RESEARCH ARTICLE OPEN ACCESS

High Availability and Faster Convergence Techniques in IP Networks

Er. Jashandeep Kaur [1], Er. Rupinder Kaur Gurm [1]
Research Scholar (M.Tech) [1], Assistant Professor [2]
Department of Computer Science and Engineering
RIMT-IET Mandi Gobindgarh
Punjab - India

ABSTRACT
This paper explains the various High Availability and Faster Convergence techniques that helps today's fast paced networks getting things done and converge within milliseconds from primary to secondary link in case of primary link failure. It expla ins techniques used in real world networks to achieve high available networks. Various techniques used and the designs in which they can be used are explained in these paper. High availability and Faster Converegence are a need for almost all the industries like Banking, Retail, Health, manufacturing, Gaming etc and network is always critical for all the industries and how to achieve this fast convergence and how to make the network always highly available is described in this paper.
Keywords **:-** Fast Reroute, Fast Hello Detection, LSA Pacing , IP Event Dampening, SPF Timers

I. INTRODUCTION

Networking is the soul of IT industry. We cannot imagine the world without Internet in today's world. Businesses are moving towards e-commerce solutions. All the data is shifting to the clouds which are nothing but lots and lots of servers at the data centers. Our calling and video solutions in industry and in our daily lives is going towards the new IP age. All these above applications have a common need to be successful - **High Availability and Faster Convergence**. While designing networks, whether we are designing "Enterprise Networks", "Service Provider Networks", "Data Center Networks", one goal which is the most important in every design is "High Availability and Faster Convergence". Convergence can be measured with the following formula:

Failure Detection + Event Propagation + Routing Process + FIB Update

High Availability and Faster Convergence both work together in a way that faster the convergence higher the availability. A great example that one can take is of an e-commerce company like Amazon.com, if Amazon.com becomes unreachable for 5 minutes from his customers, how much bad impact (financial and reputation) does it make. Sub-Second convergence is what is needed when you are using VoIP in the network as VoIP uses User Datagram Protocol(UDP) for transporting Voice and Video traffic and delay can end the Voice or Video connection instantly.

On the other hand, High Availability is measured using the following formula:

Availability = (MTBF-MTTR)/MTBF

Where **MTBF** is mean time between failure means "What, when, why and how does it fail?"and **MTTR** is mean time to repair means "How long does it take to fix ?"

	Availability	DPM	Downtime Per Year (24x365)		
Reactive?	99.000%	10000	3 Days	15 Hours	36 Minutes
	99.500%	5000	1 Day	19 Hours	48 Minutes
Proactive?	99.900%	1000		8 Hours	46 Minutes
	99.950%	500		4 Hours	23 Minutes
Predictive?	99.990%	100			53 Minutes
	99.999%	10			5 Minutes
	99.9999%	1			30 Seconds

Figure 1.1 - High Availability Measurement Table

Above high availability measurement table shows availability of networks in terms of percentage and downtime per year and in today's network era, 99.999% and 99.9999% are termed as highly available networks.

An highly available or predictive network needs to have :

- There should not be single points of failures.
- Fault, performance and workflow process tools.
- Excellent consistency is needed with Hardware, Software, Configuration and design.
- Consistent processes for fault, security and performance.

1.2 IP Event Dampening **-** It prevents routing protocol churn which is caused by constant interface state changes taking lots of CPU resources. It is supported by static routing and all the dynamic routing protocols like RIP, EIGRP, OSPF, IS-IS and BGP. It also supports HSRP. It cannot be applied on sub-interfaces and can work on physical interfaces only. IP event dampening has taken the idea from route-flap dampening feature of BGP. When applied on interfaces, it tracks interface flapping or state change, and applies penalty to the flapping interface and if the penalty reaches threshold tolerance , then it put the interface in down state and if penalty is decreased below threshold level, then it brings interface to UP state again. It is kind of a exponential backoff algorithm, therefore it actually deals with events, and IP dampening adds some value to it based on the type of event and also the time difference or frequency at which the event is occuring. Below figure shows a good example of IP dampening :

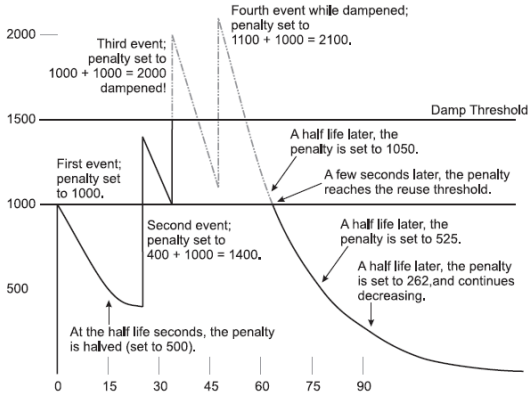

Figure 1.2 - IP Event Dampening Example

The reporting of an event is totally based on penalty, with the higher the penalty that is applied to a given interface or route, less desirable it gets. Penalty is always added when the interface or item goes up or down. Figure above shows the dampening example with time starts at 0 and also penalty of 0. After the very first event, a penalty of 1000 is added . When no other event occurs, then penalty is decreased, which is totally based on the half life. After every passing of 15 seconds, the penalty is halved, therefore we can say that after 15 seconds if the current penalty is 1000, then it is halved to 500.

In the above figure, as the penalty is decreasing after the first event, the second event occurs and another 100 is added to the current penalty making the total penalty 1400. Now with the time passes, the penalty decreases exponentially, that will make it reach 1000 before the occurence of third event. After the third event, 1000 is again added to the penalty making it reach 2000, which is above the damp threshold. so the future events are dampened and it leaves the route and interface in the down state.

As the time passes, the penalty is cut to half for each 15 seconds which is the half-life, which will make it reach 1100 before the occurence of fourth event. After the fourth event, 1000 is again added to the current penalty making it 2100 which again makes the interface in dampening state until the penalty is reduced to the resuse threshold level again. After around 60 seconds, the penalty finally drops below 1000, which is the reuse threshold level. Now the interface is again being tracked till the penalty reaches the dampen threshold in the future. Following are the values that we define with the IP event dampening feature :

- reuse-threshold
- suppress-threshold
- max-suppress-time
- half-life-period
- restart-penalty

1.3 OSPF Fast Hello Detection, LSA Group Pacing and Tuning SPF Timers :

We can detect a neighbor failure or link failure using Polling interval method by polling through fast hellos which are transmitted at Layer 2 and Layer 3. We can fasten the hello packet interval time in milliseconds, for example ospf can transmit 5 hello packets every 1 second and the dead timer is set to 1 second. If we need to configure 200 millisecond hellos with OSPF, then we can set 5 hello packets in a dead interval of 1 using the following command in Cisco IOS software : ip ospf dead-interval minimal hello-multiplier [multiplier], where multiplier is the number of hello packets that we need to send in 1 second.

OSPF produce large bursts of LSA Flooding traffic every 30 minutes, while individual aging do fragmentary re-flooding. LSA group pacing feature would help in controlled bursting. For example if we change the group pacing timer to 10 minutes, a small batch of LSAs that are close to be aged-out are processed together. A figure below shows LSA group pacing effect :

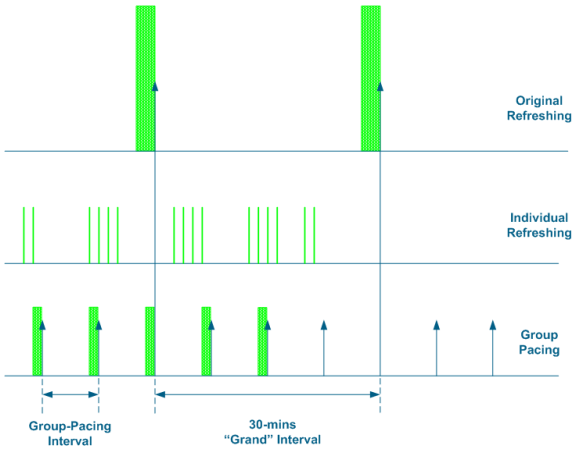

Figure 1.3 LSA Group Pacing in OSPF

Tuning SPF Timers comes under OSPF Exponential Backoff for the generation of Link State Advertisements. It is also known as LSA throttling. It includes three attributes :

- **spf-start** - It is the initial Shortest Path First schedule delay in milliseconds.
- **spf-hold**- It is the minimum hold time between two consecutive SPF calculations.
- **spf-max-wait** - It is the maximum wait time between two consecutive SPF calculations.

1.4 MPLS Fast Re-Route

Multiprotocol Label Switching is used in almost all the service providers in the world. It is the technology that is the heart of Internet Service Provider core networks. From one Provider edge to other provider edge, mostly there are two or more than two paths. For faster convergence and for link protection in case of link failure, MPLS Fast Reroute is used. MPLS FRR provides protection against link or node failures. The FRR mechanism provides sub-second convergence by having backup path pre-calculated which is used in case of primary link failure. It allows data flow to continue even when the headend router tries to create a new end-to-end Label Switch Path that is used to bypass the failure. Notification of primary link failure to headend router is sent by Interior Gateway Routing Protocol like OSPF and ISIS and through RSVP. Below figure shows the MPLS FRR process :

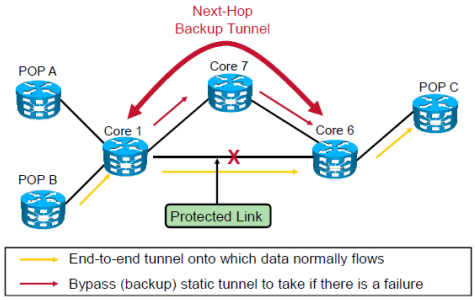

Figure 1.4 - MPLS FRR Sample Process

In the above MPLS FRR figure, it is shown that when primary link between Core 1 router of ISP and Core 6 goes down, then the backup link comes up immediately. Backup link come to use immediately within a second helps in almost zero loss of data. LSP paths are calculated at the Headend. Backup tunnels that bypasses next-hop nodes for LSP paths are known as next-next-hop backup tunnels and the reason is that they terminate at the node which is following the next-hop-node of the LSP path, therefore it bypasses the next-hop-node. Protection from failure of nodes is also made sure with MPLS FRR. Therefore if a node along the LSP path goes down, then the LSP is created over the backup with less than a second, and the convergence is very fast. Figure below shows the node failure and MPLS FRR.

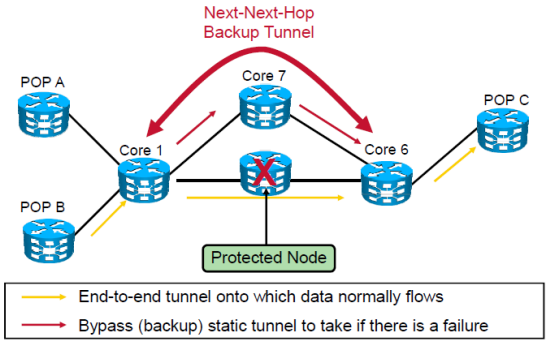

Figure 1.5 - MPLS Core Node Failure and how MPLS FRR does a sub second convergence

II. LITERATURE SURVEY

MPLS Traffic Engineering – Fast Reroute [1] by Shuguftha Naveed, S. Vinay Kumar of Vasavi College of

Engineering(Osmania University), Hyderabad in May, 2014 under IJSR – ISSN: 2319-7064 draws a conclusion that in the event of link failure, traditional recovery technologies takes unacceptable time in case of VoIP and Video based critical solutions, while MPLS traffic engineering Fast Reroute meets the requirements of real-time applications with fast recovery that facilitate high availability to converge. The research shows that MPLS Fast Reroute method provides great performance in case of link failure as compared with traditional IP networks..

Fast Reroute Extensions to RSVP-TE for LSP Tunnels [3] by P. Pan, Ed. Of Hammerhead Systems, G. Swallow, Ed. Of Cisco Systems and A. Atlas, Ed. Of Avici Systems in IETF RFC 4090 defines RSVP-TE extensions to establish backup label-switched path(LSP) tunnels for local repair of LSP tunnels. These mechanisms enable the re-direction of traffic onto backup LSP tunnels in 10s of milliseconds, in th event of failure.

Survey on the RIP, OSPF, EIGRP Routing Protocols [4] by V. Vetriselvan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1058-1065, specifies a performance evaluation of various routing protocols with certain criteria's like Jitter, Convergence Time, end to end delay.

Bidirectional Forwarding Detection (BFD) [5] by D. Katz and D. Ward of Juniper Networks in IETF RFC 5880 describes a protocol intended to detect faults in the bidirectional path between two forwarding engines, including interfaces, data link(s), and to the extent possible the forwarding engines themselves, with potentially very low latency. It operates independently of media, data protocols, and routing protocols.

Bidirectional Forwarding Detection (BFD) on Link Aggregation Group (LAG) Interfaces [6]by M. Bhatia of Alcatel-Lucent, M. Chen of Huawei Technologies, S. Boutros, M. Binderberger of Cisco Systems. J. Haas of Juniper Networks in IETF RFC 7130 defines a mechanism to run Bidirectional Forwarding Detection(BFD) on Link Aggregation Group(LAG) interfaces.

Graceful OSPF Restart by J.Moy of Symacore Networks, P. Pillay-Esnault of Juniper Networks and A . Lindem of Redback Networks in IETF RFC 3623 [7]describes where an OSPF router can stay on the forwarding path even as its OSPF software is restarted. This process is called "graceful restart" or "non-stop forwarding". In this paper, operation of the restarting router, Graceful restart advantages in unplanned outages, its format is defined.

OSPFv3 Graceful Restart by P. Pillay-Esnault of Cisco Systems and A. Lindem of Redback Networks in IETF RFC 5187[10] describes the OSPFv3 graceful restart mechanism. It is pretty much identical to OSPFv2. There are very few differences which are specified in the document. This includes the format of the grace Link State Advertisements(LSAs).

Basic Specification of IP Fast Reroute for IP Fast Reroute: Loop-Free Alternatives by A. Atlas, Ed. Of British Telecom and A. Zinin, Ed. Of Alcatel-Lucent in IETF RFC 5286[11] describes the use of loop-free alternatives to provide the local protection for unicast traffic in pure IP and MPLS networks if a failure on a link occurs. The objective of MPLS Fast Reroute is to reduce the loss of packets that happens while the routers converge after the primary link failure. It has a rapid failure repair with the pre-calculated backup next-hops towards the destination.

Fast Reroute Extensions to RSVP-TE for LSP Tunnels by P. Pan, Ed. of Hammerhead Systems, G. Swallow, Ed. of Cisco Systems, A. Atlas, Ed. of Avici Systems in IETF RFC 4090[12] in May 2005 defines two methods in this paper. First is a one-to-one backup method that creates detour LSPs for each protected LSP at each potential point of local repair. Other one is a facility backup method that creates a bypass tunnel to protect a potential failure point by using MPLS stacking, this bypass tunnel can protect the set of LSPs that have similar backup constraints. Both these methods are used in case of link or node failures.

III. PROBLEM DEFINITION

- High Availability in Networks is one of the major goals of every company. Large Downtime of network in a company creates big loss in companies(data centers, ISPs, Enterprises, E-commerce Companies). Various High availability protocols can be used, but all of them are used according to a specific network design. A specific set of protocols are used for different layers, all of them have different requirements, both economically and infrastructure wise.

- Same is the case with Faster convergence as if not properly implemented, the results can be severe.

- So designing a highly available and faster converging network is a very difficult task with set of different protocols trying to achieve the same.
- As world is diving towards VoIP and Video solutions, that needs high bandwidth and low delay, faster convergence has become much more important.

IV. OBJECTIVE

Comparative analysis of High Availability technologies and Faster Convergence Technologies MPLS Fast Reroute, Tuning SPF Timers, LSA Pacing Timers, OSPF Fast Hello Mechanism, Dampening will be done.

Selection of best High Availability and Faster Convergence method in Medium to Large Service Providers.

Selection of best High availability and faster convergence methods that generates minimum delay for VoIP based networks

Tools that will be used are Graphic Network Simulator (GNS3), Wireshark Packet Analyzer and Cisco 2821, 1841 series routers will be used .

V. RESULTS AND DISCUSSIONS

5.1 LSA Throttling/Tuning SPF Timers -

SPF is the algorithm used in Link State routing protocols, Service Provider Networks uses only Link State routing protocols for their IGP routing, so for faster convergence between Core of Service Provider, we can tune the SPF timers , in order to converge at much better speed. Three timers that are tuned with SPF tuning are spf-start, which is the initial SPF schedule delay in milliseconds, spf-hold, which is the minimum hold-time between two successive SPF calculations, spf-max-wait, which is the maximum wait time between two SPF calculations. I have used a MPLS topology of service provider to test SPF tuning in Service Provider, Topology is shown below :

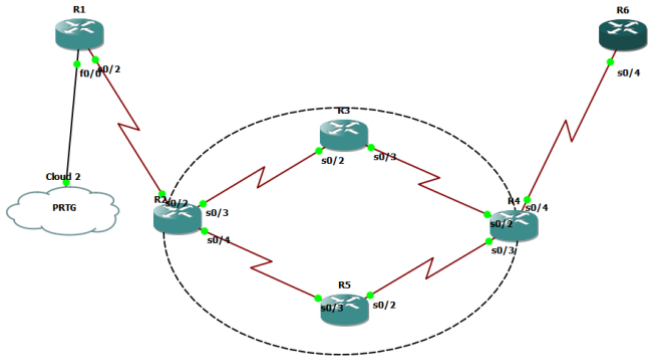

Figure 5.1 - MPLS Topology used in testing

In the above topology, we are sending data traffic from Cloud(PRTG) to R6, traffic is using the path Cloud-R1-R2-R5-R4-R6 as the primary path, when link between R2 and R5 goes down, the convergence time between default parameters is shown below:

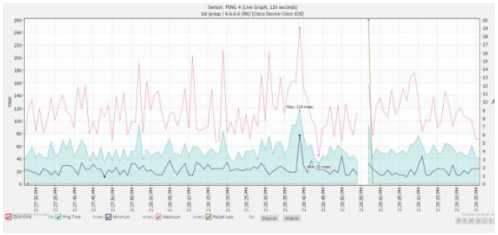

Figure 5.2 - MPLS Convergence time with default parameters.

After tuning SPF timers, the convergence time is shown below :

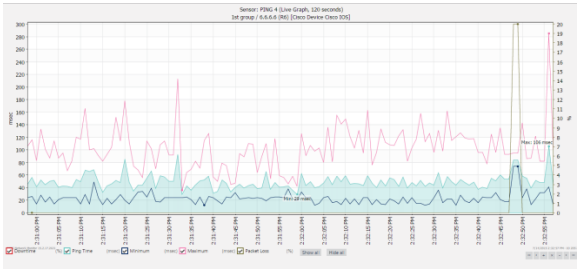

Figure 5.3 - Convergence Time after SPF tuning

After SPF tuning convergence time drops down to milliseconds, which was around 3-4 seconds with default parameters.

5.2 MPLS Fast Reroute

MPLS Fast reroute is a technology that we have used for faster convergence or LSP protection by having a backup link precalculated in case of primary link or a core node/router failure in MPLS Backbone network. Topology used for testing is given below :

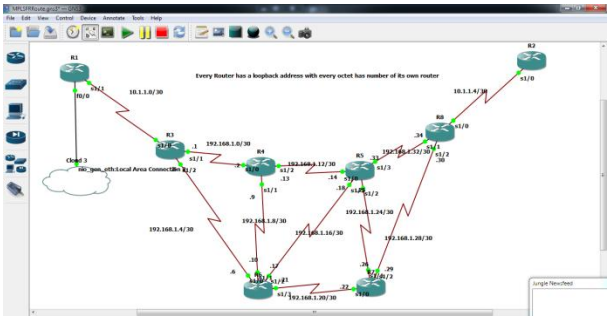

Figure 5.4 – MPLS Fast Reroute Topology

In the above topology, R1 and R2 act as Customer Edge routers or CE devices, R3 and R8 acts as Provider Edge devices while R4,R5,R6,R7 acts as Provider or Core routers. Network of 192.168.1.x/30 is used between the ISP backbone, while customer uses 10.1.1.x/30 network between customer edge and provider edge router. Every router has a loopback address created on it with the same number used in all 4 octets as the router number. For example, if i am using R4, then a loopback of 4.4.4.4 is created on it. With MPLS FRR, PE-PE primary LSP has been hardcoded with R3-R6-R5-R8 with addresses 192.168.1.6 - 192.168.1.18 - 192.168.1.34 - 8.8.8.8, and a backup path is left dynamic, and it automatically pre-calculate the backup path from PE-PE LSP. Next Figure shows the Primary path selection.

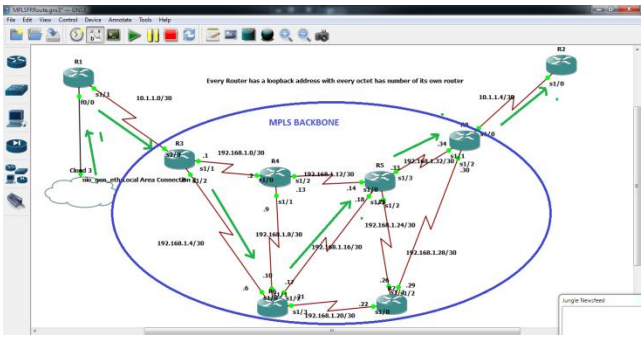

Figure 5.5 – MPLS FRR Explicit Path-1

Below is the configuration that is done for explicit path used as primary LSP path with Fast RR.

```
ip explicit-path name ABC enable
 next-address 192.168.1.6
 next-address 192.168.1.18
 next-address 192.168.1.34
 next-address 8.8.8.8
```

Figure 5.6- MPLS FRR Explicit Route Configuration

When a traceroute command is issued on a PC connected with Customer Edge Device 1 towards Customer Edge Device 2, then the result shows that the explicit path is in use :

```
Tracing route to 2.2.2.2 over a maximum of 30 hops

  1    86 ms     9 ms    92 ms  10.1.1.10
  2   113 ms     *      219 ms  10.1.1.2
  3   119 ms    97 ms    79 ms  192.168.1.6
  4   116 ms    88 ms    70 ms  192.168.1.18
  5   293 ms   246 ms   190 ms  10.1.1.5
  6   262 ms   211 ms   155 ms  2.2.2.2

Trace complete.
```

Figure 5.7 - Traffic from Source to destination via Explicit Path

Figure below shows the output that path option 1 is explicit that i have entered and path option 2 is dynamic, tunnel destination 8.8.8.8 and path is valid. It shows the PE router's outbound interface address as 192.168.1.5 and the next hop as 192.168.1.6 with all the router's address that will come under explicit route.

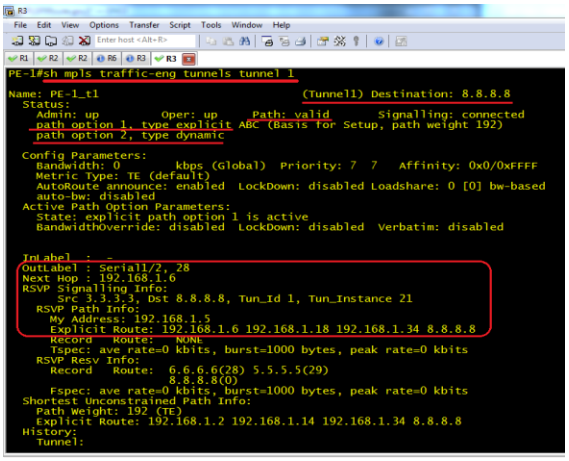

Figure 5.8– Working Traffic Engineering Tunnel Confirmation

When i break the core link between R6 and R5 to break the explicit path, then the traffic shifts to other dynamic link in a quick amount of time and only a single packet is lost with MPLS FRR which is quite good for large MPLS

backbone networks. Below is the convergence result output shown in a screenshot :

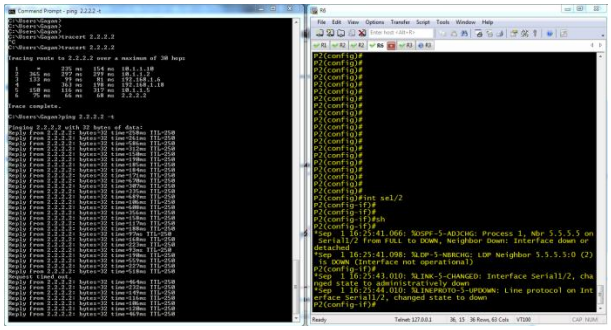

Figure 5.9– Faster Convergence Output with MPLS FRR

5.3 IP Event Dampening

I have used IP event dampening on interface level, with the following configuration:

```
R1(config)#interface fastEthernet 0/0
R1(config-if)#damp
R1(config-if)#dampening ?
  <1-30>  Half-life time for the penalty
  <cr>

R1(config-if)#dampening 15 ?
  <1-20000>  Value to start reusing an interface
  <cr>

R1(config-if)#dampening 15 1000 ?
  <1-20000>  Value to start suppressing an interface

R1(config-if)#dampening 15 1000 2000 ?
  <1-255>  Maximum duration to suppress an interface

R1(config-if)#dampening 15 1000 2000 60 ?
  restart  Enable restart penalty
  <cr>

R1(config-if)#dampening 15 1000 2000 60 re
R1(config-if)#dampening 15 1000 2000 60 restart ?
  <1-20000>  Penalty applied at restart
  <cr>

R1(config-if)#dampening 15 1000 2000 60 restart 2000
R1(config-if)#exit
```

Figure 5.10 - IP event dampening configuration.

I have configured suppress penalty to 2000 and reuse threshold to 1000. Below is the output after 4 penalties and flapping of routes four times.

```
R1#sh interfaces dampening
FastEthernet0/0
  Flaps Penalty    Supp ReuseTm   HalfL  ReuseV   SuppV  MaxSTm    MaxP Restart
     4     542   FALSE        0      15    1000    2000      60   16000    2000
```

Figure 5.11- Penalty after 4 times link flaps.

5.4 OSPF Fast Hello Mechanism

As stated before OSPF fast hello mechanism is used for faster convergence, as it sends multiple hello packets within a single packet and detects a failure of neighbor within 1 second by sending bulk of single packets in a single second as shown below :

```
R2(config-if)#ip ospf dead-interval minimal hello-multiplier ?
  <3-20>  Number of Hellos sent within 1 second
```

Figure 5.12 - ospf fast hello example config

Above figure shows the example configuration according to which we can send from 3 to 20 hello packets within a single second. We have selected 5 hello packets in a single second. Example is given below :

```
R2(config-if)#ip ospf dead-interval minimal hello-multiplier 5
```

Figure 5.13- 5 hello packets are configured to be sent within 1 second

VI. CONCLUSION & FUTURE SCOPE

High Availability and Faster Convergence are the two important factors that are taken into account while designing networks. There are various techniques that can be used to improve availability and convergence time of IP based networks according to the requirements of services deployed in the network. High Availability and Faster Convergence techniques are a need for a network to perform in best manner with predictive nature. All the convergence techniques are necessity with today's voip and video based networks and networks which are needed to be highly available and a faster converged network. In high availability, .Tuning SPF algorithm timers can make Link State Routing protocol works faster than default. LSA Pacing Timers can also make OSPF behaves faster. Fast Hello Detection is like BFD, but it sends OSPF hello packets instead of small size packets of BFD. IP event dampening is used in case of interface or route flapping and in networks where flapping can cause damage to the network and reachability. MPLS gives faster convergence by fastening Link state routing protocols. MPLS Fast Reroute mechanism is a used in Service Provider Networks where if core router or links goes down, then Label Switch path will have sub second convergence, which will be must for ISP as customer traffic is going over ISP and Voice and Video related traffic can easily converge in case of primary link failure. Every faster convergence and high availability technique is used for different network designs and are a must in today's networks where data, voice and video traffic is increasing and more and more organizations are connecting with their networks with their businesses to increase their profit and downtime of networks is unbearable for their strategies.

Network Speeds are increasing with time, over the last 15 years, network speeds have increased over 18 million times and with these increasing speeds faster convergence is the need of time. VoIP and Video traffic needs faster convergence and a constant speed in order to smoothly run its applications. Also e-commerce applications needs high availability in order to have their maximum reach on Internet as all their customers are coming to them over internet. Therefore faster convergence and high availability technologies are always in demand and will be a permanent need of the future and with Internet of Things is in implementation phase all around the world, business, health, manufacturing, retail etc industries will see the network boom in next 10 years with everything gets connected with Internet and for sub-second convergence and highly available networks become necessity.

REFERENCES

[1] MPLS Traffic Engineering – Fast Reroute by Shuguftha Naveed, S. Vinay Kumar of Vasavi College of Engineering(Osmania University), Hyderabad in May, 2014 under IJSR – ISSN: 2319-7064.

[2] Virtual Router Redundancy Protocol (VRRP) by R. Hinden, Ed. Of Nokia in Internet Engineering Task Force – RFC 3768

[3] Fast Reroute Extensions to RSVP-TE for LSP Tunnels by P. Pan, Ed. Of Hammerhead Systems, G. Swallow, Ed. Of Cisco Systems and A. Atlas, Ed. Of Avici Systems in IETF RFC 4090

[4] Survey on the RIP, OSPF, EIGRP Routing Protocols by V. Vetriselvan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1058-1065

[5] Bidirectional Forwarding Detection (BFD) by D. Katz and D. Ward of Juniper Networks in IETF RFC 5880.

[6] Bidirectional Forwarding Detection (BFD) on Link Aggregation Group (LAG) Interfaces by M. Bhatia of Alcatel-Lucent, M. Chen of Huawei Technologies, S. Boutros, M. Binderberger of Cisco Systems. J. Haas of Juniper Networks in IETF RFC 7130

[7] Graceful OSPF Restart by J.Moy of Symacore Networks, P. Pillay-Esnault of Juniper Networks and A . Lindem of Redback Networks in IETF RFC 3623

[8] Graceful Restart Mechanism for BGP by S. Sangli, E. Chen of Cisco Systems, R. Fernando, J. Scudder, Y. Rekhter of Juniper Networks in IETF RFC 4724

[9] Graceful Restart Mechanism for BGP with MPLS by Y. Rekhter and R. Aggarwal of Juniper networks in IETF RFC 4781

[10] OSPFv3 Graceful Restart by P. Pillay-Esnault of Cisco Systems and A. Lindem of Redback Networks in IETF RFC 5187

[11] Basic Specification of IP Fast Reroute for IP Fast Reroute: Loop-Free Alternatives by A. Atlas, Ed. Of British Telecom and A. Zinin, Ed. Of Alcatel-Lucent in IETF RFC 5286

[12] Fast Reroute Extensions to RSVP-TE for LSP Tunnels by P. Pan, Ed. of Hammerhead Systems, G. Swallow, Ed. of Cisco Systems, A. Atlas, Ed. of Avici Systems in IETF RFC 4090