

Dámy a pánové,

jsem velmi rád, že Vás mohu opět uvítat jménem Fakulty bezpečnostního managementu na půdě Policejní akademie ČR v Praze. Stává se to tradicí, že na PA ČR se v rámci pracovní skupiny 05 - Kybernetická bezpečnost organizujeme společně s českou pobočkou AFCEA bezpečnostní seminář. Dnešní, již 6. bezpečnostní seminář je významným počinem, který zorganizovala Fakulta bezpečnostního managementu PA ČR v Praze společně s AFCEA ČR na téma „Big data“ a zajisté přispěje ke vzájemné výměně poznatků a zkušeností v oblasti datové, informační, fyzické a další bezpečnosti organizací a státu.

Proč se zabýváme touto problematikou je každému jasné, neboť velké objemy dat a SW jsou mimo dosah klientů.

Až do nedávné doby bylo zpracování dat pro analytické účely poměrně statickou úlohou. Konkrétně podniky produkují zejména strukturovaná data ze stabilních datových modelů prostřednictvím podnikových aplikací, jako je CRM, ERP a finanční aplikace. Pomocí ETL nástrojů se data z těchto systémů přenáší do pracovní oblasti, kde se kontroluje kvalita údajů a provádí se jejich normalizace, finálně se pak ukládají do datového skladu. Tento proces obvykle běží v pravidelném cyklu – zpravidla na denní nebo týdenní bázi. Správci datového skladu pak vytvářejí reporty, které běží nad daty uloženými ve skladu.

Nástup webu, mobilních zařízení a dalších technologií zapříčinil zásadní změnu charakteru dat a způsobu jejich využití. Již nejsou centralizovaná, vysoce strukturovaná a snadno zvládnutelná, ale více než dříve jsou volně strukturovaná (pokud mají vůbec nějakou strukturu), vysoce distribuovaná a mají vzrůstající objem. Často se v této souvislosti hovoří o trojrozměrnosti velikosti a růstu dat (zkráceně také jako 3V):

- objem (volume) – množství dat vznikajících v rámci provozu firem roste exponenciálně každý rok,

- typ (variety) – různorodost typů dat vzrůstá, například nestrukturované textové soubory, semi-strukturovaná data (XML), data o geografické poloze, data z logů,
- rychlost (velocity) – rychlost s jakou data vznikají a potřeba jejich analýzy v reálném čase vzrůstá díky pokračující digitalizaci většiny transakcí, mobilním zařízení a vzrůstajícímu počtu internetových uživatelů.

Big data mají odlišné vlastnosti, které je odlišují od „tradičních“ firemních dat. Tradiční datové sklady a nástroje pro správu dat nejsou připraveny na zpracování a analýzy velkých objemů dat ve velmi krátkém čase (někdy real-time) nebo nákladově efektivním způsobem. Proto je třeba hledat nové způsoby zpracování a analýzy velkých objemů dat.

Na dnešním semináři se podělíme o zkušenosti se zajištěním bezpečnosti Big data na různých úrovních a oblastech.

Dovolte mi, abych popřál Vašemu, našemu jednání úspěch a budeme se s Vámi těšit na půdě PA ČR na další setkání.

Praha, 27. listopadu 2012 doc. RNDr. Josef Požár, CSc.
děkan