Next Generation Architecture and Capabilities
How is Big Data transforming the way organizations analyze information and generate actionable insights?
Paradigm shifts enabled by big data
Leverage more of the data being captured

TRADITIONAL APPROACH
Analyze small subsets of Information
All available information

BIG DATA APPROACH
Analyze all information
All available information analyzed
Paradigm shifts enabled by big data
Reduce effort required to leverage data

TRADITIONAL APPROACH

Carefully cleanse information before any analysis

Small amount of carefully organized information

BIG DATA APPROACH

Analyze information as is, cleanse as needed

Large amount of messy information
Paradigm shifts enabled by big data
Data leads the way—and sometimes correlations are good enough

**TRADITIONAL APPROACH**

- Hypothesis
- Question
- Answer
- Data

Start with hypothesis and test against selected data

**BIG DATA APPROACH**

- Data
- Exploration
- Insight
- Correlation

Explore all data and identify correlations
Paradigm shifts enabled by big data

Leverage data as it is captured

TRADITIONAL APPROACH

Data → Repository → Analysis → Insight

Analyze data *after* it’s been processed and landed in a warehouse or mart

BIG DATA APPROACH

Data → Analysis → Insight

Analyze data *in motion* as it’s generated, in real-time
Next generation architecture
for delivering information and insights

Data sources
- Machine and sensor data
- Image and video
- Enterprise content
- Transaction and application data
- Social data
- Third-party data

Real-time analytics

Hadoop

Enterprise warehouse

Exploration, landing and archive

Actionable insight
- Decision management
- Predictive analytics and modeling
- Reporting, analysis, content analytics
- Discovery and exploration

Analytic appliances

Data mart

Information governance
IBM Research Projects in Big Data
University of Ontario Institute of Technology (UOIT) uses big data to improve quality of care for neonatal babies

Need

• Performing real-time analytics using physiological data from neonatal babies
• Continuously correlates data from medical monitors to detect subtle changes and alert hospital staff sooner
• Early warning gives caregivers the ability to proactively deal with complications

Benefits:

• Detecting life threatening conditions 24 hours sooner than symptoms exhibited
• Lower morbidity and improved patient care
Asian Health Bureau reduces diagnostic errors

Capabilities Utilized: Hadoop System

- Telemedicine imaging diagnostics service to improve rural healthcare
- Automatically sifts and analyzes large collections looking for anomalies and disease
- Makes it possible for radiologists and Pathologists to analyze:
  
  **1000s of patient images**

Significant improvements expected:

- Reduction in diagnostic errors
- Improved outcomes by leveraging physicians treating similar cases

“Over 80% of healthcare data is medical imaging”
University of Maryland, Baltimore County (UMBC) developing accurate forecasts for wildfire smoke dispersion

Need

- Researchers needed a sophisticated analysis platform with powerful real-time processing capabilities for tracking how fire and smoke spread

Benefits

- Enables an improvement of about 16% in forecast accuracy in the first three days of an event
- Fuses data from multiple sources to provide public officials with real-time data about fire and smoke status
TerraEchos uses streaming data technology to support covert intelligence and surveillance sensor systems

Need

• Deployed security surveillance system to detect, classify, locate, and track potential threats at highly sensitive national lab

Benefits

• Reduces time to capture and analyze 275MB of acoustic data from hours to one-fourteenth of a second

• Enables analysis of real-time data from different types of sensors and 1,024 individual channels to support extended perimeter security

• Enables a faster and more intelligent response to any threat
Marine Institute Ireland, monitors buoy sensor data to detect floods sooner

Need

• To make better use of the environmental data it collected
• Compiling this data to perform trend analyses or draw conclusions, however, was particularly complicated

Benefits

• Uses maritime data to increase the accuracy of flood predictions and support automated alarm systems, improving public safety
• Fosters new businesses and overall growth by providing seafood, shipping and monitoring operations with streaming maritime data
Dublin City Centre; Robust and efficient citywide traffic awareness system, enhance rapid action on incidents

Need
- A budget effective solution to improve traffic awareness system.
- To bring accuracy in event detection, inferring traffic condition (road speed) and prediction of bus arrival.
- Challenge is to rightly analyze GPS data, which is typically high data throughput and difficult to capture

Benefits
- Monitor 600 buses across 150 routes daily
- Analyzes 50 bus location updates per second, using InfoSphere Streams
- Collects, processes, and visualizes location data of all public transportation vehicles