



Pekko Ahtiainen

Confluencen klusterointi Azure cloudiin

Metropolia Ammattikorkeakoulu

Insinööri (AMK)

Tieto- ja viestintätekniiikan tutkinto-ohjelma

Insinöörityö

13.2.2022

Tiivistelmä

Tekijä: Pekko Ahtiainen
Otsikko: Confluencen klusterointi Azure cloudiin
Sivumäärä: 31 sivua + 3 liitettä
Aika: 28.3.2022

Tutkinto: Insinööri (AMK)
Tutkinto-ohjelma: Tieto- ja viestintätekniikan tutkinto-ohjelma
Ammatillinen pääaine: Verkot ja pilvipalvelut
Ohjaajat: Ohjaava opettaja Janne Salonen
Tilaaajan edustaja Erri Javola

Insinööriyössä selvitettiin liiketoiminnalle kriittiseksi muuttuneen dokumentointijärjestelmän saatavuuden ja tietojen säilyvyyden varmistamisen keinoja. Lisäksi käydään läpi löydettyjen keinojen käyttöönotto.

Yritykselleärkevimmäksi vaihtoehdoksi todettiin Confluencen Data Center-hinnoittelumallin käyttöönotto Microsoftin Azure-pilvipalvelussa.

Dokumentointijärjestelmälle suunniteltiin ja rakennettiin uusi ympäristö, sekä siirrettiin dokumentointijärjestelmä uuteen ympäristöön. Suunnittelussa kiinnitettiin erityistä huomiota Microsoft Azuren tarjoamiin tieton ja saatavuuden varmennuskeinoihin.

Ympäristön suunnitelma ja rakentaminen sekä käyttöönotto onnistuivat. Ympäristö on toiminut virheettömästi. Käyttäjän erehdyksessä pysyvästi poistama tieto on onnistuneesti palautettu varmuuskopiosta.

Työssä tehty selvitystyö on kaikinensa dokumentoitu kyseiseen dokumentaatiojärjestelmään ja sitä voidaan käyttää järjestelmän ylläpitoon, virheenkorjaukseen ja jatkokkehitykseen.

Avainsanat: Confluence, Azure, Pilvipalvelu

Abstract

Author: Pekko Ahtiainen
Title: Clustering Confluence to Azure cloud
Number of Pages: 31 pages + 3 appendices
Date: 28 March 2022

Degree: Bachelor of Engineering
Degree Programme: Information Technology
Professional Major: IoT and Cloud Computing
Instructors: Janne Salonen, Directing teacher
Erri Javola, Employer's representative

The purpose of this bachelor's thesis was to investigate ways to ensure the availability and data preservation of a documentation system that had become mission critical. Implementation of the identified means is also part of this thesis.

Implementation of Confluence's Data Center pricing model in Microsoft's Azure cloud was identified as the most sensible option for the company.

New environment was designed and build, and the documentation system was migrated to the new environment. The design paid particular attention to availability and data durability capabilities provided by Microsoft Azure.

The environment was successfully designed, built, and deployed. The environment has worked flawlessly ever since. Data permanently deleted by a user has been successfully restored from backup.

The work carried out in this thesis has been documented in the documentation system and can be used for system maintenance, debugging and further development.

Keywords: Confluence, Azure, Cloud computing

Sisällys

Lyhenteet

1	Johdanto	1
1.1	Taustaa	1
1.2	Tavoitteet	2
2	Tarpeen tunnistaminen	2
2.1	Ympäristön haavoittuvuus	2
2.2	Tarve	3
2.3	Ratkaisu	3
3	Vaihtoehtojen selvitys	3
3.1	Vaihtoehto 1: Kuormanjakoklusteri (load balancing cluster)	4
3.2	Vaihtoehto 2: Fail over –klusteri	6
3.3	Vaihtoehto 3: Confluence Cloud	7
3.4	Vaihtoehto 4: Private Datacenter omassa konesalissa	8
3.5	Vaihtoehto 5: Private Datacenter pilvipalvelussa	9
3.6	Kustannusten arviointi	11
4	Käynnistyslupa	17
5	Projektin suunnittelu	17
5.1	Tietoliikennevaatimukset	18
5.2	Sähköpostipalvelin	18
5.3	Tunnistautumispalvelu	18
5.4	Tietokanta	18
5.5	Tiedostopalvelin	18
5.6	Nodet	19
5.7	Tietoturva	19
5.8	Tietosuoja	19
5.9	Varmuuskopiointi	19
6	Testiympäristön rakentaminen	19
7	Testaus	20

8	Tuotantoympäristön määrittely	20
8.1	Tietoliikenne	21
8.2	Sähköposti	21
8.3	Tunnistautuminen	22
8.4	Tietokantapalvelin	22
8.5	Tiedostopalvelin	25
8.6	Nodet	26
8.7	Tietoturva	26
8.8	Tietosuoja	28
8.9	Varmuuskopiointi	28
9	Käyttöönottosuunnitelma	29
10	Käyttöönotto	30
11	Virheiden käsittely	311
12	Opit	31
	Lähteet	32

Liitteet

Liite 1: Verstas Confluencen klusterointiehdotus Azure-cloudiin

Liite 2: Varmuuskopiosta palauttaminen

Liite 3: Käyttöönottoviesti

Lyhenteet

Confluence Australialaisen Atlassian-ohjelmistoyrityksen kehittämä ja markkinoima organisaatiowikiohjelmisto. Se on verkkosivusto, jonka sisältöä käyttäjät voivat itse muokata webselaimellaan, mikäli sivun luoja on antanut heille siihen oikeudet. Confluencessa on kiinnitetty huomiota erityisesti yrityskäytössä tarvittaviin ominaisuuksiin, kuten käyttöoikeuksien hallintaan ja integroitavuuteen. (Wikipedia, 2019)

Verstas Samlinkin Confluencen sisäinen nimi

Wiki Verkkosivusto, jonka sisältöä käyttäjät voivat itse muokata haluamallaan tavalla, usein ilman sivustolle kirjautumista. Vuorovaikutteisuus ja muutosten tekemisen yksinkertaisuus tekevät wikistä tehokkaan yhteisöllisen kirjoittamisen työkalun.

Azure Microsoftin pilvipalvelu

Pilvipalvelu Tietoteknisen infrastruktuurin tai datakeskusten omistamisen sijaan vuokraamista pilvipalvelun tarjoajalta.

Node Nimitys palvelimelle, joka on identtinen n-kappaletta muun palvelimen kanssa. Noden tarkoitus on tasata kuormaa ja estää käyttökatoja virhetilanteissa.

DTU Database Transaction Unit. Keskenään vertailukelpoinen yksikkö, jolla mitataan tietokannan suorituskykyä.

Active Directory (AD) Microsoftin Windows-toimialueen käyttäjätietokanta ja hakemistopalvelu, joka sisältää tietoa käyttäjistä, tietokoneista ja verkon resursseista.

SAML (engl. Security Assertion Markup Language) on XML-standardi tietojärjestelmien käyttäjien tunnistamiseen ja valtuuttamiseen liittyvien tietojen jakamiseen tietoverkossa.

Application Gateway Azuren versio kuormantasaajasta ja verkkoliikenteen ohjaajasta nodeille.

Resource Group Ryhmä, jonka alle on kasattu kaikki ratkaisuun tarvittavat resurssit Azuressa.

1 Johdanto

1.1 Taustaa

Samlink on suomalainen, kokenut finanssialan ratkaisujen kehittäjä ja palveluntuottaja. Cognizant on Yhdysvalloissa pääkonttoria pitävä monikansallinen IT-palveluita tarjoava yritys. Cognizantilla on n. 300 000 työntekijää, joista suunnilleen puolet Intiassa. Cognizant osti Samlinkin vuonna 2019.

Confluence on Australialaisen Atlassianin kehittämä yrityskäyttöön suunniteltu wikipohjainen dokumentointialusta. Se on verkkosivusto, jonka sisältöä käyttäjät voivat itse muokata webiselaimellaan, mikäli sivun luoja on antanut heille siihen oikeudet. Confluencessa on kiinnitetty huomiota erityisesti yrityskäytössä tarvittaviin ominaisuuksiin, kuten käyttöoikeuksien hallintaan ja integroitavuuteen

Confluencen käyttö lähti laajenemaan Samlinkissa, kun Jyväskylän ja Espoon toimistojen instanssit yhdistettiin vuonna 2014.

Yhdistämisen yhteydessä Confluence määritettiin sovelluskehityksen työkaluksi. Sinne oli tarkoitus tallettaa projektinaikainen dokumentaatio (pois lukien hallinnollinen dokumentaatio), järjestelmien ja ympäristöjen kuvaukset, toimintamallin ohjeistus ja tiimien työdokumentaatio.

Käyttö laajeni nopeasti ja orgaanisesti kattamaan projektien hallinnollisen dokumentaation, sovellusdokumentaation, asiakkaille tuotettavan tarjousmateriaalin, roolikohtaisen uutisoinnin, kysymys- ja vastauspalstan, työntekijäkokemuksen kehitykseen tarkoitettun ideointisivuston jne.

Vuonna 2017 alkoi olla selvää, että Confluencesta on tullut Samlinkille liiketoimintakriittinen järjestelmä, jonka käyttökatkot aiheuttavat yritykselle taloudellisia menetyksiä ja pahimmassa tapauksessa asiakasluottamuksen heikkenemistä.

1.2 Tavoitteet

Tämän opinnäytetyön tavoitteena on kuvata, kuinka Samlinkin Confluencen saatavuutta ja suorituskykyä parannettiin kahdentamalla ja käyttämällä Azure-cloudin tarjoamia varmennettuja palveluita.

Työssä käsitellään pilvipalvelun käyttöönottoprojektin eri vaiheet esiselvityksestä tuotantokäyttöönottoon ja vianselvitykseen. Työssä sivutaan myös sisäisen projektin resursointi- ja priorisointiongelmia.

Työstä jää Samlinkille lopputulokseksi järjestelmädokumentaatio ja projektikertomus Confluence sivustoina. Kyseiset sivustot ovat tämän opinnäytetyön pääasiallinen lähde.

2 Tarpeen tunnistaminen

2.1 Ympäristön haavoittuvuus

Confluence oli asennettu tuotantotestipalvelimelle yhdeksi palveluksi monien joukossa. Ongelmia alkoi ilmaantua vuonna 2016. Käyttökatkot olivat yksittäisiä. Syyt ovat moninaisia, mutta useimmiten ne liittyivät palvelimen päivityksiin. Loppuvuodesta 2016 Confluencen prosessi alkoi viedä 25–75% palvelimen prosessoriajasta. Prosessoriaikaongelma saatiin korjattua tammikuussa 2017 asentamalla uusi versio Confluencesta. Tässä versiossa oli keskitytty erityisesti suorituskyvyn parantamiseen. Myöhemmin Confluence palvelu lopetti automaattisen käynnistymisen palvelimen käynnistymisen yhteydessä. Jokainen uudelleen käynnistystä vaativa päivitys palvelimella aiheutti siis Confluencen käyttökatkon. Nämä käyttökatkot alkoivat aamuyöstä ja jatkuivat, kunnes ylläpitäjä käynnisti Confluencen käsin. Confluence saatiin lopulta käynnistymään automaattisesti siirtämällä sen automaattinen käynnistys tapahtumaan muiden käynnistystehtävien jälkeen.

Suurin ongelma oli kuitenkin varmistus. Ainoa varmistus oli päivittäin otettava levykuva palvelimesta. Se sisälsi sovelluksen, liitteet ja tietokannan. Levykuvan palauttamista ei ollut koskaan testattu.

2.2 Tarve

Confluencen käytön lisääntyminen, sinne tallennetun tiedon määrä ja kriittisyys, sekä ympäristön ongelmat ja puutteet asettivat selkeän tarpeen tilanteen parantamiseksi. Confluencen tulee olla saatavilla jatkuvasti ja mahdollisten käyttökatkojen tulee olla äärimmäisen harvinaisia ja niiden täytyy pysyä lyhyinä. Tätäkin tärkeämpää on tiedon säilyvyyden varmistaminen.

2.3 Ratkaisu

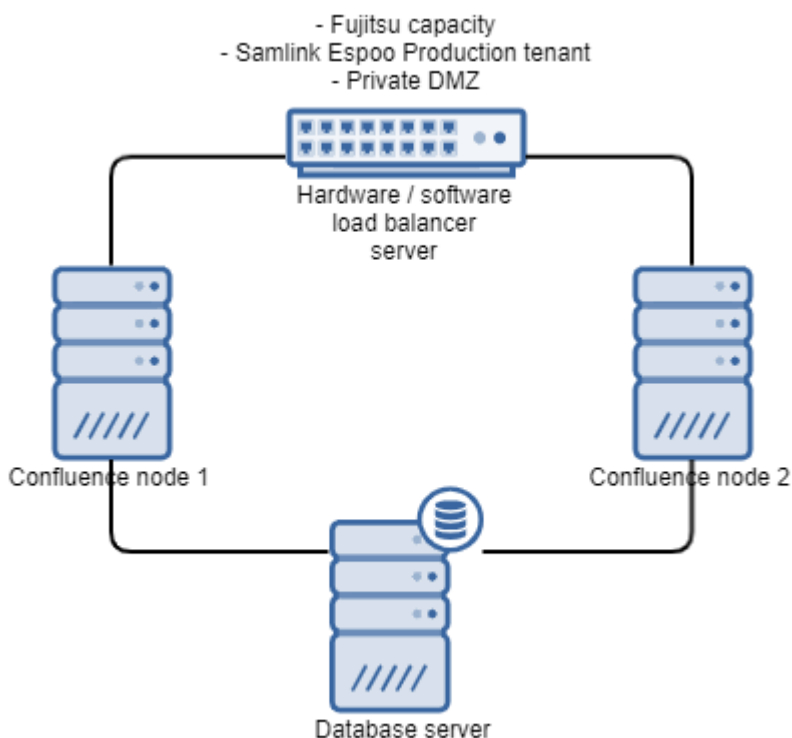
Palvelun saatavuutta voidaan parantaa kahdentamalla sen osat. Confluence päätettiin klusteroida.

Toimin tässä työssä tuotepäällikön, pääkäyttäjän, järjestelmäasiantuntijan ja projektipäällikön rooleissa. Lisäksi suunnitteluvaiheeseen osallistuivat Timo Holm ja Pasi Pekkanen. Määrittely- ja toteutusvaiheeseen osallistuivat lisäksi: Antti Häkli, joka rakensi määrittelemämme ympäristön, Jarkko Timonen Samlinkin Azuren pääarkkitehtina, Anssi Rajaniemi tietoliikennevastaavana, Keijo Korte pääsynhallinnan asiantuntijana, Pasi Sainio Linux arkkitehtina, Teijo Raatikainen ja Sami Marjokorpi Linux ja DNS asiantuntijoina, Ari Kaatonen ja Mikko Källroos tietokanta-asiantuntijoina. Lisäksi erinäisiä asiantuntijoita osallistui Confluencen ja muiden järjestelmien välisien ongelmien ratkaisuun.

3 Vaihtoehtojen selvitys

Teimme kattavan selvityksen vaihtoehtoista ja niiden kustannuksista. Eri vaihtoehtoja oli aluksi apuna selvittämässä tiimikaverini Timo Holm ja loppuvaiheessa konsultti Pasi Pekkanen Contribytelta. Kustannuslaskenta ja työmääräarviot ovat kokonaisuudessaan minun tekemiäni.

3.1 Vaihtoehto 1: Kuormanjakoklusteri (load balancing cluster)



Kuva 1 Verkkonäkökuorma kuormanjakoklusterista

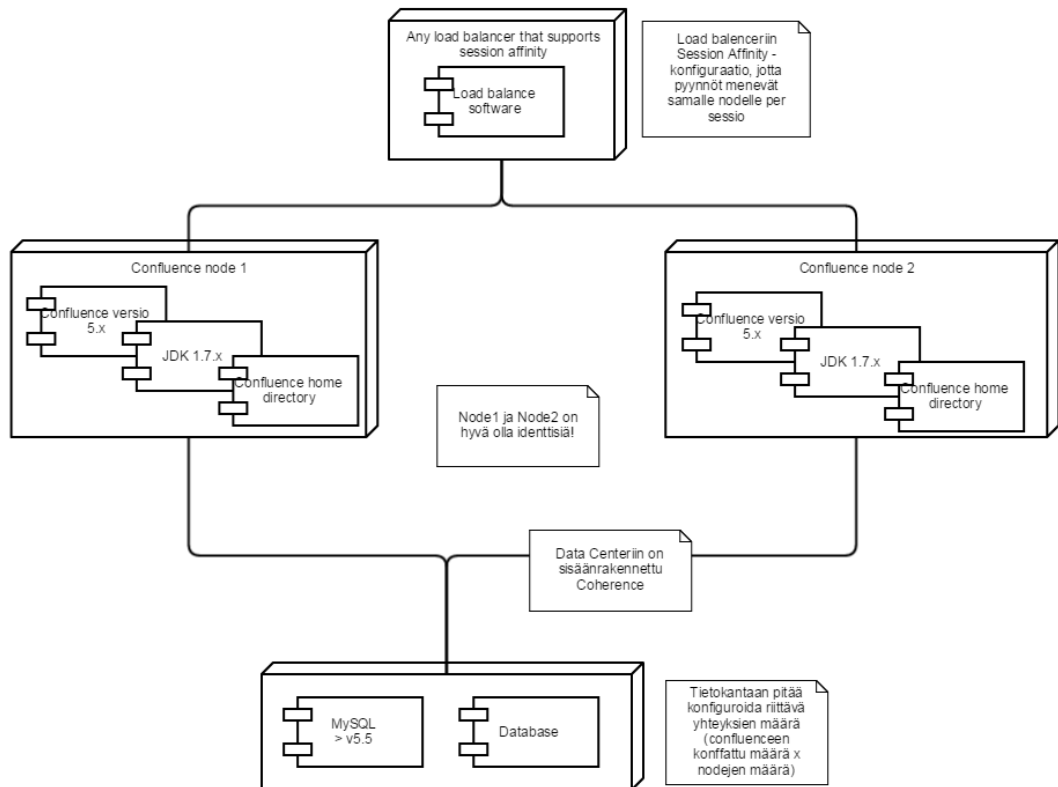
Tässä klusterissa on kaksi aktiivista Confluence-palvelinta, johon käyttäjän istuntokohtaiset pyynnöt jaetaan kuormanjakajan toimesta. Tietokanta vietään erilliseen tietokantaklusteriin (SQL Server 2014 tai mahdollisesti perustettava Linux palveluiden yhteinen tietokantaklusteri).

3.1.1 Hyödyt

- Parempi suorituskyky, jos käyttäjiä on paljon.

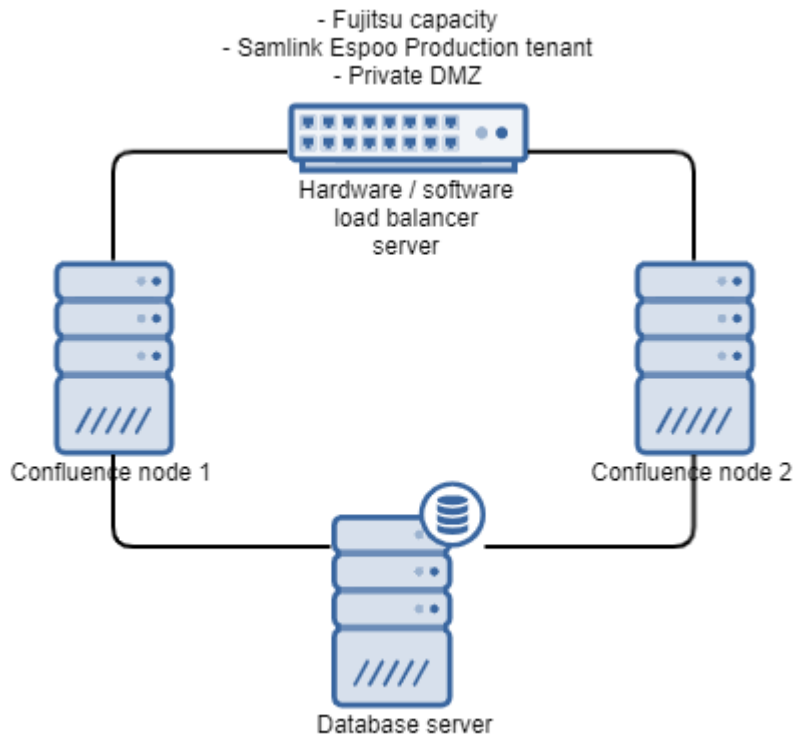
3.1.2 Haitat

- Vaatii Confluencen ja lisäosien tuplalisensoinnin ja 2 identtistä palvelinta.



Kuva 2 Sijoittelukaavio kuormanjakoklusterista

3.2 Vaihtoehto 2: Fail over –klusteri



Kuva 3 Verkkonäkökulma Fail Over –klusterista

Vaihtoehto 2.1: Tässä klusterissa on kaksi aktiivista Confluence-palvelinta, johon pyynnöt viedään vain yhdelle palvelimelle kerrallaan. Mikäli palvelin vikaantuu, pyynnöt ohjataan toiselle palvelimelle.

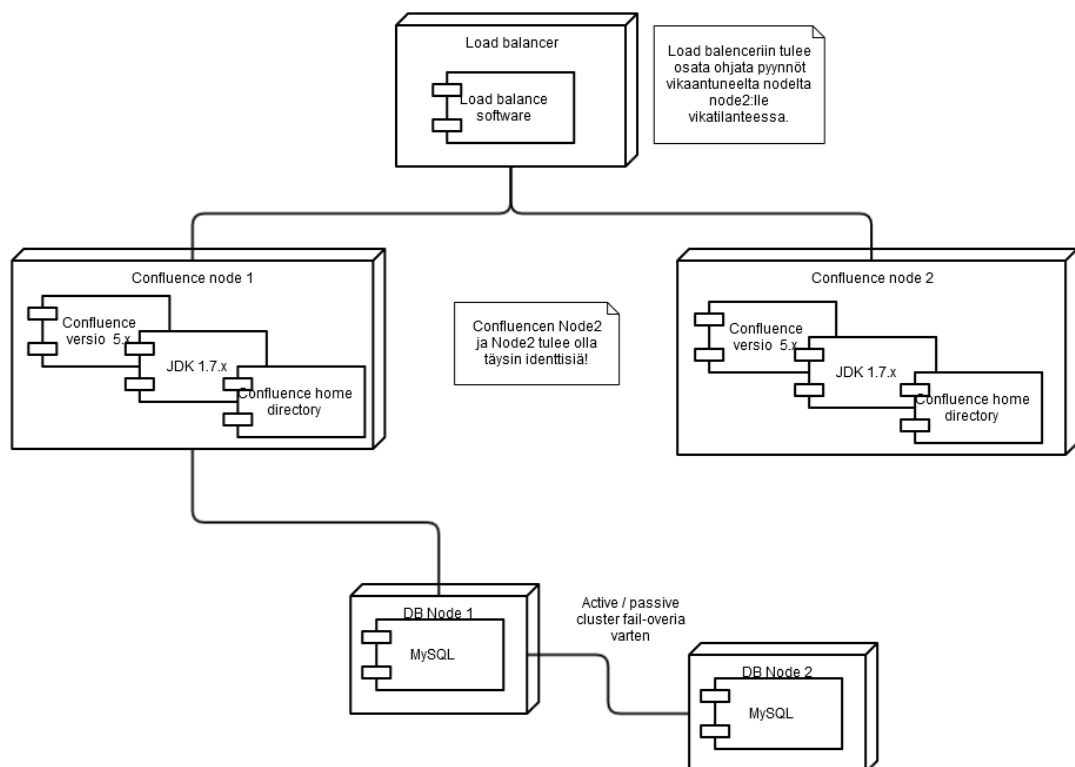
Vaihtoehto 2.2: Tässä klusterissa on kaksi Confluence-palvelinta, johon pyynnöt viedään vain yhdelle palvelimelle kerrallaan. Toinen on aktiivinen, toinen passiivinen. Mikäli aktiivinen palvelin vikaantuu, toinen nostetaan käsin tai automaattisesti aktiiviseksi ja pyynnöt ohjataan toiselle palvelimelle.

3.2.1 Hyödyt

- Palvelun saatavuus on parempi kuin nykytilassa.

3.2.2 Haitat

- Vaatii Confluencen ja lisäosien tuplalisensoinnin.
- Vaatii kaksi käynnissä olevaa palvelinta, jolloin kustannussäästö kuituu (Fail Over palvelin voi olla kapasiteetiltaan pienempi, josta saadaan jotain säästöä).
- Palvelimen Lucene indeksi tulee uudelleenrakentaa, jos palvelin on ollut alhaalla yli 3 tuntia. Lucene indeksin rakentaminen kestää kymmeniä minutteja (mahdollisesti jopa tunti), joten inaktiivisen palvelimen käynnistyminen on hidasta ja katko-aika on pitkä.



Kuva 4 Sijoittelukaavio Fail Over klusterista

3.3 Vaihtoehto 3: Confluence Cloud

Atlassianin tarjoama pilvipalvelu.

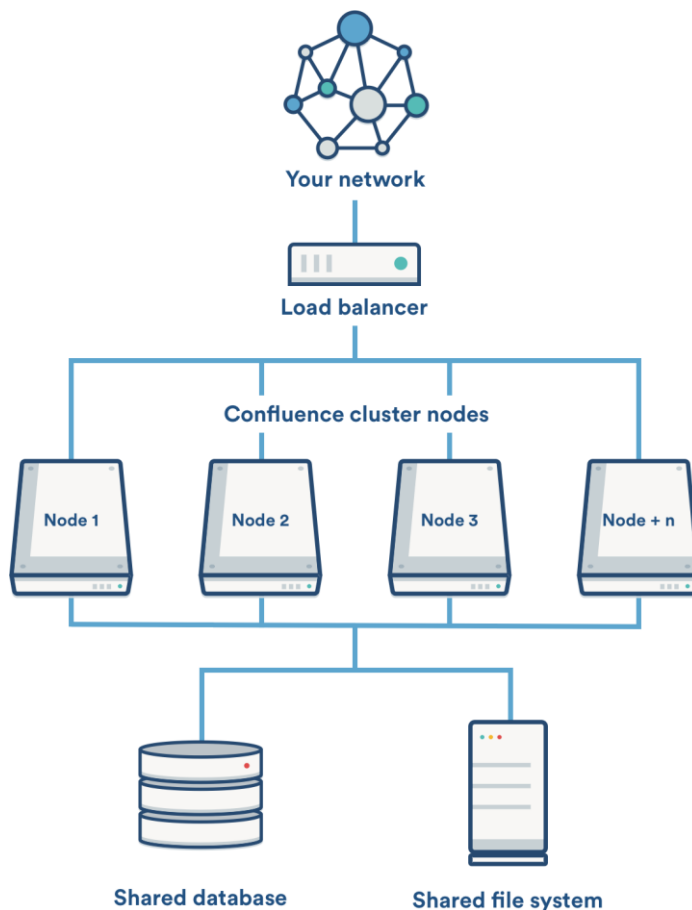
3.3.1 Hyödyt

- Ylläpito ulkoistettu. Aina uusien versio.

3.3.2 Haitat

- Versio ja ominaisuudet päivittyvät kysymättä.
- Tietojen ei luvata pysyvän EU:n sisällä.
- Suuret lisenssikustannukset etenkin lisäosien osalta.
- Kustomointimahdollisuudet ovat pienet.
- Ideation (Kehitysehdotukset), Lockpoint (liitetiedostojen lukitus) ja Balsamiq (mallikuvien piirtäminen) eivät ole saatavilla Cloud versioon.

3.4 Vaihtoehto 4: Private Datacenter omassa konesalissa



Kuva 5 Verkkonäkökuorma Private Datacenter mallista

Tässä klusterissa on kaksi tai useampi aktiivista Confluence-palvelinta (node), joihin käyttäjän istuntokohtaiset pyynnöt jaetaan kuormanjakajan toimesta, sekä

lisäksi erillinen yhteinen tiedostopalvelin. Tietokanta viedään erilliseen tietokantaklusteriin (SQL Server 2014 tai mahdollisesti perustettava Linux palveluiden yhteinen tietokantaklusteri).

3.4.1 Hyödyt

- Mahdollistaa sekä kuormanjaon, että Fail Overin samalla lisenssikustannuksella.

3.4.2 Haitat

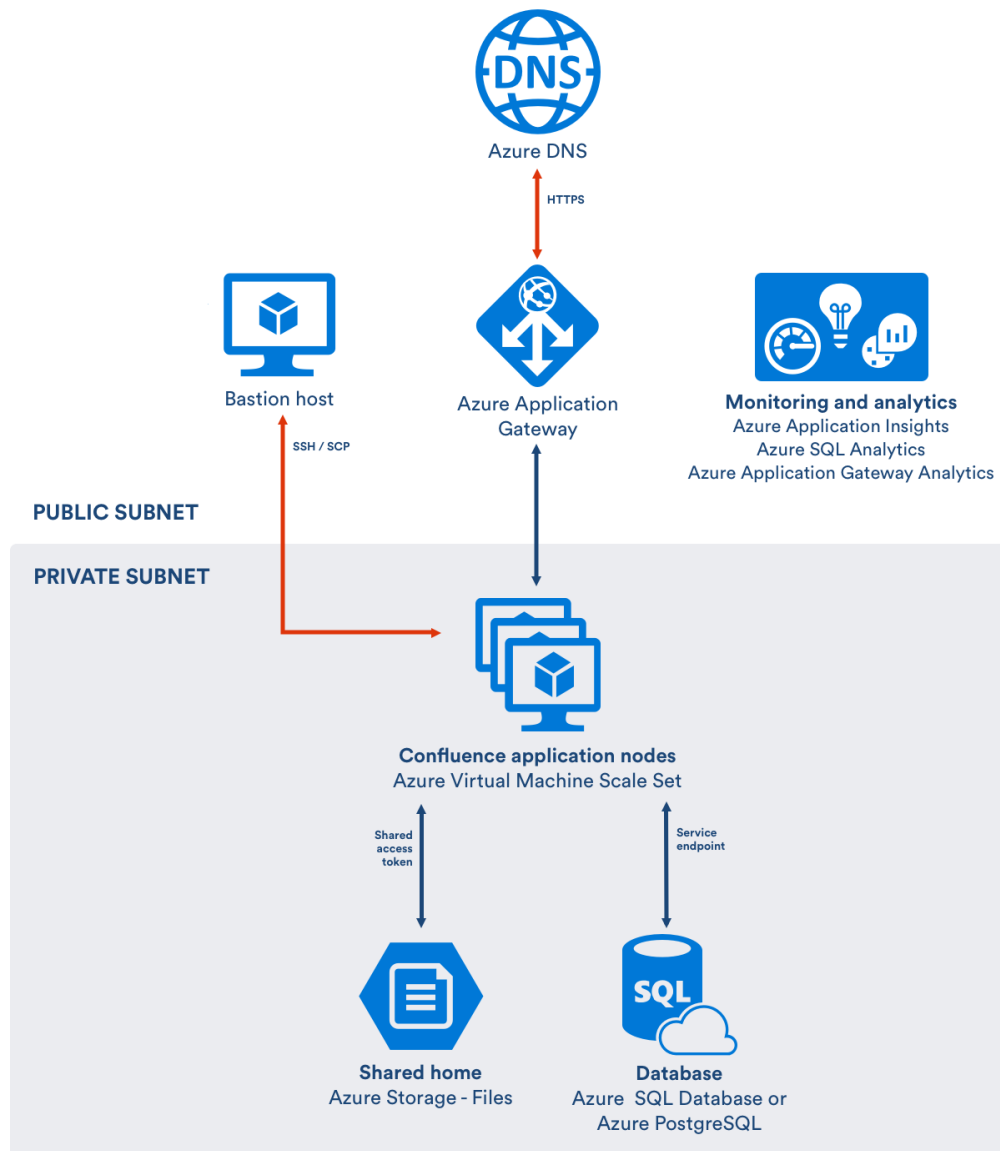
- Vaatii 2 identtistä palvelinta, tiedostopalvelimen ja tietokantapalvelimen.

3.5 Vaihtoehto 5: Private Datacenter pilvipalvelussa

Pilvipalvelut tulivat mukaan, kun alettiin selvittää klusteroinnin palvelinkustannuksia. Ne nousevat massiivisiksi, jos joudutaan kahdentamaan sovelluspalvelinten lisäksi tietokantapalvelin ja tiedostopalvelin. Pilvipalvelu vaihtoehtoina harkittiin Amazon Web Serviceä (AWS), sekä Microsoftin Azurea. Azure valittiin, koska siihen ympäristöön löytyi Samlinkista parempi osaaminen.

Pilvipalvelut tarjoavat valmiiksi varmennettuja palveluja kuormantasaajan, tiedostopalvelimen ja tietokantojen osalta. Näiden palvelujen käyttö on huomattavasti halvempaa, kuin omien rakentaminen.

Datacenter ympäristö rakennetaan konsultin avustuksella siten, että ylläpito-osaaminen jää Samlinkin asiantuntijoille.



Kuva 6 Private Datacenter Azure-pilvessä verkkoäkökulma (Atlassian, 2021)

3.5.1 Hyödyt

- Mahdollistaa sekä kuormanjaon, että Fail Overin samalla lisenssikustannuksella.
- Malli on dokumentoitu kattavasti Atlassianin toimesta. Valmis implementointimalli.

- Samlinkin infra ei rajoita implementointimahdollisuuksia, vaan voidaan toimia suositusten mukaan.
- Palvelinkustannukset ovat IBM:n tai Fujitsun kapasiteettia pienemmät.
- Nodet voidaan jakaa fyysisesti eri sijainteihin.
- Azureen on olemassa Express Gateway, joka mahdollistaa monien asioiden toiminnan kuten sisäverkossa.
- Azure osaaminen lisääntyy

3.5.2 Haitat

- Pilviosaaminen Samlinkissa vähäistä.

3.6 Kustannusten arviointi

Kustannuksia arvioitaessa huomiotiin pääasiallisesti jatkuvat palvelin- ja lisenssikustannukset. Myös projektin kustannuksia arvioitiin karkealla tasolla.

3.6.1 Lisensointikustannukset

Confluence vaatii 500 käyttäjän lisenssin jokaista ajossa olevaa Confluence-instanssia kohden, eli kahden aktiivisen palvelinklusteroinnin tapauksessa pitäisi ostaa 2 kpl 500 hengen Confluence-lisenssejä, yhteensä 12 000 dollaria alkavalta vuodelta. Lisäksi vaihtoehdossa 2 tulee ostaa kaikkien lisäosien lisenssit molempiin aktiivisiin instansseihin. Jos käyttäjämäärää tai lisäosien määrää joudutaan lisäämään, tulee hintaero Data Center vaihtoehtoon vain kasvamaan.

Alla esitetään alkuperäiset laskelmat. Tuotantoon otettiin kuitenkin 1000 hengen lisenssit, jotta käyttäjätunnuksia riittää myös Samlinkin projekteihin osallistuville Cognizantin Offshoren henkilöille.

Taulukko 1 Lisensointikustannukset eri vaihtoehtoissa

	Nykykus- tannus	Vaihtoeh- dot 1 & 2	Vaihto- ehto 3	Vaihtoeh- dot 4 & 5
Confluence	\$ 9 650	\$ 19 300	\$ 15 600	\$ 12 000
Calendars	\$ 1 565	\$ 3 130	\$ 7 800	\$ 1 250
Questions	\$ 4 815	\$ 9 630	\$ 7 800	\$ 3 850
Gliffy	\$ 3 125	\$ 6 250	\$ 12 300	\$ 5 000
Balsamiq	\$ 2 000	\$ 4 000		\$ 1 000
Ideation	\$ 750	\$ 1 500		\$ 750
Scroll PDF	\$ 1 950	\$ 3 900	\$ 2 700	\$ 1 950
Lockpoint	\$ 675	\$ 1 350		\$ 675
Yhteensä \$	\$ 24 530	\$ 49 060	\$ 46 200	\$ 26 475
Yhteensä €	21 765 €	43 531 €	40 993 €	23 491 €

Huomaa, että Cloud-versiosta puuttuu kolmen lisäosan kustannus, koska niitä ei ole saatavilla Cloud-versioon.

3.6.2 Palvelinkustannukset

Taulukko 2 Palvelinkustannukset eri vaihtoehdoilla

	Nykykus- tannus	Vaihto- ehto 1	Vaihto- ehto 2	Vaihto- ehto 3	Vaihto- ehto 4	Vaihto- ehto 5
Nodet	400 €	1 090 €	890 €	- €	1 090 €	400 €
Tiedosto- palvelin		- €	- €	- €	330 €	50 €
Tietokanta- palvelin		- €	- €	- €	- €	200 €
Yht / kk	400 €	1 100 €	1 100 €	- €	1 400 €	650 €
Yht / vuosi	4 800 €	13 080 €	10 680 €	- €	17 000 €	7 800 €

Kuormanjako (1) / Fail Over (2) vaihtoehdossa ei ole erillisiä tietokanta- ja tiedostopalvelimia, vaan ne on rakennettu samoille palvelimille Confluencen kanssa. Private Data Center omassa konesalissa (4) on ajateltu tietokannan rakennettavan Samlinkin Microsoft SQL tietokantaklusteriin, joten siitä ei tulisi lisäkustannuksia.

Palvelinkustannusten arvioitiin seuraavien ominaisuuksien pohjalta:

- 2 x Confluence-node + kuormanjakaja:
 - 4 vCPU, 16 Gt Ram, 100 Gt HDD
- 1 x MySQL-kantapalvelin
 - vCPU, 8 Gt Ram, 200 Gt HDD

- 1 x tiedostopalvelin
 - vCPU, 8 Gt Ram, 200 Gt HDD(?)

3.6.3 Kokonaisvuosikustannukset

Taulukko 3 Kokonaiskustannukset eri vaihtoehtoilla

	Nykykus- tannus	Vaihtoeh- dot 1 & 2	Vaihto- ehto 3	Vaihto- ehto 4	Vaihto- ehto 5
Palvelinkus- tannus	4 800 €	13 100 €	- €	17 000 €	7 800 €
Lisenssikus- tannus	21 800 €	43 600 €	40 100 €	23 500 €	23 500 €
Yhteensä	26 600 €	56 700 €	40 100 €	40 500 €	31 300 €
Ero	- €	30 100 €	13 500 €	13 900 €	4 700 €

Taulukosta on selkeästi nähtävissä Data Center Azuren (5) edut.

3.6.4 Investointikustannukset

Investointikustannuksia syntyy kaikissa vaihtoehtoissa lähinnä sisäisestä ja ulkoisesta työstä. Lisäksi Kuormanjako (1) / Fail Over (2) vaihtoehdossa joudutaan ensimmäisenä vuonna maksamaan 21 800€ ylimääräisiä lisensointikustannuksia.

Taulukko 4 Investointikustannukset

	Nykykus- tannus	Vaihtoehto 1 & 2	Vaihtoehto 3	Vaihtoehto 4	Vaihtoehto 5
Lisenssit	- €	21 800 €	- €	- €	- €
Palvelimet	- €	3 000 €	3 000 €	- €	- €
Konsulttityö	4 100 €	8 300 €	8 300 €	8 300 €	8 300 €
Yhteensä	4 100 €	33 100 €	11 300 €	8 300 €	8 300 €

Taulukko 5 Työmääräarvio

	Sisäinen htp	Ulkoinen htp
Sovellustuki		2
Azure	3	
Linux	2	
Windows	1	
Tietokanta	1	
Pääsynhallinta	1	
Tietoliikenne	2	

Testiympäristön rakentaminen	2	2
Tuotantoympäristön rakentaminen	2	2
Käyttöönotto	2	2
Testaus ja virheen korjaus	20	
Projektin johtaminen	4	
Yhteensä	40 htp	8 htp

3.6.5 Sopivan vaihtoehdon valinta

Fail Over (2) hylättiin nopeasti, kun tajuttiin, kuinka pitkiksi käyttökätkot venyisivät indeksin rakentamisen takia.

Load Balancing (1) hylättiin korkeiden lisenssikulujen takia. Lisäsyynä oli se, että Data Center malli on käytännössä parannettu versio kuormantasauksesta.

Atlassianin tarjoama Confluence Cloud (3) hylättiin tietosuoja- ja lisäosapuutteiden sekä korkeiden lisenssihintojen takia.

Data Center Azure (5) oli lopulta helppo valinta. Se oli selvästi halvin vaihtoehto. Iso osa palveluista tarjotaan valmiina ja niiden käyttövarmuus on tasolla, mihin ei itse rakentamalla helposti päästä. Palvelun haavoittuvin osa lienee ”Express Route” Samlinkin konosalin ja Azure Cloudin välillä. Se on 200/200Mbps dedikoitu kaista konosalien välillä. Toiseksi haavoittuvin ovat palvelin-nodet, koska ne on vain kahdennettu.

4 Käynnistyslupa

Koska pohjatyö oli tehty huolella ja operointikustannusten nousu oli pieni, ei projektin käynnistysluvan saaminen ollut vaikeaa. Lupa saatiin kesällä 2017. Kaksi päivää myöhemmin lupa kuitenkin peruttiin käynnissä olleiden yritysjärjestelyjen vuoksi.

Syksyllä 2018 saimme uudelleen luvan käynnistää projekti. Tammikuussa 2019 julkistettiin, että Cognizant on ostanut Samlinkin. Muutama kuukausi kaupan jälkeen kävi selväksi, että Intiasta täytyy päästä Samlinkin Confluenceen. Pääsyn edellytyksenä oli, että Confluence ei saa sijaita Samlinkin tuotantoverkossa. Saimme viimeinkin riittävän prioriteetin projektin läpisaattamiseen järkevässä ajassa.

Tässä vaiheessa hyväksyttäväksi annettiin yksinkertaistettu ehdotus (liite 1).

5 Projektin suunnittelu

Projektin suunnittelu aloitettiin ”kick-off” -palaverilla, jonka tarkoituksena oli karhottaa, mitä projektin läpiviemiseen tarvitaan. Palaveriin kutsuttiin edustajat Windows-, Unix-, tietokanta-, pääsynhallinta- sekä tietoliikennetiimistä.

Palaverin tärkein oppi oli, että pilvipalveluiden yhdistäminen Samlinkin infraan oli hyvin alkuvaiheessa eikä meillä ollut mitään valmiita käytänteitä, joita voisimme käyttää apuna.

Antti lupasi rakentaa esimerkkiympäristön seuraavaan palaveriin. Asian käynninen läpi esimerkin kautta osoittautui huomattavasti hedelmällisemmäksi.

Aloimme toisessa palaverissa kirjaamaan ylös vaatimuksia ja määrittelemään projektia.

5.1 Tietoliikennevaatimukset

- SSH yhteys Samlinkin toimistoverkoista Confluence palvelimille (Node)
- Salattu HTTPS yhteys Samlinkin toimistoverkoista Azuren yhdyskäytävään (Application Gateway).
- Portti 8090 auki yhdyskäytävältä palvelinten aliverkkoon.
- Portit 5701, 5801, 25500 auki palvelinten aliverkon sisällä.
- Yhteys tietokannan ja palvelinten välille.
- Yhteys tiedostopalvelun ja palvelinten välille.
- DNS ohjaus Samlinkin nimipalvelimelle ja siihen liitetty SSL sertifikaatti.
- Vanhojen linkkien ohjaus uuteen osoitteeseen.

5.2 Sähköpostipalvelin

Palveluviestien lähetykseen käyttäjille. Samlinkin oma sähköpostipalvelin tai Microsoft Office 365 mailer.

5.3 Tunnistautumispalvelu

Samlinkin Active Directorysta tai Azure Domain Services palvelun kautta. LDAPS tai SAML tekniikkaa käyttäen.

5.4 Tietokanta

Azure SQL tietokanta. Suorituskykytarpeeksi arvioitiin 50 DTU tuotantoon ja 10 DTU testiin.

5.5 Tiedostopalvelin

Azure Storage account. Minimitilan tarve arvioitiin olemassa olevan datan määrän perusteella olevan 150GB. Tuotantoon otettiin 300GB.

5.6 Nodet

Kaksi kappaletta.

5.7 Tietoturva

Tietoturvasoksi määritettiin luottamuksellinen tieto laajennettuna henkilötietolain vaatimuksilla. Käytännössä tämä tarkoittaa tietojen salaamista, tietoihin pääsyn rajaamista sekä Azure alueen (region) rajaamista Eurooppaan.

5.8 Tietosuoja

Tietosuojariskit tulee arvioida ja hyväksyä.

5.9 Varmuuskopiointi

Palautus katastrofaalisen virheen jälkeen ja yksittäisten työtilojen, sivujen tai liitetiedostojen palauttamiseen.

6 Testiympäristön rakentaminen

Testiympäristö rakennettiin Atlassianin referenssiohjeistuksen mukaisesti (Atlassian, 2021).

Ajatus oli lisätä ominaisuuksia toteutukseen iteroiden. Testiympäristön pystytys mahdollisti monen asian samanaikaisen edistämisen ja testauksen.

Aluksi testiympäristöön oli pääsy vain SSH-tunnelointia käyttäen.

7 Testaus

Testiympäristössä testattiin määritettyjen vaatimuksien lisäksi tietokantamigraatio, Confluencen päivitys, liitetiedostojen siirto sekä varmuuskopiosta palauttaminen.

Testauksessa ilmeni, että joudumme tietokantamigraatiota varten ottamaan käyttöön suunniteltua uudemman version Confluencesta. Uudemmassa versiossa oli korjattu migraatiota vaivannut bugi.

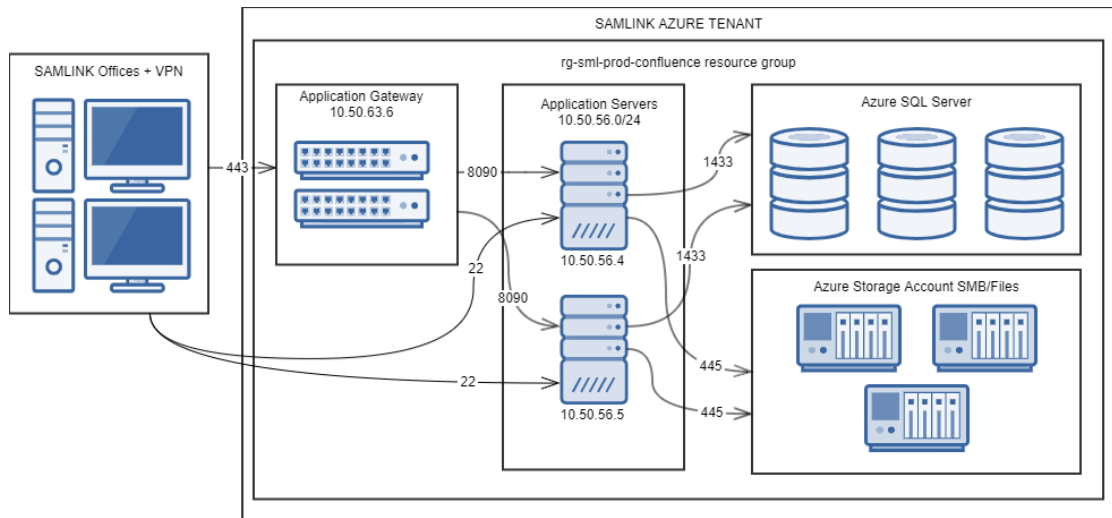
Merkittävästi testauskapasiteettia käytettiin oman kustomoidun ulkoasun, sekä lisäosien testaamiseen uuden Confluence-version kanssa.

Koska ympäristön ja Confluencen versiopäivitykset on jatkossakin tarkoitus ostaa konsultointina, dokumentoitiin pelkästään varmuuskopiosta palautus (liite 2).

8 Tuotantoympäristön määrittely

Tuotantoympäristöön oli tarkoitus monistaa testiympäristöön määritetty toimiva konfiguraatio. Suunnittelussa keskityttiin varmistamaan, että jokainen osa ympäristöstä on saatavuudeltaan, tietoturvaltaan ja varmuuskopioinniltaan vaaditulla tasolla.

8.1 Tietoliikenne



Kuva 7 Azure tuotantoympäristön verkkonäkökulma (Samlink 2019)

Kuvassa näkyvät aiemmin määritellyt yhteydet.

Vanhassa osoitteessa oli mukana palvelimen nimi. Koska kyseisellä palvelimella oli käytössä muitakin palveluita, haluttiin uudelle ympäristölle uusi osoite. Uusi osoite helpotti myös käyttöönottoa, sillä se voitiin asettaa käyttökuntoon ennen käyttöönottoa.

Koska työntekijöiden sähköpostiin on vuosien saatossa lähetetty miljoonia linkkejä, rakennettiin ohjaus vanhasta osoitteesta (verstas.saminet.fi/confluence) uuteen (confluence.saminet.fi).

8.2 Sähköposti

Sähköpostipalvelimeksi konfiguroitiin Samlinkin Microsoft Office 365 Mailer. Valinnalla varmistetaan yhteensopivuus tulevaisuudessa ja ehkäistään tehokkaasti järjestelmän lähettämien viestien merkintä roskapostiksi.

8.3 Tunnistautuminen

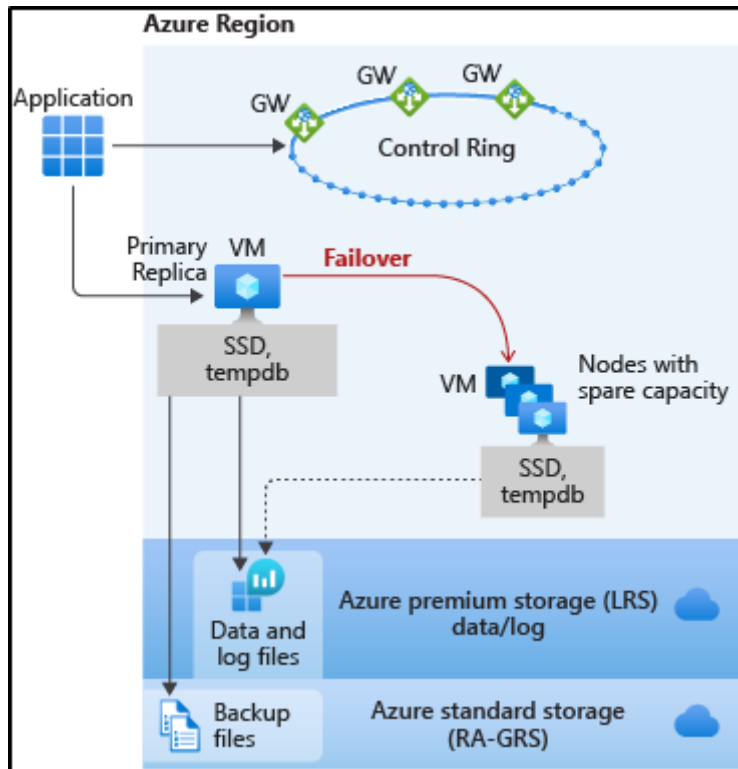
Samlinkin Azure Domain Servicen käyttäjäryhmiin liittyvien rajoitusten vuoksi päädyimme käyttämään vanhempaa Active Directorya.

Tunnistautumista varten hankittiin Confluence SAML SSO -lisäosa. Se mahdollistaa muitakin lisäominaisuuksia, kuten käyttäjien automaattisen passivoinnin. Kyseinen ominaisuus pudottaa merkittävästi aktiivisten käyttäjien määrää, joka taas mahdollistaa merkittäviä säästöjä pienemmän lisenssin käyttämisen kautta.

8.4 Tietokantapalvelin

Tietokantapalvelimeen päätettiin ottaa 50 DTU:n sijaan 100 DTU:ta. Muutos johtuu palvelutason vaihtumisesta S2 tasosta S3 tasoon. S3 taso on ensimmäinen taso, joka sisältää vähintään yhden kokonaisen virtuaalisen suorittimen (vCore). (Microsoft, 2019)

Microsoft lupaa perustason (Standard) tietokantapalvelimille 99,99 % saatavuuden.



Kuva 8 Azure tietokannan varmennus (Microsoft, 2020)

Pääasiallista (Primary Replica) tietokantaa ajetaan virtuaalisella koneella (VM). Tällä koneella on SSD levyllä ja muistissa vain väliaikaistiedostoja. Häiriötilanteissa ja päivityksien aikana prosessi siirretään toiselle virtuaalikoneelle, jolla on vapaata kapasiteettia. Siirto aiheuttaa hetkellisen suorituskyvyn alentumisen, koska uusi virtuaalikone aloittaa tyhjällä välimuistilla. Virtuaalikoneen vaihto ei vaikuta tietokannan tietoihin millään tavalla.

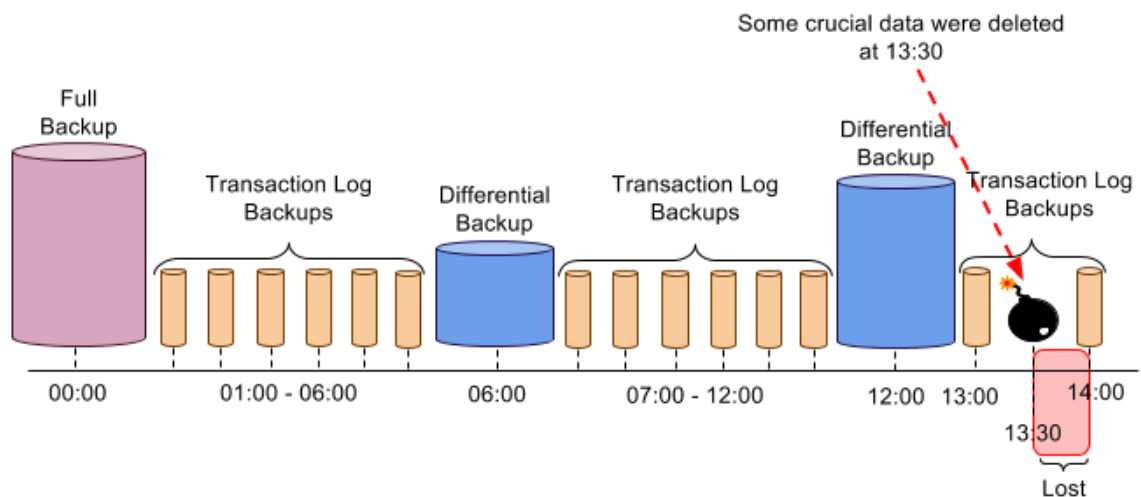
Itse tiedot on tallennettu paikallisesti varmennettuun (Locally Redundant Storage) nopeaan (Premium) tiedostopalveluun. Paikallisesti varmennettu tieto on monennettu useaan eri päivitysryhmään ja konesaliin. Katko aiheutuisi vain, jos häiriö on koko valitun alueen (Zone) laajuinen.

Tällä menettelyllä tieto on kirjoitettu kuuteen eri sijaintiin. Se säilyy vähintään 99,99999999999999 % varmuudella ehjänä. Luku- ja kirjoituspyynnöt onnistuvat vähintään 99,99 % varmuudella.

Tämän lisäksi tiedot varmuuskopioidaan perustason tiedostopalveluun, joka on varmennettu kahden fyysisesti eri saatavuusalueen välille (Geo Redundant Storage) joista toissijaiselle alueelle on vain lukuoikeudet (Read Access).

Tietosuojasyistä alueiksi on valittu West Europe ja North Europe.

Yllä kuvatulla menettelyllä katetaan järjestelmästä johtuvat ja katastrofaalisen virheen aiheuttamat virhetilanteet. Vahingossa poistettujen tietojen osalta käytössä on lisäksi alla kuvatun mukainen varmuuskopiointi.



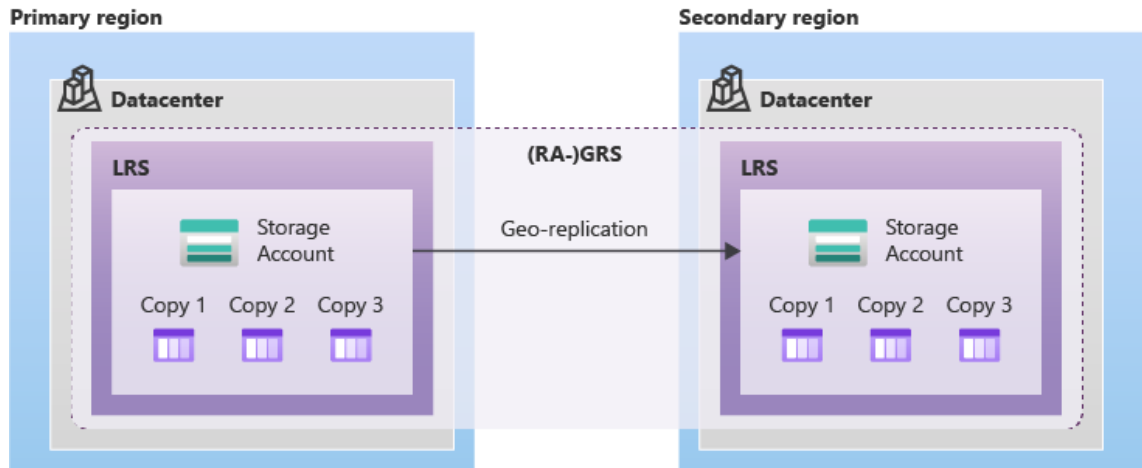
Kuva 9 tietokannan aikapisteeseen palautuksen toimintaperiaate (Omelchenko, 2015)

Azure tekee viikoittain täyden varmuuskopion tietokannasta. 12–24 tunnin välein tehdään eroavaisuusvarmuuskopiointi. Niiden välissä tehdään 5–10 minuutin välein transaktioloki-varmuuskopio.

Tietokanta voidaan palauttaa 5–10 minuutin tarkkuudella mihin tahansa tilaan viimeisen 28 päivän ajalta.

8.5 Tiedostopalvelin

Azure Storage Account. Tiedostopalvelimella sijaitsevat liitetiedostot.



Kuva 10 Azure tiedostopalvelimen varmennus (Microsoft, 2022)

Tiedostopalvelimen tiedot on tallennettu paikallisesti varmennettuun (Locally Redundant Storage) tiedostopalveluun. Paikallisesti varmennettu tieto on monennettu useaan eri päivitysryhmään ja konesaliin. Katko aiheutuisi vain, jos häiriö on koko valitun alueen (Zone) laajuinen.

Tämän lisäksi tiedot varmuuskopioidaan fyysisesti eri saatavuusalueelle (Geo Redundant Storage) joista toissijaiselle alueelle on vain lukuoikeudet (Read Access).

Tällä menettelyllä tieto on kirjoitettu kuuteen eri sijaintiin. Se säilyy vähintään 99,99999999999999 % varmuudella ehjänä. Luku- ja kirjoituspyynnöt onnistuvat vähintään 99,99 % varmuudella.

Yllä kuvatulla menettelyllä katetaan järjestelmästä johtuvat ja katastrofaalisen virheen aiheuttamat virhetilanteet. Vahingossa poistettujen tietojen osalta käytössä on lisäksi alla kuvatun mukaiset erilliset varmuuskopiot.

Tiedostopalvelin varmuuskopioidaan päivittäin. Päivittäiset varmuuskopiot säilytetään seitsemän päivän ajan. Sunnuntaisin ajetaan viikoittainen varmuuskopio, jota säilytetään 4 viikon ajan.

Koko tiedostojaon varmuuskopioiden lisäksi Azuressa on päällä pehmeä tiedostojen poisto (Soft delete). Poistetut tiedostot siirtyvät 30 päivän ajaksi roskakoriin, josta ne 30 päivän jälkeen poistuvat automaattisesti.

8.6 Nodet

Nodet (palvelimet) joilla ajetaan Confluence palvelua on monennettu. Samlinkin liikennemäärät ovat sen verran pieniä, että yksikin node voisi riittää resurssien puolesta. Projektin tarkoitus oli kuitenkin varmistaa Confluencen saatavuus. Saatavuuden parantamiseksi käyttöön otettiin 2 nodea.

Atlassianin suosituksia noudattaen valitsimme nodeiksi Standard D4s v3 (4 vcpus, 16 GiB memory).

Kustannus- ja käyttövarmuussyistä valitsimme käyttöjärjestelmäksi Linuxin. Tarkoituksena oli saada Unix tiimiltä käyttöön Samlinkin oma Red Hat julkaisu. Se sisältäisi valmiiksi pääkäyttäjien luvituksen, tietoturvatyökalut, automaattipäivitykset ja valvontatyökalut.

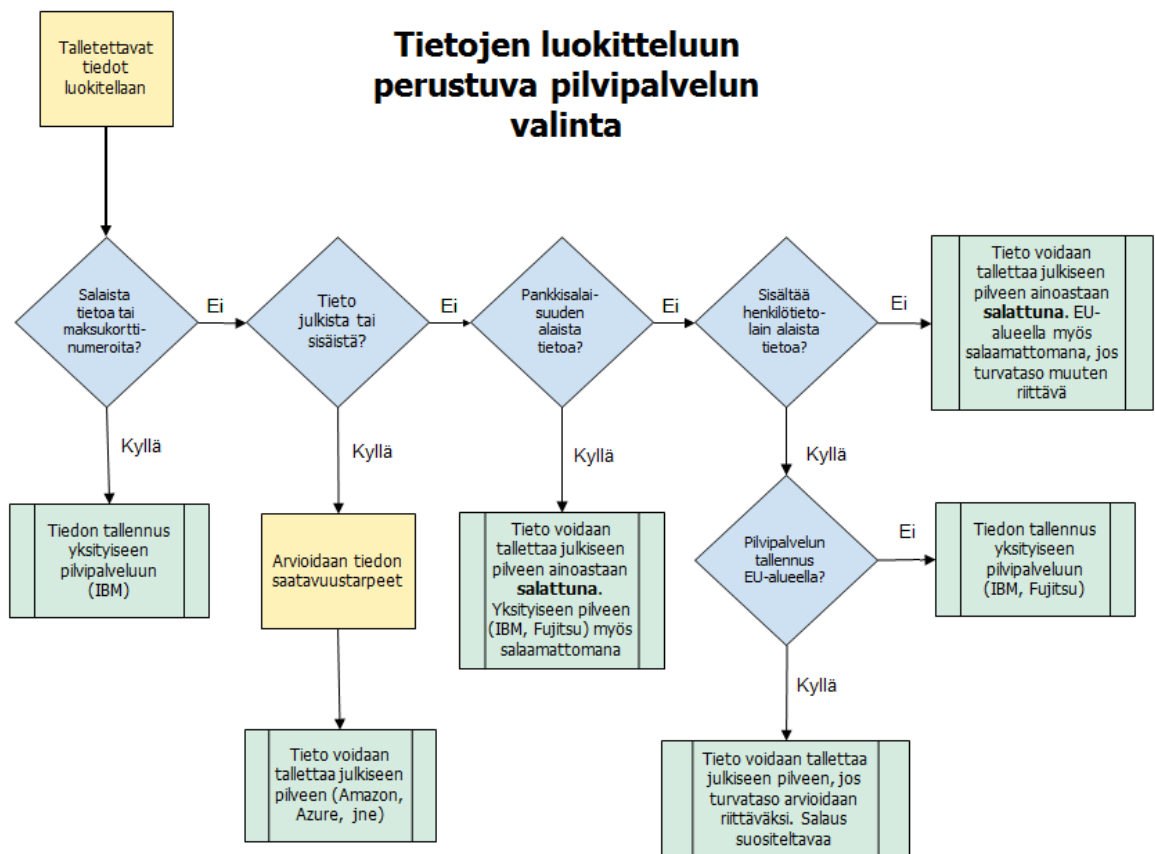
Valitettavasti sellaista ei ollut valmiina Azure ympäristöön, emmekä saaneet sitä projektin puitteissa toteutettua. Nodeille päädyttiin asentamaan Azuren vakio Red Hat julkaisu.

8.7 Tietoturva

Tietoturva-vaatimuksiin vaikuttavat sekä tiedon luokittelu että lait ja standardit. Samlinkissa tiedot luokitellaan julkiseksi, sisäiseksi, luottamukselliseksi tai salaiseksi. Luokittelun lisäksi vaatimuksia tulee esimerkiksi PCI-DSS luottokorttistandardista.

Samlinkin tietoturva-vaatimusten mukaan julkiseen pilvipalveluun ei saa tallettaa maksukorttinumeroita, eikä salaiseksi luokiteltua tietoa riippumatta pilvipalvelun suojaustasosta.

Pankkien asiakas- ja varallisuustietoja (pankkisalaisuus) voi tallentaa vain vahvasti salattuna ja salausavainten on oltava Samlinkin hallinnassa ja hallinnasta on huolehdittava Samlinkin ohjeiden mukaisesti.



Kuva 11 Tietojen luokitteluun perustuva pilvipalvelun valinta (Samlink, 2019)

Confluenceen ei ole koskaan saanut tallettaa maksukorttinumeroita, pankkisalaisuuden alaista tietoa tai ylipäätään mitään loppukäyttäjän tietoja. Ympäristö rakennettiin siis täyttämään luottamuksellisen tiedon sekä henkilötietolain vaatimukset. Päätimme kuitenkin rakentaa option pankkisalaisuuden alaisten tietojen säilytykseen. Se on toteutettavissa yksinkertaisesti ottamalla salausavaimet Azuren automaattihallinnasta Samlinkin hallintaan.

Cognizant Securityn lisävaatimus tietoturvaan oli mahdollisuus ilmiantaa salaiseksi luokiteltu tai pankkisalaisuuden alainen sisältö. Se toteutettiin lisäämällä päävalikkoon Report-nappi, jota painamalla muodostuu sivun osoitteen sisältävä sähköposti, jonka vastaanottajana on tiketöintijärjestelmä.

8.8 Tietosuoja

Tietosuojariskeiksi arvioitiin seuraavat:

- Käyttäjän toimesta järjestelmään on mahdollista viedä mitä vain. Esimerkiksi loppuasiakkaan henkilötietoja asettamatta tällaisten tietojen vaatimia käyttöoikeusrajoituksia ja tietojen poistomenettelyä.
- Yllä mainitut tiedot päätyvät asiattomaan käyttöön tahallisesti tai tahattomasti
- Käyttäjän toimesta on mahdollista toimittaa asiakkaalle materiaalia, joka voi sisältää toisen asiakkaan ratkaisun tietoja.
- Tietojen tahaton pysyvä poisto (tiedot joudutaan palauttamaan varmuuskopiosta)

Riskit pisteytettiin todennäköisyyden ja vaikuttavuuden mukaan. Pisteytyksistä laskettiin riskiarvo kaavalla $\text{todennäköisyys} \times \text{vaikuttavuus} \times \text{vaikuttavuus}$.

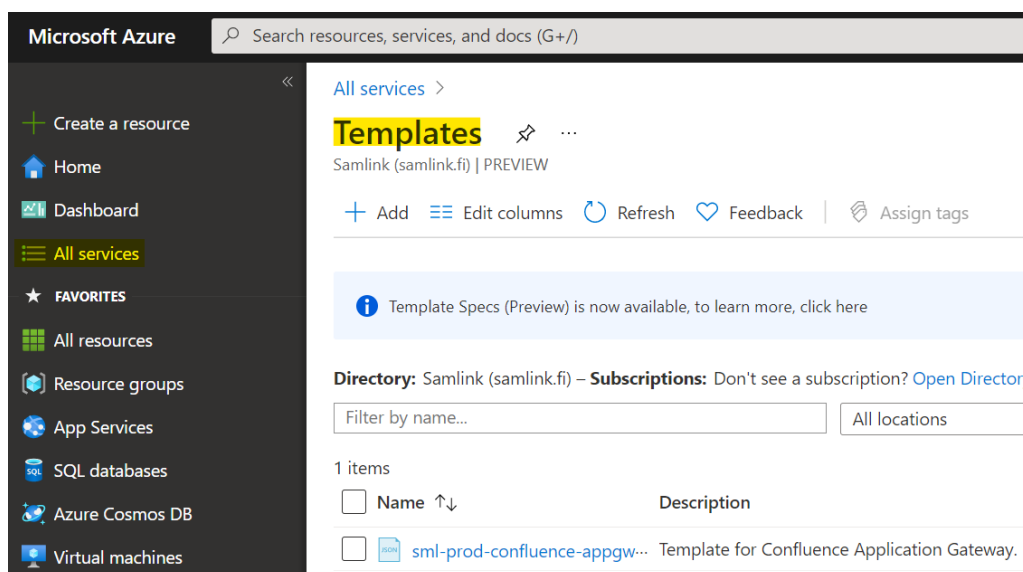
Kyseisellä kaavalla riskien vertailuluvut olivat järjestyksessä 4, 9, 8 ja 1. Alle 32 riskiarvot ovat matalia. Koska riskiarvot olivat näin matalia, hyväksyttiin riskit ilman erityisiä minimointitoimenpiteitä.

8.9 Varmuuskopiointi

Varmuuskopiointia sisällön osalta on käsitelty tietokantapalvelimen ja tiedostopalvelimen osalta.

Tiedon lisäksi myös Azure konfiguraatiot on tallennettu pohjiksi Azure-portaaliin. Application Gateway, Resource Group ja Nodet voidaan kaikki luoda uudelleen

käyttämällä valmiita pohjia (Template).



Kuva 12 Pohjat Azure Portaalissa (Samlink, 2019)

Noden hajotessa on kuitenkin suositeltavaa monistaa toimiva node käyttöön. Jos molemmat nodet hajoavat on alkuperäisestä toimivasta nodesta olemassa Snapshot joka on toissijainen palautusmenetelmä. Vasta kolmantena turvaututaan pohjasta palauttamiseen.

9 Käyttöönottosuunnitelma

Testivaiheessa havaittiin liitetiedostojen kopiointin ja tietokantamigraation alentavan merkittävästi tuotantoympäristön suorituskykyä.

Käyttöönotto suunniteltiin aloitettavan torstaina viiden aikaan iltapäivällä siirtämällä liitetiedostot uuteen ympäristöön. Seuraavassa aikataulu:

Perjantaina viiden aikaan tuotantoympäristö laitetaan vain-luku tilaan ja aloitetaan tietokantamigraatio, sekä liitetiedostojen synkronointiajo, jossa verrataan torstaina kopioituja tiedostoja tuotantoympäristön tiedostoihin ja kirjoitetaan muuttuneet tiedot uuteen ympäristöön.

Lauantaina aamupäivällä validoidaan tietokannan siirtyminen ja testataan uuden ympäristön toiminta.

Kun kaikki toimii, uutisoidaan asiasta ja laitetaan vanhojen linkkien ohjaus päälle.

Virhetilanteissa käyttöönottoa jatketaan tarvittaessa sunnuntain puolelle. Jos uutta ympäristöä ei saada toimimaan perutaan käyttöönotto palauttamalla vanhan tuotantoympäristön muokkausoikeudet.

10 Käyttöönotto

Käyttöönotto osoittautui tietokantamigraation osalta oletettua ja testattua haastavammaksi. Löysimme useita virheitä palatessamme lauantaina toimistolle tarkastamaan siirtojen onnistumista.

Iso osa lauantapäivästä meni yrittäessämme saada epäonnistunut migraatio korjattua ilman uutta ajoa. Päädyimme lopulta kuitenkin uuteen ajoon. Loppupäivän selvitimme mahdollisuuksia vähentää migraation virhealttiutta.

Pitkällisen selvityksen jälkeen päädyimme tilanteeseen, jossa pystyimme hallitsemaan virheet omalla interventiolla. Se valitettavasti tarkoitti koneella istumista pitkälle aamuyöhön.

Sunnuntai aamuna migraatio oli onnistuneesti ohi ja saimme lopputuloksen valitua onnistuneesti.

Julkaisimme tiedotteen uuden ympäristön käyttöönotosta (Liite 2) ja ajoimme vanhan ympäristön alas.

11 Virheiden käsittely

Käyttöönoton jälkeen löytyi kymmenkunta pienempää virhettä. Yksi virhe johtui tietokantamigraatiosta, yksi konfiguraatiovirheestä toisella nodella. Kaksi virhettä liittyi omaan kustomoituun ulkoasuun. Neljä virhettä johtui lisäosien yhteensopivuusongelmista. Virheiden korjausta seurattiin etusivun uutisessa.

12 Opit

Työmääräarvio osoittautui optimistiseksi. Sovellustukea kirjattiin projektille lopulta 12 päivän edestä ja ulkoista työtä 10 päivän edestä, mutta etenkin projektin johtamiseen meni huomattavasti oletettua enemmän aikaa. Koska projekti oli pieni ja tekeminen hyvin henkilöitynyttä, en toteuttanut kirjauserittelyä. Minun osuuteni (testi- ja tuotantoympäristön rakentaminen, testaus ja virheen korjaus ja projektin johtaminen) kasvoi lopulta 60 päivän kokoiseksi. Arviolta kolmannes siitä kului projektin johtamiseen.

Tärkein henkilökohtainen oppi oli, että ei kannata ottaa liian montaa tehtävää projektiin. Jos olet pääasiallinen resurssi toteutuksen osalta ei kannata sen lisäksi lähteä projektipäälliköksi.

Lähteet

Atlassian, 2021. *Getting started with Confluence Data Center on Azure*. [Online]

Available at: <https://confluence.atlassian.com/doc/getting-started-with-confluence-data-center-on-azure-937176452.html>

[Haettu 30 January 2022].

Microsoft, 2019. *Resource limits for single databases using the DTU purchasing model*. [Online]

Available at: <https://docs.microsoft.com/en-us/azure/azure-sql/database/resource-limits-dtu-single-databases>

[Accessed 14 Maaliskuu 2021].

Microsoft, 2020. *High availability for Azure SQL Database and SQL Managed Instance*. [Online]

Available at: <https://docs.microsoft.com/en-us/azure/azure-sql/database/high-availability-sla>

[Haettu 14 Maaliskuu 2021].

Microsoft, 2022. *Azure Storage redundancy*. [Online]

Available at: <https://docs.microsoft.com/en-us/azure/storage/common/storage-redundancy>

[Haettu 30 January 2022].

Omelchenko, A., 2015. *Point-in-time recovery*. [Online]

Available at: <https://academy.sqlbak.com/point-in-time-recovery/>

[Haettu 30 January 2022].

Samlink, 2019. *Samlink*. Espoo: Samlink.

Wikipedia, 2019. *Wikipedia - Confluence*. [Online]

Available at: <https://fi.wikipedia.org/wiki?curid=834536>

[Haettu 13 February 2022].

Liite 1: Verstas Confluencen klusterointiehdotus Azure-cloudiin

Nykytilanne

Verstas on aikanaan perustettu paljon pienemmälle käyttäjäkunnalle ja huomattavasti nykyistä kapeampaan käyttöön. Ajan myötä käyttötarkoitukset ovat laajentuneet kattamaan kaiken asiakasdokumentaatiosta järjestelmädokumentaatioon, toimintamalliin, ohjeistukseen ja jopa roolikohtaiseen tiedottamiseen, sekä tiedostonhallintaan.

Käyttötapausten ja käyttäjämäärän lisääntyessä on myös käyttökatkojen aiheuttama häiriö ja kustannus kasvanut. Käyttökatko voi pahimmillaan aiheuttaa lisää viivästystä pankkiympäristön korjaukseen, jos järjestelmädokumentaatioon ei päästä käsiksi. On myös mahdollista, että tarjous myöhästyy käyttökatkon takia. Lyhyelle puolen tunnin käyttökatkolle voidaan laskea hintaa muutamissa tuhanissa euroissa, mutta katkon venyessä, vaikka koko päivän mittaiseksi, aletaan puhua merkittävästä haitasta satojen ihmisten työntekoon. Tällöin puhutaan kertaluokkaa suuremmista tappioista työajan menetyksinä, jotka vielä kertaantuvat aikataulujen venymisenä myöhemmin.

Verstas on nykyään ei dedikoidulla palvelimella, jonka resurssit alkavat olla sille hyvin rajalliset. Vika Confluencen moottorissa, tietokannassa tai huonossa tapauksessa muussa palvelimen tarjoamassa palvelussa aiheuttaa auttamatta käyttökatkon.

Verstaasta on tullut meille liiketoimintakriittinen järjestelmä.

Klusterointi eli palvelun kahdentaminen

Ainoa järkevä tapa klusteroida Confluence on Atlassianin luoma Datacenter-malli.

Private datacenter malli

Tässä klusterissa on kaksi aktiivista Confluence-palvelinta, johon käyttäjän istuntokohtaiset pyynnöt jaetaan kuormanjakajan toimesta. Lisäksi erillinen yhteinen tiedostopalvelin. Tietokanta viedään erilliseen tietokantaklusteriin (SQL Server 2014)

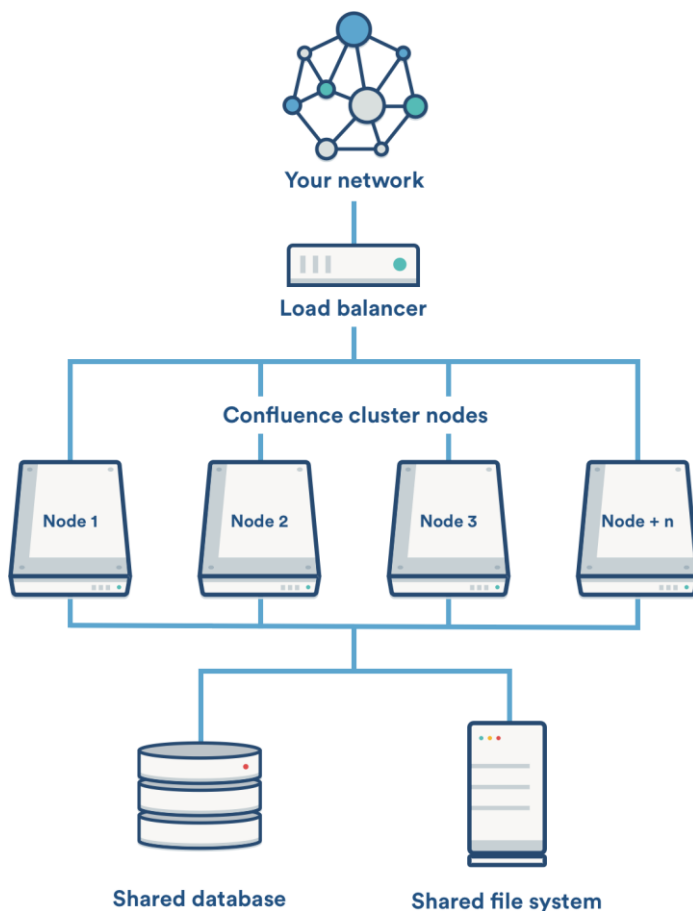
Hyödyt:

- mahdollistaa sekä kuormanjaon, että Fail Overin samalla lisenssikustannuksella
- malli on dokumentoitu kattavasti Atlassianin toimesta. Valmis implementointimalli.

Haitat:

- Vaatii 2 identtistä palvelinta, tiedostopalvelimen ja tietokantapalvelimen

Verkkonäkökulma Private datacenter



Datacenter Azure -pilvipalvelussa

Rakennetaan Datacenter ympäristö Azure-pilveen. Pilvi rakennetaan konsultin avustuksella siten, että ylläpito-osaaminen jää Samlinkin asiantuntijalle.

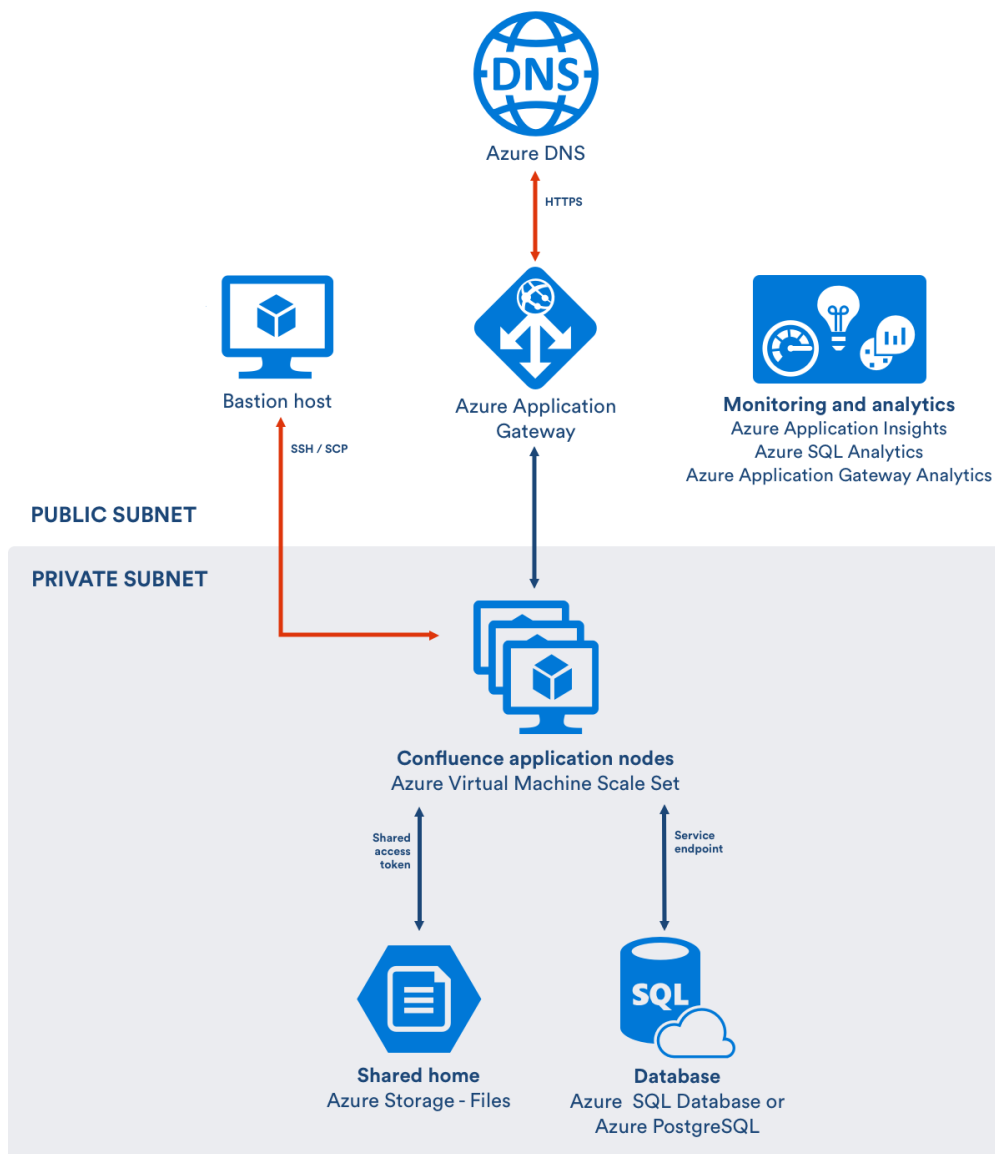
Hyödyt:

- mahdollistaa sekä kuormanjaon, että Fail Overin samalla lisenssikustannuksella.
- malli on dokumentoitu kattavasti Atlassianin toimesta. Valmis implementointimalli.
- Samlinkin infra ei rajoita implementointimahdollisuuksia, vaan voidaan toimia suositusten mukaan.
- Palvelinkustannukset ovat IBM:n tai Fujitsun kapasiteettia pienemmät.
- Nodet voidaan jakaa fyysisesti eri sijainteihin.
- Azureen on olemassa Express Gateway, joka mahdollistaa monien asioiden toiminnan kuten sisäverkossa.
- Azure osaaminen lisääntyy

Este:

- Pilviosaaminen Samlinkissa vähäistä.

Verkkonäkökulma Datacenter Azure Cloud



Työmääräarvio

Infra-asiantuntijat 10 htp, pääkäyttäjä 26 htp. projektipäällikkö 4 htp, konsultti 8 htp. Yhteensä 48 htp.

Kokonaisvuosikustannukset

	Nykykus- tannus	Private Datacenter	Azure Datacenter
Palvelinkus- tannus	4 800 €	17 000 €	7 800 €
Lisenssikus- tannus	21 800 €	23 500 €	23 500 €
Yhteensä	26 600 €	40 500 €	31 300 €
Ero	- €	13 900 €	4 700 €

Henkilöresurssit

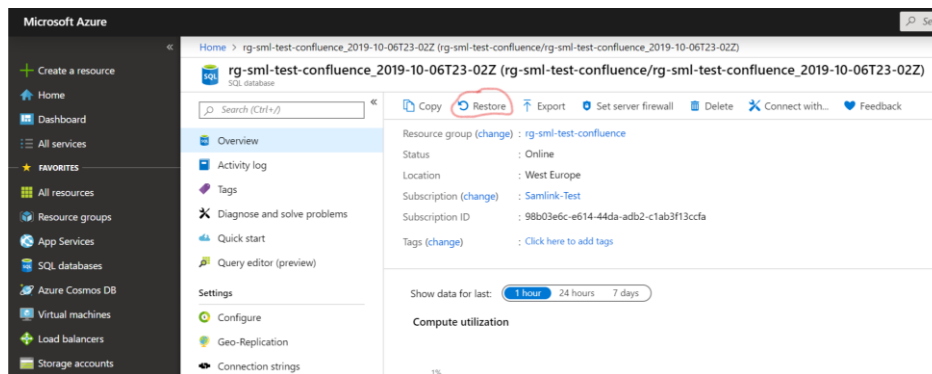
- Pasi Sainio / Teijo Raatikainen / Sami Marjokorpi Linux/DNS
- Jarkko Timonen Windows/Azure
- Anssi Rajaniemi Tietoliikenne
- Ari Kaatonen / Mikko Källroos Tietokannat
- Antti Häkli Ympäristöjen rakentaminen, konsultti

Liite 2: Varmuuskopiosta palauttaminen

Palautuminen katastrofisesta virheestä

Create a new database from backup

Navigate to portal.azure.com and find the correct DB.



Click Restore

Select the correct time. (Usually, the latest time known to work)

Restore

rg-sml-test-confluence_2019-10-06T23-02Z

Select source
Point-in-time

* Database name
rg-sml-test-confluence_2019-10-06T23-02Z_...

Point-in-time
Oldest restore point
2019-10-09 11:22 UTC

Restore point (UTC)
10/10/2019 10:54:00 AM

* Target server
rg-sml-test-confluence westeuro...

Elastic pool
None

* Pricing tier
Standard S3: 100 DTUs, 250 GB s...

Set Confluence to use the new database

Shut down Confluence on both nodes

```
systemctl stop confluence
```

Copy the new DB name. Restore will create a duplicate DB with a new name.

On each node open:

```
/var/atlassian/application-data/confluence/confluence.cfg.xml
```

Edit database name on this line:

```
<property name="hibernate.connection.url">jdbc:sqlserver://rg-sml-test-confluence.database.windows.net:1433;database=DATABASE_NAME_HERE; ...
```

Start confluence on 1 node first. When Confluence responds start the second node:

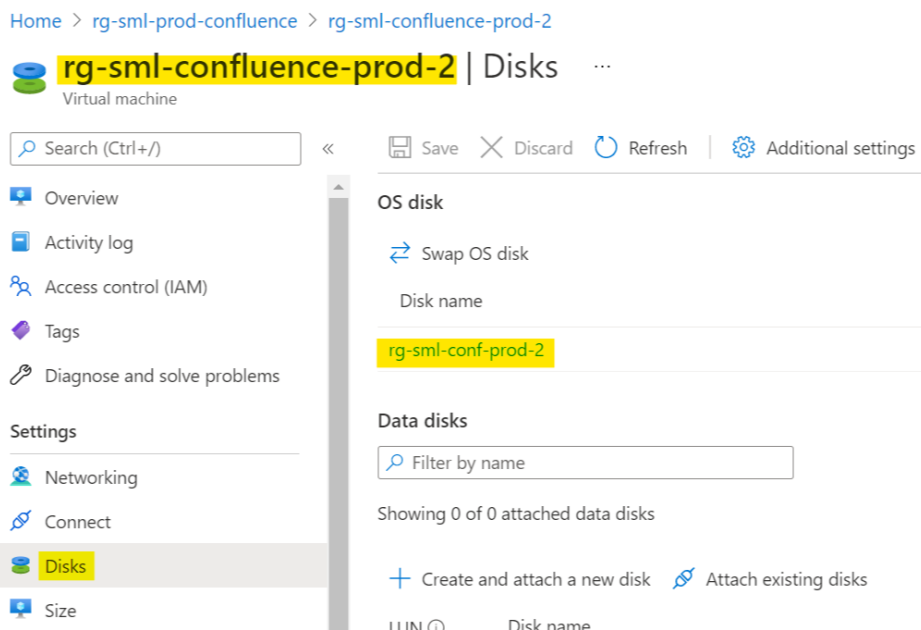
```
systemctl start confluence
```

Yksittäisten tietojen palauttaminen

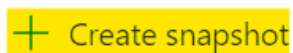
Clone VM

Navigate to the disks of the VM you are gonna clone.

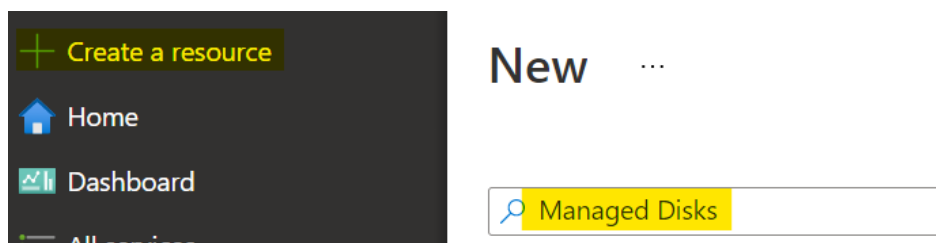
Click on the disk name.



Click:



Create a new Managed Disk:



Select:

Disk size : 128 GiB
Disk sku : Premium SSD LRS

Set the following:

[Home](#) > [New](#) > [Managed Disks](#) >

Create a managed disk ...

[Basics](#) Encryption Networking Advanced Tags Review + create

Select the disk type and size needed for your workload. Azure disks are designed for 99.999% availability. All disks encrypt your data at rest, by default, using Storage Service Encryption. [Learn more about encryption](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize your resources.


Subscription * ⓘ	<input type="text" value="Samlink-Production"/>
Resource group * ⓘ	<input type="text" value="rg-sml-prod-confluence"/> Create new

Disk details

Disk name * ⓘ	<input type="text" value="rg-sml-confluence-prod-2_MNGDISK_17_3_2021"/>
Region * ⓘ	<input type="text" value="(Europe) West Europe"/>
Availability zone	<input type="text" value="None"/>
Source type ⓘ	<input type="text" value="Snapshot"/>
*Source snapshot ⓘ	<input type="text" value="rg-sml-confluence-prod-2_SNAPSHOT_17_3_2021"/>
Size * ⓘ	<input type="text" value="128 GiB"/> Premium SSD LRS Change size

Go to created resource:

✓ Your deployment is complete

 Deployment name: Microsoft.ManagedDisk-202103
Subscription: [Samlink-Production](#)
Resource group: [rg-sml-prod-confluence](#)

∨ Deployment details ([Download](#))

∧ Next steps

[Go to resource](#)

Click:



Set the following:

[Basics](#) [Disks](#) [Networking](#) [Management](#) [Advanced](#) [Tags](#) [Review + create](#)

Create a virtual machine that runs Linux or Windows. Select an image from Azure marketplace or use your own customized image. Complete the Basics tab then Review + create to provision a virtual machine with default parameters or review each tab for full customization. [Learn more](#)

Project details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription ⓘ	Samlink-Production
Resource group * ⓘ	rg-sml-prod-confluence

[Create new](#)

Instance details

Virtual machine name * ⓘ	rg-sml-confluence-prod-TEMP-VM-17-3-2021
Region ⓘ	(Europe) West Europe
Availability options ⓘ	No infrastructure redundancy required
Image * ⓘ	rg-sml-confluence-prod-2_MNGDISK_17_3_2021 - Gen1
Azure Spot instance ⓘ	<input type="checkbox"/>
Size * ⓘ	Standard_D4s_v3 - 4 vcpus, 16 GiB memory (147,46 €/month)

[See all images](#)
[See all sizes](#)

Basics Disks **Networking** Management Advanced Tags Review + create

Define network connectivity for your virtual machine by configuring network interface card (NIC) settings. You can control ports, inbound and outbound connectivity with security group rules, or place behind an existing load balancing solution. [Learn more](#)

Network interface

When creating a virtual machine, a network interface will be created for you.

Virtual network * ⓘ [Create new](#)

Subnet * ⓘ [Manage subnet configuration](#)

Public IP ⓘ [Create new](#)

NIC network security group ⓘ None
 Basic
 Advanced

i The selected subnet 'subnet-sml-prod-int-linux-01 (10.50.56.0/24)' is already associated to a network security group 'nsg-sml-prod-int'. We recommend managing connectivity to this virtual machine via the existing network security group instead of creating a new one here.

Configure network security group * [Create new](#)

Accelerated networking ⓘ The selected image does not support accelerated networking.

Load balancing

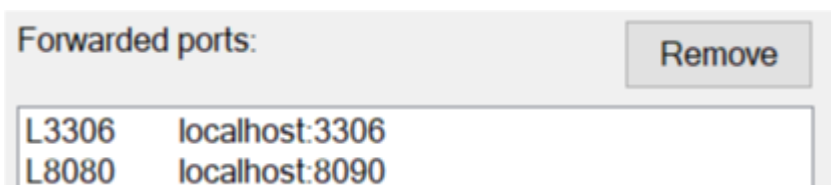
You can place this virtual machine in the backend pool of an existing Azure load balancing solution. [Learn more](#)

Place this virtual machine behind an existing load balancing solution?

Done.

Set DB

Make sure you have the following tunnels set in your SSH client:



As Superuser do the following:

Stop confluence service

```
systemctl stop confluence
```

Edit:

```
/var/atlassian/application-data/confluence/confluence.cfg.xml
```

Edit database name on this line:

```
<property name="hibernate.connection.url">jdbc:sqlserver://rg-sml-test-confluence.database.windows.net:1433;database=DATABASE_NAME_HERE; ...
```

Start confluence service

```
systemctl start confluence
```

Keep SSH session active and navigate to <http://localhost:8080/login.action?nosso>

Liite 3: Käyttöönottoviesti

Verstas siirretty pilveen ja päivitetty

Verstaan päivitys saatiin päätökseen sunnuntai iltana. Voit kirjautua Verstaalle uudesta osoitteesta: <https://confluence.saminet.fi/>

Vanhat linkit toimivat ohjauksen ansiosta jatkossakin.

Huomioitavaa

Käyttäjät näkyvät Unknown Usereina kunnes kirjautuvat Verstaalle ensimmäisen kerran.

Jos huomaat Verstaalla poikkeavaa toimintaa päivityksen jälkeen ota yhteyttä verstas@samlink.fi Voit käyttää yläpalkin **Report**-nappia. Muistathan sitä käyttäessäsi muuttaa viestin otsikon.

Muutoksia

- Verstaalle ei jatkossa tarvitse erikseen kirjautua.
- Verstaan ulkoasu ja muutamat toiminnallisuudet tulevat muuttumaan jonkin verran. Siirrymme 5.10.6 versiosta 7.1.0 versioon.
- Hakukentän toiminta muuttuu. Entisen pikahaun otsikkotietojen sijaan se hakee kaikesta sisällöstä ja hakutuloksien rajaaminen suodattimilla muuttuu entistä tärkeämmäksi. Täysi haku (Advanced search) toimii entiseen tapaan.
- Sivujen muokkaus muuttuu yhteisöllisemmäksi. Muiden käyttäjien muutokset näkyvät Edit tilassa lähes reaaliajassa.
- "Link to Page" toiminnallisuuden lyhytlinkki löytyy jatkossa "Share" napin takaa.