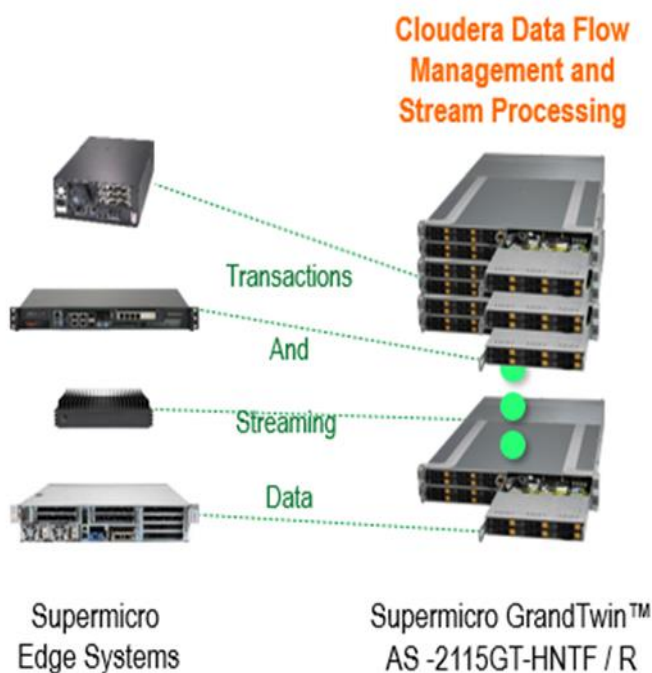




# SUPERMICRO AND CLUDERA DELIVER RELIABLE DATA COLLECTION FOR AI & ANALYTICS

*Supermicro GrandTwin™ Servers Running Cludera Flow Management and Stream Processing*



## Executive Summary

Businesses create incredible value using data they have on their systems, including edge systems. Retail, financial transactions, machine time series data, security, or manufacturing video feeds bring valuable insight to the corresponding business, enabling process improvements and the creation of new products and services.

Supermicro offers a complete line of edge systems to collect data at business edge locations, such as storefronts. Supermicro also offers a range of multi-node GrandTwin systems to run Cludera Data Flow, the Data-in-Motion platform. The Cludera Data Flow reliably collects structured and unstructured data (i.e., transactions, time-series data, images, videos)

## Cludera Data Flow Management and Stream Processing

Reliable Data Pipelines



Large File Transfers



Reliable Transactions Tracking



from data sources and edge systems. The data in motion is reliably collected and pipelined to downstream analytics and AI systems. These systems then identify issues or defects while learning real-time changes. The learning could turn into new opportunities for the business.

Figure 1. Open source protocols to collect data.

## Cludera Data Flow Management and Stream Processing

- [Cludera Data Flow\(CDF\)](#) is a scalable, real-time streaming data platform that makes it easy to collect, curate, and analyze data to gain key insights for immediate actionable intelligence.

- CDF does this through Flow Management, where customers can build scalable data movement flows from the edge to the data center using a low-code/no-code interface. The Flow Management is powered by open source Apache NiFi, MiniFi, and Edge Flow Manager.
- Stream Processing, which provides an enterprise-grade stream management and stateful processing capability to build real time applications. The stream processing is powered by SQL Stream Builder, Apache Flink, and Apache Kafka.
- All of this is secured and governed through the Cloudera Shared Data Experience, called SDX.

### Supermicro Data Center H13 GrandTwin Servers and Edge Systems

- [Supermicro H13 GrandTwin](#) systems provide a multi-node rackmount platform for cloud data centers. They come in 2U with four nodes for optimal deployment. These systems offer the latest AMD 4<sup>th</sup> Gen EPYC processors that support DDR-5 4800 memory and PCI Express Gen 5 I/O.
- Supermicro Edge systems come in various form factors to provide optimal deployment in edge locations to collect business data.

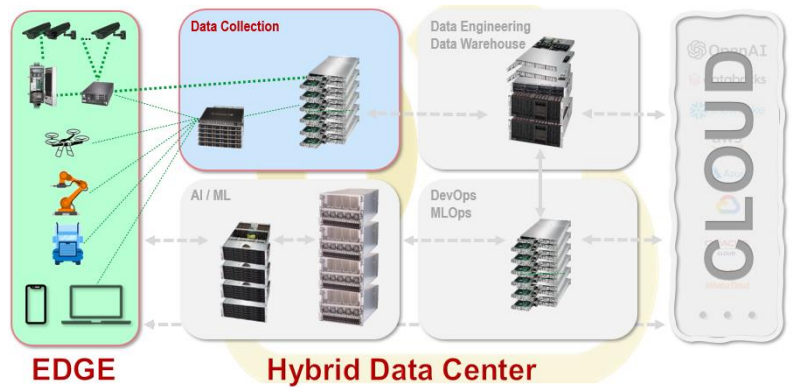


Figure 3. Hybrid data center architecture

### Implementation of Cloudera Data Flow on Supermicro H13 GrandTwin Servers

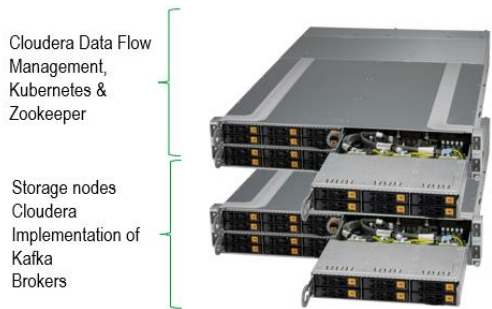


Figure 4. Cloudera Data Flow on Supermicro GrandTwin servers

<b>8 Nodes</b>	<b>AS -2115GT-HNTR</b>	
	<b>H13 GrandTwin, 2 x 2U/4 nodes, Rear I/O</b>	
	<b>Description</b>	<b>Qty</b>
	SPR 6430 2P 32C 2.1G 270W 60MB BI(1000)	2
	32GB DDR5 4800 ECC REG---MEM-DR532L-SL06-ER48	12
	Micron 7450 PRO 960GB NVMe PCIe 4.0 M.2 22x80mm	2
D3 S4520 7.68TB SATA6Gb/s 3D TLC 2.5" 7.0mm	4	
AIOM 2-port 100GbE QSFP28,Mellanox CX-6 DX	1	

or

<b>8 Nodes</b>	<b>AS -2115GT-HNTF</b>	
	<b>H13 GrandTwin, 2 x 2U/4 nodes, Front I/O</b>	
	<b>Description</b>	<b>Qty</b>
	SPR 6430 2P 32C 2.1G 270W 60MB BI(1000)	2
	32GB DDR5 4800 ECC REG---MEM-DR532L-SL06-ER48	12
	Micron 7450 PRO 960GB NVMe PCIe 4.0 M.2 22x80mm	2
D3 S4520 7.68TB SATA6Gb/s 3D TLC 2.5" 7.0mm	4	
AIOM 2-port 100GbE QSFP28,Mellanox CX-6 DX	1	

Streaming	Transactions	Time Series	Video
Energy	✓	✓	✓
Finance	✓		
Healthcare	✓		✓
Manufacturing	✓	✓	✓
Media	✓		✓
Retail	✓		✓
Telecom	✓		
Others	✓	✓	✓

## Conclusion

Supermicro systems running on the edge and data center, running the Cloudera data platform help businesses collect their data and turn the data into reliable business operations, new insights, and business opportunities.

For more information:

[Supermicro GrandTwin](#)

[Cloudera Data Flow Management](#) and [Stream Processing](#)

## SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

## CLOUDERA

Cloudera helps all industries tackle transformational use cases and extract real-time insights from an ever-increasing amount of data to drive value and competitive differentiation. Using open source and open standards, Cloudera delivers a hybrid data platform with secure data management and portable cloud-native data analytics. Cloudera has offices around the globe and is headquartered in Silicon Valley, California..