



Fact Sheet

Big Data Intelligence Begins with Intel

The number of connected devices is equal to the current global population, and is expected to double by 2015. The increase is spurred largely by billions of networked sensors and intelligent systems – also known as the Internet of Things (IoT). From the proliferation of mobile devices such as smart phones and tablets to RFID readers and sensors, people and machines are producing data at exponential rates.

In fact, Intel estimates that the world generates 1 petabyte of data every 11 seconds or the equivalent of 13 years of high definition video. The term “big data” commonly refers to this explosion of data – characterized by its volume, variety and velocity – that offers the potential to enrich our lives through new scientific discoveries, business models and consumer experiences.

Expanding Software Portfolio to Advance Big Data Analytics

Intel’s position is that every individual and organization in the world should be able to unlock the intelligence available in big data. The company aims to address the cost, complexity and confidentiality concerns associated with managing, storing and securing the massive amounts of data.

Intel is addressing these concerns by delivering open data management and analytics software platforms including the Intel Distribution of Apache Hadoop* software (Intel® Distribution) and the Intel Enterprise Edition for Lustre* software.

- ***Intel Distribution of Apache Hadoop software*** - Hadoop is an open source framework for storing and processing large volumes of diverse data on a scalable cluster of servers. The Intel Distribution is the first to provide complete encryption with support of Intel® AES New Instructions in the Intel® Xeon® processor. By incorporating silicon-based encryption support of the Hadoop Distributed File System*, organizations can now more securely analyze their data sets without compromising performance. The optimizations made for the networking and IO technologies in the Intel Xeon processor platform also enable new levels of performance. Analyzing one terabyte of data, which would previously take more than 4 hours to fully process, can now be done in 7 minutes¹ thanks to the data-crunching combination of Intel’s hardware and the Intel Distribution.
- ***Intel Enterprise Edition for Lustre software*** - Lustre is an open source parallel distributed file system and key storage technology that ties together data and enables extremely fast access. The Intel Enterprise Edition for Lustre software is a validated and supported distribution of Lustre featuring management tools as well as a new adaptor for the Intel Distribution. When paired with the Intel Distribution, the Intel Enterprise Edition for Lustre software allows Hadoop to be run on top of Lustre, significantly

improving speed in which data can be accessed and analyzed. This allows users to access data files directly from the global file system at faster rates and speeds up analytics time, providing more productive use of storage assets as well as simpler storage management.

Intel is committed to contributing its enhancements made to both the Apache Hadoop and Lustre code back to the open source community. The aim is to provide the industry with a better foundation from which it can push the limits of innovation and realize the transformational opportunity of big data.

Simplifying the Internet of Things to Maximize Big Data Potential

According to IDC*, by 2015 more than 15 billion devices will be connected to the Internet, rapidly accelerating opportunities for the Internet of Things (IoT)¹. As these devices connect and generate data, they are generating a new wave of big data in industries such as manufacturing automation, energy and transportation.

End-to-end analytics defines how these billions of networked sensors and intelligent systems will become connected in the next few years. Today those applications often refer to monitoring a single machine and sending out an alert when something goes wrong. However, as many more machines become connected to each other and the cloud, there are opportunities not only to address single machine, but also entire systems. For example, networks of Intel-based systems, called synchrophasors, are being deployed in the electrical grid to solve some of the largest challenges faced by electric utilities. Time synchronized analysis of the massive amount of data collected by these grid sensors provides utility operators a real-time perspective of the performance and health of the electrical grid.

In order to fully realize these types of gains, the technology has to be easy to implement and automated. By providing a standardized and open platform for advanced analytics from the device to the cloud, Intel aims to unlock the intelligence available in big data in order to enrich our lives through new scientific discoveries, business models and consumer experiences.

Driving Innovation in Big Data

Investments, led by Intel Labs, are driving academic research in data-intensive computing platforms, machine learning, parallel algorithms, visualization and computer architecture. Intel has established an [Intel Science and Technology Center \(ISTC\)](#) for Big Data. It is hosted by the Massachusetts Institute of Technology's (MIT) Computer Science and Artificial Intelligence Laboratory (CSAIL). The objective of the research center is to encourage new data-intensive user experiences that accelerate the pace of discoveries across science, medicine and industry.

Intel has also released [GraphBuilder](#), a beta, open software that helps scientists in industry and academia to rapidly develop new applications by constructing graphs that outline relationships within data. Additionally, Intel continues to invest in research and capital to advance the big data ecosystem. Intel Labs is at the forefront of advanced analytics research which includes the development of Intel® Graph Builder for Apache Hadoop* software, a library to construct large data sets into graphs to help visualize relationships between data.

Intel Capital continues making major investments in disruptive big data analytics technologies including MongoDB company [10gen](#) and big data analytics solution provider [Guavus Analytics](#).

About Intel

Intel (NASDAQ: INTC) is a world leader in computing innovation. The company designs and builds the essential technologies that serve as the foundation for the world's computing devices. Additional information about Intel is available at newsroom.intel.com and blogs.intel.com.

– 30 –

Intel and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

* Other names and brands may be claimed as the property of others.

¹Industry Developments and Models, Intelligent Systems: The Next Big Opportunity, IDC 2011

CONTACT: Krystal Temple
480-242-6943
Krystal.temple@intel.com