

# Reason For Outage Report (RFO):

Regarding power loss Tulegatan site October 11<sup>th</sup> 2011

**Place:** Stockholm

**Date:** October 12<sup>th</sup>-

**Author:** Jørgen Qvist, Chief Network Operating Officer NORDUnet

**Approved:** René Buch, CEO NORDUnet

## Affected Services:

- **Network Services:**

- 15 minutes Loss off redundancy for all SUNET dual connected customers, down time (07:14-07:29)
  - Högskolan Borås(HB) and Karlstad Universitet(KAU) experienced an extended loss of redundancy due to transponder issues
- 15 minutes loss off service for the below Single connected customers, down time (07:14-07:29)
  - Verket för Högskoleservice (VHS)
  - VINNOVA (VINNOVA) - Forskning och innovation för hållbar tillväxt
  - Nordiska Museet (NORDISKA)
  - Svenska Akademien (SVEAK)
  - Moderna Museet (MODM)
  - Naturhistoriska riksmuseet (NRM)
  - MDH Sickla
  - Försvarshögskolan (FHS)
  - Institut Mittag-Leffler (MLI)
  - Riksantikvarieämbetet i Stockholm (RAA)
  - Etnografiska museet (ETNOGR)
  - Sophiahemmet (SOPH)
  - Carl Malmstens Centrum för Träteknik & Design (CTD)
  - Sveriges Utbildningsradio (UR)
  - Sveriges Utbildningsradio (UR)
  - Talboks- och punktskriftsbiblioteket (TPB)
  - Gymnastik- och idrottshögskolan (GIH)
  - KK-stiftelsen (KKS) - Stiftelsen för kunskaps- och kompetensutveckling
  - Institutet för språk och folkminnen (SOFI)
  - Tillväxtverket - Verket för näringslivsutveckling (TILVAXTVERKET)
  - Riksbankens Jubileumsfond (RJ)
  - Nationalmuseum (NATM)
  - Operahögskolan (OHS)
  - Skansen (SKANSEN)
  - Arkitekturmuseet (ARKM)
  - Statens kulturråd (KUR)
  - Kungl. Ingenjörsvetenskapsakademin (IVA)
  - Stockholm International Peace Research Institute (SIPRI)
  - Statens musiksamlingar (SMUS)
  - Danshögskolan (DANS)
  - Kungl. Konsthögskolan (KKH)
  - Tekniska museet (TEKM)
  - Livrustkammaren och Skoklosters slott med Stiftelsen Hallwylska museet (LSH)
  - Armémuseum (ARMEM)
  - Sjöhistoriska museet (SJOM)
  - Kungl. Vetenskapsakademin (KVA)
  - Konstnärsnämnden (KONSTN)

*No dual connected customers had loss of service due to this incident*

- **Other Services**

- E-meeting (Adobe Connect PRO) down time (07:14-12:06)
- SUNET Eduroam (Radius) downtime (07:14-12:06)
- SUNET mail, down time (07:14-12:12)
- University Hosting (Örebro, KTH, TDC DNS, SU), down time (07:14-13:02)
- www.sunet.se and other web services, down time (07:14-13:05)
- SWAMID, down time (07:14-13:05)

## Sequence of events October 11<sup>th</sup> 2011

07:14:03	Utility power lost - UPS load on batteries.
07:14:38	Site Power loss.
07:16:04	The Generator completes start-up and takes load Total power loss on site ~120 seconds.
07:31:20	Utility power restores
07:35:39	The UPS recognises the utility power as being stable for more than 4 minutes and switches back
07:30	All network connections for single connected customers and redundancy for all dual connected customers except KAU and HB are restored.
08:00	At around it became clear that none of the stand-a-lone servers and the VMware cluster servers were able to recognise the SAN. The procedure to re-start the EQL SAN was initiated.
08:30	The procedure was completed, but the servers were not able to see all LUN's on the SAN. The Vcenter was not able to come up and control the VMware setup as it could not see its LUN. The stand-a-lone ESX SQL servers were not able to detect the SAN at all. The procedure to restart all of the ESX servers one at a time was initiated.
09:30	The procedure was completed but the result was still the same that the servers were not able to recognise all LUN's. The Vcenter was not able to come up and control the VMware setup as it could not see it's LUN. The stand-a-lone SQL servers were not able to detect the SAN at all.
10:00	The case was escalated to our vendor and together we went through the same procedures. This work was completed with basically the same result. A complete manual rescan was initiated on one server that finally made that server able to recognise all LUN's
11:30	Vcenter is now up and running and can control the VMware setup. The remaining servers are still not able to see all LUN's. The rescan procedure is now executed on each of the remaining servers.
12:00	Access to all LUN's have been restored and the hosts and services is brought up one by one as follows
12:06	E-meeting (Adobe Connect PRO) services restored SUNET Eduroam (Radius) services restored
12:12	SUNET mail services restored
13:02	The University hosting (Örebro, KTH, TDC DNS, SU), was restored following the replacement of a broken switch that was up, but not passing data
13:05	www.sunet.se and other web services restored SWAMID services restored
14:08	Redundant network services to Karlstad universitet were restored as below after resetting the transponders manually
14:16	Redundant network services to Högskolan Borås were restored as below after resetting the transponders manually
14:16	All services were verified as operational. End Event.

## Root cause and corrective actions

### *Root causes*

- 1) The root cause of the problem was loss of site power, caused by the UPS batteries not being able to take the load long enough for the generator to start (120 sec Start Up Time).  
The battery setup is based on 2 separate strings (A/B) of 33 batteries each, delivering a redundant 400V input to the UPS.

The battery banks are designed to be able to take the load for +60 minutes with a capacity of 150ah and an operational load of 135A.

Based on the physical findings, the theory is that one bad battery in the B string caused that string to short circuit. The full current was then pulled from the A string causing that battery string to overheat and collapse.

Investigations of the batteries are on-going to establish the actual cause of the failure.

- 2) Our vendor is still investigating the root cause for the lengthy outages for the servers dependent on the SAN. All indications points towards issues with the ICSSI based SAN but no plausible cause or explanation has been identified so far.

### *Corrective actions*

All UPS batteries were replaced on October 13<sup>th</sup> 2011. Further investigations regarding the battery incident will be performed by an external party, which will also perform a complete audit of the design and installation. The audit will also include existing maintenance and operational procedures.

Our Vendor for the VMware, Servers and SAN is continuing to investigate the cause of the SAN not being able to communicate correctly with the servers. A full Audit of the setup will be performed and corrective actions will be done to ensure that a similar situation will not occur.

NORDUnet Optical Team will together with the equipment vendor investigate the cause for the required manual reset of the transponders.

In addition to the above the following corrective actions have been initiated:

- For essential services where there is a redundant server setup where both servers are currently virtualised, one of the servers will be changed to a physical machine, where as the other remains virtualised, but by affinity rules locked to a site different to where the physical server is located.
- A number of issues where affinity rules had not been done correctly, has been identified, and procedure put in place to apply affinity rules on all redundant server pairs.
- The Vcenter machine will be changed to a physical machine with local storage.
- A system for off line access to documentation will be implemented.
- Incident manager appointment to be done according to mgt. escalation list.
- Incident manager has been made responsible for customer communication.