

Where Data and Discovery Converge

Cray® Urika®-XC analytics

Unlock sophisticated big data analytics at the speed of supercomputing with Cray® Urika®-XC analytics, a software solution that brings you powerful analytics tools, engineered specifically to take advantage of the unique, scalable Cray® XC™ series architecture and Cray's Aries™ network interconnect. The Urika-XC framework is managed with familiar Cray XC system tools, while the analytics applications benefit from supercomputer power and scale.



DO IT ALL

Gain new insights and solve the toughest problems across all your data on one machine — without additional hardware, and without moving data from machine to machine. Urika-XC software allows you to run analytics and simulation workloads simultaneously on Cray XC series systems, using familiar system tools and schedulers. With supercomputer speed and efficiency, you can run more types of analytics faster, yielding unprecedented breadth of results and maximizing your opportunity for transformative discoveries.



GET MODERN

Use modern tools and languages to power richer insights and reduce the need for scarce talent. The Urika-XC package includes the Cray Graph Engine for sophisticated graph analysis that scales to billions of edges, as well as the Apache® Spark™ processing engine and familiar development languages such as R and Python.



IMPROVE TCO

With Cray Urika-XC technology, you can make the most of your investment in Cray XC series supercomputers and reduce the need to maintain separate systems. Meet increased demand for analytics without capital investment or retraining.

A Look Inside

Cray Graph Engine	In-memory semantic graph database, implemented using HPC technologies. Based on W3C industry standards, implemented as an RDF triplestore with SPARQL 1.1 query language. Designed to scale to 256 nodes and 73 billion triples.	
Open-Source Analytics	Spark distributed in-memory processing for analytics and deep learning	<ul style="list-style-type: none"> ▪ Apache Spark 2.1.1 ▪ BigDL distributed deep learning library for Spark
	PyData Python analytics environment	<ul style="list-style-type: none"> ▪ Anaconda 4.1.1 with Python 2.7, 3.4, 3.5, 3.6 ▪ Distributed Dask parallel computing for Python environments
	Analytics compilers and tools	<ul style="list-style-type: none"> ▪ Languages: Java 1.8.0, Scala 2.11.8, R 3.3.2 ▪ Tools: sbt 0.13.9 build tool, Maven 3.3.9 project management