Ultra-Reliable NFV-based 5G Networks using Diversity and Network Coding

Nazli Siasi, *Student Member*, *IEEE*, Nabeel I. Sulieman, *Student Member*, *IEEE* and Richard D. Gitlin, *Life Fellow*, *IEEE*
Department of Electrical Engineering
University of South Florida
Tampa, FL, USA
Email: {nazlisiasi, nis}@mail.usf.edu, richgitlin@usf.edu

***Abstract*—This paper presents a robust solution for link and node failures in network function virtualization for 5G and beyond system. Namely, the synergistic combination of Diversity and Network Coding schemes is leveraged here for robust link failure recovery. The scheme offers near-instantaneous packet recovery without feedback requests with enhanced throughput. Hence, it eliminates latency associated with retransmissions and rerouting by using error control across spatially diverse paths.**

***Index Terms*—5G, network function virtualization, Diversity Coding, Network Coding, link failure, link recovery**

I. Introduction

The unprecedented increase in data traffic imposes various challenges for cellular operators for 5G and Beyond Networks. Such as requirement for efficient resource utilization, reliable transmission, lower latencies, reduced power and energy consumption for green communications.

In addition, cellular operators aim to decrease Operation Expenditure (APEX) and Capital Expenditure (CAPAX) [1]. In particular, the high traffic demand and increased connectivity result in magnified costs, i.e., attributed to site-acquisition, densification, and hence increased network equipment installation, and maintenance requirements.

One viable solution here is the adoption of network function virtualization (NFV) technology for cellular systems. This has been manifested by the VNF Task Group launched recently by the European Telecommunications Standards Institute (ETSI) [2]. It aims to virtualize cellular network functionalities, operations, coordination and management.

Here NFV decouples software functionalities from physical equipment, i.e., it deploys virtual network functions (VNFs) as instances running on virtual machines (VMs). Moreover, these VMs are located on VNF-enabled high volume hardware (e.g., servers, switches and routers) distributed at different geographical areas. Therefore, the various VNFs perform the hardware equipment functionalities in the cellular network. Overall, NFV provides flexible and adaptive approach for network functions management. It also accelerates service deployment compared to dedicated hardware installation, in addition, NFV offers operators huge flexibility, scalability, capacity, and resiliency. Nevertheless, it enables network slicing for that enables a single physical network to affiliate various radio access networks (RANs) that are virtually managed.

Despite the merits of NFV in 5G networks, this paradigm imposes significant challenges that need to be addressed prior to its deployment. One major challenge here is link and node failure in the physical layer of the VNF infrastructure (NFVI). This is mainly attributed to the link and node failures, i.e., main constituents to network failures. Hence, effective link/node recovery schemes need to be delivered for NFV-based 5G network reliability. In particular, schemes that feature low latencies, fast recovery times, and minimum resource utilization (e.g., reduced protection links).

Now various methods have been proposed for network reliability, foremost synchronous optical networking (SONET), *p*-cycle ring [4]-[5], and mesh-based link protection [6]. Although the methods here significantly enhance network reliability and robustness, they still suffer from prolonged restoration times that contradict control and data latencies for 5G and beyond.

In light of the above, this paper leverages a robust failure recovery method to address physical link/node failure for ultra-reliable NFV 5G networks, i.e., Diversity Coding-Network Coding (DC-NC) self-recovery scheme [7]. Mainly, DC-NC features near-instantaneous (very low latency) feedforward recovery, and thus eliminates rerouting and retransmission requests. It also provides high throughput and reduces number of redundant facilities that inherent in DC by 30%-40% [7].

This paper is organized as follows. Section II presents the virtual network model, with emphasis on service function chaining. Then VNF link failure is introduced in Section III. This is followed by DC-NC recovery scheme for NFV-based 5G network in Section IV presents, along with conclusions in Section V.

II. Network Function Virtualization Model

As mentioned earlier, this paper adopts VNF to virtualize network functions (NF). In particular, the decomposition of the radio remote head (RRH) and baseband unit (BBU), i.e., BBU functions are performed via VNFs. Namely, evolved virtual packet functions that include mobility management entity (MME), serving gateway (SGW), and packet data network gateway (PGW), and session management and Policy and Charging Rules Function (PCRF) [8].

In light of the above, cellular system and NFV models are next presented, i.e., to perform BBU functions virtually.

Now consider cellular network consists of b=1,2,…,B base stations (BSs), m=1, 2, …,M mobile stations (MSs), and h=1,2,…,H radio remote heads (RRHs). Here RRHs are connected to the NFV via fronthaul links.

Now in order to introduce resource pooling, consider a network graph $G(N, V)$ of N physical nodes and V virtual links.

Here N is composed of I_k intermediate nodes, and N_v nodes that dispose VNFs, i.e., $N=I_k+N_v$, $I_k \subseteq N$, $N_v \subseteq N$. Also, V virtual links are mapped to L physical links, composed of l_p=1,2,...,L_p primary and l_r=1,2,…,L_r protection (redundant) links, i.e., $L=L_p+L_r$, $l_p \subseteq L_p \in L$, $l_r \subseteq L_r \in L$. Furthermore, the physical links L here connect N physical nodes (e.g., servers) that host VNFs, given by:

$$F = \begin{bmatrix} F_{11} & \cdots & F_{1j} \\ \vdots & \ddots & \vdots \\ F_{i1} & \cdots & F_{ij} \end{bmatrix}, \tag{1}$$

where rows present VNFs of same functionally, meanwhile columns present different VNFs. Now $p=p_1, p_2, \ldots, p_d$, data packets (streams) flow on L_p primary links. Meanwhile $c=c_1$, c_2, …,c_r coded packets flow on L_r protection links.

Note that service function chaining (SFC) is adopted here for F. Namely; SFC here determines the paths, by which service function requests (SFR) must take, i.e., path that connects sequence between dependent required F VNFs [3]. In other words, SFC defines logical route between F chained together through virtual links. The priority dependence between F in SFC compels that data packets should flow through functions in an ordered way. Particularly, when SFR such as $F_{11} \rightarrow F_{21}$ exists (e.g., SGW → PGW), the precedence requirements in a SFC means that F_{21} has a priority dependency on F_{11}. Therefore, the traffic flow should pass through F_{11} before reaching F_{21}. Meanwhile, when there is no precedence requirement in a SFC, then packet traffic flow can pass through functions randomly.

III. VNF Link Failure

Now consider a single RRH that receives D SFRs from M MSs in active-mode. These SFRs are then transferred to the virtual network via fronthaul to be processed at the NFV gateway. This paper considers multiple SFRs (D=3), where each SFR consists of source node *src* and destination node *dst*, and set of requested F VNFs, i.e., $src \subseteq N$, $dst \subseteq N$. Mainly, when a particular SFR is received, a designated physical path is constructed based upon the SFC. Overall, RRH performs its functions virtually via the VNF that run on VMs on the physical layer. Here, this layer is the major constituent of network failures in NFV. This can be attributed to various factors, such as intentional attacks, excessive interference, noisy environment, software malfunction, false routing, and random failures. In spite of that, intermediate links and nodes become vulnerable to failures. Hence, this leads to disconnection between *src* and *dst* access nodes.

Therefore, work here leverages DC-NC to address aforementioned failure vulnerability in NFV physical layers. This scheme exhibits ultra-reliable recovery at near-instantaneous restoration times. In addition to the robust reliability here, DC-NC also features network throughput.

IV. DC-NC Recovery for NFV-based 5G Network

Diversity coding [9] is an open loop coding technique that utilizes a feedforward error control mechanism across disjoint paths at the expense of redundant transmission facilities. Once the failure is detected by deploying this technique, the lost data can be retrieved instantaneously (once the faults have been identified) without requiring any rerouting or retransmission. This saliency further magnified by the enhanced throughput, attributed to NC [10] enhances throughput problem by enhancing the network throughput. The key idea of NC is combining packets by coding them at network nodes and transmitting the result to *dst* through different paths.

Hence, a synergistic combination of Diversity Coding and Network Coding (DC-NC) [7] is applied here for NFV as shown in Fig. 1. In turn, this simultaneously improves network reliability with very low latency in network recovery (restoration), along with high throughput for broadcasting applications. Here, the nodes are connected to each other in a mesh topology by bi-directional fiber optic cables.

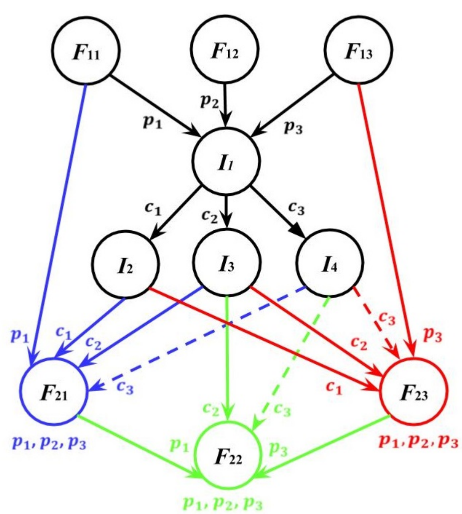

Fig. 1. DC-NC coding scheme for NFV-based 5G Network

Figure 1 illustrates DC-NC application for NFV-based 5G networks. Now three packets p_1, p_2, and p_3 are transmitted on the primary links, along with three coded packets c_1, c_2, and c_3 transmitted on the protection links. Data packets here are initiated from source nodes hosting F_{11}, F_{12}, F_{13} directly transmitted to the destination nodes F_{21}, F_{22}, and F_{23}, or indirectly through intermediate nodes. In this scenario, F_{11} and F_{13} use their direct links with F_{21} and F_{23} to transmit p_1, and p_3 respectively. In addition, the DC-NC encoding node I1 receives p_1, p_2 and p_3 from F_{11}, F_{12}, and F_{13} respectively and encodes them to produce three coded packets c_1, c_2 and c_3 as follows:

$$c_1=\beta_{11}p_1+\beta_{21}p_2+\beta_{31}p_3, \tag{2}$$
$$c_2=\beta_{12}p_1+\beta_{22}p_2+\beta_{32}p_3, \tag{3}$$
$$c_3=\beta_{13}p_1+\beta_{23}p_2+\beta_{33}p_3, \tag{4}$$

where $\begin{bmatrix}\beta_{11} & \beta_{21} & \beta_{31}\\ \beta_{12} & \beta_{22} & \beta_{32}\\ \beta_{13} & \beta_{23} & \beta_{33}\end{bmatrix}$ is the parity generator matrix for c_1, c_2 and c_3. The parameters in this matrix are linearly independent to each other to achieve matrix inversion. Since the multiplication and addition are performed in $\mathrm{GF}(2^{\mathrm{m}})$, they correspond to "AND" and "XOR" operations, respectively. Coded packets c_1, c_2 and c_3 will be sent to I_2, I_3, and I_4 respectively. The throughput gains will be obtained by sending c_1 through I_2 to F_{21} and F_{23} respectively to and c_2 through I_3 to all destinations. F_{22} will receive p_1 and p_3 from F_{21} and F_{23} respectively. F_{21} will decode p_1, c_1 and c_2 and obtain p_2 and p_3 as follows:

$$\tilde{c}_1 = c_1 + \beta_{11}\, p_1, \tag{5}$$
$$\tilde{c}_2 = c_2 + \beta_{12}\, p_1. \tag{6}$$

Now by applying equations (2)-(3) to (5)-(6) respectively,

$$\tilde{c}_1 = \beta_{11}p_1 + \beta_{21}p_2 + \beta_{31}p_3 + \beta_{11}p_1 = \beta_{21}p_2 + \beta_{31}p_3, \tag{7}$$
$$\tilde{c}_2 = \beta_{12}p_1 + \beta_{22}p_2 + \beta_{32}p_3 + \beta_{12}p_1 = \beta_{22}p_2 + \beta_{32}p_3. \tag{8}$$

Equations (7) and (8) can be expressed in a matrix form as follow as:

$$\begin{bmatrix}\tilde{c}_1\\ \tilde{c}_2\end{bmatrix} = \begin{bmatrix}\beta_{21} & \beta_{31}\\ \beta_{22} & \beta_{32}\end{bmatrix}\begin{bmatrix}p_2\\ p_3\end{bmatrix}. \tag{9}$$

By simple matrix inversion p_2 and p_3 can be easily obtained by:

$$\begin{bmatrix}p_2\\ p_3\end{bmatrix} = \begin{bmatrix}\beta_{21} & \beta_{31}\\ \beta_{22} & \beta_{32}\end{bmatrix}^{-1}\begin{bmatrix}\tilde{c}_1\\ \tilde{c}_2\end{bmatrix}. \tag{10}$$

Note that the coefficients β_{ij} are fixed and known at all nodes. The destination node F_{23} will decode p_3, c_1 and c_2 in similar way to get p_1 and p_2. In this way, the throughput gains will be improved by at least 20% [7].

The reliability of the network can be enhanced by sending c_3 from I_4 to all destinations. In case of none link failure existence, destinations will ignore c_3. However, when there is a link failure, for example if the link between F_{11} and F_{21} is failed and once F_{21} detects the failure, it uses c_1, c_2, and c_3 to recover all packets p_1, p_2, and p_3 as follows:

$$\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix} = \begin{bmatrix}\beta_{11} & \beta_{21} & \beta_{31}\\ \beta_{12} & \beta_{22} & \beta_{32}\\ \beta_{13} & \beta_{23} & \beta_{33}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix}. \tag{11}$$

The required packets are obtained by performing simple matrix inversion

$$\begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} = \begin{bmatrix}\beta_{11} & \beta_{21} & \beta_{31}\\ \beta_{12} & \beta_{22} & \beta_{32}\\ \beta_{13} & \beta_{23} & \beta_{33}\end{bmatrix}^{-1}\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix}. \tag{12}$$

In addition, if c_1 or c_2 are lost through link failure at F_{21}, still p_1, c_3 and (c_2 or c_1) are available at F_{21} to easily and quickly recover the required packets in the same way as explained in equations (5)-(10). Similarly, F_{23} can recover the lost packets. However, since F_{22} does not have any access to c_1, therefore if p_1 is lost, then it can easily and near-instantly be recovered with p_2 as follows:

$$\tilde{c}_1 = c_1 + \beta_{31}\, p_3, \tag{13}$$
$$\tilde{c}_2 = c_2 + \beta_{32}\, p_3, \tag{14}$$

and applying (2)-(3) to (13)-(14) respectively,

$$\tilde{c}_1 = \beta_{11}p_1 + \beta_{21}p_2 + \beta_{31}p_3 + \beta_{31}p_3 = \beta_{11}p_1 + \beta_{21}p_2, \tag{15}$$
$$\tilde{c}_2 = \beta_{12}p_1 + \beta_{22}p_2 + \beta_{32}p_3 + \beta_{32}p_3 = \beta_{12}p_1 + \beta_{22}p_2. \tag{16}$$

Equations (15) and (16) can be expressed in a matrix form as following:

$$\begin{bmatrix}\tilde{c}_1\\ \tilde{c}_2\end{bmatrix} = \begin{bmatrix}\beta_{11} & \beta_{21}\\ \beta_{12} & \beta_{22}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\end{bmatrix}. \tag{17}$$

By simple matrix inversion, p_1 and p_2 are easily obtained as:

$$\begin{bmatrix}p_1\\ p_2\end{bmatrix} = \begin{bmatrix}\beta_{11} & \beta_{21}\\ \beta_{12} & \beta_{22}\end{bmatrix}^{-1}\begin{bmatrix}\tilde{c}_1\\ \tilde{c}_2\end{bmatrix}. \tag{18}$$

Also, if c_2 fails, F_{22} still have access to p_1,p_3 and c_3 and it can quickly and easily recover p_2 as follows:

$$\tilde{c}_3 = c_3 + \beta_{13}\, p_1 + \beta_{33}\, p_3, \tag{19}$$

and applying equation (4) to (19) result in

$$\tilde{c}_3 = \beta_{13}p_1 + \beta_{23}p_2 + \beta_{33}p_3 + \beta_{13}\, p_1 + \beta_{33}\, p_3, \tag{20}$$
$$p_2 = \tilde{c}_3/\beta_{23}. \tag{21}$$

Hence, one link failure at each destination node can be recovered by DC-NC coding scheme. Consequently, three simultaneous link failures can be tolerated, one for each destination node. Furthermore, except DC-NC encoding node (I_1), one intermediate node failure is also tolerable. For example, if I_2 fails, F_{21} and F_{23} will loss c_1, which can quickly and easily be recovered in similar way that is shown in (5) - (10).

Here only four protection links are used to improve the reliability of the network, i.e., by protecting it from single link failure at each destination node. To further enhance the network reliability and protect it from dual-link failures at each destination node, four more protection links are required. Generally, L_f link failures can be tolerated for each destination node at the expense of EL_f+L_f protection links, where E is the number of destination nodes, i.e., $L_f \leq$ number of broadcasted packets over primary links.

V. Conclusions

The paper addresses the problem of link and node failure in NVF-based 5G and beyond networks. Namely, Diversity and Network Coding (DC-NC) are exploited to achieve ultra-reliable failure recovery. Namely, recovery is achieved here in near-instantaneous restoration times, without requirement for feedbacks and re-transmissions. Hence, this significantly decreases latencies for future cellular networks. Future efforts will investigate the efficiency of DC-NC for multiple failure occurrences at each destination, addressing increase of protection links.

Acknowledgment

Nabeel I. Sulieman is supported by the Higher Committee for Education Development in Iraq (HCED-IRAQ).

References

[1] E. Hernandez-Valencia, S. Izzo, and B. Polonsky, “How Will NFV/SDN Transform Service Provider OPEX?” *IEEE Network*, vol. 29, no. 3, 2015, pp. 60–67.

[2] R. Guerzoni, “Network functions virtualization: an introduction, benefits enablers, challenges and call for action, introductory white paper,” in *Proc. of SDN and OpenFlow*, World Congress,

2012.

[3] J. Halpern et al., "Service Function Chaining (SFC) Architecture", Internet Requests for Comments RFC Editor RFC 7665, October 2015.

[4] S. Avci and E. Ayanoglu, "Recovery from link failures in networks with arbitrary topology via Diversity Coding," in *Proc. IEEE Globecom*, Houston, TX, Dec. 2011, pp. 1-6.

[5] J.-P. Vasseur, M. Pickavet, and P. Demeester, Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS. Elsevier, 2004.

[6] C.-W. Chao, P. M. Dollard, J. E. Weythman, L. T. Nguyen, and H. Eslambolchi, "FASTAR - A robust system for fast DS3 restoration," in *Proc. IEEE Global Communications Conference*, December 1991, pp. 1396–1400.

[7] N. I. Sulieman, E. Balevi, K. Davaslioglu, and R. D. Gitlin, "Diversity and Network coded 5G fronthaul wireless networks for ultra-reliable and low latency communications," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Montreal, QC, Canada, Oct. 2017.

[8] A. Checko et al., "Cloud RAN for Mobile Networks, A Technology Overview," IEEE Commun. Surveys & Tutorials, vol. 17, no. 1, 2014, pp. 405–26.

[9] E. Ayanoglu, C.-L. I, R. D. Gitlin, and J. E. Mazo, "Diversity Coding for transparent self-healing and fault-tolerant communication networks," *IEEE Trans. Commun*., vol. 41, pp. 1677–1686, Nov. 1993.

[10] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.