

Bijlage 2 Reactieformulier marktconsultatievragen

Graag uw antwoorden ook voorzien van een toelichting.

De IST en de SOLL

1. Wat heeft de gemeente volgens u nodig aan nieuwe hardware, gegeven de beschreven huidige en gewenste situatie (zowel Front-end als Back-end)?

Front-End

Een VDI oplossing heeft als belangrijkste basis de gebruikerservaring en we zien dan ook de toegevoegde waarde van GPU's, ook voor de standaard werkplek. Daarnaast CPU's met hoge clock snelheden doordat veel desktop applicaties single threaded eigenschappen hebben. In de IST opzet is sprake van de volgende capaciteitsvraag verdeeld over beide datacenters:

- *3.153 vCores, 10.330GB vMEM, 1.057GB vGPU mem.*

Het IST capaciteit aanbod is (op basis van 18 actieve nodes in beide datacenters):

- *576 cores, 13.824GB Memory, 1.152GB GPU mem*
- *Minus 25% CPU en 10% mem HCI en hypervisor overhead: 432 cores, 12.442GB Memory, 1.152GB GPU mem*

Hieruit volgt dat de IST omgeving voldoende geheugen heeft met een vCPU oversubscription van 6:1.

Voor de SOLL gaan we uit van de nieuw gedefinieerde werkplek profielen en de aantallen van jaar 5, uitgaande van een standaard werkplek met 1GB aan videogeheugen (vGPU) en de overige (Adobe en CAD) elk met 2GB. Dan hebben we de volgende capaciteit vraag:

- *5.148 vCores, 21.056 vMEM, 1.345GB vGPU mem.*

Een server configuratie ziet er dan als volgt uit:

*2 x AMD EPYC 9474F, 48 cores, 3.6Ghz
16 x 128GB
2 x NVIDIA A16, 64GB*

Uitgaande van het kunnen uitvallen van één server binnen een cluster (locatie) hebben we in totaal 14 servers nodig met bovenstaande specificaties (7 per datacenter, totaal 14U per datacenter).

2. Is een vGPU benodigd op de standaard werkplek voor Teams en/of YouTube? G

GPU's gebruiken veel stroom, zijn kostbaar in aanschaf en vereisen kostbare licenties, toch denken wij dat de gemeente er goed aan doet vGPU's aan te bieden als onderdeel van de standaard werkplek. De onderbouwing hiervan is als volgt:

Gewenning gebruikers

De huidige oplossing voorziet in een vGPU voor de standaard werkplek waardoor de beeldopbouw en het bekijken van videostreams soepel verloopt. Een vGPU zorgt ervoor dat de ervaren latency consistent is en ongeveer 3x verbeterd wordt, dat er 1,5x meer beelden per seconden worden weergegeven en dat er geen concessies worden gedaan aan de beeldkwaliteit. Zonder deze vGPU zal het beeld trager opbouwen en videostreams een stuk minder vloeiend verlopen. Door gebruikers zal dit als een achteruitgang worden ervaren.

Flexibiliteit

Door de Adobe en CAD werkplekprofielen zijn er hoe dan ook servers nodig die voorzien in GPU capaciteit. Als er wordt gewerkt met verschillende serverconfiguraties, al dan niet in aparte clusters, is voor elke variant reserve capaciteit nodig om aan de high availability eisen te kunnen voldoen. Ook dit resulteert in extra servers en aparte silo's. We adviseren het gebruik van gestandaardiseerde servers met GPU capaciteit die voor alle profielen inzetbaar zijn. Dit scheelt minimaal 2 servers en heeft bovendien voordelen voor de beheersbaarheid, beschikbaarheid en flexibiliteit.

Licenties

Waar echter wel op bespaard wordt zijn de benodigde vGPU licenties (vPC) als de standaard werkplek geen gebruik maakt van GPU rekenkracht. Hoe groot dit voordeel is hangt af van het al dan niet kunnen inzetten van bestaande perpetual licenties.

Toekomst gericht

Zoals bij het volgende punt (3) beschreven zien we dat de toepassingen en de beeldkwaliteit ervan enkel toenemen bij beeldverwerking. Nu een stapje terug doen, lijkt dan onverstandig.

Samengevat zorgt het aanbieden van GPU rekenkracht voor alle profielen voor een gestandaardiseerde, efficiënte, schaalbare en flexibele oplossing. De daarbij behorende NVIDIA licentie kosten zorgen voor een goede en toekomst gerichte gebruikerservaring en voorkomt extra complexiteit in de architectuur en beheersbaarheid.

3. Moet, met de toekomst in acht genomen, ondersteuning voor 4k of Full HD voor YouTube worden vereist? En hoe gaan andere organisaties hier mee om?

Wat wij zien is dat op dit moment de meest gangbare resolutie Full HD is bij organisaties, maar er komen steeds meer devices en toepassingen met hogere resoluties. Daarnaast wordt er meer remote en thuisgewerkt, bijvoorbeeld via Teams. Grotere schermen met hogere resoluties zoals 4k dragen hier positief aan bij. Deze ontwikkeling zal zich de komende jaren voortzetten. Dit kan gevolgen hebben voor de VDI omgeving en het GPU gebruik. Voor Full HD kunnen met 1GB aan video geheugen twee schermen worden aangestuurd, terwijl voor 4k het dubbele (2GB) aan video geheugen nodig is. Ons advies is om te starten met 1GB aan video geheugen voor Full HD support en de mogelijkheid tot uitbreiden in de toekomst als eis aan de geboden oplossing te stellen. Concreet kan overwogen worden een VDI server initieel uit te rusten met de helft van de capaciteit aan video geheugen, met de optie om dit later, mocht het nodig zijn, te verdubbelen.

4. Hoe zou moeten worden omgegaan met verschillende hardware-vereisten van Software over de verschillende profielen, maar ook binnen hetzelfde profiel?

Voorbeeld ter verduidelijking: CUDA en OpenCL zijn nodig voor sommige Adobe pakketten, wat dure videokaarten vereist. Moeten dan alle servers worden uitgerust met zware videokaarten of wordt aanbevolen een cluster met gemixte hardware te maken of juist

aparte clusters?

Zoals bij punt 2 aangegeven adviseren wij gestandaardiseerde VDI servers, elk voorzien van voldoende GPU capaciteit voor zowel standaard werkplekken als ook Adobe en CAD werkplekken. Stel er zijn 10 servers beschikbaar voor werkplekken met elk 64GB aan video geheugen, dan kunnen bijvoorbeeld op elke server 50 standaard werkplekken (1GB) actief zijn samen met enkele Adobe en CAD werkplekken (2GB). Hiermee is de oplossing hardware onafhankelijk en biedt flexibiliteit, ook als profielen in de toekomst veranderen.

5. Is de beschrijving van de IST en SOLL duidelijk?

Ja, de IST en SOLL is wat ons betreft duidelijk en helder beschreven. Wel zijn er enkele kleine punten waarbij we aanvullende vragen of opmerkingen hebben:

- **100% Memory uitbreiding**
Wij begrijpen de behoefte aan flexibiliteit en de daaraan gekoppelde uitbreidingsmogelijkheden van o.a. het geheugen en opslag capaciteit..
- **CPU Clocksnelheid**
Bij de SOLL van de Front-End is het Adobe en het CAD profiel voorzien van een 4Ghz+ vCPU, terwijl bij de hardware eisen sprake is van een 3,5Ghz+ CPU. Dit lijkt strijdig te zijn.

6. Ontbreken er nog zaken?

Afhankelijk van de manier van uitvragen kan het van waarde zijn de huidige CPU en memory belasting van de systemen toe te voegen. Met deze informatie kan de aanbieder bepalen welk type CPU en hoeveel geheugen het beste past en daadwerkelijk nodig is. Deze informatie en de daaruit volgende vertaalslag is vooral zinvol als de bestaande systemen niet geheel of maar beperkt belast zijn.
Verder zien we dat er om 1GbE BASE-T wordt gevraagd voor iLO of iDRAC, wij vermoeden (en adviseren) dat de gemeente daarnaast nog 1GbE BASE-T voor gescheiden management in gebruik wil nemen zodat beheer (Nutanix, Hypervisor, Citrix, enz.) vanuit een apart security domein kan plaatsvinden.

7. Raad u aan nog andere functionele / technische eisen op te voeren? Zo ja, welke en waarom?

Ons beeld over de aankomende uitvraag is dat het voornamelijk om vervanging gaat van een bestaande architectuur, waarbij de bestaande capaciteit als basis dient voor de vervanging, rekening houdend met een bepaalde groeiverwachting de komende jaren. Vanuit dat oogpunt denken wij dat er afdoende technische en functionele eisen zijn opgevoerd.

De komende aanbesteding

De gemeente is voornemens de aanbesteding in twee onderdelen te splitsen:

Aanbesteding 1:

- Fysieke servers, inclusief support;
- GPU kaarten (t.b.v. grafische performance), inclusief licenties;
- Nutanix Hyper Converged Infrastructure (HCI) licenties;
- VMWare (deze licentie wordt waarschijnlijk vervangen door Nutanix)
- Momenteel zijn er uitsluitend HP-servers in het gebruik

*De gemeente ziet zich genooddaakt om voor de HCI oplossing weer hetzelfde merk (Nutanix) uit te vragen vanwege de technische verwevenheid met de installed base en dat een wijziging van deze hardware-software slechts uitgevoerd zou kunnen worden tegen erg hoge kosten met het risico dat de **continue beschikbaarheid** van o.a. de werkplekomgeving voor de medewerkers niet gegarandeerd kan worden. Dit laatste zou leiden tot een onacceptabel lange downtime van deze werkplek.*

8. Voor deze aanbesteding overweegt de gemeente te gunnen op basis van Kwaliteit en Prijs. Welke beoordelingsaspecten voor kwaliteit vindt u goed passen bij de opdracht en waarom?

De beoordelingsaspecten voor kwaliteit hangen m.n. af van de mate waarin, naast de levering van hardware en software, de gemeente ruimte wil geven voor creatieve oplossingen, het verwezenlijken van doelstellingen en ambities en de verwachte dienstverlening van de aanbieder. In alle gevallen zal de expertise van de aanbieder de kwaliteit van de oplossing verhogen.

Creativiteit

In plaats van exact, op line-item niveau de benodigde hardware en software uit te vragen, kan de gemeente ervoor kiezen om wel duidelijke kaders en eisen te stellen, maar enige mate van ruimte te laten voor de invulling ervan. Op die manier kunnen niet eerder aangedragen ideeën, b.v. getriggerd door de manier van uitvragen of de laatste technologische ontwikkelingen, worden meegewogen als dit een voordeel voor de gemeente oplevert.

Dienstverlening

Afhankelijk of de gemeente behoefte heeft aan additionele dienstverlening kan hiervan de kwaliteit beoordeeld worden. Denk bijvoorbeeld een aan single point of contact voor support, beheerdiensten (b.v. hardware beheer, beheer tot aan hypervisor, enz.) en implementatie en migratiediensten. De gemeente kan beschrijvingen en het plan van aanpak beoordelen.

Expertise

Een kwalitatieve aanbieder zal niet alleen een product aanbieden, maar zal ook actief meedenken en adviseren om de beste oplossing te vinden die aansluit bij de specifieke behoeften, uitdagingen en doelstellingen van de gemeente. Dit vereist een hoog niveau van expertise en betrokkenheid niet alleen tijdens het adviseren en aanbieden, maar zeker ook bij het realiseren ervan.

Samengevat zien wij kwaliteit als de meerwaarde voor de gemeente bij alle stappen in het proces, over de gehele looptijd en bij optionele ondersteunende diensten.

9. Voor de Prijs overweegt de gemeente om nettoprijzen uit te vragen voor de servers en een opslagpercentage voor de Nutanix-licentie (op een fictief bedrag dat resultante is van een Listprice en uw dealerkorting). Vindt u dit een passende methodiek, zo niet; waarom niet en wat is volgens u een betere methodiek?

Kortingen zijn afhankelijk van verschillende factoren zoals volume, dollarkoers, status en certificeringen bij de fabrikant, technische kennis in relatie tot onderhandelingsvaardigheden enz. In de RFP-fase is het moeilijk om een consistent en eerlijk opslagpercentage vast te stellen dat niet alleen initieel maar ook voor een langere periode hanteerbaar is.

Verder is het belangrijk om op te merken dat de kwaliteit van een leverancier niet eenvoudigweg kan worden uitgedrukt in de hoogte van de toegekende korting. Juist de vindingrijkheid, diepgaande kennis, de ervaring om te weten wat wel en niet werkt, zijn cruciale aspecten die bepalen hoe effectief een leverancier en daardoor het project zal zijn. Het vermogen om risico's te identificeren en te adresseren en om oplossingen te bieden, die nauw aansluiten bij de behoeften van de gemeente. Denk hierbij aan BIO, duurzaamheid, wat wil en kan de gemeente t.a.v. beheer, type organisatie, cultuur etc. Dit heeft veel meer impact en valt niet uit te drukken in het aantal punten korting.

Aanbesteding 2:

Tot de serverinfrastructuur behoren ook softwarelicenties voor de gebruikersomgeving. Deze worden apart ingekocht via een andere inkoopprocedure, omdat er voor de

gebruikersomgeving nog keuzes gemaakt moeten worden, de levertijd van bedoelde licenties veel korter is dan die van de hardware, bovendien zal het gebruik van deze licenties de komende jaren zal worden afgebouwd. Door nieuwe mogelijkheden voor de werkplekomgeving worden deze licenties overbodig gemaakt.

De volgende softwarelicenties zijn momenteel in gebruik:

- Citrix (werkomgeving medewerkers);
- Ivanti (software voor beheer applicatie en werkomgeving);
- Imprivata (authenticatie voor applicaties).

Andere merken kunnen worden aangeboden mits volledig gelijkwaardig aan genoemde merken, waarbij de gelijkwaardigheid bij Inschrijving aangetoond dient te worden én te worden gegarandeerd.

10. Voor deze aanbesteding overweegt de gemeente te gunnen op basis van prijs. Vindt u dit een passende methodiek, zo niet; waarom niet en wat is volgens u een betere methodiek?

Op het moment dat er sprake is van vooraf vastgestelde software licenties, zowel in merk, type en hoeveelheid, dan is gunnen op basis van prijs de meest van de hand liggende optie. Hier wordt echter een opening gegeven voor gelijkwaardige alternatieven. Om te borgen dat er op geen enkele manier twijfel kan bestaan over wat precies met “gelijkwaardig” wordt bedoeld, is het verstandig dit kwalitatief te toetsen. Naast functionele en kwalitatieve toetsing o.a. de vraag of een alternatief van invloed is op het onderliggende platform dat onderdeel is van Aanbesteding 1. M.a.w. is het in basis een goed idee om op prijs te gunnen, maar dan zonder de optie voor alternatieven en vooropgesteld dat de leverancier geen rol vervult bij het support, implementatie en de migratie ervan.

Algemeen

11. Heeft u nog aanvullende op- en aanmerkingen?