CONSULT21

BT wholesale

CONDOC CLOSURE DOCUMENT

Final Conclusion to BT's Consultation regarding MSIL Resilience (C21-NS-016)

Document number: C21-NS-027

Issue: Issue 2.0

Date: 31st January 2008

Author: Martin Wade and Ian Lewis, BT Design

IMPORTANT NOTE

This document is supplied by BT to Communications Providers, being customers of BT Wholesale and or Openreach, as part of the Consult21 consultation process.

Details within it may be subject to change as a result of that process and / or for technical, commercial or operational reasons.

1 Introduction

BT wishes to formally place on record that industry engagement has been undertaken and concluded for the activities outlined below.

This document forms part of the overall consultation and closure process.

2 Industry Consultation Areas

2.1 Overview

BT has previously stated its intention to launch a new 21CN Multi- Service Interconnect Link – MSIL. The MSIL is intended to support traffic from different services over the same physical link. Services are separated using VLAN tags. Services may use many tags themselves for different flows (e.g. one or more for Voice Media and one or more Voice Signalling) over the same MSIL. This consultation is only concerned with Ethernet MSIL.

MSIL will be available in 1G and 10G interface variations. 100M MSILs are subject to a Consult21 process (see C21-PP-024). 1G MSILs will be available from Metro or Core node locations. 10G MSILs will be available from Core node locations.

This consultation is managed under the auspices of Consult21. Details may be found at: www.btwholesale.com (paste link into browser) and following the "21CN & Consult 21" link

This consultation regarding 21CN MSIL has covered:

- The resilience functions needed on the Multi-Service Access Port/s that connects to the MSIL in the BT site.
- Resilience mechanisms to protect the MSIL and services provided across it.

This document captures the responses of industry and provides a statement of BT's intentions after consideration of those responses.

2.2 What did we consult on?

This consultation document described BT's proposals and sought responses from industry on the options for MSIL resilience in the following areas:

- Options for In-Building, In-Span and Customer Sited handover with resilience
- The use of LAG and LACP to enable resilience
- The use of MSIL using Openreach RO1 and RO2 options
- The various architectural options for using multiple MSILs to enable a resilient service

2.3 Summary timetable and of consultation related documents

Consultation document issued to CPs	9th November 2007
Discussion / Questions on consultation at Point to Point WG and by audio	21st November 2007

Bilateral discussions Begin (Face to Face and Audio)	19th November 2007
Bilateral Discussions End	3rd December 2007
Deadline for Responses from Communication Providers – note date extended to 14th Dec at CPs request	14th December 2007
BT Publishes Consolidated Responses & Conclusions – draft for comment	18th January 2008
BT Publishes final issue and consultation closure	1st February 2008

2.4 Number of Industry Responses Received

Received within the due date:	1
Received after due date but in time for consideration:	4
Total Received/Progressed:	5

2.5 Summary of Industry Responses and BT initial response

Industry responses have been summarised to retain their anonymity and to shorten the length of this document.

Question 1. Do CPs see this method of using LAG as beneficial as a method to make an MSIL resilient? Do you agree with this approach to using LAG?
(This relates to using LAG without LACP)

CP1

We would question whether "each makes its own decision" causes monitoring and reporting issues for the services – eg NGNCC and WBC – that traverse the MSIL. It may be more appropriate to nominate a primary link and inform the other party. Otherwise agreed.

CP2

Providing a LAG without any control protocol can result in unexpected behaviour in failure scenarios such as unidirectional failures. This is especially true for 1GE which does not directly have a remote defect notification capability. This option would not be preferred.

CP3

No Comment

CP4

Additional clarity needs to be provided with regards to how many MSILs can be part of a LAG. If it is only two to allow auto fall-over (ie 1+1 resilient configuration) in general it will come down to cost against exposure to customer requirements. Also due to BT's changes to network topology and stranded asset issues closure examination of viable LAG groups would have to be considered.

If more that two MSILs can be part of a LAG then this becomes a n+1 resilience and whilst is still not desirable is a better option but again this will depend on pricing differences between 100m to 1Gig and 1Gig to 10Gig. However it is unclear if this could be implemented in an auto fall-over situation.

CP5

The lack of auto-negotiation and auto-configuration features is unattractive. Manual configuration increases the chance of misdelivery through misconfiguration. This option is seen as less desirable than the LACP version.

BT Response:

Each LAG peer can make its own independent choice on how it distributes traffic between the ports. Determining which port is actually used for traffic can depend on the traffic content and the network equipment distributing the traffic. We do not see that it is practical to nominate a primary link when LAG is working in a load-sharing mode since the link used by the traffic cannot necessarily be determined.

Two MSILs are being proposed here to be grouped together to provide resilience.

The difficulty with operating with more than two links in a LAG is how traffic is shared among them. The load-sharing is performed at the Ethernet layer and traffic within one VLAN uses one link. A VLAN's size is then limited by the size of one link. Further, with a few large VLANs the chance of evenly distributing the VLANs among the links is small. To achieve better load-sharing the size of each VLAN should be small compared to the size of the links. This does not present a useful model if a single VLAN may be used when the MSIL is just a single link, but needs to be broken down into many smaller ones to achieve load sharing when additional links are added.

Additionally see the response to Question 2 for potential impact to QoS features where load sharing LAG is used.

Question 2. **Do CPs see this method of using LAG as beneficial as a method to make an MSIL resilient and as an option they would use?**
(This relates to using LAG with LACP in a 1:1 mode)

CP1

It's not clear whether BT maintaining the worker/standby control has a unidirectional or bidirectional impact. Please clarify whether traffic from the CP to BT would be via the BT nominated link. Given that BT's rationale with this method is to improve QoS features. We would like more detail of the QoS impact on supported services of the non-LACP and LACP approaches. We see value in this method.

CP2

LAG with LACP in active/standby does seem like a logical option. We agree that the implications of using LAG in load-sharing mode would not be advisable, because of the added complexity of controlling QoS and Admission Control.

CP3

We would support this method of using LAG (LAG with LACP SYNC flag control option) above the method without LACP. This could then be combined with BFD. This would allow the BFD packets to always flow on the active link. This will increase the physical failure detection on the MSIL link.

CP4

Whilst we would have no issue with who controls worker/standby for automated fall-over the main issue for use is cost.

CP5

LACP auto-negotiation and auto-configuration render this option preferable to the preceding option.

BT Response:

This intention is to achieve a worker/standby operation end to end. The most robust way we see that this can be done is if one LAG peer controls the worker/standby nature of the links and the other peer listens to it. In this way the peer controlling the operation implements worker/standby LAG while the other peer uses load-sharing LAG. Since LACP is used, the peer controlling the operation signals which link is to be used by setting the SYNC flag in the LACP PDUs. The peer performing the load-sharing LAG is then forced to use only one of the two links. This avoids the possible problem of each peer trying to control the worker/standby operation and each peer selecting a different link, which would result in zero throughput. Hence traffic from the CP to BT would also be forced to use the nominated link.

Network equipment generally performs rate limiting and traffic shaping functions in a distributed fashion. This is to enable QoS features to be applied to traffic operating at a high rate. This can mean that these functions are implemented separately for each port in a LAG and not on the aggregate traffic for all ports in the LAG. The result is that more traffic may be allowed into the network by a policing function and that traffic cannot be shaped accurately to the contracted rate by a shaping function. For instance, a 100Mbps policing function may allow 100Mbps in on each port rather than 100Mbps in total. The intention of using LACP to restrict the ports in a LAG to operate in a 1:1 mode, is to overcome this limitation by ensuring that only one port is in use at any one time. Any load-sharing solution needs to consider the potential impact to the network by admitting more than the contracted traffic.

Question 3. **Do CPs see this method of using LAG as beneficial as a method to make an MSIL resilient? Do you agree with this approach to using LAG?**
(This relates to using LAG without LACP)

CP1

No comment.

CP2

As Q1.

CP3

No Comment

CP4

This in theory is the same as Answer1 but for CSI delivered service and therefore has the same issues. However it is further restricting the CP by not allowing the use of 100m MSIL.

CP5

The lack of auto-negotiation and auto-configuration features is unattractive. Manual configuration increases the chance of misdelivery through misconfiguration. This option is seen as less desirable

than the LACP version.

BT Response:

The ability to use LAG is restricted to 1G and 10G MSILs since the LAG feature is not universally available on the network equipment used to terminate 100M MSILs. Support of LAG on 100M MSILs requires development and is dependent on volumes and the outcome of C21-PP-024.

Question 4.	**Do CPs see this method of using LAG as beneficial as a method to make an MSIL resilient and as an option they would use?** (This relates to using LAG with LACP in a 1:1 mode)

CP1

No comment.

CP2

As Q2

CP3

We would support this method of using LAG (LAG with LACP SYNC flag control option) above the method without LACP. This could then be combined with BFD. This would allow the BFD packets to always flow on the active link. This will increase the physical failure detection on the MSIL link.

CP4

Again same as Answer 2

CP5

LACP auto-negotiation and auto-configuration render this option preferable to the preceding option.

It is difficult to express a preference between 1) the Dual Link LAC with LACP and 2) the OpenReach LAG with LACP as this will depend upon pricing differentials, and availability figures for OpenReach RO2. We would also be concerned about convergence times for these solutions, which we believe likely to be of the order of 1 second. We would like this decreased as far as possible.

Does BT also plan to control LACP global addressing/naming?

BT Response:

See response to Question 2

The convergence time achieved will depend on the type of failure and the support of Link Loss Forwarding between the Openreach product and the network equipment. With LLF support we believe convergence can occur within a few hundred milliseconds. BT has no plans to control the global addressing or naming of LACP. It is expected that the MAC addresses used for LACP are automatically allocated by the network equipment and so no global control of this address space is necessary.

Question 5.	**Are there any other resilience options for a single MSIL that CPs would see a benefit in having? If so please identify them.**

CP1

We would welcome extension of LAG to reflect the inherent loadsharing capability, however we recognise that this is not a resilience issue.

CP2

We could see a case for 1G or 10G being a mix - IBH/ISH on one leg, and one leg CSH – using LACP. This permits a level of physical diversity between a CP's own-fibre route to the BT node, and an Openreach WES service.

CP3

A suitable N x 1G, N x 10G solution needs to be addressed. At the moment the only option proposed is BGP load balancing based on two MSIL's. This would not provide a granular load-sharing mechanism, just site/node resilience with some crude load sharing based on subnets.

CP4

No Comment

CP5

No

BT Response:

Regarding load-sharing LAG see the response to Question 1 and Question 2. BT have not pursued a general load-sharing solution due to the inability to easily utilise multiple links and the potential restrictions to QoS features.

It shall be possible to order two MSILs each using a different delivery mechanism (e.g. IBH and CSH). However, it will be difficult to implement and maintain the diversity of the MSILs since the operators' infrastructure is not apparent to one another (e.g. Openreach's CSH and the CP's own fibre route).

Question 6. Are there any other resilience options for a pair of MSILs that CPs would see a benefit in having? If so please identify them and indicate for which services they would be applicable.

CP1

No comment.

CP2

The 3 options offered highlight that wBC and NGN CC are fundamentally different – wBC demands same Metro node; NGN CC expects different. It would be useful to have an option for WBC that allowed handoff at different metros, and backhauled the broadband traffic to the main PoSI.

We do not have enough information yet on IPConnect and WBMC to understand how these services would use 2 or more MSILs on different nodes.

However, we would point out that "Layer 1" resilience using the LAG/LACP as per section 2 on individual MSILs increases the availability substantially, and can be invisible to any service running over it.

CP3

No comment

CP4

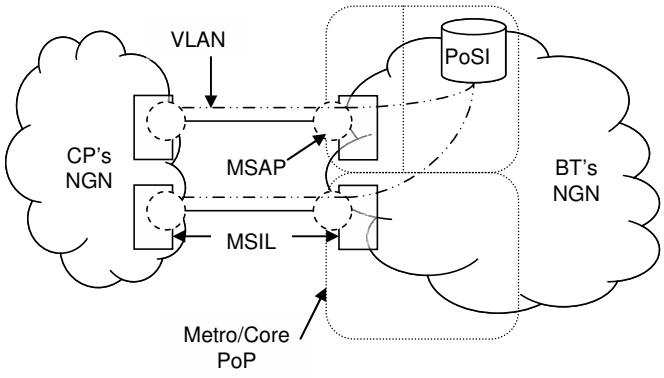
No additional comments or alternative considerations

CP5

No, the options for multiple MSILs seem satisfactory.

BT Response:

The following diagram illustrates the option to have two MSILs delivered to two different Metro Nodes and connect to the same PoSI.



Technically this can be constructed but there are limitations to the amount of bandwidth that can be provided across the core between the Nodes. Further work is required to determine which products can effectively use this resilience option.

General Comments

CP1

We would want to understand more about how BT will ensure, perhaps guarantee, path diversity of resilient MSIL?

CP2

Whilst the document talks about technical options for resilience, there is no discussion on the commercials. We have pointed out for some months that for resilience, we would expect MSIL ports to be charged at a lower cost than a main port, rather than the same. BT's pessimistic assumption on availability of an MSIL port makes resilience between BT and CPs highly desirable by both parties. Where ports are being used as 1+1, then costs should be less than x 2 - we're willing to discuss the trade-off of one-off vs recurring.

There is also nothing on how upgrading from 1G to N*1G or 10G can be best handled especially when resilience is added – for example moving from 1G unprotected to 10G protected.

There are also no comments on the cost impacts on different service requirements and their reaction to different MSIL resilience. For voice services, the architecture suggested follows 3.1, figure 9; for broadband the architecture is 2.1.3/2.2.3. Yet in voice we look at reciprocal flows of traffic to derive costs falling where they lie – and hence rebates would be based on unprotected MSILs. However the incremental benefit to each company if voice were put over a protected 10G used for broadband as well would end up with the rebate not covering the protection port use – the CP would end up paying the BT end as well.

CP3

No comment

CP4

No comment

BT Response:

Migration from a 1G to a 10G MSIL would require significant network configuration. If the CP site is within range of a Core Node (which delivers 10G MSILs) this migration is best performed as a provide

and cease with specific per service migrations from one MSIL to another. The 10G MSIL CSH delivery is distance limited by the underlying Openreach product and CPs may find that they are unable to upgrade to resilient 10G due to this distance limitation.

Introducing resilience in the form of LAG will result in some disruption when it is first enabled. It requires coordination between BT and the CP to minimise disruption.

A resilient MSIL (single parent with LAG) can provide diversity. Diversity to two parents cannot in general be guaranteed. Though use of a CSH MSIL can provide diversity with two parents.

3 BT Closing Statements

After consideration of the Industry feedback, BTWholesale can confirm the following conclusions:

Summary of CP comments

- LAG without LACP is not preferred
- LAG with LACP working in 1:1 mode is preferred
- CPs would like options for LAG N
- CPs would like more information on the diversity that can be achieved between two MSILs
- CPs would like more information on migrating between MSILs

BT's technical proposals for solutions.

- BT will develop a resilient MSIL variant using 1:1 LAG with LACP. This will support 1G and 10G MSILs.
- Developments for 100M MSILs will depend upon the outcome of the Consult21 process (C21-PP-024).
- BT will require volume forecasts for the use LAG N on MSIL before embarking on development.
- BT will produce recommendations on the process for adding LAG protection to an unprotected MSIL.

4 Next Steps

BT Wholesale will progress the design of the resilient MSIL option as defined by the technical recommendations and with continued input from Industry..

The development of the resilient MSIL option will depend on the low level design, cost and CP volume forecast. Pricing will be devised when the final development costs are known.

The development plan and timescales will be communicated to the Industry via the Consult 21 P2P Work Group.