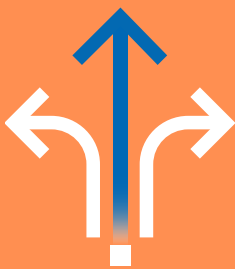


Achieve up to 90% More Throughput for Apache Kafka® on Confluent Platform™ with AWS EC2 R5 Instances Featuring 2nd Gen Intel® Xeon® Scalable Processors

Kafka



Get up to 90% more Apache Kafka throughput with AWS r5.xlarge instances with 2nd Gen Intel Xeon Scalable processors

Vs. AWS r4.xlarge instances



Get up to 23% more Apache Kafka throughput with AWS r5.xlarge instances with 2nd Gen Intel Xeon Scalable processors

Vs. AWS c5.xlarge instances

R5 Instances Outperformed Both R4 Instances and C5 Instances with the Same or Previous-Generation Processors

Organizations seeking to run real-time data streaming workloads on the Amazon Web Services (AWS) EC2 cloud may find it difficult to discern which instance type offers the Kafka performance they need. To show how popular AWS instances stack up, we compared the Apache Kafka® on Confluent Platform™ performance of three instance types in a Kubernetes containerized environment:

- AWS R5 instances with 2nd Gen Intel® Xeon® Scalable processors
- AWS R4 instances with Intel Xeon E5-2686 v4 processors
- AWS C5 instances with 1st or 2nd Gen Intel Xeon Scalable processors

Using the Kafka Producer Performance test built into Apache Kafka 2.7.0, AWS R5 instances with 2nd Gen Intel Xeon Scalable processors improved throughput for Kafka on Confluent Platform over previous-gen R4 instances by 90%. This shows that for real-time data platform workloads, selecting AWS R5 instances with newer processor technology can offer better Kafka on Confluent Platform streaming performance than instances built on previous processor generations. Additionally, the memory-optimized R5 instances offered 23% more throughput than the compute-optimized C5 instances. By processing and moving data through the pipeline at a faster rate, these instances can handle more customer requests per instance without noticeable latency.

Get Better Kafka Performance from AWS R5 Instances

The first round of testing compared 12-node clusters of two memory-optimized instance types: the AWS R5 instances with 2nd Gen Intel Xeon Scalable processors to older R4 instances (see Figure 1). Due in part to the newer processors, the R5 instances delivered nearly double the Kafka throughput of R4 instances with older processors.

Normalized Kafka Performance

Normalized throughput | Higher is better

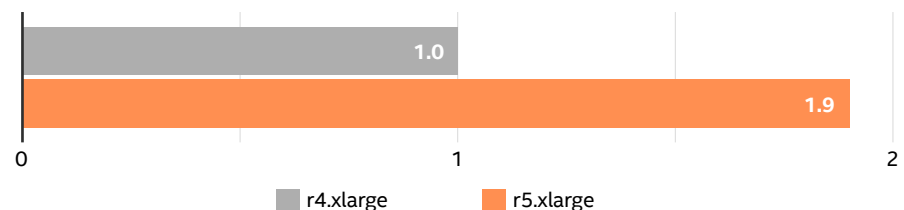


Figure 1. Relative Apache Kafka throughput for AWS R4 instances vs. AWS R5 instances with 2nd Gen Intel Xeon Scalable processors. Higher numbers are better.

Comparing Kafka Performance Across Instance Types

As Figure 2 shows, 12-node memory-optimized AWS R5 instance clusters with 2nd Gen Intel® Xeon® Scalable processors also outperformed 12-node compute-optimized AWS C5 instance clusters, which run on 1st or 2nd Gen Intel Xeon Scalable processors. Compared to the C5 instances, AWS R5 instances offered 23% more Kafka throughput—a strong increase in the number of events that each instance can handle.

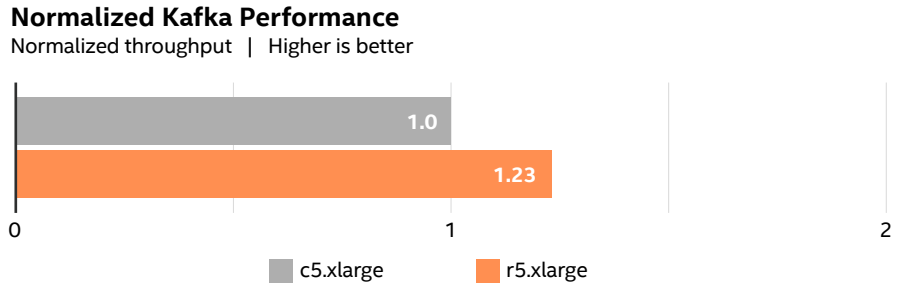


Figure 2. Relative Apache Kafka throughput for AWS C5 instances vs. AWS R5 instances with 2nd Gen Intel Xeon Scalable processors. Higher numbers are better.

Test configurations

Figure 3 shows how we configured each Apache Kafka cluster in our tests for each VM type.

Apache Kafka on Confluent Platform Environment					
Name	Instance Type	# of Instances	vCPU Count/ Instance	Memory (GiB)/ Instance	Network Performance
Ansible-Inventory-Node	t2.micro	1	1	1	Low to Moderate
Confluent-Control-Center	m4.large	1	2	8	Moderate
Confluent-Schema-Registry	m4.large	1	2	8	Moderate
Zookeeper	m4.large	3	2	8	Moderate
Producer	m5.2xlarge	3	8	32	Up to 10 Gigabit
Broker type config 1	r4.xlarge	12	4	30.5	Up to 10 Gigabit
Broker type config 2	c5.xlarge	12	4	16	Up to 10 Gigabit
Broker type config 3	r5.xlarge	12	4	32	Up to 10 Gigabit

Figure 3. The Apache Kafka cluster VMs for each configuration.

Conclusion

With distributed event streaming platforms, organizations want to process customer transactions and interactions in real time—without significant delays. These tests show that businesses running Apache Kafka on Confluent Platform on the AWS cloud can handle more events by selecting AWS R5 instances with 2nd Gen Intel Xeon Scalable processors, enabling them to provide faster and scalable performance for customers.

Learn More

To begin running Kafka workloads on Amazon EC2 R5 instances, visit <https://aws.amazon.com/ec2/instance-types/r5/>.

12-VM cluster tests by Intel Jul-Aug 2021. All configs EBS storage on CentOS 7 3.10.0-1160.6.1.el7.x86_64 Confluent Platform 6.0.0-post openjdk version "1.8.0_292" Apache Kafka 2.7.0 (Kafka Producer Performance test) Producer Settings: Ingestion Rate: 120000 records/sec, Record Size: 1KB, Run Duration: 10mins, Number of topics: 1, Topic Partitions: 24, Broker Settings: log.dirs:/dev/sda1(EBS), num.io.threads:16, num.network.threads:8, num.partitions:1. VMs: r4.xlarge, 4vcpus, Intel Xeon® E5-2686 v4, 30.5 GB total DDR4 memory; r5.xlarge, 4vcpus, Intel Xeon® Platinum 8000 series processors, 32 GB total DDR4 memory; c5.xlarge, 4vcpus, Intel Xeon® Platinum 8000 series processors, 8 GB total DDR4 memory



Performance varies by use, configuration and other factors. Learn more at www.intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure. Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Printed in USA 0422/JO/PT/PDF US001

