

Testrapport
Kiezen op Afstand
Inhoudelijke Stresstest

September 2006
Dit document heeft 10 pagina's

Document historie

Versie	Datum	Bijzonderheden	Autorisatie
0.1	20-09-2006	Opzet	
0.2	22-09-2006	Wijzigingen doorgevoerd	

Inhoudsopgave

1	Algemeen.....	4
1.1	Inleiding	4
1.2	Referenties.....	4
1.3	Opmerkingen.....	4
2	Testplan	5
2.1	Scope van het testen	5
2.2	Organisatie/ rollen/ verantwoordelijkheden	5
2.3	Testomgeving(en).....	6
3	Testresultaten.....	7
3.1	Testresultaten Prepare Run 1.....	7
3.2	Testresultaten Prepare Run 2.....	7
3.3	Testresultaten Prepare Run 3.....	8
3.4	Testresultaten Stem invoering Run 1	8
3.5	Testresultaten Stem invoering Run 2	8
3.6	Testresultaten Tally Run 1	9
4	Conclusie	10

1 Algemeen

1.1 Inleiding

In dit document worden de resultaten van een inhoudelijke stresstest op de verschillende deelsystemen voor Kiezen op Afstand beschreven.

1.2 Referenties

Nr	Referentie	Versie
1	Master Test Plan Kiezen op Afstand	0.26
2	Testplan Deelsystemen Test Kiezen op Afstand	0.3
3	Testcases Deelsystemen Test Stemdienst	div.
4	FO Kiezen op Afstand	0.34
5	Ries – 2008 Crypto, Cryptographic architecture for RIES-2008 and RIES-KOA	0.95

1.3 Opmerkingen

Enkele algemene opmerkingen aangaande het uitvoeren van de test.

- Tijdens het testtraject zijn enkele blokkerende bevindingen direct opgelost door het bouwteam in overleg met het testteam, waarna verder getest kon worden.
- De tests zijn uitgevoerd op verschillende versies van de programmatuur. Na een blokkerende bevinding binnen een Prepare-componenten van versie 2.2, is door het bouwteam een oplossing geleverd in de vorm van versie 2.3 (zowel Prepare als Tally). Verdere tests zijn uitgevoerd op versie 2.3.

2 Testplan

2.1 Scope van het testen

Doel van het uitvoeren van een inhoudelijke stresstest is om de functionele werking, betrouwbaarheid en juistheid van de stemdienst te toetsen bij een grote hoeveelheid kiezers. In tegenstelling tot het uitvoeren van verschillende door Surfnet uitgevoerde performance – en stresstesten op Voting Window, zal nu ook een aanzienlijke belasting worden uitgevoerd op de onderdelen Prepare als Tally.

De opzet van de test bestaat uit 200.000 fictieve kiezers die deels geldige en deels dubbele (ongeldige) stemmen zullen uitvoeren. De verwachte hoeveelheid kiezers voor de verkiezing van 2006 wordt geschat op 15.000.

De inhoudelijke stresstest zal gebruik maken van de volgende deelcomponenten:

- *Onderdeel Prepare:*
 - 1) SaltpartsGenVoterKey
 - 2) SaltpartsMK
 - 3) SaltpartsReceiptKey
 - 4) GenVoterKey
 - 5) ReceiptKey
 - 6) AcceptData / DataCheck
 - 7) GenDuplicates / GenTest
 - 8) GenC10
 - 9) RTCheck
 - 10) RICOR
 - 11) GenVoteWndwData
- *Onderdeel Tally:*
 - 12) Helpdesk mutation tool
 - 13) RTCheck
 - 14) RICOR
 - 15) Tally tool
 - 16) Receipt conversion tool
- *Administration tool*
- *Voting Window*

2.2 Organisatie/ rollen/ verantwoordelijkheden

In deze paragraaf wordt een overzicht gegeven van de verschillende verantwoordelijkheden binnen de teststrategie:

Testsoort	Verantwoordelijke rol
Inhoudelijke stresstest	Testteam Kiezen op Afstand
Applicatiebeheer	Functioneel beheerder internetstemdiest
Technisch beheer	Surfnet

2.3 Testomgeving(en)

1) Testomgeving testteam

De testactiviteiten door het testteam Kiezen op Afstand zijn wat betreft het invoeren van de grote hoeveelheid stemmen uitgevoerd via de eigen omgeving. De omgeving is ingericht met vooraf gegenereerde set stemdata, welke middels de applicatie Apache Jmeter geautomatiseerd zijn ingevoerd op het Voting Window systeem.

2) omgeving beheerder internetstemdienst

De testactiviteiten met betrekking tot de beheerconsole en de verwerking van de Prepare en Tally processen zijn uitgevoerd door het testteam Kiezen op Afstand in samenwerking met een specialist van de beheerder van de stemdienst, gebruik makende van de werkplekomgeving die uiteindelijk ook door de beheerders van de stemdienst als productie omgeving gebruikt zullen worden.

3) Surfnets omgeving

Door Surfnets is de technische (productie) omgeving van Voting Window nauwlettend gemonitord.

3 Testresultaten

Enkele algemene opmerkingen aangaande het uitvoeren van de inhoudelijke stresstest.

- Er bleek een fout te zijn gemaakt door het testteam bij het genereren van de 200.000 fictieve kiezers in het WV-STUF-K10 bestand. Uit de eerste tests bleek dat het volgnummer niet uniek was. Hiervoor is een nieuwe versie van het K10 bestand opgeleverd.
- Het automatisch invoeren van de stemmen (Voting Window) via Jmeter is op twee manieren uitgevoerd. In eerste instantie was gekozen om kiezergedrag te simuleren door alle pagina's (en de bijhorende javascripts en CSS-bestanden) op te vragen. Hieruit bleek dat de target van 200.000 stemmen niet gehaald kon worden binnen de daarvoor bestemde tijd. Om deze reden is gekozen om direct de technische stem te versturen naar de server.
- Het uitvoeren van de Prepare-fase is uitgevoerd in meerdere runs, vanwege een aantal blokkerende bevindingen.
- De steminvoer is uitgevoerd in twee runs, waarbij in de eerste run de HTTP-connecties waren ingesteld om deze aan te houden gedurende de sessie. Tijdens deze run bleek dat Jmeter de connecties niet meer vrijgaf, waardoor het aantal mogelijke nieuwe connecties drastisch afnam en de testuitvoer hinderde. In run 2 is gekozen om de http-requests onafhankelijk van elkaar uit te voeren, waardoor de connecties vrijgegeven werden.
- De opgeleverde uitslag uit het Tally proces kan niet worden gevalideerd. Dit omdat er geen inzicht is in welke stemmen er niet succesvol zijn doorgegeven. Dit was niet het doel van de test.

3.1 Testresultaten Prepare Run 1

Aan de hand van een STUF-K10 bestand, waarin de kiezergegevens van 200.000 fictieve personen opgenomen waren, is getracht de Prepare-fase in te gaan met applicatieversie 2.2. Bij het uitvoeren van de eerste controle op de data door middel van de AcceptData-applicatie, bleek dat het geheugengebruik binnen de applicatie niet optimaal was. De java-applicatie liep door de grote hoeveelheid kiezers buiten haar beschikbare geheugenhoeveelheid, waardoor de applicatie crashte. Dit probleem is verholpen in Run 2 door een maximale hoeveelheid werkgeheugen toe te kennen aan alle java-applicaties en hiervoor een minimale hoeveelheid te reserveren.

3.2 Testresultaten Prepare Run 2

Tijdens Run 2 werden de minimale en maximale hoeveelheid werkgeheugen vastgelegd voor elke applicatie. Wederom werd getracht de Prepare-fase in te gaan met applicatieversie 2.2. Het uitvoeren van de eerste controle op de data door middel van de AcceptData-applicatie gaf dit maal geen applicatiecrash, toch bleek dat het gebruikte algoritme binnen de applicatie niet optimaal was. De bepaling van de uniciteit van de volgnummers binnen K10 werd aanzienlijk trager naarmate de applicatie meer records had verwerkt. Omdat er geen inzicht was in de hoeveelheid verwerkte records en dus in de status van de applicatie is dit besproken met het bouwteam.

Het bouwteam heeft een teller ingebouwd in de applicatie en het algoritme herzien. Deze oplossingen zijn ingesloten in versie 2.3 van Prepare en Tally.

3.3 Testresultaten Prepare Run 3

Run 3 verliep verder zonder problemen. De AcceptData applicatie en alle overige applicaties uit het Prepare proces verwerkten het K10 bestand correct en relatief snel. De relevante (mogelijk tijdrovende) stappen in het prepare proces zijn:

- AcceptData, verwerkingstijd van 200.000 kiezers = 10 seconden;
- GenC10, verwerkingstijd van 200.000 kiezers = 16 minuten.

3.4 Testresultaten Stem invoering Run 1

Tijdens run 1 was Jmeter, zoals eerder vermeld, ingesteld om HTTP-connecties aan te houden gedurende de sessie. Tijdens de sessie werd kiezergedrag gesimuleerd door alle pagina's en bijhorende javascripts en CSS-bestanden in logische volgorde op te vragen en relevante data naar de server te versturen.

Voor de uitvoering van Run 1 zijn 4 laptops ingesteld, elk met 50.000 vooraf berekende stemmen.

Uit de run bleek dat het aantal mogelijke connecties naar de Voting Window server relatief snel in gebruik werden genomen door het testscript, en dat de connecties niet meer vrijgegeven werden door de Voting Window server. Direct gevolg hiervan was het ontstaan van een wachtrij aan de kant van de testmachines en zodoende werden er te weinig stemmen aan de kant van de Voting Window server ontvangen. Het resultaat was dat na 15 uur er slechts 5000 stemmen doorgekomen en opgeslagen waren. Besloten is om de test run af te breken.

3.5 Testresultaten Stem invoering Run 2

Na het afbreken van de automatische invoer tijdens Run 1 is besloten om de Keep-Alive instelling voor de HTTP-connecties uit te schakelen en het aantal requests te beperken tot het direct invoeren van de stem.

Voor de uitvoering van Run 2 zijn 2 laptops ingesteld, elk met 100.000 vooraf berekende stemmen.

Bovengenoemde wijzigingen hadden als gevolg dat er 50-60 stemmen per seconde ingevoerd konden worden vanuit de testomgeving, waardoor de target van 200.000 stemmen in relatief korte tijd gehaald werd.

Uit de test bleek dat ongeveer 1 uit 2000 requests (0.5%) faalde. De reden hiervan is naar alle waarschijnlijkheid te vinden in de proxyserver in de testomgeving. Naar vermoeden kon de proxyserver het grote aantal simultane connecties niet aan (was ingesteld op 512, is gedurende de test verhoogd naar 1024), waardoor voor enkele request een time-out optrad en de stem niet correct werd doorgegeven aan de Voting Window server.

In de gebruikelijke live-situatie van een kiezer zal deze situatie niet voorkomen, daar de grote hoeveelheid stemmen niet verwacht wordt. Daarnaast kan de kiezer zelf reactief handelen op een eventuele time-out, door op een later tijdstip nogmaals zijn stem uit te brengen.

3.6 Testresultaten Tally Run 1

Nadat het testteam meer dan 200.000 stemmen ingevoerd had, is de verkiezing gesloten en kon het Tally proces worden uitgevoerd. Hierbij is gebruik gemaakt van versie 2.3, welke de geheugen-allocatie oplossing uit hoofdstuk 3.1 bevat.

Bij de uitvoering van het Tally-proces zijn geen fouten opgetreden.

De uitslag van Tally bevatte 204939 stemmen.

Het tellen van de stemmen en genereren van de uitslag neemt de volgende hoeveelheid tijd in beslag:

- 1) RT check: 2 minuten;
- 2) Ricor: 4 minuten;
- 3) Shipment: 7 minuten;
- 4) Tally: 25 minuten.

4 Conclusie

Uit de inhoudelijke stresstest blijkt dat de componenten, na hertstel en hertest, correct functioneren en binnen acceptabele performance criteria blijft bij een hoeveelheid van 200.000 kiezers. Uit de test kwam wel het vraagstuk over het aantal toegestane connecties naar de Voting Window server bij Surfnet, welke wellicht meer onderzoek vereist.