Cluster Executor Concept

Workflows created with KNIME Analytics Platform can specify local or distributed execution for each node. Select your cluster execution engine of choice from within the preference settings and use node-specific settings for more refined cluster use.

Data may be split into multiple subsets (see picture) and sent to the master server in the cluster. The master then coordinates task execution. Third-party nodes not normally offering cluster support may also be distributed in this way.

Parallel nodes in a single workflow can be executed on separate resources in the cluster. When remote execution of all tasks is finished, KNIME Analytics Platform collects the results from the remotely-executed nodes and completes workflow execution locally.

Applications

KNIME Cluster Executor is used by power-users across industries to provide high-speed calculation when mission critical results are time sensitive.

Workflows that process large data sets or require compute-intensive calculations are ideal candidates for KNIME Cluster Executor. This extension has been used successfully for 3-D modelling calculations and virtual docking experiments in pharmaceutical research.

Another area of use is the analysis and mining of huge data volumes for customer relationship management, telecom, and financial databases.

KNIME Software Integration

KNIME Cluster Executor is a commercial extension of KNIME Analytics Platform.

KNIME Server complements Cluster Executor by adding shared repositories, advanced access management, enhanced execution (scheduling), and web enablement (web portal and REST services). On completion, results may be downloaded in popular document formats including PDF, Word, Excel, or PowerPoint.

KNIME Cluster Executor is fully compatible with other KNIME Collaboration Extensions, including Server Lite, WebPortal, and TeamSpace.
Extending KNIME Analytics Platform

The full featured, unrestricted, open source, and free KNIME Analytics Platform is the perfect environment for unleashing the potential of a single data scientist. When you are ready to take your analytics to the next level, KNIME software extends those capabilities with commercial extensions to enhance collaboration, performance, and productivity.

Whatever your data needs, KNIME Commercial Software can take you there.

Software Requirements

KNIME Cluster Executor is compatible with Grid Engine derivatives supporting DRMAA (including Sun/Oracle/Univa Grid Engine). It requires Linux running on both Submit Clients and Cluster Slaves.

About KNIME

At KNIME®, we build software for fast, easy, and intuitive access to advanced data science, helping organizations drive innovation.

Our KNIME Analytics Platform is the leading open solution for data-driven innovation, designed for discovering the potential hidden in data, mining for fresh insights, or predicting new futures. Organizations can take their collaboration, productivity and performance to the next level with a robust range of commercial extensions to our open source platform.

For over a decade, a thriving community of data scientists in over 60 countries has been applying our platform to every kind of data: from numbers to images, molecules to humans, signals to complex networks, and simple statistics to big data analytics.

KNIME’s headquarters are in Zurich, with additional offices in Konstanz, Berlin, and Silicon Valley. We’re open for innovation®, so visit us at KNIME.com.

Example workflow using KNIME Cluster Executor