



# オブジェクトの格納方法（レプリケーションまたはイレイジャーコーディング） StorageGRID 11.5

NetApp  
April 11, 2024

# 目次

オブジェクトの格納方法（レプリケーションまたはイレイジャーコーディング）	1
レプリケーションとは	1
シングルコピーレプリケーションを使用しない理由	2
イレイジャーコーディングとは	5
イレイジャーコーディングスキームとは	7
イレイジャーコーディングのメリット、デメリット、および要件	10

# オブジェクトの格納方法（レプリケーションまたはイレイジャーコーディング）

StorageGRID では、レプリケートコピーを格納するかイレイジャーコーディングコピーを格納することで、オブジェクトを損失から保護できます。作成するコピーのタイプは、ILMルールの配置手順で指定します。

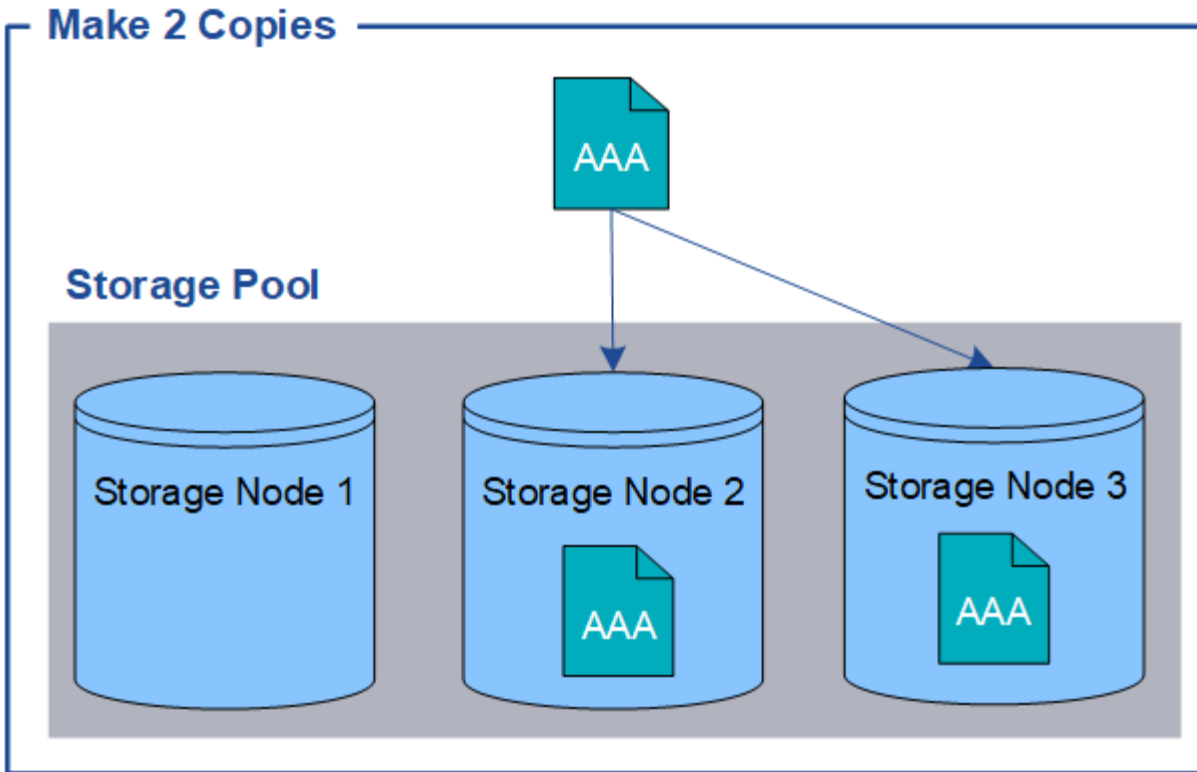
- ["レプリケーションとは"](#)
- ["シングルコピーレプリケーションを使用しない理由"](#)
- ["イレイジャーコーディングとは"](#)
- ["イレイジャーコーディングスキームとは"](#)
- ["イレイジャーコーディングのメリット、デメリット、および要件"](#)

## レプリケーションとは

レプリケーションは、StorageGRID がオブジェクトデータを格納するために使用する 2 つの方法のうちの一つです。レプリケーションを使用する ILM ルールにオブジェクトが一致すると、オブジェクトデータの完全なコピーが作成され、ストレージノードまたはアーカイブノードに格納されます。

レプリケートコピーを作成するように ILM ルールを設定する場合は、作成するコピーの数、コピーを配置する場所、およびそれぞれの場所にコピーを格納する期間を指定します。

次の例の ILM ルールは、各オブジェクトのレプリケートコピーを 2 つずつ、3 つのストレージノードからなるストレージプールに配置するように指定されています。



このルールにオブジェクトが一致した場合、StorageGRID はオブジェクトのコピーを 2 つ作成して、ストレージプール内の別々のストレージノードにそれぞれのコピーを配置します。この 2 つのコピーは、使用可能な 3 つのストレージノードのうちいずれか 2 つに配置されます。この場合、ストレージノード 2 と 3 に配置されています。コピーは 2 つあるため、ストレージプール内のいずれかのノードで障害が発生した場合でもオブジェクトを読み出すことができます。



StorageGRID が任意のストレージノードに格納できるレプリケートコピーは 1 つのオブジェクトにつき 1 つだけです。グリッドにストレージノードが 3 つあり、4 コピーの ILM ルールを作成した場合、作成されるコピーはストレージノードごとに 1 つだけになります。ILM placement unAchievable \* アラートがトリガーされ、ILM ルールを完全に適用できなかったことを示します。

#### 関連情報

["ストレージプールとは"](#)

["複数のストレージプールを使用したサイト間レプリケーション"](#)

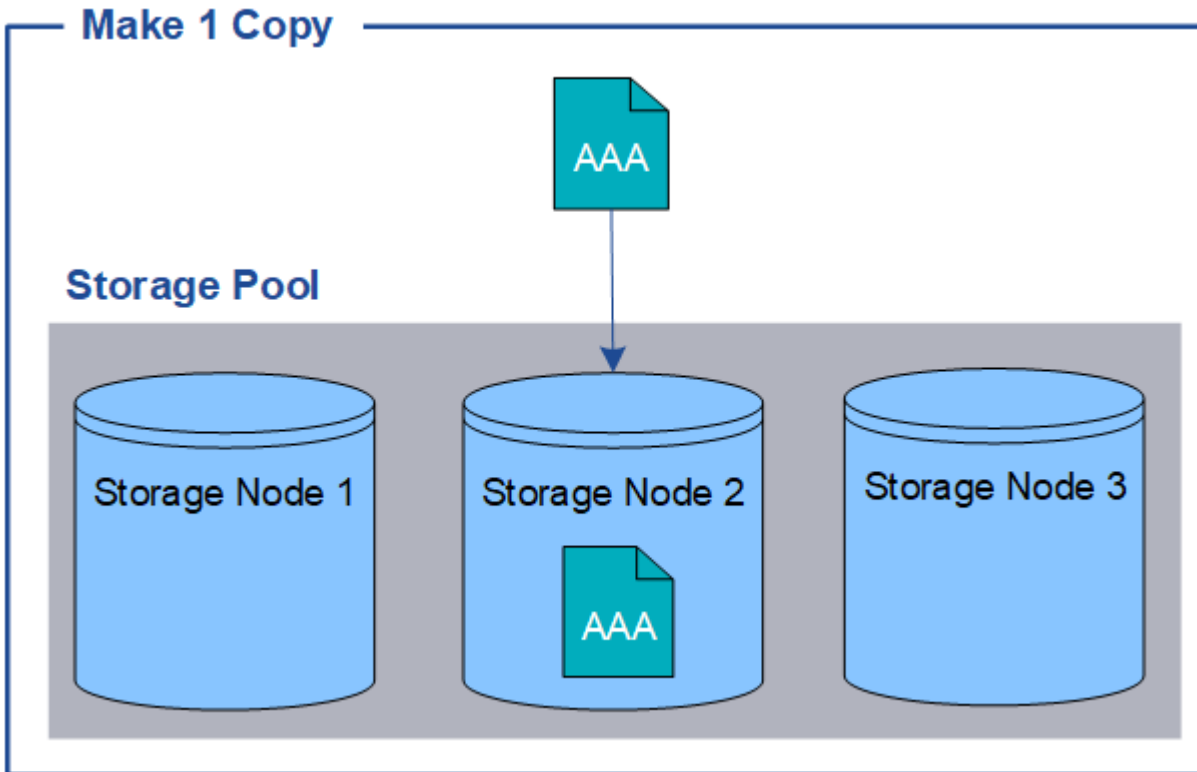
## シングルコピーレプリケーションを使用しない理由

レプリケートコピーを作成する ILM ルールを作成するときは、配置手順の任意の期間に少なくとも 2 つのコピーを指定する必要があります。

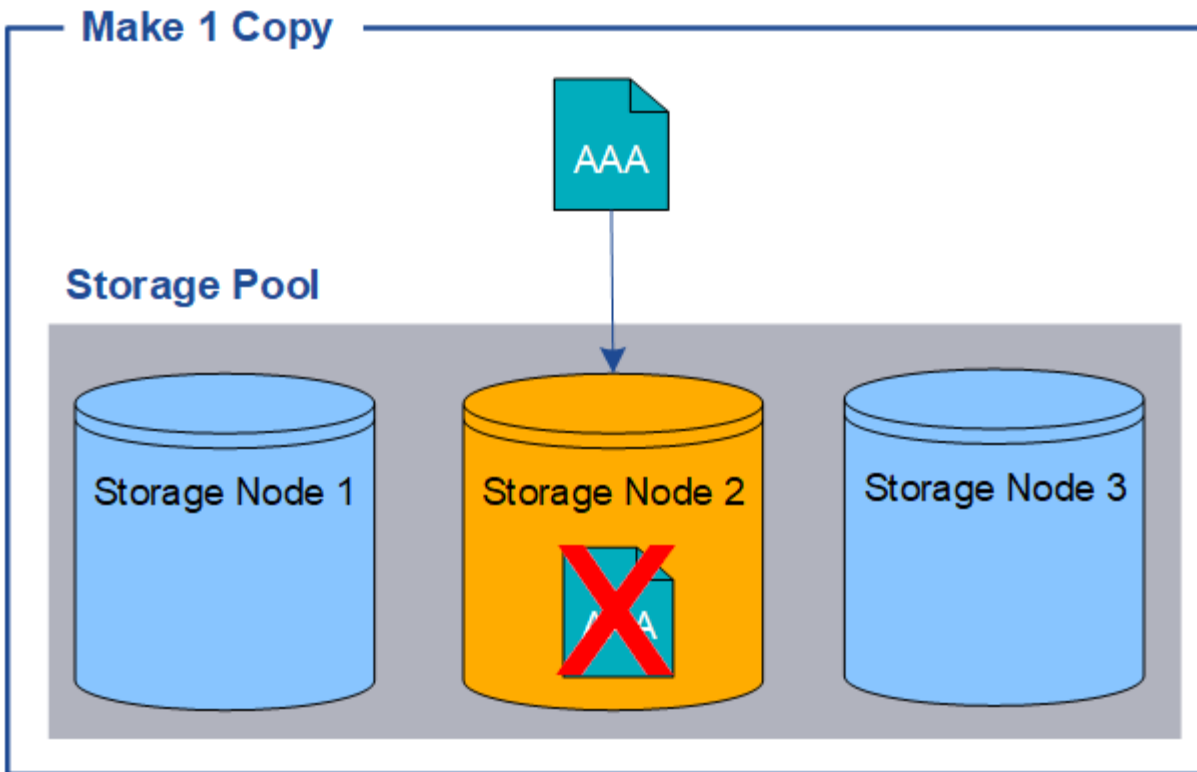


レプリケートコピーを 1 つだけ作成する ILM ルールは、どの期間も使用しないでください。オブジェクトのレプリケートコピーが 1 つしかない場合、ストレージノードに障害が発生したり、重大なエラーが発生すると、そのオブジェクトは失われます。また、アップグレードなどのメンテナンス作業中は、オブジェクトへのアクセスが一時的に失われます。

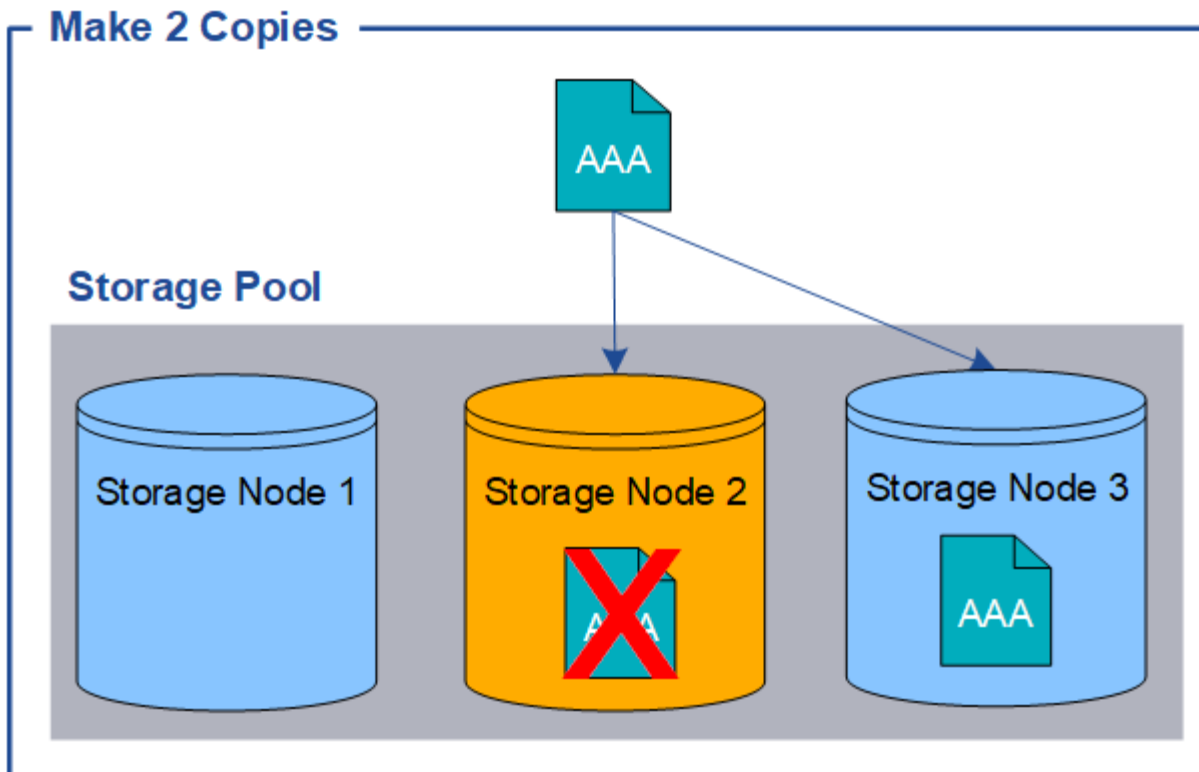
次の例では、Make 1 Copy ILM ルールによって、1つのオブジェクトのレプリケートコピーを3つのストレージノードからなるストレージプールに配置するように指定しています。このルールに一致するオブジェクトが取り込まれると、StorageGRID は1つのストレージノードにのみコピーを配置します。



ILM ルールにオブジェクトのレプリケートコピーが1つしか作成されていない場合、ストレージノードが使用できなくなるとオブジェクトにアクセスできなくなります。この例では、アップグレードやその他のメンテナンス手順の実行中など、ストレージノード2がオフラインになるとオブジェクトAAAへのアクセスが一時的に失われます。ストレージノード2で障害が発生すると、オブジェクトAAAが完全に失われます。



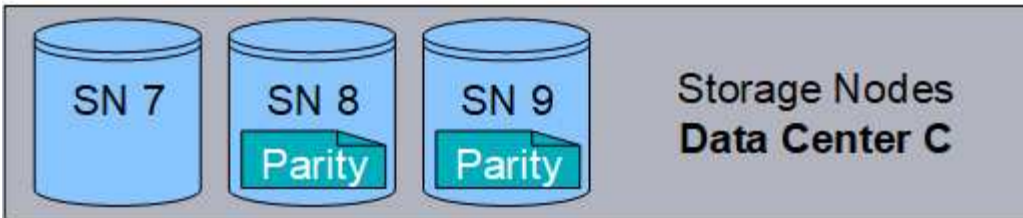
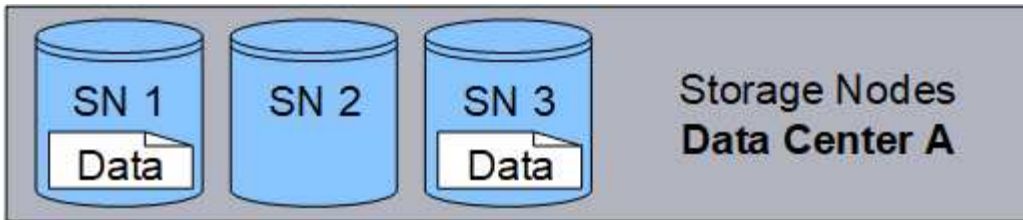
オブジェクトデータの損失を防ぐには、レプリケーションで保護するすべてのオブジェクトのコピーを常に2つ以上作成する必要があります。コピーが複数ある場合も、1つのストレージノードに障害が発生した場合やオフラインになった場合でもオブジェクトにアクセスできます。



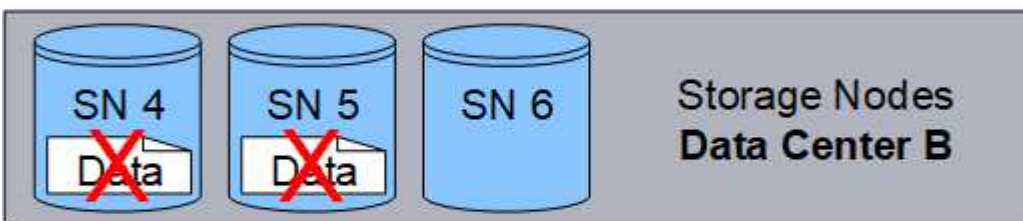
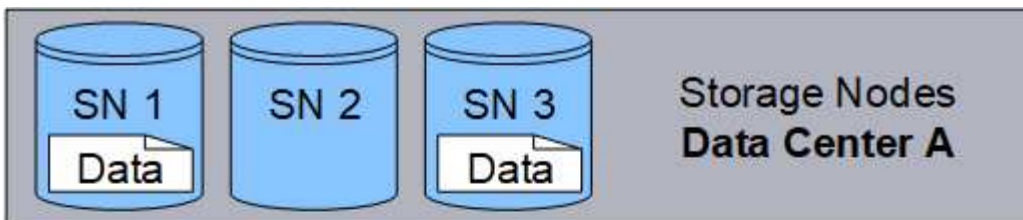
## イレイジャーコーディングとは

イレイジャーコーディングは、オブジェクトデータを格納するために StorageGRID で使用される 2 つ目の方法です。StorageGRID がイレイジャーコーディングコピーを作成するために設定された ILM ルールとオブジェクトを照合する場合は、オブジェクトデータを複数のデータフラグメントに分割し、追加のパリティフラグメントを計算して、各フラグメントを別のストレージノードに格納します。アクセスされたオブジェクトは、格納されたフラグメントを使用して再アセンブルされます。データフラグメントまたはパリティフラグメントが破損したり失われたりしても、イレイジャーコーディングアルゴリズムが残りのデータフラグメントとパリティフラグメントを使用してそのフラグメントを再作成します。

次の例は、オブジェクトのデータに対するイレイジャーコーディングアルゴリズムの使用方法を示しています。この例の ILM ルールでは 4+2 のイレイジャーコーディングスキームを使用します。各オブジェクトは 4 つのデータフラグメントに等分され、オブジェクトデータから 2 つのパリティフラグメントが計算されます。ノードやサイトの障害時にもデータが保護されるよう、6 つの各フラグメントは 3 つのデータセンターサイトの別々のノードに格納されます。

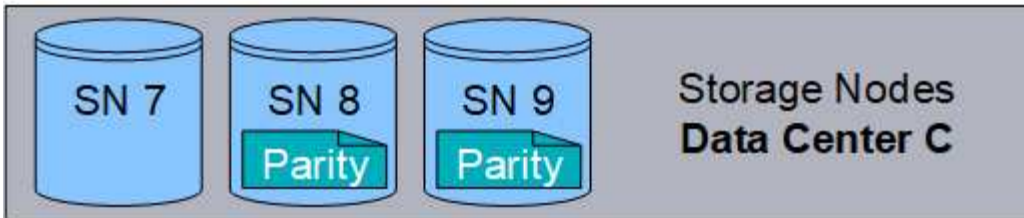
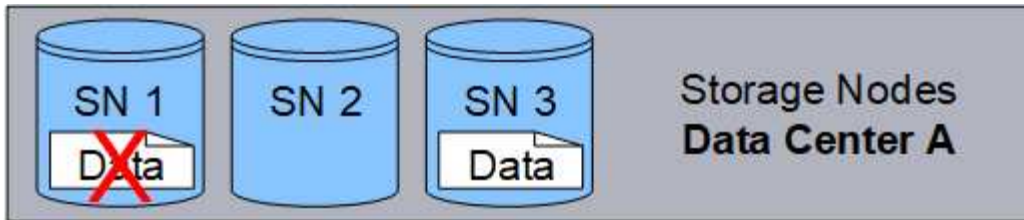


4+2 のイレイジャーコーディングスキームでは、少なくとも 9 個のストレージノードが必要です。このノードには、3 つのサイトそれぞれに 3 個のストレージノードが必要です。6 つのうちのいずれか 4 つのフラグメント（データまたはパリティ）が使用可能であれば、オブジェクトを読み出すことができます。最大 2 つのフラグメントが失われても、オブジェクトデータが失われることはありません。データセンターサイト全体で障害が発生した場合でも、他のすべてのフラグメントに引き続きアクセスできれば、オブジェクトの読み出しまたは修復が可能です。



3 つ以上のストレージノードが失われると、オブジェクトを読み出せなくなります。





関連情報

["ストレージプールとは"](#)

["イレイジャーコーディングスキームとは"](#)

["イレイジャーコーディングプロファイルの設定"](#)

## イレイジャーコーディングスキームとは

ILM ルールにイレイジャーコーディングプロファイルを設定する場合は、使用するストレージプールを構成するストレージノードとサイトの数に基づいて、使用可能なイレイジャーコーディングスキームを選択します。イレイジャーコーディングスキームは、各オブジェクト用に作成されるデータフラグメントとパリティフラグメントの数を制御します。

StorageGRID システムは、Reed-Solomon イレイジャーコーディングアルゴリズムを使用します。このアルゴリズムは、オブジェクトを  $k$  個のデータフラグメントに分割して、 $m$  個のパリティフラグメントを計算します。 $k + m = n$  個のフラグメントが  $n$  個のストレージノードに分散され、データ保護を提供します。失われたフラグメントまたは破損したフラグメントは、オブジェクトが保持できる最大  $m$  個です。 $k$  個のフラグメントがオブジェクトの読み出しまたは修復に必要です。

イレイジャーコーディングプロファイルを設定する場合は、ストレージプールについて次のガイドラインに従ってください。

- ストレージプールには 3 つ以上のサイト、または 1 つのサイトだけが含まれている必要があります。



ストレージプールにサイトが2つ含まれている場合、イレイジャーコーディングプロファイルは設定できません。

- 3つ以上のサイトを含むストレージプールのイレイジャーコーディングスキーム
- 1サイトのストレージプールのイレイジャーコーディングスキーム

- デフォルトのストレージプール、すべてのストレージノード、またはデフォルトサイトであるすべてのサイトを含むストレージプールは使用しないでください。
- ストレージプールには少なくとも  $k + m + 1$  ストレージノードを含める必要があります。

必要なストレージノードの最小数は、 $k + m$  です。ただし、必要なストレージノードが一時的に使用できない場合に、少なくとも1つのストレージノードを追加することで、取り込みエラーやILMバックログが発生するのを防ぐことができます。

イレイジャーコーディングスキームのストレージオーバーヘッドは、パリティフラグメントの数 ( $m$ ) をデータフラグメントの数 ( $k$ ) で割ることによって計算されます。ストレージオーバーヘッドを使用して、各イレイジャーコーディングオブジェクトに必要なディスクスペースを計算できます。

$$\text{disk space} = \text{object size} + (\text{object size} * \text{storage overhead})$$

たとえば、4+2スキームを使用して10MBのオブジェクト（ストレージオーバーヘッドが50%）を格納すると、そのオブジェクトが消費するグリッドストレージは15MBです。6+3のストレージオーバーヘッドを含む6+2スキームを使用して同じ10MBのオブジェクトを格納すると、オブジェクトが消費するサイズは約13.3MBになります。

合計値が最も小さいイレイジャーコーディングスキーム ( $k + m_{\text{that}}$ ) をニーズに合わせて選択します。フラグメント数が少ないイレイジャーコーディングスキームは全体的に計算効率が高く、1つのオブジェクトに作成されて分散（または取得）されるフラグメント数が少なくて済むため、フラグメントサイズが大きい場合パフォーマンスが向上し、ストレージの追加が必要になった場合に拡張時に必要なノード数が少なくて済みます。（ストレージ拡張の計画については、StorageGRIDの拡張手順を参照してください）。

### 3つ以上のサイトを含むストレージプールのイレイジャーコーディングスキーム

次の表に、3つ以上のサイトを含むストレージプールについて、StorageGRIDで現在サポートされているイレイジャーコーディングスキームを示します。これらの方式はいずれもサイト障害からの保護を提供します。1つのサイトが失われてもオブジェクトには引き続きアクセスできます。

サイト損失の保護を提供するイレイジャーコーディングスキームの場合、ストレージプールに推奨されるストレージノードの数は各サイトに少なくとも3つのストレージノードが必要なため  $k + m + 1$  を超えています。

イレイジャーコーディングスキーム ( $k + m$ )	サイトの最小数	各サイトで推奨されるストレージノードの数	推奨されるストレージノードの総数	サイト障害からの保護	ストレージオーバーヘッド
4+2	3.	3.	9.	はい。	50%
6+2	4.	3.	12.	はい。	33%
8+2	5.	3.	15	はい。	25%

イレイジャーコーディングスキーム ( $k+m$ )	サイトの最小数	各サイトで推奨されるストレージノードの数	推奨されるストレージノードの総数	サイト障害からの保護	ストレージオーバーヘッド
6+3	3.	4.	12.	はい。	50%
9+3	4.	4.	16	はい。	33%
2+1	3.	3.	9.	はい。	50%
4+1	5.	3.	15	はい。	25%
6+1	7.	3.	21	はい。	17%
7+5	3.	5.	15	はい。	71%



StorageGRID では、サイトごとに少なくとも 3 つのストレージノードが必要です。7+5 スキームを使用するには、各サイトに少なくとも 4 つのストレージノードが必要。サイトごとに 5 つのストレージノードを使用することを推奨します。

サイト保護を提供するイレイジャーコーディングスキームを選択する場合は、次の要素の相対的な重要性を調整します。

- \* フラグメント数 \* : フラグメントの総数が少ないほど、一般にパフォーマンスと拡張の柔軟性が向上します。
- \* フォールトトレランス \* : パリティセグメントの数を増やすことでフォールトトレランスが向上します (  $_m_$  の値が大きい場合 ) 。
- \* ネットワーク・トラフィック \* : 障害から回復する場合、フラグメント数の多いスキーム (つまり、  $k+m$  ) を使用すると、より多くのネットワーク・トラフィックが生成されます。
- \* ストレージ・オーバーヘッド \* : オーバーヘッドの大きいスキームでは、オブジェクトごとにより多くのストレージ・スペースが必要です。

たとえば、4+2 と 6+3 のどちらかのスキーム (どちらも 50% のストレージオーバーヘッドがある) を選ぶ場合、フォールトトレランスをさらに高める必要がある場合は 6+3 のスキームを選択します。ネットワークリソースが制限されている場合は、4+2 のスキームを選択します。他のすべての要素が等しい場合は、フラグメントの合計数が少ないため、4+2 を選択します。



使用するスキームが不明な場合は、4+2 または 6+3 を選択するか、テクニカルサポートにお問い合わせください。

## 1 サイトのストレージプールのイレイジャーコーディングスキーム

1 サイトのストレージプールでは、サイトに十分な数のストレージノードがある場合、3 つ以上のサイト用に定義されたすべてのイレイジャーコーディングスキームがサポートされます。

必要なストレージノードの最小数は  $k+m_1$  ですが、  $_k+m+1$  ストレージノードを含むストレージプールを推奨します。たとえば、2+1 イレイジャーコーディングスキームには少なくとも 3 つのストレージノードが

異なるストレージプールが必要ですが、推奨されるストレージノード数は4つです。

イレイジャーコーディングスキーム ( $k+m$ )	ストレージノードの最小数	推奨されるストレージノードの数	ストレージオーバーヘッド
4+2	6.	7.	50%
6+2	8.	9.	33%
8+2	10.	11.	25%
6+3	9.	10.	50%
9+3	12.	13.	33%
2+1	3.	4.	50%
4+1	5.	6.	25%
6+1	7.	8.	17%
7+5	12.	13.	71%

関連情報

["グリッドを展開します"](#)

## イレイジャーコーディングのメリット、デメリット、および要件

レプリケーションとイレイジャーコーディングのどちらを使用してオブジェクトデータを損失から保護するかを決定する前に、イレイジャーコーディングのメリット、デメリット、および要件を理解しておく必要があります。

### イレイジャーコーディングのメリット

イレイジャーコーディングは、レプリケーションに比べて信頼性、可用性、ストレージ効率に優れています。

- **\* 信頼性 \***：信頼性はフォールトトレランス、つまり同時にデータを失うことなく維持できる障害の数によって判断されます。レプリケーションでは、複数の同一コピーが異なるノード上およびサイト間に格納されます。イレイジャーコーディングの場合、オブジェクトはデータフラグメントとパリティフラグメントにエンコードされ、多数のノードとサイトに分散されます。この分散によってサイトとノード両方の障害からの保護を提供します。イレイジャーコーディングは、同等のストレージコストでレプリケーションよりも優れた信頼性を提供します。
- **\* 可用性 \***：可用性は、ストレージノードに障害が発生した場合や、ノードにアクセスできなくなった場合にオブジェクトを読み出すことができるかどうかによって定義されます。イレイジャーコーディングは、同等のストレージコストでレプリケーションよりも優れた可用性を提供します。

- \* Storage Efficiency \* : 可用性と信頼性が同等レベルの場合、イレイジャーコーディングで保護されたオブジェクトが消費するディスクスペースは、同じオブジェクトをレプリケーションで保護する場合よりも少なくなります。たとえば、10MB のオブジェクトを 2 つのサイトにレプリケートするとディスクスペースを 20MB (2 つのコピー) 消費しますが、6+3 のイレイジャーコーディングスキームを使用して 3 つのサイトにイレイジャーコーディングされたオブジェクトが消費するディスクスペースは 15MB のみです。



イレイジャーコーディングオブジェクトのディスクスペースは、オブジェクトサイズにストレージオーバーヘッドを加えたものです。ストレージオーバーヘッドの割合は、パーティフラグメント数をデータフラグメント数で割って算出します。

## イレイジャーコーディングのデメリット

レプリケーションと比較した場合のイレイジャーコーディングのデメリットは次のとおりです。

- より多くのストレージノードとサイトが必要です。たとえば、6+3 のイレイジャーコーディングスキームを使用する場合は、3 つのサイトに少なくとも 3 つのストレージノードが必要です。一方、オブジェクトデータをレプリケートする場合は、各コピーに必要なストレージノードは 1 つだけです。
- ストレージの拡張にかかるコストと複雑さが増大します。レプリケーションを使用する環境を拡張するには、オブジェクトコピーを作成するすべての場所にストレージ容量を追加するだけです。イレイジャーコーディングを使用する環境を拡張する場合は、使用中のイレイジャーコーディングスキームと、既存のストレージノードの使用率の両方を考慮する必要があります。たとえば、既存のノードが 100% フルになるまで待つ場合は、少なくとも  $k + m$  Storage ノードを追加する必要があります。ただし、既存のノードが 70% フルになった時点で拡張する場合は、サイトごとに 2 つのノードを追加し、使用可能なストレージ容量を最大化できます。詳細については、StorageGRID の拡張手順を参照してください。
- 地理的に分散したサイトでイレイジャーコーディングを使用する場合は、読み出しのレイテンシが上昇します。イレイジャーコーディングされてリモートサイトに分散されたオブジェクトのフラグメントを WAN 接続経由で読み出す場合、レプリケートされてローカル (クライアントの接続先と同じサイト) で利用可能なオブジェクトよりも時間がかかります。
- 地理的に分散したサイトでイレイジャーコーディングを使用する場合は、特に WAN ネットワーク接続経由でオブジェクトを頻繁に読み出ししたり修復したりするケースでは読み出しと修復の WAN ネットワークトラフィックが増大します。
- サイト間でイレイジャーコーディングを使用する場合は、サイト間のネットワークレイテンシの上昇に伴ってオブジェクトの最大スループットが大幅に低下します。この最大スループットの低下は TCP ネットワークのスループットが低下したことによるもので、StorageGRID システムによるオブジェクトフラグメントの格納 / 読み出し速度に影響します。
- コンピューティングリソースの利用率が向上します。

## イレイジャーコーディングを使用する状況

イレイジャーコーディングは次の要件に最適です。

- 1MBを超えるオブジェクト



イレイジャーコーディングコピーに関連付けられているフラグメント数の管理でオーバーヘッドが発生するため、200KB以下のオブジェクトにはイレイジャーコーディングを使用しないでください。

- 頻繁に読み出されないコンテンツの長期保存またはコールドストレージ。
- 高いデータ可用性と信頼性。
- サイトやノードの障害に対する保護
- ストレージ効率
- 複数のレプリケートコピーではなく 1 つのイレイジャーコーディングコピーのみを使用して効率的にデータを保護する必要のある単一サイト環境
- サイト間レイテンシが 100 ミリ秒未満の複数サイト環境

#### 関連情報

["グリッドを展開します"](#)

## 著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用权を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用权については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

## 商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。