THORIUM

SCR 2531

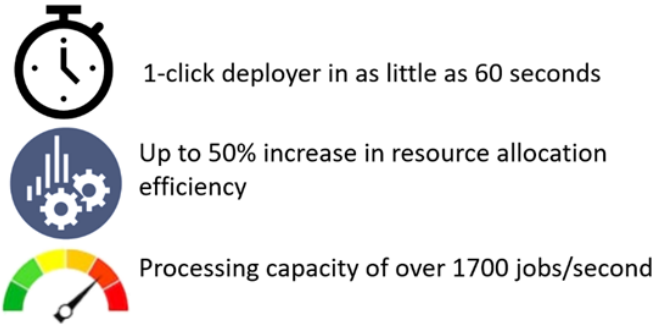
Kubernetes

Job Scheduler

High Performance

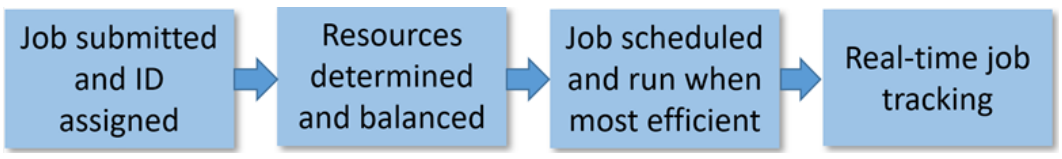
As a simple, architecture-agnostic job scheduler that allows users to run data processing jobs with zero deployment files, Thorium saves money by automatically and efficiently managing data infrastructure.

Modern business needs for data management have resulted in a surge in the fields of data science, machine learning, and big data platforms. These business needs (e.g. marketing, advertising, and hiring) are essential to both end-products and business critical operations and require large amounts of computational resources. Improper allocation and scheduling of these resources can cost companies millions of dollars per year. Sub-optimal resource allocation is often time consuming to diagnose and requires experienced engineers many weeks to solve. While modern infrastructure-as-a-service platforms, including Amazon Web Services and Google Cloud provide easy interfaces for scaling computational needs, efficient allocation and scheduling of these resources are still necessary to efficiently utilize them. Thorium fills these data management gaps.



The Thorium job scheduler is an architecture-agnostic job scheduler that allows users to run data processing jobs with zero deployment files. System users such as data engineers and business development teams can submit computational tasks to perform and Thorium provisions and allocates resources on a compute environment for the task to run. Thorium then ensures that the job is created, scheduled, executed, and completed successfully. Thorium can scale the compute environment from a laptop to a large compute cluster.

Optimized Process Flow:

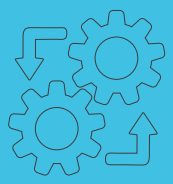


Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2021-10360 M

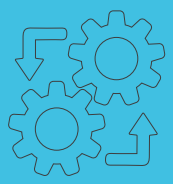
Sandia National Laboratories

THORIUM

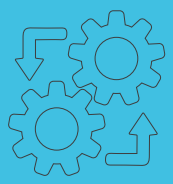
SCR 2531



Kubernetes



Job Scheduler



High Performance

Thorium allows you to dedicate more of your compute cluster to completing the jobs you want to do and need to do

	Thorium	Slurm	Celery	Resque	PBS Pro	Apache Airflow
Easy to Build Pipelines	✓	✗	✗	✗	✗	✗
Auto-handles Scaling	✓	✓	✓	✓	✓	✓
Auto-handles Deployment	✓	--	✗	✗	--	✓
Requires Minimal Resources	✓	✓	✓	✓	✗	✗
Native Container Support	✓	✗	✗	✗	✗	✓

Thorium's Four Main Components

Scheduling API

The job scheduling API backend defines pipelines to string containers together to complete jobs. Pipelines include naming, resource specification, and priority specification.

Deployment Operator

Oversees scheduling operations, including software deployment to new resources with a one-click deploy. When resources are available, they can be automatically provisioned by the deployment operator to run jobs available in the queue and are prioritized by the service level agreement (SLA).

Authentication

Provides group-based access controls and authentication to the scheduler ensuring your jobs and data are secure.

Graphical UI

Provides access to job queues, prior execution logs, and current reaction statuses. Also enables users to rapidly onboard without familiarity with API or requiring any code writing.

Technical Benefits

- Efficient provisioning and allocation of resources on a compute environment
- Does not require complex coding or engineering to scale – eliminates the need for large engineering support staff
- Deploys on top of any Kubernetes cluster
- Ensures that the job is created, scheduled, executed, and completed successfully

Industries & Applications

- Job scheduler
- Handling of large data sets and processing infrastructure

ip.sandia.gov
ip@sandia.gov

U.S. DEPARTMENT OF ENERGY | NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2021-10360 M