

BIG DATA IS A WORLD OF POSSIBILITY

Analytics solutions for big data to scale as you grow

Stay agile in a dynamic data world with the combined power of Cloudera and AMD

THE STATE OF DATA

In order to scale, enterprises must find a way to manage growing quantities of data — which is expected to reach 463 exabytes globally by 2025.¹

For perspective, some estimate that all the words ever spoken by humankind equal only five exabytes.² An exabyte is the equivalent of one quintillion bytes, or one billion gigabytes.

2.5

quintillion bytes of data is generated by Internet users alone every day³

99%

of mainstream companies have made big data investments⁴

39.3%

of mainstream companies are effectively managing data as an asset⁴

1 quintillion would be equal to streaming 25.6 billion hours of music on Spotify.

BECOMING A DATA-DRIVEN ORGANIZATION MEANS FINDING A WAY TO ACCELERATE THE PROCESSING OF HIGH-VOLUME DATA IN REAL TIME.

CLUDERA DATA PLATFORM (CDP), RUNNING ON AMD EPYC™ CPU-BASED SERVERS, FUELS ANALYTICS PERFORMANCE ACROSS ORGANIZATIONS.

ANALYTICS CHALLENGES

Cloudera and AMD: Big Data analytics solutions that scale as businesses grow.



Performance

Peak performance with Cloudera hybrid data platform and AMD EPYC™ processors.

- Faster processing
- Comprehensive data services
- Optimized insights



Security

Comprehensive, always-on security with Cloudera Shared Data Experience (SDX) and AMD Infinity Guard.

- Protection at rest, in transit and in use
- Encryption keys for virtual machines
- Compliance and governance across cloud environments



Sustainability

Avoid tech sprawl and reduce footprint with Cloudera hybrid data platform and AMD EPYC™ processors.

- Reduce total cost of ownership (TCO)
- Support IT teams with single-point solutions
- Lower energy costs and greater energy efficiency

Better for business, better for our world.

30x

Increase in energy efficiency for processors and accelerators powering servers for AI and HPC from 2020-2025⁵

60%

Higher performance on Apache Hadoop Performance vs. Intel Processors⁶

127%

Faster analytics using just one AMD EPYC CPU vs. two Intel XEON Gold CPUs⁶

72%

Lower cost per sort using one AMD EPYC CPU vs. two Intel XEON Gold CPUs⁶

POWERING POSSIBILITY

The technical partnership between AMD and Cloudera opens the door for organizations to accelerate high-volume data processing in real time.

It's a world of possibility with a world-class partnership. Let's go.

Learn more about what AMD and Cloudera can do for you by visiting: cloudera.com/partner/amd

CLUDERA

¹Vuleta, B. (2022, May 25). How much data is created every day? [27 powerful stats]. SeedScientific. Retrieved June 1, 2022, from <https://seedscientific.com/how-much-data-is-created-every-day/>

²The exponential pull of Innovation Googlisation 2.0: SEI. SEI. (n.d.). Retrieved June 8, 2022, from <https://seic.com/en-gb/read-online-exponential-pull-innovation-googlisation-203Data>

³Data Never Sleeps 5.0 <https://www.domo.com/learn/infographic/data-never-sleeps-5>

⁴Bean, R. (2021). *Big Data and AI Executive Survey 2021*. Boston, MA; NewVantage Partners.

⁵Includes AMD high-performance CPU and GPU accelerators used for AI training and High-Performance Computing in a 4-Accelerator, CPU hosted configuration. Goal calculations are based on performance scores as measured by standard performance metrics (HPC: Linpack DGEMM kernel FLOPS with 4k matrix size. AI training: lower precision training-focused floating-point math GEMM kernels such as FP16 or BF16 FLOPS operating on 4k matrices) divided by the rated power consumption of a representative accelerated compute node including the CPU host + memory, and 4 GPU accelerators.

⁶MLN-070 TPC Benchmark Express™ HSpH@3TB comparison based on highest system results published as of 03/12/2021. Configurations: 17-node, 1x AMD EPYC 75F3 (34.52 HSpH@3TB, \$35,615.50/HSpH@3TB, <http://www.tpc.org/5548>, available 3/15/2021, Framework 1 MapReduce) versus 2x Intel Xeon Gold 6262V (15.21 HSpH@3TB, \$129,143.92/HSpH@3TB, <http://www.tpc.org/5546>, available 11/09/2020, Framework 1 MapReduce) for 127% higher [2.3x the] Hadoop Sort throughput performance at 72% better price/performance. <https://www.amd.com/en/claims/epyc>