Oma server või avalik pilv? Tehisaru-buum puhub tolmu vanalt dilemmalt

8 min tagasi - 24.11.2026 Autor: AM

Tehisaru on murdnud end viimase paari aastaga pea igasse kontorisse ja koos sellega on IT-juhtide lauale naasnud üks igipõline küsimus. Kui arvutusvõimsust on vaja aina rohkem, siis kas rentida seda mugavalt avalikust pilvest või laduda riiulisse päris oma raud?

Äsja teatas LHV, et on otsustanud investeerida oma tehisaruserveritesse. Kas see tähendab, et nüüd peaksid kõik ettevõtted tormama kalleid serverikappe kokku ostma? Telia IKT-teenuste juht Alar Kurvits tõmbab siinkohal veidi hoogu maha.

„Esmalt tuleb selgeks teha, mida ettevõte AI abil teha tahab, milliseid andmeid kasutab ja kui palju arvutusvõimsust tegelikult vajab," selgitab Kurvits. Tema sõnul pole küsimus selles, kas pilv või oma server on põhimõtteliselt parem, vaid selles, milline töökoormus millisesse keskkonda sobib. Sama ettevõtte jaoks võivad õiged olla lausa mõlemad.

Kontrollil on hind – ja see ei lõpe ostutšekiga

Oma tehisarutaristu suurim trump on täielik kontroll: ettevõte otsustab ise, kus andmeid töödeldakse ja kes neile ligi pääseb. Kogu selle masinavärgi südameks on võimsad graafikaprotsessorid ehk GPU-d. Kui nende koormus on suur ja püsiv, tasub oma riistvara end ära.

Kui aga kallis masin seisab suurema osa ajast jõude, meenutab see sportauto ostmist pelgalt selleks, et kord nädalas nurgapoes piima järel käia.

„Oma server ei tähenda tasuta arvutusvõimsust. Ostuhinnale lisanduvad elekter, jahutus, ühendused, turvalisus ja haldus. Võrrelda tuleb lahenduse kogukulu kogu elutsükli jooksul, mitte serveri hinda ühe kuu pilvearvega," rõhutab Kurvits.

Lisaks jääb oma keskkonnas kogu vastutus andmete kaitsmise, ligipääsude, varunduse ja igapäevase tööshoidmise eest ettevõtte enda õlule.

Miks LHV oma kliendiandmeid pilve ei luba?

LHV mõtles panga IT-juhi Kain Kalju sõnul enne Telialt tehisaruserverite soetamist iga sammu põhjalikult läbi. Finantsasutuse roll seab andmete liikumisele lihtsalt väga ranged piirid.

„Kliendiandmetega päringuid ei tohi me teha väljapoole Euroopat ning üldjuhul püüame vältida, et need üldse meie perimeetrist välja liiguksid," ütleb Kalju. Oma riistvara kasuks kallutas kaalukausi ka karm reaalsus: suure jõudlusega GPU-sid on avalikust pilvest keeruline rentida, sest neid lihtsalt napib. „Klientidele pakutavat teenust ei saa me rajada nii ebakindlale platvormile," lisab ta.

Panga jaoks tähendab oma raud kolme selget eelist: andmete asukoht on teada, võimsus on alati käepärast ja kulud on ette ennustatavad. „Boonusena on kulubaas fikseeritud, mis ei ole tänases maailmas sugugi ebaoluline," märgib Kalju.

LHV enda serverites töötavad mudelid teevad näiteks kokkuvõtteid klientide pöördumistest ja analüüsivad laenutaotluste materjale, ilma et tundlik info panga keskkonnast lahkuks. Pilvest pank siiski täielikult ei loobu – sinna jäävad uute mudelite katsetamine ja üldotstarbelised tööriistad, mis kliendiandmeid ei puuduta. Oma serverid paigutas pank aga Telia Tallinna andmekeskusesse, pääsedes nii ise serveriruumi ehitamisest.

Pilv on parim liivakast katsetamiseks

Kui Sinu ettevõtte tehisaruprojekt on alles katsetamisjärgus, ei ole Kurvitsa hinnangul suur serveriinvesteering mõistlik. Pilv toimib suurepärase liivakastina, kus saab proovida eri lahendusi ning võimsust vastavalt vajadusele juurde keerata või maha tõmmata. See sobib hästi ka hooajalise ja raskesti ennustatava koormuse puhul – eeldusel, et andmeid tohib pilves töödelda.

„Pilv annab võimaluse kasutusjuht läbi proovida ja tegelik vajadus selgeks saada. Kui lahendus jõuab igapäevasesse kasutusse, tasub kulud ja arhitektuur uuesti üle vaadata," soovitab Kurvits.

Siiski tasub meeles pidada, et pilvearve ei koosne ainult GPU tunnihinnast. Lõppsummale lisanduvad salvestusmaht, andmete liigutamine, võrguühendused, turvalisus ja haldus. Pealegi pole iga GPU-tüüp pilves alati kohe saadaval, mistõttu tuleb suurem võimsus ette broneerida. Kuldseks keskteeks osutubki sageli hübriidlahendus: tundlikud andmed ja püsiv töökoormus jooksevad oma või kohaliku teenusepakkuja juures, arendus ja ajutised sutsakad jäävad pilve.

Väikesele tiimile piisab vahel ühest võimsast tööjaamast

Kõik ettevõtted ei vaja siiski oma serveriparki. Väikese või ebaregulaarse tehisaruvajadusega tiimile on pilv sageli märksa soodsam. Üksikule arendajale või pisemale meeskonnale võib aga vabalt piisata ühest korraliku GPU-ga tööjaamast, kus kohalikke mudeleid jooksutada ja prototüüpe ehitada.

„Alustada ei tasu küsimusest, millist GPU-d osta. Alustada tuleb sellest, kes seda kasutama hakkab ja kui sageli. Mõnele ettevõttele on õige vastus oma serverid, teisele pilv ning kolmandale üks tööjaam," võtab Kurvits kokku.

6 põhiküsimust

Enne kui rahakotirauad lahti teha, võiks juhtkond ja IT-tiim ühiselt läbi arutada kuus põhiküsimust:

1. Milliseid andmeid hakkab tehisaru töötlema ja kus tohivad need paikneda?
2. Kas arvutuskoormus on pidev või tekib vaid aeg-ajalt?
3. Kui palju inimesi või äppe lahendust kasutama hakkab?
4. Kui palju andmeid peab eri keskkondade vahel liikuma?
5. Kui kiiresti ja töökindlalt peab lahendus äriprotsessi jaoks töötama?
6. Kui suur on lahenduse kogukulu kolme kuni viie aasta jooksul?

- Uudised
- Lahendused

- Serverid

Pilt

