Service Function Chaining
Internet-Draft
Intended status: Informational
Expires: October 1, 2014

S. Kumar
C. Obediente
Cisco Systems, Inc.
M. Tufail
Citi

March 30, 2014

Service Function Chaining Use Cases In Data Centers
draft-kumar-sfc-dc-use-cases-01

Abstract

Data center operators deploy a variety of layer 4 through layer 7 service functions in both physical and virtual form factors. Most traffic originating, transiting, or terminating in the data center is subject to treatment by multiple service functions.

This document describes use cases that demonstrate the applicability of Service Function Chaining (SFC) within a data center environment and provides SFC requirements for data center centric use cases.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on October 1, 2014.

Table of Contents

1. Introduction

Data centers -- enterprise, cloud or service provider -- deploy service nodes at various points in the network topology. These nodes provide a range of service functions and the set of service functions hosted at a given service node may overlap with service functions hosted at other service nodes.

Often, data center topologies follow a hierarchical design with core, aggregation, access and virtual access layers of network devices. In such topologies service nodes are deployed either in the aggregation or access layers. More recent data center designs utilize a folded CLOS topology to improve scale, performance and resilience while ensuring determininstic hop count between end points. In such spine-leaf topologies, service nodes are often deployed at compute or virtual access layers as well as physical access layers.

The primary purpose of deploying service functions at different points in the network is to apply service functions to different types of traffic:

a. Traffic originating at physical or virtual workloads in the data center and destined to physical or virtual workloads in the data center; for example three-tiered deployment of applications: web, application, and database tiers, with traffic flowing between the adjacent tiers.

b. Traffic originating at a location remote to the data center and destined to physical or virtual workloads in the data center; for example traffic originating at a branch or regional office, destined to one of the primary data centers in an Enterprise, or traffic originating at one of the tenants of a Service Provider destined to that tenants applications in the Service Provider data center. Yet another variant of this type of traffic includes third party vendors and partners of the data center operator remotely accessing their applications in the data center over secure connections.

c. Traffic that is originating at a location remote to the data center and destined to a location remote to the data center but transiting through the data center; for example traffic originating at a mobile device destined to servers in the Internet routed through the data center to in order to service it.

Servicing of traffic involves directing the traffic through a series of service functions that may be located at different places in the network or within a single device connected to the network or any

combination in between. Delivery of multiple service functions in a sequence, in a datacenter, thus creates many requirements on the overall service delivery architecture. Such architectures may be termed service function chaining architectures while the list of service functions applied to the traffic is a Service Function Chain (SFC).

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Definition Of Terms

Additional terms are defined in [I-D.ietf-sfc-problem-statement], which the reader may find helpful.

End Point (EP): A device or an application that is the ultimate origination or destination entity of specific traffic. Mobile devices, desktop or server computers and applications running on them are some examples.

Workload (WL): A physical or virtual machine performing a dedicated task that consumes compute, storage, network and other resources. This may include web servers, database servers, storage servers and a variety of application servers.

Service Function (SF): A function that is responsible for specific treatment of received packets. A Service Function can act at the network layer or other OSI layers. A Service Function can be a virtual instance or be embedded in a physical network element. One of multiple Service Functions can be embedded in the same network element. Multiple instances of the Service Function can be enabled in the same administrative domain. A non-exhaustive list of Service Functions includes: firewalls, WAN and application acceleration, Deep Packet Inspection (DPI), server load balancers, NAT44 [RFC3022], NAT64 [RFC6146], HOST_ID injection, HTTP Header Enrichment functions, TCP optimizer, etc.

Service Node (SN): A virtual or physical device that hosts one or more service functions, which can be accessed via the network location associated with it.

Deep Packet Inspection (DPI): service function that performs stateful inspection of traffic, identification of applications and policy enforcement, among others.

Intrusion Detection and/or Prevention System (IDS/IPS): Is a DPI SN with additional capabilities to recognize malware and other threats and take corrective action.

Edge Firewall (EdgeFW): SN hosting service functions such as VPN, DHCP, NAT, IP-Audit, Protocol Inspection, DPI etc with policies primarily focussing on threats external to the data center.

Segment Firewall (SegFW): SN hosting a subset of the functions in the EdgeFW not including VPN and is deployed to protect traffic crossing segments, such as VLANs.

Application Firewall (AppFW): service function that isolates traffic within a segment or protects from application specific threats. This falls into the same class as DPI but deployed much closer to the applications. It is an intra-segment firewall.

Application Delivery Controller (ADC): service function that distributes traffic across a pool of servers (applications) for efficient resource utilization, application scaling as well as to provide high availability among others.

Web Optimization Control (WOC): SN hosting service functions to optimize the use of WAN link bandwidth, improve effective user throughput and latencies leading to overall improved user experience. WOC includes various optimizers such as compression, de-duplication, congestion control, application specific optimizers, etc. WOC requires peers at either end of the WAN link to perform optimizations. The scope of this document is limited to the DC side of the WAN link.

Monitoring (MON): SN hosting service functions to obtain operational visibility into the network to characterize network and application performance, troubleshoot performance issues, optimize resource utilization, etc.

Note: The above definitions are generalized. Actual implementations may vary in scope and in a lot of cases the actual service functions hosted on SNs overlap. For instance, DPI function is not only implemented as a standalone service function but is also implemented in EdgeFWs. Likewise EdgeFw functions, such as VPN, are implemented in routers. The terms used are representative of common usage and not absolute deployment.

3. Use Cases

The following sections highlight some of the most common data center use case scenarios and are in no way exhaustive.

3.1. Traffic Types

IT assets in an enterprise are consolidated into few data centers located in the main office. This consolidation stems from regulatory compliance regarding security, control on the enterprise assets, operational cost savings, disaster recovery strategies, etc. The data center resources are accessible from any geographic location whether inside or outside the enterprise network. Further, enterprise data centers may be organized along businesses, with each business treated as a tenant, thereby supporting multi-tenancy.

Service provider data centers have similar requirements as the enterprise. Data centers may be distributed regionally and globally to support the needs of their tenants. Multi-tenancy underlines every consideration in such data centers: resources and assets are organized & managed on tenant boundaries, policies are organized along tenant boundaries, traffic is segregated and policies enforced on tenant boundaries, etc. This is true in all "as a service" models: IaaS, PaaS and SaaS.

This leads to two primary types of traffic: North-South and East-West, both with different service requirements.

3.2. North-South Traffic

North-South traffic originates from outside the data center and is typically associated with users - onsite, remote and VPN - conducting their jobs. The traffic may also be associated with consumers accessing news, email, social media and other websites. This traffic is typically destined to applications or resources hosted in the data centers. Increasing adoption of BYOD and social networking applications requires traffic be analyzed, application and users be identified, transactions be authorized, and at the same time security threats be mitigated or eliminated. To this end, various service functions, as illustrated in Figure 1, are deployed in different SNs and in many instances of those SNs at various topological locations in the network. These SNs are selected based on the policy required for the specific use case.

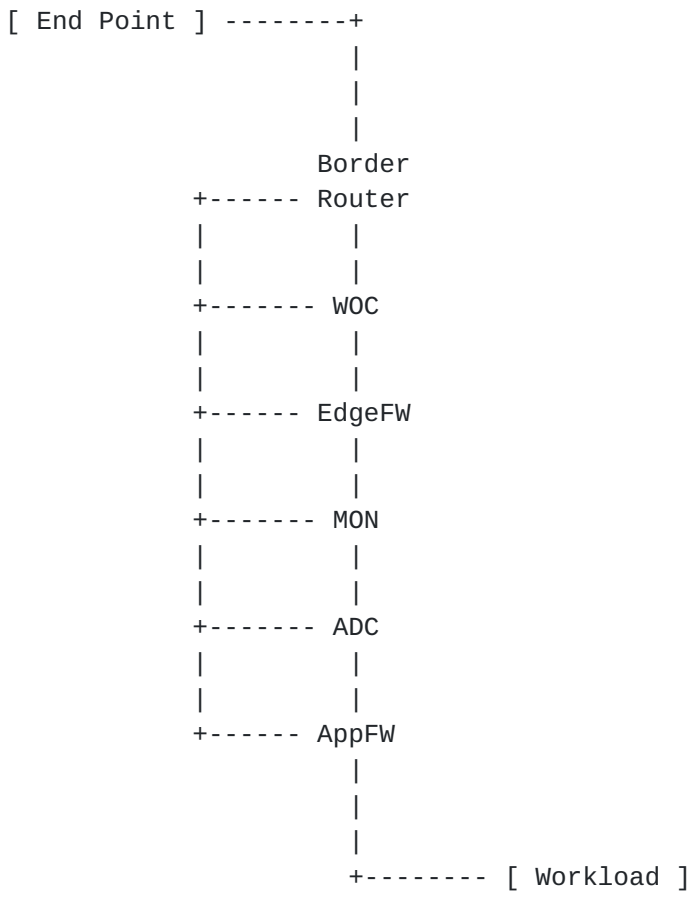

Figure 1: Service functions applied to North-South traffic

3.2.1. Sample north-south service function chains

SFC-1. EdgeFW

SFC-2. EdgeFW : ADC

SFC-3. EdgeFW : ADC : AppFW

SFC-4. WOC : EdgeFW : ADC : AppFW

SFC-5. WOC : EdgeFW : MON : ADC : AppFW

3.2.2. Sample north-south SFC description

Sample service chains numbered SFC-1 through SFC-5 capture the essence of services required on the north-south traffic.

SFC-1: This represents the simplest of use cases where a remote or mobile worker accesses a specific data center server. Traffic comes into the data center on VPN and is terminated on the EdgeFW. EdgeFW subjects the traffic to other service functions such as DPI, IPS/IDS, which may be hosted on the EdgeFW or off and reachable via VLAN stitching. Policy permitting, the traffic is allowed to its destination.

SFC-2: This is an extension of SFC-1. Traffic instead of destined to a specific server is destined to a data center application that is front ended by an ADC. The EdgeFW performs its function as before and the traffic is allowed, policy permitting. This traffic reaches its virtual destination, the ADC. ADC, based on local policy, which includes among other things predictors to select the real destination, determines the appropriate application instance. ADCs are stateful and ensure the return traffic pass through them by performing source NAT. Since many applications require the original source address, ADC preserves the original address in extension headers of the HTTP protocol. Traffic is then forwarded on to the ultimate destination - the real application workload.

SFC-3: This extends SFC-2. The segment where the application server resides may be shared with other applications and resources. To segregate these applications and resources further fine grain policies may be required and are enforced via a security appliance such as the AppFW. As a consequence AppFW first services the traffic from the load balancer before it is forwarded to its ultimate destination, the application server.

SFC-4: This is a variant of SFC-3 with WOC being part of the chain. This represents the use case where users at a branch office access the data center resources. The WOC SNs located at either end of the WAN optimize the traffic first. The WOC located in the datacenter requires a mechanism to steer traffic to it while not deployed inline with the traffic. This is achieved either with PBR or VLAN stitching. WOC treated traffic is subject to firewall policies which may lead to the application of SFs such as protocol inspection, DPI, IDS/IPS and then forwarded to its virtual destination, the ADC.

SFC-5: This is similar to SFC-4. An additional service - MON, is used to collect and analyze traffic entering and leaving the data center. This monitoring and analysis of traffic helps maintain performance levels of the infrastructure to achieve service level agreements, particularly in SP data centers.

3.3. East-West Traffic

This is the predominant traffic in data centers today. Server virtualization has led to the new paradigm where virtual machines can migrate from one server to another across the datacenter. This explosion in east-west traffic is leading to newer data center network fabric architectures that provide consistent latencies from one point in the fabric to another.

The key difference with east-west from the north-south traffic is in the kind of threats and the security needs thereof. Unlike north-south traffic where security threats may come from outside the data center, any threat to this traffic comes from within the data center.

Figure 2: Service functions applied to East-West traffic

3.3.1. Sample east-west service function chains

SFC-6. SegFW : ADC : MON : AppFW

3.3.2. Sample east-west SFC description

SFC-6: In a typical three tiered architecture, requests coming to a webserver trigger interaction with application servers, which in turn trigger interaction with the database servers. It has to be noted that each of these tiers are deployed in their own segments or zones for isolation, optimization and security. SegFW enforces the security policies between the tiers and facilitates isolation at the segment level or addres space re-use via NAT deployment. ADC provides the distribution, scale and resiliency to the applications while the AppFW protects and isolates traffic within the segement in addition to enforcing application specific security policies. Finally, monitoring service enables visibility into application traffic which in turn is used to maintain application performance levels.

3.4. Multi-tenancy

Multi-tenancy is relevant in both enterprise as well as service provider data centers although it is the primary differentiator between service provider (SP) and enterprise datacenter. Enterprises treat organizations or business units within the enterprise as tenants and thus require tenant aware service models.

Multi-tenant service delivery is achieved in two primary ways: a) SNs themselves are tenant aware - every SN is built to support multiple tenants. b) SN instances are dedicated for each tenant. In both the cases, the SP manages the SNs.

To support multi-tenant aware service functions or SNs, traffic being serviced by a service function chain has to be identified by a tenant identifier. A tenant identifier has to be carried along with the traffic to be serviced. It is typical of tenant assets to be deployed in an isolated layer2 or layer3 domain such as VLAN, VXLAN or VRF. It has to be noted that the SNs themselves maybe deployed in different domains to suit the deployment needs of the SP and hence using the domain in which the SN is deployed is not an option. Although such a model is feasible it removes the deployment flexibility for the service providers.

3.5. SFCs in data centers

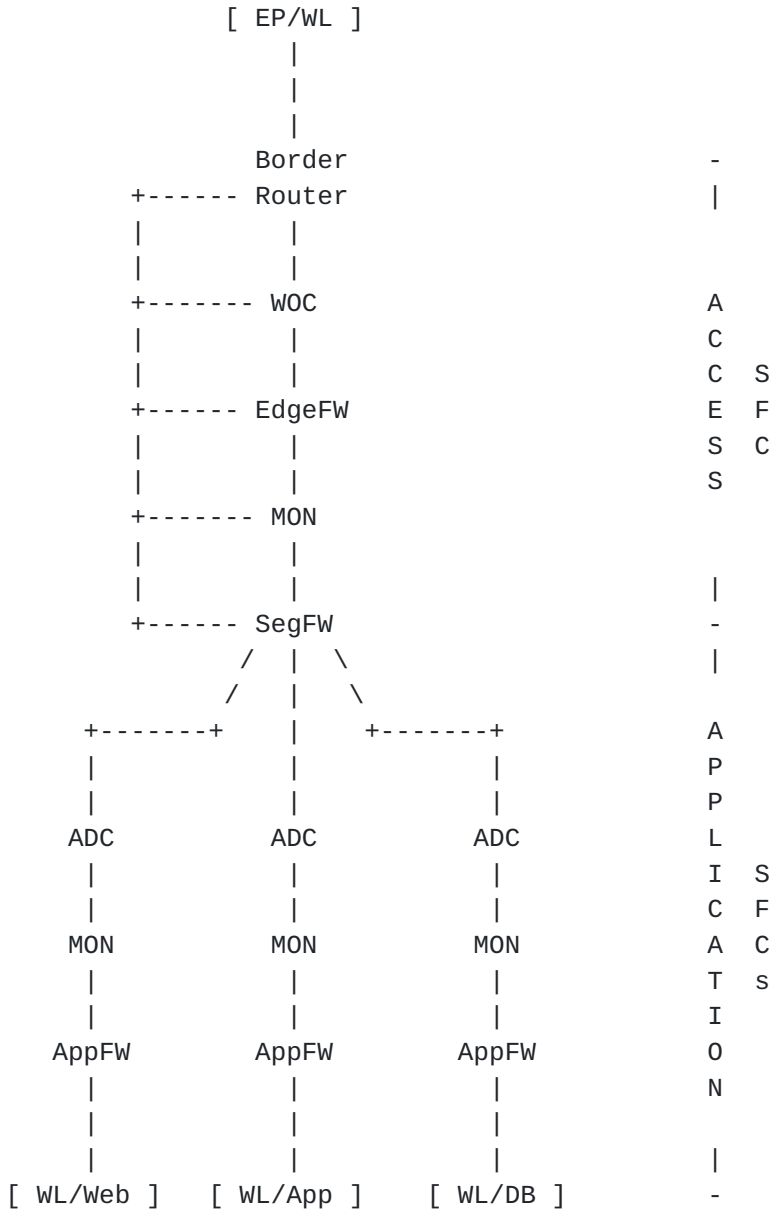

Figure 3: Service function chains in data center

Figure 3 shows the global view of SFCs applied in an enterprise or service provider data center. At a high level the SFCs can be broadly categorized into two types:

1. Access SFCs

2. Application SFCs

Access SFCs are focused on servicing traffic entering and leaving the data center while Application SFCs are focused on servicing traffic destined to applications.

Service providers deploy a single "Access SFC" and multiple "Application SFCs" for each tenant. Enterprise data center operators on the other hand may not have a need for Access SFCs depending on the size and requirements of the enterprise. Where such Access SFCs are indeed needed, such as large enterprises, the operator may deploy a bare minimum Access SFC instead. Such simple Access SFCs include WOC and VPN SFs to support the branch and mobile user traffic while at the same time utilizing the security policies in the application SFCs. The latter is the case in de-perimetrized network architectures where security policies are enforced close to the resources and applications as opposed to the WAN edge.

4. Drawbacks Of Existing Service Chaining Methods

The above use cases are realized in a traditional fashion and are not viable in the evolving hybrid data centers with virtual and physical assets. The following are some of the obvious short comings of existing SFC methods exposed by the above use cases.

DB-1. Policy based purely on VLANs is no longer sufficient. Connecting SNs to each other to construct a service chain thus makes it very static and removes deployment flexibility. As can be seen from the sample north-south service chains, a large number of VLANs not only have to be stitched in a certain fashion to achieve a basic SFC, it is simply not flexible to share the SNs among different SFCs as even simple sharing among a few SNs becomes intractable from basic configuration perspective let alone future changes or manageability aspects.

DB-2. Traffic does not always have to be steered through all the SNs of a traditional VLAN stitched service chain. In Figure 1, traffic from the border router is not always necessary to flow through the WOC as remote or mobile worker may not have a WOC peer deployed. Connecting multiple VLANs among service nodes to overcome to achieve this only aggravates the problem of deployment and manageability. Truly, there exists a need for dynamically determining the next sub SFC at such branching points to avoid forcing all

traffic through the same SFC.

DB-3. Virtual environments require the virtual SNs be migration capable just like the compute workloads. As a consequence it is simply not feasible to continue VLAN stitching in the hybrid data centers. Every time a virtual SN migrates, such as the AppFW in Figure 1 and Figure 2, the operator has to ensure the VLANs are provisioned in the destination. Further, stretching the VLANs across the network may not be an option for the operator or even worse the virtual SN may be L3 hop away from the previous SN.

DB-4. Policy Based Routing (PBR) to move traffic to SNs although provides a much better granularity than VLAN stitching it suffers from the requirement to configure such policies all along the path to the SNs. In Figure 1, if WOC is multiple hops away from the border router, all network elements in between border router and WOC need to be configured with consistent policies.

DB-5. Source NAT (SNAT) is required by some SNs, such as ADC in Figure 1, in order to ensure traffic sent to the load balanced servers pass through the ADC in reverse direction. However, SNAT removes the ability to detect the originator of the traffic. Using HTTP extension header to pass originator information is not only an overhead but addresses only one specific protocol.

DB-6. Static service chains do not allow for scaling the SFCs as they require the ability to add SNs or remove SNs to scale up and down the service capacity. Likewise the ability to dynamically pick one among the many SN instance is not available. For instance, WOC must scale to support the high data rate of traffic flowing to the data center. Likewise, AppFWs must scale up to not impact the workload throughput. Further they may be required to sacle within tenant boundaries.

DB-7. Static SFCs constructed over the under lay network cannot pass metadata to the SNs. Border Router in Figure 1 cannot pass policy based tags derived locally at the start of the SFC all the way through the SFC. Such metadata is necessary to enforce consistent security policies across the network, as one example.

DB-8. In multi-tenant deployments, the segment on which the SN is deployed may not correspond to the segment assigned to the tenant in which the workloads are hosted. In Figure 2, AppFW may be deployed on a different segment than the Workload. As a consequence, it is not viable to derive the tenant segment simply based on the tag associated with the incoming traffic at the AppFW. This ultimately prevents the ability to have the same SN serve multiple tenants. Forcing the SN to be on the same segment as the tenants' workload limits deployment flexibility.

DB-9. Traffic may originate in a physical or virtual network or transit these networks before being delivered to the SNs for servicing. The following is very complex to achieve with the existing SFC mechanism.

A. Physical SN servicing traffic originating in the virtual access network.

B. Virtual SN servicing traffic originating in the physical network.

DB-10. Although SNs are purpose built service appliances, it is neither a requirement nor an indication of how service functions are implemented in emerging data centers with commodity compute and storage capabilities. AppFW in Figure 1, for instance, may be built and deployed as a virtual SN. Further, SFCs are limited to exclusively physical or virtual SNs and not a mix. This excludes the ability to combine the benefits offered by physical SNs with the flexibility and agility of the virtual SNs. The EdgeFW in Figure 1, for instance, may be a purpose built SN to take advantage of SFs implemented in hardware while the AppFW may be a virtual SN deployed to be close to the virtual workload and may even move with the workload in the virtual environment.

5. General Requirements

The above use cases and the drawbacks thereof lead to the following general requirements in today's evolving hybrid datacenters to apply SFCs to traffic.

GR1. SFC polices MUST be applicable at the edges - network elements as well as the workloads.

GR2. SFC policies MUST be applicable to either Ingress or Egress traffic.

GR3. SFC MUST support virtual as well as physical SNs.

GR4. SFC SHOULD support the ability to mix virtual and physical SNs in the same SFC.

GR5. SFC SNs MUST be deployable L2 or L3 hop away from each other or from the SFC starting entity.

GR6. SFC traffic MUST be allowed to follow paths free from underlying network topology.

GR7. SFC SNs MUST be able to derive the tenant identification without being tied to the underlying topology

GR8. SFCs MUST support the ability to pass metadata among the SNs or between the SNs and the network elements.

GR9. A composite SFC SHOULD be achievable by way of joining sub SFCs, branching to sub SFCs where necessary.

GR10. SFCs SHOULD NOT require SNAT inside the SFs to attract traffic back to them

GR11. SFCs SHOULD have the ability to choose SN instances dynamically, at the time of forwarding traffic to them.

6. Acknowledgements

The authors would like to thank Paul Quinn, Jim Guichard, Jim French and Nagaraj Bagepalli for their review and comments.

A special thanks to Abhijit Patra for his guidance.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

Security of traffic being serviced is very important in the use cases described in this document. The SNs deployed as part of the SFC are expected to include SFs specifically addressing the security aspect

either individually or in concert with other SFs. In this regard organizational security policies are expected to drive the security posture adapted in the SFCs. However, securing the traffic moving between the SFs or SNs is not a consideration beyond the methods used for moving such traffic.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[I-D.ietf-sfc-problem-statement] Quinn, P. and T. Nadeau, "Service Function Chaining Problem Statement", draft-ietf-sfc-problem-statement-02 (work in progress), February 2014.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

Authors' Addresses

Surendra Kumar
Cisco Systems, Inc.
170 W. Tasman Dr.
San Jose, CA 95134

Email: smkumar@cisco.com

Cesar Obediente
Cisco Systems, Inc.
7200-10 Kit Creek Rd.
Resarch Triangle Park, NC 27709-4987

Email: cobedien@cisco.com

Mudassir Tufail
Citi
238 King George Rd
Warren, NJ 07059-5153

Email: mudassir.tufail@citi.com