Project Start Architectuur
RIVM Infectieradar

Document informatie

Auteur	:	5.1.2e
Datum	:	3-9-2020
Versie	:	1.6
Status	:	Definitief

Documentversies

Versie	Datum	Opmerkingen	Goedgekeurd
0.1		Initiële versie	
0.5	22-6-2020	Review door 5.1.2e, 5.1.2e, 5.1.2e 5.1.2e, 5.1.2e.	
0.8	23-6-2020	Input mbt netwerk verwerkt.	
0.9	30-6-2020	Input van 5.1.2e, 5.1.2e verwerkt.	
1.0	1-7-2020	Als in overleg met 5.1.2e op 1-7-2020 besproken en vastgesteld als de 1.0	
1.1	31-7-2020	Input risicoinventarisatie verwerkt.	
1.2	20-8-2020	Input 5.1.2e verwerkt. Overige updates verwerkt.	
1.3	28-8-2020	Input mbt performance, e-mail-afhandeling, migratie en beheer verwerkt.	
1.4	30-8-2020	Aanpassingen functionaliteit vragenlijst, afhandeling e-mail.	
1.5	2-9-2020	Tekstuele aanpassingen.	
1.6	3-9-2020	Review 5.1.2e verwerkt.	

Referenties

Documenttype	Documentnaam / Verwijzing	Datum
Requirementsdocument	InfluenzaNet2.0-Requirements-v1.5.pdf	9-1-2019
Architectuurbeschrijving door Coneno	InfluenzaNet2.0-Architecture-v1.3.pdf	9-1-2019
Systeemoverzicht	2020-06-01_survey-overview.pdf	01-06-2020
Project flyer	InfluenzanetVsVirus.pdf	Q1 2020
Beschrijving Influenzanet	2020-04_study-system-decription	17-4-2020
Beschrijving ontkoppeling persoons en medische gegevens	2020-06-03_id-connections.pdf	03-06-2020
Toelichting coneno als leverancier	Referentie onderzoek Coneno.docx	2-9-2020
Beschrijving Dataflows	Dataflow infectieradar beschrijving – vanuit alle actoren 1.3.docx Infectieradar-flows-v2.vsdx	2-9-2020
Systeemdecompositie	Systeemdecompositie Influenzanet-v4.docx	25-8-2020

Inhoudsopgave

1 Inleiding

Deze Project Start Architectuur (PSA) is een projectdocument dat als hulpmiddel bij het project wordt ingezet om veranderingen te faciliteren. De PSA richt zich daarbij op kaders die op dit project van toepassing zijn en de impact van deze kaders op de beoogde verandering. De PSA maakt concreet wat architectuur voor dit project betekent. Deze PSA beschrijft de toepassing van RIVM Infectieradar voor het monitoren van COVID-19.

1.1 Context van het project

Het RIVM houd op verschillende manieren de verspreiding van infectieziekten in de gaten. Een van de voorzieningen die daarvoor worden ingezet is Infectieradar.nl: een webapplicatie waarbij middels wekelijkse vragenlijsten deelnemers aangeven of en welke klachten zij hebben. In tijden van Covid-19 is het belang van deze voorziening toegenomen. Het onderhavige project ziet op een optimalisatie van de bestaande voorziening.

1.2 Doelstelling van het project

In dit project gaat het om de ontwikkeling van de applicatie 'RIVM Infectieradar' en de overdracht ervan naar de beheer-organisatie bij SSC-Campus en ten slotte de ingebruikname. Onderdeel van dit proces is het uitvoeren van risico-inventarisatie, risicoanalyse en privacy impact analyse.

2.1 Wat is RIVM Infectieradar

Het doel van RIVM Infectieradar is het monitoren van (de verspreiding) klachten mogelijk veroorzaakt door infectieziekten zoals COVID-19 . Deelnemers vullen daarvoor periodiek een vragenlijst in middels een web-applicatie. Aan de hand van verzamelde data kunnen analyses uitgevoerd worden over verspreiding van infectieziekten.

Deelnemers doen vrijwillig mee. Deelnemers registreren zich bij de landelijke webapplicatie RIVM Infectieradar. Met een bevestigings-email met link bevestigt de deelnemer deelname (en juistheid van het opgegeven e-mailadres). Daarna kan de deelnemer inloggen en meedoen. Een deelnemer kan zelf in zijn account de gezinsleden (jonger dan 16 jaar) opvoeren en voor die gezinsleden vragenlijsten invullen.

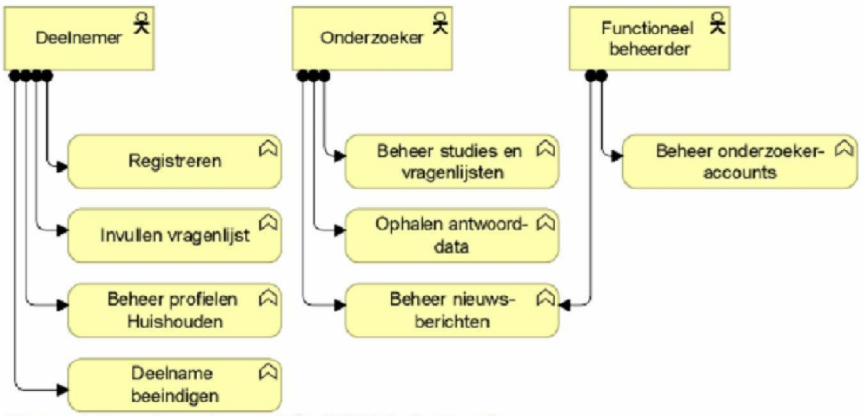
Deelnemers worden uitgenodigd via e-mail om een vragenlijst in te vullen. De e-mail bevat dan een link met daarin een token. Via de link wordt de deelnemer gevraagd in te loggen. Daarvoor is een wachtwoord vereist. Na succesvolle inlog wordt direct de vragenlijst geopend. In Bijlage A is een flow-chart opgenomen waarin het proces dat de deelnemers die aan Infectieradar.nl deelnemen is beschreven.

Nieuwsberichten worden getoond in een apart venster waarvan de inhoud door een zgn. Drupal-instantie gevoed wordt. Onderzoekers die als redacteur zijn opgevoerd in Drupal kunnen de inhoud van de berichten aanpassen.

Een functioneel beheerder beheert accounts voor onderzoekers.

Onderzoekers kunnen de inhoud van de vragenlijsten configureren afhankelijk van waar de focus van onderzoek op een gegeven moment ligt. Onderzoekers kunnen de antwoorden van vragenlijsten downloaden. De antwoorden van vragenlijsten in de download zijn niet tot personen herleidbaar. De download bevat geen account-gegevens.
Zowel deelnemer-accounts als antwoorden van de vragenlijsten blijven in Nederland.

RIVM Infectieradar wordt gehost bij het RIVM. De partij waarmee wordt samengewerkt voor de ontwikkeling van de RIVM Infectieradar-applicatie, Coneno, heeft geen toegang tot RIVM-systemen.



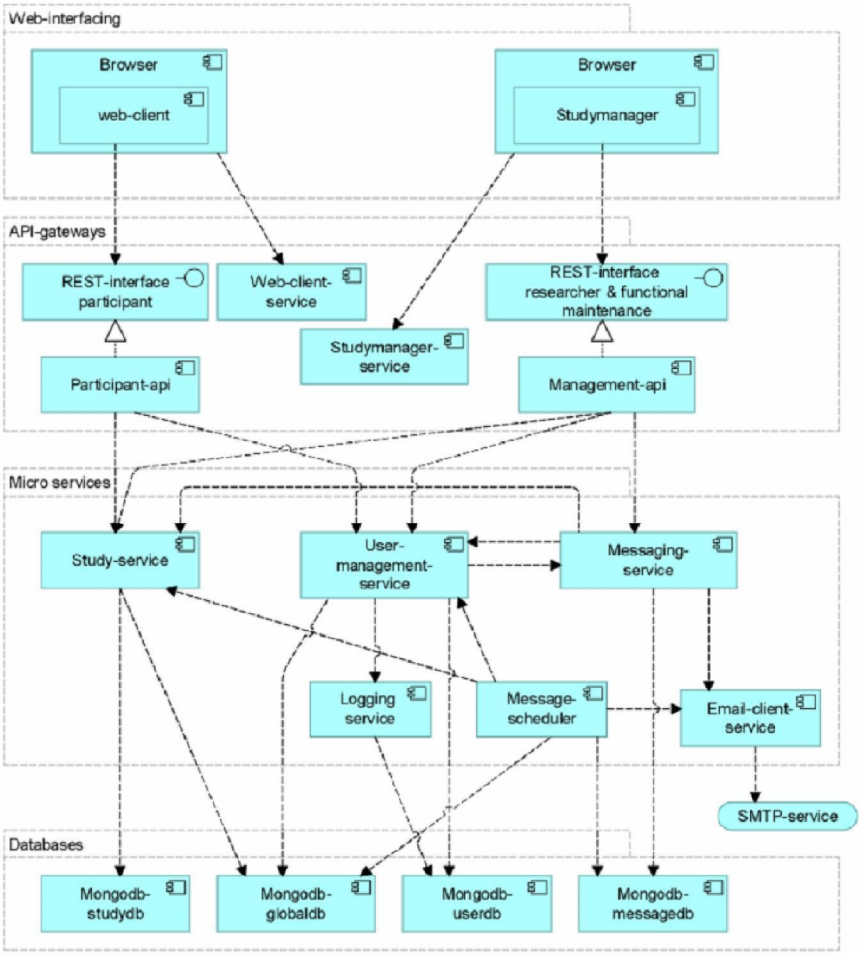
Figuur 1 Functioneel overzicht RIVM Infectieradar

RIVM Infectieradar wordt ontwikkeld door coneno GmbH[1].

2.2 **Ontwerp RIVM Infectieradar**

RIVM Infectieradar is een op microservices gebaseerde applicatie. De functionaliteit van de applicatie wordt gerealiseerd door een aantal onderling verbonden deel-applicaties (zie figuur 2). In Bijlage C is een gedetailleerd systeemoverzicht opgenomen.

[1] Een toelichting op de rol en achtergrond van Coneno alsmede de keuze voor deze partij is vermeld in "Referentie onderzoek Coneno.docx".



Figuur 2 Applicatiecomponenten waaruit RIVM Infectieradar is opgebouwd.

De deelnemer logt in via de web-applicatie met usernaam en wachtwoord. Via e-mail wordt een 6-cijferige code toegestuurd die vereist is om de inlog af te ronden. Authenticatie vindt plaats met behulp van de User-management-service. Als authenticatie slaagt krijgt de deelnemer een *Access Token* waarmee de ingelogde persoon bij volgende requests wordt geïdentificeerd. De geldigheidsduur van het access token is beperkt tot 5 minuten. Een uitwerking van het inlog-proces is weergegeven in Bijlage B. RIVM Infectieradar heeft een eigen account-database. RIVM-gebruikers loggen conform de architectuur SSC-Campus[2] in, met het RIVM e-account. Toegangsrechten worden dan via de Active Directory ingeregeld.

[2] Zie "Architectuur SSC-Campus", januari 2018

De diverse services communiceren middels het gRPC-protocol. Hier is geen sprake van authenticatie. Services authenticeren bij de databases middels SCRAM (Salted Challenge Response Authentication Mechanism).

Onderzoeker en beheerder gebruiken de REST-api met tooling, bijv. Postman of Insomnia. De API's zijn gedocumenteerd op Postman-website op de volgende locatie:
5.1.2h
De Management-API is alleen te bereiken vanuit het RIVM-netwerk.

De code van RIVM Infectieradar is beschikbaar als open source op
5.1.2h

2.3 Bouwblokken

RIVM Infectieradar is opgebouwd uit een aantal RIVM architectuur-bouwblokken. In onderstaande tabel wordt aangegeven welke bouwblokken het betreft en of deze zijn hergebruikt, of nieuw zijn opgevoerd in dit project.

SOLL → / ↓ Bouwblok	Hergebruik	Aanpassen	Bestaand bouwblok introduceren	Nieuw	Elimineren	Toelichting
Openshift platform	X					Het openshift platform is operationeel.
RIVM Infectieradar				X		Het ophalen van sources, bouwen van docker images en deployment in Openshift
E-mail server			X			Nieuwe E-mail-service voor RIVM Infectieradar vanwege groot e-mail-volume. Whitelisting van mailadres RIVM Infectieradar.
F5	X					Gebruik van de F5 als reverse proxy.
MongoDB				X		Alle data wordt opgeslagen in 4 instanties van MongoDB.

2.4 Autorisatie

RIVM Infectieradar kent 3 rollen:

- Deelnemer : invullen vragenlijsten, beheer eigen profiel en dat van gezinsleden jonger dan 16 jaar , afmelden.
- Onderzoeker: beheer van vragenlijst-templates, ophalen ingevulde vragenlijsten, beheer van tekst van emails.

- Admin: beheer onderzoeker-accounts, beheer vragenlijst-templates, ophalen ingevulde vragenlijsten, beheer van tekst van emails.

Vragenlijsten worden beheerd door de onderzoeker die de vragenlijst aanmaakt en het onderzoek uitvoert.

Het eerste account wordt met een MongoDB-query in de Userdatabase aangemaakt door een applicatiebeheerder en voorzien van de admin-role. Met dit account worden via de management-API de onderzoeker-accounts aangemaakt. Onderzoeker-accounts kunnen geen accounts wijzigen.

De autorisaties behorend bij een rol zijn verwerkt in de applicatiecode en niet configureerbaar.

2.5 **Autorisatie ten behoeve van beheer**

Applicatiebeheerders en beheerders van de infrastructuur (waar onder OpenShift) loggen in op systemen met het account dat is gekoppeld aan het RIVM IAM. Daar wordt een 2-factor login toegepast (of 1 factor wanneer op RIVM-terrein).
Toegang tot RIVM Infectieradar binnen Openshift is middels RIVM IAM alleen mogelijk voor 6 applicatie- en infrastructuur-beheerders. Deze beheerders hebben toegang tot alle aspecten van de applicatie, ook de databases en de sleutels die gebruikt worden voor encryptie. Dat deze beheerders toegang tot alle aspecten van de applicatie hebben is omdat zij voor het onderhoud en beheer op de applicatie in moeten kunnen ingrijpen op alle onderdelen van de applicatie en het onderliggende systeem. In verband met vereiste hoge beschikbaarheid van RIVM Infectieradar zijn beheerfuncties ook meervoudig ingevuld.

Door de gebruikte versleutelingen zijn accountgegevens en ingevulde vragenlijsten niet zonder meer te koppelen door beheerders. Daarvoor is het nodig om het algoritme dat gebruikt wordt in de applicatie voor de pseudonimisering, te reproduceren buiten RIVM Infectieradar.

In de logging van Openshift kunnen de gelogde acties niet herleid worden tot de identiteit van degene die ze uitvoert.

2.6 **Technologiestack**

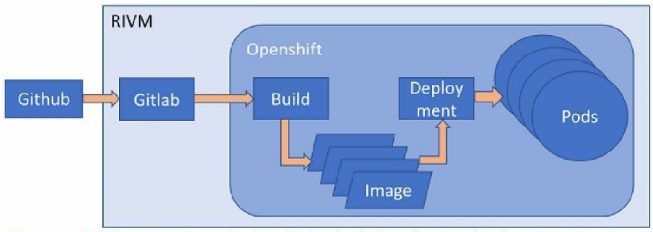
De applicatie gebruikt de volgende technologieën

- Ontwikkel-taal: Go
- Frontend: React, Flutter (UI toolkit for mobile, web, desktop)
- Communicatie tussen microservices: gRPC
- Database: MongoDB
- Deployment en hosting: Docker en Kubernetes.
- Operating system: RHEL 7 base image.

2.7 **Hosting**

RIVM Infectieradar wordt gehost bij SSC-Campus (onderdeel van RIVM). De huisvesting van systemen vindt plaats bij het Rijksoverheid Datacenter bij Equinix.

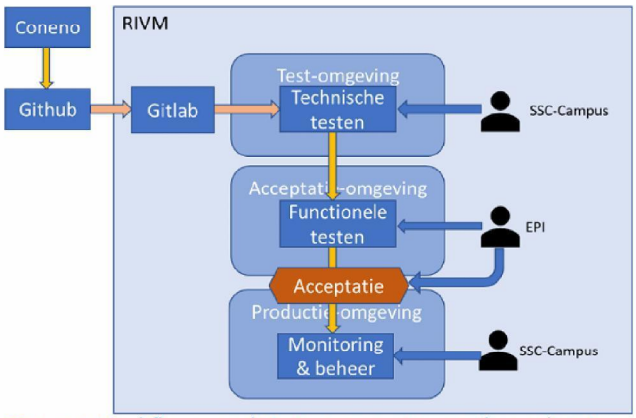
RIVM Infectieradar wordt gehost op het Openshift containerplatform. Docker images worden gebouwd in Openshift. De Build-configuration beschrijft de bouw-stappen. Deployment kan daarna plaats vinden. In Deployment-configuraties wordt de runtime-omgeving ingesteld. Iedere component/service wordt in een aparte pod gestart.



Figuur 3 Flow van code in Github (cloud-service) naar werkende applicatie in Openshift-pods: een verzameling van 1 of meer Docker-containers.

Voor Test, Acceptatie en Productie zijn aparte Openshift-projecten ingericht. Op basis van tags in de Git-repository wordt een specifieke versie van de code van RIVM Infectieradar opgehaald in Test, Acceptatie en Productie.

Coneno publiceert de sources op Github. RIVM haalt de sources op van github en plaatst deze in de interne gitlab gitlab.int.ssc-campus.nl.



Figuur 4 Workflow waarbij nieuwe versie van de applicatie-code via Test- en Acceptatie-omgevingen naar de Productie-omgeving worden gebracht.

Testen vinden plaats in test- en acceptatieomgeving. Nieuwe releases worden aan geautomatiseerde regressietesten onderworpen. Na succesvolle uitvoering van deze testen wordt de nieuwe release beschikbaar gesteld op de acceptatieomgeving. Daar vinden functionele testen plaats door de systeem-eigenaar (EPI). Pas na acceptatie door de

systeem-eigenaar wordt een nieuwe versie van de software in productie genomen (zie figuur 4)

Wanneer de belasting hoog is kunnen de services in RIVM Infectieradar onafhankelijk van elkaar worden opgeschaald. Voor het opschalen van MongoDB-servers zijn specifieke MongoDB-technieken vereist, bijv. Sharding. Deze technieken zijn (voor MongoDB) nog niet eerder toegepast binnen RIVM. Performancetesten op de acceptatieomgeving geven aan dat 7500 ingezonden vragenlijsten per minuut langdurig kunnen worden verwerkt. Korte pieken tot 20.000 ingezonden vragenlijsten worden eveneens goed verwerkt. Dit blijkt uit de uitgevoerde performancetesten.

2.8 **E-mail-afhandeling**

Deelnemers worden wekelijks via e-mail uitgenodigd een vragenlijst in te vullen. Vooralsnog wordt met 100.000 deelnemers gerekend. Gestart wordt met het uitnodigen van de bestaande 55.000 deelnemers aan Infectieradar.
Om deze mail-load te kunnen verwerken zijn 3 nieuwe Exchange-servers ingericht specifiek voor RIVM Infectieradar. 2 Servers zijn voor het versturen van de herinnerings-emails. Een derde server is ingericht om de e-mails met inlogcodes te versturen. E-mails met inlogcodes kunnen daardoor snel verstuurd worden onafhankelijk van de queue van herinnerings-emails.

In onderstaande tabel is het aantal wekelijks te versturen e-mails gespecificeerd bij een mogelijk scenario voor het uitnodigen van bestaande deelnemers van Infectieradar, en twee piekscenario's met veel nieuwe aanmeldingen bij ruim 55.000 en 90.000 bestaande deelnemers.

Tabel 1 Specificatie van het mail-volume tijdens uitnodigen van bestaande deelnemers van Infectieradar en doorgroei.

	Aantal nieuwe deelnemers	Aantal bestaande deelnemers	Confirmatie email	Email inlog-code (50% B)	Email wachtwoord vergeten (3% C)	Wekelijkse reminder	Wekelijkse reminder verdeeld over 7 dagen	totaal	load veroorzaakt door wekelijse reminder
termijn voor versturen			binnen 12 uur	direct na inlog	direct	binnen 7 dagen	binnen 1 dag		
moment 1	100	0	100	50	0	100	14	250	40%
moment 2	1.000	100	1.000	500	3	1.100	157	2.603	42%
moment 3	15.000	1.100	15.000	7.500	33	16.100	2.300	38.633	42%
moment 4	20.000	16.100	20.000	10.000	483	36.100	5.157	66.583	54%
moment 5	20.000	36.100	20.000	10.000	1.083	56.100	8.014	87.183	64%
stationaire	1.000	56.100	1.000	500	1.683	57.100	8.157	60.283	95%
piekmoment 1	10.000	57.100	10.000	5.000	1.713	67.100	9.586	83.813	80%
piekmoment 2	10.000	90.000	10.000	5.000	2.700	100.000	14.286	117.700	85%

Er zijn twee fasen en bijbehorende aandachtspunten:

1. *Opvoeren van de bestaande deelnemers aan Infectieradar*. De verwachting is dat deelnemers na aanmelden, inloggen om hun profiel te bekijken en te wijzigen. Daardoor moeten login-codes verstuurd worden die niet vertraagd kunnen worden.

2. *Autonome groei in het aantal deelnemers*. Mogelijk pieken in het aantal aanmeldingen direct na media-aandacht. Als het aantal deelnemers is opgebouwd moeten wekelijks 100.000 uitnodigings-mails verstuurd kunnen worden.

Bij 100.000 deelnemers zal continue zo'n 600 e-mail's per uur verstuurd worden (exclusief RIVM-email). Tijdens opschalen wordt grofweg de helft van de e-mails veroorzaakt door inloggen. Dit zal minder verspreid zijn over de week en dag.

Mitigatie voor fase 1:
Uitnodigingen voor bestaande deelnemers van Infectieradar verspreid versturen door de week. Daardoor zullen vervolg-acties van deelnemers in RIVM Infectieradar ook verspreid raken.

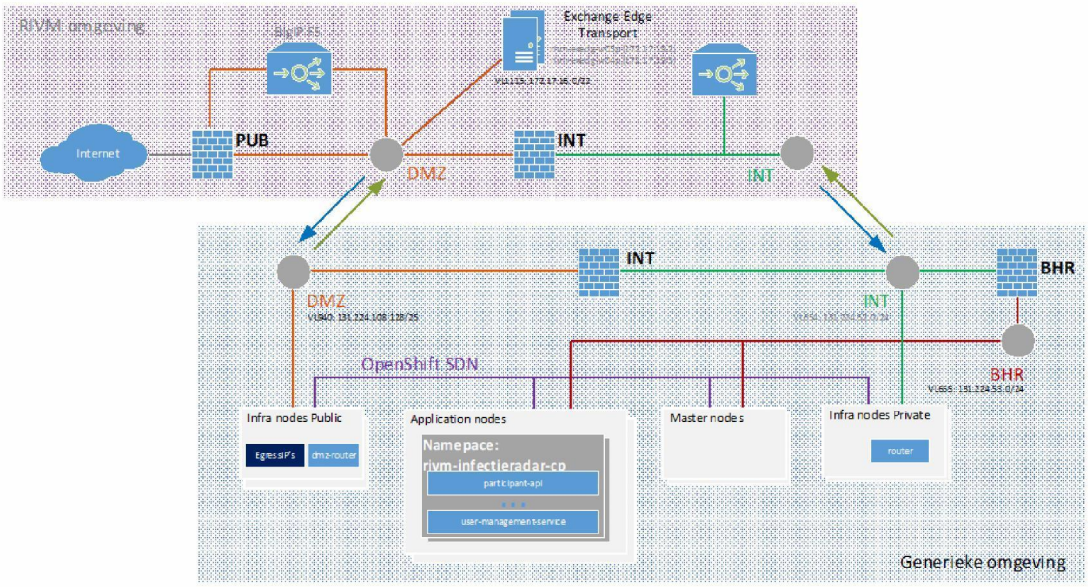
Mitigatie voor fase 2:
Aan nieuwe deelnemers wordt in de applicatie een dag van de week gekoppeld waarop de deelnemer de herinneringsmail zal ontvangen. De dagen van de week worden gelijk verdeeld over nieuwe deelnemers.

Mitigatie black-listing in beide fasen:
- De stroom uitgaande mail wordt langzaam opgevoerd in de test-periode.
- Rate-limiting wordt toegepast op het aantal nieuw te maken accounts in een korte tijdspanne en op het (opnieuw) verzenden van bevestigings-e-mails.

2.9 **Netwerkinrichting**

De Openshift-projecten voor Acceptatie en Productie zijn gedefinieerd als 'extern'. Dat houdt in dat de applicatie in de semi-vertrouwde netwerkzone is geplaatst. Deze zone is afgeschermd naar buiten door de F5 reverse proxy. De Test-instantie is gedefinieerd in het interne netwerk.

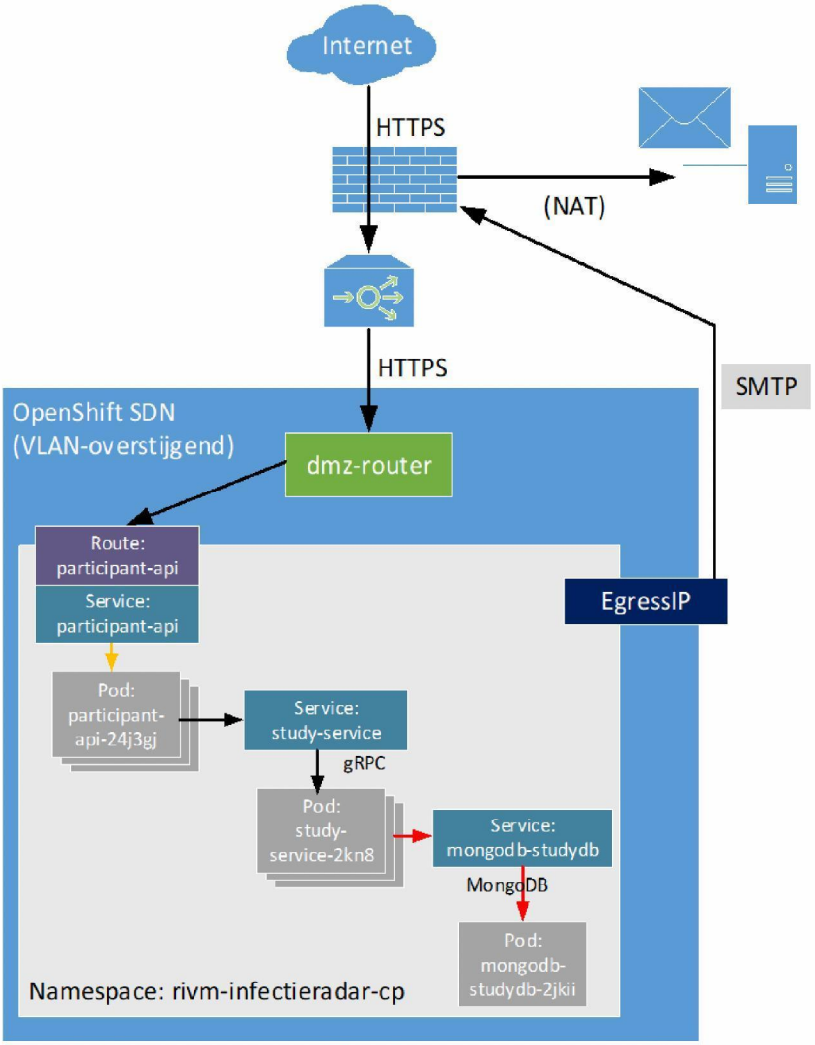
Het Openshift Software Defined Network (SDN) word ontsloten via de DMZ en via het interne netwerk. Via tags die aan een namespace zijn gekoppeld wordt bepaald of de namespace intern of extern wordt ontsloten. Zie figuur 5.



Figuur 5 Schematische weergave van de netwerk-ontsluiting van Openshift.

Processen binnen een pod zijn niet van elkaar afgeschermd (hier wordt geen gebruik van gemaakt). Processen in verschillende pods kunnen elkaar (alleen) benaderen via in het Openshift-project gedefinieerde 'services'. Processen in verschillende Openshift-projecten kunnen elkaar niet bereiken, tenzij routering buiten Openshift is ingesteld.
Processen kunnen alleen van buiten Openshift benaderd worden als een 'route' is gedefinieerd voor de 'service' behorend bij het proces.
Daarnaast is routering binnen met RIVM-netwerk vereist.
Deze inrichting bewerkstelligd dat alleen voor de applicatie noodzakelijke verbindingen tot stand gebracht kunnen worden. Mocht een onderdeel van de applicatie gecompromitteerd worden, dan nog is het onmogelijk om verbindingen naar andere onderdelen van de RIVM-infrastructuur of daarbuiten te leggen.

Buiten de beschreven mogelijkheden voor communicatie zijn pods volledig van elkaar afgeschermd.



Figuur 6 Ontsluiting van openshift-pods via services en ontsluiting buiten Openshift via een route. Via een gedefinieerd egress-ip adres kan een pod verbinding leggen met externe services.

Een overzicht van de gebruikt URL's:

url	extern zichtbaar	interne url	service	poort
webclient.acc.infectieradar.nl	Ja	5.1.2h	web-client-server-a	8080
participantAPI.acc.infectieradar.nl	Ja	5.1.2h	participant-api	3231

		5.1.2h		
5.1.2h	Nee	5.1.2h	management-api	3231

Productie

url	extern zichtbaar	interne url	service	poort
webclient.infectieradar.nl	Ja	5.1.2h	web-client-server-p	8080
participantAPI.infectieradar.nl	Ja	5.1.2h	participant-api	3231
5.1.2h	Nee	5.1.2h	management-api	3231

2.10 **Performancetesten**

Een aantal testen zijn uitgevoerd om het gedrag van de applicatie en infrastructuur onder hoge belasting te onderzoeken. Een samenvatting van de uitgevoerde testen:

Geteste actie	Intensiteit en duur	Responsetijden (95%)
Opvragen van de aanmeldpagina Aanmelden	10.000 per minuut gedurende 10 minuten	lager dan 700 ms
Inzenden wekelijkse vragenlijst	1500 – 7500 per minuut gedurende 10 minuten	lager dan 70 ms
Inzenden wekelijkse vragenlijst	10.000-20.000 per minuut gedurende 5 minuten	lager dan 150 ms
Inzenden wekelijkse vragenlijst	40.000 per minuut gedurende 5 minuten	lager dan 1800 ms tot 1.9% fouten bij ophalen inlogpagina.

De resultaten geven aan dat de applicatie vragenlijsten die 7500 gebruikers per minuut insturen goed en voor langere tijd kan verwerken. Korte pieken van 10.000 – 20.000 worden goed verwerkt. Bij hogere belasting nemen de responsetijden toe tot ongewenste niveaus en treden fouten op. Inmiddels is rate-limiting ingesteld op de aanmeldpagina waardoor het aantal nieuwe aanmeldingen is te maximeren. Een gedetailleerd testrapport is beschikbaar.

2.11 **Migratie**

De bestaande deelnemers aan Infectieradar worden uitgenodigd om deelname voort te zetten in het nieuwe systeem (RIVM Infectieradar). De accounts voor deze deelnemers worden middels scripting in de database

aangemaakt. De deelnemers krijgen een uitnodiging via E-mail met daarin een link waarmee het wachtwoord ingesteld kan worden (zelfde functionaliteit als 'wachtwoord-vergeten'). Als de deelnemer het wachtwoord instelt is daarmee het nieuwe account geverifieerd en actief. Deelnemers worden in en aantal tranches uitgenodigd.

Om de gegevens in de bestaande Infectieradar te kunnen correleren aan de gegevens in RIVM Infectieradar wordt de deelnemer-ID uit de oude Infectieradar opgeslagen in de SurveyDB in RIVM Infectieradar. De koppeling met de nieuwe ID vindt plaats op basis van het E-mailadres. Bij de download van ingevulde vragenlijsten door de onderzoeker wordt de 'oude Infectieradar-ID' meegegeven bij iedere ingevulde vragenlijst. Zodra de analyse is opgezet en samenvoeging van de oude en nieuwe data is voltooid wordt de koppeltabel (E-mailadres, oude Infectieradar-ID) definitief verwijderd. Daardoor is het niet meer mogelijk om ingevulde vragenlijsten, via een oude Infectieradar-ID, te koppelen aan E-mailadres en daarmee aan een persoon.

2.12 **Beheer**

Het Technisch applicatiebeheer wordt uitgevoerd bij RIVM (organisatieonderdeel SSC-Campus) door de afdeling Applicatie en Functionaliteiten Management. Uitgevoerde activiteiten:

- Monitoring van de applicatie: performance, resource-gebruik;
- Uitrol releases;
- Afhandeling afwijkingen en storingen in de applicatie;
- Coordineren wijzigingen (infrastructuur, applicatie);
- Leveranciermanagement.

Het volledige takenpakket is beschreven in de Product Dienst Catalogus van SSC-Campus. Specifieke beheer-afspraken worden vastgelegd in een beheer-plan.

Technisch applicatiebeheer beheert de onderzoeker-accounts in RIVM Infectieradar.

Het functioneel beheer, dwz. het beheer van de vragenlijst-templates en berichten-templates, ligt bij de afdeling EPI.

Voor het onderhoud (preventief, correctief, blijvend voldoen aan IB- en privacy-eisen) en de doorontwikkeling van RIVM Infectieradar worden afspraken gemaakt tussen RIVM en Coneno.

Geadviseerd wordt de volgende punten te agenderen voor doorontwikkeling van RIVM Infectieradar op korte termijn:

- Koppeling van de RIVM Infectieradar-accounts van onderzoekers en functioneel beheerders met RIVM-Identity and Access Management;
- Het verwijderen van onderzoeker-rechten bij de 'admin'-rol in de applicatie;
- Verwijderen van log-informatie na een instelbare bewaarperiode;
- Aanvullende maatregelen voor rate-limiting op de aanmeldpagina;

- Niet-geverifieerde accounts verwijderen nadat de verificatie-periode is verstreken.

3 Informatiebeveiliging en Privacy

Van toepassing zijnde wet&regelgeving, kaders en richtlijnen:

- Wet-en regelgeving: voldoen
  - VIR: Voorschrift informatiebeveiliging Rijk;
  - VIR-BI: Voorschrift Informatiebeveiliging Rijksdienst-bijzondere informatie;
    - Uitwerking BIO: Baseline Informatiebeveiliging Overheid;
  - AVG: Algemene verordening gegevensbescherming;
  - UAVG: Uitvoeringswet Algemene Verordening Gegevensbescherming;
  - Archiefwet.

- Aanvullende kaders: uitwerking van regelgeving
  - ICT-Beveiligingsrichtlijnen voor Webapplicaties, NCSC;
  - Webrichtlijnen / Rijkshuisstijl ;
  - Lijst verplichte open standaarden forumstandaardisatie;
  - Architectuur overheid NORA.
  - Centrum Informatiebeveiliging en Privacy:
    - Serie documentatie "Grip op Secure Software Development"

- RIVM-specifiek
  - Privacy-beleid van het RIVM (Vertrouwd omgaan met persoonsgegevens[3]);

De exacte set aan gegevens die verzameld wordt binnen RIVM Infectieradar geconfigureerd binnen een 'studie' middels een vragenlijst-template. De PSA beschrijft de toepassing van RIVM Infectieradar voor het monitoren van COVID-19, met een daar op toegesneden vragenlijst. Deze vragenlijst bevat geen open vragen. Er zijn twee varianten van de vragenlijst: de basis-vragenlijst voor deelnemers die tot nu toe geen klachten hebben gemeld en een vragenlijst voor deelnemers die in de vorige vragenlijst klachten hebben gemeld. Dit kenmerk (wel/geen klachten) wordt als bijzonder persoonsgegeven verwerkt.

RIVM Infectieradar, geconfigureerd met de Covid-19-vragenlijst, verwerkt persoonsgegevens, demografische persoonsgegevens en bijzondere persoonsgegevens betreffende de gezondheid van personen. De verwerkte gegevens zijn in detail beschreven in de Gegevensbeschermings-effectbeoordeling (PIA).

Privacy-maatregelen

De volgende privacy-beschermingsmaatregelen zijn genomen:

- Accountgegevens (e-mailadres, wachtwoord) worden apart opgeslagen van de overige persoonsgegevens.
- De Studie-ID wordt gehashed en versleuteld opgeslagen bij de vragenlijst-antwoorden (zie figuur 7). De vragenlijst-antwoorden zijn daardoor niet meer herleidbaar tot een persoon.
- Ingevulde vragenlijsten die opgevraagd worden door onderzoekers zijn niet herleidbaar tot een persoon.
- De bewaartermijn van accountgegevens (waaronder e-mailadres): accountgegevens worden bewaard voor een periode van 2 jaar na de laatst ingevulde vragenlijst. Na deze bewaartermijn wordt het betreffende account verwijderd. Daarmee zijn de ingevulde vragenlijsten niet meer herleidbaar naar een persoon.
- Er is geen inzage door de deelnemer mogelijk in eerder ingevulde vragenlijsten als een deelnemer inlogt. Alleen e-mail adres wordt getoond.
- E-mails worden geautomatiseerd verstuurd vanuit de applicatie; onderzoekers hebben geen inzage in de e-mailadressen.
- Logging van events rond het instellen, opheffen en wijzigen van RIVM Infectieradar-accounts.

Toelichting op pseudonimisering

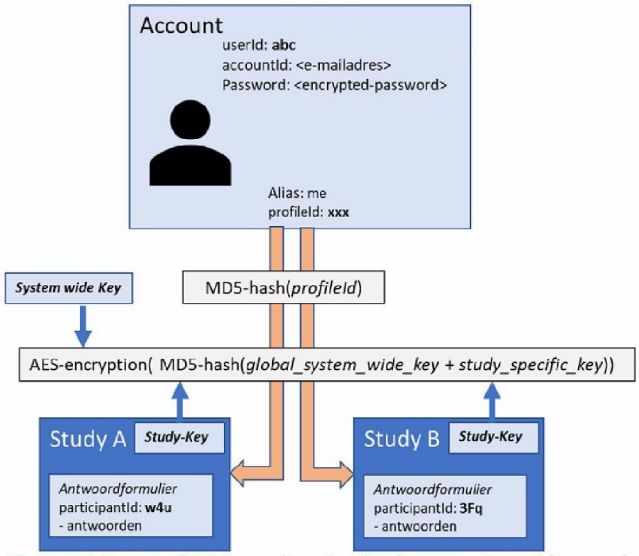
Iedere deelnemer en aangemeld gezinslid heeft een unieke 'profile-ID'. Een profile is gekoppeld aan een account, waarin het E-mailadres, wachtwoord en account-ID is opgeslagen. Per studie wordt de profile-ID vertaald in een unieke participant-ID. Voor een deelnemer of gezinslid is de participant-ID in verschillende studies ook verschillend.

De participant-ID komt als volgt tot stand (zie figuur 7):

1. Van de profile-ID wordt een md5-hash bepaald.
2. Een encryptie-sleutel wordt bepaald door een md5-hash uit te voeren op een aaneenschakeling van een globale sleutel en een studie-specifieke sleutel (strings).
3. Op de in stap 1 bepaalde hash-waarde wordt een AES-encryptie toegepast met de in stap 2 bepaalde encryptie-sleutel. De AES-encryptie is in CTR-mode. De uitkomst van deze stap is de participantID.

Om de link tussen profile-ID en profile-ID te herstellen zijn de volgende stappen nodig:

- Toegang krijgen tot de account- en profiel-gegevens in de UserDB.
- Toegang krijgen tot de studie-data en studie-sleutel in de StudyDB.
- Toegang krijgen tot de systeem-globale sleutel in de GlobalDB.
- Kennis hebben van het gebruikte algoritme voor pseudonimisering, bijv. door bestuderen van de source-code.
- Buiten RIVM Infectieradar: van profile-ID's de participant-ID's bepalen.
- Nu pas kan een profiel-ID gerelateerd worden aan een participant-ID en daarmee met de bijbehorende ingevulde vragenlijsten.



Figuur 7 Het profileId wordt gehashed en versleuteld opgeslagen. Daarbij wordt een systeem-key gebruikt en een key per studie. Daarmee is de relatie tussen antwoordformulier en profiel en account niet meer te leggen.

Salting is niet toegepast omdat de applicatie op meerdere momenten de participant-ID moet kunnen bepalen op basis van de profile-ID en deze transformatie deterministisch moet zijn.

Maatregelen voor Informatiebeveiliging

De BBN-classificatie van RIVM Infectieradar in combinatie met de Covid-19 intake- en wekelijkse vragenlijst is BBN2. Een risico-inventarisatie en analyse op de applicatie (dus niet op de standaard dienstverlening) heeft plaats gevonden waarin te mitigeren dreigingen zijn bepaald.

De volgende maatregelen zijn genomen in het kader van informatiebeveiliging, aanvullend aan de maatregelen die vanuit de standaard-dienstverlening van SSC-Campus (Managed Server Hosting, Container Hosting, Toegang) plaats vinden :

- Plaatsing van de applicatie-servers in de semi-vertrouwde netwerkzone achter de centrale proxy.
- Aparte Openshift namespaces voor Test, Acceptatie en Productie, afgeschermd met accounts onder regiem van RIVM Identity&Access Management.
- De Management-API is alleen bereikbaar vanuit het RIVM-netwerk.
- Versleuteling in database van e-mailadres en wachtwoorden.
- Versleuteld versturen van gegevens (https) met TLS 1.2 of hoger.
- Inlog- en refresh-tokens hebben een beperkte geldigheidsduur van 5 minuten (hiermee wordt ook de 'sessie-duur' beperkt) .

- Beperken van het aantal nieuwe accounts dat in een korte periode (10 minuten) aangemaakt kan worden.
- Beperken van achtereenvolgende inlog-pogingen.
- Beperken van achtereenvolgende verzoeken om bevestigings-email te sturen.
- RIVM maakt gebruikt van dienstverlening van SURF, SURFcert, voor bescherming tegen DDOS-aanvallen.
- Wachtwoord moet minimaal uit 6 karakters bestaan; kleine letters, hoofdletters, cijfers of speciale tekens. Minimaal 3 type karakters moeten aanwezig zijn.
- Extra 6-cijferige code benodigd bij inloggen. Deze wordt via e-mail verstuurd.
- De link in de herinneringsmail die leidt naar de vragenlijst bevat een code die is gebaseerd op de huidige tijd en random waarden. Deze code is niet voorspelbaar.
- Toepassen Content Security Header in de webapplicatie.
- Testen tijdens de applicatieontwikkeling (bron: Coneno):
    - Unit tests for the simple methods
    - Integration tests for testing the database and application layer integration, and
    - Testing the API endpoints from client side
    - Code-review bij code-merge.
- Onafhankelijke PEN-test voor life-gang.
- Onafhankelijke code-review voor life-gang.
- Performance-testen van RIVM Infectieradar op de RIVM-infrastructuur.

Toelichting op logging

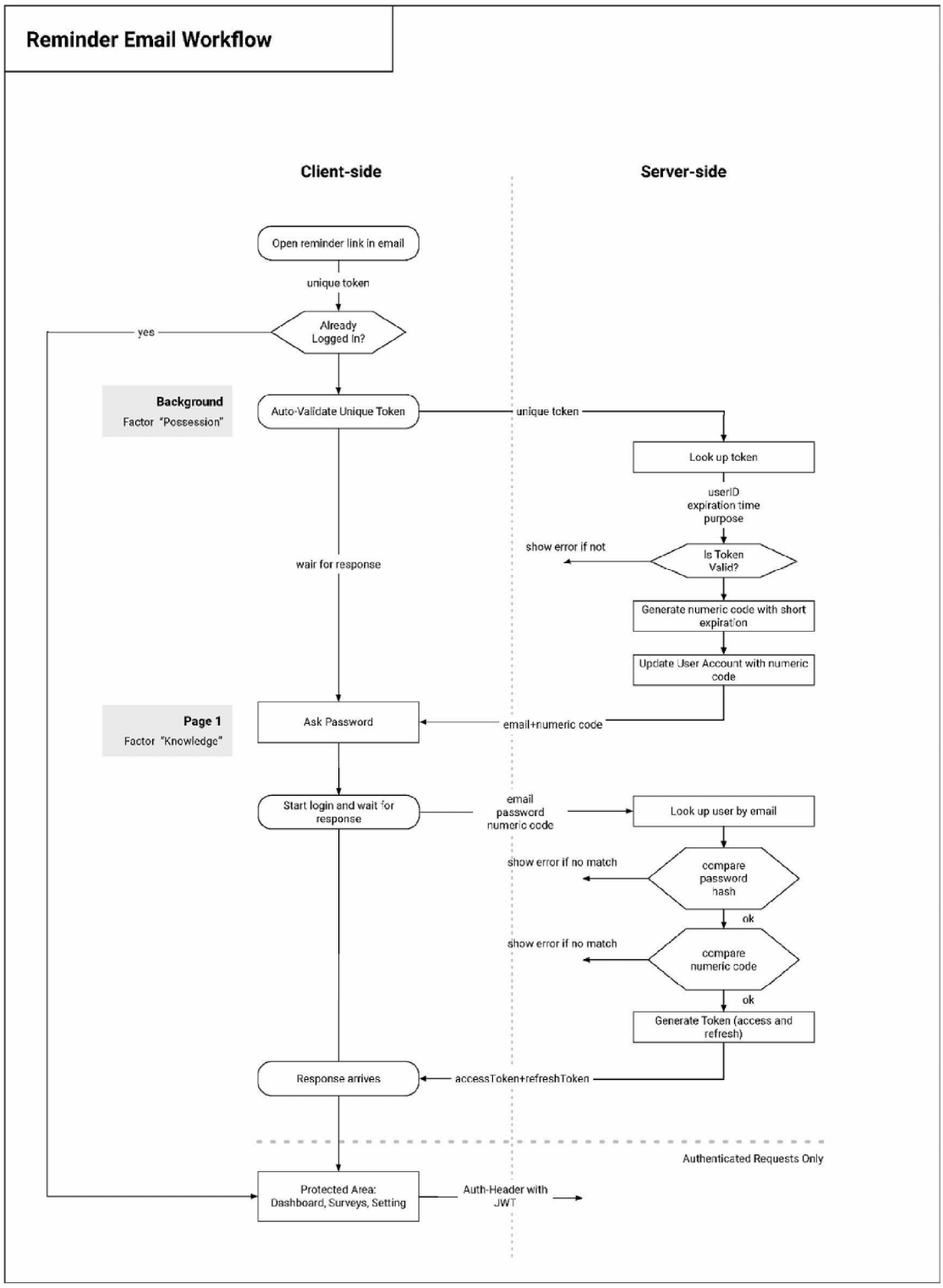
Logging van gebeurtenissen is vereist vanuit AVG:
- aanmaken en verwijderen van accounts;
- aanmaken en wijzigen van profielen;
- wijzigen van wachtwoord;
- wijzigen van e-mailadres;
- het wijzigen van de rol van een account.

Referentie: 5.1.2h 5.1.2h 5.1.2h

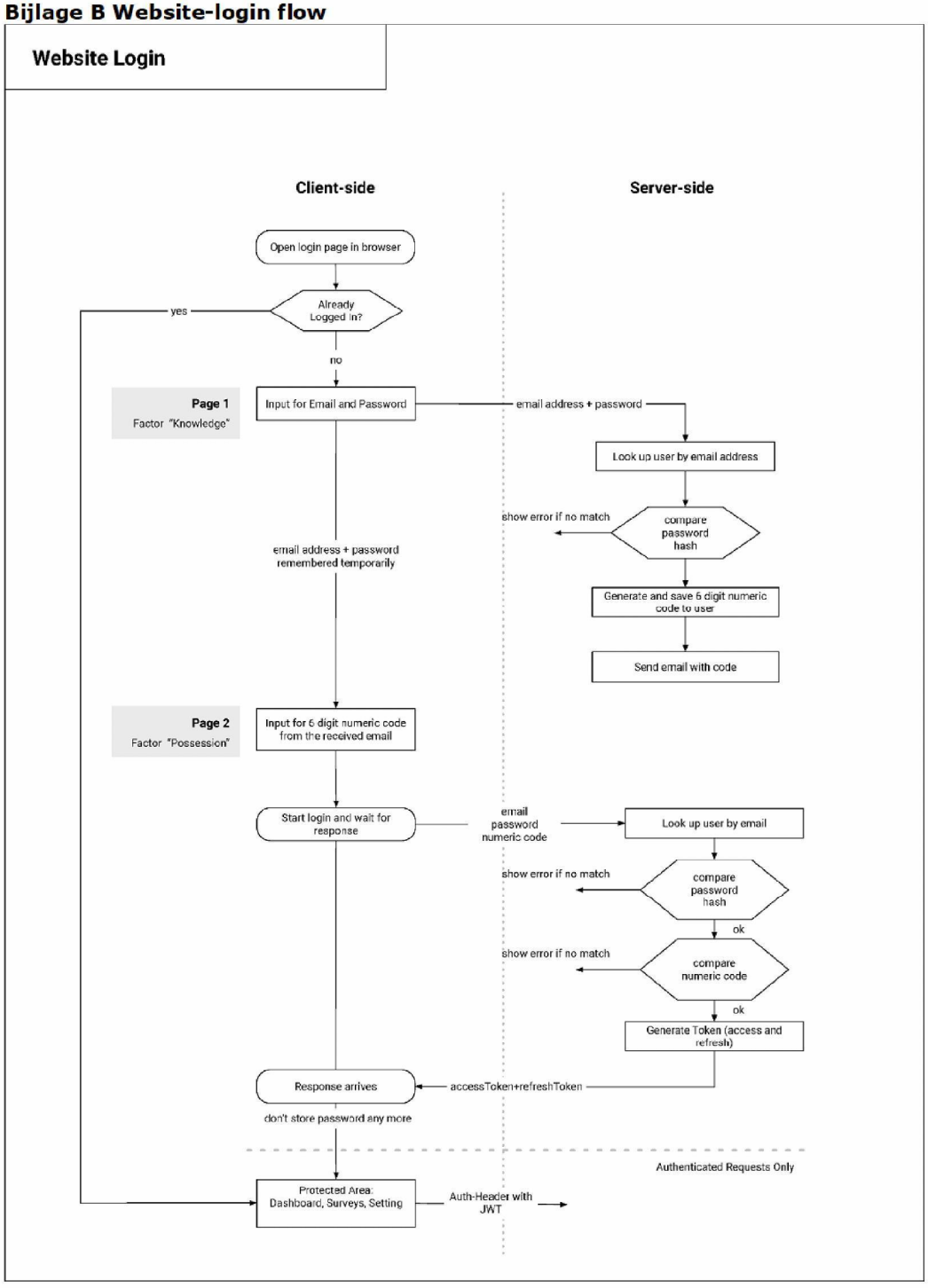
De logregel moet het volgende bevatten:
- account van de ingelogde gebruiker
- datum/tijd
- id van het betrokken object
- actie (opvragen/wijzigen/verwijderen)
- resultaat van de actie: is opvragen/wijzigen/verwijderen geslaagd?
- Waar mogelijk de identiteit van het werkstation of de locatie

Logging vindt plaats in een database-tabel. Daarmee is de doorzoekbaarheid gegarandeerd. De logregels moeten na een instelbare bewaartermijn verwijderd te worden.

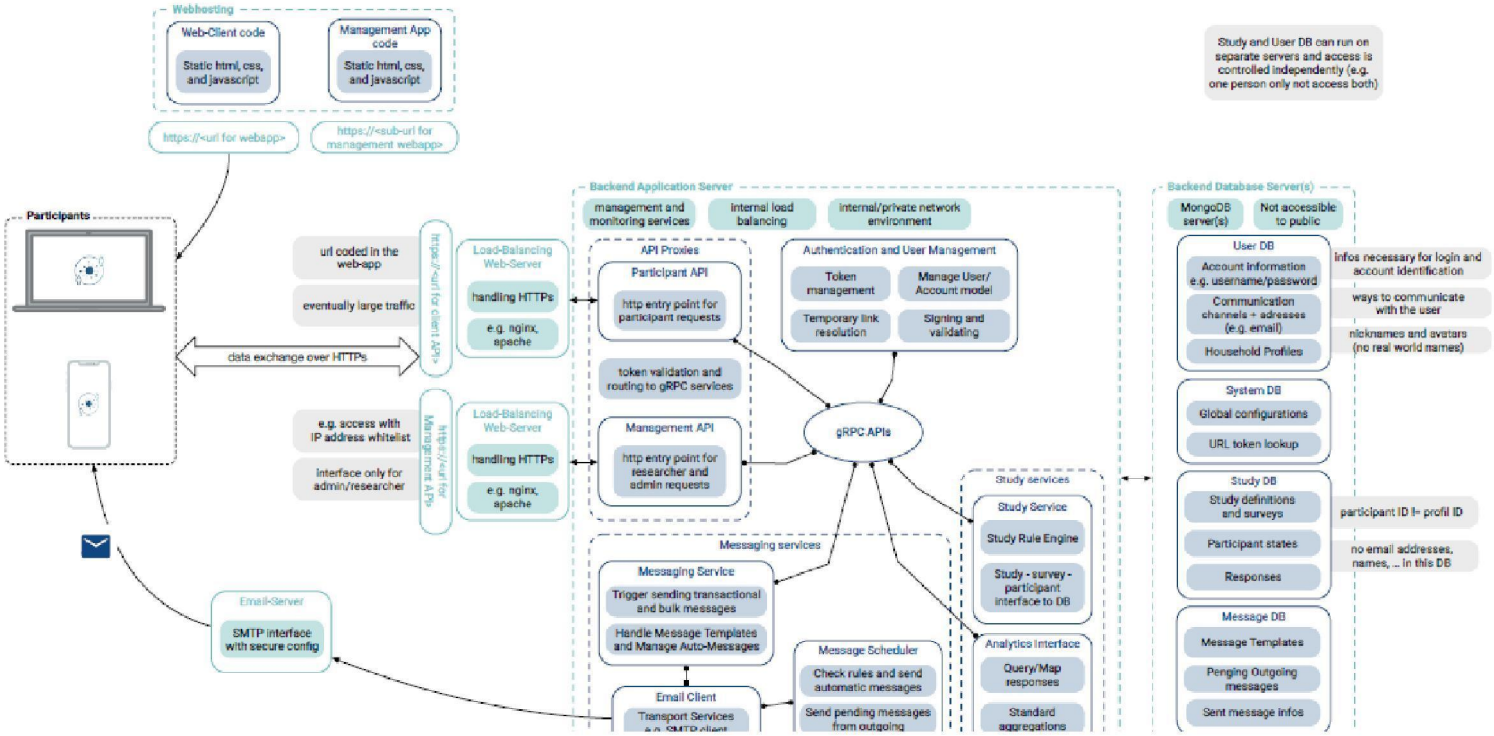
Bijlage A E-mailflow uitnodigingsmail vragenlijst



Bijlage B Website-login flow



Bijlage C Systeemoverzicht


Webhosting
Web-Client code
Static html, css, and javascript
Management App code
Static html, css, and javascript
https://<url for webapp>
https://<sub-url for management webapp>
Study and User DB can run on separate servers and access is controlled independently (e.g. one person only not access both)
Participants
url coded in the web-app
eventually large traffic
data exchange over HTTPs
https://<url for client APIs>
Load-Balancing Web-Server
handling HTTPs
e.g. nginx, apache
e.g. access with IP address whitelist
Interface only for admin/researcher
https://<url for Management APIs>
Load-Balancing Web-Server
handling HTTPs
e.g. nginx, apache
Backend Application Server
management and monitoring services
internal load balancing
internal/private network environment
API Proxies
Participant API
http entry point for participant requests
token validation and routing to gRPC services
Management API
http entry point for researcher and admin requests
Authentication and User Management
Token management
Manage User/ Account model
Temporary link resolution
Signing and validating
gRPC APIs
Study services
Study Service
Study Rule Engine
Study - survey - participant interface to DB
Analytics Interface
Query/Map responses
Standard aggregations
Messaging services
Messaging Service
Trigger sending transactional and bulk messages
Handle Message Templates and Manage Auto-Messages
Message Scheduler
Check rules and send automatic messages
Send pending messages from outgoing
Email Client
Transport Services
Email-Server
SMTP interface with secure config
Backend Database Server(s)
MongoDB server(s)
Not accessible to public
User DB
Account information e.g. username/password
Communication channels + adresses (e.g. email)
Household Profiles
Infos necessary for login and account identification
ways to communicate with the user
nicknames and avatars (no real world names)
System DB
Global configurations
URL token lookup
Study DB
Study definitions and surveys
Participant states
Responses
participant ID != profil ID
no email addresses, names, ... in this DB
Message DB
Message Templates
Pending Outgoing messages
Sent message infos