SAND2020-1649C

ExaGraph: Parallel Partitioning and Coloring for Exascale Applications

Seher Acer, Erik Boman, Siva Rajamanickam

Sandia National Labs*

Ian Bogle (RPI/Sandia), George Slota (RPI/Sandia)

ECP Annual Meeting, Houston, Feb. 2020

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

ExaGraph: Sandia focus

Problems/topics:

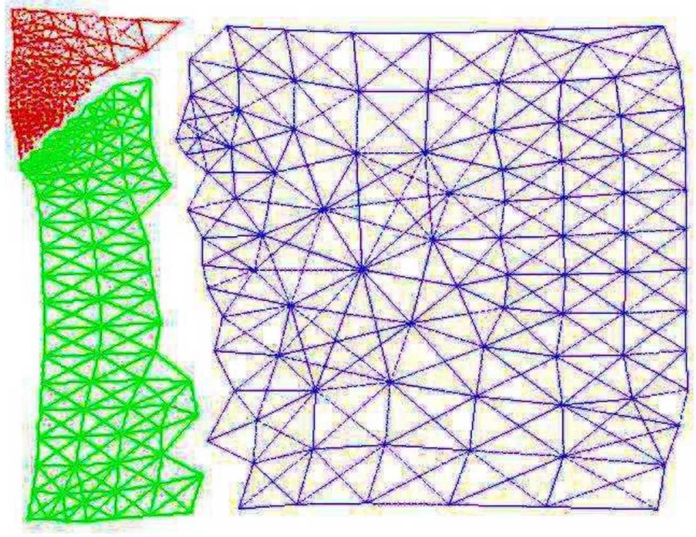
- Graph coloring
- Graph partitioning
- Machine learning using graphs

Software: We will deliver through existing software

- Trilinos/Zoltan2 (distributed-memory)
- KokkosKernels (shared-memory)

Load Balancing and Partitioning

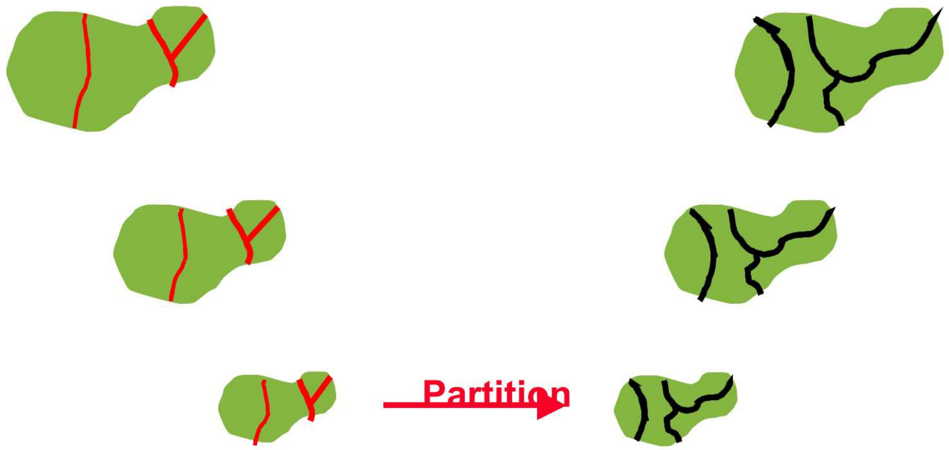
- Partitioning:
  - Assignment of application data to processors.
    - For example: mesh points, elements, matrix rows.
- Ideal partition:
  - Work (load) is well balanced among proc.
  - Inter-processor communication is kept low.
    - Low communication volume, few messages, etc.



Partition of an unstructured finite element mesh for three processors

Exascale Graph Partitioning

- Multilevel: Highly successful graph partitioning method
  - Bui & Jones (1993); Hendrickson & Leland (1993); Karypis and Kumar (1995)
  - Construct smaller approximations to graph.
  - Perform graph partitioning on coarse graph.
  - Propagate partition back, refine as needed (typically each level)
- Software
  - Graphs: (Par-)Metis, Scotch, KaHip/Kaffpa, …
  - Hypergraphs: PaToH, hMetis, Zoltan/PHG, Mondriaan
- Problems
  - Algorithms are hard to parallelize
  - Quality and run-time often deteriorate for large #proc (cores)
  - **No software for multi-GPU** (or similar accelerators)!

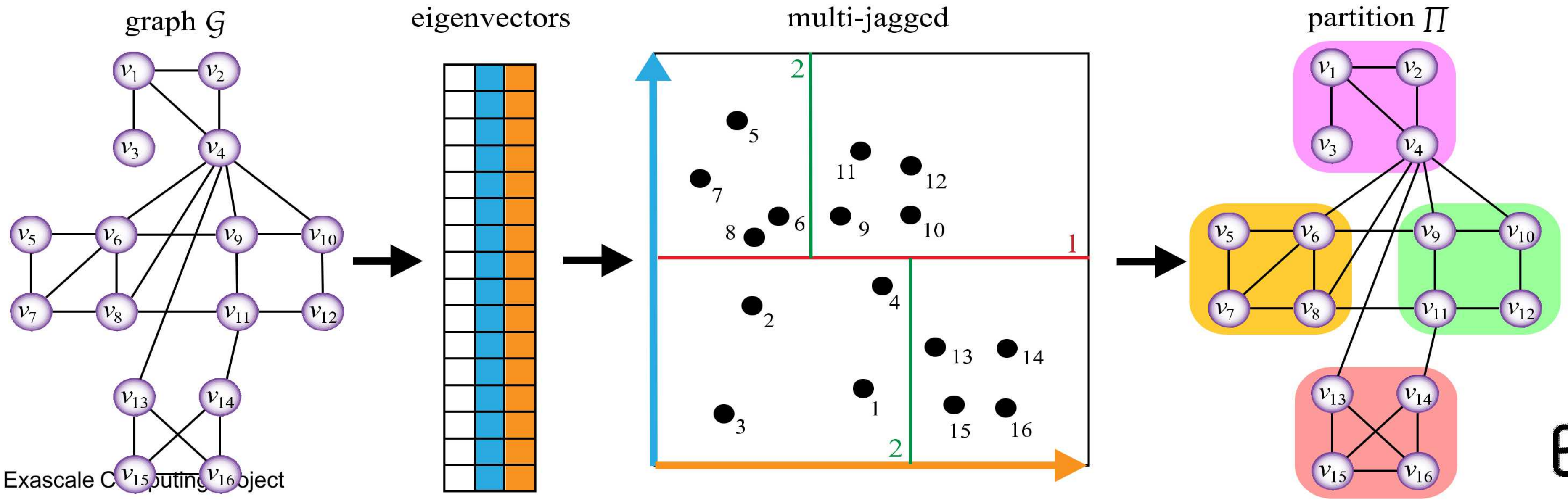


Spectral partitioning

- Algebraic algorithm based on graph Laplacian: *min* $\frac{x^T L(G) x}{x^T x}$
  – Partitioning: Fiedler ('73), Donath, Hoffman ('73), Pothen, Simon, Liou ('90)
  – Clustering: Hagen, Kahng ('92), Shi, Malik ('00), Ng et al. ('02)
- Pros:
  – Matrix-based, can reuse linear algebra software (e.g. Trilinos)
  – In parallel, quality is almost independent of # processors
  – Well suited to GPU and accelerators
- Cons:
  – Computationally expensive
  – Poor quality in some cases

Exagraph Partitioning

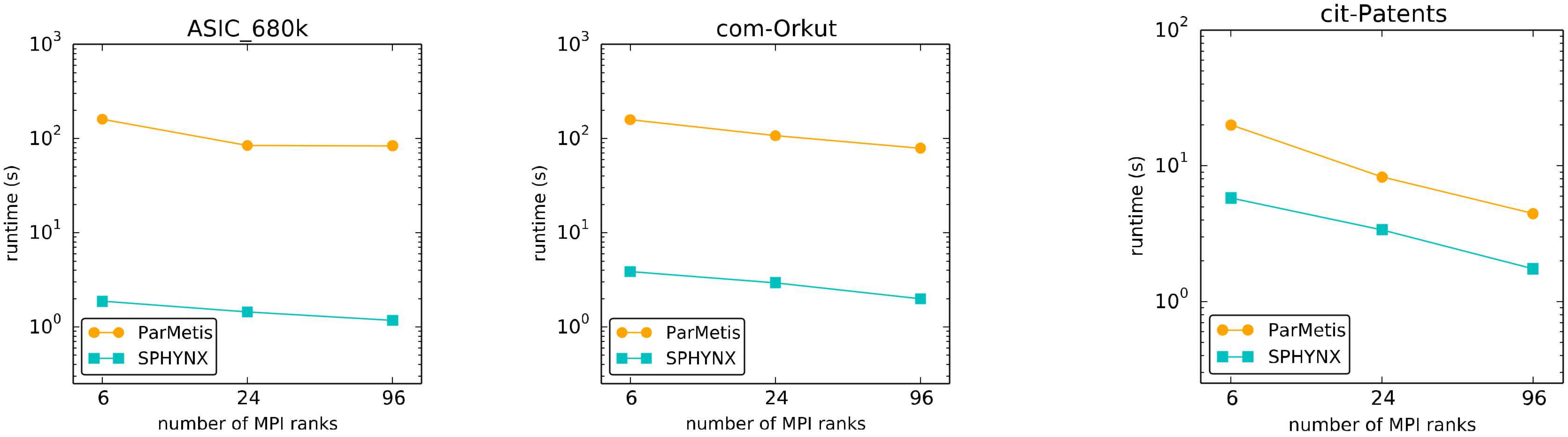
- SPHYNX: New spectral partitioner
  - Based on Trilinos, uses Kokkos for portability (GPU)
  - Uses Anasazi for eigenvector computations
  - Uses Zoltan2 MJ for geometric partitioning
  - Runs on multi-GPU systems



Sphynx: Results

Comparison on 24 GPUs against ParMETIS (24 MPI ranks) on Summit:

- On ASIC_680k, it is 54x faster with 1.66x increase in cutsize
- On com-Orkut, it is 27x faster with 1.95x increase in cutsize
- On cit-Patents, it is 2.3x faster with 3.79x increase in cutsize



EXASCALE COMPUTING PROJECT

Sphynx: Work in Progress

- Integration into Trilinos/Zoltan2
- Improve performance on meshes
  - Need multilevel method or scalable preconditioner
  - Promising results using MueLu in LOBPCG

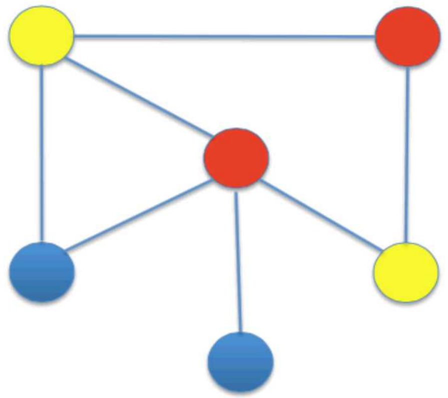
Graph coloring

- Problem:
  - Label vertices in a graph so no two neighbors have the same color.
  - Many variations
- Applications:
  - Find independent tasks for parallel computing
  - Compress sparse matrices
  - Finite differences & AD
  - Independent sets for aggregation in AMG

Graph Coloring: Problem

- Given a graph G = (V , E),
  - With vertices $v \in V$
  - Edges $(v_1, v_2) \in E \quad v_1, v_2 \in V$

Image courtesy of
Sariyuce, Saule, Catalyurek,
SIAM PP, 2012



  - Distance-1 graph coloring: assign colors to vertices so that each vertex have different color from all of its neighbors $C : V \rightarrow N \quad C(v_1) \neq C(v_2) \quad for\ all\ (v_1, v_2) \in E$
- The distinct number of colors assigned to vertices: $|C|$
  - Minimize $|C|$ is NP-Hard, not practical
  - Fast greedy heuristics work well in practice
- Trade-offs:
  - Speed vs quality
  - Deterministic or non-deterministic (parallel)

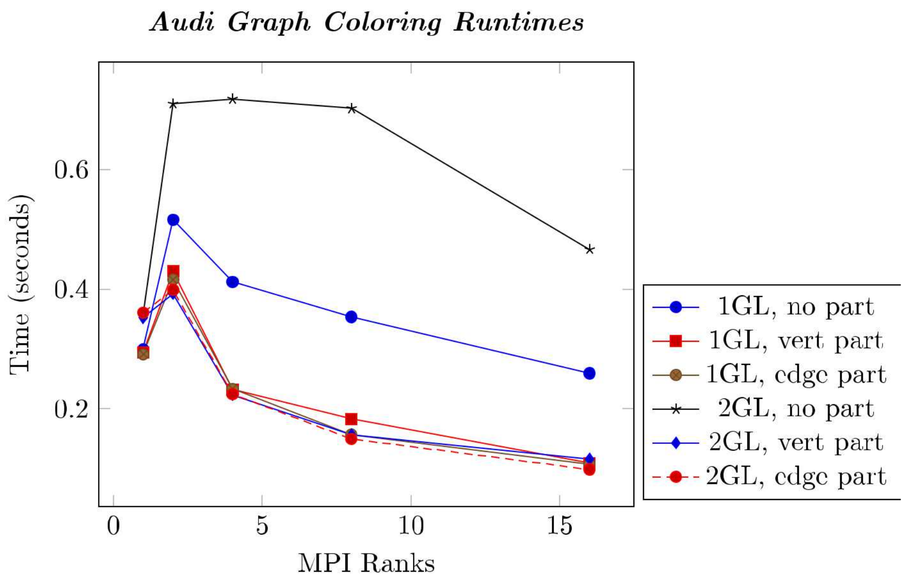
EXASCALE COMPUTING PROJECT

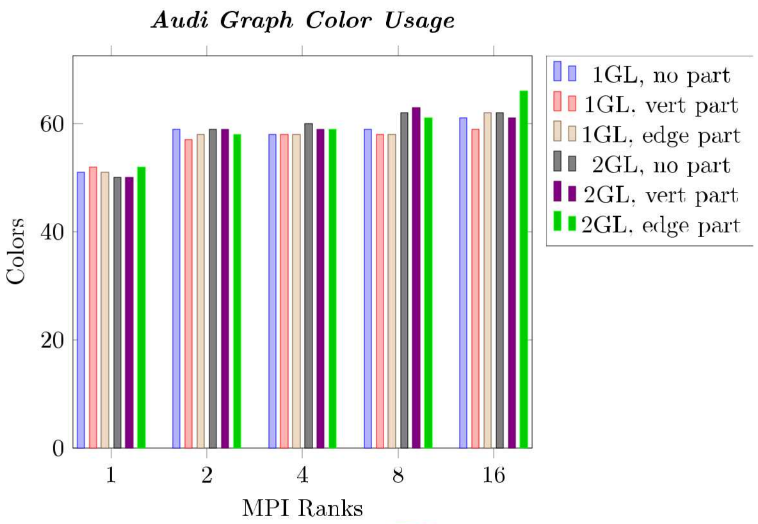
Graph Coloring Software

- ColPack (Pothen et al., Purdue)
  - Serial code (some OpenMP)
  - Wide variety of coloring problems
- Zoltan
  - MPI parallel
  - Dist-1, dist-2, partial dist-2
- KokkosKernels
  - On-node parallel using Kokkos (CPU, GPU)
  - Dist-1 and dist-2
- Zoltan2
  - Basic serial method
  - Coming soon: MPI+Kokkos coloring based on KokkosKernels

ExaGraph Coloring

- Collaboration with RPI (Bogle, Slota)
- Applications: SNL/ATDM (Empire, Sparc)
  - Want coloring of sparse Jacobian matrices (distributed)
  - Must run on both CPU and GPU
- Approach: Use KokkosKernels coloring on-node
  - Speculative method with conflict detection/resolution
  - Works with MPI+Kokkos (multi-GPU systems)
  - Novel communication-avoiding version with multiple ghost layers





EXASCALE COMPUTING PROJECT

Backup

Zoltan2 Overview

- Zoltan2: Trilinos package. Toolkit of combinatorial algorithms for parallel computing on emerging architectures
  - Supersedes the Zoltan package
- Goals:
  - Provide graph algorithms needed on exascale systems
    - Load-balancing and task placement for supercomputers (hierarchical systems)
    - Node-level coloring for multi-threaded parallelism
    - Highly parallel (eventually GPU) partitioning algorithms
  - Support very large application problem sizes
    - Templated data type for global indices (64-bit integers work)
  - Greater integration with Trilinos' next-generation solver stack

Zoltan2 Capabilities: Parallel Partitioning

- Provides common models and uniform interface to collection of partitioners
  - Easy to try & compare different methods
- Multi-Jagged Geometric Partitioning
  - Fast; scalable; enforces geometric locality of data
  - Uses geometric coordinates; generalizes RCB
- Graph partitioning
  - Interfaces to PT-Scotch (INRIA, France), ParMETIS (U. Minnesota), PuLP and XtraPuLP (SNL, Penn St., RPI)
  - Connectivity-based; explicitly models communication costs (not exact)
  - Widely used, proven track record
- Zoltan partitioning algorithms
  - Supports RCB, RIB, HSFC, hypergraph (PHG)

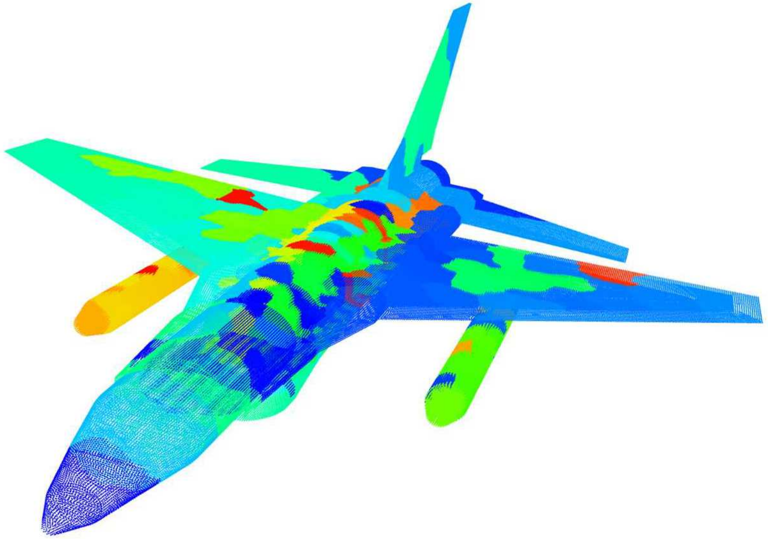


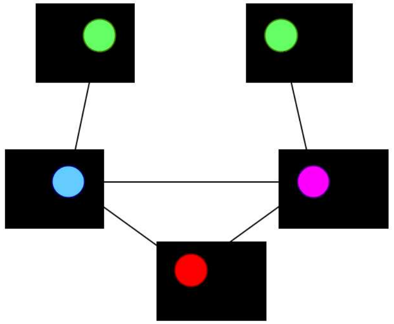
Image courtesy of Manoj Bhardwaj, SNL.

Coloring and Structural Orthogonality

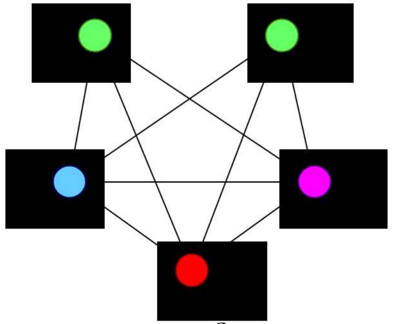
structurally orthogonal partition | distance-2 coloring | distance-1 coloring

symmetric case

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 & a_{15} \\ 0 & a_{22} & a_{23} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & a_{35} \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ a_{51} & 0 & a_{53} & a_{54} & a_{55} \end{bmatrix}$$



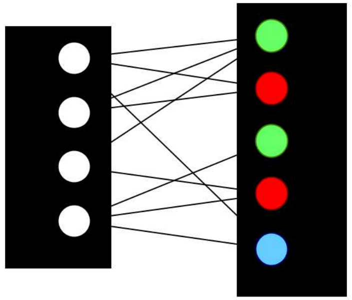
$\mathcal{G}_a$



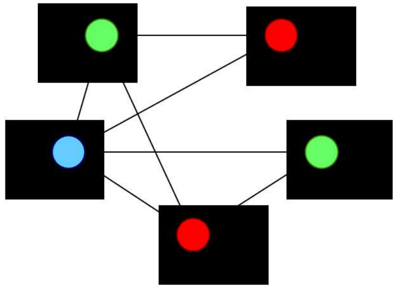
$\mathcal{G}_a^2$

nonsymmetric case

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & a_{15} \\ a_{21} & a_{22} & 0 & 0 & 0 \\ a_{31} & 0 & 0 & a_{34} & 0 \\ 0 & 0 & a_{43} & a_{44} & a_{45} \end{bmatrix}$$



$\mathcal{G}_b$



$\mathcal{G}_b^2[\mathcal{V}_2]$