

# 6.6 Ressursside kasutust mõjutavad näitajad ja ressursinõuded

## 1. Näitajad

- Avaliku kasutaja tõlkemälude suurus ja tõlkemälude umbkaudne hulk**
  - Avalike kasutajate kõikide tõlkemälude kogumaht on 2TB.
  - Maksimaalne lubatud üheaegne kasutajate arv avalikkuse jaoks on 27 000 (u 2% Eesti rahvaarvust).
- Ametniku tõlkemälude suurus ja umbkaudne hulk**
  - Ametnike kõikide tõlkemälude kogumaht on 2TB (PPA-s kasutaja ühe mälu maht on keskmiselt 30MB).
  - Maksimaalne lubatud üheaegne kasutajate arv avaliku sektori jaoks 13 000 (10% ametnike koguarvust).
- Terminikogude suurus ja umbkaudne hulk**
  - Ametnikel võib olla maksimaalselt üks terminikogu mahuga 100MB.
  - Ekilexi maht - PostgreSQL andmebaas võtab kettal ruumi hetkel 23GB. Pakitud *dump* ca 1.3 GB
- Tõlketööde hulk ajaühiku peale - keskmine ja tipp.**
  - Maksimaalne lubatud üheaegne toimingute arv avaliku sektori jaoks 13 000 toimingut/s
  - Maksimaalne lubatud üheaegne toimingute arv avalikkuse jaoks 27 000 toimingut/s
  - Üldlevinud päevane nominaaltõlkemaht on 5 lk/päev, kuid olenevalt teksti olemusest ja tõlkijast võib see ulatuda ka mahuni 10–12 lk /päev.

## 2. Ressursinõuded

Suur osa komponente on sisuliselt lihtsalt päringute vahendajad andmebaasi ja veebilehitseja vahel. Seetõttu on paljudel moodulitel pigem madalad ressursinõuded. Allpool on välja toodud mõttekäigud komponentide kohta, kus see ei pruugi ilmselge olla.

### 2.1. Komponentid

#### 2.1.1. Andmebaas

Andmebaasis on kõige mahukamad tõlkemälud. Täpne kasutajate arv platvormil on teadmata, ent võttes arvesse seda, et teadaolevalt on harilik tõlkemälu maht ühe ametniku kohta ca 30MB, siis on mõistlik arvestada tõlkimismälude mahuks 2TB tavakasutajate ja 2TB ametnike jaoks. Sellest peaks piisama arvestatava varuga.

Aastas juurde tekitatava mahu arvestamiseks kasutasime järgnevaid eeldusi:

|                       |      |
|-----------------------|------|
| Tähti lehel           | 1800 |
| Sõnu lehel            | 250  |
| Sõnu lauses           | 15   |
| Baite tähemärgi kohta | 2    |

Arvutuste tegemiseks kasutasime antud andmeid tõlkide arvu ja tõlgitavate lehekülgede kohta. Samuti on arvesse võetud, et salvestada tuleb nii originaaltekst kui ka tõlge.

|              | Min     | Max      |
|--------------|---------|----------|
| Tõlke kokku  | 150     | 300      |
| Lehti/päev   | 5       | 12       |
| Sõnu/päev    | 187500  | 900000   |
| Lauseid/päev | 12500   | 60000    |
| B/päev       | 5400000 | 25920000 |
| MB/päev      | 5.15    | 24.72    |
| GB/aasta     | 1.22    | 5.84     |

#### 2.1.2. Sõnumiteenus

Sõnumiteenuse puhul eeldame, et kuna valdav osa tööst nõuab kiireid vastuseid, siis sõnumijärjekorrad ei saa kuigi pikaks muutuda. Samas soovitame kasutada püsivaid sõnumijärjekordi, kus järjekorra sisu salvestatakse ka diskile, et mitte kaotada olulist osa süsteemi seisust sõnumijärjekorra teenuse häire puhul.

Sõnumiteenuse koormuse hindamisel lähtusime sellest, et lõviosa koormusest tuleb tõlketööst endast. Siin lähtusime samadest lähteandmetest tõlkimismahtude kohta, mida on kasutatud ka andmebaasi mahtude hindamiseks. Arvutatud on keskmine tõlkepäringute hulk sekundis ning, eeldades et päringud jaotuvad normaaljaotuse järgi, ka päringute arv alla mille on päringuid vastavalt 99 ja 90% juhtudest.

|                   | Min  | Max  |
|-------------------|------|------|
| Lauseid sekundis  | 0.43 | 2.08 |
| 99% alla väärtuse | 4.16 | 5.81 |
| 90% alla väärtuse | 2.48 | 4.13 |

Arvestades eeldatavat nõudlust teenuse järgi tundub igati mõistlik paigutada sõnumiteenus ja andmebaas samale masinale.

### 2.1.3. Tellija moodul

Tellijä mooduli pakutav funktsionaalsus on peamiselt andmebaasi päringud, andmete laadimine ja suhtlus sõnumiteenusega. Ehk peamine tegevus on I/O ning protsessorile erilist koormust ei lange. Et I/O koormust vähendada, on võimalik rakendusse ehitada puhverdussüsteem. See suurendab rakenduse mäluarvet, aga vähendab koormust andmebaasile, mis on mõistlik, arvestades potentsiaalseid kasutusmustrid rakenduses.

### 2.1.4. Tõlketöö moodul

Tõlketöö moodulis on suhteliselt palju funktsionaalsust, ehk ka rakendus ise on suurem. Samuti on siin palju rohkem ruumi puhverdussüsteemi kasutamiseks. Suurem osa rakenduse funktsionaalsusest teostatakse kas kasutaja sirvikus - tõlketöö, toimetamine, töövoogude haldus, jne - või teostatakse läbi andmebaasi päringute.

Tõenäoliselt suurimat koormust tekitab siin tõlketöö ise oma interaktiivse olemuse tõttu. Pärast iga lause tõlkimist või automaattõlke kinnitamist tuleb see salvestada andmebaasi.

### 2.1.5. Kvaliteedikontrolli moodul

Mooduli peamine ressursikulu on terminibaaside vahemälu hoidmine. Et kiirendada protseduuri ja vältida andmebaasi asjatut koormamist, on mõistlik viimati kasutatud terminibaase hoida puhvris.

Kuna töö ei jaotu mitte lause, vaid dokumendi peale, siis oodatav kasutuskoormus ei ole kuigi kõrge.

### 2.1.6. Teksti liigi tuvastamise moodul

Moodul töötab küll masinõppemudeliga, aga seda tööd teostav mudel on suhteliselt lihtne ning ei kasuta kuigi palju protsessorit ega vahemälu. Samuti rakendub see toiming korra dokumendi kohta, ehk instantsi korraga käigus hoida ei ole vaja kuigi palju.

### 2.1.7. Automaattõlkemoodul

Automaattõlkemoodul on kõige ressursinõudlikum komponent süsteemis kahel põhjusel:

1. Masintõlkemootor vajab palju ressursse.
2. Tõlkemälusid peab hoidma vahemälu ning kasutatavate tõlkemälude hulk tõlkes on piiramata.

Samuti arvestatakse siin seda, et tõlkemootorit kasutatakse protsessoril ja mitte graafikakaardil. Ühest küljest tähendab see, et tõlkimise kiirus on märkimisväärselt aeglasem, aga ka seda, et lausete tõlkimine paralleelselt ei anna mingit eelist. See aga tähendab, et piisava tõlkekiiruse tagamiseks on vaja hoida töös mitut instantsi, mis omavahel vahemälu ei jaga.

Arvestades ülal välja toodud oodatavat tõlkepäringute arvu, on tõenäoliselt võimalik enamikul juhtudest teenindada rakendust ühe käigusoleva instantsiga iga tõlkevaldkonna kohta. Samas on mõistlik hoida ühe instantsi tõlkemälude puhvrit mõistlikul tasemel, sest konteiner reserveerib endale kogu määratud mälu sõltumata sellest, kas seda on kohe vaja või mitte. Kui tõlkeks on tarvis tõlkemälu, mida puhvris ei ole, laetakse sealt vanim välja ja uus sisse. See tähendab, et juhul kui samas valdkonnas on tõlkepäringute arv suur ja samas tõlkemälud väga erinevad, võib tõlke kiirus kannatada, aga see on eeldatavasti üpris haruldane juhtum.

Masintõlkemootorite ressursinõudlikkus on mudeliarhitektuuride vahel üpris varieeruv ning samuti on võimalik ressursinõudlust vähendada, kui teha järeleandmisi kvaliteedis. Sellest ka suur kõikumine allpool olevas tabelis.

### 2.1.8. Kõnetuvastuse moodul

Erinevalt masintõlkest, on kõnetuvastust võimalik väga edukalt kasutada tavalisel protsessoril. Ressursinõudlikkus on aga samuti väga varieeruv ning sõltub kasutatavast mudelist ning võimalikust lisafunktsionaalsusest, nagu punktuatsiooniparandus ja täiendav keelemudel.

Allpool olevad hinnangud on antud eestikeelse kõnetuvastuse kogemusest lähtuvalt.

## 2.2. Nõuded

Nõuded komponendi kaupa on toodud allpool olevas tabelis. Tabelis esitatud hinnangud on antud ühe konteineri kohta.

| Komponent                       | Protsessor | Vahemälu (MiB) | Mälu (GiB) | GPU (tk) | Instants | Kommentaar   |
|---------------------------------|------------|----------------|------------|----------|----------|--|
| Rollihaldus                     | 0.5        | 256            | 0.5        | 0        | 1        |  |
| Tellija moodul                  | 2          | 1024           | 0.5        | 0        | 1        | Reaalne vahemälu vajadus sõltub puhverdamise lahendusest.  |
| Tõlketöö moodul                 | 2          | 2048           | 2          | 0        | 1        | Reaalne vahemälu vajadus sõltub puhverdamise lahendusest.  |
| Töövoogude teostusmoodul        | 0.5        | 128            | 0.5        | 0        | 1        |  |
| Teavitusmoodul                  | 0.2        | 128            | 0.5        | 0        | 1        |  |
| Kvaliteedikontrolli moodul      | 1          | 1024           | 1          | 0        | 1        | Peamine vahemälu kulu on terminibaasid.  |
| Teksti liigi tuvastamise moodul | 1          | 1024           | 2          | 0        | 1        |  |
| Andmebaasi teenindusmoodul      | 1          | 256            | 0.5        | 0        | 1        |  |
| Avaandmete eksportimise moodul  | 0.2        | 128            | 0.5        | 0        | 1        |  |
| Automaattõlkemoodul             | 1          | 6144-9216      | 2-5        | 0        | 5        | Mälukasutus sõltub tõlkemootorist ja tõlkemälude puhvri suurusest. Instantside arv valdkondade arvust. |
| Kõnetuvastuse moodul            | 0.5-2      | 512-5120       | 2          | 0        | 1        | Mälukasutus sõltub mudelist, mis sõltub paljudest parameetritest.                                      |
| Tõlkemälu haldusmoodul          | 1          | 1024           | 2          | 0        | 1        |  |
| Statistika kogumise moodul      | 0.2        | 128            | 0.5        | 0        | 1        |  |
| Väliste teenuste tõlkemoodul    | 0.2        | 128            | 0.5        | 0        | 1        |  |
| X-tee turvaserver               | 0.2        | 0.15           | 0.5        | 0        | 1        |  |
| Logiserver                      | 1          | 2              | 180        | 0        | 1        |  |
| Tõlkemootori treeningmoodul     | 4          | 16             | 100        | 1        | 1        |  |
| Andmebaas                       | 4          | 20480          | 5120       | 0        | 1        |  |
| Sõnumiteenus                    | 1          | 2048           | 4          | 0        | 1        |  |

### 3. Kokkuvõte, ressursinõue kogu süsteemi toimimiseks

vCPU - 26

RAM (GB) - 66 - 86

GPU - 1

Disk standard (GB) - 0

Disk ssd (GB) - 5400