

Redundantny system plików

Kacper Pieniążek

December 4, 2019

Cel pracy

- System plików zapewniający ochronę danych nawet w przypadku ich uszkodzenia.
- W oparciu o interfejs FUSE.

Założenia pracy

Założenia pracy

- Wykrywanie niezgodności danych
 - ▶ Sumy kontrolne
 - ▶ Kody korekcyjne

Założenia pracy

- Wykrywanie niezgodności danych
 - ▶ Sumy kontrolne
 - ▶ Kody korekcyjne
- Odzyskiwanie danych
 - ▶ Duplikacja danych
 - ▶ Kody korekcyjne

Założenia pracy

- Wykrywanie niezgodności danych
 - ▶ Sumy kontrolne
 - ▶ Kody korekcyjne
- Odzyskiwanie danych
 - ▶ Duplikacja danych
 - ▶ Kody korekcyjne
- Obsługa wybranych funkcjonalności systemu plików

Założenia pracy

- Wykrywanie niezgodności danych
 - ▶ Sumy kontrolne
 - ▶ Kody korekcyjne
- Odzyskiwanie danych
 - ▶ Duplikacja danych
 - ▶ Kody korekcyjne
- Obsługa wybranych funkcjonalności systemu plików
- Wygodne użytkowanie
 - ▶ Podział i duplikacja danych jak najbardziej transparentna dla użytkownika
 - ▶ Rozwiązywanie konfliktów w przypadku niezgodności danych
 - ▶ Dostosowanie systemu do własnych potrzeb

Analiza problemu

- Jak zapewnić redundancję?

Analiza problemu

- Jak zapewnić redundancję?
- Jak zapewnić spójność danych?
 - ▶ W przypadku braku synchronizacji między replikami, wybór poprawnych danych

Analiza problemu

- Jak zapewnić redundancję?
- Jak zapewnić spójność danych?
 - ▶ W przypadku braku synchronizacji między replikami, wybór poprawnych danych
- Jak wykrywać rozbieżność danych?
 - ▶ Podczas operowania na uszkodzonych danych, błędny odczyt, kody korekcyjne, sumy kontrolne

Analiza problemu

- Jak zapewnić redundancję?
- Jak zapewnić spójność danych?
 - ▶ W przypadku braku synchronizacji między replikami, wybór poprawnych danych
- Jak wykrywać rozbieżność danych?
 - ▶ Podczas operowania na uszkodzonych danych, błędny odczyt, kody korekcyjne, sumy kontrolne
- Jak naprawiać rozbieżność danych?

Rozwiązanie

Definicja

Replika - jeden z podsystemów zawierający kopię chronionych danych

Definicja

Replika - jeden z podsystemów zawierający kopię chronionych danych

- Cały system może być podzielony na warstwy współpracujące ze sobą.
 - ▶ Każdy rodzaj repliki to nowa warstwa

Definicja

Replika - jeden z podsystemów zawierający kopię chronionych danych

- Cały system może być podzielony na warstwy współpracujące ze sobą.
 - ▶ Każdy rodzaj repliki to nowa warstwa
- Podział na warstwy umożliwia rozwiązania typu RAID

Definicja

Replika duplikująca kopiuje chronione dane bez dodatkowych informacji

Rozwiązanie

Definicja

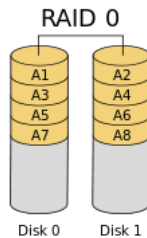
Replika duplikująca kopiuje chronione dane bez dodatkowych informacji

Definicja

Replika blokowa rozdziela chronione dane na bloki zależne od wybranej implementacji; Może zawierać dodatkowe informacje potrzebne w przypadku uszkodzenia

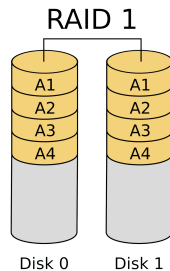
Replika typu 0

Lorem ipsum



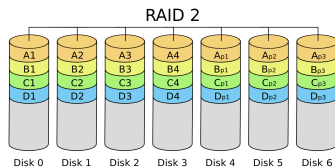
Replika typu 1

Lorem ipsum



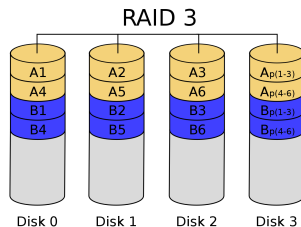
Replika typu 2

Lorem ipsum



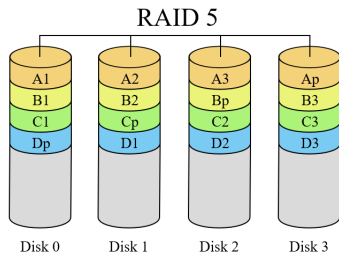
Replika typu 3

Lorem ipsum



Replika typu 4

Lorem ipsum



Przykłady

Przykład

Dysk, na którym zamontowana była jedna z replik duplikujących dane został odłączony.

Przykłady

Przykład

Dysk, na którym zamontowana była jedna z replik duplikujących dane został odłączony.

Przykład

Wystąpił błąd podczas zapisu do jednej z replik blokowych.

Przykłady

Przykład

Dysk, na którym zamontowana była jedna z replik duplikujących dane został odłączony.

Przykład

Wystąpił błąd podczas zapisu do jednej z replik blokowych.

Przykład

Obrócone bity na danych w jednej z replik.

Konfiguracja oraz komunikacja