Get a Farm-to-Table View of Your Data

Tracking data quality and lineage on-premises and in the cloud, on and off the cluster

Dr. Tendü Yoğurtçu, Chief Technology Officer
Who is Syncsort?

The global leader in Big Iron to Big Data

>7,000 Customers

84 of Fortune 100 are Customers

500+ Experienced & Talented Data Professionals

100+ Countries We Do Business In

3x Revenue Growth In Last 12 Months
**Customer Use Cases & Strategic Partnerships**

### Data Infrastructure Optimization
- Mainframe Optimization
- Cross-Platform Capacity Management
- EDW Optimization
- Application Modernization

### Data Availability
- High Availability & Disaster Recovery
- Mission-Critical Migration
- Cross-Platform Data Sharing
- IBM i Data Security & Audit

### Data Integration
- Mainframe Access & Integration for Machine Data
- Mainframe Access & Integration for App Data
- High-performance ETL

### Data Quality
- Data Governance
- Customer 360
- Big Data Quality & Integration
- Data Enrichment & Validation

---

**Big Iron to Big Data**

A fast-growing market segment composed of solutions that **optimize** traditional data systems and **deliver** mission-critical data from these systems to **next-generation analytic environments**.
Farm to Table
Technology Trends Advancing Data

- Advanced Business & Operational Analytics
- Cloud
- IoT & Streaming Data
- Data Governance
- Data Science & Artificial Intelligence
Technology Trends Advancing Data

- Advanced Business & Operational Analytics
- Cloud
- IoT & Streaming Data
- Data Science & Artificial Intelligence
- Data Governance
Technology Trends Advancing Data

Advanced Business & Operational Analytics

- Data Governance
- Cloud
- IoT & Streaming Data
- Data Science & Artificial Intelligence

Syncsort Confidential and Proprietary - do not copy or distribute
Technology Trends Advancing Data

- Advanced Business & Operational Analytics
- Cloud
- Data Governance
- IoT & Streaming Data
- Data Science & Artificial Intelligence
Data Governance

GOALS
- Regulatory compliance
- Understand data context, meaning
- Accuracy, completeness, consistency, relevancy, timeliness, validity of data

CHALLENGES
- Multi-platform, data volume and complexity
- Diversity and consistency of sources
- Compliance demands: broader & deeper

• Business imperative across platforms and deployment models, on-premise and in the cloud
Data Governance

• Requires a multi-faceted approach

QUALITY
• Discover sources of, relationships between, data
• Apply business rules to measure data quality continuously

SECURITY
• Protect the confidentiality, integrity and availability of data

LINEAGE
• Get insights into where data came from, what changes were made and where it lands
End to End Data Lineage in Cloudera Navigator
End to End Data Lineage in Cloudera Navigator

Data Sources

Syncsort accesses data from sources outside cluster.
End to End Data Lineage in Cloudera Navigator

Data Sources

Syncsort accesses data from sources outside cluster.

Syncsort onboards data, modifies on-the-fly to match Hadoop storage model.
End to End Data Lineage in Cloudera Navigator

Data Sources

- Syncsort accesses data from sources outside cluster.
- Syncsort onboards data, modifies on-the-fly to match Hadoop storage model.
- Syncsort changes, enhances, joins data in cluster with MapReduce or Spark.

Data Hub
End to End Data Lineage in Cloudera Navigator

Data Sources

- Syncsort accesses data from sources outside cluster.
- Syncsort onboards data, modifies on-the-fly to match Hadoop storage model.
- Syncsort changes, enhances, joins data in cluster with MapReduce or Spark.
- Syncsort passes source-to-cluster data lineage info to Navigator.

Data Hub

- Syncsort
- Cloudera Navigator

Syncsort Confidential and Proprietary - do not copy or distribute
End to End Data Lineage in Cloudera Navigator

Data Sources

- Syncsort accesses data from sources outside cluster.
- Syncsort onboards data, modifies on-the-fly to match Hadoop storage model.
- Syncsort changes, enhances, joins data in cluster with MapReduce or Spark.
- Syncsort passes source-to-cluster data lineage info to Navigator.

Data Hub

- Navigator gathers any other changes made to data on cluster.

Data changes made by MapReduce, Spark, HiveQL.
Syncsort accesses data from sources outside cluster. Syncsort onboards data, modifies on-the-fly to match Hadoop storage model. Syncsort changes, enhances, joins data in cluster with MapReduce or Spark. Syncsort passes source-to-cluster data lineage info to Navigator. Navigator gathers any other changes made to data on cluster. Analytics and visualizations get complete data.

Data analyst gets end-to-end data lineage info from Navigator.
Syncsort DMX-h + Cloudera Navigator for End-to-End Lineage
Data Lineage + Data Quality = Foundations of Data Governance

Data Sources

Multi-field fuzzy matching, de-duplication, cleansing, enrichment, standardization, business rule enforcement.

Data Lineage

Data Hub

Analytics, Visualization

Discovery and Profiling

Analytics and visualizations on clean, complete data you can trust.
Anti-Money Laundering Solution on CDH at Large Global Bank

- **Challenge:** Meet AML transaction monitoring and FCA compliance demands
  - Data too large, diversely scattered to analyze
  - Disparate data sources -- Mainframe, RDBMS, Cloud, etc

- **Requirements:**
  - Consolidated, clean, verified data for all analytics and reporting.
  - MUST have complete, detailed data lineage from origin to end point
  - MUST be secure: Kerberos and LDAP integration required
  - Need unmodified copy of mainframe data stored on Hadoop for backup, archive
Anti-Money Laundering Solution on CDH at Large Global Bank

Solution:

• Syncsort DMX-h to create “Golden Record” on CDH for compliance archiving
• Trillium Quality for Big Data for cluster-native data verification, enrichment, and demanding multi-field entity resolution on Spark framework
• Full end-to-end lineage to Cloudera Navigator, from all sources, through transformations, to data landing, including HiveQL changes

Benefits:

• New financial crimes data hub produces high performance results at massive scale
• Bank meets stringent Anti-Money Laundering compliance requirements
THANK YOU

Learn More & See For Yourself!
Visit Us at Booth 1022