

Adaptec ASR2120S

Stručný manuál

Autor: František Ryšánek <rysanek@fccps.cz>

FCC Průmyslové Systémy s.r.o.

Řadič Adaptec ASR2120S je SCSI řadič, na kterém lze nakonfigurovat "diskové pole", anglicky RAID (Redundant Array of Inexpensive Disks) - virtuální disk skládající se z více disků.

Jedná se o solidní RAID řešení: umí RAID level 0, 1, 5 a JBOD a lze vytvářet více logických disků na jedné sadě fyzických disků - viz též samostatný dokument o typech polí.

Jedná se o hardwarový RAID, založený na procesoru Intel IOP302, čímž jsou dány následující parametry: PCI64/66MHz (533 MBps), jádro i960 taktované na 66 MHz, on-chip hardware pro akceleraci XOR (pro RAID5). Řadič je dále vybaven 64MB SDRAM s možností instalovat baterii pro zálohu diskové write-back cache. Řadič disponuje pouze jedním SCSI kanálem, jedná se ovšem o U320 SCSI, což je v tuto chvíli nejrychlejší disková sběrnice na trhu. K dispozici je také dvoukanálový sourozenec 2200S. Čtyřkanálový 3400S umí pouze U160.

Vedle provozu s trvale namontovanými pevnými disky řadič umí spolupracovat i s hot-swapovými storage skříněmi. V podmínkách firmy FCC Průmyslové Systémy a u jejích zákazníků je praxí důkladně prověřena bezchybná funkce s SCA backplanem založeným na SAF-TE procesoru GEM318 ve skříní Advantech RS-200-RP. Lze pochopitelně uvažovat i o připojení disků umístěných v externích expanzních jednotkách.

Řadič má ve srovnání s konkurencí poměrně kvalitní softwarovou výbavu pro Windows a Linux.

Celkově se jedná o jeden z nejlepších PCI SCSI RAID řadičů na trhu. Co do počtu kanálů, taktu procesoru a velikosti cache jde přinejmenším o solidní střední třídu a navíc není drahý.

Většina zákazníků si pořizuje pole kvůli spolehlivosti, proto **důrazně doporučujeme RAID 1 (mirror)! případně RAID5.**

S nastavením řadiče a vytvořených polí lze pracovat jednak z nástroje, který je součástí BIOSu řadiče a lze ho vyvolat klávesou Ctrl+A při startu počítače (ještě než nastartuje operační systém), jednak pomocí utility nainstalované v operačním systému Windows či UNIX (Linux, FreeBSD).

Grafický webový správce „Adaptec Storage Manager“ pro Windows a Linux je na pohled hezký, ale poměrně těžkopádný. Příkazový řádek *aaccli* působí spartánsky, ale je maličký, mocný a funguje svižně. Oba druhy utilit jsou použitelné pro obnovu degradovaného pole za provozu. Nástroj v BIOSu lze použít pro počáteční nastavení a v případech, kdy operační systém není k dispozici.



Společné poznámky (nezávislé na operačním systému)

Počáteční nastavení pole v BIOSu

K vytvoření pole jsou potřeba alespoň dva disky.

Následující postup popisuje vytvoření pole typu mirror ze dvou disků, které zabírá celou kapacitu obou disků.

proběhnou úvodní hlášky BIOSu a vzápětí se objeví se zpráva

```
<<< Press <Ctrl>+<A> for Adaptec RAID Configuration Utility >>>  
Booting the Controller Kernel...
```

kteřou již vypisuje BIOS řadiče Adaptec ASR2120S. Řadiči chvíli trvá, než se zinicizuje, najde disky atp., máte tedy trochu času. V průběhu této inicializace **stiskněte Ctrl+A**. BIOS řadiče potvrdí, že příkaz přijal a po skončení inicializace se spustí konfigurační utilita, která je součástí BIOSu řadiče.

V konfigurační utilitě vytvoříme pole, a to následujícím způsobem:

- > zvolte "Array Configuration Utility" (šipky, <Enter>)
- > zvolte "Initialize Drives" (šipky, <Enter>)
- > mezeríkem označte oba disky a stiskněte <Enter>
- > přijměte varování v červeném rámečku, že inicializace smaže všechna data na označených discích – stiskněte klávesu 'y'
- > řadič bude chvíli pracovat na obou discích a poté se utilita vrátí do hlavního menu.
- > zvolte "Create Array" (šipky, <Enter>)
- > mezeríkem označte oba disky a stiskněte <Enter>
- > otevře se tabulka s parametry nově vytvářeného pole ("Array Properties"), kterou utilita prochází odshora dolů a u každé volby se ptá na výběr:
 - > Array Type: zvolte "Raid1 (Mirror)"
 - > Array Label: zadejte nějaké jméno, třeba "Pole c.1"
 - > Array Size: ponechte default (celá dostupná kapacita) – dvakrát <Enter>
 - > Stripe Size: u mirroru nemá smysl
 - > Read Caching: ponechte default (ano) - <Enter>
 - > Write Caching: pokud jste si neobjednali záložní baterii, bezpečnější je volba "Disable".
 - > Create RAID via: zvolte default (Build/Verify) – ostatní volby jsou méně praktické.
 - > tlačítko "Done" potvrďte klávesou <Enter>.
- > přijměte informativní hlášku, že pole bude k dispozici okamžitě, ale po nějakou dobu ještě bude na pozadí (v režii hardwaru řadiče) probíhat vytváření mirroru – stiskněte <Enter>
- > utilita se vrátí do hlavního menu a řadič se dá do práce (disky se rozsvítí).
- > nyní je možno utilitu opustit a začít s polem pracovat – například lze ihned přikročit k instalaci operačního systému. (2x escape, vybrat "Yes", chceme ukončit utilitu, <Enter>)
- > po opuštění utility se počítač včetně řadiče resetuje, ale vytváření pole poté automaticky pokračuje tam, kde řadič před resetem skončil.
- > nyní můžete pokračovat v instalaci podle kapitoly určené pro Váš operační systém.



Jak se pozná porucha pole

Když něco není v pořádku, počítač začne akusticky signalizovat poruchu. Samotný řadič vydává táhlé píštění a pokud má Váš počítač hot-swap backplane, řadič typicky ještě vyvolá alarmovou signalizaci backplanu, takže se přidá druhý zvukový signál – například přerušované pípání. Na poli LED kontrolkek, které je umístěno na čelním panelu počítače a interně ovládáno SCA backplanem, se také červeně rozsvítí nebo rozbliká kontrolka disku, který má potíže. Případně, podle typu závady, začnou blikat kontrolky všech disků na zasaženém kanálu nebo v zasaženém poli.

Lepší přehled si lze vytvořit na základě bootovací hlášky BIOSu řadiče – řadič automaticky hlásí, který disk chybí nebo zhavaroval (SCSI kanál : ID : LUN)

Ještě nějaké podrobnosti navrch lze zjistit v BIOS utilitě nebo ve správcovských utilitách pod Windows či pod Unixem (příkazový řádek `accli`, GUI storage manager).

Také je možno "identifikovat" havarovaný disk záměrným blikáním LEDdiodou na SCA panelu.

Obnova redundance pole po havárii disku

Obnovit lze pouze pole typu RAID 1 (mirror) nebo RAID 5.

Vždy je třeba nejprve vytáhnout havarovaný disk a nahradit ho novým. Pokud jsou disky k řadiči připojeny prostřednictvím hot-swap mechaniky, řadič toho umí využít a výměna disků za provozu funguje – při výměně disku není třeba vypínat počítač, ani restartovat operační systém. Při obnovitelné degradaci pole by měl počítač nastartovat a běžet i s degradovaným polem.

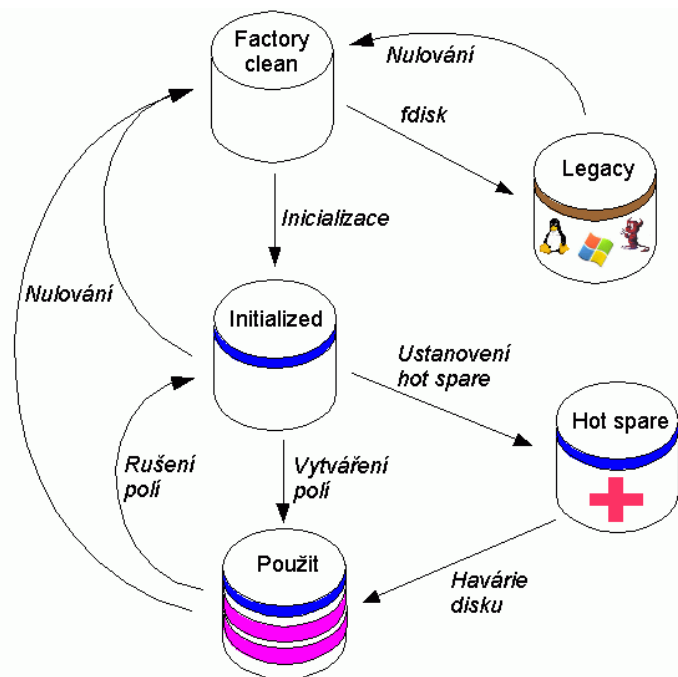
Postup pro obnovu je zhruba následující:

- 1) náhradní disk je třeba uvést do stavu "initialized"
- 2) inicializovaný náhradní disk je třeba označit jako "hot spare"
- 3) v tu chvíli automaticky nastartuje obnova redundance. Obnovu není třeba spouštět explicitně.

Disk připojený k řadiči ASR2120 se může nacházet v jednom z následujících stavů:

- "**factory clean**" neboli "**not initialized**" – čistý, obsahující ve všech sektorech samé nuly (včetně MBR) – do tohoto stavu lze použitý disk znovu uvést přepsáním nulami.
- "**legacy**" nebo "**DOS**" – na disku byla objevena PC/BIOS/DOS tabulka rozdělení. Řadič má podezření, že na disku jsou užitečná data, takže bude odmítat disk inicializovat. Tuto ochranu lze "přehlasovat" potvrzením inicializace (BIOS) nebo parametrem `/always` příkazu "`disk initialize`" (`accli`).
- "**initialized**" – připraven pro RAID. Na takovém disku lze vytvářet pole, nebo ho lze označit jako hot spare.
- "**hot spare**" – v tomto stavu disk vydrží pouze v případě, že řadič nemá žádné degradované pole. Jakmile se objeví opravitelná degradace pole, první vhodný "hot spare" disk je automaticky natrvalo zařazen do pole namísto zhavarovaného disku a nastartuje obnova.
- **použit** – zařazen v poli. Utility Adapteců bohužel většinou tento disk zobrazují jako "initialized" a je třeba se podívat do seznamu polí, které disky jsou obsazené. Souvisí to patrně s tím, že na jednom disku lze nakonfigurovat více polí a může na něm zbývat volné místo pro konfiguraci dalšího pole.





Náhradní disk pro obnovu by měl být v ideálním případě “factory clean”, popřípadě inicialized. Pokud již byl použit samostatně nebo v poli (které nebylo zrušeno), může to znamenat práci navíc.

Pokud řadič detekuje, že nově zasunutý disk je původní disk, patřící do momentálně živého a degradovaného pole, který byl znovu připojen po dočasné poruše (vypadlý kabel, vypnuté napájení, simulace defektu vytažením z hot-swap mechaniky), nastartuje obnova pole automaticky. Lhostejno, zda byl disk připojen ve vypnutém stavu nebo za provozu (hot-swap). Automatickou obnovu v tomto případě spustí a provádí samotný řadič – není ani třeba mít nainstalovány a spuštěny správcovskou aplikaci, dokonce ani ovladače.

Za určitých okolností se řadič sám chopí obnovy i na “factory clean” disk – podle zkušeností a poznámky v manuálu se zdá, že se tak stane, pokud řadiči sebereme za běhu disk z živého pole a po chvíli mu namísto něj předložíme čistý disk jako náhradu.

Obecně je ovšem třeba spustit obnovu pole ručně – buď v BIOSu nebo ve správcovské utilitě v některém z podporovaných operačních systémů.

Nově zasunutý disk je třeba napřed inicializovat (pokud již nebyl dříve inicializován a poté uskladněn v tomto stavu). Řadič si při inicializaci na disku vytvoří čistý MBR a své režijní datové struktury, čímž ho připraví k zařazení do pole.

Pokud byl disk již dříve použit v poli a toto pole nebylo před vymontováním disku zrušeno, řadič toto pole nadetekuje. Před použitím takového disku pro obnovu je potřeba “fantóma” starého pole správcovskou utilitou zrušit a pro jistotu disk třeba ještě zinicializovat.

Alternativně se lze fantóma pole nebo “legacy” dat zbavit tak, že disk smažeme (vynulujeme) na obyčejném SCSI řadiči - napříkazem pod Unixem příkazem

```
cp /dev/zero /dev/sd<X> (uvedený název diskového zařízení je z Linuxu)
```

případně

```
dd if=/dev/zero of=/dev/sd<X> (pomalejší metoda – sync() po každém bloku)
```

Poté se disk bude tvářit jako “factory clean”. Viz samostatný dokument obecně o RAIDu.

Obnova pole nastartuje automaticky poté, co nově vložený a ručně inicializovaný disk označíme jako náhradní, tj. “spare drive”. Jakmile řadič zjistí, že má k dispozici náhradní disk, zmocní se ho, přiřadí ho do pole a začne na pozadí obnovovat.

Obnova tedy úspěšně začala. Pokud obnovu pole přeručíme, což lze provést prakticky jedině vypnutím či resetem řadiče, po novém startu firmwaru řadič automaticky pokračuje v obnově pole tam, kde před resetem skončil. Reset řadiče je vyvolán studeným i teplým resetem celého počítače a lze ho vyvolat i programově, což mohou ve vyjímečných případech provést ovladače i za běhu operačního systému.

Obnova pole v BIOS utilitě (pouze na platformě x86 PC)

Zjistili jsme problém s polem, opravili jsme hardwarovou příčinu degradace. Nyní budeme obnovovat pole. Nastartujeme tedy počítač. (Pokud máme hot-swap rámečky, lze vyměnit vadný disk i za chodu.)

Po úvodních hláškách BIOSu se objeví varování, že něco není v pořádku – konkrétně které pole selhalo (číslo a popis) a co se s ním stalo, tj. který fyzický disk (člen pole) má problémy. Pokud se jedná o opravitelnou degradaci, řadič čeká 30 sekund na potvrzení a pak se pokusí nastartovat.

Pokud v této době stiskneme CTRL+A, spustí se BIOSová utilita pro správu RAID řadiče a polí.

-> zvolte Array Configuration Utility

-> utilita si postěžuje, že se “změnila konfigurace” (tj. vypadl disk) – zvolte “Accept”

-> v tuto chvíli je možné nově vložený disk inicializovat v menu “Initialize Drives” (vybrat disky mezerníkem, potvrdit enterem) – zdá se ovšem, že tento krok není třeba, pokud je nově vložený náhradní disk “factory clean”.

-> přistupte k přidělení nového disku jako “hot spare”. Zvolte “Manage Arrays”,

-> šipkami přesuňte kurzor v seznamu na degradované pole, které chcete opravit

-> stiskněte kombinaci Ctrl+S – otevře se seznam volných disků a seznam “hot spare” disků

-> šipkami přesuňte kurzor v seznamu na volný disk, který chcete použít jako náhradní. V tomto seznamu se nabízejí jenom disky, které dosud nejsou použity v poli, resp. jim zbývá dostatečná volná kapacita pro obnovu výše vybraného degradovaného pole.

-> klávesou <Insert> nebo mezerníkem zvolený disk přesunete ze seznamu volných disků do seznamu přiřazených záložních disků (“Assigned Hotspare Drives”). Přiřazení disku můžete v tuto chvíli ještě vrátit klávesou Delete v seznamu přiřazených disků, všechny změny provedené v této obrazovce můžete vzít zpět stiskem Escape.

-> řekněme, že jste přesunuli jeden disk ze seznamu volných disků do seznamu záložních disků.

Pokud se právě snažíte obnovit redundanci degradovaného pole a vyměnili jste havarovaný disk, jedná se pravděpodobně v tuto chvíli o jediný disk, který v těchto seznamech dohromady figuruje. Pokud jste s výběrem spokojeni, stiskněte klávesu <Enter> a ještě jednou potvrďte (“y”, <Enter>).

-> během několika sekund by se měla rozběhnout obnova výše ošetřeného pole.

Obnova probíhá na pozadí, BIOSovou správcovskou utilitu lze opustit a nastartovat operační systém.



Obnova pole v utilitě aaccli (Windows, UNIX)

Utilita *aaccli* používá nativní názvosloví Adaptec FSA Raidu, které úplně přesně neodpovídá obecné terminologii, ani terminologii jiných RAID produktů od Adaptecu, dokonce ani terminologii aktuálního RAID BIOSu dodávného s řadičem. Proto níže uvádíme stručnou definici pojmů:

- **controller** = řadič Adaptec 2120S
- **container** = logický disk, pole (RAID 0, 1 nebo 5)
- **disk** = fyzický disk
- **task** = úloha. V zásadě proces, prováděný kontrolérem, např. obnova pole.
- **failover** (failover drive) = “hot spare”, tj. náhradní disk.

V UNIXu se utilita spouští jednoduše příkazem “aaccli”, pod Windows přes menu *Start -> Programy -> SMBE -> CLI*. Po spuštění se octnete v poměrně spartánském příkazovém řádku. K dispozici je online nápověda (otazníková konvence ala Cisco) a také je možno opakovat předešlé příkazy (listuje se v nich šipkou nahoru a dolů). Nezbytným prvním příkazem je “otevření” kontroléru.

```
sh# aaccli
CLI> ?
CLI> controller list
CLI> open aac0 # “otevři” kontrolér – kupodivu i pod Windows
CLI> container list
CLI> disk list
CLI> task list
CLI> disk blink 1 10 # blikaj LEDdiodou fyzického disku č.1 po dobu 10 sekund
CLI> container delete 1 # smaž nepoužité pole
CLI> disk initialize (0,1,0) # inicializuj fyzický disk se SCSI ID=1
CLI> disk list # už je inicializovaný?
CLI> container set failover 0 (0,1,0) # tumáš failover drive
# teď by měla začít obnova
CLI> task list
```

Jmenujme ještě tři zajímavé příkazy:

```
CLI> container scrub 0
```

Tento příkaz provede kontrolu konzistence na funkčním poli – nejedná se o příkaz pro manuální spuštění obnovy, při pokusu o kontrolu degradovaného pole se ohradí chybovou hláškou.

```
CLI> container restore RAID5 0
```

Tento příkaz se pokusí zachránit data z havarovaného pole RAID5 – tj. v případě, kdy zhavaroval víc než jeden disk. Dokumentace neuvádí podrobnosti – mohlo by se jednat o funkci, která prochází disky (včetně “porouchaných”) sektor po sektoru a kompletuje data v “sadách proužků”, kde se podaří poskládat dostatek dat k obnově. Jedná se v každém případě o naprosto nouzovou funkci bez záruky rozumného výsledku.



```
CLI> disk show defects (0,1,0)
```

Tento příkaz vypíše počet sektorů, které jsou vadné již z výroby a počet vadných sektorů, které se objevily v průběhu používání disku. Moderní disky zvládají určitý objem vadných sektorů včas detekovat a transparentně přemapovat do náhradního prostoru.

Instalační CD obsahuje extrémně podrobnou referenční příručku k použití aaccli (350 stran).

Zajímavou vlastností aaccli je možnost skriptování. Skript je textový soubor, který obsahuje prostou posloupnost příkazů, řazených přesně tak, jak by je zadával operátor v interaktivním režimu. Každý příkaz je na samostatném řádku. Skript lze natáhnout a provést v rámci CLI relace příkazem

```
CLI>@jmeno_skriptu.txt
```

Tato zavináčová konvence je popsána v manuálu.

Je tu ovšem ještě další možnost: skript lze externě nasměrovat na standardní vstup utility:

```
sh# aaccli <jmeno_skriptu.txt
```

Pozor, v tomto případě je třeba obsloužit skutečně celou relaci CLI: první příkaz musí být *open* a poslední příkaz musí být *exit*. Zejména humornou chybou je vypuštění závěrečného příkazu „exit“: v tom případě se CLI neukončí, nevrátí se do shellu, nelze mu ručně zadat další příkazy, nereaguje na Ctrl+C, pomůže mu jedině ‚kill‘.

Lze se domnívat, že by utility aaccli šlo ovládat i přes „rouru“ například z Perlu nebo C.

Referenční příručka k aaccli poněkud záhadně tvrdí, že toto je nejjednodušší podoba skriptu – žádné pokročilejší schopnosti ovšem nerozvádí.

Obnova pole po havárii řadiče

Stačí vyměnit řadič za nový a připojit disky na původní kanály a SCSI ID (v SCA mechanice pokud možno nepoplést pořadí). Řadič nicméně většinou dokáže vzájemnou záměnu disků odhalit a vyrovnat se s ní (příslušně upraví konfiguraci).

Konfigurace pole je uložena na discích a měla by se automaticky načíst a spustit. Když řadič detekuje změnu, čeká třicet sekund na potvrzení/zrušení (default = změna potvrzena) a pak pokračuje v bootování.

Čištění disku po použití v poli - pozor, záلودnost

Dobrý administrátor, poté co dostal od doavatele nový typ SCSI RAID řadiče, si před ostrým provozem nejprve předem nacvičí ovládání – nejlépe tak, že simuluje výpadek.

Mějme stroj s řadičem ASR2120 a v něm dva disky. Na discích vytvoříme pole typu mirror. Na pole nainstalujeme operační systém. Vyjmeme jeden z obou disků – řadič signalizuje degradaci pole, pole běží dál v degradovaném režimu. Rozhodneme se, že vytažený disk vymažeme, abychom mohli simulovat obnovu na prázdný disk. Proto korektně vypneme operační systém i počítač, vyjmeme zbylý disk (který tímto zůstal „v pořádku“) a dáme ho stranou. „Havarovaný“ disk připojíme zpátky k RAID řadiči, spustíme počítač a BIOSovou utility. Řadič nadetkoval degradované pole – pochopitelně, vytažený disk je jednou ze dvou polovin mirroru (nyní tou neaktuální). Takže v BIOSové utilitě zrušíme pole a vypneme počítač. Vložíme i „dobrý“ disk – takže nyní jsou v počítači opět oba disky, jeden dobrý s nakonfigurovaným polem a jeden prázdný (leč



inicializovaný), který do pole také původně patřil. Nastartujeme počítač – a **OUHA. Řadič nenašel pole. JAKTO?!?**

Došlo patrně k tomu, že řadič poznal, že oba disky patří k sobě. A za aktuálnější považoval ten disk, kde pole bylo zrušeno – tuto operaci jsme skutečně provedli časově jako poslední. Takže konfiguraci synchronizoval tím způsobem, že zrušil pole i na “dobrém” disku - na kterém jsme měli korektně nainstalovaný a vypnutý operační systém.

Z toho plyne poučení: pokud si hrajete s polem a vzájemně vyměňujete disky, a pak je před uložením zpět do skladu čistíte, nespolehejte na to, že “odebráním pole” je disk opět čistý! Disky je třeba buď inicializovat, nebo vynulovat – buď v utilitách řadiče, nebo lépe pod UNIXem na samostatném SCSI kanále kopíí dat z /dev/zero – viz obecný dokument o RAIDu.

Je sice málo pravděpodobné, že by se Vám oba disky časem opět sešly v jednom stroji, a kromě toho pokud původní disk nějakou dobu poběží, bude považován za aktuálnější, ale čert nikdy nespí a do firmwaru řadiče není vidět, co a jak v kritický okamžik provede. Pro maximální míru determinismu je jediné vhodné použití pro obnovu pole disk, který je “factory clean”!

Umístění režijních dat

Pro rychlejší čištění fyzických disků po použití je dobré vědět, kde jsou na disku umístěna režijní data – stačí totiž vynulovat sektory, o kterých víme, že je řadič používá a není třeba čekat, až se důkladně vynuluje celý disk.

Zdá se, že řadič ASR2120S si zapisuje své režijní údaje do prvních několika sektorů fyzického disku. Poslední sektor, který po inicializaci a zřízení mirroru metodou „clear“ obsahuje nenulová data, je sektor č.17, počítáno od nuly => je třeba přepsat prvních 18 sektorů po 512 bajtech.



Windows 2000

Instalace Windows2000

V průběhu instalace budeme potřebovat disketu s ovladači pro SCSI RAID řadič. Ovladače pro řadič Adaptec ASR2120S se ovšem nedodávají na disketě, ale na CDROM. Takže je potřeba ovladače napřed zkopírovat na prázdnou, zformátovanou disketu. Konkrétně je třeba do kořenového adresáře diskety zkopírovat z CD-ROM obsah adresáře \Packages\WINDOWS_2000\Driver\ . Pokud disketu vytváříte z archivu ovladačů staženého z webu, do kořene diskety je třeba nakopírovat obsah zabaleného adresáře .\W2K_driver\ .

Pozor, instalátor Windows 2000 nepozná disketu zformátovanou pod Windows95! Při formátování ovladačové diskety je třeba použít novější verzi Windows.

Alternativně lze použít volbu "Create driver disk" z menu, které otevře autorun.exe z CDčka od Adaptecu – ani tuto možnost nelze použít pod Windows95. Pokud tímto způsobem (create driver disk) potřebujete vytvořit disketu s ovladači pro Windows 2003 Server, použijte variantu „Windows 2000“ (nikoli Windows XP!)

Windows 2000 nepoužívají pro přístup k disku BIOS a neobsahují ovladače pro řadič Adaptec ASR2120S. Proto Windows při standardní instalaci pole nenajdou a nahlásí, že se nemají kam nainstalovat.

Je třeba instalátoru ve vhodný okamžik podstrčit disketu s ovladači - to se dělá takto:

- vložte instalační CD-ROM Windows2000 do mechaniky
- nastartujte počítač - pokud je správně nastaveno pořadí bootovacích zařízení v BIOS setupu, nastartuje z CDčka instalátor Windows2000
- jakmile se začne rozbíhat instalátor Windows2000, čekejte pozorně na výzvu ke stisku F6 (tato výzva po několika sekundách zmizí beze stopy). Stiskněte včas klávesu F6.
- po stisku F6 instalátor chvíli pracuje na disku a pak se zeptá na disketu s ovladači. Stiskněte "S", vložte disketu, stiskněte <ENTER> - a pokračujte v instalaci.

Instalace aplikací pro správu pole v systému Windows

Aplikace "Adaptec Storage Manager – Browser Edition" (ASM-BE) se nachází na CD od výrobce. Pro správnou funkci pole (včetně obnovy degradovaného pole po havárii disku) není tato aplikace striktně nutná – nicméně doporučujeme ji nainstalovat, protože poskytuje uživatelsky přívětivý přehled o stavu řadiče a polí.

Po vložení CD do mechaniky v systému Windows se pomocí funkce autorun automaticky spustí základní obrazovka instalátoru, ve které lze spustit vlastní instalaci (Install Adaptec Storage Manager...). Teprve po spuštění instalace je zcela zřejmé, že ve skutečnosti se nainstaluje jednoúčelový webserver. A nejen webserver - k dispozici je i SNMP služba pro automatizovaný



vzdálený dohled. Doporučujeme zvolit v počáteční fázi instalace “custom” skladbu, kde lze zvolit potřebné součásti. Celkem nakonec v systému běží až čtyři služby od Adaptec.

Po skončení instalace je potřeba provést restart.

Po restartu se již instalátor znovu nevynoří, v systému běží čtyři nové služby a zvědavý správce zjišťuje, kudy dál. Ve Start menu lze objevit novou složku SMBE a v ní dvě položky: aaccli a také zástupce na webové URL:

<https://localhost:3154/adaptec> (Zde je vidět, že webserver je zabezpečený pomocí SSL.)

Na zmíněném URL Vás přivítá klasický login dialog. Pro přihlášení použijte jméno a heslo, kterým se na tomto počítači hlásíte do systému Windows – tedy např. Administrator a příslušné heslo.

Vlevo se zobrazují fyzická zařízení (disky, řadiče), vpravo logické svazky (pole). Seznam zařízení se po startu zobrazí “sbalený”. Poklepem na šipečku vlevo v každém panelu lze rozbalit podrobnosti.

Pokud označíte nějaké zařízení, automaticky se zvýrazní zařízení k němu příslušející. Například pokud označíte pole, automaticky se zvýrazní disky, které do něj patří.

U každé z obou kategorií (fyzická vs. logická zařízení) se v hlavičce panelu zobrazují použitelné operace. Po zvolení operace se v protějším panelu rozběhne “průvodce”, který se (poněkud nadbytečně) dále vyptává.

Protože uživatelské rozhraní běží v browseru, vykazuje typické nedostatky tohoto typu rozhraní: pokud přidáte disk, musíte ještě v browseru stisknout “reload”, aby se zobrazil v příslušném panelu. Atd. Když v browseru běží počítadlo sekund uplynulých od začátku operace, není vůbec jasné, co hardware na pozadí dělá, zda vůbec něco dělá, zda náhodou nevytuhl. A není možné operaci zrušit.

Obnova degradovaného pole pomocí SMBE

Doporučený postup vypadá takto:

-> vložte nový disk

-> počkejte dvacet sekund, stiskněte v browseru “reload”. Pokud se nový disk nezobrazí, klikněte na tlačítko “rescan”

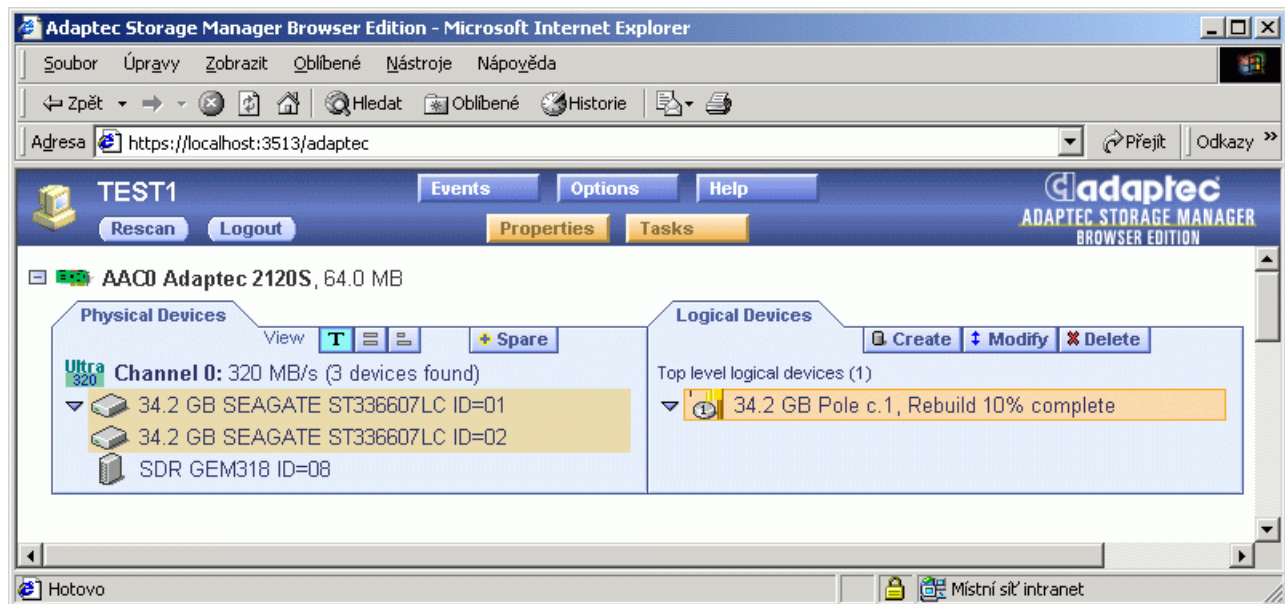
-> pokud se ukáže, že přidaný disk nese stopy konfigurace dalšího pole, smažte toto pole – označte ho na seznamu vpravo a stiskněte tlačítko “delete”.

-> pokud je disk “factory clean” nebo obsahuje standardní tabulku rozdělení (“legacy”), je třeba ho ještě inicializovat. Označte disk, klikněte na tlačítko “tasks”, zvolte záložku “new task”, vyberte typ úkolu “clear”, zvolte “start now” a klikněte “apply”. Ještě jednou potvrďte, že si přejete disk smazat. Na pozadí se rozběhne úkol “mazání disku” – v tabulce fyzických disků začnou naskakovat procenta. Řadič provádí úlohu typu “zero” čili nulování (což lze zjistit pomocí aaccli), nicméně nakonec se provede také inicializace, tj. přepsání režijních datových struktur do výchozí podoby.

-> v panelu vlevo stiskněte tlačítko “spare+” – budete dotázáni, který disk se má použít jako náhradní. Pokud toto skončí chybou, je třeba disk napřed inicializovat – viz předchozí bod.

-> za několik sekund po přidání náhradního disku by se měla rozběhnout obnova.





Poznámka autora tohoto českého manuálu: grafické rozhraní SMBE při zkoušení tohoto jednoduchého postupu ukázalo jako těžkopádné. Především neumožňuje inicializovat disk bez zdlouhavého nulování. Kromě toho Storage Manager při každé operaci nesmyslně dlouho čeká na výsledek a dokonce několikrát zatuhl při odstraňování starého pole.

Zdá se, že Adaptec nechal vyvinout SMBE jako jednotné prostředí pro správu RAID řadičů z několika rodin, které získal do svého portfolia vlastním vývojem a akvizicemi (FSA RAID, DPT RAID, HostRaid...). Každá tato rodina používá jiné nativní názvosloví a koncepty a umí jiné věci. Nejvíce možností má asi FSA RAID, jehož nejmladšími zástupci jsou právě ASR2120 a ASR2200. Proto také SMBE používá jinou terminologii než aaccli (nativní aplikace pro FSA RAID). SMBE je na pohled krásná aplikace, ale zdá se, že se příliš nepovedlo její provázání s nativním back-endem pro FSA RAID. To je daň za univerzálnost a nezávislost na hardwaru.

Z výše uvedených důvodů doporučujeme používat pro změny konfigurace příkazový řádek aaccli.

Linux

Softwarové vybavení pod Linuxem se skládá z následujících součástí:

- **ovladač v jádře** (povinný)
- device node **/dev/aac0** (nepovinný - je zapotřebí pro správčovské utility)
- **aaccli** – správčovská utilita na bázi příkazového řádku (nepovinná)
- Storage Manager – správčovská utilita (nepovinná) – dva až čtyři démoni a nepovinný grafický front-end pro Xwindows (upravená Mozilla)

Informace o Linuxu uváděné v tomto dokumentu se týkají verzí jádra okolo 2.4.20 – řekněme 2.4.18 – 2.4.22 (v tuto chvíli nejnovější jádro v řadě 2.4).

Instalační CD od Adaptecu obsahuje jakési multiplatformní instalační skripty – nicméně doporučujeme provést instalaci ručně, není složitá.

Ovladač

Ovladač se jmenuje “aacraid” a je součástí linuxového jádra již poměrně dlouho, protože hardwarové API “Adaptec FSA RAID” má poměrně dlouhou tradici. Řadiče ASR2120 a ASR2200 bohužel vykazují určité drobné hardwarové chyby v oblasti MMIO (PCI DMA), které mohou způsobit nefunkčnost se staršími verzemi ovladačů. Podle autorových zkušeností tyto řadiče fungují bez nejmenších problémů přinejmenším v RedHat Linuxu 8.0, tj. s jádrem 2.4.18-14, jakož i s novějšími jádry včetně “vanilkových”. Na webu Adaptecu jsou k dispozici backportované zdrojáky i binární RPM balíčky pro některé starší verze některých velkých distribucí.

Pokud si budete kompilovat jádro s upravenou konfigurací, ovladač pro AAC RAID se nachází v Menuconfigu ve větvi “SCSI support” -> ”SCSI low-level drivers”, tj. mezi SCSI řadiči.

Soubor zařízení (device node)

Standardní distribuce Linuxu, byť obsahují a instalují ovladač, bohužel obvykle při instalaci nevytvářejí device nody /dev/aac*, takže pokud chceme používat správčovské utility, musíme si device node vytvořit sami. Ovladač aacraid má bohužel dosud statut ”experimentální”, takže nemá úředně přiděleno stabilní ”major number” – patrně proto ho nezná ani skript /dev/MAKEDEV. Proto také budeme muset major a minor number zjistit jinak a ručně spustit program mknod s příslušnými parametry.

Experimentální zařízení si nechávají major number přidělit jádrem dynamicky. Dynamicky přidělovaná major numbers jsou alokována odshora počínaje číslem 254 – zařízení aac obvykle dostane právě toto číslo. Jedná se o ”znakové” zařízení (character class device). Třídou a major number zařízení, jejichž ovladače jsou momentálně aktivní, lze zjistit z pseudosouboru /proc/devices, třeba příkazem

```
cat /proc/devices | less
```



Linuxové zařízení pro ovládání řadiče ASR2120 se jmenuje "aac". Každý řadič instalovaný v systému dostane své vlastní minor number. Minor numbers jsou přidělována po jednom od nuly nahoru: /dev/aac0 = minor 0, /dev/aac1 = minor 1 atd.

Čili v typickém případě použijeme následující příkaz:

```
mknod /dev/aac0 c 254 0
```

Instalace aaccli

Správcovská utilita aaccli je k dispozici pouze v binární podobě. Lze ji nalézt na CDrom disku s ovladači, který Adaptec dodává s každým řadičem, případně stáhnout v podobě RPM z webu <http://www.adaptec.com>.

Pokud nepotřebujete Storage Manager, doporučujeme provést instalaci aaccli do systému ručně – například takto:

```
mount /mnt/cdrom
cp /mnt/cdrom/bootcd/sbin/aaccli /usr/sbin
chmod 700 /usr/sbin/aaccli
umount /mnt/cdrom
```

Storage Manager

Adaptec Storage Manager – Browser Edition existuje i pro Linux. Vlastní Storage Manager tvoří dva až čtyři démoni. Dodává se s upravenou Mozillou, ve které se zobrazuje uživatelské rozhraní. K samotnému běhu démonů (a vzdálený přístup přes HTTP) ovšem nejsou Xwindows na spravovaném počítači potřeba.

Na první pohled zajímavý je démonek "anotifyd", který umí posílat alarmy mailem. Konfigurační soubor je ve formátu XML, program ovšem může být spuštěn s parametrem -c, který způsobí spuštění v konfiguračním režimu a umožní zařídit konfiguraci přítulnějším způsobem než je editace XML kódu s pouze tušenou sémantikou.

Člověka hned napadne, jestli by se démonek anotifyd nedal použít samostatně. K jeho fungování je ovšem zapotřebí ještě démonek aiomgrd, a co je horší, také HTTPS démon arcpd. Démonek anotifyd totiž nekomunikuje přes interní API rovnou s démonem aiomgrd (natožpak rovnou s řadičem), ale připojuje se na uživatelský TCP port HTTPS démona (3513). Což znamená, že chceme-li upozornění mailem, musíme nainstalovat trio démonů, kteří potřebují SSL, aby si spolu mohli hrát. To je kvůli upozornění mailem docela hromada softwaru.

Automatizovaná instalace celého balíku se spouští skriptem install.sh, který se nachází v kořenovém adresáři CD. Instalace je šitá na míru distribuci RedHat, kde proběhne bez problémů (pod Xwindows i bez nich) – v jiných distribucích může být zapotřebí jistá improvizace.

Webové rozhraní SMBE chce login – použijte nějaký účet založený v hostitelském systému, např. root a rootovské heslo. Pokud se dostanete až sem, řiďte se dále návodem k verzi pro Windows.



Linux – shrnutí

Jedná se o jeden z nejvýkonnějších a nejkomfortnějších PCI SCSI RAID řadičů, které jsou pod Linux k dispozici. Což prakticky znamená, že řadič bez problémů funguje a pole lze spravovat přímo z operačního systému – tj. i na dálku. Systém nashartuje bez problémů i z degradovaného pole. Instalovat lze i na čerstvě vytvořené pole, které se na pozadí teprve vytváří.

Instalace softwaru Adaptec Storage Manager – Browser Edition je díky standardnímu instalačnímu skriptu jednoduchá (RedHat 9) a funguje v Xwindows i bez nich. Storage Manager má ale jisté nevýhody: zakrývá podstatu věcí, je lehce nepřehledný a hrozně pomalý. Linuxovému správci nelze než doporučit příkazový řádek aaccli, který lze bez problémů nainstalovat rukama a funguje poměrně slušně.



FreeBSD

Softwarové vybavení pod FreeBSD se skládá z následujících součástí:

- **ovladač v jádře** (povinný)
- device node **/dev/aac0** (nepovinný - je zapotřebí pro správčovské utility)
- **emulace Linuxu** (přesněji Linux ABI) a změna konfigurace jádra (+= aac_linux.o)
(tento bod je nepovinný - je zapotřebí pro správčovské utility)
- **aaccli** – správčovská utilita na bázi příkazového řádku (nepovinná)
- Storage Manager – správčovská utilita (nepovinná) – dva až čtyři démoni a nepovinný grafický front-end pro Xwindows (upravená Mozilla)

Informace o FreeBSD uváděné v tomto dokumentu se týkají verzí 4.7 – 4.9-RC1 a 5.0 – 5.1.

FreeBSD není výrobcem řadiče (Adaptec) přímo podporováno – na instalačním CD není po FreeBSD ani stopy. Na webu Adaptecu se vyskytuje starší ovladač pro FreeBSD 4.4 a binární podoba CLI utility pro tutéž verzi, oboje lze najít v podpoře pro řadič ASR5400S.

Nicméně u Adaptecu má vazby jeden ze členů FreeBSD týmu jménem Scott Long, který ovladač “aac” udržuje ve FreeBSD 4 i ve FreeBSD 5. Díky němu fungují ovladače i aaccli (které pravda není udržováno open-source komunitou, protože nebyly zveřejněny zdrojové kódy, a ani Scott Long se mu nevěnuje).

Ovladač

Ovladač se jmenuje “aac” a je součástí jádra FreeBSD již poměrně dlouho, protože hardwarové API “Adaptec FSA RAID” má poměrně dlouhou tradici. V konfiguračním souboru je tedy třeba odkomentovat

```
device aac
```

plus samozřejmě “device scbus”, “device sd” a případné další relevantní SCSI součástky.

Řadiče ASR2120 a ASR2200 bohužel vykazují oproti starším FSA RAIDům určité drobné hardwarové odlišnosti (údajně chyby) v oblasti MMIO (PCI DMA), které způsobují naprostou nefunkčnost těchto karet ve FreeBSD 4.7 a nižších (s původním aac ovladačem).

Ve FreeBSD 5.0/5.1 a ve FreeBSD 4.8 ovladače v zásadě fungovaly, vykazovaly ovšem několik drobných i zásadnějších much, které se projevovaly v kritických situacích, tj. při degradaci pole.

- ve FreeBSD 5.0 a ve FreeBSD 4.8-pre je ve standardní distribuci zapnut ovladač (device) “aacp”, tj. AAC PassThrough, který umožňuje operačnímu systému vidět zařízení na “privátní” SCSI sběrnici RAIDu. Tento ovladač při své inicializaci resetuje “privátní” SCSI sběrnici, čímž často rozhodí fungování RAIDu natolik, že se řadič už nevzpamatuje (a systém nenastartuje) – toto platí zejména v případě, kdy se obnovuje degradované pole. Typicky se tato závada projevuje při instalaci na čerstvě vytvořené pole, které se ještě nestihlo “dobudovat”.

Ve verzích 4.8-RELEASE a 5.1 byla tato chyba odstraněna, instalátorový i GENERIC kernel mají “device aacp” standardně vypnutý.



- ve FreeBSD 4.8 a 5.0/5.1 se projevoval jakýsi další problém v okamžiku degradace pole (“zhavaroval disk”) a při bootu z degradovaného či obnovujícího se pole. Problém způsobuje zatuhnutí počítače – je potřeba restartovat a obnovit pole pomocí BIOSové utility, pak teprve lze opět zprovoznit operační systém. Konkrétně ve 4.8 se problém neprojevuje v GENERIC kernelu (uniprocessorový, zapnuto “device isp”) – projeví se ve chvíli, kdy zapneme SMP+APIC_IO nebo vypneme “device isp” (případně obojí). Na nejnižší úrovni je za těchto podmínek dohodnuta jiná adresa pro MMIO, se kterou má řadič v nouzovém režimu patrně problémy.
Tento problém byl odstraněn ve FreeBSD 4.9-RC1 a novějších 5.2-PRE.

Při správné konfiguraci kernelu (aacp) funguje řadič dnes již bezchybně, problémy v degradovaném režimu se v nejnovějších RELEASEch již neprojevují – autor tohoto manuálu důkladně trápil řadič ve FreeBSD 4.9, 5.2 a 5.2.1 a reportoval výsledky Scottu Longovi. Ani na 4.8/5.1 by při degradaci pole nemělo dojít k závažnější ztrátě dat – po obnově pole v BIOSu (nebo v GENERIC/UP režimu) a po fsck lze pokračovat v normálním provozu.

Soubor zařízení (device node)

Standardní instalace FreeBSD nevytváří device nody pro AAC RAID, takže pokud

Standardní instalace FreeBSD, byť obsahuje a instaluje ovladač, bohužel při instalaci nevytváří device nody /dev/aac*, takže pokud chceme používat správčovské utility, musíme si device node vytvořit sami. Lze použít skript /dev/MAKEDEV, který tomuto zařízení přiřadí major number 150:

```
cd /dev
./MAKEDEV aac0
```

Jinak též

```
mkknod /dev/aac0 c 150 0
```

Každý řadič instalovaný v systému dostane své vlastní minor number. Minor numbers jsou přidělována po jednom od nuly nahoru: /dev/aac0 = minor 0, /dev/aac1 = minor 1 atd.

Emulace Linuxu

Aktuální verze správčovských utilit není k dispozici pro FreeBSD – ale je k dispozici pro Linux. Proto lze s úspěchem provozovat aktuální verze softwaru pod “linuxovým emulátorem” – což je nesprávné označení pro tzv. Linux ABI (linuxové nativní aplikační rozhraní k jádru FreeBSD).

Instalaci podpory pro linuxové binární programy lze povolit již při instalaci FreeBSD standardním instalátorem – v pokročilejší fázi instalace se sysinstall ptá na tuto možnost samostatným dotazem (okénkem). Touto volbou se nainstalují linuxové knihovny a některé systémové programy do sandboxu /compat/linux/ a zajistí se natahování linuxového ABI do jádra při startu (ve formě modulu linux.ko).

Dále je třeba zařídit, aby se linuxová IOCTL volání dostala až k ovladačům “aac” a jejich prostřednictvím k řadiči. K tomu je bohužel zapotřebí ještě několik úderů do klávesnice. Konkrétně je třeba upravit konfiguraci kernelu (a překompilovat). Do konfiguračního souboru je třeba přidat volby




```
options AAC_COMPAT_LINUX      # pouze FreeBSD 4.x, nikoli FreeBSD 5.x
options COMPAT_LINUX
```

Jedině obě tyto volby dohromady způsobí, že se zkompileje objektový soubor /usr/src/sys/dev/aac/aac_linux.o a přilinkuje se do ovladače aac a potažmo do kernelu.

Wrapper uložený v aac_linux.o se registruje do Linux ABI – ošetřuje překlad linuxového formátu IOCTL syscallů pro zařízení aac do BSD formátu. Vlastní IOCTL kódy (příkazy) zařízení aac se mezi Linuxem a FreeBSD neliší.

Hlavním posláním volby COMPAT_LINUX je mimochodem monolitické linkování Linux ABI do kernelu (takže není třeba jej při bootu natahovat rc skriptem jako modul).

Podrobnější informace o Linux ABI naleznete v dokumentaci FreeBSD (např. Handbook).

Instalace aaccli

Správcovská utilita aaccli je k dispozici pouze v binární podobě.

Na webu Adaptecu existuje starší verze aaccli zkompilevaná pro FreeBSD 4.4 – lze ji stáhnout ze stránek s ovladači pro Adaptec 5400S. Pro tuto verzi není zapotřebí Linux ABI a běhá pod 4.x i pod 5.x, nicméně Scott Long a další lidé tvrdí, že tato verze aaccli je zastaralá a potenciálně nekompatibilní s novějšími kontroléry.

Aktuální verze aaccli je k dispozici pouze pro Linux. Lze ji nalézt na CDrom disku s ovladači, který Adaptec dodává s každým řadičem, případně stáhnout v podobě RPM z webu <http://www.adaptec.com>.

Linuxová verze očekává knihovnu ncurses v5, ve standardním “linuxovém emulátoru” je verze 4 – nejsnazším řešením je zkopírovat knihovnu z CDčka pod pseudo-chroot /compat/linux.

Doporučujeme použít Linuxovou verzi aaccli. Instalaci do systému lze provést například takto:

```
mount /cdrom
cp /mnt/cdrom/bootcd/sbin/aaccli /usr/local/sbin
chmod 700 /usr/local/sbin/aaccli
cp /mnt/cdrom/bootcd/usr/lib/libncurses.* /compat/linux/usr/lib
umount /cdrom
```

Storage Manager

Linuxový Storage Manager by se zřejmě také dal použít. Zejména démonek “anotifyd” je zajímavý, protože umí posílat alarmy mailem. Nicméně i kvůli tomuto rozesílání mailů je třeba mít nainstalovaný kompletní SMBE včetně HTTPS démona (SSL). Rozběhávat tohle rukama v emulaci může být nemilé, zejména device nody pro sdílenou paměť nebudí důvěru. Že jsou instalační balíky ve formátu RPM, to by nebyl až takový problém – instalačnímu skriptu se BSD nelíbí a nejjednodušší způsob instalace je zřejmě beztak ručním kopírováním součástí z Linuxové instalace, dokud to nebude chodit. Alternativou je instalace RPM do FreeBSD a úprava skriptu, aby nekontroloval platformu a použil linuxovou instalaci.

Samotný grafický Storage Manager s jeho nepovedeným GUI ovšem za tu námahu patrně nestojí.

Sečteno a podtrženo: autor manuálu tuto aplikaci nezkoušel pod FreeBSD instalovat.



FreeBSD – shrnutí

Jedná se o jeden z nejvýkonnějších a nejkomfortnějších PCI SCSI RAID řadičů, které jsou pod FreeBSD k dispozici. Což prakticky znamená, že řadič bez problémů funguje a pole lze spravovat přímo z operačního systému – tj. i na dálku. Systém nashartuje bez problémů i z degradovaného pole. Instalovat lze i na čerstvě vytvořené pole, které se na pozadí teprve vytváří.

Utilita aaccli funguje (nejlépe je použít aktuální linuxovou verzi), grafický webový Storage Manager patrně nestojí za námahu.

