



Yellowbrick 

WHITEPAPER

Andromeda Optimized Instances

Table of Contents

INTRODUCTION	3
INSTANCE DESIGN FOR DATA WAREHOUSING	4
Compute	4
Network	4
Storage	4
KALIDAH ACCELERATOR	5
SERVER HARDWARE	6
REMOTE SUPPORT	6
SUMMARY	7



Introduction

The Yellowbrick Data Warehouse is a cloud-native, parallel SQL database designed for the most demanding batch, ad hoc, real time, and mixed workloads. Yellowbrick innovates in three key areas:

- Enabling a more efficient business model by allowing customers to consume a modern, elastic, SaaS user experience in their own cloud account with predictable cost.
- Optimizing price/performance for new use cases that require more concurrent users and more ad hoc analytics.
- Providing deployment flexibility by offering an identical data warehouse for on-premises use cases.

For on-premises use cases, Yellowbrick has developed the “Andromeda” server hardware instance, and our new Kalidah processor, which together drive new efficiencies in price/performance. Andromeda contains a unique mix of ingredients delicately balanced and paired to provide optimal price/performance and high availability data warehousing. With our database and Andromeda, it's not uncommon to find one server node providing the equivalent query throughput of a dozen or more nodes of competitive cloud and on-premises databases, at a fraction of the total cost. For other details about Yellowbrick database software architecture, including our storage engine, see our “Inside the Yellowbrick Data Warehouse” whitepaper.



Instance Design for Data Warehousing

Parallel data warehouse workloads place substantial stress on servers, networks, and storage, somewhat like supercomputer applications. Unlike storage systems that just read or write data from discs and send it over a network, MPP database servers require large amounts of compute to process and transform the data before it's read or written, and as much memory bandwidth as possible to support random lookups of data for operations such as aggregates and joins. Furthermore, all the servers in a cluster need to continually coordinate query processing (requiring ultra-low network latency to rapidly execute short queries) and exchange data (requiring massive amounts of streaming bandwidth for large queries). During query processing, throughput will be bound by the network (latency or bandwidth), computation (cores or memory channels), or storage (reads or writes for spilling), depending on the operators in use.

Data warehouses are becoming Tier 1, business-critical applications, requiring instances to be highly available at the hardware and system level, fully resilient to hardware components (fans, power supplies, drives, adapters, etc.) failure, network failure, server node failure, and partial power failure.

Compute

For compute, we care about the cost of each CPU core, which largely dictates how fast we can go on executing instructions, and the cost per memory channel, which largely dictates how fast we can do large aggregates, joins, and sorts. With the introduction of AMD's EPYC processors, it is affordable to acquire 64 cores of compute with eight memory channels to result in the lowest possible price per core and memory channel.

Network

100Gb networks are now the sweet spot in cost per unit of bandwidth. Since a redundant network architecture is required for high availability, each server node has access to two network interfaces running over two separate switches. In addition, we have made use of features on the EPYC processor and the network interface to closely couple the fabric and query processing, enabling us to drive an incredible 200Gb/sec per node of data across the network – roughly 20GB/sec per node, full duplex, or 400GB/sec per chassis. To make this process efficient, we use a remote direct memory access (RDMA) fabric that allows direct movement of data – typically cache-resident – between nodes, with no TCP/IP or Linux kernel in the way to slow things down.

Storage

Each Andromeda server supports 8x 7mm NVMe U.2 drives, offering 24GB/sec of read bandwidth per node and 16GB/sec of write bandwidth. Because data is compressed, the effective read bandwidth per node is over 3x higher, sometimes peaking at over 100GB/sec of user data scanned per server node. To scan data at this rate, we need a hardware accelerator.

Kalidah Accelerator

The gap between CPU performance and storage throughput continues to grow, with storage throughput doubling every couple of years but CPU core counts and clock rates increasing far slower. The industry has tackled this problem by developing dedicated special purpose “accelerator processors.” This trend started with GPUs for graphics and has evolved into special variants now widely used in large-scale data processing for machine learning, searching the web, recognizing the environment in autonomous vehicles, or taking better photos on mobile phones.

In data warehousing, we have observed that substantial amounts of CPU time are spent just finding on disc the data on which we want to operate – combing through a haystack looking for needles. We have started Yellowbrick’s accelerator journey by offloading this effort from the host CPU to a new, dedicated processor, designed to do that at higher rates than software alone can accomplish.

The Kalidah processor core accelerates bandwidth-oriented data processing tasks used during table scans such as data validation, decompression, filtering, compaction, and reorganization. Each Kalidah core receives instructions on command/completion queues modeled on the NVME protocol. Like other Yellowbrick code, the Kalidah driver is asynchronous, reactive, and polls for completions. Kalidah contains instructions for the following block-oriented operations:

- Moving and reorganizing data via DMA.
- Parsing on-disc file formats.
- Decompressing data with multiple decompression codecs.
- Applying range filters and bloom filters to data.
- Recompacting data to remove rows that do not meet filter criteria.

Kalidah can support these operations on all data types currently supported by the Yellowbrick database. The cores are aggressively pipelined and operations can be chained together and executed one after the other on data as it’s streamed from disc. This means that in one shot, on-disc data can be parsed, decompressed, validated, range filtered, and bloom filtered – all without the need to write any decompressed data to memory.

The memory attached to each dual-core Kalidah is uniformly addressable by the SSDs, the host CPU, and the Kalidah cores themselves, enabling Kalidah to do bandwidth-intensive operations on large amounts of data without consuming the host CPU’s memory bandwidth (See Figure 1).

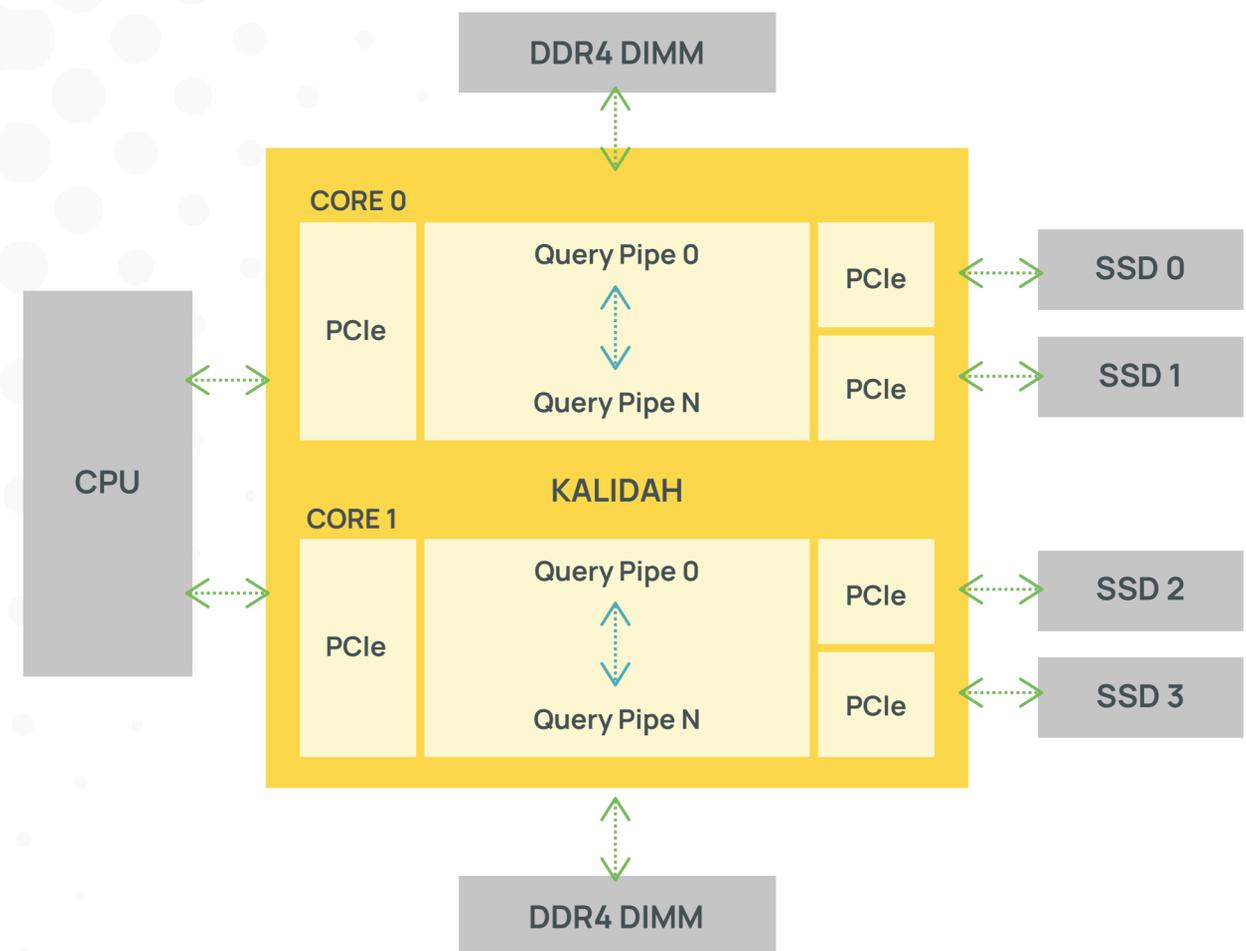


Figure 1

Kalidah offers a substantial increase in performance compared with running optimized software on the CPU, resulting in each Andromeda node having the following capabilities:

- **Stream throughput:** Each Andromeda server node can decompress, filter, compact, and reorganize data at 64GB/sec, regardless of the data type in use. To put this in perspective, that throughput is equivalent to 32 CPU cores.
- **Bloom filtering:** Bloom filtering is one of the most intensive operations used to look for specific data in table scans and to accelerate joins. It involves hashing data and matching the hash against multiple bitmaps. Each Andromeda server node can accomplish 8 billion bloom filter lookups per second, regardless of the data type in use. This is equivalent to about 24 CPU cores of performance for fixed-length data types, and 96 CPU cores of performance for variable-length data types.

Yellowbrick continually tunes and optimizes its database software to make more use of Kalidah.

Server Hardware

Andromeda is a blade server based on an existing design customized for Yellowbrick by a large server original design manufacturer (ODM) that supplies several public cloud vendors. The server motherboards are manufactured and tested on the same assembly lines that produce a large number of the world's servers for many major original equipment manufacturers (OEMs) and ODMs.

On-premises Yellowbrick customers receive Andromeda instances along with their software subscriptions. The hardware alone is not available for standalone purchase and is fully serviced and supported by Yellowbrick. Within the Andromeda chassis, all the following components are both hot-swappable and redundant:

- **SSDs**
- **Whole server blades**
- **Network switches**
- **Power supplies**
- **Fans**

Andromeda has been tested to scale efficiently from 3 blades to 80 blades (8x chassis) per data warehouse instance. The table below lists key Andromeda specifications; see the appendix for full details about available configurations.

Category	Specification
CPU	AMD EPYC with 64 cores, 8 memory channels
Network	200Gb (2x100Gb) RDMA fabric and 2x switches
Storage	8x U.2 hot-swap NVMe SSDs
Storage Capacity Per Blade	16TB, 32TB, 64TB (128TB on roadmap)
Memory Per Blade	512GB, 1TB, 2TB
Scan Accelerator	2x dual-core Kalidah processors
Minimum Blades	3 (in 1 chassis)
Maximum Blades	80 (8 chassis; 10 blades each)
Minimum vCPUs	384 (3 blades, 1 chassis)
Maximum vCPUs	10,240 vCPU (80 blades, 8 chassis)

Remote Support

Andromeda is designed with a cloud user experience even when running on-premises. By default, Yellowbrick maintains remote, unidirectional encrypted connectivity to the Yellowbrick SaaS monitoring platform. Key events are sent back over this Phone Home connection which our Customer Support team monitors 24/7.

In the event of any issues, chances are Yellowbrick knows about them first and will take corrective action. Optionally, different levels of data scrubbing are present to guard personal identifiable information (PII) and other confidential data. No customer data is ever shared. For customers with strict security requirements, the Phone Home feature can be disabled completely.

Summary

This whitepaper describes how Andromeda-optimized instances are designed to bring significant performance, efficiency, and economic advantages to customers deploying Yellowbrick inside private clouds. The result is a new kind of cloud-compatible data warehouse that provides the best economics in the industry, along with all other expected features and functions of a mature product that can be trusted to help run your business faster and more efficiently.

This document is provided for information purposes only and not warranted to be error free. It is subject to change without notice.
Yellowbrick is a registered trademark of Yellowbrick Data, Inc. Other names may be trademarks of their respective owners.

