



The Gathering – Galskap satt i system

Kristian Lyngstøl, Fornebu, Desember 2018

The hvaffornoe?



- Digital festival, eller dataparty, hver påske
 - Spilling, sosialisering, konkurranser, sosialisering, konserter, sosialisering, mekking, sosialisering, osv.
- Arrangert i Vikingskipet i cirka 20 år-ish, fra Onsdag til Søndag
- Cirka 5000 deltakere, snittalder er rundt 19 år
 - Typisk 6000 aktive dhcp-leaser
 - Cirka 10500 unike mac-adresser i løpet av påsken i DHCP-loggen
- 100% non-profit
- Datasfærens svar på Norway Cup

TG i bilder



TG i bilder



Tidligere...



300-400 frivillige



Tech crew



- Ansvarlig for planlegging, anskaffelse, oppsett og drift av nettverket
- 100% frivillig
- Stort spillerom for å teste ting
 - Trenger bare å “komme oss igjennom uka”
- Hardt arbeid I forkant og under opprigget
- Rolig under avvikling – om vi har gjort en god jobb
 - Dette er tiden noen gjerne finner på et “artig prosjekt”, som å skrive om hele NMS’et i fart.
 - Som strengt tatt er bedre enn det året man byttet fra OSPF til IS-IS i fart.

Tech crew



- Tech:Support ~18 folk
 - Kabler og monterer alle kantsvitsjer og APer på stativer
 - 1. linje mot deltakere
 - Mellom 15km og 20km TP til kantsvitsjer og APer
 - “Stasjonert” på gulvet, i informasjonsskranken
- Tech:Net ~14 “folk”
 - Gjør resten
 - 2.-3. linje, design, konfig, fiber, bygge infrastruktur, dhcp, nms, serverdrift, ip-telefoni, mm.
 - De første som møter mannsterke i Vikingskipet
 - “Stasjonert” i “NOCen” - en VIP losje øverst på tribunen



Kjapp gjennomgang av nettet, fra kant til core

Nettet



- Todelt: “Gulvet” (deltakere) er standard stjerne-nett, og “resten” er ring.
- Det ble gjort et større innkjøp av Juniper-utstyr i 2015, og dekker mye, men ikke alt av det som er i bruk.
- Resten lånes/sponses av Juniper, via nLogic
- Ganske likt design fra 2015, men en del viktige justeringer.
- Under TG18 hadde vi ~15 minutter nedetid rett etter dørene åpnet.

Kantsvitsjer



- Kantsvitsjer autoprovijonerer med ZTP og option 82.
 - Switch spør om DHCP, første “hop” (dhcp-relay) legger på info om hvilken port/vlan/distro det er fra, dhcp-server svarer med URL til relevant config.
- Deltakere kobler seg direkte inn i kant.
- Og kobler regelmessig ut power/uplink ved uhell.
- Strømbrudd og uplinks som går ned er svært utbredt det første døgnet.

~180 Juniper EX200 i kant



Eier vel strengt tatt
noen færre nå, i
hvert fall 2
eksploderte under
TG18 da noen
(ikke oss!) hadde
klart å koble opp
400V der det
skulle vært 240V

Distrobusjonssvitsjer

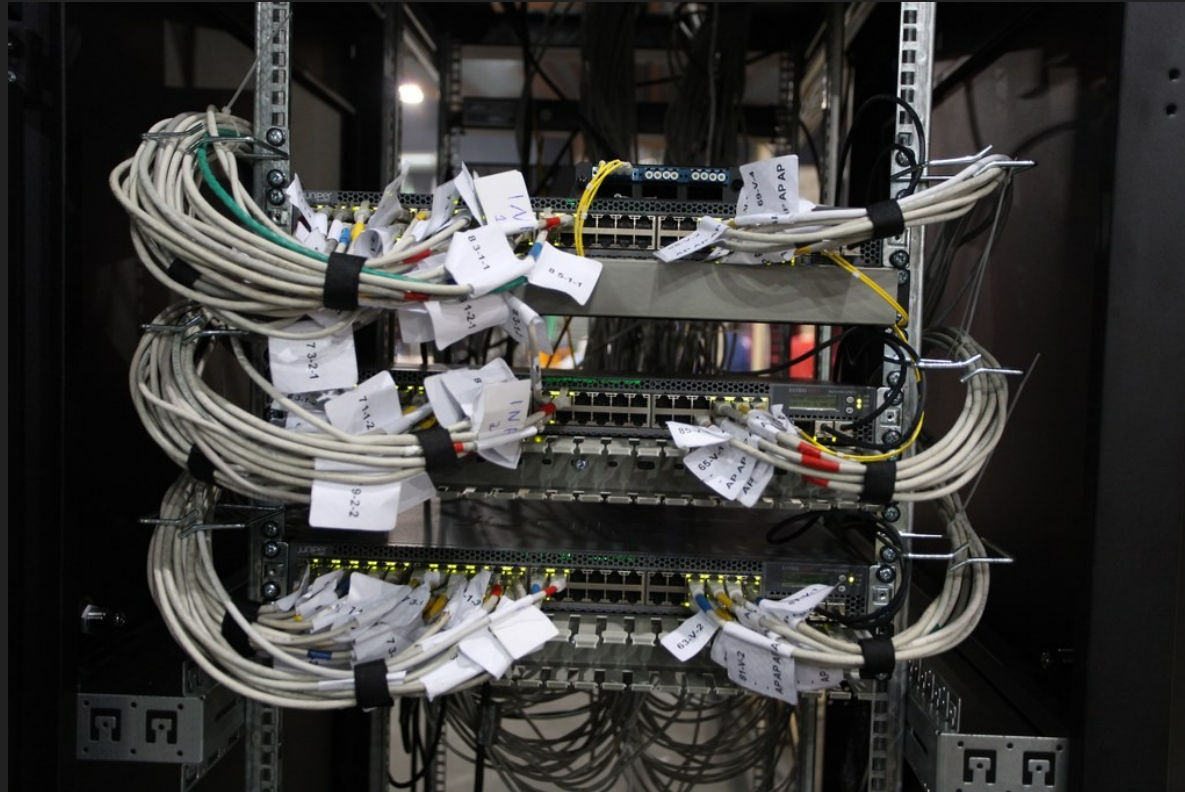


- En “distro” betjener cirka 300 til 700 deltakere.
- Alle kantsvitsjer er koblet i trippel til distro, en 1Gbit link til hvert member, i virtual chassis (VC).
- Hver distro har to 10Gbit uplinks.
- Bokstavelig talt. De går opp i taket.



3x EX3300 i VC som distribusjonssvitsj

9 slike rack
plassert langs
midtgangen

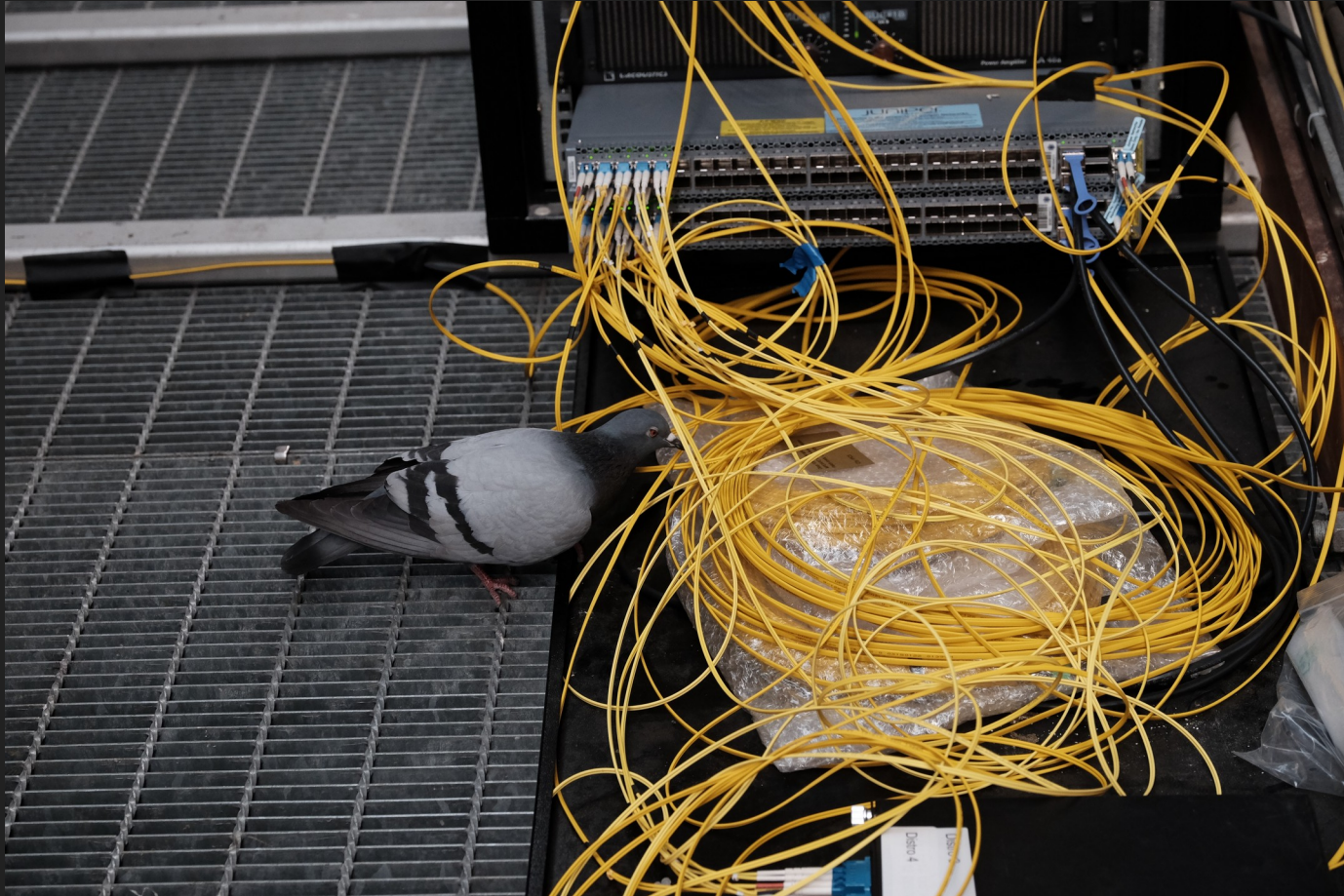




TG18: Collapsed core

- TG18 hadde “collapsed core” etter dårlig erfaring med EX3300 som ruter i TG15, 16 og 17.
- All lag 3 ble terminert på core, ikke distro.
- Core ble “oppgradert” fra to QFX5100 til en MX480 (og cold spare)

Utfordringer i høyden... 2xQFX som core i taket, TG17

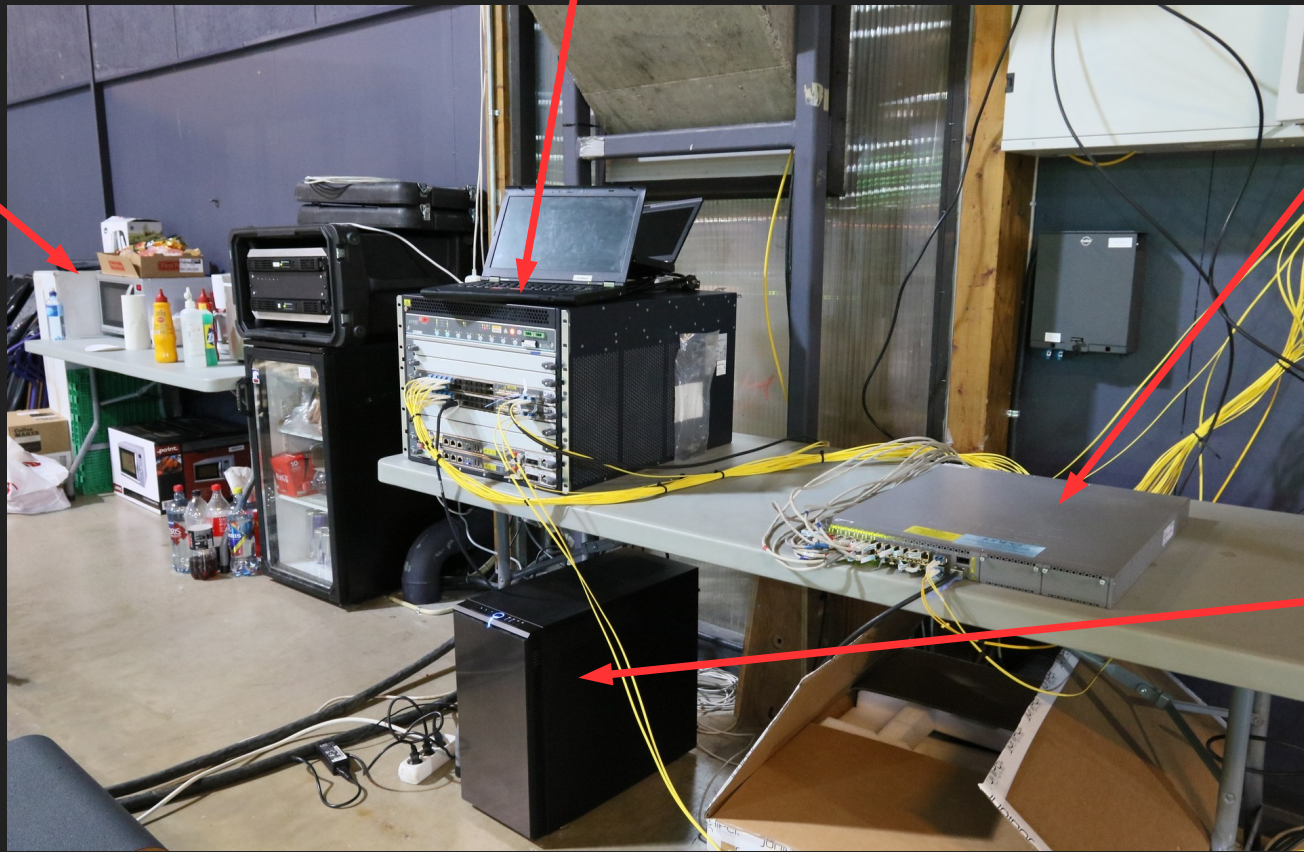




TG18

MX480 core, rett utenfor NOCen

Pølser, brus og
microbølgeovn
er kritisk
infrastruktur.



Ring-node

Og PC'en til Sjur rett
på core?!?

“Ringen”



- “Ringen” er en fiberring, betjent av EX4300/EX4600 i Virtual Chassis
- 6 noder plassert rundt hele Vikingskipet for å betjene “ymse”. Typisk crew-funksjoner, NOC, auditoriet, scenen, sponsorstands, inngangsparti, med mer.
- 40Gbit linker i redundant ring.

Ringnode i "sving": betjente scenen og streamingstudio under TG18



Fiberringen



Ring-noden

Internett!

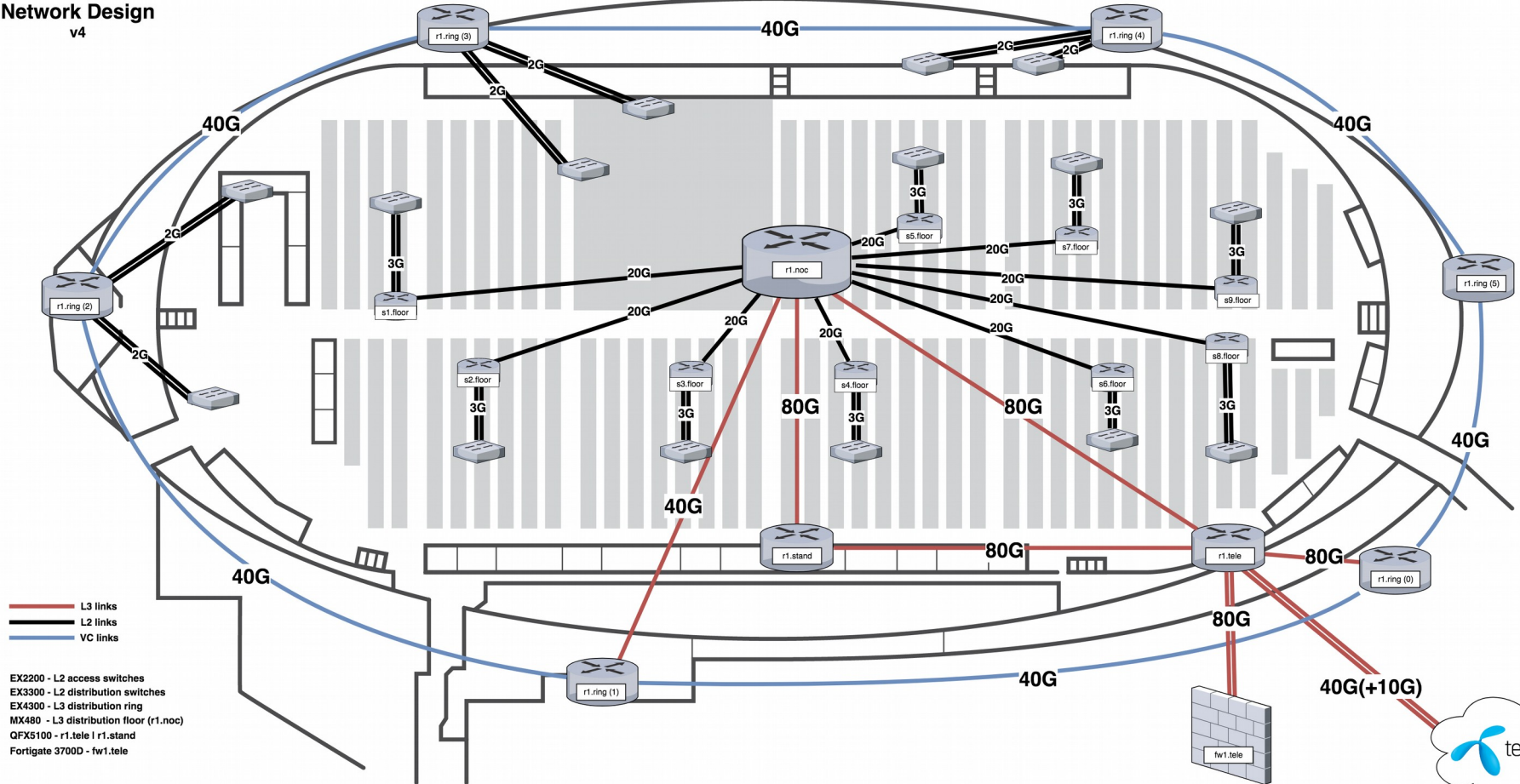


- Vært mye forskjellig, inkludert 200Gbit/s via Altibox
- De siste årene: Telenor!
- 50Gbit/s, delt opp i 40Gbit/s + 10Gbit/s for “sensitive ting”
 - Konkurranser, streaming fra scenen, mm
- Snitt-bruk mot internett er typisk 10-20Gbit/s

Designkart



The Gathering 2018 Network Design v4



Hvordan?



- Planlegging for TG19 er godt i gang
- Design legges typisk i oktober/november, arbeid med leverandører begynner deretter
- Mye forarbeid – logistikk og legge plan A, B og C
- Helga før palmesøndag, halvannen uke før åpning, begynner vi rigging av deler av kjernenettet. Typisk internett-akksess og ringen.

Hvor NMS kommer inn...



- Mest arbeid foregår FØR dørene åpnes.
- Fra fredag kveld til onsdag morgen er det “around the clock” jobbing, inkludert mye småjustering på Gondul
- Trenger oversikt over fremdrift
- Generer konfig
- “Test-driven network deployment”: Knot med dingsen til Gondul sier den er frisk.



Gondul i et nøtteskall

- NMS
- SNMP-polling, ping, tail av dhcp-log, operative logger
- Målet er at alt skal ligge i database og vi genererer konfig derfra
 - 100% sant for alt av kantsvitsjer, distro og core
 - Foreløpig ikke sant for ring, gateway på telematikk, firewall
 - Sant for DHCP-config, DNS-config, med mer.
- Åpent API der deler eksponeres direkte mot publikum
- Friprog, 90% skrevet av undertegnede



Drilldown

Health:

- Bestemmer farge

Modulbasert:

- Knøttsmå JS-moduler parser info og leverer et array med info og en "score/forklaring" - høyeste score = health.

Health (lower is better)	500 (LAG member (ge/xe/et) speed is 3000 but logical (ae) is 3000)
Management IP (v4)	88.92.88.10
Management IP (v6)	2a06:5844:88a::10
Subnet (v4)	88.92.4.128/26
Subnet (v6)	2a06:5844:4c::/64
Distro	s1.floor
Client VLAN	1113
Management VLAN	666
Uplinks (live/configured)	3 / 3
Client ports (live/total)	6 / 44
Non-uplink ports (live/total)	23 / 64
Temperature	20°C
IPv4 latency(ms)	19.98
IPv6 latency(ms)	21.647
DHCP age	1 minutes
DHCP clients	7
SNMP data	Available
System uptime	18 hours 26 minutes
CPU utilization	19%
Memory utilization	63%
SNMP uplink data	LAG member (ge/xe/et) speed is 3000 but logical (ae) is 3000



Logging – av mennesker

- “Oplog” er generisk. Ekstremt viktig med lav terskel for logging.
- Viser siste X entries I kart (bildet til venstre)
- Logentries kan kobles mot system(er). Vil da vises ved drill-down. (bildet til høyre)
- (Bilder fra den mest travle timen vår)

09:59	53-1 [Kirø] stræmbrudd, support melder til HOA	Distro-port	Distro port is live
09:55	[sjurtf] s5.sponsor levert på MX og s1.floor. Confer kant manuelt	DHCP age	No DHCP data
09:54	53-1 [Kirø] support on it	SNMP data	No data
09:52	71-3 [Kirø] OK	Relevant log entries	
09:48	e59-3 [Kristian] Mistenker strømbrudd. Avventer etter at support har sjekka.	2018-03-28 09:48	e59-3[Kristian] Mistenker strømbrudd. Avventer etter at support har sjekka.
09:48	67-1 [Kirø] stræmbrudd, HOA på saken	2018-03-28 09:43	59-3[Kirø] nede, rapportert til support
09:47	67-1 [Kirø] nede, support på saken	2018-03-26 14:31	floor[sjurtf] All L2 aksess på floor er inne nå. Tar meg en FAP test i RACK
09:45	71-3 [Kirø] support er på saken	2018-03-26 05:10	floor[sjurtf] Mangler portgrupper og access config for border. Simen jobber med dette.
09:44	53-1 [Kirø] Lasse+ er på stedet	2018-03-22 03:09	"e59-3"[system] [internal]Network added: e59-3
09:43	59-3 [Kirø] nede, rapportert til support	2018-03-22 02:06	"e59-3"[system] [internal]Switch added: e59-3

Søk



- Søk er “live”: Du ser hva som matches umiddelbart mens du skriver – 100% frontend
- Samme logikk brukes for å knytte log-entries til utstyr
- Regex på switch-navn, distro-navn, snmp description, antall aktive porter, diverse IP-adresser, mm
- Brukes til alt fra “dette er dit du skal gå” til å enkelt sjekke hva som henger på hvilken distro til søking på serienummer eller versjonsnummer

s1.noc



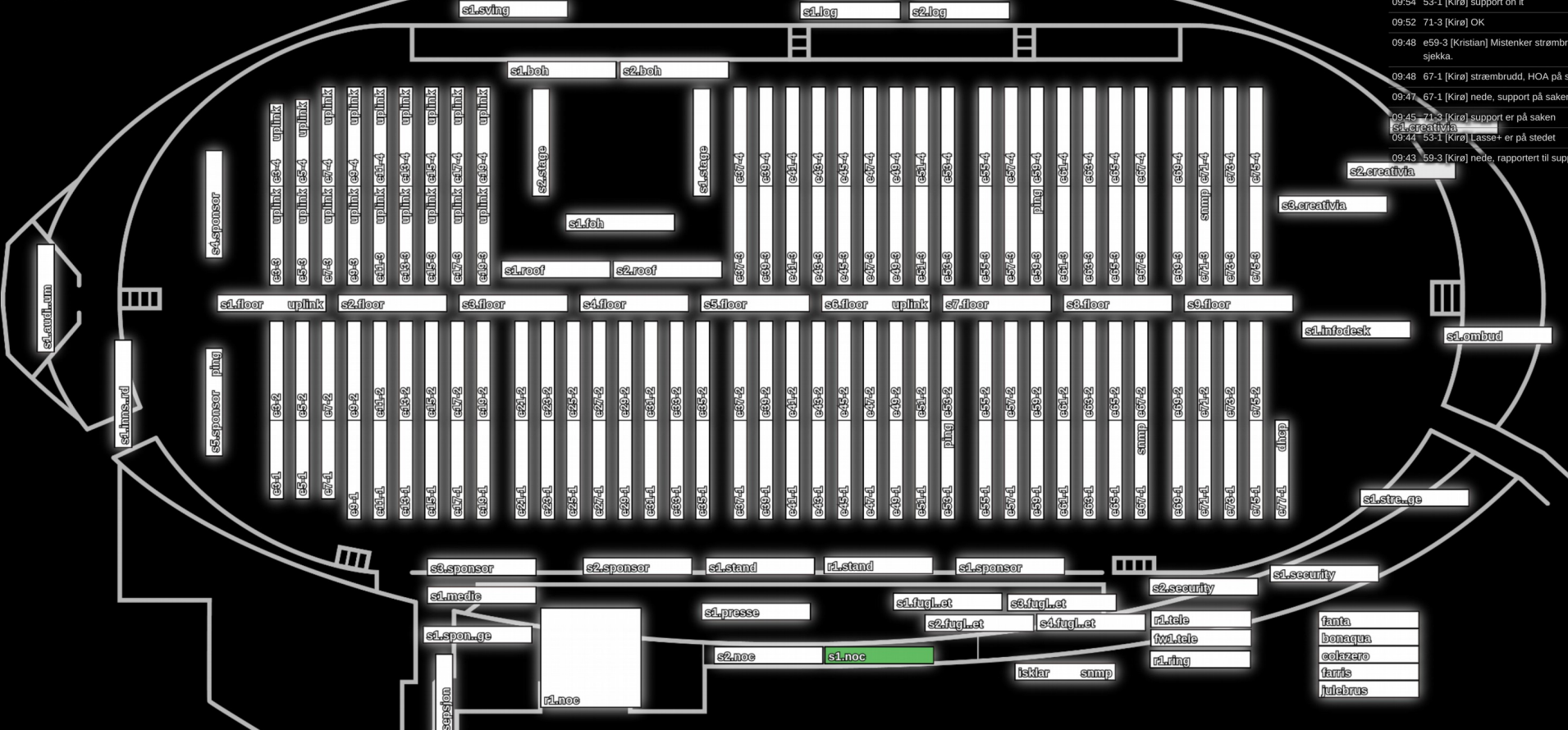
M A T O M

s1.noc X ? Go! 696 clients

System(s) Log entry

Health

Wed Mar 28 2018 10:00:00



- 09:59 53-1 [Kira] stræmbrudd, support n
- 09:55 [sjurt] s5.sponsor levert på MX og
- 09:54 53-1 [Kira] support on it
- 09:52 71-3 [Kira] OK
- 09:48 e59-3 [Kristian] Mistenker stræmbr
- 09:48 67-1 [Kira] stræmbrudd, HOA på s
- 09:47 67-1 [Kira] nede, support på saker
- 09:45 71-3 [Kira] support er på saken
- 09:44 53-1 [Kira] Lasse+ er på stedet
- 09:43 59-3 [Kira] nede, rapportert til sup

JUNOS 1

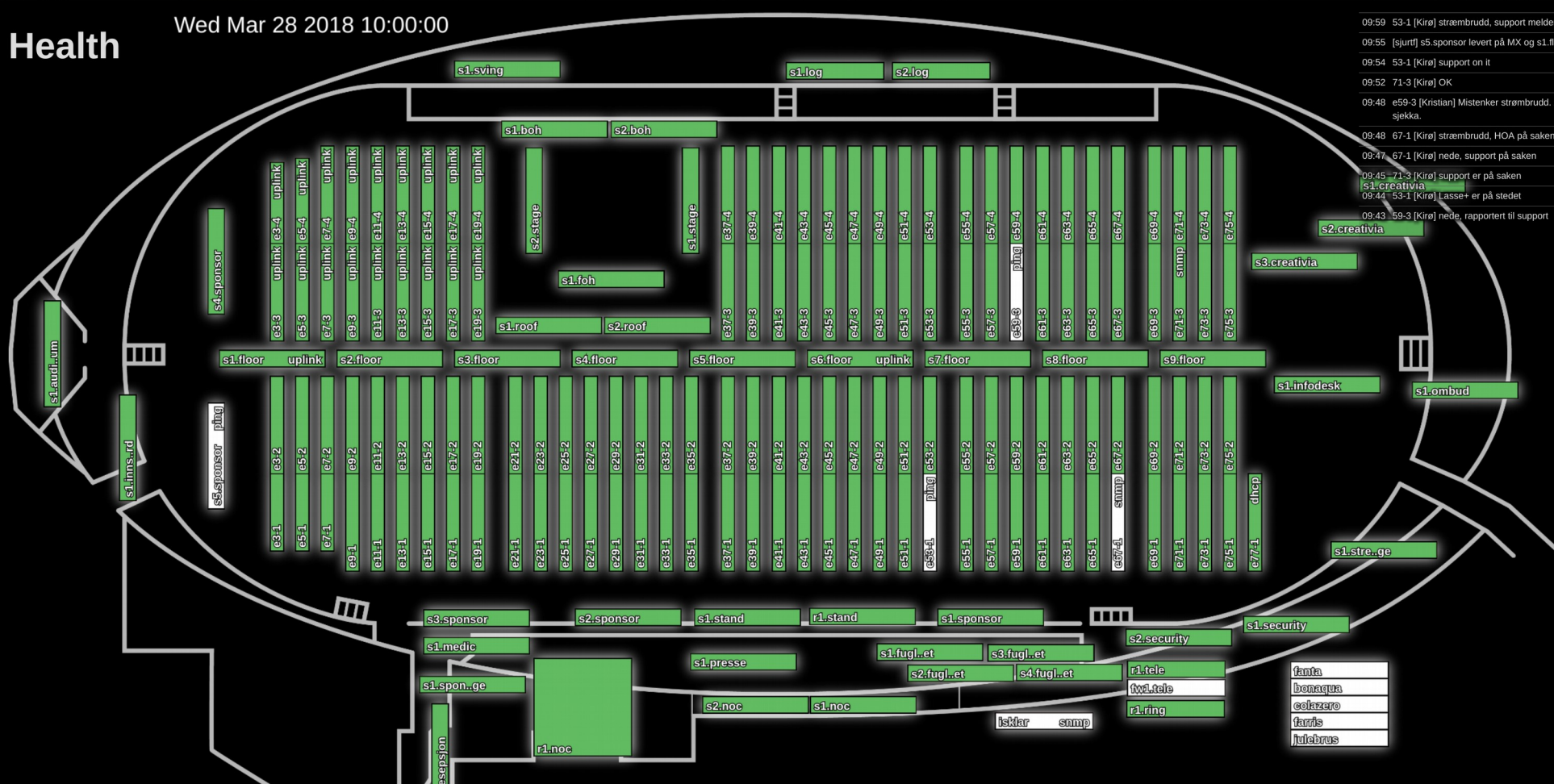


M A T O M JUNOS 1 X 7 7 Gor 696 clients

System(s) Log entry

Health

Wed Mar 28 2018 10:00:00



- 09:59 53-1 [Kiro] stræmbrudd, support melder
- 09:55 [sjurt] s5.sponsor levert på MX og s1.fl
- 09:54 53-1 [Kiro] support on it
- 09:52 71-3 [Kiro] OK
- 09:48 e59-3 [Kristian] Mistenker strømbrudd. sjekka.
- 09:48 67-1 [Kiro] stræmbrudd, HOA på saken
- 09:47 67-1 [Kiro] nede, support på saken
- 09:45 71-3 [Kiro] support er på saken
- 09:44 53-1 [Kiro] Lasse+ er på stedet
- 09:43 59-3 [Kiro] nede, rapportert til support

- fanta
- bonaqua
- colazero
- fantis
- julebrus



Provisjonering

- Steg 1: IP-Plan – så langt gjort for hånd, minus “gulvet”
- Gulvet: IP-Plan og plassering av distro kalkuleres av “planning.cpp”
 - Kalkulerer plassering og kabling
 - Forteller at til svitsj e35-2 trenger du en 50-meter og en 20-meter TP, og at det gir deg X meter slack.
 - Prøver å finne optimal kombinasjon av kabling og plassering av distrosvitsjer for å spare kabler, skjøter, mm.



Planning output

(Kabelmengde/skjøter ganges med antall uplinks: 3)

```
10m: 133
30m: 79
50m: 11
Extensions: 91
Horizontal gap crossings: 0

Total cable: 4250.0m (cost = 3532.8m)
Total slack: 717.2m (16.88%)
Remaining ports on distro 1: 13
Remaining ports on distro 2: 13
Remaining ports on distro 3: 15
Remaining ports on distro 4: 13
Remaining ports on distro 5: 13
Remaining ports on distro 6: 19
Remaining ports on distro 7: 19
Remaining ports on distro 8: 21
Remaining ports on distro 9: 21

kly@poe:~/src/tgmanage/planning$
```

```
kly@poe:~/src/tgmanage/planning$ ./planning -4 6 14 -23 24 -30 32 35 -37
Finding optimal layout for 132 switches
Augmented using 132 paths.
1 ( 1- 2)
2 ( 3- 4) 2 (50 ) (41.0) 2 (30 ) (26.6) 1 (10+10) (19.4) 1 (30+10) (33.8)
3 ( 5- 6) 2 (30+10) (37.4) 2 (30 ) (23.0) 1 (10+10) (15.8) 1 (30+10) (30.2)
4 ( 7- 8) 2 (30+10) (33.8) 2 (10+10) (19.4) *1 (10+10) (12.2) 1 (30 ) (26.6)
5 ( 9-10) 2 (30+10) (30.2) 2 (10+10) (15.8) 1 (10+10) (15.8) 1 (30+10) (30.2)
6 (11-12) 2 (30 ) (26.6) 2 (10+10) (12.2)* 1 (10+10) (19.4) 1 (30+10) (33.8)
7 (13-14) 2 (30+10) (30.2) 2 (10+10) (15.8) 1 (30 ) (23.0) 1 (30+10) (37.4)
8 (15-16) 2 (30+10) (33.8) 2 (10+10) (19.4) 1 (30 ) (26.6) 1 (50 ) (41.0)
9 (17-18) 2 (30+10) (37.4) 2 (30 ) (23.0) 1 (30+10) (30.2) 1 (50 ) (44.6)
10 (19-20) 2 (50 ) (41.0) 2 (30 ) (26.6) 1 (30+10) (33.8) 1 (50 ) (48.2)
11 (21-22) 3 (30+10) (37.4) 3 (30 ) (23.0)
12 (23-24) 3 (30+10) (33.8) 3 (10+10) (19.4)
13 (25-26) 3 (30+10) (30.2) 3 (10+10) (15.8)
14 (27-28) 3 (30 ) (26.6) 3 (10+10) (12.2)*
15 (29-30) 3 (30+10) (30.2) 3 (10+10) (15.8)
16 (31-32) 3 (30+10) (33.8) 3 (10+10) (19.4)
17 (33-34) 3 (30+10) (37.4) 3 (30 ) (23.0)
18 (35-36) 3 (50 ) (41.0) 3 (30 ) (26.6)
19 (37-38) 5 (50 ) (44.6) 5 (30+10) (30.2) 4 (30 ) (26.6) 4 (50 ) (41.0)
20 (39-40) 5 (50 ) (41.0) 5 (30 ) (26.6) 4 (30 ) (23.0) 4 (30+10) (37.4)
21 (41-42) 5 (30+10) (37.4) 5 (30 ) (23.0) 4 (10+10) (19.4) 4 (30+10) (33.8)
22 (43-44) 5 (30+10) (33.8) 5 (10+10) (19.4) 4 (10+10) (15.8) 4 (30+10) (30.2)
23 (45-46) 5 (30+10) (30.2) 5 (10+10) (15.8) *4 (10+10) (12.2) 4 (30 ) (26.6)
24 (47-48) 5 (30 ) (26.6) 5 (10+10) (12.2)* 4 (10+10) (15.8) 4 (30+10) (30.2)
25 (49-50) 5 (30+10) (30.2) 5 (10+10) (15.8) 4 (10+10) (19.4) 4 (30+10) (33.8)
26 (51-52) 5 (30+10) (33.8) 5 (10+10) (19.4) 4 (30 ) (23.0) 4 (30+10) (37.4)
27 (53-54) 5 (30+10) (37.4) 5 (30 ) (23.0) 4 (30 ) (26.6) 4 (50 ) (41.0)
28 (55-56) 7 (50 ) (41.0) 7 (30 ) (26.6) 6 (10+10) (19.4) 6 (30+10) (33.8)
29 (57-58) 7 (30+10) (37.4) 7 (30 ) (23.0) 6 (10+10) (15.8) 6 (30+10) (30.2)
30 (59-60) 7 (30+10) (33.8) 7 (10+10) (19.4) *6 (10+10) (12.2) 6 (30 ) (26.6)
31 (61-62) 7 (30+10) (30.2) 7 (10+10) (15.8) 6 (10+10) (15.8) 6 (30+10) (30.2)
32 (63-64) 7 (30 ) (26.6) 7 (10+10) (12.2)* 6 (10+10) (19.4) 6 (30+10) (33.8)
33 (65-66) 7 (30+10) (30.2) 7 (10+10) (15.8) 6 (30 ) (23.0) 6 (30+10) (37.4)
34 (67-68) 8 (30+10) (30.2) 8 (10+10) (15.8) 9 (30 ) (23.0) 9 (30+10) (37.4)
35 (69-70) 8 (30 ) (26.6) 8 (10+10) (12.2)* 9 (10+10) (19.4) 9 (30+10) (33.8)
36 (71-72) 8 (30+10) (30.2) 8 (10+10) (15.8) 9 (10+10) (15.8) 9 (30+10) (30.2)
37 (73-74) 8 (30+10) (33.8) 8 (10+10) (19.4) *9 (10+10) (12.2) 9 (30 ) (26.6)
38 (75-76) 8 (30+10) (37.4) 8 (30 ) (23.0) 9 (10+10) (15.8) 9 (30+10) (30.2)
```



Provisjonering del 2

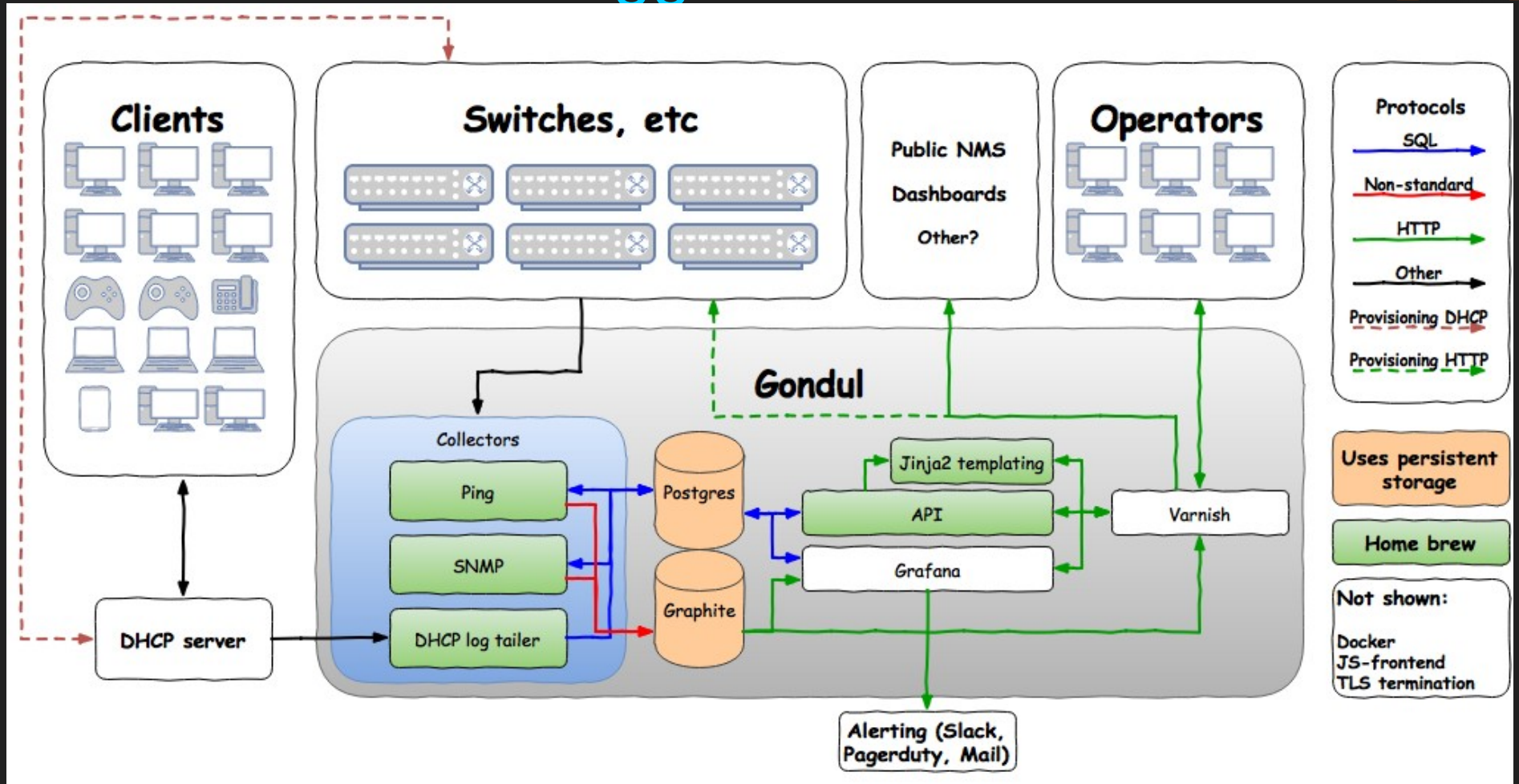
- Gondul mates med output fra planning og semi-manuelt av resten.
- Gondul sitt API er veldig script-vennlig. Finnes svitsjen oppdateres den, finnes den ikke så opprettes den med default på parametere som ikke er oppgitt...
- Gjøres selvsagt også i front-end



Provisjonering del 3

- Config skrives i jinja2 templates, som gondul kompilerer.
- Gondul kan compilere templates som er lagret på server, eller ta i mot fra POST, så man kan skrive “one-off” om man vil.
- Kantsvitsjer er helautomatisk provisjonert etter initiell konfigurering er satt opp:
 - Lag svitsjen i Gondul, si hvor du skal koble den til
 - Koble den til
 - Funk
- Resten gjøres normalt ved at man henter konfigurering “manuelt” fra gondul.

Grunnleggende arkitektur



Design



- All magi ligger i JS-frontend og i at APlet er svært simpelt
- Ekstremt simple collectors – Perl
- Enkle konvensjoner for config som Gondul tolker:
 - Uplinks som ikke er i bruk må settes i Admin Down – vi konfer typisk opp 4 uplinks men kabler 2 til 3.
 - ifAlias settes til “Fysisk”, “Gruppe”, “AP”, “Client” avhengig av bruk.
- Time Series er adskilt, tg17: graphite, tg18: influx
- Grafana var nesten ikke brukt....



Designføring: Utvikling

- Brukt mye tid på simpelt design som lar seg endre “i fart”
- Legges alltid til mer logikk i frontend under arrangementet
 - “Vi trenger oversikt over CPU-ytelse på distroene”
 - På plass 15 minutter senere.
- Store løft tas før påske, “business logikk” kommer typisk under arrangementet og en uke eller to etter.

API



- Prinsippet er: Send all info om infrastruktur over til klient
- Heftig bruk av moderne cache-mekanismer
- Generisk data (rå SNMP-data), samt logger, er passordbeskytta
- Normalisert data som ikke kan knyttes til en deltakerport er “public” - temperatur, latens, båndbredde på uplinks, mm.

Frontend: Native JavaScript



- Forholdsvis bra modularisert
 - En modul henter data og kjører event-handlere når det endres
 - En annen modul registrerer slike event-handlere og bygger opp info om hva det betyr. Denne modulen består av masse mikro-moduler på 10-40 linjer kode, per “business case”.
 - Kart-tegning er separat fra det igjen
 - Osv osv



```
function tempInfo(sw)
{
    var ret = new handlerInfo("temp", "Temperature");
    ret.why = "Temp";
    ret.score = 0;
    ret.data[0].value = "N/A";
    if (testTree(nmsData, ['switchstate', 'switches', sw, 'temp'])) {
        var temp = nmsData.switchstate.switches[sw].temp;
        if (temp == undefined) {
            ret.data[0].value = "N/A";
        } else {
            temp = parseInt(temp);
            ret.data[0].value = temp + "°C";
        }
        if (temp > 30 && temp < 50) {
            ret.score = temp;
        } else if (temp > 50 && temp < 100) {
            ret.score = temp * 5;
            ret.why = "Very high temperature";
        }
    }
    return ret;
}
```



```
function snmpUpInfo(sw) {
  var ret = new handlerInfo("uplink","SNMP uplink data");
  ret.why = "No SNMP data"; ret.score = 0;
  if (testTree(nmsData,['snmp','snmp',sw, 'ports'])) {
    var total_up = 0; var seen_up = 0;
    for (var port in nmsData.snmp.snmp[sw].ports) {
      var x = nmsData.snmp.snmp[sw].ports[port];
      if (x["ifAlias"].match(/Gruppe/i) && x["ifOperStatus"] == "up"){
        total_up += parseInt(x["ifHighSpeed"]);
      }
      if (x["ifAlias"].match(/Fysisk/i) && x["ifOperStatus"] == "up"){
        seen_up += parseInt(x["ifHighSpeed"]);
      }
    }
    ret.data[0].value = "LAG member speed and total speed is " + seen_up;
    if (total_up != seen_up) {
      ret.score = 500;
      if (tagged(sw,'ignoreuplink')) {
        ret.score = 0;
      }
      ret.why = "LAG member (ge/xg/et) speed is " + seen_up +
        " but logical (ae) is " + total_up;
      ret.data[0].value = ret.why;
    }
  }
  return ret;
}
```

Replay



- Vi lagrer all data – alle SNMP-pollinger, alle ping-svar (eller mangel derpå), osv.
- Kan spørre APIet om å gi oss verdensbildet slik det var klokka X
- Kan brukes til å “replaye” historisk data
- **SVÆRT** nyttig under utvikling

Spørsmål?



PETITION FOR INTERNET ~~BACK~~ BACK (P/s)

Olav B
Lars V. Myhre
Sigmund Øre
Benjamin M. Simonsen
De Jakobsen
vegar
Simone
Line
Alexander
JESU
Knut
Ludvig S. Hagberg
Johannes L. E. Eriksen
Markus T. H.
Knut J. J. J.
Tomas K. Braut
Benedikt Tøndevoldshagen
Morten Rognerød
Søren M. V. V.
Eirik M. S. S.
Eirik M. S. S.

Jørgen J
Hans Petter Austråttne
Espen Fremstad
Perk Røpstad
Jørgen Kleuson
Leo Vindenes
Sigmund Rognerød
Sebastian K
Aleksander
Knut
Arne
marcus
Tor
Bjørn
Dyroland
ERLEND
Jonathan
Knut
Mark
Jennis
Ole

Danier V
Petter H
Johann
Simon Ørning
Kristian
Kaia
Jonas Viken
Witt & Toren

PETITION FOR INTERNET BACK LINK 1/4 2

Mats Viken
Børge Østby
Håvard Steenslid
Lasse Brekken
Nathias Holm
Jens H. H. H.
Kristoffer H. H.
Johann Skals
Anders H. H. H.
CHRISTIAN H. H. H.
MARTIN ANDRASEN