Qonnections

TE38 QlikView Performance

9 April, 2014

Lars Skage, Magnus Berg, Jens Argentzell

Who are we?

Lars Skage
- Services Scalability Manager
- Lund, Sweden

Magnus Berg
- Master Principal Enterprise Architect
- Solna, Sweden

Jens Argentzell
- Manager, Scalability Center
- Lund, Sweden

Legal disclaimer

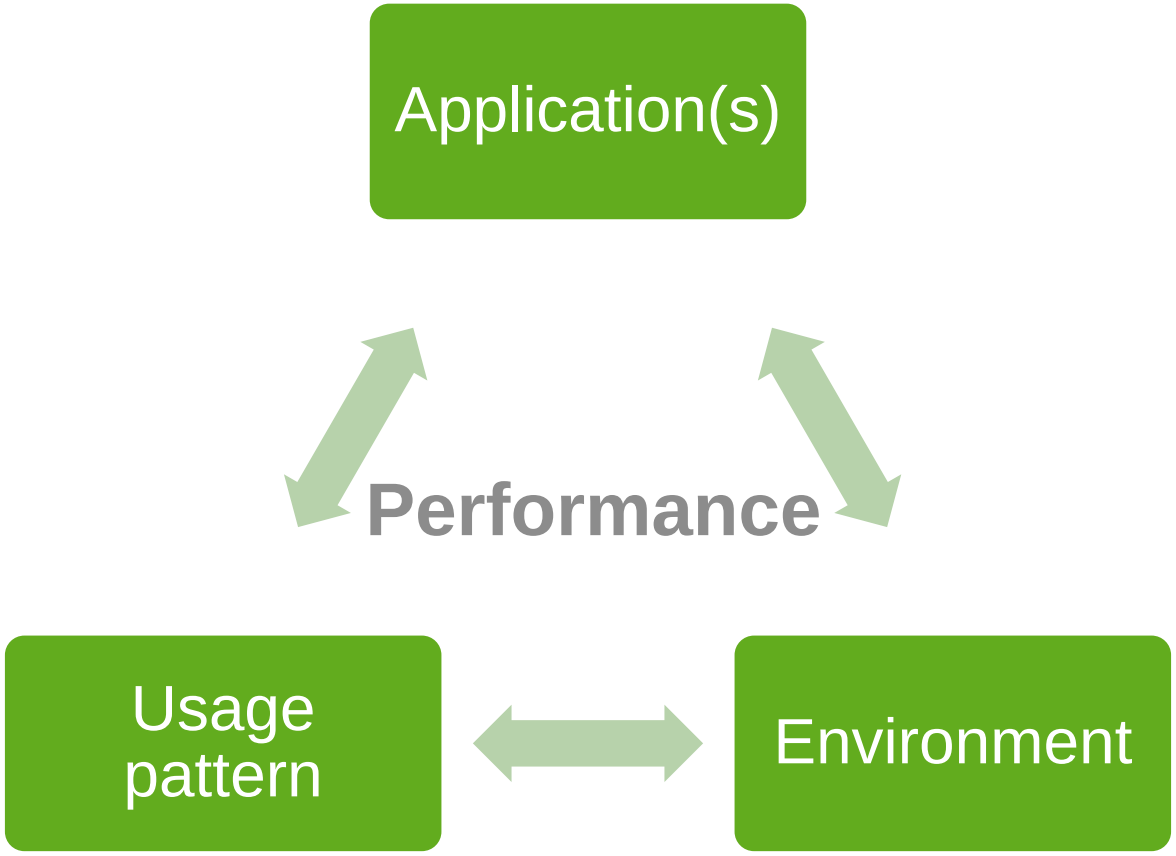
This Presentation contains forward-looking statements, including, but not limited to, statements regarding the value and effectiveness of Qlik's products, the introduction of product enhancements or additional products, Qlik's partner and customer relationships, and Qlik's growth, expansion and market leadership, that involve risks, uncertainties, assumptions and other factors which, if they do not materialize or prove correct, could cause Qlik's results to differ materially from those expressed or implied by such forward-looking statements. All statements, other than statements of historical fact, are statements that could be deemed forward-looking statements, including statements containing the words "predicts," "plan," "expects," "anticipates," "believes," "goal," "target," "estimate," "potential," "may", "will," "might," "could," and similar words. Qlik intends all such forward-looking statements to be covered by the safe harbor provisions for forward-looking statements contained in Section 21E of the Exchange Act and the Private Securities Litigation Reform Act of 1995. Actual results may differ materially from those projected in such statements due to various factors, including but not limited to: risks and uncertainties inherent in our business; our ability to attract new customers and retain existing customers; our ability to effectively sell, service and support our products; our ability to manage our international operations; our ability to compete effectively; our ability to develop and introduce new products and add-ons or enhancements to existing products; our ability to continue to promote and maintain our brand in a cost-effective manner; our ability to manage growth; our ability to attract and retain key personnel; the scope and validity of intellectual property rights applicable to our products; adverse economic conditions in general and adverse economic conditions specifically affecting the markets in which we operate; and other risks and uncertainties more fully described in Qlik's publicly available filings with the Securities and Exchange Commission. Past performance is not necessarily indicative of future results. The forward-looking statements included in this presentation represent Qlik's views as of the date of this presentation. Qlik anticipates that subsequent events and developments will cause its views to change. Qlik undertakes no intention or obligation to update or revise any forward-looking statements, whether as a result of new information, future events or otherwise. These forward-looking statements should not be relied upon as representing Qlik's views as of any date subsequent to the date of this presentation.

This Presentation should be read in conjunction with Qlik's periodic reports filed with the SEC (SEC Information), including the disclosures therein of certain factors which may affect Qlik's future performance. Individual statements appearing in this Presentation are intended to be read in conjunction with and in the context of the complete SEC Information documents in which they appear, rather than as stand-alone statements. This presentation is intended to outline our general product direction and should not be relied on in making a purchase decision, as the development, release, and timing of any features or functionality described for our products remains at our sole discretion.

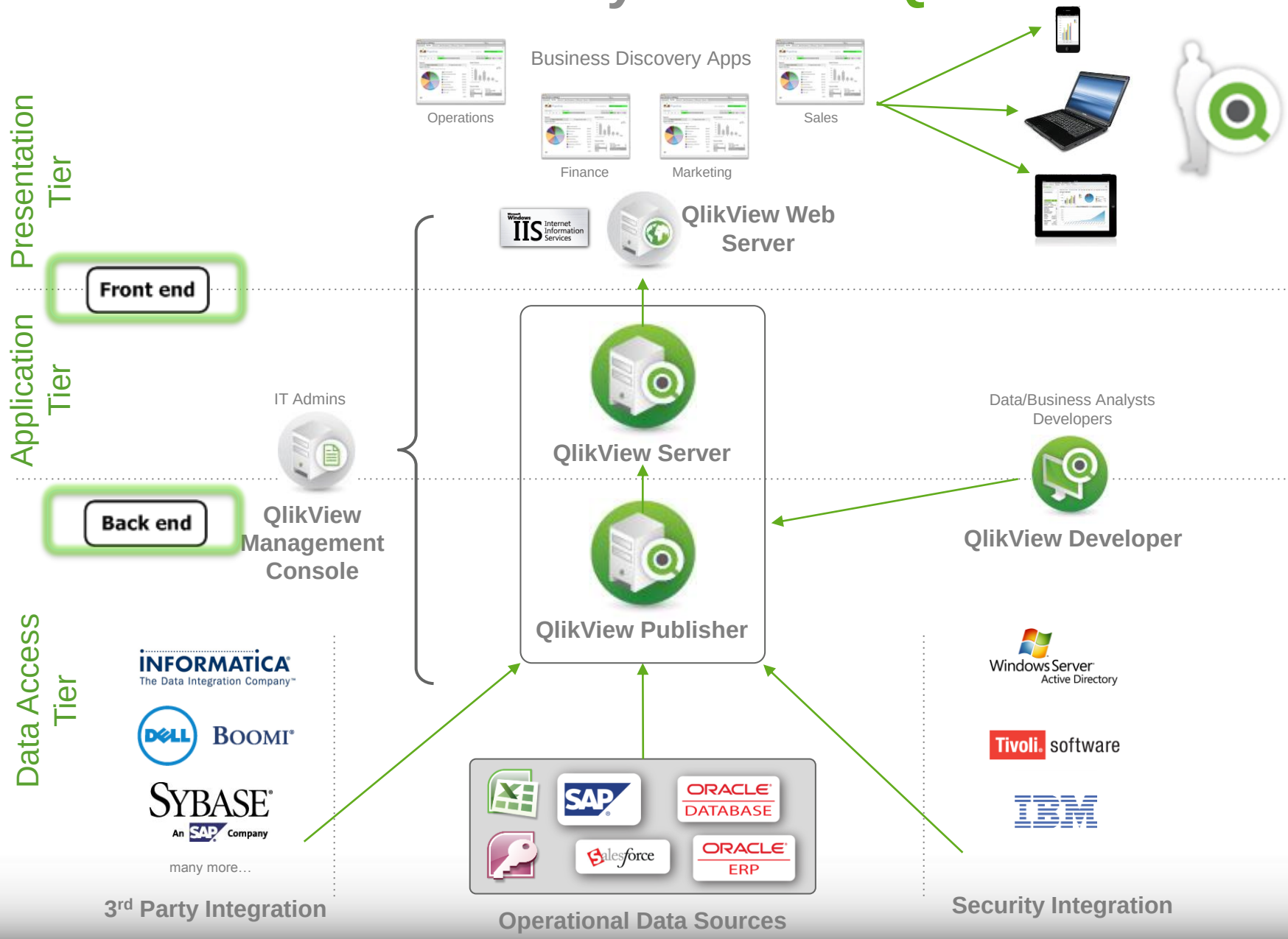
Agenda

- Introduction
  - Performance triangle
- Reload Performance
  - Results
  - Key Takeaways - Selecting Hardware
- QlikView Server performance
  - How QlikView Server utilizes Hardware and proof points
  - Selecting Hardware

On QlikView Scalability



Business Discovery Platform QlikView 11



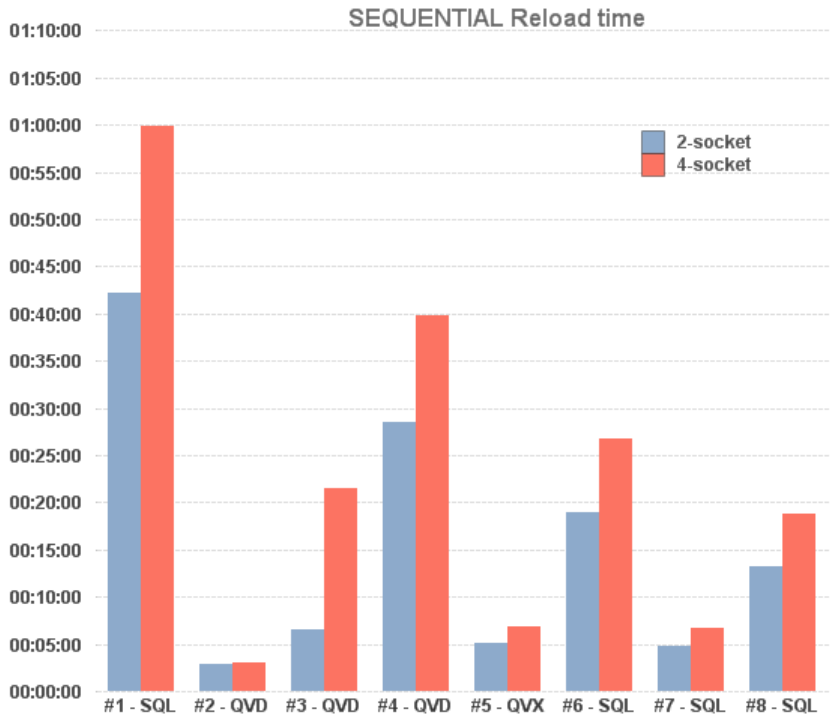
Reload Performance

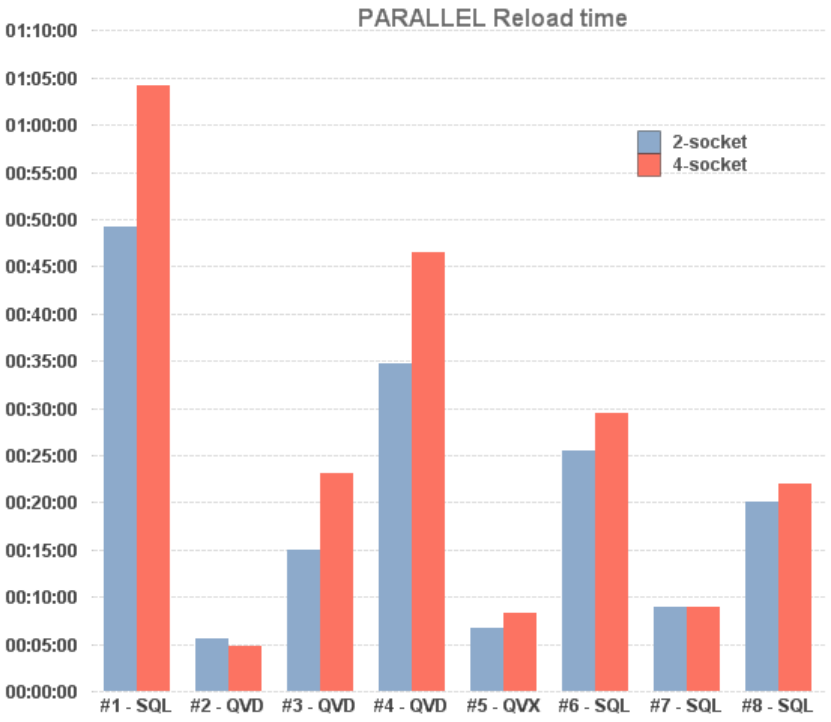
Reload Performance

- We will present several case studies focusing on:
  - Environment
    - Reloads and Architectures
    - Reloads and Clock Frequency
    - Reloads and Disk Speed
  - Application/Script
    - Reloads and Data Sources (QVD, SQL)
    - Reloads - Scaling With Data
    - Database Aggregations
  - Usage pattern
    - Reloads - Multiple Reloads

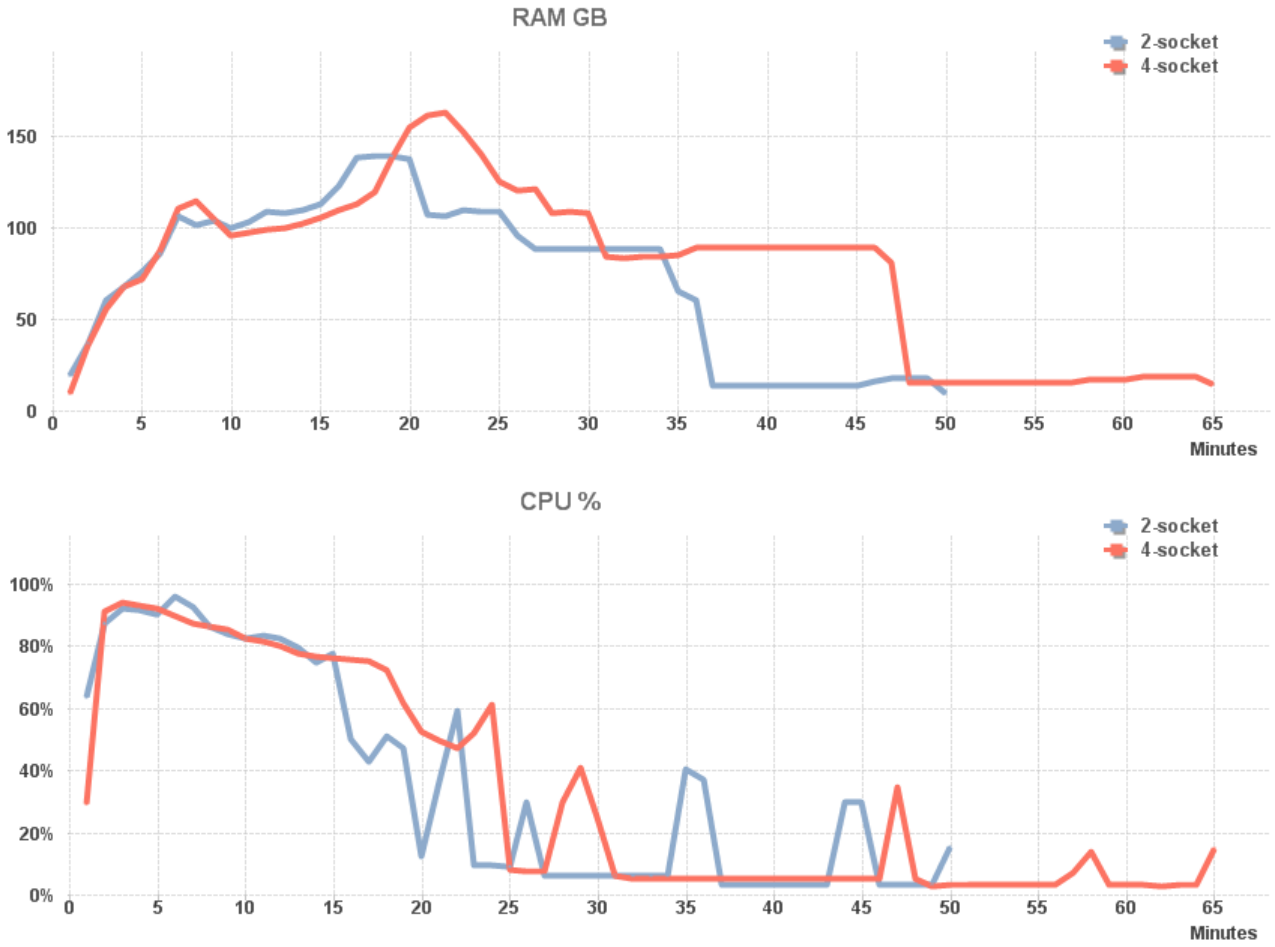
Reloads and Architectures

Tests with different sources, tasks and content performed on two white listed architectures in sequence and in parallel (8 QVBs).





Reloads and Architectures



Key takeaway:

2 socket E5 solutions fit QlikView reload workloads better than 4 socket E7 solutions.

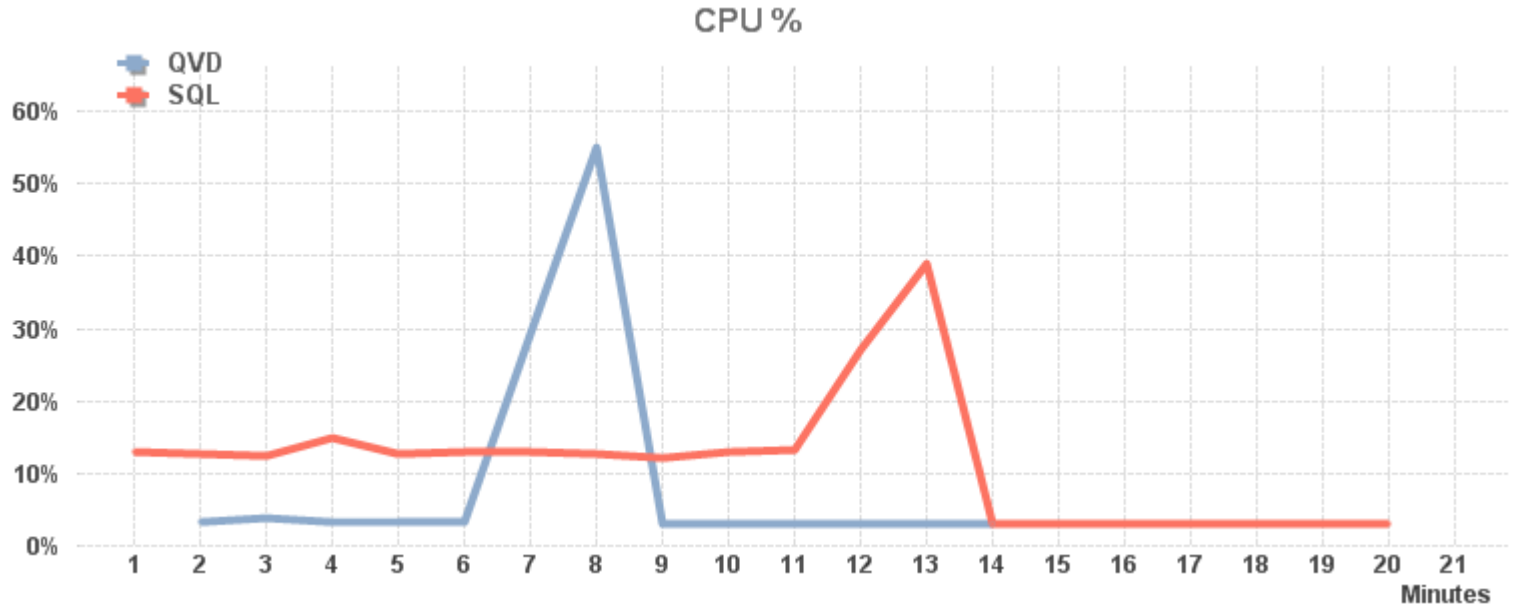
Reloads and Clock Frequency

Difference	
Difference in capacity	+11.5 %
QVD	+11.5 %
QVX	+13.6 %
SQL	+ 0 %

Key takeaway:

Clock frequency is one of the main factors to consider for reloads

Reloads and Data Sources (QVD and SQL)



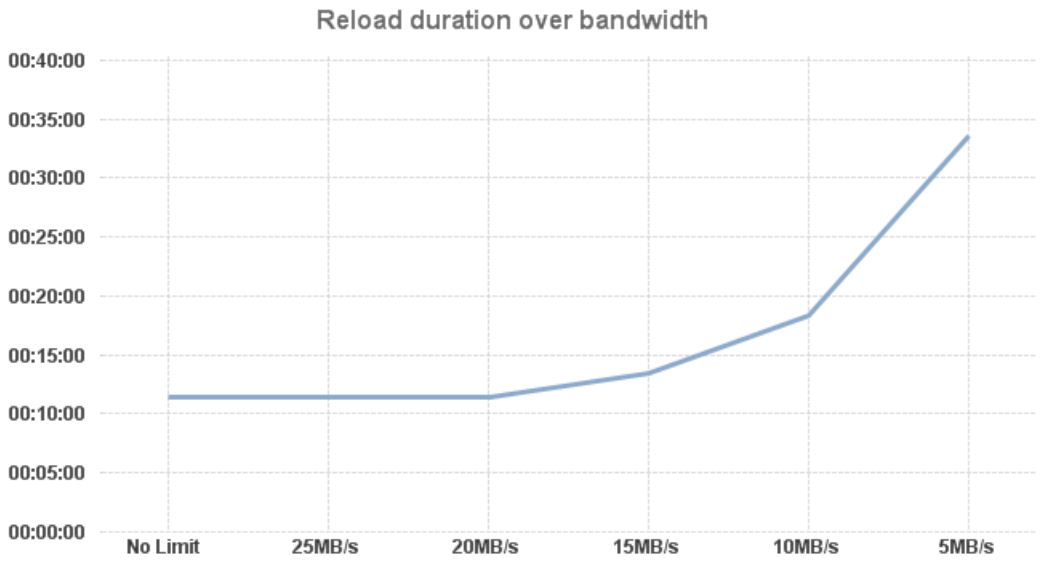
Source	Time Reload (mm:ss)	Time Save (mm:ss)	Total Time (mm:ss)
QVD	5:37	8:05	14:39
SQL	10:56	8:05	19:46

Key takeaway:

QVD is more resource friendly when loaded optimized. SQL loading data continuously requires computational power to transform and validate data.

Reloads From Remote Sources

Source	Reload Time
SQL non cached	25:40
SQL cached	19:55

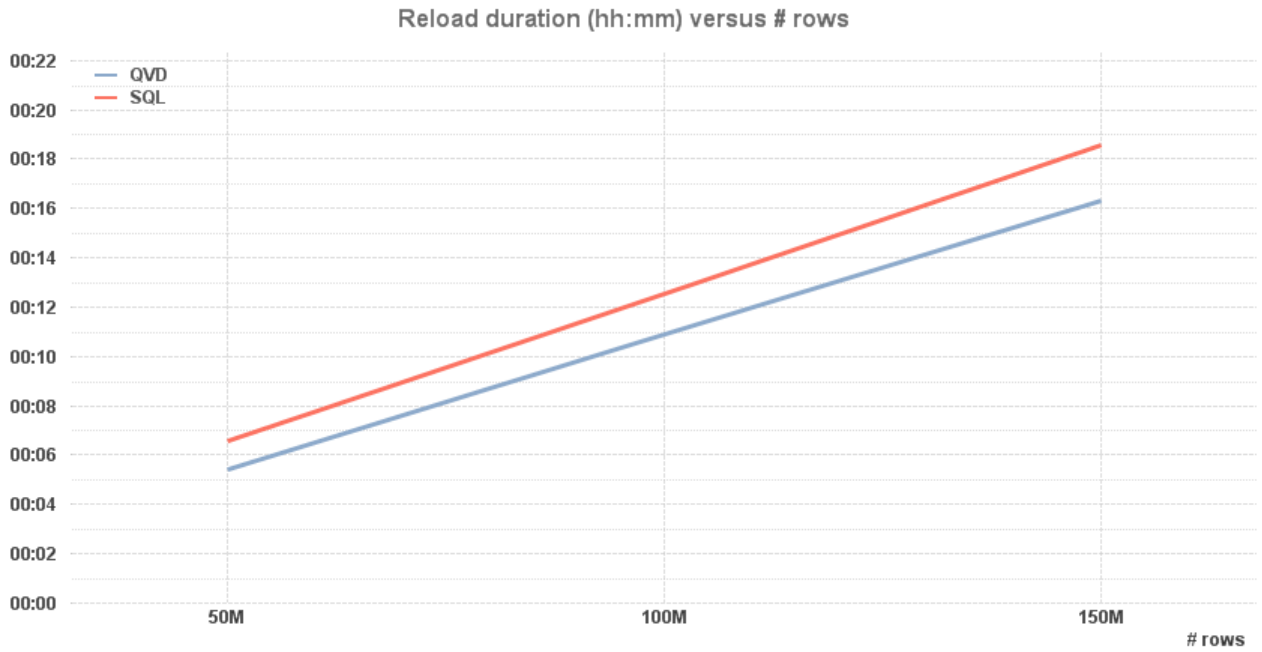


Key takeaway:

Optimizing remote DB (e.g. SQL server) might yield a big improvement.

Ensure bandwidth is sufficient for the pace the remote system can deliver data.

Reloads – Scaling With Data

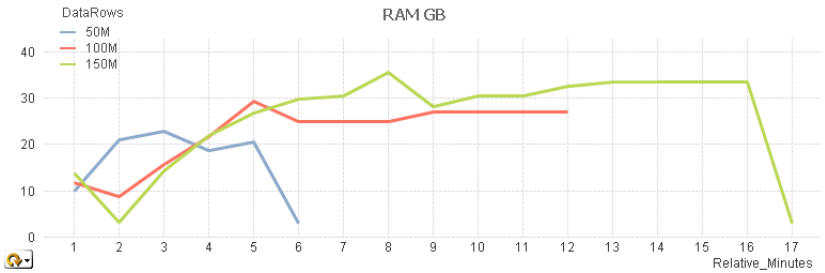


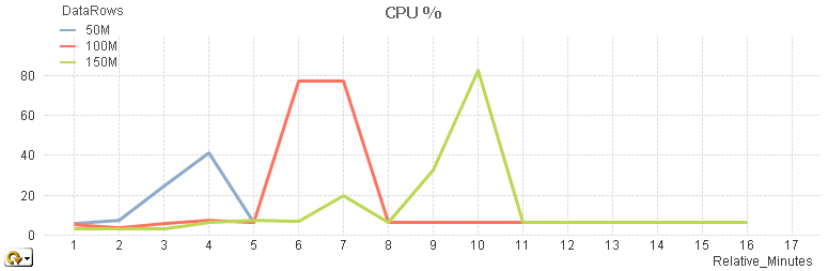
Key takeaway:

Resource consumption and reload duration is very predictable and scales linearly with amount of data.

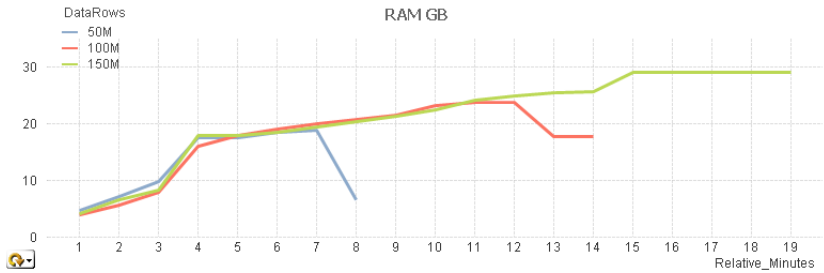
Reloads – Scaling With Data

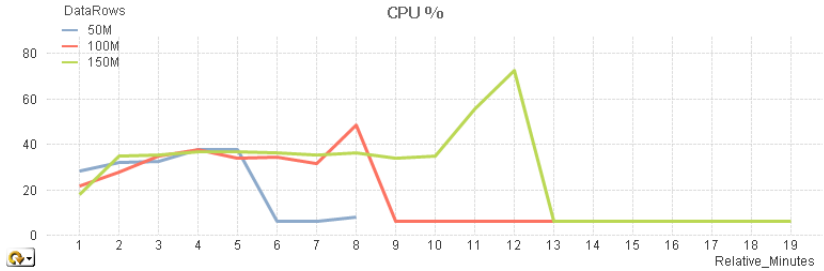
QVD





SQL





Key takeaway:

Resource consumption and reload duration is very predictable and scales linearly with amount of data.

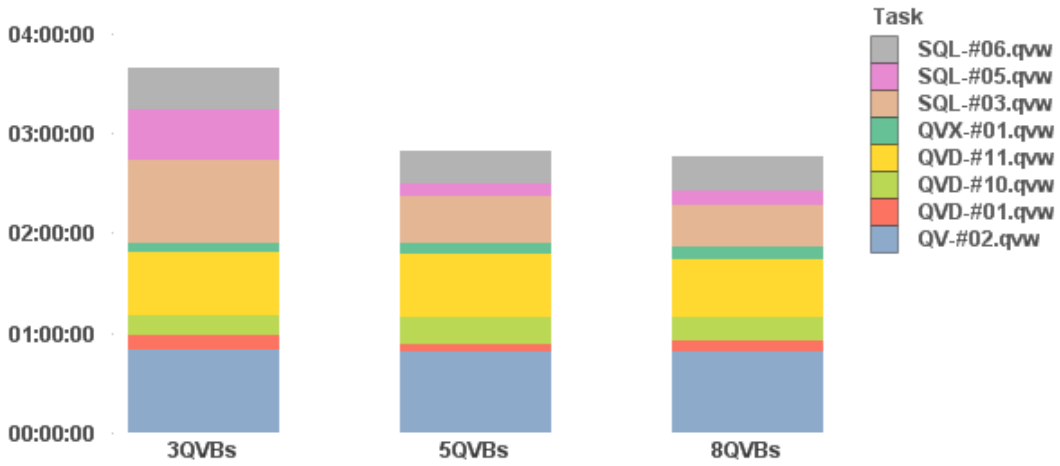
Database Aggregations

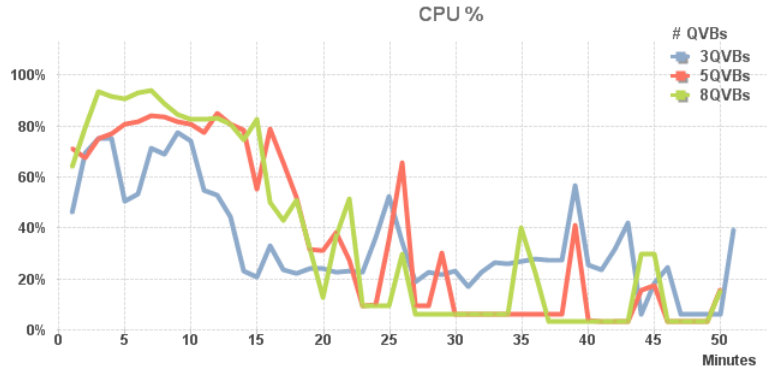
Data	QlikView	SQL	Combination
Load tables	7:17	7:22	7:20
Aggregations	14:35	0:08	1:32
Total duration	22:01	7:38	8:59
Script	Facts: **SQL** SELECT CustomerID, TaxAmount FROM SDB.dbo."50M"; **LOAD** CustomerID, avg(TaxAmount) as AVGtax RESIDENT Facts GROUP BY CustomerID;	**SQL** SELECT CustomerID, avg(TaxAmount) as AVGtax FROM SDB.dbo."50M" as Detail GROUP BY Detail.CustomerID;	**LOAD** CustomerID, avg(TaxAmount) as AVGtax GROUP BY CustomerID; **SQL** SELECT CustomerID, TaxAmount FROM SDB.dbo."50M";

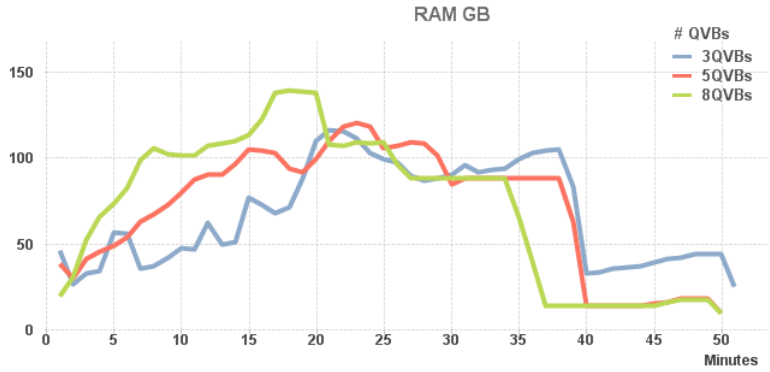
Key takeaway:

Let each component in the chain perform the task it is best suited for. Here it is allowing database to perform a part of or the whole aggregation needed.

Reloads – Multiple Reloads







Key takeaway:

QlikView Publisher is very efficient with parallel reloads and resource consumption is very predictable.

Reloads – Your Own Investigations

- Reload in QlikView Desktop will calculate initial state and keep document in RAM
- Each reload from Publisher will start a new QVB process.
- The QVB process will terminate when finished and release RAM as soon as the task is complete
- Every QVB is its own process
  - Multiple QVB does not exchange information/coordinate resource consumption
- Use Document Logfile for troubleshooting reload performance per document
- Use Task log for troubleshooting reload performance per machine

Key Takeaways Reloads - Selecting Hardware

- The “reload window” and other circumstances will define the importance of resource utilization and optimization.
  - In many cases a virtual environment might be sufficient or even the QlikView Server machine, if nightly reloads are applicable.

- When there are constraints:
  - Reading a stream of data is often very single threaded and this create a high clock frequency dependency.
  - Parallelizing reloads might improve unless contention reduce the total throughput

- White Listed 2-socket E5 are well suited for reload resource demand

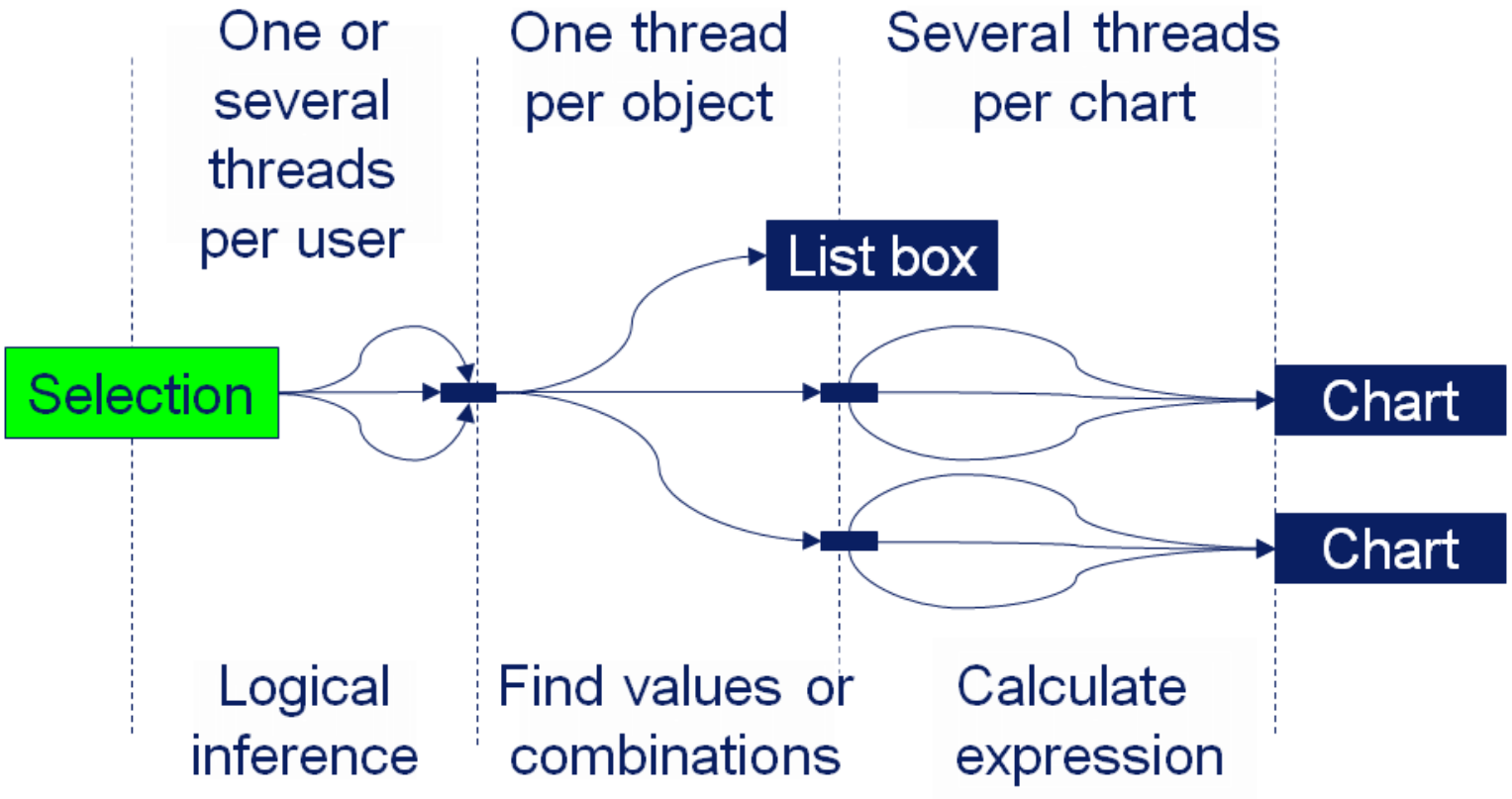
QlikView Server Performance

Behind The Scenes

Key components:

- The data, Symbol tables, State vectors, State space, Hyper cubes
- The data model is our view of the underlying structure of the data.
- We will focus on how this effects the hardware

Behind The Scenes

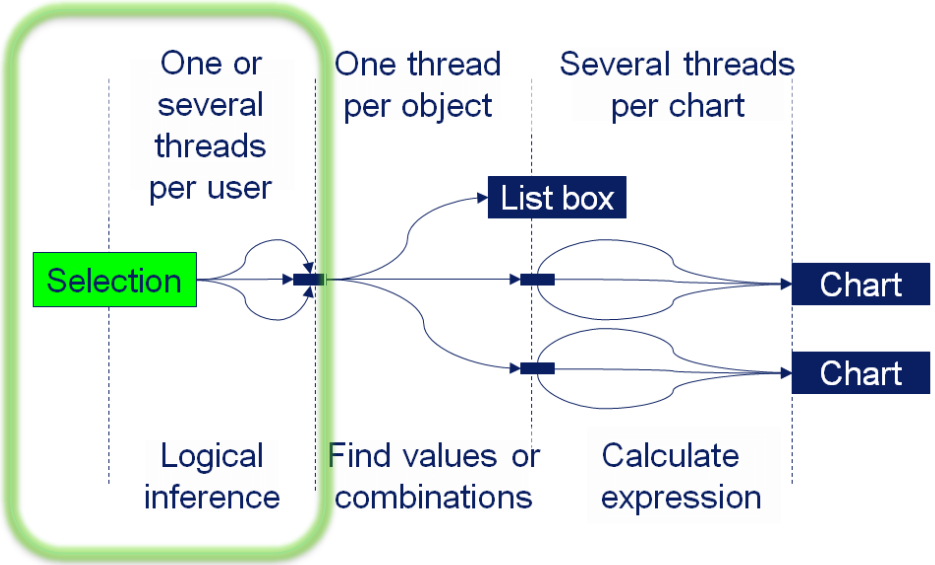


From QlikCommunity -> QlikView Design Blog -> The Calculation Engine
by Henric Cronström, August 20th 2013

Behind The Scenes

Selection phase, multi-threaded

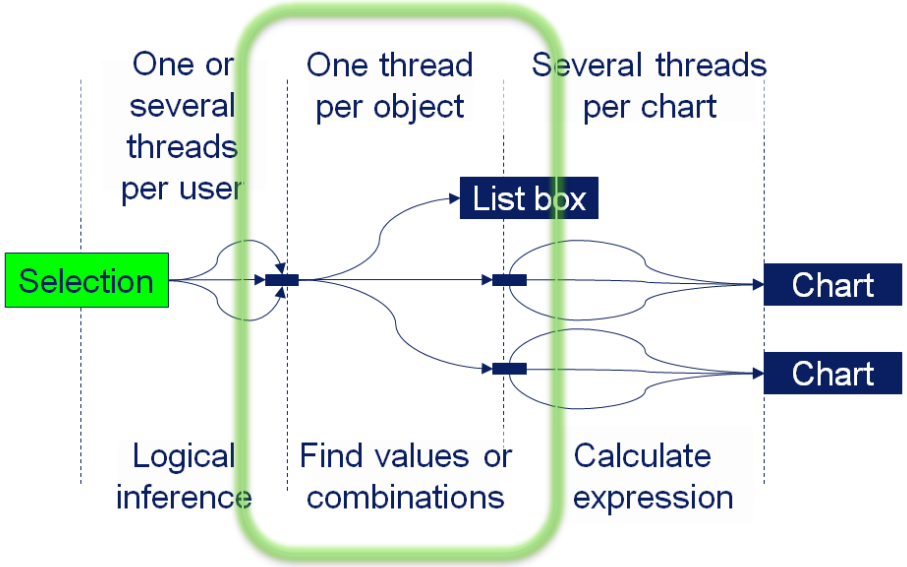
- Performed by the “inference engine”
- Receiving the selected value/s
- Going through the row-based data tables to create the “current state”
- Store the state space in the shared cache



Behind The Scenes

First Calculation phase, single-threaded

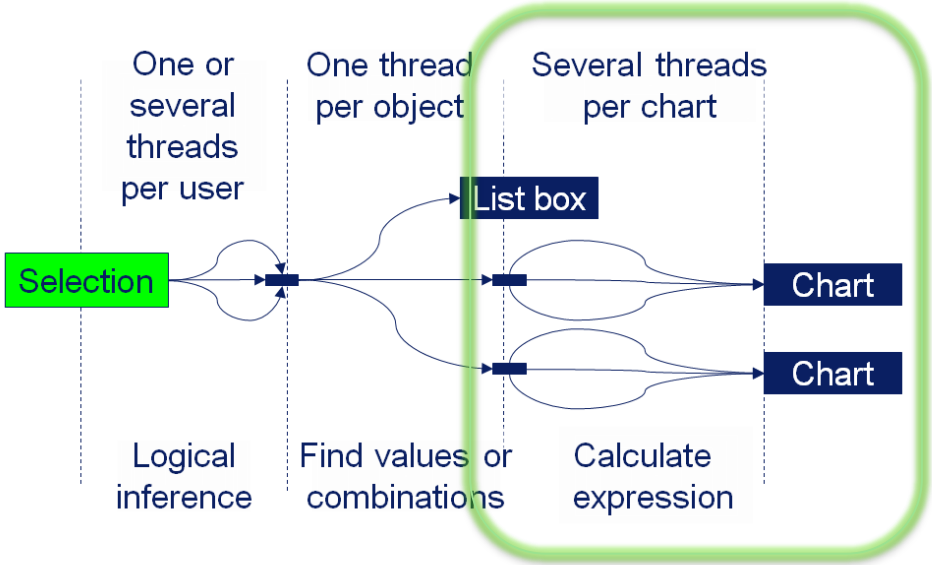
- Find values or combinations
- Per object, get the involved fields, dimensions, measures and use current state as input to create the dataset
- This include creating temporary tables, if needed



Behind The Scenes

Second Calculation phase, multi-threaded

- The actual calculation
- Per object, use the output from the previous phase and iterate over the dimensions and perform the actual calculation
- Perform the presentation of the result-table
- Store the result matrix in the cache

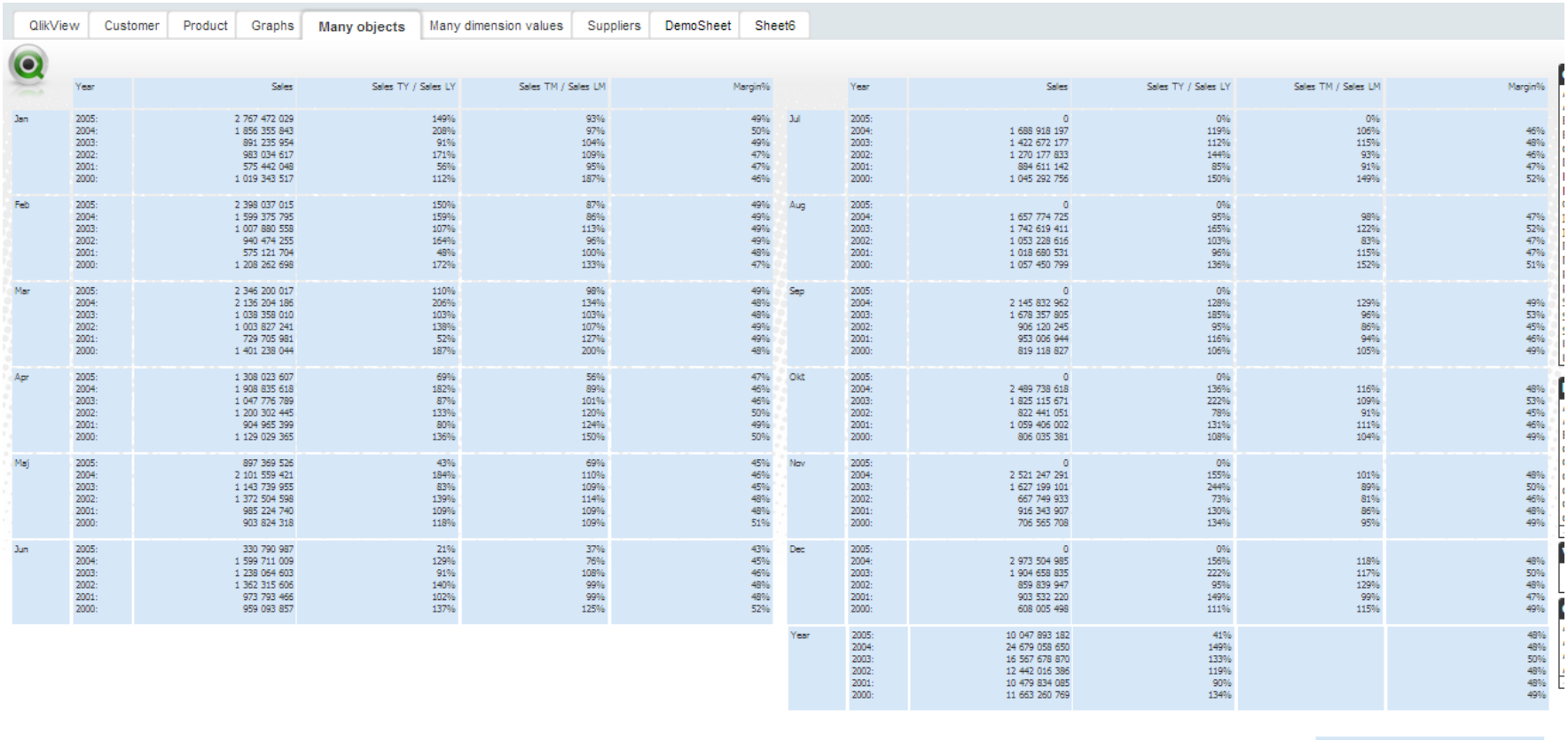


QlikView Server In Action

- 200m records fact table
- Sales data
- A quite common data model

Many Objects

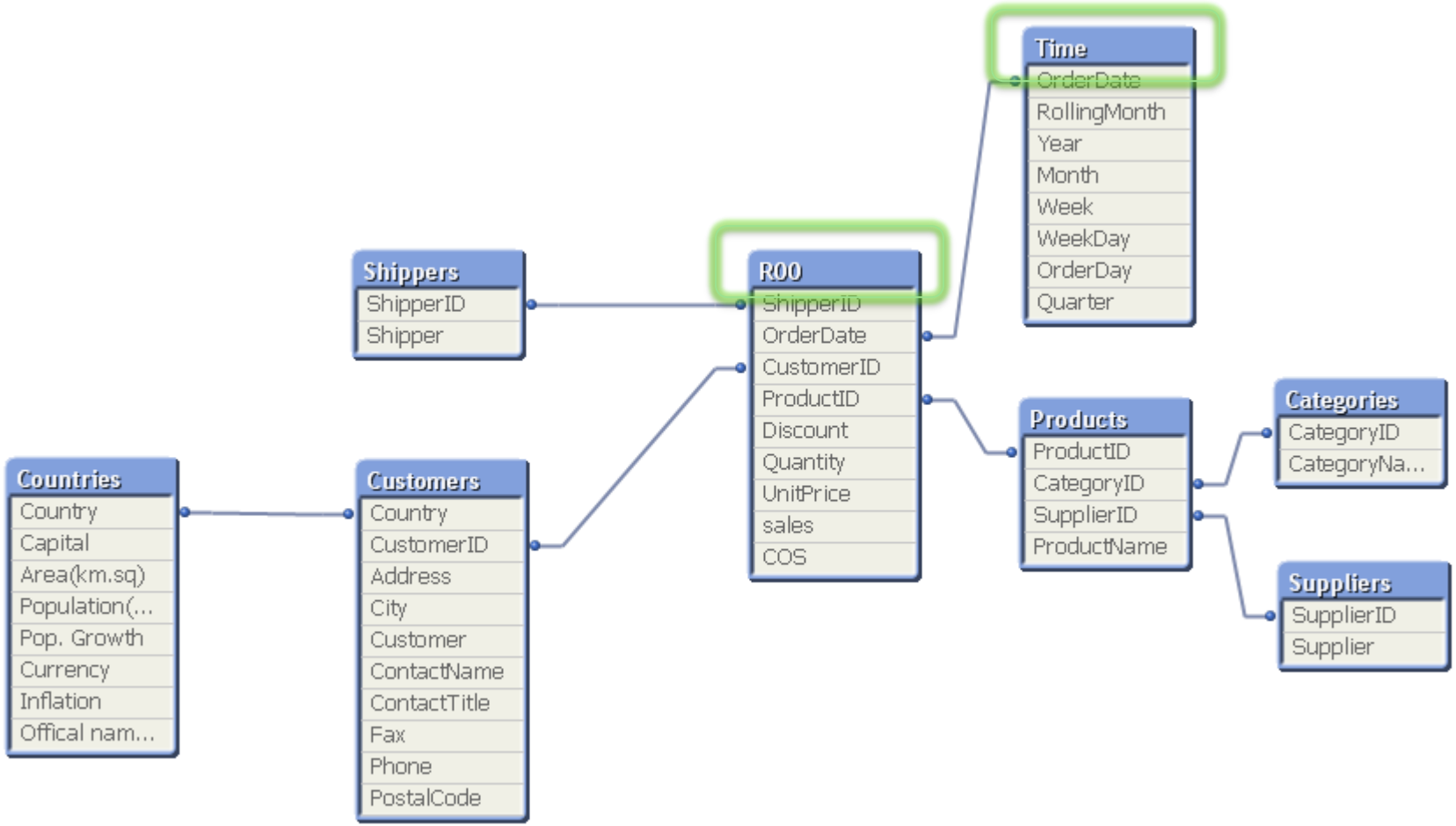
- Analysis of sales over time, many separate expressions and objects



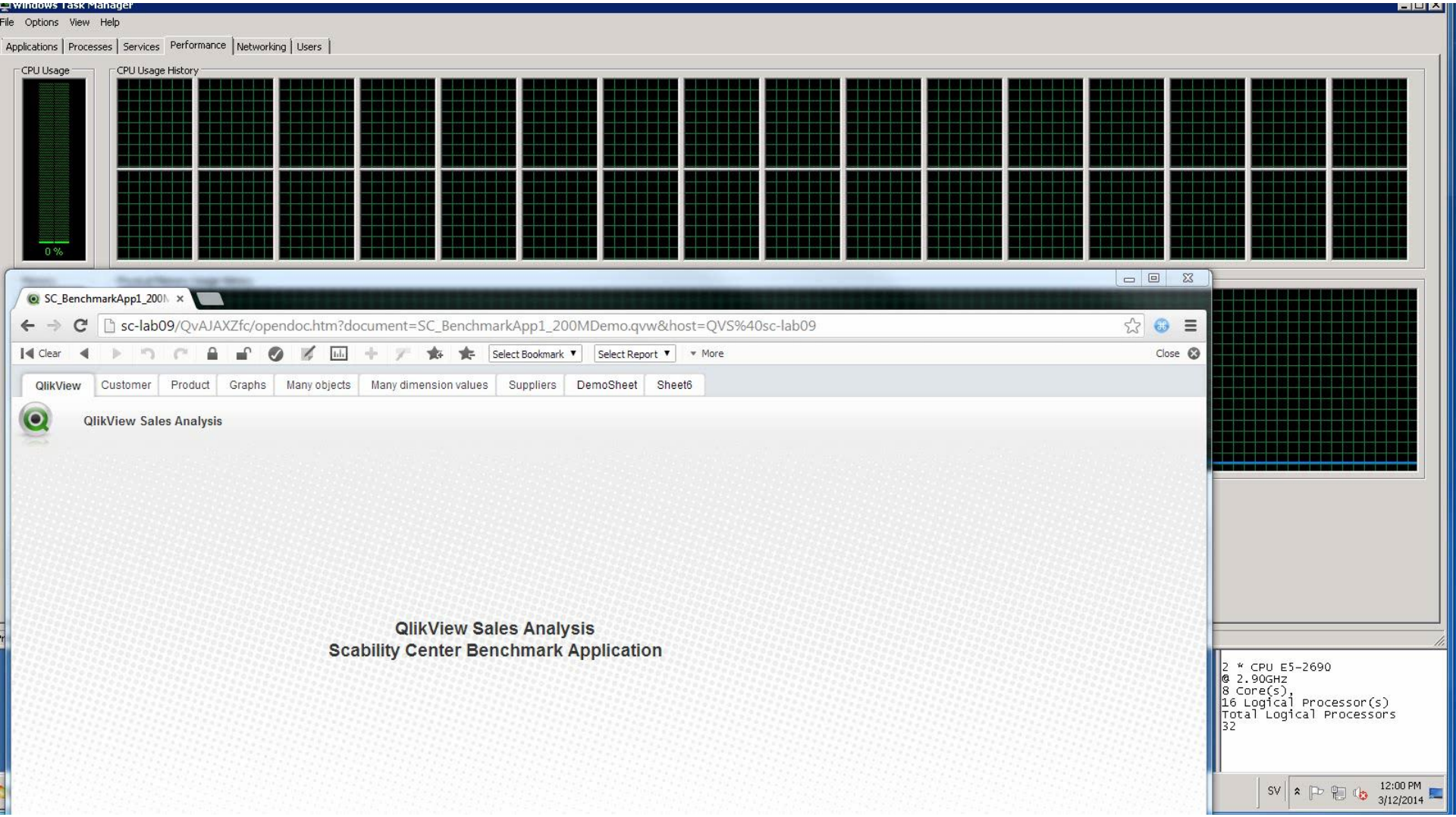
	Year	Sales	Sales TY / Sales LY	Sales TM / Sales LM	Margin%
Jan	2005:	2 767 472 029	149%	93%	49%
	2004:	1 856 355 843	208%	97%	50%
	2003:	891 235 954	91%	104%	49%
	2002:	983 034 617	171%	109%	47%
	2001:	575 442 048	56%	95%	47%
	2000:	1 019 343 517	112%	187%	46%
Feb	2005:	2 398 037 015	150%	87%	49%
	2004:	1 599 375 795	159%	86%	49%
	2003:	1 007 880 558	107%	113%	49%
	2002:	940 474 255	164%	96%	49%
	2001:	575 121 704	48%	100%	48%
	2000:	1 208 262 698	172%	133%	47%
Mar	2005:	2 346 200 017	110%	98%	49%
	2004:	2 136 204 186	206%	134%	48%
	2003:	1 038 358 010	103%	103%	48%
	2002:	1 003 827 241	138%	107%	49%
	2001:	729 705 981	52%	127%	49%
	2000:	1 401 238 044	187%	200%	48%
Apr	2005:	1 308 023 607	69%	56%	47%
	2004:	1 908 835 618	182%	89%	46%
	2003:	1 047 776 789	87%	101%	46%
	2002:	1 200 302 445	133%	120%	50%
	2001:	904 965 399	80%	124%	49%
	2000:	1 129 029 365	136%	150%	50%
Maj	2005:	897 369 526	43%	69%	45%
	2004:	2 101 559 421	184%	110%	46%
	2003:	1 143 739 955	83%	109%	45%
	2002:	1 372 504 598	139%	114%	48%
	2001:	985 224 740	109%	109%	48%
	2000:	903 824 318	118%	109%	51%
Jun	2005:	330 790 987	21%	37%	43%
	2004:	1 599 711 009	129%	76%	45%
	2003:	1 238 064 603	91%	108%	46%
	2002:	1 362 315 606	140%	99%	48%
	2001:	973 793 466	102%	99%	48%
	2000:	959 093 857	137%	125%	52%

	Year	Sales	Sales TY / Sales LY	Sales TM / Sales LM	Margin%
Jul	2005:	0	0%	0%	
	2004:	1 688 918 197	119%	106%	46%
	2003:	1 422 672 177	112%	115%	48%
	2002:	1 270 177 833	144%	93%	46%
	2001:	884 611 142	85%	91%	47%
	2000:	1 045 292 756	150%	149%	52%
Aug	2005:	0	0%		
	2004:	1 657 774 725	95%	98%	47%
	2003:	1 742 619 411	165%	122%	52%
	2002:	1 053 228 616	103%	83%	47%
	2001:	1 018 680 531	96%	115%	47%
	2000:	1 057 450 799	136%	152%	51%
Sep	2005:	0	0%		
	2004:	2 145 832 962	128%	129%	49%
	2003:	1 678 357 805	185%	96%	53%
	2002:	906 120 245	95%	85%	45%
	2001:	953 006 944	116%	94%	46%
	2000:	819 118 827	106%	105%	49%
Okt	2005:	0	0%		
	2004:	2 489 738 618	136%	116%	48%
	2003:	1 825 115 671	222%	109%	53%
	2002:	822 441 051	78%	91%	45%
	2001:	1 059 406 002	131%	111%	46%
	2000:	806 035 381	108%	104%	49%
Nov	2005:	0	0%		
	2004:	2 521 247 291	155%	101%	48%
	2003:	1 627 199 101	244%	89%	50%
	2002:	667 749 933	73%	81%	46%
	2001:	916 343 907	130%	86%	48%
	2000:	706 565 708	134%	95%	49%
Dec	2005:	0	0%		
	2004:	2 973 504 985	156%	118%	48%
	2003:	1 904 658 835	222%	117%	50%
	2002:	859 839 947	95%	129%	48%
	2001:	903 532 220	149%	99%	47%
	2000:	608 005 498	111%	115%	49%
Year	2005:	10 047 893 182	41%		48%
	2004:	24 679 058 650	149%		48%
	2003:	16 567 678 870	133%		50%
	2002:	12 442 016 386	119%		48%
	2001:	10 479 834 085	90%		48%
	2000:	11 663 260 769	134%		49%

Many Objects

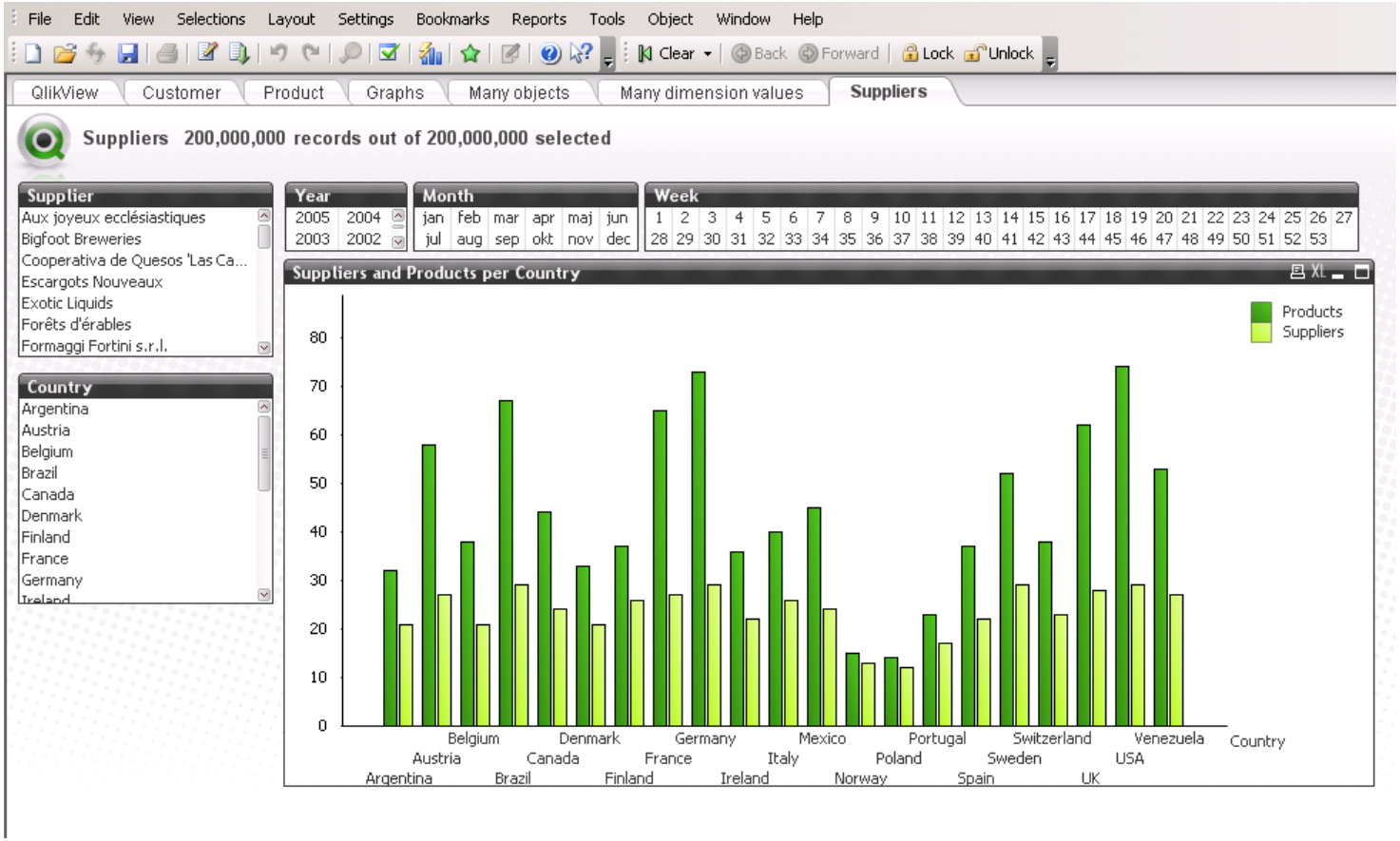


Many Objects

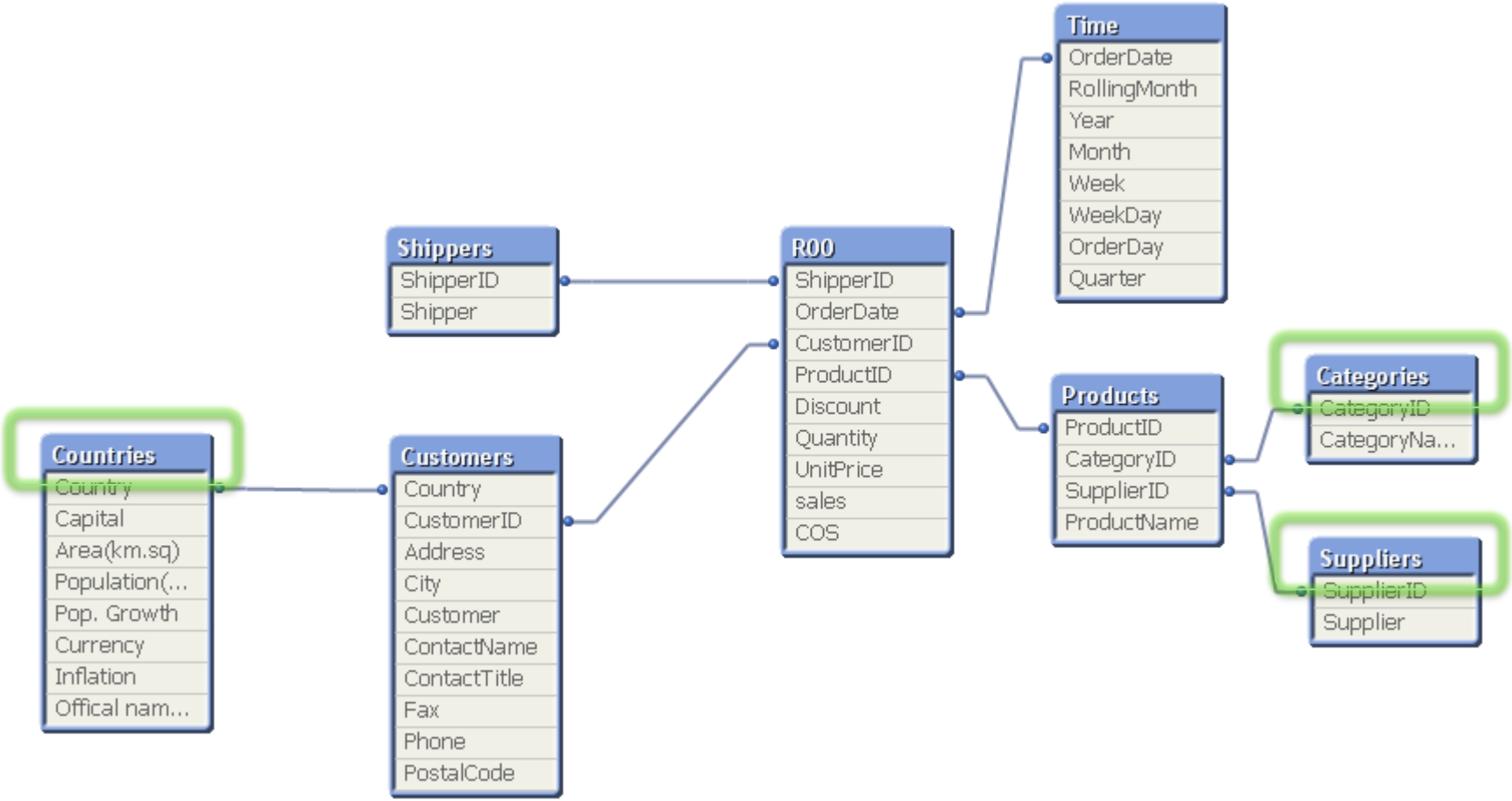


Heavy in the “First Calculation” Phase

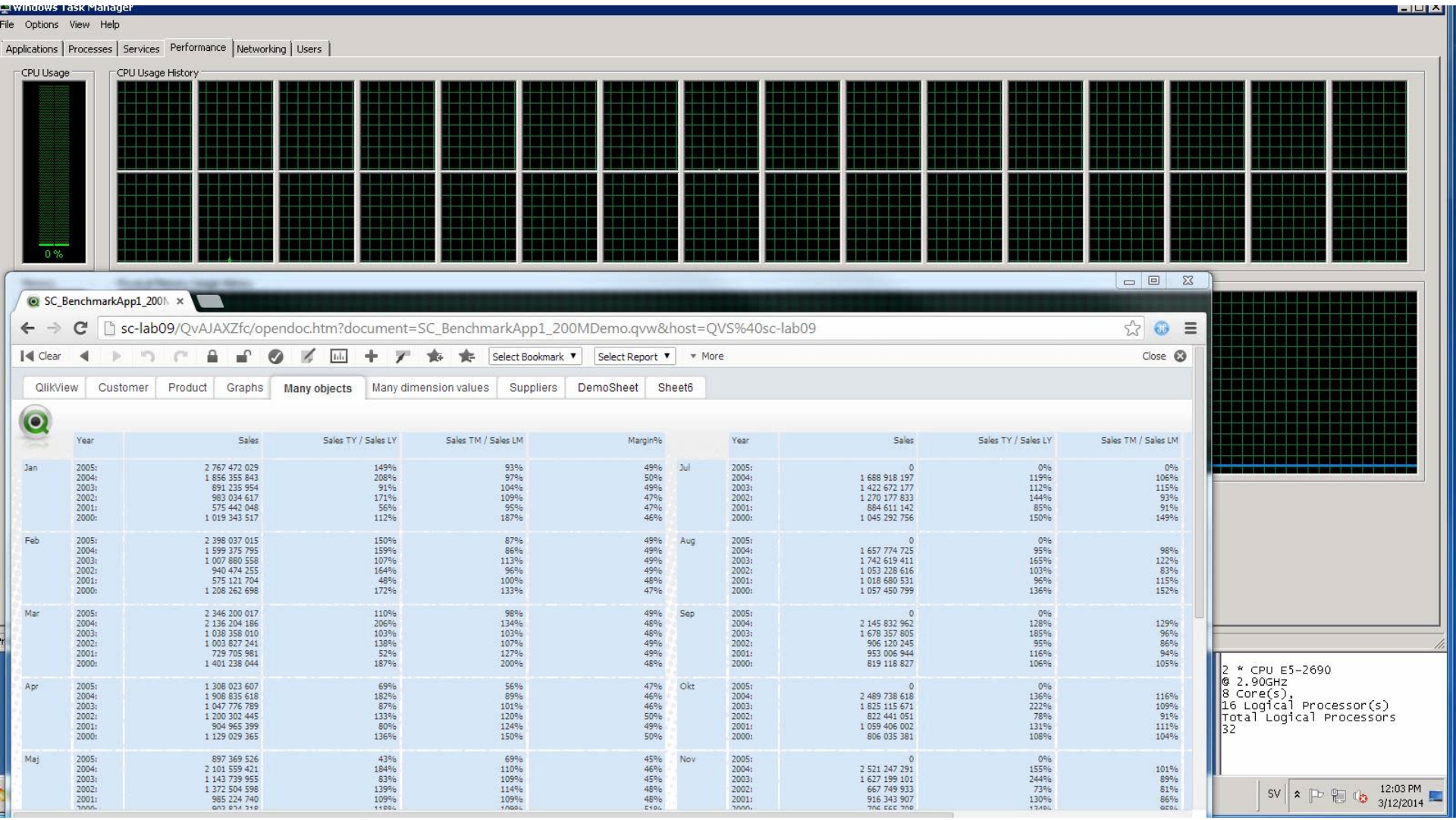
- Discoveries around suppliers over countries and their products



Heavy in the “First Calculation” Phase

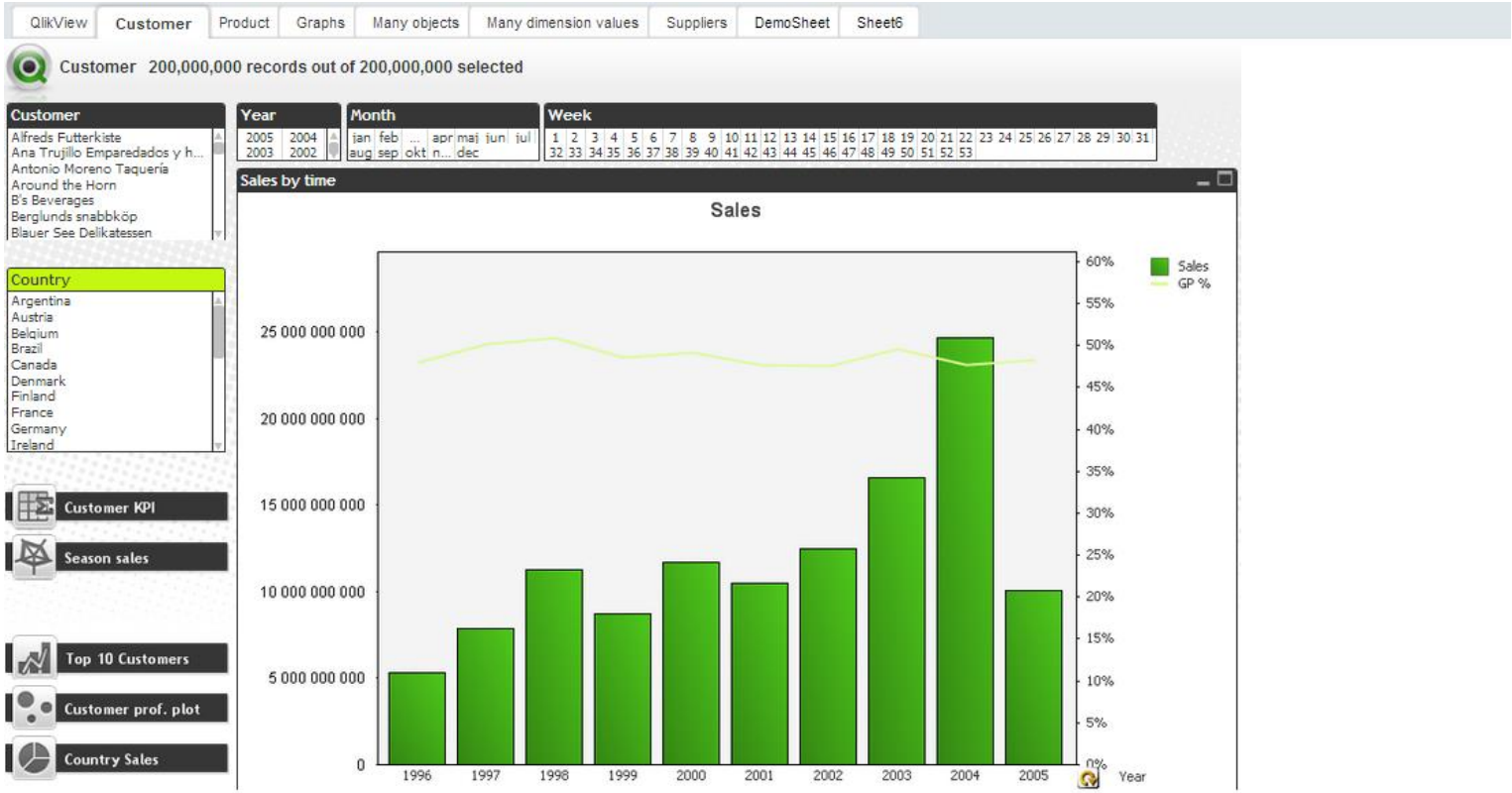


Heavy in the “First Calculation” Phase

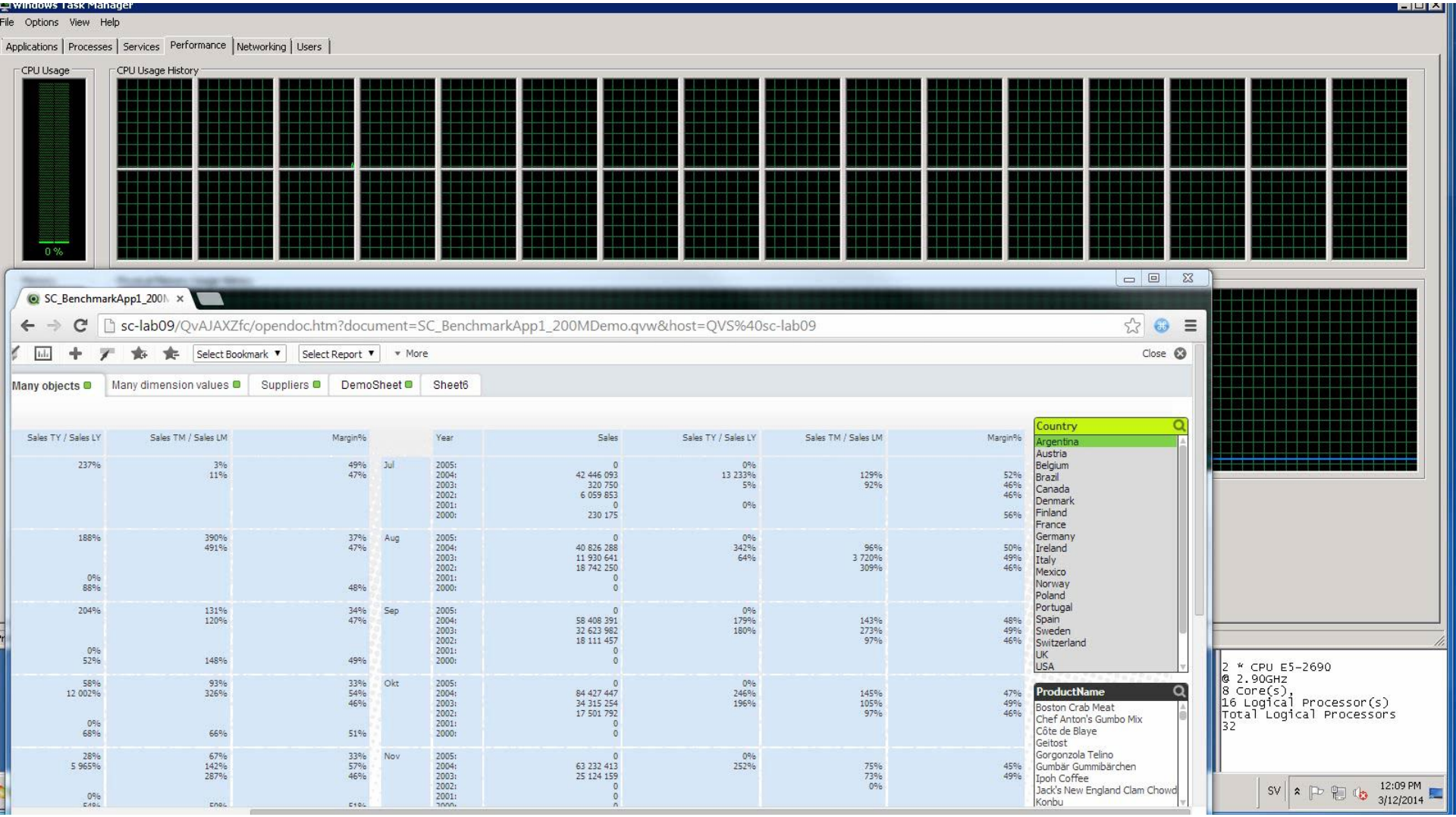


Common Resource Usage

- Discoveries around sales in different countries



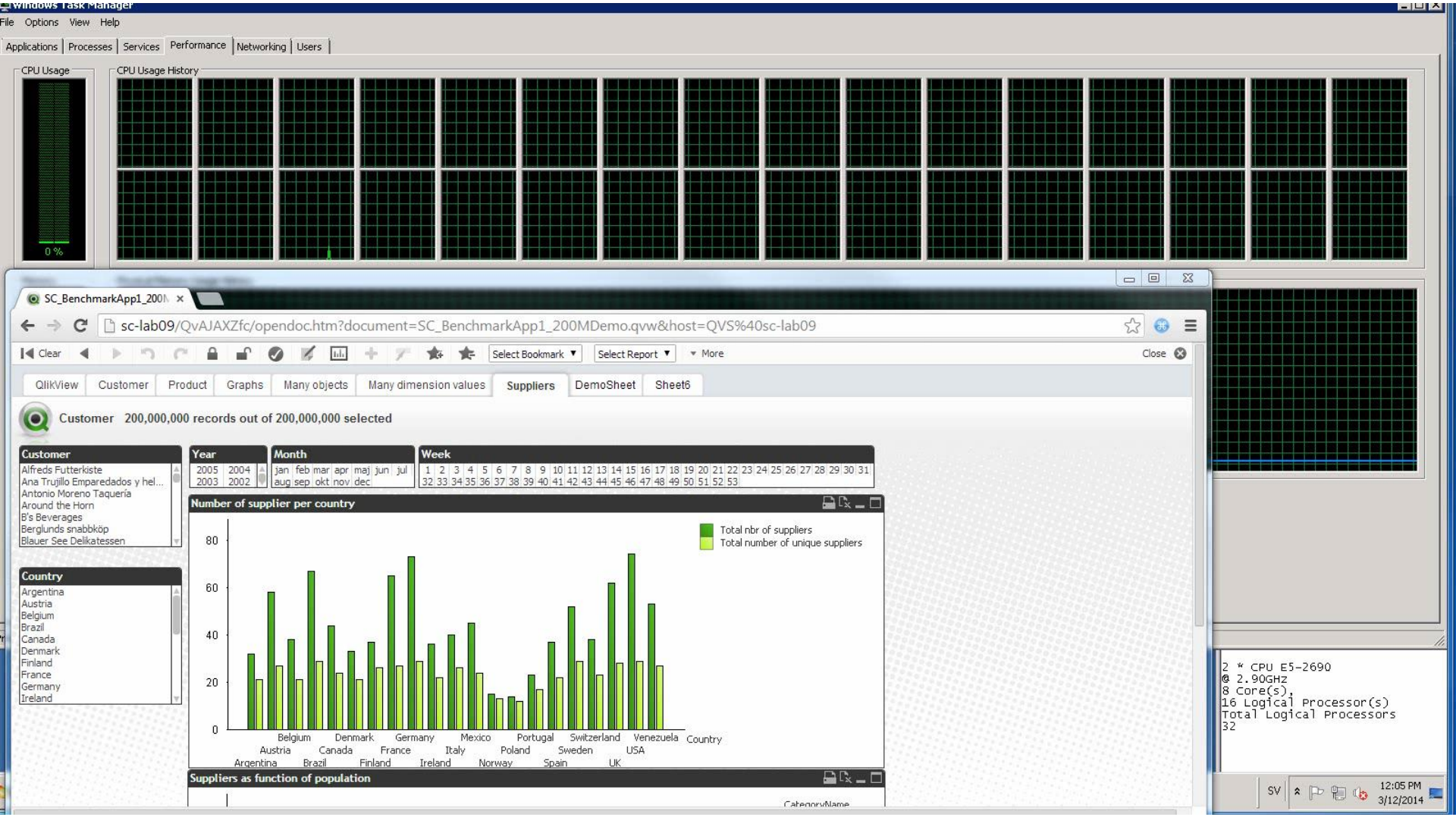
Common Resource Usage



Cached Calculations

- Once something is calculated, it is stored in the cache and reused, when needed, instead of going through the same phases all over again.

Cached Calculations



Summary and key takeaways

- It is not the functions themselves that appear to be single threaded but how long each object stay in the first calculation phase, when gathering data before performing the actual calculations.
- Key focus areas:
  - Identify the object/s that exhibit the unwanted behavior
  - Identify the fields used in these objects and their position in the data model
  - Understand the data involved when traversing the data model
  - Assess all of the expressions that are triggered in these objects

Selecting Hardware

Architecture

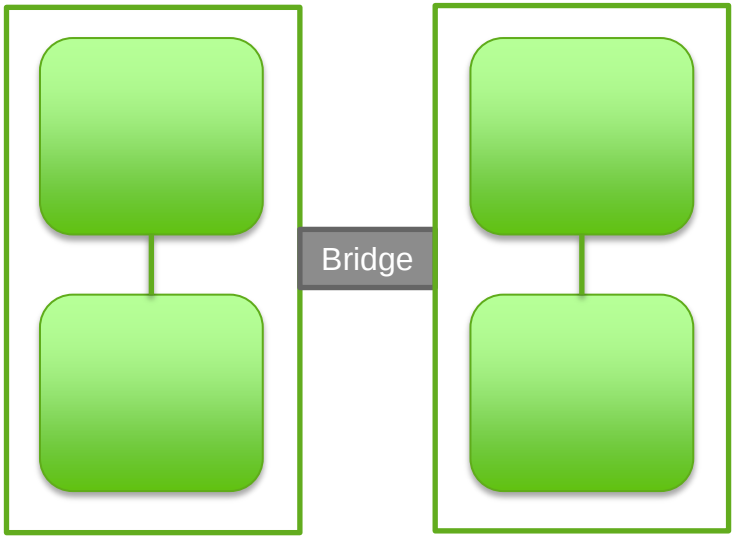
2-socket machine

Architecture

2-socket machine

- What is a good fit for QlikView Performance?

Good fit:
E5-26XX



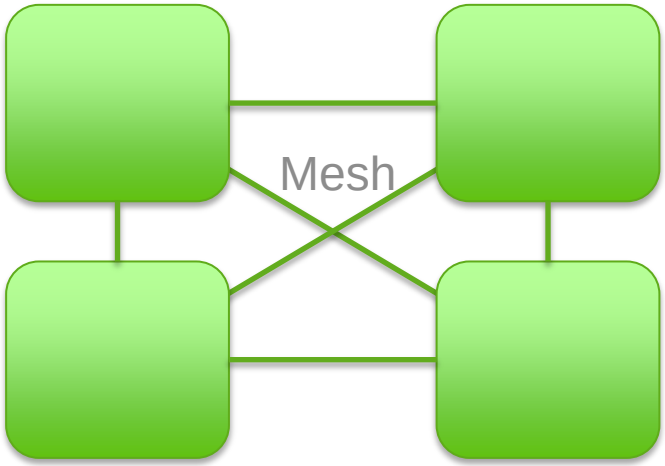
Not a good fit:
E7-28XX

Architecture

4-socket machine

- What is a good fit for QlikView Performance?

Not a good fit:
E5-46XX

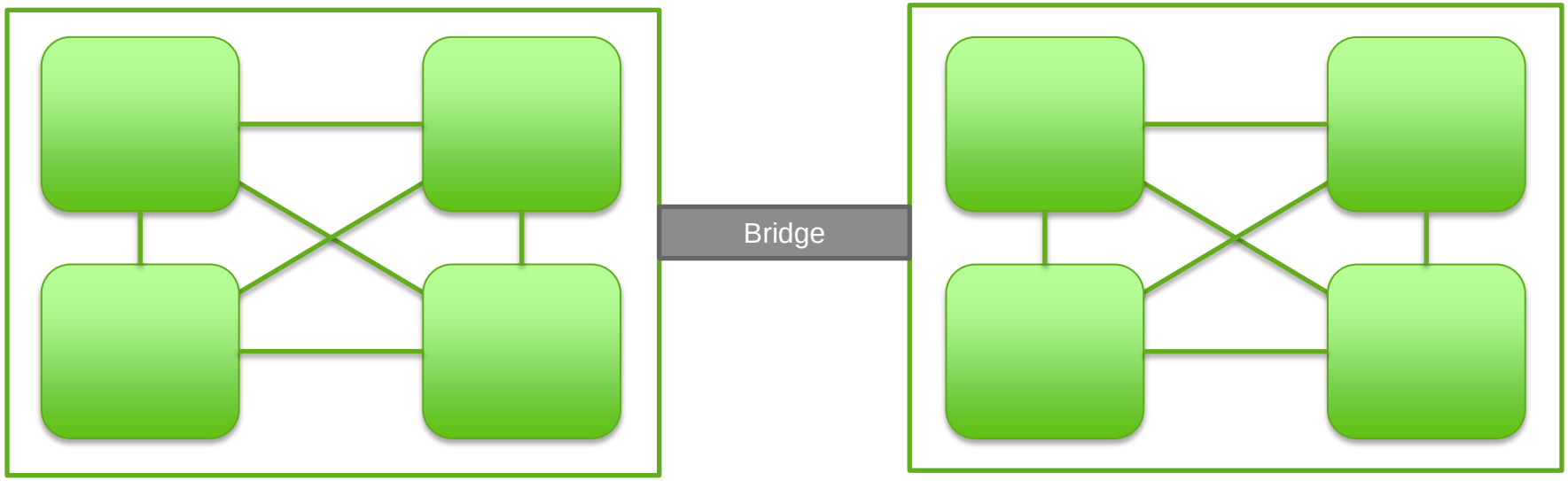


Good fit:
E7-48XX

Architecture

8-socket machine

- 8-sockets can be assembled in many different ways, bridged or “glue less”
- No 8-socket machines have made it to the Server White List



Selecting Hardware

A good starting point

- Whitelisted
  - http://www.qlik.com/us/explore/resources/brochures-datasheets
- Configured according to recommendations
  - http://community.qlik.com/docs/DOC-2362
- Leverage QlikView Hardware Benchmarking test to validate that server performs as expected
  - http://community.qlik.com/docs/DOC-2942

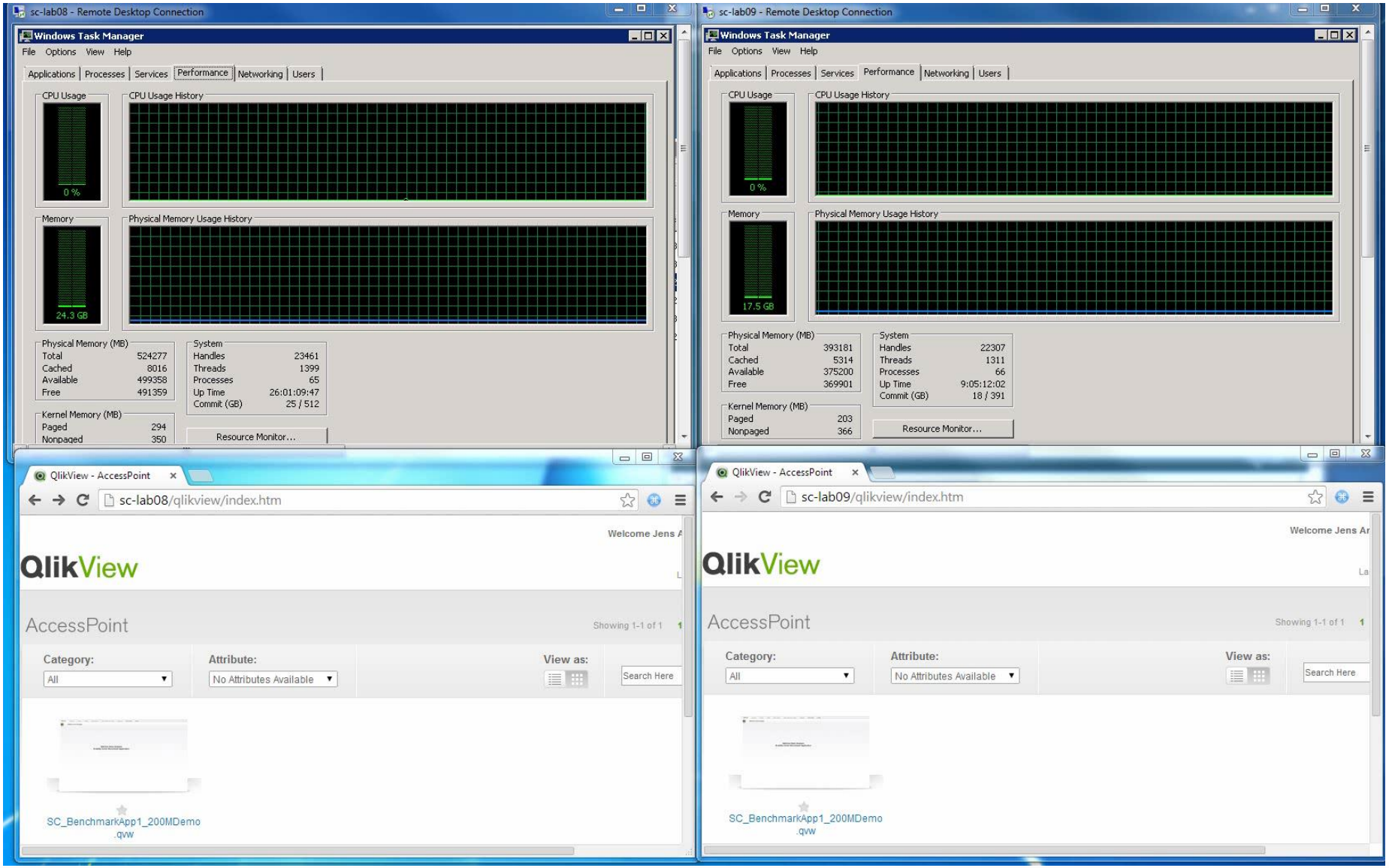
But which and how many of the whitelisted servers to select?

QlikView on 4- and 2-socket

A case study

- Same application
- Side by side comparison
  - 2 socket E5-2690@2.9/3.8 GHz (white listed)
  - 4 socket E7-4870@2.4/2.8 GHz (white listed)
- When does performance benefit from clock speed and when does it benefit from the total processing capacity (cores AND clock speed)?

QlikView on 4- and 2-socket



4x10 E7-4870@2.4GHz

2x8(HT) E5-2690@2.9GHz

Selecting Hardware

- What is the difference between servers?
  - Vendors
  - Model
  - Architecture
  - # Sockets
  - # RAM
  - Chipsets
    - Clock frequency
    - Cores
    - QPI speed
    - …

The whitelist is a “cheat sheet” with good combinations of the above

Selecting Hardware

- How to think about these attributes when selecting servers
- What type of resources are needed
  - Performance, single user scenarios
    - For a large well formed application -> Total capacity
    - For applications with many demanding calculations -> Total capacity
    - For less demanding calculations -> Clock frequency
    - For a less than optimal application -> Clock frequency
  - Performance, multi user scenarios
    - See above
    - Once approaching saturation -> Total capacity is key

- To estimate demand for total capacity
  - QlikView Scalability Tools, http://community.qlik.com/docs/DOC-2705

When Is My Hardware Saturated?

- CPU
  - Saturation can be detected via average CPU utilization, but what % will depend on time frame and pattern
    - E.g. > 70% during busy hour yields saturation
    - E.g. 80% average over a certain minute does not
- RAM
  - QlikView Server is having a hard time to stay below Working Set Low means saturation. Seen by eventlog messages
    - “Warning WorkingSet: Virtual Memory is growing beyond parameters”
    - “Warning WorkingSet: Virtual Memory is growing CRITICALLY beyond parameters”

Summary and key takeaways

- Leverage the knowledge of how QlikView utilize resources to better:
  - Shorter reload times
  - Make better data strategies
  - Select hardware for reloads
  - Optimize applications
  - Select hardware for QlikView Servers
  - Troubleshoot performance issues

Feedback

Please complete the track session survey via the mobile app

- Access the track session survey through the mobile app
- Enter track session code TE38
- Provide your feedback

Question
&
Answer

Thank You