

Informatikai technológiák laboratórium 2.
(BMEVIMIA429)

Felhőalapú szolgáltatások vizsgálata
(Mérési segédlet)

Szatmári Zoltán, Tóth Áron
Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs Rendszerek Tanszék

2014. november 19.

Contents

1	Bevezetés	2
2	Szükséges előismeretek	3
3	Felhőalapú technológia	3
4	Amazon Web Services	4
4.1	Szolgáltatások	5
4.1.1	Elastic Compute Cloud (EC2)	5
4.1.2	Simple Storage Service (S3)	6
4.1.3	CloudFront	6
4.1.4	Relational Database Service (RDS)	7
4.1.5	Auto Scaling	7
4.1.6	Elastic Beanstalk	8
4.2	A hozzáférés szabályozása	8
5	A mérésről	8

1 Bevezetés

Az informatikai szolgáltatások komplexitása és architektúráis igénye folyamatosan növekszik a technológia és a piaci igények fejlődésével. A növekvő technológiai komplexitás egyre mélyebb és szerteágazóbb szaktudást követel meg egy-egy végfelhasználókat is elérő szolgáltatótól.

Tekintsük a Google vagy akár Facebook szolgáltatók példáját: már maga csak az online felület, a böngészőnkben megjelenő HTML, JavaScript és CSS kód sem egyszerű, de vajon milyen infrastruktúra szolgálja ezeket ki? Milyen számítógépek, milyen informatikai hálózat, milyen tároló rendszerek állnak a háttérben?

Emlékezzünk vissza az Intelligens rendszerfelügyelet tantárgy kereteiben tanult különböző témakörökre! Milyen szintű és rendszerességű biztonsági mentést alkalmaznak? Hogyan biztosítják a szolgáltatásaik monitorozását, a tartalmak minden földrészben történő elérhetőségét? Hogyan működik a felhasználók hitelesítése? Vajon ezek a nagy szolgáltatók milyen szinten foglalkoznak az egyes hardverhibákkal?

Lássuk be, hogy ha nem vagyunk a példában emlegetett szolgáltatókkal egy súlycsoportban, de mégis valamilyen téren velük összemérhető szolgáltatást szeretnénk kiépíteni önerőből, akkor nagyon szerteágazó ismeretekre és sok-sok pénzre van szükségünk. Vegyük észre, hogy még mindig csak rendszertervezési és üzemeltetési kérdéseket feszegetünk, de a cégvezetést, az emberi erőforrások kezelését, technikai terméktámogatást és hasonló területeket nem is érintettük.

A szolgáltatók korábban a szolgáltatások kiépítéséhez és üzemeltetéséhez a teljes informatikai infrastruktúrát felépítették, infrastrukturális bemeneti szükségletként többnyire csak az elektromos áram, az internetkapcsolat és az üzemeltetés fizikai helyszíne jelent meg.

Ez az infrastrukturális alapszükséglet alakul át napjainkban egyre komplexebb szolgáltatásokra való igényné, azaz jellemzően már nem az áram és az internetkapcsolat, hanem magasabb szintű szolgáltatások, kiszolgálók vagy akár különböző platformok szolgálnak alapul egy-egy jelenleg népszerű online szolgáltatásnak. Megfigyelhető, hogy ilyen formában több szintre tagolódik, azaz specializálódik az informatikai szolgáltatások összessége, és ezek egymásra épülésével alakul ki a végfelhasználóhoz eljutó (többnyire online) tartalom.

Ezen a piacon jelentek meg a felhőalapú szolgáltatást nyújtó iparági szereplők, melyekre jellemző, hogy az előző értelemben vett összetett, de más szolgáltatások számára alapul szolgáló, hálózaton keresztül elérhető szolgáltatásokat kínálnak.

A felhőalapú szolgáltatások igénybevétele legtöbb esetben rugalmasságot és függetlenséget biztosít, de gazdasági oldalról tekintve egy jó üzleti modellel pénzügyileg is megtérülhet.

Gondoljunk bele, hogy egy induló vállalkozás számára mekkora pénzügyi kockázatot és terhet jelent a fizikai infrastruktúra felállítása vagy egy hűségidővel terhelt bérleti konstrukció vállalása! A felhőszolgáltatók hagyományos értelemben vett szolgáltatást nyújtanak, és az igénybevétel ideje, mértéke vagy éppen jellege alapján számlázzák ki a használati díjat. Egy-egy ötlet ebben a formában viszonylag olcsón kipróbálható, nem igényel konkrét beruházást, elköteleződést.

Továbbgondolva a korábbi üzleti modelleket, nem csak az induló vállalkozások, hanem akár régóta működő szolgáltatók is profitálhatnak a szolgáltatás jellegéből. Vegyük példaként Forma 1-hez kapcsolódó médiavállalkozást, amelynek kizárólagos joga a világba-jnokság futamainak élő közvetítése! A világbajnokság jellemzően 18-20 versenyhétvégéből áll, azaz az év 365 napjából mindössze maximum 60 napon van élő közvetítés. Érdeemes-e erre egy kb. 17% kihasználtságú infrastruktúrát fenntartani, vagy igény jelleggel (on demand) inkább csak szolgáltatásként igénybe venni azt?

Ahogy az előző példából is látszik, nagyobb méretű szolgáltatások architektúráis igénye dinamikus jelleget mutathat, melyhez statikus infrastruktúrát fenntartani ma már nem biztos, hogy gazdaságos tevékenység.

Felhőalapú szolgáltatások esetén az igénybe vett erőforrások mértéke legtöbb esetben nagyon finom mértékben hangolható az aktuális terhelésnek megfelelő módon. A gyakorlatban ez azt is jelenti, hogy az előbbi Forma 1-es példában emlegetett heti, kétheti gyakoriságnál lényegesen sűrűbben, akár perces frekvenciával szabályozható az igénybe vett erőforrások mennyisége és az azokra fordított költség.

Az Amazon Web Services, a Microsoft Azure, a Google App Engine és a Rackspace napjaink talán legismertebb „infrastructure as a service” (IaaS) és „platform as a service” (PaaS) szolgáltatói. Ezen szolgáltatásokat veszik igénybe az olyan, jelenleg is népszerű szolgáltatások mint a Reddit, Foursquare, Netflix, Pinterest, TMZ, Instagram és a magyar gyökerekkel rendelkező Prezi.com.

2 Szükséges előismeretek

A mérés során a következő tárgyak során elsajátított előismeretekre építünk, melyek gyakorlati alkalmazása szükséges a mérés elvégzéséhez.

- Számítógép hálózatok:
A TCP/IP alapú hálózatok működése (IP címek, IP cím osztályok, alhálózatok [1])
- Mérés laboratórium 4:
TCP/IP mérés (hálózati diagnosztika eszközök), Alkalmazási réteg mérés (HTTP protokoll, Wireshark eszköz), UNIX/Linux mérés (alapvető Linux ismeretek)
- Intelligens rendszerfelügyelet:
Szolgáltatásbiztonság, rendszermonitorozás, virtualizáció és cloud computing témakörök
- Adatbázisok, Szoftver laboratórium 5:
SQL, PHP alapismeretek

3 Felhőalapú technológia

A „számítási felhők” egy modell, amely lehetővé teszi a hálózaton keresztül való kényelmes és széles körű hozzáférést konfigurálható számítási erőforrások egy megosztott

halmazához (pl. hálózati erőforrások, szerverek, tárolók, alkalmazások és szolgáltatások).

A méréshez való felkészüléshez további **kötelező olvasmány** a *The NIST Definition of Cloud Computing* [2] című dokumentum második fejezete.

Alapvető szolgáltatási típusok:

- *Software as a Service (SaaS)*: a szolgáltató alkalmazásaink használata, jellemzően vékony kliens segítségével. (Pl. Google Apps, Microsoft Office 365)
- *Platform as a Service (PaaS)*: a szolgáltató futtatókörnyezetének (pl. Java, .NET) használata saját alkalmazás futtatására. (Pl. Amazon Elastic Beanstalk, Cloud Foundry, Google App Engine, Windows Azure Compute)
- *Infrastructure as a Service (IaaS)*: alapvető számítási erőforrások igény szerinti foglalása és használata. (Pl. Amazon EC2, Rackspace Cloud, Terremark, Windows Azure Virtual Machines, Google Compute Engine)
- *Datastore as a Service (dSaaS)*: szolgáltató által nyújtott elosztott tárhely használata. (Pl. Dropbox, Google Drive, iCloud, Microsoft SkyDrive, Chubby, Tre-sorit)

Hozzáférés szerint megkülönböztetjük a következőket

- *Public cloud*: publikus számítási felhő, melynek szolgáltatásait bárki igénybe veheti,
- *Private cloud*: privát számítási felhő, melyet tulajdonosai használnak,
- *Hybrid cloud*: hibrid számítási felhő pedig az előző kettő kombinációja

A felhőalapú technológiák használata számos előnnyel rendelkezik, de számos aggódó hangot is hallani a technológiákkal kapcsolatban. A felmerülő problémák jellemzően adatbiztonság és rendelkezésre állás kapcsán kerülnek elő, de a vállalkozások által kezelt érzékeny adatok felhőalapú feldolgozása akár jogi problémákat is felvethet. Fontos megjegyezni a megbízhatósággal, teljesítménnyel és más egyéb SLA-ban foglalt jellemzővel kapcsolatban, hogy a legtöbb esetben oszttott erőforrásokon dolgozunk, így a különböző felhasználók egymásrahatása sehol sem kizárható.

4 Amazon Web Services

Az Amazon napjaink egyik legmeghatározóbb felhőalapú szolgáltatója, termékpalettája lefedi az előzőekben ismertetett IaaS, PaaS, dSaaS területeket. Szolgáltatásait egy jól definiált interfészen keresztül vehetjük igénybe, melyhez a legerterjedtebb alkalmazásfejlesztési környezetekhez (Java, .Net, PHP, SOAP, Bash) biztosítanak programkönyvtárakat. A programozott igénybevételén túl rendelkezésünkre áll egy webes menedzsment felület, melyen keresztül egy böngésző segítségével a legtöbb szolgáltatást grafikus felületen is használhatjuk.

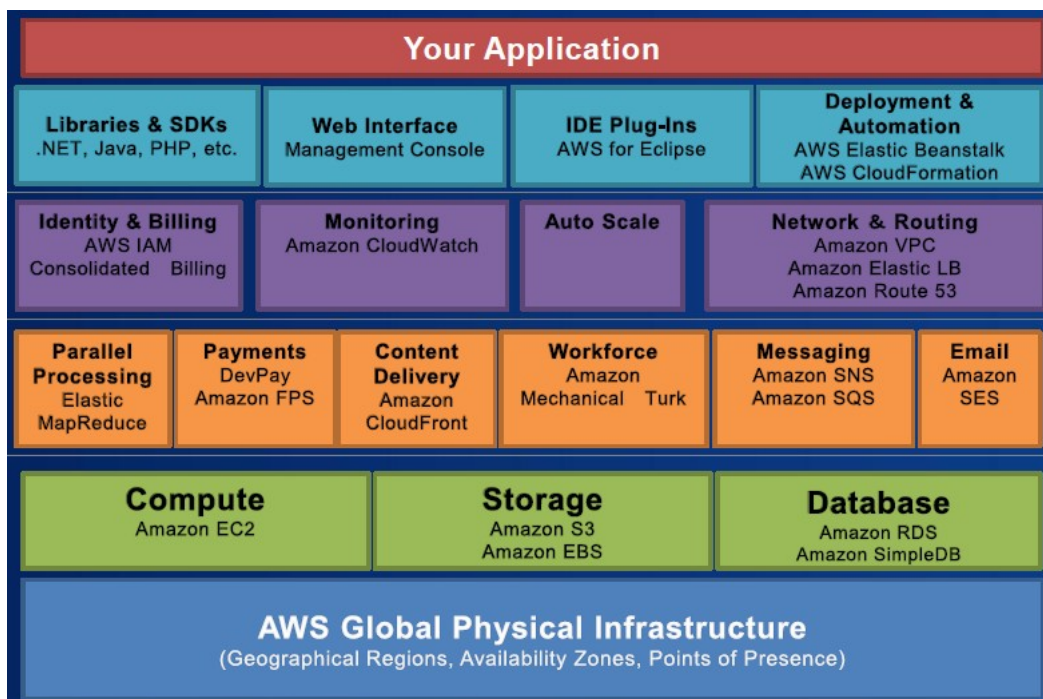


Figure 1: Az AWS szolgáltatások egymásra épülése

4.1 Szolgáltatások

Az alábbiakban röviden áttekintjük a mérésen is felhasznált szolgáltatásokat. A mérésre való felkészüléshez az egyes szolgáltatásoknál hivatkozott külső szakirodalom elsajátítása is szükséges.

Az itt hivatkozott irodalmak megismerését a beugróban is ellenőrizzük. A felsorolt irodalmaknál minden esetben feltüntetésre kerültek a megismerendő fejezetek vagy alfejezetek címei, melyek az adott irodalom PDF verziójában találhatóak meg (a HTML verzió picit más struktúrájú lehet). A hivatkozott irodalmakban található további hivatkozásokon elérhető tartalmak ismerete a beugróhoz nem szükséges, de mélyebb érdeklődés esetén hasznos lehet, vagy a megértést segítheti.

4.1.1 Elastic Compute Cloud (EC2)

Az Amazon IaaS szolgáltatása, melyen keresztül többek között virtuális gépeket igényelhetünk általunk megadott paraméterek alapján. Különböző előre definiált erőforrás-sablonok alapján (instance type), adott régióban van lehetőségünk a számos előre telepített virtuális gép (amazon machine image - AMI) közül példányosítani, és azt használni.

A hálózati beállítások megfelelő megadása (security group) után a Linux-alapú virtuális gépekre ssh kapcsolaton keresztül tudunk belépni ssh kulcsok használatával (key pair).

Lehetőségünk van hálózati terheléelosztó (load balancer) definiálására. Ennek segítségével az egy erre a célra használt IP-címre érkező hálózati kéréseket továbbíthatjuk a beállított terheléelosztási stratégia (pl. round robin) alapján több különböző korábban elindított virtuális gép felé.

További irodalom az Amazon EC2 User Guide [3] következő szakaszai:

- *What Is Amazon EC2?:* alapfogalmak
- *Setting Up:* alapvető objektumok definiálása
- *Getting Started:* EC2 alapvető használata

4.1.2 Simple Storage Service (S3)

Az Amazon dSaaS szolgáltatása, melyen keresztül stabil adattárolót vehetünk igénybe, melyen akár az EC2 virtuális gépeinek életciklusain túlmutatóan is tárolhatjuk adatainkat. Lehetőségünk van tárolók (bucket) létrehozására, melyben fájlokat tárolhatunk, majd a tárolót többek között felcsatolhatjuk az EC2 virtuális gépek fájlrendszerébe (mount).

További irodalom az Amazon S3 Getting Started Guide [4] következő szakaszai:

- *Amazon S3 Basics*
- *Create a Bucket*
- *View An Object*
- *Delete an Object and Bucket*
- *Common Use Scenarios*
- *Considerations Going Forward / Naming Strategy*

4.1.3 CloudFront

Content Delivery Network-nek (CDN) nevezzük azokat az elosztott rendszereket, melyek számos, a világ különböző adatközpontjaiban elhelyezett szerverből épülnek fel, és biztosítják online elérhető tartalmak nagy rendelkezésreállású és teljesítményű kiszolgálását szerte a világon. Lehetővé teszi, hogy mindig a klienshez legközelebbi adatközpontból kerüljön letöltésre az adott tartalom.

A CloudFront az Amazon CDN szolgáltatása, melynek segítségével lehetőségünk van az S3 tárolóiban elhelyezett tartalmakat ilyen formában elérhetővé tenni.

További irodalom az Amazon CloudFront Developer Guide [5] következő szakaszai:

- *What Is Amazon CloudFront and Why Do I Need It?*
- *Getting Started with CloudFront*

4.1.4 Relational Database Service (RDS)

Az Amazon ezen szolgáltatásán keresztül egy általunk kiválasztott implementációjú (MySQL, Oracle, MsSQL, ...) relációs adatbázis szolgáltatást vehetünk igénybe mindenféle szerverkonfigurálás nélkül.

További irodalom az Amazon Relational Database Service Getting Started Guide [6] következő szakaszai (A Getting Started Guide rejtélyes okból lekerült a dokumentáció nyitólapjáról, de alul az archívum linken elérhető):

- *Create a Security Group*
- *Creating a MySQL DB Instance and Connecting to a Database on a MySQL DB Instance*

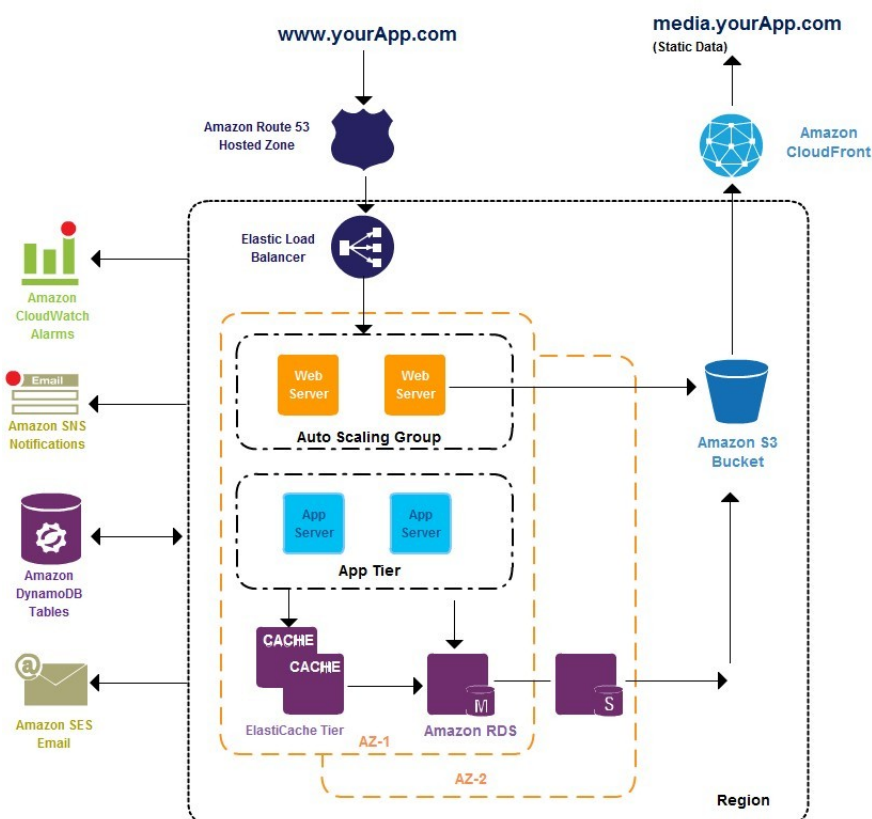


Figure 2: Az AWS szolgáltatásinak felhasználása egy komplex szolgáltatás kialakítására

4.1.5 Auto Scaling

Az Auto Scaling az Amazon egy olyan kiegészítő szolgáltatása, melynek segítségével EC2 virtuális gépeket indíthatunk és állíthatunk le adott felhasználói szabályok, ütemezések és monitoring információk szerint. Tipikus felhasználási eset a terhelésváltozás

dinamikus követése, és például egy terhelésnövekedés esetén a terheléelosztó mögé további szerverek automatikus felállítása.

Az Auto Scaling nem rendelkezik webes menedzsment felülettel, jelenleg csak valamelyik programkönyvtáron keresztül érhető el.

4.1.6 Elastic Beanstalk

Az Amazon PaaS szolgáltatása, melynek segítségével különböző, előre beállított platformokat (Java, PHP, .NET, Ruby, Python) vehetünk igénybe alkalmazásaink futtatására. Valójában az Elastic Beanstalk a háttérben a többi Amazon szolgáltatás felhasználásával automatikusan elindítja és beállítja a kiszolgáló infrastruktúrát, és ezek után a felhasználóra csak az alkalmazás telepítése hárul.

4.2 A hozzáférés szabályozása

Az Amazon Web Services használata során több különböző hozzáférési módszer szóba kerül, ezek a kezdő felhasználók számára könnyen összekeverhetőek. Röviden áttekintjük a használt hozzáférési módszereket:

- E-mail cím és jelszó: belépés a webes menedzsment felületre.
- Access Keys: a programkönyvtárak ezt használják azonosításra. Egy felhasználói fiókhoz több ilyen kulcs is tartozhat.
- Key Pair: SSH kulcs, melynek segítségével az EC2 virtuális gépre léphetünk be.

5 A mérésről

A mérés során az Amazon Web Services IaaS szolgáltató ismertett szolgáltatásait vesszük igénybe egy összetett webes alkalmazás kiszolgálásához. Áttekintjük, hogy milyen lehetőségek adódnak a szolgáltatások kézi, vagy akár programozott igénybevételére, hogyan használhatjuk a mai online szolgáltatásokhoz szükséges korszerű módszereket.

References

- [1] Andrew S. Tanenbaum. *Számítógép-hálózatok*. PANEM, 2006.
- [2] The NIST Definition of Cloud Computing. <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>.
- [3] Amazon Elastic Compute Cloud (EC2) Documentation. <http://aws.amazon.com/documentation/ec2/>.

- [4] Amazon Simple Storage Service (S3) Documentation. <http://aws.amazon.com/documentation/s3/>.
- [5] Amazon CloudFront Documentation. <http://aws.amazon.com/documentation/cloudfront/>.
- [6] Amazon Relational Database Service (RDS) Documentation. <http://aws.amazon.com/documentation/rds/>.