



DNS64 IMPLEMENTÁCIÓK TELJESÍTMÉNYÉNEK ÖSSZEHAISONLÍTÁSA MÉRÉSEKKEL

Dr. Lencse Gábor
egyetemi docens
SZE Távközlési Tanszék
lencse@sze.hu



Tartalom

- A Japánban található NICT StarBED rendszer
- A mérésről
 - vizsgált DNS64 implementációk
 - mérési módszer (RFC 8219)
 - mérési összeállítás
 - mérés végrehajtása
- A mérés eredményei és kiértékelésük
- DNS64 szolgáltatás számítási hatékonysága



A JAPÁNBAN TALÁLHATÓ NICT STARBED RENDSZER

NICT StarBED

- Hogy kerültem én oda?
 - Vendég kutató voltam: 2017. 06. 15. – 12. 15.
 - NAIST: Nara Institute of Science and Technology
 - „IPv6 áttérési technológiák biztonsági kérdéseinek vizsgálata”
 - A DNS64 teljesítőképesség-vizsgálat csak „ráadás” volt 😊
- NICT: National Institute of Information and Communications Technology



- Egyik intézménye:

Hokuriku StarBED Technology Center

NICT StarBED

- Mi van ott?

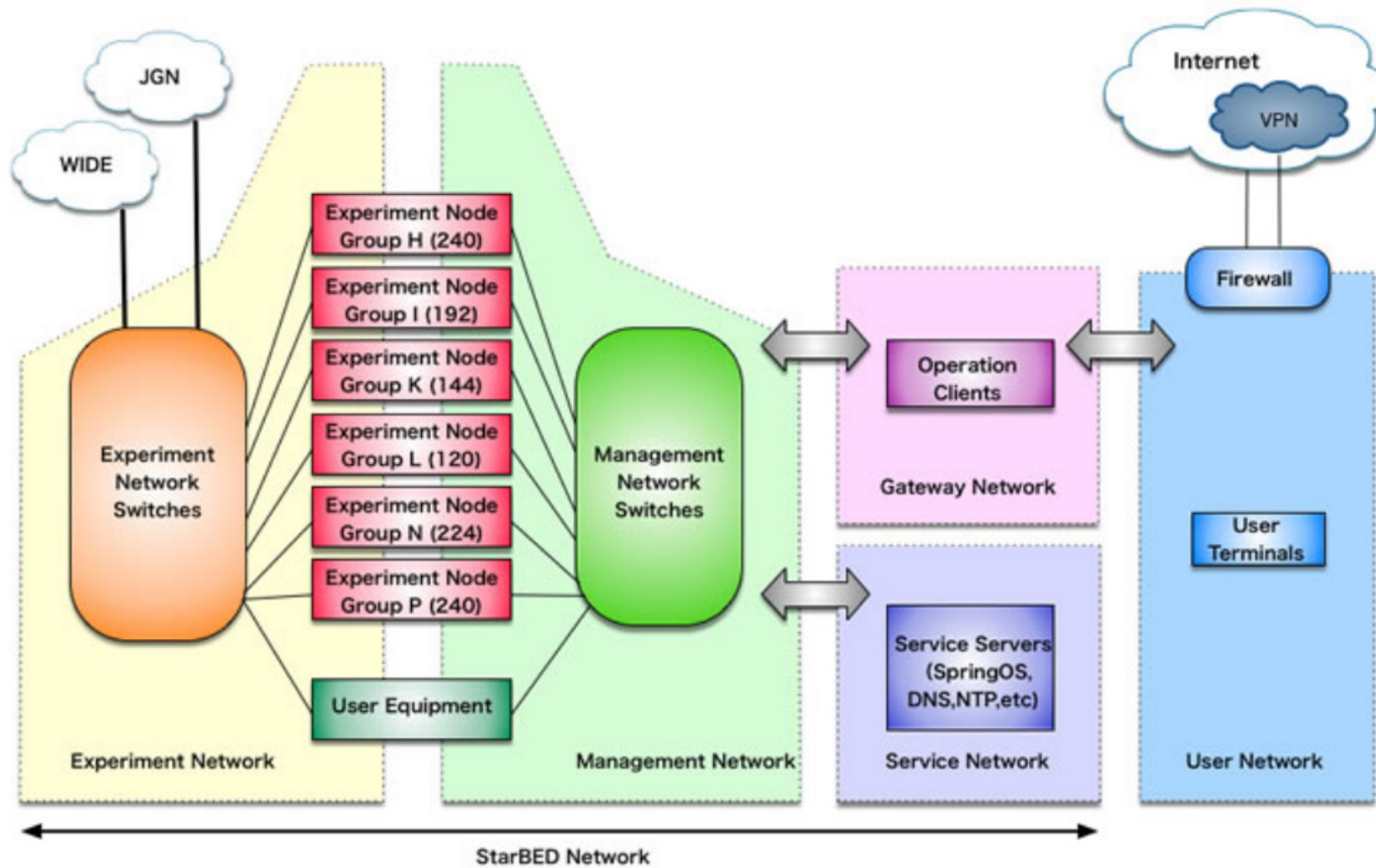
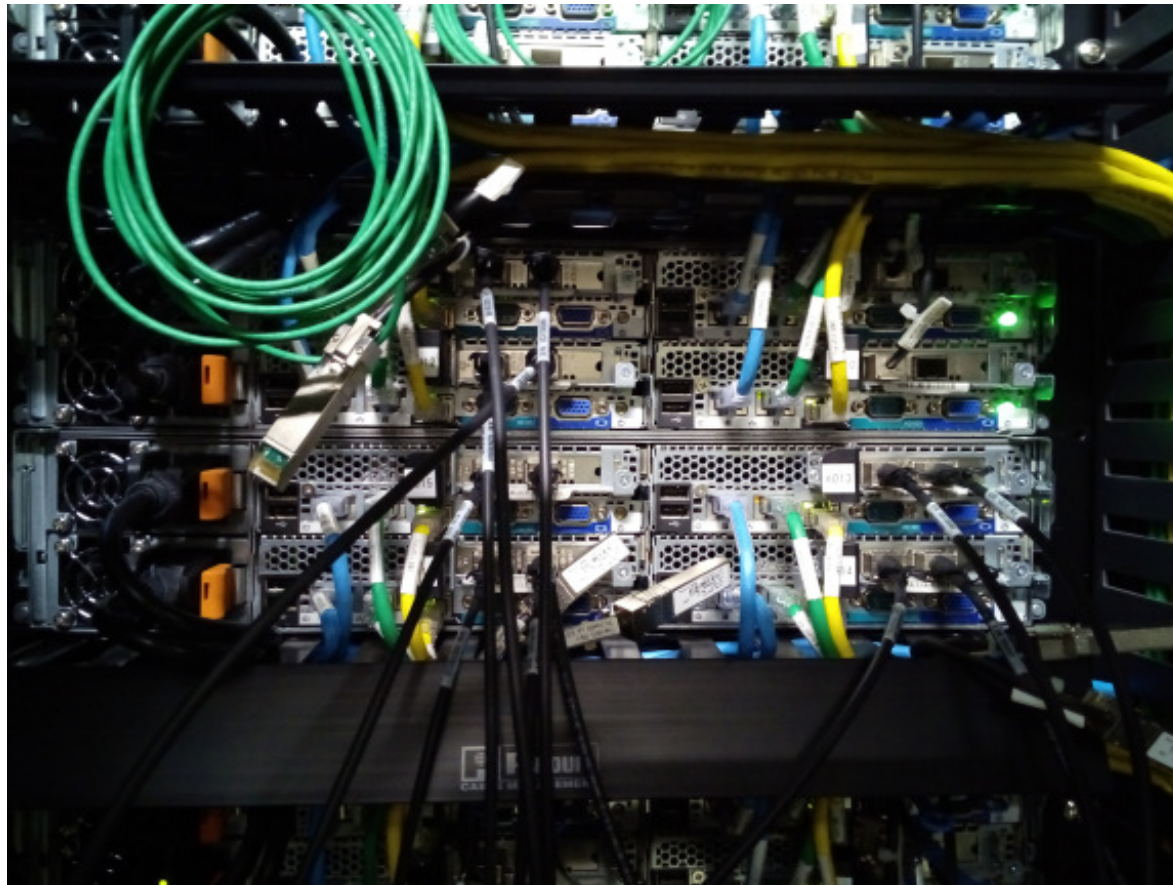
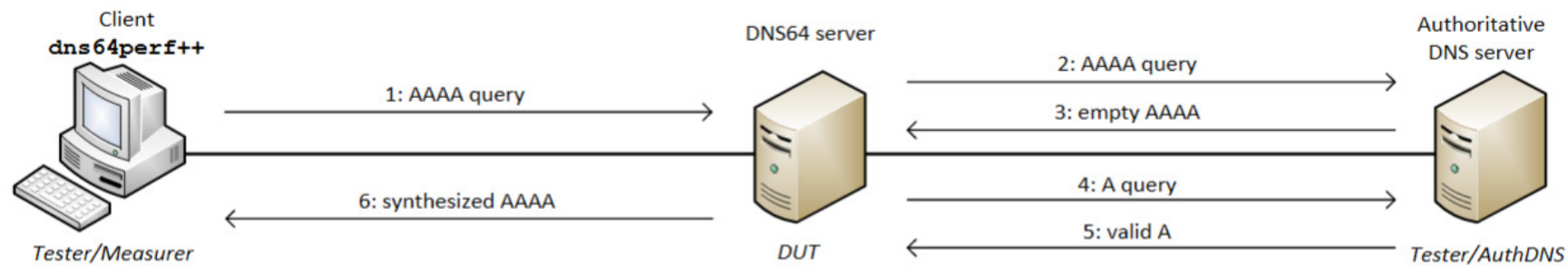


Image of StarBED Equipment

NICT StarBED

- 18 db Dell PowerEdge C6220 szervert kaptam
 - 2x Intel Xeon E5-2650 CPU (2x8 mag)
 - 16x8GB 1333MHz DDR3 RAM

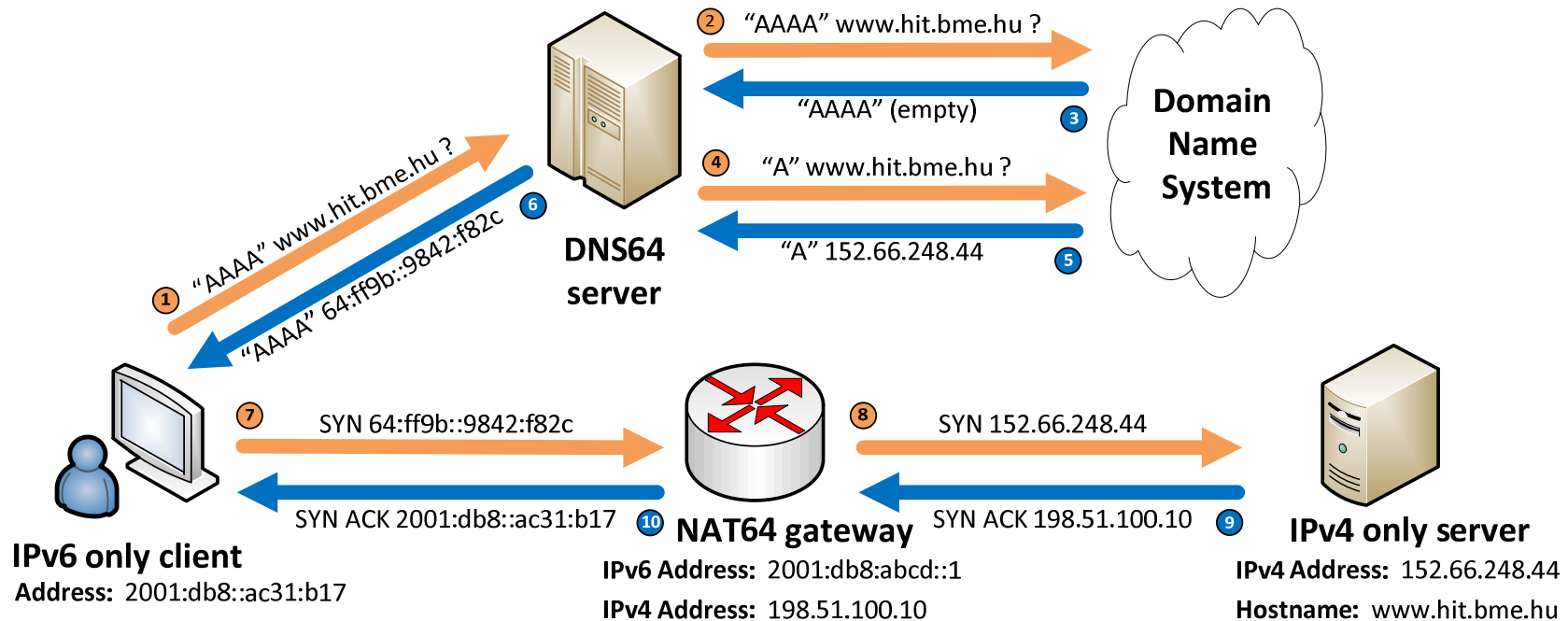




A MÉRÉSRŐL

Emlékeztető: DNS64+NAT64

- Kliensnek csak IPv6, szervernek csak IPv4 címe van
 - A publikus IPv4 címek kifogyása miatt várható tipikus megoldandó probléma
 - Az elvi megoldást az alábbi példán mutatjuk be



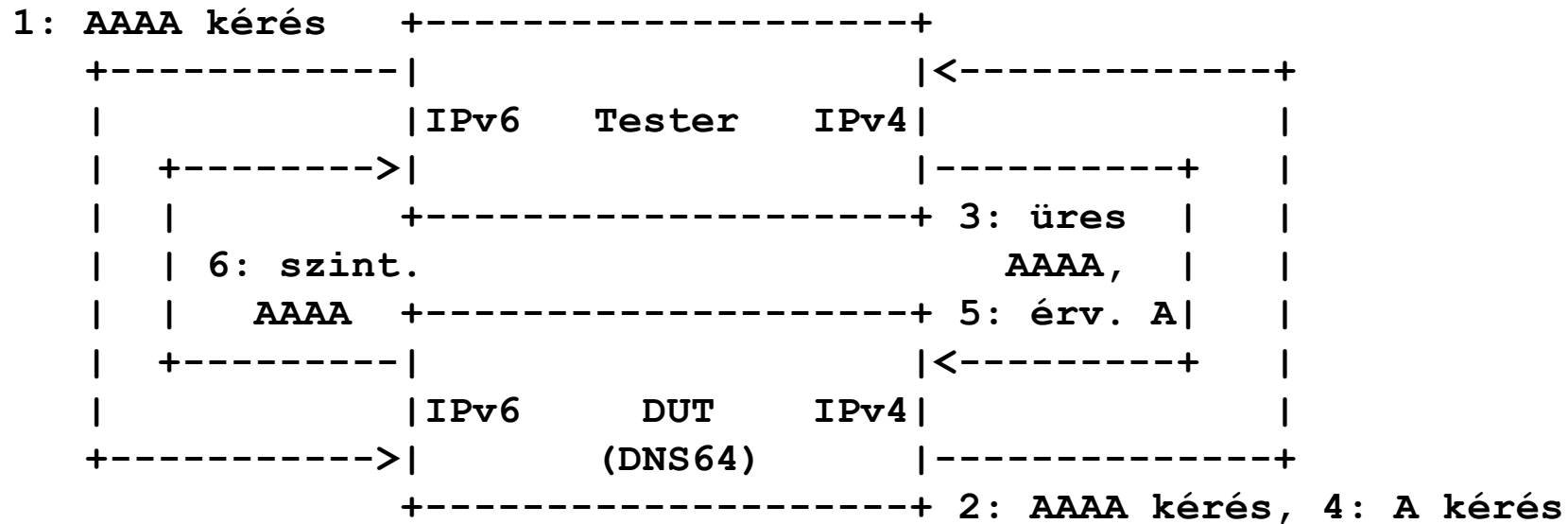
DNS64 implementációk

- BIND – „de facto” DNS szerver szabvány
- PowerDNS
- Unbound

- Nincs több, mert:
 - TOTD: nem fejlesztik, még a javított verziója is elszállt
 - Ecdysis: csak BIND/Unbound patch, illetve perl script
 - mtd64-ng: csak kísérleti fejlesztés, szolgáltatás nyújtására még nem használható
 - tesztelésre azért hasznos volt

RFC 8219: DNS64 benchmarking

- Mérési összeállítás és forgalom



- Szerepek:

- **Tester**: DNS64 kliens + Authoritative DNS szerver
 - **DUT** (Device Under Test): DNS64 szerver
- A Tester megvalósítható egy vagy két fizikai eszközzel is.

Névfeloldás teljesítményének mérése

- Cél
 - A DNS64 teljesítmény meghatározása a *sikeresen kiszolgált DNS kérések másodpercenkénti számával*.
- Alap esetben: „worst case”, azaz minden kérés különböző
- Opcionális kiegészítések:
 - A kérések 20%-a, 40%-a, 60%-a, 80%-a, illetve 100%-a olyan szimbolikus nevet tartalmaz, amelyekhez létezik AAAA rekord
 - A kérések 20%-a, 40%-a, 60%-a, 80%-a, illetve 100%-a olyan szimbolikus nevet tartalmaz, amelyek cache-elve vannak

Követelmények a Testerrel szemben

- Cél
 - A Tester elégtelen teljesítménye ne befolyásolhassa az eredményeket.
- Megoldás
 - Mérés előtt a Tester öntesztje
 - A Testert visszahurkoljuk (DUT kihagyásával)
 - Az authoritative DNS alrendszer most AAAA rekodokat ad
 - r kérés/másodperc sebesség és t timeout melletti méréshez
 - sikeres önteszt feltétele: $2 \cdot r \cdot (1 + \delta)$ sebességgel és $0,25 \cdot t$ timeouton belül kell AAAA rekordokat szolgáltatnia, ahol $\delta \geq 0,1$.
 - Magyarázat
 - DNS64 mérésnél minden AAAA kéréshez max. 2 kérés, ezek kiszolgálása a timeout időnek max. felét használja el.
 - A δ célja a teljesítménytartalék.

Tester öntesztje

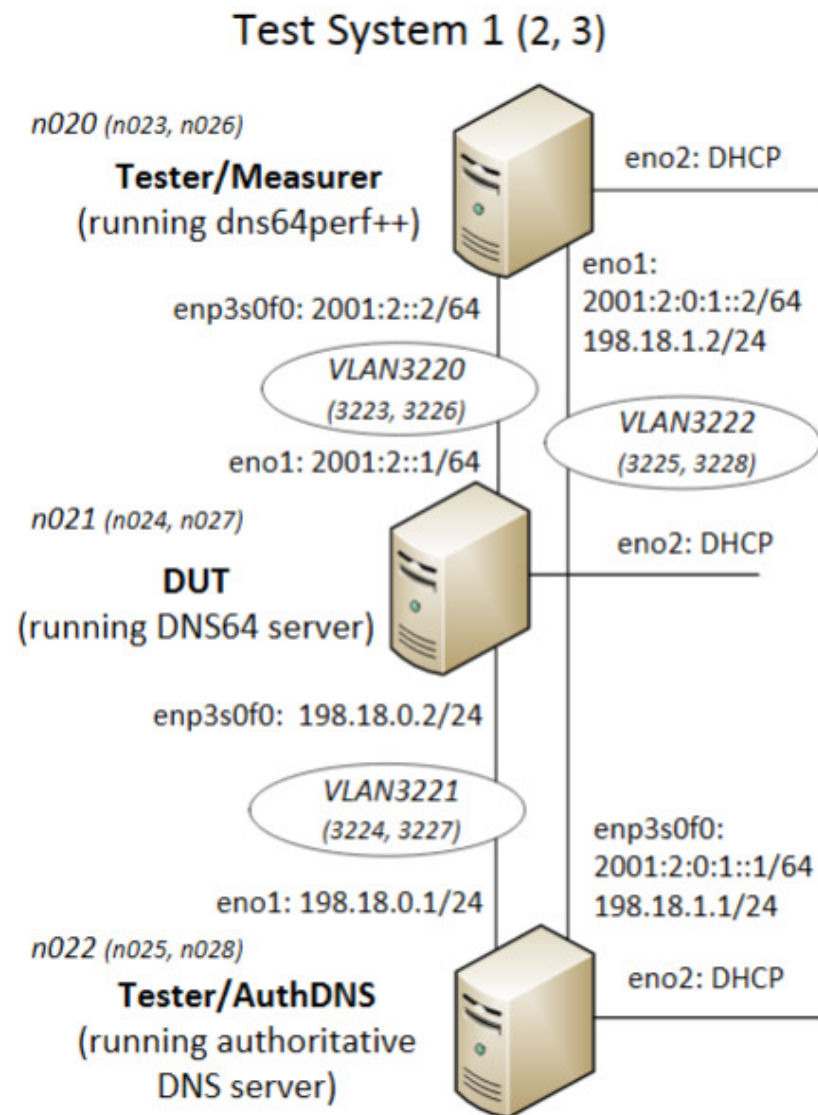
- 3 darab tesztrendszer
 - elvileg azonosak
 - de az eredmények nem teljesen
- önteszt eredmények

TABLE I
AUTHORITATIVE DNS PERFORMANCE OF YADIFA OVER IPV4

Test System	1	2	3
Median (queries per second)	180140	163641	180200
Minimum (queries per second)	163839	162303	176127
Maximum (queries per second)	182915	164641	185345

TABLE II
AUTHORITATIVE DNS PERFORMANCE OF YADIFA OVER IPV6

Test System	1	2	3
Median (queries per second)	162129	160003	159546
Minimum (queries per second)	147391	155133	157055
Maximum (queries per second)	166017	164357	163857

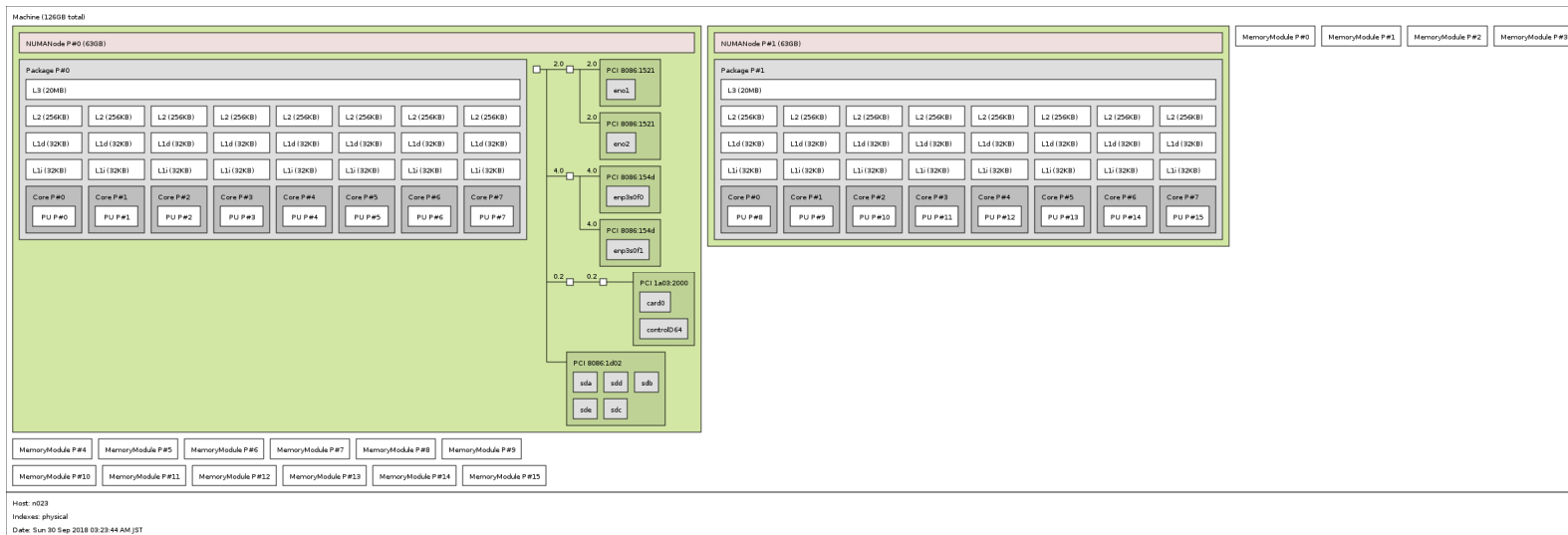


Mérőrendszerek használata

- Funkcionális felosztás
 - 1. mérőrendszer (n020, n021, n022)
 - Kötelező „worst case” mérések
 - 2. mérőrendszer (n023, n024, n025)
 - Opcionális mérés: különféle arányban van AAAA rekord
 - 3. mérőrendszer (n026, n027, n028)
 - Opcionális mérés: különféle arányban van cache találat
- Skálázhatóság vizsgálata (kötelező méréseknél)
 - 1, 2, 4, 8, 16 CPU mag használata
 - Figyelem! 2 db fizikai CPU
 - 0-7 magok, 8-15 magok
 - NUMA: Non Uniform Memory Access

NUMA helyzet a szervereken

- A használt Dell PowerEdge C6220 szervereken
 - A memóriát megfelezték
 - Az összes periféria a 0-s NUMA node-hoz csatlakozik



Nagy számú mérés

- A kötelező méréseknél
 - 3 db DNS64 szerver (PowerDNS, Unbound, BIND)
 - 5 munkapont (1, 2, 4, 8, 16 CPU mag)
 - 20-szor kell mindent elvégezni
 - A mérés is bináris keresést használ (kb. 16 iteráció)
 - Egy elemi mérés 60s, köztük 10s „pihenő idő”
- Kivitelezés scriptekkel
- CPU magok ki/be kapcsolgatása
 - i.* CPU mag **ki/be** kapcsolása:
`echo 0/1 > /sys/devices/system/cpu/cpui/online`
(Már nem így csinálnám!)

Kitérő: néhány jó tanács – 1

- Ha valaki számítógépes teljesítményméréseket végez, érdemes ezekre figyelni, mert megbízhatatlan mérést, erősen szóródó eredményeket okozhatnak
 - Automatikus CPU frekvencia állítás (pl. Intel SpeedStep)
 - Linux alatt `cpufrequtils` csomaggal fix értékre állítani
 - BIOS Setupból letiltani
 - HyperTreading (2x annyi CPU mag látszik)
 - Le szoktam tiltani
 - Asztali PC nem alkalmas, hálózati kártyáknál missed interrupt
 - Szerver gépeken interrupt moderation
 - Nem okoz minden egyes csomag megszakítást

Kitérő: néhány jó tanács – 2

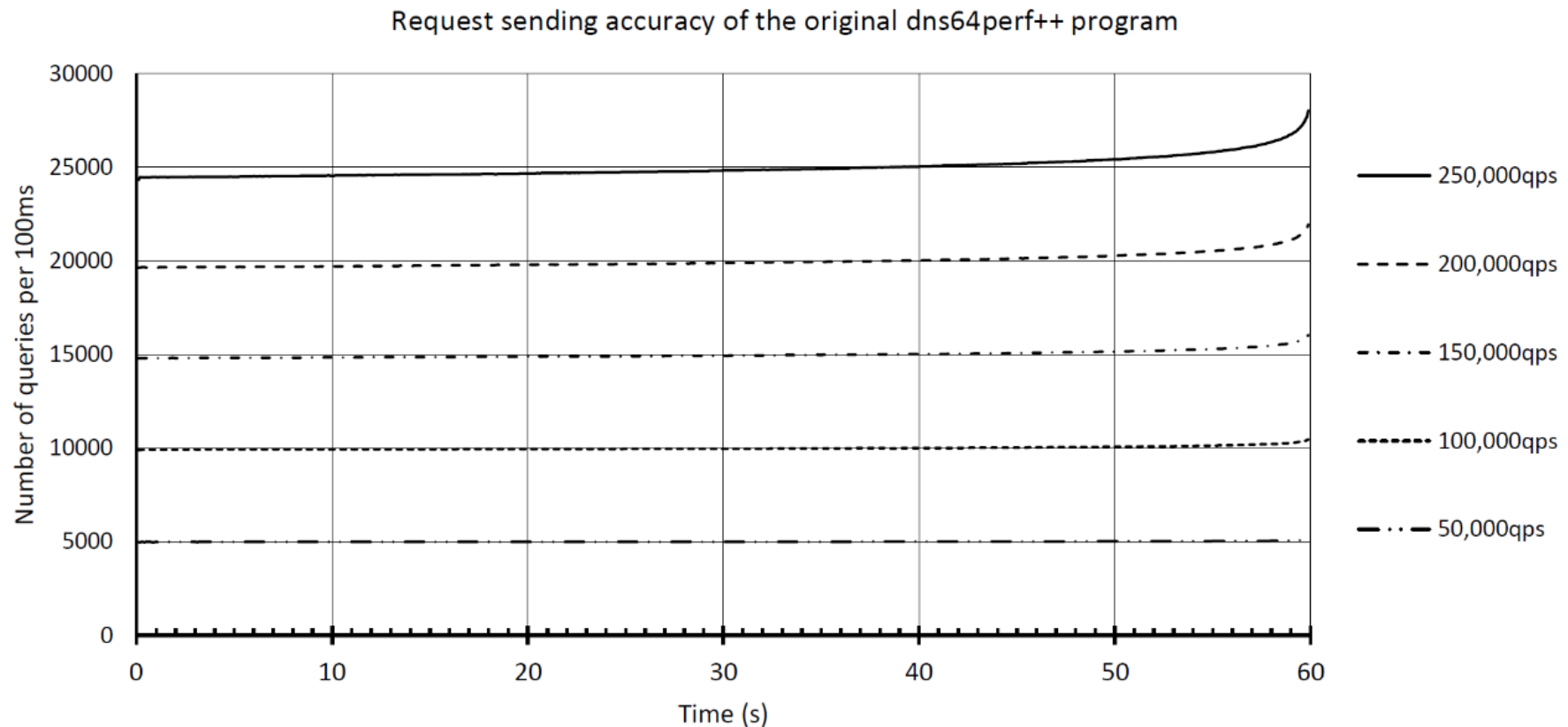
- Teljesítmény vizsgálata a CPU magok számának függvényében (CPU magok ki/be kapcsolgatása)
 - `/sys/devices/system/cpu/cpui/online` módszer
 - Naplófájlban: megszakítások CPU affinitása sérül
 - Néhányszor 10 ezer csomag/s esetén valószínűleg nem gond
 - Néhányszor 100 ezer csomag/s esetén már gond
 - Én nem ajánlom!
 - `maxcpus=n` kernel paraméter
 - Még így is feszültségmentesítés szükséges!
 - Újraindítás NEM elég
 - A Dell PowerEdge gépeken grafikus menedzsmnt felületen: „Hard Reset (Restart)”
- További tapasztalatokat én is örömmel fogadok!

Mérőrendszer pontossága?

- A `dns64perf++` programot senki sem hitelesítette
 - Nincsen mivel/mihez
 - Néhány éve itt az NMHH-ban próbálkozunk a Spirent N4U teszterrel DNS64 teljesítményt mérni.
 - A bináris keresés automatizálását nem sikerült megoldani
- Mit jelent, ha a mérési eredmények szórnak? Oka lehet
 - A `dns64perf++` program
 - A vizsgált implementáció
 - Hardver (gépek, hálózat), szoftver (operációs rendszer)
- A mérőprogrammal korábban voltak gondjaink
 - Míg néhányszor 10 kqps esetén minden rendben volt
 - 100-200 kqps esetén az eredmények durván szórtak

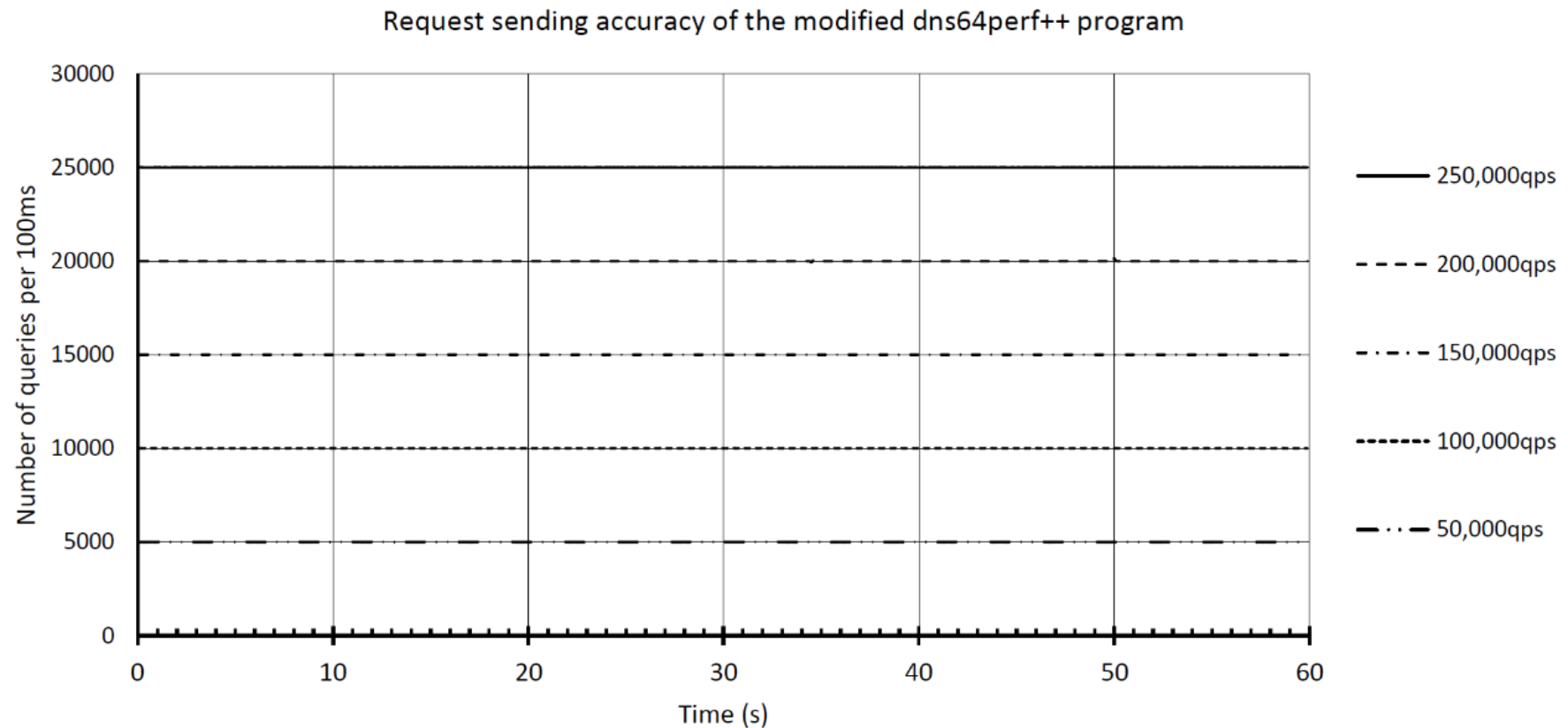
A hibás időzítés hatása

- Az időzítés
 - ns pontosságú órát használ
 - önkorrekciós képlet célja: érzékelt elcsúszás „szétterítése”



Az időzítés kijavítása

- Az időzítés
 - az érzékelt összes elcsúszás azonnali, teljes korrekciója



- A küldési időbélyegek szerint a korrekció sikeres 😊

Mérőrendszer pontossága

- Az eredmények minőségének mértékéül az eredmények diszperzióját így definiáltam

$$dispersion = \frac{\max - \min}{median} \cdot 100\%$$

- Az mtd64-ng DNS64 implementáció teljesítménye 8 magon

TABLE III

DNS64 PERFORMANCE OF MTD64-NG USING 8 CPU CORES

Test System	1	2	3
Median (queries per second)	63303	63093	63003
Minimum (queries per second)	62271	62343	61951
Maximum (queries per second)	64121	64001	64531
Dispersion (%)	2.92	2.63	4.10

Mérőrendszer pontossága

- Az NSD autoritatív DNS szerver teljesítménye
 - 1 aktív CPU mag esetén
 - IPv4 fölött elérve

TABLE IV

AUTHORITATIVE DNS PERFORMANCE OF NSD, SINGLE CPU CORE, IPV4

Test System	1	2	3
Median (queries per second)	168623	163904	167137
Minimum (queries per second)	167935	163447	165887
Maximum (queries per second)	168961	164361	167489
Dispersion (%)	0.61	0.56	0.96

➔ Van értelme mérni ezzel a rendszerrel. 😊

Eredmények

- A **PowerDNS** DNS64 teljesítménye a CPU magok számának függvényében

TABLE V

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, POWERDNS

Num. CPU cores	1	2	4	8	16
Median (qps)	3077	5498	10030	17290	26620
Minimum (qps)	3061	5439	9855	16603	24447
Maximum (qps)	3105	5633	10153	18563	27665
Dispersion (%)	1.43	3.53	2.97	11.34	12.09

- Nem lineáris, de azért elég jól skálázódik! 😊
- 8 és 16 CPU mag mellett már jelentős a diszperzió

Eredmények

- Az **Unbound** DNS64 teljesítménye a CPU magok számának függvényében

TABLE VI

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, UNBOUND (SCATTERED RESULTS)

Num. CPU cores	1	2	4	8	16
Median (qps)	8708	15623	19993	31502	17131
Minimum (qps)	8511	11121	16223	24575	11955
Maximum (qps)	8865	16897	25665	32753	22017
Dispersion (%)	4.07	36.97	47.23	25.96	58.74

- 1 CPU mag esetén kiváló teljesítmény 😊
- 2 CPU magtól kezdve elfogadhatatlanul nagy a diszperzió ☹️
- 16 CPU mag esetén kisebb a teljesítménye, mint 4 vagy 8 CPU mag esetén ☹️

Eredmények

- Az **BIND** DNS64 teljesítménye a CPU magok számának függvényében

TABLE VII

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, BIND

Num. CPU cores	1	2	4	8	16
Median (qps)	2425	4788	6659	6659	6661
Minimum (qps)	2303	4731	6659	6659	6661
Maximum (qps)	2441	4897	6659	6659	6661
Dispersion (%)	5.69	3.47	0	0	0

- A három közül a BIND-nak a leggyengébb a teljesítménye
- 4, 8 és 16 CPU mag esetén az eredmények meglepően azonosak ☹

Mi lehet a nagy diszperzió oka?

- A mérési naplófájlokban feltűnt, hogy sok esetben 99% fölötti a sikeres kérések aránya, de mégsem 100%
- Az RFC 8219 (az RFC 2544 mintájára) csak 0 csomagvesztést fogad el
- Bár az RFC 8219-nek nem felel meg, kísérletként legyen 99% az elfogadási kritérium

Eredmények

- A **PowerDNS** DNS64 teljesítménye a CPU magok számának függvényében 99% elfogadási kritérium esetén

TABLE VIII

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, POWERDNS

ACCEPTANCE CRITERION: 99%, NOT RFC 8219 COMPLIANT!

Num. CPU cores	1	2	4	8	16
Median (qps)	3092	5527	10146	19541	30954
Minimum (qps)	3071	5481	9949	17887	30393
Maximum (qps)	3121	5633	10321	20485	31793
Dispersion (%)	1.62	2.75	3.67	13.30	4.52

- 1-2-4 CPU mag esetén nincs érdemi különbség
- 8 és 16 CPU mag esetén a teljesítmény kis mértékben nőtt
- 16 CPU mag esetén a diszperzió OK, de 8-nál megmaradt

Eredmények

- A **Unbound** DNS64 teljesítménye a CPU magok számának függvényében 99% elfogadási kritérium esetén

TABLE IX

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, UNBOUND

ACCEPTANCE CRITERION: 99%, NOT RFC 8219 COMPLIANT!

Num. CPU cores	1	2	4	8	16
Median (qps)	8710	16624	33852	66552	43790
Minimum (qps)	8447	16379	32767	65407	40959
Maximum (qps)	8793	16949	34305	67585	45089
Dispersion (%)	3.97	3.43	4.54	3.27	9.43

- 4-8-16 CPU mag esetén lényegesen nőtt a teljesítmény
- A diszperzió lényegében rendbe jött
- 8 → 16 CPU mag esetén most is csökken a teljesítmény

Eredmények

- A **BIND** DNS64 teljesítménye a CPU magok számának függvényében 99% elfogadási kritérium esetén

TABLE X

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND
THREADS, BIND,

ACCEPTANCE CRITERION: 99%, NOT RFC 8219 COMPLIANT!

Num. CPU cores	1	2	4	8	16
Median (qps)	2448	4852	9467	17508	15748
Minimum (qps)	2367	4727	9403	17135	15577
Maximum (qps)	2477	4993	9601	17731	16385
Dispersion (%)	4.49	5.48	2.09	3.40	5.13

- 1-től 8 CPU magig most jól skálázódom a teljesítmény
- 8 → 16 CPU mag esetén kis mértékben csökken a teljesítmény

További észrevétel

- Az 1 CPU mag (= 1 program szál) mellett mért teljesítményt a 99%-os elfogadási kritérium **NEM** változtatta meg
 - ➔ A problémák gyökere a program szálak közötti együttműködésben keresendő
- Teljesítmény modell: a teljesítmény két összetevője:
 - 1 CPU mag melletti teljesítmény
 - ebben az Unbound kiváló
 - program szálak közötti kooperáció hatékonysága
 - ebben a PowerDNS kiváló

NUMA problémák vizsgálata

- Az egyes CPU magok nem egyenértékűek
 - 0-7: NUMA node 0
 - 8-15: NUMA node 1
- Vizsgáljuk meg, hogy van-e különbség a következő kettő egyaránt 8 CPU magos összeállítás teljesítménye között
 - 0-7 CPU magokat használjuk
 - 0-3 + 8-11 CPU magokat használjuk

PowerDNS: $\approx 3\%$ az eltérés

TABLE V

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, POWERDNS

Num. CPU cores	1	2	4	8	16
Median (qps)	3077	5498	10030	17290	26620
Minimum (qps)	3061	5439	9855	16603	24447
Maximum (qps)	3105	5633	10153	18563	27665
Dispersion (%)	1.43	3.53	2.97	11.34	12.09

TABLE XI

DNS64 PERFORMANCE USING 4 CORES FROM EACH CPU, POWERDNS

Num. CPU cores	4+4
Median (qps)	16771
Minimum (qps)	16095
Maximum (qps)	18529
Dispersion (%)	14.51

Unbound: $\approx 10\%$ az eltérés

TABLE IX

DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, UNBOUND

ACCEPTANCE CRITERION: 99%, NOT RFC 8219 COMPLIANT!

Num. CPU cores	1	2	4	8	16
Median (qps)	8710	16624	33852	66552	43790
Minimum (qps)	8447	16379	32767	65407	40959
Maximum (qps)	8793	16949	34305	67585	45089
Dispersion (%)	3.97	3.43	4.54	3.27	9.43

TABLE XII

DNS64 PERFORMANCE USING 4 CORES FROM EACH CPU, UNBOUND

ACCEPTANCE CRITERION: 99%, NOT RFC 8219 COMPLIANT!

Num. CPU cores	4+4
Median (qps)	59832
Minimum (qps)	58879
Maximum (qps)	60681
Dispersion (%)	3.01

-27%

DNS64 SZOLGÁLTATÁS SZÁMÍTÁSI HATÉKONYSÁGA

Mibe kerül nekünk...?

- Megvizsgáltuk, hogy melyik DNS64 implementációnak mi a maximális teljesítménye
- De mennyi számítási munka árán végzi, amit végez?
 - Technikai: 1 CPU mag: 100%, 16 CPU mag: 1600%
- **Számítási teljesítmény relatív DNS64 teljesítmény**
 - **Egy adott DNS64 implementáció hány kérést tud kiszolgálni másodpercenként egy CPU mag teljes számítási kapacitása árán?**
- További kérdés ennek a számítási teljesítménynek az ára
 - számítógép ára
 - elektromos áram ára (+ a hűtésé!)

Ezekkel már nem tudunk foglalkozni.

Kezdeti tranziens meddig tart?

- A tranziens hossza függhet számos tényezőtől
 - DNS64 szerver típusa
 - Magok száma
 - Terhelés mértéke
- Empirikus vizsgálat: CPU kihasználtság mérése
 - Viszonylag hosszabb ideig (180s)
 - Idő függvényében
 - 1s felbontással
- A mérőscript lényegi utasítása
top -b -n 180 -d 1 -p \$(pidof \$DNS64server)
- Néhány releváns esetet ábrázolunk grafikusán

Kezdeti tranziens meddig tart?

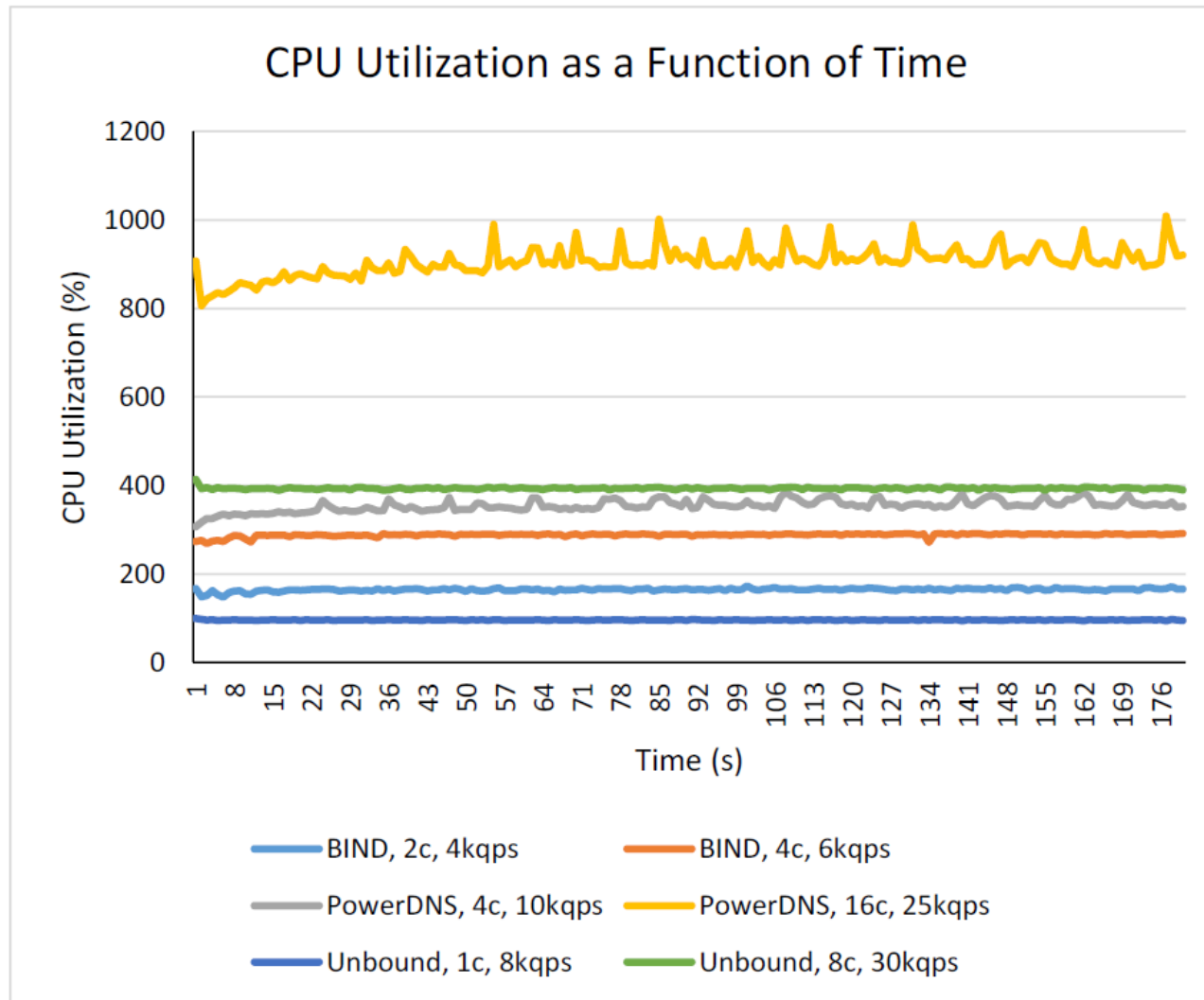


Fig. 3. Study of the initial transient of the CPU utilization.

Hogyan mérjük?

- Kezdeti tranziens hosszára egy jó felső korlát 60s
 - Utána legyen 120s időtartamú mérés
- A terhelés mértékétől függően is különbség lehet a hatékonyságban
 - Ezt is megvizsgáltam néhány esetben

A munkapont hatása

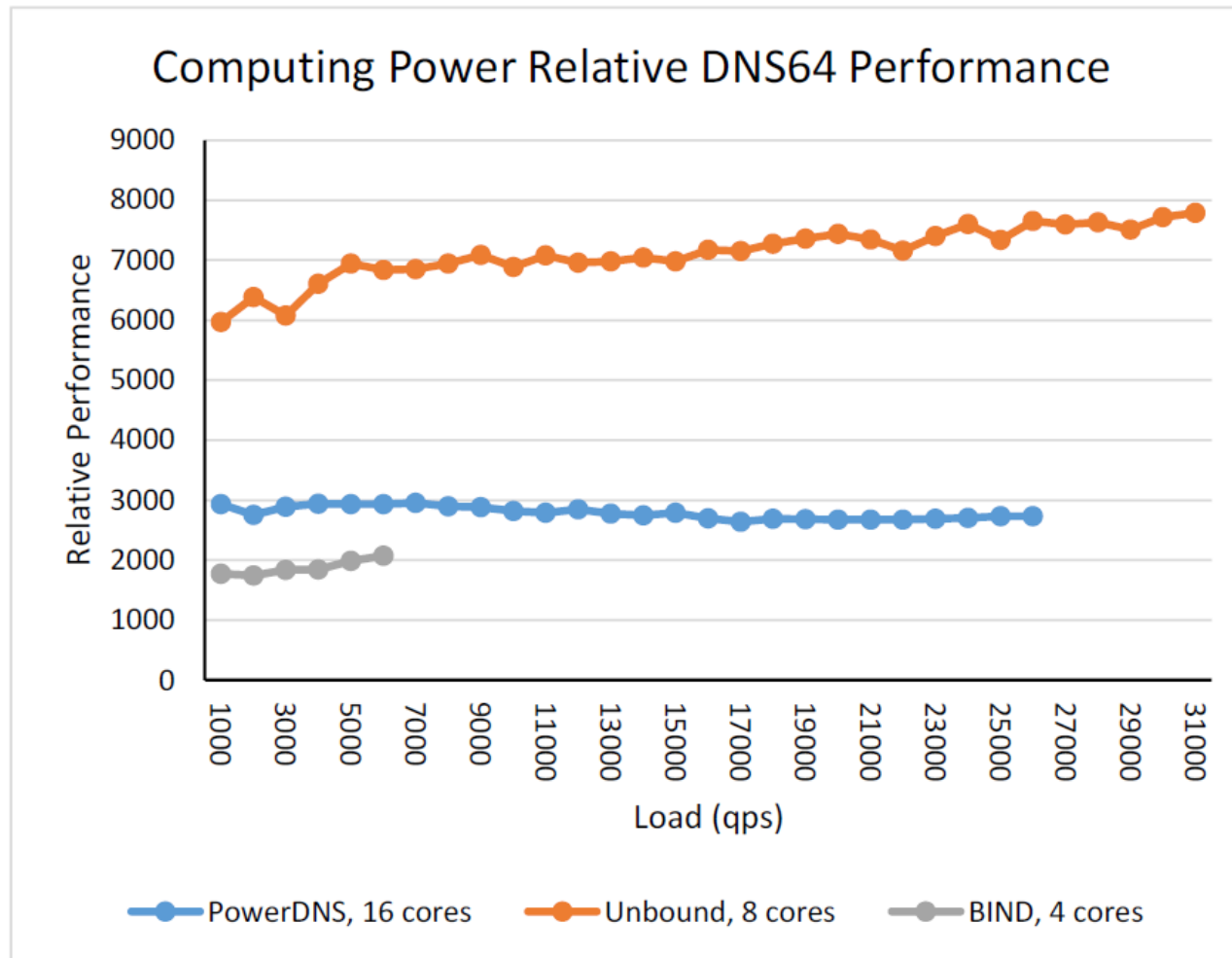


Fig. 4. Computing power relative performances of PowerDNS, Unbound and BIND under different load conditions.

Eredmények

- A **PowerDNS** számítási teljesítmény relatív DNS64 teljesítménye a CPU magok számának és a terhelésnek a függvényében

TABLE XIX
COMPUTING POWER RELATIVE DNS64 PERFORMANCE (CPRDP) UNDER
DIFFERENT LOAD CONDITIONS, POWERDNS

Num. CPU cores	1	2	4	8	16
Full rate (qps)	3077	5498	10030	17290	26620
CPRDP @25%	2960	2855	2938	2973	2932
CPRDP @50%	2947	2830	2806	2831	2808
CPRDP @75%	2932	2751	2780	2717	2655
CPU util. @100%	99.3%	196.6%	360.7%	617.1%	974.6%

- 25% terhelésnél a magok számával NEM romlik 😊
- 75 % terhelésnél és több magnál is csak alig romlik 😊

Eredmények

- Az **Unbound** számítási teljesítmény relatív DNS64 teljesítménye a CPU magok számának és a terhelésnek a függvényében

TABLE XX

COMPUTING POWER RELATIVE DNS64 PERFORMANCE (CPRDP) UNDER DIFFERENT LOAD CONDITIONS, UNBOUND

Num. CPU cores	1	2	4	8	16
Full rate (qps)	8708	15623	19993	31502	17131
CPRDP @25%	7735	7524	7598	6855	5369
CPRDP @50%	7735	7273	7793	7141	5667
CPRDP @75%	7749	7786	7805	7517	5847
CPU util. @100%	98.8%	188.2%	247.5%	407.8%	291.3%

- Kifejezetten magas, tehát gazdaságos 😊
- 16 CPU mag mellett romlik, de ennél sokkal nagyobb baj az abszolút teljesítmény romlása ☹️

Eredmények

- A **BIND** számítási teljesítmény relatív DNS64 teljesítménye a CPU magok számának és a terhelésnek a függvényében

TABLE XXI
COMPUTING POWER RELATIVE DNS64 PERFORMANCE (CPRDP) UNDER
DIFFERENT LOAD CONDITIONS, BIND

Num. CPU cores	1	2	4	8	16
Full rate (qps)	2425	4788	6659	6659	6661
CPRDP @25%	1861	1735	1785	1786	1538
CPRDP @50%	1980	1929	1826	1645	1550
CPRDP @75%	2062	2452	2005	1696	1333
CPU util. @100%	99.3%	196.6%	312.8%	387.6%	565.9%

- Legrosszabb a három közül ☹
- 16 magon ugyanannyi teljesítmény több munka, mint 4/8 magon ☹
- A legnagyobb baj, hogy az abszolút teljesítmény nem skálázódik ☹

További információk

- A **BIND** furcsa viselkedését okozó problémáról szóló hipotézisem, és további részletek megtalálhatók:

G. Lencse and Y. Kadobayashi, "Benchmarking DNS64 Implementations: Theory and Practice", *Computer Communications* (Elsevier), vol. 127, no. 1, pp. 61-74, September 1, 2018, DOI: 10.1016/j.comcom.2018.05.005

bírálati verzió elérhető:

<http://www.hit.bme.hu/~lencse/publications/ECC-2018-DNS64-BM-for-review.pdf>

További terveim

- Autoritatív DNS szerverek teljesítőképességének vizsgálata
 - BIND, YADIFA, Knot DNS, (NSD, FakeDNS)
- Több autoritatív DNS szerverből álló rendszer teljesítményének összegzése
- Többszálú dns64perf++ mérőszoftver vizsgálata
- BIND súlyos DNS64 teljesítmény korlátjának mi az oka?
- Újabb DNS64 mérés
 - StarBED P node-okon
 - PowerDNS, Unbound, (BIND, mtd64-ng)
- NAT64 mérőszoftver kifejlesztése
- NAT64 átjárók teljesítményének mérése
 - Itt az NMHH-ban is Pilisi Istvánnal 😊

Összefoglalás

- A Japánban található NICT StarBED rendszer
- A mérésről
 - vizsgált DNS64 implementációk
 - mérési módszer (RFC 8219)
 - mérési összeállítás
 - mérés végrehajtása
- A mérés eredményei és kiértékelésük
- DNS64 szolgáltatás számítási hatékonysága

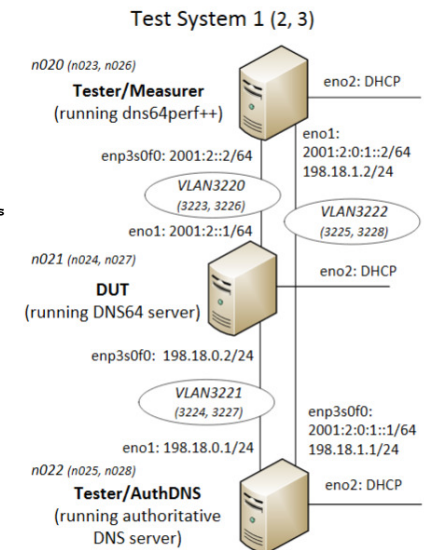
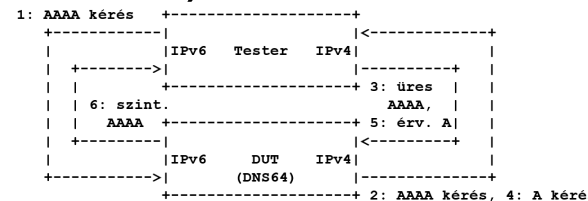


TABLE V
DNS64 PERFORMANCE AS A FUNCTION OF THE NUMBER OF CPU CORES AND THREADS, POWERDNS

Num. CPU cores	1	2	4	8	16
Median (qps)	3077	5498	10030	17290	26620
Minimum (qps)	3061	5439	9855	16603	24447
Maximum (qps)	3105	5633	10153	18563	27665
Dispersion (%)	1.43	3.53	2.97	11.34	12.09

TABLE XX
COMPUTING POWER RELATIVE DNS64 PERFORMANCE (CPRDP) UNDER DIFFERENT LOAD CONDITIONS, UNBOUND

Num. CPU cores	1	2	4	8	16
Full rate (qps)	8708	15623	19993	31502	47131
CPRDP @25%	7735	7524	7598	6855	5369
CPRDP @50%	7735	7273	7793	7141	5667
CPRDP @75%	7749	7786	7805	7517	5847
CPU util. @100%	98.8%	188.2%	247.5%	407.8%	291.3%

Köszönöm a figyelmet!

Kérdések?

Dr. Lencse Gábor
egyetemi docens
SZE Távközlési Tanszék
lencse@sze.hu

