HADMÉRNÖK

XIII. Évfolyam 4. szám – 2018. december

ELOSZTOTT ADATTÁROLÁS EGYES KÉRDÉSEI

CERTAIN ASPECTS OF DISTRIBUTED DATA STORAGE

KRALOVÁNSZKY Kristóf

(ORCID ID: 0000-0002-5560-3525)

Kralovanszky.kristof@d-systems.hu

Absztrakt

A Nemzeti Adatvagyon várható jelentős növekedése, illetve az ezzel azonos megítélés alá kerülő egyéb adatbázisok védelme és integritásának folyamatos biztosítása egyre komolyabb kihívás elé állítja a vonatkozó rendszerek üzemeltetőit. Lehetséges sérülékenységük az általuk hordozott érték bővülése és az okozható kár nagysága miatt egyre jelentősebb. Jelen írás célja megvizsgálni, hogy egy elosztott adattárolási koncepció miként növelheti mindezen adatok racionális védelmét, illetve a rendszer kritikus részei vonatkozásában hogyan képes magasabb szintű redundanciát biztosítani - megőrizve a dinamikus méretezhetőség lehetőségét és a prioritási rendszer rugalmasságát.

Kulcsszavak: *elosztott adattárolás, kritikus infrastruktúra, kritikus információs infrastruktúra, adatmentés, redundancia*

Abstract

The expected rapid growth of Hungary's National Databases and the protection of similar databases result in complex challenges for systems operators. Their vulnerability poses a growing risk due to their increasing value plus the possible damage that can be caused by their failure. The goal of this article is to show how distributed data storage can increase the protection of such data and also to demonstrate how higher levels of redundancy can be assets to critical data infrastructure – with maintaining the possibility of dynamic scaling and priority flexibility.

Keywords: *distributed data storage, critical infrastructure, critical information technology infrastructure, data backup, redundancy*

A kézirat benyújtásának dátuma (Date of the submission): 2018.06.02.
A kézirat elfogadásának dátuma (Date of the acceptance): 2018.06.26.

BEVEZETÉS

Mai világunkban egyre jelentősebb az infokommunikációs eszközökre való utaltságunk, így az azoknak való kiszolgáltatottságunk. Közhely ugyan, de gondoljuk végig, mi lenne akkor, ha a különböző közszolgáltatók a fontosabb számítógépeiket (értsük itt az ipari célszámítógépeket és szervereket is) lekapcsolnák. Sokat nem kell időznünk, kiderülne, hogy sem áram nem lenne, melynek következtében sem víz-, sem gázszolgáltatás nem lenne, sem a pénzünkhöz nem férnénk hozzá és erről elektronikusan (telefonon, e-mailben, stb.) senkit nem tudnánk tájékoztatni. Nyilvánvaló ebből a rövid gondolatból is, hogy ezen alapszolgáltatásoknak a fenntartása megkérdőjelezhetetlen feladat. Nehezebb kérdés már, hogy kinek a feladata? Rövid választ erre nem lehet adni, különösen nem jelen írás keretei között, de leszögezhető, hogy az egyik komponens feltétlenül a megfelelő törvényi szabályozás, melynek alapja a fogalmak tartalmi rendezése. Olyan egyszerűnek tűnő fogalmak értelmezése, mint a kibertér, komoly problémát okoz, hiszen nincs rá nemzetközileg elfogadott, (közel) egységes meghatározás. Hazánkban létezik természetesen törvényi meghatározása a kibertér fogalmának [1], de ahogy Munk Sándor is írja, a szakmán belül használt fogalmi értelmezések jelentősen eltérnek.[1] [2] [2]

A lényegesen komolyabb probléma, hogy amíg a fogalmi meghatározásokat és a jogi környezetet igyekszünk létrehozni, addig az a valóság, melyeket így leírni próbálunk, létezik, működik, jelentős hatással van ránk és működési rendellenességeit percek, másodpercek alatt érzékeljük és elszenvedjük. Nincs tehát időnk arra, hogy a hagyományos értelemben először a fogalmakat tisztázzuk, aztán megvizsgáljuk az aktuális működést, majd értelmezzük a ránk gyakorolt hatását, utána pedig szabályrendszert hozzunk létre mindennek kezelésére. Átgondoltan, de párhuzamosan kell mindezt megtennünk és folyamatosan idomulnunk kell a valóság rendkívül dinamikus változásaihoz, mely változásokat részben a fogalmi meghatározások és az általunk létrehozott szabályok szinte valós időben befolyásolnak. Legfontosabb a folyamatokat megismernünk, mert csak akkor leszünk képesek irányítani. Ellenkező esetben a folyamatok fognak uralkodni rajtunk és akkor a csatát biztosan, de lehet, hogy a háborút is elveszítettük.

Írásunk célja, hogy a kritikus infrastruktúrák infokommunikációs rendszereinek és szolgáltatásainak egyik fontos hátterét, az adattárolást és annak biztonságát (elérhetőségét) vizsgálja, feltételezve, hogy a jelenleg érvényes törvényi szabályozása, kategóriákba sorolása felülvizsgálatot igényel és koncepcionális döntésekre van szükség a valódi biztonságuk megteremtéséhez.

FOGALMI ÁTTEKINTÉS

Rendkívül sokszor és sok helyen lehet olvasni elosztott adattárolásról, felhő alapú modellekről, kritikus infrastruktúrákról és kritikus adatállományokról (adatbázisokról). Számtalanszor előfordul az is, hogy a szerzők mást és mást értenek e fogalmakon, így annak érdekében, hogy egyértelmű legyen az írás tartalmi része, szeretném definiálni a fogalmakat, hogy saját értelmezésemben mit takarnak.

---

[1] A cikk szerzője is arra a megállapításra jut, hogy nincs kizárólagosan megfelelő meghatározás a kibertér fogalmára.

[2] Haig Zsolt megfogalmazásában a kibertér: „Ember által mesterségesen létrehozott dinamikusan változó tartomány, amelyben az információ gyűjtését, tárolását, feldolgozását, továbbítását és felhasználását végző, egymással hálózatba kapcsolt, és az elektromágneses spektrumot is felhasználó infokommunikációs eszközök és rendszerek működnek, lehetővé téve ezzel az emberek és a különféle eszközök közötti folyamatos és globális kapcsolatot.” [18]

Kritikus infrastruktúra tartalmi meghatározása is rendkívül nehéz feladat, számos tanulmány és szakmai cikk foglalkozott a kérdéssel. Definícióik azonban megegyeznek abban, hogy ezen infrastruktúrák működése a társadalom egészének működőképessége szempontjából alapvető és elengedhetetlen.[3] Kritikus infrastruktúrák mellett fontos említést tenni sebezhető infrastruktúrákról is, melyek meghatározott elv szerint csoportba rendezve már kritikusnak tekinthetőek. [3]

Tekintsünk azonban a fogalom határaira és vizsgáljuk meg az infrastruktúra fogalmát. Itt az előbbinél is nehezebb helyzetbe kerülünk, hiszen általában fizikai eszközt, intézményt és/vagy azok rendszerét írja le egy-egy meghatározás. [4] 2017. november 1-el elindult EESZT[4] rendszer strukturált adatbázisban gyűjti és tárolja Magyarország lakosságának egészségügyi adatait. Kérdés, hogy önmagában egy ilyen adatbázis tekinthető-e kritikus információs infrastruktúrának? A válasz véleményem szerint egyértelműen igen, ha egy fizikai szerveren lévő adatbázisról beszélünk. Mivel ekkora mennyiségű adat elhelyezése egy szerveren nyilvánvalóan nem lehetséges, így szerverek csoportját kell értenünk az üzemelés alapjaként, de még ebben az esetben is igaz az állítás, hogy természetesen ez is kritikus információs infrastruktúra.

Annak érdekében, hogy a rendszer üzemelésének biztonságát tovább növeljük, logikusan adódik a kiszolgáló szerverek fizikai elkülönítésének szükségessége is, hiszen ahogy a fentiekben is láttuk, a kritikus infrastruktúrák működőképességének biztosítása alapvető és elengedhetetlen. Tételezzük fel tehát, hogy egy „A” és egy „B” jelű helyen, egymással adatkapcsolatban állva, de földrajzilag elkülönülve működik egy-egy szerver csoport, melyek feladata az EESZT adatbázisának tárolása és futtatása. Tekinthetünk-e az „A” és „B” helyekre, mint kritikus információs infrastruktúrákra? Igennel kell válaszolnunk, ha „A” és „B” egyidejű működése szükséges a rendszer üzeméhez. Eddigi meghatározásaink alapján nem egyértelmű a válasz azonban akkor, ha „A” és „B” képes önállóan is működésre és meghatározott szabály alapján dől el, hogy melyik az, amelyik a felhasználókat kiszolgálja (de egyidejűleg csak egy működik).

Vizsgáljuk meg tehát, hogy hogy „A” és „B” helyszínek mitől tekinthetőek kritikus információs infrastruktúrának? Elsősorban attól, hogy olyan rendszert (szoftvert és adatbázist) tárolnak és futtatnak, melynek működése a társadalom egészének működőképessége szempontjából alapvető és elengedhetetlen. A futtatást megelőzően azonban „csak” rendkívül értékes hardverek csoportjáról beszélhetünk, melyek paramétereik alapján meghatározott feladat végrehajtására képesek. Vizsgáljunk meg egy olyan lehetőséget, ahol a már működő rendszeren egy fizikai sérülés miatt az adatbázis használhatatlanná válik. Érték marad-e a hardver? Önmagában nem, hiszen adattárolásra alkalmatlanná vált és a rajta lévő adatok nem elérhetőek.[5]

Megállapítható tehát (az előbbi példa vonatkozásában), hogy elsődlegesen a valódi kritikus információs infrastruktúra maga az EESZT adatbázis és az ahhoz való hozzáférést biztosító

---

[3] Hatályos honi szabályozás szerint a létfontosságú rendszerelem (nemzetközileg elfogadott elnevezés szerint kritikus infrastruktúra) meghatározása jogi értelemben a 2012. évi CLXVI. törvény alapján történik.

[4] EESZT - Elektronikus Egészségügyi Szolgáltatási Tér. Az egészségügyi és a hozzájuk kapcsolódó személyes adatok kezeléséről és védelméről szóló 1997. évi XLVII. törvény (Eüak.) III/A. fejezet, valamint a személyazonosító jel helyébe lépő azonosítási módokról és az azonosító kódok használatáról szóló 1996. évi XX. törvény 23. § k), l), m) pont a TAJ szám működtető általi több célból történő kezeléséről szóló törvények alapján létrehozott, az Elektronikus Egészségügyi Szolgáltatási Térrel kapcsolatos részletes szabályokról szóló 39/2016. (XII. 21.) EMMI rendeletben meghatározott módon működő komplex rendszer

[5] Természetesen a példa rendkívül leegyszerűsített és kisarkított, hiszen egy ilyen adattároló rendszer esetében minden egyes merevlemez tartalmazhat személyes és különleges adatot, melyet a törvény véd. Ebből fakadóan egyértelmű, hogy elemekként továbbra is a legszigorúbb adatvédelmi előírások szerint kezelendőek.

szoftver egysége. Nyilvánvaló azonban az is, hogy mindaz a hardver, ami ennek üzemelését biztosítja, ugyanúgy kritikus információs infrastruktúra, de nem önmaga okán, hanem a „rajta" üzemelő szoftver és adatbázis miatt, vagyis hordozóként. Jelen példánál már indokolt a kritikus infrastruktúra megnevezés is, a lakosság nagy részének érintettsége miatt. [5]

A NEMZETI ADATVAGYON ÉS ANNAK VÉDELMÉNEK TÖRVÉNYI HÁTTERE

Magyarországon a 2010. évi CLVII. törvény[6] bevezette a Nemzeti Adatvagyon fogalmát, ám rendkívül sajnálatos módon semmilyen formában nem rendelkezik annak fizikai- és kiber vonatkozásban értelmezett strukturális védelméről. Igaz ugyan, hogy a 2013. évi L. törvény[7] ezt a hiányosságot részben kezeli, de míg a 2010. évi CLVII. törvény az adatbázis és a rendszer oldaláról közelíti a kérdést, addig a 2013. évi L. törvény egészen kitűnően, de szervezet oldaláról teszi ugyanezt. Messzemenően kívánatos lenne az előbbi okokból a Nemzeti Adatvagyon fogalmi bővítése és védelmének teljesebb rendezése.[8] Sajnos sem az EESZT, sem bizonyos egyéb (akár minősített) adatbázisok nem részei ezen adatvagyonnak, így jelenleg jogszabályi szinten az adatkezelő oldaláról történő megközelítés dominál.

Megvizsgálva a közigazgatásban az egyre növekvő ütemben keletkező és sokszor csak elektronikusan rendelkezésre álló dokumentumokat, elgondolkodtató, hogy azok újbóli előállítása egy esetleges adathordozó, vagy mentés sérülése esetén mennyiben lesz lehetséges. Annak az oka, hogy ez egyáltalán kérdésként merülhet fel sajnos jóval mélyebb, mint a dokumentum készítőjének információbiztonsági képzetlensége. A Nemzeti Adatvagyon példáján is jól látszik, hogy a jogalkotó konkrét rendszereket nevez meg és nem adattartalom alapján végzi a kategóriákba sorolást. Egyértelműen látszik, hogy a jogszabályi háttér lemaradása az információs társadalom rendkívül dinamikus fejlődéséhez képest egyre növekszik, így mind égetőbb kérdéssé válik az adatbázis rendszerek megfelelő jogi oltalmának biztosítása is. [6]

TÁMOGATÓ RENDSZEREK

Adatbázisok csak egy részei azonban a kritikus információs infrastruktúráknak. Tekintsünk a létfontosságú rendszerek és rendszerelemek felé, ott is különösen a kliens oldali üzemeltető szoftverekre. Egyre inkább elfogadottá válik, hogy rendszerekről is (nem csupán adatokról) mentéseket kell végezni, egy esetlegesen bekövetkező sérülésből való mihamarabbi üzemképes visszaállás érdekében. Példaként vizsgáljunk egy kórházi információs rendszert, mely a betegek adatait kezeli és szakmai vonatkozásban üzemelteti az intézményt. Itt jellemzően az adatbázis rész és a kezelőfelületi rész különválik egymástól, de az adatbázis integritásának megmaradása még nem jelenti azt, hogy ahhoz a hozzáférés a felhasználói oldalról biztosított marad. Egy nem célzott, hanem felhasználói gondatlanságra visszavezethető zsarolóvírus támadás következtében az adott intézmény több napig üzemképtelen marad, de nem elsősorban azért, mert az adatbázisai sérültek, hanem mert a felhasználói oldalon üzemelő szoftvereket újra kell telepíteni, általában az operációs rendszerrel (és annak összes biztonsági javításával) együtt. E feladat végrehajtása biztonságos formában tárolt munkaállomás mentésekről az idő töredékét veszi igénybe és napok helyett órákon belül elindulhat az intézményben a legfontosabb munka.

[6] A nemzeti adatvagyon körébe tartozó állami nyilvántartások fokozottabb védelméről
[7] Az állami és önkormányzati szervek elektronikus információbiztonságáról
[8] Reményeink szerint azonban a Törvényalkotó az elmaradást felismerve a közeli jövőben elvégzi majd a szükséges módosításokat, hogy a valóban kritikus adatbázisok teljes köre megfelelő törvényi védettséget élvezzen.

Tegyük vizsgálat tárgyává a rendszereket összekötő gerinchálózati átviteltechnikai eszközöket, melyek már önmagukban is kritikus információs infrastruktúrának minősülnek. Nézzük meg, hogy lehetséges-e olyan állapot, amikor a kiszolgált infrastruktúrák rendben működnek, de képtelenek kommunikálni egy magasabb vezetési/vezérlési szinttel? A válasz itt is igen, hiszen a piacon domináns szerepet betöltő Cisco aktív hálózati eszközökre általában 4-6 havonta érkezik olyan javítás, mely valamilyen közepesnél magasabb sebezhetőséget orvosol.[9] Ezek telepítése során előfordulhat olyan hiba, mely a konfigurációs beállítások (statikus útvonalak, port továbbítások, IP cím engedélyek, stb.) elvesztésével jár, de mely nélkül az eszköz az adott rendszerben alapfeladatát ellátni képtelen. A megoldás itt is az ismerten jó konfigurációk mentése. [7] [8]

Megállapítható tehát, hogy a védendő adattartalomba, mint kategóriába elengedhetetlenül bele kell foglalni a kritikus információs infrastruktúraként viselkedő adatbázisokat, a kritikus infrastruktúrák üzemelését biztosító rendszerek adat- és rendszer mentéseit, illetve az átviteltechnikai eszközök beállításainak mentéseit is.

ELOSZTOTT ADATTÁROLÁS

Hogyan lehetséges ezen adattartalom leghatékonyabb védelme? Mielőtt a technikai megoldás lehetőségeit vizsgálnánk nagyon fontos, hogy e vizsgálat csak gazdasági racionalitás mentén lehetséges. Nyilvánvaló ugyanakkor az is, hogy a védelem minden esetben a védendő értékkel áll kapcsolatban. Nehéz ennek az értéknek a meghatározása, mert az nem mérhető piaci erőben, (újra)előállítási költségben, vagy egyéb a gazdaságban általában elfogadott módon. Kivétel nélkül olyan adatokról beszélünk, melyek részben, vagy egészben történő kompromittálódása hazánk nemzetbiztonságát veszélyezteti – akár önmagukban, mint adatok, akár a kritikus infrastruktúrák üzembiztonsági kockázataként. [6]

Tekinthetjük-e az adattárolás egy adott (egyedüli) helyét biztonságosnak, függetlenül attól, hogy földrajzi helyszínét tekintve hol valósul meg? Nem, hiszen szinte bármilyen objektumban lehetséges a külső, erőszakos fizikai károkozás, ezért jött létre a redundáns adattárolás, mint rendszer. A lehetséges megoldás egyik alapvetése tehát az adattárolás többszörözése – ez az, ami az adattárolást biztonságossá teszi. A redundancia minden esetben valamilyen formájú elosztott adattárolást jelent, ahol a valódi különbséget az elosztás térbeni (tágabb értelemben földrajzi) megvalósulása, a konfiguráció szervezési elve és az ezekből adódó (számítható) redundancia szint fogja jelenteni.

Az elosztott adattárolás tehát az adatok sértetlenségének (integritásának) és rendelkezésre állásának, így az információbiztonságnak is alappilére.[10] [9] [10] Maga az elosztás egyik lehetséges, de talán legfontosabb kérdése, hogy eszközön belüli, tárolási helyszínen belüli, vagy tárolási helyszínek közötti elosztásról van-e szó. Az eszközön belüli elosztásról Big Data[11]

[9] Az adott eszköz 5 éves életciklusában, a bevezetéstől számítva, a kiadott javítások átlagában. A példa a Cisco 5500-as sorozatú tűzfal/router eszközére vonatkozik - https://tools.cisco.com/security/center/publicationListing.x?product=Cisco&title=5500&sort=-day_sir#~Vulnerabilities (letöltés ideje: 2017. november 5.)

[10] Az információbiztonság harmadik alappillére, a bizalmasság itt csak áttételesen jelenik meg, mert az elsősorban szabályozási és authentikációs kérdés. Sérülhet ugyanakkor a bozalmasság elve, ha az adott adatbázis hozzáférési / jogosultsági szabályai sérülnek, illetve a támadónak kifejezetten az a célja, hogy egy jóval a megkívánt jogosultsági szintet alulmúló belső személy számára tegyen adatokat hozzáférhetővé.

[11] Nehéz pontos definíciót találni a Big Data fogalomra, számos tanulmány foglalkozik csak a téma megfogalmazásával. Illés Noémi őrnagy „A Nagy Adat; a Nagy Adat nemzetbiztonsági aspektusai” című írásában átfogóan foglalkozik a fogalmi kérdésekkel és azok kapcsolataival.

esetekben nem lehet beszélni, hiszen egy RAID[12] tömb adattárolási képessége a 100 TB-os nagyságrendnél jellemzően megáll. [11] [12] Fölötte már dedikált adattároló eszközöket alkalmazunk, onnantól pedig nem eszközön belüli adattárolásról van szó. Helyszínen belüli redundancia esetünkben szintén nem képezi további vizsgálat tárgyát, hiszen megállapítottuk, hogy az egyedüli adattárolási helyszín nem tekinthető biztonságosnak. [13] A továbbiakban tehát a több helyszínű, vagyis georedundáns tárolást, mint az elosztott modell legfelső szintjét vizsgáljuk koncepcionálisan. [14]

A földrajzilag elkülönülő adattárolások hátrányait érdemes először számba venni, hiszen csak azok ismeretében értékelhető a valós előnyük. Az adatok konzisztenciájának különböző tárhelyek közötti biztosítása egy plusz feldolgozási rétegként jelenik meg, mely réteget minden helyszínen üzemeltetni kell. Adatmozgás függvényében többszörözni kell a rendelkezésre álló adatátviteli csatornákat, hiszen a helyszíneket egymással szinkronban kell tartani, illetve egy két-helyszínű modell esetében bármelyik helyszínnek képesnek kell lennie a teljes rendszer kiszolgálására. A többszörözött infrastruktúra megnövekedett kezelőszemélyzetet, megnövekedett üzemeltetési költségeket és jelentősen megnövekedett kezdeti költségeket jelent. Számos esetben a hátrány azonban előnyként viselkedik, mert csak e hátrányokon keresztül érhető el a redundancia mint cél. Ezért helyénvalóbb költség oldalról közelíteni és azt mondani, hogy vannak költségesebb és kevésbé költséges megoldások. A bekerülési költség azonban mindig csak a konkrét védendő adattömeg értékének vizsgálatát és a lehetséges fenyegetettségek átfogó kockázatelemzését követően értelmezhető, nélkülük soha.

TÁROLÁSI MODELLEK

Mindezek függvényében eldöntendő az is, hogy két, vagy több helyszínes modellt választunk-e. Vezessünk be erre egy új fogalmat[13], a GRAIS-t, ami a Geo-Redundant Array of Interdependent Sites[14] kezdőbetűit veszi alapul és részben vegyük át a RAID tömbökben ismert számozási rendszert: különböztessünk meg tehát GRAIS-1 és GRAIS-N és GRAIS-ML rendszereket:

- GRAIS-1 legyen az a rendszer ahol két földrajzi helyszín egymás „tükreként” működik, vagyis mindkettő képes önálló működésre. Beállítás kérdése, hogy egy vezérlő dönti-e el, hogy a két helyszín közül van-e elsődleges és a másik folyamatosan hozzá szinkronizál, vagy intelligens terheléselosztás alapján történik a helyszínekről a kiszolgálás. Ez utóbbinak egy nagyobb (országos) rendszer esetén lehet hálózati útválasztás szempontjából előnye.
- GRAIS-N[15] rendszer esetében három, vagy annál több földrajzi helyszínről van szó, ahol egyik sem képes egyedül, önállóan működni és a helyszínek számának, valamint a redundancia logikájának szintjéből adódik, hogy mennyi helyszín kiesését viseli még el a rendszer.
- GRAIS-ML[16] megoldásnál pedig rétegezett modellről van szó, vagyis több adatközpont létezik, melyek több, egymástól független rendszer tárolását végzik, ahol bizonyos részek lehetnek GRAIS-1, mások GRAIS-N típusúak.

[12] RAID – Redundant Array of Inexpensive Disks – több olcsó fizikai merevlemez egy logikai lemezzé szervezése, melyen utána ugyanúgy partíciók hozhatók létre, mint egy fizikai lemezen.
[13] A fogalom bevezetésének célja, hogy a későbbiekben a különböző georedundáns tárolási topológiákról egyértelműbb és egyszerűbb leírást lehessen adni.
[14] A helyszínek között interdependencia úgy értelmezhető, hogy egészséges rendszer esetén a szinkronban tartás egymástól függ, GRAIS-1 rendszereket kivéve pedig csak egymás segítségével tud a kiszolgálás megtörténni.
[15] N – a GRAIS-1-től való megkülönböztetést jelenti és többféle modellt követhet, pl. RAID-5, RAID-6 típust.
[16] ML – Multi Level – az adattárolás többszintű megvalósítását jelezve

Mindhárom rendszernek rendkívül nagy előnyei lehetnek egy adott megvalósítás során. GRAIS-1 rendszer szerint működik egy megnevezni nem kívánt ország komplex határrendészeti rendszere, míg GRAIS-N modell alapján üzemelt indulásakor Magyarország egyik mobil távközlési szolgáltatójának SIM[17] nyilvántartó rendszere, három központtal. Rendkívül hosszan lehetne értekezni a különböző modellek előnyeiről és hátrányairól, mely jelen írásnak (terjedelmi okokból) nem célja.

GRAIS-ML rendszerek további magyarázatot igényelnek, hiszen eddig ritkán alkalmazott struktúráról van szó. Példaként tekintsünk egy olyan kormányzati rendszert, mely több, a teljes lakosság adatait tartalmazó adatbázist kezel (népességnyilvántartás, közlekedési nyilvántartás, társadalombiztosítási alapnyilvántartás). Mivel rendkívül nagy adattömegről van szó, a vélelmezhető optimális tárolási modell a GRAIS-1 lesz, az adatok értéke és nemzetbiztonsági jelentősége miatt rendkívül magas fizikai biztonságú tárolással kerül megvalósításra. Ez azt is jelenti, hogy létezik két olyan (földrajzilag eltérő helyen lévő) objektum, ahol e tárolás megvalósul. Mindezzel párhuzamosan gondoljunk egy olyan erőműre, mely az országos villamosenergia szolgáltatásban vesz részt. Ezen erőmű működés-kritikus rendszereinek mentését, mely egy kiber-sérülés utáni visszaálláshoz szükségesek, az adott erőművön kívül is tárolni kell. [15] [16] Ez megoldható több másik erőműben elosztva (GRAIS-N rendszerben) és azon felül a központi kormányzati adattárolás GRAIS-1 rendszerében (részben, vagy egészben).

Így a kormányzati szintű rendszernél a GRAIS-1 adatbázisok mellett megjelenik egy GRAIS-N adatbázis is. Ez történhet akár úgy is, hogy Mindezek mellett a teljes erőművi adatállomány is egy alacsonyabb rendelkezésre állási szinten tárolásra kerül, ezek azonban már tervezési kérdések, melyek igen sok egyéb változó függvényei. [17] [14]

ÖSSZEGZÉS - KÖVETKEZTETÉSEK

A fentiekben leírt struktúrák képesek jelenleg kiszolgálni azt számos, önmagában jelentősen különböző rendszert, melyek adattárolását központosítani, de kialakítását tekintve elosztani kell. Ezek a modellek biztosítják azt a dinamikus növekedési lehetőséget, melynek során további adatterületek bekapcsolása, illetve a tárolás felfelé történő skálázása nem okozza a teljes adattárolási rendszer leállítását vagy ideiglenes szüneteltetését. Ugyanígy lehetővé teszi azt is, hogy a felhasználói oldalon új kliensek hozzáadásával zökkenőmentes üzem biztosítható minden már meglévő felhasználó számára is. A tárolási/kiszolgálási erőforrás oldalon biztosítja az ideiglenesen megnövelt kapacitási igények biztosítását (például választások idején) további kiszolgáló fürtök (és szükség esetén átmeneti adattárak) bekapcsolásával.

Jelen írás célja néhány elvi koncepció lehetőségeinek ismertetése volt. A példaként használt kritikus (információs) infrastruktúrákon keresztül bemutattam, hogy maga az adatbázis önmagában is tekinthető kritikus infrastruktúrának és e minőségénél fogva vonatkoznak rá mindazon védelmi intézkedések, melyek a klasszikus, fizikai valójukban is megjelenő kritikus infrastruktúrákat megilletik. Mivel az ilyen adatbázisok folyamatosan és dinamikusan változnak, a sértetlenség és rendelkezésre állás feltételeit racionális keretek között csak elosztott adattárolás segítségével lehet biztosítani. Ezt követően bevezettem a GRAIS adattárolási modellek fogalmát és igazoltam, hogy kormányzati Big Data rendszereknél a jelenleg ismert és alkalmazható utak közül heterogén környezetben használható alternatíva jellemzően nem található.

[17] SIM – Subscriber Identification Module – mobiltelefonok üzemeltetéséhez elengedhetetlen, az előfizetőt azonosító chipkártya (ami újabb készülékekben már a telefonba beépítetten is megtalálható). Tárolja az előfizető International Mobile Subscriber Identity (IMSI) azonosítót.

FELHASZNÁLT IRODALOM

[1] 2013. évi L. törvény az állami és önkormányzati szervek elektronikus információbiztonságáról. Available at: https://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=a1300050.tv. (Accessed: 19th November 2017)

[2] MUNK, S: A kibertér fogalmának egyes , az egységes értelmezést biztosító kérdései. *Hadtudomány*, (2018), 113–131 doi:10.17047/HADTUD.2018.28.1.113

[3] KOVÁCS, L.: Kritikus információs infrastruktúrák Magyarországon. *Hadmérnök*, Robothadviselés konferencia különszám (2007), 1–20

[4] NAGYNÉ TAKÁCS, V.: Adatvagyon-gazdálkodás divatszó vagy szükségszerűség? *Hadmérnök*, 4. (2015), 166–176

[5] NYITRAI, M.: Összehasonlító tanulmány az Európai Unió és az Egyesült Államok kritikus infrastruktúra védelem szabályozása és megvalósítása területein. *Hadtudományi Szemle*, 2 (2017), 232–253

[6] FLEINER, R.: *Az adatbázis-biztonság szerepe és megvalósításának feladatai a kritikus információs infrastruktúrák védelmében*. Budapest: Zrínyi Miklós Nemzetvédelmi Egyetem, 2011

[7] FEHÉR, J.: A rendőrségi informatikai hálózatok információbiztonsági hátterének meghatározása. *Hadmérnök*, 2. (2016), 133–144

[8] JOBBÁGY, S.: A Magyar Honvédség kormányzati célú elkülönült hírközlő hálózata. *Hadmérnök*, XII (2017), 223–236

[9] SZÁDECZKY, T.: E-kormányzati szolgáltatások kommunikációbiztonsága. *Hadmérnök*, XII (2017), 280–289

[10] RÓZSA, T.: *Az információs műveletek vizsgálata, különös tekintettel a befolyásolási képességek alkalmazásának lehetőségeire a Magyar Honvédség feladatrendszerében*. Budapest: Nemzeti Közszolgálati Egyetem, 2016

[11] A Big Data, mint a kritikus infrastruktúra része. in *Báthori-Brassai nemzetközi konferencia előadásai I.- kötet* (eds. Prof. Dr. Rajnai, Z., Dr. Fregan, B., Marosné, K. Z. & Dr. Ozsváth, J.)

[12] ILLÉS, N.: A Nagy Adat, a Nagy Adat nemzetbiztonsági aspektusai. *Szakmai szemle*, 4 (2016), 101–135

[13] What Is A Distributed Database? And Why Do You Need One? *Cloud Strategy* Available at: https://dzone.com/storage/assets/6180392-why-you-need-a-distributed-database.pdf. (Accessed: 19th November 2017)

[14] The Guiding Principles for Cloud-scale, Geo-distributed Databases. Available at: http://www.databasejournal.com/sqletc/the-guiding-principles-for-cloud-scale-geo-distributed-databases.html. (Accessed: 19th November 2017)

[15] KOVÁCS, L., KRASZNAY, C.: Digitális Mohács 2.0: kibertámadások és kibervédelem a szakértők szerint. *Nemzet és Biztonság*, 1 (2017), 3–16

[16] SÁGI, G.: Megvédhetőek-e a kritikus információs infrastruktúrák? *Hadmérnök*, 2 (2016), 154–169

[17] FALEIRO, J.M., ABADI, D.J.: FIT: A Distributed Database Performance Tradeoff. *Data Engineering*, 1 (2015), 10–17

[18] HAIG, ZS.: *Információs műveletek a kibertérben*. kézirat,