Open. Together.
SSD with Compression for the Compute and Storage Infrastructure: Implementation, Interface and Use Cases

Erich F. Haratsch, Managing Technologist, Seagate
Data reduction techniques such as compression and deduplication have been employed in some storage systems, but are not widely available inside SSDs yet.

Confusion exists about benefits, use cases and data integrity when SSDs implement compression.
Compressibility of Data

Data bases, OS files, application data are typically highly compressible.

Image and video files may have some small compressibility left.

Encrypted data is not compressible.

Compressibility = (1 – OutBytes / InBytes) * 100%

<table>
<thead>
<tr>
<th>Compression Algorithm</th>
<th>MySQL</th>
<th>Oracle</th>
<th>Win8</th>
<th>Linux VM</th>
</tr>
</thead>
<tbody>
<tr>
<td>gzip</td>
<td>60%</td>
<td>70%</td>
<td>50%</td>
<td>60%</td>
</tr>
</tbody>
</table>
SSD with Compression

Compression algorithm needs to be lossless

Compression needs to run inline at full data rate: low impact to write and read latencies

Compression needs to be done before encryption and ECC encoding

Compression reduces data written to media

Write original data if data is incompressible
SSD Flash Translation Layer with Compression

Traditional FTL writes data chunks with equal physical size that fit into a flash page.

With compression, FTL needs to have ability to manage physical data units of variable size.

LU: Logical unit
PU: Physical unit
Use Compression to Increase Effective Overprovisioning

Logical capacity does not change
Reduces write amplification
Increases random write and mixed read/write performance
Increases endurance
Write Amp, OP and Compression

For random write workloads, write amplification increases as OP decreases.

Compression increases available OP.

Compression reduces WA and therefore increases endurance and performance.
Use Compression to Increase Logical Capacity

Report higher logical capacity to host

Actual logical capacity depends on data entropy

Host needs to monitor free physical space
QLC SSD with Compression

QLC NAND media has typically low endurance and performance characteristics.

Compression can make QLC SSDs more attractive by increasing:
- endurance
- performance
- user capacity
SSD Product with Compression

Nytro® 1000 SATA SSD series

Seagate DuraWrite™ lossless data reduction technology is designed to increase performance and deliver high-power efficiency

Tunable capacity for performance- or capacity-optimized SSD solutions

Seagate Secure technology with secure supply chain, SD&D, Seagate Instant Secure Erase, and SED options

Easy deployment in legacy storage infrastructures with SATA 6Gb/s interface

Consistent IOPS performance with low latency for faster random access

Won Best of Show award at Flash Memory Summit 2018
Call to Action

Add compression-enabled SSD devices to the Open Compute Project by defining use cases, workloads and interfaces

Define guidelines to specify performance, endurance and QoS of compression-enabled SSDs

Define workloads to benchmark compression-enabled SSDs

Educate OCP Community about benefits

This will give the industry an additional attribute for optimizing SSD specifications such as performance, endurance and capacity
Open. Together.

OCP Global Summit | March 14–15, 2019