

High Performance Computing (HPC) Lab

Risk-free, rapid evaluation of cloud HPC benefits using fourTheorem HPC and HTC accelerators.



What is HPC/HTC?

High Performance Computing (HPC) and High Throughput Computing (HTC) involves the use of clusters of processors, working in parallel, to process large datasets and solve complex problems at high speed.

Typical On Premises HPC Challenges

- Long execution times
- Capex and capacity planning
- High TCO
- Maintenance burden
- Low agility / innovation

Why now?

The growth in compute capacity and services on AWS means that HPC workloads, traditionally running on on-premises grids can now run more effectively and at lower cost on cloud.

What does it cost?

By leveraging AWS Funding programs, we minimize or eliminate all customer costs for the Lab.

HPC Lab

HPC Lab is fourTheorem's timeboxed HPC accelerator. Leveraging AWS scalable, low-maintenance compute services including:

- AWS Lambda, AWS Fargate, Amazon ECS, AWS ParallelCluster, AWS Batch, Amazon EC2

Who is it for?

- Financial services such as Insurance, Banking, Investments, Capital Markets
- Life Sciences
- Energy
- Automotive and Aerospace
- Weather and Climate modelling

Typical On-Premises Technology

- Bespoke clusters
- Tibco
- Symphony
- Slurm
- Hadoop/Spark

Key Benefits of HPC/HTC on AWS

- **Elastic scaling on demand** - compute resource scales in line with business needs.
- **Unconstrained parallel execution** - Remove resource contention issues and capacity planning.
- **Cost Flexibility** - Transparent insights into costs; optimize price/performance and only pay for what you use.
- **Dramatically Reduce code and infrastructure** - fewer bugs, less maintenance, reduced TCO.
- **Compliance** - Reduce the effort of security and regulatory compliance by leveraging the AWS Shared Responsibility Model.
- **Pulse of the Future** - The technology behind HPC Lab evolves with you, making obsolescence a thing of the past.
- **Sustainability** - Reduce carbon footprint by moving from always on grids to compute on demand.

How it works



The lab begins with an initial discovery workshop, which typically takes one day. Following the workshop, we assess the suitability of the proposed workload, determine a scope of work, technical architecture, data security requirements and Key Performance Indicators. In total this phase takes approximately one week.

During the evaluation phase, we migrate and modernize key elements of the workload, ensuring that the required KPIs, cost and scalability metrics are measured. This phase typically takes between 6 to 12 weeks depending on the nature of the workload and is iterative and transparent.

Finally, we produce an analysis report that details the results of the lab, measured KPIs and a go forward plan to move the workload fully to production. Post the lab engagement, we support the full build and scale out to production. A dedicated team is assigned full time for the duration of the lab.

Case Study - RenaissanceRe

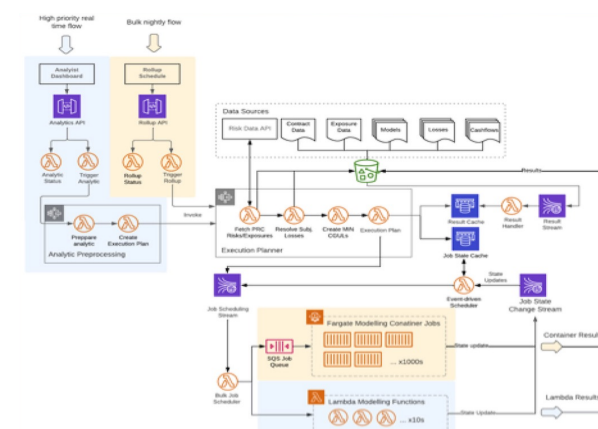
RenaissanceRe is a global provider of reinsurance and insurance that specializes in matching well-structured risks with efficient sources of capital.

Problem

- Global Reinsurance risk analysis platform
- Legacy on premises compute grid
- Contention and Execution times constraining business growth

Goals

- Cut execution times by an order of magnitude
- Deliver horizontal scale
- Reduce costs



Solutions and Outcomes

- Transform to run on AWS
- Utilize commodity, cloud native infrastructure.
- Execution time cut by 90% (12h to < 1h)
- Code base reduced by 70%
- On demand scaling
- Pay per use
- CO2 emissions cut in half
- Supports bulk and real-time in parallel

