



Find the pearls.

Locate patterns, hidden relationships
and opportunities in Big Data.

HPCC Systems
Open-Source, Enterprise-Proven
Big Data Processing and Analytics



Risk Solutions
HPCC Systems

Linking and Analyzing Big Data Provides Pearls of Insight

Organizations today are dealing with a wealth of data, amassing to gigabytes, terabytes to even petabytes of information. By linking and analyzing current and historical data, organizations can create a complete view of their customers to gain knowledge of customer preferences and behaviors to better increase customer loyalty. Insights from Big Data can also help organizations grow the top line by exploring new sources of revenues and new types of services. Organizations can also leverage Big Data to take advantage of real-time trends, manage revenue and costs, quickly detect fraud and collusion, and launch new products into markets faster. All of these benefits create a greater competitive advantage for the enterprise.

However, organizations struggle to cope in an environment where their ability to collect information is outpacing their ability to use it effectively. In addition, most organizations realize that their information infrastructure cannot hold, link, analyze and process the data collected. Industry reports show that organizations that can't manage and analyze all their current and historical data are putting themselves at a competitive disadvantage and will lag behind their peers.

Today, companies need the processing power to pull data from disparate sources to process and analyze.

LexisNexis has Decades of Experience with Processing, Linking and Analyzing Big Data to Glean Insights and Target Opportunities

Customers such as leading banks, insurance companies, utilities, law enforcement and Federal government depend on LexisNexis technology and information solutions to help them verify identity, assess risk, predict fraud, comply with legislation and know customers better.

LexisNexis Has Four Petabytes of Information:

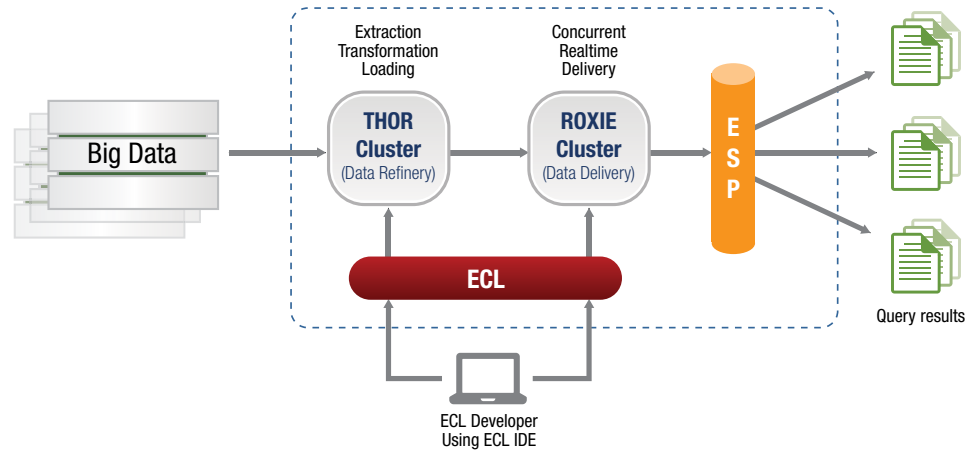
- 50 billion records
- 20,000 sources
- Several million records added daily
- 7.5 billion unique name and address combinations
- 250 million unique identities
- Over 1.1 billion unique business contacts

LexisNexis Big Data Processing Capabilities are Now Open Source

To manage, sort, link, and analyze billions of records, LexisNexis developed a data intensive supercomputer to run its own operations and grow the LexisNexis Risk Solutions division to a \$1.4 billion business. The supercomputer, a high performance computing cluster (HPCC) has been proven with enterprise customers, through LexisNexis products and services, who need to process large volumes of data in mission-critical 24/7 environments. In 2011, LexisNexis decided to open source the proprietary data intensive supercomputer and launch it into the marketplace as HPCC Systems for Big Data analytics processing.

HPCC Systems helps organizations gain competitive advantages by leveraging all the data to help scale for innovation and growth. The streamlined platform needs fewer resources to operate and eliminates expensive legacy technology to lower the Total Cost of Ownership (TCO) of Big Data solutions. The launch of HPCC Systems represents a paradigm shift to enterprise companies open-sourcing their proprietary and proven technology assets. LexisNexis has disrupted the Big Data industry by offering a proven asset to the open source community, while continuing to remain committed to developing the most innovative products and services for customers.

High Performance Computing Cluster platform known as HPCC Systems.



The HPCC Systems platform has three distinguishing factors that make it an effective choice for big data analytics and processing:

- The HPCC Systems **Data Refinery engine (Thor)** helps clean, link, transform and analyze Big Data. Thor supports ETL (Extraction, Transformation and Loading) functions like ingesting unstructured/structured data out, data profiling, data hygiene, and data linking out of the box. In addition, Thor supports flexible record oriented data structures.
- The HPCC Systems **Data Delivery engine (Roxie)** provides highly concurrent and low latency real time query capability. The Thor processed data can be accessed by a large number of users concurrently in real time fashion using the Roxie. The Roxie queries are typically complex and could include embedded rules logic.
- The programming language is called Enterprise Control Language (ECL) and is used to program both the data processing jobs on Thor and the queries on Roxie. ECL is a declarative, implicitly parallel and data flow oriented programming language that abstracts complex data processing tasks by providing a simple programming interface.

Proven Linking Technology is Critical for Successful Entity Resolution

Big Data processing capabilities are not enough to ensure entity resolution success. Data must be linked quickly and accurately. LexIDSM is the proprietary ingredient in our products that turns disparate information into meaningful insights. This technology enables customers to identify, link and organize data associated with a record, so that identities and entities can be disambiguated quickly with a high degree of accuracy. LexID helps our customers:

Get a More Complete Picture.

Make intelligent information connections beyond the obvious by drawing insights from both traditional and new sources of data.

Better Results, Faster.

Use the fastest technology for processing large amounts of data to help you solve cases more quickly and confidently.



Benefits of HPCC Systems and LexID

- Enables **entity resolution** on a **scale** not previously available
- Offers a **single architecture**: two data platforms (query and refinery) and a consistent data-intensive programming language, **ECL**
- Enables **massive joins, merges, sorts and data transformations**
- Accelerates creation of new services via **rapid prototyping capabilities**
- **Cost Savings**: Commodity hardware and fewer people can do much more in less time



HPCC Systems Community Edition is a free version supported by an active community of developers and enthusiasts via the online discussion forums.

HPCC Systems Enterprise Edition is a subscription-based offering for enterprises. It includes software, services and support to deliver the highest levels of reliability, security, and uptime proven over the years in mission-critical industrial settings.

For more information:

Call 678.694.2200
or visit <http://hpccsystems.com>

About HPCC Systems®

HPCC Systems® from LexisNexis® Risk Solutions offers a proven, data-intensive supercomputing platform designed for the enterprise to process and deliver Big Data analytical problems. As an alternative to Hadoop and mainframes, HPCC Systems offers a consistent data-centric programming language, two processing platforms and a single architecture for efficient processing. Customers, such as financial institutions, insurance carriers, insurance companies, law enforcement agencies, federal government and other enterprise-class organizations leverage the HPCC Systems technology through LexisNexis® products and services. For more information, visit <http://hpccsystems.com>.

About LexisNexis Risk Solutions

LexisNexis® Risk Solutions (www.lexisnexis.com/risk/) is a leader in providing essential information that helps customers across all industries and government predict, assess and manage risk. Combining cutting-edge technology, unique data and advanced scoring analytics, we provide products and services that address evolving client needs in the risk sector while upholding the highest standards of security and privacy. LexisNexis Risk Solutions is part of Reed Elsevier, a leading publisher and information provider that serves customers in more than 100 countries with more than 30,000 employees worldwide.

