RAID based archives: threat or menace?

Executive Summary:
There’s a serious issue behind the humorous title: RAID technology is dangerously unsuited for archive applications. For archives, a newer storage paradigm — scale-out object storage — is far superior. Here’s why.

Introduction
Although object storage is not well known in enterprise datacenters, it has had the highest growth rate of any storage for the last decade. Hyperscale cloud data centers run on object storage, giving them tremendous advantages, enabling them to transform datacenter economics.

Archives have unique requirements that are poorly served by RAID arrays. These include

- **Durable storage**
- **In-place upgrades**
- **Minimal management**
- **Cost-effective capacity**

Object storage adds metadata to files (which is then called an object) including a unique identifier (UID). The object is accessed by calling the UID. Think of it as valet parking: you hand the ticket to the valet, and your car is retrieved.

Durable storage

**Durability is the most important requirement for archives.**
There are two basic availability strategies: replication; and advanced erasure codes. In Hadoop, for example, triple replication is the default. For archives, advanced erasure codes offer unsurpassed durability and economy over triple replication.

Archive systems today define data durability as 15 nines — as in 99.9999999999999% availability. In practice that means the data will survive five simultaneous hardware failures. Out of 1,000 trillion objects stored, only one will be unreadable.
In-place upgrades

Your data must outlast the hardware.
How can we achieve a 50 year life with 3-7 year life components? Short answer: rolling upgrades where new components are automatically included in the archive infrastructure and data is distributed to use them without compromising availability policies.

One of the advantages of object stores is that, unlike RAID arrays, they do not necessarily require that all drives offer the same capacity. That makes it possible to replace an existing — perhaps failing — drive with a newer and larger capacity drive, which also enables rolling upgrades as components wear out. Your archive vendor can tell you more.

Using the same process as recovery from failures ensures that the code paths are regularly tested, proving critical data distribution software performs as promised.

Minimal management

Personnel costs are 70% of data center expense.
Cutting management overhead pays significant dividends. Highly automated self-management is the right model for long-term archives as well. But that’s not the RAID model.

Cost-effective capacity

Advanced erasure codes protect against five hard failures with 60% overhead.
In addition, there are no data rebuild times as long as the hard failures are less than the maximum. The data remains available while any needed replication is carried out in the background. For example, in a system with 18/5 encoding, data will be available, as long as 13 of the 18 shards remain. This is instead of the hours, even days, that large capacity drives require in RAID. And, of course, RAID systems face a significant chance of recovery failure with large capacity disks.

Risk of recovery failure with RAID 5 and 6 as drive capacities grow.

<table>
<thead>
<tr>
<th>Disk Size GB</th>
<th># of HDDs</th>
<th>Array Size GB</th>
<th>Bits/Array</th>
<th>Error Rate</th>
<th>RAID 5</th>
<th>RAID 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>6</td>
<td>2500</td>
<td>2.00E+13</td>
<td>1.00E-14</td>
<td>18.11%</td>
<td>1.74%</td>
</tr>
<tr>
<td>1000</td>
<td>6</td>
<td>5000</td>
<td>4.00E+13</td>
<td>1.00E-14</td>
<td>32.95%</td>
<td>6.13%</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>10000</td>
<td>8.00E+13</td>
<td>1.00E-14</td>
<td>55.04%</td>
<td>19.07%</td>
</tr>
<tr>
<td>4000</td>
<td>6</td>
<td>20000</td>
<td>1.60E+14</td>
<td>1.00E-14</td>
<td>79.78%</td>
<td>47.44%</td>
</tr>
<tr>
<td>6000</td>
<td>6</td>
<td>30000</td>
<td>2.40E+14</td>
<td>1.00E-14</td>
<td>90.92%</td>
<td>69.10%</td>
</tr>
</tbody>
</table>

RAID for archives: threat or menace?
The RAID model

RAID arrays were a massive improvement over the major data protection tool — disk mirroring — in use 25 years ago. They offered higher performance, simpler management, lower cost, and, in theory, equal data protection.

But the last decade of hyperscale object storage systems development has proved that there are better ways to manage and protect data especially for archives.

Compared to purpose-built archive systems, RAID arrays fail on multiple dimensions.

**Limited data protection.** RAID 5 and RAID 6 protect against the loss of only one or two drives respectively. The likelihood of an Unrecoverable Read Error (URE) during a RAID 5 rebuild has RAID vendors recommending RAID 6, (see Table 1) which makes data capacity more costly. In addition, as drive capacities continue to grow, rebuild times grow as well, extending data vulnerability to other failures.

**Costly maintenance.** Academic research has shown that when one drive in a RAID group fails, another drive failure is much more likely, so prompt drive replacement reduces exposure to another failure during rebuilds. Given that need 24/7 maintenance is almost a requirement — and that costs money.

**Forklift upgrades.** RAID arrays are tightly integrated and closed, offering limited expandability, technical obsolescence, disruptive upgrades, and data transfers that can take days or even weeks.

Conclusion

RAID technology was a massive improvement over earlier storage models, but for a great many workloads has been superseded by object storage that is more scalable, lower cost, simpler to manage and with higher data integrity. RAID protects disk drives; erasure codes protect data. The success of the cloud vendors is proof that critical 24/7 data services don’t require traditional RAID arrays.

Scale-out archive object storage systems with low-cost hardware and smart software, and advanced erasure codes, offer extraordinary data protection at a cost far below that of legacy RAID systems.

Data protection is an archive’s fundamental requirement and the RAID array’s low durability is a threat to data preservation. The ease of management, long-term upgradeability, increased hardware redundancy, and configuration flexibility of object storage systems and their extraordinary data protection, make archives one application where RAID systems are a menace to archive data.