



Analysis of Large Scale Data Volumes

HighQSoft's Analysis Server "Merlin 2G"

ASAM US Workshop, Novi

Dr. Ralf Nörenberg
ralf.noerenberg@highqsoft.de





HighQSoft's Analysis Server "Merlin 2G"

Content

- 1 The basic ideas of Merlin
- 2 The basic ideas of an analysis
- 3 Layout of an analysis
- 4 Merlin as an infrastructure
- 5 Setup of Big Data Use Case II setup



The basic ideas of Merlin

Any test/measurement is subject to analysis

Domain-Experts lose time and resources developing duplicate analysis programs which also run next door.

The analysis programs

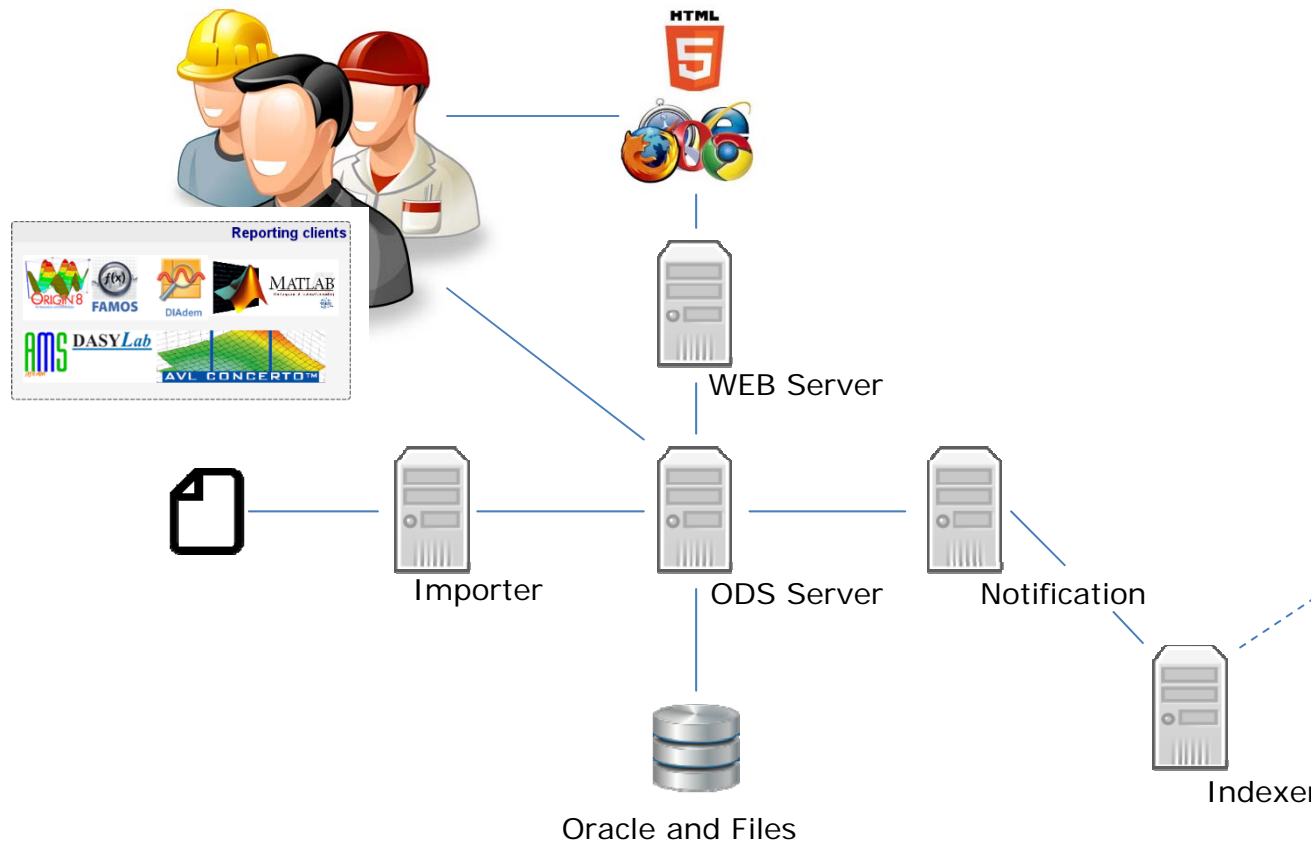
- contain a high level on domain specific know-how
- will only run once (locally) and can not be contributed to a automated process





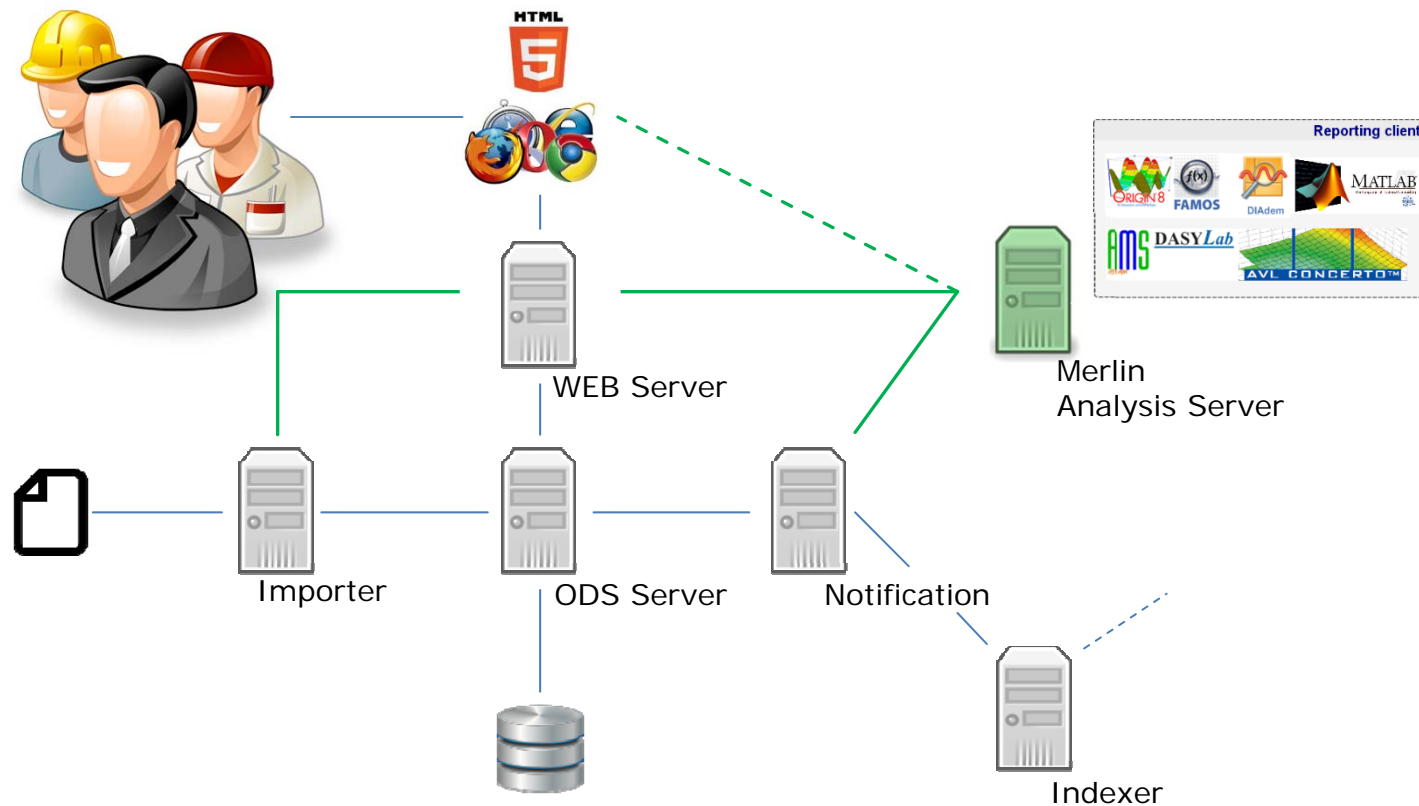
The basic ideas of Merlin

Basic Setup



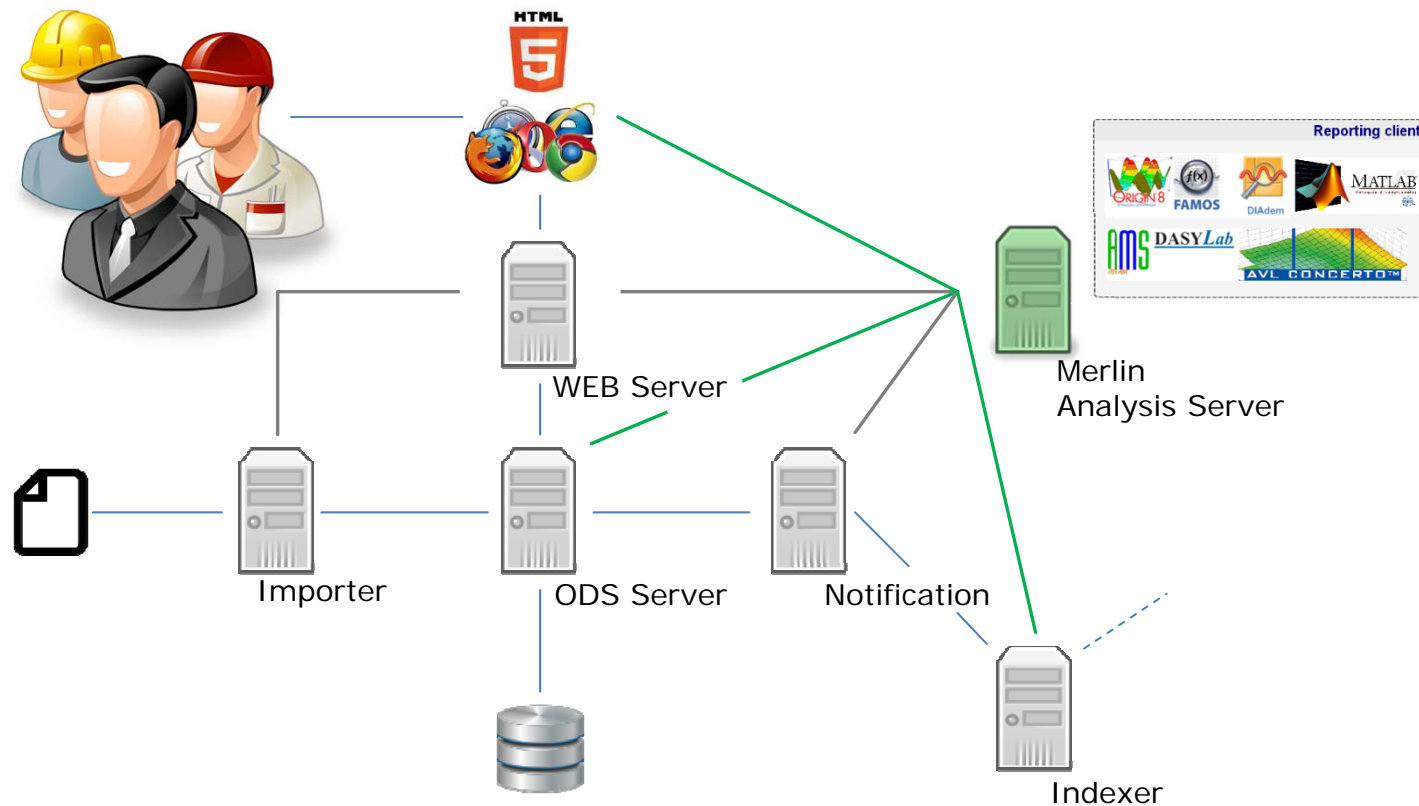


The basic ideas of Merlin Automated Analysis





The basic ideas of Merlin Automated Report / Results





HighQSoft's Analysis Server "Merlin 2G"

Content

- 1 The basic ideas of Merlin
- 2 The basic ideas of an analysis
- 3 Layout of an analysis
- 4 Merlin as an infrastructure
- 5 Setup of Big Data Use Case II setup



Basic ideas of an analysis

Analyses shall be small and manageable
and include domain know-how, no IT know-how



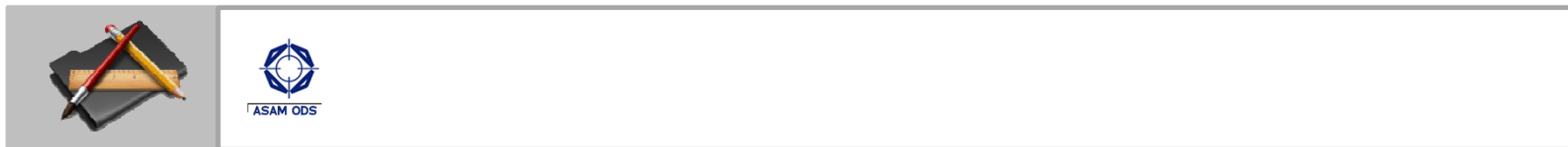


Basic ideas of an analysis

Analyses have read- and write access on ODS data

Read and write of meta- and mass-data

Calculation of missing attributes, consistency checks of data





Basic ideas of an analysis

Analyses may be affected by defined parameters

Example: Channel names, definition of scopes, arguments for calculations, internationalization

End-Users only have limited and defined possibilities of interaction





Basic ideas of an analysis

Analyses will deliver results based on data file MIME-types

Example: generated images, movies, sounds, PDF reports, ATF/XML

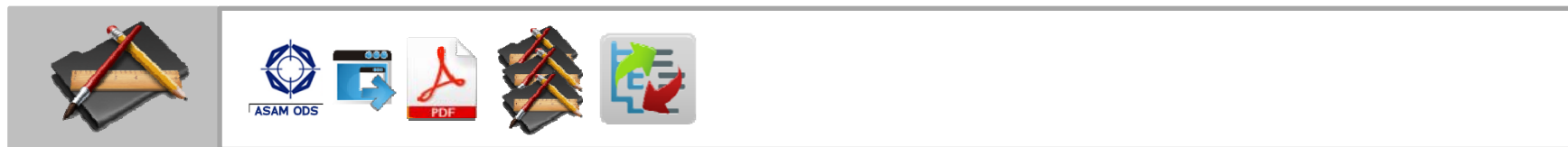
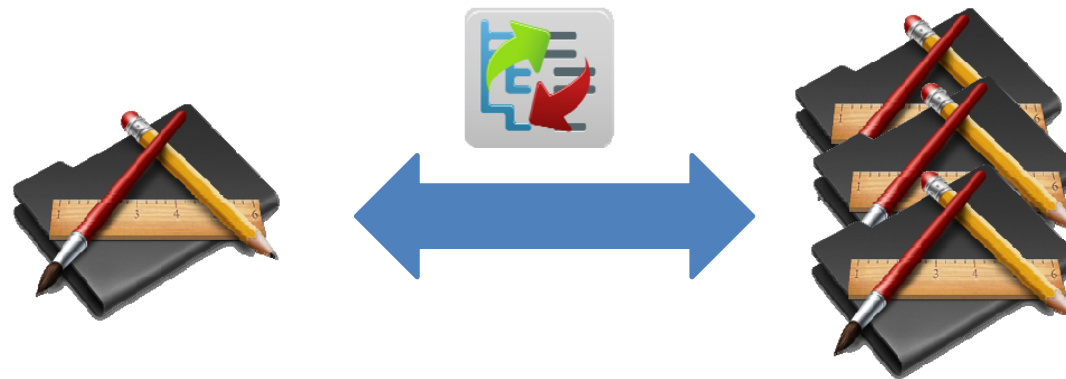




Basic ideas of an analysis

Analyses may use other analyses and are interchangeable

Example: „Statistics“ uses „Min“ and „Max“

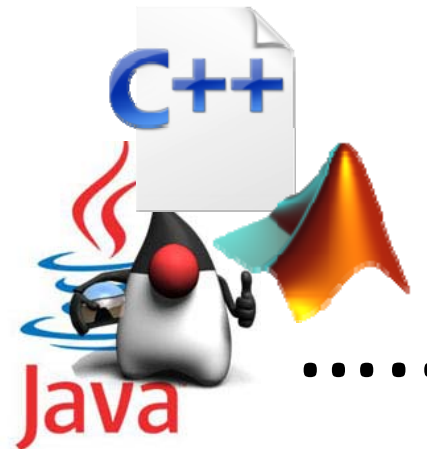




Basic ideas of an analysis

Analyses may be integrated in different languages

Examples: JAVA, MATLAB, DIAdem, C/C++, System Scripts, ...

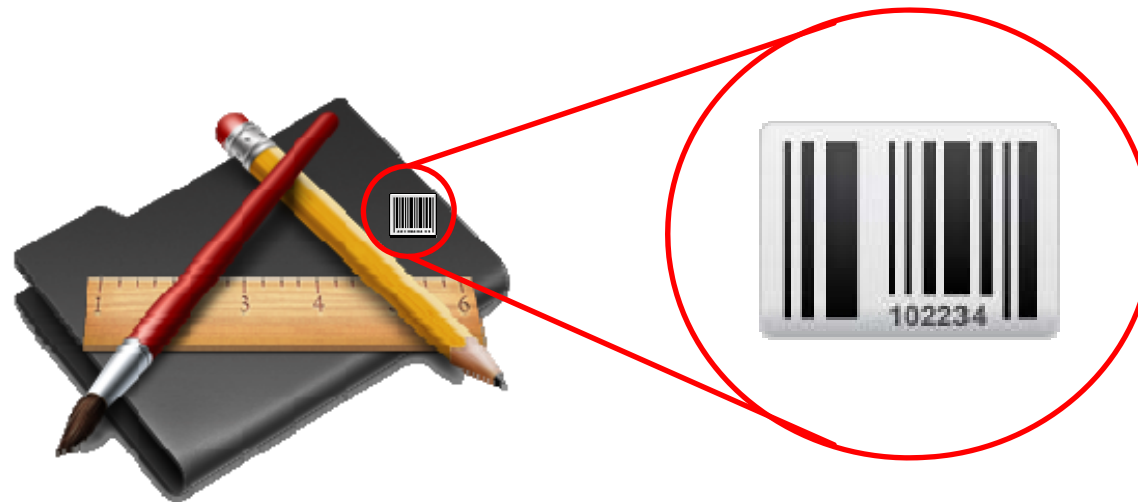




Basic ideas of an analysis

Analyses are to be understood as „standard analyses“

Example: Every measurement of certain test stands

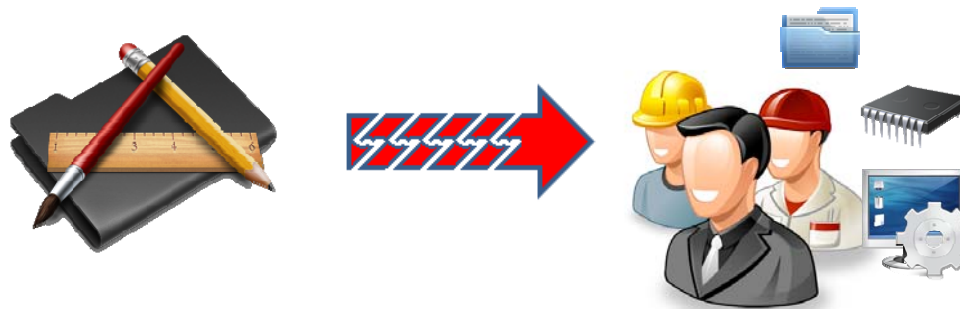




Analyses: **NOT** intended

Analyses having access on local resources of the end-users

Examples: local files, system settings, CPU or memory....

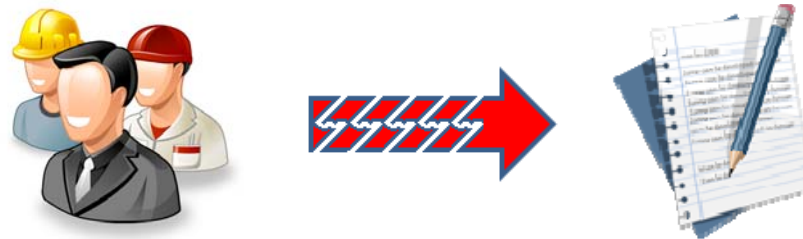




Analyses: **NOT** intended

End-users having access to the source-code or the scripts of analyses

Example: Editing of DIAdem scripts or MATLAB Code





HighQSoft's Analysis Server "Merlin 2G"

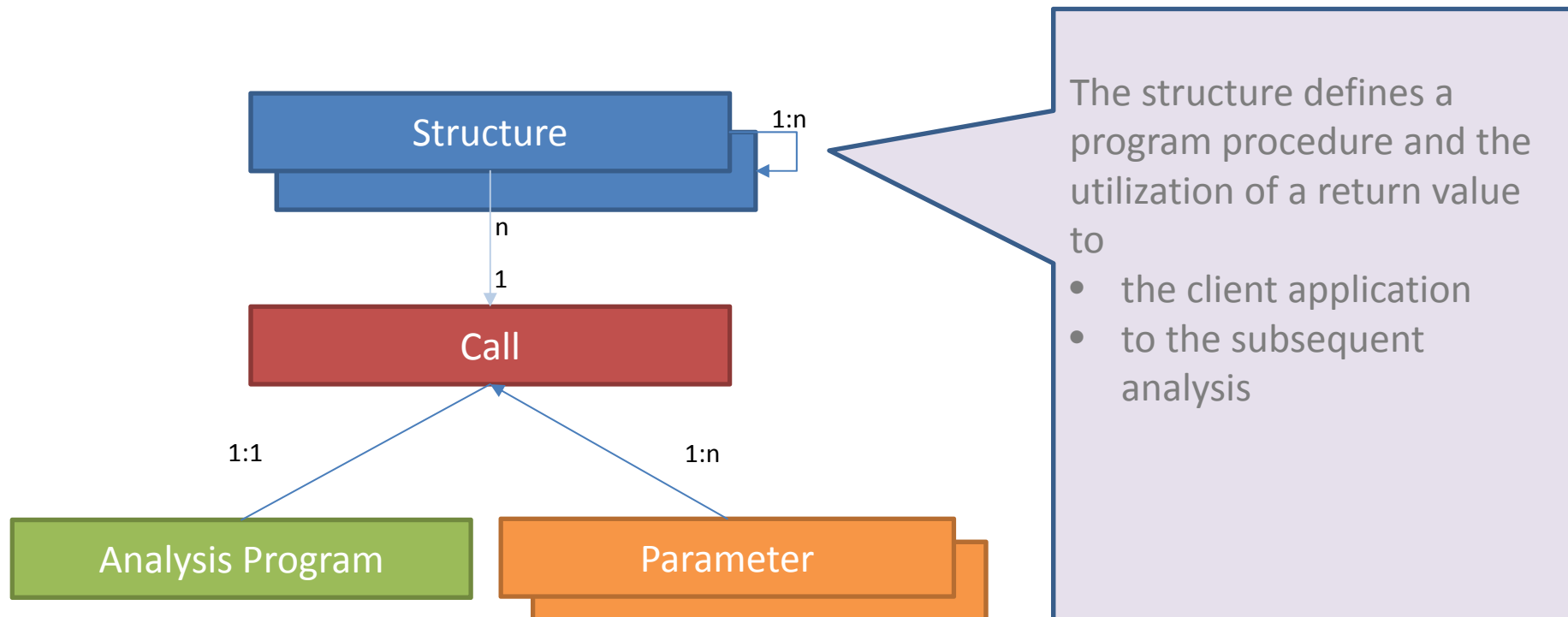
Content

- 1 The basic ideas of Merlin
- 2 The basic ideas of an analysis
- 3 Layout of an analysis
- 4 Merlin as an infrastructure
- 5 Setup of Data Use Case II



The layout of an analysis

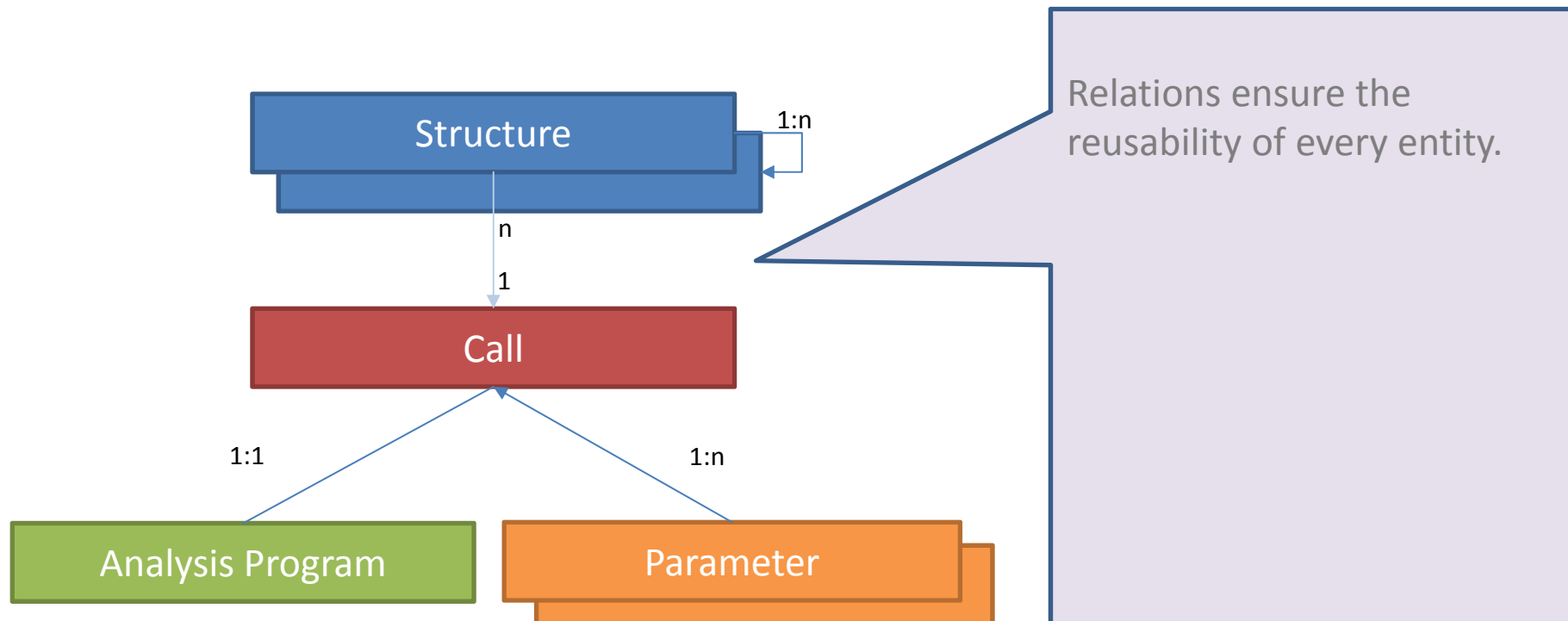
3-Level architecture





The layout of an analysis

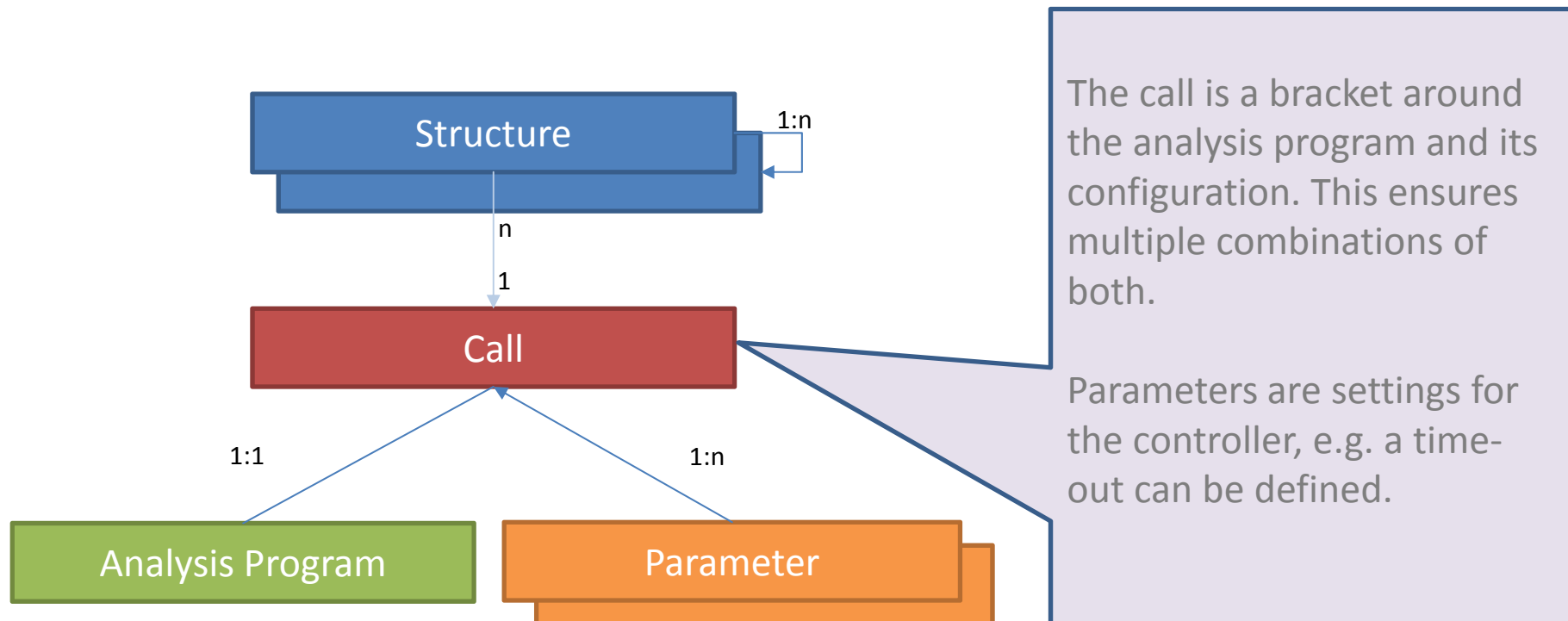
3-Level architecture





The layout of an analysis

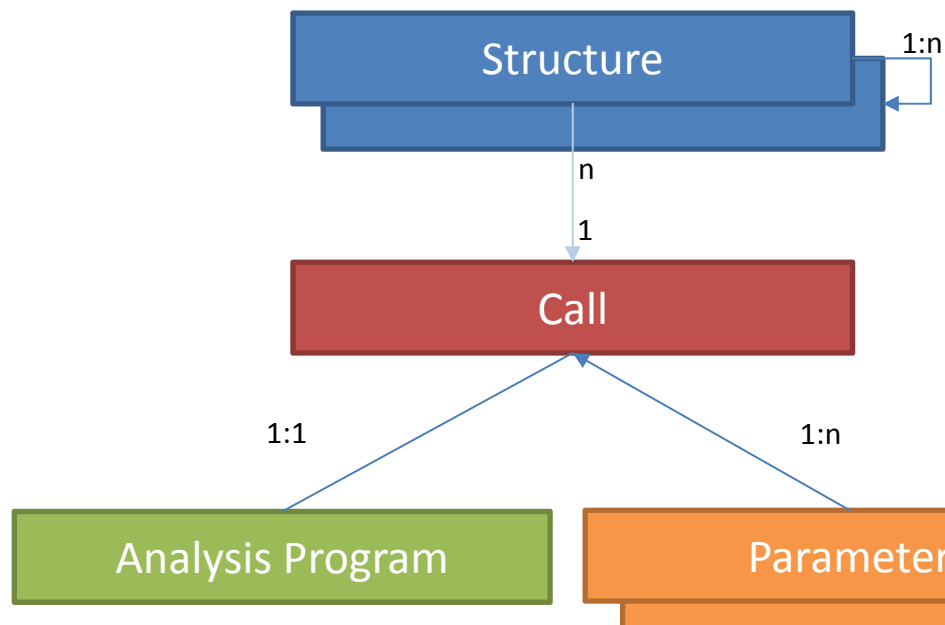
3-Level architecture





The layout of an analysis

3-Level architecture



Analysis programs may be combined with multiple parameters.

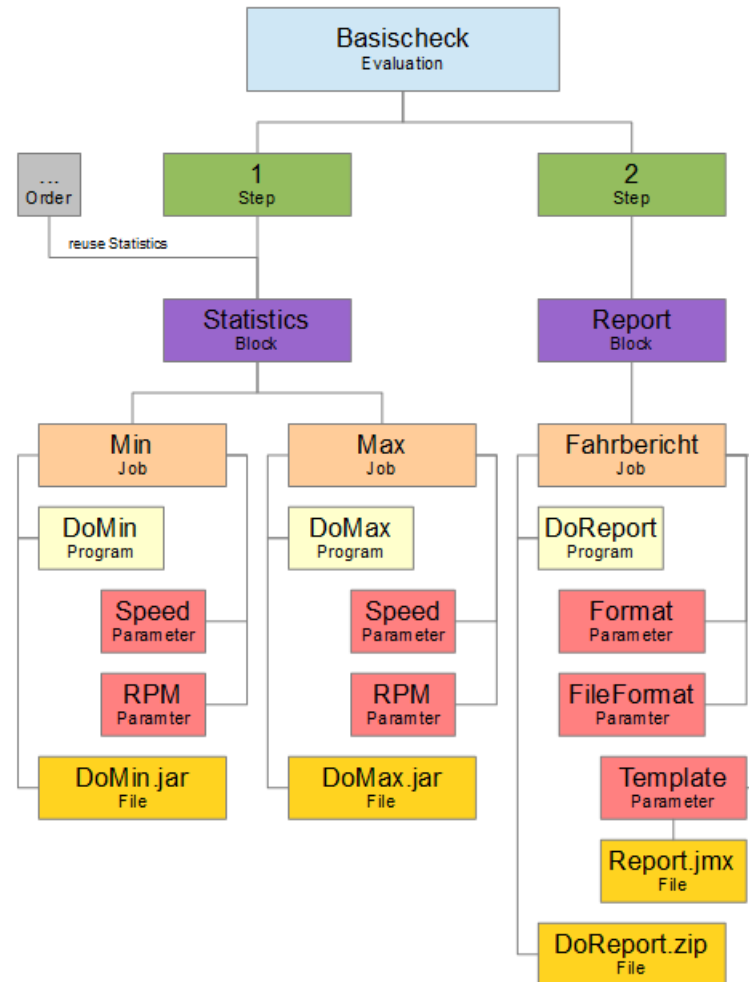
Parameters will be passed on from the Merlin Engine.

Parameters may contain complex information, e.g. be a XML file.



The layout of an analysis

Example





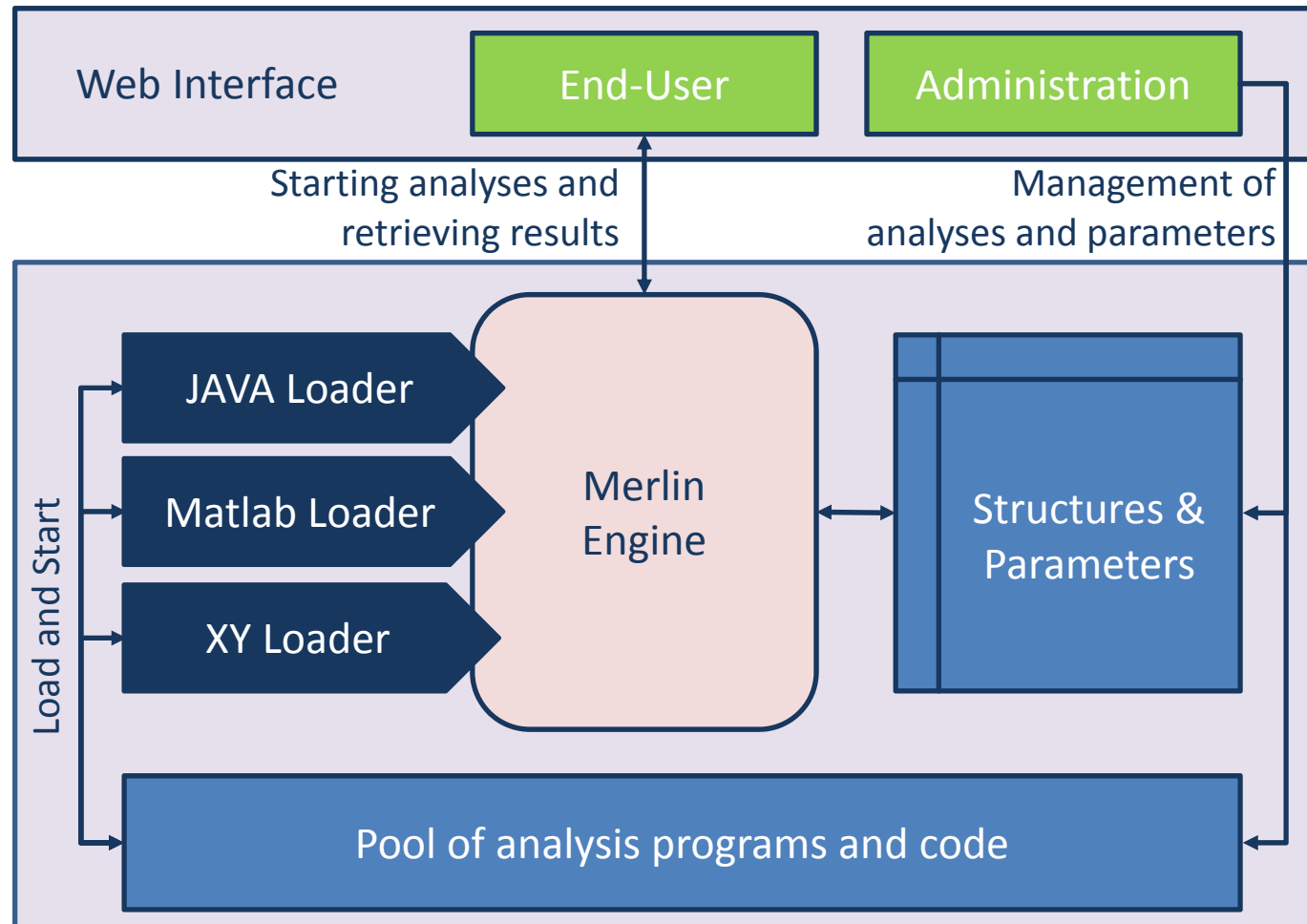
HighQSoft's Analysis Server "Merlin 2G"

Content

- 1 The basic ideas of Merlin
- 2 The basic ideas of an analysis
- 3 Layout of an analysis
- 4 Merlin as an infrastructure
- 5 Setup Big Data Use Case II

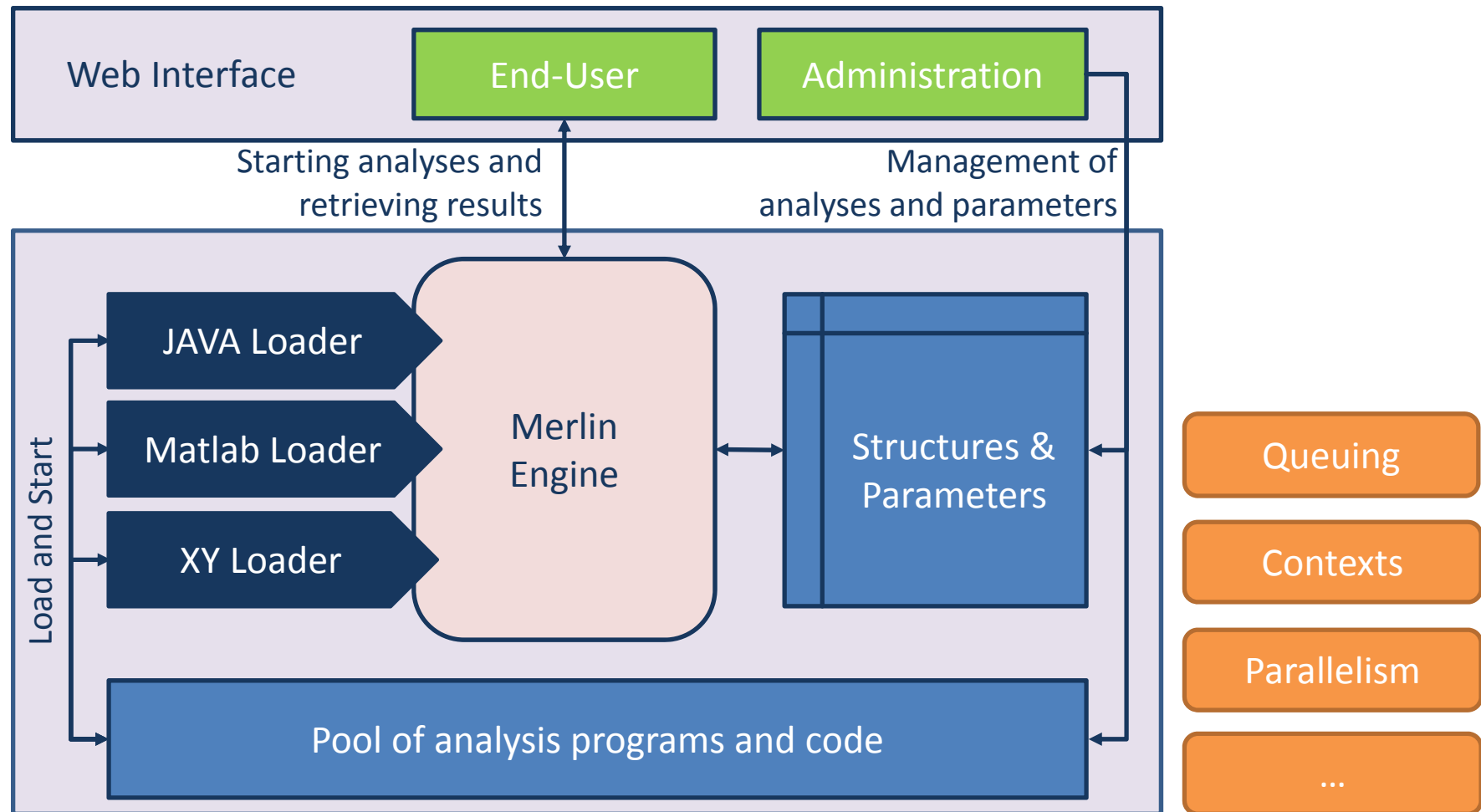


Merlin as an infrastructure





Merlin as an infrastructure

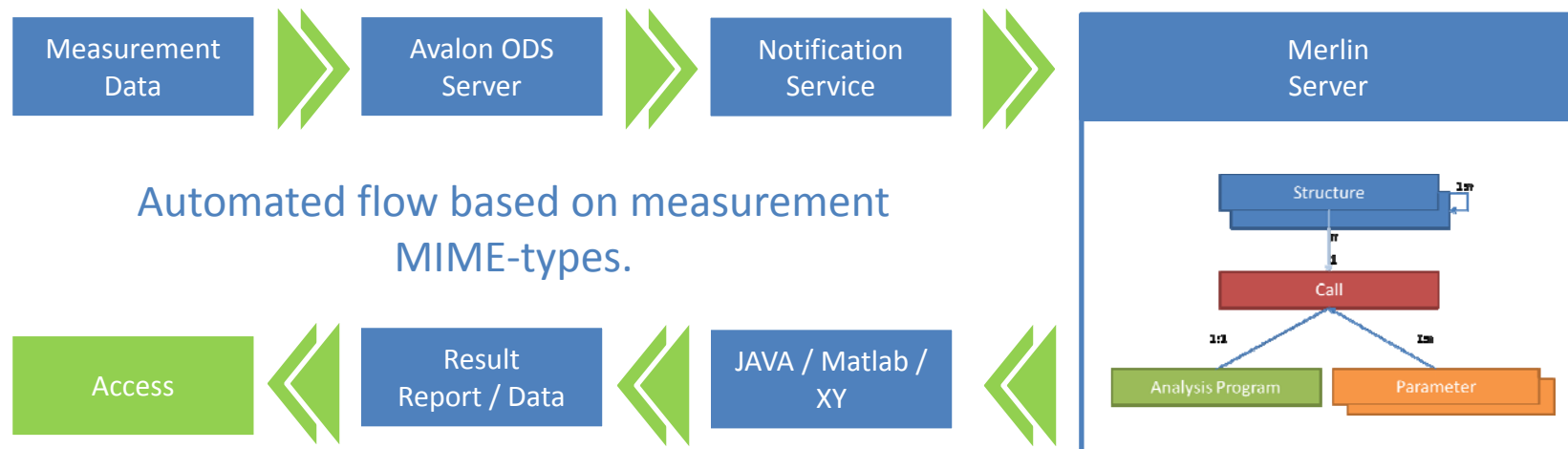




Merlin as an infrastructure

Integration of the whole process

- Analysis is a task that is standardized and strictly repetitive (>80%)
- Analysis can be done on-site, only results need to be forwarded

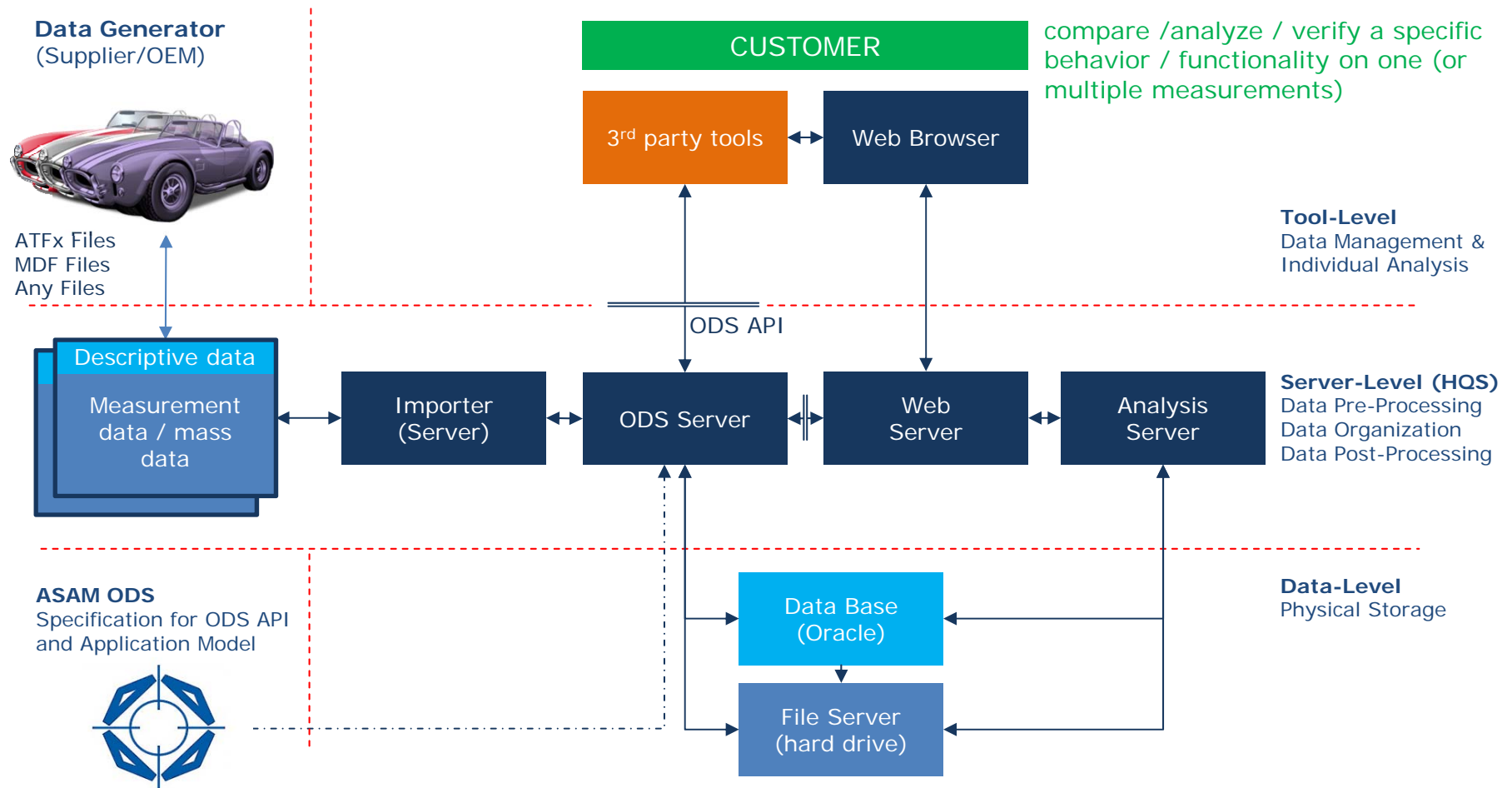


(Individual analysis by specialists will always require a download)



Merlin as an infrastructure

Integration of Merlin in a state-of-the-art setup





HighQSoft's Analysis Server "Merlin 2G"

Content

- 1 The basic ideas of Merlin
- 2 The basic ideas of an analysis
- 3 Layout of an analysis
- 4 Merlin as an infrastructure
- 5 Setup of Big Data Use Case II setup



Big Data Use Case II setup

Vehicle Fleet Testing (planned tests)

This **road-load data project** is set up in an environment of a car manufacture who has all vehicle fleet testing data (and most activities) centralized within one project.

The scope of this projects contains

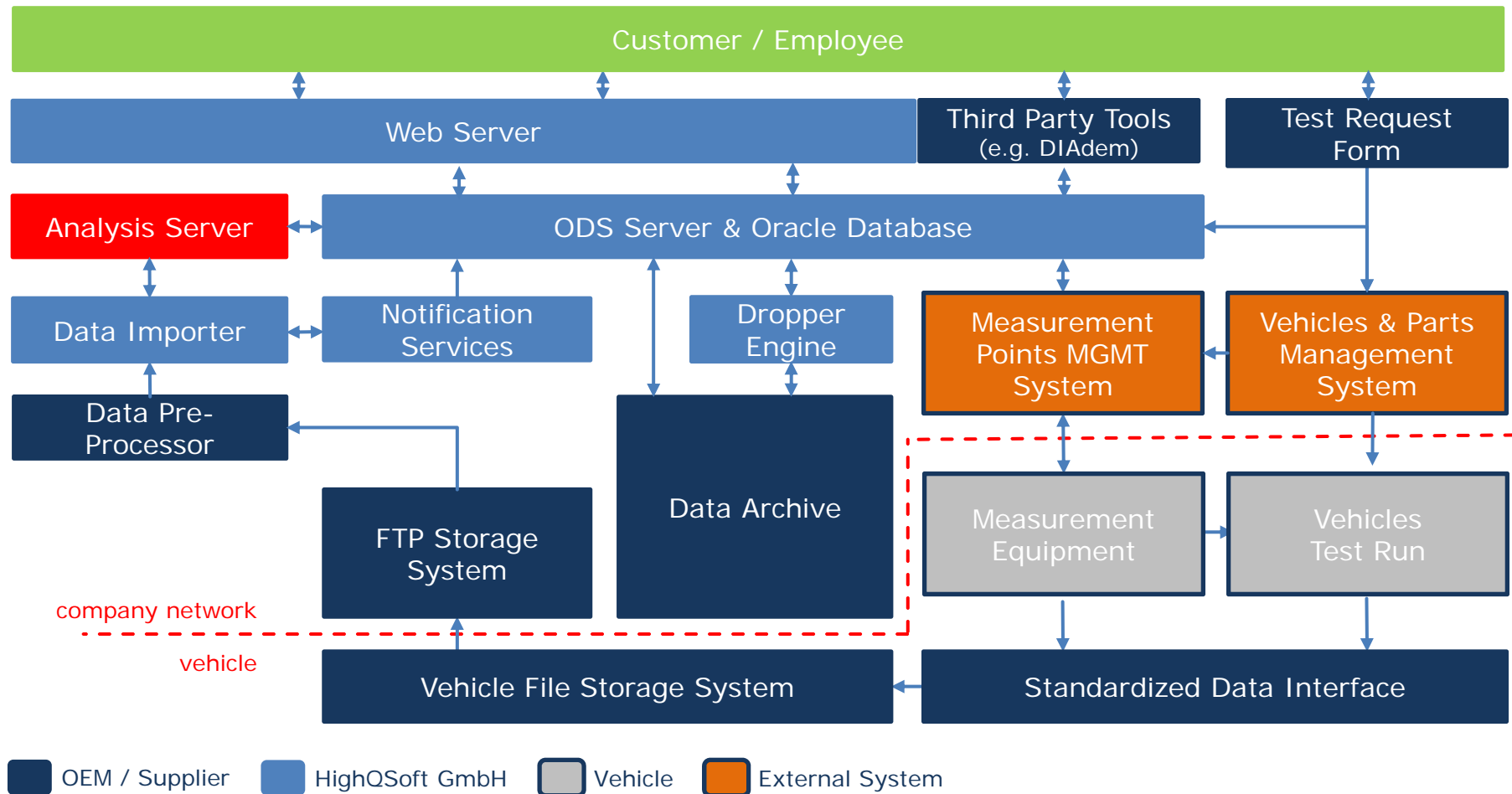
- many customers (all internal development departments)
- all passenger vehicle types and almost all ECUs available
- all domains and data types
- over **500 testing vehicles** to be managed

Within ODS, this is the upmost state-of-the-art system to manage the test execution flow and standardized measurement data.



Big Data Use Case II setup

Overall System Architecture (detailed)





Big Data Use Case II setup In Numbers

• Data Base Size (Oracle)	900 GB
• Files on File Server(s)	
• Files	39.346.529 Files
• Volume	5.400 GB
• Absolute growth	1.5-1.8 GB per day
• Files in tape archive	
• Files	58.462.941 Files
• Volume	7.700 GB
• Analysis Servers	
• No. of Servers	1 in total
• No. of Analyses	210.000 per month
• ODS Servers	
• Productive ODS Servers	3 in total

► Project is state-of-the-art and meets performance requirements



Thank you!

Any questions, suggestions and ideas?

HighQSoft GmbH
ralf.noerenberg@highqsoft.de