Technical Overview

Alan G. Labouseur, Ph.D.
Marist College
3NF Consulting, Inc.
alan.labouseur@Marist.edu
alan@3NFconsulting.com
G* is a dynamic graph database system with many features.

- graph distribution
  - multi-core scale up
  - multi-server scale out
- deduplicated disk storage for very large graphs
- in-memory compact indexing
- shared computation
- easy use of sophisticated parallel graph-theoretic queries
- integrates with Relational databases and other stores

Analyzing evolving graphs enables applications in many areas.

- social media analysis
- network traffic threat assessment
- fraud detection
- marketing
- transportation
- epidemiology
- pharmacology
- ...and many other areas
G* studio

Browser Application

G* Studio

Graph Editor

- Edge from 46 to 23 added to graph 2.0.
- Edge from 47 to 23 added to graph 2.0.
- Edge from 48 to 23 added to graph 2.0.
- Edge from 49 to 23 added to graph 2.0.
- Edge from 50 to 23 added to graph 2.0.
- Edge from 51 to 23 added to graph 2.0.
- Edge from 52 to 23 added to graph 2.0.
- Edge from 53 to 23 added to graph 2.0.
- Edge from 54 to 23 added to graph 2.0.
- Edge from 55 to 23 added to graph 2.0.
- Edge from 56 to 23 added to graph 2.0.
- Edge from 57 to 23 added to graph 2.0.
- Edge from 58 to 23 added to graph 2.0.
- Edge from 59 to 23 added to graph 2.0.
- Edge from 60 to 23 added to graph 2.0.
- Edge from 61 to 23 added to graph 2.0.
- Edge from 62 to 23 added to graph 2.0.
- Edge from 63 to 23 added to graph 2.0.
- Checkpoint succeeded.
- Draw
- Drawing complete.
- Query node degree 5
- Drawing complete.
- *

Visualizer

2-hop neighbors of vertex 1: 2, 3, 4, 5, 6, 7
3-hop neighbors of vertex 1: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Top 5 vertices by total degree: 10, 11, 12, 13, 14
3-hop neighbors of vertex 3: 6, 7, 12, 13, 14, 15, 2, 24, 25, 26, 27, 28, 29, 30, 31, 4, 5
3-hop neighbors of vertex 3: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Graph Editor

```
130  -- Evolution: 4 Incremental Graphs (with cloning)
131  create graph 10.0
132  add vertex 1 with attributes (color=black)
133  add vertex 2 with attributes (color=black)
134  add vertex 3 with attributes (color=black)
135  add edge 1-2
136  add edge 2-3
137
138  clone graph 11.0 from 10.0
139  add vertex a with attributes (color=white)
140  add vertex b with attributes (color=white)
141  add vertex c with attributes (color=white)
142  add edge 1-a
143  add edge 1-b
144  add edge 1-c
145
146  clone graph 12.0 from 11.0
147  add vertex d with attributes (color=white)
148  add vertex e with attributes (color=white)
149  add vertex f with attributes (color=white)
150  add edge 2-d
151  add edge 2-e
```

Console

- Evolution
- Common
  - 8-vertex Full
  - 32-vertex Ring
  - 32-vertex Bipartite (16 pairs)
  - 63-vertex Tree (branch factor = 2)
  - 64-vertex Star
  - 64-vertex 72-edge Erdos-Renyi Random
- Other
G* studio

Interactive Console

```
Console

5  Graph 2.0 : 
6   Vertices: 63 
7    Edges : 62 
8  Graph 1.0 : 
9   Vertices: 4 
10  Edges : 2 
11 Graph 0.0 : 
12   Vertices: 2 
13  Edges : 1 
14  G* create graph 4 
15       New graph 4.0 was created. 
16  G* add vertex Kirk 
17  Vertex Kirk added to graph 4.0. 
18  G* add vertex Spock 
19  Vertex Spock added to graph 4.0. 
20  G* add vertex McCoy 
21  Vertex McCoy added to graph 4.0. 
22  G* add edge Kirk-Spock 
23  Edge from Kirk to Spock added to graph 4.0. 
24  G* add edge Kirk-McCoy 
25  Edge from Kirk to McCoy added to graph 4.0. 
26  G* draw 
27  Drawing complete. 
28  G* |
```

Information

Queries

Degree Distribution

Top-k vertices by degree

Top-k vertices with the largest change in degree over consecutive graph snapshot pairs
Making Graphs

- **add graph <graph-id>**
  creates a graph with the given <graph-id>

- **clone graph <graph-id> from <graph-id>**
  creates a new graph as a clone of an existing graph

- **add vertex <vertex-id> [with attributes<attributeName>=<AttributeValue>,...]**
  creates a vertex with id specified by <vertex-id> in the active graph. Can optionally add attributes, with one or more attribute pairs.

- **add edge <from-vertex-id> -- <to-vertex-id>**
  creates an edge from <from-vertex-id> to <to-vertex-id> in the active graph.

- **update <vertex-id> with attributes<attributeName>=<AttributeValue>,...**

Data

3-hop neighbors of vertex 1:
2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Erdős-Rényi random graph

Degree Distribution

---

**Degree Distribution Query**

Dataset: G*studio
Rising/Falling Stars Query

Top 20 vertices with the largest change in degree over consecutive graph snapshot pairs from 6 to 8:

<table>
<thead>
<tr>
<th>snapshotPairs</th>
<th>vertexID</th>
<th>change</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-&gt;6</td>
<td>1</td>
<td>+3</td>
</tr>
<tr>
<td>6-&gt;7</td>
<td>2</td>
<td>+5</td>
</tr>
<tr>
<td>7-&gt;8</td>
<td>3</td>
<td>+3</td>
</tr>
<tr>
<td>5-&gt;6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5-&gt;6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>6-&gt;7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6-&gt;7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>6-&gt;7</td>
<td>a</td>
<td>0</td>
</tr>
<tr>
<td>7-&gt;8</td>
<td>2</td>
<td>-2</td>
</tr>
</tbody>
</table>
Distributed Architecture

The Dynamic Graph Database

G* master

Query Parser -> Query Optimizer -> Query Coordinator

Communication Layer

Communication Layer

Query Execution Engine
Graph Manager
Index  Memory Buffer
Disk

HA

β

α
Snapshot Management

$G_1$ snapshot

$G_2$ snapshot

$G_3$ snapshot

Time passes

The Dynamic Graph Database
Deduplicated Graph Distribution

The Dynamic Graph Database
Implementation Overview

- **Browser**
  - G* Graph Editor
  - Console
  - Visualizer

- **API**
  - Graph Manager
  - Memory Buffer
  - Disk

- **Servers**
  - Query Parser
  - Query Optimizer
  - Query Coordinator

- **Communication Layer**
  - REST
  - JSON

- **Java**
  - NanoHTTPD
  - custom RMI, BSP, and messaging

- **JavaScript**
  - Bootstrap
  - jQuery
  - D3

- **Implementation**
  - Overview

Languages:
- Java
- JavaScript
We can build other applications on top of the G* API.
and IBM’s Knowledge Anyhow Workbench

---

**G* Playpen**

A simple exploratory notebook that demonstrates the use of the [G* API on AWS](https://www.amazonaws.com) within a Jupyter based interactive notebook.

```python
import pandas as pd

try it

Query a list of all graphs from the database.

```

```python
In [5]: url = "http://ec2-54-172-65-147.compute-1.amazonaws.com:8080/graphs"
def = pd.read_json(url)
df
```

```
Out[5]:
<table>
<thead>
<tr>
<th>edges</th>
<th>graph</th>
<th>vertices</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>83</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
```

---

**REST** ↓ **JSON**

16
G* REST API Server version gs.0.63

Copyright © 2014-2015 Alan G. Laboureur. All Rights Reserved.

Home

Method = GET

URL / API command = /home

API help

API commands: GET [action], POST [action], PUT [action], DELETE [action]

--- GET /version
--- GET /time
--- GET /workers
--- GET /graphs
+ GET /graphs/[gid]
- GET /graphs/[gid]/vertices
+ GET /graphs/[gid]/edges
- GET /graphs/[gid]/summary
--- GET /currentgraph
--- GET /shutdown
--- GET /checkpoint
--- GET /query/gid/degree/distribution/[gid]
--- GET /query/gid/topkdegrees/[gid]/[k]
--- GET /query/gid/topkdegreechangesbydelta/[start gid]/[stop gid]/[k]
--- POST /graphs/[new-gid]
--- POST /graphs/[new-gid]/[old-gid]
--- POST /graphs/[gid]/vertices/[vid]
--- POST /graphs/[gid]/edges/[vidFrom]/[vidTo]
--- POST /currentgraph/[gid]
--- POST /graphs/[gid]/vertices/[vid]/[attr_type]/attr_name/attr_value
--- POST /graphs/[gid]/edges/[vidFrom]/[vidTo]/[attr_type]/attr_name/attr_value

API Testers

GET /  
POST /
G* integrates with Relational databases. Supports SQL to Graph and Graph to SQL transformations.
More integrations coming soon.


G*studio
and
G*The Dynamic Graph Database

Contact

Alan G. Labouseur, Ph.D.

phone   1-845-440-1102
web     www.Labouseur.com
        www.3NFconsulting.com
        www.marist.edu/compscimath/faculty.html

e-mail  alan@Labouseur.com
        alan@3NFconsulting.com
        alan.labouseur@Marist.edu