

# Kluczowe komponenty odzyskiwania po awarii w AWS RDS

2020 SOFTPROM [softprom.com](https://softprom.com) | [info@softprom.com](mailto:info@softprom.com) Amazon Web Services Partner  
Network ADVANCED CONSULTING PARTNER

---

## O usłudze Amazon RDS

Amazon RDS (Relational Database Service) to zarządzana przez AWS usługa, która pozwala na łatwą konfigurację, obsługę i skalowanie relacyjnych baz danych w chmurze. Amazon RDS oferuje do wyboru sześć popularnych silników bazodanowych, w tym: Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database oraz SQL Server. Usługa zapewnia efektywne kosztowo i skalowalne wykorzystanie zasobów, automatyzując jednocześnie czasochłonne zadania administracyjne, takie jak alokacja sprzętu, bazy danych, instalowanie poprawek (patching) oraz odzyskiwanie po awarii (DR).

W niniejszym dokumencie przedstawiono trzy funkcje Amazon RDS, które wspierają odzyskiwanie po awarii: automatyczne kopie zapasowe, ręczne kopie zapasowe (migawki) oraz repliki odczytu.

---

## Dlaczego należy posiadać plan odzyskiwania po awarii?

W głównym środowisku produkcyjnym ważne jest zachowanie środków ostrożności, aby móc przywrócić działanie systemu w przypadku nieoczekiwanej awarii. Amazon RDS zapewnia wysoką niezawodność w konfiguracji wielostrefowej (Multi-AZ), ale nie może zagwarantować ochrony przed wszystkimi możliwymi czynnikami, takimi jak katastrofy naturalne, działania cyberprzestępców czy logiczne uszkodzenie bazy danych. W celu zapewnienia ciągłości działania biznesu (Business Continuity) kluczowe jest opracowanie i przetestowanie planu odzyskiwania po awarii.

## Zrozumienie RTO i RPO

Czas przywracania (RTO) oraz punkt przywracania (RPO) to dwa kluczowe parametry, które należy wziąć pod uwagę przy opracowywaniu planu DR:

- **RTO (Recovery Time Objective):** Wskazuje, ile czasu zajmie przywrócenie systemu do działania po wystąpieniu awarii.
- **RPO (Recovery Point Objective):** Wyrażane w czasie, wskazuje maksymalny okres, za który dopuszczalna jest utrata danych w przypadku awarii. Na przykład RPO wynoszące 1 godzinę oznacza, że w przypadku awarii nastąpi utrata danych z ostatniej godziny pracy.

Różne typy tworzenia kopii zapasowych w Amazon RDS mają zastosowanie przy odpowiednich wartościach RTO/RPO oraz zróżnicowanych kosztach:

Funkcja (Feature)	RTO	RPO	Koszt (Cost)	Zasięg (Scope)
<b>Automated backups</b>	Dobry (Good)	Lepszy (Better)	Niski (Low)	Jeden region (Single Region)
<b>Manual snapshots</b>	Lepszy (Better)	Dobry (Good)	Średni (Medium)	Międzyregionowy (Cross-Region)
<b>Read replicas</b>	Najlepszy (Best)	Najlepszy (Best)	Wysoki (High)	Międzyregionowy (Cross-Region)

Jak pokazano w tabeli, automatyczne kopie zapasowe mają zastosowanie tylko dla jednego regionu AWS, podczas gdy ręczne migawki danych (snapshots) oraz repliki odczytu mogą być stosowane w wielu regionach.

## Wskaźniki RPO/RTO dla różnych typów niezawodności w Amazon RDS

Funkcja (Feature)	RPO (przybliżone)	RTO (przybliżone)
<b>Multi-AZ dla wysokiej dostępności</b>	0	1 do 2 minut
<b>Odtworzenie z migawki (Snapshot restore)</b>	Godziny	< 1 godzina
<b>Odtworzenie do określonego momentu (PITR)</b>	Minuty	< 1 do kilku godzin
<b>Promowanie repliki odczytu (w regionie)</b>	Minuty	< 5 minut
<b>Promowanie repliki odczytu (między regionami)</b>	Minuty	< 5 minut

**Uwaga dotycząca Amazon Aurora:**

Amazon Aurora posiada inne funkcje odzyskiwania po awarii i wymaga osobnego rozpatrzenia, jednak wiele z opisanych poniżej podejść ma zastosowanie również do klastrów baz danych Aurora.

## Kopie zapasowe Amazon RDS

Kopie zapasowe są kluczowym komponentem planu odzyskiwania bazy danych po awarii. Amazon RDS obsługuje dwa różne typy kopii zapasowych: automatyczne tworzenie migawek danych (snapshots) oraz kopie tworzone ręcznie.

Rezerwowe kopiowanie Amazon RDS podlega następującym zasadom:

- Instancja bazy danych musi znajdować się w stanie **AKTYWNYM** (ACTIVE), aby możliwe było wykonanie kopii zapasowej.
- Jednoczesne automatyczne kopiowanie zapasowe i automatyczne tworzenie migawek w tym samym regionie dla tej samej instancji bazy danych jest niemożliwe.
- Migawki instancji bazy danych mają charakter przyrostowy (inkrementalny) — pierwsza migawka jest pełna, a kolejne rejestrują jedynie ostatnie zmiany danych.
- B przypadku konfiguracji bazy danych w kilku strefach dostępności (Multi-AZ), kopia zapasowa jest wykonywana na instancji rezerwowej (standby), aby zmniejszyć obciążenie instancji głównej.

## Automatyczne kopie zapasowe (Automated Backups)

Funkcja automatycznego kopiowania zapasowego w Amazon RDS jest włączona domyślnie. Amazon RDS tworzy migawkę wolumenu pamięci masowej instancji bazy danych, wykonując kopię zapasową całej instancji i wszystkich zawartych w niej baz danych. Każda kolejna kopia zapasowa ma charakter przyrostowy. Każda migawka zawiera wskaźniki do wszystkich bloków danych potrzebnych do jej odtworzenia. Usunięcie wcześniejszej migawki nie prowadzi do utraty danych, dopóki referencje do tych danych znajdują się w co najmniej jednej z późniejszych migawek.

Zalety automatycznego kopiowania zapasowego:

- Migawki są przechowywane w kubełku AWS S3, który jest automatycznie tworzony i zarządzany przez usługę AWS RDS.
- Całkowicie eliminuje się czynnik ludzki.

- Można wybrać dogodny harmonogram tworzenia kopii: codziennie, cotygodniowo lub comiesięcznie.
- Zmniejsza się wpływ czynników związanych zarówno z działalnością człowieka, jak i katastrofami naturalnymi (np. wirusy, błędy oprogramowania czy awarie zasilania).
- Do minimum sprowadzona zostaje utrata cennych danych.

### Harmonogram i czas przechowywania kopii

- **Harmonogram (Backup Window):** Jest to okno czasowe w ciągu tygodnia wykorzystywane do tworzenia automatycznych kopii. Czas jest wybierany losowo z 8-godzinnego bloku czasu dla każdego regionu AWS. Zdecydowanie zaleca się ustawienie czasu kopii zapasowej na godziny niskiego obciążenia, aby zapobiec nadmiernemu obciążeniu serwera.
- **Okres przechowywania (Backup Retention Period):** To przedział czasu, w którym można wykonać przywrócenie bazy danych do wymaganego momentu (Point-In-Time Restore - PITR). Wymagany okres przechowywania można ustawić podczas tworzenia instancji bazy danych i zmienić go w razie potrzeby później.

### Różnice między migawkami ręcznymi a automatycznymi:

- Ręcznie można utworzyć maksymalnie 100 migawek na region. Przy automatycznych kopiach nie ma takich ograniczeń.
- Okres przechowywania kopii zapasowych nie dotyczy migawek wykonanych ręcznie.
- Migawki ręczne nie są usuwane automatycznie; mogą zostać usunięte wyłącznie przez użytkownika.

### Przywracanie do określonego momentu (PITR)

Przywracanie do określonego momentu (Point-In-Time Restore) to proces odtwarzania bazy danych do stanu, w jakim znajdowała się w wymaganym dniu i o określonej godzinie. Gdy automatyczne kopie są włączone, Amazon RDS tworzy codzienne migawki. Dodatkowo AWS RDS zapisuje dzienniki transakcji (transaction logs) do Amazon S3 co 5 minut. Archiwizacja dzienników transakcji to kluczowa część procesu DR i PITR. Podczas inicjowania PITR dzienniki transakcji są aplikowane do najbardziej odpowiedniej codziennej kopii zapasowej, aby przywrócić instancję do żądanego stanu.

### Прехводы автоматических копий после удаления базы

Przy usuwaniu instancji bazy danych istnieje możliwość zachowania automatycznych kopii zapasowych. Jest to przydatne, jeśli w przyszłości zajdzie potrzeba odtworzenia instancji. Zachowane kopie zawierają automatyczne migawki stanu, dzienniki transakcji oraz właściwości instancji (takie jak rozmiar pamięci i parametry), które są niezbędne do

odtworzenia. Zarządzanie nimi (odtworzenie lub usunięcie) jest możliwe z poziomu Konsoli zarządzania AWS, API Amazon RDS oraz AWS CLI.

---

## Ręczne migawki bazy danych (Manual Snapshots)

Migawki bazy danych to ręczne (inicjowane przez użytkownika) pełne kopie zapasowe całej instancji bazy danych. Są one przechowywane w Amazon S3 i pozostają tam do momentu jawnego usunięcia przez użytkownika. Migawki te można kopiować i przysyłać do różnych regionów oraz na inne konta AWS.

Czas tworzenia migawki zależy od rozmiaru pamięci masowej instancji. Tworzenie migawki w instancji działającej w jednej strefie dostępności (Single-AZ) prowadzi do krótkotrwałego zawieszenia operacji wejścia/wyjścia (I/O), trwającego od kilku sekund do kilku minut. Przy architekturze wielostrefowej (Multi-AZ) zawieszenie I/O nie występuje, ponieważ kopia jest wykonywana z instancji rezerwowej.

## Współdzielenie ręcznych migawek między kontami AWS

Amazon RDS umożliwia udostępnianie ręcznych migawek bazy danych innym kontom AWS. Współdzielenie migawek z innymi kontami pozwala skutecznie zabezpieczyć krytyczne dane przed działaniem intruzów na koncie produkcyjnym.

Zasady udostępniania migawek:

- Ręczne migawki można udostępniać maksymalnie 20 kontom AWS.
- Automatyczne migawki nie mogą być bezpośrednio udostępniane innym kontom. Należy najpierw utworzyć kopię migawki automatycznej (co przekształci ją w wersję ręczną), a dopiero potem ją udostępnić.
- Ręczne migawki instancji bazy danych, które używają niestandardowych grup parametrów z funkcjami takimi jak Transparent Data Encryption (TDE) lub niestandardowa strefa czasowa, nie mogą być przesyłane na inne konto.
- Migawki korzystające z domyślnych kluczy szyfrowania Amazon RDS nie mogą być udostępniane bezpośrednio. Należy najpierw skopiować migawkę, wybierając własny (niestandardowy) klucz szyfrowania, a następnie udostępnić ten klucz oraz skopiowaną migawkę drugiemu kontu AWS.

## Odtwarzanie bazy danych z migawki

W przypadku awarii można utworzyć nową instancję bazy danych poprzez odtworzenie jej z migawki. Należy wtedy wybrać żadaną migawkę oraz wskazać nazwę dla nowej instancji.

Ważne aspekty procesu odtwarzania:

- Nie można przywrócić danych z migawki do istniejącej już instancji bazy danych. Podczas odtwarzania zawsze tworzona jest nowa instancja. Jeśli chcesz użyć starej nazwy instancji, musisz najpierw usunąć lub zmienić nazwę istniejącej instancji.
- Możliwe jest odtworzenie bazy danych z migawki na instancję o innym typie pamięci masowej niż oryginalna, jednak proces ten zajmie więcej czasu.
- Nie można odtworzyć bazy danych bezpośrednio z zaszyfrowanej migawki bazy danych. Najpierw należy wykonać kopię zaszyfrowanej migawki bazy danych, a następnie przeprowadzić proces odtwarzania z tej kopii.
- Zaleca się zapisanie i zachowanie grup parametrów (parameter groups) dla tworzonych migawek, co pozwoli na poprawne odtworzenie instancji z właściwymi ustawieniami. Modyfikacja grupy ustawień jest również możliwa bezpośrednio podczas odtwarzania.

#### *Odtwarzanie w innych regionach AWS:*

*Przy odtwarzaniu migawki w innym regionie AWS, należy najpierw skopiować migawkę do docelowego regionu, a dopiero stamtąd przeprowadzić proces przywracania danych do nowej instancji bazy danych.*

## Integracja z AWS Backup

Migawki stanu baz danych w Amazon RDS mogą być zintegrowane z usługą *AWS Backup*. Jest to w pełni zarządzana usługa, która pozwala na scentralizowanie i automatyzację tworzenia kopii zapasowych danych w usługach AWS w chmurze oraz lokalnie. Usługa pozwala na centralną konfigurację polityk oraz monitorowanie aktywności backupu dla działających zasobów.

---

## Repliki odczytu (Read Replicas)

Amazon RDS dla systemów MariaDB, MySQL, PostgreSQL oraz Oracle wspiera możliwość tworzenia replik odczytu źródłowej bazy danych. Przy tworzeniu repliki odczytu Amazon RDS najpierw wykonuje migawkę instancji źródłowej, a następnie tworzy nową instancję przeznaczoną wyłącznie do odczytu.

W celu aktualizacji repliki odczytu po każdej zmianie w bazie źródłowej, Amazon RDS wykorzystuje metodę **asynchronicznej replikacji**. Aplikacje mogą łączyć się z repliką odczytu w taki sam sposób, jak do każdej innej instancji bazy danych, przekierowując tam zapytania typu READ. Dla każdej instancji bazy źródłowej można utworzyć do pięciu replik odczytu.

Oprócz odciążania bazy głównej, repliki odczytu mogą być wykorzystywane do realizacji planu odzyskiwania po awarii (DR). W przypadku awarii serwera źródłowego, replika odczytu może zostać przekształcona (wypromowana) w nowy samodzielny serwer bazy danych.

Repliki odczytu mogą być tworzone w innym regionie AWS niż baza główna. Zastosowanie repliki międzyregionowej pozwala na przywrócenie działania aplikacji biznesowej, jeśli wystąpią problemy z dostępem do całego regionu podstawowego.

### Ważne parametry replik odczytu pod kątem DR:

- **Opóźnienie replikacji (Replication Lag):** Wpływa bezpośrednio na kompletność odzyskiwania danych (RPO). Opóźnienie to zależy od szybkości transferu danych między regionami AWS oraz od ilości przesyłanych danych.
- **Czas odzyskiwania (RTO):** Ponieważ repliki odczytu zawierają już aktualne dane z głównej instancji bazy danych, czas potrzebny na odzyskanie sprawności systemu (RTO) jest minimalny.
- **Koszty:** Wykorzystanie replik odczytu w planie DR jest opcją znacznie droższą niż korzystanie z automatycznych kopii zapasowych lub migawek bazy danych.

### Przeniesienie głównej instancji bazy danych na replikę odczytu (Promoting a Read Replica)

W przeciwieństwie do konfiguracji wielostrefowej (Multi-AZ), proces wypromowania repliki odczytu do rangi serwera głównego **nie następuje automatycznie**.

Przy korzystaniu z repliki w innym regionie AWS należy upewnić się, czy przeniesienie głównej bazy danych do innego regionu jest uzasadnione operacyjnie. Podczas przesyłania danych między regionami mogą występować opóźnienia, co może wiązać się ze skomplikowaną rekonfiguracją działającej aplikacji. Po przeniesieniu głównej bazy na replikę odczytu w innym regionie, powrót do regionu pierwotnego (failback) również wymaga ręcznego utworzenia nowej repliki odczytu i powtórzenia procedury przeniesienia.

---

## Testowanie planu odzyskiwania po awarii

Plan odzyskiwania po awarii jest przydatny tylko wtedy, gdy jest okresowo testowany, co udowadnia jego skuteczność. Testowanie pozwala na wczesne wykrycie i wyeliminowanie potencjalnych problemów. Kompleksowy plan DR obejmuje nie tylko same bazy danych, ale całą infrastrukturę aplikacji. Pełny test planu DR może zająć znaczną ilość czasu i

zasobów, ale jest jedynym sposobem na upewnienie się o jego rzeczywistej efektywności w warunkach bojowych.