A grid hálózatok biztonsági kérdései

KŐVÁRI KÁLMÁN

KFKI Részecske- és Magfizikai Kutatóintézet
dalion@sunserv.kfki.hu

Lektorált

Kulcsszavak: adatbiztonság, betörés, incidens, sebezhetőség, X.509, grid, proxy, klaszter

Számítógépes klasztereket egységes nemzetközi hálózatba kapcsolva, megfelelő bróker- és menedzserrétegek közbeiktatásával látszólag egyszerűen létre lehet hozni egy grid rendszert. Viszont a biztonsági szempontból könnyen védhető, zárt és gyakran a világhálóról is leválasztott helyi klaszterekkel szemben a sokszor publikus internet kapcsolatokat használó grid rendszerek komoly biztonsági kérdéseket vetnek fel. A cikk erre a kérdéskörre igyekszik rávilágítani.

1. Bevezetés

„A biztonság folyamat, nem pedig termék." Bruce Schneier ezen mondatát rengetegen idézték már [1], hiszen rámutat a fogalom körül keringő legnagyobb tévhitre. Egy terméket meg lehet vásárolni, ezután lehet használni vagy nem használni és amennyiben már elavultnak érezzük, és újabb verziót látunk megjelenni, frissíthetjük, vagy maradhatunk a régi és jól megszokott változatnál. A biztonsági kérdések tekintetében efféle viselkedést sajnos nem engedhetünk meg magunknak. A biztonság egy folyamat, tehát állandó aktív figyelmet igényel, folyamatosan és éberen kell kezelnünk az érintett eszközöket és szolgáltatásokat. A tapasztalat azt mutatja, hogy az a biztonsági szakember, aki még soha nem volt érintett incidensben, vagy szerencsés, vagy nagyon fiatal. A jó szakember ismérve, hogy minden esetben törekszik a problémák és incidensek megelőzésre, de képes ezek elhárítására és utólagosan a megfelelő reakcióra is. Az utóbbi kettő legalább annyira fontos, mint maga a megelőzés, hiszen ezzel kerülhető el az esetlegesen még súlyosabb másodrendű károkozás és a probléma esetleges megismétlődése.

A biztonság szót sokszor, sok esetben teljesen különböző jelentéstartalommal használjuk, viszont minden esetben valamilyen értelemben a bizalomhoz kapcsolódik. Lehet szó például adatbiztonságról, ezen belül sem mindegy, hogy az a fontos, hogy az adat megmaradjon akár atomtámadás esetén is, vagy az, hogy ténylegesen korlátozott és szabályozott legyen a hozzáférhetősége. Lehet szó biztonságos azonosításról, ahol a fő feladat egy kliens minél megbízhatóbb azonosítása, bizonyos feladatokra való autorizálása, illetve egy személy által delegált másik személy vagy szolgáltatás azonosítása.

A kapcsolatok biztonságának megítélése a két végpont közti hálózati összeköttetés megbízhatóságát, kódoltságát és kívülállók számára való hozzáférhetőségét jelenti. A szolgáltatási biztonság a kliens oldalon felmerülő kérdés: mennyire bízhat meg a felhasználó abban, hogy az általa kért művelet ténylegesen végrehajtódik a rendszerben.

A biztonságot csökkentő tényezők lehetnek fizikai jellegűek, például egy hálózati kapcsolat korlátozott sebessége, de többnyire inkább logikai eredetűek, és vagy vezérlési, vagy biztonsági megfontolásból vannak jelen. A felhasználónak általában csak két apró igénye szokott lenni: mindent és azonnal. Ezzel pedig mindig szemben áll a helyi biztonsági szakértő, aki rögtön azt mondja, hogy semmit és soha, mert úgy biztonságos. Természetesen egyik megközelítés sem helyes, mindig valamiféle kompromisszum a megoldás. Ezen kérdések nem csak a gridek világában aktuálisak, de ezek nagy komplexitása miatt itt hangsúlyozottan jelennek meg. A jelen cikkben ezeket a kérdéseket tekintjük át.

2. Azonosítás, jogosultsági kérdések

A grid világában is szükséges a személyazonosságunk igazolása. Ennek megvalósítására a legtöbb grid-implementáció az X.509 típusú digitális tanúsítványok által biztosított PKI-t (Public Key Infrastructure) választotta. Röviden annyit kell tudni ezekről a tanúsítványokról, hogy egy titkos és egy nyilvános kulcsból állnak, a titkos kulcsot a felhasználó saját feladata titokban tartani, ezért többnyire jelszóval vagy PIN kóddal is védik. A titkos kulcs kompromittálódása vagy annak gyanúja esetén a tanúsítványt azonnali hatállyal vissza kell vonni. A nyilvános kulcsot egy CA (Certification Authority – tanúsítványhitelesítő) a felhasználó azonosítása után hitelesíti, azaz a saját titkos kulcsával aláírja. Ennek a digitális aláírásnak a hitelessége a CA nyilvános kulcsának birtokában ellenőrizhető. A felhasználónak így csak néhány CA nyilvános kulcsának valódiságában kell bíznia, ezek többnyire az operációs rendszerrel együtt, annak részeként letölthetőek. A módszerrel nem csak személyek, hanem szolgáltatások vagy számítógépek kilétének megbízható ellenőrzése is megoldható.

A grid-felhasználók virtuális szervezetekbe (VO, Virtual Organization) tömörülnek. A virtuális szervezetek taglistája nyilvános információ, a tagok DN-jével (Distin-

guished Name – a személyi szám digitális megfelelője) együtt. Minden erőforrás-szolgáltató maga döntheti el, hogy mely virtuális szervezeteket támogatja. Ezekből csak néhány darab van, így ezek felsorolásával könnyen szabályozható a hozzáférés. Alapvetően egy támogatott virtuális szervezet minden tagja jogosult a kérdéses erőforrások valamilyen mértékű felhasználására, viszont minden szolgáltatónak lehetősége van akár személyre szabottan is módosítani az erőforrásaira vonatkozó jogosultságokat.

A pontos erőforrás-hozzáférési beállítások már erősen függnek az alkalmazott grid-rendszertől, de általánosan jellemző, hogy lehetőség van virtuális szervezetek számára dedikálni erőforrásokat – így csak ők használhatják, – illetve úgynevezett *fair share* elven megosztani, ami azt jelenti, hogy bizonyos idő-intervallumonkénti átlagban adható meg, hogy melyik VO milyen mértékben veheti igénybe az adott szolgáltatást.

3. Adatbiztonság

A feladat egyszerű: a felhasználó azt szeretné, hogy az adatai olyan helyen legyenek, ahol nem veszhetnek el egy esetleges diszkhiba, áramkimaradás, vagy rosszindulatú cselekedet hatására. Erre elég egyszerű megoldás, ha az adatot több példányban tároljuk, azaz replikáljuk. A módszer nehézsége egyrészt abban rejlik, hogy a másolatokat folyamatosan frissíteni kell, ha az eredeti adaton változtatunk, másrészt az egyes példányokat az egységes elnevezés mellett is meg kell tudnunk különböztetni egymástól. A replikáció előnye viszont, hogy a skálázott elérhetőséget is megoldja: minél több replika áll rendelkezésre egy adathalmazról különböző földrajzi illetve hálózati pontokon, annál többen férhetnek hozzá egyszerre az adatokhoz. Ezeket a szolgáltatásokat valósítják meg általában az úgynevezett Data Grid rendszerek.

A nehezebb feladat a hozzáférhetőség problémája. Ezt több szinten lehet megoldani. Ha elég a klienseknek az, hogy ők maguk meghatározhatják, hogy a rendszer kinek adjon engedélyt az adat olvasására, akkor ezt egy viszonylag egyszerű ACL (Access Control List – jogosultság lista) alapú hozzáférési sémával meg lehet oldani. Viszont egyes kliensek – kiemelhető példaként a biomedika, konkrétan az orvosi adatok – megkövetelik azt, hogy maga a rendszergazda se férhessen hozzá az adathoz az ő engedélyük nélkül.

Ilyen feladatokat oldanak meg a Hydra tárolórendszerek. Működésük lényege, hogy az adat kódolt formában kerül feltöltésre és replikálásra. A dekódoló kulcsot három részre osztják és a darabokat három különböző kulcs-szerveren helyezik el, szigorú szabályozással, illetve további kódolással biztosítva, hogy csak az arra jogosultak férhessenek hozzá. Ezt a korábban említett PKI használatával lehet megoldani oly módon, hogy az összes hozzáférésre jogosult személy digitális tanúsítványának nyilvános kulcsával titkosítjuk a dekódoló kulcs megfelelő darabját. Ezek után csak a kérdéses tanúsítványok titkos kulcsának birtokában lehet hozzáférni a dekódoló kulcs egy-egy darabjához. Mivel a három kulcs-szerver mindegyike, egyenletesen elosztva 2-2 kulcs-darab birtokában van, így egyetlen kulcsszerver kompromittálódása nem jelenti a kulcs kompromittálódását, másrészt egyetlen szerver elérhetetlensége – esetleg megsemmisülése – sem vezet az adat elvesztéséhez. Természetesen a kulcs-tördelést ennél több szerverre is el lehet végezni úgy, hogy a tördelésnek ezen tulajdonsága megmaradjon, de akár több szerver kiesését is tolerálja a rendszer.

Meg kell azonban jegyeznünk, hogy a módszernek két jelentős hátránya is van: az egyik, hogy a dekódoló kulcsot mindenkinek ki kell adni, aki az adatot olvasni akarja, s ha akár egyetlen kulcs is kompromittálódik, veszélybe kerül az adat. A másik probléma, hogy nem lehetséges az olvasási jog visszavonása, ha egy kliens egyszer már megszerezte a kulcsot. Ezekre a nehézségekre születtek már megoldási ötletek, ilyen például az, hogy a dekódoláshoz szükséges szimmetrikus kulcsot az adatot tároló szerver kapja meg közvetlenül, a felhasználók pedig ideiglenes, egyedi kulcsokat használnak a hozzáféréshez. Ebben az esetben a tároló szervernek is nagyon megbízhatónak kell lennie és mivel minden ilyen biztonsági szintű adatfolyamhoz két kódolást is el kell végezni az adott gépen, nagyon nagy számítási kapacitásra is képesnek kell lennie. Ennek a módszernek a megvalósítása és tesztelése még folyamatban van.

4. Proxy-k, delegáció

Amikor a griden dolgozunk, elemi számítási egységeket, úgynevezett feladatokat (job) küldünk be. A feladatok a küldés pillanatától fogva a terminálásig többnyire 4-5 számítógépet is érintenek. Ezek közül többnek meg kell győződnie a feladat tulajdonosának személyazonosságáról. Ezt nem közvetlenül a digitális tanúsítvánnyal oldjuk meg, hanem egy köztes hírvivő entitást delegálunk, így nem kell kiadnunk a kezünkből a tanúsítványunk titkos kulcsát.

Természetesen egy nem-elsőgenerációs grid szolgáltató programcsomag (middleware) az alábbi lépéseket már automatikusan megteszi helyettünk, de fontos, hogy értsük a háttér-mechanizmust, hogy rálátást nyerjünk a felmerülő biztonságtechnikai kérdésekre. A titkos kulcsot nem adhatjuk ki a kezünkből, a tanúsítványról márpedig csak annak segítségével lehet eldönteni, hogy tényleg hozzánk tartozik-e.

A megoldás ilyen esetekben az, hogy egy új kulcspárt generálunk és ennek segítségével létrehozunk egy rövid – tipikusan 12-24 óra – érvényességű helyettesítő tanúsítványt, úgynevezett *proxy*-t. A proxy nyilvános kulcsát aláírjuk a saját titkos kulcsunkkal, ez a CA aláírásához hasonlóan ellenőrizhető, viszont a titkos kulcsot nem lehet belőle rekonstruálni, így ha mellékeljük hozzá a mi nyilvános kulcsunkat, amelyet a CA aláírt és a szolgáltató gép is bízik az adott CA-ban, akkor a CA nyilvános kulcsával ellenőrizheti a mi tanúsítványunk hitelességét és a mi tanúsítványunk nyilvános kulcsával a proxy hitelességét. A proxy titkos kulcsát jelszavas vé-

delem nélkül, de a lehetőségekhez mérten – például az operációs rendszer által biztosított kizárólagos olvasási joggal – védetten lehet tárolni, így a proxy létrejötte után jelszó gépelése nélkül tudjuk magunkat azonosítani, illetve az általunk létrehozott feladat tudja saját magát azonosítani az ezt igénylő szolgáltatások felé.

Felmerül a kérdés: mi értelme ennek a köztes lépésnek? Maga a proxy kompromisszum a biztonság és a használhatóság között: ésszerű kockázat mellett mentesít a jelszó sokszori begépelése alól. A proxy érvényességi idejének rövidre vételével csökkenthető a vállalt kockázat, hiszen a proxy birtokosa teljes mértékben gyakorolhatja a proxy kibocsátójának jogait; másrészt viszont, ha a proxy lejár, mielőtt a feladat lefutna, az ütemező törli a munkánkat, mivel érvényes proxy nélkül nem vagyunk jogosultak a rendszer használatára.

Ennek a kockázatnak a mértékét tovább lehet csökkenteni speciális proxy tanúsítványokkal, amelyek csak bizonyos állomásokra vagy bizonyos feladatokra érvényesek, de ezek használatával az általános felhasználhatóságot is jelentősen megszorítja az adott grid-rendszer. Ebben a témakörben az aktuális fejlesztések a megszorítások és a biztonság elfogadható kompromisszumát keresik.

5. Szolgáltatási biztonság

Attól a ponttól, hogy beküldtünk egy munkát egy grid-rendszerbe, nincs módunk a további sorsának irányítására. Egyetlen módosítási lehetőség van csak: a feladat törlése a rendszerből (noha bizonyos gridek lehetővé teszik egy feladat szüneteltetését is, ez nem változtat a későbbi lefutásán). Az elvárás tehát az, hogy ha egy feladat bekerült a rendszerbe, akkor fusson is le.

Mivel a rendszer jóval összetettebb, mint egy helyi klaszter vagy szuperszámítógép, egy-egy hiba esetenként előfordulhat. Ezen ritka hibák kiszűrésére a „production” színvonalú grid-rendszerek egytől egyig megvalósítottak valamilyen monitorozó rendszert, amely akár a különböző szolgáltatások logjait elemezve, akár bizonyos időközönként beküldött teszt feladatokkal, vagy a komponensek működését külön-külön ellenőrző funkcionális tesztekkel vizsgálja, hogy mennyire megbízható az infrastruktúra.

Érdekességként megemlítendő, hogy az esetlegesen észlelt hiba sokszor nem a grid-infrastruktúra rendellenességét mutatja, hanem gyakran a hálózat ideiglenes zavarából vagy épp a tesztelő program helytelen működéséből származik.

6. Összefoglalás

Összességében elmondhatjuk, hogy a grid-rendszerekben, mivel természetüknél fogva igen komplexek, mindig könnyen bukkanhat fel hiba, amely biztonsági kockázatot jelenthet még akkor is, ha a felhasználók és a kliensgépek tökéletes viselkedését feltételezzük. Ezért hangsúlyozzuk, hogy minden grid-implementációban nagyon komoly figyelmet kell fordítani a biztonsági kérdésekre, hogy elkerülhetőek legyenek az esetenként igen súlyos incidensek.

Irodalom

[1] Bruce Schneier,
“Secrets & Lies: Digital Security in a Networked World”,
John Wiley & Sons, 2000.