



# Greining á framkvæmd samræmdra könnunarprófa

3.4.2018

Hannes Pétursson

R.ice ehf

Heiðvangur 14

220 Hafnarfjörður

## Inngangur

Samræmd könnunarpróf í íslensku, stærðfræði og og ensku voru haldin dagana 7., 8. og 9. mars síðastliðinn fyrir um 4300 nemendur. Á fyrsta og þriðja próftökudegi komu upp vandamál við framkvæmd þeirra sem urðu til þess að margir nemendur gátu ekki þeytt þau eða gerðu það við óviðunandi kringumstæður. Tilgangur þessarar skýrslu er að greina þau vandamál sem upp komu og leiða í ljós hvað hefði mátt betur fara í undirbúningi stofnunarinnar.

Menntamálastofnun (MMS) sér alfarið um undirbúning og framkvæmd samræmdra könnunarprófa fyrir grunnskóla og kaupir aðgang að tölvukerfinu FastTest auk hýsingar frá fyrirtækinu Assessment Systems Corp. (ASC) sem notað er við framkvæmd prófanna. Þetta kerfi hefur verið notað við samræmd próf frá árinu 2016 og er því komin ágæt reynsla af notkun þess hér á landi. Sem dæmi vor prófin lögð fyrir 9. og 10. bekk samtímis árið 2017 fyrir um tvöfalt fleiri nemendur á sama tíma. Auk þess hefur kerfið verið notað víða um heim við margs konar skilyrði. Kerfið er hannað fyrir mikið álag og hafa vandamál sem þessi ekki komið upp áður þó álagið hafi verið mun meira í fyrri prófum þar sem það var notað.

ASC er fyrirtæki sem hefur boðið upp á hugbúnað og þjónustu í tengslum við prófafyrirlagnir frá árinu 1979 og FastTest er sérstaklega hannað fyrir aðstæður þegar mikið liggur undir (e. high-stake tests) þar sem mikið álag er á kerfinu. Þegar bilanir verða eins og ræddar eru í þessari skýrslu, eru oftast en ekki fleiri en ein villa sem koma upp á sama tíma. Margföldunaráhrif þeirra verða til þess að kerfið bregst, þrátt fyrir stöndugleika þess. Tilgangurinn með greiningunni sem lýst er í þessari skýrslu er:

- A. að fá skilning á hvað orsakaði tæknilegar kerfisbilanir í þessum tilfellum og
- B. draga af því lærdóm sem nýtist til að bæta ferli við tæknilegar uppsetningar svo þessar aðstæður þurfi ekki að koma upp aftur.

## Helstu niðurstöður

Nokkrum dögum eftir kerfisbilunina, nánar tiltekið 16. mars, kom yfirlýsing frá ASC til Menntamálastofnunar, þar sem þeir lýsa yfir fullri ábyrgð á því sem gerðist þegar samræmdu prófin voru lögð fyrir hér á landi. Í henni stendur m.a.:

“We sincerely apologize to the Iceland Directorate of Education and all the parents and students that were affected by the failure experienced during test sessions on March 7th & 9th. We take full responsibility for the inconvenience, delay, and frustration that this has caused. **There was nothing the Directorate of Education could have done to prevent this from happening.** Our system was used appropriately however, our QA process failed. When the issue was discovered, we worked diligently to ensure it was immediately corrected.”

Það er alveg skýrt að ASC tekur fulla ábyrgð á þeim bilunum sem upp komu og samræmist það niðurstöðunum sem birtast hér. Meginástæðan fyrir þessum rekstrarbilunum voru villur í kerfum ASC og það er fátt, ef nokkuð, sem MMS hefði geta gert til að koma í veg fyrir þær.

Þó ASC beri ábyrgð á villunum sjálfum þá er ýmislegt sem má bæta í ferlum og vinnubrögðum MMS hvað samræmd könnunarpróf varðar. Ekkert formlegt gæðaferli var til staðar og allt traust var sett þriðja aðila. Nánar verður farið í það í kaflanum um rót vandans.

Þar sem annar greiningaraðili fer yfir ferlana sérstaklega þá mun þessi skýrsla ekki fara djúpt í þann hluta vandans heldur einblína frekar á tæknilegu hliðina.

## Upplýsingaöflun og greining

Niðurstöður þessarar greiningar eru byggðar á eftirfarandi atriðum:

- Viðtölum við lykilstarfsmenn MMS
- Tölvupóstsamskiptum milli MMS og ASC í aðdraganda fyrirlagnar og á prófdögum
- Skjölum frá MMS í tengslum við vinnuferli og áhættugreiningu
- Greiningu ASC á rekstraróhöppunum (e. Incident report)
- Niðurstöðum úr álagsprófum frá ASC
- Lýsingum á rekstrarumhverfi og nánari lýsingu á rekstraróhöppum frá ASC
- Ýmsum opinberum upplýsingum um starfsemi MMS og ASC

Veitt var fullt aðgengi að lykilstarfsfólki MMS og að tengiliðum hjá ASC. Samskiptin voru góð og komu svör við öllum spurningum og aðgengi veitt að öllum þeim gögnum sem óskað var eftir.

## Bilanagreining

Eins og nefnt var hér að framan þá er algengt í tilfalli stórkostlegra rekstrarbilana að upp koma nokkrir samverkandi þættir sem gerir greiningu vandans erfiðari og bilanagreiningu torveldari.

Niðurstaða greiningar á orsökum rekstarbilana er að ástæðurnar megi rekja til fjögurra þátta, sem eru:

1. Álagsþolið umhverfi “High volume engine” (HVE)
2. Auðkenningarskjár “Pretest screen”
3. Slæm fyrirspurn “Bad query”
4. Stjórnendavirkni

Þessir þættir eru útskýrðir nánar hér að neðan.

### Álagsþolið umhverfi High Volume Engine (HVE)

Þegar próf eru lögð fyrir fjölda nemanda á sama tíma þarf að setja þau upp í umhverfi sem þolir mikið álag, svokallað High Volume Engine (HVE). Nánari lýsing á því í hverju það felst er í viðauka 1.

Ferlið hjá ASC til að merkja próf í álagsþolinum flokk er frumstætt og býður upp á mannleg mistök. Það fer fram á eftirfarandi hátt:

- MMS sendir lista í tölvupósti á sinn tengilið hjá ASC með prófum sem eiga að fara í álagsþolið umhverfi.
- ASC gerir breytingar á kerfinu **handvirkt** og lætur MMS vita þegar prófin eru komin í álagsþolið umhverfi.

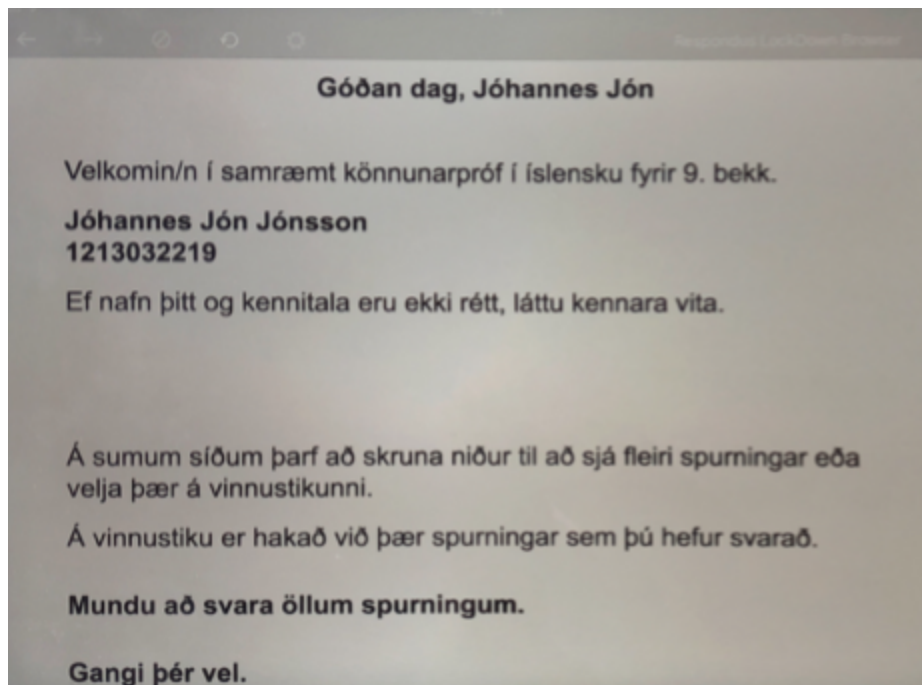
Vegna þessara handvirku stillinga í kerfinu er engin leið fyrir MMS að ganga úr skugga um hvort búið sé að setja próf í álagsþolinum flokk. Í svari frá ASC kemur fram að þetta sé áhættusöm aðgerð og mikil hættu sé á mannlegum mistökum. Einnig fullyrða þeir að það sé forgangsmál að gera þessar merkingar sýnilegar fyrir kaupendum kerfisins í nýrri útgáfu kerfisins sem er væntanleg á næstu vikum.

Eftir fyrsta prófdag kom í ljós að prófin höfðu ekki verið sett í álagsþolinum flokk í kerfinu.

## Auðkenningargluggi (Pretest screen)

Eftir að nemandi hefur slegið inn sinn auðkenningarkóða kemur upp gluggi sem birtir nafn og kennitölu nemandans og býður hann velkominn. Þetta er gert til þess að nemandinn fái staðfestingu á því að hann hafi slegið inn réttan auðkenningarkóða.

Hér er dæmi um auðkenningarglugga:



Mynd 1 Auðkenningargluggi

Villa reyndist vera í kerfinu sem kom fram þegar þessi gluggi birtist. Það lýsti sér í auknu álagi á gagnagrunnsþjóninn. Kaupendur kerfisins geta valið hvort þeir nota þennan skjá eða ekki. Álagsprófanir sem ASC hefur framkvæmt reglulega á kerfinu hafa ekki notað þennan valmöguleika í keyrslu kerfisins. Prófanirnar sýndu þar af leiðandi ekki álagið miðað við þær stillingar sem notaðar voru þessa umræddu daga sem bilunin varð hér á landi.

## Slæm fyrirspurn (Bad query)

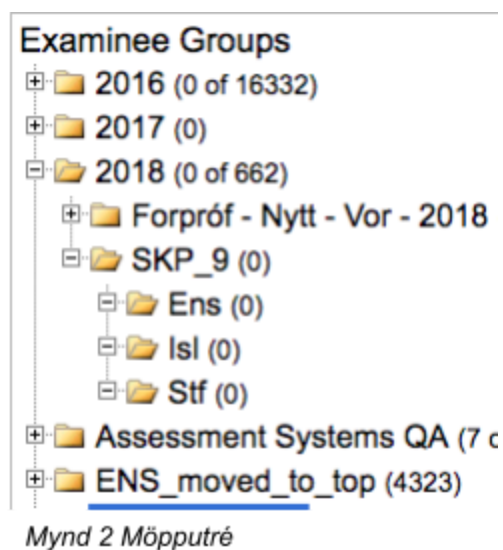
Þá villu sem að öllum líkindum hefur valdið mestu álagi á kerfið kölluðu starfsmenn ASC *slæmu fyrirspurnina* (bad query). Þessi villa varð til þess að gerð var fyrirspurn á gagnagrunninn sem leitaði í um 8000-12000 línur í grunninum fyrir innskráningu hjá *hverjum* nemanda. Í hvert skipti sem nemandi skráði sig inn í kerfið fór *slæma fyrirspurnin* af stað og bjó til mikið álag á gagnagrunninn. Álagið var það mikið að við framkvæmd enskuprófsins varð kerfið algjörlega óstarfhæft í um 45 mínútur. Það er mjög líklegt að hver

og einn nemandi hafi reynt mörgum sinnum að skrá sig inn í kerfið, þegar villur komu upp í innskráningu, sem margfaldaði áhrif þessarar fyrirspurnar fyrir hvern nemanda.

Villan tengdist því hvernig prófin voru geymd í prófkerfinu. Hægt er að búa til möpputré til að skipuleggja betur utanumhald prófa. Þar sem MMS hefur notað þetta kerfi í nokkur ár og er komin með umtalsverðan fjölda af prófum var ákveðið að flokka þau eftir árum og svo eftir bekkjum (sem dæmi: SKP\_9, samræmt könnunarpróf 9. bekkur) og svo fagi (Ens, Isk, Stf).

*Slæma fyrirspurnin* var framkölluð vegna þess að prófin voru geymd dýpra í trénu en vanlegt er hjá öðrum notendum. Við álagsprófanir hjá ASC hafa prófin ávallt verið staðsett í rótarmöppunni og því kom vandamál *slæmu fyrirspurnarinnar* ekki upp á yfirborðið í þeim (sjá viðauka 4).

Starfsmenn ASC áttuðu sig ekki á *slæmu fyrirspurninni* fyrr en á föstudeginum. Þegar þeir voru búnir að greina villuna var hægt að koma í veg fyrir hana tímabundið með því að færa prófin efst í möpputréd. Á myndinni hér til hliðar (*mynd 2*) má sjá að búið er að færa enskuprófið upp í rótina (ENS\_moved\_to\_top).



ASC hefur sagt í tölvupóstum eftir að villan kom upp að önnur fyrirtæki sem lagt hafa fyrir álagspólin próf (HVE) hafi haft prófin í rótarmöppunni en MMS var ekki gert kunnugt um þá vinnureglu.

Þessi villa olli e.t.v. mesta álaginu á kerfið og liggur ábyrgð ASC í því að hafa ekki upplýst MMS um nauðsyn þess að hafa prófin í rótarmöppunni til að fyrirbyggja *slæmar fyrirspurnir*.

## Stjórnendavirkni

Í kerfinu eru sérstakar aðgerðir sem eru ætlaðar fyrir stjórnendur til að hafa eftirlit með framkvæmd prófanna, m.a. til að sjá hversu margir nemendur eru búnir að skrá sig inn í próf og fleira.

Þessar aðgerðir geta verið þungar í keyrslu og var ákveðið að slökkva á þessum tengingum þegar hið mikla álag á gagnagrunninn var orðið ljóst.

Í upphafi var talið að þessar tengingar hefðu getað valdið álagi á grunninn en eftir að búið var að fara í nánari greiningar og átta sig á *slæmu fyrirspurninni* þykir ólíklegt að þetta hafi haft teljanleg áhrif.

## Atburðarás

### Aðdragandinn

Í aðdraganda prófanna fóru margir tölvupóstar á milli ASC og MMS. Smávægilegir hnökrar komu upp og MMS leitaði til ASC í þeim tilgangi að fá ráðleggingar við lausn þeirra.

Atburðarrásin var þessi:

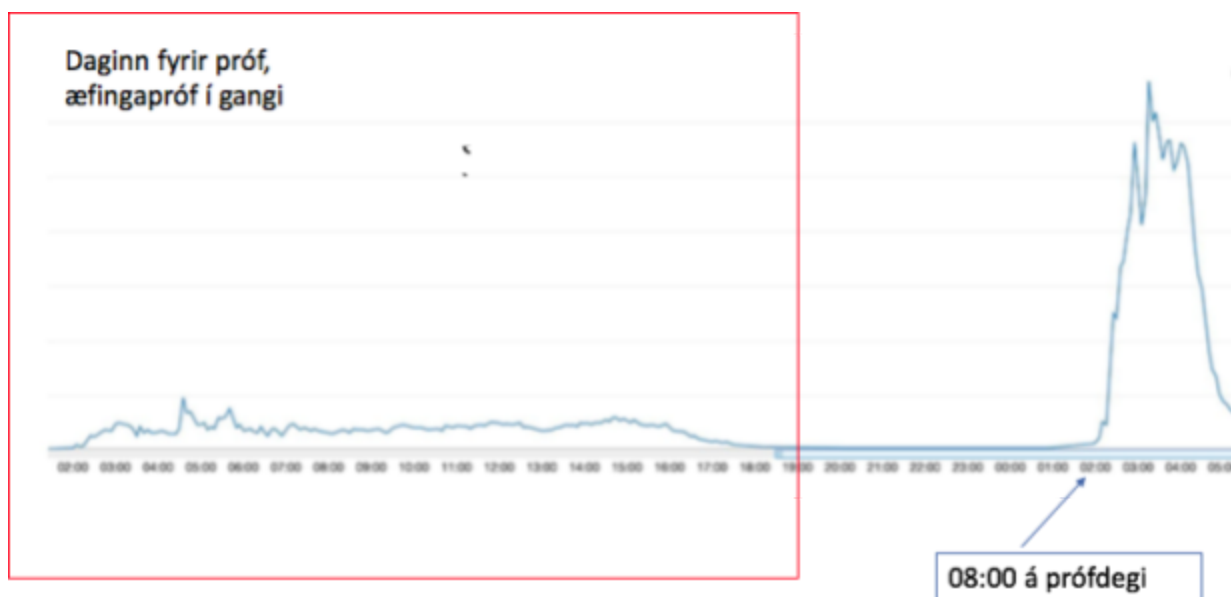
1. MMS sendir póst 1. mars og spyr hvort eitthvað þurfi að gera áður en prófin verði lögð fyrir. ASC óskar þá eftir lista yfir prófin svo hægt sé að flokka þau sem álagsþolin (HVE). Formlegt svar frá MMS við þessari beiðni ASC var ekki sent til ASC.
2. Þann 2. mars eru samskipti MMS og ASC tengd villu í notkun iPad spjaldtölva. Í þeim samskiptum svarar MMS póstinum frá 1. mars með lista yfir þau próf sem á að leggja fyrir og ASC segist ætla að setja þau í álagsþolinn flokk. MMS biður um staðfestingu á þeirri flokkun en fær ekki svar við því.
3. Þann 5. mars sendir MMS póst á ASC og lætur vita að öll prófin séu tilbúin til fyrirlagnar.  
Það er mitt mat að þar sem þessi samskipti fara ekki fram formlega sem spurning og svar undir sama efni (e. subject) í tölvupóstum virðist sem þessar upplýsingar hafi misfarist hjá ASC (sjá tölvupósta í viðauka 2).
4. Starfsmenn ASC fylgjast með álagi daginn fyrir prófdag þegar nemendur hafa kost á að framkvæma æfingapróf og sjá að álagið var lítið (sjá mynd). Þeir átta sig á að sum prófin eru ekki flokkuð sem álagsþolin en meta það í lagi í ljósi lítils álags daginn fyrir próf.  
Þetta tel ég vera einkennilega ályktun þar sem álag fyrir prófdag dreifðist jafnt yfir allan daginn þar sem nemendum er frjálst að sækja þessi próf þegar þeim hentar, en skrá sig ekki allir inn á sama tíma eins og gerist á prófdögum sjálfum. Auk þess var hvorki notast við auðkenningarkóða né auðkenningarglugga (Pretest screen) sem var notaður á prófdaginn og síðast en ekki síst voru prófin ekki vistuð í þrískiptu möppuskipulaginu eins og gert var prófadagana og því kom vandamál *slæmu fyrirspurnarinnar* ekki upp í æfingaprófunum.

### Prófdagur 1 (7. mars Íslenskupróf)

Niðurstaða bilanagreiningar ASC eftir fyrsta dag prófanna var sú að að álagið hefði verið mikið vegna þess að prófin voru ekki flokkuð sem álagsþolin. Þetta tel ég vera eðlilega ályktun hjá þeim í ljósi þess að framkvæmd prófanna hylur þær villur sem komu fram seinni tvo dagana. Hins vegar hefði auk þess verið eðlilegt að ASC hefði skoðað fleiri þætti í stað þess að gera ráð fyrir að orsök álagsins lægi í þessu eina atriði, sérstaklega í ljósi þess

að kerfið hefur verið notað með góðum árangri áður fyrir mun stærri hópa og það ætti að gefa kerfisstjórum tilefni til að fara yfir uppsetninguna í heild. Enn fremur tel ég það eðlilega kröfu að ASC hefði álagsprófað kerfið miðað við uppsetningu MMS eftir að svona rekstrarstopp kom upp fyrsta prófdaginn.

Hér má sjá yfirlitsmynd (*mynd 3*) yfir álag á gagnagrunnsþjóninn daginn fyrir próf og á prófdegi.

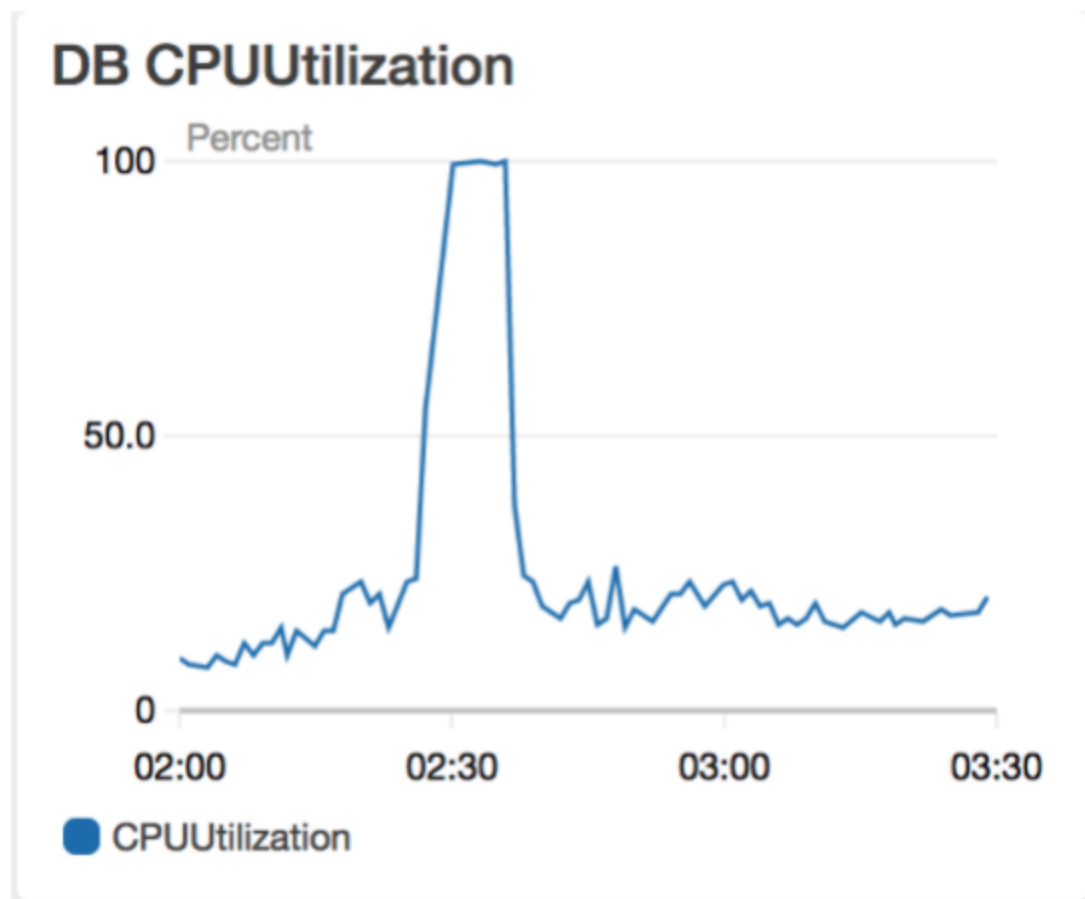


Mynd 3 Álag á gagnagrunnsþjón á prófdegi 1.

## Prófdagur 2 (8. mars Stærðfræðiþróf)

Talað hefur verið um að framkvæmd stærðfræðiþrófsins á öðrum prófdegi hafi gengið snurðulaust fyrir sig en þegar rýnt er í gögn sem óskað var eftir frá ASC má sjá að það kom upp mikið álag í upphafi próftímans (klukkan 2:30 í USA en þá var klukkan einmitt 8:30 á Íslandi, eða tíminn sem stærðfræðiþrófið byrjaði).





Mynd 4 Álag á gagnagrunnsþjón á prófdegi 2.

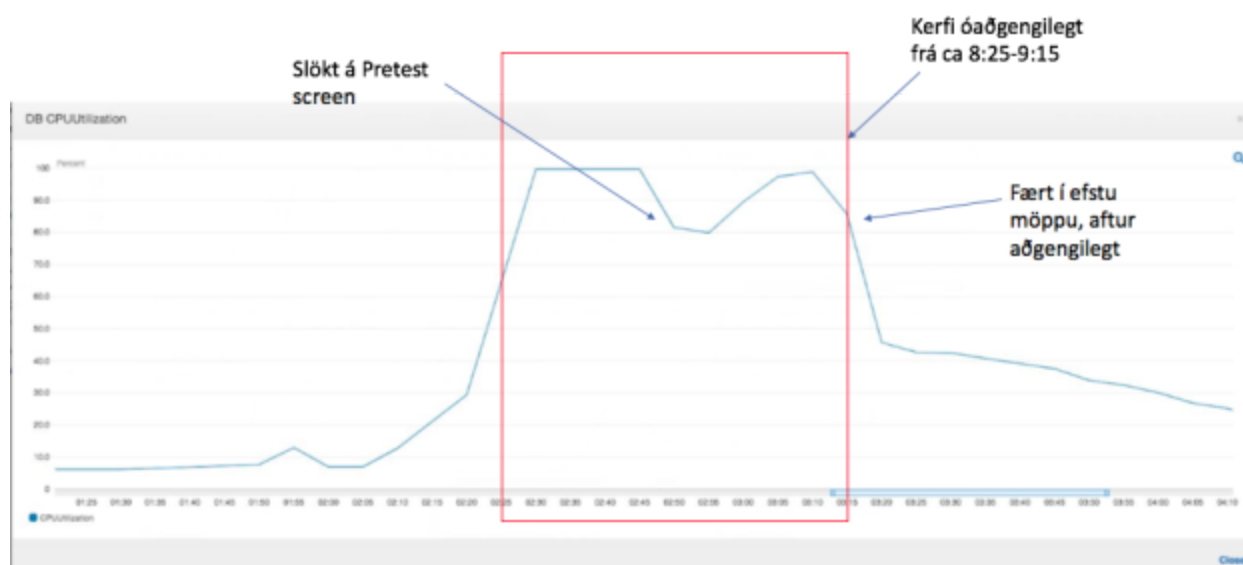
Á mynd 4 má sjá að um klukkan hálf níu, fyrir hádegi, fer álagið á gagnagrunnsþjóninum í 100% í nokkrar mínútur og er því nokkuð víst að einhverjir nemendur hafi lent í vandræðum á þessum tíma með að komast inn í prófið en líklega geta skráð sig aftur inn þegar búið var að koma böndum á gagnagrunnsþjóninn.

Á þessum tímamarki var búið að merkja öll próf til að keyra í álagsþolnu umhverfi (HVE).

Starfsmenn ASC gátu komið álaginu í lag með því að slökkva tímabundið á auðkenningaglugganum (Pretest screen) og stoppa allar tengingar frá stjórnendum við gagnagrunninn. Við þá aðgerð stoppuðu þeir tengingar sem urðu út af *slæmu fyrirspurninni* (Bad query) án þess að gera sér grein fyrir því. Þess vegna uppgötvast hún ekki á þessum tíma. Það tók starfsmenn ASC stuttan tíma að koma böndum á gagnagrunninn og allt gekk eðlilega eftir það þennan prófdag. Þetta stutta álagstímabil virtist ekki hafa haft slæm áhrif á próftöku þennan dag. **Hinsvegar tel ég að starfsmenn ASC hefðu átt að kanna þetta betur og upplýsa MMS um að það hefði orðið tímabundið álag eftir að þetta kom upp á degi tvö.**

## Prófdagur 3 (9. mars Enskupróf)

Þriðja prófadaginn var mikið álag á gagnagrunnsþjóninum vegna auðkenningarglugga (pretest screen) og út af *slæmri fyrirspurn*. Gagnagrunnsþjóninn hætti að svara og prófakerfið varð óstarfhæft. Þær ráðstafanir sem ASC gerðu deginum áður dugðu ekki til að minnka álagið. *Slæma fyrirspurnin* olli helmingi fleiri fyrirspurnum fyrir hverja innskráningu. Fyrirspurnirnar voru 12.000 en daginn áður voru þær 8.000. Það var fyrst þá sem starfsmenn ASC áttuðu sig á vandamáli *slæmu fyrirspurnarinnar*. Fram að þessu héldu þeir að það væri nóg að slökkva á *auðkenningarglugganum*. Kerfið var orðið óstarfhæft áður en hægt var að reyna þær fyrirbyggjandi aðgerðir sem voru framkvæmdar deginum áður.



Mynd 5 Álag á gagnagrunnsþjón á prófdegi 3.

Með því að færa prófið í rótarmöppu í prófakerfinu var hægt að koma álaginu á gagnagrunnsþjóninum í eðlilegt ástand og kerfið varð aftur starfhæft en þá hafði verið hætt við að reyna áfram og nemendur sendir heim.

**Ég tel að ASC hefði átt að greina þetta vandamál betur og vera með annað viðbragðsplan færi gagnagrunnsþjóninn aftur upp í 100% álag.**

**Ef kerfið hefði auk þess verið álagsprófað án auðkenningargluggans hefði *slæma fyrirspurnin* komið skýrt fram.**

## Niðurstöður

Bæði fyrsta og þriðja prófdaginn kom upp sama vandamálið, þ.e. of mikið álag varð á kerfinu og það varð óstarfhæft. Er eitthvað sem MMS hefði getað gert til að koma í veg fyrir þessi vandamál? Til að svara því þarf að kafa aðeins dýpra í vandamálin og þá er gott að spyrja sig nokkrum sinnum *af hverju?*

### Rót vandans, lærdómur

Vandamál 1 (fyrirlögn á íslenskuprófi)

**Einkenni:** Prófakerfið ræður ekki við álag og verður óstarfhæft

Greining á rót vandans:

#### 1. Af hverju voru prófin ekki sett í álagspolinn flokk?

Ferlið við að merkja próf í álagspolinn flokk er mjög frumstætt og treystir á formleg tölvupóstsamskipti. Þau samskipti voru ekki skilvirk og því má segja að um samskiptavandamál hafi verið að ræða.

#### 2. Af hverju var samskiptavandamál?

Það er lítil áhersla í verkferlinu á umsýslu og rekstur tölvukerfisins. Það kom ekki fram í verkferli að óska þyrfti formlega eftir að setja próf í álagspolinn flokk.

#### 3. Af hverju er ekkert í verkferlinu um þessar stillingar?

Það verkferli sem unnið er eftir hjá MMS er nánast eingöngu um gerð prófsins á pappír en ekkert um tölvukerfið. Ferlið er því pappírsmiðað.

#### 4. Af hverju er verkferlið enn pappírsmiðað?

Aðeins nokkur ár eru síðan öll prófin voru framkvæmd á pappír. Þegar framkvæmdin var færð yfir í tölvur var ferlið ekki endurskoðað með tilliti til þess í samráði við hugbúnaðarsérfræðinga og hafa engir hugbúnaðarsérfræðingar komið að ferlinu hingað til.

**Niðurstaða: Það vantar aðkomu hugbúnaðarsérfræðinga að ferlinu.**

Vandamál 2 (fyrirlögn á enskuprófi)

**Einkenni:** Prófakerfið ræður ekki við álag og verður óstarfhæft

Greining á rót vandans:

#### 1. Af hverju komu upp villur í kerfinu?

Mjög alvarlegar villur voru í kerfinu tengdar innskráningu og auðkenningarglugga. Við álagsprófun á kerfinu miðað við uppsetningu og stillingar MMS hefðu þessir veikleikar átt að koma upp.

## 2. Af hverju var kerfið ekki prófað miðað við þessa uppsetningu?

Það er ekkert í verkferli MMS sem lýtur að því að prófa tölvukerfið.

## 3. Af hverju er ekkert í verkferlinu um að álagsprófa kerfið?

Verkferlið hjá MMS snýr nær eingöngu að prófgerð á pappír og hefur ekkert breyst eftir að prófin voru færð á tölvutækt form.

## 4. Af hverju er verkferlið enn pappírsmiðað?

Ferlið er pappírsmiðað vegna þess að það koma engir hugbúnaðarsérfræðingar að ferlinu.

**Niðurstaða: Engir hugbúnaðarsérfræðingar koma að ferlinu.**

Eins og má sjá eru einkenni þessara tveggja rekstrarbilana þau sömu og við nánari greiningu má sjá að rót vandamála er sú sama.

## Pappírsmiðað verkferli

Í verkferlalýsingunni fyrir framkvæmd prófanna eru ítarlegar lýsingar á undirbúningi prófanna en þegar kemur að hugbúnaðinum er einungis einn verkþáttur. Í ferlinu virðist einu atriði hafa verið skipt út þegar prófin voru tölvukeyrð og þar sem áður stóð *ljósritun* stendur *staðfesting á virkni tölvukerfis*. Það er því enn pappírsmiðað og hefur ekki tekið mið af breytingum sem hafa átt sér stað í próftökuferlinu og uppsetningarferlinu.

**Ég tel að í verkferlinu hefðu átt að vera verkþættir tengdir því að setja kerfið í álagsþolið umhverfi, vísun í formlegar prófanir þar sem væri a.m.k. listi yfir þá hnökra sem hafa komið upp síðustu ár og síðast en ekki síst álagsprófanir.**

Eins mætti vera listi um fyrirbyggjandi aðgerðir tengdar tölvukerfinu í áhættugreiningunni.

## Verkferlið

Í verkferlalýsingum er aðeins einn verkþáttur sem snýr að virkni kerfisins. Þar þyrfti að vera tilgreindur ábyrgðaraðili, eins áætlaður tími í framkvæmd kerfisins og leið til að fylgjast með framgangi þess.

247	Lokaprófun rafrænna könnunarprófa	50%	02.03.2018	02.03.2018	1d		35 NK
248	Prófun á borðtölvu og fartölvu	100%	02.03.2018	02.03.2018	2h		Skjalprófun, Not
249	Prófun á iPad	100%	02.03.2018	02.03.2018	2h	47	Skjalprófun, Not
250	Prófun á önnur spjöld		02.03.2018	02.03.2018	2h	48	Skjalprófun, Not
251	Prófun á Chromebook		02.03.2018	02.03.2018	2h	49	Skjalprófun, Not
252	Staðfesta virkni kerfis		02.03.2018	02.03.2018		50	
253	Prófdagar		09.03.2018	12.03.2018	3d		

Mynd 6 Hluti úr verkferlalýsingum MMS

Eftir viðtöl við starfsmenn hefur komið fram að farið var yfir virkni kerfisins en þær prófanir voru óformlegar (handahófskenndar). Það er nokkuð ljóst að þessi hluti fékk hverfandi athygli hjá stofnuninni.

### Áhættugreining

Í áhættugreiningunni er aðeins ein áhætta tengd tölvukerfinu og engar fyrirbyggjandi aðgerðir hvað tölvukerfisáhættu varðar. Ekkert mat er lagt á líkur eða alvarleika afleiðinga hugsanlegra bilana. Með það að leiðarljósi að búið er að nota þetta kerfi frá árinu 2016 með góðum árangri má segja að það hafi ekki verið miklar líkur á rekstrarbilun en afleiðingar hugsanlegra bilana eru alvarlegar og í ljósi þeirra hefði verið eðlilegt að leggja meiri áherslu á fyrirbyggjandi aðgerðir til að tryggja rekstrarsamfellu.

Ef þetta gerist:	Þá bregðumst við svona við:	Athugasemdir- hvað þarf að gera? Ráðstafanir til að minnka áhættu:	Ábyrgð:
Veikindi / slys	Hafa varafólk. Vera ávallt með fleiri en einn sem kann verkin.		
Skólagáttarvandamál	Reynum að leysa útfrá <a href="#">Ferli</a> .	verður hér á miðvð og fimmtud. mætir snemma á föstud.	
Fast-Test vandamál	Þá seinkum við prófi.	Gera viðvart.	

Mynd 7 Hluti úr áhættugreiningu MMS

### Engir hugbúnaðarsérfræðingar

Daginn sem ákveðið var að færa samræmd próf yfir á tölvutækt form breyttust aðferðir við próftöku og þetta varð að hluta að hugbúnaðarverkefni. Tækni við að fjölfalda prófin fór úr því að vera ljósrítunarvél í flókin miðstýrð tölvukerfi. Tækni til að dreifa prófunum fór úr höndum kennara í hendur aðkeyþtra tölvunarfræðinga og tækni við að taka prófin úr blýöntum og strokleðrum í tölvur og lyklaborð. Samfara þessari breytingu hefði samsetning starfsfólks átt að breytast og verkferlarnir í takti við nýjar þarfir. Það er mjög mikilvægt að fá sérfræðinga í hugbúnaðargerð og rekstri með innsýn á verkefni sem byggja að stórum hluta á keyrslu á flóknum tölvubúnaði, sérstaklega þar sem afleiðingar bilana eru svo

stórkostlegar. Ljóst er að þá nálgun vantar algjörlega í allan undirbúning og fyrirlögn prófanna.

**Það koma engir hugbúnaðarsérfræðingar að prófakerfinu og það sama á við um stofnunina í heild. Allar nefndir og fagráð sem tengjast MMS eru skipaðar sérfræðingum á menntasviði án aðkomu hugbúnaðarsérfræðinga. Sjá nánar um nefndir og ráð í viðauka.**

## Samantekt

Hér fyrir neðan má sjá samantekt á helstu orsakapáttum bilananna:

- Mjög ógagnsætt og viðkvæmt ferli er hjá ASC til að setja kerfi í álagspolið umhverfi (HVE).
- Bilanagreining var ekki nægilega ítarleg hjá ASC og hefði mögulega verið hægt að greina rót vandans fyrr með ítarlegri greiningu eftir fyrsta prófdaginn.
- Álagsprófanir hjá ASC voru ekki nægilega ítarlegar. Auðkenningarglugginn (pretest screen) var ekki prófaður né uppsetningin þar sem próf voru geymd á mismunandi stöðum í möpputrénu.
- Ferlin hjá MMS í kringum rekstur og umsýslu með prófakerfi eru mjög takmörkuð hvað varðar prófanir á hugbúnaðinum.
- Engar fyrirbyggjandi aðgerðir voru til staðar hjá MMS gegn tæknilegum misskilningi. Engin viðbragðsáætlun var til staðar ef tölvukerfi ASC virkaði ekki sem skyldi.
- Engar formlegar prófanir voru framkvæmdar hjá MMS á lokaútgáfum prófa í kerfi ASC til að ganga úr skugga um að kerfið virkaði sem skyldi.

## Lokaorð

Þrátt fyrir að þessi fyrirlögn á samræmdum könnunarprófum hafi gengið mjög illa og gagnrýna megi ASC fyrir að ná ekki að greina vandamálið strax eftir fyrsta prófdag, þá hefur ASC áralanga reynslu og fjölda stórra viðskiptavina um allan heim. Þeir eru með áreiðanlegt rekstrarumhverfi sem undir eðlilegum kringumstæðum ræður við mikið af samtímanotendum og geta stillt upp umhverfi sem tekur við miklu álagi. Þær villur sem komu upp voru þess eðlis að vinnslugeta skiptir litlu máli. Það er því ekki hægt að gagnrýna MMS fyrir val á samstarfsaðila. Eins er fátt sem MMS hefði getað gert til að koma í veg fyrir þessar villur og ábyrgðin því að langmestu leyti hjá ASC.

Aftur á móti má gagnrýna margt í ferlinu hjá MMS. Ástæðurnar má rekja til þess að þar er lítil þekking á rekstri hugbúnaðarlausna og því hvernig hægt er að fyrirbyggja að svona aðstæður komi upp.



**Álagsprófanir á lokaútgáfu af samræmdu prófunum þremur er sú aðgerð sem hefði lágmarkað áhættuna við fyrirlögn prófanna.**

Það er umdeilanlegt hvort MMS ætti að standa að eigin álagsprófun eða sjá til þess að þær væru framkvæmdar. Í ljósi mikilvægi þess að kerfið ráði við þetta skyndilega álag og hversu afleiðingarnar eru miklar þá ætti MMS að sjá til þess að þetta sé framkvæmt á faglegan hátt.

**ACS gefur reglulega út nýjar útgáfur og það þýðir að MMS verður að framkvæma prófanir fyrir hverja fyrirlögn, þar sem kerfið er ekki undir stöðugri notkun og afleiðingar rekstrarbilunar eru gríðarlegar.**

Það er ekki hægt að setja allt traust á þriðja aðila.

---

Hannes Pétursson



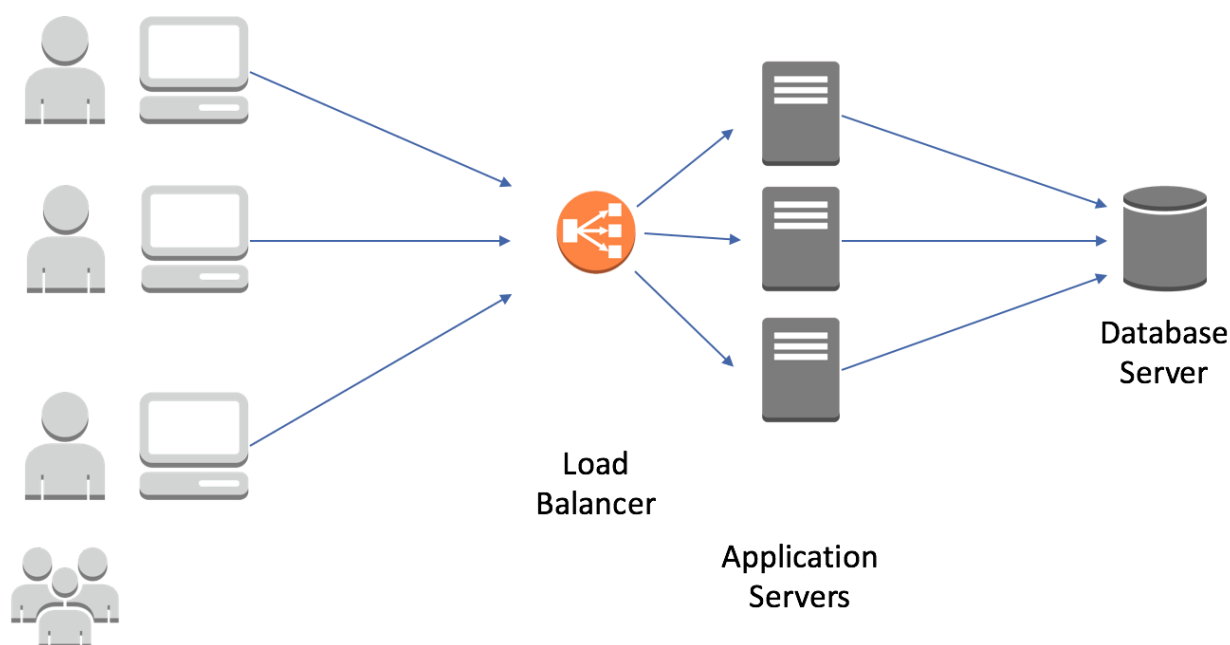
# Viðaukar



## Viðauki 1. Álagspolið umhverfi (“High volume engine”)

ASC hýsir kerfið hjá AWS (<https://aws.amazon.com/>) og er með uppsetningu bæði í Bandaríkunum og á Írlandi. Uppsetning fyrir MMS er hýst á Írlandi. Yfir 100 fyrirtæki og stofnanir nota FastTest kerfið og á síðasta ári voru lögð fyrir um 1.200.000 próf. Stærsta einstaka fyrirlögnin var 35 þúsund próf.

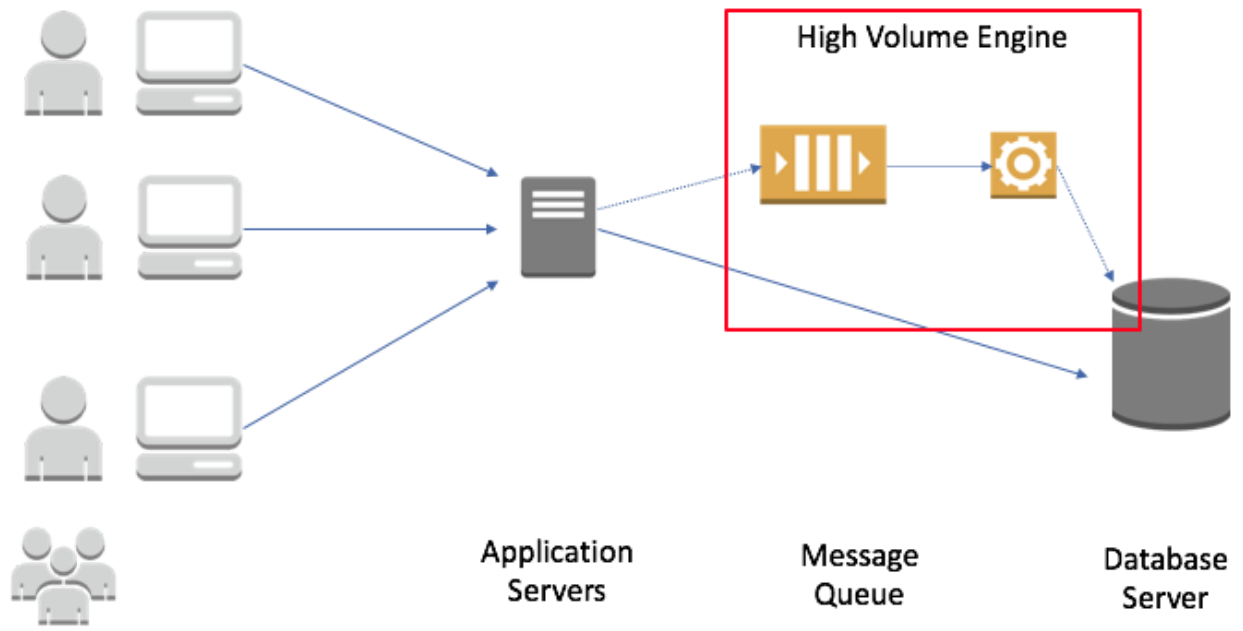
Högun kerfisins er uppbyggð til að geta tekið við miklu álagi og geta þeir bætt við fleiri vefþjónum ef á þarf að halda (Application Servers). Yfir prófdagana voru 8 þjónar að keyra.



Það er verið að nota Amazon Aurora gagnagrunnsþjón (<https://aws.amazon.com/rds/aurora>) sem er hraðvirkur grunnur og byggður til að taka við miklu álagi. Ólíkt vefþjónum þá er ekki möguleiki á að bæta við fleiri vélum til að þola aukið álag og því þarf að stýra álaginu á gagnagrunnsþjóninn.

Til að stýra betur álaginu er boðið upp á að setja próf í álagspolið umhverfi (High volume engine) þegar vitað er að það verður mikið álag á sama tíma eins og tilfelli samræmdu könnunarprófanna.

Það er útfært þannig að allar niðurstöður úr prófum fara í skilaboðabiðröð (Message queue) og þannig er bæði hægt stjórna betur álagi á gagnagrunnsþjóninn og tryggja betur að öll prófasvör séu geymd þó svo að það komi upp tímabundin bilun í gagnagrunni.



## Viðauki 2. (Tölvupóstsamskipti)

Hér eru klippur úr þeim tölvupóstum sem fóru á milli MMS og ASC í aðdraganda prófanna.

### Subject: Tests to be administered next week (1. mars)

MMS:

We will be administering the tests next week, Wednesday through Friday. We will start the tests at 7:30 am local time, or 1:30 am your time.

Is there anything you need to do on your side before we administer the tests. **I think there was something that had to be done last year but nobody is sure about it.**

ASC:

Hello! Hope all is well! Before you create test codes, We need to mark tests as high volume. **Please send names of test sessions and we will do this ASAP**

### Subject :iPad problems with the tests (2. mars)

MMS:

I have also discovered that I cannot close those test manually. (mark tests as completed) I get an error:

Invalid data: Unexpected end of JSON input

**These tests are -> test scheduler -> 2018 -> SKP\_9 -> ENS | ISL | STF each has 10 test students on them.**

ASC:

These new test sessions are not set to deliver in our High Volume test engine. **I have converting all of them to do so.**

This should prevent the error from ever happening, because the answers will be **recorded asynchronously**. Reentry of the test will now not cause any issues

MMS:

And everything on the high volume test engine? If so we will early Monday input all examinees

ASC:

We found and fixed one other problem. A couple of your subscores were set to GRM scoring instead of Dichotomous IRT and/or GPCM.

### Subject: :We are go, go for lunch (5. Mars)

MMS:

I just wanted to let you know that we have launched the codes into the cosmos. From our end, Everything is ready.

We have had minor funky issues with the weblock app that we have bypassed. I will send you more detailed report about that after the assessments.

Otherwise, we are doing well and everything is 10 by 10.

ASC:

We'll keep our eyes on things here. Thanks,

## Viðauki 3 (nefndir og ráð í tengslum við Menntamálastofnun)

Forstjóri hefur sér til ráðgjafar sjö manna ráðgjafanefnd sem ráðherra skipar til fjögurra ára í senn. Í henni eru:

- Þórólfur Þórlindsson
- Arna Hauksdóttir
- Svandís Ingimundardóttir
- Guðrún Hrefna Guðmundsdóttir
- Jón Torfi Jónasson
- Aðalheiður Steingrímsdóttir
- Sigrún Edda Eðvarðsdóttir

Svo eru starfandi eftirfarandi fagrað (skv. lögum, forstjóri setur á fót)

- Náms- og gæðamati
- Gerð og miðlun námsgagna
- Upplýsingagjafar og þjónustu

Í Náms- og gæðamati eru:

- Anna Lind Pétursdóttir
- Fjóla Þorvaldsdóttir
- Guðbjörg Ragnarsdóttir
- Jón Pétur Zimsen

- Þóroddur Helgason

Ráðherra skipaði svo sérfræðihóp um samræmd könnunarpróf í júní 2017 en í honum eiga sæti:

- Amalía Björnsdóttir sem er formaður
- Ingileif Ástvaldsdóttir tilnefnd af KÍ
- Vigfús Hallgrímsson tilnefndur af SÍS

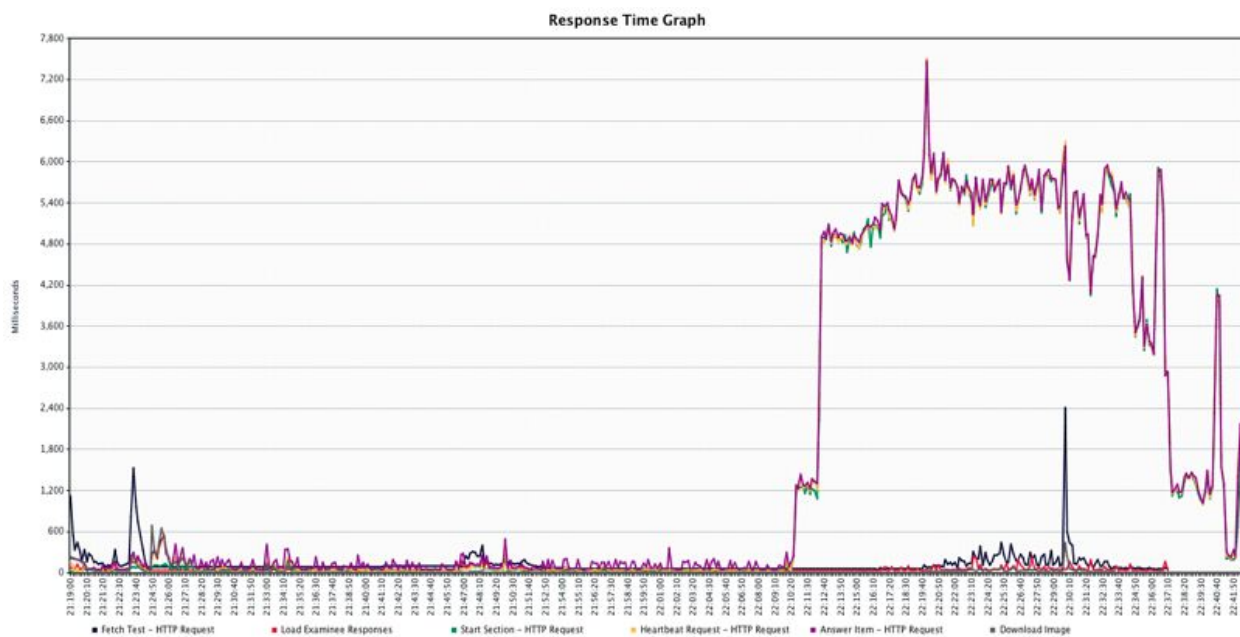
Ekkert af þessu fólki hefur menntun eða reynslu í þróun eða rekstri hugbúnaðarkerfa.

## Viðauki 4. Álagspróf

Hér má sjá niðurstöður úr álagsprófi sem gert var af ASC þar sem 12 þúsund samtímanotendur framkvæmdu aðgerðir í kerfinu.

ATH. Þarna var ekki verið að hlaða upp auðkenningarglugga (Pretest screen).

ATH. Þarna eru prófin staðsett í rótarmöppu og valda því ekki *slæmri fyrirspurn* (Bad query).



Þrátt fyrir lengri svartími undir lokin þá voru allar fyrirspurnir árangursríkar.