



NAIAD

K.M.D.M Karunaratna

University Of Cambridge – 1st Dec 2015

Significance

- New Computation Model, differential dataflow
- Operating over collections of differences rather than collections of records
- Resulting in very efficient implementations of programming patterns that are expensive in existing systems

Similar Projects

Batch
Processing

Spark

Dryad

Hadoop

Graph
Processing

Giraf

Graphlab

Stream
Processing

Storm

S4

Actually Naiad...

- Bring together - **Batch computation, Streaming computation, and Graph computation** into a **common platform**, while retaining the **performance** of specialized systems.
- To that end, Naiad provides libraries for,
Streaming analytics
 - I. **Graph computation**
 - II. **Machine learning**built using low-level primitives for data flow construction and execution.
- The result is a system that supports iterative, incremental, and interactive data analysis; and combinations of all three.

Naiad Demo

- Word Count
- Throughput
- Key Value Lookup
- Latency
- Connected Components

Word Count

```
C:\Users\MUTHU\Desktop\lectures uk\R212\Naiad\Backups\2\Naiad-release_0.5\
Examples\bin\Debug>Examples.exe wordcount
Logging initialized to console
Start entering lines of text. An empty line will exit the program.
Naiad will display counts (and changes in counts) of words you type.
Hello World
[Hello 1]
[World 1]
Hello Muthu
[Hello 2]
[Muthu 1]
Muthu
[Muthu 2]
[ 1]
Hello
[Hello 3]
[ 2]
```

Lookup

```
C:\Users\MUTHU\Desktop\lectures uk\R212\Naiad\Backups\2\Naiad-release_0.5\
Examples\bin\Debug>Examples.exe lookup
Logging initialized to console
Enter two strings to insert/overwrite a (key, value) pairs.
Enter one string to look up a key.
R Red
B Blue
G Green
R
value["R"]:      "Red"
G
value["G"]:      "Green"
```

Connected Components

```
C:\Users\MUTHU\Desktop\lectures uk\R212\Naiad\Backups\2\Naiad-release_0.5\  
Examples\bin\Debug>Examples.exe connectedcomponents  
Logging initialized to console  
Computing components of a random graph on 1000 nodes and 2000 edges  
labeled 982 nodes in 00:00:00.3472100
```

Connected Components

```
C:\Users\MUTHU\Desktop\lectures uk\R212\Naiad\Backups\2\Naiad-release_0.5\
Examples\bin\Debug>Examples.exe dd-connectedcomponents
Logging initialized to console
Running connected components on a random graph (1000000 nodes, 750000 edge
s)
For each size, the number of components of that size (may take a moment):
Time to process: 00:00:42.1775505
[ 2, +37779 ]
[ 3, +12352 ]
[ 4, +5455 ]
[ 5, +2898 ]
[ 6, +1655 ]
[ 7, +1062 ]
[ 8, +687 ]
[ 9, +450 ]
[ 10, +324 ]
[ 11, +213 ]
[ 12, +184 ]
[ 13, +128 ]
[ 14, +112 ]
[ 15, +86 ]
[ 16, +57 ]
[ 17, +43 ]
[ 18, +32 ]
[ 19, +25 ]
```

What I am Going to do...

- A Naiad Graphical visualization in Visual Studio 2013
 - I. Search Index
 - II. Word Count
 - III. Key Value Lookup
 - IV. Throughput
 - V. Latency
 - VI. Connected Components
- Comparison Naiad performance with default .Net methods
- Simulating Naiad in distributed Environment

Suggestions?

Thank You

Appendix

Naiad Version History

- **Version 1** - Serial, multicore, and distributed implementations of differential dataflow.

Example programs that illustrate how to use differential dataflow for a variety of graph and non-graph computations.

- **Version 2** - The license was changed to the Apache 2.0 open-source license.

The Naiad runtime was split into the Naiad.dll runtime, and multiple frameworks. This release contained the Lindi (LINQ with declarative iteration) and Differential Dataflow frameworks.

- **Version 3** - Added the Cluster Submission solution, which contains support for launching Naiad programs in a variety of distributed settings, including: Microsoft Azure HDInsight 3.0/Hadoop 2.0 (YARN) on Windows.

The Azure Support framework makes it easier to use Naiad programs with Microsoft Azure Storage Blobs and Tables.

Naiad Version History

- **Version 4** - Comprehensive documentation of every public class, method, and property.

The Graph LINQ framework contains data types and extension methods that add optimized graph-specific operators for streams.

- **Version 5** - Frameworks for reading and writing data in HDFS

Thank you...