Kryptografická ochrana osobných údajov v cloude

Lenka Stúpalová

Školiteľ: Michal Rjaško

Cloud computing

- Všadeprítomný, pohodlný a praktický prístup cez sieť ku zdielanej kôpke konfigurovateľných výpočtových zdrojov
- (napr. siete, servery, úložný priestor, aplikácie a služby)

Cloud computing

- Všadeprítomný, pohodlný a praktický prístup cez sieť ku zdielanej kôpke konfigurovateľných výpočtových zdrojov
- (napr. siete, servery, úložný priestor, aplikácie a služby)

- Dec 2016 - Skyhigh (McAfee) – 30 mil. zamestnancov, 600 firmách

18.1%

OF FILES IN THE CLOUD

CONTAIN SENSITIVE DATA

Ochrana osobných údajov

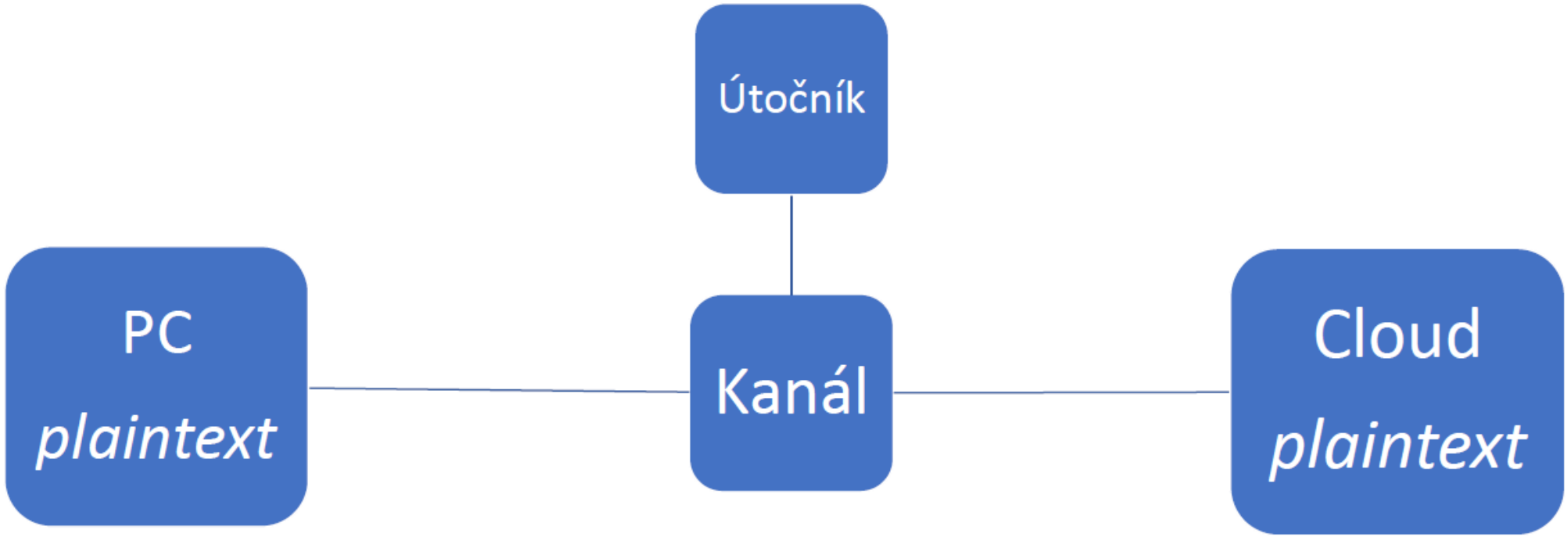
- 25.mája 2018 - GDPR
- OÚ – všetko, podľa čoho vie niekto nejakú osobu identifikovať

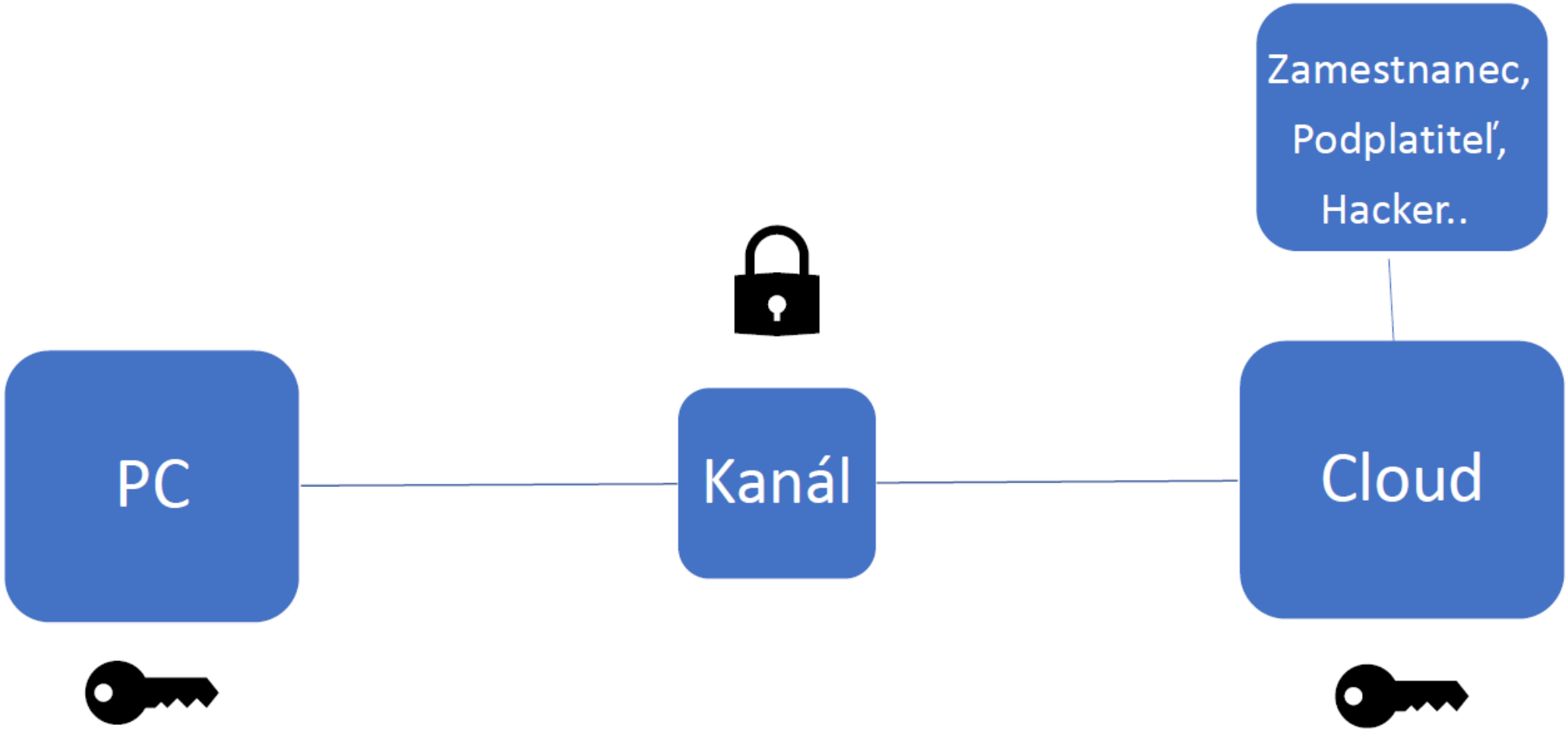
- Zabezpečenie trvalej dôvernosti, integrity, dostupnosti a odolonosti systémov spracúvania osobných údajov

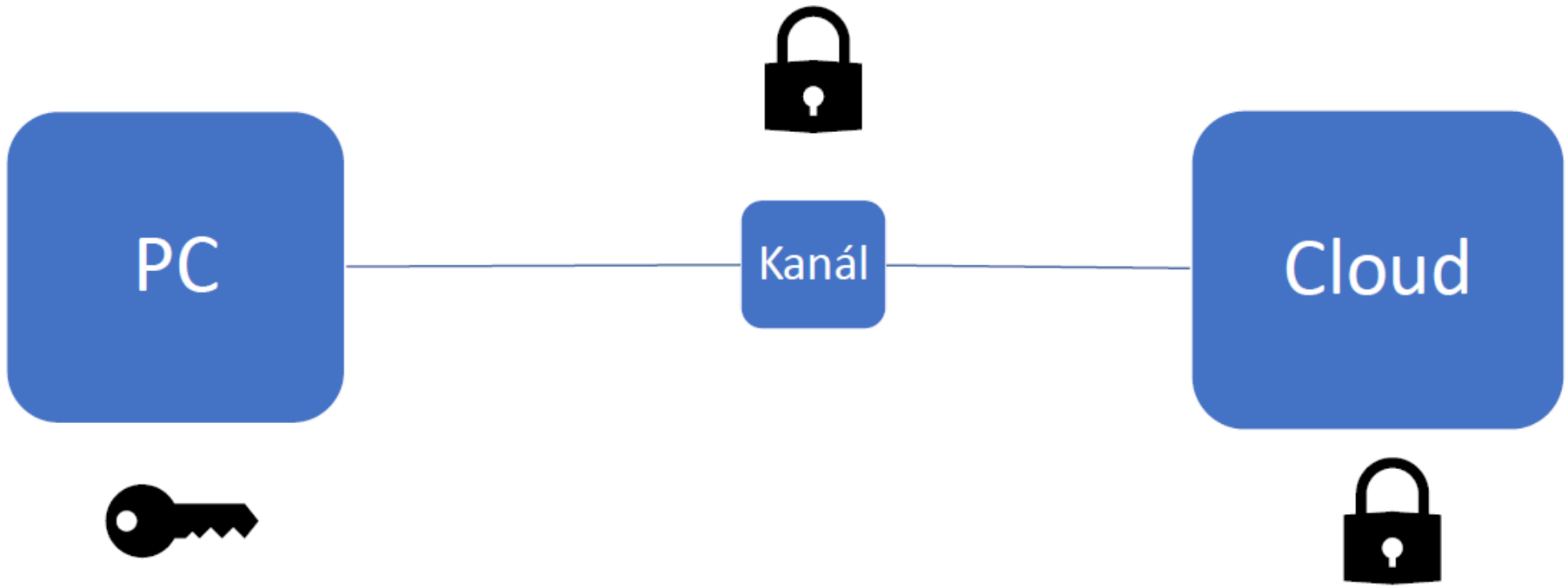
- §9 Zásada správnosti.
  - OÚ musia byť správne a podľa potreby aktualizované
  - OÚ nesprávne bez odkladu vymazať alebo opraviť (prípadne označiť, že sú zlé)
- §11 Zásada integrity a dôvernosti – primeraná bezpečnosť OÚ
- §12 Zásada zodpovednosti – prevádzkovateľ zodpovedný (sankcie)
- §21 Právo na prístup k OÚ – povinnosť poskytnutia spracovávaných OÚ na vyžiadanie dotknutej osoby (opakované môže spoplatniť)
- §22 Právo na opravu OÚ, §23 Právo na výmaz OÚ

- Čiže potrebné funkcie, ktoré musíme zo zákona vedieť robiť s OÚ – nájsť/vyhľadať osobu, zmeniť, vymazať,..

- Ako teda použiť cloud na spracovanie osobných údajov?


Útočník
PC
plaintext
Kanál
Cloud
plaintext


Zamestnanec,
Podplatiteľ,
Hacker..
PC
Kanál
Cloud


PC
Kanál
Cloud

- Šifrovací kľúč osbných údajov by nemal byť zdielaný s cloudom

- Lokálne šifrovanie – len ja mám kľúč

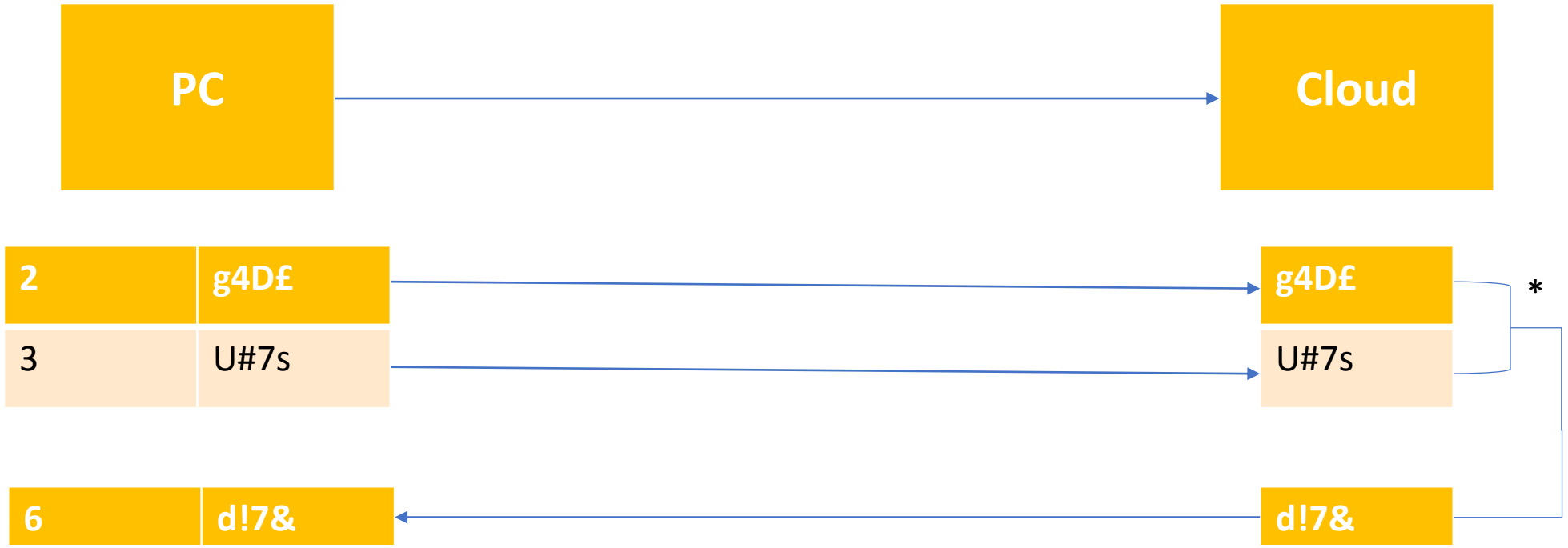
(takto sa to bežne robí ak máme nejaké tajnejšie informácie)

Nevýhoda lokálneho šifrovania

- Prístup k dátam/nájsť/zmeniť - napr. databáza ľudí, zmena priezviska/bydliska/poisťovne/oprava preklepu
  - Stiahnuť celý súbor
  - Odšifrovať celý súbor
  - Vykonať zmenu/akciu
  - Zašifrovať celý súbor
  - Poslať/nahrať na cloud

- 1000 záznamov ľudí možno cca 20MB,
  miliarda záznamov –TBs (pr. eBay)

- Chceli by sme - vykonávať operácie na zašifrovanom texte bez nutnosti odšifrovávania



Homomorfizmus

- Univerzum a operácia (R,*)
- Funkcia f pr. f(x)=|x|
- Vlastnosť **f(x*y) = f(x)*f(y)**

- Príklady homomorfizmov
  - (R,*): $f(x) = x^2$
  - (Z, +): f(x) = 3x
  - (Z mod p, +): f(x)= x + p

Homomorfické šifrovanie

- Štvorica (K, E, D, A) pravdepodobnostných polynomiálnych algoritmov
- Pravdepodobnosť $[D(k, c) \neq m]$ je zanedbateľná
- Homomorfická vlastnosť – A – vstup ke, c1, c2, výstup c3, nech $m3 = m1 \odot m2$ (c1, c2 sú zašifrované m1 a m2), potom Pravdepodobnosť $[D(A(ke, c1, c2))] \neq m3$ je zanedbateľná

- Čiastočne a plne homomorfický kryptosystém

- Výhoda – RLWE – odolné voči kvantovým pc

Homomorfické šifrovanie

- Univerzum a operácie : (Z mod 2, +) && (Z mod 2, *)
  - T.j. množina {0,1} - bity

- E(k,M1* M2) = E(k, M1) * E (k, M2)
- E(k,M1+ M2) = E(k, M1) + E (k, M2)

- Skúsme zašifrovať (1 + a*b)

  Enc(k, 1+a*b)= Enc(k, 1) + Enc(k, a)* Enc(k, b)

a	b	1 + a*b
0	0	1
0	1	1
1	0	1
1	1	0

- Máme NAND
- NAND bloky –> ľubovoľné logické hradlo - logický obvod -> ľubovoľná funkcia -> ľubovoľný program

  (Každý program vieme zredukovať až na logické hradlá)

- Spravíme program z NANDov, zašifrujeme, pošleme na cloud, cloud vykoná neznámu funkciu na neznámych dátach a vráti výstup

  (napr. nájdenie a zmena údaju)

Prečo teda nemáme HŠ všade?

- Nepraktické
- Veľkosť programu len z NAND operácií
- Prvý nárvh schémy – 2009 Gentry +IBM - biliónkrát pomalší ako operácie na plaintexte
- 2013 – zlepšenie na milión – 16 jadrový server – ak trvala operácia na plaintexte 1s, s hom. šif. bude trvať 12 dní
- 2018 – zrýchlenie 15-75 krát

- Nepodporuje prístup viacerých používateľov

- Snažíme sa navrhnúť ako chrániť osobné údaje v cloudových službách
- PRAKTICKY možné riešenie
- Podpora viacerých používateľov – rôznych skupín ľudí v rôznym skupinám dát

- Iný prístup – jednoduchšie a efektívnejšie symetrické/asymetrické šifrovanie

Id študenta	Meno, priezvisko	Rodné číslo...
nezašifrované	ostatné tu všetko zašifrované náhodným K pre každý záznam	

Accessor ID Kto pristupuje	Accessed ID Kam pristupuje	K k danému riadku zašifrované jeho heslom

Problémy

- Čítať / aj písať dáta? (pr. študent si nevie zapísať známku)
  - Riešenie – namiesto jedného K – asymetrické šifrovanie – verejná a súkromná časť kľúča K
- Odchod zamestnanca – pozná K
  - Prešifrovanie všetkého, nové K, ku ktorým mal prístup – príliš dlhé
  - Spolupráca servera pri dešifrovaní
- Neoprávnená zmena dát
  - Digitálny podpis riadkov – súčasť verejného K
- Pr. nového učiteľa k predmetu – veľa ďalších riadkov v tabuľke s K
  - Pridanie skupín

Id študenta	Meno, priezvisko	Rodné číslo...	Digitálny podpis / hash
nezašifrované	ostatné tu všetko zašifrované náhodným K pre každý záznam		šifrovaný odtlačok riadku

Accessor ID Kto pristupuje	Accessed ID Kam pristupuje	Verejné K k danému riadku	Súkromné K k danému riadku
Užívateľ alebo grupa	Grupa alebo id študenta resp. odkaz na inú tabuľku	Zašifrované náhodným s, ktoré je zložené xorom hesla používateľa a časťou servera	

- Server – tabuľka „serverových tajomstiev“ k jednotlivým záznamom

Priebeh

- Riadok šifrovaný náhodným stringom s
- Serverové tajomstvo = s ⊕ heslo používateľa

Užívateľ Cloud

heslo ⊕ náhodný string -> šifrované dáta ⊕ časť servera ⊕ (heslo ⊕ náhodný string)

<-

Odxoruje svôj náhodný string

heslo ⊕ časť servera = s (ktorým boli zašifrované dáta)

s ⊕ šifrované dáta = odšifrované dáta

Užívateľ sa nedozvie ako vyzerá serverové tajomstvo ani ako vyzerajú dáta v šifrovanej podobe, cloud sa nikdy nedozvie užívateľovo heslo ani dáta v plaintexte ani kľúč na ich odšifrovanie

Záver

- Snažíme sa zanalyzovať rôzne možnosti ochrany osobných údajov v cloude, ich bezpečnosť a efektívnosť a navrhnúť riešenie, ktoré by bolo dostatočne bezpečné aj efektívne

Ďakujem za pozornosť!