

Microsoft Feedback to Energy Star Storage Specification Framework

Building block # 1

d. Power Supplies:

Microsoft feedback: Power supplies definitions for both Computer Server and Storage Products should be the same.

e. Operational States:

1. Idle:

Microsoft feedback: To minimize confusion about different idle and active states for enterprise storage products, we'd like to propose a few distinct and well defined definitions to better manage the power management aspect as follows:

Proposed definition of Idle states:

- **Host/Initiator Idle:** An operational state in which the OS and application software have completed loading and the Storage Product is capable of completing workload transactions, but no additional active workload transactions are pending by the host.
- **Target Idle:** An operational state in which the target / Storage Product has no active workload transactions, either by the host or maintenance tasks initiated by the target itself e.g. RAID rebuilding, hot sparring etc.

Note: Storage Product such as RAID may need to perform basic house-keeping functions as background tasks without informing the host/application.

- **System idle:** when both initiator and target are idle.

2. Active:

Proposed definition of Active states:

- **Host/Initiator induced Active:** An operational state in which the OS and application software have pending transactions and the Storage Product has no completed all the active /pending transactions requested by the host.
- **Target induced Active:** An operational state in which the target / Storage Product has active workload maintenance transactions initiated by itself.
- **System Active:** when both initiator and target are active.

d. Questions for Discussion (page 4)

1. How are Active, Idle, Maximum, and Full Load states defined in the industry?

Microsoft feedback: Assuming that a rough representative “enterprise storage” workload was defined, it would have to be parameterized so that it would be usable for storage products of all different sizes and capabilities. But no matter what workload mix is chosen, it will be decried as unfair by someone. So our advice would be to keep it incredibly simple. For example:

- 8KB requests
- Cover the entire space of storage capacity
- 25% random reads; 25% random writes; 25% sequential reads; 25% sequential writes
- No writes tagged as write-through
- No cache flush requests
- Run for several hours continuously before capturing the power consumption
- Constant number of outstanding requests of each type (rr/rw/sr/sw)
 - This is the tricky one, as it determines the intensity of the workload. It could end up being 1 request per N gigabytes of usable capacity, or 1 request per N storage media devices, or?

Alternately, you could use one of the workload definitions that SNIA or SPEC has defined or will define for performing power tests optimizing for Storage Products.

3. Are there other Operational States specific to Storage Products that will need to be defined in this Specification? For example, how should EPA address data maintenance functions that may occur in the background while a product is in an idle state?

There are disk drive manufacturers proposing specifications allowing enterprise class drives to be put into low power states. Whether under the control of hardware or software, these power states have ramifications regarding the latency imposed on requests that cannot be satisfied without bringing the devices back to full active functionality. Therefore, there may need to be additional power states specified with specific latency tolerances, e.g. Idle or Low-Power States 1, 10, and 30 – where the designating number indicates the maximum number of latency in seconds required to service any 4 KB request that arrives when a storage product is in a low-power low-performance state.

Building block #3:

d. Q1

Microsoft feedback:

Idle and Max are achievable without too much contention. Active would be difficult for the industry to come to consensus on, since whatever workload is chosen will work well for some products and not as well for others. Idle may become a problem in the future, when enterprise storage devices implement multiple low-power states, but in the Tier-1 timeframe it shouldn't be a big issue.

Idle is probably the most important potential power savings area, since individual storage media is idle more often than it is active.

Building block # 4

c.3

Microsoft feedback: “Utilization measurements” – the standard two throughput metrics would be “average requests per second” and “average bytes transferred per second,” and the standard response time metric would be “average response time per request.” But none of these necessarily capture the true utilization level of the storage media or product. Nevertheless, they will probably have to do. It might be useful to also report the average request size. And splitting up the data in read-specific and write-specific statistics could be helpful for informing an administrator who knows what his or her target workload is comprised of.