Microsoft Support for Large-Sector Drives

Paul Luber, Lead Program Manager

Windows Clusters, Filesystems & Storage
Microsoft Corporation

Microsoft's Commitment

Our Views on Large-Sector Drives

- The industry's need for large-sector drives is clear.
- This is a good opportunity to work with you on a long-term plan for our partners and joint customers.
- We agree on the need to support these drives at the earliest opportunity.

New OS Support

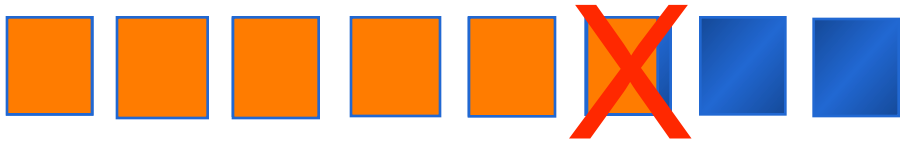
- For ATA drives, extending ATA alignment structure of IOCTL_STORAGE_QUERY_PROPERTY to include logical and physical sector size

```
typedef struct _STORAGE_ATA_ACCESS_ALIGNMENT_DESCRIPTOR {
   …
   ULONG BytesPerLogicalSector;
   ULONG BytesPerPhysicalSector;
   …
   }
```

- Populated from the ATA IDENTIFY DATA as defined in ATA7 Volume 1 Section 6.17.57
- Since physical = logical on SCSI, there are no plans for an equivalent SCSI method.
- Backporting this change is under consideration.
  - Apps could determine the physical sector size on any version and act appropriately by possibly blocking their use.

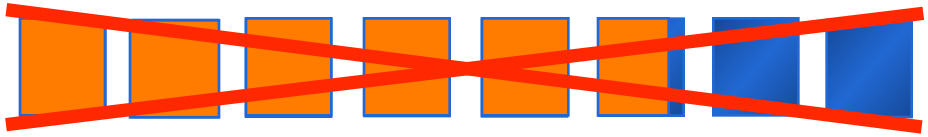
Why is this important?

- Today, atomic log writes are assumed to be 512 bytes by many database engines and filesystems



8 - 512 byte logical / physical sectors

- 512 logical / 4K physical disks break these assumptions.



1 - 512 byte logical / 4K physical sector

- On every dirty shutdown / power loss, transactions previously considered committed could be unpredictably aborted creating inconsistency and corruption.
- Backporting the ATA alignment method under consideration
  - Apps could determine this condition (512/4K) on all versions and act appropriately by possibly blocking their use.

Planned Support Matrix

	Downlevel OS		Longhorn / Vista	
	512/4K*	4K/4K	512/4K*	4K/4K
Windows				
Boot Support	+	✗	✓	✓
NTFS	+	++	✓	✓
AD	✗	✗	✓	✓
SQL	Under Investigation			
Exchange	Under Investigation			

* ATA only. ✓ Supported ✗ Not Supported

+ After a dirty shutdown / power loss, NTFS may have errors that would normally be recovered through journaling that would only be recoverable though a chkdsk on 512/4K drives. Basic disks only.

++ May be supportable. Need more investigation and testing.

Call to Action and Asks

- Feedback, Feedback, Feedback!
  - Send feedback to longsec@microsoft.com
- More engineering samples for testing
- Consider the use of NVRAM or logging techniques to mitigate the risks of read-modify-write for 512/4K drives.

Microsoft®

Your potential. Our passion.™