COMMS COUNCIL UK

THE VOICE OF ADVANCED COMMUNICATIONS

Recommendations for ITSP Resilience

Version 1.0
November 2019

Contact: team@commscouncil.uk

COMMS COUNCIL UK

THE VOICE OF ADVANCED COMMUNICATIONS

Contents

COMMS COUNCIL UK

THE VOICE OF ADVANCED COMMUNICATIONS

Summary

This document gives guidance on what measures an ITSP should take to ensure that the resilience of network and systems are adequate.

The document is designed to be read in order but the final section contains a checklist for the ITSP to run through to ensure that the major points have been covered in their own implementation.

This document is not exhaustive on the subject and is provided as guidance only.

Introduction

This document has been created as an advisory for Comms Council UK members. It has been produced by members of Comms Council UK to encourage best practices to be developed and used within the industry as a whole.

Purpose

This document aims to outline a series of Resilience measures to assist ITSPs to:

- Comply with legal obligations
- Use technical mechanisms to produce a secure network

Legal Aspects

The legal aspects are defined by Ofcom's General Conditions of Entitlement, and in particular the following 2 conditions;

General Condition A3 - Availability of services and access to emergency services

"*This condition aims to ensure the fullest possible availability of public communications services at all times, including in the event of a disaster or catastrophic network failure, and uninterrupted access to emergency organisations. It requires providers of call services to ensure that calls can be made to emergency organisations free of charge and to make caller location information available to emergency organisations where technically feasible. It also includes specific rules relating to providers of VoIP outbound call services which aim to ensure that users of those services are aware of any potential limitations on making calls to emergency organisations and that accurate and up-to-date caller location information can be provided to the emergency organisations where possible*"

This requires ITSPs to ensure the following:

- The fullest possible availability telephone services provided in the event of catastrophic network breakdown or in cases of force majeure
- Uninterrupted access to emergency organisations, and the supply of Caller Location Information for these (999/112) calls.

General Condition A4 - Emergency Planning

"*This condition requires all communications providers who provide publicly available telephone services or a public telephone network over which these services are provided, to agree arrangements with emergency organisations and other public authorities to ensure the provision or rapid restoration of networks and services in the event of a disaster*."

This requires ITSPs to provide the following:

- Provision or restoration of services following an emergency
- And implicitly requires ITSPs to provide networks resilient against emergencies and failures

COMMS COUNCIL UK

THE VOICE OF ADVANCED COMMUNICATIONS

Technical Aspects

Geographic Resilience

A geographic redundant system intends to safeguard the system against geographic disperse disasters, therefore redundant systems should be geographically diverse.

Most telecommunication systems will have a "worker / standby" architecture, to allow the standby side to take over if the worker side is unavailable (due to maintenance or failure). To ensure that the maximum redundancy is maintained these 2 seperate sides should be in different data centres in different geographic locations.

What should be considered when selecting geographic locations?

- Are power supplies separate?
- Is the latency between the 2 sides low enough to allow the system to function correctly?
- Is connectivity to upstream suppliers and downstream users available at both locations?

End User Experience

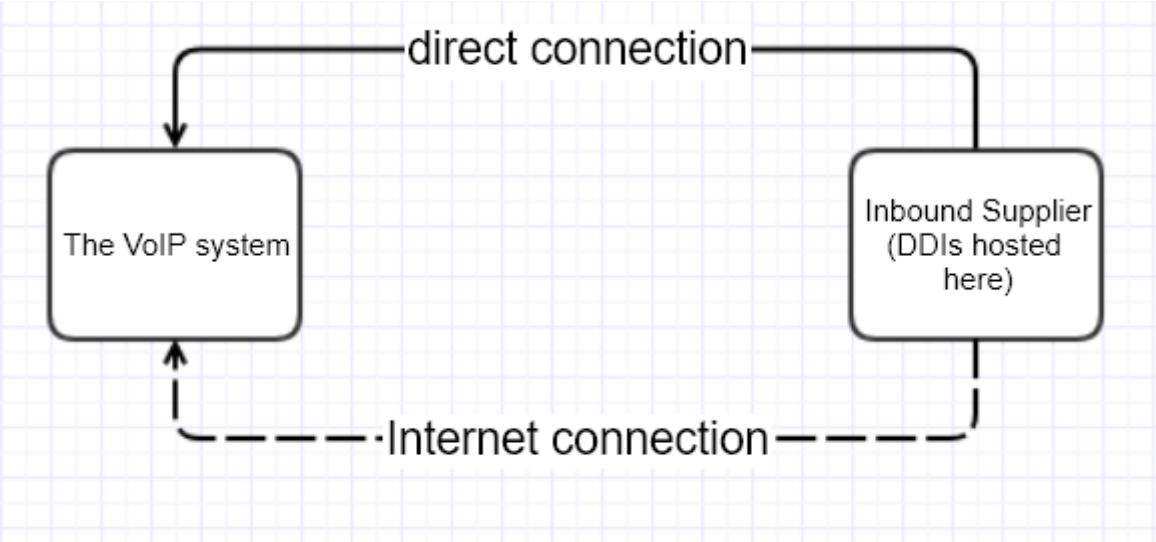
The end users should not be able to tell which geographic location is handling either their outbound or inbound calls at any time. Ideally this will include calls in progress when the worker / standby sides reverse roles, i.e. when the standby takes over the handling of calls from the original worker.

Geographic redundancy of upstream suppliers

It is important to ensure that connectivity to upstream suppliers is also geographically redundant, either by connecting to individual suppliers from more than one geographic location or by connecting to different suppliers from different locations.

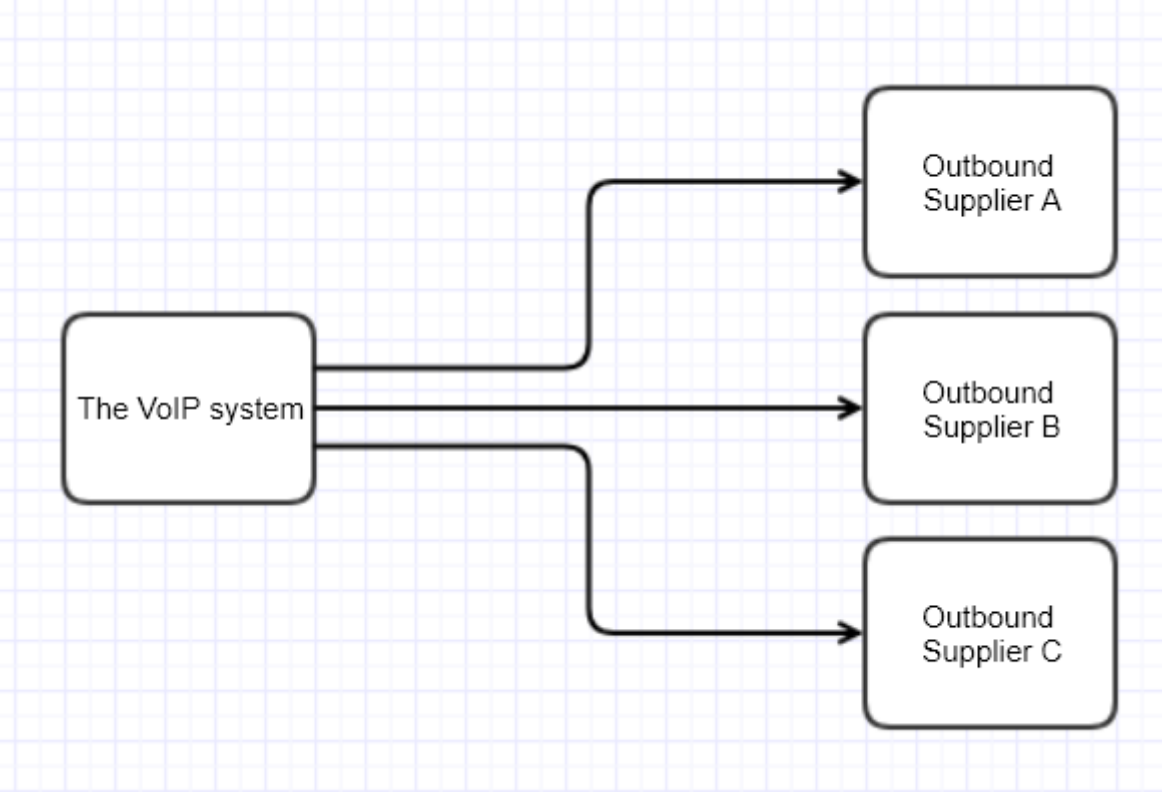
Upstream suppliers or "interconnects" are critical to providing "breakout" capability for any VoIP offering; so ensuring that they are available at all times is essential.

Suppliers can provide both inbound and outbound call capability, but inbound (hosted) telephone numbers will normally only be provided by a single supplier. This means that multiple diverse routes to an inbound number supplier is essential, otherwise loss of connectivity to the supplier means that inbound service is lost.



In the figure above inbound traffic would normally be delivered by the direct connection, but in the event of failure, calls would be delivered across the internet.

For outbound service multiple and diverse suppliers is always preferable to a single supplier, and should be programmed into dial-plans to "failover" to alternate suppliers in the event of becoming isolated from an individual supplier.



The figure above shows a possible routing plan where 3 different suppliers are available, which supplier is chosen for a call is likely to depend on cost and quality issues but if a supplier becomes unavailable, then other suppliers are available to handle the call.

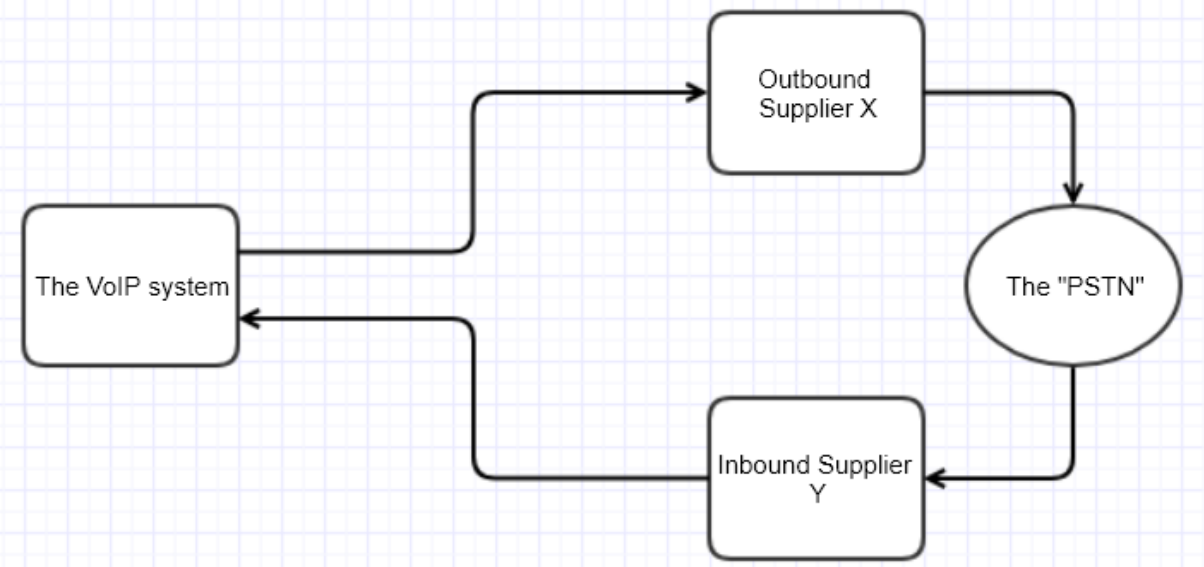
End user experience

Although one call may be routed through one supplier and the next through another supplier, it is important to ensure that the user is unaware of this. The end user is only concerned with the call being successfully set up, in a timely manner and the quality of the voice over the call.

Routine Calls

The use of (OPTION) Pings to verify the connectivity of a supplier is the most basic monitoring action that can be done; but in reality the only way to test the quality of a supplier is to pass calls through them. In order to monitor suppliers continuously "routine" calls should be made and the MOS values for the calls identified; there are many monitoring systems that are available for this action.

It should be noted that if you try to loop a call through a supplier that handles both your inbound and outbound calls they may reject it (BT IPX for instance) as they do not expect to send a call back to where it came from.



The figure above shows a typical call route for monitoring the quality of calls; the call is sent out on supplier X and is subsequently received from supplier Y. By use of a combination of routine calls the connection that is causing a quality failure can be established.

Automation of the calls allows for regular (routine) monitoring of call quality (a similar approach of making a selection of looped and simple calls can be done for specific diagnostics in the event of a problem being detected).

Emergency Call Handling (999/112)

Your obligations under the Ofcom conditions are perhaps the most serious due to the obvious implications if the service fails, therefore Ofcom do regularly monitor industry to ensure compliance. The following guidance is provided to help members review the resilience of their solution and identify any areas that should be improved as a matter of priority.

End user experience

Systems must support dialling of 999 and 112 and should be able to transit your servers in the same way as any other call to enable simple integration with end user devices. It is generally considered good practice to only pass authenticated calls that have been subject to usual account checks (i.e. calls originated from existing clients only, authorised Caller ID used or inserted etc) but calls should not be restricted based on lack of credit.

Using another ITSP

For Service Providers using another ITSP to carry your calls you must ensure that you have carried out initial tests with them to verify that emergency calls will pass without restriction. The PECN will insist on a number of conditions such as using a unique and valid CLI for the call and the requirement to update the location information database regularly, however they should not block live emergency calls if these conditions are not met.

It is important to note that the Emergency Handling Authority, Ofcom and the ITSP will carry out routine checks in an effort to maintain the effectiveness and reliability of the service and associated data and therefore any ongoing breach of the conditions may result in your service being suspended. It is vital that all members take their responsibilities in this matter seriously and work with other parties in the call path to improve service where possible.

If using a transit network you must also satisfy yourselves that they too offer a robust network solution, following the standards listed below, as a failure within their own network will cause call failures for you and your end users. Likewise, if your account with the network should be disabled or suspended for any reason all calls are likely to fail so you must take alternative steps to continue uninterrupted service.

Network resilience

Calls to emergency service numbers should be identified immediately by border authentication services, and further downstream routing choices made having already predetermined that the final termination will be to your designated ITSP (or in the case of the ITSP, to BT as no other carrier is used for emergency services calls).

A resilient solution would require multiple call processing units to handle these calls, all of which should be monitored for availability and removed from routing automatically if

problems are detected. It is essential to assess for any single points of failure that may prevent a call reaching the emergency call handling authority and putting alternatives in place.

This may mean hosting equipment on a separate site, interconnecting to two PECN networks or having resilient interconnects with your chosen PECN. The PECN will need multiple handover points to the EHA(BT) and monitoring in place to ensure those points remain available. Automatic failovers that can detect breaks in service and route around them would be a sensible minimum standard.

In addition to addressing single points of failure, it is considered good practice to have monitoring in place which will notify your team in the event of any failures within the network. For many networks any failure is likely to put them in a vulnerable position which a single point in service therefore it's crucial to be able to fix and restore full service as a matter of urgency.

Of course the extent to which a network can be resilient must be balanced with technical and financial feasibility but we would recommend that a full assessment be carried out and where potential improvements are not feasible these should be documented with the steps being taken to mitigate the risks. In the event of an outage Ofcom are likely to investigate all parties in the call path and will expect to see that all reasonable steps have been taken to make the service meet your obligations under the GCs.

Back up

ITSPs should also consider what alternative methods of dialling the emergency services the end user will have should their service fail or if there is a power outage. VoIP providers specifically are required to notify their users of the limitations of the service in the event of a power failure or loss of broadband service during the sales process.

Traditionally the advice has been to recommend that users maintain a PSTN line for this purpose, however this is becoming less appropriate and therefore SPs should offer advice relevant to the users situation. This may include ensuring they have a mobile in their location or an alternative service provider if needed. As of November 2019 *Ofcom are consulting on guidance in this area which may result in a requirement for a wider range of providers to offer a battery back-up solutions.*

Distributed denial-of-service attack (DDoS)

A distributed denial-of-service attack (DDoS) is one of the toughest foes a service provider can face and currently one of the most high-profile cyber security threats.

Every organisation that depends on Internet-facing assets should have a DDoS protection plan. When fully prepared for a DDoS attack with DDoS protection already in place, and a DDoS

response runbook, it is easy to respond effectively to DDoS incidents and quickly mitigate operational, financial, regulatory and reputational damage.

Planning ahead and being prepared is a good practice for business operations and your best defence against DDoS attacks.

ITSPs should consider the following as vulnerable to DDOS attacks:

- Session Border Controllers (SBCs) & Registrars:
- Domain Name System (DNS) Servers: DNS infrastructure is a ripe target for malicious actors, because it provides a necessary service for end users devices and browsers to find your services.
- Websites: Even a simple DDoS attack can flood an unprotected website with a high volume of requests that exceeds its capacity.
- Web Applications: A web application can't easily tell the difference between a DDoS attack and legitimate user requests. Login pages are often targeted, because they trigger back-end processes that consume CPU cycles on the web server, such as fraud prevention, database access, and authentication routines.
- Application Programming Interfaces (APIs): APIs are being targeted more frequently, in part because more and more websites are enabling communications through APIs. APIs may feed information to mobile apps or pull content from third-party content sources into a web application.
- Data Center & Network Infrastructure: Network and data center infrastructure, and network bandwidth into a data center, are targets. If an attacker can fill your network pipes or overwhelm routers and switches, legitimate traffic can't get through.

DDOS mitigation

- Determine and ensure that your infrastructure has sufficient balance with headroom above and beyond legitimate peak requirements.
- Determine and ensure that your network capacity is in excess and has enough headroom, in excess of normal peak traffic, to withstand a DDOS attack.
    - For example, if your network is hit with 10 Gbps of traffic from a reflection DDoS attack with hundreds of sources, how long will it take you to block it using an access control list (ACL)?
- How large of a DDoS attack will you attempt to mitigate before you decide to blackhole traffic?
- If traffic is blackholed due to a DDoS attack, what requirements do you have before you restore service?
- Consider blocking known fraudulent IP ranges that will never need to access to the system.

DNS - For Device Redundancy

Device / User Agent (UA) & SIP Trunk redundancy can be achieved using hierarchical, weighted DNS records, for example:

1. NAPTR (Name Authority Pointer Record), which resolves to:
2. SRVs (Service Records), which resolves to:
3. A Records, Prioritised, which resolve to:
4. Asset IPs

- Deployments using DNS SRV must be tested as behaviour varies between different vendors.

A worked example being:

1. NAPTR: Access.foo.com resolves to prioritised SRV records:

2. SRVs: Configured as a high priority TLS SRV record and a lower priority UDP SRV record to fallback to in case there's a problem with TLS.
    a. _sips._tcp.access.foo.com (TLS, Order 10 / Preference 100)
    b. _sip._udp.access.foo.com (UDP, Order 20 / Preference 1000)

3. DNS A Records associated with the IPs: The SRVs resolve to prioritised DNS A Records
    a. sbc1.foo.com
    b. sbc2.foo.com
    c. sbc3.foo.com
    d. sbc4.foo.com

Testing and Monitoring

Testing is the exceptional actions to ensure that network components and the whole service works as desired. Monitoring is the regular activities that report on the behaviour of the components and service.

Testing

Regular and effective testing of all network components is essential. A detailed test plan should be developed as part of the initial network design process. This plan must be regularly reviewed and updated. The test plan must cover both normal operation and the operation of resilience and fall-back features. Designing and implementing a test plan to cover resilience features is difficult as the tests should exercise those features while minimising the impact of live network traffic. It is essential that the resilience features are regularly tested. Designing a

test plan but failing to run regular tests generates a false sense of security and is arguably worse than having no test plan.

A test plan must be comprehensive and must cover every aspect of network resilience. The other sections of this document are a good starting point to develop a plan. The plan should be sufficiently detailed to enable personnel not familiar with the details of the network design and operation to run the tests and interpret the result. Where possible tests should be automated allowing them to be started manually or automatically. The test report should highlight any test failures.

As testing resilience and fall-back features has the potential to degrade or disrupt service, consideration should be given to establishing a test network which replicates the live network but on a smaller scale. The test network can be used to develop and validate the test plan and exercise the resilience features. It also enables testing of any planned changes prior to implementation. However, using a test network in this way is no substitute for testing on the live network, so live network tests should be run at suitable intervals and after every significant change. The interval between scheduled live network tests is down to individual preference, but every 6 months is a realistic goal. The test schedule should be planned for at least 18 months in advance. A scheduled test must not be postponed unless there is a compelling business reason for doing so.

Live network tests should be scheduled outside of peak hours and should be run only when resources are available to quickly identify and rectify any problem that may arise. The availability of suitable resources must be defined as a prerequisite in the test plan. The tests should be run under a range of load conditions including normal and peak operational loads. If, as recommended, tests are run outside of peak hours, a traffic generator should be used to simulate peak load.

At a minimum resilience testing should include:

- Failure of one of more key components
- Failure of one or more external software services (e.g. DNS)
- Failure of one or more external service providers
- Failure of one or more external network links
- Failure of one or more environmental services (power, air conditioning)
- Ability of the network to withstand or limit the effect of a DoS attack

The test plan should detail how these failures can be simulated and to validate that the designed resilience features operated correctly.

Monitoring

Monitoring is fundamental to sustaining a proactive approach to network resilience.

Networks should operate a 24/7 NOC responding to upstream network mailing lists (IXPs as an example)

Networks should regularly benchmark and maintain records of so that relevant alerts can triggered by monitoring. A general rule of thumb is that any link should have the capacity to service the entire load under normal conditions.

Software automation in monitoring allows NOCs to respond positively before degradation in, or a total loss; of service occurs. Monitoring software is widely available in combinations of free/paid and open/closed source.

Monitoring could be as basic as a collection of MRTG graphs or a comprehensive package such as Nagios, OpenNMS, Zabbix, PandoraFMS, PRTG, Solarwinds or Splunk

Common protocols like SNMP (when implemented securely - v3 only) are supported by all vendors and offer comprehensive metrics.

Most vendors implement proprietary/open protocols which can be leveraged with along with SNMP.

- Cisco (Netflow)
- Juniper (Jflow)
- sFlow, or even SDN integrated applications (Arista, BSN)

Monitoring the quality of telephony routes can be simply achieved either via support from the packages listed above (Solarwinds) or indeed specific packages (HOMER) or via scripted SIP calls via both Off-net and On-net origination to capture a MOS average at required intervals. Using a widely supported codec (G.711) with a benchmark of 4.4 as an acceptable MOS. External SIP call monitoring companies exist (VoiPSpear).

Risk Assessment

Network resilience will always have risks associated with it, for example, even if you have two diverse networks to carry your traffic, it is possible that both networks will have failures at the same time from independent incidents.

It is important to assess the risks that can impact on your network and international standards such as ISO 27001 give a framework for doing just that. Five simple steps can be used to do a risk assessment:

- Establish a framework for risk assessment
  - This is the mechanism you are going to use to assess risks on a regular basis
- Identify risks
  - Find out what risks can impact your network
- Analyse risks
  - Find out what the impact of the risk is
- Evaluate risks
  - Find out what the likelihood of the risk happening is
- Decide how to deal with the risk

What can you do to minimise the likelihood of it happening and how can you minimise the impact if it does happen. Remember to take into account cost as well as if a risk is low impact and unlikely to happen is it worth spending lots of money in guarding against?

COMMS COUNCIL UK

THE VOICE OF ADVANCED COMMUNICATIONS

Checklist

Item	OK?
Do a risk assessment and use this to decide what actions your business needs to take	
Check the geographic resilience both of your systems and of the networks of your upstream providers	
How are emergency calls going to be supported in the event of network failures?	
Protection against DDOS attacks, can your upstream supplier protect you?	
Are the end devices, telephones, protected by use of DNS and alternative destinations for both outbound and inbound calls?	
Create a plan for testing and run it	
Check that your monitoring of network and service performance is adequate and will allow diagnostics to be run in case of problems	
This is an ongoing process. Agree a regular calendar review date.	