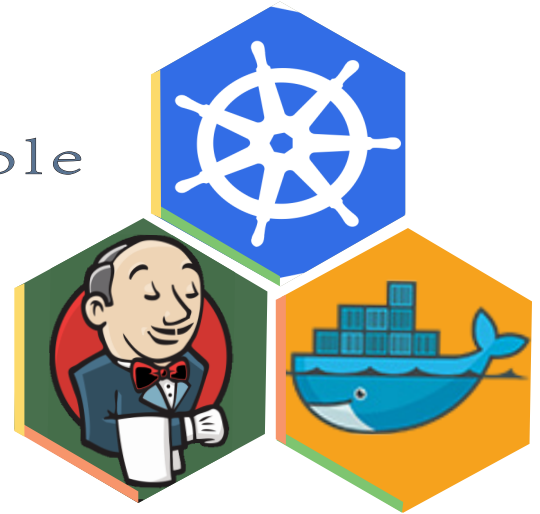
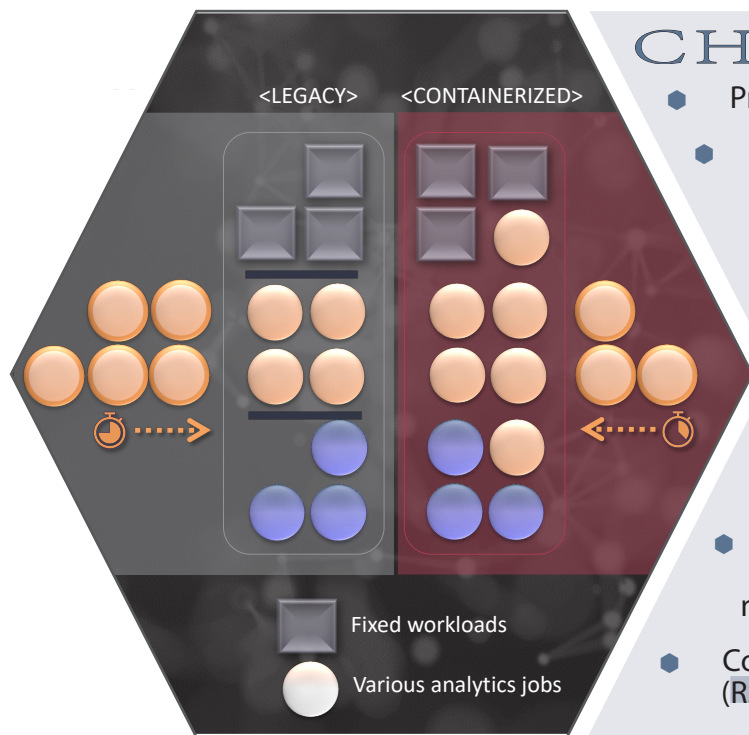


Kubernetes together with Docker enables an efficient, reliable Continuous Delivery workflow.

In the context of moving a custom analytics platform to managed, containerized components, we designed and implemented a modern solution to provide reproducible, on-demand analytics batch jobs in a resource-efficient manner.



Significantly increased capacity for running analytics on similar infrastructure.



CHALLENGES

- Providing support for heterogeneous analytics workloads
- Using the same infrastructure for long-running applications - composed of several simulation, modeling and reporting steps - and on-demand transient analytics workloads
- Ensuring numerical reproducibility across development, CI and production environments

SOLUTION

- Continuous integration builds with Jenkins producing container images based on versioned code packages
- Replaced commercial solution for submitting job requests to a dedicated cluster with a custom component for managing analytics job requests on Kubernetes.
- Configure containerized development environments (RStudio sessions running on Kubernetes)

BENEFITS

- **Flexible resource allocation:** Kubernetes allows pooling resources across multiple machines and between all components. All free resources can be used by analytics jobs, instead of only a dedicated cluster.
- **Numerical reproducibility:** Using container images enables better control over underlying dependencies and specific configurations, ensuring that calculations are reproducible across all environments.



"We were quickly able to run a prototype in parallel with the legacy solution, and validate our approach, which gave us the needed confidence to build the complete solution."

