

Automatische testbed monitoring optimalisatie voor toekomstig internet onderzoek met jFed.

**Masterproef voorgedragen tot het behalen van het diploma van
Master in de industriële wetenschappen: informatica**

Andreas DE LILLE

*Promotoren: Geert VAN HOOGENBEMT
Piet DEMEESTER*

*Begeleiders: Wim VANDENBERGHE
Brecht VERMEULEN
Wim VAN DE MEERSSCHE
Thijs WALCARIUS*

Abstract

Door de huidige verschuiving naar cloudgebaseerde technologieën zal het belang van netwerk-protocollen en beschikbaarheid van netwerken alleen maar toenemen. Om deze verschuiving vlot te laten verlopen is er meer en meer onderzoek nodig naar netwerktechnologieën. Voor dit onderzoek wordt er dan ook veelvuldig gebruik gemaakt van testbeds. Testbeds worden aangestuurd via jFed (java Federation) en worden gebruikt om netwerken te simuleren. Fed4FIRE (federation for future internet research experimentation). De situatie heeft enkele nadelen. Het is voor een onderzoeker soms erg moeilijk om te bepalen of een bepaald gedrag in hun experiment te wijten is aan de eigen ontwikkelingen, of aan het testbed zelf.

Deze masterproef zal trachten dit probleem te verhelpen door de testbed monitoring te automatiseren. Via een webservice zal deze informatie dan aangeboden worden. Deze webservice zal weergeven hoe betrouwbaar een testbed is. Hiervoor komen verschillende ontwikkelingstools aan bod. Uiteindelijk zal deze service ervoor zorgen dat een onderzoeker via zijn primaire gebruikersinterface op de hoogte gebracht wordt van eventuele problemen.

Abstract

Due to the new cloud-based technologies, network protocols and network reachability are now more important than ever. To make this change quick and clean, we need more and more network research. This research heavily relies on testbeds to simulate networks. These testbeds are controlled via software tools. In this thesis we will look at jFed, a java tool developed in the European FED4FIRE project. Unfortunately, the current situation has a major downside in that it is very hard for researchers to determine if a certain behavior is caused by the testconfiguration or by the testbed.

This thesis will try to solve the aforementioned problem by automating the testbed monitoring. A webservice will then share this information and show how reliable a testbed is. After having determined the reliability of the testbed, the webservice will notify the researcher of any problems through the primary user interface.

Lijst van afkortingen

AM	Aggregate Manager
API	application programming interface
bash	bourne again shell
Fed4FIRE	federation for future internet research experimentation
GUI	graphical user interface
html	HyperText Markup Language
jFed	java Federation
json	javascript object notation
OS	operating system
php	hypertext preprocessor
sql	structured query language
SSH	secure shell
SSL	secure sockets layer
XML	extensible markup language

Lijst van figuren

1.1	Opbouw van jFed	2
2.1	Internationale testbed monitoring	7
2.2	Lokale testbed monitoring	7
2.3	Geen webservice in de huidige versie	8
2.4	De huidige scenario/stitching test monitoring	8

Lijst van tabellen

Inhoudsopgave

Lijst van figuren	v
Lijst van tabellen	vi
1 Inleiding	1
1.1 Bestaande situatie	1
1.2 Probleemstelling	3
2 Analyse	4
2.1 Bestaande uitwerking	4
2.1.1 Databanken	4
2.1.2 Webpagina's	7
2.2 Wat kan anders	9
2.2.1 Samenvoegen databases	9
2.2.2 Structuur database	9
2.2.3 Webservice uitbouwen	10
2.3 Besluit	10
3 Ontwerp service	11
3.1 Api Calls	11
3.2 Calls - Results	11
3.2.1 Functies	11
3.2.2 Parameters	12
3.2.3 Voorbeelden	13
3.3 Calls - configuration	16
3.3.1 Functies	16
3.3.2 Parameters	16
3.3.3 Voorbeelden	16
4 Ontwerp Databank	20
4.1 Tabellen	20
4.2 Testbeds	20
4.3 TestDefinitions	21
4.4 ParameterDefinitions	21

4.5	ReturnDefinitions	21
4.6	TestInstances	21
4.7	ParameterInstances	22
4.8	Results	22
4.9	SubResults	22
5	Keuzes	23
5.1	Formaat Returnwaarden	23
5.2	Structuur returnwaarden	23
5.3	Databank technologie	24
5.4	webservice technologie	25
	Bibliografie	26

Hoofdstuk 1

Inleiding

1.1 Bestaande situatie

iMinds is een onafhankelijk onderzoekscentrum dat opgericht werd door de Vlaamse overheid. Het is voornamelijk bezig met onderzoek omtrent ICT-innovatie. Bij dit onderzoek wordt veelvuldig gebruik gemaakt van testbeds. Een testbed is een verzameling van nodes waarmee netwerkopstellingen kunnen gesimuleerd worden.

Een goed voorbeeld van een testbed is de Virtual Wall. De Virtual Wall is ontworpen voor het experimenteren met nieuwe netwerkoplossingen voor de volgende generatie van het Internet. De kracht van deze infrastructuur is dat ze de mogelijkheid biedt om op een eenvoudige wijze specifieke netwerkopstellingen op te zetten volgens de noden van het onderzoek.

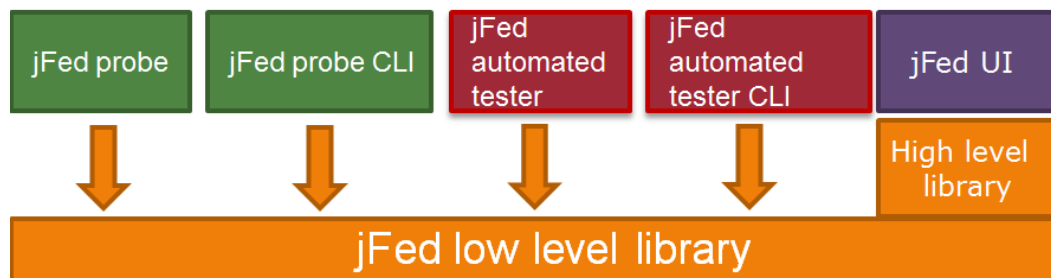
Een voorbeeld waarvoor dit testbed wordt gebruikt is dat men een server en een aantal cliënts definieert. Deze worden verbonden met een aantal tussenliggende routers. Vervolgens wordt een videostream opgestart. Op deze videostream kan men storing introduceren door pakketten te droppen. Deze storing zal ervoor zorgen dat het beeld aan de client-side hapert. Er kunnen technieken ingebouwd worden aan client-side om deze storing op te vangen. Zo kan er overgeschakeld worden naar een lagere kwaliteit indien blijkt dat de beschikbare bandbreedte onvoldoende is. Testen van degelijke technieken verloopt dan ook aan de hand van testbeds.

Onderzoekers bij iMinds hebben niet alleen toegang tot hun Virtual Wall, maar door samenwerkingsakkoorden met diverse onderzoekscentra hebben ze ook wereldwijd toegang tot andere en soms erg verschillende testbeds. Deze zijn elk apart ontwikkeld. Hierdoor maakt elk testbed gebruik van een eigen interface. Deze interface wordt aangeboden door middel van een federation API (application programming interface) . Een federation API is een interface die alle mogelijk functionaliteiten van een specifiek testbed aanbiedt. Hij staat in voor de communicatie tussen een programma en het testbed. Onderzoekers die met meerdere testbeds werken, moeten dan ook de werking van elk testbed apart bestuderen alvorens ermee

aan de slag te kunnen. Ook rechtstreekse verbindingen leggen tussen verschillende testbeds wordt hierdoor moeilijker.

Daarom werd in het Europese onderzoeksproject Fed4FIRE beslist om een aantal federation API's vast te leggen die gelijk moeten zijn op elk testbed. Dit maakt het mogelijk om tools te ontwikkelen die bruikbaar zijn op alle testbeds verenigd in dit project. Hiervoor werd de tool jFed ontwikkeld door iMinds. jFed biedt een wrapper aan rond de specifieke federation API's die in dat project worden toegepast. Daarenboven voorziet jFed in verschillende applicaties om enerzijds te kunnen testen of testbeds deze API's correct ondersteunen, en anderzijds om als onderzoeker op een eenvoudige wijze een experiment te kunnen opzetten.

De kernmodule waar alle andere modules gebruik van maken is de "jFed library". De voornaamste functie van deze library is om de communicatie met de testbeds te vereenvoudigen. De hiervoor bruikbare testbeds implementeren enkele gestandaardiseerde API's, die de benodigde functionaliteiten aanbieden. De jFed library implementeert de client kant van deze API's (application programming interface), en voorziet Java interfaces voor de API-calls. De gebruiker zal hierdoor niet geconfronteerd worden met deze details. Zo zal de library o.a. de achterliggend SSL-connectie (secure sockets layer) en de bijhorende certificaten en private keys beheren. Verder zullen ook de nodige omzettingen gebeuren. Daarnaast bevat deze library vele hulpklassen en methoden om de API's aan te spreken.



Figuur 1.1: Opbouw van jFed

Een goed voorbeeld van een API die door jFed wordt geïmplementeerd is de AM (Aggregate Manager) API, versie 2 en 3. Een voorbeeld van een call, is de "CreateSliver" call in de AMv2 API. De AMv2 API specificeert dat alle calls gedaan worden over een SSL verbinding met client-side certificaten. Verder specificeert de API de exacte argumenten die vereist zijn en het formaat van het resultaat. De jFed library biedt een Java methode "createSliver" in de klasse "AggregateManager2" aan. De argumenten worden opgegeven in een Java formaat. Zo zal bijvoorbeeld een datum-argument als een Java Date-object doorgegeven worden. De jFed library zal de datum naar het RFC3339 formaat omzetten voor het naar de server verstuurd wordt. Ook het resultaat van de call wordt verwerkt door de jFed library terug omgezet.

Steunend op deze library zijn enkele tools zoals jFed probe en jFed compliance tester gebouwd. Deze hebben als doel het valideren of testbeds de federation API's correct ondersteunen. Tenslotte is er de jFed graphical user interface (GUI) tool. Deze tool maakt het mogelijk om als eindgebruiker een experiment op een intuïtieve manier op te zetten. De jFed GUI heeft momenteel echter een relatief complexe interface die gebruiksvriendelijker gemaakt kan worden.

1.2 Probleemstelling

Er is echter nog ruimte voor verbetering in jFed. Zo worden er op geregelde tijdstippen verschillende testen automatisch uitgevoerd op testbeds. Deze testen geven o.a. weer of testbeds online zijn en of het aanmelden via een SSH-verbinding (secure shell) gelukt is. Deze resultaten worden momenteel nog redelijk ad hoc bijgehouden. Men maakt geen gebruik van een databank, maar van bourne again shell (bash) scripts om de resultaten te archiveren. Resultaten op deze manier bijhouden, is uiteraard inefficiënt. Bovendien staat deze informatie momenteel enkel op een enkele specifieke website vermeld.

Een onderzoeker wordt dus niet via zijn primaire gebruikersinterface op de hoogte gebracht van eventuele storingen met de testbeds opgenomen in zijn experiment. Dit maakt het voor onderzoekers soms moeilijk om te identificeren of een bepaald gedrag in hun experiment te wijten is aan de eigen ontwikkelingen, of aan de testbeds zelf. Hierdoor kan veel tijd onnodig verloren gaan tijdens de uitvoering en de analyse van het experiment. Een betere oplossing zou zijn dat bij het selecteren van een testbed de betrouwbaarheid weergegeven wordt. Zo kan een betrouwbaar testbed gekozen worden, om dergelijke situaties zoveel mogelijk te vermijden.

Niet enkel de betrouwbaarheid is van belang, ook het 'comfort' is van belang. Een testbed met een snellere verbinding geniet de voorkeur op een testbed met een trage verbinding. Naast de betrouwbaarheid en comfort moet ook gekeken worden naar de vereisten. Indien een betrouwbaar testbed niet beschikt over voldoende resources kan het niet gebruikt worden voor een experiment.

Door rekening houden met de betrouwbaarheid, het comfort en de vereisten kan jFed de onderzoeker melden dat een gekozen testbed niet betrouwbaar is. Hiervoor is een monitoring service nodig die bijhoudt hoe vaak en hoelang een testbed offline is.

Hoofdstuk 2

Analyse

2.1 Bestaande uitwerking

De ontwikkeling van de huidige situatie is door de sneller ontwikkeling, minder gestructureerd verlopen. Hierdoor bestaat de software uit een basis versie gevolgd door een aantal ‘quick and dirty’ toevoegingen. Er moet opgemerkt worden dat de huidige situatie werkt, maar het kan beter. Het gemist van structuur in opbouw zal op lange termijn leiden tot code die zeer moeilijk aan te passen is.

De oplossing van dit probleem en tevens het onderwerp van mijn masterproef is een webservice die de monitoring automatiseert. Hierbij komen een aantal vragen naar boven. Welke informatie moet bijgehouden worden? Wat bepaald de betrouwbaarheid van een testbed? Hoe nauwkeurig moet deze informatie bijgehouden worden?

2.1.1 Databanken

De huidige situatie voorziet niet in een centrale webservice (Partners, 2014). Wat er wel bestaat is een verzameling websites die rechtstreeks verbinding maken met een of meerdere databanken. Er zijn 3 databanken voorzien:

1. flsmonitoring
2. flsmonitoring-international
3. scenarios

Flsmonitoring en flsmonitoring-internation databank

In de eerste en de tweede databank bestaan uit een tabel waarin de laatste resultaten van elke test bijgehouden worden. Het verschil tussen deze 2 databanken komt overeen met de toegevoegde categorie waarin het testbed zich bevindt. De eerste databank bevat de lokale testbeds, de tweede bevat de internationale testbeds. Deze tabellen bevatten volgende kolommen:

- testbedid
- testbedname
- testbedurl
- pinglatency
- getversionstatus
- aggregatetestbedstate
- last-check

De eerste 3 parameters zijn duidelijk. ‘Pinglatency’ houdt de waarde van de pingtest bij. De kolommen ‘getversionsStatus’ en ‘aggregatetestbedstate’ worden gebruikt om de uitkomst van de getVersion test bij te houden. Deze test bevat o.a. het versie nummer van de aggregate manager. Doordat er geen ssl authenticatie nodig is voor deze test, wordt hij vaak gebruikt om de status van een server op te vragen. De kolom ‘last-check’ bevat een timestamp om bij te houden wanneer de lijn laatste werd aangepast.

Scenario databank

De laatste databank bestaat uit 3 tabellen. Het doel ervan is het bijhouden van informatie over de scenariotesten. Scenariotesten of stitchingtesten zijn complexe testen die uit meerdere subtesten bestaan. Eenvoudig gezegd zal een stitching test de verbinding tussen verschillende testbeds testen. Hiervoor worden op elk testbed meerdere resources aangevraagd. Deze zullen dan trachten naar elkaar te pinggen. Indien een testbed offline is wordt de volledige test afgebroken. Hieronder staan de opeenvolgende stappen die een stitching of scenariotest doorloopt.

1. setUp
2. getUserCredential
3. generateRspec
4. createSlice

5. initStitching
6. callSCS
7. callCreateSlivers
8. waitForAllReady
9. loginAndPing
10. callDeletes

De inhoud van elke subtest wordt hier buiten beschouwing gelaten. Wat wel opgemerkt kan worden is dat we de tests kunnen opdelen in 3 groepen. Zo zijn testen 1-6 voorbereidende testen. Ze dienen om de configuratie voor testen 7-9 klaar te zetten. Test 10 is de cleanup die de opgebouwde configuratie van stappen 1-6 terug ongedaan maakt. Elke subtest heeft een resultaat. Een stitching test zou dus minstens 10 resultaten hebben. In de huidige versie zijn er slechts 3 statussen gedefiniëerd. De stitching test is volledig gelukt, dit komt overeen met 10 geslaagde subtesten. De status is gedeeltelijk gelukt, dit komt overeen met de voorbereiding die wel gelukt is, maar de stappen 7-9 zijn niet of slechts gedeeltelijk gelukt. De laatste status geeft aan dat alle subtesten mislukt zijn.

De database is opgebouwd uit 3 tabellen.

- test-results
- test-context
- testbeds

De eerste resultaat houdt informatie bij over de testresultaten. De tweede tabel houdt de context van de test bij. De laatste tabel houdt de testbeds bij. De concrete invulling van de tabellen is van minder belang en wordt hier niet vermeld.

2.1.2 Webpagina's

In de huidige versie zijn een aantal webinterfaces voorzien (Partners, 2014). Er is een webinterface die enkel de status van de lokale testbeds weergeeft. Daarnaast er een webinterface die de internationale testbeds weergeeft. Deze 2 webinterfaces geven de naam van het testbeds met de ping, een veld dat aangeeft of de getVersion call gelukt is en het aantal beschikbare resources.

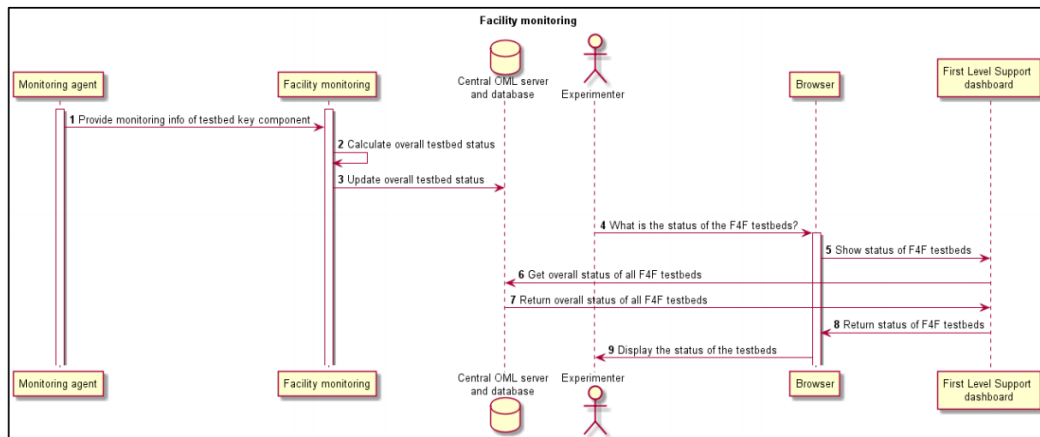
International Federation Monitoring			
Testbed Name	Ping latency (ms)	GetVersion Status	Free Resources
BBN emulab	86.75	ok	0
FIU PGV1	Unreachable	not ok	not ok
InstaGeni BBN	86.74	ok	0
InstaGeni Clemson	124.02	ok	0
InstaGeni GATech	135.69	ok	0
InstaGeni Illinois	156.33	ok	0
InstaGeni Kettering	Unreachable	not ok	not ok
InstaGeni MAX	100.89	ok	0
InstaGeni Missouri	160.29	ok	0
InstaGeni NorthWestern	138.77	ok	0
InstaGeni Nysernet	99.43	ok	0
InstaGeni Princeton	85.65	ok	0
InstaGeni SOX Atlanta	135.5	ok	0
InstaGeni Uky	133.62	ok	0
InstaGeni Utah DDC	171.6	ok	0
ION internet2 AM	130.94	ok	0
KetteringU emulab	Unreachable	not ok	not ok

Figuur 2.1: Internationale testbed monitoring

Fed4FIRE First Level Support Monitoring					
Testbed Name	Ping latency (ms)	GetVersion Status	Free Resources	Internal testbed monitoring status	Last check internal status
BonFIRE	31.17	N/A	N/A	ok	2014-03-25 11:16:09+01
Fuseco	16.83	ok	1	ok	2014-03-25 11:16:03+01
Koren	284.47	N/A	N/A	N/A	N/A
NETMODE	67.33	ok	20	ok	2014-03-25 11:13:22+01
NITOS Broker	73.03	ok	38	ok	2014-03-25 11:15:01+01
NITOS SFAWrap	31.29	ok	65	N/A	N/A
Norbit	N/A	N/A	N/A	ok	2014-03-25 11:10:29+01
Ofelia (Bristol island)	11.89	N/A	N/A	ok	2014-03-25 11:10:02+01
Ofelia (I2CAT island)	N/A	N/A	N/A	ok	2014-03-25 11:10:02+01
Planetlab Europe	32.11	ok	285	ok	2014-03-25 11:15:02+01
SmartSantander	53.82	ok	0	ok	2014-03-25 11:10:01+01
Virtual Wall	0.21	ok	23	ok	2014-03-25 11:14:39+01
w-ILab.t 2	20.33	ok	68	ok	2014-03-25 11:14:58+01

Figuur 2.2: Lokale testbed monitoring

Zowel de webinterfaces als het aanpassen van monitoringinformatie verloopt rechtstreeks via de databank. Er komt geen webservice bij kijken.



Figuur 2.3: Geen webservice in de huidige versie

De laatste webinterface geeft de resultaten van de scenariotesten weer. Er is ook de mogelijkheid om de log files van een test te bekijken. Tevens is het ook mogelijk om de geschiedenis van een scenariotest te bekijken.

jFed Automated Scenario Tests								
Options: Show code tag								
Category	Test Name	Last Test Start Time (CET)	Last Test Duration	Last Partial Success	Last Full Success	Time since last Failure	Last Log	History
login	Fuseco	2014-03-25 09:17:47	14 seconds	FAILURE	FAILURE		log	history
login	InstaGeni BBN	2014-03-25 09:18:01	3 minutes and 6 seconds	SUCCESS	SUCCESS	9 days and 13 hours	log	history
login	InstaGeni Clemson	2014-03-25 09:21:08	45 seconds	WARN	WARN		log	history
login	InstaGeni GATech	2014-03-25 09:21:54	2 minutes and 24 seconds	SUCCESS	SUCCESS	27 days and 14 hours	log	history
login	InstaGeni Illinois	2014-03-25 09:24:20	3 minutes and 27 seconds	SUCCESS	SUCCESS	5 months and 11 days	log	history
login	InstaGeni Kettering	2014-03-25 09:27:50	10 seconds	FAILURE	FAILURE		log	history
login	InstaGeni MAX	2014-03-25 09:28:00	2 minutes and 55 seconds	SUCCESS	SUCCESS	2 months and 9 days	log	history

Figuur 2.4: De huidige scenario/stitching test monitoring

2.2 Wat kan anders

2.2.1 Samenvoegen databases

Een eerste situatie die beter kan is het bijhouden van de laatste resultaten. Deze resultaten zitten in een databank, die voor elke testbed een lijn bevat. Nieuwe waarde overschrijven die lijn. Meteen kan opgemerkt worden dat het niet mogelijk is om statistieken over een lange termijn weer te geven. Het aanpassen van de monitoring waarden gebeurt door shell scripts. Deze worden periodiek uitgevoerd en lezen een configuratiefile in. Indien het testbed waarop de test uitgevoerd moet worden al in de databank zit is het id vermeld in de configuratiefile. Indien er geen id staat zal het script zelf een nieuwe lijn aanmaken en vervolgens de id wegschrijven naar de file. Eenmaal uitgevoerd, geeft de test een resultaat terug in de vorm van een xml-file (Extensible Markup Language) die door geïnstalleerde commando's geparset wordt. Vervolgens wordt de data weggeschreven naar de databank.

Door nieuwe resultaten toe te voegen op een nieuwe lijn kunnen we wel statistieken over langere termijn bijhouden. Ook de configuratiefiles kunnen opnemen in de databank. Toevoegen van een nieuwe test is dan eenmalig een entry toevoegen aan de database. Deze entry bevat dan het commando en de benodigde commando's.

Een tweede punt is de opdeling flsmonitoring en flsmonitoring-international. Deze 2 databanken zijn eigenlijk gelijk en kunnen gemakkelijk ondergebracht worden in een database. Dit kan door een extra punt toe te voegen dat weergeeft in welke categorie het testbed zich bevindt. Een andere mogelijkheid is de internationale testbeds hardcoderen op de monitoring site.

Door de databanken scenario's en flsmonitoring in een databank onder te brengen, kunnen we het beheer vereenvoudigen. Beide gegevens zijn eenmaal resultaten van testen. Deze meer generische aanpak zal er toe leiden dat testen eenvoudiger gedefiniëerd kunnen worden.

2.2.2 Structuur database

De huidige opbouw van de databanken is niet flexibel genoeg. Als we de databanken zouden samenvoegen is er een complexere structuur nodig die meer flexibiliteit toelaat. Het moet mogelijk zijn dat een tests meerdere argumenten meekrijgt. Dit aantal is verschillend per type tests, werken met extra kolommen is dus niet aangeraden. Ook geven sommige testen meerdere resultaten terug. Dit aantal is ook variabel en afhankelijk van het type test dat gebruikt wordt. De structuur van de databank wordt later in uitgewerkt.

2.2.3 Webservice uitbouwen

Doordat de webinterfaces rechtstreeks contact maken met de databank is er minder overhead. Als de databank echter ook vanuit jFed bereikbaar moet zijn, is het beter om een webservice tussen te voegen. Dit vermijdt duplicatie van code. Deze service zal complexere zaken zoals filteren van resultaten voorzien. Daardoor kunnen zowel jFed als de webinterfaces en eventuele toekomstige toepassingen met een eenvoudige call de informatie ophalen en moet niet elke mogelijke applicatie terug zelf de filtering voorzien.

Eenmaal de webservice uitgebouwd is kunnen er sites gemaakt worden die deze webservice contacteren om informatie op te halen. De huidige webinterfaces moeten dus ook vervangen worden door nieuwe versies die wel gebruik maken van de webservice. Er kan ook integratie in jFed voorzien worden. Dit kan als een onderdeel van de masterproef, of als een latere uitbreiding voor jFed.

2.3 Besluit

Om aan de verwachten te voldoen zijn er 2 grote veranderingen nodig. Er zullen twee grote veranderingen gebeuren. De eerste eerst is het samenvoegen van de databanken. Om het beheer van de data te vereenvoudigen zou 1 grote databank het werk van de 3 kleinere databanken over nemen. Deze zou zo opgebouwd zijn dat we testen met een variabel aantal argumenten en resultaten kunnen toevoegen, zonder iets te wijzigen aan de structuur van de databank. Zowel de configuratie van de testen als de resultaten zouden in deze databank opgenomen zijn.

Hoofdstuk 3

Ontwerp service

In dit hoofdstuk wordt de interface van de webservice besproken. Daarna worden er enkele voorbeelden van calls uitgewerkt.

3.1 Api Calls

De calls kunnen opgedeeld worden in 2 grote groepen.

- Results: Het opvragen van resultaten van een test.
- Configuration: Het opvragen, aanpassen en aanmaken van testen en de bijhorende configuratie.

3.2 Calls - Results

Voor het opvragen van resultaten wordt gebruik gemaakt van een generische functie.

3.2.1 Functies

Er zijn een aantal functies voorzien voor het opvragen van de resultaten.

- last: Deze functie geeft het laatste/ de laatste resultaat/resultaten terug (per test,testbed).
- list: Deze functie geeft een lijst resultaten die voldoen aan de opgegeven parameters. Zo kan deze lijst bijvoorbeeld beperkt worden tot een resultaten die overeenkomen met een bepaalde status of vallen tussen bepaalde data.
- detail: Deze functie geeft een detailweergave terug van een test. Deze detail weergave kan bijvoorbeeld ook logfiles bevatten om debugging te vergemakkelijken.

3.2.2 Parameters

Hier worden de parameters besproken. Deze parameters zullen het resultaat van de functies hierboven filteren.

- testbed: Deze parameter duidt een of meerdere testbeds aan.
- testtype: Deze parameter duidt het type van de test aan.
- status: Deze parameter duidt de status aan. Let op dit is per subtest.
- count: Deze parameter geeft het aantal laatste resultaten aan dat teruggegeven moet worden. Ook wanneer er enkel een till opgegeven is, worden de laatste resultaten voor de opgegeven tijd teruggegeven. Wanneer er enkel een from parameter opgegeven is, dan worden de eerste x resultaten startend vanaf de from terug gegeven.
- from: Geeft aan vanaf welke timestamp er gezocht moet worden.
- till: Geeft aan tot welke timestamp de resultaten gezocht moeten worden.
- resultid: Geeft een of meerdere id's van een testresultaat.
- testname: De naam van de test. Voornamelijk nodig bij stitching tests omdat er daar meerdere testbeds in betrokken zijn.
- instanceId: Filteren op de testinstantie, gelijkaardig aan de testname.
- format: Deze parameter dient niet om de resultaten te filteren, maar om het formaat van het resultaat te bepalen. Standaard is dit json, later kan eventueel xml en/of csv toegevoegd worden.

Indien bij tests of bij testbeds de waarde all wordt gegeven, worden respectievelijk alle tests of alle testbeds teruggegeven. Ook opgeven van lijsten is mogelijk. Zo kunnen de resultaten van pingtest voor 2 verschillende testbeds opvraagd worden met een call, zie ook voorbeelden.

3.2.3 Voorbeelden

Hieronder worden een aantal voorbeelden van calls en bijhorende antwoorden geven. Voorlopig zijn de voorbeelden beperkt tot het opvragen van data met behulp van HTTP GET requests.

last?testbed=urn-testbed0&testtype=ping&count=2

Deze call geeft voor testbed0 de laatste 2 ping resultaten terug.

```
{
  "177": {
    "testinstanceid": "1",
    "testtype": "ping",
    "testname": "ping voor testbed0",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\1\\1386",
    "timestamp": "2014-03-18 19:29:12.876569",
    "testbeds": [
      "urn-testbed0"
    ],
    "results": {
      "pingValue": "166"
    }
  },
  "155": {
    "testinstanceid": "1",
    "testtype": "ping",
    "testname": "ping voor testbed0",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\1\\1615",
    "timestamp": "2014-03-18 19:29:09.138309",
    "testbeds": [
      "urn-testbed0"
    ],
    "results": {
      "pingValue": "223"
    }
  }
}
```

last?testbed=ALL&testtype=ping

Deze call geeft voor elk testbed de laatste ping resultaten terug.

```
{  "177": {
    "testinstanceid": "1",
    "testtype": "ping",
    "testname": "ping voor testbed0",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\1\\1386",
    "timestamp": "2014-03-18 19:29:12.876569",
    "testbeds": [ "urn-testbed0" ],
    "results": { "pingValue": "166" }
  }, "178": {
    "testinstanceid": "2",
    "testtype": "ping",
    "testname": "ping voor testbed1",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\2\\9739",
    "timestamp": "2014-03-18 19:29:12.889286",
    "testbeds": [ "urn-testbed1" ],
    "results": { "pingValue": "112" }
  }, "187": {
    "testinstanceid": "11",
    "testtype": "ping",
    "testname": "ping voor testbed10",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\11\\4280",
    "timestamp": "2014-03-18 19:29:13.039074",
    "testbeds": [ "urn-testbed10" ],
    "results": { "pingValue": "106" }
  }, "188": {
    "testinstanceid": "12",
    "testtype": "ping",
    "testname": "ping voor testbed11",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\12\\2792",
    "timestamp": "2014-03-18 19:29:13.055714",
    "testbeds": [ "urn-testbed11" ],
    "results": { "pingValue": "156" }
  }
  ...
}
```

last?testname=stitching2

Deze functie geeft een detail resultaat terug van de stitchingtest met naam 'stitching2'.

Voorbeeld invoegen

list?from=2014-03-18T19:29:00&till=2014-03-18T19:29:10&testtype=ping&testbed=urn-testbed1

Deze call geeft een lijst van de pingtests tussen 2014-03-18 19:29:00 en 2014-03-18 19u29:10 op testbed1 terug.

```
{
  "156": {
    "testinstanceid": "2",
    "testtype": "ping",
    "testname": "ping voor testbed1",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\2\\2306",
    "timestamp": "2014-03-18 19:29:09.152291",
    "testbeds": [
      "urn-testbed1"
    ],
    "results": {
      "pingValue": "66"
    }
  },
  "134": {
    "testinstanceid": "2",
    "testtype": "ping",
    "testname": "ping voor testbed1",
    "log": "http:\\\\f4f-mon-dev.intec.ugent.be\\logs\\2\\1991",
    "timestamp": "2014-03-18 19:29:05.413795",
    "testbeds": [
      "urn-testbed1"
    ],
    "results": {
      "pingValue": "24"
    }
  },
  ...
}
```

3.3 Calls - configuration

Deze calls hebben het doel om de configuratie van een test op te vragen.

3.3.1 Functies

Er zijn (uiteraard) ook een aantal functies voor het opvragen van de testconfiguratie.

- **TestDefinition** : Geeft de definitie van een testweer. Deze bevat informatie over de naam, het commando. Verder duidt deze definitie ook aan welke parameters er nodig zijn en welke waarden er terug gegeven worden. Deze waarden zijn echter niet ingevuld. Het is de bedoel om hier bijvoorbeeld te definiëren dat een pingtest bijvoorbeeld bestaat uit een ping commando en dat deze een time-out en een testbed moet meekrijgen. Samengevat kan gesteld worden dat dit de beschrijving van een type test is.
- **Testbed**: Geeft een of meerdere testbeds aan waarop de testinstantie gedefiniëerd moet worden.

Deze opsplitsing is noodzakelijk om flexibiliteit aan te bieden. Een ping test wordt op meerdere testbeds uitgevoerd. Door deze opbouw moet een ping test maar eenmalig gedefiniëerd worden. Als we de uitwerking vergelijking met het model-view-controller pattern, geeft de definition de klasse, de instance kan gezien worden als een object van die klasse.

3.3.2 Parameters

- **testtype**: Deze parameter duidt aan welke testtype opgevraagd wordt.
- **testinstanceId**: Deze parameter duidt aan welke instance opgevraagd wordt. Er is geen equivalente variant voor testdefinition, omdat het testtype daar de unieke id voorstelt.

Het is ook mogelijk om als waarde ALL of een lijst van id's op te geven.

3.3.3 Voorbeelden

Hieronder worden een aantal mogelijke calls weergegeven. Voorlopig zijn deze voorbeelden beperkt tot het opvragen van data via http get requests.

testDefinition?testtype=stitch

Deze call geeft de beschrijving van een stitching test terug.

```
{  "stitch": {
    "testcommand": "stitch",
    "parameters": {
      "topology": {
        "type": "string",
        "description": "ring | line"
      }, "testbedId": {
        "type": "testbedId[]",
        "description": "multiple testbeds for ping test"
      }
    }, "return": {
      "callDeletes": {
        "type": "string", "description": "status of subtest"
      }, "loginAndPing": {
        "type": "string", "description": "status of subtest"
      }, "waitForAllReady": {
        "type": "string", "description": "status of subtest"
      }, "callCreateSlivers": {
        "type": "string", "description": "status of subtest"
      }, "callSCS": {
        "type": "string", "description": "status of subtest"
      }, "initStitching": {
        "type": "string", "description": "status of subtest"
      }, "createSlice": {
        "type": "string", "description": "status of subtest"
      }, "generateRspec": {
        "type": "string", "description": "status of subtest"
      }, "getUserCredential": {
        "type": "string", "description": "status of subtest"
      }, "setUp": {
        "type": "string", "description": "status of subtest"
      }
    }
  }
}
```

testInstance?testbed=urn-testbed1,urn-testbed2

Hierbij worden alle testinstanties van testbed1 en testbed2 teruggegeven. Dit zijn dus de testen die effectief draaien op testbed1 en/of testbed2.

```
{  "2": {
    "testname": "ping voor testbed1",
    "testtype": "ping",
    "frequency": "60",
    "parameters": [
      {"testbedId": "urn-testbed1"},
      {"timeout": "300"}
    ]
  }, "3": {
    "testname": "ping voor testbed2",
    "testtype": "ping",
    "frequency": "60",
    "parameters": [
      {"testbedId": "urn-testbed2"},
      {"timeout": "300"}
    ]
  }, "18": {
    "testname": "stiching0",
    "testtype": "stitch",
    "frequency": "3600",
    "parameters": [
      {"testbedId": "urn-testbed2"},
      {"testbedId": "urn-testbed1"},
      {"testbedId": "urn-testbed0"},
      {"topology": "ring"}
    ]
  },
  ...
}
```

testInstance?testtype=ping

Deze call geeft alle geplande tests terug.

```
{
  "1": {
    "testname": "ping voor testbed0",
    "testtype": "ping",
    "frequency": "60",
    "parameters": [
      { "testbedId": "urn-testbed0" },
      { "timeout": "300" }
    ]
  },
  "2": {
    "testname": "ping voor testbed1",
    "testtype": "ping",
    "frequency": "60",
    "parameters": [
      { "testbedId": "urn-testbed1" },
      { "timeout": "300" }
    ]
  },
  "3": {
    "testname": "ping voor testbed2",
    "testtype": "ping",
    "frequency": "60",
    "parameters": [
      { "testbedId": "urn-testbed2" },
      { "timeout": "300" }
    ]
  },
  ...
}
```

Hoofdstuk 4

Ontwerp Databank

In dit hoofdstuk wordt de interface van de databank besproken.

4.1 Tabellen

De database zal bestaan uit volgende tabellen.

- testbeds: Deze tabel zal informatie over de testbeds bijhouden.
- testDefinitions: Deze tabel zal de testbeschrijving bijhouden.
- parameterDefinitions: Deze tabel houdt de parameters bij die de test nodig heeft.
- returnDefinitions: Deze tabel houdt de returnwaardes bij die een test genereert.
- testInstances: Hier worden de instanties van de testdefinities bijgehouden. Deze tabel zorgt voor de link tussen de testdefinitie en de ingevulde parameters.
- parameterInstances: Deze tabel houdt de ingevulde waarden van de parameters bij.
- results: Deze tabel zorgt voor de link tussen de testinstantie en de subresults.
- subResults: Deze tabel houdt de verschillende resultaten van een testbij.

Merk hier ook de opsplitsing `testDefinition <=> testInstance`. `TestDefinition` is een definitie van een test, `testinstantie` is de test met de ingevulde waarden.

4.2 Testbeds

Deze tabel houdt de verschillende testbeds bij en bevat volgende kolommen :

- testbedId: Het id van het testbed. Aangezien elk testbed een uniek urn heeft zal dit waarschijnlijk gebruikt worden als testbedid. Een nadeel is wel dat een urn moeilijk te onthouden is, wat gebruik van de service zal bemoeilijken.

- name: de naam van het testbed.

4.3 TestDefinitions

Deze tabel houdt de definities van de tests bij. Hierbij wordt een test abstract omschreven als een commando met een aantal parameters en een of meerdere return waarden. De tabel bevat volgende kolommen:

- testtype : Dit is de naam van een testtype en moet tevens uniek zijn.
- testcommand : Het commando (of script) dat uitgevoerd moet worden.

4.4 ParameterDefinitions

Deze tabel houdt bij welke parameters een test nodig heeft. Hierbij wordt de naam, het type en een beschrijving van een parameter opgeslagen. De tabel bevat volgende kolommen:

- testtype: Houdt bij voor welke testtype de parameters van toepassing zijn.
- parameterName: De naam van de parameter.
- parameterType: Het type van de parameter.
- parameterDescription: De beschrijving van een parameter.

4.5 ReturnDefinitions

Deze tabel bevat de returnwaarden die bij een testtype horen. De tabel bevat volgende kolommen:

- testtype: Houdt bij voor welke testtype de returnwaarden van toepassing zijn.
- returnName: Naam van de returnwaarde.
- returnType: Type van de returnwaarde.
- returnDescription: Beschrijving van de returnwaarde.

4.6 TestInstances

Deze tabel zal later de testdefinition koppelen aan de tabel met parameters. In de eerste versie zal hij de waarden hardgecodeerd bevatten. In deze tabel wordt er bijgehouden der er om de 5 minuten een ping test uitgevoerd moet worden op een testbedX met een timeout Y. De tabel bevat volgende kolommen:

- testInstanceId: Een id om de testinstance mee aan te duiden.
- testtype: Geeft het testtype aan.
- testname: Geeft de naam van een test aan.
- Frequency : De frequentie waarmee de test uitgevoerd moet worden.

4.7 ParameterInstances

Deze tabel houdt de ingevulde waarden van de parameters bij voor een testinstantie. De tabel bevat volgende kolommen:

- testInstanceId: Geeft aan voor welke instantie de waarde geldt.
- parameterName: De naam van de parameter.
- parameterValue: De waarde van de parameter.

4.8 Results

Deze tabel vormt de link tussen de testinstantie en de subResults. De tabel bevat volgende kolommen:

- resultid : Later gebruikt om detail weergave op te vragen van een resultaat.
- testInstanceId: Geeft aan van welke testinstantie deze resultaten afkomstig zijn.
- log : De link naar de link van de log file. Deze zal beschikbaar staan op het testbed, via de url is het mogelijk om hem op te vragen.
- timestamp : Geeft aan wanneer die resultaat is bepaald.

4.9 SubResults

Deze tabel bevat de ingevulde resultaten.

- resultId: Geeft aan bij welk resultaat deze waarden horen.
- name: Geeft de naam van het resultaat aan.
- value: Geeft de waarde voor het resultaat aan.

Hoofdstuk 5

Keuzes

In dit hoofdstuk worden belangrijke keuzes bij de uitwerking uitgelegd en gemotiveerd.

5.1 Formaat Returnwaarden

Er kan gekozen worden tussen xml en json. Het verschil tussen xml en java script object notation (json) is dat xml een boomstructuur beschrijft. De json voorstel komt meer overeen met een hashmap. Xml is meer geschikt voor grote geavanceerde structuren. Aangezien het hier om eenvoudige monitoring informatie gaat, is het gebruik van xml af te raden. Bovendien is de xml-notatie van een object langer dan de overeenkomstige json-notatie. Ook is het parsen van json eenvoudiger. Bijgevolg zal hier dus geen xml, maar json gebruikt worden. Door de model-view-controller opbouw die gebruikt zal worden, kan xml-functionnalité achteraf makkelijk toegevoegd worden. Dit is al voorzien door de parameter format, die standaard ingevuld wordt met json.

5.2 Structuur returnwaarden

De teruggegeven data zal ook generiek gemaakt worden als een soort geneste hash/array waarin alle waarden zitten. Dit geeft het voordeel dat we ons geen zorgen moeten maken over attributen die niet ingevuld zijn, deze worden dan gewoon weggelaten. Het dan ook aan de client om de teruggegeven data om te zetten naar klassen, indien nodig.

5.3 Databank technologie

Het systeem van Oracle is al een aantal jaren de meest gebruikte structured query language (sql) versie (Fontecchio, 2012). De grootste open source variant is mysql (Mysql, 2008). Het huidige systeem draait echter op postgresql. Postgresql is een geavanceerd open source database management systeem (PostgreSQL, 2014a). Het heeft een zeer goede reputatie op gebied van snelheid, en robuustheid. Deze reputatie wordt ondersteund door de vele prijzen die het programma al in de wacht heeft kunnen slepen (PostgreSQL, 2014b). De huidige versie werkt met een PostgreSQL database (Partners, 2014).

Voordat er zonder meer verder gewerkt wordt met PostgreSQL, moet er toch een vergelijking gemaakt worden tussen PostgreSQL en MySQL. Er bestaan uiteraard ook andere databank systemen, toch komen deze 2 het meest in aanmerking door o.a. hun schaalbaarheid, snelheid en robuustheid (wikipedia, 2014).

Het voornaamste verschil is integriteit van data (Mikoluk, 2014). Doordat PostgreSQL minder verspreid is, zijn er minder hacks gekend. Daardoor wordt PostgreSQL als veiliger beschouwd. MySQL is sneller dan PostgreSQL, maar die laatste ondersteunt asynchrone calls. Hierdoor kunnen er meerdere calls tegelijk uitgevoerd worden. Hierdoor is het moeilijk om de prestatie tussen beide systemen te vergelijken.

Beide systemen ondersteunen atomicity, consistency, isolation en durability (ACID)(Lane, 2000) (MySQL, 2014). Atomicity beschrijft de ‘Alles of niets’ techniek die gebruikt wordt bij databanken. Complexe databanken bestaan vaak uit meerdere tabellen die met voorwaarden verbonden zijn. Als we data toevoegen, moet dat vaak in meerdere tabellen gebeuren. Dit kan uiteraard tot problemen leiden als er bij het toevoegen van een stuk data A aan tabelA een fout optreedt terwijl het toevoegen van data B aan tabelB wel lukt. Vermits dit tot inconsistenties leidt, is het ontoelaatbaar dat data B blijft staan. Om dit probleem op te lossen worden atomicity's gedefinieerd. Deze definiëren een set instructies die ofwel allemaal ofwel niet uitgevoerd moeten worden.

Consistency zorgt ervoor dat de datatypes gerespecteerd worden. Zo kan je geen stuk tekst steken in een kolom van het type date. Isolatie zorgt ervoor dat een database management systeem meerdere query's tegelijk kan laten lopen zonder dat deze interfereren met elkaar. Durability staat in voor back-uppen van data.

Beide technologieën gaan redelijk gelijk op. Een keuze maken is dan ook moeilijk. Mysql is sneller als we gebruik maken van de MyISAM engine. Deze engine is sneller doordat hij minder integriteitschecks uitvoert. Dit is geen probleem indien het om een website gaat waar voornamelijk gelezen wordt. PostgreSQL daarin tegen is zeer betrouwbaar en zal meer integriteitschecks uitvoeren. Deze webservice zal op zeer korte intervallen tests uitvoeren

en zo data binnen krijgen. Voor deze situatie geniet PostgreSQL de voorkeur. De huidige versie werkt ook met PostgreSQL. Deze masterproef zal ook met PostgreSQL gebruiken om de databank aan te sturen.

5.4 webservice technologie

Voor de programmeertaal mochten er voorstellen gedaan worden. Deze moesten echter voldoen aan een aantal vereisten. Zo moet technologie kunnen draaien op Linux-gebaseerde operating system (OS) . Verder moet de taal die gebruikt wordt wijd verspreid zijn.

De gekozen technologie, hypertext preprocessor (php), voldoet hieraan. PHP is een open source scripting taal gemaakt voor webdevelopment (PHP, 2014b). PHP draait volledig op de server en gegenereerd hypertext markup language (html) pagina's voor ze naar de client gestuurd worden. Doordat ze open source is, kan iedereen ze zonder meer gebruiken. Doordat ze vrij te gebruiken is, wordt PHP aangeboden door bijna alle hostingsbedrijven. Hierdoor is er een grote community van mensen die met PHP bezig is. Dat zorgt op zijn beurt voor uitgebreide documentatie en tutorial op het internet. PHP is een makkelijke taal om te leren, terwijl ze toch zeer krachtig blijkt te zijn. PHP heeft zeer veel ingebouwde functies. Dit alles zorgt ervoor dat PHP meer dan 80 procent van het markt aandeel bezig (w3Techs, 2014). Een vaak gebruikt tegen PHP is dat PHP niet goed is voor grote websites. Er zijn echter zeer veel grote websites die er gebruik van maken (Knop, 2010).

Doordat PHP ontworpen is voor webdevelopment, is het zeer eenvoudig om de opgegeven GET en POST parameters op te vragen. Deze parameters worden immers direct ingeladen. Ook is PHP voorzien om samen te werken met een hele reeks gekende databanken. Hiervoor moeten vaak geen externe drivers geïnstalleerd worden (PHP, 2014a). Tot deze reeks van databases behoort ook PostgreSQL, het database management systeem dat in deze masterproef gehanteerd wordt.

Een persoonlijk nadeel van PHP is dat ik het nog niet geleerd heb. Er zal dus een tijd nodig zijn om PHP te leren, daar tegenover staat wel dat de huidige versie in PHP is gemaakt, en deze dus als een uitstekend voorbeeld kan dienen. Ook de uitdaging om een nieuwe en veel gebruikte taal te leren spreekt mij aan. Zoals eerder vermeld is PHP een makkelijke taal die snel te leren is.

Bibliografie

- M. Fontecchio (2012). Oracle the clear leader in the \$24 billion rdbms market. URL <http://itknowledgeexchange.techtarget.com/eye-on-oracle/oracle-the-clear-leader-in-24-billion-rdbms-market/>.
- R. Knop (2010). Some really large php websites. URL <http://blog.richardknop.com/2010/03/some-really-large-php-websites/>.
- T. Lane (2000). Transactions processing in postgresql. URL <http://www.postgresql.org/files/developer/transactions.pdf>.
- K. Mikoluk (2014). Mysql vs postgresql: why mysql is superior to postgresql. URL <https://www.udemy.com/blog/mysql-vs-postgresql/>.
- Mysql (2008). Market share. URL <http://www.mysql.com/why-mysql/marketshare/>.
- MySQL (2014). Mysql and the acid model. URL <https://dev.mysql.com/doc/refman/5.6/en/mysql-acid.html>.
- W. Partners (2014). Fp7-ict-2011-8. URL http://www.fed4fire.eu/fileadmin/documents/public_deliverables/D2-4_Second_federation_architecture_Fed4FIRE_318389.pdf.
- PHP (2014a). Vendor specific database extensions. URL <http://www.php.net/manual/en/refs.database.vendors.php>.
- PHP (2014b). What is php? URL <http://www.php.net/manual/en/intro-what-is.php>.
- PostgreSQL (2014a). About. URL <http://www.postgresql.org/about/>.
- PostgreSQL (2014b). Awards. URL <http://www.postgresql.org/about/awards/>.
- w3Techs (2014). Usage statistics and market share of php for websites. URL <http://w3techs.com/technologies/details/pl-php/all/all>.
- wikipedia (2014). Database-comparison. URL http://en.wikipedia.org/wiki/Comparison_of_relational_database_management_systems.