Abelian: A Compiler and Runtime for Graph Analytics on Distributed, Heterogeneous Platforms

Gurbinder Gill and Keshav Pingali
Department of Computer Science
The University of Texas at Austin

Distributed Heterogeneous Graph Analytics

- **Programming model:**
  - Generalization of vertex programming model
  - Nodes and edges have labels, which are iteratively updated
  - Labels updated by applying `operator` to `active nodes` (activity)
- **Finding active nodes:**
  - Topology driven
  - Data driven:
    - Worklist based
    - Filter based
  - Bulk synchronous execution

Graph Partitioning

- An example of partitioning a graph for 2 hosts:
  - **Communication Patterns:**
    - Reduce: values at mirrors are combined on the master using a reduction operation
    - Broadcast: value at the master is broadcast to the mirrors

Gluon Sync API

- Compiler generated synchronization structures:
  - **Reduce API:**
    - `struct` `reduceField`
    - `static` `ValTy` `extract(NodeData& n)`
  - **Broadcast API:**
    - `struct` `broadcastField`
    - `static` `ValTy` `extract(NodeData& n)`

Graph-Data Access Analysis

- **Type of field access:**
  - Reduction: Read and updated
  - Read-only
  - Write-only
- **Where is field accessed:**
  - At source of edge
  - At destination of edge
  - At any

Restructuring Computation

- **Fine-grain iteration level parallelism to Bulk synchronous parallelism**
- Operator splitting
- Worklist elimination

Inserting Communication

- **Fine-grained communication:**
  - Communicate only updated fields using field specific bitVectors

Distributed Heterogeneous Execution Model

Overview of Abelian System

- **Supports heterogeneity in Programming model:**
  - D-Galois = Galois + Gluon
  - D-Liga = Liga + Gluon
- **Supports heterogeneity in Architecture:**
  - D-IrGL = IrGL + Gluon (GPU)

References


Results

- **Gluon based systems (Hand-Tuned):**
  - *CPU*
  - *GPU*
- **Gluon based systems scale well**
- D-Galois is ~3.9x faster than Gemini
- D-IrGL is ~4.9x faster than Gemini

- **Compiler generated versions:**
  - *CPU*
  - *GPU*

- "23x reduction in communication volume over UO"
- FO matches performance HT (within 12%)
- Increased productivity without performance loss