

arcserve Backup

データ デデュプリケーション

効果測定

Arcserve Japan

arcserve

Assured recovery

はじめに……

この資料は、データ デデュプリケーション機能を検討いただくにあたり、ディスク使用率とバックアップ パフォーマンスについて、データ デデュプリケーション デバイス（以降 DDD と記述）とファイル システム デバイス（以降 FSD と記述）を比較した資料になります。

- FSD は、ローカル マシンまたはネットワーク上のリモート マシン フォルダをバックアップ先としてテープと同様に活用するためのディスク上の仮想テープです。
- DDD は、データ デデュプリケーションを実行するためのデバイスです。FSD と同様にローカル マシンまたはリモート マシン上に作成し、バックアップ 先として活用するためのディスク上の仮想テープです。

※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

この資料は情報提供を目的としており、ここに記載されている明示的・暗示的・黙示的な情報にかかわらず、いかなる内容についてもArcserve Japan 合同会社は一切の責任を負わないものとします。この資料に掲載されている情報は将来予告なしに変更することがあります。内容の一部またはすべてをArcserve Japan 合同会社の許諾なく無断で販売、転載することは禁止されています。すべての製品名、サービス名、会社名およびロゴは、各社の商標、または登録商標です。

Agenda

- 1. 測定環境**
- 2. フル バックアップ**
- 3. 増分 / 増分バックアップ**
- 4. まとめ**

1. 測定環境

測定環境

マシン環境	
オペレーティング システム	Windows Server 2008 Enterprise Edition (64bit)
サービス パック	Service Pack 1
プロセッサ	Intel Xeon (E5320) 1.86GHz (4core CPUx2) HT有効
メモリ	8 GB
ディスク	SAS RAID5 15000rpm

ディスク構成	ドライブ	格納データ
ディスク 0	C:¥	システム、arcserve Backup r12.5※
	E:¥	DDD インデックス ファイル
ディスク 1	F:¥	バックアップ対象データ
	G:¥	FSDまたはDDD データ ファイル

※ 本資料はarcserve Backup r12.5の測定結果を元に作成していますが、製品動作や導入効果は r15/r16/r16.5 においても違いはございません。いずれの環境においても参考資料としてご活用いただけます。

2. フル バックアップ

同じバックアップ対象データを
以下の2パターンで実施

1. FSDへ毎日フル バックアップ
2. DDDへ毎日フル バックアップ

フル バックアップ

■ 7日間 (日～月) 以下のバックアップを実施

バックアップ先	バックアップ詳細
1. FSD (ファイル システム デバイス)	ローテーション スキーマ : 週7日 フル バックアップ
2. DDD (データ デデュプリケーション デバイス)	

■ バックアップ対象データ

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
データ量 (GB)	10.02	10.37	10.41	10.96	11.90	12.46	13.16
ファイル数	2,960	2,293	2,840	3,053	3,538	3,765	3,980

- 合計約 10GB のデータを準備
 - 毎日約 2GB の大きなファイルの編集
 - バックアップ対象全体量の約 10% の複数ファイルを編集、およびファイルの追加/削除を実施
- データは、主に Microsoft Word や Microsoft Excel など Microsoft Office 系のファイル

※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

フル バックアップ測定結果

<1. FSD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	10,254	20,905	31,564	42,782	54,970	67,732	81,211
ディスク使用量 (MB)	10,254	10,651	10,659	11,218	12,188	12,762	13,479
所要時間 (秒)	431	324	444	457	548	579	639

<2. DDD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	7,207	7,354	7,752	8,093	8,454	8,843	9,268
(A) ディスク使用量 (MB)	7,168	118	369	312	332	355	391
(B) インデックス ディスク使用量 (MB)	39	29	29	29	29	34	34
(C) ディスク使用量合計 (MB) (A+B)	7,207	147	398	341	361	389	425
圧縮割合	33%	98%	96%	97%	97%	97%	97%
所要時間 (秒)	378	119	138	138	151	151	161

- ① FSD は毎回バックアップ容量とほぼ同量のディスクを使用するが、DDD は2回目以降のバックアップではFSDの5%程度のディスク使用量となっていることが解る
- ② FSD はバックアップ時間もバックアップ容量に比例するが、DDD は2回目以降はバックアップ時間が短縮された

※ ここで挙げている測定結果は参考値です。実際のデータバックアップをデータデデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

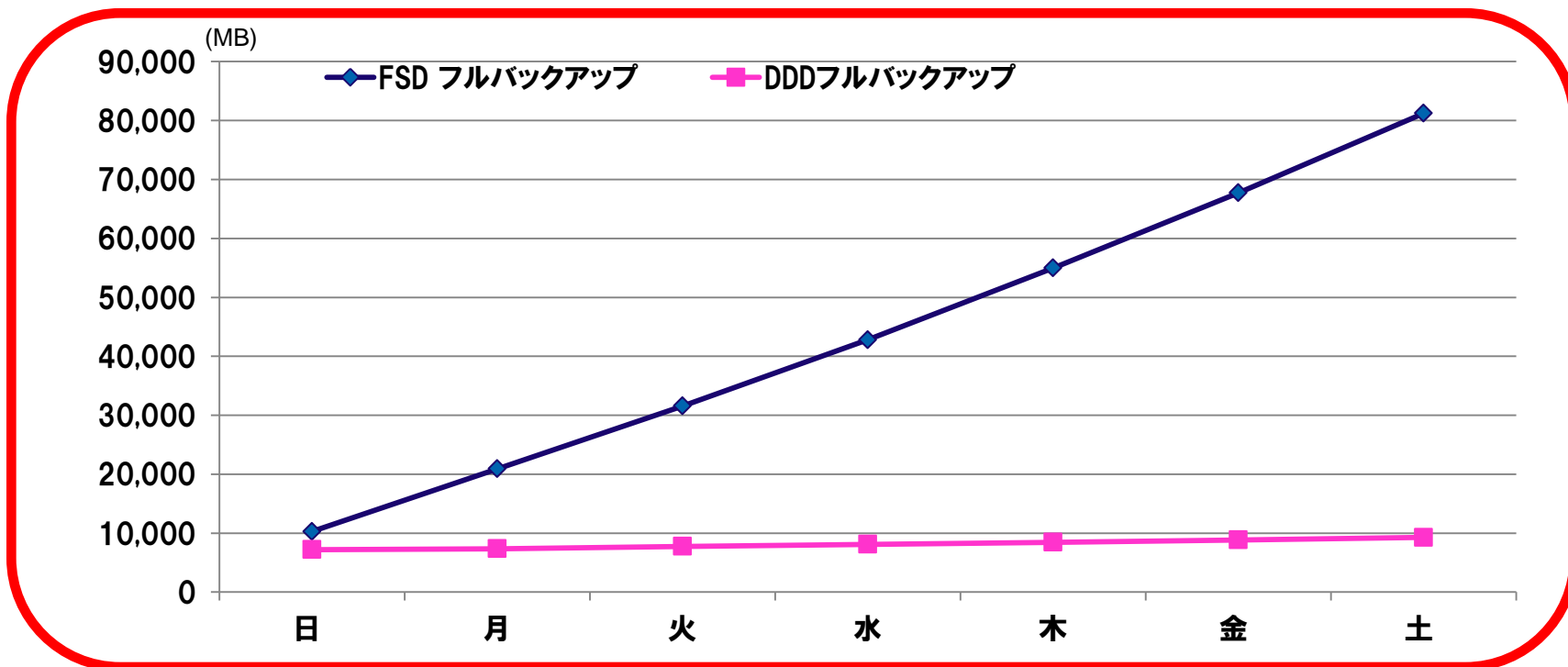
(参考)フル バックアップのディスク使用量比較 ①

<1. FSD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	10,254	20,905	31,564	42,782	54,970	67,732	81,211

<2. DDD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	7,207	7,354	7,752	8,093	8,454	8,843	9,268



※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

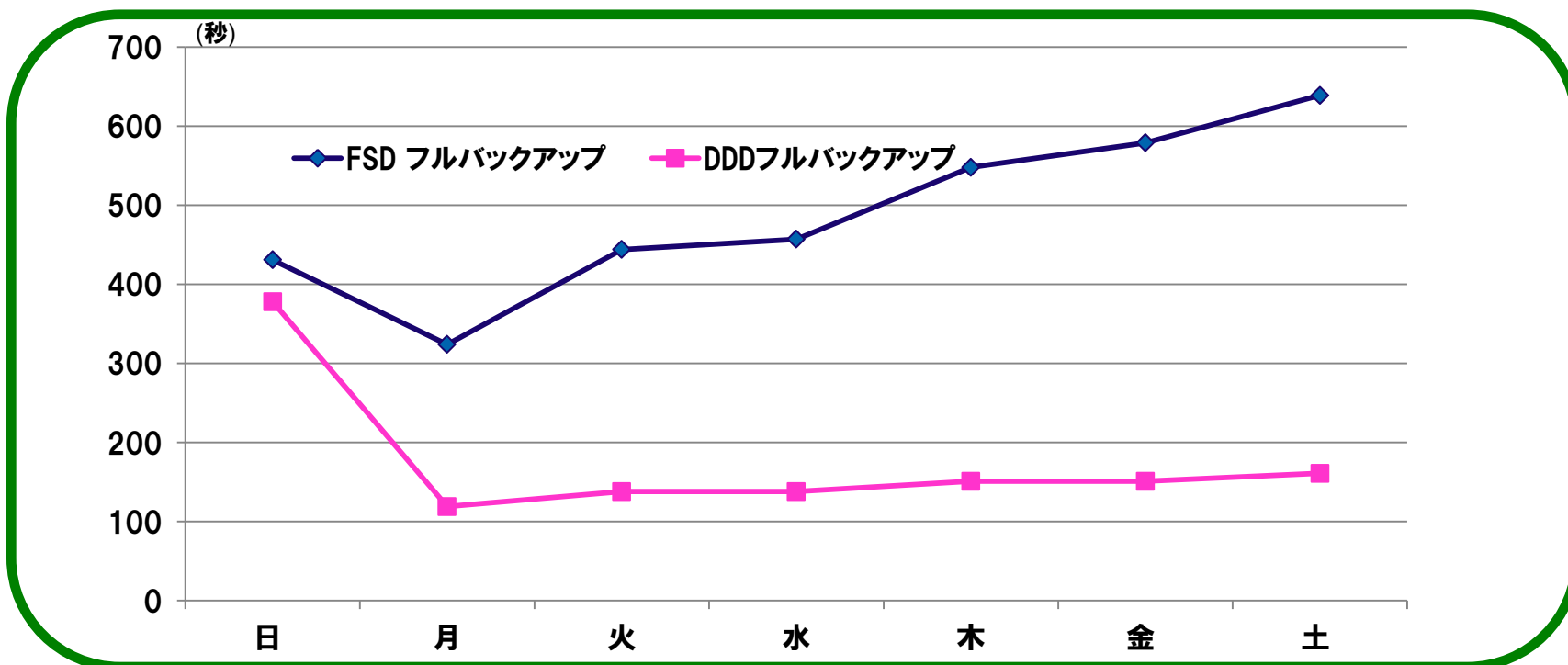
(参考)フル バックアップのディスク使用量比較 ②

<1. FSD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
所要時間 (秒)	431	324	444	457	548	579	639

<2. DDD フル バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
所要時間 (秒)	378	119	138	138	151	151	161



※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

(参考) リストア時間の比較

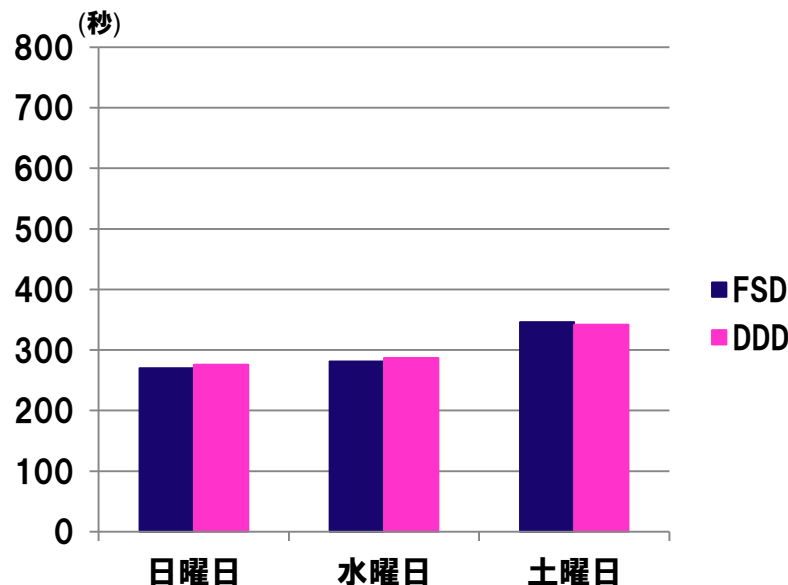
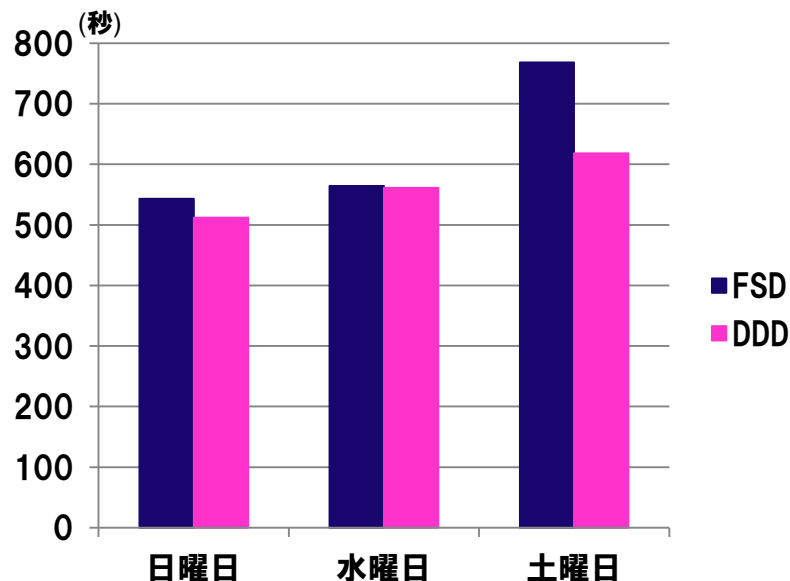
- 前ページで実行したバックアップ データを利用し、空のフォルダへ曜日ごとにすべてのデータをリストア

<リストア所要時間 (秒)>

	リストア先	日曜日	水曜日	土曜日
FSD (ディスク0)	F:\ (ディスク0)	543	564	768
DDD (ディスク0)	F:\ (ディスク0)	512	561	618

	リストア先	日曜日	水曜日	土曜日
FSD (ディスク0)	E:\ (ディスク1)	270	281	346
DDD (ディスク0)	E:\ (ディスク1)	276	287	342

I/O が重なり、リストア時間が長くなっています。

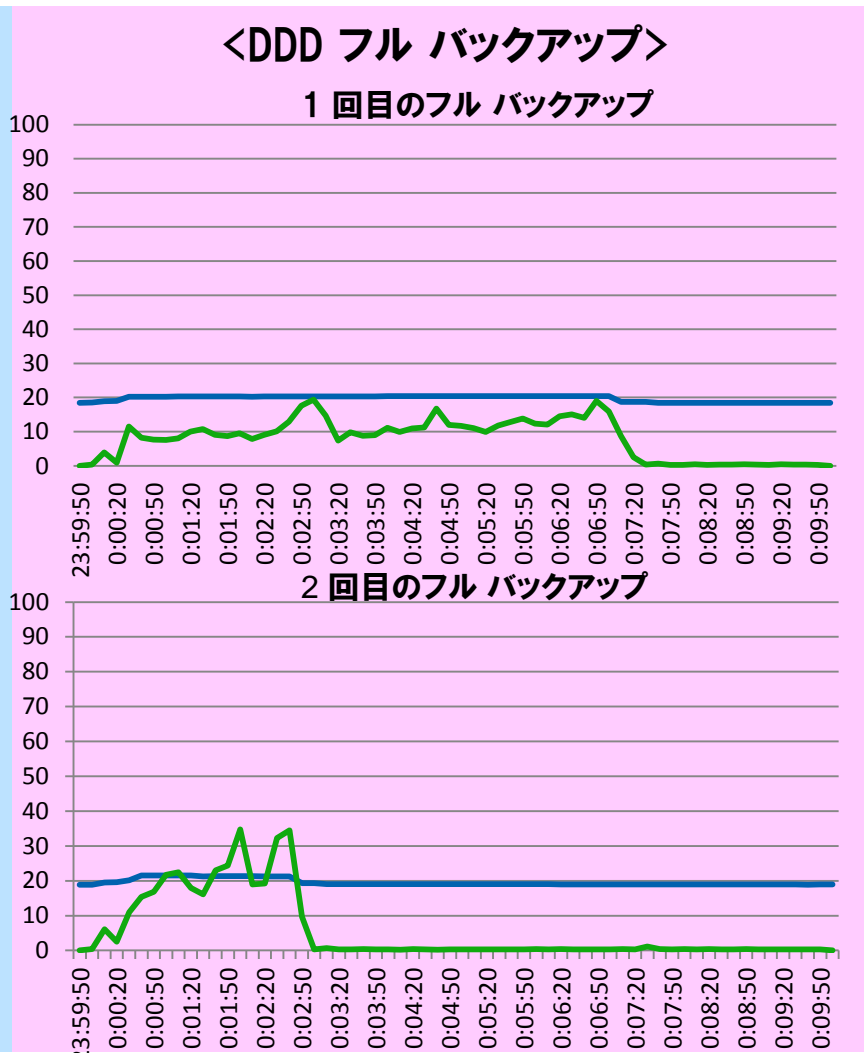
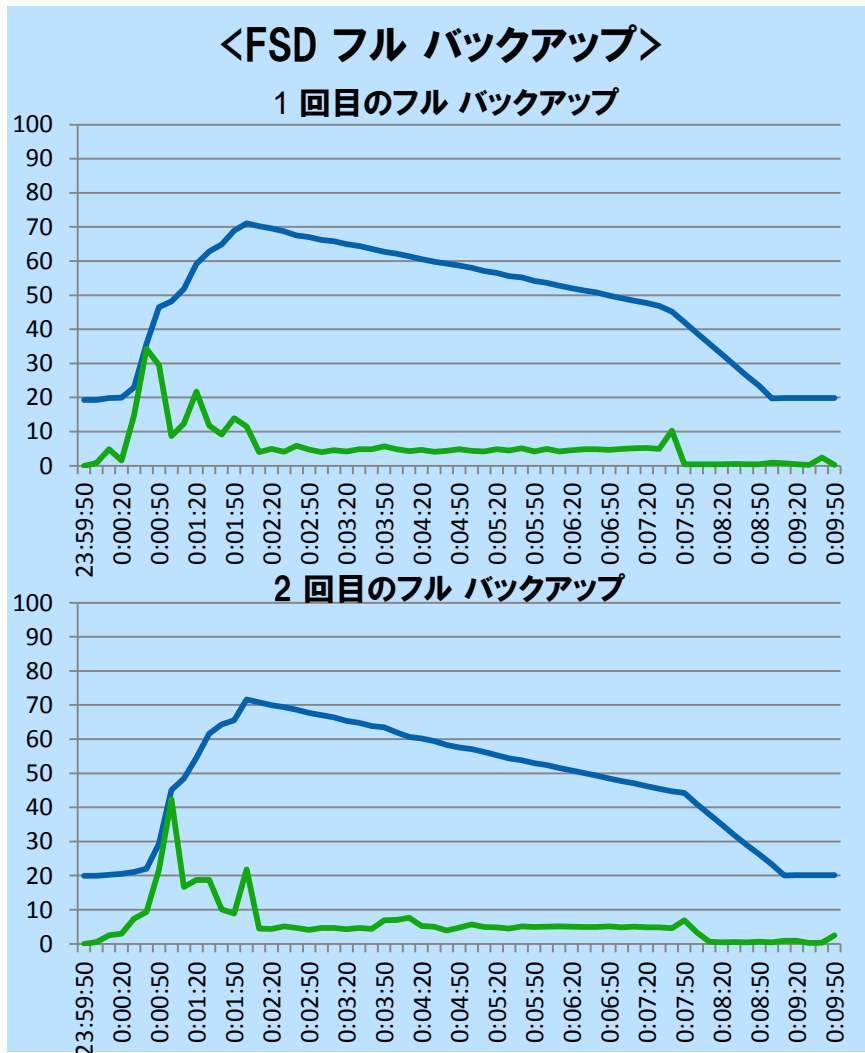


※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュープリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

(参考) フル バックアップのパフォーマンス

— メモリ使用量(%)
— プロセッサ処理時間(%)

バックアップ中のメモリ使用量およびプロセッサの処理時間の比較



※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュープリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

3. 増分/差分 バックアップ

同じバックアップ対象データを以下の
パターンで実施

1. FSD へ差分および増分バックアップ
2. DDD へ差分および増分バックアップ

差分バックアップ

■ 7日間（日～月）以下のバックアップを実施

バックアップ先	バックアップ詳細
FSD (ファイル システム デバイス)	ローテーション スキーマ: 週7日 差分バックアップ、日曜日にフル バックアップ
DDD (データ デデュプリケーション デバイス)	

■ バックアップ対象データ

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
差分データ量 (GB)	10.02	2.79	3.87	4.98	6.07※1	7.15	8.33
ファイル数	2,960	329	1,338	1,757	2,346	3,035	3,315

※1 20ページのまとめを参照

- 合計約 10GB のデータを準備
 - 毎日約 2GB の大きなファイルの編集
 - バックアップ対象全体量の約 10% の複数ファイルを編集、およびファイルの追加/削除を実施
- データは、主に Microsoft Word や Microsoft Excel など Microsoft Office 系のファイル

※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

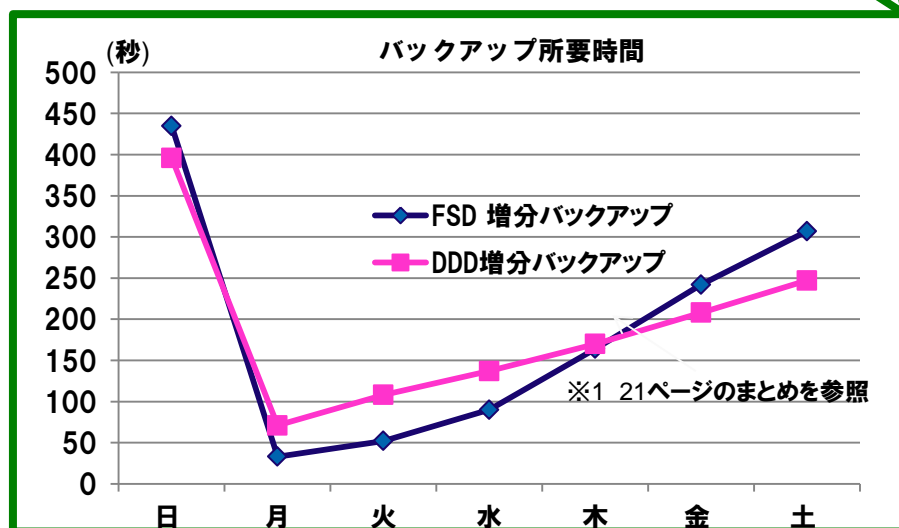
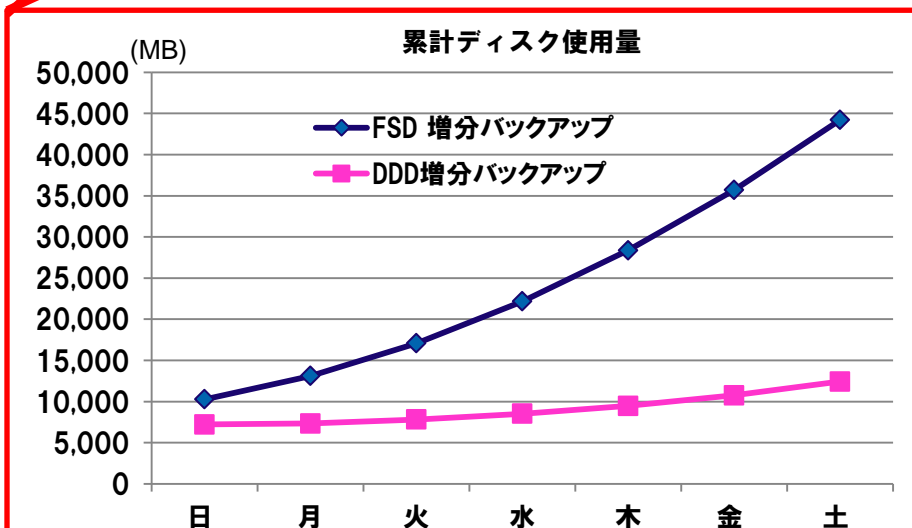
差分バックアップ測定結果

<FSD 差分バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	10,254	13,105	17,063	22,160	28,372	35,689	44,211
ディスク使用量 (MB)	10,254	2,851	3,958	5,097	6,212	7,317	8,522
所要時間 (秒)	435	33	52	90	164	242	307

<DDD 差分バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	7,207	7,335	7,815	8,517	9,459	10,749	12,424
(A) ディスク使用量 (MB)	7,168	118	470	687	922	1269	1649
(B) インデックス ディスク使用量 (MB)	39	10	10	15	20	21	26
(C) ディスク使用量合計 (MB) (A+B)	7,207	128	480	702	942	1290	1675
圧縮割合	33%	95%	88%	86%	85%	83%	81%
所要時間 (秒)	396	71	108	137	170	208	247



※ここで挙げている測定結果は参考値です。実際のデータバックアップをデータデデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

増分バックアップ

■ 7日間（日～月）以下のバックアップを実施

バックアップ先	バックアップ詳細
FSD (ファイル システム デバイス)	ローテーション スキーマ: 週7日 増分バックアップ、日曜日にフル バックアップ
DDD (データ デデュプリケーション デバイス)	

■ バックアップ対象データ

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
増分データ量 (GB)	10.02	2.79	2.95	3.01	3.02	3.04	3.17
ファイル数	2,960	329	1,010	420	590	690	281

- 合計約 10GB のデータを準備
 - 毎日約 2GB の大きなファイルの編集
 - バックアップ対象全体量の約 10% の複数ファイルを編集、およびファイルの追加/削除を実施
- データは、主に Microsoft Word や Microsoft Excel など Microsoft Office 系のファイル

※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

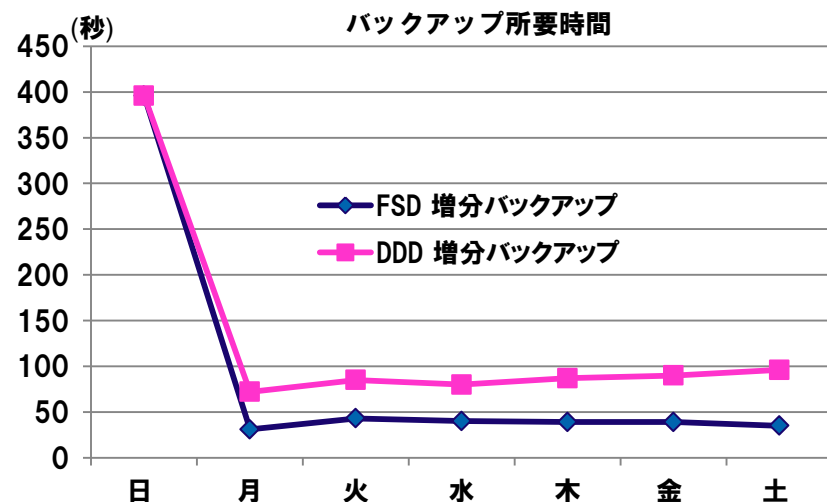
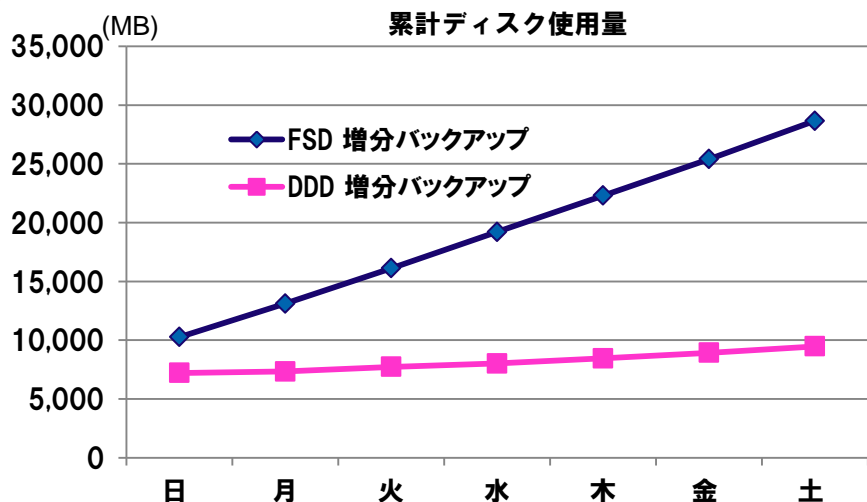
増分バックアップ測定結果

<FSD増分 バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	10,254	13,105	16,119	19,199	22,289	25,400	28,642
ディスク使用量 (MB)	10,254	2,851	3,014	3,080	3,090	3,111	3,242
所要時間 (秒)	396	31	43	40	39	39	35

<DDD増分 バックアップ>

	日曜日	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日
累計ディスク使用量 (MB)	7,207	7,335	7,726	8,021	8,444	8,926	9,467
(A) ディスク使用量 (MB)	7,168	118	381	285	413	472	531
(B) インデックス ディスク使用量 (MB)	39	10	10	10	10	10	10
(C) ディスク使用量合計 (MB) (A+B)	7,207	128	391	295	423	482	541
圧縮割合	33%	95%	87%	90%	86%	85%	84%
所要時間 (秒)	396	72	85	80	87	90	96



※ ここで挙げている測定結果は参考値です。実際のデータバックアップをデータデデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

(参考) リストア時間の比較

- 前ページまでに実行したバックアップ データを利用し、空のフォルダへそれぞれの曜日の対象となるすべてのデータをリストア

<差分バックアップ データのリストア所要時間(秒)>

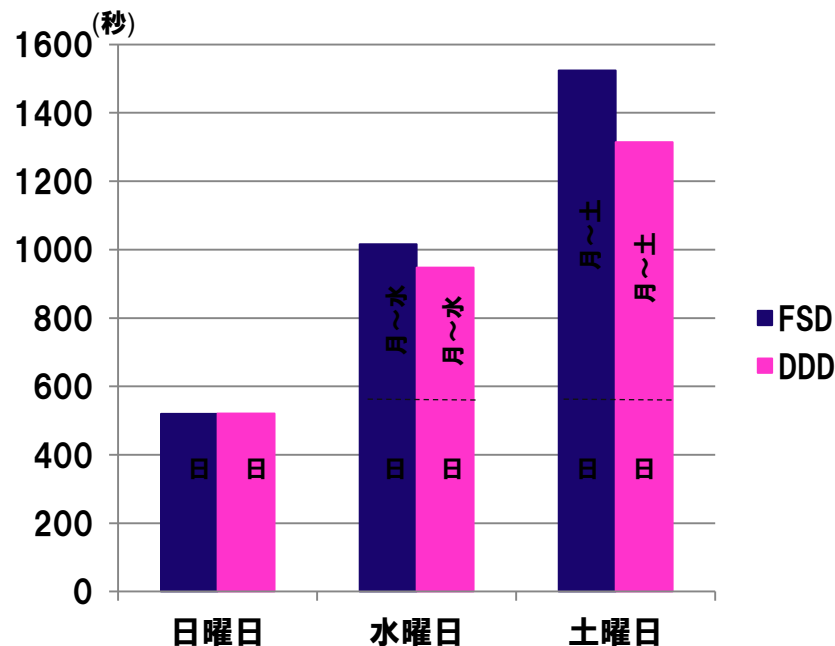
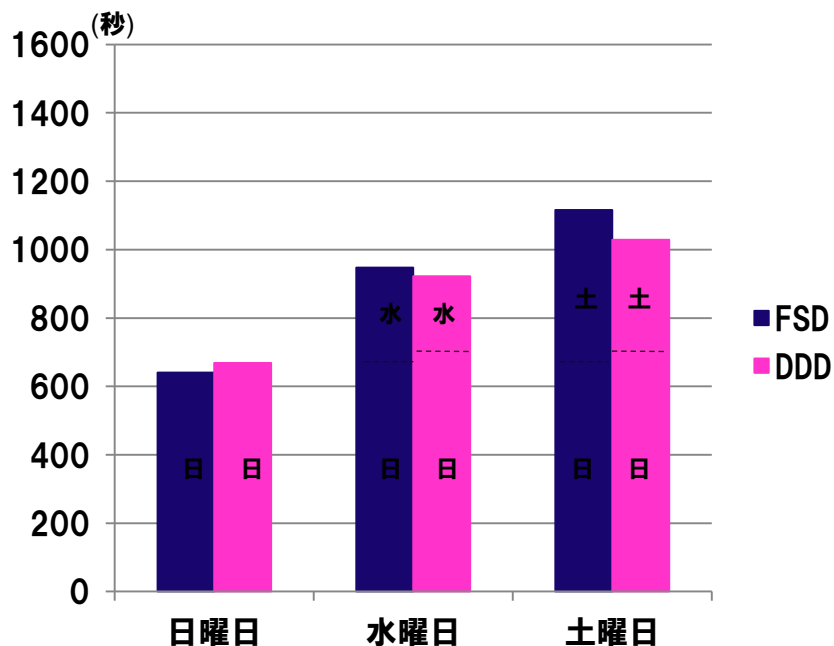
	リストア先	日曜日	水曜日	土曜日
FSD	F:\	640	947	1,116
DDD	F:\	669	922	1,029

※ 差分:水曜日=日(フル)+水、土曜日=日(フル)+土

<増分バックアップ データのリストア所要時間(秒)>

	リストア先	日曜日	水曜日	土曜日
FSD	F:\	520	1,016	1,524
DDD	F:\	521	947	1,314

※ 増分:水曜日=日(フル)から水まで、土曜日=日(フル)から土まで



※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュープリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

4. まとめ

まとめ

■ ディスク使用量

– フル バックアップ

1週間の累計ディスク使用量で比較すると、DDD を利用したバックアップでは、重複排除が行われ、FSD のおよそ1/8 以下のディスク使用量になっていることが確認できます。2日目以降のみを比較すると、およそ 1/30 のディスク使用量になっています。

– 差分/増分

DDD のディスク使用量は、およそ1/3 以下となり、フル バックアップと同様に少ないことがわかります。バックアップ対象の重複データの量と比例するためフル バックアップよりも重複排除率が低くなっています。

ディスクスペース有効活用

効果 大

<デデュプリケーションを利用したバックアップ運用の注意点>

デデュプリケーションは、デフォルトで4週間のパーシポリシーが設定されています。運用によってパーシ期間を変更してください。
参照 :「デデュプリケーション デバイスの作成方法」

<http://www.arcserve.com/~media/Files/TechnicalDocuments/asbu-r165-dedupe-definition.pdf>

まとめ

■ 処理時間

– フル バックアップ

DDD は、ディスク使用量だけではなく、ディスクに書き込む量が少ないため、バックアップ時間も短く、効率的なバックアップ運用が実現できます。

– 差分/増分フル バックアップ

今回の環境では、バックアップ対象データ量がおよそ 6 GB ※1 を超えたあたりから、DDD を利用したバックアップ時間の方が FSD よりも短くなります。そのため、今回の増分バックアップでは FSD を利用したバックアップの方が高速に感じられますが、環境およびデータ量 (数) によって変化します。

※1 15ページのグラフ参照

■ リストア

今回の測定では、ディスクが2つのためリストア時にファイル I/O が重なり、リストア時間が長くなっています。デデュプリケーションでは、ディスクが3つある場合には、インデックスおよびバックアップ対象データ保存先およびリストア先を、それぞれが物理的に違うディスクに設定することで、さらに時間短縮ができ、効率化すると考えられます。

時間短縮

効果 中

※ ここで挙げている測定結果は参考値です。実際のデータ バックアップをデータ デデュプリケーション機能を利用して実施した場合、結果が異なります。この結果を保証するものではありません。

arcserve Backup

データ デデュプリケーション

効果測定

Arcserve Japan

arcserve

Assured recovery