Method – GAPP Overview

Januari 2011

Gerwin Hendriksen
Information Architect

http://method-gapp.com

V1.0

Who I am…

- Started as DBA at Dedicate (later called Parity) for 1.5 years
- Worked for 5 years at Oracle The Netherlands
  - First 1.5 year as RDBMS analyst
  - Rest worked as ARE EBS Performance and RAC
    - This included Content Lead of EBS with RAC world wide
- Worked for 2 years at IBM
  - Business Consultancy Services
- Worked for 3 years at AMIS Services BV
- Currently working for IT-eye as Information Architect

Agenda…

- Situation
- End-User processes and infrastructure
- What is GAPP and how can it help us
- Conclusions
- Q/A

Situation

Situation…

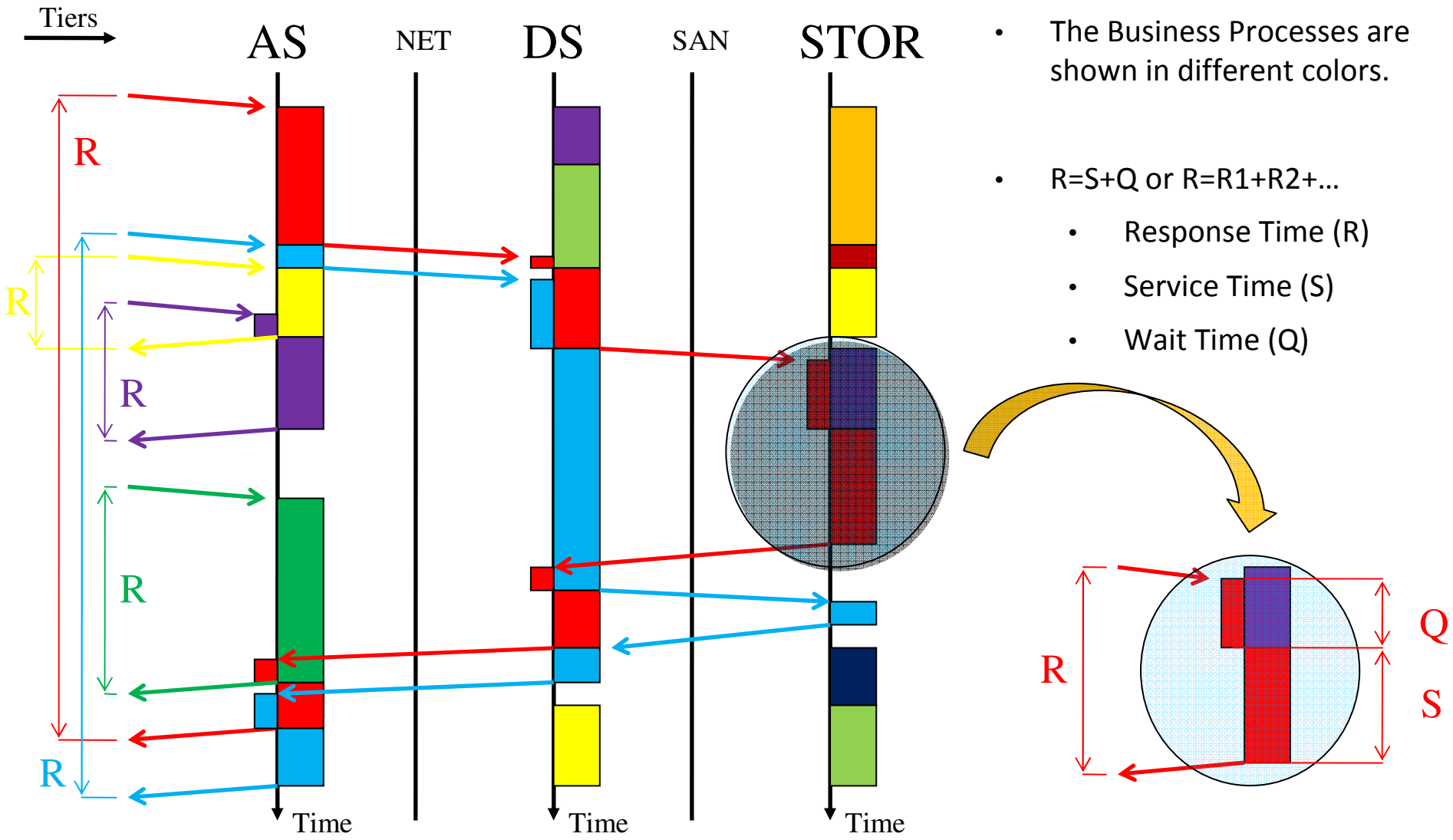
- End user processes are sometimes too slow and the cause hard to be determined.
- The technical infrastructure has shared components like a SAN, Server, etc.
- The application is hard to be debugged, because it is a third party application.

It would be great if…

- You could find the cause of such problems without
  - Installing agents
  - Changing code (hooks)
  - Technology dependency
- You could even find out the influence of other applications on our end user processes.
- You could predict what the performance impact would be before doing investments.

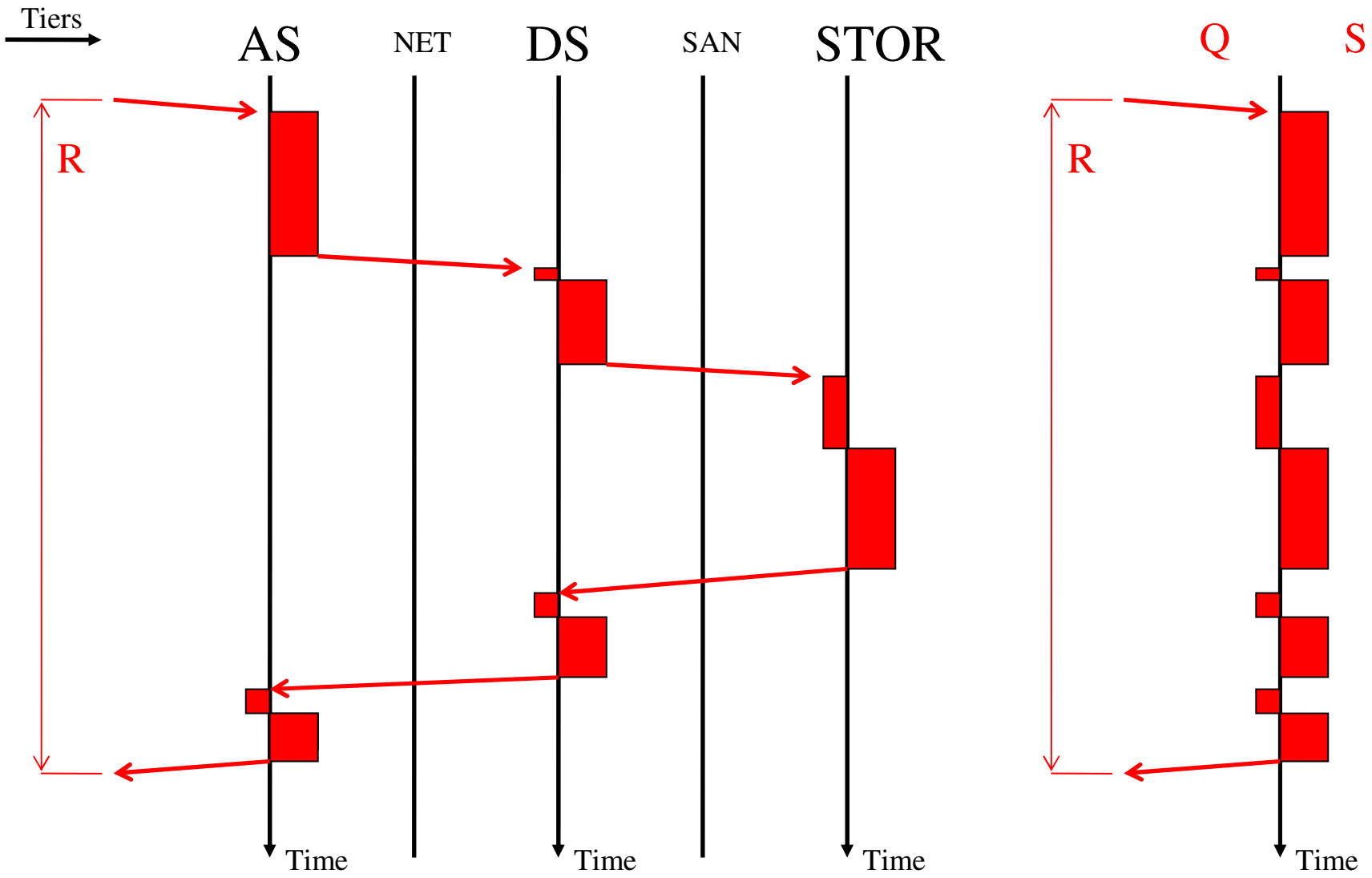
End-User processes and infrastructure

Method-R and the sequence diagram…

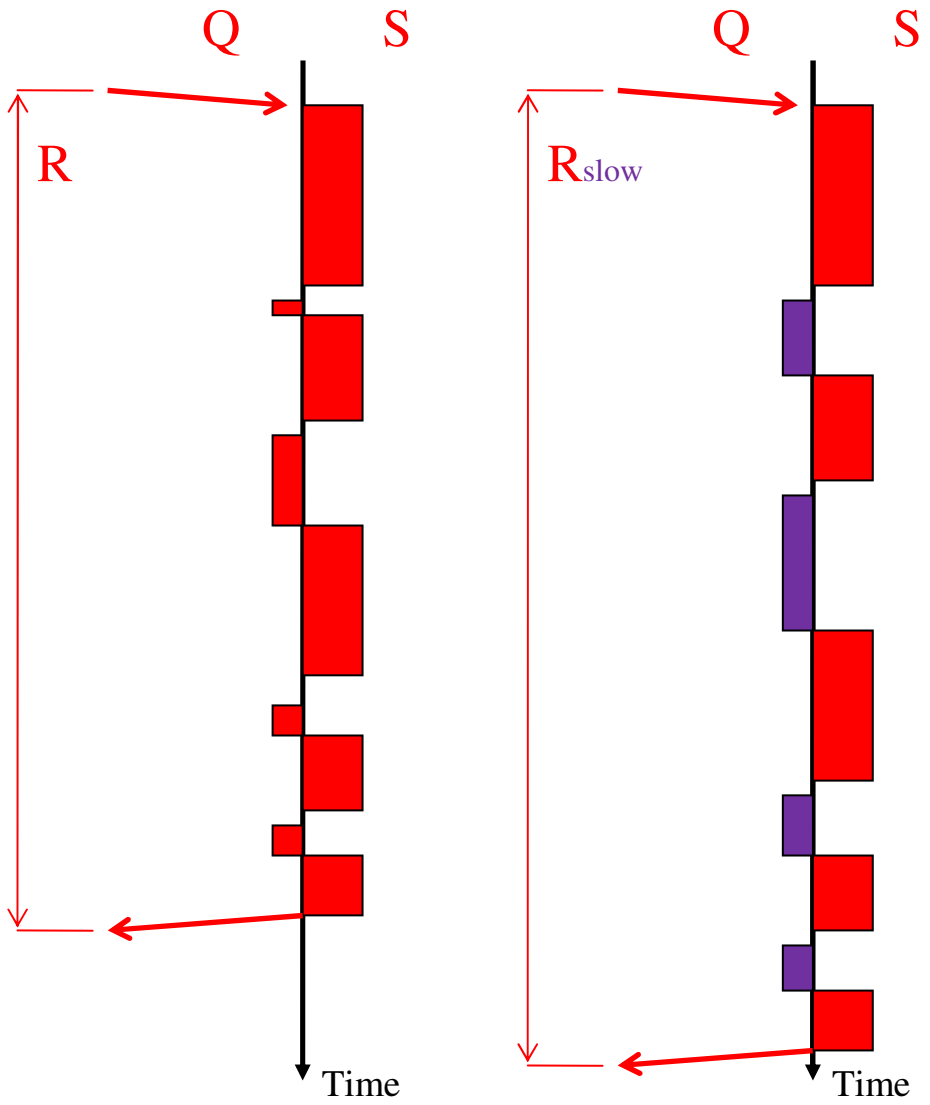


- The Business Processes are shown in different colors.
- R=S+Q or R=R1+R2+…
  - Response Time (R)
  - Service Time (S)
  - Wait Time (Q)

Highlighted the “Red” process…



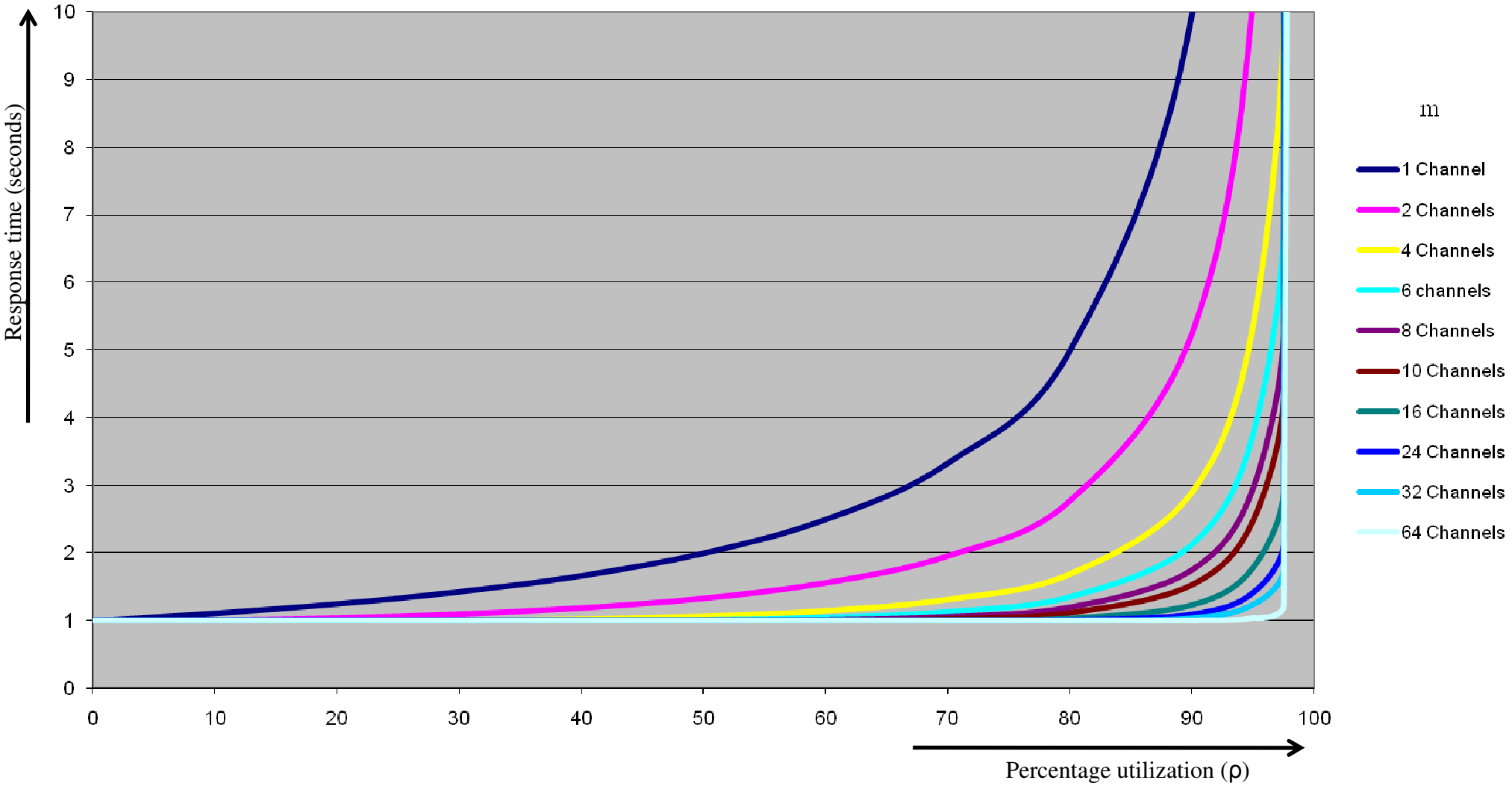
Increasing Response Time…



- "Service Time" stays within limits the same*, given the fact that the process is doing every time something similar.
- "Wait Time" is caused by resource business
- "Wait Time" is causing the significant part of the variance in R (under normal load conditions)

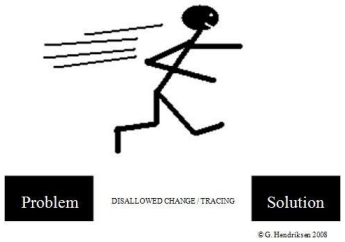
* Coherency effects

Illustrative M/M/m Model Simplified…



What is GAPP

What is Method-GAPP…



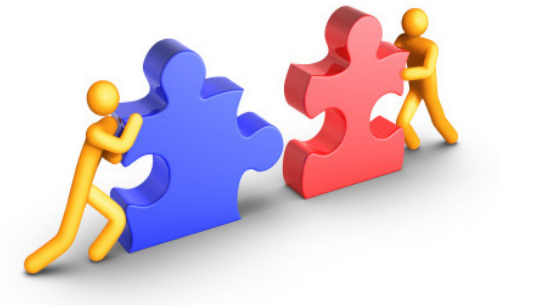
Method-GAPP is a performance method which makes smart use of underlying queuing models and data mining to find bottlenecks in complex architectures, for specific end-user processes within an enterprise.

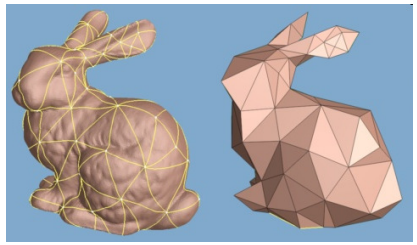
"GAPP" is an abbreviation of:

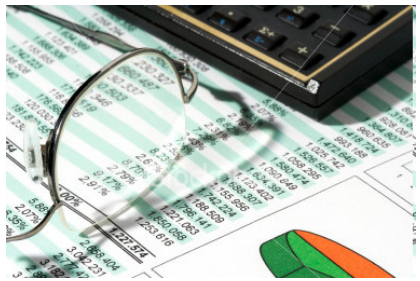
"**General Approach Performance Profiling**"

How does Method-GAPP Work…

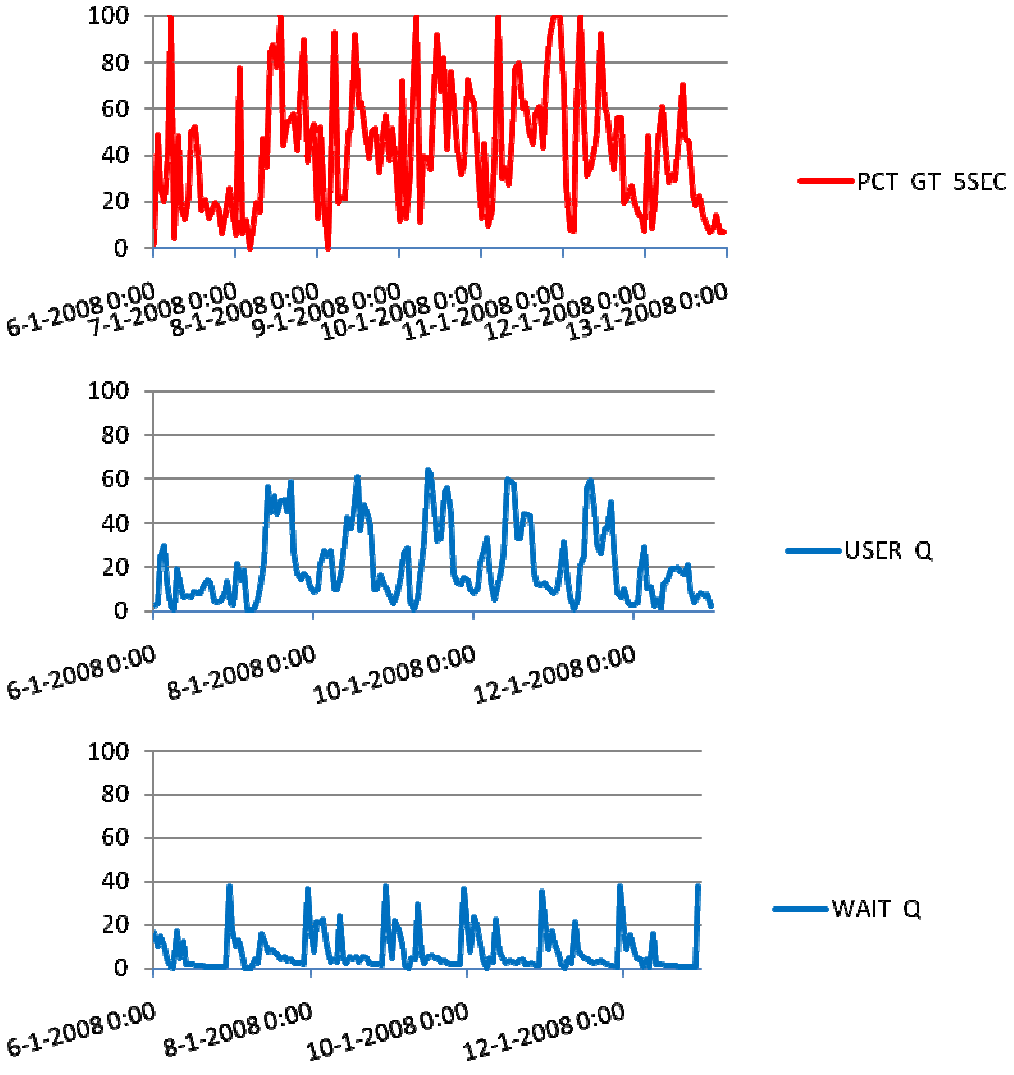
- Data Collection
- Data Synchronization
- Data Modeling
- Data Mining
- Data Interpretation

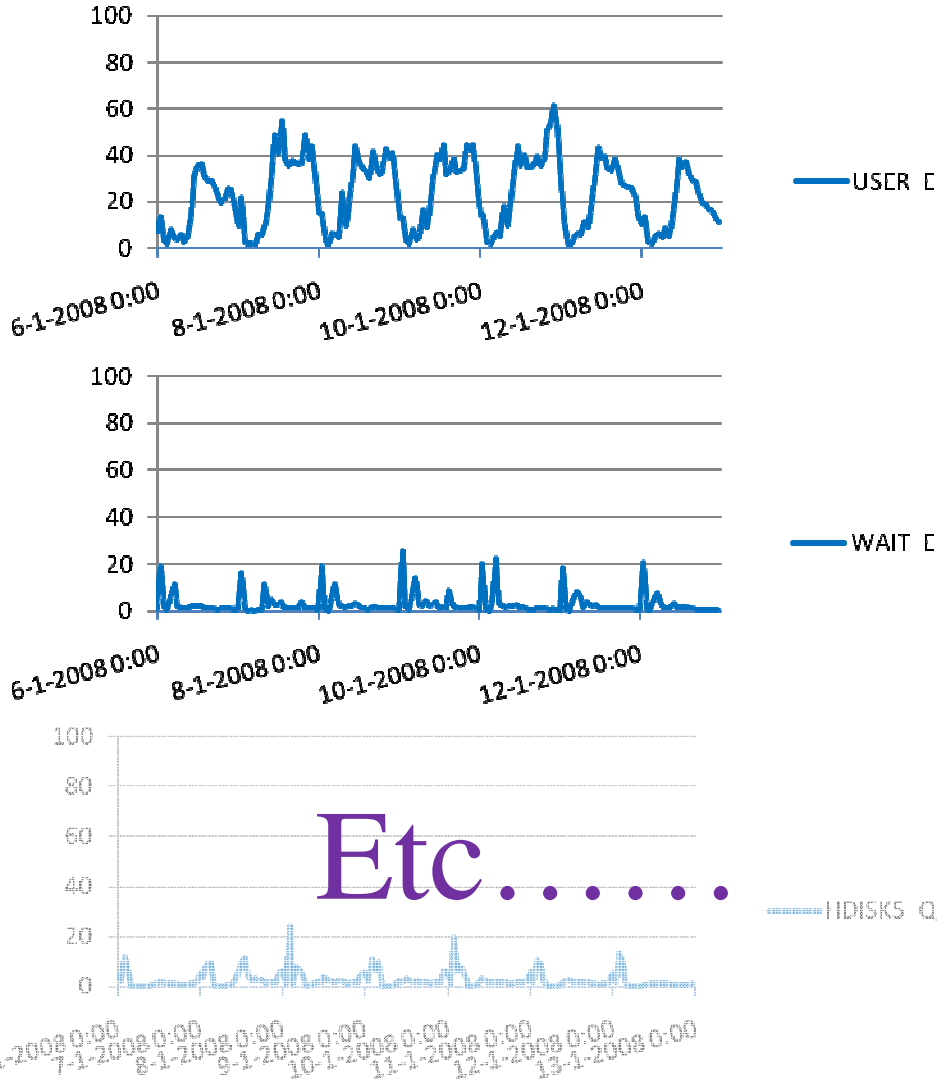






Data Collection and Synchronization…

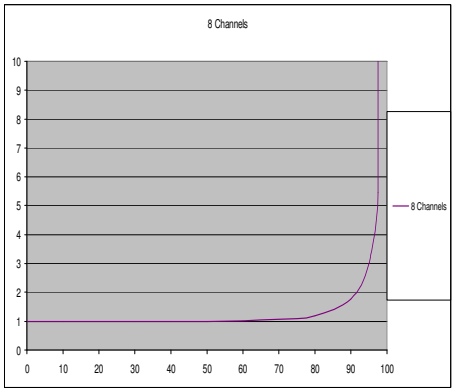


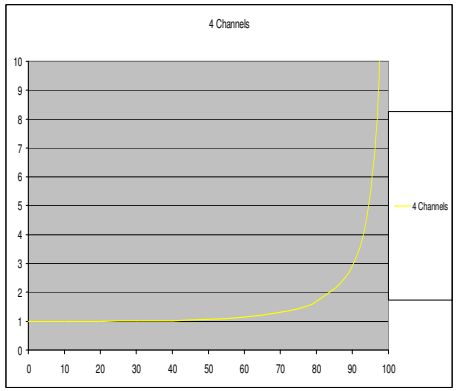


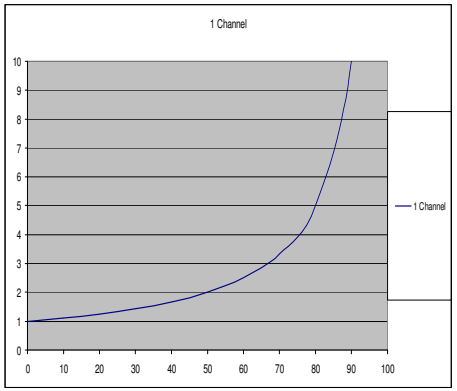
Data Mining and data modeling…

For Example:

$$R_{tot} = c_1 R_{1,n=8} + c_2 R_{2,n=4} + c_3 R_{3,n=1} + etc.$$







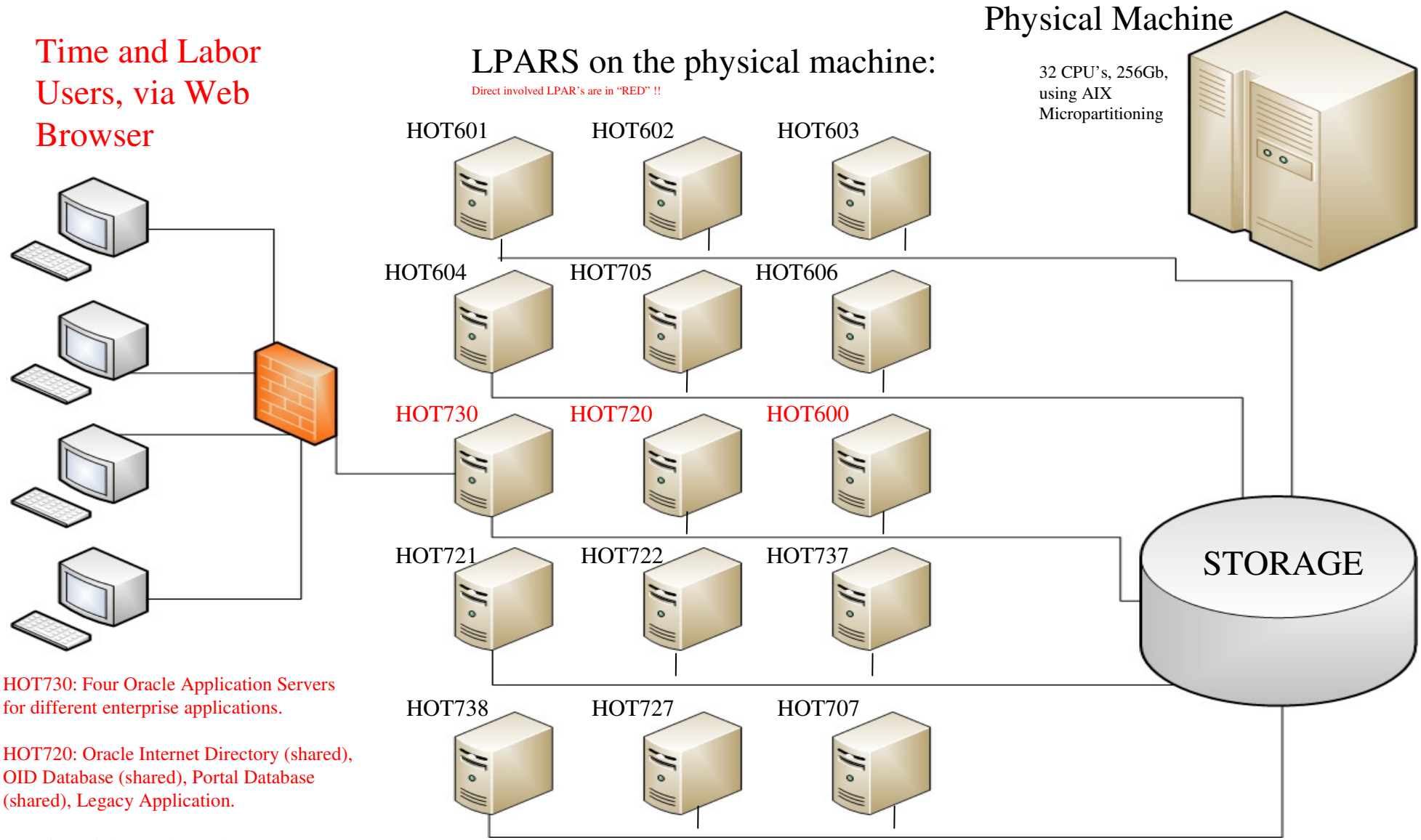
Gives the best fit, determined by data mining…

The Case…

- Response time of a very important business process in a time and labor system has sometimes very unpredictable performance

- The Used system is complex and the direct involved LPAR machines are:
  - HOT600 running several databases.
  - HOT720 running Portal and Oracle Internet Directory
  - HOT730 running Application Server for Time and Labor System

GAPP In A Complex Architecture…

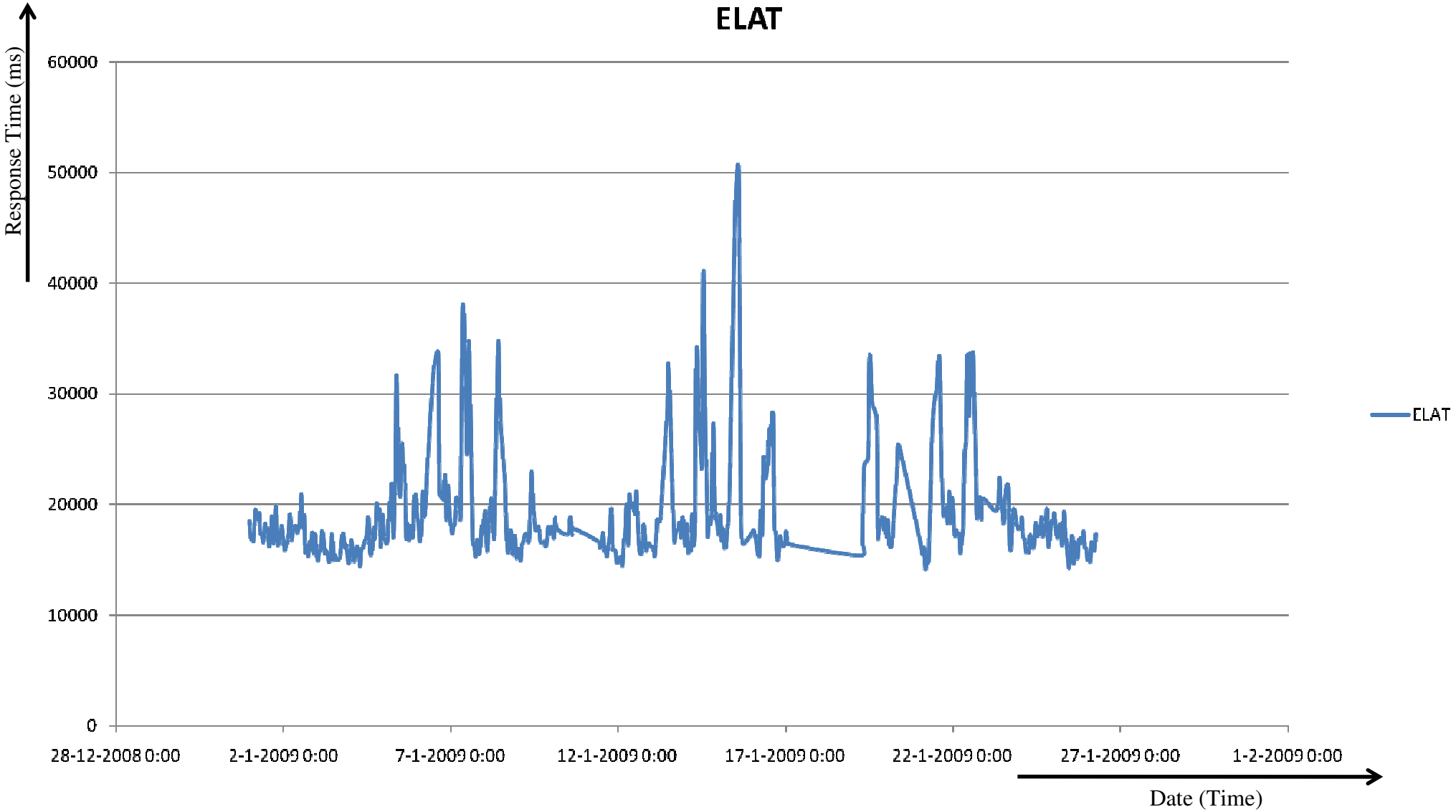


HOT730: Four Oracle Application Servers for different enterprise applications.

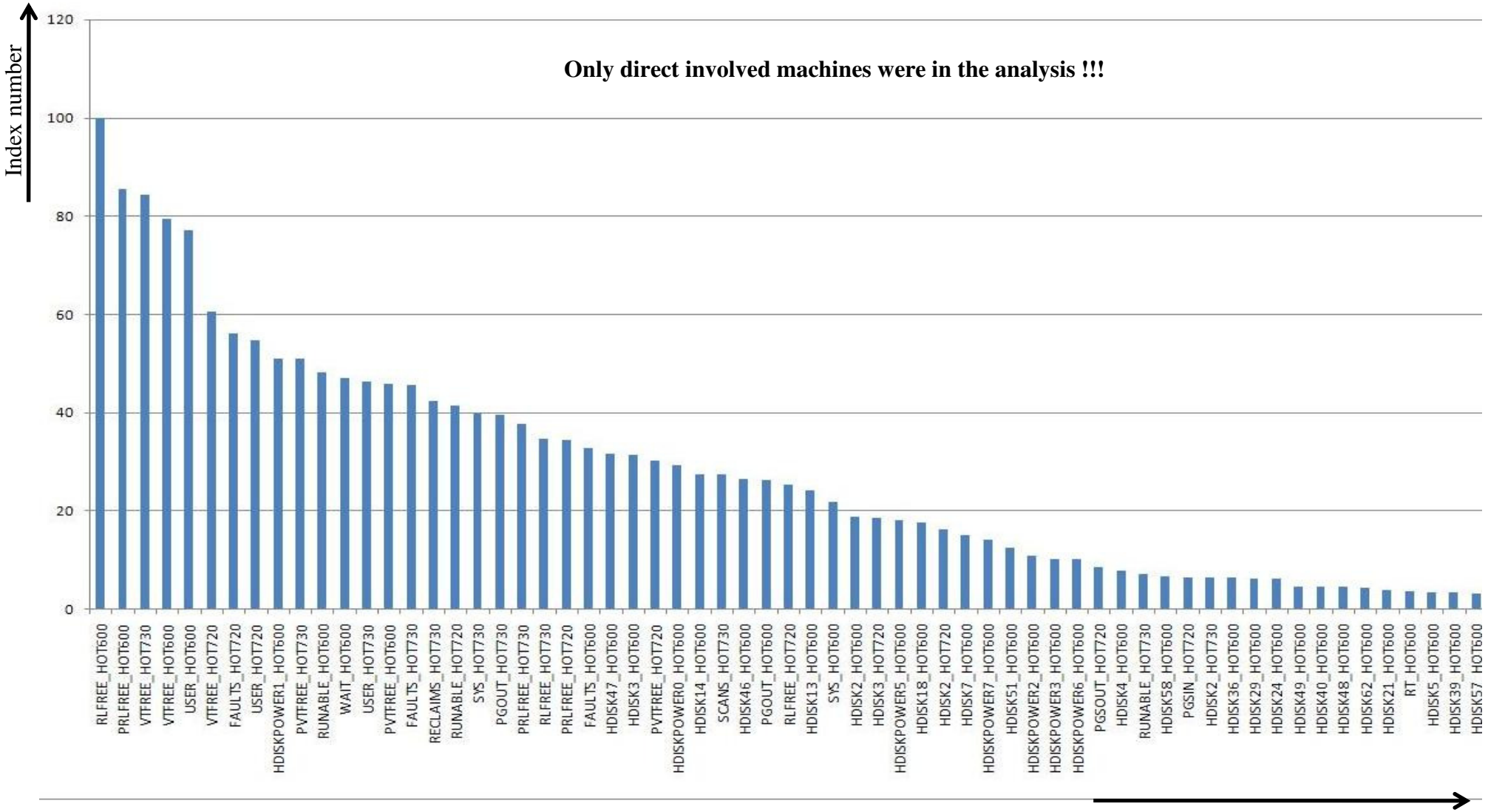
HOT720: Oracle Internet Directory (shared), OID Database (shared), Portal Database (shared), Legacy Application.

HOT600: Eight Oracle Databases

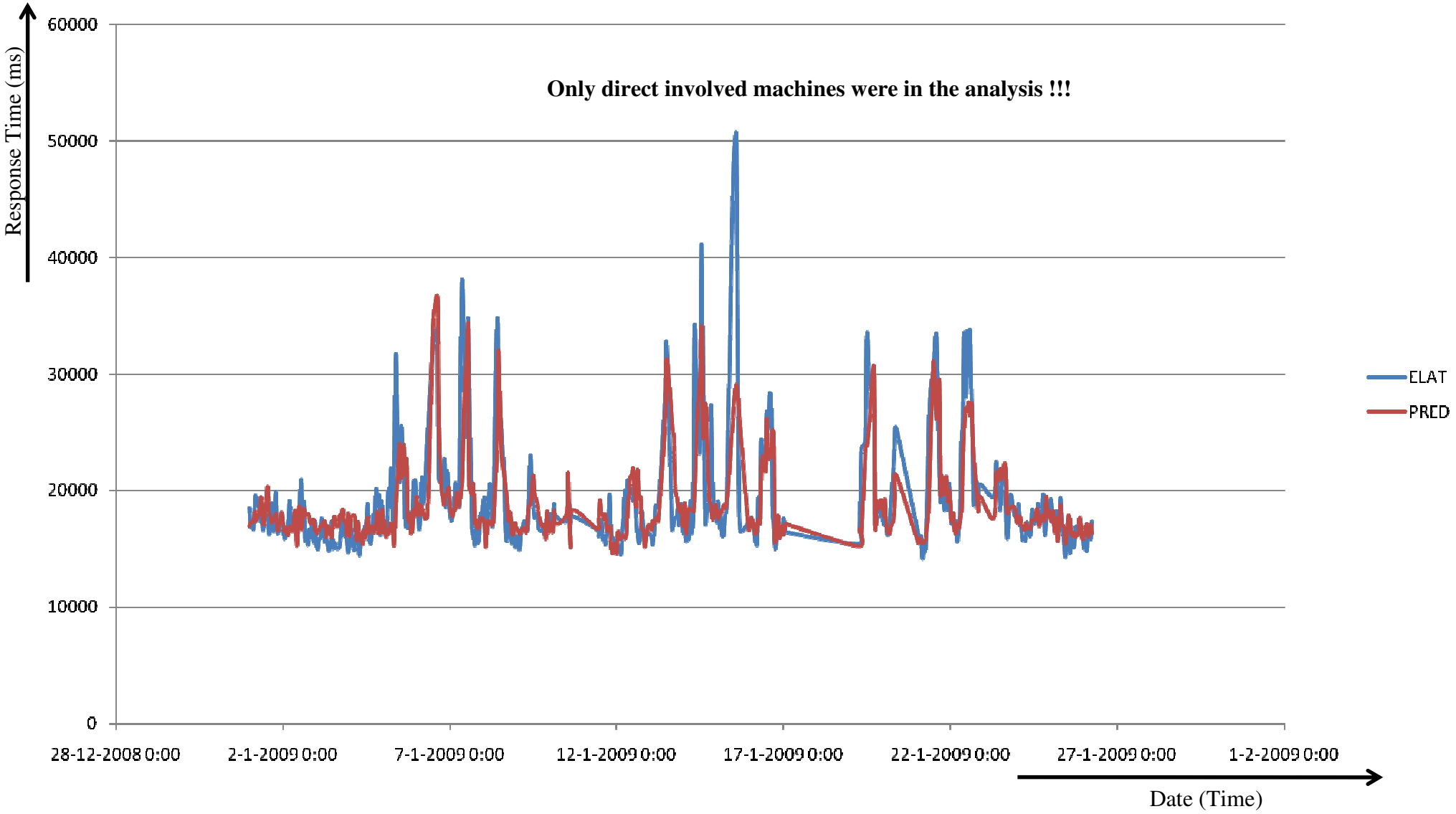
The Variance in Response Time…



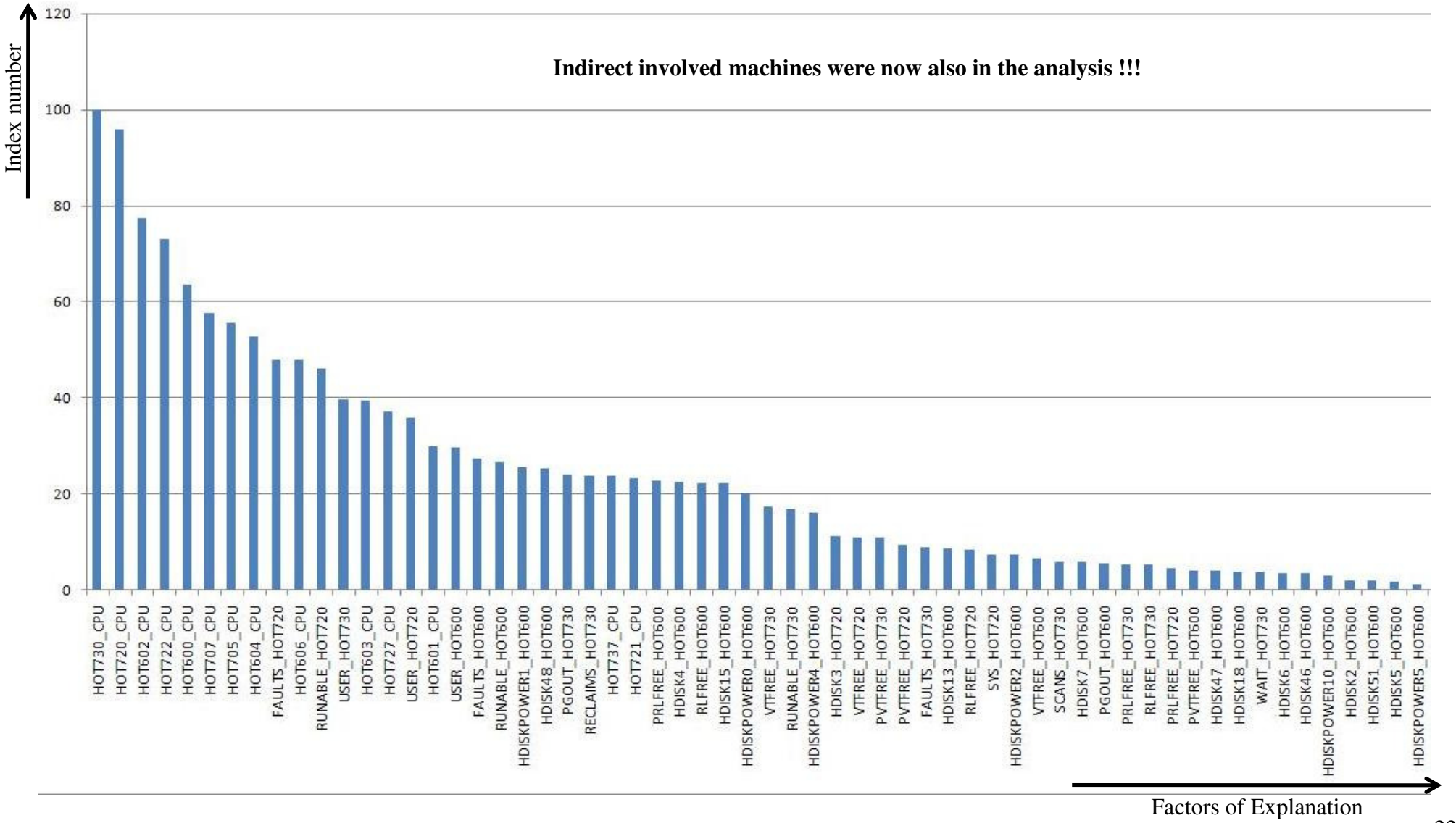
The Factorial Analyses (Explain)…



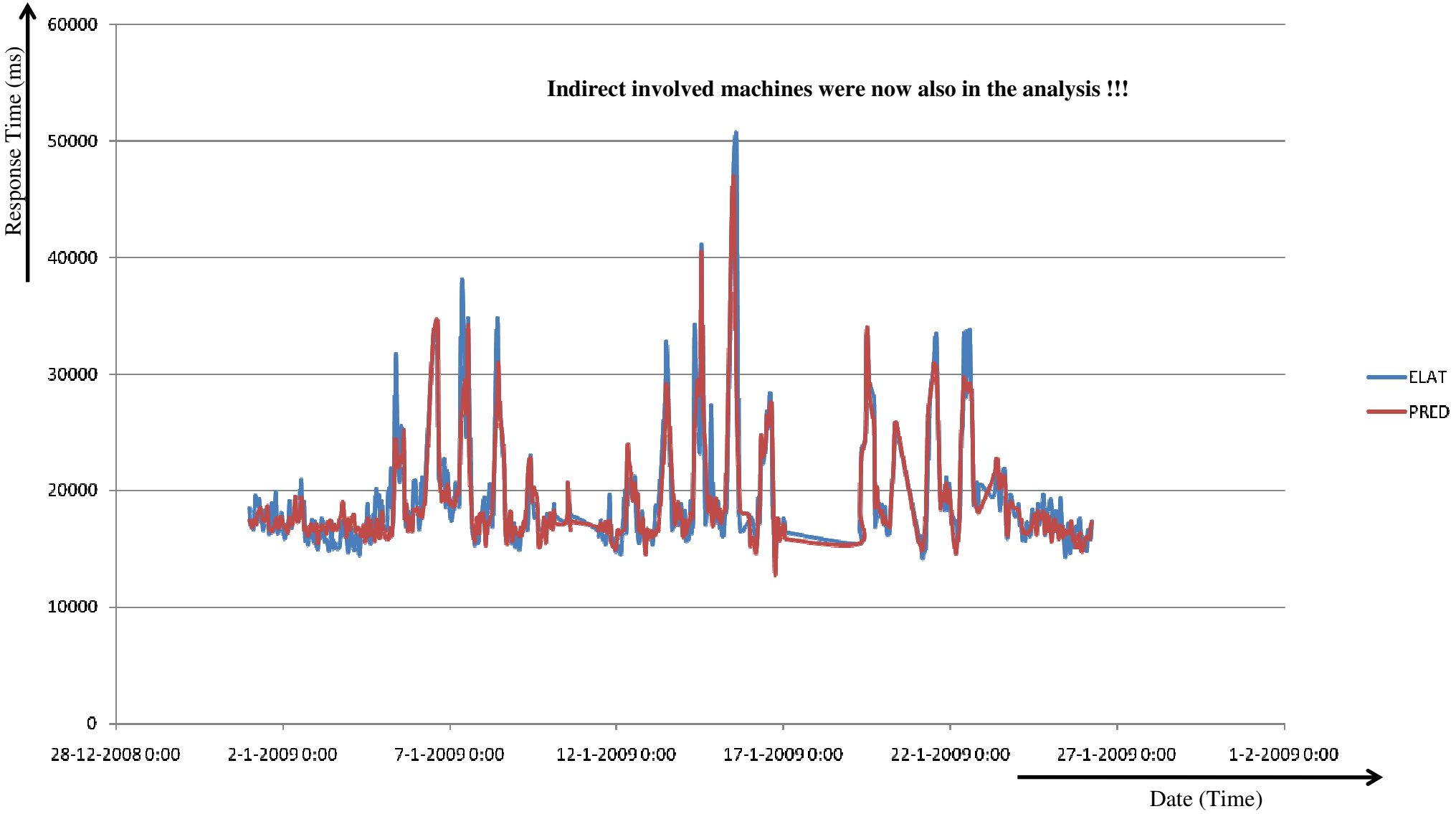
The Model (Predict)…



The Factorial Analyses (Explain)…



The Model (Predict)…



Interpretation…

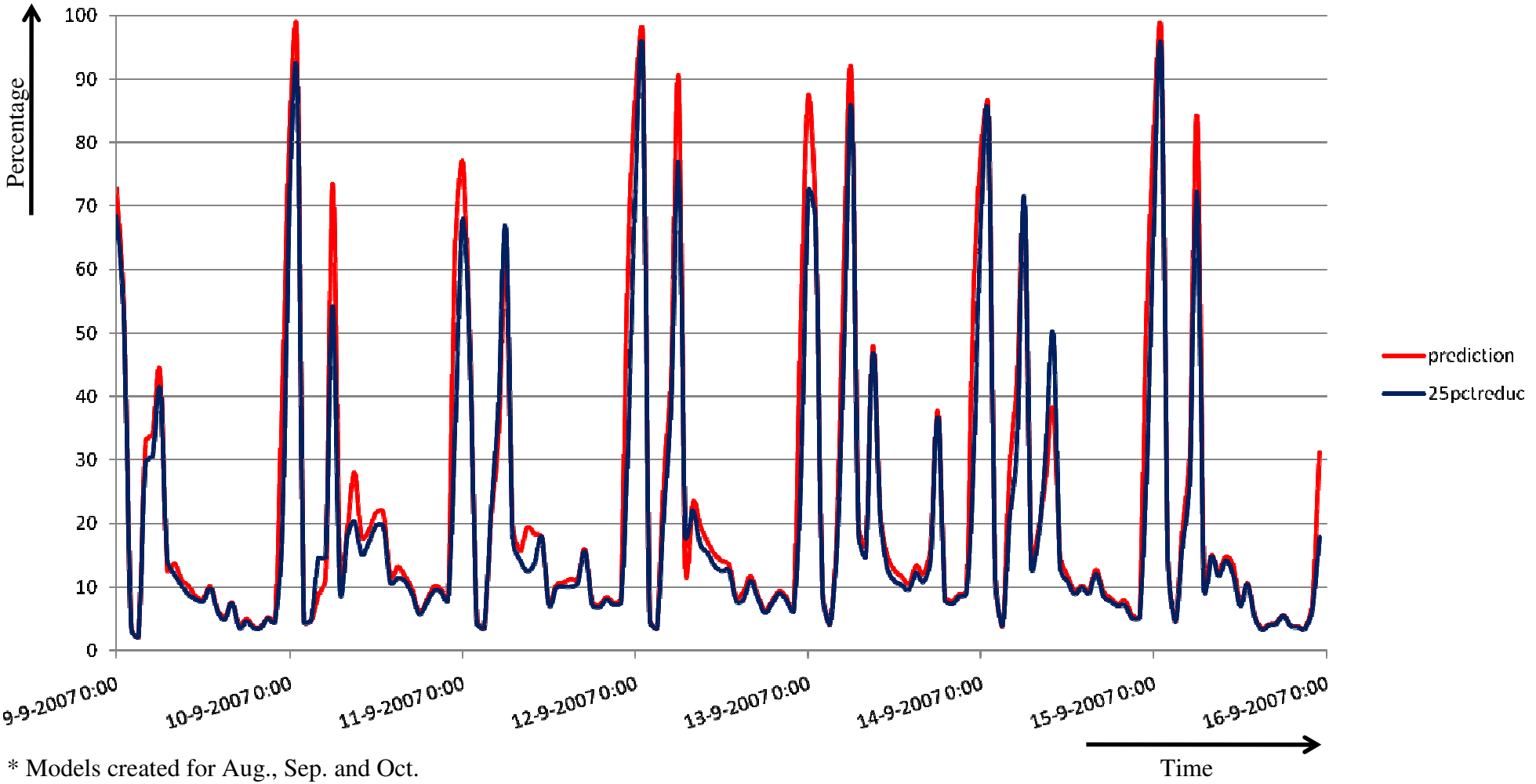
- Most of the peaks are very well explained:
  - The performance of the involved business process was very depending on the machine's overall resources.
  - Beside the HOT730 and HOT720, the HOT602 and HOT722 had also much impact.

- The GAPP analyses showed that at the moment the physical machine gets short on physical Cpu all the different machines will have high impact on each other due to competition in hardware resources.
  - By Workload Scheduling the biggest problems could be solved

Prediction

When 25% I/O Reduction…

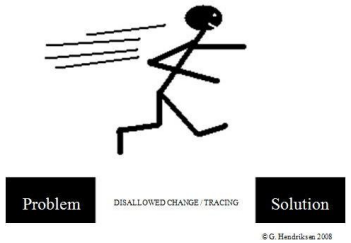
Percentage logins (PCT_GT_5SEC) which took longer than 5 seconds:



* Models created for Aug., Sep. and Oct.

Software used for analyses

What could you use for Data Gathering…



- Self-made API's
  - Very flexible
  - Needs investment
  - Nmon, Nimbus and moniforce API already build.

- Splunk
  - Based on Hadoop software
  - License fee per amount of data gathered.

splunk>

- Hadoop (Chukwa)
  - Open Source

What could you use for the mining…

- Oracle Data Mining (ODM)
  - Cost 15K per CPU above EE
  - Makes method-GAPP dependent on ORACLE
  - Data can be stored in Oracle DB and be mined

- Hadoop (Mahout):
  - Is Open Source
  - Makes method-GAPP completely platform independent
  - Needs Database (which could be open source)

Conclusions

Conclusions…

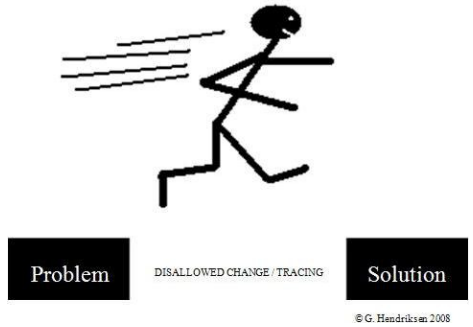
- Although you didn't made "hooks" in the application or installed agents you were able to find the problem, independent of technology. Nevertheless agents would be necessary, when used on continuous base.

- After your analyses was done you found out that also components in the infrastructure not part of the application were responsible for the encountered problem.

- The model you made, made it possible to do predictions of the impact of a possible investment on the end-user process performance.

Reference…

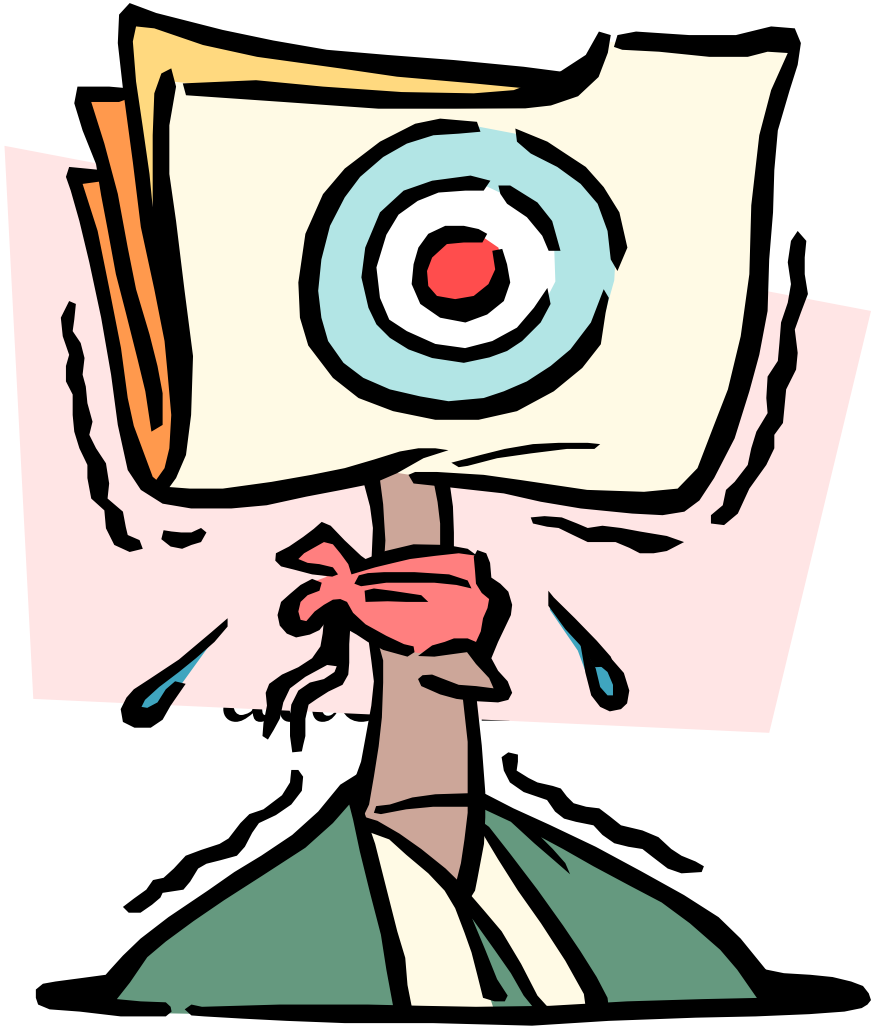
- Personal blog:
  - http://blog.gerwinhendriksen.com
  - http://method-gapp.com



- Method-R:
  - Oracle Performance by Cary Millsap (http://method-r.com/)

- Oracle Data Mining:
  - http://www.oracle.com/technology/products/bi/odm/index.html

- Hadoop
  - http://hadoop.apache.org/



Q/A