories

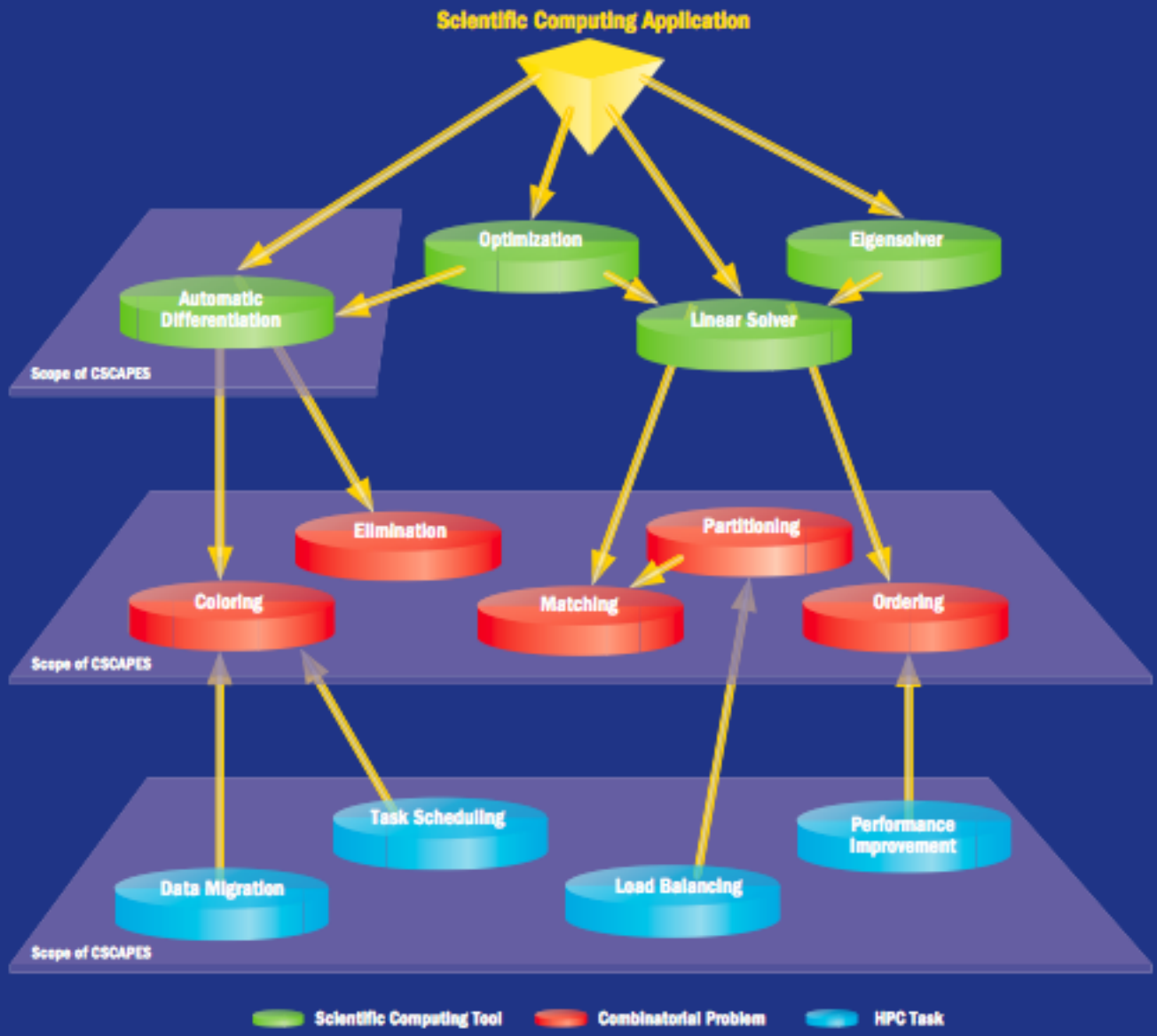
Argonne National Laboratory

The Ohio State University

Colorado State University

CSC Community Thanks CERFACS

- **CSCAPES Institute is a natural outgrowth of the CSC community, which has been fostered by CERFACS for the past 20 years.**
- **Thank you to CERFACS researchers, who have led the way on sparse solvers, parallel computing, matchings, mathematical software, and many other topics!**
- **Thanks for organizing the CSC05 Workshop at CERFACS!**
- **Thanks for organizing the annual Sparse Days conferences, a remarkable forum for discussions, good food and wine!**


Scientific Computing Application
Optimization
Eigensolver
Automatic Differentiation
Linear Solver
Scope of CSCAPES
Elimination
Partitioning
Coloring
Matching
Ordering
Scope of CSCAPES
Task Scheduling
Performance Improvement
Data Migration
Load Balancing
Scope of CSCAPES
Scientific Computing Tool
Combinatorial Problem
HPC Task

Parallelization,

Load Balancing

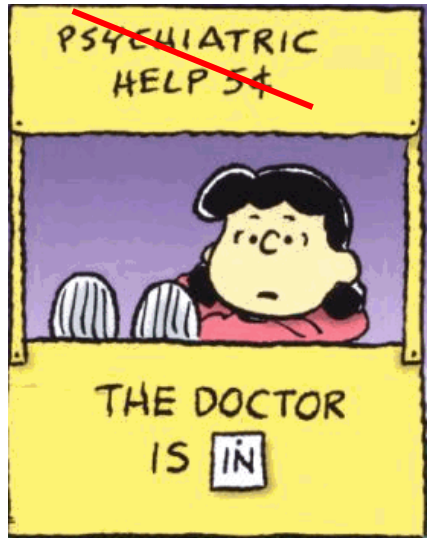
Graph Coloring

Performance

Automatic Differentiation

Graph Matching

Combinatorial problems?

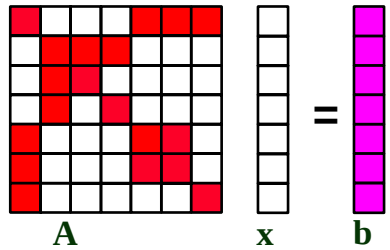


Load Balancing

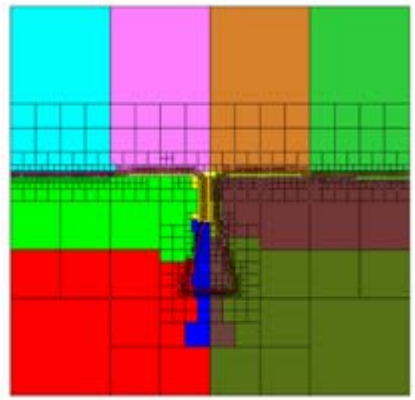
…enabling parallelization and fast run-times for irregular applications

Partitioning and Load Balancing

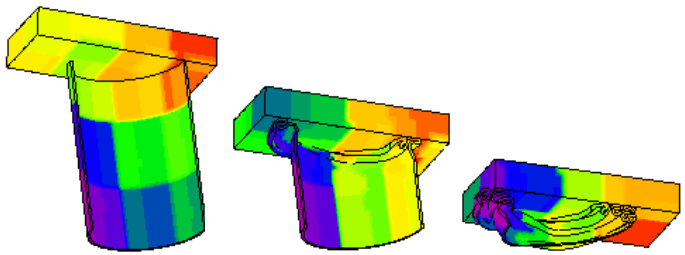
- Goal: **assign data (and tasks) to processors to**
  - minimize application runtime
  - maximize utilization of computing resources
- Metrics:
  - minimize processor idle time (balance workloads)
  - keep inter-processor communication costs low
- **Impacts performance of a wide range of simulations**
  - Accelerator code speeded up 3X with a geometric partitioner
- **Several partitioning and load balancing algorithms**



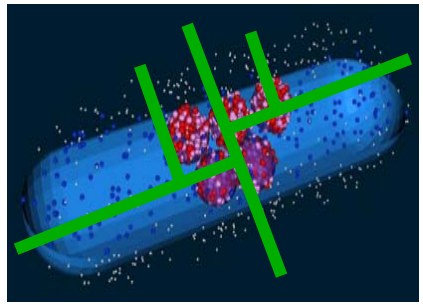
Linear solvers & preconditioners



Adaptive mesh refinement



Contact detection

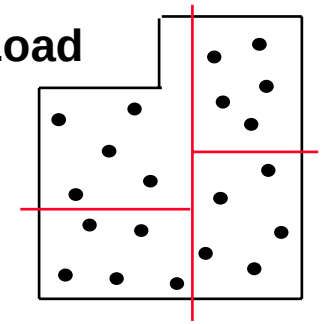


Particle simulations

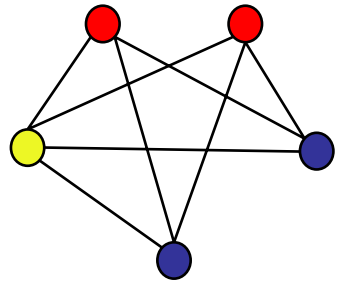
cscapes

Zoltan Toolkit: Data Services for Dynamic Applications

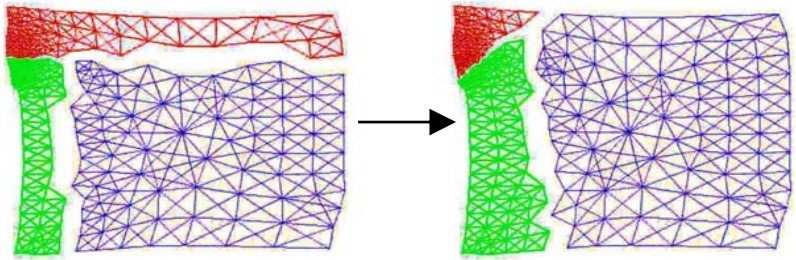
Dynamic Load Balancing



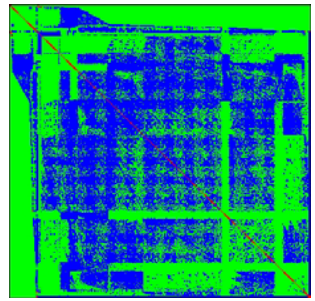
Graph Coloring



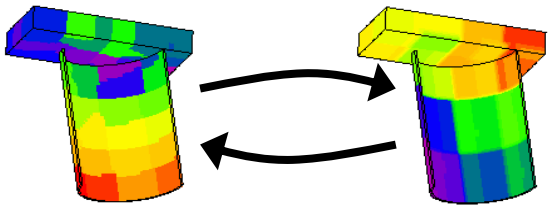
Data Migration



Matrix Ordering



Unstructured Communication



Distributed Data Directories

A	B	C	D	E	F	G	H	I
0	1	0	2	1	0	1	2	1

cscapes

Zoltan 3.0

Zoltan 3.0 is now available (www.cs.sandia.gov/Zoltan). New features use hypergraphs for modeling communications accurately:

- **Hypergraph repartitioning**
  - Reduces total communication in dynamic applications.
- **Hypergraph refinement**
  - Quickly improves an existing parallel distribution (partitioning)
- **Hypergraph partitioning with fixed vertices**
  - Allows application to fix certain data to specific processors.
- **Hierarchical partitioning**
  - 2-level partitioning, possibly using different algorithms, cost metrics
  - Useful for complex computer architectures (e.g., multi-core)

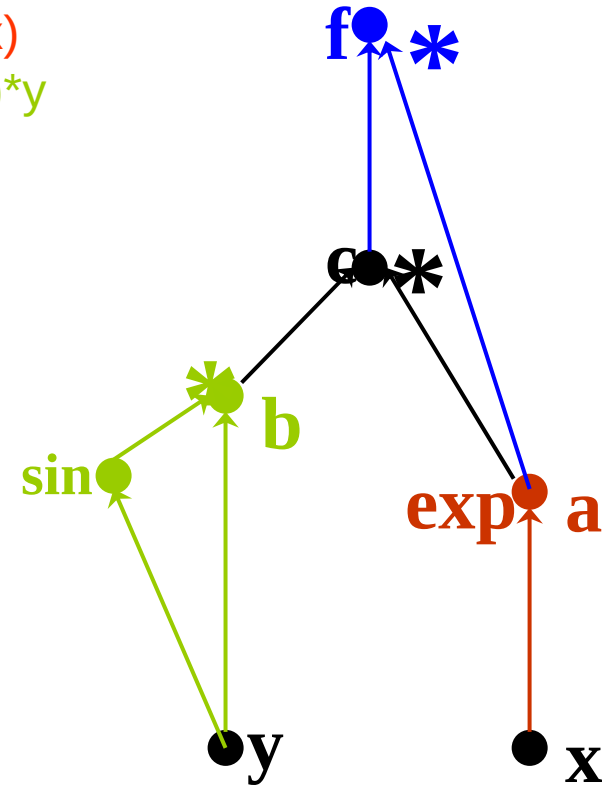
Automatic Differentiation

…enabling the solution of nonlinear differential equations, optimization, sensitivity analysis, uncertainty quantification, etc.

AD: Introduction

- **Transforms code for computing a function into code for differentiating it**
- **Function computed from intrinsic operations, and modeled by a directed acyclic graph (DAG)**
- **Compute derivatives by composing partial derivatives for each operation, using the chain rule on the DAG**
- **Efficiency of generated code depends on sophistication of compiler analysis and combinatorial algorithms**

a = exp(x)
b = sin(y)*y
c = a*b
f = a*c

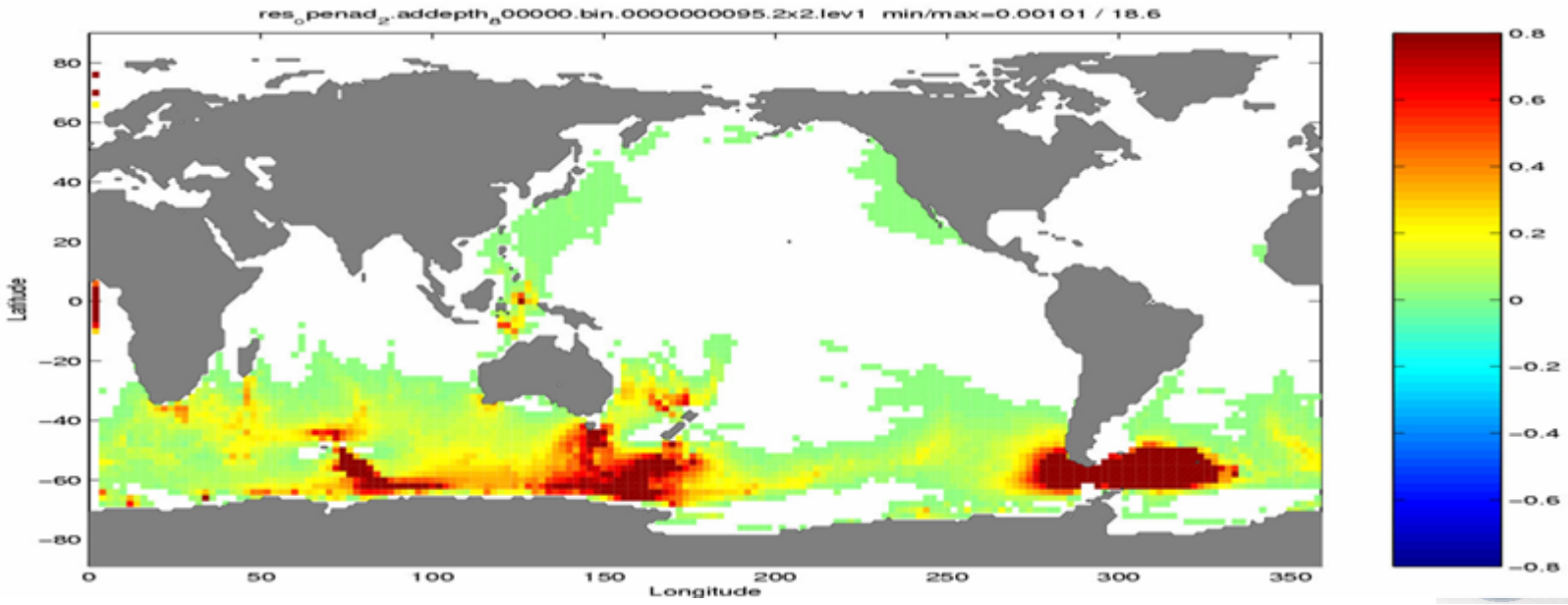


AD: Combinatorial Problems

- **Parallel algorithms for differentiating reduction operations**
- **Reduce operations and storage needed to compute the derivatives by evaluating the DAG in suitable orders**
  - Two extreme modes: Forward and Reverse
  - Modeled as vertex and edge elimination in DAG
  - Stop at some intermediate stage to find minimum storage
- **Location of checkpoints in reverse mode**
- **Graph coloring for computing many derivatives in one AD pass through the DAG**
- **Integration with PETSc and Zoltan toolkits**

Sensitivity analysis in climate model

- **Sensitivity of flow through Drake Passage to ocean bottom topography (P. Heimbach, MIT)**
  - Finite difference approximations: 23 days
  - Naïve automatic differentiation: 2 hours 23 minutes
  - Smart automatic differentiation: 22 minutes



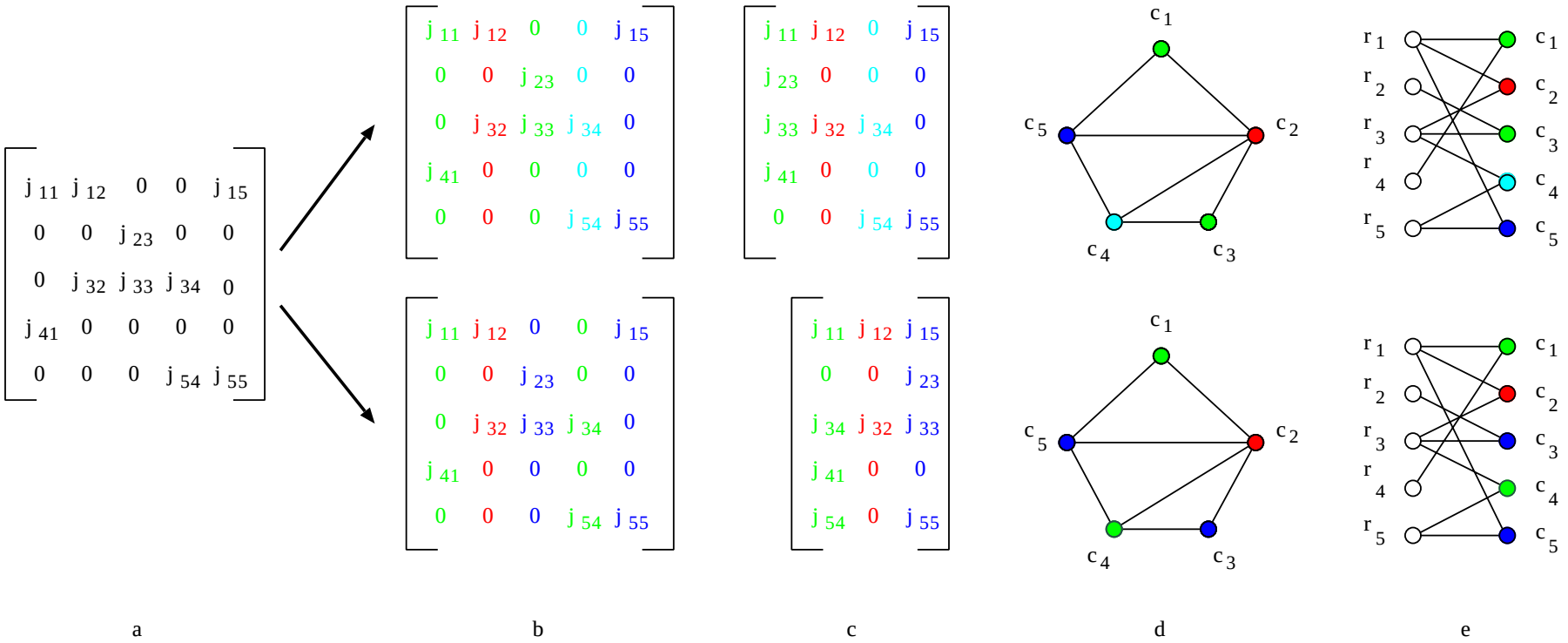
AD: Current Capabilities

- **Fortran 77: ADIFOR 2.0/3.0**
  - Robust, mature tool with excellent language coverage
  - Excellent compiler analysis
  - Efficient forward mode; adequate reverse mode
- **C/C++: ADIC 2.0**
  - Semi-mature tool with full C language coverage
  - Sophisticated differentiation algorithms
  - Efficient forward mode
- **Fortran 90: OpenAD/F**
  - New tool with partial language coverage
  - Sophisticated differentiation algorithms
  - Accurate and novel compiler analysis
  - Innovative templating mechanism
  - Efficient forward and reverse modes

Graph Coloring

…reducing work in Automatic Differentiation; and discovering parallelism in computations

Coloring and Jacobian Computation



Original Jacobian

Compressed **representation (*Structurally orthogonal* columns packed together)**

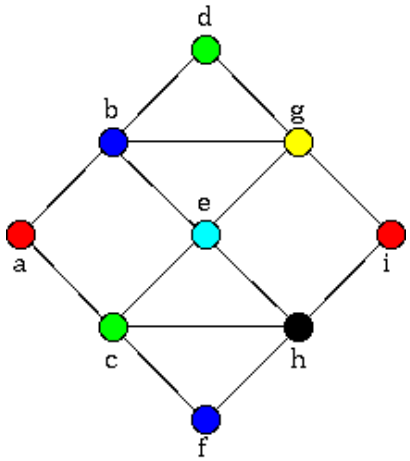
D1 coloring **formulation on column inter. graph**

D2 coloring **bipartite graph**

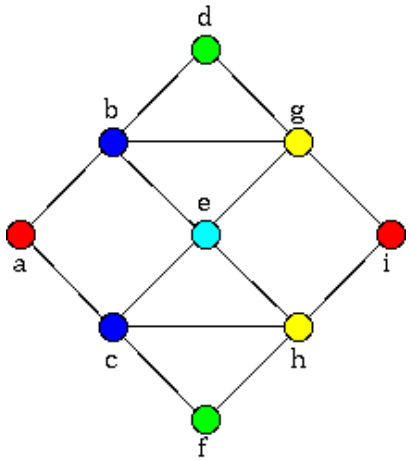
Coloring and Derivatives: The Big Picture

- **Scenarios and coloring models:**
  - unsymmetric vs symmetric matrix
  - direct vs substitution method
  - uni- vs bi-directional partitioning
- **Developed novel sequential algorithms**
- **Future plans**
  - Develop parallel versions
  - integrate with AD tools

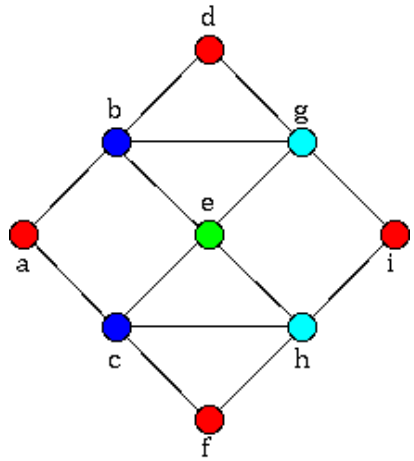
	1d partition	*2d partition*	
Jacobian	*Distance-2 coloring*	*Star bicoloring*	*Direct*
Hessian	*Star coloring*	*NA*	*Direct*
Jacobian	*NA*	*Acyclic bicoloring*	*Subst*
Hessian	*Acyclic coloring*	*NA*	*Subst*



D2 coloring



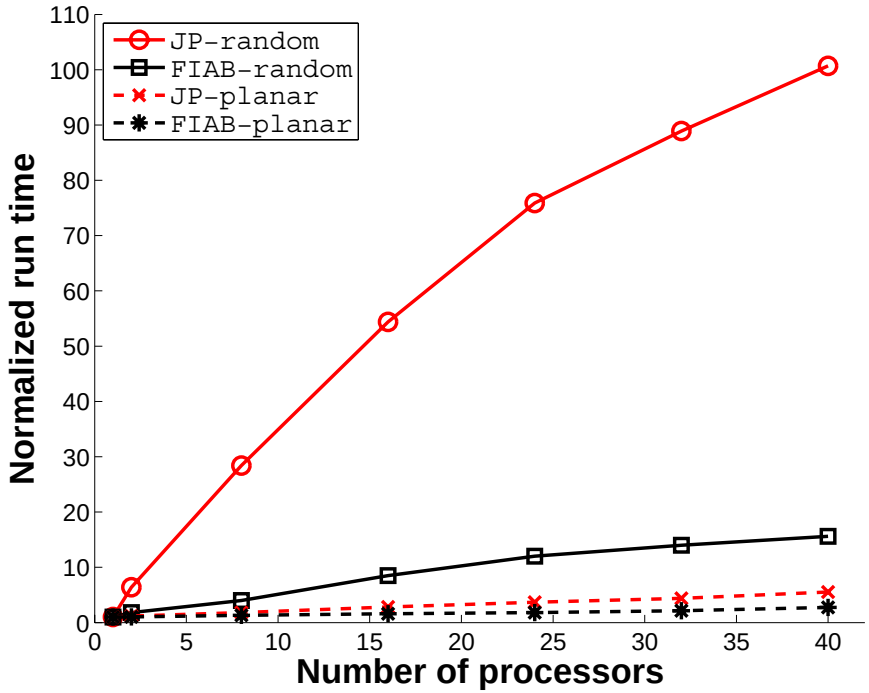
Star coloring



Acyclic coloring

Framework for parallel coloring

- **Essential ingredients of framework:**
  - Partition graph on processors, and speculatively color subgraphs in rounds
  - Exchange color info after a superstep (coloring a specified no. of vertices)
  - Detect conflicts after each round, resolve using randomization, recolor when needed
- **Applied to D-1 and D-2 coloring, implemented in MPI; available in Zoltan**
- **Extending the framework to**
  - Tera- and peta-scale machines
  - Other graph problems

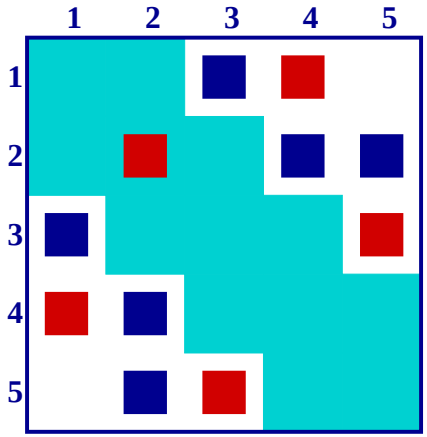


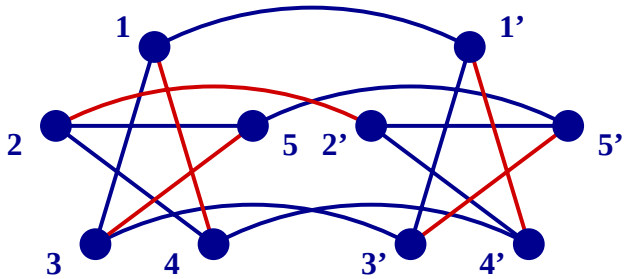
Weak scalability **on two families of graphs: random (unstructured); planar (structured).**

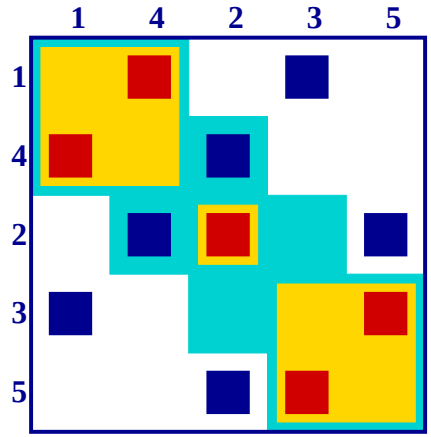
Matchings in Graphs

…enabling load balancing and linear solvers

Matchings in Graphs

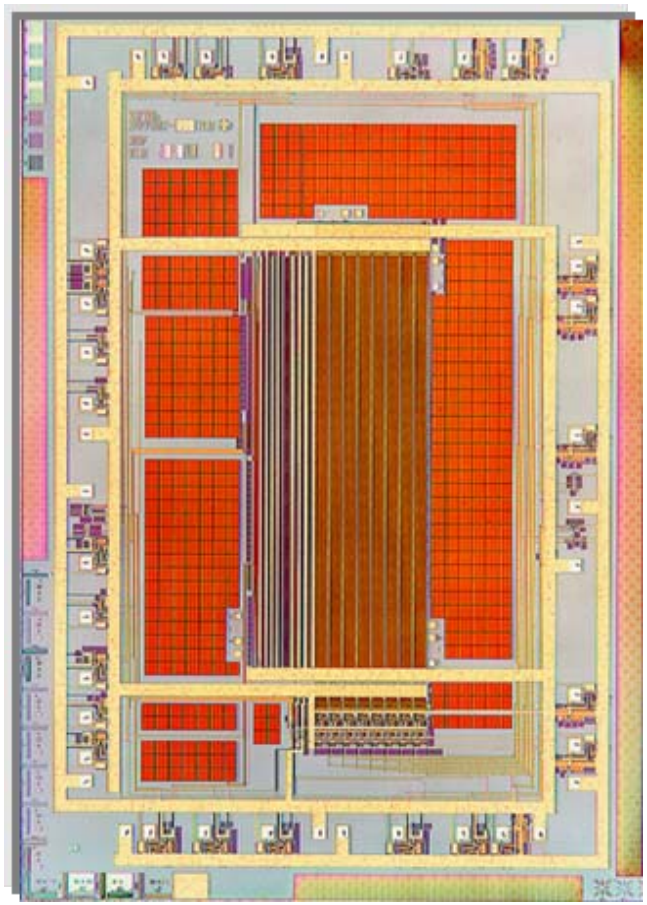


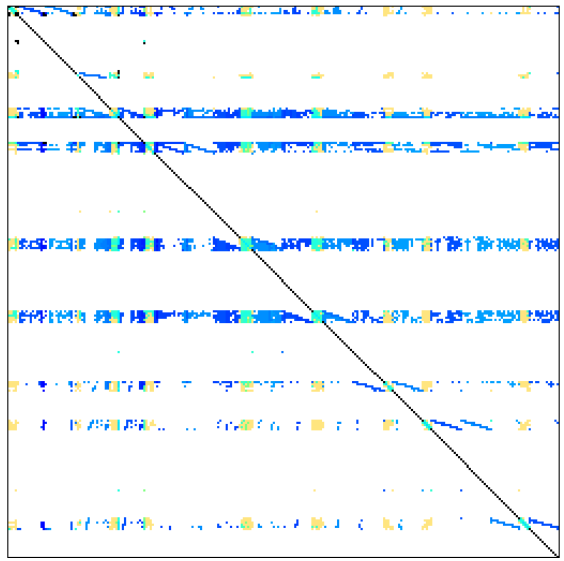


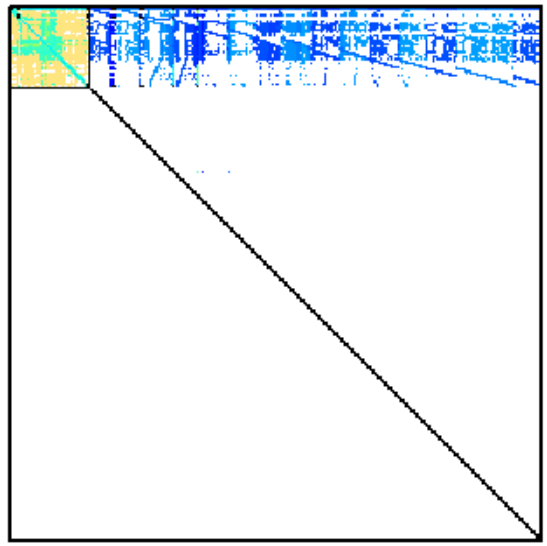


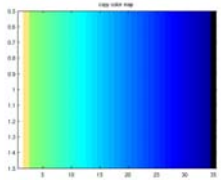
- **Matching is a pairing of vertices; a vertex is paired with one neighboring vertex or none**
- **Applications**
  - Place large elements on diagonals of matrices for solvers
  - Block triangular form to reduce work in solvers, improve condition number
  - Coarsening step in multilevel graph and hypergraph partitioners

Block Upper Triangular Form (BTF)









Circuit model from Xyce (Hoekstra, Day; Sandia) 683K rows, 2M nnz, 584K diag blocks
Solved 200 times faster! 100M problem waits.

Outreach and Training

- **Organized the SIAM Workshop on CSC in Feb. 2007. 100 attendees, 12 early career researchers supported. SIAM News article in May 2007. URL: www.cscapes.org, click on CSC07**
- **International collaborations with CERFACS, AD groups in Germany, CSC groups in Norway, Netherlands, and others.**
- **5 Postdoctoral researchers, 4 PhD students, and an undergraduate are involved in CSCAPES research, and are co-mentored by Lab scientists.**
- **Working with several enabling technology and applications groups to integrate CSC software and solve their combinatorial problems.**
- **We welcome application kernels where CSC issues are significant; tell us about your combinatorial problems!**

CSCAPES Institute

www.cscapes.org

Erik Boman, Karen Devine, Bruce Hendrickson;

Paul Hovland, Boyana Norris, Jean Utke;

Alex Pothen, Assefaw Gebremedhin, Florin Dobrian, Sudip Seal;

Umit Catalyurek, Doruk Bozdag;

Michelle Mills Strout

Old Dominion University

Sandia National Laboratories

Argonne National Laboratory

The Ohio State University

Colorado State University