SOLUTION BRIEF

SUPERMICRO AND CLOUDERA DELIVER RELIABLE DATA COLLECTION FOR AI & ANALYTICS

Supermicro CloudDC Servers Excel at Running Cloudera Flow Management and Stream Processing

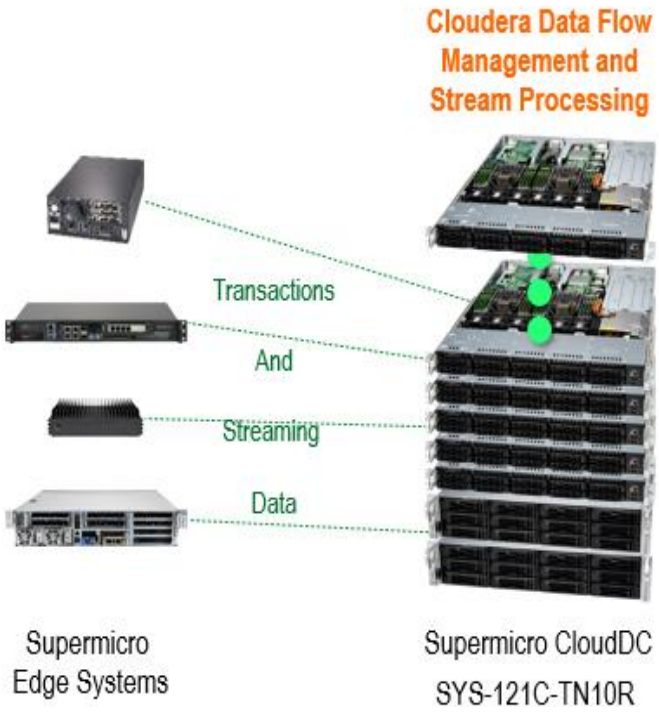

Figure 2. Data flow architecture using Supermicro CloudDC

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.

Figure 1. Open source protocols to collect data.

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

Cloudera Data Flow Management and Stream Processing

- Cloudera 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.
- This is secured and governed through the Cloudera Shared Data Experience, called SDX.

Supermicro Data Center X13 CloudDC Servers and Edge Systems

- Supermicro X13 CloudDC systems provide an all-in-one rackmount platform for cloud data centers. They come in either 1U or 2U for optimal deployment. These systems offer the latest Intel 4th Gen Xeon Scalable processors that support DDR-5 4800MHz memory and PCIe 5.0 I/O.
- Supermicro Edge systems come in various form factors to provide optimal deployment in edge locations to collect business data.

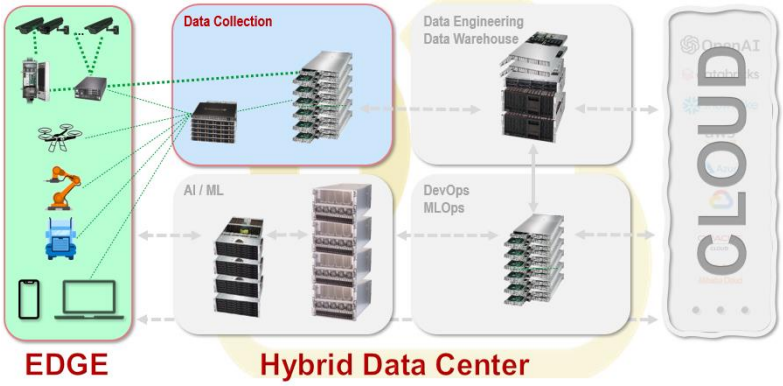

Figure 3. Hybrid data center architecture

Implementation of Cloudera Data Flow on Supermicro X13 CloudDC Servers

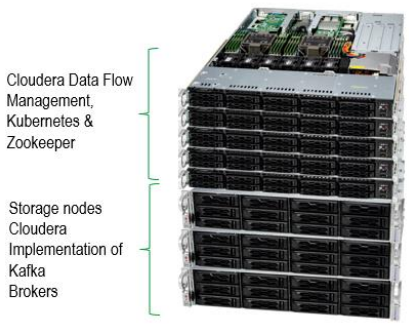

Figure 4. Cloudera Data Flow on Supermicro CloudDC servers

5 Nodes	SYS-121C-TN10R X13 CloudDC 1U, 10x 2.5" ,	
	Description	Qty
P4X-SPR6430-SRM7A-XCC	SPR 6430 2P 32C 2.1G 270W 60MB BI(1000)	2
	32GB DDR5 4800 ECC REG---MEM-DR532L-SL06-ER48	16
	Micron 7450 PRO 960GB NVMe PCIe 4.0 M.2 22x80mm	2
	D3 S4520 1.92TB SATA6Gb/s 3D TLC 2.5" 7.0mm	1
	AIOM 2-port 100GbE QSFP28,Mellanox CX-6 DX	1

3 Nodes	SYS-621C-TN12R X13 CloudDC 2U, 12x 3.5"	
	Description	Qty
	SPR 6430 2P 32C 2.1G 270W 60MB BI(1000)	2
	32GB DDR5 4800 ECC REG	16
	Micron 7450 MAX 400GB NVMe PCIe 4.0 3D TLC M.2 22x80 mm, 3DWPD	2
	5400 PRO 7.68TB, SATA, 2.5", 3D TLC, 0.6DWPD,7mm,TCG OPAL 2.0	4
	AIOM 2-port 100GbE QSFP28,Mellanox CX-6 DX	1

© 2023 Copyright Super Micro Computer, Inc. All rights reserved

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 CloudDC
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 exact 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..

© 2023 Copyright Super Micro Computer, Inc. All rights reserved