THE GOAL

MOVE DATA FASTER

IMPROVE APPLICATION PERFORMANCE WITH APPLICATION DEVICE QUEUES (ADQ)

IMPROVE PACKET PROCESSING EFFICIENCY WITH DYNAMIC DEVICE PERSONALIZATION (DDP)

IMPROVE DATA MOVEMENT WITH CONCURRENT SUPPORT FOR IWARP AND ROCE V2 RDMA
SCALE-OUT APPLICATION PERFORMANCE PARAMETERS

PREDICTABILITY

LATENCY

THROUGHPUT
Why Application Response Predictability Matters

Meeting the Scale-out Challenge
- Reducing variability in application response time (jitter) improves throughput and reduces latency

Benefits of Reducing Jitter
- More servers can be added to parallelize task
- Support more end-users with existing hardware

Higher Predictability Enables More Servers Working in Parallel Within a Desired Response Time

1. “The Tail at Scale” – Communications of the ACM. February 2013
Jeffrey Dean – Google Senior Fellow and Luiz André Barroso – Google Fellow / VP of Engineering
https://cseweb.ucsd.edu/~gmporter/classes/fa17/cse124/post/schedule/p74-dean.pdf
+ SLA = Service-Level Agreement
HOW TO IMPROVE PREDICTABILITY
ANALOGY: TIME TO REACH AIRPORT FOR FLIGHT

APPLICATION DEVICE QUEUES (ADQ) IMPROVES PREDICTABILITY WITH DEDICATED LANES AND RATE LIMITING
What is ADQ?
- An application specific queuing and steering technology

How does ADQ work?
- Filters application traffic to a dedicated set of queues
- Application threads of execution are connected to specific queues within the ADQ queue set
- Rate limits application egress (Tx) network traffic

What are the benefits of ADQ?

With ADQ
Application traffic to a dedicated set of queues

Without ADQ
Application traffic intermixed with other traffic types

1. Features & schedule are subject to change. All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
2. See Demo or Performance Testing Application Device Queues (ADQ) with Redis® Solution Brief
INTEL® ETHERNET – APPLICATION DEVICE QUEUES (ADQ)¹
LATEST NETWORK TECHNOLOGY INNOVATION FOR INTEL® ETHERNET 800 SERIES

ADQ Hardware Components
Packet Pipeline and Traffic Classification

ADQ Basics
- Filters application traffic to a dedicated set of queues
- Application threads of execution are connected to specific queues within the ADQ queue set
- Rate limits application egress (Tx) network traffic

Inbound Traffic Flows
Outbound Traffic Flows

Intel® Ethernet 800 Series Controllers and Network Adapters

ADQ Basics
- Filters application traffic to a dedicated set of queues
- Application threads of execution are connected to specific queues within the ADQ queue set
- Rate limits application egress (Tx) network traffic

Inbound Traffic Flows
Outbound Traffic Flows

Intel® Ethernet 800 Series Controllers and Network Adapters

ADQ Hardware Components
Packet Pipeline and Traffic Classification

ADQ Basics
- Filters application traffic to a dedicated set of queues
- Application threads of execution are connected to specific queues within the ADQ queue set
- Rate limits application egress (Tx) network traffic

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

ADQ Basics
- Filters application traffic to a dedicated set of queues
- Application threads of execution are connected to specific queues within the ADQ queue set
- Rate limits application egress (Tx) network traffic

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Inbound Traffic Flows
Outbound Traffic Flows

Component | Requirement
---|---
APPLICATION | Case 1: No change (e.g. single instance) Case 2: Intel provided reference code
CONFIGURATION | Standard OS Tools (Traffic Configuration (TC), ethtool)
OS | Linux 4.19 or later
ETHERNET DRIVER | Intel® Ethernet 800 Series Driver
ETHERNET NIC | Intel® Ethernet 800 Series

¹ Features & schedule are subject to change. All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
“Getting useful insights in real-time out of Big Data comes with a set of major challenges such as predictable low latency and maximum throughput at the network layer. Aerospike, as always, is at the forefront of addressing these challenges. We expect the Intel® Ethernet 800 Series with Application Device Queues (ADQ) coupled with Aerospike Enterprise will help get predictable performance, higher data throughput and lower latency. We are pleased to work closely with Intel to bring this exciting new technology to our customers.”

Srini Srinivasan, Founder and Chief Product Officer
Aerospike real-time database ecosystem

- Throughput > 1M TPS
- Latency < 1ms
- Reliability > five 9s
- Scale up to petabytes
- Strong consistency
- TCO < 2-5X of other DBs

Consumers require customized real-time user experience

DB must deliver predictable performance at scale

- Location A (SOE)
- Location B (SOE)
- Location C (SOE)
- Location D (SOE)

Real-time System of Record

Single Source of Truth

Petabyte scale

Terabyte scale

Query & Reporting Store (SOR)

Legacy Data Store

AI/ML Engines

Predictive Analytics

WHAT IS THE NEXT BEST ACTION? (in milliseconds)

WHAT IS THE BEST WAY TO AUGMENT MODEL? (in seconds or minutes)

Spark Streaming

XDR

Consumers require customized real-time user experience
Aerospike delivers speed at scale using Hybrid-Memory Architecture

**Highlights**

1. Direct device access
2. Large Block Writes
3. Indexes in DRAM
4. Highly Parallelized
5. Log-structured FS “copy-on-write”
6. Fast restart with shared memory
Aerospike is designed for wire-line speed

Multi-core Architecture

Optimized C based DB kernel
1. Multi-threaded data structures (NUMA pinned)
2. Nested locking model for synchronization
3. Lockless data structures
4. Partitioned single threaded data structures
5. Index entries are aligned to cache line (64 bytes)
6. Custom memory management (arenas)

Memory Arena Assignment
AEROSPIKE DEMO
INTEL® ETHERNET 800 SERIES BASELINE VS. WITH APPLICATION DEVICE QUEUES (ADQ) (NETPERF* - THROUGHPUT)

Higher is better

Throughput (Transactions/sec)

1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93

Throughput (Transactions/sec)

1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87 89 91 93

1

3

5

7

9

11

13

15

17

19

21

23

25

27

29

31

33

35

37

39

41

43

45

47

49

51

53

55

57

59

61

63

65

67

69

71

73

75

77

79

81

83

85

87

89

91

93

Intel® Ethernet 800 Series Baseline

Intel® Ethernet 800 Series with ADQ

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

Source: Performance results are based on Intel internal testing as of March 2019, and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure. Tests performed using Netperf 2.7.0 Open Source on 2nd Generation Intel® Xeon® Scalable processors with Intel® Ethernet 800 Series 100GbE on RedHat 7.5 Linux (4.19.15 stable kernel) (see backup).

*Other names and brands may be claimed as the property of others.
Higher is better

Lower is better

**INTEL® ETHERNET 800 SERIES BASELINE VS. W/ADQ**

**(NETPERF* - MEAN LATENCY)**

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

Source: Performance results are based on Intel internal testing as of March 2019, and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure. Tests performed using Netperf 2.7.0 Open Source on 2nd Generation Intel® Xeon® Scalable processors with Intel® Ethernet 800 Series 100GbE on RedHat 7.5 Linux (4.19.15 stable kernel) (see backup).

*Other names and brands may be claimed as the property of others.
INTEL® ETHERNET 800 SERIES BASELINE VS. W/ADQ
(NETPERF* – STD DEV LATENCY & P99 LATENCY)

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

Source: Performance results are based on Intel internal testing as of March 2019, and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure. Tests performed using Netperf 2.7.0 Open Source on 2nd Generation Intel® Xeon® Scalable processors with Intel® Ethernet 800 Series 100GbE on RedHat 7.5 Linux (4.19.15 stable kernel) (see backup)

*Other names and brands may be claimed as the property of others.
NETPERF DEMO
SUMMARY
INTEL® ETHERNET 800 SERIES WITH APPLICATION DEVICE QUEUES (ADQ) DELIVERS

- PREDICTABILITY
- LATENCY
- THROUGHPUT

FAST, PREDICTABLE NETWORKING FOR FAST COMPUTE AND MEMORY
NOTICES AND DISCLAIMERS

Intel technologies’ features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Performance results are based on testing as of February 2019 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or computer system can be absolutely secure. For more complete information visit http://www.intel.com/benchmarks.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document. Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, the Intel logo and others are trademarks of Intel Corporation in the U.S. and/or other countries. Apache, Spark, SQL, Kafka, the Apache Logos and others are trademarks of Apache Software Foundation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

© 2019 Intel Corporation.
**INTEL® ETHERNET ARCHITECTURE EVOLUTION**

**FOUNDATIONAL NICS**

<table>
<thead>
<tr>
<th>More</th>
<th>Less</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Pipeline</td>
<td>Partially Programmable Pipeline</td>
</tr>
<tr>
<td>Intel® Ethernet 500 Series Niantic</td>
<td>Intel® Ethernet Adaptive Virtual Function (Intel AVF)</td>
</tr>
<tr>
<td>Intel® Ethernet 700 Series Fortville</td>
<td>Table definition modifications with a Dynamic Device Personalization (DDP) profile package</td>
</tr>
<tr>
<td>Intel® Ethernet 800 Series Columbiaville³</td>
<td>Storage</td>
</tr>
<tr>
<td></td>
<td>RDMA (iWARP* &amp; RoCE*v2)</td>
</tr>
</tbody>
</table>

Queue and Steering Hardware Assists
- Application Device Queues (ADQ)
- Fully Programmable Pipeline
- Table definition with DDP profile packages

Features & schedule are subject to change. All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.
## INTEL(R) 800 SERIES BASELINE VS. INTEL® ETHERNET 800 SERIES
### WITH APPLICATION DEVICE QUEUES (ADQ) NETPERF® TEST CONFIGURATION
#### (PRE-PRODUCTION COLUMBIAVILLE AND PRE-PRODUCTION CASCADE LAKE)

**SUT Configuration**

<table>
<thead>
<tr>
<th>Platform</th>
<th>Intel® Server Board S2600WFTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>Red Hat Enterprise Linux Server 7.5 (Maipo) (Kernel 4.19.15.stable)</td>
</tr>
<tr>
<td>Processor</td>
<td>Cascade Lake (Pre-production)</td>
</tr>
<tr>
<td>QPI/UPI Speed</td>
<td>10.4 GT/s</td>
</tr>
<tr>
<td>Memory Amt/Type/Speed</td>
<td>536 GB, 2400 MHz</td>
</tr>
<tr>
<td>Chipset</td>
<td></td>
</tr>
<tr>
<td>BIOS Version</td>
<td>SE5C620.868.00.00.01.0134.091920181249</td>
</tr>
<tr>
<td>NIC/DUT</td>
<td>Intel® Ethernet X810 (&quot;Tacoma Rapids&quot;) A1</td>
</tr>
<tr>
<td>NIC/DUT Driver filename/version</td>
<td>ice.0.8.15, ice.0.8.15</td>
</tr>
<tr>
<td>TIC</td>
<td>521059</td>
</tr>
<tr>
<td>NIC EEPROM/NVM</td>
<td>0.00 0x800000941 255.65535.255</td>
</tr>
<tr>
<td>PCIe Generation</td>
<td>Gen 3</td>
</tr>
<tr>
<td>PCIe Lane Width</td>
<td>+16</td>
</tr>
</tbody>
</table>

**Traffic Generator Configuration**

<table>
<thead>
<tr>
<th>Platform</th>
<th>Intel® Server Board S2600WFTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>Red Hat Enterprise Linux Server 7.5 (Maipo) (Kernel 4.19.15.stable)</td>
</tr>
<tr>
<td>NIC (Port/Speed)</td>
<td>Intel® Ethernet X810 (&quot;Tacoma Rapids&quot;) A1</td>
</tr>
</tbody>
</table>

**SUT Virtualization Configuration**

<table>
<thead>
<tr>
<th>Virt. Technology</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host VMs</td>
<td>N/A</td>
</tr>
<tr>
<td>RAM &amp; VCPU per VM</td>
<td>N/A</td>
</tr>
<tr>
<td>Guest OS</td>
<td>N/A</td>
</tr>
<tr>
<td>VCPU Affinity Configuration</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**BIOS Version**

<table>
<thead>
<tr>
<th>OS/Kernel Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>stopped and disabled: irqbalance, cpupower, firewalld, rngd, NetworkManager ptables - F</td>
</tr>
<tr>
<td>SELINUX disabled</td>
</tr>
<tr>
<td>scaling_governor set to performance for every core</td>
</tr>
</tbody>
</table>

**Notes (Configuration to enabled ADQ)**

<table>
<thead>
<tr>
<th>Notes (Configuration to enabled ADQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back-to-back setup with two X810-Tacoma Rapids A1. Ice driver with compiled with ADQ flags. make -j $(nproc) CFLAGS_EXTRA=-DADQ_PERF=-DADQ_PERF_COUNTERS install BKC ADQ configuration for Netperf (Revision 1.0, February 11, 2019)</td>
</tr>
</tbody>
</table>

*Other names and brands may be claimed as the property of others.*
# NETPERF* SYSTEM UNDER TEST AND CLIENT CONFIGURATIONS

<table>
<thead>
<tr>
<th></th>
<th>Config1 (Client)</th>
<th>Config2 (Server)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test by</td>
<td>Intel</td>
<td>Intel</td>
</tr>
<tr>
<td>Test date</td>
<td>02/27/2019</td>
<td>02/27/2019</td>
</tr>
<tr>
<td>Platform</td>
<td>Intel® Server Board S2600WFTF</td>
<td>Intel® Server Board S2600WFTF</td>
</tr>
<tr>
<td># Nodes</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td># Sockets</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CPU</td>
<td>CLX K0 3.1 GHz 18C 24.75MB Cache, 205W, QQZ0</td>
<td>CLX K0 3.1 GHz 18C 24.75MB Cache, 205W, QQZ0</td>
</tr>
<tr>
<td>Cores/socket, Threads/socket</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>ucode</td>
<td>0x400000a</td>
<td>0x400000a</td>
</tr>
<tr>
<td>HT</td>
<td>On</td>
<td>On</td>
</tr>
<tr>
<td>Turbo</td>
<td>On</td>
<td>On</td>
</tr>
<tr>
<td>BIOS version</td>
<td>SE5C620.86B.0D.01.0134.091920181249</td>
<td>SE5C620.86B.0D.01.0134.091920181249</td>
</tr>
<tr>
<td>System DDR Mem Config: slots / cap / run-speed</td>
<td>2 slots / 568GB / 2400</td>
<td>2 slots / 536 GB / 2400</td>
</tr>
<tr>
<td>System DCPMM Config: slots / cap / run-speed</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total Memory/Node (DDR+DCPMM)</td>
<td>568 GB</td>
<td>536 GB</td>
</tr>
<tr>
<td>OS</td>
<td>RedHat 7.5</td>
<td>RedHat 7.5</td>
</tr>
<tr>
<td>Kernel</td>
<td>4.19.15.stable</td>
<td>4.19.15.stable</td>
</tr>
<tr>
<td>IBRS (0=disable, 1=enable)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>eIBRS (0=disable, 1=enable)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Retpoline (0=disable, 1=enable)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IBPB (0=disable, 1=enable)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PTI (0=disable, 1=enable)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mitigation variants (1,2,3,3a,4, L1TF)</td>
<td>1,2,3,3a,4, L1TF</td>
<td>1,2,3,3a,4, L1TF</td>
</tr>
<tr>
<td>Workload &amp; version</td>
<td>netperf 2.7.0</td>
<td>netperf 2.7.0</td>
</tr>
<tr>
<td>Compiler</td>
<td>GCC</td>
<td>GCC</td>
</tr>
</tbody>
</table>

*Other names and brands may be claimed as the property of others.