Analysis and practice of SPC-1 Benchmark

Sun Hao

- The introduction to SPC-1 Benchmark

◆ SPC and SPC-1

◆ Test model

◆ Test process

- **What is SPC**
  - Storage Performance Council (SPC) is a non-profit organization with high authority and wide recognition in the field of storage performance testing.
  - SPC members are open to all companies, academic institutions and individuals, and the current core members include many well-known enterprises in the storage field, such as IBM, HP, fujitsu, TTA, inspur, HUAWEI, Lenovo and Infortrend.

- **How to test performance**
  - Due to the increasingly complex application environment of storage system, the test of a single feature is often insufficient to reflect the performance of storage system. Thus, SPC has conducted a large amount of investigation and analysis for various typical application services in the industry, and launched corresponding test specifications and test tools for the common online services, large-scale file processing services, and file system-oriented backup and recovery services. Meanwhile, it provides realistic and feasible theoretical basis and test tools for evaluating storage system performance and cost.

- **What is SPC-1**
  - SPC-1 benchmark is a test standard introduced by SPC to measure the performance of storage subsystems when performing business-critical applications, such as OLTP, database operations, and mail server. These applications include query and update operations, characterized by random I/O operations. As mid/high-end storage systems often serve as data foundation platforms for critical business applications in the user's system environment, the SPC-1 benchmark results have a strong reference significance for mid/high-end clients and are currently the most authoritative and convincing performance standards.

4

Test model

❯ The SPC-1 benchmark, which measures throughput (IOPS) of a storage system under **random I/O** loads, reflects the basic performance of the storage system in handling complex requests and large amounts of data.

❯ In order to test the performance of the storage system in typical business applications, SPC-1 specifically designed a load model，which is mainly composed of random I/O and continuously queries and updates the business system concurrently. These random I/O operations, which involve database-type OLTP applications and E-mail system applications, provide a good measure of storage system throughput (IOPS) metrics.

❯ Due to the online business has the following features: sensitive to the time of reading and writing; the dynamic change of the workload; the high reliability of storage space; the diversification of user access behavior; the requirement of continuous storage of data without damage or loss. SPC has designed eight types of I/ O stream simulating user data access behavior, and through the SSU (Stimulus Scaling Units) to control the scale of the I/O request, through the ASU (Application Strorage Units) to describe the storage configuration of the corresponding I/O request, these elements together form the SPC-1 benchmark model.

Test model

ASU	IO stream	IO size	Access mode	Read fraction	Transfer address
ASU1	I/O stream 1	8k	random	0.5	Uniform: 0.0 - 1.0
	I/O stream 2	8k	random	0.5	R1/W1: 0.15 - 0.2
	I/O stream 3	4k 8k 16k 32k 64k	sequential	1	R1: 0.2 - 0.7
	I/O stream 4	8k	random	0.5	R1/W1: 0.7 - 0.75
ASU2	I/O stream 5	8k	random	0.3	Uniform: 0.0 - 1.0
	I/O stream 6	8k	random	0.3	R1/W1: 0.47 - 0.52
	I/O stream 7	4k 8k 16k 32k 64k	sequential	1	R1:0.2 -0.7
ASU3	I/O stream 8	4k 8k 16k 32k 64k	sequential	0	Uniform: 0.0 - 1.0

INSPUR 浪潮

➤ the complete SPC-1 benchmark process consists of Pre_Fill, Verify, Metric, and Persistence, which are described in detail below.
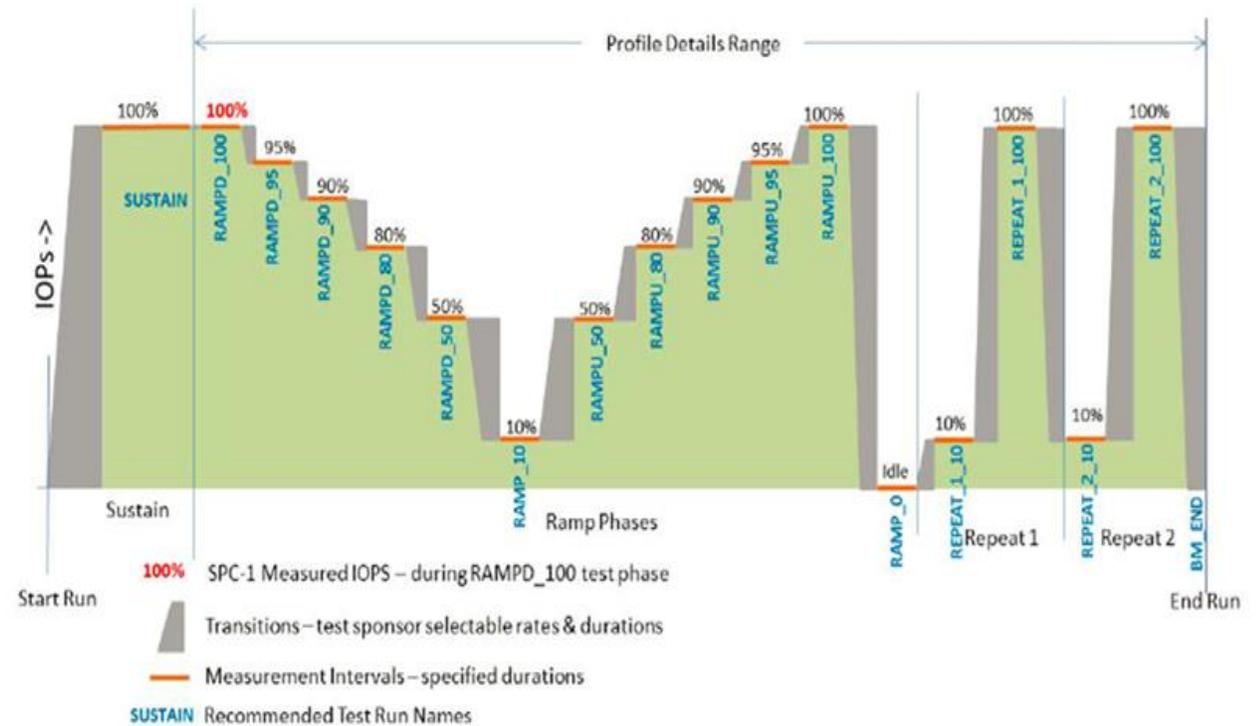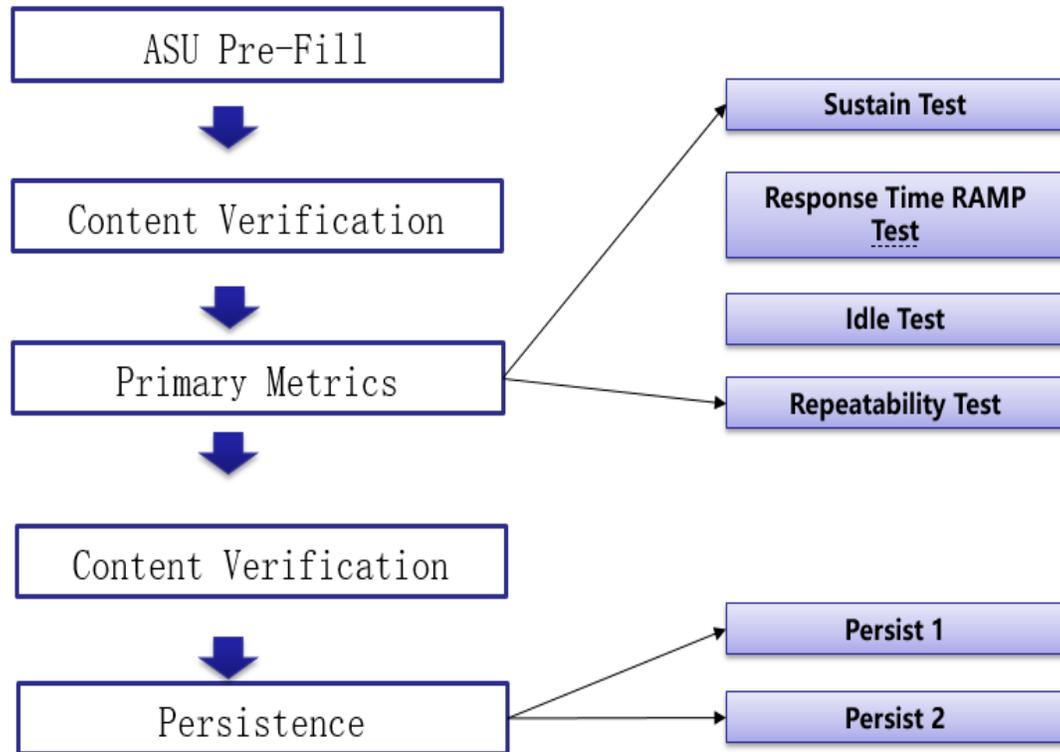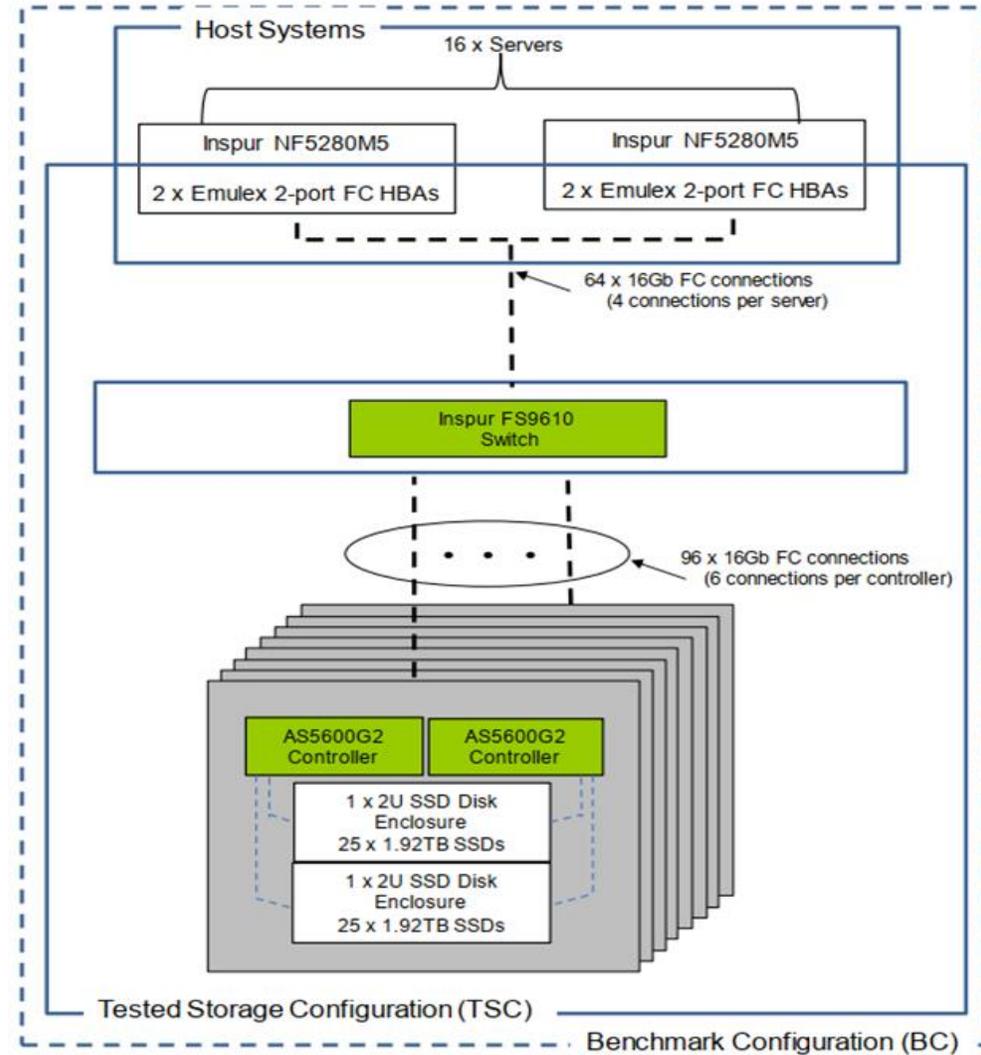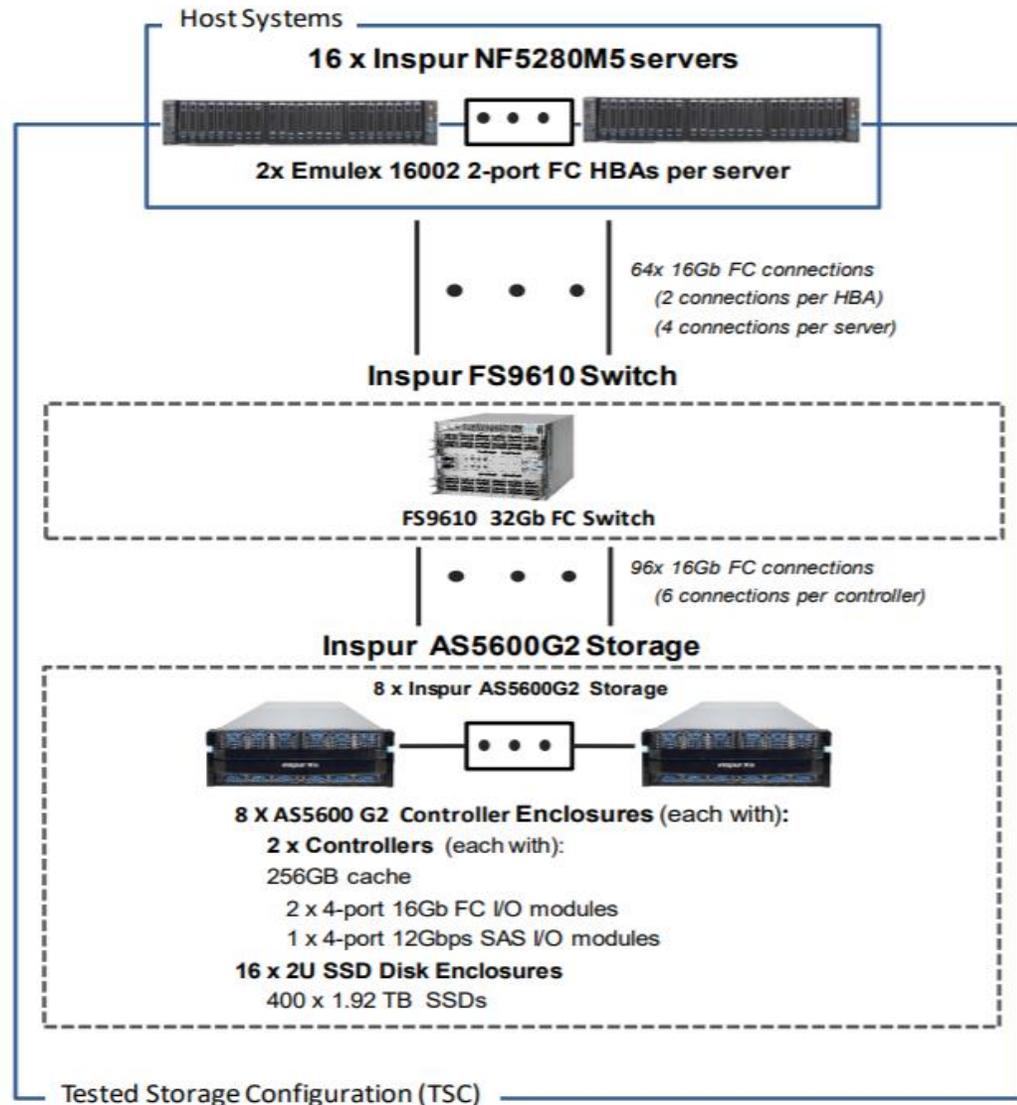
- practice of SPC-1 benchmark

◆ AS5600G2

◆ Configuration and TSC for AS5600G2 data model

◆ Advanced technology in AS5600G2

◆ result of AS5600G2 SPC-1 Benchmark

AS5600G2

❯ Inspur AS5600G2 is a mid/high-end hybrid flash storage system for medium/large-sized enterprises and provides both SAN and NAS. With the storage operating system especially developed for cloud computing and big data, rich software features, and industry-leading hardware platform, AS5600G2 satisfies the data storage and disaster recovery requirement of various applications, such as medium/large-sized OLTP/OLAP databases, virtualization and file sharing.

Product	ASAS5600G2
Controller	2-16
Controller Cabinet	4U
Processor	Multi-core
Cache	128GB-4TB
supported Storage Protocols	FC, FCoE, iSCSI, NFS, CIFS, HTTP, NDMP, FTP
Frontend Port Protocols	16Gb FC,8Gb FC, 10Gb FCoE, 1/10/25/40Gb Ethernet
Type of Hard Drive	SSD, SAS, NL-SAS, SATA
Number of Hard Drive	1800-14400
System IO Port	12-96
Type of System Backend Port	SAS3.0, single port 4*12Gbps, maximum 6144Gbps
Expansion Enclosure	2U12, 2U25, 3U48, 5U92
RAID Level	0, 1, 3, 5, 6, 10, 50, 60
Boost Resource Efficiency	Intelligent Thin Provisioning (InThin) Intelligent virtualization RAID (InRAID) Intelligent data migration (InMigration) Intelligent online compression (InCompression) Intelligent online deduplication (InDedupe) Intelligent volume conversion (InTune) Intelligent heterogeneous virtualization (InVirtualization) Intelligent file service (InFileService) Intelligent tiering (InTier) Intelligent multi-tenant (InMulti-tenant)
Data Protection Software	Intelligent snapshot (InSnapShot) Intelligent cloning (InClone) Intelligent backup (InBackup) Intelligent disk mirroring (InVdiskMirror) Intelligent remote replication (InRemoteCopy) Intelligent active-active (InMetro) Intelligent cloud tiering (InCloudTier) Intelligent encryption (InEncryption) Intelligent data destruction (InErase)
Mission-Critical Guarantee	Intelligent quality of service (InQoS) Intelligent automatic cache partition (InAutoPartition) Intelligent Cache Acceleration (InFlashCache)

10

ınspur 浪潮

Host Systems

16 x Inspur NF5280M5 servers

2x Emulex 16002 2-port FC HBAs per server

64x 16Gb FC connections
(2 connections per HBA)
(4 connections per server)

Inspur FS9610 Switch

FS9610 32Gb FC Switch

96x 16Gb FC connections
(6 connections per controller)

Inspur AS5600G2 Storage

8 x Inspur AS5600G2 Storage

8 X AS5600 G2 Controller Enclosures (each with):
   2 x Controllers (each with):
   256GB cache
      2 x 4-port 16Gb FC I/O modules
      1 x 4-port 12Gbps SAS I/O modules
16 x 2U SSD Disk Enclosures
   400 x 1.92 TB SSDs

Tested Storage Configuration (TSC)

Host Systems

16 x Servers

Inspur NF5280M5

2 x Emulex 2-port FC HBAs

Inspur NF5280M5

2 x Emulex 2-port FC HBAs

64 x 16Gb FC connections
(4 connections per server)

Inspur FS9610 Switch

96 x 16Gb FC connections
(6 connections per controller)

AS5600G2 Controller

AS5600G2 Controller

1 x 2U SSD Disk Enclosure
25 x 1.92TB SSDs

1 x 2U SSD Disk Enclosure
25 x 1.92TB SSDs

Tested Storage Configuration (TSC)

Benchmark Configuration (BC)

11

Advanced technology in AS5600G2

▶ Lock-free design

- Thread binding technology, as shown in Figure, is used to bind all tasks to different CPU cores. When the task is executed, it does not cross the CPU cores, there is no need to protect data from conflicts, and there is no need to use locks to protect conflicts, which can avoid the CPU context switch caused by locking. As a result, CPU efficiency is improved
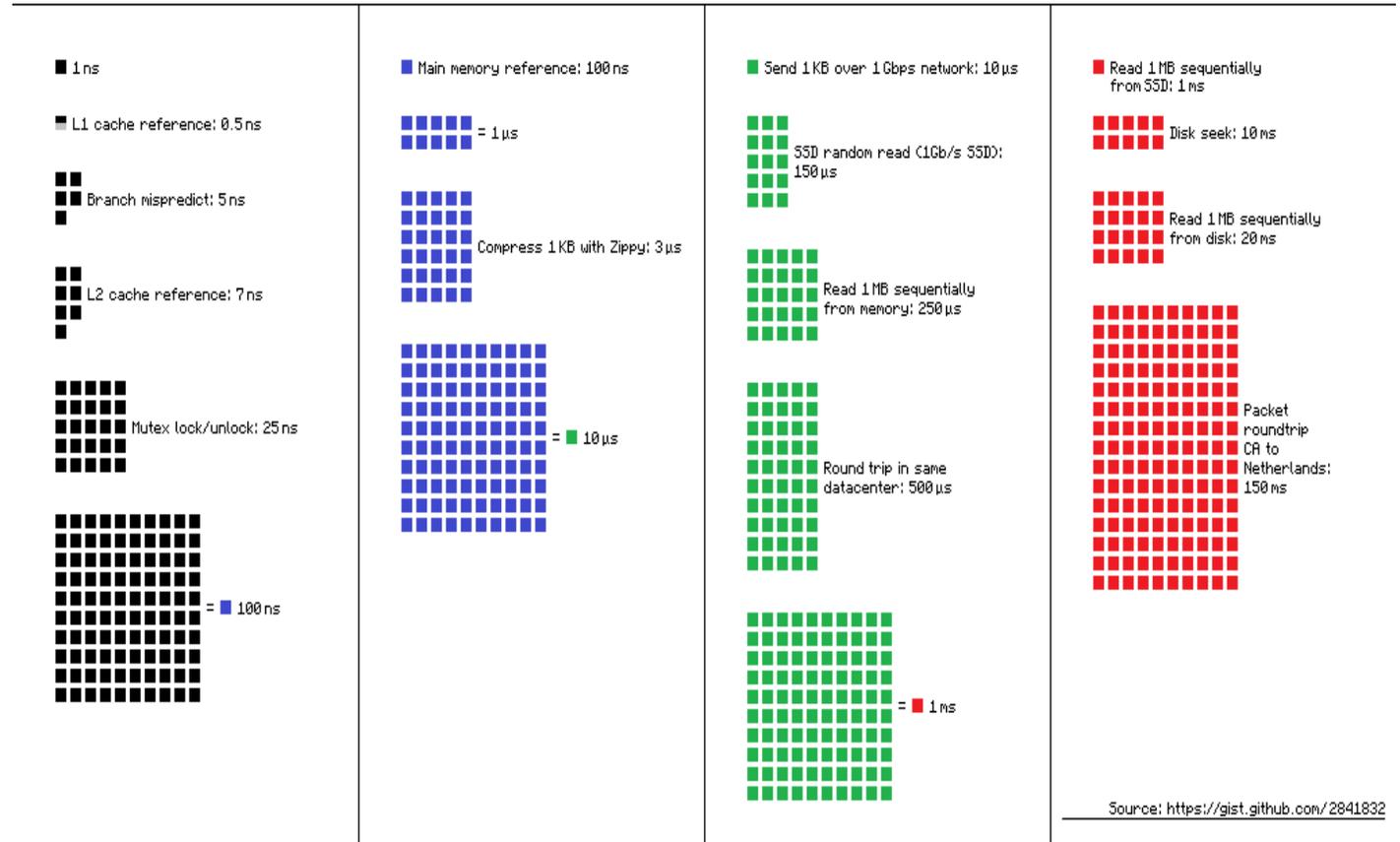
Advanced technology in AS5600G2
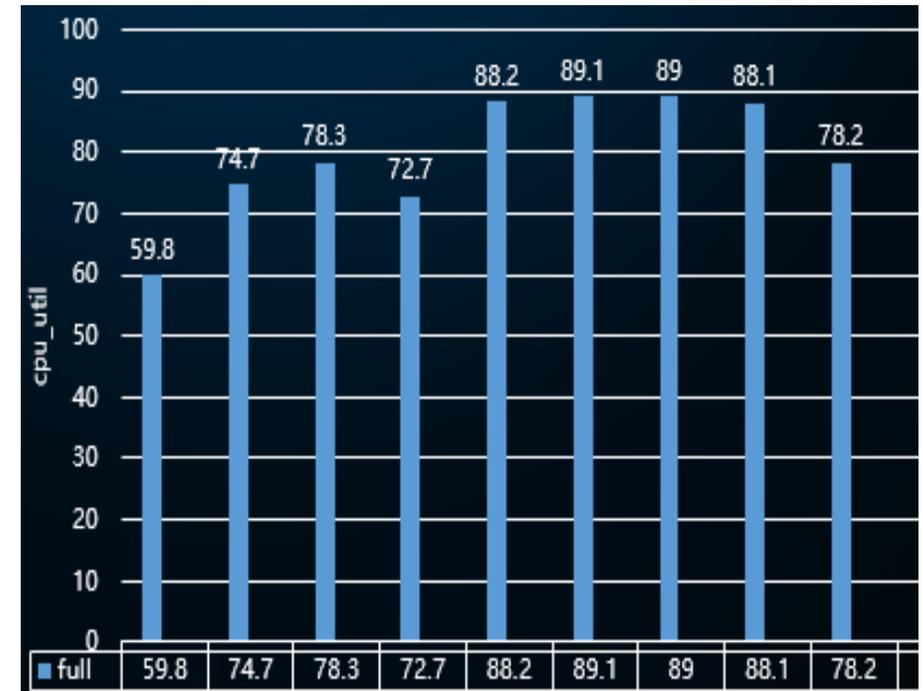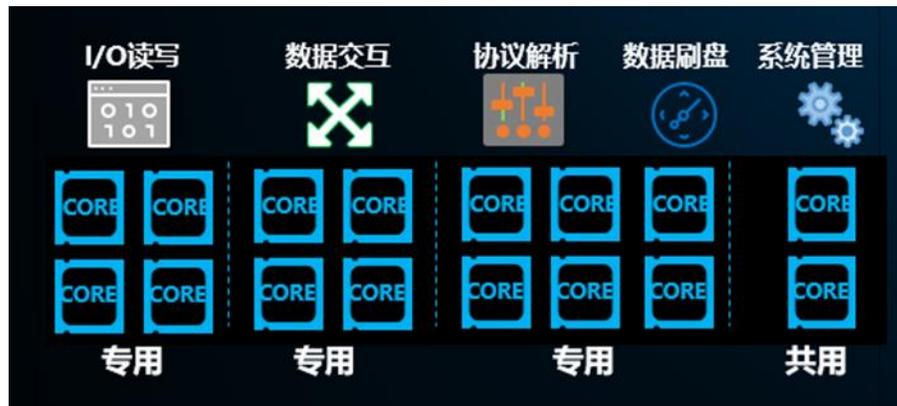
> **Pre-allocation of memory**
>
> - In order to handle IO of different sizes on the host side, frequent memory requests are usually required to transfer data over the FC network. There are many ways to apply for memory, among which the memory management function provided by the operating system has poor performance, needs to consume a large number of CPU computing resources, and after running for a long time, it is easy to cause fragmentation of memory management
>
> - AS5600G2 uses pre-allocated memory to provide services uniformly in the form of memory pools, avoiding the overhead of repeatedly requesting memory from the operating system and reducing memory fragmentation. At the same time, the memory managed by AS5600G2 is large enough to have enough caches to improve performance.

Cache

- AS5600G2 reasonably arranges the content of the data structure to align the Cache Line of the data structure and reduce the number of memory accesses

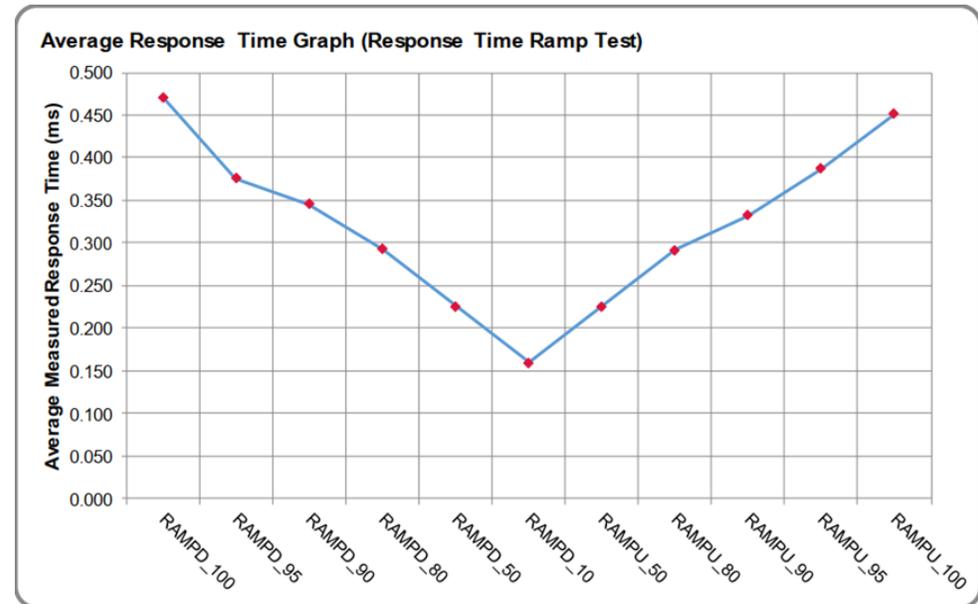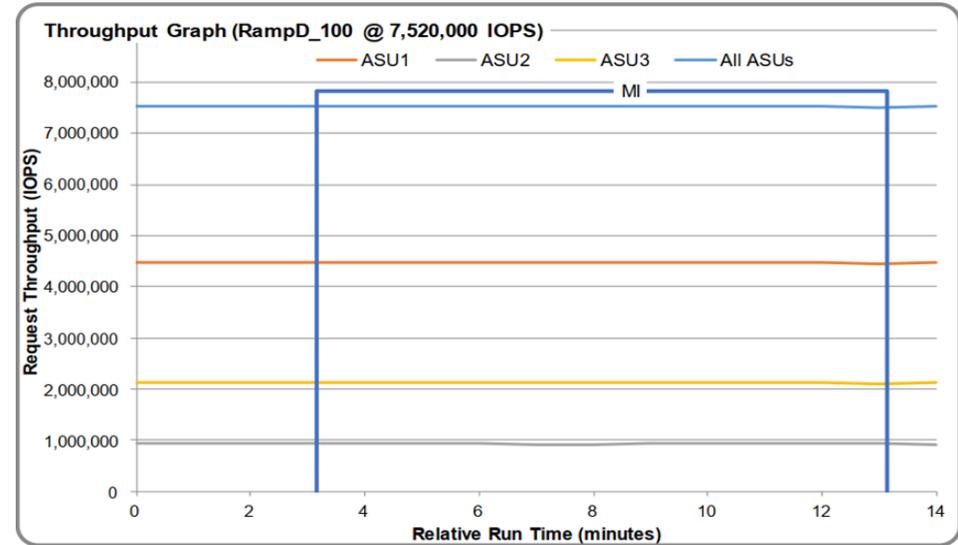### Latency Numbers Every Programmer Should Know

- 1 ns
- L1 cache reference: 0.5 ns
- Branch mispredict: 5 ns
- L2 cache reference: 7 ns
- Mutex lock/unlock: 25 ns
- = 100 ns

- Main memory reference: 100 ns
- = 1 μs
- Compress 1 KB with Zippy: 3 μs
- = 10 μs

- Send 1 KB over 1 Gbps network: 10 μs
- SSD random read (1Gb/s SSD): 150 μs
- Read 1 MB sequentially from memory: 250 μs
- Round trip in same datacenter: 500 μs
- = 1 ms

- Read 1 MB sequentially from SSD: 1 ms
- Disk seek: 10 ms
- Read 1 MB sequentially from disk: 20 ms
- Packet roundtrip CA to Netherlands: 150 ms

Source: https://gist.github.com/2841832

Intelligent scheduling technology is mainly based on data characteristics (IO priority, etc.) during IO processing to automatically select a path (fast path/slow path). Priority is given to ensuring fast path scheduling resource requirements. By dividing the critical path into slices, it is decomposed into multiple uncoupled processing steps, and then combined with cpu affinity binding technology to achieve true concurrent pipeline processing. This in turn ensures that the consumption of average IO latency is reduced, and the overall concurrency of the software is further improved, which is ultimately reflected in the overall performance of the system.
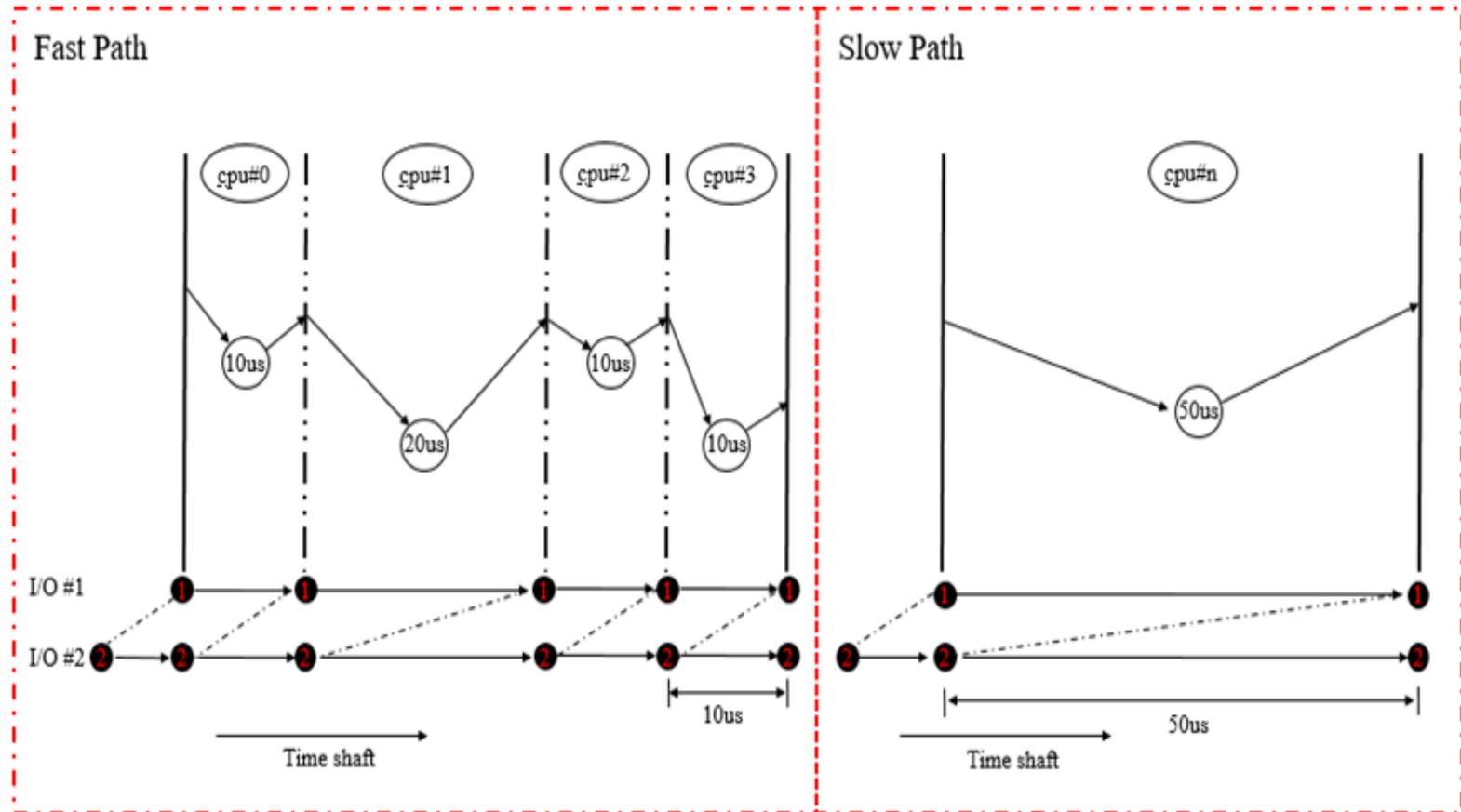
SPC-1 IOPS™	7,520,358
SPC-1 Price-Performance™	$386.50/SPC-1 KIOPS™
SPC-1 IOPS™ Response Time	0.472 ms
SPC-1 Overall Response Time	0.251 ms
SPC-1 ASU Capacity	274,877 GB
SPC-1 Space Effectiveness Ratio	NA
SPC-1 ASU Price	$10.58/GB
SPC-1 Total System Price	$2,906,551.87
Data Protection Level	Protected 2 (RAID-10)
Physical Storage Capacity	768,000 GB
Pricing Currency / Target Country	U.S. Dollars / China

Throughput Graph (RampD_100 @ 7,520,000 IOPS)

Average Response Time Graph (Response Time Ramp Test)

Thanks for your watching!