



Below list of acronyms and definitions is gathered to provide a common terminology and understanding benefitting the high-performance scalable parallel file and storage systems community.

Acronyms

<u>ACRONYM</u>	<u>DEFINITION</u>
AFSB	Average File System Bandwidth
ANSI	American National Standards Institute
ASB	Average Storage Bandwidth
API	Application Programming Interface
BER	Bit Error Rate – the number of erroneous bits received divided by the total number of bits transmitted
CRC	Cyclic Redundancy Check
DARPA	Defense Advanced Research Projects Agency
DB	Database
DDR	Double Data Rate
DIF	ANSI T10 Data Integrity Field (aka Protection Information)
DIMM	Dual In-line Memory Module
DR	Disaster Recovery
DRAM	Dynamic Random Access Memory
DREN	Defense Research and Engineering Network
ECC	Error-Correcting Code
EEDC	Enhanced Ethernet for Data Center
EOL	End of Life
EOS	End of Service
FC	Fibre Channel
FTE	Full-time Equivalent
GbE	Gigabit Ethernet
GB	Gigabytes (10 ⁹ Bytes)
GPFS	General Parallel File System
HBA	Host Bus Adapter
HD	High Definition
HPCS	High Productivity Computing Systems
HPSS	High Performance Storage System
IB	InfiniBand
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IOPS	I/O requests per second
iSCSI	Internet Small Computer System Interface
LAN	Local Area Network
LBA	Logical Block Addressing

LTO	Linear Tape-Open
LUN	Logical Unit Number
MAID	Massive Array of Idle Disks
MHz	Megahertz
MIB	Management Information Base
MTBF	Mean Time Between Failures
NDA	Non-Disclosure Agreement
NIST	National Institute of Standards and Technology
OC-n	Optical Carrier, where the speed will equal $n \times 51.8 \text{ Mb/s}$
OSD	ANSI T10 Object-based Storage Device
PB	Petabytes (10^{15} Bytes)
PI	T10 Protection Information (ANSI T10 standard)
POSIX	UNIX interface standard
QDR	Quad Data Rate
RAID	Redundant Array of Independent Disks
RPM	Revolutions per minute
SAN	Storage Area Network
SAS	Serial Attached SCSI
SATA	Serial ATA (Advanced Technology Attachment)
SCSI	Small Computer System Interface
SHA	Secure Hash Algorithm
SLA	Service Level Agreement
SMI-S	A SNIA standard for management of devices
SNIA	Storage Networking Industry Association
SNMP	Standard Networking Management Protocol (IETF defined)
SSD	Solid-State Drive
T10	Technical committee responsible for SCSI storage interface standards
T11	Technical committee responsible for Fibre Channel interface standards
T13	Technical committee responsible for AT Attachment interface standards
TB	Terabytes (10^{12} Bytes)
TBD	To be determined
U	Unit Rack Height
UER	Undetectable Error Rate
WORM	Write once read many
XAM	eXtensible Access Method

Definitions:

Aggregate Bandwidth: In a file system composed of multiple underlying data stores (eg. multiple object storage targets), this is the sum of the bandwidth across all components (both reads and writes).

Average File System Bandwidth (AFSB): The calculated average bandwidth over multiple iterations of the same benchmark for either a fixed period of time or a fixed amount of data.

Average Storage Bandwidth (ASB): The average read and write bandwidth, calculated as follows:
Average Bandwidth = (Read Bandwidth+Write Bandwidth)/2, in GB/s.

Bandwidth: The data rate (a read rate or a write rate) of a file system or component, in bytes/second, as measured in some specific way. The value could be the result of a benchmark or it could be an observation of embedded performance counters (see *Monitoring*). The bandwidth of a file system is often meant to indicate the maximum achievable data rate under ideal circumstances. In other contexts it means the actual data rate observed under specific circumstances.

Bandwidth Density: The Aggregate Bandwidth per Unit Rack Height designated as (U), calculated as follows: Bandwidth Density = Aggregate Bandwidth/Aggregate unit rack height, in GB/s/U.

Benchmark: A test (software application) to measure a file system or file system component (the *target*) according to specific standards.

Benchmark Testing (benchmarking): To run (i.e. gather the results of) a benchmark or, more generally, a series of benchmarks. For example, the series might include: a) multiple instances of the same test on the same target, thus gathering statistics about average behavior and variability; b) the same test on a variety of targets, where the targets may be different hardware or the same hardware with different configuration details; c) a variety of tests on the same target, where the tests differ by, for example, varying one parameter; or a specific combination of many of the foregoing. See also *Performance Testing*.

Capacity Density: The storage capacity per U, calculated as follows: Capacity Density = Aggregate raw disk capacity/U, in TB/U.

DARPA HPCS I/O Scenarios: The I/O workloads defined for the DARPA HPCS program and the tests designed to measure the performance of a file system and hardware configuration on these I/O scenario to demonstrate the scalability of a given storage solution. (See *Workload*). (<http://sourceforge.net/p/hpcs-io/home/Home/>)

Defensive I/O (checkpointing): Inserting fault tolerance into computing applications and systems by regularly and periodically writing some portion of the current application state to a file system, with the intent to use it for restarting the execution of the application to recover from a failure.

File System Appliance: An integrated file system product that combines computer and storage hardware with file system software.

Floor Space Capacity Planning: The estimation of the space, computer hardware, software and connection infrastructure resources that will be needed over some future period of time.

High Availability: A system designed or configured to recover quickly from a failure by means of redundancy features built into the system.

Performance Monitoring: Gathering and preserving operational data about the behavior of a file system. This might include the results of a standardized and regularly scheduled *Benchmark* test. It may also include observations of file system performance counters and statistics.

Performance Testing (evaluation): A process of assessing how a file system or component performs under a particular workload. This might include gathering the results of a series of benchmarks (see “Benchmark Testing”) that have been designed to present a representative workload. It may also involve other activities, such as gathering operational data about the system (see *Monitoring*) while undergoing an actual or synthetic workload.

Power Density: The total power consumed (in Watts) (not including the ambient cooling) per GB/s of the Aggregate bandwidth, calculated as follows: $\text{Power Density} = \text{Total power consumed} / \text{Aggregate bandwidth}$, in W/GB/s.

Scalability: The ability to retain performance levels when adding additional resources and to take full advantage of the new resources to increase performance in a near linear manner.

Scalable File System: A file system in which the capacity and/or the performance can modularly be increased while maintaining a level manageability and access. (e.g. Lustre, Panasas, GPFS)

Storage Unit (system): A storage system consisting I/O controllers (commonly in redundant configuration) and raw storage capacity.

Workload: The actual, anticipated, or desired I/O activity on a file system. This could include the amount of data, the frequency of requests, the characteristics of how the data is read or written, or the spectrum of possible values of the foregoing. Taken together the details of type of I/O are the *workload characteristics*. A *synthetic workload* might consist of a single type of I/O (See *DARPA HPC I/O Scenarios*). A particular application might have a few specific workload characteristics. A particular site might have a collection of applications with a rich mix of workload characteristics. A *production workload* characterizes the I/O of a system while carrying out its normal operations, as opposed to during benchmark testing.