

発表概要

POWER8 プロセッサ内蔵データ圧縮アクセラレータの性能評価

緒方 一則^{1,a)} 堀井 洋¹

2017年11月1日発表

大量データを扱う計算処理は、ビッグデータ解析や機械学習などの普及により、ますます重要になっている。大量データを扱う際の課題の1つは、それらを保存するストレージを確保するために圧縮が必要となるので、その圧縮されたデータの使用時に伸張する必要があることである。これらの圧縮・伸張処理は通常 CPU を使って行われるので、本来の計算処理に使える CPU リソースが削減される。この課題を解決する手法の1つは、圧縮・伸張処理をアクセラレータにオフロードすることである。POWER7+以降のPOWER プロセッサにはデータ圧縮アクセラレータがCPU に組み込まれており、CPU を使わずに圧縮・伸張を行うことが可能である。したがって、このアクセラレータを使うことで、大量データの処理速度向上が期待できる。本発表では、TPC-DS 入力データや Spark SQL のシャッフル作業ファイルの圧縮・伸張を、POWER8 プロセッサ内蔵アクセラレータにオフロードすることの効果を検証する。

Evaluation of the Built-in Compression Accelerator of the POWER8 Processor

KAZUNORI OGATA^{1,a)} HIROSHI HORII¹

Presented: November 1, 2017

There is growing demand for computing a large amount of data in the areas such as big data analysis and machine learning. One of the issues for handling a large amount of data is that such data often needs to be compressed before saving into storage, and then decompressed before using it. These compression and decompression are usually done by using CPU, and it reduces the amount of CPU resource available for the computation of the data. One approach for this issue is to offload the compression and decompression to accelerators. The POWER7+ processor and its successors have a built-in compression accelerator, so the system can compress and decompress data without using CPU. We can expect performance improvement by offloading compression/decompression tasks to the accelerator. This presentation shows effectiveness of the compression accelerator in POWER8 processor by offloading compression and decompression of the TPC-DS input data and the shuffle work files of Spark SQL.

¹ 日本 IBM 株式会社東京基礎研究所
IBM Research – Tokyo, IBM Japan Ltd., Chuo, Tokyo 103-8510, Japan

a) ogatak@jp.ibm.com