

Platfora Big Data Analytics and Cisco Unified Computing System



Solution
Partner

Platfora, the leading enterprise big data analytics platform built natively on Hadoop and Spark, delivers business outcomes and competitive advantage.

Platfora enables business users and data scientists to visually interact with petabyte-scale data in seconds, allowing them to work with even the rawest forms of transaction, customer interaction, and machine data. When combined with high-performance Cisco® Unified Computing System™ (Cisco UCS®) servers, the solution encourages more use cases and unlocks the enterprise's competitive edge by providing deeper insights into much larger data sets.

Analytics for Business Users

Platfora's solution runs on a separate dedicated cluster alongside the Apache Hadoop cluster. A Platfora cluster is typically configured with large amounts of RAM to support terabytes of in-memory analysis performed directly by business users, delivering exceptional performance while still allowing users to analyze all the data inside the Hadoop cluster. Platfora's solution is optimized to run on Cisco UCS, which can be seamlessly integrated with Cisco UCS Hadoop deployments.

See and Work with 100 Percent of Your Data

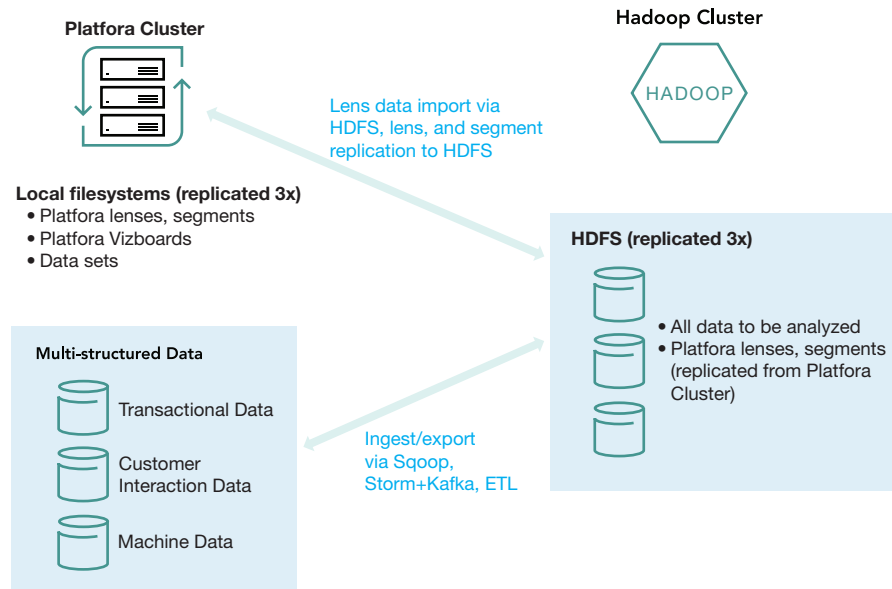
Business analysts can dig into all the multistructured data in your organization—transactions, customer interactions, and machine data—via Platfora's self-service access platform. With Platfora, analysts have access to all the data they need without needing to involve IT staff.

Improved Time to Value

Platfora enables business analysts to access and iterate their workloads in minutes, rather than hours. This faster time-to-insight drives faster reaction to changing business conditions and better support for business users whose budget drives IT adoption.



Platfora on Cisco UCS



How Platfora Works

- Platfora’s Interest Driven Pipeline™ enables business users to derive insight directly vs. engaging IT staff to condition and structure data required for analysis, an approach required by traditional business intelligence tooling. This approach reduces time-to-insight by orders of magnitude.
- Platfora performs behavioral analysis and iterative segmentation across 100 percent of the data in Hadoop. Traditional business intelligence tools cannot access all the data in Hadoop because they typically throttle through a SQL-over-Hadoop approach and squeeze the data into a single server, precluding exploratory analysis on very large data sets.
- Platfora delivers a scale-out, in-memory MPP-based accelerator engine, ensuring high-performance, low-latency response for business users.
- Platfora automatically generates MapReduce and Spark jobs to populate an in-memory acceleration layer, removing the need for business analysts to write SQL, Hive, or MapReduce code.

Use Cases

Customer Analytics and Insights: Understand Your Audience Better Than Ever Before

Platfora encourages you to ask new questions about your data, making it easy for marketing professionals to follow hunches, test theories, and continuously refine analysis until they find exactly what they are looking for—all with no coding required. Frequently implemented use cases include golden path to purchase, segmentation analysis, and customer churn drivers.

Internet of Things: Combine New Data Sets in Ways Never Before Possible

Platfora enables analysts to combine large amounts of data in multiple formats—from product telemetry via connected devices to user experience, traffic and parking control, and telemedicine. Frequently implemented use cases include product/feature utilization, sensor analytics, and supportability analysis.

Security and Compliance: Practical, Proactive Protection

Network security analytics with Platfora uses the power of Hadoop to spot subtle breach patterns across billions of events—without waiting two or three months for analysis. Platfora delivers full situational awareness via the ability to rapidly interrogate data to investigate incidents and improve understanding of network-borne threats and adjust the questions being asking based on rapidly changing behaviors. Advanced persistent threat identification is the most frequently-implemented use case.

Platfora on Cisco UCS Integrated Infrastructure for Big Data

Cisco UCS is the first converged data center platform that combines industry-standard x86-architecture servers with networking and storage access into a single converged system. The Cisco UCS innovations of unified fabric and unified management for all connected devices offer simplified management, world-class performance, and exceptional scalability needed to support the Platfora analytic workloads.

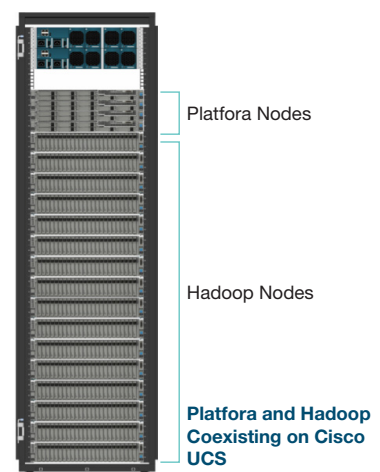
Cisco UCS Integrated Infrastructure for Big Data is the third generation of Cisco's solution for big data. It extends the Cisco UCS Common Platform Architecture (CPA) for Big Data with improvements in performance and capacity. The solution has been widely adopted for a variety of workloads across all enterprise segments. It accelerates deployment, delivers predictable performance, and reduces total cost of ownership (TCO).

Reference Configurations

The Platfora on Cisco UCS reference configurations are based on the latest Cisco solution for big data. The Platfora cluster is designed to deploy alongside Hadoop environments (shown in the figure), taking full advantage of the lossless 10Gbps unified fabric connectivity for Platfora's lens-building traffic. The solution extends Hadoop deployments on Cisco UCS CPA for Big Data with Platfora's scale-out, in-memory MPP-based accelerator engine to deliver business intelligence that meets the need of business users.

The reference configurations are built with the following components:

Cisco UCS 6200 Series Fabric Interconnects provide high-bandwidth, low-latency connectivity for servers, with Cisco UCS Manager providing integrated, unified management for all connected devices. Deployed in redundant pairs, Cisco fabric interconnects offer the full active-active redundancy, performance, and exceptional scalability needed to support the large number of nodes that are typical in clusters serving big data applications. Cisco UCS Manager enables rapid and consistent server configuration using service profiles, automating ongoing system maintenance activities such as firmware updates across the entire cluster as a single operation. Cisco UCS Manager also offers advanced monitoring with options to raise alarms and send notifications about the health of the entire cluster.



Cisco UCS C220 M4 Rack Servers are designed for performance and density over a wide range of business workloads in a 1-rack unit (1RU) form factor. Cisco UCS C220 M4 servers are powered by dual Intel Xeon E5-2600 v3 series CPUs, and they support up to 768 GB of main memory. These servers support four or eight SAS/SATA/SSD drives as well as Cisco UCS virtual interface cards (VICs) optimized for high-bandwidth and low-latency cluster connectivity, with support for up to 256 virtual devices. Cisco UCS C220 M4 servers are ideal for building Platfora clusters.

Cisco UCS C240 M4 Rack Servers are enterprise-class servers that deliver an outstanding combination of performance, flexibility, and efficiency for storage. The Cisco UCS C240M4 servers are 2-socket, 2-rack-unit (2RU) servers based on Intel Xeon E5-2600 v3 series processors supporting up to 768 GB of DDR4 main memory. These servers support up to 24 SFF SAS/SATA/SSD drives or 12 LFF SAS/SATA drives, plus 2 SFF SSD drives. Their expandability and exceptional performance makes them an ideal fit for big data analytics, virtualization, and graphics-rich and bare-metal applications. Cisco UCS C240 M4 servers are ideal for Hadoop deployments.

The following table lists the reference configurations of Platfora on Cisco UCS. The options include deploying Platfora with Cisco’s performance optimized or capacity optimized solution for big data. The typical ratio of Hadoop nodes to Platfora nodes in deployment ranges from 4:1 to 8:1,

Reference Configurations

| SOLUTION | PLATFORA WITH PERFORMANCE OPTIMIZED HADOOP | PLATFORA WITH CAPACITY OPTIMIZED HADOOP |
|--|---|--|
| Platfora Servers | 4 Cisco UCS C220 M4 Rack Servers, each with: <ul style="list-style-type: none"> • 2 Intel Xeon processors E5-2620 v3 at 2.4 GHz • 256 GB of memory • Cisco 12G-2GB RAID Controller • 8 1.2TB 10K SFF SAS drives per server (38 TB total) | |
| Cisco UCS Solution Accelerator Paks | Hadoop–Performance Optimized* | Hadoop–Capacity Optimized* |
| Connectivity | 2 Cisco UCS 6296UP 96-Port Fabric Interconnects | 2 Cisco UCS 6296UP 96-Port Fabric Interconnects |
| Hadoop Servers | 16 Cisco UCS C240 M4 Rack Servers, each with: <ul style="list-style-type: none"> • 2 Intel Xeon processor E5-2680 v3 CPUs • 256 GB of memory • Cisco 12-Gbps SAS Modular Raid Controller with 2-GB FBWC cache • 2 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs • 24 1.2-TB 10K SFF SAS drives • Cisco UCS VIC 1227 (with 2 10GE SFP+ports) | 16 Cisco UCS C240 M4 Rack Servers, each with: <ul style="list-style-type: none"> • 2 Intel Xeon processor E5-2620 v3 CPUs • 128 GB of memory • Cisco 12-Gbps SAS Modular Raid Controller with 2-GB FBWC cache • 2 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs • 12 4-TB 7.2K Large Form Factor (LFF) SAS drives • Cisco UCS VIC 1227 (with 2 10GE SFP+ports) |
| Storage | 460 TB with 42 GBps of bandwidth | 768 TB with 16 GBps of bandwidth |
| Optional Software from Cisco | Red Hat Enterprise Linux, SUSE Linux Enterprise Cloudera, MapR, or Hortonworks Cisco UCS Director Express for Big Data | Red Hat Enterprise Linux SUSE Linux Enterprise Cloudera, MapR, or Hortonworks Cisco UCS Director Express for Big Data |

* Base rack solution is available as single SKU bundles. Performance Optimized Rack: UCS-SL-CPA3-P. Capacity Optimized: UCS-SL-CPA3-C.

depending on workload. Therefore, these reference configurations can be further customized to meet a variety of big data and analytics workload requirements. Up to 160 servers are supported in a single management domain with Cisco Nexus® 2232PP 10GE Fabric Extenders. These configurations can be further scaled to thousands of servers using Cisco Nexus 7000 or 9000 Series Switches.

Conclusion

Platfora Big Data Analytics on the Cisco UCS solution is designed to help organizations derive instant value from their big data deployments. Without the expense entailed in designing and building custom solutions, this solution can help organizations quickly and easily deploy the Platfora analytics platform on a powerful, and secure Hadoop environment with 100 percent of the data available for business users, without additional developer support.

More Information

www.platfora.com

www.cisco.com/go/ucs

For more information about Cisco UCS big data solutions, visit www.cisco.com/go/bigdata.

For more information about the Cisco CPA for Big Data, visit blogs.cisco.com/datacenter/cpav3.