

重複排除  
Inline と Post-Process  
の比較

2013年7月  
ワイヴァンアイエス研究所  
<http://www.y-vern-is.com>

# InlineとPost-Process プラス面とマイナス面の一般論

- › InlineとPost-Processの面からのみの比較
  - 他の、ソース型対ターゲット型などの観点からの検討ではない
- › 重複排除処理を高速化するためには、高性能CPUと大容量メモリーを必要とし、それはコストを押し上げることになっていた
- › ところが技術の進歩によって、CPUの高速高性能化と大容量メモリーの低コスト化が実現し、これらのリソースを豊富に使えるようになった
- › インライン方式採用に躊躇するマイナス要因が少なくなった

# 重複排除の処理タイミング

## Inline と Post Process のプラス面とマイナス面

	Inline	Post Process
	Ingest（バックアップサーバからのデータの取り込み）と同時に重複排除処理を行う、ディスク容量の効率化を優先する方式	バックアップサーバからのデータを特定のディスク領域に格納してから重複排除処理を行う、データのIngestを優先する方式
プラス面	<ul style="list-style-type: none"> <li>バックアップイメージを保持するlanding zoneが不要なため、トータル物理ディスク容量はPost-Process製品の約半分</li> <li>ラック搭載の高さ（U）がPost-Process製品より小さい</li> <li>レプリケーション開始時間が明確（バックアップ時のingest処理時間に左右されない）</li> <li>技術が最新のイメージ</li> </ul>	<ul style="list-style-type: none"> <li>データの取り込みに重複排除処理の影響を受けない（バックアップサーバから見て重複排除処理のオーバーヘッドを気にせずに完了できる）</li> <li>バックアップサーバから見たとき、バックアップ終了はInlineに比べると高速*1</li> <li>最新バックアップデータのリストアは高速（バックアップイメージがそのまま使える場合はブロックから再生成するrehydrate処理が不要）*2</li> <li>重複排除処理に必要なCPUやメモリーにかかるコストが少なく済む</li> </ul>
マイナス面	<ul style="list-style-type: none"> <li>Ingest時に重複排除処理のオーバーヘッド</li> <li>バックアップサーバから見たとき、バックアップ終了はinlineに比べると低速*1</li> <li>バックアップデータのリストアはInlineに比べて低速*2</li> <li>重複排除処理に高速なCPUとメモリー容量が必要・重要（CPUやメモリーにコストがかかる）</li> </ul>	<ul style="list-style-type: none"> <li>バックアップイメージを保持するlanding zoneが別途必要（少なくとも2倍のディスク容量）</li> <li>レプリケーションの開始時間が不明確（ingestに時間がかかるとレプリケーションの開始に遅延が発生）</li> <li>ラック搭載の高さ（U）がInline製品より大きくなる</li> <li>技術が古いイメージ</li> </ul>

\*1 実製品では、Inlineの方が高速の製品がある

\*2 実際のパフォーマンス値で確認が必要

ありがとうございます。

本文書に記載のすべての商標、商号、または製品名は、該当する各社の商標、商号、または製品名です。本文書は情報提供のみを目的として提供されており、ここに記載される内容は予告なく変更されることがあります。本文書は一切間違いがないことを保証するものではなく、いかなる他の保証や条件も提供するものではありません。この文書の使用によって直接的または間接的に生じた損害について、一切責任を負いません。