

Outi Sorsa

Langattomien tapahtumaverkkojen käyttöasteen raportointi

Metropolia Ammattikorkeakoulu

Insinööri (AMK)

Tietotekniikan koulutusohjelma

Insinööriytyö

6.5.2016

Tekijä Otsikko	Outi Sorsa Langattomien tapahtumaverkkojen käyttöasteen raportointi
Sivumäärä Aika	32 sivua + 4 liitettä 6.5.2016
Tutkinto	Insinööri (AMK)
Koulutusohjelma	Tietotekniikka
Suuntautumisvaihtoehto	Tietoliikennetekniikka
Ohjaajat	Lehtori Sami Sainio Puheenjohtaja Matti Antila
<p>Insinööriyön tarkoituksena oli luoda langattomien lähiverkkojen käytön esittämiseen helpokäyttöinen raportointijärjestelmä, joka on helpokäyttöinen ja sovellettavissa erilaisten tapahtumien tarpeisiin. Tavoitteena oli automatisoida raporttien luonti.</p> <p>Työn ensimmäisessä vaiheessa tutustuttiin aiemmin manuaalisesti tuotettuihin käyttöraportteihin, ja selvitettiin valmiudet tuottaa käytettävissä olevista lähtötiedoista vastaavia raportteja automatisoidusti. Seuraavassa vaiheessa selvitettiin vaihtoehtoiset raportointiohjelmistot ja valittiin niistä tarkoitukseen parhaiten soveltuva. Osana työn suorittamista tutustuttiin erilaisiin tiedon esitysmuotoihin ja selvitettiin niiden soveltumista erityyppisten tietojen esittämiseen. Lisäksi työssä automatisoitiin lähtötietojen muokkaaminen raportointiohjelmiston ymmärtämään muotoon.</p> <p>Työn tuloksena syntyi valmis raporttipohja, johon voidaan syöttää määritellyssä muodossa olevaa verkkostatistiikkaa. Lisäksi työn tilaajalle laadittiin ohjeistus raporttien luomiseksi. Valmis raporttipohja mahdollistaa raporttien toimituksen tilaajan yhteistyökumppaneille nopeasti ja helposti. Lisäksi raporttien visuaalinen ilme tekee ne helppolukuisiksi. Helppo raportointi vähentää manuaalista työtä ja vapauttaa aikaa itse verkon suunnitteluun.</p>	
Avainsanat	langaton verkko, Qlik Sense, statistiikka, raportointijärjestelmä

Author Title	Outi Sorsa Wireless network statistics in event networks
Number of Pages Date	32 pages + 4 appendices 6 May 2016
Degree	Bachelor of Engineering
Degree Programme	Information and Communications Technology
Specialisation option	Communication Networks
Instructors	Sami Sainio, Lecturer Matti Antila, Chairman of the board
<p>The purpose of the study was to create an easy-to-use reporting system for introducing the use of wireless local area networks. The system was to be applicable to the needs of various events. The purpose of the study was to automate the creation of usage reports.</p> <p>The first phase of the study was to get familiar with previously used manually produced reports and initial data. The possibility of producing similar output with the available information in an automated manner was investigated. The next phase was to find possible reporting software and select the most suitable for the purpose. It was also part of this study to explore various forms of information presentation and examine which of them is suitable for the different types of data. In addition, initial data handling and manipulation to a software usable form was automated.</p> <p>As a result, a reporting template, which takes a predetermined statistics file as an input, was created. Documentation on how to produce reports was also created as part of the study. This system makes it easier to deliver reports to partners and other involved parties. In addition, the visual appearance of the reports makes them easier to read. Easy reporting reduces manual effort and frees up time for network planning.</p>	
Keywords	wireless, network, qlik sense, statistics

Sisällys

Lyhenteet ja käsitteet

1	Johdanto	1
2	Raportoinnin suunnittelu	2
2.1	Asiakkaan tarpeet	2
2.2	Saatavilla olevan datan läpikäynti	4
2.3	Raportin ulkoasun suunnittelu	4
3	Tiedon kerääminen ja esittäminen	5
3.1	Verkkoympäristön rakenne	5
3.2	Vaihtoehtoiset raportointiohjelmistot	8
3.3	Tietojen tuominen ja esittäminen	9
4	Qlik Sensen käyttäminen tiedon esittämisessä	11
4.1	Uuden raportin luonti	11
4.2	Tiedon lataus ja koostaminen	12
4.3	Ongelmien käsittely	15
4.4	Tuotettu raportti	24
5	Testaus	25
5.1	Testauksen merkitys	25
5.2	Testausmenetelmät	25
6	Raportin viimeistely ja toimittaminen	27
7	Yhteenveto	30
	Lähteet	32
	Liitteet	
	Liite 1. Aiemmin käytetty raportti	
	Liite 2. Lähtötiedot (esimerkki)	
	Liite 3. Raporttien käyttöohje	
	Liite 4. Valmis raportti (esimerkki)	

Lyhenteet ja käsitteet

CSV	Tekstimuotoinen tietorakenne, jossa tietueet erotetaan kenttäerotinmerkkiä käyttäen
MAC-osoite	Verkkolaitteen yksilöllinen laiteosoite
MSE	Langattomaan verkkoon paikannuspalvelua tuottava järjestelmä
Raportti	Insinööriyön tuloksena laaditun järjestelmän tuottama dokumentti tapahtumaverkon käytöstä
SSID	Langattoman verkkopalvelun tunniste
SQL	Relaatiotietokantojen kyselykieli
Tilaaaja	Insinööriyön tilaaja, Netcrew ry
Tapahtumajärjestäjä	Tapahtuman järjestävä organisaatio, jonka toimeksiannosta Netcrew ry toteuttaa verkon
Yhteistyökumppani	Tapahtumaa tukeva organisaatio

1 Johdanto

Työn tarkoituksena on luoda langattomien lähiverkkojen käytön esittämiseen helppo-
käyttöinen raportointijärjestelmä, joka on sovellettavissa erilaisten tapahtumien tarpei-
siin.

Työn tilaajana on Netcrew Ry, joka toteuttaa erilaisten tapahtumien langattomia ja lan-
gallisia tietoverkkoja tapahtumajärjestäjien toimeksiannosta. Tilaajan toiminta on voittoa
tavoittelematonta vapaaehtoistyötä. Toimeksiannoista valtaosa liittyy verkkopelitapahtu-
miin. Haastavat tapahtumaverkkoympäristöt ovat tarjonneet toimeksiantajalle mahdolli-
suuden kerätä kokemusta ja kehittyä tapahtumaverkkojen toteuttajana. Tapahtumajär-
jestäjiltä saatu hyvä palaute on osaltaan avannut ovia uusiin tapahtumiin. Verkkojen to-
teuttaminen mainittuihin muihin tapahtumiin ei kuulu tilaajan pääasialliseen toimintaan.
Kaupallisten toimijoiden rajalliset valmiudet tarjota ja toteuttaa laajoja langattomia tapah-
tumaverkkoja on tuonut tilaajalle mahdollisuuksia toimia yhteistyössä muutamien suur-
ten yleisötapahtumien kanssa ja toteuttaa niiden verkkoympäristöjä.

Toteutettavat verkot räätälöidään aina vastaamaan tapahtuman tarpeita. Tämän vuoksi
verkkototeutukset eroavat merkittävästi toisistaan tarjottavien verkkopalveluiden, katta-
vuuden ja kuuluvuusalueiden osalta. Tapahtuman järjestäjälle toimitetaan tyypillisesti ta-
pahtuman jälkeen raportti verkon käytöstä. Tuotettava raportti kertoo tapahtuman järjes-
täjälle tietoa verkon käyttäjämääristä, verkon kuormituksesta ja käytön jakautumisesta
ajan mukaan. Raportointi tulee pystyä esittämään kullekin langattomalle verkkopalvelulle
erikseen ja tarvittaessa koontiraporttina.

Aiemmin haasteena on ollut raportoinnin vaatima työmäärä. Raportointi on tehty pitkälle
käsityönä, useista eri lähteistä koottuna. Syynä tähän on, että vaihtuviin tarpeisiin ei ole
löydetty yhtä soveltuvaa järjestelmää, joka mahdollistaisi tarvittavan tiedon esittämisen
ja joustavan raportoinnin erilaisista verkkoympäristöistä.

Raportointitarve on tyypiltään erityislaatuinen, koska toimeksiantajien vaatimukset rapor-
toinnin sisällön suhteen vaihtelevat suuresti. Yleisötapahtumissa kävijöille tarjottava
verkkoyhteys on konseptina verrattain nuori, ja yhteistyökumppanit ovat vasta löytä-
mässä mahdollisuutta saada näkyvyyttä tarjoamalla yhteyksiä tapahtumien kävijöille.

Työn tilaajan käyttämä raportointimalli on muotoutunut kokemusten kautta vastaamaan eri toimeksiantajien vaatimuksia. Verkon käyttöä kuvaava raportti on tapahtuman yhteistyökumppanille ainoa konkreettinen todiste siitä, kuinka hyvin tai huonosti tapahtuman kävijät ovat käyttäneet tarjottua palvelua ja minkälaista näkyvyyttä yhteistyökumppani on tällä tavalla saanut.

2 Raportoinnin suunnittelu

2.1 Asiakkaan tarpeet

Tässä luvussa käsitellään raportointia yleisellä tasolla ja lisäksi paneudutaan käsiteltävän aiheen erityispiirteisiin. Raportoinnissa yleisesti on tärkeää tietää mitä tietoa asiakas haluaa raportillaan nähdä ja mitkä tiedoista ovat erityisen merkityksellisiä. Työtä helpotti se, että tilaaja on jo aiemmin tehnyt raportointia toimeksiantajilleen, joten minulla oli varsin kattavasti lähtötietoja käytössäni. Liitteessä 1 on esitelty aiemmin käytetty raporttipohja. Raportit on aiemmin joka kerta koostettu käsin ja tiedot on kerätty useista eri lähteistä. Käsityön haittana ovat myös mahdolliset inhimilliset virheet, joita on vaikea huomata. Tämän vuoksi raportoinnin automatisoinnilla saadaan selkeää hyötyä tapahtumien lisääntyessä, kun manuaalinen työmäärä vähentyy. Automatisointi mahdollistaa myös raporttien nopean toimittamisen ja esimerkiksi tapahtuman aikaisten väliraporttien tuottamisen toimeksiantajalle.

Olemassa olevan raporttipohjan puuttuminen lisäisi työmäärää merkittävästi. Raporttipohjaa laadittaessa tulee ymmärtää, mitä ja miten raportin lukijalle halutaan esittää. Hyvän lopputuloksen kannalta raportin laatijan tulee sisäistää käytettävissä olevien lähtötietojen rakenne ja hyödyntämismahdollisuudet. Hyvän raporttipohjan luomisvaiheessa voidaan tarvita useita iteraatiota tilaajan ja raportin laatijan kesken.

Kun on selvitetty, mitä tietoa asiakas haluaa raporteilla nähdä, tulee seuraavaksi selvittää mistä tarvittavat lähtötiedot saadaan. Tässä työssä lähtötiedot saadaan suoraan olemassa olevasta järjestelmästä CSV-muodossa. Esimerkki lähtötiedoista on esitetty liitteessä 2. Vaihtoehtoisesti tieto voitaisiin saada myös esimerkiksi Excel-tiedostoista tai tietokannoista riippuen käytettävän järjestelmän tarjoamista mahdollisuuksista ja rajapinnoista.

Raportoinnin hyödyt

Toimeksiantajalle, tai sen yhteistyökumppaneille, on tärkeää nähdä miten, missä ja paljonko verkkoa on käytetty. Tämän avulla voidaan perustella verkon tarve esimerkiksi seuraavaan tapahtumaan ja mahdollisesti jopa kattavamman verkon toteutustarve. Toimeksiantajilla on usein sopimuksia yhteistyökumppanien kanssa, minkä vuoksi tarjottava verkkopalvelu voidaan nimeätä yhteistyökumppanin tai tuotteen mukaan. Käytännössä kysymyksessä on siis näkyvyyden ostaminen yhteistyökumppanille yleisötapahtumassa. Jos verkko ei toimi kunnolla tai sitä ei vain käytetä, ei siitä ole yhteistyökumppaneille hyötyä, ja taas jos verkko on ollut kovassa käytössä ja toiminut hyvin, ovat yhteistyökumppanit saaneet hyvää mainosta. Tuotettavan raportin avulla toimeksiantaja tai sen yhteistyökumppani pystyy muodostamaan käsityksen siitä, miten paljon tarjottua palvelua on käytetty. Vastaavasti yhteistyökumppanin ollessa kyseessä voidaan raportin avulla mitata saavutettua näkyvyyttä.

Tapahtumien jälkeen raportointi antaa tilaajalle hyödyllistä tietoa tulevien tietoverkkojen suunnittelua ja rakentamista varten. Toimeksiantajat voivat olla hyvin erilaisia tahoja, kuten esimerkiksi verkkopelitapahtumat, musiikkifestivaalit tai messutapahtumat. Ympäristön ja verkon käyttäjäkunnan vaihtuminen tekee verkon suunnittelusta haasteellisen. Mitä enemmän pohjatietoa on käytettävissä, sen laadukkaampaan toteutukseen päästään. Raporteista voidaan jälkikäteen tarkastella, kuinka moni kävijä on verkkoa käyttänyt, kuinka pitkiä aikoja käyttäjät ovat verkkoa käyttäneet ja mihin aikaan. Raportointi antaa kuvan mahdollisista pullonkauloista verkossa, jos niitä on ollut. Tulevaisuudessa ongelmia voidaan korjata ja verkot suunnitella paremmin vastaamaan senhetkisiä tarpeita.

Lisäksi jokainen raportti auttaa ymmärtämään, millaista lähtötietoa tarvitaan tilaajalta, jotta verkon mitoitus saadaan vastaamaan tarpeita mahdollisimman hyvin. Sekä yli- että alimitoitus ovat ongelmia. Ylimitoitus tarkoittaa, että verkkoon tarvitaan arvokkaampia laitteita ja/tai sen rakentamiseen kuluu turhan paljon aikaa saatavaan hyötyyn nähden. Alimitoitus taas tarkoittaa, että verkon käyttäjillä on ollut selkeitä ongelmia ja osa ei ole ehkä saanut verkkoa toimimaan lainkaan.

2.2 Saatavilla olevan datan läpikäynti

Kun asiakas tai tilaaja on toimittanut lähdemateriaalin tai pääsyn sopivaan tietokantaan, voidaan käydä läpi saatavilla oleva aineisto. Lähdemateriaalien läpikäynnin perusteella voidaan päätellä, onko lähtötiedoista mahdollista koostaa kaikki kaaviot tai taulukot, joita tilaaja haluaa nähdä. Tässä yhteydessä on myös aiheellista selvittää, ovatko lähtötiedot jatkokäsittelyyn soveltuvassa muodossa. Lähtötietojen esitysmuodon muokkaamistarpeen toteamisen jälkeen tulee selvittää, missä pisteessä tietojen muokkaaminen on optimaalisinta suorittaa. Tällaisia tilanteita on esimerkiksi silloin, kun samoja lähtötietoja käytetään monessa eri järjestelmässä ja ne kaikki tarvitsevat tiedot yhtenäisessä muodossa.

Tietojen käsittelyssä on otettava huomioon myös voimassa oleva lainsäädäntö. Tämän työn toimeksiannon tapauksessa lähtötieto sisältää yksilöllisiä laitetunnisteita (MAC-osoite), joiden perusteella on mahdollista seurata yksittäistä päätelaitetta tapahtumassa. Tämän tyyppisten tietojen käytöstä on säädetty tietoyhteiskuntakaareissa [1, 142 §]. Esitettäessä tietoja yhteenvetoina lain vaatimus täytetään eikä tuloksista ole mahdollista tunnistaa yksittäistä käyttäjää.

2.3 Raportin ulkoasun suunnittelu

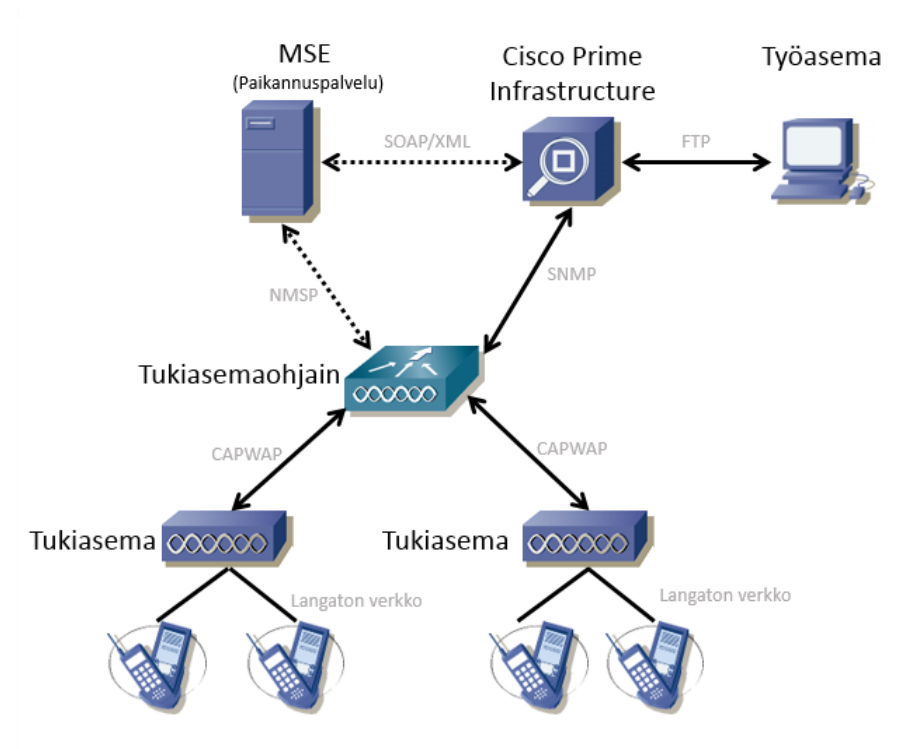
Raportin ulkoasu tulee suunnitella niin, että tiedot ovat helposti luettavia ja ymmärrettäviä. Etusivulla tulisi esitellä yhteenvetona kokonaisuuden kannalta tärkeät asiat, joita muilla sivuilla tai välilehdillä käsitellään tarkemmin. Sen lisäksi kaavioista tulee selkeästi käydä ilmi, mitä tietoa ne sisältävät ja millaisella mitta-asteikoilla. Samaan aihepiiriin liittyvät kaaviot tulisi esittää samalla välilehdellä mahdollisuuksien mukaan.

3 Tiedon kerääminen ja esittäminen

3.1 Verkkoympäristön rakenne

Insinööriyön tilaajalla on käytössään valmis infrastruktuuri, joka mahdollistaa verkon käyttötietojen keräämisen CSV-muotoon ja siirtämisen edelleen muihin järjestelmiin tietojen jatkokäsittelyä varten. Infrastruktuuri koostuu kontrolleripohjaisesta langattomasta verkkoympäristöstä ja sen oheispalveluista.

Ympäristöön kuuluvat langattoman verkon tukiasemat yhdistetään tukiasemaohjaimen (kontrolleri). Ratkaisu mahdollistaa langattoman verkkoympäristön keskitetyn hallinnan sekä käyttötietojen ja hälytysten välittämisen edelleen muihin järjestelmiin. Kuvassa 1 on esitetty järjestelmän komponentit ja niiden väliset yhteydet.



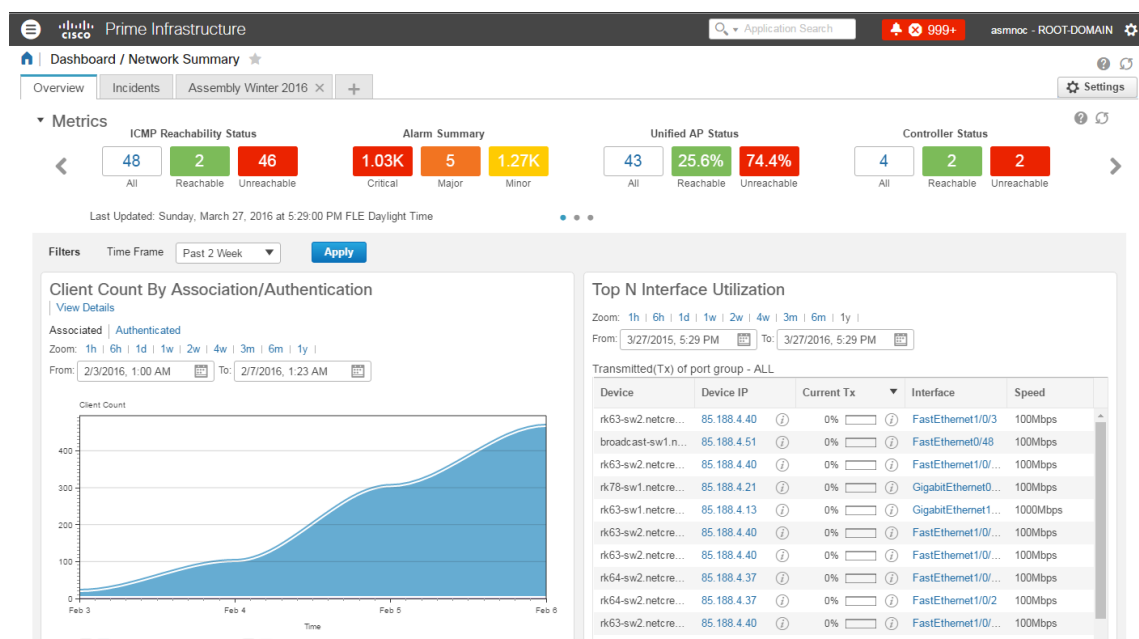
Kuva 1. Netcrew'n verkon rakenne raportoinnin kannalta [2].

Langattoman verkon käyttäjät liittyvät verkkoon tukiasemien kautta. Tukiasemilla on jatkuva yhteys niiden toimintaa ohjaavaan tukiasemaohjaimen, jonka kautta myös käyttäjien liikenne välitetään.

Tukiasemaohjain on liitettyä sekä Ciscon MSE -paikannuspalveluun että Cisco Prime Infrastructure -hallintajärjestelmään, jota käytetään kerättävän tilastotiedon taltioimiseen. MSE -paikannuspalvelua en käytä tässä työssä, mutta se on esitetty mukana, koska se kuuluu olemassa olevaan verkkoon. Cisco Prime Infrastructure voi olla joko fyysinen laite tai virtuaalipalvelin. Tilaajalla on käytössään virtuaalipalvelinversio Cisco Prime Infrastructure -järjestelmästä.

Cisco Prime Infrastructure -hallintajärjestelmä

Cisco Prime Infrastructure on tuote, jonka avulla on mahdollista ylläpitää langattomia ja langallisia verkkoja. Sen avulla voidaan seurata muun muassa verkosta tulevia hälytyksiä ja muuttaa langattomien tukiasemien asetuksia [3]. Sen lisäksi Cisco Prime Infrastructure tarjoaa mahdollisuuden luoda raportteja verkon toiminnasta. Sen tuottamat raportit ovat oletuksena CSV-muotoisina, mutta ne on mahdollista saada myös PDF-muodossa. Kuvassa 2 Cisco Prime Infrastructure -järjestelmän etusivu.



Kuva 2. Cisco Prime – palvelun etusivu.

Uuden raportin luonti on suhteellisen helppoa. Valikosta löytyy suoraan kohdat "Reports" ja "Report Launch Pad". Koska tarkoitus oli kerätä nimenomaan yhteyskohtaisia tietoja valitsin erilaisista raporttimahdollisuuksista vaihtoehdon "Client sessions". Samalta riviltä kanssa löytyi myös linkki uuteen raporttiin (New). Kuvassa 3 on esitetty uuden raportin luontinäköymä.

The screenshot shows the Cisco Prime Infrastructure web interface for configuring a new report. The interface is divided into two main sections: 'Settings' and 'Schedule'.

Settings:

- Create reports in current and each sub Virtual Domains
- View sub Virtual Domains
- Report Title: [Text Input]
- Report By: All
- Report Criteria: All [Edit]
- Reporting Period: Select a time period...
- From: [Text Input]
- To: [Text Input]
- Customize Report: [Customize] Customize the data for this report

Schedule:

- Scheduling: Enable
- Export Format: CSV
- Destination:
 - File: /localdisk/ftp/reports/ClientSessions/<ReportTitle Name>_<yyyyymmdd>_<HHMMSS>.csv
 - Email: [Text Input]
- Start Date/Time: 03/27/2016 05:25 PM (GMT+2:0) EET
- Current Server Time: 27-Mar-2016,17:23:14 EEST
- Recurrence:
 - No Recurrence
 - Hourly
 - Daily
 - Weekly
 - Monthly

Buttons at the top and bottom of the interface include: Run, Save, Run and Save, Save and Export, Save and Email, and Cancel.

Kuva 3. Cisco Prime Infrastructure: Uuden raportin luonti.

Uuden raportin asetukset ovat yksiselitteisiä. Tässä tapauksessa raportteja ei tarvitse ajastaa, vaan ne ajetaan vain kerran tapahtuman jälkeen. Tällöin raportti sisältää koko tapahtuman tiedot.

Uutta raporttia luotaessa sille annetaan nimi. Raportin nimellä ei ole jatkotoimenpiteiden kannalta merkitystä, mutta nimen on hyvä olla kuvaava, jotta tulevaisuudessa raportti pystytään yhdistämään oikeaan tapahtumaan. Sen lisäksi voidaan valita esimerkiksi, halutaanko tietoihin langattomat verkot vai langalliset verkot tai rajataanko tietoa esimerkiksi SSID:n perusteella. Lopuksi on mahdollista määrittää, miltä aikaväliltä raportit ajetaan. Yleensä tapahtumien raportoinnissa käytetään tapahtuman alku- ja loppuaikoja.

Cisco Prime Infrastructuren asetuksissa on määritelty tiedon säilymisaika, joka on oletuksena 90 päivää. Raportit tulee luoda tämän ajan sisällä, muussa tapauksessa yhteyksistä kerätyt tiedot menetetään. Myös valmiiden CSV-muotoisten tiedostojen säilytysaika on rajallinen. Tämän vuoksi luodut tiedostot tulee tallentaa erikseen sovittavaan yhteiseen tallennuspaikkaan jatkotoimenpiteitä ja mahdollisia myöhemmin tehtäviä analysointeja varten.

3.2 Vaihtoehtoiset raportointiohjelmistot

Raporttien tuottamiseen soveltuvien ohjelmistojen selvittäminen ja vertailu olivat varsinaisen työn suorittamiseen liittyviä ensimmäisiä toimenpiteitä. Vaihtoehtoina olivat Cisco Prime Infrastructuren oma raportointityökalu, Qlik Sense, Chrystal Reports ja Tableau. Nämä vaihtoehdot valittiin, koska niistä oli joko ennestään kokemusta tai tiedettiin niiden olevan toiminnaltaan sopivia tähän käyttöön. Työtä tehdessä tuli huomioida rajallinen aika uuden ohjelmiston opetteluun ja rajallinen budjetti.

Cisco Prime Infrastructuren raportointityökalu oli ensimmäinen vaihtoehto, koska se oli jo valmiiksi olemassa ja tiedot käyttäjien yhteyksistä kerätään automaattisesti tähän järjestelmään. Monipuolisista raportointimahdollisuuksista huolimatta Cisco Prime Infrastructure ei kykene tuottamaan raportteja toivotussa muodossa. Esimerkkinä tästä voidaan mainita usean langattoman verkkopalvelun käyttäjämäärien esittäminen eriteltyinä yhdessä kuvaajassa. Järjestelmässä on tarjolla hyviä raportteja, mutta tietoja ei voida rajata halutulla tarkkuudella.

Toinen vaihtoehto oli Qlik Sense. Siitä oli valmiiksi eniten kokemusta, ja se on helppokäyttöinen verrattuna muihin tutkittuihin vaihtoehtoihin. Ohjelmistoon tutustuttaessa varmistui, että Qlik Sensen ominaisuuksien avulla tarvittavien raporttien luominen on mahdollista. Tilaajan rajallisen budjetin vuoksi Qlik Sense -ohjelmiston käyttämisen ilmaisuus – mikäli käytetään uusinta kulloinkin tarjolla olevaa versiota ohjelmistosta – todettiin merkittäväksi eduksi tilaajan kanssa käydyissä keskusteluissa. Palvelun mukana tulee myös käyttöön pilvipalvelu, joka mahdollistaa raporttien jakamisen muille pilvipalveluun kirjautuneille käyttäjille.

Kolmas vaihtoehto oli Chrystal Reports. Aiemman kokemuksen perusteella ohjelmisto voisi sopia tähän käyttöön hyvin. Chrystal Reports on monipuolinen, mutta käytettävyys ei ole samalla tasolla Qlik Sensen kanssa. Poiketen aiemmin mainituista ohjelmistoista Chrystal Reports on maksullinen, eikä siitä ollut saatavilla ilmaista kokeiluversiota.

Viimeinen vaihtoehto oli Tableau. Muista vaihtoehdoista poiketen siitä ei ollut aiempaa kokemusta, mutta se pääsi listalle, koska se on työyhteisössä käytetty työkalu ja saatujen kommenttien perusteella se voisi sopia tähän käyttöön. Tableausta oli saatavilla vain

kuukauden ilmainen kokeilujakso. Lisäksi sen käyttö vaatisi huomattavan paljon enemmän uuden opiskelua muihin vaihtoehtoihin verrattuna. Selkeitä etuja muihin ohjelmistoihin verrattuna ei ollut.

Käytettäväksi työkaluksi valikoitui Qlik Sense, koska siinä että raporttipohjan laatiminen onnistui kohtuullisella työmäärällä, tarjolla oleva pilvipalvelu mahdollistaa tuotoksien jakamisen sekä hinnoittelupolitiikka on insinööriyön tilaajan budjettiin sopiva. Lisäksi ohjelmistosta entuudestaan olevat käyttökokemukset puolsivat valintaa.

3.3 Tietojen tuominen ja esittäminen

Cisco Prime Infrastructure -järjestelmästä saatavat lähtötiedot ovat CSV -muotoisessa tiedostossa, josta on esimerkki liitteessä 2. Taulukossa jokainen käyttäjän yhteys langattomaan verkkoon on tilastoitu omalla rivillään. Riviltä löytyvät tiedot muun muassa yhteyden aloitusajasta, verkon SSID, käyttäjän päätelaitteen valmistaja ja muut yksilöivät tiedot. Qlik Sense käsittelee tietoja muistinvaraisessa tietokannassaan, joten tiedot on erikseen tuotava ohjelmaan. Tässä yhteydessä tietoja on mahdollista muotoilla uudestaan raportoinnin edellyttämään muotoon. Tämä tapahtuu hyvin paljon SQL: ää muistuttavalla kielellä ohjelman omassa käyttöliittymässä. Esimerkiksi tässä työssä kellonaikainfo tieto täytyy muuttaa tekstitiedosta aikaleimaksi, jotta myöhemmin sen avulla voidaan tehdä kaavioita ajan perusteella.

Kun tiedot on ladattu ohjelmaan, on mahdollista aloittaa itse analyysin tuottaminen. Ensimmäisessä vaiheessa tulee pohtia, mitä tietoa halutaan esittää. Tilaajan aiemmin käytetty raportointimalli on esitetty liitteessä 1, ja vähintään vastaavien tietojen esittämistä tilaajan toivoo. Kun esitettävä tieto on päätetty, pitää pohtia, missä muodossa tieto on järkevintä esittää. Vaihtoehtoja on useita, esimerkiksi pylväskaavio, viivakaavio, ympyräkaavio tai paljas teksti [4]. Seuraavaksi esitellään erilaisia kuvaajia ja niiden käyttökohteita.

Ympyräkaavio

Ympyräkaaviossa jokainen kategoria esitetään omana sektorinaan. Ympyräkaavio havainnollistaa kategorioiden osuuksia kokonaisuudesta, kunhan kategorioita on määrällisesti vähän. Osuudet voidaan näyttää kaaviossa kuvan lisäksi joko prosentteina kokonaisuudesta ja/tai suoraan lukumäärinä. Ympyräkaaviolla voidaan esimerkiksi esittää, kuinka suuret prosenttiosuudet eri jäätelövalmistajat saivat kaupan jäätelön myynnistä edellisenä kuukautena.

Pylväskaavio

Pylväskaavio soveltuu hyvin erilaisten määrien havainnollistamiseen. Y-akselilla ilmoitetaan määrä ja X-akselilla joko esimerkiksi aika tai kategoria. Pylväskaaviolla voidaan esittää kaupan jäätelön myynti eri päivinä.

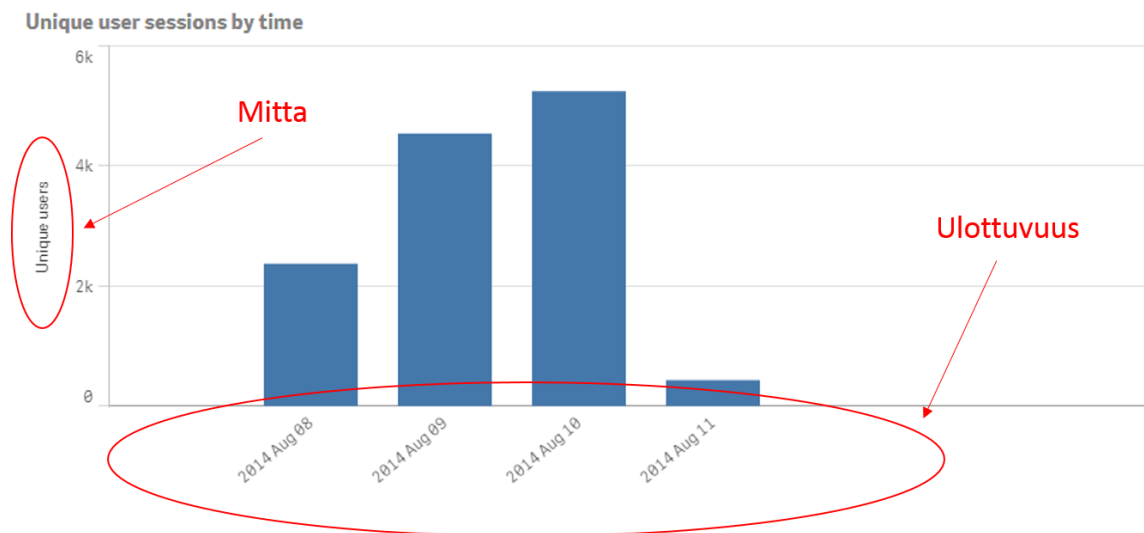
Viivakaavio

Viivakaaviot ovat käytännöllisiä, kun halutaan havainnollistaa jatkumoa. Saman tiedon voi usein näyttää myös pylväskaaviona, mutta se ei ole yhtä selkeää. Viivakaaviolla voidaan esittää esimerkiksi, kuinka paljon jäätelöä on myyty eri päivinä kesän aikana. Yhdistämällä samaan viivakaavioon useita tietoja, voidaan tuoda esiin asioiden välistä riippuvuutta: esimerkiksi esittää jäätelömyyntikaaviossa lämpötila, jolloin lukija voi kaaviosta nähdä, kuinka lämpötilanvaihtelut vaikuttavat jäätelömyyntiin.

Kaaviotyypin valintaan vaikuttaa myös se, halutaanko kaaviossa esittää tietojen jakaumaa vai vertailla eri arvoja keskenään ja halutaanko kaaviossa esittää tietoa ajan funktiona vai ei ja kuinka monta asiaa halutaan samassa kaaviossa esittää. Kaikkia tietoja ei voida esittää kaikilla kaaviotyypeillä tai tiedot eivät ole selkeästi luettavissa väärästä kaaviotyypistä. [5.]

Kun kaaviotyyppi on valittu, pitää rakentaa itse kaavio. Hankalin vaihe kaavioiden tekemisessä on ymmärtää erilaiset kaaviot ja miten ne muodostuvat. Qlik Sense käyttää termejä ulottuvuus (dimension) ja mitta (measure). Ensimmäinen kysymys on: "Mitä haluan esittää?" Tässä tapauksessa vastaus voi olla "yhteyksien määrää tukiasemalle x". Tämä tieto on mitta. Toinen kysymys on: "Kuinka usein haluamme tiedon laskea?" Vastauk-

senä tässä tapauksessa olisi "kerran tunnissa". Tämä tieto on ulottuvuus. Kaavioon halutaan siis mitaksi yhteyksien määrä ja ulottuvuudeksi aika. Kuvan 4 kaaviossa mitta (measure) on uniikkien käyttäjien lukumäärä (unique users) ja ulottuvuus (dimension) on päivämäärä (date).

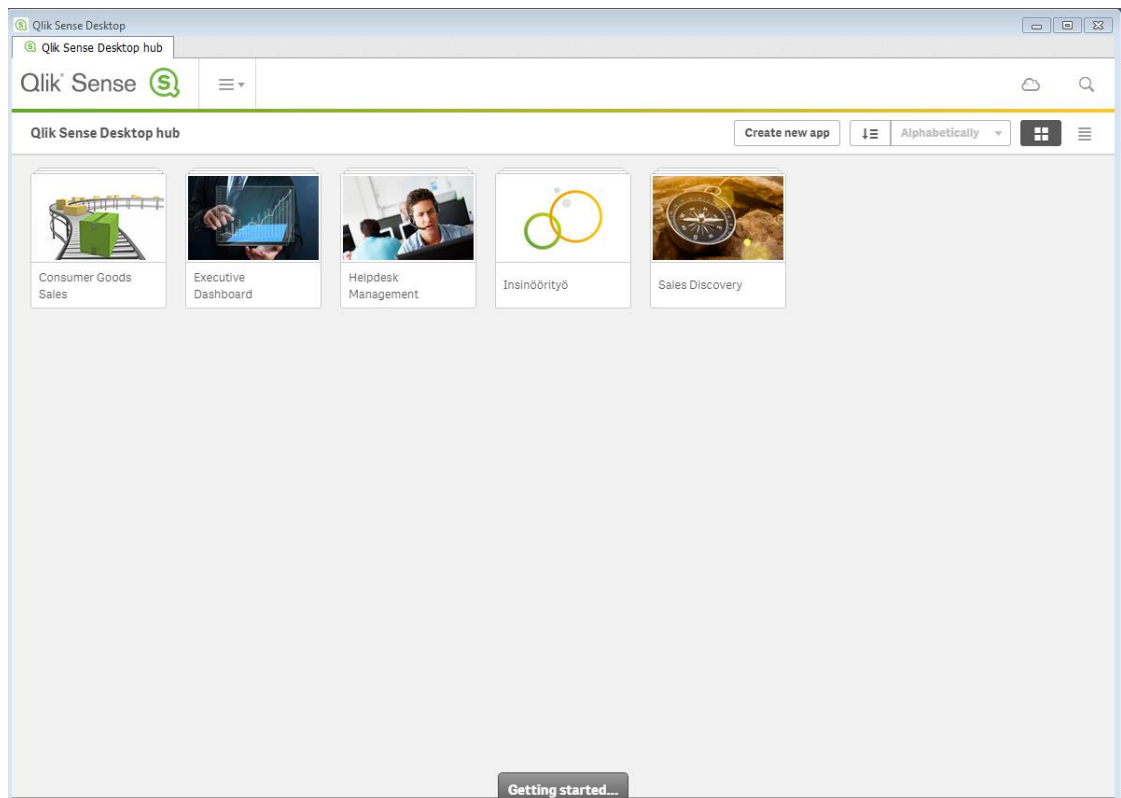


Kuva 4. Esimerkki ulottuvuuksien ja mittojen käyttämisestä kaaviossa. [6]

4 Qlik Senseen käyttäminen tiedon esittämisessä

4.1 Uuden raportin luonti

Tässä luvussa käsitellään työkaluksi valitun Qlik Sense -ohjelmiston käyttämistä uuden raportin luomisessa. Ohjelma käynnistyy aloitusikkunaan nimeltä "Qlik Sense Desktop Hub". Tällä sivulla näytetään kaikki ohjelman tunnistamat projektit (kuva 5).



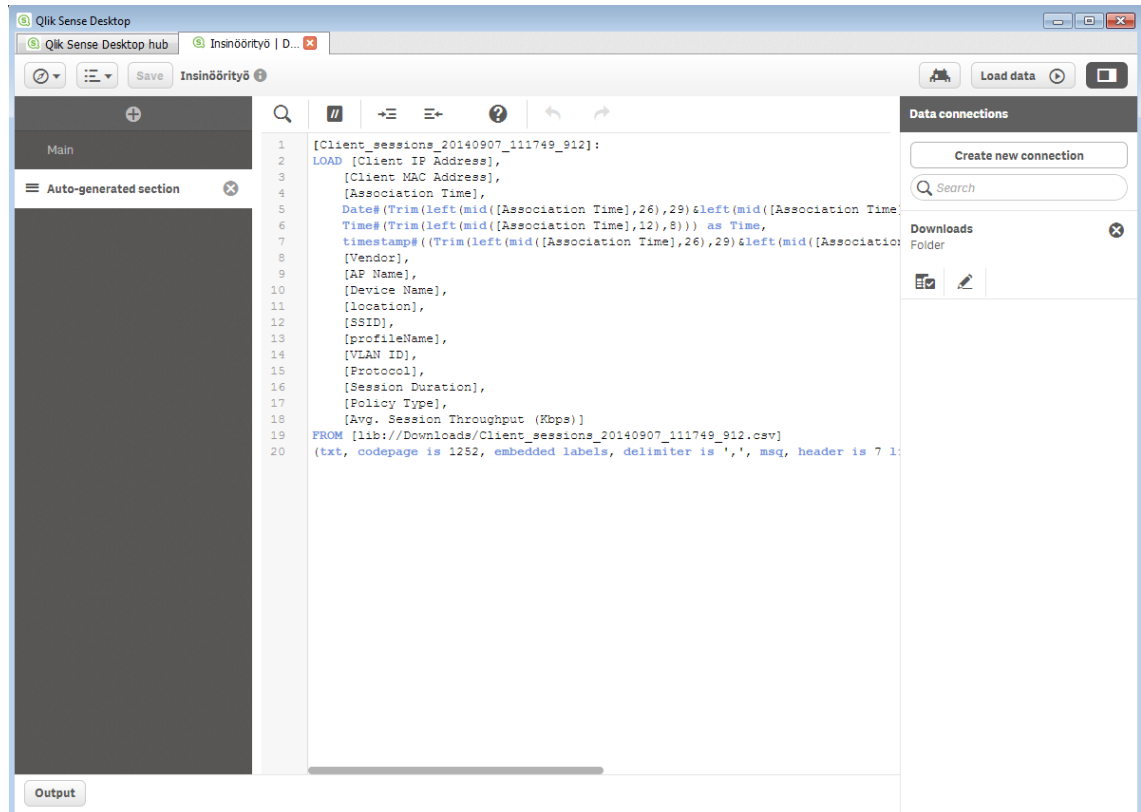
Kuva 5. Qlik Sense Hub (aloitusikkuna).

Uuden analyysin teko aloitetaan painamalla "Create new app" -nappia. Ohjelma pyytää antamaan nimen analyysille, ja sen jälkeen analyysin voi avata painamalla "Open"-nappia. Ohjelmaan on mahdollista tuoda tietoja kahdella tapaa, joko käyttämällä ohjattua latausta (Add data) tai tekemällä latauksen käsin Data load managerilla. Jälkimmäinen vaihtoehto on järkevää valita vain, jos Qlik Sense -ohjelman käyttämä skriptikieli on tuttua ennestään.

4.2 Tiedon lataus ja koostaminen

Analyysi tehtiin käyttämällä ohjattua latausta. Tiedon lähteeksi (data source) valittiin työaseman levyllä tallennettu, Cisco Prime Infrastructure -ohjelmistosta tuotettu CSV-muotoinen tiedosto. Qlik Sense pyrkii aluksi tunnistamaan, missä muodossa tieto on, ja tekemään automaattisesti skriptin, jolla ladataan tiedot muistiin. Ennen varsinaista tietojen lataamista on mahdollista muuttaa esimerkiksi solujen nimiä ja sitä, mistä riviltä varsinaisen tieto alkaa. Kun skripti on valmis, painetaan "Load data" -nappia. Tietojen latausta

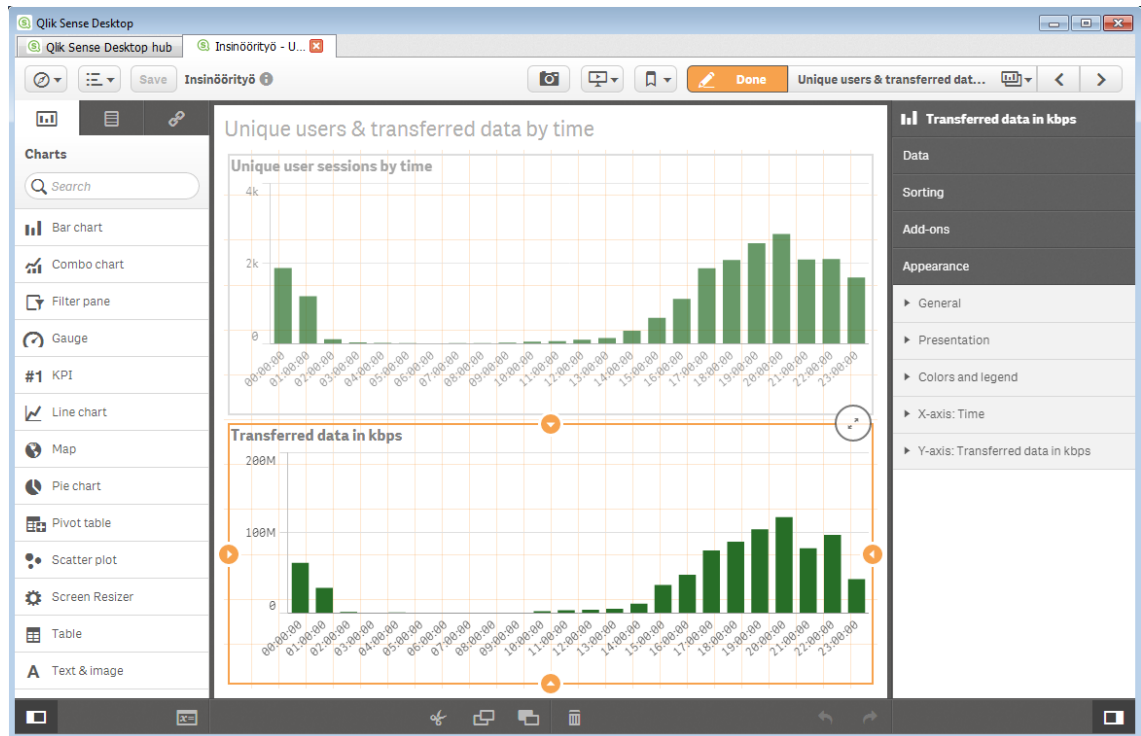
voi seurata päälle avautuvasta ikkunasta. Tässä ikkunassa näytetään myös tietojen latauksessa tapahtuneet virheet. Tiedot ladataan muistinvaraiseen Qlik Senseen omaan tietokantaan. Kuvassa 6 on esitetty Data Load Manager –ikkuna, jossa voidaan määrittellä ladattavat tiedot ja esimerkiksi niiden esitysmuoto.



Kuva 6. Data load manager -ikkuna, jossa määritellään, mitä tietoa ladataan.

Jos ladattavaa tietoa on enemmän ja useista eri lähteistä kannattaa, muistaa että samannimiset solut yhdistetään automaattisesti keskenään. Tämän vuoksi kahteen eri asiaan liittyvien solujen nimet tulee muuttaa erilaisiksi tietojen latausvaiheessa. Jos kuitenkin useassa taulussa on samanniminen sarake, ne voivat muodostaa renkaan (loop) ja ohjelma jää jumiin latausvaiheeseen. Uusimmat Qlik Senseen versiot tunnistavat tällaiset virheet ja antavat virheilmoituksen.

Kun tiedot on ladattu, voidaan tehdä itse analyysi kaavioineen. Analyysiä pääsee tekemään painamalla vasemman yläkulman kompassin näköistä nappia (navigation) ja valitsemalla "App overview". Aluksi on aloitettava uusi välilehti painamalla "Create new sheet" ja antamalla välilehdelle nimi. Raportin muokkausnäkö on esitetty kuvassa 7.



Kuva 7. App overview, jossa raportin ulkoasu ja kaaviot luodaan.

Tämän jälkeen aukeaa ikkuna, jossa näkyy itse raportti. Sivun aluksi on tyhjä tila, jossa raporttia voi katsoa, mutta ei muokata. Muokkaus tapahtuu painamalla "Edit" -nappia, joka tuo esiin työkalut ja ruudukon, jonka avulla eri elementtejä, esimerkiksi pylväsdiagrammia, on helppo kohdistaa. Vasemmalla laidalla on erityyppisiä kaaviota, joita voi raahata sivulle haluamaansa paikkaan. Kun kaaviotyyppi on valittu, voi kaavion tietoja muokata oikealla olevan valikon kautta. Muokattavia asioita ovat esimerkiksi ulottuvuudet, mitat, kaavion teksti ja värit. Itse kaaviot toimivat hyvin samaan tyyliin kuin Microsoft Excelin kaaviot.

Tiettyissä elementeissä lähtötiedoista saatavaa dataa ei voida käyttää sellaisenaan, vaan tietoa on koostettava ennen esittämistä. Qlik Sense tarjoaa seuraavat mahdollisuudet tietojen koostamiselle: summaus (Sum), määrä (Count), keskiarvo (Avg), minimiarvo (Min) tai maksimiarvo (Max).

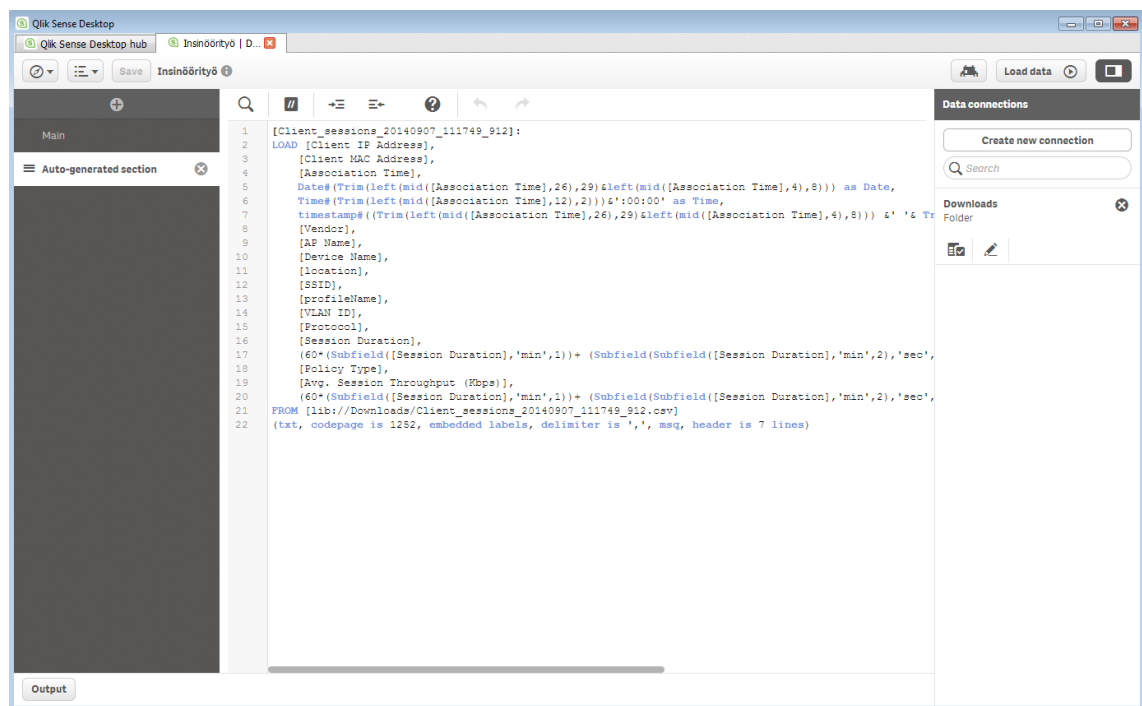
Osassa kaavioista tietoa koostetaan useammalta riviltä yhteen, ja tällöin tulee raporttia laatiessa pohtia, mitä tietoa halutaan näyttää. Raportissa on esimerkiksi mahdollista laskea sarakkeiden tietoja yhteen: monellako rivillä tiettyssä sarakkeessa on teksti 'Koira',

tai voi laskea keskiarvon tuotteen käyttäjien iästä. Jos koostamistavan unohtaa määrittellä elementtiä luotaessa, järjestelmä ei kykene tuottamaan kuvaajaa, virheellisen määrittelyn takia, vaikka lähtötiedot ovat saatavilla.

4.3 Ongelmien käsittely

Ongelmat tietojen tuonnissa

Mikäli raporttia laatiessa huomaa jonkin tiedon olevan väärässä muodossa, sitä voi muokata navigaatio-valikosta sijaitsevan Data load editor -toiminnon avulla. Data load editor on sama ikkuna, jota käytettiin alussa tiedon tuomiseen ohjelmaan. Esimerkiksi ylimääräiset merkit voidaan poistaa skriptien avulla sarakkeista. Kuvassa 8 esitetään muokattu latausskripti.



Kuva 8. Data load manager avoinna skriptin muokkausta varten.

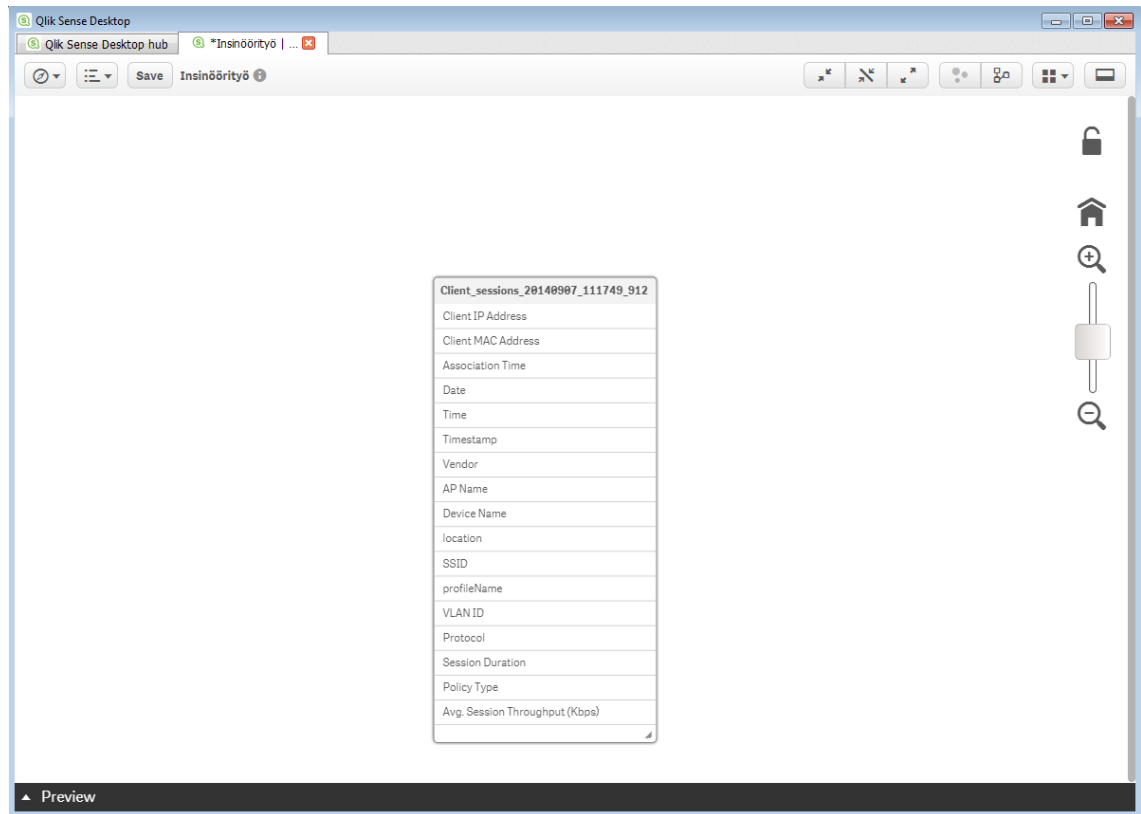
Navigaatio -valikosta saa esiin myös "Data manager" -työkalun, joka on hyödyllinen, jos ei ole varma, missä muodossa jokin tieto on tuotu tai onko tietoja ladattu lainkaan ohjelmaan. Siellä näkyvät kaikki ohjelmaan tuodut solut ja se millaista tietoa ja missä muodossa ne pitävät sisällään. Jos aiemmin on pyrkinyt skriptillä muuttamaan jonkin sarakkeen sisältöä, tältä sivulta voi tarkistaa, onko muutos onnistunut. Huomionarvoinen

seikka on se, että sivulla esitetään yleensä vain pieni osan tiedoista, ei siis aivan kaikkia rivejä. Kuvassa 9 on esitetty näkymä data load manageriin, kun työhön tarvittavat tiedot on ladattu ohjelmaan.

Client IP Address	Client MAC Address	Association Time	Date	Time	Timestamp
172.16.16.157	00:03:7e:04:c0:9d	Sat Aug 09 19:56:24 EEST 2014	2014 Aug 09	19:56:24	2014 Aug 09 19:56:24
172.16.16.157	00:03:7e:04:c0:9d	Sat Aug 09 20:01:24 EEST 2014	2014 Aug 09	20:01:24	2014 Aug 09 20:01:24
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 18:35:09 EEST 2014	2014 Aug 08	18:35:09	2014 Aug 08 18:35:09
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 19:40:23 EEST 2014	2014 Aug 08	19:40:23	2014 Aug 08 19:40:23
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 19:45:24 EEST 2014	2014 Aug 08	19:45:24	2014 Aug 08 19:45:24
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 20:05:25 EEST 2014	2014 Aug 08	20:05:25	2014 Aug 08 20:05:25
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 20:20:35 EEST 2014	2014 Aug 08	20:20:35	2014 Aug 08 20:20:35
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 22:05:55 EEST 2014	2014 Aug 08	22:05:55	2014 Aug 08 22:05:55
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 22:36:08 EEST 2014	2014 Aug 08	22:36:08	2014 Aug 08 22:36:08
172.16.17.118	00:03:aa:c6:8c:84	Fri Aug 08 23:01:11 EEST 2014	2014 Aug 08	23:01:11	2014 Aug 08 23:01:11
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 17:30:53 EEST 2014	2014 Aug 09	17:30:53	2014 Aug 09 17:30:53
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:21:07 EEST 2014	2014 Aug 09	18:21:07	2014 Aug 09 18:21:07
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:31:08 EEST 2014	2014 Aug 09	18:31:08	2014 Aug 09 18:31:08
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:41:09 EEST 2014	2014 Aug 09	18:41:09	2014 Aug 09 18:41:09
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:46:10 EEST 2014	2014 Aug 09	18:46:10	2014 Aug 09 18:46:10
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:51:10 EEST 2014	2014 Aug 09	18:51:10	2014 Aug 09 18:51:10
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 18:56:11 EEST 2014	2014 Aug 09	18:56:11	2014 Aug 09 18:56:11
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 19:01:12 EEST 2014	2014 Aug 09	19:01:12	2014 Aug 09 19:01:12
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 19:06:12 EEST 2014	2014 Aug 09	19:06:12	2014 Aug 09 19:06:12
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 20:41:39 EEST 2014	2014 Aug 09	20:41:39	2014 Aug 09 20:41:39
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 20:46:40 EEST 2014	2014 Aug 09	20:46:40	2014 Aug 09 20:46:40
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 20:56:42 EEST 2014	2014 Aug 09	20:56:42	2014 Aug 09 20:56:42
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 22:01:56 EEST 2014	2014 Aug 09	22:01:56	2014 Aug 09 22:01:56
172.16.27.41	00:03:aa:c6:8c:84	Sat Aug 09 22:32:07 EEST 2014	2014 Aug 09	22:32:07	2014 Aug 09 22:32:07
172.16.27.41	00:03:aa:c6:8c:84	Sun Aug 10 00:32:33 EEST 2014	2014 Aug 10	00:32:33	2014 Aug 10 00:32:33
172.16.27.41	00:03:aa:c6:8c:84	Sun Aug 10 00:37:33 EEST 2014	2014 Aug 10	00:37:33	2014 Aug 10 00:37:33
172.16.16.31	00:03:aa:c6:8c:84	Sun Aug 10 19:16:43 EEST 2014	2014 Aug 10	19:16:43	2014 Aug 10 19:16:43

Kuva 9. Data Manager avoinna tuodun tiedon tarkistamista varten.

"Data model viewer" on hyödyllinen työkalu kompleksisimmassa projekteissa, joissa tietoa tuodaan useista lähteistä ja taulukoita yhdistellään toisiinsa. Sieltä näkee eri taulujen riippuvuudet toisistaan. Tässä työssä on vain yksi taulu, joten tälle työkalulle ei ole käyttöä. Tämä ominaisuus on esitetty kuvassa 10.



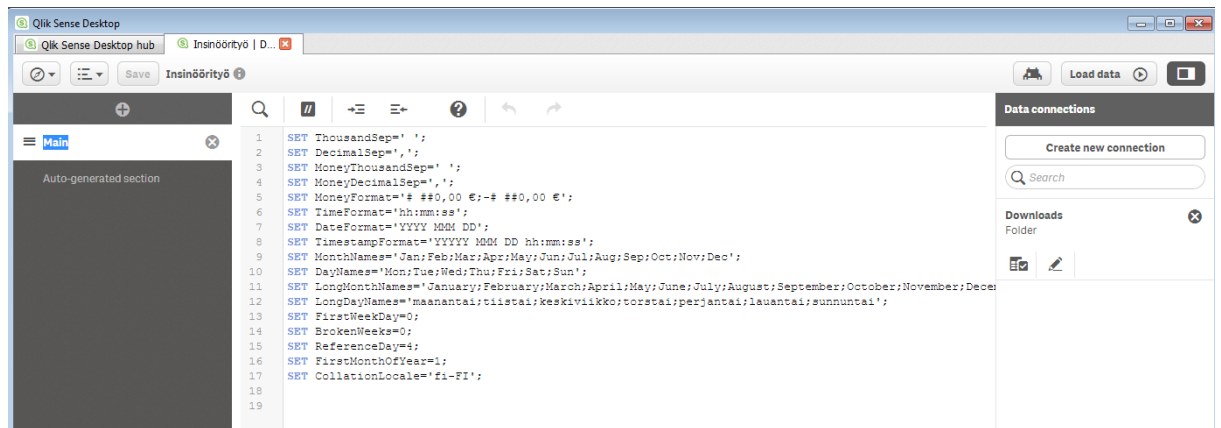
Kuva 10. Data model viewer, jossa voi katsella, miten Qlik Sense on mallintanut tietokannan muistiin.

Ongelmia raporttia tehtäessä

Kaaviota tehtäessä suurimman osan mitoista ja ulottuvuuksista saatiin suoraan taulukosta. Mutta joissakin tapauksissa tieto ei ollut halutussa muodossa, vaan jouduin muuttamaan alkuperäistä latausskriptiä. Esimerkiksi aikaleima oli alkuperäisessä CSV-taulukossa pelkkä teksti. Qlik Sensessä on myös mahdollisuus tehdä taulukoita dynaamisiksi, eli kun käyttäjä valitsee tietyn päivämäärän jossain kohdin raporttia, osa taulukoista muuttuu siten, että niistä näkee datan purettuna tuntitasolle valittuna päivänä. Tästä käytetään termiä porautuminen (drilling). Tällä tavalla saadaan havainnollisesti esitettyä tieto skaalautuvassa muodossa.

Tuotettavissa raporteissa halutaan hyödyntää porautumisominaisuutta, joten ensin on luotava uusi ulottuvuus -tietue. Luotava uusi tietue sisältää kaksi muuta ulottuvuutta: päivämäärä ja kellonaika. Alkuperäisissä lähtötiedoissa on ainoastaan ulottuvuus nimeltä ”Association Time” ja sen sisältämä tieto on tekstimuodossa, joten sitä ei voida käyttää

suoraan, vaan tieto tulee muuttua muotoon, jonka Qlik Sense ymmärtää kellonajaksi ja päivämääräksi. Alkuperäinen teksti on muotoa ”Sat Aug 09 19:56:24 EEST 2014”. Tästä tekstistä pitää siis erottaa erikseen päivämäärä ja aika. Sen lisäksi pitää miettiä, missä muodossa Qlik Sense käsittelee kellonaikaa, eli millaisessa muodossa se odottaa kellonajan olevan. Sen voi määrittää Data load manager -sivulla. Tästä esimerkki kuvassa 11.



Kuva 11. Data load managerin perusasetukset.

Kuvassa on esitetty, millaisia asetuksia tietojen latauksessa käytetään. Sieltä voidaan poimia kaksi asetusta:

```

SET TimeFormat='hh:mm:ss';
SET DateFormat='YYYY MMM DD';

```

Näitä muotoja voi siis muuttaa haluamallaan tavalla. Tässä tapauksessa pidetään nämä asetukset ja siirrytään käsittelemään sitä, miten tiedot ladataan alkuperäisestä dokumentista Qlik Sensen sisäiseen muistiin. Alla on esitetty kolme riviä, jotka lisätään alkuperäiseen lautausskriptiin. Niiden avulla aikatiedot saadaan ladattua oikeassa muodossa.

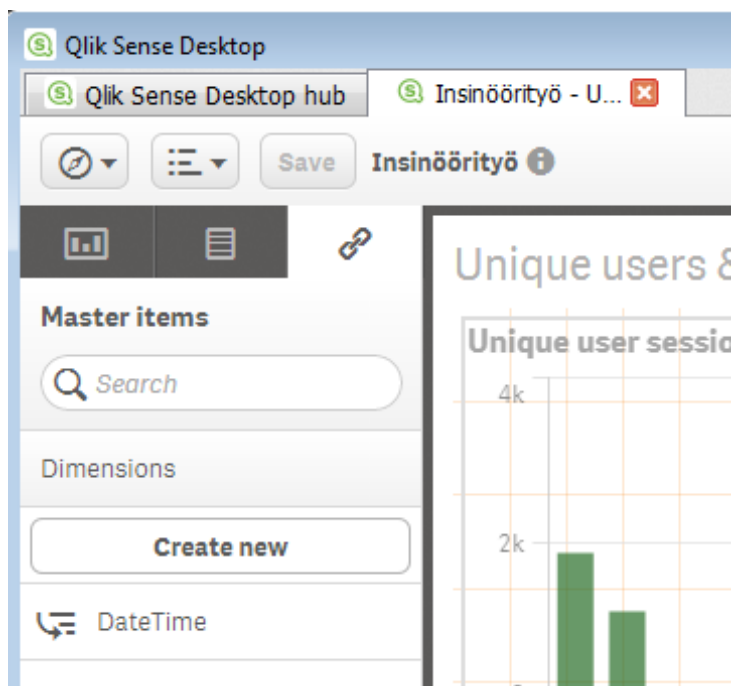
```

Date#(Trim(left(mid([Association Time],26),29)
&left(mid([Association Time],4),8))) as Date,
Time#(Trim(left(mid([Association Time],12),8))) as Time,
timestamp#(Trim(left(mid([Association Time]
,26),29)&left(mid([Association Time],4),8))) &' '&
Trim(left(mid([Association Time],12),8))) as Timestamp,

```

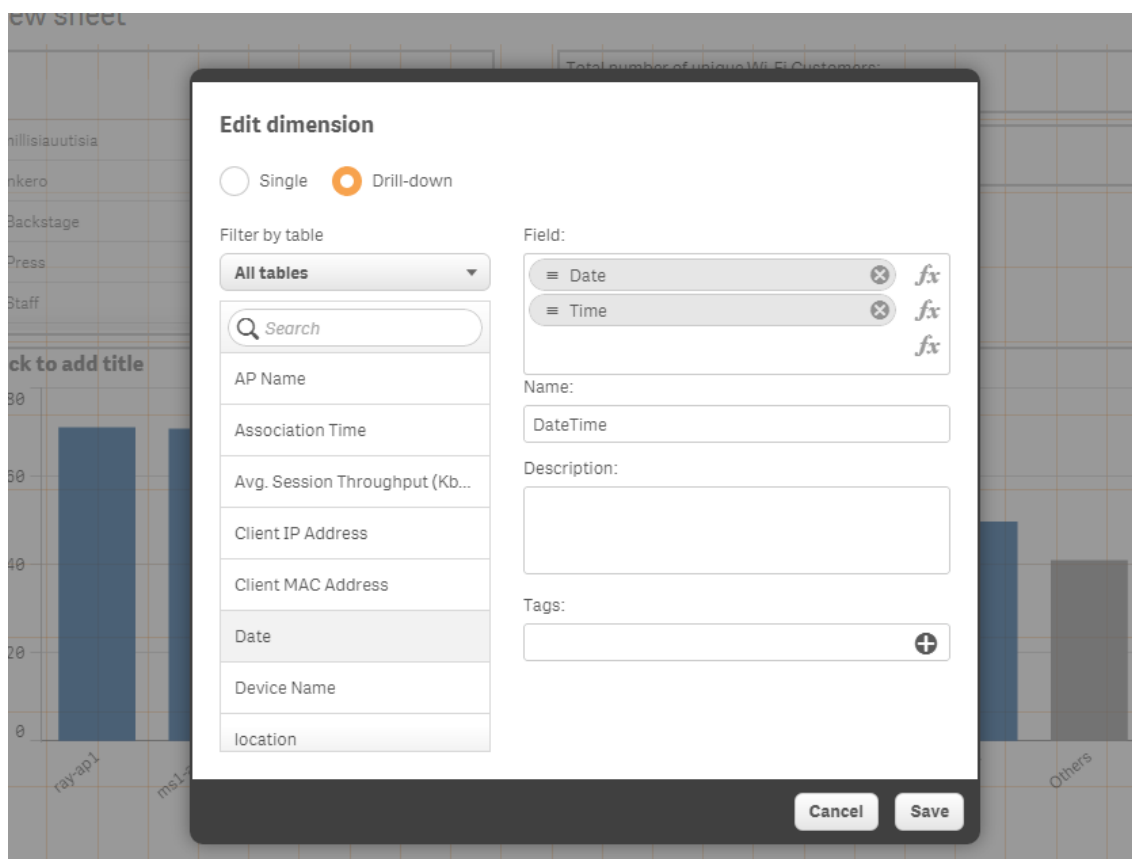
Edellä olevassa esimerkissä on esitetty sinisellä värillä tietueet, joihin tuodaan päivämäärä tai kellonaika halutussa muodossa. Qlik Sense -ohjelmistolle kerrotaan tuotavan datan tyyppi tarkoituksenmukaista funktiota käyttäen. Funktiot on esimerkissä kuvattu vihreällä värillä. Tietosisällön muokkaamiseen käytetään Trim-funktiota, jonka avulla voidaan poimia halutut osat alkuperäisestä merkkijonosta. Qlik Sense odottaa saavansa päivämäärän määrätyssä muodossa, joka saadaan järjestelemällä alkuperäisen päivämäärätiedon sisältö uudestaan. Qlik Sense odottaa ensin vuosilukua (kuva 11), joka poimitaan Trim-funktiota käyttäen. Ensin poimitaan merkkijonosta vasemmalta laskettuna 26–29, jolloin tässä tapauksessa saadaan tulokseksi ”2014”. Trim-funktiota käyttäen poimitaan seuraavaksi vasemmalta laskettuna merkit 4–8, ja tässä tapauksessa tuloksena saadaan ” Aug 09”. Trim-funktio on kuvattu esimerkissä oranssilla värillä. & -merkkiä käyttäen, joka esimerkissä on selkeyden vuoksi kuvattu punaisella värillä, yhdistetään edellä poimitut vuosiluku ja kuukausi sekä päivän järjestysnumero toisiinsa. Lopputuloksena saadaan Date -tietueen sisällöksi ”2014 Aug 09”. Samanlainen muotoilu tehdään myös tietueille Time ja Timestamp.

Kun skriptiä on muutettu, täytyy tiedot tuoda uudelleen raporttiin. Tämä tehdään painamalla ”Load data” -painiketta ”Data Load Editor” -ikkunassa. Tämän jälkeen on mahdollista luoda kaavioita, jotka muuttuvat valintojen mukaan joko niin, että käyttäjä näkee koonnin päivän tapahtumista, tai jos yksi päivä on valittuna, hän näkee tiedot kellonajan mukaan siltä päivältä. Jotta kaavio toimii näin, on ensin luotava uusi ulottuvuus -muuttuja, joka sisältää sekä Date- että Time- tietueet. Uusi ulottuvuus voidaan luoda valitsemalla vasemman laidan valikosta ketju-kuvake ja painamalla nappia Luo uusi (Create new), joka on esitetty kuvassa 12.



Kuva 12. Uuden ulottuvuuden luontiin tarvittava valikko.

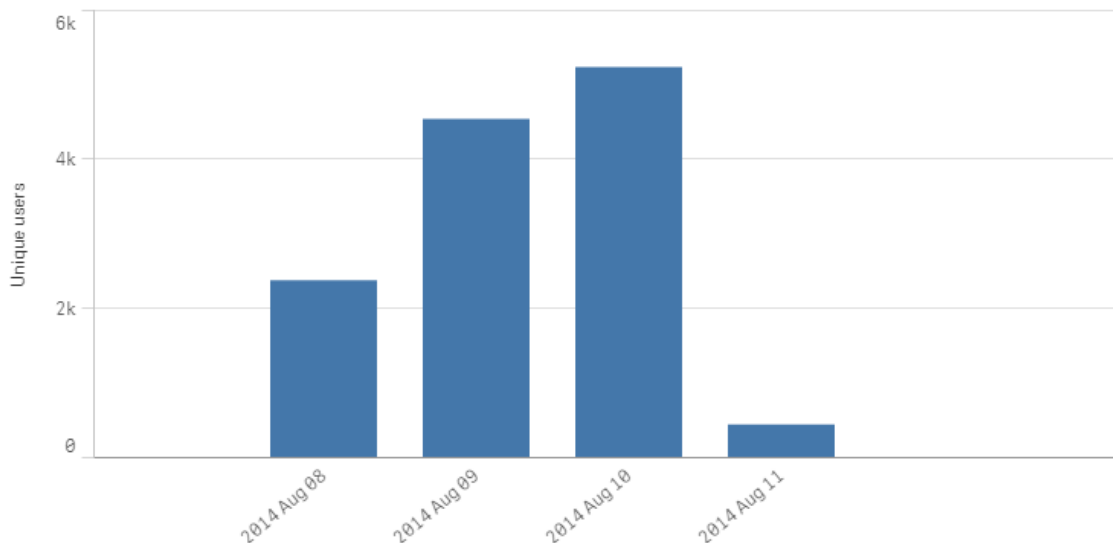
Uuden ikkunan auetta valitaan vaihtoehdoksi Poraus (Drill-down) ja lisätään kenttä- valikkoon halutut ulottuvuudet, tässä tapauksessa Date ja Time. Kuvassa 13 on tästä esi- merkki. Lopuksi uudelle ulottuvuudelle annetaan nimi, jota järjestelmä käyttää sisäisesti raportissa. Tässä raportissa ulottuvuuden nimeksi on määritelty "DateTime", jota ei käy- tetä muualla.



Kuva 13. Uuden ulottuvuuden asetukset.

Kun uusi ulottuvuus on käyttövalmis, voidaan sen avulla luoda uusi kaavio. Kuvassa 14 on esimerkissä kaavio, jossa ulottuvuutena on aika eli juuri luotu DateTime. Mittana on käytetty uniikkeja verkon käyttäjiä.

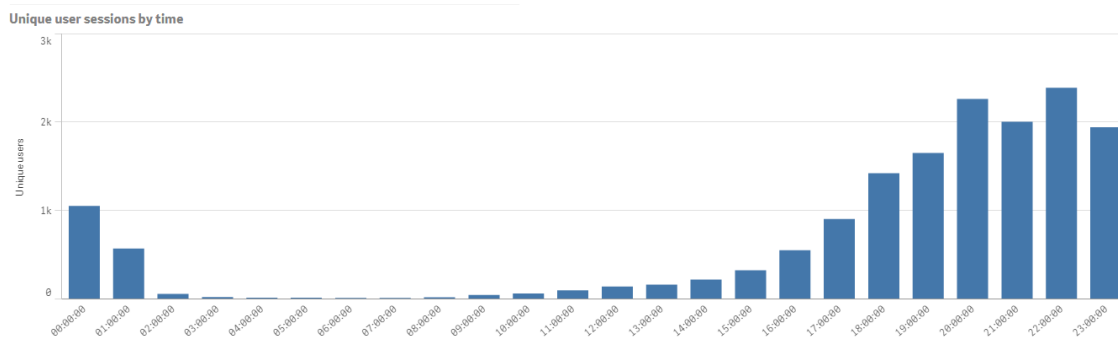
Unique user sessions by time



Kuva 14. Porautuva kaavio päivämäärän mukaan.

Kuvassa 14 nähdään summa kaikista uniikeista käyttäjistä eri päivinä. Klikkaamalla kaavion tolppaa ja sen jälkeen vihreää valintamerkkiä kaavion oikeassa yläkulmassa on mahdollista porautua kyseisen päivän tuntikohtaisiin tietoihin.

Kun tietty päivä on valittu, muuttuu alkuperäinen kuvaaja niin, että uniikit kävijät esitetään kellonajan mukaan kyseiseltä päivältä. Esimerkki on kuvassa 15.



Kuva 15. Porautuva kaavio kellonajan mukaan valittuna päivänä.

Siirretyn datan määrä

Raporttiin haluttiin kaavio, jossa on vaaka-akselilla aika ja pystyakselilla siirretyn tiedon määrä. On keksittävä, miten voidaan laskea siirretyn tiedon määrän edes kohtuullisella

tarkkuudella ajan funktiona. Käytettävissä ei ole mittaria siirretystä datamäärästä, ainoastaan keskiarvo sekuntia kohden, "Avg. Session Throughput (Kbps)". Jotta saataisiin jonkinlainen kuva, millaisia datamääriä verkossa liikkuu, työssä käytettiin olemassa olevaa tietoa yhteyden kestosta, "Session duration". Kun nämä kaksi lukua kerrotaan keskenään, saadaan tulokseksi siirretyn datan määrää kilobitteinä. Tämä luku ei ole tarkka, mutta tarkkuus on riittävä tässä yhteydessä.

Seuraava ongelma oli, että yhteyden kesto ilmoitettiin muodossa "20 min 1s". Tämä tieto tuli muuttaa niin, että se olisi pelkkä lukuarvo sekunteina. Muunnos tehdään datan latausvaiheessa. Tähän muuntamiseen käytin apuna Subfield-funktiota, joka ottaa kaksi argumenttia: erotinmerkin tai tekstin ja sarakkeen numeron.

```
Subfield([Session Duration], 'min', 1) as DurationMin,
```

Edellä esitetty lauseke luo uuden sarakkeen nimeltä DurationMin, jonne se tallentaa ennen "min"- erotinta olevan tiedon, tässä tapauksessa luvun 20. Vastaavasti voitaisiin ottaa myös sekunnit omaan sarakkeeseensa.

Kun molemmat luvut ovat olemassa, pitää minuutit kertoa kuudellakymmenellä, jotta niistä saadaan sekunteja, ja sen jälkeen laskea molemmat sarakkeet yhteen. Mutta koska datan latauksen aikana ei ole mahdollista käyttää juuri luotua uutta saraketta (esimerkissä DurationMin), on laskutoimitukset ja tietomuunnokset tehtävä yhdessä lauseessa. Esimerkiksi näin:

```
(60*(Subfield([Session Duration], 'min', 1)) +  
(Subfield(Subfield([SessionDuration], 'min', 2), 'sec', 1))) as  
DurationInSeconds,
```

Edellä kuvattujen toimenpiteiden avulla saadaan laskettua kunkin yhteyden kesto sekunteina. Tehtävän toinen osa on kertoa tämä siirretyn datan määrän keskiarvolla. Koska siirretty määrä saattaa olla pienempi kuin 0,1 kilobittia sekunnissa, se merkitään lähtötiedoissa '<0.1'. Tämä tuottaa ongelman, koska <-merkin takia arvo luetaan aina tekstiarvoksi ja sillä ei voi kertoa paljasta lukua. Tämän vuoksi merkkijonosta on poistettava erikoismerkki. Tähän käytetään funktiota nimeltä PurgeChar, joka poistaa merkkijonosta sille argumenttina annetun merkin tai merkit. Tämä aiheuttaa tilastovääristymää, mutta

datamäärän pienuuden vuoksi sillä ei ole käytännön merkitystä. Seuravaassa on esimerkki PurgeChar-funktion käytöstä.

```
Purgechar([Avg. Session Throughput (Kbps)], '<') as
SessionThroughput,
```

Lopuksi pitää vielä laskea kertolasku yhteyden kestosta ja siirretyn datan määrästä. Sen jälkeen luotua tietuetta voi käyttää laskennassa.

```
(60*(Subfield([Session Duration], 'min', 1))+(Subfield(Subfield(
[Session Duration], 'min', 2), 'sec', 1)))*(Purgechar([Avg. Session
Throughput (Kbps)], '<')) As DataAmount,
```

4.4 Tuotettu raportti

Liitteessä 4 on esitetty työn tuloksena tuotettu raportti, jonka sisältö esitellään seuraavissa kappaleissa.

Välilehti 1 – Event overview

Ensimmäisellä sivulla kerrotaan numeraalisesti koko tapahtuman aikana otetut yhteydet ja uniikit käyttäjät. Tämä tieto vaihtelee esimerkiksi sen mukaan, mitä päivämääriä tai verkkoja on valittu tarkastelun kohteeksi.

Sivun alalaidassa näytetään viivakaaviona, kuinka monta uniikkia käyttäjää on milläkin ajanhetkellä. Kuvaajaksi valittiin viivakaavio, koska siinä haluttiin esittää käyttäjämäärien kehitystä ajan suhteen. Lisäksi viivadiagrammin lukemisen helpottamiseksi väritettiin myös viivan alle osuva alue.

Välilehti 2 – Unique users & transferred data by time

Toisella sivulla on rinnakkain uniikkien käyttäjien määrä ja siirretyn tiedon määrä ajan funktiona. Molempiin kaavioihin valittiin selkeyden vuoksi pylväskaaviot. Yleensä tärkein tieto on kokonaismäärä käyttäjistä, ja aikajatkumon esittäminen on toissijaista.

Välilehti 3 – Unique service users by time and Vendors

Kolmannella sivulla esitetään ylhäällä uniikit käyttäjät eri aikoina yhteistyökumppanien verkoissa. Tässä kuvaajassa haluttiin jälleen korostaa aikajatkumoa, joten kaaviotyypiksi valittiin viivakaavio.

Sivun alapuolella esitetään eri päätelaitevalmistajien suhteelliset osuudet joko yhteysmääristä laskettuna tai uniikkien käyttäjien mukaan laskettuna. Koska kategorioita oli suhteellisen vähän ja tarkoitus oli esittää niiden suhteita toisiinsa, esitystavaksi valikoitui ympyräkaavio.

5 Testaus

5.1 Testauksen merkitys

Testaaminen on tuotteen toiminnan ja luotettavuuden kannalta oleellista. Tuotettavien raporttien tarkoitus on tuoda hyödyllistä tietoa tilaajalle, tapahtumajärjestäjille ja mahdollisille yhteistyökumppaneille. Jos tuotetta ei ole testattu, ei sen tuottamaan raporttiin voida luottaa ja raportin tuottama hyöty on kyseenalainen.

Testaamistapoja on olemassa useita. Tässä työssä testauksella viitataan toiminnalliseen testaukseen (functional testing), jossa ohjelmistolle annetaan tietty syöte ja tutkitaan, käsitteleekö ohjelmisto tiedot kuten pitää. Testaus on tarkemmin sanottuna laatikkotestausta. Laatikkotestauksella tarkoitetaan testausmenetelmää, jossa ohjelmaan syötetään ennalta valitut arvot ja saatuja tuloksia verrataan vertailutuloksiin. Koska testaavalla henkilöllä on pääsy skripteihin ja raporttipohjiin, luetaan testaus white-box -testaukseksi.

5.2 Testausmenetelmät

Raportteja luotaessa tulee tarkkailla lukujen summittaisia arvoja. Niiden tulisi olla vähintään oikeassa kokoluokassa. Jos lukujen summittaisetkaan arvot eivät vastaa todellisuutta, on melko varmasti jossain laskentakaavassa ongelma. Silloin on ehdottoman tärkeää, että jo raportin luontivaiheessa virheet korjataan. Monesti samoja muuttujia käytetään monessa eri taulukossa, joten on helpompi korjata virhe heti kun se huomataan

kuin projektin loppuvaiheessa ja sitten mahdollisesti korjata monta muutakin taulukkoa ja kuvaajaa. Usein virhe on jo datan latausvaiheessa, ja silloin voi olla, että koko ladattavan taulun nimi muuttuu korjausprosessin aikana. Myöhemmin on työlästä käydä käsin läpi jokainen kuvaaja ja muuttaa yksitellen käytettyjen taulujen nimet oikeiksi.

Vertaaminen aikaisempiin dokumentteihin

Paras tapa tutkia, saammeko oikeita tuloksia, on verrata uutta raporttia samasta lähtödatasta aiemmin laadittuihin raporteihin, mikäli sellaisia on olemassa. Tässä tapauksessa käytettävä lähtödata raportin luontiin valittiin sen mukaan, että vertailuraportit ovat olemassa. Vertailtaessa raporteja on huomioitava myös mahdolliset virheet käsin tuotetuissa alkuperäisissä raporteissa.

Raporttia testattaessa ja verrattaessa sitä alkuperäiseen käsin tehtyyn huomio kiinnittyi siihen, että aivan kaikki taulukot eivät anna samoja tuloksia. Esimerkiksi se, kuinka paljon tietoa on siirretty minäkin päivänä, poikkesi hiukan raporttien kesken. Tämä ei kuitenkaan ollut varsinainen virhe kummassakaan raportissa, vaan vuorokausi oli määritelty eri tavalla eri raporteissa. Aiemmin käsintehdyssä raportissa vuorokausi alkoi aamulla kello 4. Tämä siksi, että lähdetieto on peräisin musiikkifestivaaleilta, joissa tavallisesti ihmiset tulevat paikalle alkuillasta ja lähtevät aamuyöstä. Näin saatiin selkeämmin näkyviin yhden illan ja yön tietoliikennemäärät. Tässä työssä tuotetussa raportissa vuorokausi vaihtuu keskiyöllä, jolloin ei ole tarpeen tapahtumakohtaisesti arvioida vuorokauden vaihtumisajankohtaa.

Käsiteltävien tietojen tarkkuus

Koska raporteista halutaan saada vain suuntaa antavaa tietoa verkon käytöstä, voi virhetoleranssi olla suurikin. Kun lähtötiedot ovat epätarkkoja, myöskään raporteissa olevien tietojen tarkkuutta ei voida vaatia.

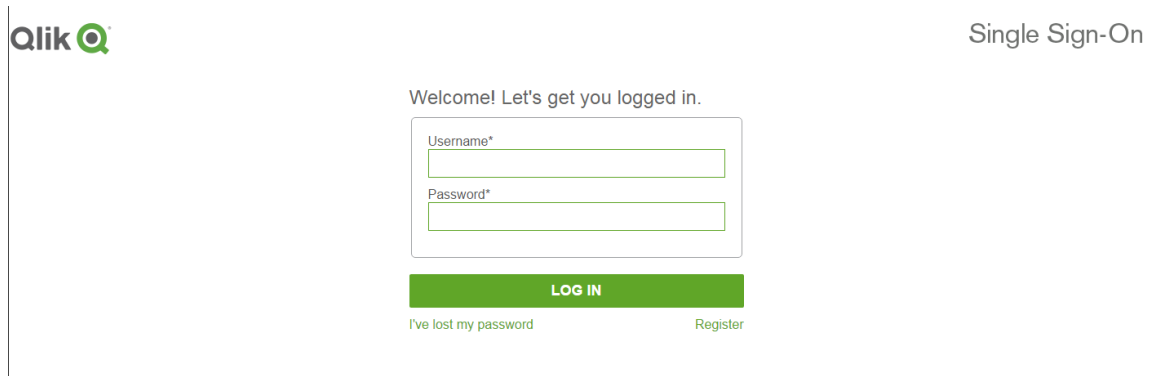
6 Raportin viimeistely ja toimittaminen

Kun raportti on muutoin valmis, se on hyvä käydä läpi kaavio kaaviolta ja katsoa, että kokonaisuus näyttää siistiltä ja ymmärrettävältä. Seuraavat asiat on hyvä tarkistaa:

- Onko kaikilla kaavioilla järkevät nimet (jos ne näytetään)?
- Näkyvätkö raportissa tarvittavat selitteet asianmukaisilla paikoilla?
- Onko tiedot kaavioissa järjestetty oikein, eli onko ne aakkosjärjestyksessä tai kelonajan mukaan?
- Saako kaavion koosta selvää?
- Näkyyhän kaavioissa vain asiaan kuuluvaa tietoa?
- Ovatko välilehtien nimet oikein?
- Ovatko kaikki kaaviot näkyvissä?

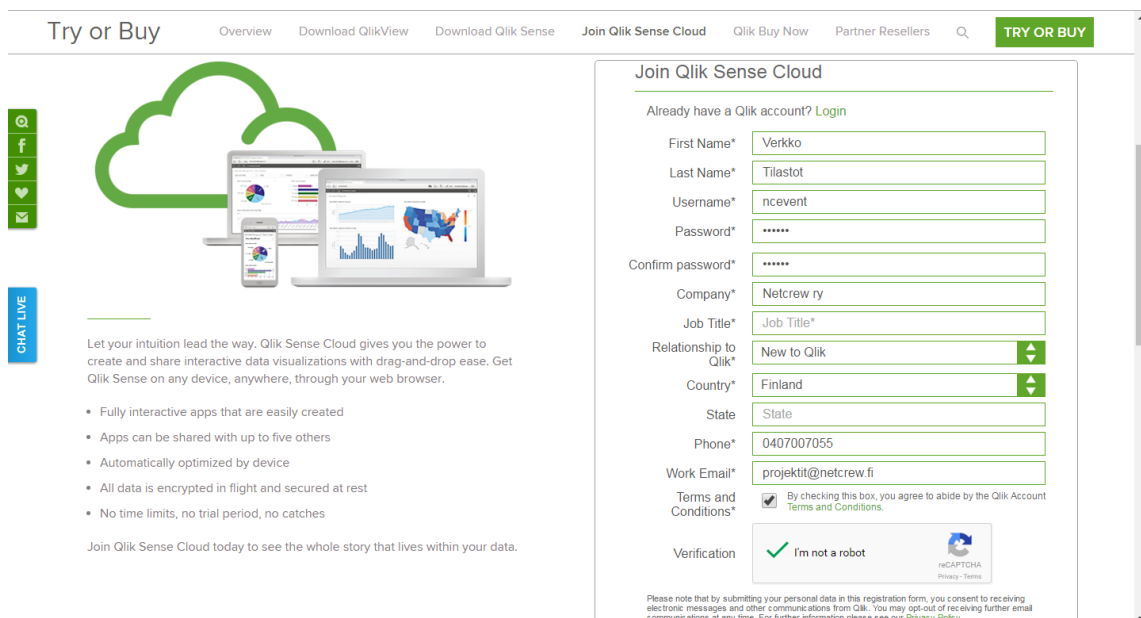
Tunnuksen luonti Qlik-pilvipalveluun tilaajalle

Qlik-pilveen kirjautuminen tapahtuu sivulta <https://qlikid.qlik.com/#/signin> . Mikäli käyttäjällä ei ole tunnusta entuudestaan, uuden tunnuksen voi luoda klikkaamalla ”Register”-linkkiä kirjautumistietojen alapuolelta. Kuvassa 16 on kuvakaappaus Qlik-palvelun kirjautumissivulta.



Kuva 16. Qlik-pilvipalvelun kirjautumissivu.

Rekisteröitymissivulla kysytään tietoja käyttäjästä ja kerrotaan lisätietoja palvelusta. Kuvasssa 17 esimerkki kirjautumissivun täytöstä.

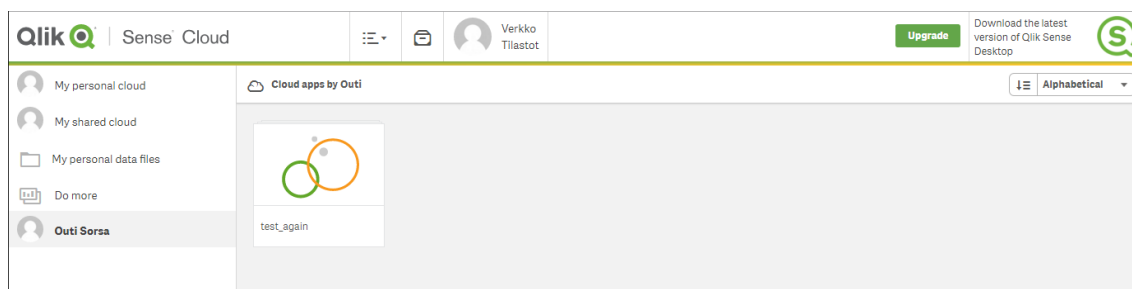


Kuva 17 Palveluun rekisteröityminen.

Kun käyttäjä on rekisteröitynyt palveluun, hänelle voidaan jakaa toisen käyttäjän tekemiä ja omaan pilveensä tallentamia raportteja. Tätä työtä tehdessä loin oman tunnukseni palveluun ja Netcrew'ille omansa. Tallensin testitiedoston omaan pilveen ja jaoin sen käyttäjälle nc_events. Tiedoston jakaja voi muuttaa tiedostoa vielä pilvessäkin, mutta käyttäjä, jolle tiedosto jaetaan, voi vain tarkastella tietoja.

Qlik-pilvipalvelussa tai Qlik Sensessä katseltu raportti on interaktiivinen, eli käyttäjä voi valita tarkasteltavaksi esimerkiksi vain haluamansa SSID:t, jolloin taulukoiden tiedot

muuttuvat käyttäjän valintojen mukaan. Kuvassa 18 on esitetty Qlik-palvelun etusivu, jossa näkyy käyttäjälle nc_events jaettu testitiedosto ”test_again”.



Kuva 18. Jaettu tiedosto Qlik Sense Cloud –palvelussa.

Kun tilaajalle on luotu omat tunnukset, voidaan käyttöön ladata myös Qlik Sense Desktop -ohjelma. Desktop-versiota tarvitaan, kun luodaan uusia raportteja tai päivitetään vanhoja. Tilaajalle toimitettiin ohjeet ohjelman latausta ja asennusta varten sekä käyttöohjeet, joiden avulla uusia raportteja voidaan luoda. Ohje on liitteessä 3.

Raportin laatiminen tapahtumaan

Laadittaessa raporttia uudesta tapahtumasta, pitää raportteja laativalla henkilöllä olla käytössään kone, jossa on asennettuna Qlik Sense -ohjelmisto ja käytettävissään tässä työssä tehty .qvf-tiedosto. Niiden lisäksi hänellä pitää olla valmiina Cisco Primesta ladattu CSV-tiedosto käyttäjien yhteyksistä kyseisen tapahtuman ajalta. Tiedosto tulee sijoittaa skriptissä määriteltyyn hakemistoon. Tiedoston nimi pitää myös muuttaa muotoon ”Client_sessions.csv”, koska alkuperäinen Qlik Sense -raportti on luotu etsimään tämän nimistä tiedostoa.

Uuden raportin luonti aloitetaan avaamalla qvf-tiedosto ja navigoimalla Data load manager -sivulle ja sen jälkeen painamalla ”Load data” -nappia. Kun lataus on valmis, tulos hyväksytään painamalla ”Ok”-nappia. Sen jälkeen uutta raporttia voi käydä tutkimassa painamalla navigaatio -nappia ja valitsemalla ”App overview”.

Raportin toimitus toimeksiantajalle

Toimeksiantajan tarpeisiin riittää tyypillisesti PDF -muotoinen raportti, jonka Qlik Sense -ohjelmisto pystyy tuottamaan. PDF -muotoisen raportin voi tehdä sekä ohjelmiston työpöytäversiolla että pilvipalvelussa. Huomioitavaa on, että tämän voi tehdä pilvipalvelussa vain raportin alkuperäinen jakaja. Siksi kaikki tiedostot tulisivat luoda tästä eteenpäin tilaajan omilla tunnuksilla. Hyvää tässä on se, että dokumentti voidaan jakaa muillekin käyttäjille tutkittavaksi ilman pelkoa, että joku muuttaa sitä.

PDF -raporttia laadittaessa tehdään aluksi halutut valinnat eli valitaan esimerkiksi päivät, jotka halutaan raporttiin. Tämän jälkeen valitaan ylälaidan menu-valikosta ”Export sheet to PDF”. Ennen varsinaista tiedoston luomista on mahdollista määritellä esimerkiksi sivun koko ja se, onko sivu vaaka- vai pystysuorassa. Käytännössä muita asetuksia ei ole. Ohjelma ilmoittaa, kun tiedosto on luotu.

Aina raporttia luotaessa kannattaa huomioida, että erilaiset valintaikkunat eivät tulostu PDF-versioon, vaan ne jäävät pois. Hyvän asian tästä tekee se, että jos lähtötiedoissa on kahden eri sponsorin asioita, voidaan tehdä kaksi erillistä dokumenttia, ilman että kumpikaan tiedostaa toisen sponsorin tietojen olemassaoloa.

Tapahtumajärjestäjästä riippuen voidaan raportit käydä myös läpi niin, että tilaaja esittelee Qlik Sense -raportit tapahtumajärjestäjälle ja/tai yhteistyökumppaneille.

7 Yhteenveto

Insinööriyön tavoitteena oli helpottaa Netcrew’n asiakkailleen toimittamien raporttien luomista. Aiemmin raportit luotiin käsin Exceliä apuna käyttäen.

Työssä tutustuttiin aluksi jo olemassa oleviin, manuaalisesti koostettuihin, raportteihin ja saatavilla oleviin lähtötietoihin. Sen jälkeen valittiin työhön ja tilaajan budjettiin soveltuva raportointiohjelmisto. Raporttejen luomista varten oli tutustuttava erilaisiin kaaviotyyppihin ja esitystapoihin, sekä valita niistä sopivimmat raporttipohjaa varten.

Raportointia varten tuli muodostaa skripti, joka käsitteli lähtötiedot haluttuun muotoon automaattisesti. Lopuksi valmis raportti piti testata ja tuottaa ohjeistus sen käytöstä insinööriyön tilaajalle.

Raporttien luomista varten pitää ymmärtää jollain tasolla teknisiä asioita kuten, SQL-tyyppistä skriptausta ja tietokantoja. Sen lisäksi pitää myös osata asettua asiakkaan asemaan ja kuunnella asiakasta tarkasti.

Halusin nostaa esille, että raporteja tarvitsevat pääosin tapahtumajärjestäjät. Heille niillä on hyvin paljon taloudellista merkitystä. Tilaajalla yhdistyksenä on paljon kokemusta erilaisista verkoista, ja saadut raportit ovat lähinnä hyvää lisätietoa.

Työssäni saavutin mielestäni tavoitteen hyvin. Pystyin luomaan raporttipohjan, jota voidaan käyttää tulevaisuudessakin niin kauan, kuin järjestelmästä saatavat lähtötiedot pysyvät saman muotoisena. Raporttien jakaminen hoituu myös hyvin kätevästi eri käyttäjien kesken. Yksi hyvä ominaisuus uusissa raporteissa on niiden visuaalisuus ja joustavuus. Pelkkä numeerinen tieto ei aina kerro tarpeeksi, ja ihmiset sisäistävät tiedon helpommin kaavioiden avulla.

Tilaajan edustajien kanssa käydyn esittelyn perusteella tuotetusta raporttipohjasta on selkeä hyöty ja raportointijärjestelmälle nähtiin myös muita käyttökohteita alkuperäisen toimeksiannon ulkopuolelta.

Lähteet

1. Tietoyhteiskuntaaari. 7.11.2014/917. Verkkodokumentti. <<https://www.finlex.fi/fi/laki/ajantasa/2014/20140917>>. Luettu 27.3.2016.
2. Connected mobile experience. 2013. Verkkodokumentti. Cisco Systems India. <<http://www.slideshare.net/CiscoIndia/connected-mobile-experience-28590281>> Luettu 14.3.2016
3. Cisco Prime Infrastructure 3.0 Data sheet. 2015. Verkkodokumentti. Cisco Systems. <<http://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/prime-infrastructure/datasheet-c78-735696.html>>. Luettu 27.3.2016.
4. Lane, David M. Introduction to statistics. Verkkodokumentti. <http://onlinestatbook.com/Online_Statistics_Education.pdf>. Luettu 10.4.2016
5. Agarwal, Amit. 2016. Choose the right chart type for your data. Verkkodokumentti. <<http://www.labnol.org/software/find-right-chart-type-for-your-data/6523/>>. Luettu 28.3.2016.
6. Cronström, Henric. 2013. Dimensions and Measures. Verkkodokumentti. <<https://community.qlik.com/blogs/qlikviewdesignblog/2013/03/25/dimensions-and-measures>>. Luettu 3.3.2016.

Aiemmin käytetty raportti

Example Event

Wi-Fi Client Statistics

Number of public Wi-Fi services	4
Number of private Wi-Fi services	3
Total number of unique Wi-Fi Clients	6982
Number of User Sessions	100362
Average Traffic per Session (MB)	3,5
Total Traffic (GB)	329,8

Daily distribution

	Fri	Sat	Sun
Number of Unique Clients	2500	4468	4516
Number of User Sessions	20864	39244	40254
Average Session Time per Client (Minutes)	111	112	109
Average Traffic per Session (MB)	4,4	2,9	3,1
Total Traffic (GB)	92,1	113,8	123,9

Per service Client statistics

	Unique clients	Maximum concurrent clients
SSID1	4349	1455 Sun 19:00-20:00
SSID2	3599	919 Sat 20:00-21:00
SSID3	2358	594 Sat 20:00-21:00
SSID4	1566	312 Sun 20:00-21:00
Event-Press	482	175 Sun 18:00-19:00
Event-Backstage	402	133 Sun 20:00-21:00
Event-Staff	290	102 Sun 20:00-21:00

Device Vendor summary

	Devices	Percentage
Apple	3259	46,68 %
Nokia	806	11,54 %
Samsung	800	11,46 %
Murata	270	3,87 %
Htc	136	1,95 %
LG	105	1,50 %
Sony	77	1,10 %
Other/Unknown	1529	21,90 %


Raporttien käyttöohje

QLIK SENSE RAPORTIN TEKO


QLIKVIEW:IN LATAUS JA ASENNUS

1. Kirjaudu <https://qlikid.qlik.com/#/signin>
2. Klikkaa "Try or Buy" nappia
3. Klikkaa "Download Qlik Sense"
4. Jos sivu ei tunnista kirjautumista heti, paina oikean laidan lomakeessa nappia login, jolloin yleensä sivu tunnistaa kirjautumistiedot
5. Hyväksy käyttöehdot ja klikkaa "Download Now"
6. Asenna paketti oletusasetuksilla

RAPORTIN PÄIVITYS

1. Tallenna csv-tiedosto Cisco Primesta Downloads-hakemistoon
2. Muuta tiedoston nimeksi Client_sessions.csv
3. Käynnistä Qlik Sense Desktop
4. Avaa erikseen toimitettu .qvf-tiedosto Qlik Sensellä
5. Avaa projekti nimeltä wifi_statistics
6. Avaa navigaatiovalikko  → Data Load Manager
7. Klikkaa oikeassa ylä laidassa nappia "Load Data". Latauksen valmistuttua klikkaa "Ok"
8. Avaa navigaatiovalikko → App Overview

RAPORTIN VIEMINEN PDF-TIEDOSTOKSI

1. Valitse käytettävä välilehti
2. Tee mahdolliset valinnat valmiiksi (Käytettävä SSID jne.)
3. Avaa menuvalikko  → Export sheet to PDF
4. Tarkista asetusten sopivuus ja klikkaa "Export"
5. Klikkaa "Click here to download your PDF file" tarkastellaksesi tiedostoa tai tallentaaksesi sen

Number of user sessions: 100377
Total number of unique Wi-Fi devices: 6969

Unique device refers to a single device such as mobile phone, tablet or computer using the wireless network.
User session is single connection from device to network.

