Paper AD10

How to organize your SCE to digest Big Data related to Digital?

Alain Vasseur, Sanofi, Chilly-Mazarin, France
Christelle Le Beaudour, Sanofi, Chilly-Mazarin, France

ABSTRACT

Since 2006, WISE is the clinical statistical computing environment of Sanofi Pharma mainly uses to prepare and compute statistical study analysis for submission. SAS® was the unique software used in this purpose.

INTRODUCTION

When in 2012, the Sanofi strategy has been to unify all the clinical information systems between Sanofi, Genzyme and Pasteur to have unique CTMS, CDMS, eTMF, and SCE, WISE had been identify has the target for the clinic within the Sanofi group. Genzyme and Pasteur joined the WISE community. Of course it has been necessary to re-design the application to onboard new users, new needs and new constraints related to health authorities and to the increasing workload. In addition to the user's needs, we decided to use WISE Version 3 project to prepare the future and to follow the new SAS roadmap by replacing the SAS® program editor within deploying SAS® enterprise guide. This change was the first step of a global roadmap to prepare the future of our WISE platform.
Since 2013, WISE version 3 is the unique Sanofi SAS® clinical platform on which all clinical analysis are done.

WISE 4 PROJECT

Maintaining this kind of platform up to date is requiring a continuing effort in order to keep an acceptable level of performance, quality security and alignment with the new rules and request.
New trends are forcing us to prepare the foundation to our future and to be able to provide an evolutionary service to WISE consumers. We continuously need to adapt our system to the evolution of technology but as well to the trends in term of data volume, variability and velocity, the famous three V of big data. Moreover, Data are recognized as the new Eldorado, and it is becoming obvious in our industry that if our internal data are very rich it is as well obvious and very important to provide our analysts access to external data (Such as Hospital, social network…). In parallel, we can observe that analytic is consumed by many new user profiles. The consequence is that SAS® is no more the unique tool use for analytic and statistics in clinic. R, Python, Scala are competitive language communally used to support analytics needs around exploration, simulation, prediction or modelization. By continuing, to improve our visualization capabilities we also provide more friendly tools to our user community, R Studio, SPOTFIRE, SAS Visual Analytics or Qlikview. Our current architecture does not allow us to follow this evolution.
We also need to adapt our strategy to new user profile that have analytics agile requirement. Moreover, used traditionally by statisticians and programmers, the needs around the platform have been extended to clinical data manager in order to simplify the data quality process and accelerate the submission dossier preparations. The sourcing strategy required, as well, to revisit our security model to be able to ensure the access to the platform to external partners with limited data transfer.
The increase usage, the increase volume of data, the increase number of users and the increase of constraints are creating more and more slowness and more difficulties to support the needs. We perfectly know that at the end of 2018 we will be above the WISE 3 limits.
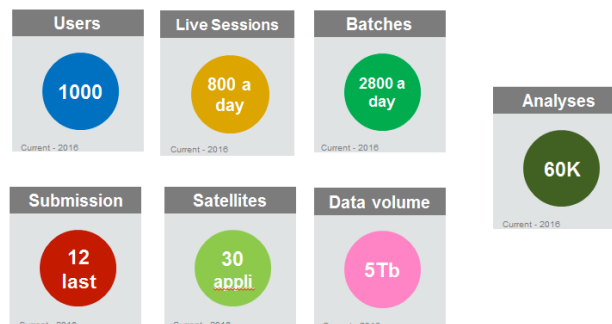
In 2016, we conduct an analysis of our platform in order to be able to ensure the support of clinics appropriately. The conclusion was that an improvement of our current infrastructure will be very expensive without the possibility to

endorse new analytics facilities (R, Python, Spotfire…), without providing us scalability and without the possibility to secure data access without a very painful effort. The analysis put as evidence the fact that to move forward we need to change totally our infrastructure.

It is why from March to September 2016 we conducted a proof of concept on the SAS® grid technology in partnership with SAS® and HPE to assess what will be possible to tackle with this approach.

Our major objectives during this proof of concept were to ensure that a new architecture based on SAS® Grid technology will be able to provide us a good scalability, an easy way to prioritize clinical analysis, the capability to add new open source language on the platform as well of course as better performances with a good control of the cost. We decided has well to test a new storage technology to improve performances regarding the fact that with the current technology we met some I/O issues.
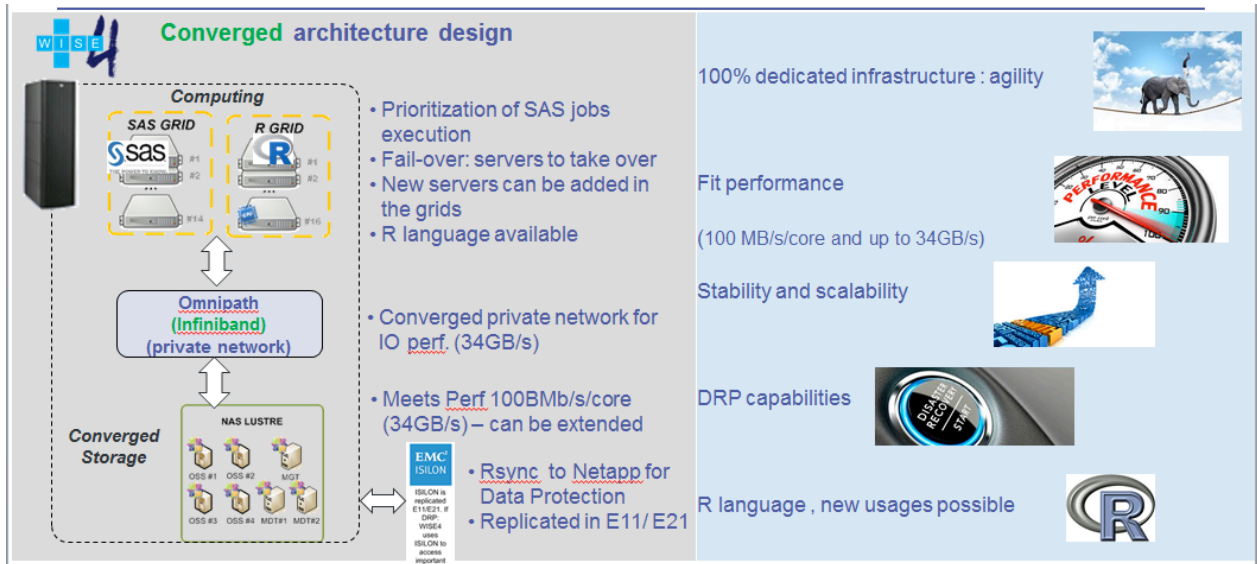
The presentation objective is to present you the result of this Proof of concept, the project initiated and the future of WISE generation 4.

You will understand why we think that the GRID but not only the SAS® Grid is our future.

CONCLUSION

In conclusion, The SAS GRID solution has been adopted in order to tackle current issues but as well to prepare the ground for Sanofi SCE new ambition and Big data Strategy. The idea is to provide a unique platform to access data, internal and potentially external. This technology will allow us to be agile and to increase our analytic capabilities in quick mode.

Contact Information

Alain Vasseur (In case a reader wants to get in touch with you, please put your contact information at the end of the paper.)

Your comments and questions are valued and encouraged. Contact the author at:

Alain Vasseur
Sanofi
1 avenue Pierre Brossolette
Chilly Mazarin 91385
+33 1 60 49 63 30
Alain.Vasseur@sanofi.com
Christelle Le Beaudour
Sanofi
1 avenue Pierre Brossolette
Chilly Mazarin 91385
+33 1 60 49 63 49
Christelle.le-Beaudour@sanofi.com

Brand and product names are trademarks of their respective companies.