



Aan: Stuurgroep DBDAAP
Van: Datalab

JIVC
JIVC/AFD KIXS

Locatie
Utrecht - Kromhoutkazerne
Herculeslaan 1
Utrecht

Postadres
Postbus 90125
3509 BB UTRECHT
MPC 58A

nota

Acceptatieverslag MVP DBDAAP

Datum
20 januari 2022

Aantal bijlagen
0

Bij beantwoording, datum en onderwerp vermelden

Management samenvatting

De Stuurgroep DBDAAP heeft het Datalab bij KIXS om een acceptatietest van de minimum viable product van DBDAAP gevraagd. Gebruikmakend van een breed scala van 14 data-gedreven applicaties heeft het Datalab deze test vervolgens tussen 21 december 2021 en 18 januari 2022 uitgevoerd. Dit acceptatieverslag beschrijft deze acceptatietest en de daaruit voortkomende resultaten.

Het Datalab concludeert op basis van haar testresultaten dat de MVP DBDAAP nog niet werkbaar is en dat de migratie van geen enkele applicatie volledig is geslaagd. Dit wordt verklaard doordat fundamentele functionaliteit voor DBDAAP nog niet goed is ingeregeld en dat het daardoor thans te prematuur is om van een productiewaardig platform te spreken. Geïdentificeerde aandachtspunten die leiden tot beperkte werkbaarheid van het platform betreffen de automatisering (DNS, TLS certificaten en firewall), toegang tot MRIN vanaf DBDAAP, alsmede het onverwachte hergebruik van een bestaande Power BI Report Server van BOO-IT.

Hoewel de scope van onderhavige acceptatietest beperkt was tot een technische appreciatie van de MVP DBDAAP, volgt in de conclusie-paragraaf een korte reflectie op bevindingen van niet-technisch aard.

Aanleiding

Op 24 april 2020 wordt in een PID met de CIO Office overeengekomen dat de afdeling Generieke IT & Infrastructuur (hierna GIT&Infra) bij JIVC een MVP voor een centraal en schaalbaar advanced analytics platform voor big data ontwikkeld ([REDACTED] 2020). Dit platform krijgt de naam Defensie Big Data Advanced Analytics Platform, afgekort tot DBDAAP. Daarnaast wordt een beheer- en serviceteam ingericht.

Met de memo van 13 november 2020 wordt bevestigd dat de officiële oplevering van het MVP DBDAAP, namelijk versie 1.0, door JIVC op 18 januari 2022 plaatsvindt ([REDACTED] , 2021). Reeds een maand voorafgaand aan deze oplevering zijn op 21 december 2021 de producten en diensten die op een bijgevoegde lijst staan technisch gereed. Vervolgens bevestigt deze memo dat DBDAAP op 21 december 2021 voor het Datalab van KIXS beschikbaar komt ten einde tot 18 januari 2022 kennis en ervaring met het platform te verwerven.

Vervolgens geeft de Stuurgroep DBDAAP het Datalab opdracht tot een acceptatietest voor DBDAAP en haar bevindingen in een acceptatieverslag neer te leggen. Daarbij benadrukt de CIO Office dat het Datalab dient te accepteren dat het platform als MVP nog volop in ontwikkeling is. Derhalve dient haar focus te liggen op een beoordeling van de meest minimale bruikbaarheid van DBDAAP.

Scope van acceptatietest

De scope van voorgenoemde acceptatietest beperkt zich tot de technische beoordeling van het platform aan de hand van het migreren van bestaande data-gedreven applicaties, oftewel de applicaties, naar DBDAAP. De scope is daarmee het bepalen van de werkbaarheid van dit platform, wetende dat het een minimal viable product betreft. De inzichten die deze migraties door eindgebruikers opleveren zijn belangrijke input voor de CIO Office.

De verslaglegging in dit acceptatieverslag is gebaseerd op testresultaten, kennis en inzichten die tot en met 18 januari 2022 zijn opgedaan.

Niet tot de scope van deze acceptatietest behoort een appreciatie van niet-technische zaken omtrent DBDAAP, waaronder de inrichting van het DBDAAP ontwikkelteam, het ontwikkelproces door GIT&Infra alsmede alle zaken van organisatorisch aard. Ook buiten de scope vallen alle overige zaken die in de memo van 13 november 2021 werden toegezegd, zoals de data(science) ontwikkelwerkplekken, het zelfserviceportaal, het vaststellen van de delta tussen de huidige functionaliteit van DBDAAP die in het voorgenoemde PID en memo worden opgesomd, alsook het beheer- en serviceteam voor DBDAAP. Het Datalab blijft in onderhavig acceptatieverslag, mede vanwege onvoldoende inzicht in het BCP platform, zeer terughoudend met het doen van adviezen of het bieden van eventuele oplossingen.

Procedure voorafgaand aan acceptatietest

Voorafgaand aan de acceptatietest wordt de volgende procedure doorlopen.

Het Datalab maakt een eerste selectie van 15 data-gedreven applicaties die intern door KIXS zijn ontwikkeld, waarmee op een pragmatische en hands-on wijze de acceptatietest wordt uitgevoerd, zonder daarbij te worden gehinderd door zaken omtrent de inrichting en functioneren van deze applicaties. Dit bevordert dat migraties zich primair op het platform richten. Deze selectie bevat een ruim scala

aan applicaties die diverse functionaliteit vereist, waaronder een eenvoudig Power BI dashboard met een Python script, tot aan applicaties die uit een groter aantal componenten en functionaliteit bestaan, zoals een front-end applicatie, back-end api-service, een database en S3 object storage.

De voorgenoemde eerste selectie van applicaties vereisen gezamenlijk de volgende functionaliteit welke op basis van het PID en de mutaties op het PID zijn afgesproken ([REDACTED], 2020, pp. 60-62):

- databases: MySQL, PostgreSQL, Elasticsearch en MongoDB;
- storage: S3 object storage, File storage en Hadoop;
- workloads: Kubernetes, DNS en TLS certificaten;
- shared services: Power BI en Gurobi; en
- processing: Apache NiFi, Apache Zeppelin en Apache Spark.

Na inhoudelijke screening en bespreking van deze 15 applicaties met de lead architect van DBDAAP, blijkt dat de volgende big data functionaliteit thans ontbreekt (Kramer, 2021):

- Hadoop;
- Apache Zeppelin; en
- Apache Spark.

Met deze verworven inzichten selecteert het Datalab vervolgens 14 bestaande data-gedreven applicaties (zie onderstaand) die wat betreft de technische requirements voldoen aan de nota d.d. 13 november 2021 en de meegestuurde lijst met technische producten en diensten (cf. [REDACTED] 2021). Hiermee wordt ook vermeden dat deze applicaties die naar DBDAAP worden gemigreerd (hierna: applicaties) andere functionaliteiten vereisen dan DBDAAP op dat moment levert. Deze applicaties staan centraal in de migratie naar DBDAAP; met andere woorden wordt de acceptatietest door de migratie van deze applicaties

vormgegeven.

Alle applicaties werden eerder door KIXS / Datalab ontwikkeld, waarmee wordt bevorderd dat kennis over de inrichting en functioneren van deze applicaties in het migratieteam aanwezig is en dit team zich primair op het testen van het platform kan richten.

Ter verduidelijking, ten einde DBDAAP uitgebreid te ervaren en testen wordt met een breed scala aan 14 applicaties geselecteerd die (a) alleen door KIXS zijn ontwikkeld zodat wordt gegarandeerd dat het Datalab voldoende kennis en ervaring met onderliggende data-gedreven applicaties heeft, (b) worden redelijk eenvoudige applicaties geselecteerd die zonder veel complexe configuraties of aanpassingen worden gemigreerd en (c) applicaties geselecteerd die in ruime mate de functionaliteit die op 21 december 2021 op DBDAAP beschikbaar is benutten.

Vervolgens ontvangt de lead architect van DBDAAP op 14 december 2021 een overzicht van deze 14 applicaties met uitgebreide opdrachtdefinities. Op 17 december 2021 worden deze applicaties inhoudelijk met de lead architect besproken en als geschikt bevonden voor de acceptatietest door middel van migraties. Dit betekent dat de applicaties de bouwstenen MySQL, PostgreSQL, Elasticsearch, MongoDB, S3 object storage, file storage, Kubernetes, DNS, TLS certificaten, Power BI, Gurobi en Apache NiFi vereisen, welke op 21 december 2021 inderdaad door DBDAAP worden geleverd.

In verband met teamverplichtingen bij GIT&Infra spreekt zij met het Datalab af dat de kick-off voor de migraties niet op 21 december maar op 22 december 2021 zal plaatsvinden. Deze virtuele kick-off vindt vervolgens met een grote groep van ongeveer 13 personen plaats.


Geselecteerde applicaties

Onderhavige acceptatietest is gebaseerd op 14 applicaties. In Tabel 1 volgt een overzicht van deze applicaties alsmede een korte beschrijving.

Tabel 1. Overzicht van 14 applicaties inclusief beschrijving

Drones en Veiligheid Vliegverkeer







Mobiliteit Defensie – Analyse van reisbewegingen civiele dienstauto's

Het doel van dit project is om inzicht te krijgen in de mate van duurzaam gebruik van de dienstauto's. Hoe vaak worden de dienstauto's bijvoorbeeld gebruikt voor korte afstanden? Of voor trajecten waar ook een goede OV-verbinding is? En komt het vaak voor dat dienstauto's langdurig stilstaan terwijl er op andere locaties een tekort aan auto's is om aan de gebruikersvraag te kunnen voldoen?

Vastgoed analyse



Inzicht in de gevaren van Deep Fakes in het counter-IED domein



Kwaliteit in Beeld

Hoe kun je van 643 verschillende functies zelf competenties onderscheiden zonder menselijke fouten of veel personeel? Het antwoord op die vraag is een kortdurend project dat een Defensie-interne applicatie aanbiedt om twee lijsten van tekst met elkaar te correleren - één is een grote Excel-file met functienamen, -codes en -omschrijvingen; de ander een lijst met competenties. De applicatie zoekt dan overeenkomstige woorden en geeft die woorden die overeenkomen terug met een matchingspercentage.

GeoStack

Het GeoStack project is een zelfstudie cursus die bestaat uit een 2-daagse GeoStack Workshop en een 10-daagse GeoStack Cursus. Het doel is om alle softwarecomponenten en datasets in één logische workflow te leren downloaden, installeren, programmeren en gebruiken om aan het einde van de dag werkende software te krijgen.

IPMS Verkenner

Het doel van dit project is het inzetten van Integrated Platform Management System (IPMS) data voor predictive maintenance van marine schepen.

JIVC Dashboard Kritieke Dienstverlening

Het Datalab heeft het mogelijk gemaakt om de laatste personeels-status van systemen online door te geven. Via een set aan vooraf gedefinieerde regels worden deze meldingen real-time verwerkt tot een totaaloverzicht. Deze data worden gevisualiseerd in een interactief dashboard. Zo wordt het voor leidinggevendenden snel inzichtelijk gemaakt welke kritieke systemen onder druk staan en waar geacteerd moet worden om te zorgen dat deze blijven draaien. Power BI wordt gebruikt voor de visualisatie van deze data.

COVID-19 inzichten Dashboard CLSK

Een Power BI-dashboard met actuele veld en squadron status in relatie tot afwezigheid door Corona/COVID-19. Commandanten kunnen iedere dag hun status in een webformulier doorgeven en dit wordt vol-automatisch verwerkt in het dashboard op de Power BI Report Server van JIVC.

Datalab Dashboard – Datalab platform in getallen

Dit project bestaat uit een Power BI dashboard waarin de statistieken over alle projecten van het Datalab worden gepresenteerd.

Gitlab dashboard

Het Gitlab dashboard geeft een overzicht van welke gebruikers op Gitlab actief zijn en wanneer deze gebruikers voor het laatst hebben ingelogd. Power BI wordt gebruikt voor de visualisatie van deze data.

ATFLOW HR IDU-voorspelling CLSK

ATFLOW is een voorspellend model voor in- en uitstroom van personeel en is ontwikkeld door de DSC bij CLSK. Power BI wordt gebruikt voor de visualisatie van deze data.

Datalab rapportages

De Datalab rapportages bestaan uit verschillende Power BI dashboards die informatie over het Datalab samen brengen in één omgeving. Het vormt daarmee een managementtool.

Matlogco Voertuigonderhoud

Dit project bestaat uit verschillende Power BI dashboards en dataverwerking ten

behoefte van voertuigonderhoud door Matlogco bij CLAS.

Doel van de migraties is om deze applicaties op de OT (ontwikkel en test)-omgeving op DBDAAP te laten laten. Daarmee krijgen deze applicaties, in afwachting op een eventuele DPIA en accreditatie, voorlopig nog geen productie-status.

Voorgenoemde applicaties variëren in status, waarbij deze bestaande data-gedreven applicaties reeds zijn gecontaineriseerd voor gebruik op het Kubernetes-platform waarop DBDAAP draait, nog niet zijn gecontaineriseerd, of dashboards betreffen die gebruik maken van Microsoft Power BI.

Tabel 2 presenteert deze applicaties met de status van iedere data-gedreven applicatie. Hieruit blijkt wederom de diversiteit aan geselecteerde applicaties waarmee het werkelijke gebruik van DBDAAP wordt weerspiegeld.

Tabel 2. Classificatie van applicaties naar gecontaineriseerd, niet gecontaineriseerd of Power BI

	Gecontaineriseerd	Niet Gecontaineriseerd	Power BI
Drones en Veiligheid Vliegverkeer	X		
Mobiliteit Defensie – Analyse van reisbewegingen civiele dienstauto's	X		
Vastgoed analyse	X		
Inzicht in de gevaren van Deep Fakes in het counter-IED domein	X		
Kwaliteit in Beeld	X		
GeoStack		X	
IPMS Verkenner		X	
JIVC Dashboard Kritieke Dienstverlening			X

COVID-19 inzichten Dashboard CLSK	X
Datalab Dashboard – Datalab platform in getallen	X
Gitlab dashboard	X
ATFLOW HR IDU-voorspelling CLSK	X
Datalab rapportages	X
Matlogco Voertuigonderhoud	X

De vijf applicaties die reeds zijn gecontaineriseerd vergen naar verwachting de minste inspanning om naar DBDAAP te migreren. Dit is in tegenstelling tot twee applicaties die nog niet zijn gecontaineriseerd; deze vergen namelijk nadere configuratie om haar componenten in separate containers te plaatsen. De zeven applicaties die gebruikmaken van Power BI zijn door het gebruiksgemak van de Power BI Report Server relatief eenvoudig te migreren, maar vereisen dat een Python script dat periodiek draait in een container wordt geplaatst en dat de gegenereerde databestanden vervolgens in een S3-bucket worden opgeslagen. De CronJob functionaliteit van Kubernetes kan daartoe worden gebruikt.

Daarnaast vereisen de applicaties een variëteit aan functionaliteit die door DBDAAP wordt geleverd. Deze functionaliteit wordt gegroepeerd als databases, storage, workloads, shared services en processing. Tabel 3 presenteert de functionaliteit die door iedere applicatie wordt gebruikt.

Tabel 3. Overzicht van 14 applicaties met bijbehorende functionaliteit

	Databases	Storage	Workloads	Shared Services	Processing							
	MySQL	PostgreSQL	Elasticsearch	MongoDB	S3 Object Storage	File Storage	Kubernetes	DNS	TLS certificaten	Power BI	Gitrobi	Apache NiFi
Drones en Veiligheid Vliegverkeer	X		X	X	X	X	X	X	X	X		
Mobiliteit Defensie – Analyse van reisbewegingen civiele dienstauto's	X		X	X	X	X	X	X				
Vastgoed analyse	X		X	X	X	X	X	X				
Inzicht in de gevaren van Deep Fakes in het counter-IED domein			X	X	X	X	X	X				

Kwaliteit in Beeld					X	X	X
GeoStack	X		X	X	X	X	X
IPMS Verkenner		X			X	X	X
JIVC Dashboard Kritieke Dienstverlening				X	X		X
COVID-19 inzichten Dashboard CLSK				X	X		X
Datalab Dashboard – Datalab platform in getallen				X	X		X
Gitlab dashboard				X	X		X
ATFLOW HR IDU-voorspelling CLSK				X	X		X
Datalab rapportages				X	X		X
Matlogco Voertuigonderhoud				X	X		X

Allereerst volgt een korte beschrijving van iedere functionaliteit. MySQL en PostgreSQL zijn beide populaire en ANSI-compatible relationele databases, waarbij PostgreSQL meer geavanceerde functionaliteit biedt dan MySQL. In veel gevallen is het relatief eenvoudig om MySQL uit te wisselen voor PostgreSQL. Elasticsearch en MongoDB zijn NoSQL databases, waarbij Elasticsearch tevens geavanceerde zoekfunctionaliteit levert. S3 Object Storage en File Storage maken de opslag en bevraging van databestanden mogelijk. Kubernetes is een open-source platform waarmee de deployment, schaalbaarheid en management van gecontaineriseerde applicaties wordt geautomatiseerd. DBDAAP draait op Kubernetes; meer precies is DBDAAP een namespace op Kubernetes. DNS maakt het mogelijk om numerieke IP adressen die aan applicaties worden toegekend via een mensvriendelijk webadres te benaderen. TLS certificaten zijn noodzakelijk om communicatie tussen containers en met gebruikers versleuteld te laten verlopen. Power BI is een dashboard-platform ontwikkeld door Microsoft; deze dashboards draaien op Power BI Report Server. Gurobi levert predictive analytics waarmee wiskundige modellen voor beslissingen worden geoptimaliseerd. Apache NiFi maakt het mogelijk om ETL-pijplijnen¹ te ontwikkelen en op een cluster uit te voeren, monitoren en beheren.

Uit deze analyse in Tabel 2 blijkt dat de functionaliteit S3 Object Storage, Kubernetes, DNS en TLS certificaten voor iedere applicatie benodigd is. Hoewel geen applicaties expliciet PostgreSQL gebruiken, kunnen de bestaande applicaties

1 ETL staat voor Extract, Transform en Load. Namelijk het ontsluiten van data, het transformeren van data en vervolgens het wegschrijven van deze data.

die MySQL gebruiken relatief eenvoudig naar PostgreSQL worden gewijzigd. Geen applicaties maken gebruik van Gurobi of Apache NiFi.

Methodiek voor acceptatietest

Hierna wordt de methodiek voor onderhavige acceptatietest beschreven.

Reeds gecontaineriseerde applicaties

De applicaties die zijn onderverdeeld volgens de in Tabel 2 aangegeven classificatie, vormen de basis voor de uitgevoerde acceptatietest. Er wordt ervan uitgegaan dat applicaties die reeds zijn gecontaineriseerd eenvoudiger zijn te migreren dan applicaties die niet reeds zijn gecontaineriseerd. "Verwerkt zijn in containers" wordt in deze context geïnterpreteerd als een softwareproject waarbij het eindproduct een container-image betreft waarin een aantal randvoorwaardelijke configuratie en software is inbegrepen om het softwareprogramma van het project succesvol uit te kunnen voeren. Dit betekent dat reeds gecontaineriseerde applicaties nauwer bij het container-platform DBDAAP aansluiten.

Nog niet gecontaineriseerde applicaties

Niet gecontaineriseerde applicaties in Tabel 2 betreft applicaties die, als uitkomst van het bijhorende bouwproces, géén container-image opleveren. Deze applicaties – en/of hun bouwstraat – moeten derhalve worden aangepast om tot een container-image te komen, alvorens dergelijke applicaties op de container-managementomgeving zoals Kubernetes kunnen landen. Kubernetes wordt namelijk gebruikt voor het Basic Cloud Platform (hierna BCP). Het ontwikkelen van dergelijke container-images is een tijdrovend proces. Dit "vercontaineriseren" van de applicaties zal een nog onbekende surplus aan tijd, inspanning en haalbaarheid met zich meebrengen en wordt derhalve gezien als een minder te makkelijk categorie voor het migreren van applicaties. Mogelijk is een containerplatform überhaupt niet een oplossing voor een dergelijke applicatie, afhankelijk van de technische aard van de applicatie in kwestie.

Power BI

De categorie Power BI in Tabel 2 bevat applicaties die niet gecontaineriseerd zijn en tevens bijzonder afhankelijk zijn van Microsoft Power BI. Derhalve wordt deze groep als gelijke of als lastigere categorie beschouwd wanneer het aankomt op de migratie naar DBDAAP 1.0. Dit heeft er mee te maken dat deze Power BI-dashboards een Python-script vereisen die periodiek als CronJob op Kubernetes wordt gedraaid, ten einde via ETL-processen data op te halen, te transformeren en naar een S3-bucket weg te schrijven. Deze functionaliteit dient daarom in Kubernetes te worden geconfigureerd.

Cloud native

Hoewel we spreken van gecontaineriseerde applicaties, onderstrepen we dat gecontaineriseerde applicatie niet per definitie tot de 'cloud native' applicaties worden gerekend. Dat wil zeggen dat de gecontaineriseerde applicaties uit onderhavige migratietest geen microservices architectuur impliceren. Dit neemt niet weg dat bij sommige applicaties een onderscheid kan worden gemaakt tussen back-end en front-end, dit als aparte container-image wordt opgeleverd en dat dit als aparte container door het platform kan worden beheerd, evenals het stateless karakter van een container en het berusten op de capaciteiten van het platform ten behoeve van persistentie. Zij het in de zin van databases, bestandsopslag of anders.

Applicaties in een spagaatsituatie

Spagaatapplicatie is een term die is ontstaan tijdens de migratiewerkzaamheden voor een applicatie die gedeeltelijk succesvol draait op de DBDAAP tenant van het Basis Cloud Platform. Dat wil zeggen dat van de onderdelen A, B en C, draait onderdeel B (succesvol) op het BCP maar onderdelen A en C draaien nog elders. Bijvoorbeeld, een MySQL Database en Minio S3 Object Storage bucket draaien nog in de KIXS proeftuin, maar de applicatie draait wél al in een container op het Kubernetes platform van het BCP. Wij erkennen dat dit resultaat onwenselijk is, omdat het doel van deze migratietest is om vast te stellen in welke mate de MVP DBDAAP werkbaar is en kan voldoen aan de minimale gebruikersbehoefte van eindgebruikers. Dat betekent ook dat DBDAAP dan als de enige landingsplaats voor data-gedreven applicaties kan functioneren.

Uitvoering

Zoals eerder beschreven werd in deze migratietest gestart met het migreren van applicaties uit de bouwstraten die reeds een container-image als softwareproduct opleveren. Hierbij is het 'learn on the go' principe gehanteerd. Dit houdt in dat de applicaties uitgerold worden op de DBDAAP tenant van het BCP en gedurende dit proces de problemen zich presenteren. Wanneer de voorgedane problemen zijn opgelost, in samenwerking met het BCP-team, is de verwachting dat éénzelfde oplossing toegepast kan worden als een soortgelijk probleem zich voordoet bij een volgende applicaties die worden gemigreerd. Desalniettemin is het streven om zo spoedig mogelijk tot een (deel)resultaat te komen. Als voorbeeld, tijdens het migratieproject DP-1194 Drones en Veiligheid Vliegverkeer is gebleken dat de applicatie niet kon opstarten zonder verbinding met een ander component, namelijk de database. Om zo min mogelijk gehinderd te worden in de voortgang, beproeving en tevens het uitsluiten van problemen met een nieuw op te zetten database, werd er verbinding gelegd met een bestaande database waarvan een correct functioneren bekend staat bij het migratieteam. Hierdoor konden opvolgende problemen met de ingress sneller aangekaart en opgelost worden, maar bevindt de applicatie zich – weliswaar tijdelijk – in de eerder genoemde spagaatsituatie.

In deze acceptatietest wordt de migratie van een applicatie als "geslaagd" beschouwd indien (a) alle componenten van deze data-gedreven applicatie op DBDAAP zijn geland, (b) de applicatie op DBDAAP naar technisch behoren functioneert en (c) dankzij deze migratie de applicatie uit haar vorige plaats, zoals de KIXS Proeftuin, kan worden verwijderd zonder dat dit enig effect heeft op de werking of continuïteit van de applicatie die naar DBDAAP werd gemigreerd.

Resultaten van individuele migraties

De resultaten van de migraties van de applicaties worden in Tabel 4 weergegeven. Per applicatie wordt het eindresultaat, namelijk geslaagd of niet geslaagd, van de migratie gepresenteerd, volgt een overzicht van de componenten van de data-gedreven applicatie die wel en niet naar DBDAAP zijn gemigreerd, alsmede een evaluatie van wat vanuit technisch oogpunt ten aanzien tot DBDAAP goed en niet goed ging. Daarbij wordt de nadruk gelegd op welke technische problemen werden verholpen en welke technische problemen ertoe hebben geleid dat de migratie is mislukt.

Tabel 4. Resultaten van migraties per applicatie

Drones en Veiligheid Vliegverkeer

Eindresultaat	Niet geslaagd
Componenten	De applicatie bestaat uit de componenten: <ul style="list-style-type: none">● 1 relationele database (MySQL)● 2 containers● 1 S3-bucket
Evaluatie	<p>Wat ging technisch goed?</p> <ul style="list-style-type: none">● 2 containers zijn gemigreerd en gestart● 1 container is te benaderen via de ingress <p>Wat ging technisch mis?</p> <ul style="list-style-type: none">● Geen MySQL database beschikbaar, waardoor huidige instantie in de KIXS proeftuin gebruikt moest worden.● Verbinding tussen KIXS proeftuin en Kubernetes was niet mogelijk.● MySQL container image, helm chart of operator was niet beschikbaar voor ons via het u- of ud-account.● DNS entry voor ingress kon niet door ons worden aangemaakt.● TLS certificaat werkte niet.● Rapportageverzoek vanuit GUI geeft foutmelding van NSX Loadbalancer. <p>Hoe zijn technische problemen verholpen?</p> <ul style="list-style-type: none">● Firewall regels zijn aangemaakt door het BCP-team om MySQL en S3 verkeer mogelijk te maken tussen KIXS proeftuin en DBDAAP Kubernetes worker nodes.● DNS entry is handmatig door het BCP-team aangemaakt.● BCP-team heeft handmatig een wildcard TLS certificaat toegevoegd. <p>Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?</p> <ul style="list-style-type: none">● Niet volledig functioneren van de webapplicatie,

vanwege foutmelding bij dronerapportage-aanvraag van loadbalancer.

- Door technische problemen, die vervolgens opgelost moesten worden door BCP-team en/of Datalab-team, is er na twee weken nog geen volledige migratie. Verdere analyse van probleem bij rapportage-aanvraag zou kunnen leiden tot een werkende oplossing.

Mobiliteit Defensie – Analyse van reisbewegingen civiele dienstaauto's

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- 1 relationele database (MySQL)
- 1 container

Evaluatie Wat ging technisch goed?

- n.v.t.

Wat ging technisch mis?

- n.v.t.

Hoe zijn technische problemen verholpen?

- n.v.t.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Deze applicatie vereist een MySQL Database, TLS certificaat en DNS entry voor de ingress. De migratie van deze applicatie kent dezelfde uitdagingen als het Drones Project. Derhalve is de migratie voor deze applicatie (nog) niet gestart.

Vastgoed analyse

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- SQLite (in container)

Evaluatie Wat ging technisch goed?

- n.v.t.

Wat ging technisch mis?

- n.v.t.

Hoe zijn technische problemen verholpen?

- n.v.t.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Deze applicatie heeft vereist een TLS certificaat en DNS entry voor de ingress. De migratie van deze migratie kent dezelfde uitdagingen als het Drones Project. Derhalve is de migratie voor deze applicatie (nog) niet gestart.

Inzicht in de gevaren van Deep Fakes in het counter-IED domein

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- S3-bucket (nu bij KIXS)

Evaluatie Wat ging technisch goed?

- De datawetenschappers kunnen zelfstandig een S3-bucket bij DBDAAP aanvragen.

Wat ging technisch mis?

- De firewall stond geen verkeer tussen de Jupyter omgeving en de S3-bucket toe.

Hoe zijn technische problemen verholpen?

- Firewall regels zijn aangemaakt door het BCP-team om verkeer mogelijk te maken tussen Jupyter nodes en S3-buckets van DBDAAP.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Geen dealbreaker meer, kwestie van bucket aanvragen maar dit is afhankelijk van een handmatige verwerking door het BCP-team.

Kwaliteit in Beeld

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- Container

Evaluatie Wat ging technisch goed?

- n.v.t.

Wat ging technisch mis?

- n.v.t.

Hoe zijn technische problemen verholpen?

- n.v.t.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Deze applicatie heeft vereist een TLS certificaat en DNS entry voor de ingress. De migratie van deze applicatie kent dezelfde uitdagingen als het Drones Project. Derhalve is de migratie voor deze applicatie (nog) niet gestart

GeoStack

Eindresultaat
Componenten

Niet geslaagd

De applicatie bestaat uit de componenten:

- Leeromgeving
- Deb package repositories (o.a. voor Ubuntu en MongoDB)

Evaluatie

Wat ging technisch goed?

- Gelukt om container op platform te draaien met een xRDP ontsloten grafische XFCE desktop.

Wat ging technisch mis?

- Gnome is niet in een container te draaien.
- XFCE i.p.v. Gnome impliceert volledige herziening van het cursusmateriaal.

Hoe zijn technische problemen verholpen?

- BCP-team stelt voor om handmatig desktop VM's aan te bieden, i.p.v. automatische containers.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Gnome Desktop VM's zijn (nog) niet beschikbaar.
- Deb package repositories zijn nog niet beschikbaar gesteld.

IPMS Verkenner

Eindresultaat	Niet geslaagd
Componenten	De applicatie bestaat uit de componenten: <ul style="list-style-type: none"> • Container
Evaluatie	<p>Wat ging technisch goed?</p> <ul style="list-style-type: none"> • n.v.t. <p>Wat ging technisch mis?</p> <ul style="list-style-type: none"> • n.v.t. <p>Hoe zijn technische problemen verholpen?</p> <ul style="list-style-type: none"> • n.v.t. <p>Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?</p> <ul style="list-style-type: none"> • n.v.t. In verband met de belasting in tijd en aandacht voor problemen rondom de migratie van de applicatie "Drones en Veiligheid Vliegverkeer", werd niet aan deze migratie toegekomen. Tevens zou het oplossen van ingress (en eventueel loadbalancerproblematiek) randvoorwaardelijk zijn voor iedere container workload die ontsloten moet worden, waaronder de applicatie IPMS Verkenner.

JIVC Dashboard Kritieke Dienstverlening

Eindresultaat	Niet geslaagd
Componenten	De applicatie bestaat uit de componenten: <ul style="list-style-type: none"> • S3 Object Storage • Kubernetes • Power BI: gemigreerd
Evaluatie	<p>Wat ging technisch goed?</p> <ul style="list-style-type: none"> • De feitelijke migratie van het PBIX-bestand ging goed. <p>Wat ging technisch mis?</p> <ul style="list-style-type: none"> • Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties

van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

COVID-19 inzichten Dashboard CLSK

Eindresultaat
Componenten

Niet geslaagd
De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie

Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat

deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

Datalab Dashboard – Datalab platform in getallen

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

Gitlab dashboard

Eindresultaat
Componenten

Niet geslaagd

De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie

Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel

op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

ATFLOW HR IDU-voorspelling CLSK

Eindresultaat

Niet geslaagd

Componenten

De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie

Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

Datalab rapportages

Eindresultaat

Niet geslaagd

Componenten

De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie

Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang tot de server werd gegeven.

Matlogco Voertuigonderhoud

Eindresultaat Niet geslaagd

Componenten De applicatie bestaat uit de componenten:

- S3 object storage
- Kubernetes
- Power BI: gemigreerd

Evaluatie Wat ging technisch goed?

- De feitelijke migratie van het PBIX-bestand ging goed.

Wat ging technisch mis?

- Hoewel het PID DBDAAP doet vermoeden dat er een nieuwe Power BI Report Server wordt ingericht, blijkt dat de bestaande Report Server van BOO-IT werd gebruikt, met als gevolg dat alle bestaande Power BI dashboards die op deze server draaien onbedoeld op een productiewaardig systeem zijn komen te draaien, zonder dat daarvoor accreditaties van dashboards hebben plaatsgevonden. Dit resulteerde in de noodzaak, oftewel overhead, om 85 KIXS-dashboards negatief te migreren, omdat deze niet in productie kunnen / mogen.

Hoe zijn technische problemen verholpen?

- Deze problemen zijn verholpen door in overleg met GIT&Infra tot een tijdelijk voorstel voor een oplossing te komen, namelijk de dashboards die wel op de Report Server mogen draaien separaat op te slaan. Hiervoor moet een deugdelijkere oplossing worden gevonden.

Wat zijn de technische dealbreakers waardoor de migratie onmogelijk was?

- Er waren problemen omtrent de authenticatie en autorisatie tot de Power BI Report Server, waardoor pas op de voorlaatste dag van de migraties toegang

tot de server werd gegeven.

Resultaten van de acceptatietest

Uit Tabel 4 is gebleken dat alle 14 applicaties niet in zijn geheel naar DBDAAP werden gemigreerd, hetgeen moet worden opgevat als dat alle migraties op basis van de huidige technische inrichting van DBDAAP niet mogelijk was.

Vanuit technisch oogpunt bestaan er aanhoudende problemen met de volgende functionaliteit die door DBDAAP wordt aangeboden:

- ingress, vanwege een gebrek aan de automatisering van het aanmaken van DNS-records, uitgifte en beheer van TLS certificaten en de configuratie van de firewall;
- MRIN, waardoor het niet mogelijk blijkt om op DBDAAP benodigde software dependencies binnen de halen die door de applicaties zijn vereist;
- hergebruik van een bestaande Power BI Report Server van BOO-IT waar thans vele honderden Power BI dashboards draaien die niet op een productiewaardig systeem mogen landen (denk hierbij aan het IVD dashboard met meldingen over voorvallen dat onlangs door de CIO Office offline werd gehaald).

Conclusie

Het doel van deze acceptatietest was het vormen van een oordeel over de werkbaarheid van DBDAAP als minimum viable product. Deze test werd uitgevoerd met behulp van 14 bestaande data-gedreven applicaties die door het Datalab en KIXS werden ontwikkeld en reeds in de KIXS Proeftuin draaien.

Op basis van voorgenoemde resultaten concludeert het Datalab dat, op basis van de beschreven testresultaten, kennis en inzichten die tot en met 18 januari 2022 zijn opgedaan, het thans te prematuur lijkt om het MVP DBDAAP als productiewaardig platform te beschouwen. Ook ontvangt het Datalab graag een concreet plan waaruit blijkt hoe en wanneer de geïdentificeerde gebreken aangaande de fundamentele en essentiële functionaliteit zoals DNS, TLS certificaten, automatisering en MRIN, alsmede de openstaande actiepunten met

betrekking tot DBDAAP worden opgepakt en verholpen.

Tot deze conclusie is gekomen doordat fundamentele maar belangrijke functionaliteit, zoals de automatisering van DNS, TLS certificaten en de firewall, herhaaldelijk het slagen van de migratie van individuele applicaties heeft bemoeilijkt. Daarnaast is vanuit DBDAAP toegang tot MRIN noodzakelijk, omdat applicaties bepaalde dependencies hebben die nodig zijn om te functioneren. Voorgenoemde doet vermoeden dat het BCP-team mogelijk het functioneren van het platform niet in samenhang tot andere functionaliteit heeft getest, omdat het mislukken van voorgenoemde migraties er juist in lag dat componenten niet adequaat met elkaar konden samenwerken of communiceren. Ook bleek dat een bestaande Power BI Report Server wordt hergebruikt, wat in deze acceptatietest tot het inzicht leidde dat in plaats van de migratie van dashboards naar DBDAAP, we juist een groot aantal andere dashboards moesten verwijderen.

Voorts doen wij verslag van bevindingen die niet technisch van aard zijn en daarmee strikt genomen buiten de scope van deze acceptatietest vallen.

Ten eerste merkt het Datalab op dat DBDAAP thans eerder lijkt op een container-platform voor relatief eenvoudige taken dan een platform dat is ingericht voor big data en big data analytics. Hiermee komt DBDAAP (Defensie Big Data Advanced Analytics Platform) niet volledig tot zijn recht. Naar inschatting van het Datalab vergt dit nadere (door)ontwikkeling van DBDAAP om voor big data analytics inzetbaar te worden. Volgende stappen in de richting van big data analytics zouden in ieder geval een Hadoop-cluster, Apache Spark en een datawarehouse vereisen, waarmee grote hoeveelheden gestructureerde en ongestructureerde data wordt opgeslagen en gedistribueerd wordt gecomputeerd.

Ten tweede, vanwege het ontbreken van de selfservice en automatisering was het Datalab in grote mate afhankelijk van het BCP-team, met als gevolg dat de doorlooptijd van migraties langer werd. Deze afhankelijkheid, in relatie tot onvoorspelbaarheid wanneer een oplossing wordt geboden, was niet wenselijk. Daarnaast werd gesuggereerd dat deze automatisering reeds in DBDAAP versie 0.9 zou zijn ontwikkeld.

Ten derde erkent het Datalab het belang van ontwikkelwerkplekken voor data science-doeleinden. Uit de memo van 13 november 2021 blijkt dat deze werkplekken op 18 januari 2022 beschikbaar komen. Helaas heeft het Datalab deze werkplekken niet kunnen ervaren, wat voor deze migraties natuurlijk zeer relevant was. Het Datalab vraagt zich af hoe deze ontwikkelwerkplekken worden ingericht (e.g., een virtuele desktop of JupyterHub-omgeving), of de GPU's reeds

binnen zijn en wanneer deze werkplekken beschikbaar komen. Tegelijkertijd blijft onduidelijk wat de toekomst van de VDDI-omgeving is; gaan de ontwikkelwerkplekken op DBDAAP ook de VDDI-omgeving vervangen?

Ten vierde was de timing van de acceptatietest rond het nieuwe jaar zeer ongelegen; tussen Kerst en nieuwjaar ligt het werk bij JIVC nagenoeg stil. Het was wenselijk geweest wanneer het Datalab voldoende tijd werd gegund om de migraties uit te voeren.

Ten vijfde koos GIT&Infra voor een één-loket (█@mindef.nl), terwijl bleek dat juist één-op-één contact met deze collega's vele malen efficiënter bleek doordat dan vragen binnen tien minuten werden opgelost. Wel merkt het Datalab op dat het BCP-team waarschijnlijk volledig uit externe inhuur bestaat, terwijl het juist waardevol voor Defensie kan zijn om ook interne expertise te regelen.

Referenties

█ (2021, 13 november). *Memo: Basic Cloud Platform. Lancing Tenant DBDAAP*. Utrecht: Ministerie van Defensie.

█ (2020, 24 april). *Projectinitiatiedocument (PID): MVP Defensie Big Data Advanced Analytics Platform (MVP DBDAAP)*. Den Haag: Ministerie van Defensie.