Internet Draft

Jamal Hadi Salim
Znyx Networks
July 2001

Requirements for Separation of IP Control and Forwarding Services

draft-salim-forces-alt-jhs-00.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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.''

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Conventions used in this document

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].

1. Abstract

This document defines a set of requirements for mechanisms to logically separate the control and data forwarding IP services in an IP network element (NE).

2. Introduction

An IP NE contains two logically separated entities that cooperate to provide services to packets travessing the NE. While separate and have a very well defined logical functions, these entities provide a unified external view of the NE. The NE entities are: control-plane(CP) components and forwarding-Engine(FE) components.

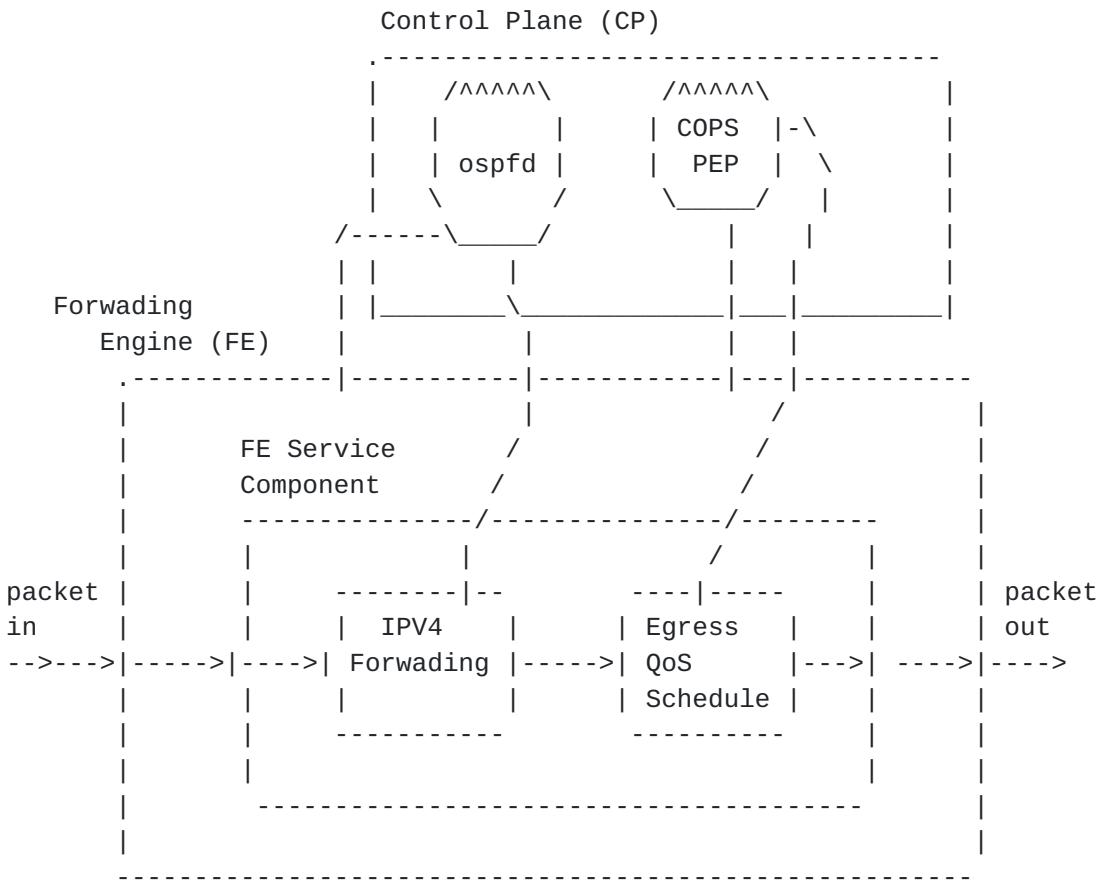
It is important to re-emphasize that the FE<->CP interaction, as well as the sole reason for the existence of the CP and FE, is to provide services to IP packets.

ForCES attempts to define a clear separation between the two entities of the NE in order to have them evolve separetely as opposed to the current monolithic evolution.

2.1. Some definitions

A CP may have several CP componets each providing control for a different IP service being executed by a FE component. This means that there will be several CP components on a physical CP if it is controlling several IP services. Likewise for the FE. In essence, the cohesion between a CP component and a FE component is the service abstraction.

In the diagram below we show a simple FE<->CP setup to provide an example of the classical IPv4 service with an extension to do some basic QoS egress scheduling and how it fits in this described model.



2.1.1. Control Plane Components

Control plane components in the ForCES context would encompass signalling protocols with diversity ranging from dynamic routing protocols such as OSPF to tag distribution protocols such as CR-LDP. Classical Management protocols and activities also fall under this category. These include SNMP, COPS or proprietary CLI/GUI configuration mechanisms.

The purpose of the control plane is to provide an execution environment for the above mentioned activities with the ultimate goal to configure and manage the second NE component: the FE. The result of the configuration would define the way packets travesing the FE are treated.

The CP components are traditionaly run in software since they tend to be very rich in syntax and are moving targets requiring ease of

modification.

In the above diagram, ospfd and COPS are distinct control plane components.

2.1.2. Forwarding Engine Components

The FE is the entity of the NE that incoming packets (from the network into the NE) first encounter.

The FE's service specific component massages the packet to provide it with a treatment to achieve a IP service as defined by the control plane components for that IP service. Different services will utilize different FE components. Service modules maybe chained to achieve a more complex service (as shown in the diagram). When built for providing a specific service, the FE service component will adhere to a Forwading Model (to use ForCES charter speak) for that service.

The FE could be implemented in software, ASICs, or Network Processors(NPs). Classical approach is to have a mixture of ASICs and software. We will not delve into design of an FE but rather focus on its purpose.

In the above diagram, the FE components include both the IPV4 Forwarding module as well as the Egress Scheduling module. Another service might just replace the IPV4 forwarder module with a web-switch forwarder. A simpler classical service would have constituted only the IPV4 forwarder.

2.1.3. IP Services

An IP Service is the treatment of an IP packet within the NE.

The time span of the service is from the moment when the packet arrives at the NE to the moment it departs. In essence an IP service in this context is a Per-Hop Behavior. A service control/signaling protocol/management-application (CP components running on NEs defining the end to end path) unifies the end to end view of the IP service. As noted above, these CP components then define the behavior of the FE (and therefore the NE) to a described packet.

A simple example of an IP service is the classical IPv4 Forwading. In this case, control components such as routing protocols(OSPF, RIP etc) and proprietary CLI/GUI configurations modify the FE's forwarding tables in order to offer the simple service of forwarding packets to the next hop. Traditionally, NEs offering this simple service are known as routers.

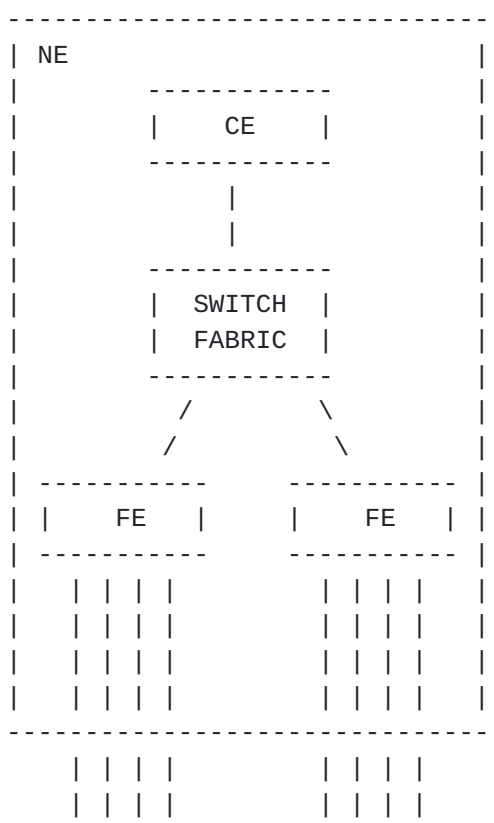
Over the years it has become important to add aditional services to the routers to meet emerging requirements. More complex services extending classical forwarding were added and standardized. These newer services might go beyond the layer 3 contents of the packet header. However, the name "router", although a misnomer, is still used to describe these NEs. Services (which may look beyond the classical L3 headers) here include firewalling, Qos in Diffserv and RSVP, NATs, policy based routing etc. Newer control protocols or management activities are introduced with these new services.

Given the observed evolution path, a very important intent is not to limit what an IP service should be. Rather leave the service definition flexible enough to not restrict future innovation. For example, one should be easily be able to integrate the services being defined by OPES within the ForCES model.

One extreme definition of a IP service is something a service provider would be able to charge for.

3. Architectural Requirements

Below is a diagram illustrating an example NE composed of one CE and two FEs connected by a switching fabric.



(1) The CP and FE (and their components) MUST communicate via the ForCES protocol in the IP service definition.

(2) The CP and FE MAY reside on different physical devices.

(3) The CP and FE MUST have a way to connect to each other. This MAY be using any mechanism of convinience. Examples of known interconnect mechanisms are Ethernet connections, proprietary backplanes, open standard buses (such as PCI), ATM (cell) fabrics, and abstractions such as sockets. Addressing to the FE is defined by access method.

(4) There is a cohesiveness between a CP component and an FE component as defined by an IP service they are trying to deliver. This is the only restriction in the architecture i.e there is no other direct linkage between an FE and CP. A CP component MAY control several FEs which may reside on different physical devices; vice-versa, there MAY be more than one CP(component) controlling a single FE (service). [Think of several routing daemons each running a different routing protocol trying to configure a Forwading service. The FE component being configured in the described service is responsible of the serialization (eg shared access of its tables).]. A direct consequence of this requirement is that several FE

components and hence several IP services can run on a single FE.

This ability to extend the number of physical CPs and FEs allows ForCES architecture to scale.

(5) There MAY be mechanisms for CEs and FEs to discover each other without apriori configuration. There MUST be a simple mechanism such as static setup to allow this.

(6) There MUST be a mechanism by which CEs and FEs can be authenticated to prevent unauthorized components from joining the network element.

(7) In the case of a CP residing on a remote device or on some proprietary device, it MAY have a proxy on an FE controller. The protocol between the proxy and FE MUST be that defined by ForCES. The protocol between the device and the CP is left to the implementation.

(8) The scope of the ForCES problem is only focussed on CP<->FE communication. CP or FE services that require to have other forms of architectures (such as HA and redundancy) MUST define their own methodology. This will help keep ForCES simple.

4. FE Services

These are services that the FE provides to either CP components or FE components. Note that we are refering to the FE-in-general here and that these are not IP services although will be communicated via the ForCES protocol. It is important to separate what the general FE provides in terms of resource and event management and the FE component in terms of IP service execution. The FE component events are only available when the FE component (or service) is active; the FE events are available at any time the FE is up. An FE model description (in addition to an FE component model which is a MUST for an IP service) MAY be required to express these simple services.

Generally, FE or CP components subscribe to listen to FE events in order to properly deliver the service value. For example, a FE component would rather not send to a downned interface and the CP component would notify its peers of the downed interface so better service connectivity decisions can be made amongst the peers.

The CP component of a IP service might also ask the FE for packet redirection to itself for the purposes of providing the IP service.

The control and management of resources within a FE is not within the mandate of ForCES. It is assumed some other mechanisms are responsible.

(1) Port Functions

When queried, the FE MUST be capable of expressing the number of ports on the device,the static attributes of each port (e.g., port type, link speed), and the configurable attributes of each port (e.g., IP address, administrative status).

(2) Event Capability discovery

The FE MAY be capable of expressing the types of asynchronous events (e.g., link up/down, redirected packet, out of memory) that a FE will generate. Common events and their templates MUST be standardized, similar to those for well known services described in the next section. Example here are those of link maintanance and those already standardized by GSMP for port events.

(3) Event Notification

The FE SHOULD be capable to deliver its events to subscribed components.

(4) Vendor-Specific Functions

The FE SHOULD be extensible so that vendor-specific event notification can be offered.

(5) Network Management capability Discovery

The FE MAY be capable of expressing the types of statistics for the resources it manages when queried.

(6) Network Management

The FE MUST be capable of delivering statistics for the resources it manages when queried.

(7) Request For Packets

The FE MUST be capable of delivering packets or copies of requested packets to the CP. Normally, these would be control packets that belong to an IP service. The CP component of the IP service would request to have the FE send all control packets to it. An example here would be all OSPF packets being passed to ospfd in diagram 1.

Another example would be all TCP SYN and FIN packets in a split-TCP webswitch to a specific service CP component on the physical CP.

5. IP Service Requirements

(1) An IP service is delivered to customer packets by the cohesive effort of the service's CP and FE components. The components MAY subscribe to FE services in order to better deliver services.

As an example in diagram 1 above, ospfd and COPS both work with their two FE counterparts to cohesively deliver an abstracted IP service which both forwards IPV4 packets and provides selective Egress QoS to customer packets.

Both will be subscribed to FE services on link events as well as FE services to deliver their respective protocol packets.

(2) Services MUST be defined using templates such as those found in GSMP[GSMP]. This will allow for simpler mechanisms for FE capability and service discovery. [Note the difference between GSMP and ForCES templates is that while GSMP's define switch connection management, ForCES' defines service management].

(3) Well known services MUST have their templates standardized. Example of well known services here includes the classical RFC1812 router and well known extensions to RFC1812 such as Diffserv and policy based routing. All standardized services MUST be issued "service Numbers".

(4) A range of service numbers MUST be reserved for the Opaque service. This is a service that could be user defined. It will allow for faster deployment of newly innovated services without requiring standardization. This would also allow for vendor specific extensions.

6. Protocol Requirements

(1) The ForCES protocol interconnects the FE service portions to their controllers (CPs). Different types of services will have different protocol requirements. It is therefore imperative to not enforce a service to a specific protocol. Rather have the service choose from a set of available mechanisms to define the protocol set.

(2) The main activities foreseen in the protocol are: discovery, configuration, event notification and statistical querying.

(3) In order for the FE and CE components of a network element to act in concert, they need to discover each other. At the minimalist level the discovery phase is hardcoded by static entries. At a slightly higher level is dynamic discovery. Using service templates allows ForCES to re-use many of the existing Service Discovery protocols and benefit from their operational experiences and wider deployment. Example of service discovery protocols include: Universal Plug-and-play, Jini, Bluetooth's SDP and SLP.

(4) After the FE discovers their CPs of choice and negotiated the service contract, the established phase is entered. In this phase the CP and FE components participate in service delivery. This includes configuration, event notification, and statistical as well as config queries.

(5) An FE component may choose at any time to terminate its contract with the CP component (and may join a different CP, for example. This is left to the specific service).

(6) There MUST be a mechanism to allow a service connection between a FE component and a CP component to have choice of either being reliable or non-reliable communication or a mixture of the two in the context of the activities in the established phase.

(7) There MUST be a mechanism by which CEs and FEs can quickly determine when a loss of connectivity between them has occurred. The policy definition MUST be left upto the IP service.

(8) Since FE configuration contains information critical to the functioning of a network any protocols defined MUST support a method of securing communication between FEs and CEs to ensure that information is delivered securely in an unmodified form in the established phase.

(9) A mechanism for Authentication, Authorization and Accounting MUST be provided.

7. Protocol Applicability statement

With the clear separation between CPs and FEs that ForCES is striving for, and more precisely with use of the IP service abstraction, the ForCES protocol becomes usable in other WG areas which have

similar setups. These range from the MIDCOM protocol to the protocol for configuring an OPES device and all sorts of layer 3 devices. One could venture into the Sub-IP area of CCAMP (but that is better suited to GSMP).

8. Security Considerations

Refer to requirement 8) of the protocol requirements.

9. References

[GSMP] A. Doria, F. Hellstrand, K. Sundell, T. Worster "General Switch Management Protocol V3", draft-ietf-gsmp-09.txt June, 2001

[RFC1812] F. Baker et al, "Requirements for IP Version 4 Routers", RFC 1812, June 1995.

10. Acknowledgements

Authors of internet draft draft-anderson-forces-req-02.txt from which this draft is derived.

11. Author's Address:

Jamal Hadi Salim
Znyx Networks
Ottawa, Canada