

Tencent Ultra-Cold Storage System Optimization with Intel® ISA-L – A Case Study



In this era of data explosion, the cumulative amount of obsolete data is becoming extremely large. For storage cost considerations, many independent Internet service providers are developing their own cold storage system. This paper discusses one such collaboration between Tencent and Intel to optimize the ultra-cold storage project in Tencent File System* (TFS). The XOR functions in [Intel® Intelligent Storage Acceleration Library \(Intel® ISA-L\)](#) successfully help TFS meet the performance requirement.

Introduction to Tencent and TFS

Tencent is one of the largest Internet companies in the world, whose services include social networks, web portals, e-commerce, and multiplayer online games. Its offerings in China include the well-known instant messenger Tencent QQ*, one of the largest web portals, QQ.com, and the mobile chat service WeChat. These offerings have helped bolster Tencent's continuous expansion.

Behind these offerings, TFS serves at the core of file services necessary for many businesses. With hundreds of millions of users, TFS is facing performance and capacity challenges. Since the Tencent Data Center is mainly based on Intel® architecture, Tencent has been working with Intel to optimize the TFS's performance.

Challenge of ultra-cold storage project in TFS

Unlike for Online Systems, procurement of processors for TFS's ultra-cold storage project is not a budget priority, so existing processors have been recycled from outdated systems. This approach does not result in powerful compute performance, with calculation performance easily the biggest bottleneck for the system.

Previously, in order to save disks capacity and maintain high reliability, the project adopted the erasure code 9+3 solution (see Figure 1).

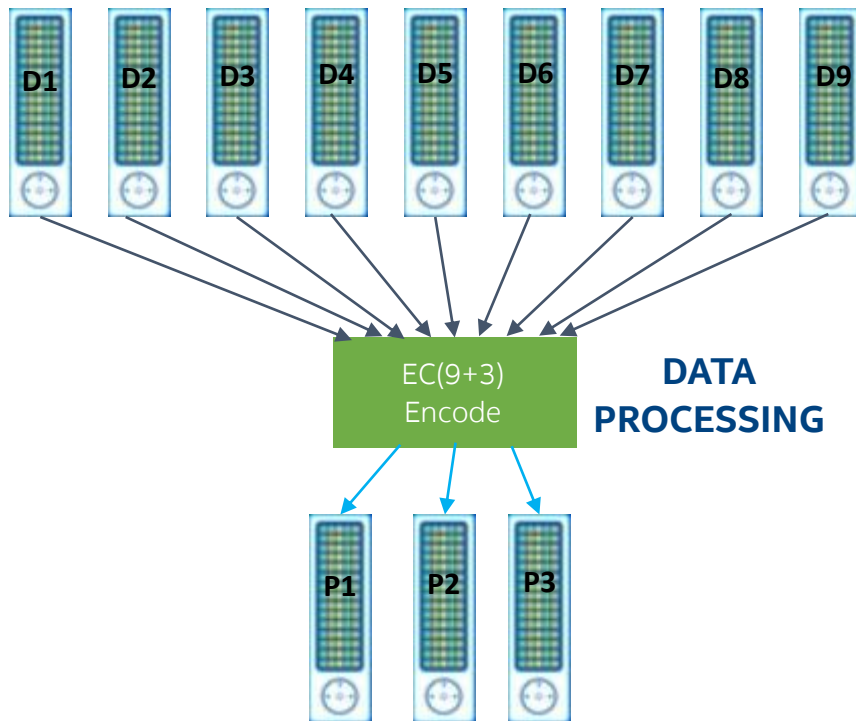


Figure 1. Original Erasure Code 9+3 solution.

Tencent has reconsidered erasure coding for several reasons:

- Much of the data stored in this ultra-cold storage system are outdated pictures. Occasional data corruptions are acceptable.
- Redundancy rate of erasure code 9+3 may be too much of a luxury for this kind of data.
- Even optimized with Intel ISA-L erasure code, it is still a heavy workload for these outdated, low-performance servers assigned to ultra-cold storage system.

In order to reduce the redundancy rate and improve performance bottlenecks, a solution that uses XOR operation on 10 stripes to generate 2 parities was adopted (see Figure 2). The first parity is horizontal processing, and the second parity is vertical processing.

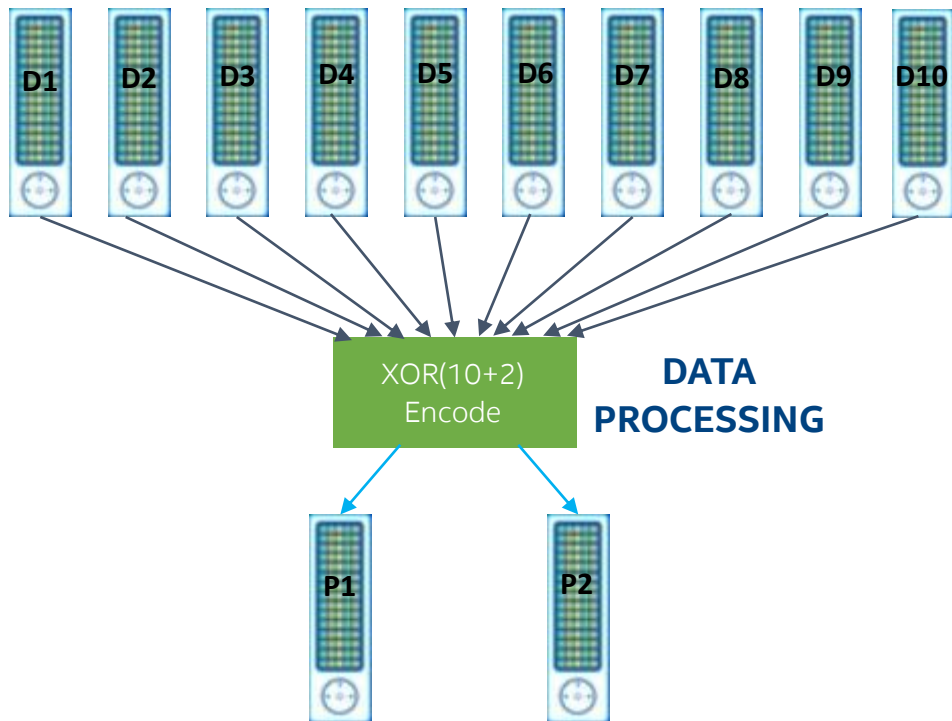


Figure 2. New XOR 10+2 solution

This new solution still had one obvious hotspot: the XOR operation limits system performance. Despite simplifying the data protection algorithm, this cost-optimized solution couldn't meet the performance requirements that Tencent Online Systems needed.

Tencent was seeking an effective and convenient way to reduce the calculation effort of the XOR operation. It needed an efficient and optimized version of XOR to alleviate the performance bottleneck and meet the design requirements for the ultra-cold storage solution.

About Intel® Intelligent Storage Acceleration Library

Intel ISA-L is a collection of optimized, low-level functions used primarily in storage applications. The general library for Intel ISA-L contains an expanded set of functions used for erasure code, data protection and integrity, compression, hashing, and encryption. It is written primarily in hand-coded ASM but with bindings for the C/C++ programming languages. Intel ISA-L contains highly optimized algorithms behind an API, automatically choosing an appropriate binary implementation for the detected processor architecture, allowing ISA-L to run on past, current, and next-generation CPUs without interface changes.

The library includes an XOR generation function, `gen_xor_avx`, as part of the Intel ISA-L data-protection functions. Intel ISA-L is highly performance optimized by Intel's Single Instruction Multiple Data instructions.

Collaboration between Tencent and Intel

Tencent and Intel have worked together using Intel ISA-L to optimize ultra-cold storage project in TFS.

The XOR function used in the ultra-cold storage project was originally coded in C Language and in Galois code format, named `galois_xor`. The first optimization proposal was to replace `galois_xor` with Intel ISA-L `gen_xor_avx` directly. The test results from this single change showed a ~50-percent performance gain.

After analyzing the parity generation method of the ultra-cold storage system, we suggested using `gen_xor_avx` in pointer array format. This second optimization proposal improved coding efficiency further, by avoiding unnecessary memory operation.

Results

The performance optimization scheme, based on the Intel ISA-L XOR function, helped solve the practical problems encountered in building an ultra-cold storage system. The test results from Tencent showed a 250-percent performance increase compared with previous method.

Method	Galois xor	Intel ISA-L <code>gen_xor_avx</code> on non-array form	Intel ISA-L <code>gen_xor_avx</code> on array form
Performance	800 MB/s	1.2 GB/s	2 GB/s

This distinct performance gain successfully met the requirements from Online Systems. Even better, since Intel ISA-L is open-source (BSD-licensed) code, there was no cost to the Tencent team for the huge improvement in system performance.

Acknowledgement

As a result of this successful collaboration with Intel, Sands Zhou, principal of the Tencent ultra-cold storage system, said: "TFS ultra-cold storage project, based on entire cabinet program, CPU became a performance bottleneck. In the meantime, the project got strong supports from Intel based on ISA-L XOR program. Thanks again, wish more collaborations with Intel in the following work."

Notices

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer or learn more at [intel.com](https://www.intel.com).

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting www.intel.com/design/literature.htm.

Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

© 2017 Intel Corporation