

令和元年6月19日現在

機関番号：12102

研究種目：基盤研究(B) (一般)

研究期間：2015～2018

課題番号：15H02674

研究課題名(和文) データストリーム伝送路におけるロスレス圧縮技術の開発

研究課題名(英文) Development of Stream-based Lossless Data Compression Technology

研究代表者

山際 伸一 (Yamagiwa, Shinichi)

筑波大学・システム情報系・准教授

研究者番号：10574725

交付決定額(研究期間全体)：(直接経費) 13,800,000円

研究成果の概要(和文)：通信・映像・センサといった情報源が発するデータ量の爆発的な増加に伴い、データ伝送路に超高周波での実装が要求され、困難を極める状況にある。一定のデータ塊を単位として圧縮・復号するブロック圧縮によるデータ量の削減は従来から行われているが、常に流れ出る高速データ源では処理が間に合わず破綻する。伝送路で直に圧縮できればよいがハードウェアで効果的に実装できる技術がない。そこで、本研究は伝送路に流れるデータの出現統計をリアルタイムに生成し、圧縮・復号するストリーム圧縮ハードウェア技術を開発する。この技術は伝送路の物理的容量限界を超えたデータ伝送が可能な高信頼コンピューティングへの基盤技術の確立を狙う。

研究成果の学術的意義や社会的意義

データストリームをリアルタイムにデータ圧縮できる新しい方式を開発することが出来た。学術的にはこれまでのデータ圧縮技術では対象としていなかったデータストリームを扱える新しい研究分野へのきっかけとなる成果をもたらすことが出来た。さらに、本研究の成果は産業界でもすでに実用化されている。データストリームを対象にするデータ圧縮技術は、Society 5.0といった将来のIoT社会を形成する上で通信量の削減や消費電力の低下、さらには、計算処理の高性能化といった多岐にわたる社会実装が可能で多くのニーズがあることがわかったことも成果の一部といえる。

研究成果の概要(英文)：Due to fast grow of the amount of data from the data sources such as network, video, sensors, etc., implementation of the fast communication data path demands. It is almost reaching the technological limit. The conventional data compression technology hires blocking approach that treats a data block stored in memory. It degrades the performance when it is applied to a situation that treats data streams. Thus, the data compression technology requires to treat directly the data stream without buffering. However, there does not exist any data compression technology that fully treats data stream without any stall. This research project will develop a new stream-based data compression/decompression technology that treats a data stream. This technology will contribute a reliable technology to computing systems.

研究分野：計算機アーキテクチャ

キーワード：データ圧縮

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

産業界では、データ伝送路に流れるデータ量の爆発的増加に対応する圧縮技術が必要とされている。PCI Express、HDMI、USB といったデータ伝送路標準は GHz に達する勢いで、それを実装するプリント基板製造の現場では複数パターンを設計し、その中から期待通りに動作するものを採用するといった非科学的開発手法を採用するまでにデータ伝送路の実装は困難を極めている。実装技術を複雑化せず転送量を増加させるには、伝送路に流れるデータを圧縮し、周波数や転送量を減少させ、実装を単純化するハードウェアでの圧縮技術の開発が喫緊の課題である。

その一方で、既存の圧縮技術をブロックング圧縮からストリーム圧縮へ技術シフトを急ぐ必要がある。つまり、これまで、メモリにデータを溜めて圧縮・復号を実行するブロックング圧縮は行われてきたが、データ伝送路の帯域が増加すると破綻する(引用文献)。さらに、ハードウェアで高速化しても、伝送路の性能向上に比例した回路の高速化を余儀なくされ、性能の“いたちごっこ”に陥り、技術破綻を予想できる。従って、伝送路の性能に比例したデータの高濃度化を実現できるストリーム圧縮に技術をシフトする必要があるが、ハードウェアにスケラブルな実装が可能な技術が未踏である。データ圧縮は最頻出で最長のデータ列(シンボル列と呼ぶ)をデータの中から探し、それを単一のデータ(シンボル)に置き換えて情報量を削減するが、LZW(引用文献)等の従来法を使うと1)シンボル列抽出処理時間予測が不可、2)シンボル変換テーブルサイズが予測不可、3)テーブルを復号側へ送信するオーバーヘッド、4)テーブル自体が圧縮率を低下、という問題によって、圧縮・復号処理を伴ったデータ伝送を連続的に実行できない。

2. 研究の目的

以上の背景から、本研究では、データ伝送路に流れるデータをリアルタイムに圧縮するストリーム圧縮技術の開発を目的とした。本研究で扱うデータは区切りがない連続したデータである。従って、シンボル変換テーブルを送るタイミングがない。そこで、リアルタイムに圧縮側と復号側で同じテーブルを保持し、圧縮・復号を行う必要がある。これをハードウェアのデータ伝送路に組み込むために本研究は、少ないリソースで、かつ、高速に動作できる、ハードウェア化が可能な、1)ストリームデータ伝送路でのリアルタイム圧縮・復号機構をもった通信プロトコルと、2)リアルタイムにデータ出現頻度をランキングするシンボル変換テーブル更新機構の開発を目指した。これらの処理をハードウェアでスケラブルに実施するためには、固定長のシンボル列をテーブルに登録し、さらに、テーブル容量が飽和した場合に最も使われていないシンボル列の規則を入れ替える方式を開発できれば良いことはわかっていた。このような動的な管理方式によって、伝送路に流れる時々刻々と出現率が変化するデータであっても、伝送路中のデータ量の高濃度化を実現でき、伝送路のピーク性能を越えるストリーム圧縮ハードウェアが実装できる。

3. 研究の方法

本研究は平成 27 年度～平成 30 年度の 4 年間で実施した。データストリームの統計的傾向が変化しても高い圧縮率を維持するシステムを開発するために、1)シンボルテーブル更新を可能にする通信プロトコルの開発、2)データストリームの統計的傾向を求めるアルゴリズムの開発、3)アプリケーションの開発、の 3 つの大テーマに分け実施した。1)に関しては、初年度に実施し、アルゴリズムの開発とその実装方式に関して議論を重ね、ソフトウェアとハードウェアでの試作を行った。2)については 2 年度の初旬に完成し、ストリームデータ圧縮技術を確立することが出来た。3 年目以降はアプリケーションを開発し成果の公表に力を入れた。特に、本研究成果をもとにした大学発ベンチャー企業を設立することに注力し、社会実装を主眼に置いて活動した。

研究組織は、システム開発チームとアルゴリズムチーム、そして、アプリケーション提供チームから構成し、スムーズに連携した研究体制をとった。

4. 研究成果

上記の研究の背景で説明した従来からのデータ圧縮における問題点を克服する、無限に続くデータストリームに適用でき、さらに、ハードウェアで高速動作させることが可能な新しい方式 LCA-DLT を提案し、ハードウェア、ソフトウェアの両面で良好な性能を示すことが出来た。この圧縮方式では、図 1 のように、2 つの単位データを 1 つに圧縮するモジュールを用意し、

そのモジュールが動的に変換テーブルを作り、圧縮されたデータの構造に のテーブルの作成手順を隠し持つことで、解凍側で同一の変換テーブルを復元し、データを元に戻す。この技術では、変換テーブルを解凍側に送る必要がなく、圧縮されたデータストリームを受信し始めると、次々と解凍を行う事ができる。つまり、最初のデータが圧縮されると、そのデータは次々と解凍側に渡されて、解凍側はそれらを順次、データ構造から連想される変換テーブルを復元し、解凍していく。従って、ストリームでの圧縮・解凍処理が可能である。

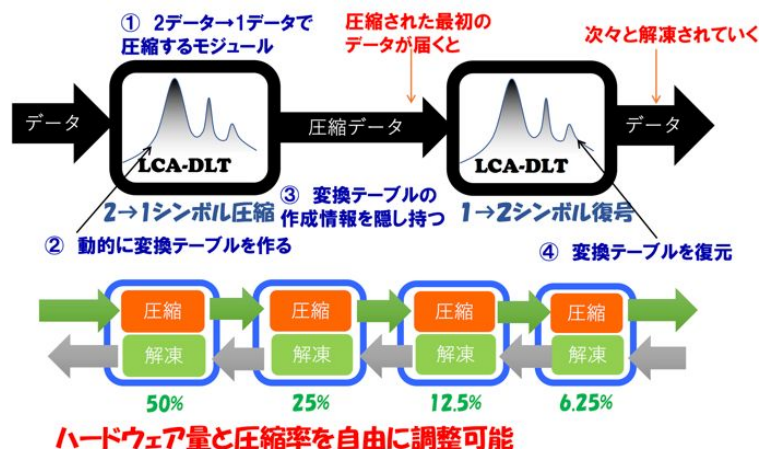


図1：ストリームデータ圧縮技術 LCA-DLT

さらに、圧縮モジュールは2データ 1データの圧縮であるため、50%しか圧縮できない。そのモジュールをカスケード接続することで、4段の場合、最大で $1/16=6.25\%$ まで理論的には圧縮できるユニバーサル性をもつ革新的な圧縮方式を開発することができた。本研究の学術的な意義として、これまで、メモリに保持したデータをプロセッサがランダムアクセスすることによるデータ圧縮方式のみが主流であったが、データストリームのようなランダムアクセスが不可能である特徴を持つデータ生成環境（例えば、センサのデータなど）においても適用できる新しいアルゴリズムが実現できることが証明でき、今後、ストリームデータ圧縮という新しい学術分野を期待できる。

本研究では、アウトリーチ活動にも後半の期間で力を入れた。LCA-DLTのハードウェア IP コアを社会実装するための大学発ベンチャー企業としてストリームテクノロジー株式会社を設立し、市場展開を進めた。この IP コアは FPGA では約 250MHz 程度での動作するため、超高速データ伝送路にも可能であることを証明した。この IP コアは画像のロスレスのリアルタイム圧縮といった産業界での応用において実用化が進んでいる。さらに、このストリームデータ圧縮技術は NEDO 事業での LSI 化を通じ、ソフトウェアでの圧縮よりも消費電力は格段に下がることを確認しており、ハードウェア実装に親和性が高く、さらに、フィルタのように滞りなく圧縮処理が行われるために、連続的にデータにアクセスする場合に有効である新しい手法として社会に大きくアピールすることが出来た。

本研究での成果は特願 2015-063449、US PAT. 9948319 として特許出願し、さらに特許査定を受けている。

本研究では組み込み分野での Embedded Technology、CEATEC、Interop といった国内最大級の展示会への成果の出展を通して社会へのアピールを積極的に行った。それらの展示会には、HDMI の映像のリアルタイム圧縮・復号化を展示し、LCA-DLT の高速実行の可能性を証明するだけでなく、産業界へのアプリケーションを来場者と議論を重ねることが出来た。さらに、CPU が扱う任意のデータをロスレスに圧縮・復号化できる PCI Express ボードを使ったデモを展示した。このボードでは 40Gbps のメモリ間伝送路における実時間データ圧縮を実証し、検査画像と行った産業利用のための映像圧縮技術に採用されるきっかけになった。新聞等のメディアにも取り上げられ、国内最大級の組込機器展示会 Embedded Technology 2015 においては、ET アワード「特別賞」を受賞した。また、本研究の成果をまとめた論文は VLDB2015 のワークショップ BPOE6 にて Best Paper Award を受賞し、学会と産業界の両面で、本研究成果が評価されている。また、LCA-DLT を事業化することに成功したストリームテクノロジー(株)は、第4回常陽ビジネスアワード「グランプリ」を受賞し、アントレプレナーシップアワード2017国内代表、JST 大学発ベンチャー表彰2018にて「科学技術振興機構理事長賞」を受賞している。

<引用文献>

Shinichi Yamagiwa, Keiichi Aoki and Koichi Wada, Performance Enhancement of Inter-Cluster Communication with Software-based Data Compression in Link Layer, Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS 2005), ACM, pp. 516-523, August 2005.

T.Welch, A Technique for High-Performance Data Compression, IEEE Computer Magazine, Volume 17 Issue 6, pp. 8-19, June 1984.

5 . 主な発表論文等

〔雑誌論文〕(計 1 件)

Koichi Marumo, Shinichi Yamagiwa, Ryuta Morita and Hiroshi Sakamoto, Lazy Management for Frequency Table on Hardware-Based Stream Lossless Data Compression, Information 2016, 7(4), 63, MDPI. 査読あり(Open Access, 16 pages)
DOI: 10.3390/info7040063

〔学会発表〕(計 4 件)

Shinichi Yamagiwa, Ryuta Morita and Koichi Marumo, Reducing Symbol Search Overhead on Stream-based Lossless Data Compression, In proceedings of ICCS 2019, LNCS, Springer, 査読あり, 2019. (Accepted for publication)

Shinichi Yamagiwa, Ryuta Morita and Koichi Marumo, Bank Select Method for Reducing Symbol Search Operations on Stream-based Lossless Data Compression, Data Compression Conference 2019, 査読あり, March 2019.

DOI: 10.1109/DCC.2019.00123

Koichi Marumo, Shinichi Yamagiwa, Time-sharing Multithreading on Stream-based Lossless Data Compression, In Proceedings of The Fifth International Symposium on Computing and Networking, pp. 305-310, IEEE, 査読あり, November 2017.

DOI: 10.1109/CANDAR.2017.42

Shinichi Yamagiwa, Koichi Marumo and Hiroshi Sakamoto, Stream-based Lossless Data Compression Hardware using Adaptive Frequency Table Management, In Proceedings of VLDB 2015/ BPOE-6, pp. 133-146, Springer, 査読あり, Sep 2015.

DOI: 10.1007/978-3-319-29006-5_11

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 2 件)

名称：データ圧縮器、データ圧縮方法、データ解凍プログラム、データ解凍器、データ解凍方法、データ圧縮プログラムおよびデータ圧縮解凍システム

発明者：山際伸一、丸茂孝一、森田隆太

権利者：筑波大学

種類：特許

番号：特願 2018-163608 号

出願年：2018

国内外の別：国内

名称：データ圧縮・解凍システム、データ圧縮方法及びデータ解凍方法並びにデータ圧縮器及びデータ解凍器

発明者：山際伸一、坂本比呂志

権利者：筑波大学

種類：特許

番号：PCT/JP2016/59372

出願年：2016

国内外の別：国外

取得状況(計 3 件)

名称：Data Compression Method and Data Decompression Method, and Data Compressor and Data Decompressor

発明者：Shinichi Yamagiwa、Hiroshi Sakamoto

権利者：University of Tsukuba

種類：特許

番号：US Patent No. 9948319

取得年：2018

国内外の別：国外

名称：データ圧縮器及びデータ解凍器

発明者：山際伸一、坂本比呂志

権利者：筑波大学、九州工業大学

種類：特許

番号：特許 6168595

取得年：2017

国内外の別： 国内

名称：データ圧縮・解凍システム、データ圧縮方法及びデータ解凍方法 並びにデータ圧縮器及びデータ解凍器

発明者：山際伸一、坂本比呂志

権利者：筑波大学、九州工業大学

種類：特許

番号：特許 6256883

取得年：2017

国内外の別： 国内

〔その他〕

(受賞) 大学発ベンチャー表彰「科学技術振興機構長賞」受賞名：ストリームデータ圧縮技術～Data Compression3.0～(2018年8月)

(受賞) 常陽ビジネスアワード 最優秀賞(成長事業部門)「Data Compression3.0～新ロスレスデータ圧縮技術がつくるスマートライフのためのテクノロジー～」(2016年5月)

(受賞) Embedded Technology Award 2015 「特別賞」(受賞技術：高性能ストリームデータ圧縮技術～ストリームを高速にコンパクトに。未来のロスレス圧縮のカタチ)(一般社団法人 組込みシステム技術協会)(2015年11月)

(受賞) Best Paper Award (VLDB 2015/ BPOE-6) (2015年9月)(Shinichi Yamagiwa, Koichi Marumo and Hiroshi Sakamoto, Stream-based Lossless Data Compression Hardware using Adaptive Frequency Table Management, In Proceedings of VLDB 2015/ BPOE-6, pp. 133-146, Springer, Sep 2015.)

(報道) 常陽新聞 2015年8月27日「新方法のデータ圧縮開発」筑波大准教授山際伸一さんら 様々な機器に実装可能

6. 研究組織

(1) 研究分担者

研究分担者氏名：坂本 比呂志

ローマ字氏名：(SAKAMOTO, hiroshi)

所属研究機関名：九州工業大学

部局名：大学院情報工学研究院

職名：教授

研究者番号(8桁)：50315123

研究分担者氏名：和田 耕一

ローマ字氏名：(WADA, koichi)

所属研究機関名：筑波大学

部局名：システム情報系

職名：教授

研究者番号(8桁)：30175145

研究分担者氏名：河原 吉伸

ローマ字氏名：(KAWAHARA, yoshinobu)

所属研究機関名：大阪大学

部局名：産業科学研究所

職名：准教授

研究者番号(8桁)：00514796

(2) 研究協力者

研究協力者氏名：丸茂 孝一

ローマ字氏名：(MARUMO, koichi)