

運用で使われているファイルサーバボリュームの データ重複率の測定と評価

荻原一隆

(株)富士通研究所

1. はじめに

日々刻々と産み出されるデータは増大し、ストレージシステムに格納されていくデータ量も増大している。ストレージデバイスなどのハードウェア故障だけでなく、ユーザの操作ミスによるファイル消失などにも備えたデータのバックアップの重要性は変わっていないが、加えて、データ量の増大に伴うバックアップ処理の長時間化は避ける必要があり、かつ、バックアップのデータ量を論理的には減らさず物理的には減らしたいという要望もある。

バックアップ時間を短縮しデータ格納量を削減する手法に差分バックアップがある。まずフルバックアップをとり、以降は更新されたデータのみをバックアップ先に保存する手法である。更新部分の検出は、ファイルの場合には、ファイル属性にあるアーカイブビットを参照することで更新ファイルを検出できるが、ファイルが大量にあるとアーカイブビットの参照だけで処理時間が長くなる。ブロックデバイスの視点ならば、デバイスへの書き込み時に更新範囲を管理すれば、更新部分の特定は短時間で行える。

最初のフルバックアップの格納に必要な容量の削減を含めて、物理的に格納するデータ量を減らす技術として重複除去がある。重複除去は、データ内容が同一のものを、論理的には別々に存在するように見せるが、実体は同じ領域を指すことで、格納データ量を減らすものである。

今回は、物理的なバックアップデータ量削減の観点から、一般業務で使用しているファイルサーバボリュームから別のストレージ装置への日次バックアップに重複除去を適用することを想定し、後述する3種類のチャンキングを用いてデータ重複率などの測定を行った。測定の目的は、ブロックに対する重複除去がファイルに対する重複除去と代替できるかどうかの確認であり、その結果と評価について報告する。

2. 関連研究

本報告の特長は、以下に概説する関連研究では用いていない容量の大きなファイルサーバのボリュームを対象にしていることである。

Park らの論文[1]では、5GB ボリュームの5世代のバックアップでセグメント構成の変化や将来必要になる格納容量の予測を行っている。

Meyer らの論文[2]では、デスクトップマシン約 800 台からディスクの内容を収集し、ファイルを対象に、ファイル全部、固定長チャンキング、可変長チャンキングの手法を用いてデータ重複率を測定している。

3. 測定対象ボリュームとチャンキング手法

データ重複率などの測定に使用したファイルサーバ用ボリューム3つの概要を表1に示す。ボリューム別に3日分の日次バックアップを想定し、連続3日分のデータを使用した。

表1. 対象ボリュームの概要

| | 総容量[TiB] /使用量[TiB] | 特徴 |
|-------|-----------------------|---------------------|
| Vol-A | 4.55/4.08 | オフィス系ファイルが多い |
| Vol-B | 1.69/1.32 | JPEG とエクセルのファイルが多い |
| Vol-C | 1.72/1.21 | JPEG と PDF のファイルが多い |

ボリューム中のデータを塊（チャンク）に区切るチャンキングは、ブロック固定長（以下 BF）とブロック可変長（以下 BV）とファイル可変長（以下 FV）の3種類を用意した。BV と FV は、[3]にあるようなローリングハッシュを用いてチャンクの境界を設定する手法を採用した。

BF はボリュームをブロックデバイスと見て64KiB 固定で区切って測定した。BV はボリュームをブロックデバイスと見て平均チャンク長が64KiB になることを期待した設定で測定した。FV は個々のファイルごとに、平均チャンク長が64KiB になることを期待した設定で測定した。

4. 測定結果と評価

前述の3種類のチャンキング(BF, BV, FV)に対応した自作の簡易ツールを使用した。簡易ツールは、切り出したチャンクに対し SHA1 を用いて

A measurement and evaluation of data duplication ratio for real fileservers volumes

Kazutaka OGIHARA

Fujitsu Laboratories LTD.

フィンガープリント (FP) を算出し、その FP の値でデータが同一かどうかを判定する。また、格納するチャンクは圧縮することを想定している。

表 2 に、Vol-A のチャンク数やデータ重複率の結果を示す。チャンクの分類は [1] にあるものを流用した。Logical (L) は論理的なチャンク数、Content (C) は物理的に格納しなければならないチャンク数を示し、L/C はデータ重複率を表す。

表 2. Vol-A のチャンク数とデータ重複率

| | | 1 日目 | 2 日目 | 3 日目 |
|------------------------|-----|------------|-------------|-------------|
| ブロック 固定長 4.55TiB | L | 76,409,616 | 152,819,232 | 229,228,848 |
| | C | 67,602,222 | 67,954,317 | 68,647,321 |
| | L/C | 1.130 | 2.249 | 3.339 |
| ブロック 可変長 4.55TiB | L | 89,473,499 | 178,963,716 | 268,477,341 |
| | C | 55,623,202 | 55,896,755 | 56,350,293 |
| | L/C | 1.609 | 3.202 | 4.764 |
| ファイル 可変長 4.08TiB | L | 87,978,655 | 176,109,390 | 264,696,380 |
| | C | 57,195,494 | 57,231,284 | 57,396,542 |
| | L/C | 1.538 | 3.077 | 4.612 |

フルバックアップ相当の処理を 3 日分するので、L は 2 日目で 1 日目の 2 倍に、3 日目で 3 倍になるが、C はほとんど変化しない。これはデータの更新がボリュームサイズと比較して非常に小さいことを示している。この傾向は Vol-B と Vol-C も同様である。また、データ重複率 L/C の日次経過によるチャンキング別の相対的な変動は見られなかった。よって、1 日目の結果のみを使いチャンキング別の差を見ることにする。

表 3. 3 つのボリュームのデータ重複率 (1 日目)

| | Vol-A | Vol-B | Vol-C |
|----|-------|-------|-------|
| BF | 1.130 | 1.251 | 1.461 |
| BV | 1.609 | 1.323 | 1.303 |
| FV | 1.538 | 1.211 | 1.123 |

表 3 から、BF と BV が FV に比べてデータ重複率が高いことが分かる。これは、FV ではバックアップ対象にならないボリュームの未使用領域が BF や BV ではバックアップ対象になることが影響していると考えている。BF や BV では特定の 1 チャンクへの重複が多く (例: Vol-A の BF では 5,469,598 の L が 1 つの C, BV では 2,733,278 の L が 1 つの C を参照)、このチャンクが未使用領域の大部分を構成していると思われる。

表 4. 補正後のデータ重複率 (1 日目)

| | Vol-A | Vol-B | Vol-C |
|----|-------|-------|-------|
| BF | 1.049 | 1.018 | 1.029 |
| BV | 1.559 | 1.200 | 1.101 |
| FV | 1.538 | 1.211 | 1.123 |

表 4 に、表 3 のデータに補正を加えて算出し直したデータ重複率を示す。補正方法は、BF と BV の L と C からデータ重複数の最も高いチャンクの情報を除く、というものである。

表 4 に示した補正の結果、FV が全体としては最も良い結果となった。また、BF は FV に比べ見劣りするが、BV は FV に差がないことが分かった。

表 5. 3 つのボリュームの予想格納容量 (1 日目)

| | Vol-A | Vol-B | Vol-C |
|----|----------|----------|----------|
| BF | 2.5[TiB] | 1.2[TiB] | 1.0[TiB] |
| BV | 1.9[TiB] | 1.0[TiB] | 1.0[TiB] |
| FV | 1.9[TiB] | 1.0[TiB] | 0.9[TiB] |

表 5 に、チャンキング別に予想されるデータ格納容量 (重複除去 + チャンク圧縮後の格納) を示す。本来であれば重複除去の管理データの考慮も必要だが、簡易ツールではチャンキングによる管理データの容量に差が見られなかったため省略した。予想格納容量が最も少ないのは FV だが、BV は FV に遜色のないことが分かった。バックアップ対象データ量は BV が FV よりも多いが、前述した未使用領域のため、BV と FV の予想格納容量には差が見られないと考えている。

5. まとめ

一般業務で使われている 3 つのファイルサーバボリューム (Vol-A/Vol-B/Vol-C) を対象に、ブロック固定長 (BF) ・ブロック可変長 (BV) ・ファイル可変長 (FV) の 3 種類のチャンキングを組み込んだ簡易ツールを用い、重複除去の効果を比較するための測定を行った。

測定した結果、データ重複率は、FV では Vol-A:1.538/Vol-B:1.211/Vol-C:1.123 となり最も高かったが、一方で、BV は Vol-A:1.559/Vol-B:1.200/Vol-C:1.101 となり、FV と同等の結果が得られた。また、予想される物理的なデータ格納容量も、BV は FV と同等の結果が得られた。

これらにより、この報告の目的としていた、BV は FV の代替になり得ることの確認、ができた。

参考文献

- [1] N. Park and D. J. Lilja: Characterizing Datasets for Data Deduplication in Backup Applications, IISWC 2010
- [2] D. T. Meyer and W. J. Bolosky: A Study of Practical Deduplication, FAST'11
- [3] "Content based slicing using Rabin-Karp hash", https://en.wikipedia.org/wiki/Rolling_hash (2015年12月21日時点)