1 slajd

Dobrý den. Jmenuju se Pavel Yadlouski. Dneska chtěl bych vám představit svou práci.

2 slajd – Úvod do problematyki

Zaprve chtěl bych vás seznamit s problémou kterou snažil jsem vyřešit.

3 slajd – Monitorování síti

Dneska pocitacová síť je absolutně všude kolem nas. A čím dal exestuje, tím vic roste. A čím víc roste, tím větší haos můžz se dít v teto sítě. Tim pádem bylo by dobře moct nějak tu síť kontrolovat a monitorovat. Jedním ze způsobu monitorování provozu na síti je tzv. monitorování toků za použití NetFlow nebo IPFIX. Architektura takového monitorování se skládá z jednoho nebo více exportérů a obvykle jednoho kolektoru. Exportér z procházejícího provozu vytváří záznamy o tocích a tyto toky následně posílá (exportuje) na sběrný bod tzv. kolektor. V teto situaci exporter muze byt napriklad router na dane site, pocitac vybavený síťovou kartou a odpovídajícm softwerem. V této prezentaci budu hovořit o IPFIX koletrou druhé generaci a jeho rozšíření

4 slajd – IPFIXcol2

Tady na obrázku můžete vídět jak ten kolektor vypadá uvnitř. Je to modulární kolektor. Moduly na vstupu můžou prijímat data přes TCP nebo UDP protokoly, a to buď přímo od exporterů nebo z datového uložiště. Pak te data můžou tect do vnitřních modulů pro zpracování dat, například do filtračního modulu. Může byt libovolný počet těch modulů. Na konci jsou moduly pro ukladání data do uložište nebo pro odesilání dalším aplikacim

Ale problem je v tom, že nejsme schopni záznamy sjednotit (nebo transformovat) podle nějakého klíče, abychom z nich získali vybrané statistiky. Ta informaci by mohla být užítečná právě na monitorování sítě, mohla by pomoct otevřít určité anomalie. z yy Prave ten problém uniku užítečních dat snažil jsem vyřešit. A to pomoci vyvýjení agregačního modulu pro IPFIX kolektor.

5 slajd – IPFIXcol2 z agregacnim modulem

Nový agregační modul slouží k vytvoření statistyk o sítě. Je to další vnitřní modul kolektora, který může být zařazen do pupliny vnitřních modulů, a bude pracovat jenom s offline daty.

6 slajd – agregace

Dál chtěl bych ukázat v čem spočívá samotná agregace

7 slajd – podzanamy

Cílem tohoto procesu je dostat určitou informaci ze všech vstupních záznamů. Jaká infomrace musí být vyextrahovaná stanoví uživatel zadávaním vstupní konfiguraci. V této konfiguraci stanoví, které položky mají být jako klíč, a které musejí být agregovany. Klíč může se setávovát z několiká položek, stejně tak i hodnotových položek může být víc než jedná.

Můžeme řict, že agregator vytváří podzáznamy z vstupních dat. Format podzáznamů můžete vídět na obrázku. V podstatě je to dvojice: klíč a hodnota.

8 slajd – funkci

Chtěl bych také uvest co ta "agregace"se predstavyje. v postate je to "provest agregaci"znamena pomoc nejake funkci provest analizu dat. takove funkce jsou nasledujici: summa hodnot, vyber maxima a minimuma. Take muzeme pouzit funkci logicky OR. to muzeme potrebovat naprkad pro flagy ze standardu IPFIX.

9 slajd – hash tabulka

Je ocevidne, ze vysledek agregaci musi byt ulozen aby byl mohli pak provest analizu vytvorenych statistik. Zvolil jsem pro ulozeni **hash tabulku s explicitne zretezenymi synonimami**. Takze kazdy podzaznam je ulozen do teto tabulki. Index do tabulky se vytvari pomoci hashovaci funkci xxHash. Na index je ulozen samotny podzaznam. V pripade kolizi, podzaznamy se vetvareji jednosmerno vazany seznam.

Snazil jsem oddelit implementaci hash tabulki od samotnej impementaci agregatoru aby v budoucnu byla moznost misto hashovaci tabulky pouzit nejakou jinou strukturu.

10 slajd – shrnuti

Tím pádem máme nasledující situaci

1. Máme problém, ze máme hodně dat, ze kterých můžeme dostat statistiky, které byly bych vhodné pro analizu síti. Ale nemám pro to nástroj v IPFIX kolektoru.
2. Řešení: vytvoření nového modulu, který bude poskitovat tuto statistzku – agregační modul
3. Modul bude dělt podzaznamy z přichozích záznamů
4. Pomocí agrigačních funkci zpacovává daty
5. Výsledek ukládá do hašovácí tabulki. Na výstupu máme vyplněnou tabulku s agregovanymí daty pro další analyzu