Targeting Intel® Compression Technology for Big Data

Overview

Data compression, as a means to reduce data volume, is targeted to better utilize storage space, or network bandwidth. Either way, **compression dramatically reduces I/O operations**. Whether structured or unstructured, compressed data brings the following advantages:

- Big data analytics – lower data footprint on storage means less I/O transactions to storage, during data access. That means better overall analytics performance;
- Big data processing – taking the compression task off the CPU into an accelerator, saves more CPU cycles for business logic processing.

Intel® QuickAssist technology is a set of Linux software drivers and libraries, providing API to access Intel® compression acceleration hardware technology, embedded in the Intel® ColetoCreek 8950 chip.

Silicom is a provider of advanced networking solutions with core value and strength in its ability to deliver tailored solutions to fit its customers' needs by using existing technology, such as Intel's quick assist, and enhancing it with API's and functionality making it easier for use. Silicom delivers the chip on a PCIe adapter, as a look aside acceleration engine. Silicom also maintains the software suite.

Data Compression – DEFLATE, LZS

The main idea of this acceleration card is to accelerate web oriented traffic. Due to the fact that present days network applications are tunneled over HTTP (Facebook, Skype, SalesForce, etc.), then it makes sense to address this segment. Hence, the compression algorithms that are supported by the hardware are **DEFLATE** (zlib format), that is abundant in web traffic, and **LZS**, that is the standard compression algorithm for IPsec, but has less dominance. Both algorithms are based on Huffman coding.

However, other types of compression are out there, like **Snappy** that is popular for its speed, or **bzip2**, that block oriented compression algorithm, that compresses very good, but is considered as a slow performing algorithm.

Although DEFLATE and LZS are stream algorithms, best suited for WAN acceleration, and web traffic, both can be efficiently used for storage compression and decompression, thus significantly optimize physical storage space.

Picture 1 – SIlicom PE3iSC02 with Intel® ColetoCreek

Features

- Compression rate – 24Gb/s
- Decompression rate – 24Gb/s
- Dynamic and static Huffman Trees
- 4 Search Depths (1, 4, 8 & 16)
- Support for Stateful compression and decompression
- Multiple history sizes with DEFLATE
- Flexible form factor
-