

Perfopticon: Visual Query Analysis for Distributed Databases

Dominik Moritz, Daniel Halperin, Bill Howe, and Jeffrey Heer
Computer Science & Engineering, University of Washington

CPSC 547
Thursday, November 12
By: Dmitry Tebaykin

Overview

1. Introduction into SQL and databases
2. Why is this paper important?
3. The 4 views of Perfopticon (with analysis and pictures)
4. Could you use Perfopticon?
5. Conclusions

1. Introduction into SQL and databases

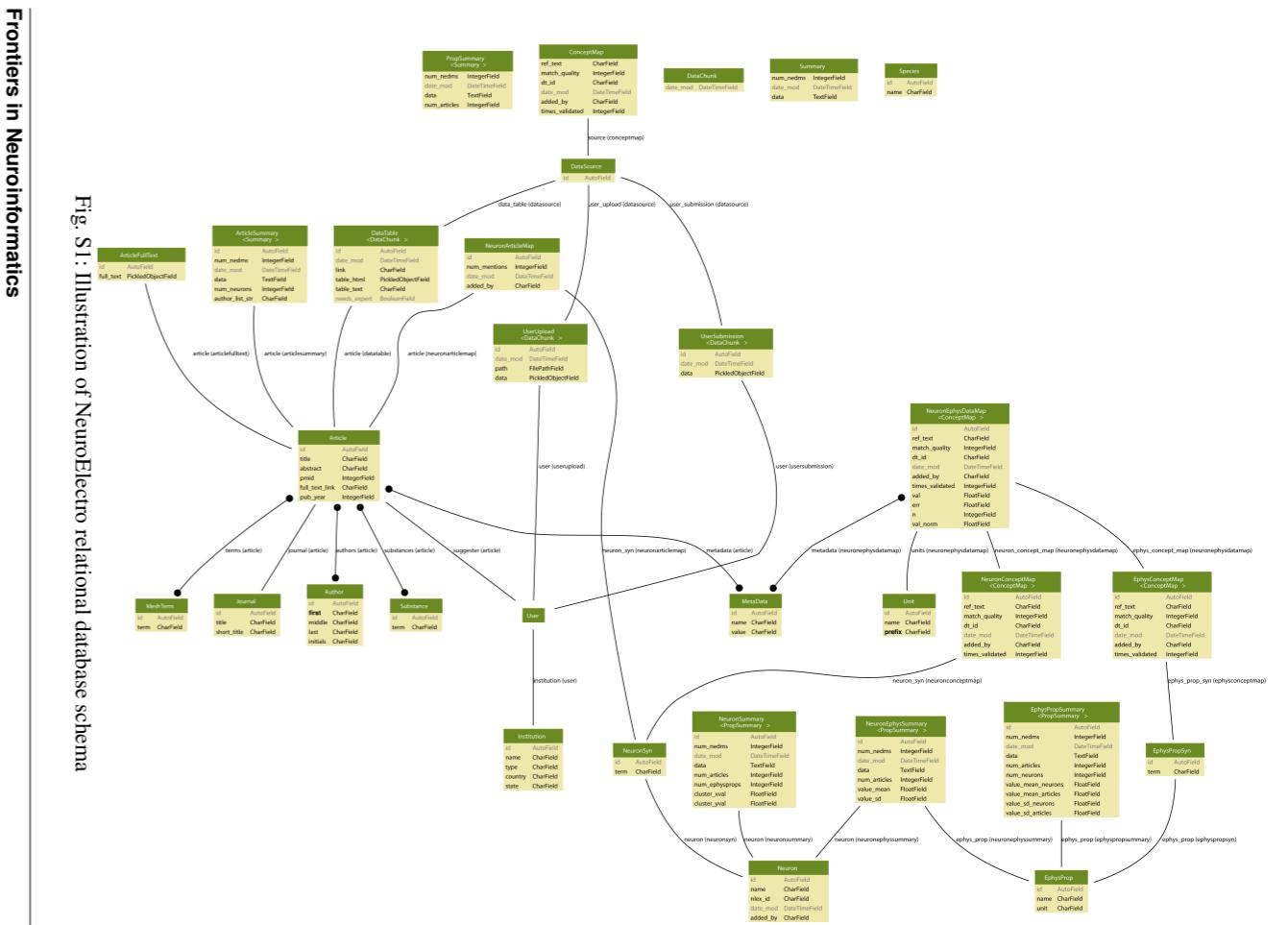
In our case:

Database - tables of data joined

SQL - language for talking to databases

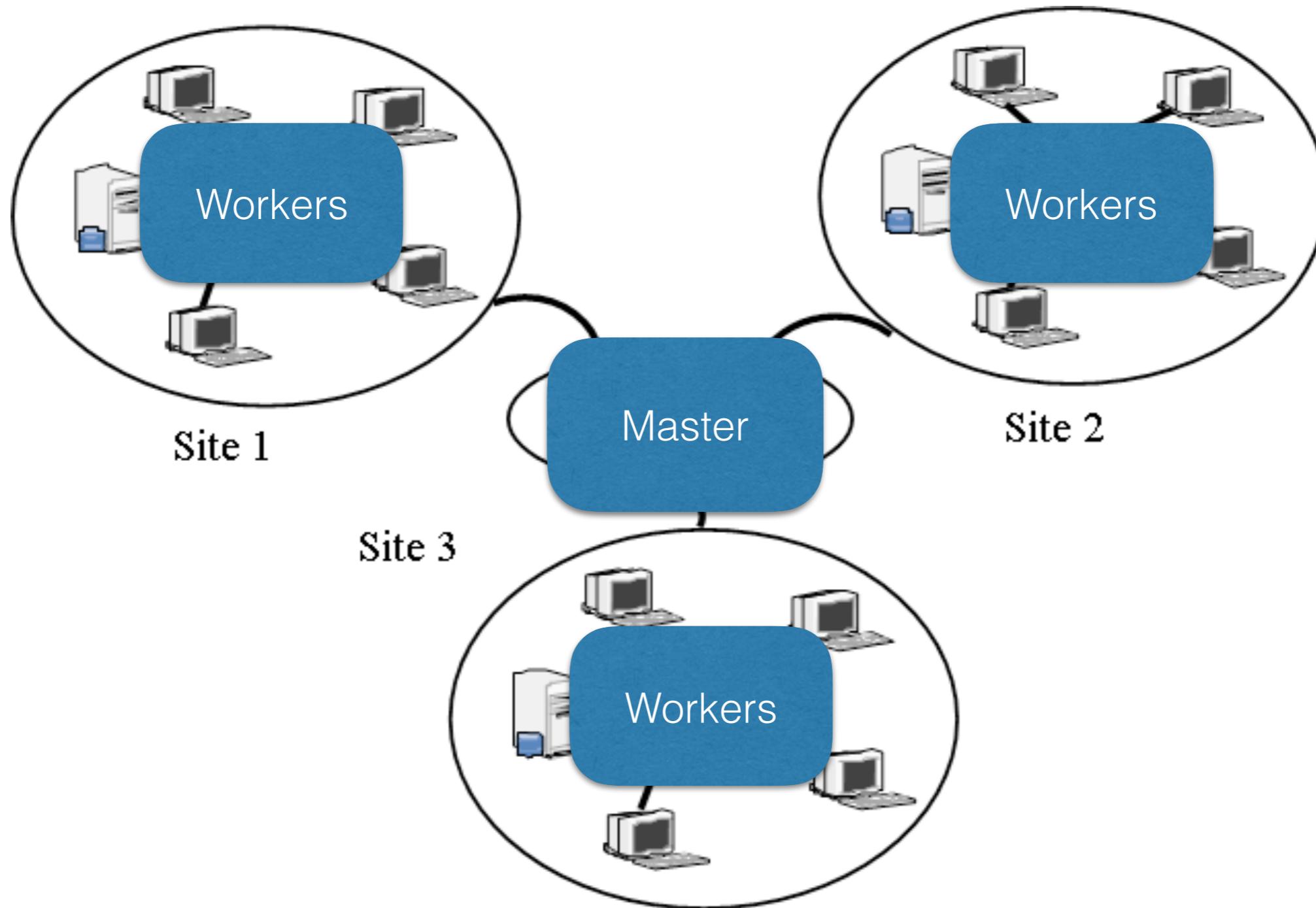
Examples of questions:

- “What is the age of every student in UBC?”
- “How many people are taking CS547 this term?”



1. Introduction into SQL and databases

Distributed database system:



<https://cnx.org/resources/0d203a416b87d2bed544825664c14614602f9385/graphics8.png>

2. Why is this paper important?

Query execution log files

```

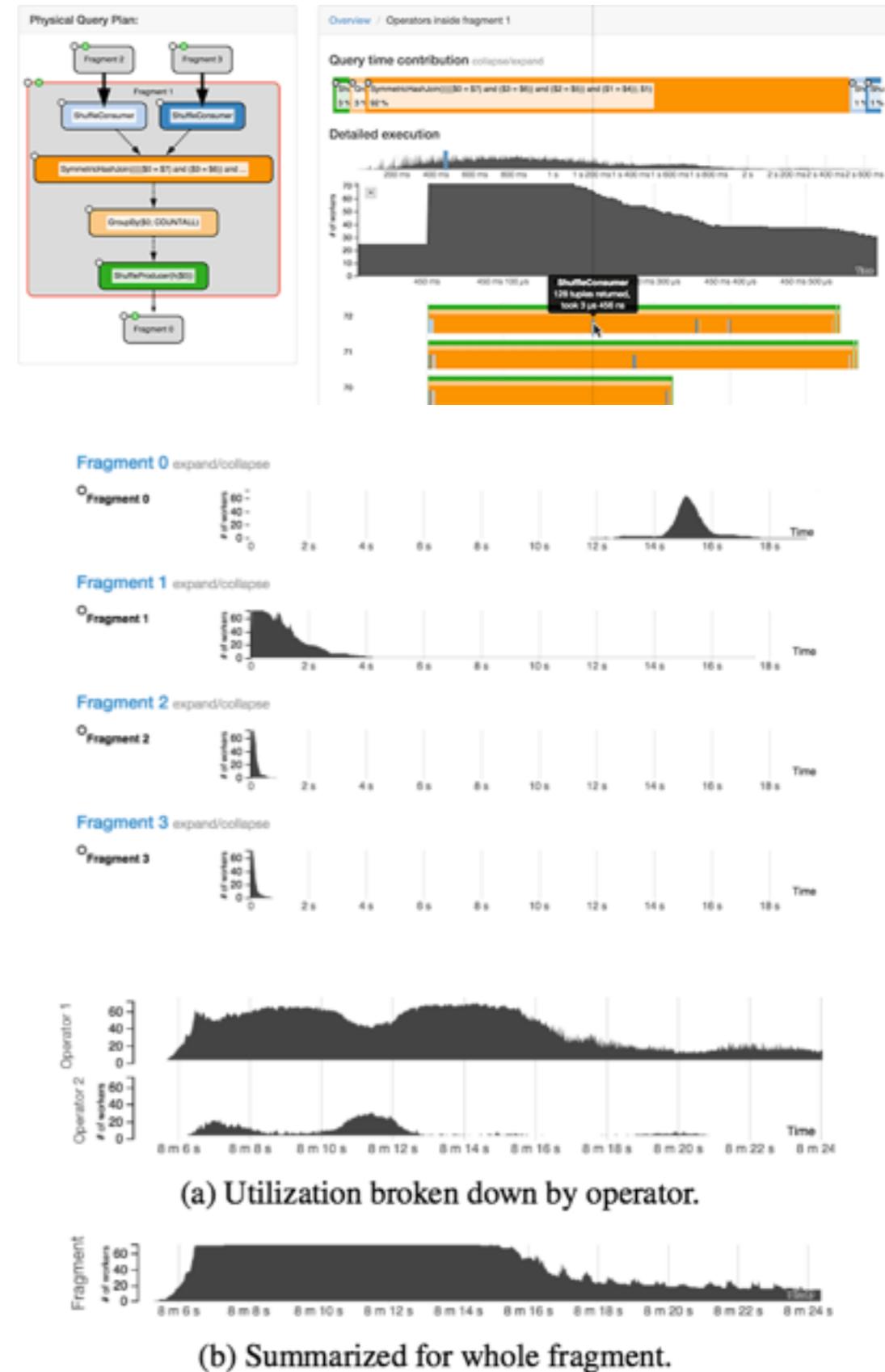
1/28/2009 1:25:00... spid101 Using Spid101.dll version 2009.100.1098 to execute extended stored procedure xp_analyzeerror
1/28/2009 1:24:53... spid254 Recovery is complete. This is an informational message only. No user action is required.
1/28/2009 1:24:53... Server SQL Server is now ready for client connections. This is an informational message; no user action is required.
1/28/2009 1:24:53... Server The SQL Server Network Interface library could not register the Service Principal Name (SPN).
1/28/2009 1:24:53... Server Dedicated admin connection support was established for listening locally on port 1434.
1/28/2009 1:24:53... Server Server is listening on [ ::1 ].1434 (tcpv4: 1434).
1/28/2009 1:24:53... Server SQL Server Network Interfaces initialized listeners on node 1 of a multi-node (NUMA) server core.
1/28/2009 1:24:53... Server Server is listening on [ 'any' ] (tcpv4: 2001).
1/28/2009 1:24:53... Server
ReportServerService_08_24_2014_21_34_30.log - Notepad

```

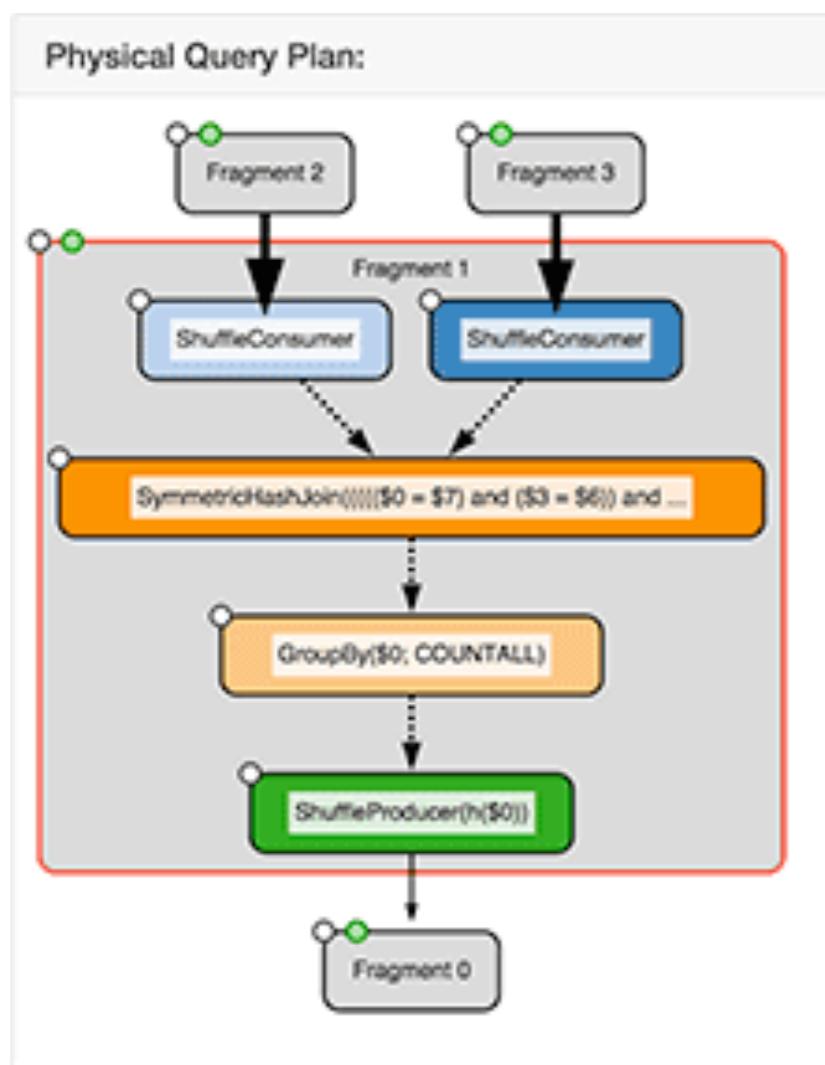
```

1/28/2009 1:24:53... Server Se
1/28/2009 1:24:53... Server S2 File Edit Format View Help
1/28/2009 1:24:53... Server S2 <Header>
1/28/2009 1:24:53... Server S2 <Product>Microsoft SQL Server Reporting Services Version 2011.0110.5058.00
1/28/2009 1:24:53... Server S2 ((SQL11_PCU_Main).140514-1820 )</Product>
1/28/2009 1:24:53... Server S2 <Locale>English (United States)</Locale>
1/28/2009 1:24:53... Server S2 <TimeZone>Eastern Daylight Time</TimeZone>
1/28/2009 1:24:53... Server S2 <Path>C:\Program Files\Microsoft SQL Server\MSRS11.SQL2012\Reporting Services
1/28/2009 1:24:53... Server S2 <LogFile>ReportServerService_08_24_2014_21_34_30.log</Path>
1/28/2009 1:24:53... Server S2 <SystemName>... ... ?C</SystemName>
1/28/2009 1:24:53... Server S2 <OSName>Microsoft Windows NT 6.2.9200</OSName>
1/28/2009 1:24:53... Server S2 <OSVersion>6.2.9200</OSVersion>
1/28/2009 1:24:53... Server <ProcessID>27800</ProcessID>
1/28/2009 1:24:53... Server <Virtualization>None</Virtualization>
1/28/2009 1:24:53... Server </Header>
1/28/2009 1:24:53... Server <ProcessorArchitecture>AMD64</ProcessorArchitecture>
1/28/2009 1:24:53... Server <ApplicationArchitecture>AMD64</ApplicationArchitecture>
1/28/2009 1:24:53... Server rsHost!rsHost!1234!08/24/2014-21:34:38:: 1 INFO: CLR runtime is initialized.
1/28/2009 1:24:53... Server rsHost!rsHost!1234!08/24/2014-21:34:38:: 1 INFO: Derived memory configuration based on
1/28/2009 1:24:53... Server physical memory as 5872500 KB
1/28/2009 1:24:53... Server appdomainmanager!DefaultDomain!1234!08/24/2014-21:34:38:: 1 INFO: Entered managed
1/28/2009 1:24:53... Server ServiceMain in DefaultDomain.
1/28/2009 1:24:53... Server library!DefaultDomain!102c!08/24/2014-21:34:36:: 1 INFO: Initializing ConnectionType
1/28/2009 1:24:53... Server to '0' as specified in Configuration file.
1/28/2009 1:24:53... Server library!DefaultDomain!102c!08/24/2014-21:34:36:: 1 INFO: Initializing
1/28/2009 1:24:53... Server SecureConnectionLevel to '0' as specified in Configuration file.
1/28/2009 1:24:53... Server library!DefaultDomain!102c!08/24/2014-21:34:36:: 1 INFO: Initializing
1/28/2009 1:24:53... Server DisableSecureFormsAuthenticationCookie to 'False' as specified in Configuration file.
1/28/2009 1:24:53... Server library!DefaultDomain!102c!08/24/2014-21:34:36:: 1 INFO: Initializing
1/28/2009 1:24:53... Server CleanupCycleMinutes to '10' minute(s) as specified in Configuration file.
1/28/2009 1:24:53... Server library!DefaultDomain!102c!08/24/2014-21:34:36:: 1 INFO: Initializing
1/28/2009 1:24:53... Server MaxActiveReqForOneUser to '20' request(s) as specified in Configuration file.

```



3. The 4 views of Perfopticon (with analysis and pictures)



View 1

Query plan view

What: data

Directed graph that represents:
query plan for data access
generated by DBMS

Why: tasks

Locate, identify, compare

How: encode

Shape marks for nodes
(execution steps), connection
marks for links

How: facet

Coordinate: linked highlighting
and navigation with other views

3. The 4 views of Perfopticon (with analysis and pictures)

	View 2	Work distribution view
Fragment 0 expand/collapse ○ Fragment 0	What: data	Tables from query log files
Fragment 1 expand/collapse ○ Fragment 1	Why: tasks	Compare, identify outliers
Fragment 2 expand/collapse ○ Fragment 2	How: encode	Histograms showing execution time of workers
Fragment 3 expand/collapse ○ Fragment 3	How: facet	Partition: multiple views for each query fragment. Coordinate: linked highlighting and navigation with other views
	How: reduce	Navigate

3. The 4 views of Perfopticon (with analysis and pictures)

View 3

Communication view

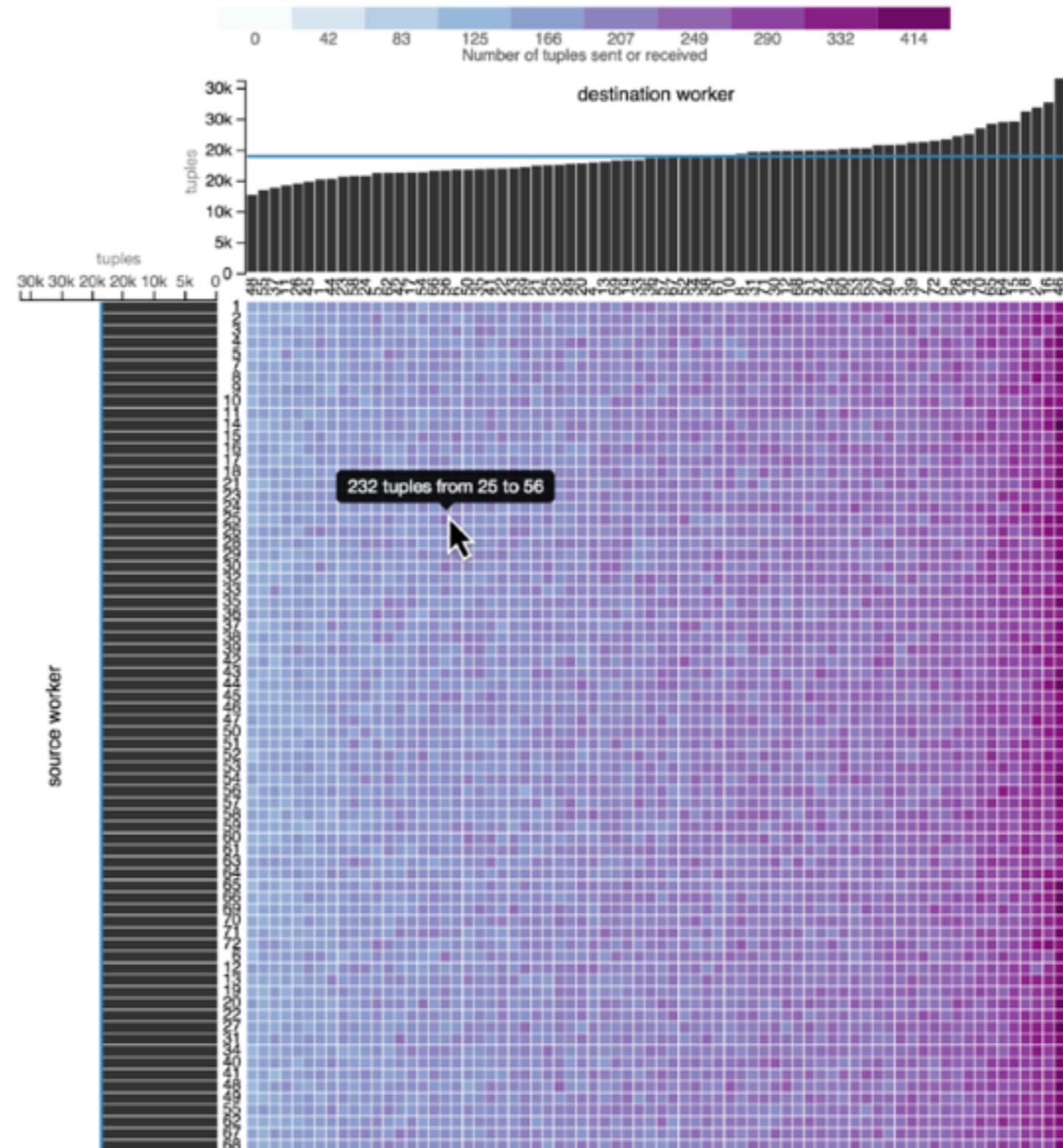
Table: two continuous variables (amount of data sent and received by workers)

What: data

Why: tasks

**How:
encode**

How: facet



3. The 4 views of Perfopticon (with analysis and pictures)

View 4	Local execution view
What: data	Tables from query log files
Why: tasks	Compare, identify outliers
How: encode	Histograms, bar charts (colour indicates active/inactive/wait states)
How: facet	Partition: multiform views. Coordinate: linked highlighting
How: reduce	Navigation



4. Could you use Perfopticon?

- Built into Myria (Giant online database), requires log files for the query executions with slight modifications.
- Their example: Myria, added 3 lines to log file per query execution step.
- The tool has a front-end component, upload your query log files and view the results.

5. Conclusions

- Perfopticon can be used effectively for query and database optimization (Emma, the oceanographer, managed to speed up her query and Chu S. et. al created a better table joining algorithm).
- Provides the ability to spot underperforming or overtasked nodes and drill down into the problem.
- Might work for non-relational databases as well.
- Needs more validation.