CLOUD SOLUTIONS

# AMD EPYC™ 7002 Series Processors in the Cloud: Cloudera® on AZURE® with Ea and Da v4-Series Instances

FEBRUARY 2020

## Power and Flexibility

AMD EPYC™ processor-based Da and Ea series instances are available with:

- Up to 96 vCPUs
- Up to 384 GiB of memory on Da series and up to 672 GiB of memory on Ea series
- Up to 2,400 GiBs of SSD-based temporary storage

## Seamless Workload Migration

x86 compatibility and fully tested software stacks enable migration of applications currently running on other cloud instances to AMD EPYC™ processor-based instances with little to no modification.

## Workload Optimization

AMD engineers have tested a variety of popular applications on Microsoft Azure® enabling you to match compute resources to workload needs to optimize cost.

## Microsoft Azure® Da and Ea v4-Series Instances Featuring AMD Processors

Microsoft Azure® Da, Ea, Das, Eas v4-series instances are powered by AMD EPYC™ 7452 processors. The Da and Das v4 series provides a balanced CPU-to-memory ratio ideal for testing and development, small to medium databases, and low- to medium-traffic web servers. The Ea and Eas v4-series offers a high memory-to-core ratio that is great for relational database servers, workloads that benefit from medium to large caches, and in-memory analytics.

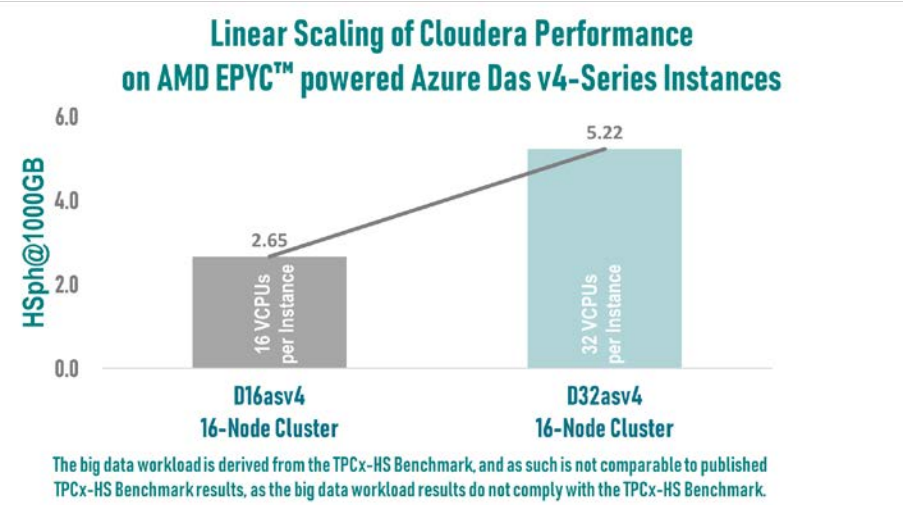



Figure 1: Cloudera Performance on Azure Das v4-series Instances[1] powered by 2nd Gen AMD EPYC™ Processors*

The Da, Ea, Das and Eas v4-series features high throughput, low latency, directly mapped local temporary storage, and runs on the AMD EPYC™ 7452 processor that features a max boost of up to 3.35GHz.

The Da and Ea v4-series virtual machines (VM)s come in sizes from 2 to 96 vCPUs in a simultaneous multi-threaded configuration. Each instance provides 4 GiB of memory per vCPU on the Da and Das v4-series, and 8 GiB of memory per vCPU on the Ea and Eas v4-series (exception 7 GiB of memory per vCPU on 96 vCPU VM), with both Ea and Da supplying up to 2400 GiB of temporary storage. The Eas and Das v4-series VMs support Premium SSD disk storage.

# Cloudera® on Azure Featuring AMD EPYC SoCs

Microsoft Azure is a growing collection of integrated public cloud services including analytics, virtual machines, databases, mobile, networking, storage, and web servers. Microsoft Azure provides a scalable computing platform that allows you to only pay for what you use, when you want it - without having to invest in on-premises server hardware. Azure is ready when you are, to scale your solutions up and out to whatever scale you require to service the needs of your business.

Cloudera® Enterprise Data Hub provides a complete multi-discipline big data solution. This solution integrates data warehousing, data science, data engineering, search, streaming, and real-time analytics into a unified platform for a single source of truth. Cloudera Shared Data Experience (SDX) creates a seamless integration of all analytic disciplines with security, governance and administration across any set of deployments: on-premise, cloud and hybrid. For more information, see : https://www.cloudera.com*

# Cloudera Reference Design on Azure Instances Powered by AMD EPYC Processors

Cloudera recommends right sizing your server configuration and deploying three or four machine types into production:

- **Master Node.** Runs the Hadoop® master daemons: NameNode, Standby NameNode, YARN Resource Manager and History Server, the HBase Master daemon, Sentry server, and the Impala StateStore Server and Catalog Server. Master nodes are also the location where Zookeeper and JournalNodes are installed. The daemons can often share single pool of servers. Depending on the cluster size, the roles can instead each be run on a dedicated server. Kudu Master Servers should also be deployed on master nodes.
- **Worker Node.** Runs the HDFS DataNode, YARN NodeManager, HBase RegionServer, Impala impalad, Search worker daemons and Kudu Tablet Servers.
- **Utility Node.** Runs Cloudera Manager and the Cloudera Management Services. It can also host a MySQL® (or another supported) database instance, which is used by Cloudera Manager, Hive, Sentry and other Hadoop-related projects.
- **Edge Node.** Contains all client-facing configurations and services, including gateway configurations for HDFS, YARN, Impala, Hive, and HBase. The edge node is also a good place for Hue, Oozie, HiveServer2, and Impala HAProxy. HiveServer2 and Impala HAProxy serve as a gateway to external applications such as Business Intelligence (BI) tools.

Note: The edge and utility nodes can be combined in smaller clusters, however in cloud environments it's often more practical to provision dedicated instances for each.

For a comprehensive discussion of Cloudera server configuration recommendation, architectures, and deployment on Azure, see: https://www.cloudera.com/documentation/other/reference-architecture/topics/ra_azure_deployment.html#azure_deployment*

## Reference Architecture

This reference architecture considers different types of workloads that are run on top of a Cloudera Enterprise Data Hub. The initial requirements focus on host types that are suitable for a diverse set of workloads. Choose the configuration based on the amount of data you have, and the workload you run on the cluster.

| CLUSTER TYPE | NODES |
|---|---|
| Starter (Up to 20 worker nodes) | • Worker nodes: Up to 20<br>• Master nodes: 2<br>• Utility/Edge node: 1 |
| Small (Up to 50 worker nodes) | • Worker nodes: Up to 50<br>• Master nodes: 2<br>• Utility/Edge node: 1 |
| Medium (Up to 200 worker nodes) | • Worker nodes: Up to 200<br>• Master nodes: 3<br>• Utility node: 2<br>• Edge node: 1+ |
| Large (Up to 500 worker nodes) | • Worker nodes: Up to 500<br>• Master nodes: 5<br>• Utility node: 2<br>• Edge node: 1+ |

# Benchmarking Cloudera® on Microsoft Azure® Das v4-Series Instances

AMD engineers ran tests derived from industry standard benchmarks across a variety of Das v4-series instance types. These results, shown below, demonstrate how effective the Das v4-series instances are at optimizing the compute power and I/O throughput for big data applications. Two instance types were tested with different vCPU and memory configurations. Five test runs were performed and the median result selected. All benchmark runs for all configurations were done with the workload at same scale factor[1].

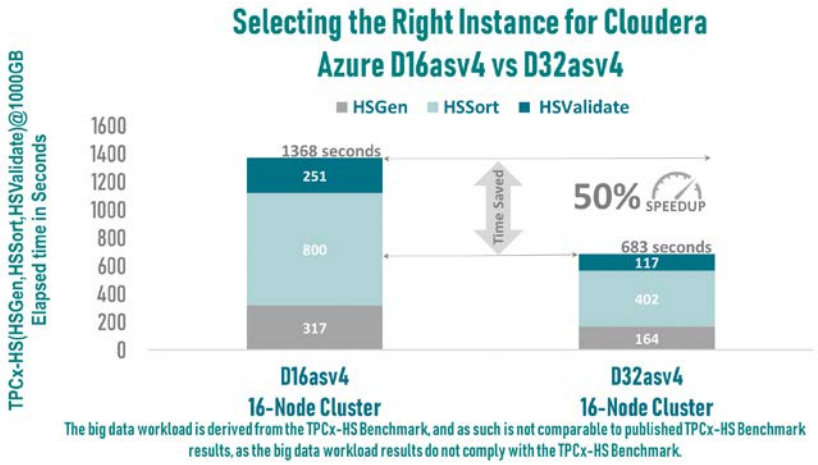



Figure 2: Cloudera Performance on Azure Das v4-series Instances[1]

Figure 1 uses time taken for three major modules of TPCx-HS Benchmark: HSGen, HSSort, and HSValidate. HSGen is a program to generate the data at a particular scale factor. HSSort is a program to sort the data into a total order. HSValidate is a program that validates the output is sorted.

Microsoft Azure® offers a variety of sizes in Da, Ea, Das, Eas v4-series instances powered by AMD EPYC™ 7452 processors. As shown in Figure 1, selecting the right instance size is very important to get best performance for big data applications. The big data workload stresses the entirety of the hardware and software stack including the CPU, memory and I/O subsystems and the network by performing a series of writes, sorts and reads across all servers in the cluster.

# Tuning Cloudera on Microsoft Azure Das v4-Series Instances

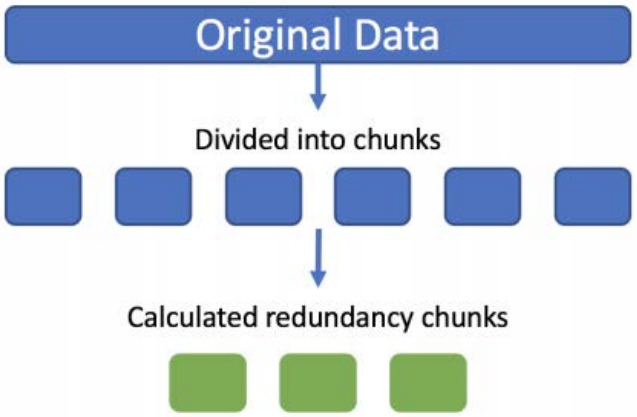
This table shows benchmark parameters and various tunable parameters on HDFS, YARN and MAPREDUCE on Cloudera to optimize big data workload performance for each cluster configuration:

| Parameter | D16asv4 16-Node Value | D32asv4 16-Node Value |
|---|---|---|
| **BENCHMARK** | | |
| NUM_MAPS | 240 | 480 |
| NUM_REDUCERS | 240 | 480 |
| HDFS | | |
| dfs.blocksize | 512 MB | 128 MB |
| dfs.replication | 3 | 3 |
| dfs.datanode.handler.count | 16 | 16 |
| dfs.namenode.handler.count | 200 | 200 |
| Erasure Coding Policy (See Erasure Coding: Data Redundancy Done Right) | hdfs ec -setPolicy -path /user/root/ -policy RS-6-3-1024k | |
| YARN/MAPREDUCE | | |
| mapreduce.reduce.memory.mb | 3 GB | 3 GB |
| mapreduce.map.memory.mb | 3 GB | 3 GB |
| mapreduce.client.java.heapsize | 2400 MB | 2400 MB |
| yarn.app.mapreduce.am.resource.mb | 3 GB | 3 GB |
| yarn.scheduler.maximum.allocation.mb | 100352 MB | 100352 MB |
| mapreduce.input.fileinputformat.split.minsize | 4166666700 | 2083333400 |
| mapreduce.map.sort.spill.percent | 0.99 | 0.99 |
| mapreduce.map.speculative | TRUE | TRUE |
| mapreduce.reduce.speculative | TRUE | TRUE |
| mapreduce.reduce.shuffle.parallelcopies | 64 | 128 |
| zlib.compress.level | BEST_SPEED | BEST_SPEED |
| mapreduce.task.io.sort.mb | 256 | 512 |
| mapreduce.task.io.sort.factor | 64 | 64 |
| mapreduce_map_java_opts | -Djava.net.preferIPv4Stack=true -Xmx1900m -XX:+UseParallelGC -XX:ParallelGCThreads=4 -XX:-UseAdaptiveSizePolicy -XX:+DisableExplicitGC | |
| mapreduce_reduce_java_opts | Same as mapreduce_map_java_opts | |

# Erasure Coding: Data Redundancy Done Right

Erasure coding is an alternative to three-way replication which follows the principles of RAID5 and RAID6. It breaks the data up into chunks and then adds some additional chunks of data which can be used to recover any data that is lost. All the chunks are then distributed across the physical media.

The primary advantage of erasure coding is that it can achieve the same or better protection against data loss as three-way replication while helping reduce the total amount of stored data. As an example, an erasure coding configuration with 6 data blocks and 3 redundancy blocks will have a storage overhead of 50%, while a replicated system with two full copies of the data (standard three-way replication) will have 200% storage overhead. Thus, in total the amount of stored data will be 50% less with the erasure coded system.




Support for erasure coding was added in version 3 of the Hadoop Distributed File System (HDFS). Erasure coding comes with some overhead. In particular, the redundancy segments need to be calculated and thus require some compute resources. However, the Cloudera performance results shown in Fig. 1 and 2 used the RS-6-3-1024K erasure coding policy, the Das v4-series VMs powered by the AMD EPYC SoC is more than up the task of handling the extra load needed to implement erasure coding.

## Conclusion

As emphasized above, best performance for big data applications is the synergic outcome of number of VCPUs, size of memory, storage with high throughput IOPS and network speed configured on the instance types used on the infrastructure side. It is very important to choose the right instance types for various components of Cloudera cluster that can handle the required load with expected response time for big queries. Microsoft Azure offers a variety of sizes in Da, Ea, Das, Eas v4-series instances powered by AMD EPYC™ 7452 processors. The high throughput and IOPS of the Premium SSD disk storage makes the Das and Eas v4-series VMs ideal for systems like the Cloudera implementation of Apache Hadoop which replicates data across multiple VMs to achieve fault tolerance in the event of the failure of a single VM.

Microsoft Azure provides a scalable computing platform that allows you to only pay for what you use, when you want it - without having to invest in on-premises server hardware. Azure is ready when you are to scale your solutions up and out to whatever scale you require to service the needs of your business.

Together AMD and Azure enable the implementation of a wide variety of application solutions. Here we have shown both excellent performance and linear scalability for Cloudera® based big data.

## Get started at azure.microsoft.com* today

*Links to third party sites are provided for convenience and unless explicitly stated, AMD is not responsible for the contents of such linked sites and no endorsement is implied

FOOTNOTES

1. D16asv4 VMs each with 16 vCPU, 64 GiB memory, Local SSD 128 GiB, Premium disk support: Yes, OS: "Red Hat Enterprise Linux Server" VERSION="7.6 (Maipo)" : Kernel version 3.10.0-957.21.3.el7.x86_64, Premium SSD – 3 x (P50, Disk size in GiB: 2048, IOPS per disk : 7500, Throughput per disk: 250 MiB/sec); D32asv4 VMs each with 32 vCPU, 128 GiB memory, Local SSD 256 GiB, Premium disk support: Yes, OS: "Red Hat Enterprise Linux Server" VERSION="7.6 (Maipo)" : Kernel version 3.10.0-957.21.3.el7.x86_64, Premium SSD – 3 x (P50, Disk size in GiB: 2048, IOPS per disk : 7500, Throughput per disk: 250 MiB/sec). Performance measured with Cloudera Distribution Hadoop 6.3.1.