



# Red Hat Ceph Storage 4.2

## リリースノート

Red Hat Ceph Storage 4.2 のリリースノート



## Red Hat Ceph Storage 4.2 リリースノート

---

Red Hat Ceph Storage 4.2 のリリースノート

Enter your first name here. Enter your surname here.

Enter your organisation's name here. Enter your organisational division here.

Enter your email address here.

## 法律上の通知

Copyright © 2022 | You need to change the HOLDER entity in the en-US/Release\_Notes.ent file |.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux<sup>®</sup> is the registered trademark of Linus Torvalds in the United States and other countries.

Java<sup>®</sup> is a registered trademark of Oracle and/or its affiliates.

XFS<sup>®</sup> is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL<sup>®</sup> is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js<sup>®</sup> is an official trademark of Joyent. Red Hat is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack<sup>®</sup> Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

## 概要

本リリースノートには、Red Hat Ceph Storage 4.2 製品リリース向けに実装された主要な機能、機能拡張、既知の問題、バグ修正を説明します。本リリースノートには、現在のリリースより前の Red Hat Ceph Storage 4.2 リリースのリリースノートが含まれます。Red Hat では、コード、ドキュメント、Web プロパティにおける配慮に欠ける用語の置き換えに取り組んでいます。まずは、マスター (master)、スレーブ (slave)、ブラックリスト (blacklist)、ホワイトリスト (whitelist) の 4 つの用語の置き換えから始めます。この取り組みは膨大な作業を要するため、今後の複数のリリースで段階的に用語の置き換えを実施して参ります。詳細は、弊社の CTO、Chris Wright のメッセージを参照してください。

## 目次

第1章 はじめに .....	3
第2章 謝辞 .....	4
第3章 新機能 .....	5
3.1. CEPH ANSIBLE ユーティリティー	5
3.2. CEPH MANAGEMENT DASHBOARD	6
3.3. CEPH ファイルシステム	6
3.4. CEPH MANAGER プラグイン	6
3.5. CEPH OBJECT GATEWAY	6
3.6. マルチサイトの CEPH OBJECT GATEWAY	7
3.7. RADOS	8
3.8. RADOS ブロックデバイス (RBD)	8
第4章 テクノロジープレビュー .....	9
4.1. ブロックデバイス (RBD)	9
4.2. OBJECT GATEWAY	9
第5章 非推奨の機能 .....	10
5.1. CEPH VOLUME ユーティリティー	10
第6章 バグ修正 .....	11
6.1. CEPH ANSIBLE ユーティリティー	11
6.2. COCKPIT CEPH インストーラー	13
6.3. CEPH MANAGEMENT DASHBOARD	14
6.4. CEPH ファイルシステム	14
6.5. CEPH MANAGER プラグイン	15
6.6. CEPH VOLUME ユーティリティー	15
6.7. CEPH OBJECT GATEWAY	16
6.8. マルチサイトの CEPH OBJECT GATEWAY	17
6.9. RADOS	17
6.10. RADOS ブロックデバイス (RBD)	18
6.11. NFS GANESHA	18
第7章 既知の問題 .....	20
7.1. CEPH ANSIBLE ユーティリティー	20
7.2. CEPH VOLUME ユーティリティー	20
7.3. マルチサイトの CEPH OBJECT GATEWAY	21
第8章 ソース .....	22



## 第1章 はじめに

Red Hat Ceph Storage は、非常にスケーラブルでオープンなソフトウェア定義のストレージプラットフォームであり、最も安定したバージョンの Ceph ストレージシステムと Ceph 管理プラットフォーム、デプロイメントユーティリティー、およびサポートサービスを組み合わせたものです。

Red Hat Ceph Storage ドキュメントは、<https://access.redhat.com/documentation/en/red-hat-ceph-storage/> から入手できます。

## 第2章 謝辞

Red Hat Ceph Storage バージョン 4.2 には、Red Hat Ceph Storage チームの数多くの貢献が反映されています。さらに Ceph プロジェクトでは、Ceph コミュニティーの個人や組織からの貢献の度合いが質と量の両面で大幅に拡大しています。Red Hat Ceph Storage チームの全メンバー、Ceph コミュニティーの個々の貢献者、および以下の組織を含むすべての方々の貢献に謝意を表します。

- Intel
- Fujitsu
- UnitedStack
- Yahoo
- Ubuntu Kylin
- Mellanox
- CERN
- Deutsche Telekom
- Mirantis
- SanDisk
- SUSE



## 第3章 新機能

本セクションでは、Red Hat Ceph Storage の今回のリリースで導入された主要な更新、機能拡張、新機能の一覧を紹介します。

### 3.1. CEPH ANSIBLE ユーティリティー

#### Playbook `ceph-ansible` が複数のノードからログを収集

今回のリリースにより、Playbook は大規模なクラスターの複数のノードからログを自動的に収集するようになりました。

#### `ceph-ansible` が、2つのサイト間で追加の接続チェックを実行

今回の更新により、`ceph-ansible` は `realm pull` の前に2つのサイト間で追加の接続チェックを実行するようになりました。

#### パージPlaybookは、未使用のCephファイルを削除します

このリリースでは、パージクラスターPlaybookは、Red Hat Ceph Storage クラスターをパージした後、`grafana-server` ノード上のすべてのCeph関連の未使用ファイルを削除します。

#### `--skip-tags wait_all_osds_up` オプションを使用して、すべてのOSDが起動するのを待機するチェックをスキップします

このリリースでは、ストレージクラスターのアップグレード中に、Ansibleランタイムで `--skip-tags wait_all_osds_up` オプションを使用することで、ユーザーはこのチェックをスキップできるため、ディスクに障害が発生した場合 `rolling_update.yml` Playbook の障害を防ぐことができます。

#### 既存プールの `crush_rule` が更新可能に

以前は、プールの作成時に特定のプールの `crush_rule` 値が設定されていましたが、後で更新できませんでした。今回のリリースで、既存プールの `crush_rule` 値を更新できるようになりました。

#### カスタムの `crush_rule` を RADOS Gateway プールに設定可能

今回のリリースで、RADOS ゲートウェイプールには、OpenStack、MDS、および Client などの他のプールに加えて、カスタムの `crush_rule` の値を使用できるようになりました。

#### `ceph_docker_http_proxy` および `ceph_docker_https_proxy` を設定して、HTTP(s) プロキシの背後でコンテナレジストリーに関するプロキシ問題を解決

以前のバージョンでは、`/etc/profile.d` ディレクトリーで定義された環境変数が読み込まれないと、ログインおよびプルレジストリー操作が失敗していました。今回の更新により、環境変数 `ceph_docker_http_proxy` または `ceph_docker_https_proxy` を設定することにより、HTTP(s) プロキシの背後にあるコンテナレジストリーが想定通りに機能するようになりました。

#### Ceph Ansible は Ansible 2.9 でのみ動作

以前のリリースでは、`ceph-ansible` は移行ソリューションとしてバージョン 2.8 および 2.9 の Ansible をサポートしていました。今回のリリースで、`ceph-ansible` は Ansible 2.9 のみをサポートするようになりました。

#### Dashboard がデフォルトで HTTPS に設定

以前のバージョンでは、Dashboard は `http` に設定されていました。今回のリリースにより、Dashboard はデフォルトでは `https` に設定されます。

## Playbook を終了する前に、ceph-mon サービスのマスクが解除

以前のリリースでは、障害の発生時に **ceph-mon** systemd サービスがマスクされ、Playbook の失敗により、サービスを手動で再起動することができませんでした。今回のリリースで、障害発生時に Playbook を終了する前に **ceph-mon** サービスのマスクを解除し、ユーザーはローリングアップデート Playbook を再起動する前に **ceph-mon** サービスを手動で再起動できるようになりました。

## 3.2. CEPH MANAGEMENT DASHBOARD

### Red Hat Ceph Storage Dashboard でユーザーのバケットクォータの使用状況を表示

今回のリリースにより、Red Hat Ceph Storage Dashboard に現在のサイズ、使用率、およびオブジェクト数など、ユーザーのバケットクォータの使用状況が表示されるようになりました。

## 3.3. CEPH ファイルシステム

### mgr/volumes CLIを使用してcephx認証IDを一覧表示できるようになりました

以前は、**ceph\_volume\_client** インターフェースを使用してcephx認証IDを一覧表示していました。このインターフェースは非推奨になりました。

このリリースでは、Manilaなどのコンシューマーは、mgr/volumeインターフェースを使用して、サブボリュームへのアクセスが許可されているcephx認証IDを一覧表示できます。

#### 構文

```
ceph fs subvolume authorized_list _VOLUME_NAME__SUB_VOLUME_NAME_ [--group_name=_GROUP_NAME_]
```

## 3.4. CEPH MANAGER プラグイン

### Ceph Manager のパフォーマンスを向上するために、C++ インターフェースへの内部 Python が変更

以前のリリースでは、**pg\_dump** により、Ceph Manager のパフォーマンスに影響するすべての情報が提供されていました。今回のリリースにより、C++ インターフェースへの内部 python が変更になり、モジュールで **pg\_ready**、**pg\_stats**、**pool\_stats**、および `'osd_ping_times'` に関する情報が提供されるようになりました。

### 進行中のモジュールをオフにすることが可能

以前のバージョンでは、**always-on** マネージャーモジュールであったため進捗モジュールはオフになっていませんでした。今回のリリースで、**ceph progress off** を使用して進捗モジュールをオフにし、**ceph progress on** を使用してオンにできるようになりました。

## 3.5. CEPH OBJECT GATEWAY

### バケットインデックス (rgw\_bucket\_index\_max\_aio) での Ceph Object Gateway のデフォルトシャード要求が 128 に増加

以前のバージョンでは、バケットインデックスの未処理のシャード要求が 8 に制限されます。これにより、バケットを一覧表示してパフォーマンスが低下しました。今回のリリースにより、バケットインデックス **rgw\_bucket\_index\_max\_aio** のシャード要求のデフォルト数が 8 から 128 に増え、バケット一覧のパフォーマンスが向上しました。

## クラスターログ情報にバケットのレイテンシー情報が含まれる

以前のバージョンでは、ログのクラスター情報にはバケット要求のレイテンシーが含まれていましたが、各バケットのレイテンシー情報は指定されませんでした。今回のリリースにより、ログの各行にはバケット名、オブジェクト名、リクエストID、操作の開始時間および操作名が含まれます。

今回の機能拡張により、ログの解析時にお客様がこの情報を収集できるようになりました。操作のレイテンシーを計算するには、**awk** スクリプトを使用して、操作の開始時点からログメッセージの時間を引きます。

## Ceph Object Gateway ログに Beast のアクセスログが追加される

今回のリリースにより、フロントエンドの Web サーバーである Beast の Ceph Object Gateway ログに、Apache スタイルのアクセスログ行が追加されました。今回のログの更新は、接続およびクライアントネットワークの問題の診断に役立ちます。

## Beast フロントエンドの明示的な要求タイムアウト

以前のバージョンでは、レイテンシーの高いネットワークで接続したクライアントなどの低速なクライアント接続は、アイドル状態の場合にドロップされる可能性がありました。

今回のリリースにより、`/etc/ceph.conf` の新たな **request\_timeout\_ms** オプションにより、Beast フロントエンドに明示的なタイムアウトを設定する機能が追加されました。**request\_timeout\_ms** のデフォルト値は 65 秒です。

要求タイムアウトが大きいと、Ceph Object Gateway に低速なクライアントに対する耐性を持たせることができます。これにより、接続がドロップされる回数が少なくなる可能性があります。

## 欠落データを含む RGW オブジェクトを一覧表示

以前のバージョンでは、誤って削除されたデータを持つ RGW オブジェクトは管理者にとって不明であったため、この問題への最適な対処方法を判断することができませんでした。今回のリリースにより、クラスター管理者は **rgw-gap-list** を使用し、データが欠落している可能性のある RGW オブジェクトの候補を一覧表示できるようになりました。

## 3.6. マルチサイトの CEPH OBJECT GATEWAY

### データ同期ロギングの処理が遅延する

以前のバージョンでは、データ同期ロギングはログエントリーの大規模なバックログを処理する際に遅延する可能性がありました。

今回のリリースにより、データ同期にバケット同期ステータスのキャッシュが含まれるようになりました。キャッシュを追加すると、バックログが存在する場合に重複するデータログエントリーの処理が高速になります。

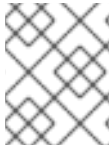
### データ同期ロギングの処理が遅延する

以前のバージョンでは、データ同期ロギングはログエントリーの大規模なバックログを処理する際に遅延する可能性がありました。

今回のリリースにより、データ同期にバケット同期ステータスのキャッシュが含まれるようになりました。キャッシュを追加すると、バックログが存在する場合に重複するデータログエントリーの処理が高速になります。

### 複数サイトの同期ロギングで FIFO を使用し、ロギングの RADOS データオブジェクトへのオフロードが可能になる

以前のバージョンでは、マルチサイトのメタデータとデータロギング設定は OMAP データログを使用していました。今回のリリースにより、FIFO データロギングが利用できるようになりました。グリーンフィールドデプロイメントで FIFO を使用するには、設定オプション `rgw_default_data_log_backing` を `fifo` に設定します。



#### 注記

設定値では、大文字と小文字が区別されます。**fifo** を小文字で使用して設定オプションを設定します。

サイトが使用するデータログバックングを変更するには、コマンド `radosgw-admin --log-type fifo datalog type` を使用します。

## 3.7. RADOS

### Ceph messenger プロトコルが msgr v2.1 に修正

今回のリリースで、新しいバージョンの Ceph messenger プロトコル msgr v2.1 が実装され、以前のバージョンである msgr v2.0 に存在する複数のセキュリティ、整合性、および潜在的なパフォーマンスの問題に対応するようになりました。すべての Ceph エンティティ (デーモンおよびクライアントの両方) で、デフォルトで msgr v2.1 に設定されるようになりました。

### Ceph の正常性に関する詳細がクラスターログに記録される

以前のバージョンでは、クラスターログに Ceph の正常性の詳細情報がないため、問題の根本原因を結論付けるのは困難でした。今回のリリースにより、Ceph の正常性の詳細情報がクラスターのログに記録されるようになったため、クラスターで発生する可能性のある問題を確認できます。

### PG 削除コードの効率の向上

以前のバージョンでは、すべての配置グループ (PG) で最後に削除されたオブジェクトへのポインターが維持されていなかったため、毎回すべてのオブジェクトに対する不要な反復が発生していました。今回のリリースにより、クライアント I/O への影響が少ない PG の削除のパフォーマンスが向上しました。`osd_delete_sleep_ssd` パラメーターおよび `osd_delete_sleep_hybrid` パラメーターは、デフォルト値 1 秒になりました。

## 3.8. RADOS ブロックデバイス (RBD)

### rbd カーネルクライアント上のカスタムネットワーク名前空間から `-o noudev` コマンドを実行する新しいオプション

以前は、rbd カーネルクライアントのカスタムネットワーク名前空間からの `rbd map`、`rbd unmap` などのコマンドは、手動で操作するまでハングしていました。今回のリリースで、`rbd map -o noudev` や `rbd unmap -o noudev` などのコマンドに `-o noudev` オプションを追加すると、想定通りに機能します。これは、OCP のネットワークにデフォルトの OpenShift SDN の代わりに Multus を使用する場合に便利です。

## 第4章 テクノロジープレビュー

本セクションでは、Red Hat Ceph Storage の本リリースで導入または更新されたテクノロジープレビュー機能の概要を説明します。



### 重要

テクノロジープレビュー機能は、Red Hat の実稼働環境でのサービスレベルアグリーメント (SLA) ではサポートされていないため、Red Hat では実稼働環境での使用を推奨していません。これらの機能は、近々発表予定の製品機能をリリースに先駆けてご提供することにより、お客様は機能性をテストし、開発プロセス中にフィードバックをお寄せいただくことができます。

Red Hat のテクノロジープレビュー機能のサポートについての詳細は、<https> を参照してください。

### 4.1. ブロックデバイス (RBD)

#### RBD イメージの NBD イメージへのマッピング

**rbd-nbd** ユーティリティーは、RADOS Block Device(RBD)イメージを Network Block Devices(NBD)にマップし、Ceph クライアントが Kubernetes 環境内のボリュームおよびイメージにアクセスできるようにします。**rbd-nbd** を使用するには、**rbd-nbd** パッケージをインストールします。詳細は、**rbd-nbd(7)** man ページを参照してください。

### 4.2. OBJECT GATEWAY

#### Object Gateway アーカイブサイト

今回のリリースにより、アーカイブサイトがテクノロジープレビューとしてサポートされるようになりました。アーカイブサイトでは、アーカイブゾーンに関連付けられたゲートウェイからのみ除外できる S3 オブジェクトのバージョンの履歴を指定できます。マルチゾーン設定にアーカイブゾーンを含めると、S3 オブジェクトのレプリカが残りのゾーン内で消費する領域を節約しつつ、1つのゾーンにのみ S3 オブジェクトの履歴の柔軟性を持たせることができます。

## 第5章 非推奨の機能

本セクションでは、Red Hat Ceph Storage の本リリースまでのすべてのマイナーリリースで非推奨となった機能の概要を説明します。

### Ubuntu がサポート対象外

Ubuntu への Red Hat Ceph Storage 4 クラスターのインストールはサポートされなくなりました。Red Hat Enterprise Linux を基礎となるオペレーティングシステムとして使用します。

### ceph-ansible を使用した iSCSI ゲートウェイの設定はサポート対象外

**ceph-ansible** ユーティリティーを使用した Ceph iSCSI ゲートウェイの設定はサポートされなくなりました。**ceph-ansible** を使用してゲートウェイをインストールし、続いて **gwcli** ユーティリティーを使用して Ceph iSCSI ゲートウェイを設定します。詳細は、『Red Hat Ceph Storage ブロックデバイスガイド』の「[Ceph iSCSI ゲートウェイ](#)」の章を参照してください。

### ceph-disk が非推奨に

今回のリリースで、**ceph-disk** ユーティリティーがサポートされなくなりました。代わりに **ceph-volume** ユーティリティーが使用されます。Red Hat Ceph Storage 4 の『[管理ガイド](#)』に含まれる「[ceph-volume が `ceph-disk` に置き換えられた理由](#)」セクションを参照してください。

### FileStore は実稼働環境でサポートされなくなる

FileStore OSD バックエンドは、新規 BlueStore バックエンドが実稼働環境で完全にサポートされるため、非推奨になりました。『Red Hat Ceph Storage インストールガイド』の「[オブジェクトストアを FileStore から BlueStore に移行する方法](#)」セクションを参照してください。

### Ceph 設定ファイルが非推奨に

Ceph 設定ファイル (**ceph.conf**) が非推奨になり、Ceph Monitor に保管された新たな集中設定が提供されるようになりました。詳細は、『Red Hat Ceph Storage 設定ガイド』の「[Ceph 設定データベース](#)」セクションを参照してください。

## 5.1. CEPH VOLUME ユーティリティー

### dm-cache のデバイス設定での OSD のデプロイはサポート対象外

**dm-cache** を使用した OSD のデプロイはサポートされなくなりました。**dm-cache** の代わりに Bluestore バックエンドを使用してください。

詳細は、『Red Hat Ceph Storage インストールガイド』の「[オブジェクトストアを FileStore から BlueStore に移行する方法](#)」セクションを参照してください。

## 第6章 バグ修正

本セクションでは、今回リリースされた Red Hat Ceph Storage で修正されたユーザーに大きな影響を及ぼすバグを説明します。また、セクションでは、以前のバージョンで見つかり修正された既知の問題を説明します。

### 6.1. CEPH ANSIBLE ユーティリティー

#### /targetsのURLは、Prometheusがダウンしていることを示している

以前は、PrometheusのターゲットURLは**localhost**値で設定されていました。**localhost**値により、ターゲットステータスがダウンしていると見なされたときに、Prometheusサービスが**localhost**アドレスをリッスンしませんでした。このリリースでは、ターゲットURL値にIPアドレスを使用するようにPrometheus構成ファイルが更新されました。その結果、Prometheusターゲットステータスは正しくレポートされています。

(BZ#1933560)

#### ceph-volumeは、Cephデプロイメント中にOSDを作成するときにメタデータの破損を引き起こす可能性がある

以前は、**ceph-volume**がボリュームグループ、論理ボリュームの作成、タグの設定などのLVMコマンドを実行すると、Cephのデプロイ時にOSDを作成するときにメタデータが破損する可能性があります。このリリースでは、ユーザーはホスト上の**group\_vars/all.yml**ファイルで**lvmetad\_disabled**パラメーターを**true**に設定することにより、条件付きで**lvmetad**サービスを無効にできるため、メタデータの破損が回避されます。

(BZ#1955040)

#### ceph-ansibleのロールceph-dashboardは、強制的に自己署名証明書のコモンネームをceph-dashboardに設定する

これまでは、**ceph-ansible**で生成された自己署名証明書を使用すると、**ceph-dashboard**にコモンネーム(CN)が強制されるため、クライアントに証明書を送信するノードのホスト名の不一致により、Prometheusなどのアプリケーションがエラーになりました。

このリリースでは、**ceph-ansible**がCNに適切な値を設定し、Prometheusが期待どおりに機能します。

(BZ#1978869)

#### Cephコンテナが配置されていると、ローリングアップグレードが失敗する

Ceph MonitorデーモンとCeph Object Gatewayデーモンがコンテナと同じ場所に配置されている、およびマルチサイトCeph Object Gatewayが有効になっていると、**rolling\_update.yml** Ansible Playbookは失敗します。この失敗は、アップグレードプロセス中にCeph Monitorコンテナが停止したために**radosgw-admin**コマンドを実行できないことが原因で発生しました。このリリースでは、**ceph-handler**ロール内のマルチサイトのCeph Object Gatewayコードは、アップグレードプロセス中にスキップされます。その結果、**rolling\_update.yml** Ansible Playbookは正常に実行されます。

(BZ#1984880)

#### コンテナ化されたデーモンに切り替えると、Ceph Monitorのクォーラムチェックが失敗する

リグレッションバグが、**switch-from-non-containerized-to-containerized-ceph-daemons.yml** Ansible Playbookで発生する。このリグレッションバグにより、現在のノードのホスト名がテストされていないため、Ceph Monitorクォーラムチェックが失敗しました。このリリースでは、現在のCeph

Monitorノードからの**ansible\_hostname**ファクトが正しく使用されます。その結果、Ceph Monitorのクォラムチェックが成功します。

(BZ#1990733)

### アップグレードに失敗した場合に新たな Ceph Object Gateway インスタンスを追加する

**radosgw\_frontend\_port** オプションでは、複数の Ceph Object Gateway インスタンスを考慮せず、全インスタンスに対してポート **8080** を設定していました。今回のリリースでは、各 Ceph Object Gateway インスタンスの **radosgw\_frontend\_port** オプションが増加し、複数の Ceph Object Gateway インスタンスを使用できるようになりました。

(BZ#1859872)

### Ceph Ansible がソケットファイルを削除し、クラスタの再デプロイを有効化

以前のバージョンでは、\*.asok ファイルはパージ Playbook の完了時に残っており、クラスタの再デプロイ時にエラーが発生していました。今回の更新により、**ceph-ansible** は存在する可能性のあるソケットファイルをすべて削除し、クラスタを安全に再デプロイできるようになりました。

(BZ#1861755)

### コンテナ化された Red Hat Ceph Storage デプロイメントの tcmu-runner プロセスのログローテーションのサポートが追加される

以前のバージョンでは、iSCSI を使用する Red Hat Ceph Storage がコンテナにデプロイされると、tcmu-runner プロセスのログローテーションがなければ、コンテナのすべての領域が使用されていました。今回のリリースにより、ログローテーションのサポートが tcmu-runner プロセスに追加され、ログファイルが定期的にローテーションされるようになりました。これにより、使用される領域が少なくなります。

(BZ#1873915)

### FileStore OSD と BlueStore OSD が混在する OSD ノードの場合に FileStore から BlueStore への移行プロセスが失敗する可能性がある

以前のバージョンでは、Red Hat Ceph Storage バージョン 3.2 よりも前のバージョンを実行するデプロイメントで、**osd\_objectstore** が **group\_vars**、**host\_vars**、または **inventory** のいずれでも明示的に設定されない場合、そのデプロイメントには FileStore OSD が含まれました。FileStore は、Red Hat Ceph Storage 3.2 より前のデフォルトでした。

デプロイされたストレージクラスタを Red Hat Ceph Storage 3.2 にアップグレードした後に、既存の OSD ノードに追加された新規 OSD は新しいデフォルトとなった BlueStore バックエンドを使用します。これにより、同じノード上に FileStore および BlueStore の OSD が混在しました。特定のケースでは、FileStore OSD が BlueStore OSD とジャーナルまたは DB デバイスを共有する場合があります。そのような場合に、パーティションが **lvm batch** で渡せないことや、GPT ヘッダーがあることにより、すべての OSD が再デプロイされると **ceph-volume** エラーが発生します。

今回のリリースにより、FileStore および BlueStore 設定の組み合わせで OSD を移行する 2 つのオプションを使用できます。

- **filestore-to-bluestore.yml** Playbook の実行時に追加の変数 **force\_filestore\_to\_bluestore** を **true** に設定します。この設定により、すでに BlueStore を使用するものであっても、Playbook によりすべての OSD が自動的に移行されます。
- **force\_filestore\_to\_bluestore** (デフォルトでは **false**) を設定せずに、**filestore-to-bluestore.yml** Playbook を実行します。これにより、Playbook は FileStore OSD と BlueStore OSD が混在するノードでの移行を自動的にスキップします。これは FileStore OSD のみを持つ



ノードを移行します。Playbook の実行の最後に、スキップされたノードを表示するレポートが表示されます。

OSD を移行する最適な方法を判断するために、Red Hat Ceph Storage 3 から 4 にアップグレードする前に、手動でスキップした各ノードを手動で確認します。

(BZ#1875777)

### 特殊文字は Docker レジストリーのパスワードで設定可能

以前のバージョンでは、Docker レジストリーのパスワードで設定された特殊文字は正しく処理されませんでした。今回のリリースにより、特殊文字が Docker レジストリーパスワードで設定されている場合に Ansible Playbook は失敗しません。特殊文字を Docker レジストリーのパスワードで使用でき、Ansible Playbook が予想通りに機能するようになりました。

(BZ#1880252)

### Ansible モジュール `ceph-volume` が、論理ボリュームおよびボリュームグループの正しい情報を報告

以前は、OSD 上で Red Hat Enterprise Linux 8 ベースのコンテナを使用する Red Hat Enterprise Linux 7 ホストで `ceph-volume lvm zap --destroy` コマンドを適用すると、そのホストの lvm キャッシュが更新されず、存在する論理ボリュームおよびボリュームグループが依然として報告されていました。今回のリリースにより、Ansible モジュール `ceph_volume` により、ホスト上のコマンドがトリガーされ、lvm キャッシュが更新され、論理ボリュームおよびボリュームグループの正しい情報を報告するようになりました。

(BZ#1886534)

### journalctl コマンドを使用して Ceph コンテナログを表示可能

以前のリリースでは、Podman がデタッチモードおよび systemd タイプのフォークでコンテナを実行するときにデフォルトのログドライバーとして Kubernetes ファイルを使用したため、Ceph コンテナログはジャーナルに存在しませんでした。今回のリリースで、Ceph コンテナは journald ログドライバーで設定され、ログは `journalctl` コマンドを使用して利用できるようになりました。

(BZ#1890439)

### Ceph Ansible がファイルおよびディレクトリーの所有権の値を `nobody:nobody` に設定する

以前のバージョンでは、Ceph Ansible はファイルとディレクトリーの所有権の値を `root:root` に設定しました。これにより、`alertmanager` サービスおよび `prometheus` サービスのパーミッションの問題が生まれました。

今回のリリースにより、Ceph Ansible が所有権を `nobody:nobody` に設定するようになりました。これにより、パーミッションの問題がなくなります。

(BZ#1901543)

## 6.2. COCKPIT CEPH インストーラー

### Cockpit を使用した Ceph インストールが失敗しなくなる

以前は、ホスト選択ページでホスト名として「`localhost`」を指定すると、Cockpit を介した Ceph インストールでは「`Systemd must be present`」エラーが表示され、失敗していました。このリリースでは、ホストページの UI に適切なエラーメッセージが表示され、ユーザーが「`localhost`」をホスト名として引き続き使用することを拒否します。

(BZ#1828246)

### Cockpit インストーラーが `rgws.yml` にバケット数の正しくない数を生成する

以前のバージョンでは、Cockpit インストーラーを使用して `rgws.yml` ファイルを生成すると、`defaults.rgw.buckets.data.pgnum` および `rgw_override_bucket_index_max_shards` に誤った値が生成されていました。

今回のリリースにより、Cockpit インストーラーが `rgws.yml` に正しい値を作成するようになりました。

(BZ#1855148)

## 6.3. CEPH MANAGEMENT DASHBOARD

### prometheus クエリーは、ダッシュボードで CPU 使用率についてリアルタイムメトリクスを報告するように修正される

以前のバージョンでは、Red Hat Ceph Storage ダッシュボードのホスト画面では、不正確な prometheus クエリーによりクラスター内のノードの CPU 使用率に関する誤ったデータが表示されました。今回のリリースにより、prometheus クエリーが修正され、CPU 使用率のデータはほぼリアルタイムのメトリクスになりました。

(BZ#1830375)

### Red Hat Ceph Storage ダッシュボードに埋め込まれた Grafana ダッシュボードからの将来のデータの使用が予想されるオプションが削除される

以前のバージョンでは、パフォーマンスグラフ全体で「Historical Data」からオプション「this week」を選択すると、将来のデータを使用する可能性のある Grafana メトリクスの履歴データオプションが少ないため、メトリクスが表示されませんでした。今回のリリースにより、将来のデータが使用される予定の履歴データオプションが削除され、メトリクスが予想通りに表示されるようになりました。

(BZ#1868638)

### CherryPy がバージョンをヘッダーおよびエラーページに表示しなくなる

以前のバージョンでは、cherryPy はヘッダーとエラーページにそのバージョンを表示しました。この情報を公開すると、潜在的なセキュリティの脆弱性が発生しました。今回のリリースにより、ダッシュボードおよび Prometheus サーバーの両方のヘッダーに、cherryPy バージョンではなく Ceph-Dashboard および Ceph-Prometheus が表示されるようになりました。この変更により、セキュリティの脆弱性がなくなります。

(BZ#1927719)

## 6.4. CEPH ファイルシステム

### Ceph fs status コマンドが `AttributeError` 例外で失敗しなくなる

以前のリリースでは、Ceph `fs status` コマンドが `AttributeError` の例外で失敗していました。これにより、再結合中のメタデータが不適切に処理されました。今回のリリースで、`NoneType` がメタデータオブジェクトタイプであれば、Ceph `fs status` コマンドが不明なオブジェクトとして予想されるステータスを返すようになりました。

(BZ#1884463)

## 6.5. CEPH MANAGER プラグイン

### Ceph Manager のアラートモジュールは、健全性警告のない状態で想定通りに機能

以前のリリースでは、Ceph Manager はデフォルト値の `smtp_ssl=true` を引き続き使用するため、デフォルト値のモジュール固有の設定オプションの一部を上書きしているため、**ALERTS SMTP\_ERROR unable to send alert email** エラーが発生するモジュールの障害が発生していました。今回のリリースで、Ceph Manager のデフォルト値処理が修正され、アラートモジュールが健全性警告がなく予想通りに機能するようになりました。

(BZ#1878145)

### restful モジュール API エンドポイントがアクセス可能に

以前は、restful モジュールは、Python 3 で利用できなくなった `dict.iteritems` を使用していました。そのため、restful モジュールの API エンドポイントの多くにアクセスできませんでした。今回のリリースにより、restful モジュールが `dic.items` を使用するようになり更新され、API エンドポイントにアクセスできるようになりました。

(BZ#1897995)

### IPv6 アドレスは ceph-mgr Prometheus エクスポートで適切に処理されない

以前のバージョンでは、ceph-mgr Prometheus エクスポートの出力に含まれる一部のメタデータメトリクスには、不完全または不正な IPv6 アドレスが表示されました。今回のリリースにより、エクスポートは IPv6 アドレスを正しく処理するようになりました。

(BZ#1929064)

## 6.6. CEPH VOLUME ユーティリティ

### ceph-volume コマンドを実行すると、デバッグ出力を /var/log/ceph/ceph-volume.log に送信されました。

以前のリリースでは、`--log-level` オプションに設定したレベルにかかわらず、**ceph-volume** コマンドを実行すると、デバッグレベルの出力を `/var/log/ceph/ceph-volume.log` に常に送信していました。今回のリリースにより、**ceph-volume** コマンドを実行すると、`--log-level` オプションで指定するレベルで出力が送信されます。

(BZ#1867717)

### ceph volume lvm batch は SSD デバイスを正しく報告し、正しい設定をデプロイ

以前は、**ceph-volume lvm batch** コマンドにより、`udev` で競合状態が発生し、SSD デバイスが HDD として誤って報告され、予期せぬデプロイメントが設定されていました。今回のリリースのコード更新により、**ceph-volume lvm batch** コマンドは SSD を正しく報告し、クラスタの予想される設定をデプロイします。

(BZ#1878500)

### 新規 SSD OSD のストレージクラスターへの追加後にスタックの更新が失敗する

以前のバージョンでは、スタックの更新を使用して新規 SSD OSD をストレージクラスターの同じボリュームグループ (VG) に追加すると、スタックの更新が失敗しました。これは、スタックの更新が、同じ LV に属するのではなく異なる論理ボリューム (LV) に属するものとして新規 SSD を誤って表示したために生じました。今回のリリースにより、スタック更新ビューが OSD を同じ LV に属するものとして新たに追加し、これにより失敗しなくなりました。

(BZ#1892441)

## 6.7. CEPH OBJECT GATEWAY

アップグレードされたOSDに対するライフサイクルの設定アクションおよびライフサイクルの削除アクションが期待どおりに機能するようになる

以前は、Red Hat Ceph Storage 3 から Red Hat Ceph Storage 4.2z2 へのアップグレード中に、アップグレードされた OSD に対してレガシーライフサイクルポリシーをインストールすると、set-lifecycle アクションは成功するものの、構造デコーディングエラーが発生しました。このリリースでは、バケットのライフサイクルステータスエントリをデコードするために必要な変更が修正され、デーモンのアップグレードが期待どおりに機能します。

(BZ#1993365)

**--reset-stats オプションが、多数のバケットを持つユーザーグループのバケットを更新**

以前のバージョンでは、radosgw-admin ユーザーの **--reset-stats** オプションは、ユーザーが所有するすべてのバケットの統計を同時に更新しました。非常に多くのバケットを持つユーザーの場合、更新を行うのに必要な時間が、関連付けられた RADOS 操作の長さを上回る場合があります。これにより、Ceph が OSD をダウン状態にあるものとしてマークする可能性があり、これにより OSD がフラッピングする可能性があります。

今回のリリースにより、**--reset-stats** オプションが 1000 バケットのグループの統計を更新するようになりました。これにより、多数のバケットを OSD のフラッピングを発生させずに更新できるようになります。

(BZ#1859257)

**gc エントリーがシステムからパージされると GC perf カウンターがインクリメントされる**

以前のバージョンでは、gc エントリーがシステムからパージされると、gc perf カウンターがインクリメントしませんでした。このコードの更新により、gc perf counter の正しい値を、システムから削除された gc エントリーの数に従って確認できます。

(BZ#1882484)

**最後の GC オブジェクトのエントリーの一覧がループを入力しない**

以前のバージョンでは、マーカーが最後の GC オブジェクトに対して毎回リセットされるため、最後の GC オブジェクトのエントリーの一覧がループに入りました。今回のリリースにより、切り捨てられたフラグが更新され、マーカーをリセットされることなく、一覧が予想通りに機能するようになりました。

(BZ#1884023)

**Ceph Object Gateway はバケット作成プロセスでバケットキャッシュ情報を同期**

以前のバージョンでは、Ceph Object Gateway はバケットの作成時にバケット情報のキャッシュを同期させませんでした。これにより、ユーザーがある RGW から存在しないバケットへのアクセスを試み、別の Ceph Object Gateway からバケットを作成し、最初の Ceph Object Gateway からバケットにアクセスして、バケットは存在するにもかかわらず存在しないことを示す 404 エラーが生じるという状況が生じました。今回の更新により、RGW はバケットの作成プロセスでキャッシュを同期し、各 Ceph Object Gateway がバケットにアクセスできるようになりました。

(BZ#1901973)

**KafkaConnect は Kafka トピックから RGW S3 バケットにオブジェクトを送信**

以前のバージョンでは、chunked-encoding オブジェクト署名が正しく計算されないため、Kafka トピックから RGW S3 バケットへのオブジェクトの送信が失敗しました。

これにより、RADOS Gateway ログに以下のエラーが発生しました。

### 20 AWSv4ComplMulti: ERROR: chunk signature mismatch

今回のリリースにより、チャンク化したエンコードオブジェクト署名が正しく計算され、KafkaConnect がオブジェクトを正常に送信できるようになりました。

([BZ#1912538](#))

## 6.8. マルチサイトの CEPH OBJECT GATEWAY

### バケットでの同期が無効になった後、バケットインデックスがエントリーを収集しない

以前のバージョンでは、マルチサイトの同期が無効になっているバケットで **radosgw-admin bucket check --fix ...** 変数を使用すると、同期が無効であることを示す誤ったフラグが設定されていました。データは、使用されない、またはトリム化されないバケットインデックスログに追加され、時間の経過とともにより多くのストレージを消費します。今回のリリースにより、**radosgw-admin bucket check --fix ...** コマンドの実行時に同期フラグが正しくコピーされるようになりました。バケットの同期が無効になると、バケットのインデックスログはエントリーを収集しません。

([BZ#1894702](#))

## 6.9. RADOS

### Progress モジュールが無限に停止しなくなる

以前のバージョンでは、Ceph ステータスの進捗イベントが無限に続くことがありました。これは、Progress モジュールが PG の状態を早期にチェックし、OSDMap のエポックと同期していないことが原因でした。今回のリリースにより、進捗イベントが想定どおりにポップアップされるようになりました。

([BZ#1826224](#))

### ProgressモジュールによるCeph Monitorのクラッシュ

バックフィルおよびリカバリ操作中に、progressモジュールが負の進行状況イベントを生成する可能性があります。大規模なストレージクラスタでは、ネガティブな進捗イベントが多すぎると、Ceph Monitorノードのメモリ割り当てが大きくなり、Ceph Monitorのカスケードクラッシュが発生します。このリリースでは、進捗イベントが負にならないようなコードになっています。その結果、progressモジュールがCeph Monitorsのクラッシュの原因になることはありません。

([BZ#1997838](#))

### Ceph Monitor が転送された OSD 障害レポートを詳細に処理する

以前のバージョンでは、Ceph Monitor は低速な操作を管理者およびロギングシステムに誤って報告していました。今回のリリースにより、転送された OSD 障害レポートを Ceph Monitor が詳細に処理するようになったため、不正確で低速な操作がはるかに少なくなります。

([BZ#1866257](#))

### モニターが osdmap を適切にトリミング

以前のバージョンでは、モニターが osdmap をトリミングする際に「in」と「out」の両方の OSD を考



慮したため、「out」OSDによって使用されていた場合でも、古い osdmap をトリミングできませんでした。今回のリリースにより、「out」の OSD は考慮されず、osdsmap が適切にトリミングされるようになりました。

(BZ#1875628)

### BlueStore および FileStore OSD が混合クラスターで同じ順序でオブジェクトを一覧表示

以前のバージョンでは、BlueStore および FileStore OSD の両方を持つクラスターでは、バックエンド内のソートメカニズムの不整合が原因で、ディープスクラブおよびバックフィルがオブジェクト不足を報告していました。今回の更新で、機能フラグ **OSD\_FIXED\_COLLECTION\_LIST** が追加され、BlueStore リストオブジェクトの **collection\_list** メソッドが FileStore と同じ順序で追加されるようになりました。

(BZ#1880188)

### 正しくないパーミッションでログファイルが作成される

以前のバージョンでは、コードの追加により、関連する関数が呼び出される順序が変更されました。これにより、一部のデーモンが正しくないパーミッションでログファイルを作成しました。今回のリリースにより、関数が正しい順序で呼び出され、デーモンは適切なパーミッションでログファイルを作成できるようになりました。

(BZ#1884469)

### bluafs\_buffered\_io を有効にするとパフォーマンスが低下する

以前のリリースでは、オプション **bluafs\_buffered\_io** が無効になり、特定のシナリオで RocksDB および OMAP の対話が遅くなり、オプション **bluafs\_buffered\_io** オプションが **True** に設定され、これによりパフォーマンスが低下します。

(BZ#1930264)

### onodes によるメモリの増加が制御されるようになる

以前のリリースでは、オプション **bluestore\_cache\_trim\_max\_skip\_pinned** のデフォルト値が 64 で、大規模なクラスターには非常に低いものでした。このオプションは、onodes のトリミング速度を制御し、現在のデフォルト値を使用しているため、onodes が蓄積してメモリーが増大する可能性があります。今回のリリースにより、**bluestore\_cache\_trim\_max\_skip\_pinned** のデフォルト値は 1000 で、メモリーの増加が制御されるようになりました。

(BZ#1947215)

## 6.10. RADOS ブロックデバイス (RBD)

### クライアント側の QoS スロットリングが有効な場合に librbd を使用するアプリケーションが機能する

以前バージョンでは、クライアント側の QoS スロットリングが librbd で有効になると、データパスがロックで適切に保護されないためにクラッシュする可能性があります。今回のリリースにより、欠落していたロックが追加され、クライアント側の QoS スロットリングが有効な場合に、IO に librbd を使用するアプリケーションが想定どおりに機能するようになりました。

(BZ#1878268)

## 6.11. NFS GANESHA

DEADDIR のおぼえのファイルレイアウトが結果を返す

**REaddir のサブディレクトリファイルレイアウトが和名で返す**

以前のバージョンでは、一部のファイルレイアウトによって REaddir のループが生じ、結果が返されなくなりました。今回の更新により、REaddir が予想通りに機能し、結果が正しく返されるようになりました。

([BZ#1845501](#))

## 第7章 既知の問題

本セクションでは、Red Hat Ceph Storage の今回リリースで見つかった既知の問題を説明します。

### 7.1. CEPH ANSIBLE ユーティリティー

CephFS 関連のプールでは、配置グループ Autoscaler のデプロイが想定どおりに機能しない

この問題を回避するには、Playbook の実行後に、CephFS 関連のプールで配置グループ Autoscaler を手動で有効にできます。

([BZ#1836431](#))

systemd ユニットテンプレートではパラメーター `Restart=always` を適用するため、Ceph OSD が `osd_max_markdown_count` パラメーターを使用することができない

OSD デーモンの systemd ユニットテンプレートは、パラメーター `Restart=always` を強制するため、`osd_max_markdown_count` パラメーターを使用しないと、サービスを再起動することになります。この問題を回避するには、`ceph_osd_systemd_overrides` 変数を使用して OSD systemd テンプレートの `Restart=` パラメーターを上書きします。つまり、次のようになります。

```
[osds]
osd0 ceph_osd_systemd_overrides="{ 'Service': { 'Restart': 'no' } }
```

([BZ#1860739](#))

Playbook の `filestore-to-bluestore` が「`osd_auto_discovery`」シナリオに対応しない

`osd_auto_recovery` のシナリオをベースとした Red Hat Ceph Storage 4 デプロイメントでは、BlueStore の移行を容易にするために Playbook `filestore-to-bluestore` を使用することができません。

この問題を回避するには、Playbook の `shrink-osd` を使用して、`osd_objectstore: bluestore` で縮小された OSD を再デプロイします。

([BZ#1881523](#))

アップグレードプロセスでは、`ceph-crash` コンテナデーモンを自動的に停止しない

アップグレードプロセスでは、`ceph-crash` ロールへの呼び出しを実行しますが、呼び出しは `ceph-crash` サービスのみを起動します。アップグレードプロセス中に `ceph-crash` コンテナデーモンが稼働したままであれば、それらはアップグレードの完了時に再起動しません。

この問題を回避するには、アップグレード後に `ceph-crash` コンテナを手動で再起動します。

([BZ#1943471](#))

### 7.2. CEPH VOLUME ユーティリティー

ユーザーが `osd.yml` または `site.yml` Playbook を実行すると、`ceph-ansible` は新しいデバイスに OSD を作成しない

ユーザーが一連の db デバイス (`-db-devices` または wal デバイス `--wal-devices`) を明示的に渡した場合に、`ceph-volume lvm` で使用できない場合は、フィルターで除外され、結果が予想とは異なります。`ceph-volume lvm batch` バッチの現在の実装では、渡された db または wal デバイスの1つが使用



できない場合は、非対話型モードで新しい OSD を追加して、予期される OSD トポロジーを防ぐことができません。この `ceph-volume` の制限により、デバイスおよび `dedicated_devices` のバッチシナリオで `ceph-ansible` が新規 OSD を追加できません。

([BZ#1896803](#))

### 7.3. マルチサイトの CEPH OBJECT GATEWAY

#### オブジェクトが Ceph Object Gateway マルチサイトセットと同期されない

一部のオブジェクトは同期に失敗し、ユーザーが Ceph Object Gateway のマルチサイトセットで **`radosgw-admin sync status`** コマンドを実行する場合にステータスの不一致が生じる場合があります。

現在、この問題に対する回避策はありません。

([BZ#1905369](#))

## 第8章 ソース

更新された Red Hat Ceph Storage ソースコードパッケージは、以下の場所から入手できます。

- Red Hat Enterprise Linux 7 の場合:  
<http://ftp.redhat.com/redhat/linux/enterprise/7Server/en/RHCEPH/SRPMS/>
- Red Hat Enterprise Linux 8 の場合:  
<http://ftp.redhat.com/redhat/linux/enterprise/8Base/en/RHCEPH/SRPMS/>