

Deployment Platform Performance Guidelines

Deployment Platform Performance Guidelines¹

Max Aggregate Throughput ³	Supportable Concurrency ⁴	WAN ⁸ Network Interface ⁵	Platform ⁷	RAM	CPU ²	Number Cores ⁶	Storage ^{11, 13}
100 Mb/s	Single Transfer Low (≤10) High (>10)	100 MbE (≤100 Mbps) 1 GbE (>100 Mbps) WiFi9 and 4G9	Laptop or PC with Win/ Linux/Mac or Mobile/ Embedded Platform	2 GB	1.7 GHz x86 or ARM processor	2 cores	5400 rpm SATA
500 Mb/s	Single Transfer Low (≤10) High (>10)	1 GbE	Laptop or Desktop with Win/Linux/Mac or Equivalent Entry Level Server	4 GB	2.2 GHz x86 processor or greater	2 cores ¹⁴ 4 cores ¹⁴	SSD or Single 10k or 15k rpm SAS or SCSI Multiple block devices (disks or SSD) in a direct attached RAID or SAN, capable of 500 Mbps throughput
1 Gb/s	Single Transfer Low (≤10) High (>10)	1 GbE	Pro Laptop or Desktop with Win/Linux/Mac Workstation or Equivalent Server	8 GB	2.7 GHz x86 processor or greater	4 cores	SSD, Spinning disk DAS with RAID; High performance NAS (separate NIC ⁸ for storage vs. WAN traffic), SAN (Fiber Channel HBA or min 1 Gbps NIC dedicated for iSCSI)
2.5 Gb/s	Single Transfer Low (≤10) High (>10) Very High (>50)	10 GbE	Server-class machine with Linux or Win	8 GB	2.9 GHz x86 processor or greater	8 cores	Single SSD or PCIe based SSD, DAS with RAID, NAS (separate 10 GbE NIC ⁸ for storage vs WAN traffic), SAN (Fiber Channel HBA rated to target speed)
Isilon S200 – 3-node cluster. Aspera transfers execute on a single node with a 10 Gbps interface. Data is striped across all nodes ¹⁵							
5 Gb/s	Single Transfer Low (≤10) High (>10) Very High (>50)	10 GbE	Server-class machine with Linux	16 GB	3.3 GHz x86 processor or greater	8 cores	Single PCIe based SSD, DAS with RAID, NAS (separate 10GbE NIC ⁸ for storage vs WAN traffic), SAN (Fiber Channel HBA rated to target speed)
7.5 Gb/s ¹⁰	Single Transfer Low (≤10) High (>10) Very High (>50)	10 Gb/s ¹²	High Performance Server-class machine with Linux	24 GB	3.5 GHz x86 processor or greater	8 cores	Single PCIe based SSD, DAS with RAID, NAS (separate 10GbE NIC ⁸ for storage vs WAN traffic), SAN (Fiber Channel HBA rated to target speed)
10 Gb/s ⁺¹⁰	Single Transfer Low (≤10) High (>10) Very High (>50)	10 Gb/s ¹²	High Performance Server-class machine with Linux	24 GB	3.5 GHz x86 processor or greater	16 cores	RAID of PCIe based SSD, DAS with RAID-ed SSD or disks, High Performance NAS or SAN (rated to target speed)

Chart provided as a general guideline and assumes that the Aspera software is the only application running on the computer during transfers. For example, Anti-Virus tools can negatively affect performance. Performance may also vary due to user's computer, storage, network equipment manufacturer and IT infrastructure differences

Deployment Platform Performance Guidelines¹

Notes for Deployment Platform Performance Guidelines Chart:

1. The server systems shown here are conservative examples to achieve the given performances with a single server box or blade. It is also possible to use lower-end (i.e. less expensive) boxes or blades with higher-end processor configurations or in a clustered, high-availability (active-active) configuration to achieve the same aggregate throughputs.
2. Transfers can be encrypted or unencrypted. Encrypted transfers incur a higher processor load. Newer Intel and AMD processors supporting AES-NI will have higher performance for encrypted transfers. A list of Intel processors with AES-NI can be found at <http://ark.intel.com/search/advanced/?s=t&AESTech=true>. Some specialized ARM processors support AES-NI, but generally ARM does not have it.
3. Single file transfers at this rate have been measured with a file size of ≥100MB. These rates have also been measured with multiple, smaller files where the mean file size was ≥500KB. In real-world terms, this can be thought of as the difference between transferring a movie file vs. a large set of individual frame files, or a single disk image vs. a large directory of document scans.
4. These values are indicative of common concurrency ranges. It is possible to support hundreds or even thousands of concurrent low-speed (e.g. 1Mbps) uploads (receives) on a single host (typically 1Gbps WAN connected) with the appropriate RAM & CPU configuration. For specialized applications, contact Aspera for additional configuration guidance.
5. Testing has shown that some 1GigE and 10GigE cards do not support the full 1Gbps and 10Gbps data rates, respectively. Intel cards tested by Aspera have consistently been able to perform at sustained data rates. Aspera has not exhaustively tested all chipsets and cards – customers have reported success with other network interface vendors and chips.
6. Virtual Machines: The same component specifications apply for equivalent performance on Virtual Machine/Hypervisor configurations. ‘Overprovisioning’ of physical resources among shared virtual instances will degrade performance, e.g. I/O contention, CPU contention, etc. Virtualized disk containers (e.g. qcow2 or vmdk) can impose a performance penalty.
7. A modern Linux is the preferred platform for performance use of Aspera. While examples of 10Gbps configurations exist with Windows and Mac, Linux is the overwhelming leader in number of deployments and understanding on how to be properly configured.
8. The WAN Network Interface is indicative of the interface used for FASP transfers. If a host is utilizing NAS storage, iSCSI, or Object storage then additional network interfaces will be necessary.
9. WiFi and 4G connectivity is subject to multiple variations not covered in this table, which affect ultimate performance. For the purposes of this table, the connectivity is assumed to support the expected data rate.
10. Due to the system-level complexities involving processing, networking and storage, customers are strongly urged to discuss specific Ultra-High Performance configurations with Aspera prior to purchasing systems.
11. The unit of measure in networking, for data rates, is expressed as –bits per second. Storage is typically expressed as –bytes per second. This table will follow networking convention and use bits when describing data rates to storage.
12. Users should qualify that network interface cards, intervening network equipment and data paths can transport the required multi-gigabit rates in a non-blocking manner.
13. Please see the sections: Storage Considerations and RAID Storage Considerations below. In general, the aggregate read and write throughput of the storage, as well as the interface/interconnect to the storage will play a big part in Aspera performance. For example, expecting 10Gbps transfers from a single 7200 rpm disk is not feasible.
14. The core count difference expressed in the split row is indicative of the added I/O necessary to support an additional NIC or HBA when using an external storage device. The lower core count is for use in environments where the storage is built-in and no software-RAID, HBA, or NIC is used.
15. Aspera ships on Isilon nodes. Multiple permutations of Isilon deployment can achieve varying Aspera targets. This example illustrates single-host performance as it scales with multiple storage nodes processing write I/O.

Deployment Platform Performance Guidelines

Guidelines Applicable Across All Products for Performing Aspera FASP® Transfers

1.1 The FASP® Protocol

- Provides reliable, high-speed transfer of file-based data
- The FASP protocol has no theoretical throughput limit
- FASP ensures determinism for each transfer aspect –
 1. Concurrency
 2. Encryption
 3. Speed/bandwidth
- Transfer speeds limited only by –
 - The source and receiver computer hardware and storage systems
 - Intervening network and equipment used for transfers

1.2 Main sources of bottlenecks

- Sender/receiver computer CPU
- The network capacity (due to b/w, architecture and equipment)
- Throughput to/from storage, primarily due to disk I/O capacity

1.3 Storage Considerations

Performance numbers are based on Aspera senders and receivers that are directly attached to the source and target storage (local disks, or SCSI-SAS/FC/SATA). Network attached storage (NAS such as CIFS/NFS) may create a bottleneck at high speeds for reading from source and/or writing to receiver-side disk if the disk I/O speed is insufficient to take full advantage of the network bandwidth.

With the most recent Linux kernels and leading NAS storage appliances, NFS can be specially configured to achieve single transfer speeds exceeding 1Gbps.

- Contact your storage vendor for server-side NFS tuning
- Internet searches will provide numerous references for client-side NFS tuning options

1.4 Storage Performance Profiling and Validation Notes

If you are interested in testing the throughput of your storage system, the following testing tools and methods may be helpful –

- IOzone – tool for random and sequential data generation and disk access performance testing. See <http://www.iozone.org/>
 - IOzone measures both sequential and random disk access
 - Profiles storage access throughput independently of network conditions
- Unix dd – A tool primarily for copying to/from STDOUT/STDIN or block devices, this tool can be used to test throughput of storage. Using specialized devices (like /dev/zero), it is possible to test single-threaded, sequential reads and writes.
 - dd should be used when IOzone is not available. dd's main strength is that is widely available on Unix and Linux systems, but it has limited testing capabilities compared to other tools.
 - Useful if trying to establish a baseline or starting point for additional testing.

For accurate results, employ large data sizes in the testing (e.g. single large file or many smaller files in aggregate). Test sets should be at least twice the size of the file cache to force actual disk access (not just cache access)

Deployment Platform Performance Guidelines

1.5 RAID Storage Considerations

There are numerous ways to configure a given type of storage, as well as choices and tradeoffs that can be made in the computing, network and security architectures. When considering storage, options include Direct-Attached Storage (DAS), Network-Attached Storage (NAS) or a Storage Area Network (SAN) depending on the required data volume, throughput rates and operational needs.

As an example, when considering DAS, the *High Throughput Performance Example Configurations* chart above notes using “DAS w/ RAID” as shorthand for saying that striped volume(s) should be implemented for optimizing disk throughput. When one moves away from a single disk configuration and begins considering volumes comprised of multiple disks, the selected RAID level can be an important factor in both the overall system performance and the fault tolerance of the storage subsystem. RAID stands for “*redundant arrays of independent disks*”. It’s a technology that allows multiple disks to be used as one, however it should be noted that each RAID “level” specifies a different strategy for achieving this, and there is a corresponding cost to each.

There are many on-line descriptions and discussions of RAID levels, configurations and their various characteristics. If you want to become familiar with basic RAID concepts, it is recommended you consult the following references:

<http://en.wikipedia.org/wiki/RAID>

http://en.wikipedia.org/wiki/RAID_storage

It can also be helpful to consult comparisons and performance tests such as the following:

<http://www.zdnet.com/blog/ou/comprehensive-raid-performance-report/484>

<http://www.pcguides.com/ref/hdd/perf/raid/levels/comp-c.html>

In reviewing these references, it will become clear that tradeoffs abound, which is why Aspera highly recommends being consulted in solution discussions with the customer, storage vendor and any additional parties involved in specifying, architecting, implementing, testing/validating and operating the system.

If your specific configuration or needs are not represented in this guide, please contact Aspera Customer Support

<https://www.ibm.com/products/aspera/support>.