Demonstration of an application-aware resilience mechanism for dynamic heterogeneous networks

Demonstration of an Application-Aware Resilience Mechanism for Dynamic Heterogeneous Networks

Chigo Okonkwo[1], Richard Martin[1], Marcos Paredes Ferrera[1], Ken Guild[1]
Mike O'Mahony[1] *Members, IEEE*, Jim Everett[2]
[1] *Department of Electronic System Engineering, University of Essex, CO4 3SQ, United Kingdom*
[2] *Ericsson, New Century Park, Coventry CV3 1HJ, United Kingdom*
Email:{cmokon, rmarti, mpared, kguild, mikej}@essex.ac.uk, jim.everett@ericsson.com

ABSTRACT
Business uptake of IP-centric services has been strong and necessitates reliable, highly-available, high-quality Internet access. Real time services such as mission-critical broadcast video, video conferencing and voice over IP (VoIP) have low tolerance for short-term network outages. However, applications like bulk data transfer maybe much more resistant to such events. Many different network resilience mechanisms can be offered to customers by service providers allowing for slower (or faster) recovery of network access. Yet in current networks, offering different resilience mechanisms for different services is complicated, involving multiple interfaces. In this work, we propose that services be offered differentiated resilience levels within a single physical interface. In order to do so, an application aware resilience mechanism is proposed based on using the supplementary IP header type of service (ToS) field to define arbitrary values in addition to the 6-bit differentiated services code points (DSCP). This mechanism allows different levels of resilience to be assigned to applications such as VoIP for emergency and mission critical services.
Keywords: Application-Aware, Differentiated Resilience, Heterogeneous Optical Networks

1. INTRODUCTION

Emerging triple play applications such as Voice over IP (VoIP), broadcast high definition video, video on demand and video conferencing possess differing requirements. For example, while VoIP is very sensitive to small-scale network outages and therefore requires rapid recovery mechanisms, some applications (such as peer-to-peer file transfer) can tolerate slower mechanisms (of the order of several hundred milliseconds). This disparity in requirements of emerging services means that, in order to be effective, different services must be given different resilience mechanisms (levels of resilience).

Currently, in order to achieve different levels of resilience multiple connections with separate SLAs are required. This increases the cost to the customer and service provider because of the multiple interfaces that are occupied by such a setup. Additionally, routing particular services to specific connections (and therefore levels of resilience) is difficult. We propose a system which allows the mapping of services to resilience levels within one physical connection (gigabit Ethernet). This is achieved with a differentiated resilience approach which protects only traffic flows requiring high level of service availability and would allow for the transport of lower grades of service over restored and unprotected network paths. This would result in economical traffic-engineering and better network utilisation.

In current networks, two concepts of resilience are predominant. The main consideration is to provide protection and restoration to whole connections in the network based on the requirements of the client. As a result, all traffic types are treated identically and are fully protected so that the end-user does not perceive the loss of a channel or fibre connection. The other concept is that of multi-layer resilience; this is based on the fact that network elements fail at different points and layers of the network. Therefore, restoration and protection can be performed at different layers. Multi-layer resilience has advanced considerably and work has been done to evaluate the performance of different layers operating simultaneously during failures. WDM protection has been compared against IP restoration in terms of minimal bandwidth provisioning costs at the WDM layer [1]. Mechanisms at different layers were combined in [1] and [2] to improve network utilisation and provide finer recovery granularity at the IP/MPLS layer. However, none of these schemes provide per-application resilience.

Recently, in [3], quality of resilience (QoR) is specified as a QoS factor to enable operators to selectively treat individual services after failure. In [4], the notion of differentiated resilience is implemented using differentiated services architecture (DiffServ), a popular mechanism for supporting QoS in IP networks. During failures DiffServ is triggered because the rate at which packets are taken out of queues is reduced. The authors in [4] use the DiffServ scheduler to provide resilience to the traffic classes according to priority. However, in order for this to happen, a transmission failure must be propagated to the IP layer. This would increase the time in which resilience occurs.

The work reported in this paper was supported by the U.K. Department of Trade and Industry (DTI) under the OSDA project PROTAGON and by the Engineering and Physical Sciences Research Council (EPSRC) funded HIPNET EP/E002382/1 and PLATFORM GR/S82695/01 projects.

In this work, a different approach is taken. The IP header type of service (ToS) field used by the differentiated service architecture has additional bits whose use has not been widely adopted. We propose that application traffic sources may use these bits (denoted level of resilience code points, LoRCP) to specify that their traffic requires a given level of resilience. At the ingress IP router, the packets are mapped to specific IEEE802.1Q virtual LAN (VLAN) tags according to the values set in the LoRCP. This scheme is based on enhanced capabilities to segregate data traffic (based on VLAN identification) in multi-service provisioning platforms (MSPP) deployed in current TDM networks. The segregated traffic is mapped to specific virtually concatenated groups (VCG) which are provided different levels of resilience (LoR). Therefore, if better resilience is required by the client application, the LoRCP bits are changed. This is a more economical and robust mechanism for providing additional service guarantees to specific applications. The aim of this work is to ensure that with few modifications, high availability is assured to mission critical IP-based applications.

2. APPLICATION AWARE RESILIENCE MECHANISM

2.1 Level of resilience code points (LoRCP)

The Type of Service (ToS) octet in the IPv4 header (Figure 1) has evolved over time and is currently used by DiffServ to classify packets into 64 possible classes in IP networks [6]. Classified packets are served (in terms of loss and delay) according to given specific Per-Hop Behaviour (PHB) markings which include the highest priority Expedited Forwarding (EF), Assured Forwarding (AF) and Best Effort (lowest priority). This classification mechanism only uses the first 6 bits of the ToS field. The remaining two bits have been defined in RFC 3168 for the TCP mechanism of Explicit Congestion Notification (ECN) but this is not widely adopted. In this work, these two bits are redefined as the LoRCP bits. Figure 1 illustrates the ToS field (8bits) within the IPv4 packet header including the 2 bits (bit 6 and 7) that is proposed for this scheme.

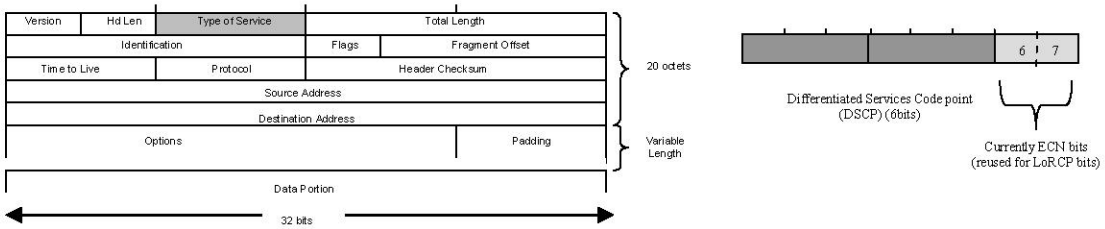

Figure 1. IPv4 packet header and LoRCP bits in ToS byte.

In our approach, the LoRCP bits are set by the client application. Therefore, incoming packets are classified at the ingress IP router according to the values in LoRCP field and mapped to network provider defined VLAN tags. The end user is not aware of the VLAN addressing scheme in the network. In the metro/core network, the VLAN tagged traffic is segregated and mapped to appropriately-resilient paths. The network operator can dynamically choose the level of resilience that is assigned to the particular VLAN tags. Using a user network interface (UNI) allows the IP router to request more bandwidth on specific connections via the GMPLS control plane in order to satisfy pre-defined SLAs. This mechanism is independent of any particular control plane and is therefore easily extensible to different transport technologies.

2.2 Network resilience mechanisms

Table 1 outlines the resilience classes offered by the hybrid electro-optical network and suggests some example mappings from service types into the resilience levels.

Table 1. Resilience Mechanisms under consideration.

Example Services	Level of Resilience (LoR)	Resilience Mechanism(s)	Restoration time	LoRCP bits
Mission Critical VoIP and video conferencing	*3*	*Electrical protection + Optical restoration*	*50 ms + 150 ms*	*xxxxxx11*
Video conferencing	*2*	*Electrical protection*	*50 ms*	*xxxxxx10*
Video on demand	*1*	*Electrical restoration*	*70 to 100 ms*	*xxxxxx01*
Best effort (pre-empted)	*0*	*unprotected*	*N/A*	*xxxxxx00*

From Table 1, there are two main types of resilience mechanism assigned to the classes:

- **Electrical and Optical Protection:** The electrical protection is a dedicated mechanism which is derived from the automatic protection switching (APS) inherent in TDM networks. Under dedicated protection (1+1) the protection path is setup along a physically diverse route. Therefore, failures in the network result in fast switchover to the protection path in approximately 50ms. This is the most expensive option as protection paths cannot be shared by other connections. The failure detection is based on loss of signal (LoS) alarms in the electrical layer. In this paper, we consider only electrical protection.
- **Electrical and Optical Restoration:** Another resilience option is restoration which, upon failure, triggers the path calculation, resource allocation and cross connection by the control plane. Constraints such as RSVP signalling time, path computation and cross connection time results in a restoration time that is more than the protection time. Under this scheme traffic is restored slower (typically around 100-500ms), which is acceptable for many applications. The failure detection is based on loss of signal (LoS) and loss of light (LoL) alarms in the electrical and optical layer respectively

3. EXPERIMENTAL SETUP AND RESULTS

In order to demonstrate the feasibility and performance of the differentiated resilience mechanism, a heterogeneous network testbed (figure 2) is employed. The testbed consists of two Ericsson OMS3240 GMPLS-enabled MSPPs with two ODU-2 (10 Gb/s) and two STM-16 (2.5 Gb/s) line cards. The multi-protocol client interface cards include several layer-2 interfaces capable of mapping the traffic into arbitrary sized virtual concatenated groups (VCG). In the data plane, the link capacity adjustment scheme (LCAS) operates with VCAT on the client cards to ensure fractional services are maintained if any member(s) of the virtual concatenated group (VCG) fails. The optical layer consists of a ring topology with working and protection paths made up of three optical cross connects (OXC) based on Micro Electro-Mechanical Systems (MEMS) switch-technology. The combination of MSPPs and OXCs constitute a possible multi-layer network-provider architecture. The service provider network consists of two Linux IP/MPLS routers with gigabit Ethernet interfaces that are connected to the network provider via the Layer 2 client cards on the MSPPs.

In the experimental setup, the traffic is generated with the LoR bits specified in *Table 1* at the application source. On arrival into the routers, the ToS byte in the application packet header is processed and assigned to the network provider designated VLAN tags. The incoming traffic at the core network is processed and segregated accordingly. The segregated traffic is mapped into arbitrary sized VCG connections which are dynamically provisioned by the GMPLS control plane signalled by the IP routers using the user network interface (UNI).

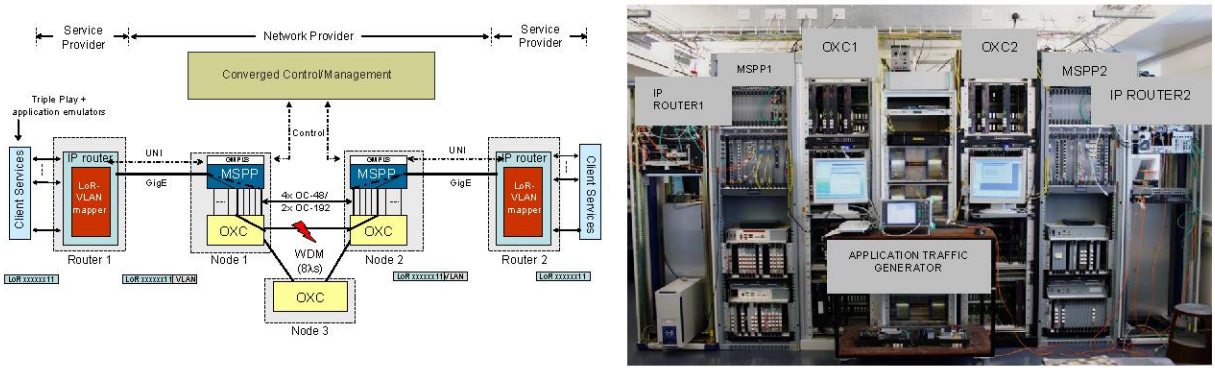

Figure 2. Heterogeneous Network Testbed.

Figure 3 shows the performance of the different levels of resilience in a scenario consisting of two sequential failures:

1. The first failure is an ODU-2 line-card failure in MSPP 1.
2. Following this, there is a fibre-break between OXC-1 and OXC-2.

The failures are measured from the application layer as a reduction in throughput using an Anritsu MD1230B Data Quality Analyser. As observed in Figure 3, the highest levels of resilience (LoR 2 & 3) are protected by the electrical TDM layer in approximately 50 ms and LoR 1 restored by the electrical layer in approximately 75 ms. LoR 0 is unprotected. Following the second failure, only LoR 3 is restored by the provisioning of an alternative path via the OXC 1 and 2 through the optical layer.

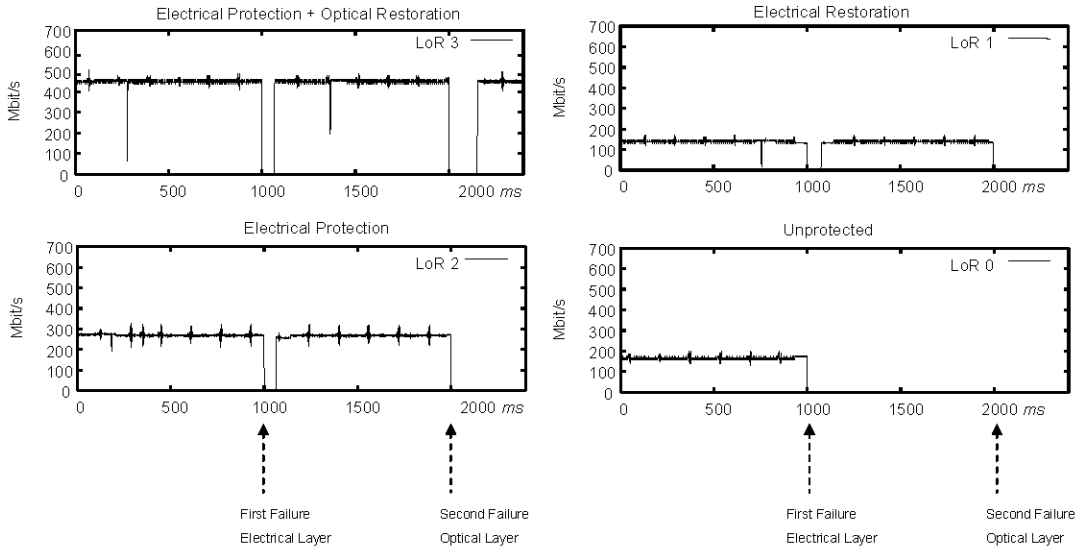

Figure 3. Core network failures results on level of resilience.

4. CONCLUSIONS

In this paper, a novel mechanism for setting the required level of resilience at the application has been proposed. This involves the use of additional code points in the IPv4 packet header ToS byte, this mechanism does not require any additional modifications and allows the mapping to be performed by any IP router utilising the LoR to VLAN tag mapping in order to assign appropriate paths with specific levels of resilience.

Presently, only simple resilience mechanisms are offered for each LoR. However, adding more sophisticated protection mechanisms (such as shared protection) is trivial within this framework, and therefore may be the subject of further research. In the case of shared protection a group of N working lines are protected by a single physically diverse protection path. The underlying assumption is that only one of the working lines will fail at any given time and that the working path is provisioned independently. The advantage of 1:N protection is that it requires much less hardware than the 1+1 dedicated protection scheme. This also ensures that fewer protection paths are dormant. Therefore, during a failure, multiple paths can be simultaneously protected (50 ms) which could result in bandwidth reduction per connection. Hitless mechanisms inherent to client interfaces at the MSPPs such as Link capacity adjustment scheme (LCAS) can be employed in the data plane to automatically provide fractional service to connections on the protection path.

Additionally, UNI can be used in conjunction with LCAS and VCAT to allow dynamic resizing of the various protection levels within the gigabit Ethernet pipe according to their respective demands. This will be a source of future work.

REFERENCES

[1] A. Fumagalli and L. Valcarenghi., IP Restoration vs. WDM Protection: Is There an Optimal Choice?, *IEEE Network,* vol. 14, no. 6, pp. 34-41, Nov 2000.

[2] L. Sahasrabuddhe, S. Ramamurthy and B. Mukherjee., Fault Management in IP-Over-WDM Networks: WDM Protection versus IP Restoration, *IEEE Journal on Selected Areas In Communications (JSAC,)* vol. 20, no.1, Jan. 2002.

[3] P.Cholda, *et al.*, Considerations about Service Differentiation Using a Combined QoS/QoR Approach, in *Proceedings of Design of Reliable Communications Networks (DRCN)*, Naples, Italy, Oct 2005.

[4] B. Sanso and C. Awad., Can DiffServ Guarantee IP QoS Under Failures, *IEEE Journal. Network.,* vol. 8. pp. 32-40, Aug 2006.

[5] S. Dong, C. Phillips and R. Friskey., Differentiated-Resilience Provisioning for the Wavelength-Routed Optical Network, *IEEE Journal of Lightwave Tech.* vol. 24, no.2, Feb 2006.

[6] K. Nichols, *et al.*, Definition of the Differentiated Services Field (DS Field)in the IPv4 and IPv6 Headers, *IETF RFC 2474,* Dec 1998.