



Forum **TERATEC** 24

Unlock the future

Migration of an HPC platform to the cloud

Thierry.Porcher@doit-now.tech



Do IT Now in a nutshell

SME engaged with the EMEA HPC community

- 100% independent capital (Self-funded company)
 - +120 HPC experts at your service
 - + 30 years of experience in the EMEA HPC Market
-
- Teratec Member
 - ETP4HPC Full Member
 - EuroHPC JU projects
 - ISC Bronze sponsor
 - Teratec Forum Gold sponsor
 - EESSI Collaborator



Our Offices - EMEA

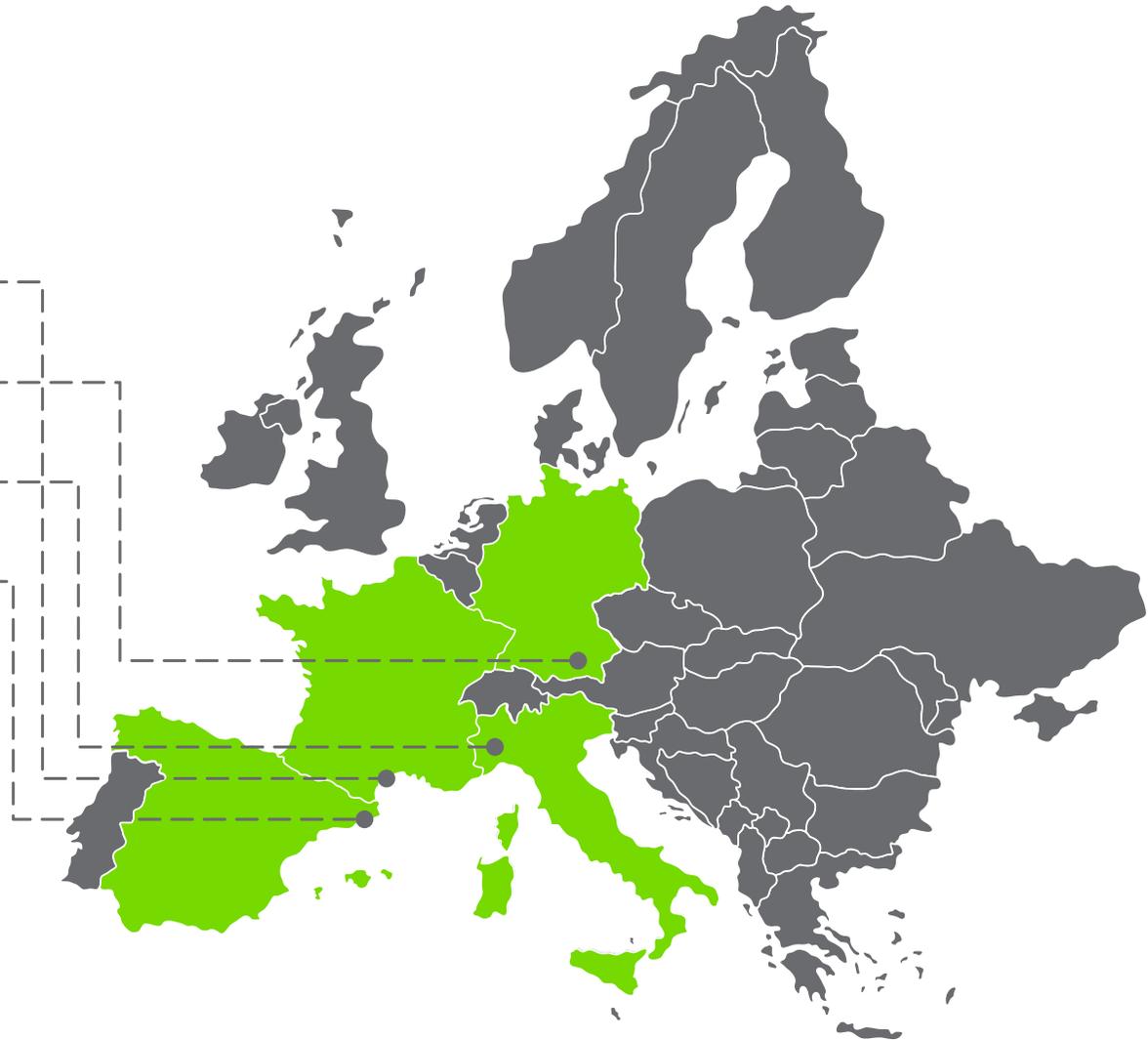
France - Montpellier

Germany - Munich

Italy - Turin

Spain - Barcelona

(APAC) New Zealand - Auckland



WISE : Worldwide Integrated Statistical Environment

- Big pharmaceutical company
- For R&D (clinic and non-clinic), CMO, IA (vaccines batches releases)
- Critical GxP Platform used for submissions to Health Authorities
- 6000 users (APAC, EMEA, NALA)
- 300-600 servers, 50 DBs, 10+ Lustre file systems



Migration to the cloud 3/2021-2/2022

- « Lift and shift » strategy (see next slide)
- 6 steps methodology
 - Audit
 - PoC -> pick the right technologies
 - Architecture design
 - Development and deployment
 - Test and validation
 - Data Migration
 - > GO LIVE

Lift and Shift ... nice concept!

Impossible to deploy obsolete technologies (unsupported or unavailable)

Network is managed by a dedicated team

Cybersecurity is managed by a dedicated team

AD is managed by a dedicated team

Admin role is managed by a dedicated team

⇒ It takes time to get things done!

Some software components had to be adapted for the cloud

Ex. Bright Computing Manager -> ISV had to be involved.

Opportunity to fix on-premises architecture limitations (single point of failures, ...)=> architecture changes

Audit phase

Understanding the on-premises platform

Identify network flows as you are going from a trusted internal network to an untrusted network

Collecting all software components versions and dependencies

- ➔ Opportunity to upgrade some components
- ➔ Fix architecture limitations of the initial platform (duplicate services + load balancer to improve robustness)

PoCs

Some examples:

- WEB infrastructure -> Kubernetes (EKS)
- File System selection: benchmarks
- BD migration: from Oracle to PostgreSQL
- VDI: Citrix or not Citrix?

Recommandations:

- => Ask advice to your cloud vendor!
- => Leverage migration tools provided by the cloud vendors (DB)
- => New solutions/services/options are available every week!

Architecture design

Infrastructure as Code: Terraform + Ansible

5 environments: DEV, INT, TEST, PROD et DR

Evaluate carefully the usage of managed services in case you want to keep open the possibility to move to another cloud provider.

- EC2, FSx for Lustre, S3, EFS, EKS, RDS, Lambda, ...

Monitoring: prometheus+grafana

Data Migration

About 200 TB of data to migrate from the data center (near Paris) to AWS

Need to provide a proof that ALL files have been transferred, that metadata has been preserved (ownership, permissions, date, ...) and content is preserved as well (checksum)

Challenge:

- 100 millions files
- Minimize service interruption (< 2 days)

Solution:

- Development of a set of tools to regroup and compress files in parallel
- Initial transfer to AWS with a Snowball (NAS that you send to AWS via UPS)
- Daily updates with rsync until GO LIVE weekend
- During the GO LIVE weekend: list files on the 2 platforms (on-prem and cloud) and compare

Results

- No major interruption of services in 2 years!
- Number of users has increased
- Users do not want any service interruption ...

New features

- Autoscaling everywhere when possible (cost control, more power)
- Kubernetes for new applications or services (Vscode, RStudio, ...)
- SAS Grid -> SAS Viya (Kubernetes)
 - Improve stability

Some challenges

- Cost optimization
 - Never ending process (new instant types, new services, ...)
 - Autoscaling also for data storage (Lustre to be replaced?)
- Resources for WISE have to be located in Europe (regulatory constraint)
 - Not in our current region
 - ⇒ Plan for multi-region from the start!
- No service interruption ever
 - Including patches, upgrades,

Conclusion

- For WISE, the cloud has been a huge opportunity
 - Increase the computing power for the most demanding users
 - Introduce flexibility in term of instance types (memory, CPU,)
 - Improve availability in a significative way
 - Ease the migration on Kubernetes for some tools (SAS, ...)