Přednáška: Problematika logování, systémy IDS a IPS

Bezpečnost ICT 2

Zdeněk Martinásek

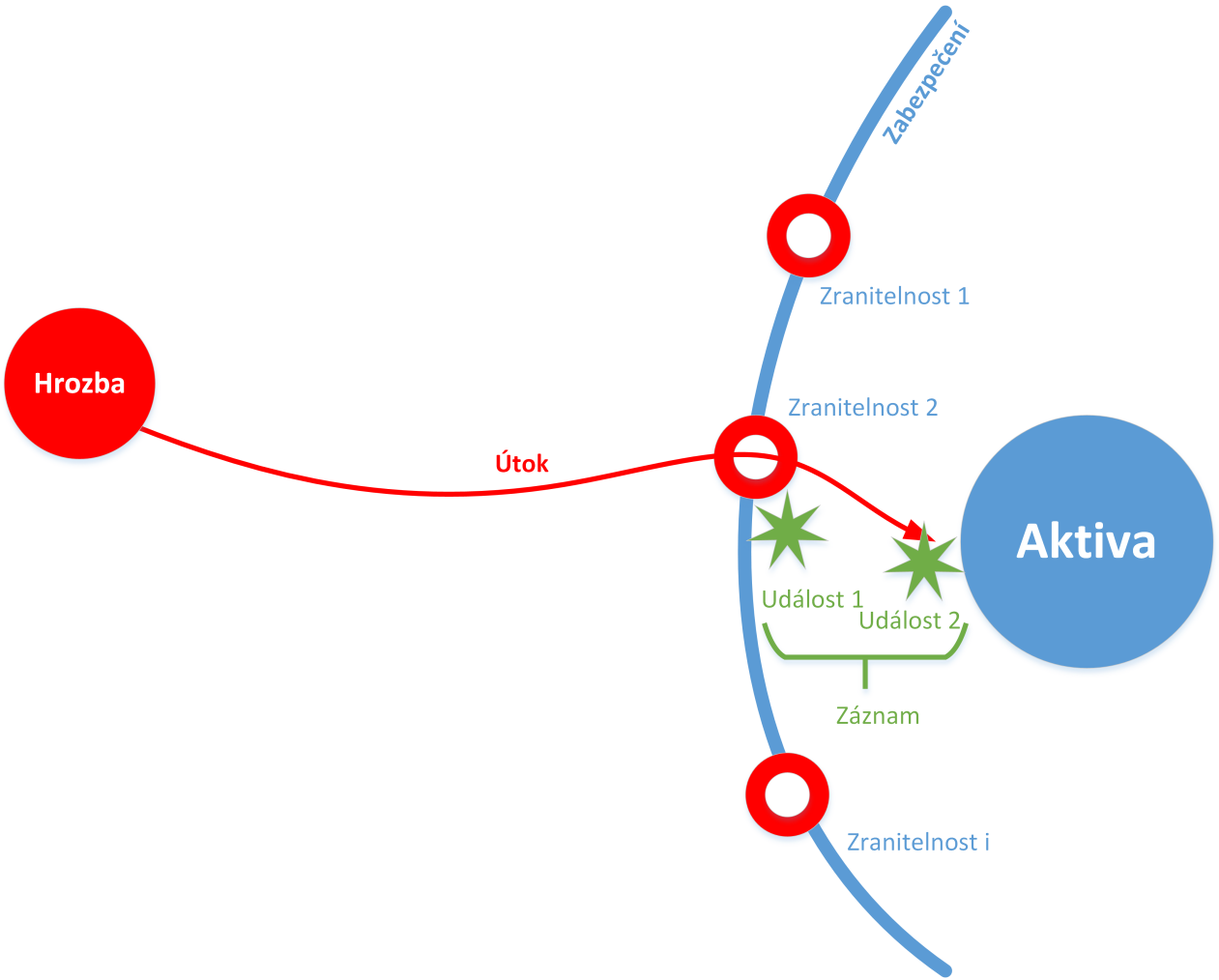
Vysoké učení technické v Brně
martinasek@vut.cz

2022

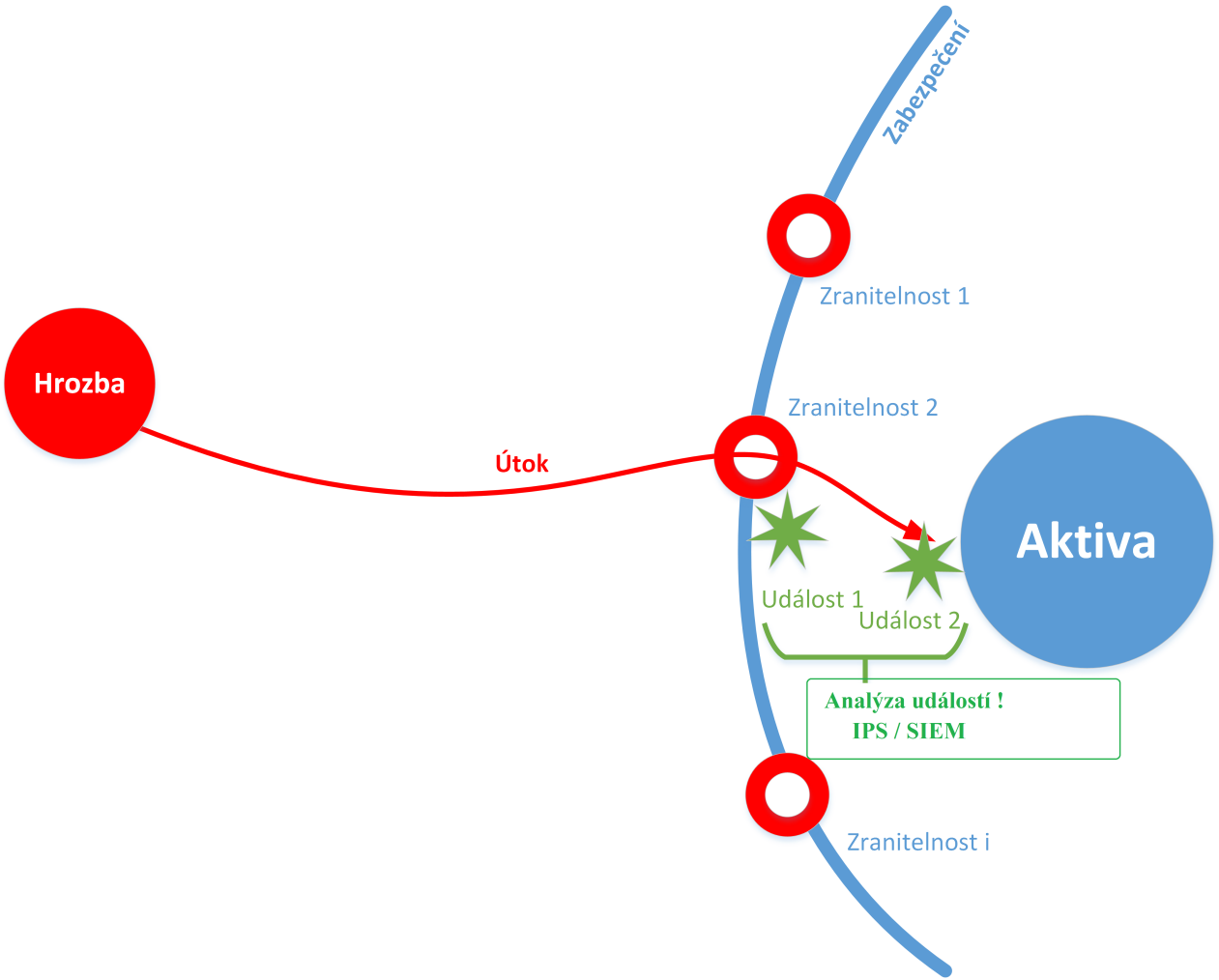
Terminologie

- **Průnik (Intrusions)** - realizace útoku při které došlo ke ztrátě aktiv, jde o využití slabiny v zabezpečení (získání neoprávněného přístupu nebo kompromitace utajení, integrity a dostupnosti).
- Průniky mají mnoho podob - malware (červy, spyware atd.), útočník pokoušející se o neautorizovaný přístup do systému z sítě Internet, autorizovaný útočník systému, který zneužije práva k získání vyššího oprávnění . . .
- Ze své podstaty je většina těchto průniků škodlivých (ohrožující bezpečnost), ale existují průniky **neškodlivé** - překlep v IP při použití RDP atd.

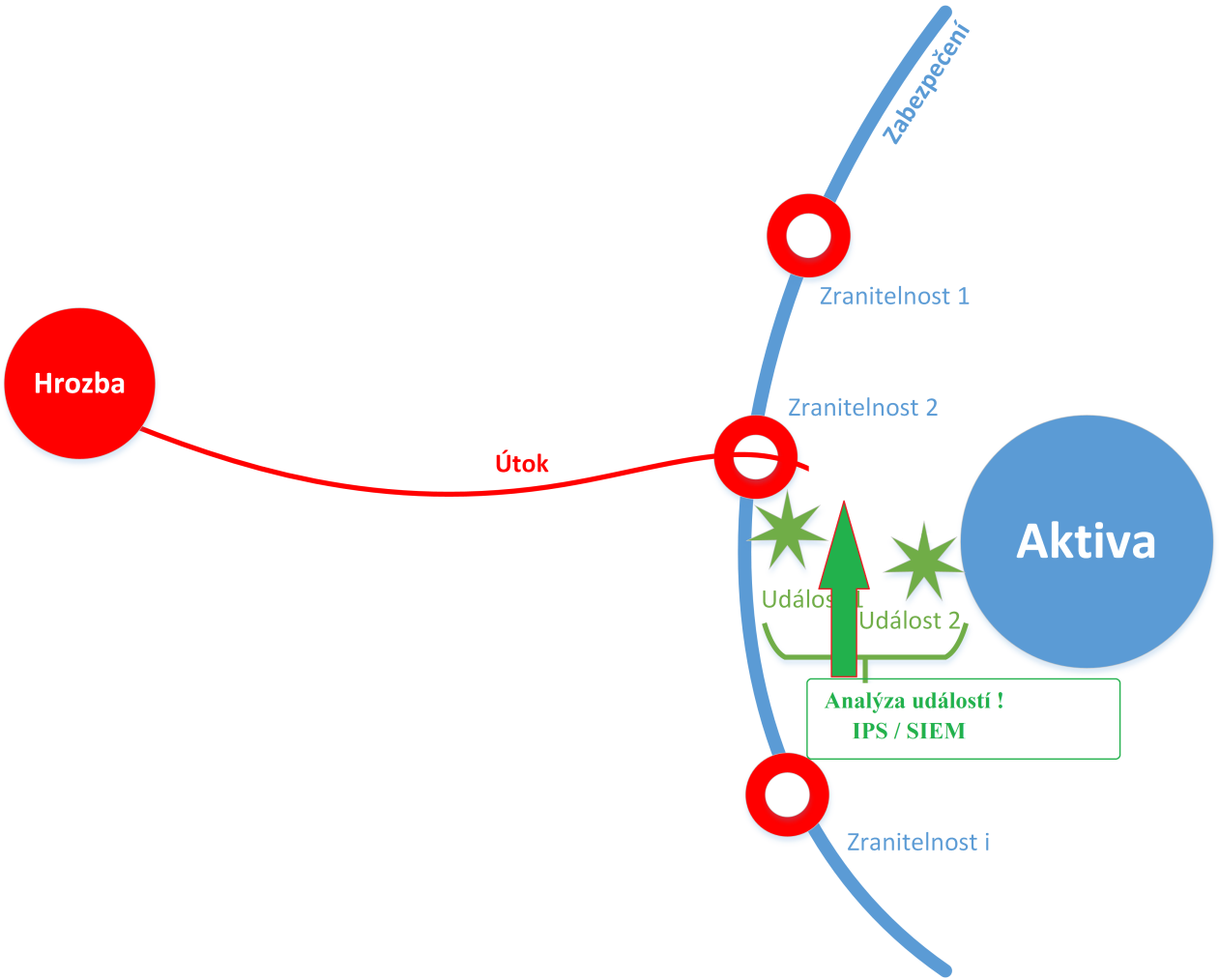
Terminologie - Průběh útoku (opakování)



Terminologie - Analýza logů



Terminologie - Výsledek analýzy logů



Terminologie

- **Detekce průniků** - proces monitorování událostí v počítačovém systému nebo síti a jejich následná analýza k detekci možných průniků.
- Výsledkem je vyhlášení **alarmu** a/nebo **aktivace ochrany** při pokusech a realizacích kybernetických útoků (incident).

Základní pojmy, terminologie

Záznamová data, logovací zpráva nebo jen *log* je soubor *záznamů* **reprezentující popis konkrétní události**, která nastala ve sledovaném systému.

Záznamy můžeme rozdělit na následující kategorie:

- **Informační**: pouze popisující stavy, události,
- **Ladící**: používané při vývoji,
- **Varovné**: označující chybějící funkci, součást systému, atp.,
- **Chybové**: označující chyby ohrožující funkčnost systému,
- **Pohotovostní**: často označující události spojené s bezpečností.

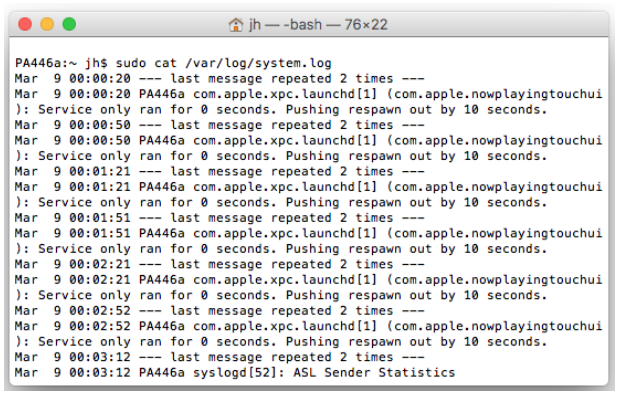
Formát logu

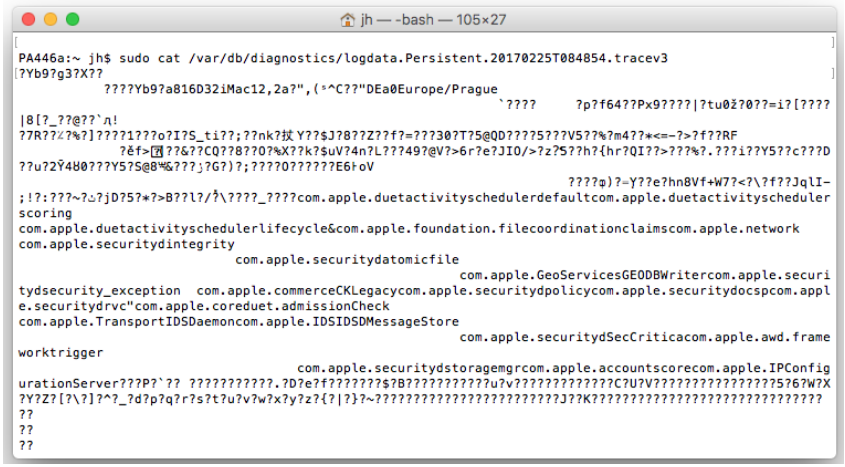
Textový formát logu:

- vytváření je nenáročné na systémové prostředky,
- lehce čitelné pomocí běžných nástrojů,
- společná syntaxe pro mnoho aplikací.

Binární formát logu:

- snadněji komprimovatelné,
- snadnější ukládání do databází, lepší optimalizace využití prostředků.





Obsah logu

- **Kdo** bude logy vytvářet?
- **Kde** se budou logy vytvářet?
  - Lokální ukládání v textovém souboru
  - Lokální ukládání v databázi
  - Vzdálené ukládání
  - Kombinace
- **Jaké události** budou vytvářet záznamy?
  - Povinné události: chyby při autentizaci, chyby systému, detekované anomálie, atp.
  - Volitelné události: nadměrné užití systému, změny zdrojového kódu, atp.
- **Co** budou záznamy obsahovat?
  - Povinné položky: čas, zdroj, uživatel, událost.
  - Volitelné položky: závažnost, návratová hodnota, atp.
  - Zakázané položky: hesla, klíče, atp.

Struktura záznamu v logu

Záznam se zpravidla skládá alespoň ze tří částí:

- **Časové razítko** – Udává datum a čas události.
- **Zdroj** – Reprezentuje název konkrétní komponenty systému, která záznam události vyvolala.
- **Vlastní data** – Zpráva obsahující popis události nebo informující o stavu problému.

Příklad záznamu v Cisco přepínači:

```
Jan 15 10:12:23 192.168.235.140 * 1 00:16:17:
%LINEPROTO-5-UPDOWN: Line protocol on Interface
FastEthernet0/0, changed state to up
```

Příklad záznamu událostí - SSH

Logování služby OpenSSH:

- zaznamenává do textových souborů,
- nejdůležitější auth.log ve `/var/log`,
- umístění logu a jeho podrobnost lze měnit v /etc/ssh/sshd_config.

Ukázka záznamu:

```
Nov 23 20:31:14 server sshd[15798]:  Failed password
for root from 188.124.3.41 port 32889 ssh2.
```

Příklad záznamu událostí - Sdílení souborů

Logování služby ProFTPd:

- zaznamenává do textových souborů pomocí syslogu,
- nejdůležitější TransferLog, SystemLog a ExtendedLog ve `/var/log`,
- umístění logu a jeho podrobnost lze měnit v /etc/proftpd/proftpd.conf.

Ukázka záznamu:

```
2017-03-09 14:46:06,862 debianServer proftpd[961]
debianServer (192.168.182.148[192.168.182.148]):
USER worker:  Login successful.
```

Příklad záznamu událostí - Sdílení souborů II

Logování služby SAMBA:

- zaznamenává do textových souborů,
- nejdůležitější smbd.log a nmbd.log ve `/usr/local/samba/var/`,
- umístění logu a jeho podrobnost lze měnit v /etc/samba/smb.conf.

Ukázka záznamu:

```
12/25/16 22:02:11 server (192.168.236.86) connect to
service public as user pcguest (uid=503,gid=100) (pid
3377).
```

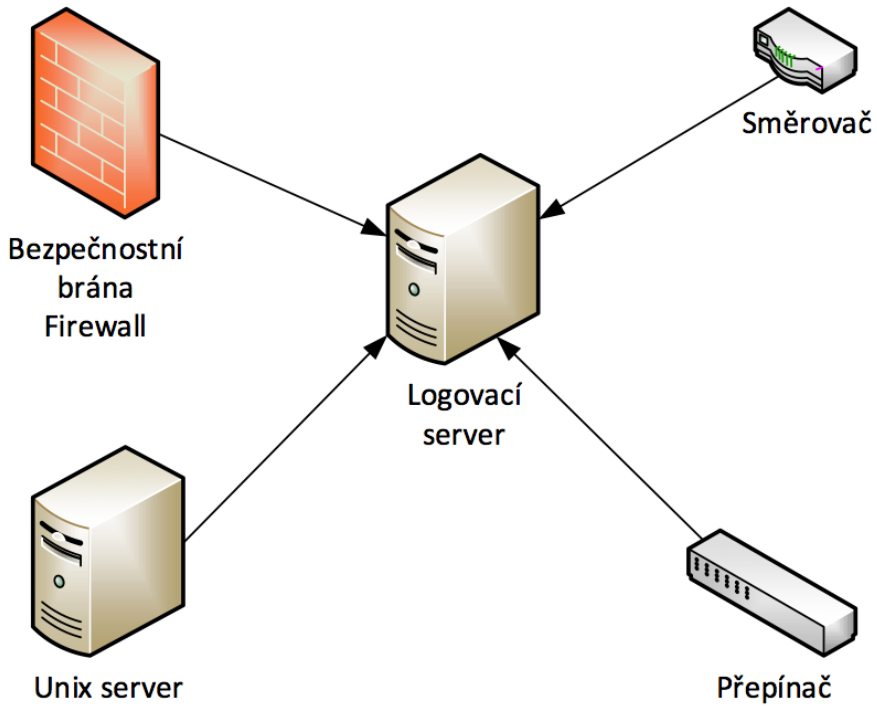
Agregace logů

V systémech většího rozsahu je nutné řešit agregaci logů a centrální správu. Možné přístupy:

- **Logovací server**: jednoduchá topologie pro malé podniky.
- **Centrální sběrné místo a logovací server**: varianta umožňující šifrování a redundanci.
- **Centrální sběrné místo, logovací servery a externí úložiště**: bezpečnější decentralizovaná varianta.

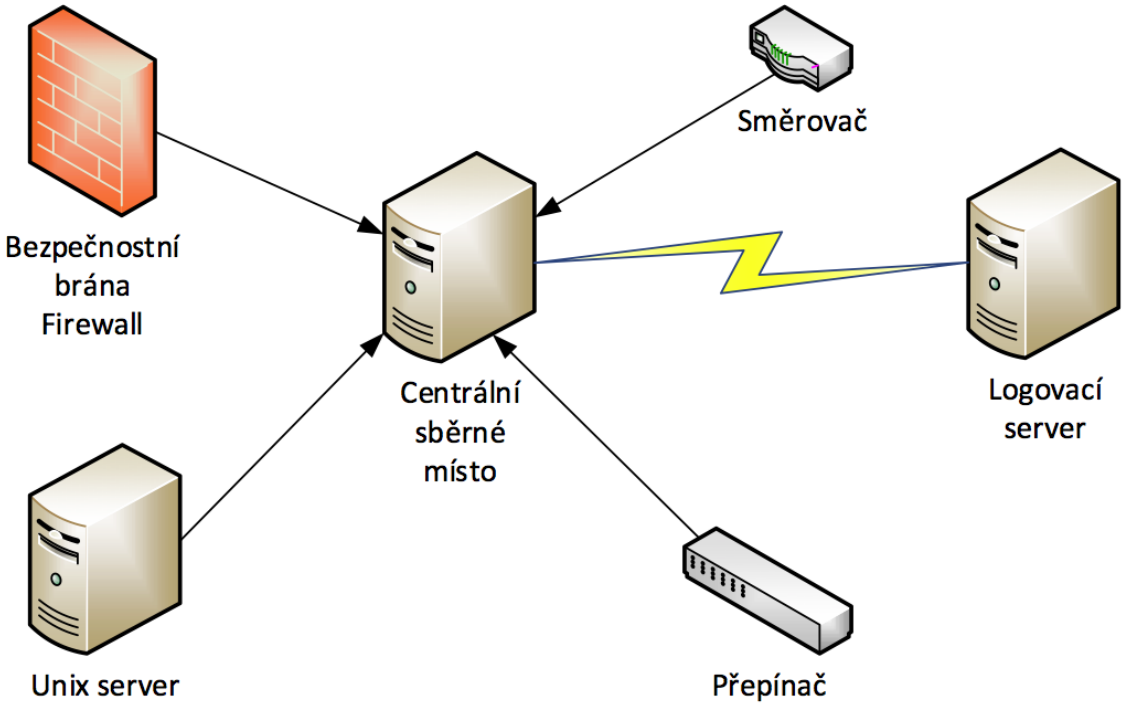
Agregace logů

Topologie s logovacím serverem.



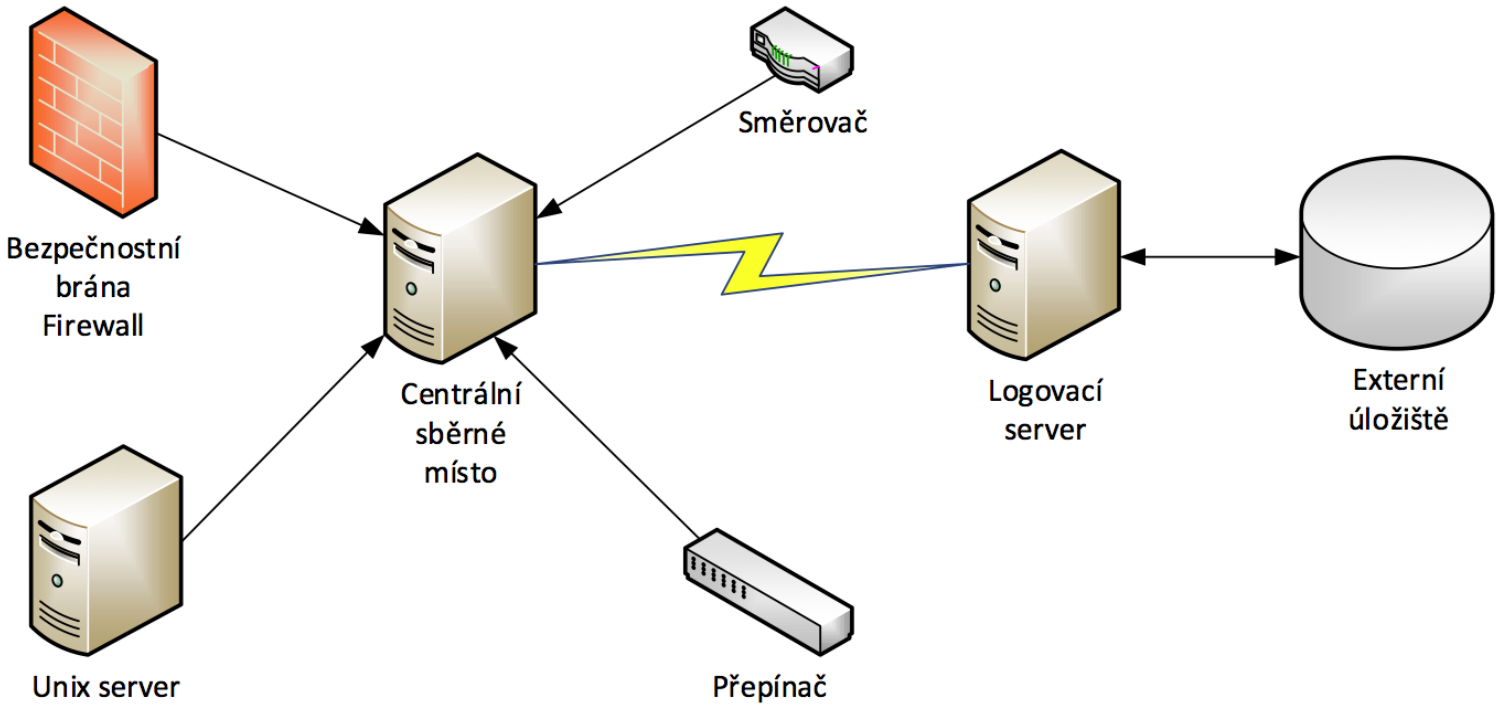
Agregace logů

Topologie s centrálním sběrným místem a logovacím serverem.



Agregace logů

Topologie s centrálním sběrným místem, logovacím serverem a externím úložištěm.



Zabezpečení logů

Soubory s logy je nutné chránit jak při ukládání, tak přenosu.

- **Uložení**: u ukládání jsou k dispozici standardní a volitelné mechanismy ochrany.
  - Standardní: **monitoring volného místa, detekce změn logovacích souborů, záloha s read-only přístupem, pokročilé řízení přístupu, striktní nastavení práv, rotace logů**.
  - Volitelné: šifrování souborů s logy, redundantní agregace, redundantní zálohy, zálohy mimo lokalitu, plánované mazání logů, archivace dle legislativních požadavků.
- **Přenos**: u přenosu logů je třeba dbát na zajištění důvěrnosti a autentičnosti dat, obvykle zajištěno pomocí mechanismů kryptografie.

Hlavní cíle analýzy logů

Cíle analýzy logů se mohou lišit v závislosti na potřebách konkrétního systému (banka, restaurace atd.).

Nicméně lze definovat dva obecné cíle analýzy logů:

- Detekce **známých** nepříznivých událostí, které se již staly.
- Detekce **neznámých** nepříznivých událostí, které zatím nikdy nebyly realizovány.

Detekce známých nepříznivých událostí

- Identifikace nepříznivé události, které se již staly v minulosti a na tyto události upozornit,
- následuje akce většinou v podobě mitigace útoku,
- z praktického hlediska **detekce již známých událostí není problematická**.

Ukázka záznamu CRS-32 útok na SSH:

```
Nov 1 18:48:15 victim sshd[9605]: fatal: Local:
Corrupted check bytes on input.
Nov 1 18:48:08 victim sshd[9588]: fatal: Local:
crc32 compensation attack: network attack detected
```

Detekce neznámých nepříznivých událostí

- Identifikace nepříznivých událostí, o kterých zatím nic nevím (nebyly viděny, možná jsou mimo detekční schopnost systému),
- tento proces je nutný k detekci neustále se vyvíjejících technik útoků,
- **není jednoduše realizovatelná** a využívá se **detekce anomálií** (statistické metody a strojové učení).

Před samotnou analýzou záznamů událostí by měli být splněny následující požadavky: věcná a časová přesnost, kontrola integrity, zajištění důvěrnosti.

Existující přístupy ke zpracování logů

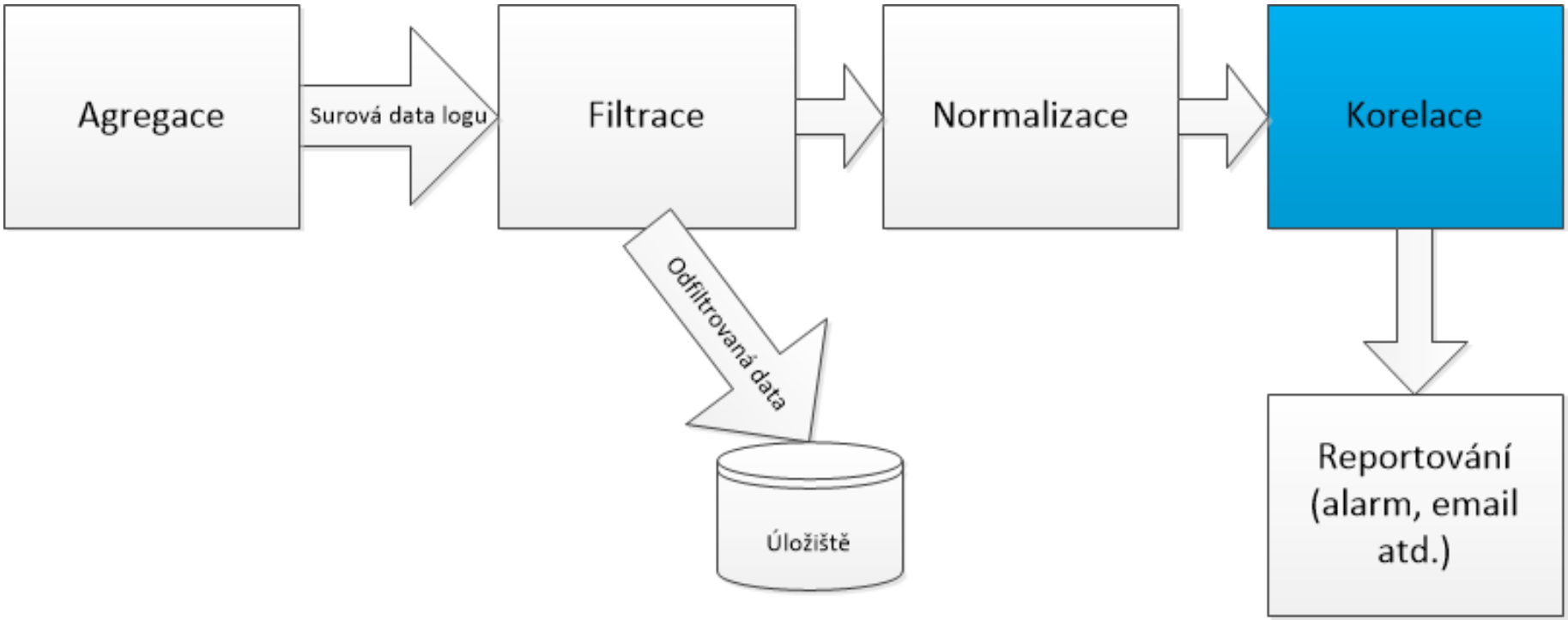
Jednoduché techniky analýz:

- Jde v podstatě o ruční analýzu záznamů (`tail, cat`, prohlížeč událostí atd.),
- vhodné pro malý počet záznamů (řádově desítky),
- aplikovatelné pro jednoduché a vypovídající záznamy (např. soubor nenalezen),
- nelze získat komplexní přehled o tom, co se děje v systému (unikají souvislosti mezi událostmi).

Automatické nástroje pro analýzu záznamů:

- Vhodné pro dnes používané informační systémy (velké množství záznamů),
- agregace, filtrace, normalizace, korelace a reportování.

Operace nutné k aplikaci automatické analýzy.



Operace nutné k aplikaci automatické analýzy.

- **Agregace** představuje proces stahování záznamů na jedno centrální místo.
- **Filtrace** je analýza surových dat v logu a rozhodnutí, která data jsou potřebná.
- **Normalizace** upravuje záznamy na společný formát (společná databáze, nad kterou bude probíhat analýza),
- V průběhu normalizace také dochází ke kategorizaci záznamu, kdy jsou některým záznamům přidány významy (např. výrobce vloží popis chyby jen ID = 75 což znamená chybné přihlášení“).
- Výše popsané operace jsou **snadno** realizovatelné.

Korelace událostí!

Korelace je spojení několika podobných nebo naprosto rozdílných událostí ve znalost o nějaké větší probíhající události, o které chceme být informování z pohledu bezpečnosti.

- Korelace představuje **nejkritičtější a nejvíce problematický blok**,
- bloky agregace, filtrace a normalizace mohou být jednoduše realizovány viz [4].

Vytváření korelací:

- na základě pravidel - **detekce signatur** (nepříznivé **známé události**),
- na základě modelu - **detekce anomálií** (nepříznivé **neznámé události**).

Detekce signatur

- Tato korelace může být realizována **vytvořením pravidla** (vzoru útoku) v nějakém **programovacím jazyce**.
- Příklad takto vytvořené korelace uvádí následující pseudokód:

```
 if nastane událost E1 == skenování portů;
která je následována událostí E2
E2 == odmítnutí požadavku FW ze zdrojové IP == E1;
udělej akci A1 == email;.
```

- Příklad ukazuje jak lze spojit dvě různé události z jiných systémů do jednoho pravidla (stavové diagramy systému).

Detekce signatur

Důležité problémy z praxe:

- Také SIEM (Security Information and Event Management) nepokrývají všechny potřeby organizace a velkou část korelací je nutné **doprogramovat**,
- představuje dodatečné náklady, protože **pravidla korelace se vkládají ručně**,
- proto společnosti po nákupu SIEM nástroje nedokáže plně využít jeho potenciál,
- problém **neschopnosti definice důležitých událostí** a korelací.

Detekce anomálií

Metody založeny na vytvoření modelů a sledování odchylek a podobností k těmto modelům:

- **Frekvenční model** - jednoduchá technika počítající výskyt definovaného jevu za pevně definovaný časový okamžik.
- **Referenční model** - vytvoření referenčního modelu normálního“ chování a sledování povolených odchylek.
- **Model strojového učení** - model je vytvořen a porovnáván s využitím algoritmů strojového učení.

Detekce anomálií - Referenční model

Referenční model:

- Základní myšlenka spočívá v definování modelu za určitý časový okamžik, který stanovuje normální fungování sledovaného systému,
- při běhu systému jsou sbírány data a porovnávány s modelem,
- **při odchýlení od normálního provozu je vyvolána akce**,
- k vytvoření modelu potřebujeme velké množství normalizovaných vzorových dat a **experta“**, který dokáže **rozlišit co je normální a škodlivé schování**,
- přesnost modelu je silně závislá na množství dat, ze kterých model vytváříme.

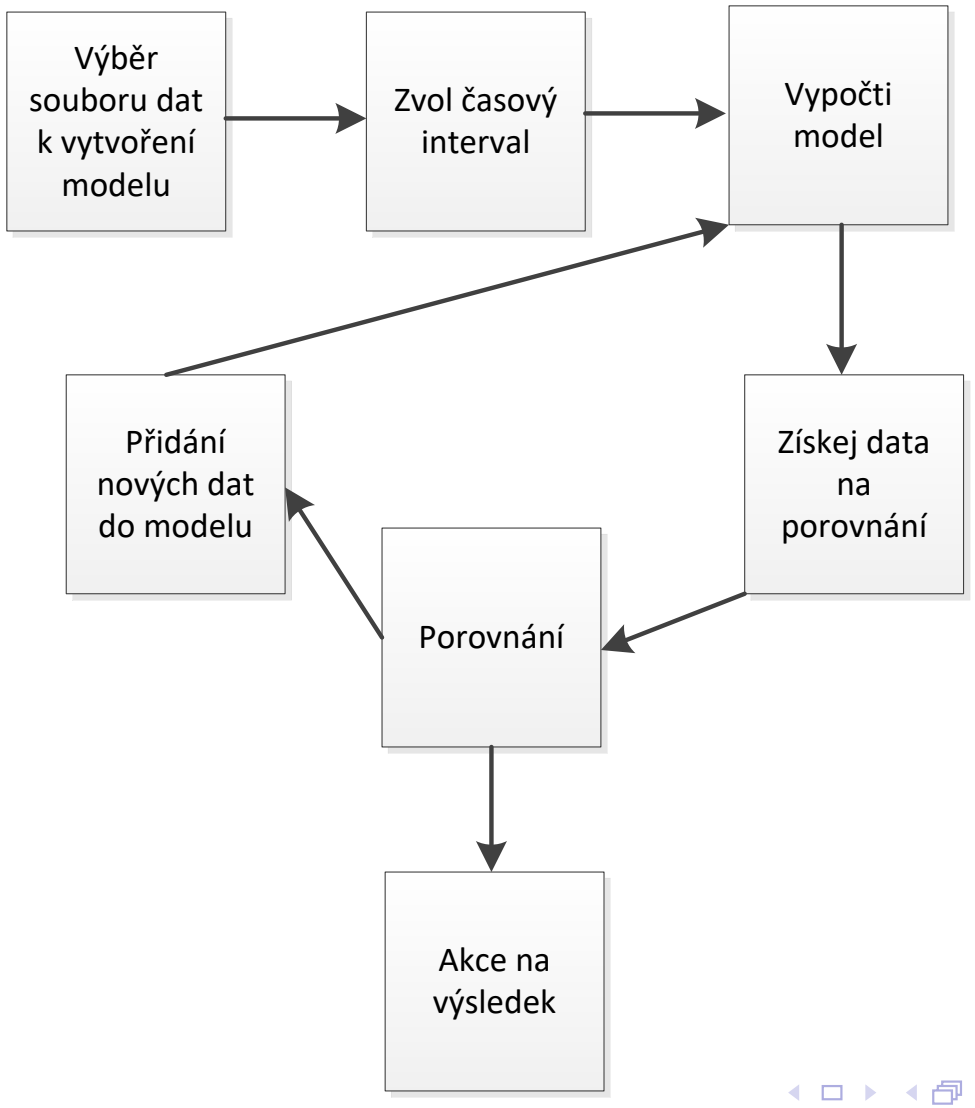
.

Detekce anomálií - Referenční model

K vypočítání modelu můžeme použít **standardní matematické vztahy** pro průměr a směrodatnou odchylku:

1. vypočítej průměr pro zvolený datový soubor ($\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$),
2. vypočítej směrodatnou odchylku ($\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})}{N}}$),
3. vypočítej standardní chybu ($e = \frac{\sigma}{N}$),
4. vynásob chybu hodnotou 1,96,
5. pomocí vypočtené hodnoty z bodu 4 urči interval spolehlivosti,
6. postupně aktualizuj model přidáváním nových dat do datového souboru.

Detekce anomálií - Referenční model



Detekce anomálií - Model strojového učení

Strojové učení se zabývá algoritmy a statistickými metodami, které umožňují změnu svého vnitřního stavu, která zefektivní schopnost přizpůsobení se změnám okolního prostředí (učení).

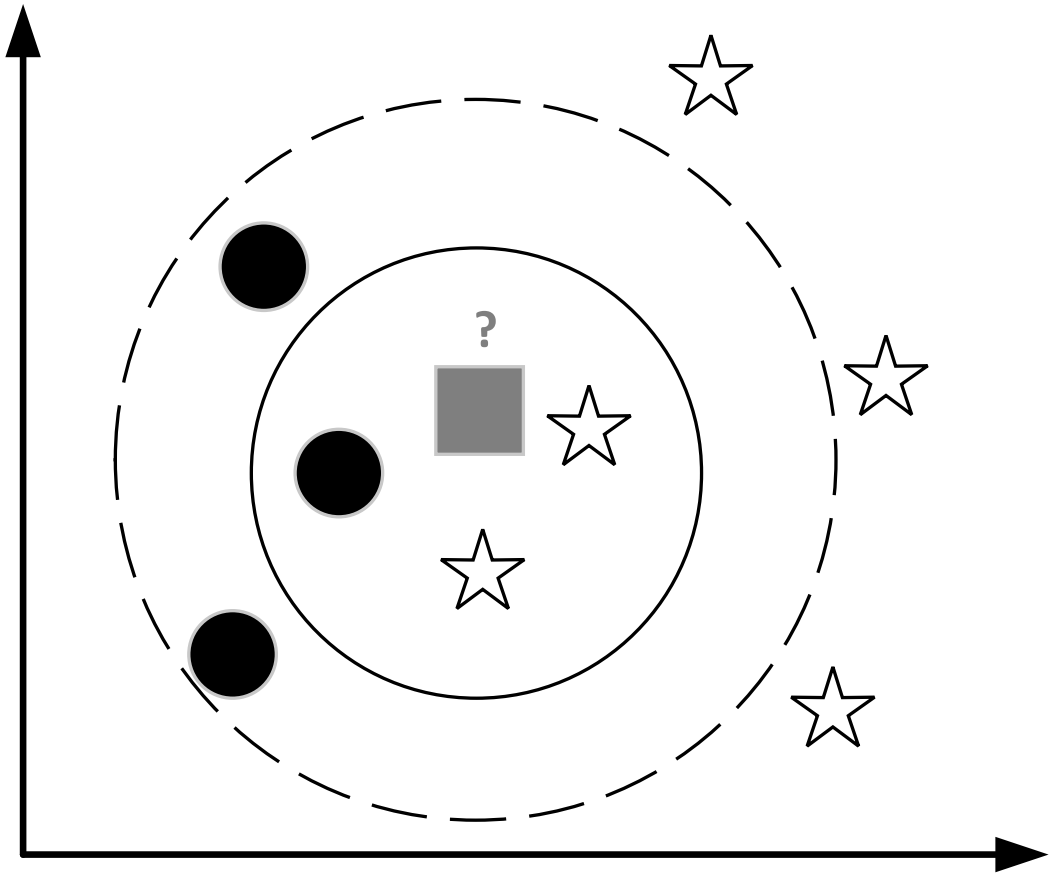
Základní druhy úloh pro strojové učení lze shrnout do následujících bodů:

- **klasifikace vstupních dat do tříd**,
- odhad číselné hodnoty výstupu podle vstupu (regrese),
- **shlukování dat s podobnými vlastnostmi**.

Detekce anomálií - Model strojového učení

- Jako příklad strojového učení byl vybrán **Algoritmus k-nejbližších sousedů (k-NN)**,
- nejjednodušších neparametrických metod používaných ke klasifikaci vzorů,
- aplikace ostatních algoritmů je obdobná s tím rozdílem, že vlastní klasifikace je na jiném principu,
- ve **fázi učení** jsou vzory pouze uloženy do paměti a označeny správnou třídou,
- přiřazení výsledné třídy dojde nalezením **k-nejbližších sousedů**.

Detekce anomálií - Model strojového učení



Detekce anomálií - Model strojového učení

- Základní princip aplikace - definování zajímavých událostí k různým vzorům útoku.
- Čím více zajímavých vlastností a vzorů, tím lépe bude model klasifikovat neznámé útoky,

Detekce anomálií - Model strojového učení

Stanovíme si následující zajímavé sledované vlastnosti:

1. nadměrný odchozí datový provoz (NOP),
2. nadměrný příchozí datový provoz (NPP),
3. FW zahazuje provoz (FWD),
4. FW akceptuje provoz (FWA),
5. VPN přihlášení mimo pracovní hodinu (VPNM),
6. více neúspěšných pokusů přihlášení (LF),
7. alespoň 1 úspěšný pokus o přihlášení (LA),
8. více cílů z jednoho zdroje (VCZ).

Detekce anomálií - Model strojového učení

Table: Vzory útoků a nalezené vlastnosti

Vzor	NOP	NPP	FWA	VPNM	LF	LA	VCZ
Útok hrubou silou			Ano		Ano	Ano	
Skenování portů		Ano					Ano
DDoS		Ano	Ano				
Ztráta dat	Ano		Ano	Ano			

Detekce anomálií - Model strojového učení

Table: Vstup Klasifikace *k*-NN algoritmem, k=3

Výsledek	NOP	NPP	FWA	VPNM	LF	LA	VCZ
??????	Ano		Ano		Ano	Ano	
??????	Ano		Ano	Ano			Ano

Výsledek Detekce anomálií - Model strojového učení

Table: Výsledek Klasifikace *k*-NN algoritmem, k=3

Výsledek	NOP	NPP	FWA	VPNM	LF	LA	VCZ
Útok hrubou silou	Ano		Ano		Ano	Ano	
Ztráta dat	Ano		Ano	Ano			Ano

Základní dělení a terminologie

- **Detekce průniků** - proces monitorování událostí v počítačovém systému nebo síti a jeho následná analýza k detekci možných průniků (vyhlášení **alarmu** při pokusech, incidentech atd.).
- **Intrusion detection system (IDS)**: je softwarové nebo hardwarové řešení, které automatizuje proces detekce průniků (hlavní úkol detekce průniků).
- **Intrusion prevention system (IPS)**: má všechny funkce IDS a navíc dokáže blokovat možné nežádoucí incidenty (např. konfigurací firewallu).

Vztah mezi IDS a IPS

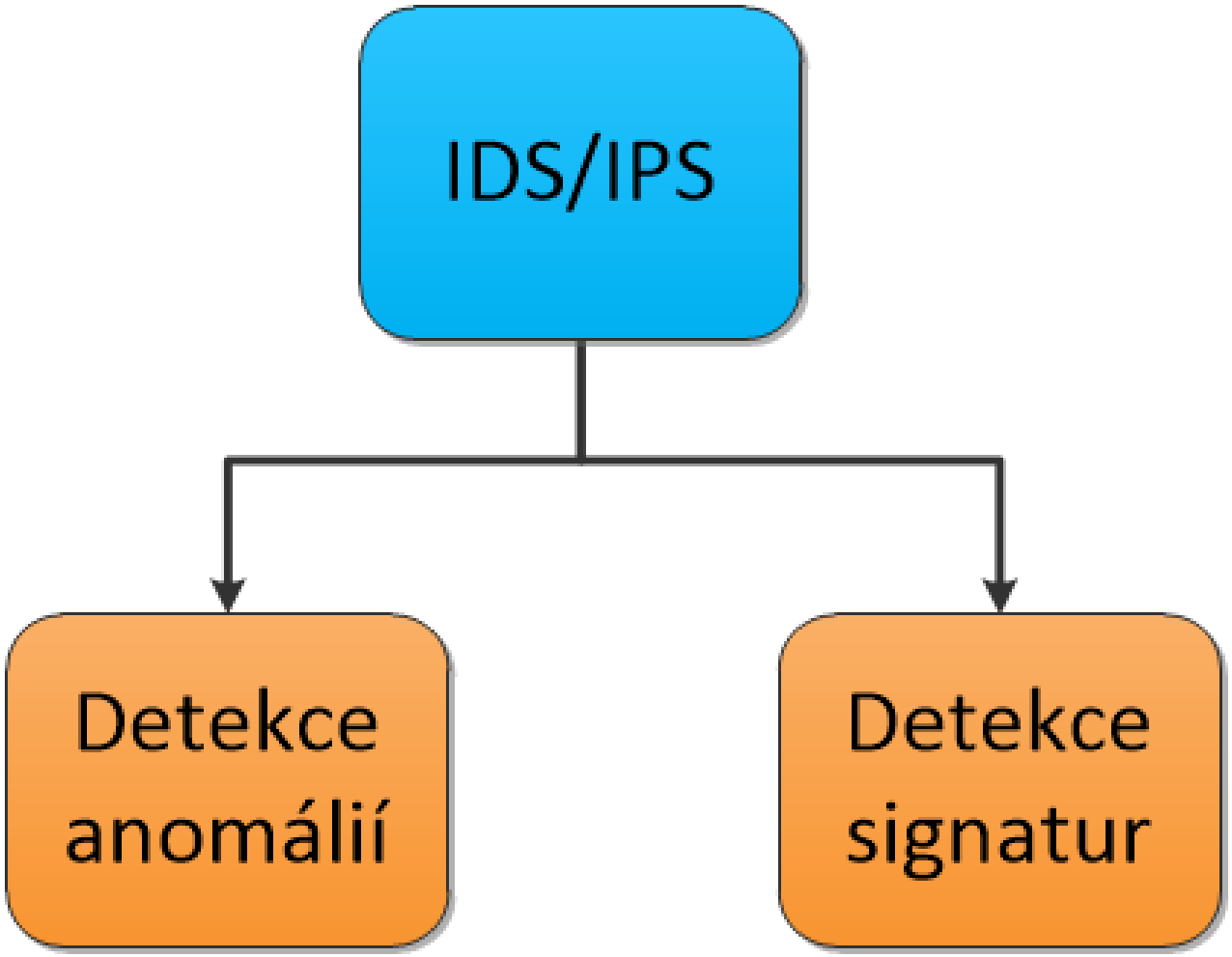
- **IDS nebo IPS**? Vzájemný vztah.
- Je zcela zřejmé, že z pohledu bezpečnosti je vhodnější blokovat možný pokus o průnik okamžitě (nedostatek IDS).
- Nemožnost okamžité blokace škodlivé síťové komunikace, vedly k vývoji nových systémů IPS.
- Technologie využití samotného IDS zastaralá, **IPS představuje prakticky vylepšenou IDS**, která dokáže na detekovanou událost reagovat (mitigace dopadu incidentu).

Hlavní cíl IDS/IPS systémů

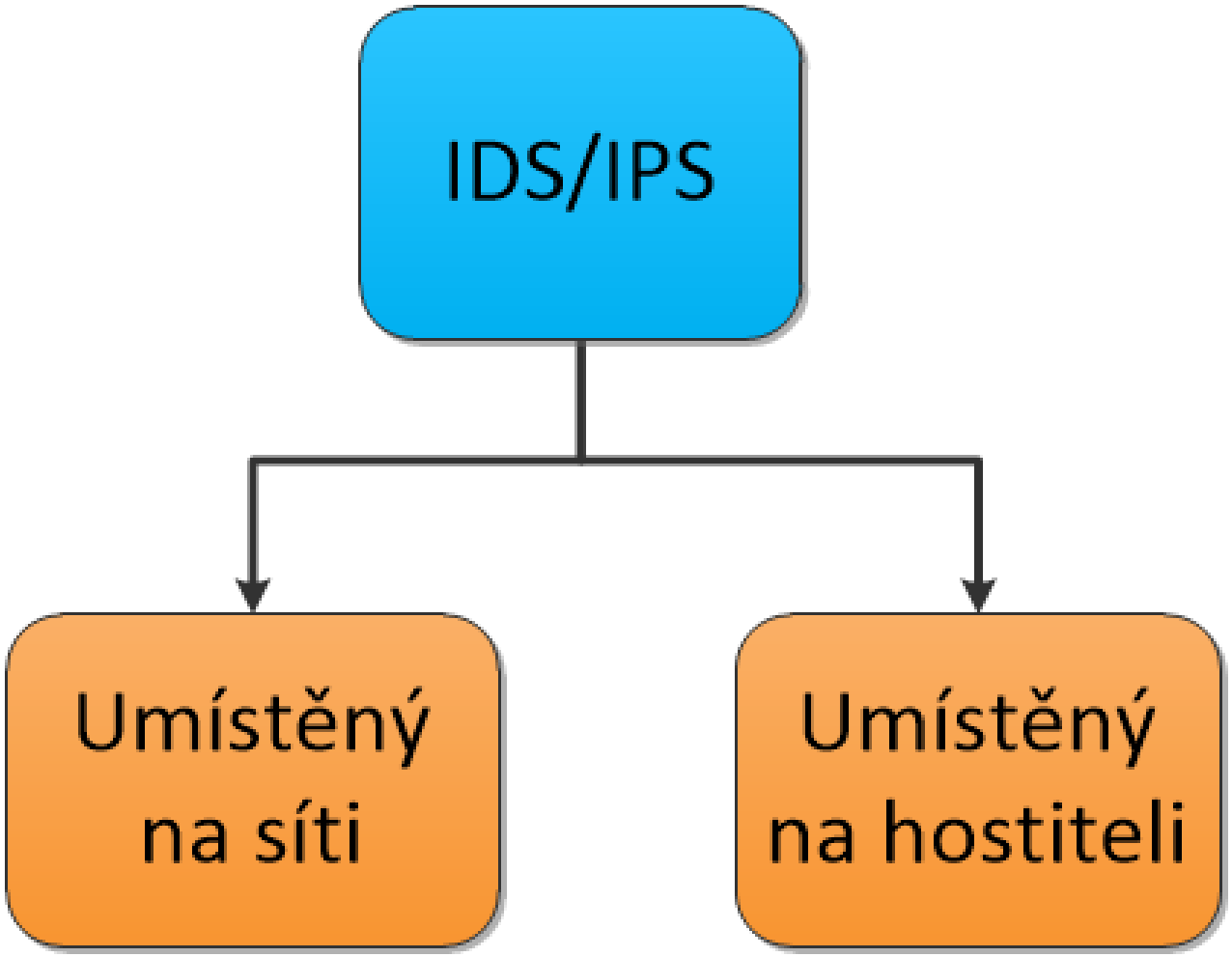
- Vyhlášení alarmu (popř. potlačení útoku) jen tehdy, pokud jde o skutečný incident.
- Vyvážit nastavení **falešné alarmy X funkčnost** (je IDS/IPS v provozu??)!!
- Ladění systému na požadovanou účinnost (tuning).

	POSITIVE	NEGATIVE
TRUE	True Positive: Alerted on intrusion attempt	True Negative: Not alerted on benign activity
FALSE	False Positive: Alerted on benign activity	False Negative: Not alerted on intrusion attempt

Klasifikace IDS/IPS



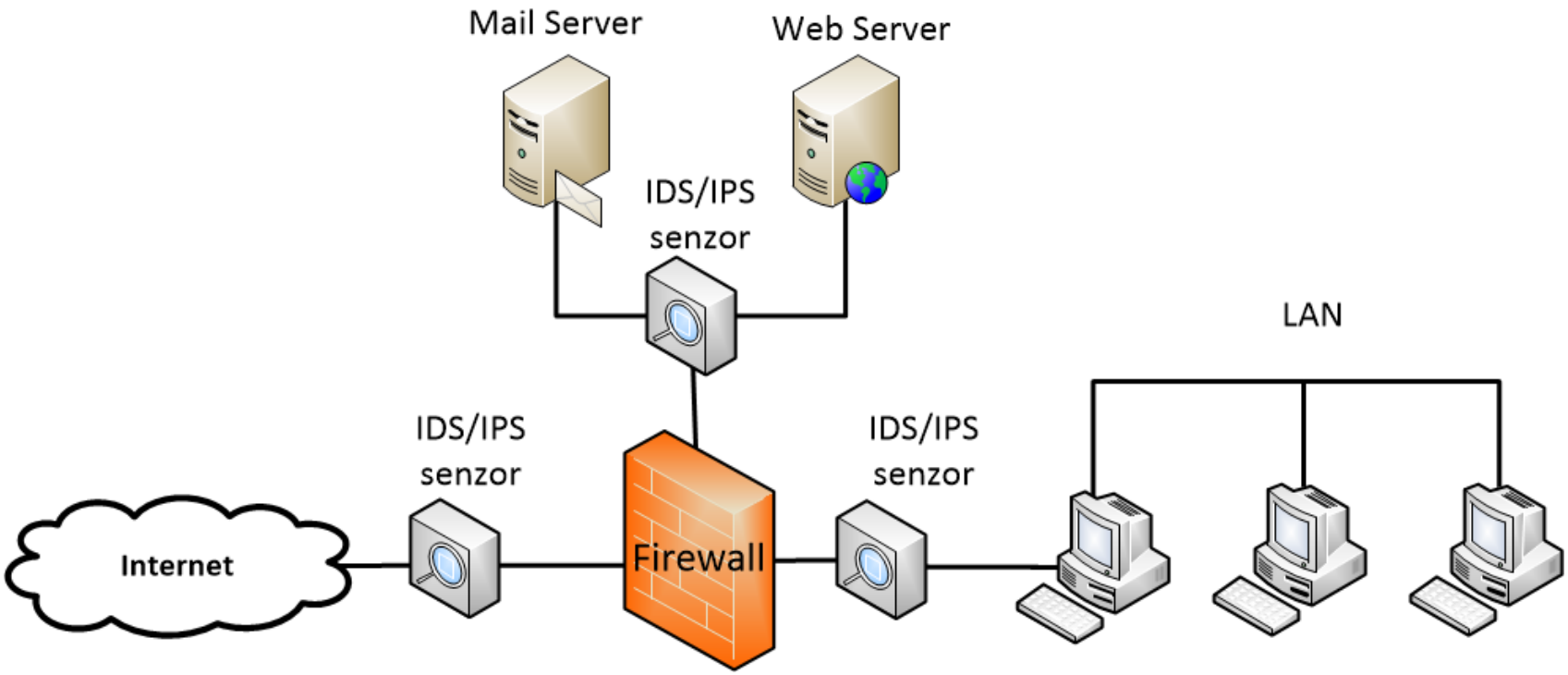
Klasifikace IDS/IPS



IDS/IPS umístěny na síti (Network Based IDS - NIDS)

- Realizují **odchytávání paketů** (packed sniffing) a následnou analýzu k detekci a potlačení podezřelé aktivity.
- Typicky jsou nasazeny inline jako síťové firewally (transparentní, 2 porty - vstup/výstup).
- Síťové rozhraní je v promiskuitním režimu.
- Umístění je rozhodující pro správnou funkci, často mezi komponenty síťové infrastruktury (úzká místa, GW, hraniční prvky).

IDS/IPS umístěny na síti (Network Based IDS - NIDS)



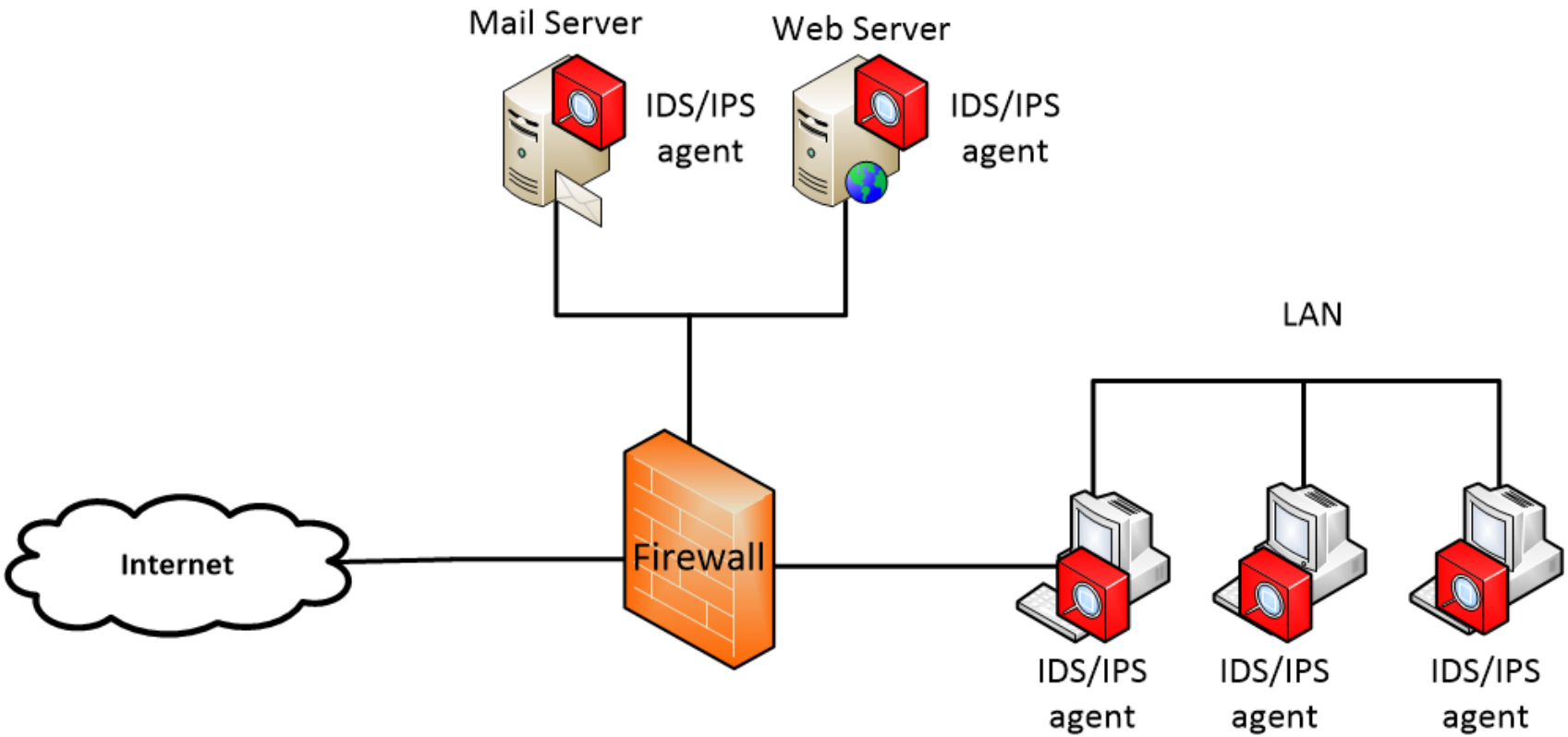
IDS/IPS umístěny na síti (Network Based IDS - NIDS)

- IDS/IPS senzory jsou k dispozici ve třech provedeních:
  - **Komplexní zařízení** - specializovaný hawdware včetně dedikovaného software k efektivnímu odchytávání a analýze provozu. Příklad: Cisco IDS4200, IBM Real Secure Network, Radware DefensePro, Radware Defense flow, F5 Network BIG-IP, McAfee Network Security Platform atd.
  - **Softwarová** řešení - software instalovaný na server, který je umístěn v síti a monitoruje provoz. Příklad: Snort, Suricata, Bro, Kismet, Cisco Security Agent atd.
  - **Cloudová** řešení - dostupné jako služba od IPS poskytovatelů (Radware, F5, McAfee).

IDS/IPS umístěny na hostiteli (Host Based IDS - HIDS)

- **Softwarový agent** instalovaný na systému, který má být monitorován (PC, server, atd.).
- Účel a princip stejný s NIDS/NIPS, ale monitoruje jedno zařízení.
- Monitorované parametry: vstupní/výstupní pakety, souborový systém (přístupy, kontrola integrity souborů), logování vybraných služeb (web server), přihlašování do systému, systémové logování, atd.
- HIDS/HIPS často nasazovány **kritické systémy** (veřejně dostupné servery, servery obsahující senzitivní informace).

IDS/IPS umístěny na hostiteli (Host Based IDS - HIDS)

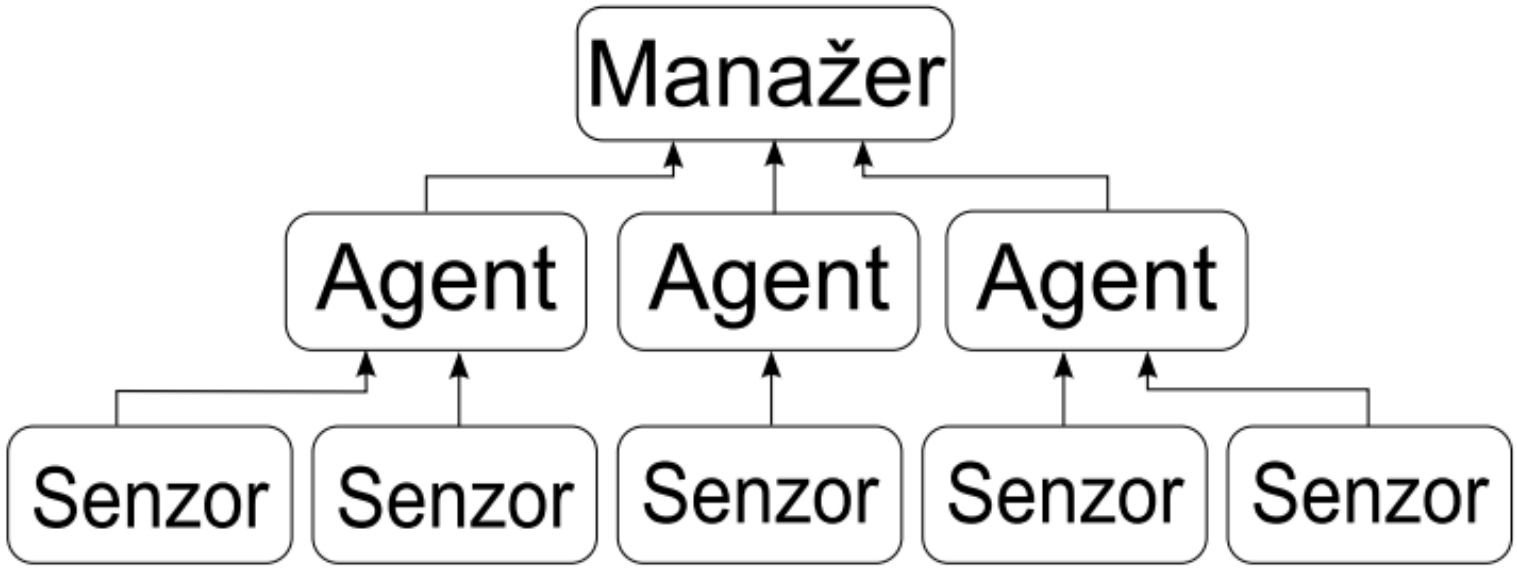


Jednovrstvá architektura

- Nejjednodušší možný případ, kdy IDS nebo IPS je tvořen pouze jednou komponentou.
- Tato komponenta obstarává všechny potřebné funkce.
- jednoduchost, cena, nezávislost X neefektivní využití komponent.

Vícevrstvá architektura

- Existence vertikální komunikace mezi entitami.
- Uspořádány do hierarchické stromové struktury.
- Zpravidla tři vrstvy.



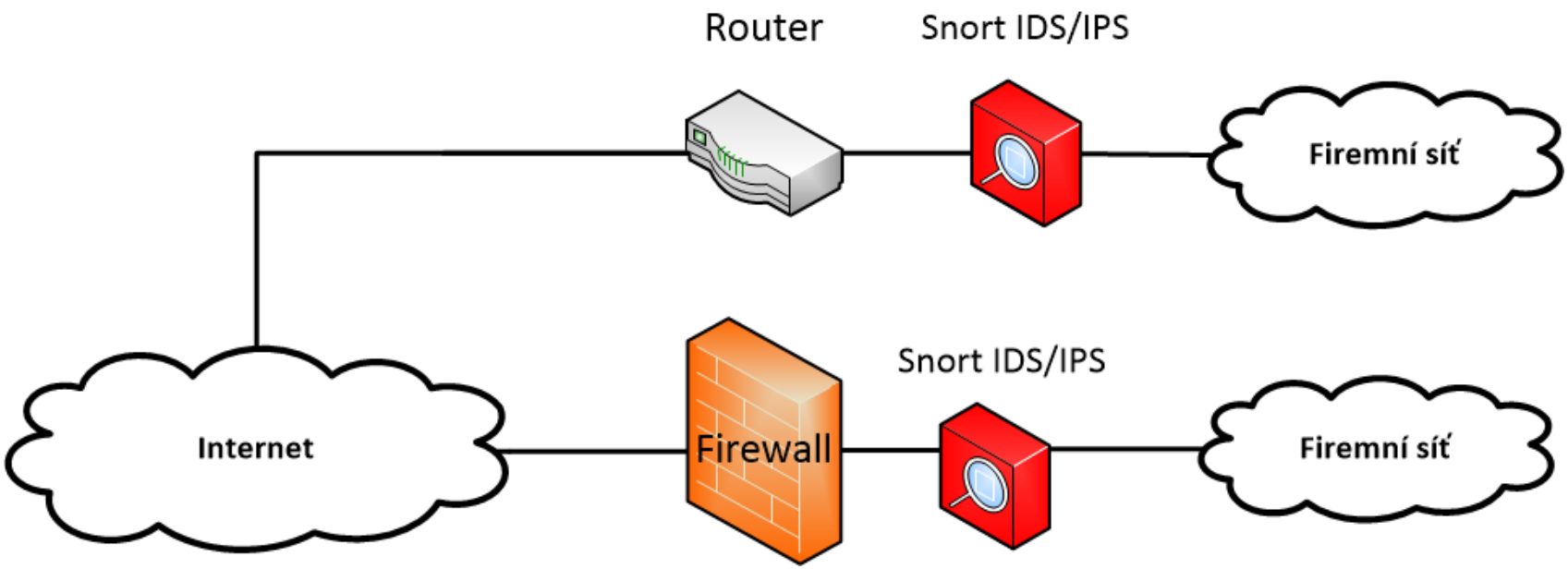
Vícevrstvá architektura

- Základem systémů jsou senzory, které monitorují síťový provoz (logování) a data předávají dále ke zpracování agentům.
- V rámci jednoho systém může být umístěno několik agentů analyzující konkrétní protokol či službu (FTP, HTTP, TCP, UDP atd.).
- Manažer analyzuje hlášení (alerts) a realizuje definovaná opatření (informuje administrátora, upravý pravidla FW, uloží incident v databázi).
- Vyšší efektivita a hloubka analýzy X vyšší náklady.

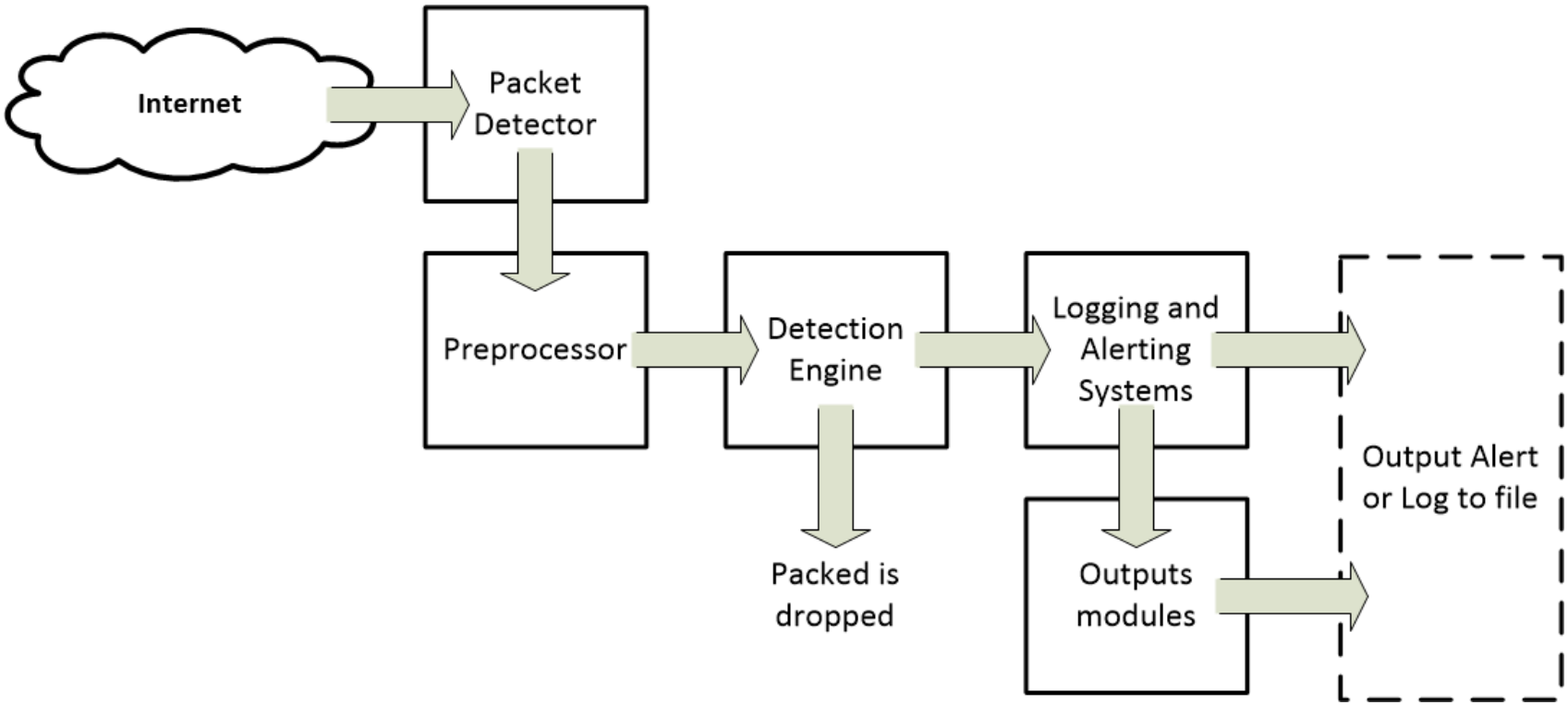
Co je Snort?

- Snort představuje open source IDS/IPS.
- Používá jazyk popisující pravidla, detekci anomálií, signatur a analýzu protokolu.
- Snort byl dříve často nasazovaný IDS/IPS systém v produkčním prostředí.

Snort - typické umístění



Snort - komponenty



Snort - komponenty

- Packet Detector - ukládá pakety z různých typů síťových rozhraní, připravuje pakety na zpracování.
- Preprocesot - připravuje data pro detekování, detekuje anomálie v hlavičkách paketu, skládá pakety (útočník skrývá signaturu rozložením útoku), dekoduje HTTP URI.
- Detection engine - nejdůležitější komponenta, aplikuje pravidla na pakety.
- Logging and Alerting Systems - logování popřípadě vyvolání alertu.
- Output Modules - zpracovávají alerty a logy do výsledného výstupu.

Suricata

- Open Source Next Generation IDS/IPS (nové technologie),
- v současné době nejpoužívanější systém (aktualizace signatur),
- na Surikatě jsou založeny komerční zařízení,
- odstraňuje nevýhody Snortu (kritické využití více vláken - rychlejší zpracování dat).

Příklad dedikovaných zařízení - McAfee a Radware

radware
Every second counts

Reference I

[1] KENT, K., SOUPPAYA, M Guide to Computer Security Log Management. NIST Special Publication [online]. 2006, , 1 - 72 [cit. 2017-02-07]. Dostupné z: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf

[2] ANSSI Recommandations de sécurité pour la mise en oeuvre d'un syst?me de journalisation. ANSSI Note technique [online]. 2013, , 1 - 23 [cit. 2017-02-07]. Dostupné z: http://www.ssi.gouv.fr/uploads/IMG/pdf/NP_Journalisation_NoteTech.pdf

[3] WATSON, C., KEARY, E., FITZGERALD, A Logging Cheat Sheet. OWASP [online]. 2016 [cit. 2017-02-07]. Dostupné z: https://www.owasp.org/index.php/Logging_Cheat_Sheet

[4] CHUVAKIN, Anton. A., Kevin J. SCHMIDT a Christopher PHILLIPS. Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management

[5] KENT, S. IP Authentication Header. RFC 4302. IETF, Fremont 2005

Reference II

[6] KENT, S. IP Encapsulating Security Payload (ESP). RFC 4303. IETF, Fremont 2005

[7] KAUFMAN, C.; HOFFMAN, P.; NIR, Y.; ERONEN, P. Internet Key Exchange Protocol Version 2 (IKEv2). RFC 5996. IETF, Fremont 2010

[8] RFC4301: Security Architecture for the Internet Protocol". Network Working Group of the IETF. December 2005

[9] NIST SP 800-30: Guide for Conducting Risk Assessments. ITL NIST. September 2012.

Děkuji za pozornost!
Dotazy ?

martinasek@vut.cz