

Instalacja programowego RAID w Arch Linuksie

Tomasz Jakub Jędrzejewski

12 grudnia 2006

RAID (ang. *Redundant Array of Independent Disks*) to specjalny tryb pracy grupy dysków twardych, w których współpracują one ze sobą, aby uzyskać korzyści nieosiągalne przy wykorzystaniu jednego HDD. Korzyści te mogą mieć wieloraki charakter, np. poprawienie bezpieczeństwa danych w razie awarii jednego z urządzeń, zwiększenie prędkości odczytu albo i przestrzeni dyskowej dostępnej jako jedna całość. Z tego powodu RAID jest szczególnie polecany na serwerach, które pracują non-stop i zazwyczaj młóćą ogromne ilości danych.

Istnieje kilka tzw. poziomów RAID numerowanych od 0 do 5. Każdy z nich niesie ze sobą nieco inne wady i korzyści, dlatego przed wyborem warto zapoznać się z ich opisem. Przykładowo, macierz RAID-0 powoduje, że miejsce na połączonych w nich dyskach widziane jest jako jedna całość. Dane przeplatane są między urządzeniami, co często daje efekt zwiększenia wydajności (jeśli dwa procesy czytają dane umieszczone na różnych dyskach, mogą to robić niezależnie od siebie). Niestety, awaria jednego z urządzeń oznacza uszkodzenie całej macierzy. Niejako odwrotnością RAID-0 jest RAID-1. Tutaj system dostaje ilość miejsca równą tylko pojemności *najmniejszego* urządzenia. Wszystkie twarde dyski są bowiem swoimi własnymi kopiami i przechowują dokładnie takie same informacje. Jeżeli nawet padnie jeden HDD, pozostałe mirrory pracują dalej. W międzyczasie admin może wymienić zepsuty dysk na nowy, a system odbuduje jego zawartość. Zwiększona zostaje także prędkość odczytu, lecz konieczność zapisania tych samych danych na wielu urządzeniach zmniejsza wydajność zapisu.

Posiadaczem macierzy RAID można zostać na kilka sposobów. Możemy wyłożyć pieniądze na odpowiedni kontroler HDD zajmujący się tym po stronie sprzętowej, lecz możemy wybrać również tańsze rozwiązanie: RAID programowy, gdzie zarządzanie macierzą należy do zadań systemu operacyjnego. W tym artykule pokażę, jak zainstalować dystrybucję Arch Linux 0.7.2 na RAID-zie programowym.

Ostrzeżenie: instalacja i konfiguracja RAID jest skomplikowana i jest więcej, niż prawdopodobne, że nie uda się za pierwszym razem. Dlatego najlepiej całą zabawę przeprowadzać na „czystym” komputerze lub przynajmniej zadbać o kopię zapasową wszystkich danych.

1 Przygotowania

Macierz RAID montuje się w zasadzie nie na całym urządzeniu, ale na pojedynczej partycji. Stąd też nie istnieją żadne przeszkody, aby część dysku pracowała np. jako RAID-5, a druga była elementem macierzy RAID-1. W naszym artykule ustawimy cztery partycje RAID-1 na dwóch 160-gigabajtowych dyskach

IDE, a następnie zainstalujemy na nich system Arch Linux. Podział przestrzeni będzie następujący:

1. / — 28 GB — główna partycja systemu
2. swap — 2 GB — partycja wymiany
3. /home — 100 GB — partycja przeznaczona na konta użytkowników
4. /var — 30 GB — Arch Linux przechowuje w katalogu /var całkiem sporo informacji, dlatego nie żałujemy dla niego przestrzeni.

Na system plików wybrałem ReiserFS, aczkolwiek warto zastanowić się, czy w Twojej konkretnej sytuacji także będzie on dobrym rozwiązaniem. Zakładam ponadto, iż jesteś obeznany nieco z Linuksem oraz procesem instalacji samego Arch Linuksa, ponieważ w niniejszym artykule pragnę skupić się wyłącznie na sprawach związanych z RAID.

2 Przebieg instalacji

Poniżej zaprezentowałem w punktach szybki przegląd kolejnych kroków instalacji.

1. Uruchomienie instalatora,
2. Partycjonowanie dysków twardych,
3. Utworzenie macierzy RAID,
4. Zamontowanie systemu plików,
5. Instalacja i konfiguracja Arch Linuksa,
6. Instalacja Gruba na pierwszym HDD,
7. Odmontowanie systemu plików i pierwsze uruchomienie,
8. Instalacja Gruba na drugim HDD,

2.1 Uruchomienie instalatora

Najwygodniejsze jest instalowanie Arch Linuksa z płyty CD. Ustawiamy w BIOS-ie start systemu z właśnie tego medium, wrzucamy płytkę do napędu i restartujemy komputer. Po włączeniu znajdującego się na płycie bootloadera naciskamy *Enter* (chyba że nasz sprzęt wymaga przekazania dodatkowych parametrów rozruchu) i czekamy na załadowanie jądra. Ujrzawszy znak zachęty, nie uruchamiamy graficznego instalatora, ponieważ na razie nie będzie on nam potrzebny.

2.2 Partycjonowanie dysków twardych

Pora na trudniejszą rzecz, czyli podział naszych dysków na partycje. Użyjemy programu `cmdisk` do utworzenia na nich identycznego zestawu partycji:

1. Partycja 1: 28 GB, typ FD, bootowalna
2. Partycja 2: 2 GB, typ FD
3. Partycja 3: 100 GB, typ FD
4. Partycja 4: 30 GB, typ FD

Najpierw uruchom:

```
# cmdisk /dev/hda
```

I utwórz każdą z partycji w następujący sposób:

1. Wybierz **New**
2. Wciśnij *Enter* do stworzenia partycji głównej (**Primary**)
3. Wpisz dla każdej z partycji rozmiar (w megabajtach). Dla ostatniej, czwartej, możesz po prostu wcisnąć *Enter*, aby wykorzystać pozostałe miejsce.
4. Wciśnij *Enter*, aby utworzyć nową partycję na początku wolnego obszaru (**Beginning**).
5. Wybierz typ partycji. Klawiszem *Enter* przejdź na drugą stronę listy i wpisz identyfikator „fd” („Linux RAID autodetect”).
6. Dla partycji 1 na każdym z dysków wybierz **Bootable**.
7. Strzałką w dół przesuń się na pozostały wolny obszar i stwórz następną w kolejności partycję.

Identyczne czynności powtarzamy także dla dysku hdb:

```
# cmdisk /dev/hdb
```

Odpowiadające sobie partycje na obu dyskach powinny mieć identyczny rozmiar. W przeciwnym razie RAID wprawdzie zadziała, ale system otrzyma obszar o rozmiarze najmniejszej z partycji wchodzących w skład macierzy.

2.3 Utworzenie macierzy RAID

Najpierw ładujemy sterowniki dla poziomów RAID, które chcemy wykorzystać. Na 18374% znajdują się one na płycie, więc jeżeli system nie może ich znaleźć, oznacza to, że masz uszkodzoną płytę albo napęd CD.

```
# modprobe raid1
```

Jesteśmy już gotowi do spięcia naszych partycji w macierz. Służy do tego program `mdadm`.

Aby utworzyć partycję /:

```
# mdadm --create /dev/md0 --level=1 --raid-devices=2 /dev/hda1 /dev/hdb1
```

Aby utworzyć partycję swap:

```
# mdadm --create /dev/md1 --level=1 --raid-devices=2 /dev/hda2 /dev/hdb2
```

Aby utworzyć partycję /home:

```
# mdadm --create /dev/md2 --level=1 --raid-devices=2 /dev/hda3 /dev/hdb3
```

Aby utworzyć partycję /var:

```
# mdadm --create /dev/md3 --level=1 --raid-devices=2 /dev/hda4 /dev/hdb4
```

Uwaga: partycja główna (/) musi ZAWSZE być zamontowana jako */dev/md0*. Jeżeli do tego urządzenia podmontujesz jakąkolwiek partycję podrzędną, nie uda Ci się ukończyć instalacji, ponieważ tak postawiony system nie będzie chciał się uruchomić.

Za pomocą dyrektywy `level` wybieramy poziom tworzonej macierzy RAID, za pomocą `raid-devices` ustalamy ilość dysków ją tworzących. Później wymieniamy fizycznie istniejące na nich partycje, które chcemy dodać do tworzonej macierzy.

Na tym etapie powinniśmy mieć już działające partycje RAID. Po ich stworzeniu system dokonuje ich wzajemnej synchronizacji tak, aby były one swoimi lustrzanymi odbiciami co do bajta. Cały proces może być monitorowany poleceniem

```
# cat /proc/mdstat
```

albo (dla pojedynczej partycji RAID):

```
# mdadm --misc --detail /dev/md0
```

Nie musisz czekać z instalacją na zakończenie procesu synchronizacji. Możesz nawet w jej trakcie zrestartować komputer i w żaden sposób go nie zakłócisz.

2.4 Zamontowanie systemu plików

Pora na utworzenie na partycjach RAID jakiegoś systemu plików. Na potrzeby artykułu wybrałem ReiserFS, aczkolwiek warto zapoznać się także z innymi oraz dobrać odpowiedni do roli, jaką dana partycja ma pełnić w systemie, aby wycisnąć z komputera jak najwięcej.

```
# mkreiserfs /dev/md0
```

```
# mkswap /dev/md1
```

```
# mkreiserfs /dev/md2
```

```
# mkreiserfs /dev/md3
```

Instalator Arch Linuksa wymaga, aby twardy dysk, na którym zainstalowany zostanie system, był zamontowany w katalogu `/mnt`. Ponieważ pracujemy z „autorskim” zestawem partycji, musimy samodzielnie zamontować partycje tam, gdzie trzeba.

```
# mount /dev/md0 /mnt
```

```
# mkdir /mnt/home
```

```
# mount /dev/md2 /mnt/home
```

```
# mkdir /mnt/var
```

```
# mount /dev/md3 /mnt/var
```

2.5 Instalacja i konfiguracja Arch Linuksa

Nie będę tutaj objaśniać szczegółowo procesu instalacji Arch, zakładam, że w podstawowym stopniu użytkownik wie, jak on się odbywa. W razie czego można zajrzeć do oficjalnego przewodnika instalacji po dodatkowe szczegóły.

Najważniejszy zestaw kroków to:

1. Wpisz `/arch/setup`, aby włączyć instalator.
2. Wybierz `<OK >` w ekranie powitalnym.
3. Wybierz `1 CD-ROM`, aby wybrać instalację z CD-ROMU.
4. W menu głównym zostaw całkowicie opcje zabawy dyskami i skocz od razu do punktu `2 Select packages`. Wybierz wszystkie pakiety z kategorii *base*, a także pakiet *mdadm* z kategorii *system* (od wersji *arch-base-0.7.2* pakiet ten został przeniesiony do kategorii *base*, aczkolwiek na wszelki wypadek upewnij się, że jest on zaznaczony).
5. Wybierz `3 Install packages` i poczekaj, aż wszystko znajdzie się na twardej dysku.

Kolejnym etapem jest wybranie `4 Configure system`. Poświęćmy mu trochę więcej uwagi, ponieważ Arch Linux przechodzi właśnie proces zmiany podstawowego mechanizmu ramdysku ze starszego *mkinitrd* na *mkinitcpio*. Dla wersji **poniżej 0.7.2** musimy tutaj wykonać następujące czynności:

Edytujemy `/mnt/etc/mkinitrd.conf`. Upewniamy się, że `REMOVE_RAID` **NIE** jest ustawiony na `1`. Następnie ustawiamy zmienną `RAID_ROOT_DEVICES`, w której wpisujemy wszystkie urządzenia wchodzące w skład głównej partycji RAID (`md0`).

```
# nie bootujemy systemu z USB ani firewire
REMOVE_USB=1
REMOVE_FW=1
# Ale potrzebujemy tego
REMOVE_RAID=
REMOVE_DM=
# Nasza główna (/) partycja RAID to md0
# (i żadna inna!)
RAID_ROOT_ARRAY="md0"
# potrzebujemy jedynie listy urządzeń dla md0, nie dla wszystkich partycji RAID
RAID_ROOT_DEVICES="/dev/hda1 /dev/hdb1"
```

Trzeba także skonfigurować bootloader Grub (plik `/mnt/boot/grub/menu.lst`):

```
# (0) Arch Linux
title Arch Linux [[/boot/vmlinuz26]]
root (hd0,0)
kernel /vmlinuz26 root=/dev/md0 ro
```

Dla wersji wykorzystującej **mkinitcpio** edytujemy plik `/mnt/etc/mkinitcpio.conf`. Dodajemy *raid* do zmiennej `HOOKS` i określamy partycję RAID, z której będziemy bootować system. Tyle wystarczy:

```
md=0,/dev/hda1,/dev/hdb1
```

Konfiguracja bootloadera Grub (plik `/mnt/boot/grub/menu.lst`):

```
# (0) Arch Linux
title Arch Linux [[/boot/vmlinuz26]]
root (hd0,0)
kernel /vmlinuz26 root=/dev/md0 ro md=0,/dev/hda1,/dev/hdb1
```

Dla wszystkich wersji: ostatnią czynnością jest edycja pliku `/mnt/etc/fstab`:

<code>/dev/md0</code>	<code>/</code>	<code>reiserfs</code>	<code>defaults</code>	<code>0</code>	<code>0</code>
<code>/dev/md1</code>	<code>swap</code>	<code>swap</code>	<code>defaults</code>	<code>0</code>	<code>0</code>
<code>/dev/md2</code>	<code>/home</code>	<code>reiserfs</code>	<code>defaults</code>	<code>0</code>	<code>0</code>
<code>/dev/md3</code>	<code>/var</code>	<code>reiserfs</code>	<code>defaults</code>	<code>0</code>	<code>0</code>

Wprowadź niezbędne dla twojej konfiguracji modyfikacje w pozostałych plikach i zainstaluj jądro 2.6.x, wybierając 5 `Install kernel`. Instalację bootloadera musimy przeprowadzić ręcznie, dlatego możesz już zakończyć działanie instalatora.

2.6 Instalacja Gruba na pierwszym HDD

Jest to ostatni krok niezbędny do wykonania pierwszego uruchomienia naszego systemu. Grub musi być instalowany na każdym dysku z osobna. Z poziomu płyty instalacyjnej wgramy go tylko na pierwsze z urządzeń, podczas gdy drugie i ewentualnie kolejne, jeśli masz więcej dysków, zostaną poddane temu procesowi już po uruchomieniu systemu. Krótkiego opisu wymaga procedura instalacji. Kopiujemy pliki instalacyjne Gruba do `/mnt/boot/grub`, montujemy `procfs` oraz drzewo urządzeń w obrębie `/mnt`, następnie przełączymy się `chrootem` na ten katalog i zainstalujemy wszystko. Na koniec zapiszemy konfigurację RAID w pliku `/etc/mdadm.conf` tak, aby była automatycznie montowana po restarcie.

Kopiowanie oraz `chrootowanie` odbywa się w następujący sposób:

```
# cp -a /mnt/usr/share/grub/i386-pc/* /mnt/boot/grub
# sync
# mount -o bind /dev /mnt/dev
# mount -t proc none /mnt/proc
# chroot /mnt /bin/bash
```

Jeżeli po `chrootowaniu` masz problem z zobaczeniem wpisywanych znaków, wpisz polecenie `reset` w konsoli — powinno pomóc. W każdym razie, gdy już będziesz widział, co piszesz, uruchom Gruba:

```
# grub
```

Po krótkim oczekiwaniu powinniśmy znaleźć się teraz w wierszu poleceń bootloadera. Wydamy tu polecenie instalacji go na pierwszym z dysków:

```
grub> root (hd0,0)
grub> setup (hd0)
grub> quit
```

Teraz trzeba zapisać konfigurację RAID do pliku. W poprzednich wersjach Arch Linuksa nie było to konieczne, ponieważ sterowniki RAID były wbudowane w jądro. Lecz kiedy ładowane są one dynamicznie w trakcie uruchamiania systemu, macierze nie są wykrywane automatycznie.

Domyślny plik `/etc/mdadm.conf` zawiera w zasadzie jedynie szereg objaśniających nie wiadomo co komentarzy. Aby go uzupełnić, niepotrzebna jest nam jakaś szczególna finezja. Wystarczy przechwycić wyjście programu `mdadm` i dokleić je na koniec tego pliku.

```
# mdadm -D --scan >>/etc/mdadm.conf
```

To wszystko. Możesz teraz powrócić do instalatora, wciskając `Ctrl+D` lub wpisując `exit`.

2.7 Odmontowanie systemu plików i pierwsze uruchomienie

Zanim zakończysz działanie płyty instalacyjnej, pamiętaj o odmontowaniu dysku twardego:

```
# cd /
# umount /mnt/dev /mnt/proc /mnt/home /mnt/var
# umount /mnt
```

Teraz usuwamy płytę z napędu i wydajemy polecenie `reboot`. Jeśli ujrzysz procedurę ładowania nowego systemu, możesz odetchnąć, aczkolwiek pamiętaj, że do pełni szczęścia wymagane jest, aby procedura ta doszła szczęśliwie i bez żadnego błędu do końca, co wcale nie jest takie oczywiste w przypadku RAID-a :).

2.8 Instalacja Gruba na pozostałych dyskach

Gruba wgrywamy także na drugi z twardych dysków. Dzięki temu nawet, gdyby padł `hda`, system wciąż będzie mógł być uruchomiony z `hdb`. Po zalogowaniu się jako `root` wykonaj następujące czynności:

```
# grub
grub> device (hd0) /dev/hdb
grub> root (hd0,0)
grub> setup (hd0)
grub> quit
```

Możesz również zarchiwizować układ partycji na każdym z urządzeń, dzięki czemu ich ewentualna wymiana będzie dziecinnie prosta:

```
# mkdir /etc/partitions
# sfdisk --dump /dev/hda >/etc/partitions/disc0.partitions
# sfdisk --dump /dev/hdb >/etc/partitions/disc1.partitions
```

3 Problemy

Podczas instalacji programowego RAID w Arch Linuksie jest całkiem prawdopodobne, iż pierwsza próba wykonania takowej nie powiedzie się. Także i później, w trakcie użytkowania, może zdarzyć się, że system nie będzie chciał z jakiegoś powodu się uruchomić. Za pomocą płyty LiveCD z instalatorem często można zlokalizować i usunąć usterkę. Jedyny problem polega na podmontowaniu się do już istniejącej macierzy RAID i nieuszkodzenia jej przy okazji. Zadanie to jest wykonalne, choć nieco topornie.

Założmy więc, że np. dodawaliśmy własne ulepszenia do plików startowych i popełniliśmy błąd składni, przez co system nie chce się uruchomić. Wkładamy do napędu płytkę Arch Linuksa i bootujemy z niej. Potrzebny nam jest dostęp do dysku, dlatego na początku ładujemy sterowniki dla wykorzystywanych poziomów RAID, po czym inicjujemy macierze... w identyczny sposób, jak przy ich tworzeniu:

```
# modprobe raid1
# mdadm --create /dev/md0 --level=1 --raid-devices=2 /dev/hda1 /dev/hdb1
# mdadm --create /dev/md1 --level=1 --raid-devices=2 /dev/hda2 /dev/hdb2
# mdadm --create /dev/md2 --level=1 --raid-devices=2 /dev/hda3 /dev/hdb3
# mdadm --create /dev/md3 --level=1 --raid-devices=2 /dev/hda4 /dev/hdb4
```

Dyrektywa `create` wygląda dość hardcore'owo, ale bez obaw. Dokładnie w identyczny sposób radziłem sobie u siebie i danych jakoś nie utraciłem. W każdym razie masz teraz stworzone urządzenia od `md0` do `md3`, które możesz użyć do zamontowania systemu plików:

```
# mount /dev/md0 /mnt
# mount /dev/md2 /mnt/home
# mount /dev/md3 /mnt/var
```

Wykonaj teraz niezbędne naprawy w konfiguracji systemu, ale wstrzymaj się z hurraoptymistycznym restartem. Bez dodatkowych czynności system dalej by się nie uruchamiał, tym razem z powodu „błędów” w macierzy RAID. Rzecz w tym, że montując je tak, jak przedstawiłem powyżej, macierze otrzymują nowe identyfikatory. Musimy zatem jeszcze raz wgrać Gruba z całą chrootową otoczką, a także zapisać nowe ustawienia w `/etc/mdadm.conf`.

```
# cp -a /mnt/usr/share/grub/i386-pc/* /mnt/boot/grub
# sync
# mount -o bind /dev /mnt/dev
# mount -t proc none /mnt/proc
# chroot /mnt /bin/bash
# grub
grub> root (hd0,0)
grub> setup (hd0)
grub> quit
```

Otwórz `/etc/mdadm.conf` jakimś edytorem i usuń z niego stare wpisy. Następnie z powrotem kontynuuj cały proces:

```
# mdadm -D --scan >>/etc/mdadm.conf
# exit
```

```
# cd /  
# umount /mnt/dev /mnt/proc /mnt/home /mnt/var  
# umount /mnt  
# reboot
```

Arch Linux powinien znów uruchomić się po staremu.

4 Zakończenie

Artykuł ten bazuje na angielskim tekście „Installing with software RAID or LVM” dostępnym na wiki Arch Linuksa. Pierwotzór objaśnia dodatkowo, jak utworzyć na macierzy RAID logiczne woluminy LVM dające większą skalowalność, niż zwykłe partycje. Doszedłem jednak do wniosku, że troszkę tego za dużo, jak na jeden tekst i dla większej czytelności skupiłem się tylko i wyłącznie na macierzach RAID. Mam nadzieję, że z LVM-em czy bez niego, artykuł spełnił swą rolę. Zatem gratulacje, zainstalowałaś właśnie Arch Linuksa z programowym RAID-em!

Artykuł udostępniony na licencji GNU Free Documentation License 2.1