

vpsAdminOS

Protože [OpenVZ](#) už pomalu dosluhuje a nové distribuce jej nepodporují, museli jsme začít řešit přechod na nějakou novější virtualizační technologii. Linux kernel už sám o sobě kontejnery do jisté míry podporuje, takže jsme se rozhodli toho využít. Dále jsme potřebovali nějakou distribuci, kterou bychom použili na nodech místo Scientific Linux 6 s OpenVZ kernelem. Volba padla na [NixOS](#), který umožňuje deklarativně definovat konfiguraci OS a služeb a pak jej opakovaně sestavit. A protože máme specifické nároky, stavíme si nad NixOS vlastní distribuci.

[vpsAdminOS](#) je založen na [NixOS](#) a [not-os](#). Je to *live* distribuce sloužící jako hypervizor pro provoz kontejnerů. vpsAdminOS je funkčností srovnatelný s OpenVZ Legacy. Jako základ pro provoz VPS (kontejnerů) slouží LXC, které spravujeme vlastní utilitou `osctl` z vpsAdminOS. vpsAdminOS umožňuje propojení s vpsAdminem, naším administračním rozhraním, nicméně je plně použitelný i bez něj a měla by to být plnohodnotná náhrada za OpenVZ Legacy, pokud jej někde používáte. Podporována je i [migrace kontejnerů](#) z OpenVZ Legacy na vpsAdminOS.

Migrace VPS z OpenVZ na vpsAdminOS

Přechod celé naší infrastruktury se všemi VPS na vpsAdminOS je rozdělen do několika fází:

1. Vývoj vpsAdminOS do použitelné podoby
2. Integrace s vpsAdminem
3. Spuštění testovacího prostředí s vpsAdminOS (🚨 zde se nacházíme 🚨)
 1. Testování, opravy chyb, doplnění chybějících funkcí, příprava na produkci
4. Nové produkční nody používají vpsAdminOS, nové VPS lze vytvářet jen na nich
5. Postupná migrace všech VPS z OpenVZ nodů na vpsAdminOS nody, jeden node po druhém
6. Pohádky je konec

Co to znamená pro členy

Snažíme se, aby migrace na vpsAdminOS byla bezproblémová, tj. aby se jednoho dne VPS vypl na OpenVZ nodu a spustil na vpsAdminOS, aniž by člen něco musel řešit. Nicméně, záleží na tom, co ve VPS provozujete. Proto všem doporučujeme vyzkoušet si VPS nad vpsAdminOS v [testovacím prostředí](#), abychom mohli případné nedostatky vyřešit a při následné migraci produkčních VPS se jim vyhnout.

Projděte si také změny chování [VPS](#) a [vpsAdminu](#).

Změny v chování VPS

User namespaces

VPS ve vpsAdminOS využívají *user namespace*, tzn. *root* ve VPS má UID 0, ale z pohledu hostitelského systému na nodu je to nějaké jiné číslo, např. 666000. Každý člen má pak přidělen svůj user namespace, což zvyšuje úroveň izolace - v případě nějaké chyby se útočník ani po úniku z kontejneru

na node nedostane k datům jiných členů.

Každý člen má přidělen user namespace o velikosti 524288 uživatelských ID. Tzn. ve VPS můžete využít UID/GID od 0 do 524287. Všechny VPS daného člena jsou umístěny do tohoto user namespace. V budoucnu přibude možnost si user namespace spravovat a nastavovat si vlastní mapování UID/GID, což umožní izolovat od sebe i VPS patřící jednomu členovi, případně vybrané UID/GID sdílet.

User namespace podstatně ovlivňují sdílení dat mezi VPS a NASem. Aktuálně **není možné** připojit NAS do VPS běžící na vpsAdminOS tak, aby měl VPS k datům přístup. Řešit se to bude tak, že u každého NAS datasetu budete mít na výběr, jaké mapování UID/GID má používat. Data pak budou přístupná jen z VPS, které mají nastaveno příslušné mapování.

Obecné

Změny týkající se VPS nezávisle na distribuci:

- /proc/cpuinfo a /proc/stat zobrazují všechny CPU, které má node k dispozici, využít však lze max 8 z nich (800% CPU)
- /proc/stat u CPU aktuálně reportuje jen user (včetně system) a idle ¹⁾
- Nastavený swap se nezobrazuje v /proc/meminfo ²⁾
- dmesg je zakázán, neboť není virtualizován

Debian/Ubuntu/Alpine

- Pro nastavení sítě při startu už není potřeba mít nainstalovaný ifconfig z net-tools, používá se ip z iproute2.
- /etc/network/interfaces.{head,tail} nejsou vkládány přímo do /etc/network/interfaces, ale načteny přes source, tzn. už neovlivňují podobu /etc/network/interfaces tak jako s vzctl.
- Pokud existuje adresář /etc/network/interfaces.d, jeho obsah je načten před /etc/network/interfaces.tail.

Změny chování vpsAdminu

- Reinstalace VPS na vpsAdminOS **nemá** subdatasety ani nevrací konfiguraci VPS do výchozího stavu, tzn. např. nastavení VPS features zůstává zachováno.
- V detailech VPS je možné změnit název síťového rozhraní, výchozí název zůstává venet0.

Testovací prostředí

Aby si všichni členové mohli vyzkoušet, jak se VPS nad vpsAdminOS chová, k dispozici je testovací prostředí, tzn. takový druhý playground node, na kterém si každý může vytvořit VPS. Ve formuláři na vytváření VPS stačí vybrat lokaci **Staging**.

Podmínky provozu jsou podobné jako pro [playground VPS](#), akorát to může být trochu divočejší, tj. nehlášené výpadky, restarty pokud potřebujeme něco aktualizovat. Každý má k dispozici 8 CPU, 4 GB

RAM, 120 GB disku, 4 veřejné IPv4 adresy, 32 IPv6 /64 adres a tyto prostředky lze rozdělit mezi 4 VPS.

Vytvořit můžete buď nové VPS, nebo si vyzkoušet naklonovat produkční VPS. Při klonování dojde k odstranění mountů, protože přístup k NASu ještě není dořešen, viz [user namespaces](#).

Podporované distribuce

- Alpine 3.6, 3.7
- Arch
- CentOS 7.5
- Debian 8, 9
- Fedora 27, 28
- Gentoo
- NixOS
- Ubuntu 16.04, 18.04

Ostatní distribuce

Pokud tvoje distribuce zatím není mezi podporovanými, můžeš nám pomoci ji zprovoznit, nebo nezbývá než počkat, až to za tebe udělá někdo jiný, viz [otevřené požadavky](#).

Šablony distribucí se vytvářejí skripty ve [vpsadminos-templates](#). Pokud tam tvá distribuce není, je potřeba ji zde přidat.

Dále je nutné vyřešit podporu dané distribuce ve vpsAdminOS tak, aby byl osctl schopen zvenku nastavovat síť a IP adresy. Stačí postupovat podle [návodu](#). Je potřeba implementovat podporu jak pro bridged veth, tak routed veth, viz [dokumentace vpsAdminOS](#) a [zdrojové kódy](#).

Features

Features lze zapínat/vypínat jednotlivě. Při jakékoliv změně dojde k restartu VPS.

Features	
Bridge	<input type="checkbox"/>
FUSE	<input type="checkbox"/>
iptables	<input type="checkbox"/>
NFS	<input type="checkbox"/>
PPP	<input type="checkbox"/>
TUN/TAP	<input type="checkbox"/>
VPS is restarted when features are changed.	
<input type="button" value="Go >>"/>	

- Docker (experimental) - Povolí zkušební podporu pro Docker.

- FUSE - „Filesystem in Userspace“ Povolí modul jádra, umožňující neprivilegovaným uživatelům vytvářet si vlastní souborové systémy.
- KVM - „Kernel-based Virtual Machine“ Povolí použití KVM, pro HW podporu virtualizace.
- LXC nesting - „Linux Containers“ Povolí vnořené kontejnery LXC.
- PPP - „Point-to-Point Protocol“ Povolí protokol používaný pro propojení dvou sítí po telefonní, případně ISDN lince.
- TUN/TAP - „TUN routing/TAP bridging“ Povolí vytváření virtuálních interface, které jsou pak bridgovány.

Doporučujeme nastavit jen features, které opravdu potřebujete.

Více o vpsAdminOS

- <https://vpsadminos.org>
- <https://github.com/vpsfreecz/vpsadminos>
- IRC chat.freenode.net #vpsadminos

Kam hlásit chyby a nápady

Podle vlastního uvážení:

- IRC: #vpsfree a #vpsadminos na chat.freenode.net
- podpora@vpsfree.cz
- vpsAdminOS issues: <https://github.com/vpsfreecz/vpsadminos/issues>
- vpsAdmin issues: <https://github.com/vpsfreecz/vpsadmin/issues>

¹⁾ <https://lists.vpsfree.cz/pipermail/community-list/2018-May/009666.html>

²⁾ mělo by stačit upravit `proc_meminfo_read()` z LXCFS:
<https://github.com/lxc/lxcfs/blob/master/bindings.c#L3174>

From:
<https://kb.vpsfree.cz/> - **Znalostní Báze**

Permanent link:
<https://kb.vpsfree.cz/navody/vps/vpsadminos?rev=1541268647>

Last update: **2018/11/03 19:10**