ForCES Working Group
Internet Draft
Document: draft-putzolu-forces-evaluation-01.txt
Expires: April 2004

D. Putzolu (editor)
Intel

October 2003

ForCES Protocol Evaluation Draft

Status of this Memo

Abstract

This document provides an evaluation of the applicability of three proposed approaches for a ForCES protocol: FACT[2], GRMP[3], and Netlink2[4]. A summary of each of the proposed protocols against the ForCES requirements[5] and the ForCES framework[6] is provided. Compliancy of each of the protocols against each requirement is detailed. A conclusion summarizes how each of the protocols fares in the evaluation.

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 [7].

Table of Contents

1. Introduction

This document provides an evaluation of the applicability of FACT, GRMP, and Netlink2 as the ForCES protocol. This evaluation provides overviews of the protocols and general statements of applicability based upon the ForCES framework and requirements documents. The format and structure as well as some of the introductory content of this document is based on and taken from a similar document being

produced in the MIDCOM working group[8].

The process for protocol evaluation found in this document consists of individuals providing sections evaluating a specific protocol. These sections are incorporated by the editor of the document, and are subject to feedback and changes based on the consensus of the ForCES working group. Some protocols that might be considered as potentially applicable as the ForCES protocol are not evaluated in this document since there where no champions to submit evaluations for them.

Section 2 of this document is a list of the proposed protocols along with background information about each of the protocols.

Section 3 of this document is an evaluation of the proposed protocols against the architectural requirements found in section 5 of the ForCES requirements. The purpose of this section is to determine how well each of the proposed protocols maps to the ForCES architecture.

Section 4 of this document is an evaluation of the proposed protocols against the model requirements found in ForCES requirements. The purpose of this section is to determine how well each of the proposed protocols can be used with FEs that meet the ForCES model requirements.

Section 5 of this document is an item level evaluation of the proposed protocols against the protocol requirements found in the ForCES requirements. The purpose of this section is to determine how well each of the proposed protocols satisfies each of the protocol requirements.

Section 6 summarizes the evaluation, and includes a table with a breakdown for each of the protocols versus the requirements. The following categories of compliance are used: Fully met, partially met through the use of extensions, partially met through other changes to the protocol, or not met. This summary is not a conclusive statement of the suitability of the protocols, but rather to provide information to be considered as input into the overall protocol decision process.

2. Protocol Proposals

The following protocols have been submitted to the ForCES WG for consideration:

- o FACT
- o GRMP
- o Netlink2

The following sections provide overviews of each of the protocols as

well as relevant background information about each protocol.

2.1 FACT

Network Elements (NE) such as routers are becoming more and more complex as they try to cope with demanding features like policy based routing, firewalls and NATs, and QoS aware routing. As a result, issues like scalability, (the ability to cost-effectively grow a network as demand increases) and extensibility (the ability to dynamically configure the network for some specific services by programming the NEs that handle those services) become very important. The ForCEs protocol has been specified to help resolve these issues by decoupling control and forwarding elements of a network element, and also adding extensibility features to the NE.

FACT (Forwarding And Control ElemenT) protocol has been designed for exchanging information between control elements (CEs) and forwarding elements (FEs) distributed in a ForCES network element (NE). The relationship between CEs and FEs is a master/slave one. The FACT protocol is logically separated into a base protocol and an extensible data model defined in [9]. It consists of a common, fixed size header and variable size payload which carries the information defined by the data model. All FACT messages are 32-bit aligned.

FACTÆs messages are grouped into six (6) classes namely:

1) Connection and Association messages, which help establish logical connections between FEs and CEs,
2) Capabilities Control messages, which the CE uses to query and configure the capabilities of the FE,
3) State Maintenance messages, which are used to track element states,
4) Traffic Maintenance messages, which are used exchanging control packets between CEs and FEs,
5) Event Notification messages used for reporting asynchronous events, and
6) Vendor Specific messages which are used to extend the protocol beyond its current capabilities.

FACT supports versioning and priority, and its unique design of separating control and data traffic into different channels helps reduce the threat of Denial of Service (DoS) attacks making the protocol more robust. It provides reliability by using a reliable transport protocol, thus simplifying the protocol design. It also provides failover mechanisms that can exploit redundant elements in the system or network element.

2.2 GRMP

General Router Management Protocol (GRMP) Version 1 is intended to be as a ForCES protocol, which acts at the Fp reference point in the

ForCES framework. GRMP is designed to meet all the requirements for the ForCES protocol at the Fp reference point.

GRMP protocol is a master-slave protocol. CEs act as masters and FEs as slaves. Slaves only have right to send to masters request or query, response, and report messages. While masters can send command messages to slaves as well as request or query, response, and report messages. GRMP protocol acts in a mode of a base-protocol associated with a data model, which is FE model in ForCES. That is, GRMP only defines basic management messages, while many of managed data are defined in the associated ForCES FE model. Most of the data types and functional descriptions relating to specific IP services such as routing service conforming to RFC 1812 [10], QoS configurations, high-tough capabilities like NAT and firewall should be expressed by Logical functional Blocks (LFBs) defined by ForCES FE model and their specific LFB topologies. GRMP application layer is responsible to configure the LFBs and the LFB topologies so as to implement specific IP services and QoS deployment.

GRMP is developed separately from General Switch Management Protocol (GSMP) protocol since current base specification version of GSMP did not support some of the basic functionality needed by GRMP. However, GRMP has been considering its possible compatibility with GSMP; with the work currently ongoing within the GSMP group that is adding flexibility to the base, GRMP is willing to work with the GSMP group to integrate this with GSMP.

GRMP protocol is composed of protocol messages. GRMP organizes these messages according to different object types the messages manage for ForCES architecture, as follows:

1. FE management messages
This type of messages is for FE layer operations, that is, to take a whole FE as a managed object. Messages of this type include that for operation of FE join or leave, FE action, FE attribute, FE event report, etc.

2. LFB management messages
This type of messages is for LFB layer operations. It takes LFBs as its managed objects. Messages of this type include that for operation of LFB action, LFB attribute, etc.

3. Datapath management messages
This type of messages is for the management of datapaths in an FE. It takes datapathes as the managed objects.

4. CE Informing messages
This type of messages takes CE as the operated object to ask CE to

send some information. Because CE acts as a master in ForCES

protocol, allowed operations toward CE from FE are only that like CE query request, CE event report, etc.

5. GRMP slave management

In GRMP, the GRMP slave part is treated as a module inside an FE. This module is outside of FE model scope, which will automatically be launched in a FE when the FE is to join a ForCES NE. Management to a GRMP slave is that like DoS protection policy, CE or FE failover policy, etc. Note that the management of GRMP slave does not take specific GRMP messages, in stead, it uses FE management messages for GRMP slave is considered as an object at FE layer.

6. Managed Object(MO) management messages

In order to meet ForCES requirement of supporting network management tools like SNMP, GRMP provides the management messages. This type of messages takes a Managed Object (MO) defined in some specific network management tools as its managed objects. Operations of MOs are that like MO get, MO set, and MO response.

There is another GRMP message that is defined out of scope of management purposes. This is GRMP ACK (acknowledge) message, which is mainly for communication control and error control purpose between CE and FE.

From perspective the message communication types between CE and FE, GRMP messages can be divided into following types:

1. Messages for query and response types. These messages can be sent from CE to FE, or from FE to CE.

2. Messages for command and configuration types. These messages are only sent from CE to FE.

3. Messages for report types. These messages can be sent from CE to FE, or from FE to CE.

In GRMP, a 5 bits "object class" identifier [3 Section 3.4.5] is applied to following GRMP managed objects:

FE capability types
FE attribute types
LFB types
LFB attribute types
CE attribute types
CE event types

Currently defined "object class" includes GRMP class, different version of ForCES FE model class, vendor class. This means the managed objects above can include that defined by GRMP itself, different versions of ForCES FE models, and different vendors.

2.3 Netlink2

<Text for this section>

3. Architectural Requirements Compliance Evaluation

This section contains a review of each protocol proposalÆs level of compliance to the ForCES architecture requirements. Many of the architectural requirements will be instantiated in some fashion in the protocol selected. Given that the architectural requirements are not direct protocol requirements, the review below will consist of prose rather than specific levels of compliance as is used in the protocol section below.

3.1 FACT

FACT fulfills all the protocol requirements listed in section 5. By doing this it in turn supports all the architectural requirements defined in the ForCES Requirements [5]. FACT supports the separation of the NE into CE and FE components, with CE handling roles such as control, signaling and routing data calculation. The CE configures the FE with all the information necessary for the FEÆs proper operation. The FEÆs functions could be layer-3 forwarding, NAT, metering, shaping, firewall, etc. Also, FACT state maintenance messages help resolve the various states of the distributed CEs and FEs to provide a unified state of the NE.

3.2 GRMP

GRMP protocol is designed based on the ForCES architecture requirements. We review its compliance to the architecture requirements according to the individual requirement items as below:

1) For architecture requirement #1
GRMP packets can be transported via any suitable mediums, such as TCP/IP, Ethernet, ATM fabrics, and bus backplanes. Because of the design consideration for GRMP to be compatible with GSMP protocol, packet encapsulations defined for GSMP protocols as in RFC 3293 can also be applied to GRMP.

2) For architecture requirement #2
ForCES requires that FEs MUST support a minimal set of capabilities necessary for establishing network connectivity (e.g., interface discovery, port up/down functions). This process is usually out of

the range of ForCES protocol, but GRMP protocol has no restriction on this functionality.

3) For architecture requirement #3

By properly configuring FEs with their LFBs in a NE via GRMP protocol, packets can arrive at one FE and depart at the other FE or FEs. In the case where more than one CE work simultaneously in a NE, the consistency and synchronization of control of the CEs is essential for above functionality, which is beyond the scope of ForCES protocol and also architecture requirements.

4) For architecture requirement #4
By properly configuring LFBs in FEs in a NE via GRMP protocol, the NE can appear as a single functional device in a network. In the case more than one CE work simultaneously in a NE, the consistency and synchronization for the CEs to control FEs is essential for this functionality, which is beyond the scope of ForCES protocol and architecture requirements.

5) For architecture requirement #5
ForCES protocol requirement #2 has comprised this architecture requirement, refer to Section 5.2.2 for details on GRMP compliance to this requirement.

6) For architecture requirement #6
Please refer to Section 5.13.2 for details.

7) For architecture requirement #7
Please refer to Section 5.8.2 for details.

8) For architecture requirement #8
Please refer to Section 5.9.2 for details.

9) For architecture requirement #9
GRMP supports RFC1812 [10] compliant router functions by means of following mechanisms in GRMP:
-Fully supporting ForCES FE model
-Packet redirection messages
-Datapath management messages
-Managed Object(MO) management messages

10) For architecture requirement #10
In GRMP, FE topology query and response messages [3 Section 4.1.3] are used for CEs to query FE topology information in a NE.

11) For architecture requirement #11
Please refer to Section 5.3.2 for details.

12) For architecture requirement #12
Please refer to Section 5.11.2 for details.

13) For architecture requirement #13

GRMP supports multiple FEs working together in a NE by using FE identifiers and by allowing CEs to be informed of FE topology information. GRMP supports multiple CEs working together in a NE by supporting CE redundancy or failover functionality.

14) For architecture requirement #14
GRMP defines Managed Object (MO) management messages [3 Section 4.5] to meet the requirement, in which it states that it must not be possible for management tools (e.g. SNMP, etc) to change the state of a FE in a manner that affects overall NE behavior without the CE being notified.

A MO is an object defined by some network management tool, such as the object defined by Object Identifier in SNMP MIBs.

MO management messages work in the way as below:

1. Query of MOs resident in an FE can be directly implemented by network management tools. To perform this, it is necessary for CE to configure some LFBs that has high touch capability in the FE.

2. Change of MOs resident in an FE can only be made via a CE. To do this, the high touch LFBs in the FE will redirect all network management protocol messages like SNMP messages concerning MO changes to the CE, then the CE will use the MO management messages to change values of MOs in the FE. Of course, if necessary, query of the MOs can also be made via the CE.

3. MOs resident in a CE can be directly queried or changed by the CE with the CE high tough capability. Before the CE can do this, network management messages are still needed to be redirected from FEs to the CE.

3.3 Netlink2

<Text for this section>

4. Model Requirements Compliance Evaluation

This section contains a review of each protocolÆs level of compliance to the ForCES model requirements. The ForCES model will indirectly relate to the protocol in that the protocol will be used to carry information that the model represents. Given that the model requirements are only indirectly related to the protocol selection, the review below will consist of prose rather than specific levels of compliance as is used in the protocol section below.

4.1 FACT

The FACT protocol is logically separated into a base protocol and an extensible payload which can be used to carry the FE, Logical Functional Block (LFB) specific data which is defined by the FE Model [9]. Thus the FACT protocol is cleanly separated from the data model that it carries. The FE Model draft [9] defines the data model for the Forwarding Element and meets all the Model requirements.

FACTÆs Configure Request and Configure Response message types under the Capabilities Control message group provide a flexible way to configure the functionality of the FE according to the FE Model [9]. The specific parameters needed to assign functionalities and behaviors to the Logical Functional Blocks (LFBs) in the FEs are dictated by the FE Model.

Vendor Specific functions are supported by VS-Data request and VS-Data response messages in the Vendor Specific message group.

4.2 GRMP

GRMP protocol is designed to use ForCES FE model as a base data model for the protocol functionality. GRMP aims to support all operations to all elements defined in ForCES FE model. Because ForCES FE model work is still in progress, following elements for ForCES FE model (including capability model and state model) with their operations are presented in current version of GRMP document:
-FE capabilities
-FE attributes, including FE statistics
-FE events
-LFBs with their attributes (including capabilities, statistics, etc), their actions, and their topologies
-Datapaths
Section 5.1.2 has described GRMP supported management for modeled elements including all above ForCES FE model elements in details. Along with the progress in ForCES FE model work, a modification of GRMP can be made to coordinate with the modification in ForCES FE model.

GRMP supports ForCES FE model to meet following model requirements without any restriction from the protocol:

1. Types of logical functions
2. Variations of logical functions
3. Ordering of logical functions
4. Flexibility

5. Minimal set of logical functions

4.3 Netlink2

<Text for this section>

5. Protocol Requirements Compliance Evaluation

This section contains a review of each protocolÆs level of compliance to the ForCES protocol requirements. Given that the protocol requirements are directly related to the protocol proposals, a very concrete method is used in reviewing compliance - the following key identifies the level of compliance for each of the following protocols to each protocol requirement in the ForCES requirements RFC:

T = Total compliance. Meets the requirement fully.

P+ = Partial compliance. Fundamentally meets the requirement through the use of extensions (e.g. packages, additional parameters, etc.)

P = Partial compliance. Meets some aspect of the requirement, however, the necessary changes require more than an extension and/or are inconsistent with the design intent of the protocol.

N = Not compliant. Does not meet the requirement.

Each subsection of this section begins with the specific protocol requirement text found in the ForCES requirements.

5.1 Protocol Requirement: Configuration of Modeled Elements

The ForCES protocol MUST allow the CEs to determine the capabilities of each FE. These capabilities SHALL be expressed using the FE model whose requirements are defined in Section 6. Furthermore, the protocol MUST provide a means for the CEs to control all the FE capabilities that are discovered through the FE model. The protocol MUST be able to add/remove classification/action entries, set/delete parameters, query statistics, and register for and receive events.

5.1.1 FACT

Compliance = T
FACTÆs Capabilities Control message class contains Configure Request and Configure Response messages that can be used to configure the FEÆs behavior from the CE. Also, the Capability request and response messages can be used by the CE to query and learn the FE

capabilities. Please see section 5.2 in [2] for more details on this.

5.1.2 GRMP

Most of GRMP protocol messages are for the management of modeled elements in ForCES FEs. They are listed as follows:

1) FE capability query and response messages [3 section 4.1.4]
The messages allow a CE to query and get response of the capabilities of a FE. The FE capabilities include all FE capability types that are defined by vendors or this GRMP protocol itself, as well as defined in FE model.

2) FE attribute manipulate message [3 section 4.1.6]
This message allows a CE to manipulate (add, delete, modify) attributes of a FE. The FE attributes should be the type of FE attributes that are allowed to be manipulated. They may be defined in FE model, by vendors or by GRMP itself.

3) FE attribute query and response messages [3 section 4.1.7]
The messages allow a CE to query and get response of FE attribute values. The FE attributes should be the types of FE attributes that are allowed to be queried. The queried FE attribute may be defined in FE model, by vendors or by GRMP itself. Note that FE attributes include FE level statistics.

4) FE event report message [3 section 4.1.8]
This message allows a FE to report events to a CE. The FE events include all events that are defined in FE model, by vendors and by GRMP itself. Some FE events may need to be registered by a CE before they are willing to send reports to the CE, but the others may need not. An FE attribute defined by GRMP itself (as a GRMP class object) is used for a CE to register its interested FE events [3 Section 4.1.6-Page29].

5) LFB action manipulate message [3 section 4.2.1].
This message allows a CE to manipulate LFB actions in a FE (LFB add, delete, modify, up, down, active, inactive, etc). The LFBs may be that defined in FE model, by vendors or by GRMP itself. Note that the LFB action manipulate message for LFB add operation has also included the operation to configure LFB topologies. In GRMP protocol, LFB topology is represented by means of PkfIDs [3 Section 3.4.6] [3 Section 4.2.1].

6) LFB topology query and response messages [3 section 4.2.2]
The messages allow a CE to query and get response of whole or some LFB topologies within a FE. The LFB topology representation is based on PkfIDs.

7) LFB attribute manipulate message [3 section 4.2.3].

This message allows a CE to manipulate (add, delete, or modify) LFB attributes in a FE. The LFB attributes include all attributes that are allowed to be manipulated in a LFB, and they may be defined via FE model, by vendors or by GRMP itself.

8) LFB attribute query and response messages [3 section 4.2.4]
The messages allow a CE to query and get response of LFB attribute values in a FE. Note that LFB attributes include LFB level capabilities, statistics, etc.

9) Datapath Manipulate Message [3 section 4.3.1]
This message allows a CE to manipulate (add, delete, modify) datapaths that interconnect LFBs in a FE. Datapaths are represented by PkfIDs.

10) Datapath query and response messages [3 section 4.3.2]
The messages allow a CE to query and get response of connection status of all or some datapaths in a LFB topology in a FE.

Protocol requirement compliance level: ( )

5.1.3 Netlink2

<Text for this section>

5.2 Protocol Requirement: Support for Secure Communication

a) FE configuration will contain information critical to the functioning of a network (e.g. IP Forwarding Tables). As such, it MUST be possible to ensure the integrity of all ForCES protocol messages and protect against man-in-the-middle attacks.
b) FE configuration information may also contain information derived from business relationships (e.g. service level agreements). Because of the confidential nature of the information, it MUST be possible to secure (make private) all ForCES protocol messages.
c) In order to ensure that authorized CEs and FEs are participating in a NE and defend against CE or FE impersonation attacks, the ForCES architecture MUST select a means of authentication for CEs and FEs.
d) In some deployments ForCES is expected to be deployed between CEs and FEs connected to each other inside a box over a backplane, where physical security of the box ensures that man-in-the-middle, snooping, and impersonation attacks are not possible. In such scenarios the ForCES architecture MAY rely on the physical security of the box to defend against these attacks and protocol mechanisms May be turned off.
e) In the case when CEs and FEs are connected over a network,

security mechanisms MUST be specified or selected that protect the

ForCES protocol against such attacks. Any security solution used for ForCES MUST specify how it deals with such attacks.

5.2.1 FACT

Compliance = T
FACT uses TLS when its endpoints are running over an IP network or in an insecure environment. For a closed box or physically secure environment, it is possible to turn off the protocol security functions. The security association between the CEs and FEs is established before any FACT association establishment messages are exchanged. Also, FACT recommends using rate limiting mechanisms on the FE to protect against DoS attacks. Please see section 8 in [2] for more details on this.

5.2.2 GRMP

1) When GRMP messages are encapsulated in a IP based medium, GRMP protocol has recommended to use IPsec or TLS(only for reliable transport protocols), which is also recommended by ForCES framework, to secure the communication between CEs and FEs to defend against possible man-in-the-middle or replay attacks and to do authentication for CEs and FEs. GRMP has no restrictions on using other approaches for secure communications.

2) When GRMP messages are transported via bus backplanes, the secure mechanism is allowed to be turned off while without affecting GRMP functions.

3) Current version of GRMP has not yet recommended secure mechanisms for GRMP messages to transmit over Ethernet or ATM mediums.

Protocol requirement compliance level: ( )

5.2.3 Netlink2

<Text for this section>

5.3 Protocol Requirement: Scalability

The ForCES protocol MUST be capable of supporting (i.e., must scale to) at least hundreds of FEs and tens of thousands of ports. For example, the ForCES protocol field sizes corresponding to FE or port numbers SHALL be large enough to support the minimum required numbers. This requirement does not relate to the performance of a NE as the number of FEs or ports in the NE grows.

5.3.1 FACT

Compliance = T
FACT can support up to 64K FEs and 64K CEs at the same time due its 16 bit addressing range of both the CE-Tag and FE-Identifier fields. Please see section 4.1 in [2] for more details on this. In addition, it uses TCP (for IP interconnection between CEs and FEs) which provides congestion control and thus helps in supporting the scalability requirement.

5.3.2 GRMP

In GRMP, a FE is identified by a 16 bits FE Identifier [3 section 3.2], which is theoretically able to identify up to 64k FEs.

Possible limitation in GRMP protocol to FE port number may be from FE port address space, maximum number of list elements in "list data format" [3 section 3.4.3], and LFB instance identifier space. The evaluations of scalability for them are as follows:

1) An Addressable Entity (AE) address data format is defined in GRMP [3 section 3.4.4], which is theoretically capable of describing any length of addresses of AEs, therefore FE port address space is not limited.

2) Element number of a list in "list data format" [3 section 3.4.3] is expressed with 16 bits data space, which theoretically limits list element number within 64k.

3) LFB instance ID [3 section 4.2.1] is expressed using 16 bits data space, which can also theoretically represent 64k instances of one kind of LFB such as a port LFB.

Protocol requirement compliance level: ( )

5.3.3 Netlink2

<Text for this section>

5.4
Protocol Requirement: Multihop

When the CEs and FEs are separated beyond a single hop, the ForCES protocol will make use of an existing RFC2914 compliant L4 protocol with adequate reliability, security and congestion control (e.g. TCP, SCTP) for transport purposes.

5.4.1 FACT

Compliance = T

FACT uses TCP as the transport protocol which is congestion aware and meets the transport requirements for multi-hop IP networks. Please see section 3.2 in [2] for more details on this.

5.4.2 GRMP

GRMP is designed in its aims to be capable of supporting remote control that allows CEs and FEs to separate multihops away, as well as supporting close or very close proximity control of CEs and FEs. GRMP has no restriction for ForCES NE administrators to use any RFC 2914 compliant L4 protocols such as TCP or SCTP as interconnection protocols to increase the control reliability, security and congestion control ability, though current document of GRMP has missed making a recommendation on this. Besides, GRMP is seeking the possibility to potentially support L3 layer QoS based traffic control between CEs and FEs with the use of scheduling mechanisms in GRMP slave module [3 section 4.6.1].

Protocol requirement compliance level: ( )

5.4.3 Netlink2

<Text for this section>

5.5
Protocol Requirement: Message Priority

The ForCES protocol MUST provide a means to express the protocol message priorities.

5.5.1 FACT

Compliance = T
FACT supports up to 8 levels of priority using the 3 priority bits in the common header. Please see section 4.1.6 in [2] for more details on this.

5.5.2 GRMP

GRMP defines a priority field (P) at GRMP message header [3 section 3.2] to express the protocol message priority. Currently only two priority levels are defined: normal priority and high priority.

Protocol requirement compliance level: ( )

5.5.3 Netlink2

<Text for this section>

5.6
Protocol Requirement: Reliability

a) The ForCES protocol will be used to transport information that requires varying levels of reliability. By strict or robust reliability in this requirement we mean, no losses, no corruption, no re-ordering of information being transported and delivery in a timely fashion.
b) Some information or payloads, such as redirected packets or packet sampling, may not require robust reliability (can tolerate some degree of losses). For information of this sort, ForCES MUST NOT be restricted to strict reliability.
c) Payloads such as configuration information, e.g. ACLs, FIB entries, or FE capability information (described in section 7, (1)) are mission critical and must be delivered in a robust reliable fashion. Thus, for information of this sort, ForCES MUST either provide built-in protocol mechanisms or use a reliable transport protocol for achieving robust/strict reliability.
d) Some information or payloads, such as heartbeat packets that may be used to detect loss of association between CE and FEs (see section 7, (8)), may prefer timeliness over reliable delivery. For information of this sort, ForCES MUST NOT be restricted to strict reliability.
e) When ForCES is carried over multi-hop IP networks, it is a requirement that ForCES MUST use a RFC 2914 [11]-compliant transport protocol.
f) In cases where ForCES is not running over an IP network such as an Ethernet or cell fabric between CE and FE, then reliability still MUST be provided when carrying critical information of the types specified in (c) above, either by the underlying link/network/transport layers or by built-in protocol mechanisms.

5.6.1 FACT

Compliance = T
FACT uses a reliable transport protocol to meet all the reliability requirements. For IP-interconnection between the protocol elements, FACT uses TCP as the transport protocol for the control channel. Please see section 3.2 in [2] for more details on this.

5.6.2 GRMP

GRMP supplies two levels of built-in error control mechanisms and several other mechanisms to improve the protocol reliability:

1) Normal level error control
In this level, GRMP protocol uses a specific GRMP ACK message [3 Section 3.4.1] associated with "Result" and "Code" fields in GRMP

message headers [3 Section 3.2] to protect against errors that may result from message transmission, message processing, or message

generation. The "Result" field can be set to "NoAck", "NoSuccessAck", "AckAll", and "SuccessAck" [3 Section 3.2] to ask the message receiver to send or not to send ACK message. According to the different importance and requirement of individual GRMP messages, recommendations have been made in GRMP for their values of ôResultö in the message header. As an example, GRMP packet redirection messages have been recommended to use "NoAck" value.

2) High level error control
If higher level of reliability is required for some protocol messages, a built-in error control based on CRC-32 checksums can furthermore be applied [3 Section 3.2]. A tag in GRMP message header is used to optionally turn on or turn off the checksum mechanism. Note that checksum error control can only improve the protocol transmission reliability, therefore it can well fit for the case when GRMP protocol is running over a comparatively unreliable medium such as Ethernet or backplane where error control may not be supplied by the medium itself.

3) Transaction identifier to control the order of messages
GRMP has defined different transaction identifiers for CE generated messages and for FE generated messages [3 Section 3.2]. This makes it possible to use protocol built-in method to order back protocol messages if in occasional cases messages are reordered.

4) QoS control of message transmission
A scheduler is applied in GRMP slave model, which is not only for protection against DoS attacks but also able to supply some sorts of QoS controls such as bandwidth and priorities for GRMP message transmission so as to improve the protocol reliability regarding the timeliness of transmission. This is especially applicable when GRMP is applied in a single hop scenario.

GRMP has no restriction on the use of any L4 layer protocols to improve the protocol reliability.

Protocol requirement compliance level: ( )

5.6.3 Netlink2

<Text for this section>

5.7 Protocol Requirement: Interconnect Independence

The ForCES protocol MUST support a variety of interconnect technologies. (refer to section 5, requirement# 1)

5.7.1 FACT

Compliance = T
FACT uses interconnect independent addressing (FE Identifier, CE tag) in its common header to provide interconnect independence. For non-IP interconnects, such as ATM, an interconnect specific encapsulation will have to be defined to carry the FACT messages. For IP interconnects, FACT uses TCP as the transport protocol. Please see section 3.1 in [2] for more details on this.

5.7.2 GRMP

GRMP packets can be transported via any suitable mediums, such as TCP/IP, Ethernet, ATM fabrics, and bus backplanes. Because of the design consideration for GRMP to be compatible with GSMP protocol, packet encapsulations defined for GSMP protocols as in RFC 3293 can also be applied to GRMP.

Protocol requirement compliance level: ( )

5.7.3 Netlink2

<Text for this section>

5.8
Protocol Requirement: CE Redundancy or CE Failover

The ForCES protocol MUST support mechanisms for CE redundancy or CE failover. This includes the ability for CEs and FEs to determine when there is a loss of association between them, ability to restore association and efficient state (re)synchronization mechanisms. This also includes the ability to preset the actions an FE will take in reaction to loss of association to its CE e.g., whether the FE will continue to forward packets or whether it will halt operations. (refer to section 5, requirement# 7)

5.8.1 FACT

Compliance = T
FACT exchanges CE and FE element states using the PE State Maintenance messages. FACT also uses Heart-Beat messages (section 5.3 in [2]) to detect protocol element (CE or FE) failure or loss of association between elements and to trigger a switch-over to a functioning redundant element (CE or FE). Please see section 7.3 in [2] for more details on the different mechanisms (Strong consistency, weak consistency) used for CE failover.

5.8.2 GRMP

GRMP meets ForCES CE redundancy or CE failover requirement by means of following mechanisms:

1) CE failover or leave policy [3 Section 4.6.4]
This policy is defined as a FE attribute and is setup via FE attribute manipulate message [3 Section 4.1.5]. The policy is usually set to a FE immediately after the FE is added to a ForCES NE and before it actually begins to provide IP service. In this policy attribute, selectable FE policies toward the CE failover are defined, which include graceful restart, going inactive, etc. In addition, CE re-association policies such as just waiting or trying to find out an alternative CE among a CE list are simultaneously defined.

2) FE heartbeat policy [3 Section 4.6.6]
The ability to determine the loss of association between a CE and a FE can be achieved in GRMP by use of this FE heartbeat policy, which is also defined as a GRMP class FE attribute. In this policy attribute, policies for the FE to receive heartbeats from a CE and to send heartbeats to a CE are individually defined. After the heartbeat policy attribute is set, the FE can then optionally send heartbeat signals to a CE or receive and process heartbeat signals from a CE. Heartbeat signals are sent by use of GRMP FE or CE event report messages.

Protocol requirement compliance level: ( )

5.8.3 Netlink2

<Text for this section>

5.9 Protocol Requirement: Packet Redirection/Mirroring

a) The ForCES protocol MUST define a way to redirect packets from the FE to the CE and vice-versa. Packet redirection terminates any further processing of the redirected packet at the FE.
b) The ForCES protocol MUST define a way to mirror packets from the FE to the CE. Mirroring allows the packet duplicated by the FE at the mirroring point to be sent to the CE while the original packet continues to be processed by the FE.
Examples of packets that may be redirected or mirrored include control packets (such as RIP, OSPF messages) addressed to the interfaces or any other relevant packets (such as those with Router Alert Option set). The ForCES protocol MUST also define a way for the CE to configure the behavior of a) and b) (above), to specify which packets are affected by each.

5.9.1 FACT

Compliance = T
FACTÆs Traffic Maintenance Message class includes Control Packet Redirect and Control Packet Forward messages to achieve packet

redirection/mirroring. These messages are sent over the separate data

channel. Please see section 5.4 in [2] for more details on this. Also, the Event Register/Deregister messages (section 5.5 in [2]) can be used to specify which packets should be redirected.

5.9.2 GRMP

GRMP supports packet redirection by packet redirection messages [3 Section 4.7]. A LFB within LFB topology in a FE should be used to filter out packets that are to be redirected. Packets to be redirected are first put in GRMP slave [3 Section 4.6.1] and then be directed to a CE via GRMP packet redirection message. The attribute of this filter LFB are set by CEs, therefore the CE has the ability to control which packets can be redirected. For example, CE may want to filter out packets that are considered from DoS attackers.

To redirect packets from CE to FE, CE just needs to encapsulate the packet to be redirected in the packet redirection message and send it to the FE. On the FE side, another LFB is used to resolve redirected packets from CE and to put the packets into datapaths in a FE LFB topology so that they can further be delivered by the FE.

By use of the packet redirection message and by properly configuring LFBs in FE, a packet can be mirrored to CE instead of purely redirected to CE. That is, the packet is duplicated and one is redirected to CE and the other continues its way in the LFB topology.

Protocol requirement compliance level: ( )

5.9.3 Netlink2

<Text for this section>

5.10 Protocol Requirement: Topology Exchange

The ForCES protocol MUST allow the FEs to provide their topology information (topology by which the FEs in the NE are connected) to the CE(s). (refer to section 5, requirement# 10)

5.10.1 FACT

Compliance = T
FACTÆs Capabilities and Control Message class includes Topology request and response messages to achieve topology information exchange between the CE and FEs. Please see sections 5.2.5, 5.2.6 in [2] for more details on this.

5.10.2 GRMP

In GRMP, FE topology query and response messages [3 Section 4.1.3] are used for CEs to query FE topology information in the NE.

Protocol requirement compliance level: ( )

5.10.3 Netlink2

<Text for this section>

5.11
Protocol Requirement: Dynamic Association

The ForCES protocol MUST allow CEs and FEs to join and leave a NE dynamically. (refer to section 5, requirement# 12)

5.11.1 FACT

Compliance = T
FACTÆs Connection and Association message class includes Join request, Join response, Leave request and Leave response messages to enable dynamic joining and leaving of protocol elements (CEs, FEs) in the NE. Please see sections 5.1.1, 5.1.2, 5.1.3, 5.1.4 in [2] for more details on this.

5.11.2 GRMP

In GRMP, specific FE join request message [3 Section 4.1.1] and FE leave request message [3 Section 4.1.2] make FEs able to dynamically join or leave a ForCES NE. While CE failover or leave policy [3 Section 4.6.4] defines the way for CEs to dynamically join or leave the NE. GRMP also defines FE failover and rejoin policy [3 Section 4.6.5] for FEs to dynamically rejoin the NE.

Protocol requirement compliance level: ( )

5.11.3 Netlink2

<Text for this section>

5.12
Protocol Requirement: Command Bundling

The ForCES protocol MUST be able to group an ordered set of commands to a FE. Each such group of commands SHOULD be sent to the FE in as few messages as possible. Furthermore, the protocol MUST support the ability to specify if a command group MUST have all-or-nothing semantics.

5.12.1 FACT

Compliance = T

FACT supports command bundling by using multiple TLVs in its message payload. For example, each TLV used in the Configure Request message could represent a different command such as Add, Delete, etc. In addition, FACT also supports 2-phase commit operations. Please see sections 5.2.3, 4.2 in [2] for more details on this.

5.12.2 GRMP

In many cases of IP services, timely execution of ForCES protocol commands are essential for properly providing the services. Command bundling is a good approach to this. GRMP supports ForCES protocol command bundling in two ways:

1) Using GRMP batch messages [3 Section 4.8]
This kind of GRMP messages allow GRMP application layers to pack several different GRMP message bodies into one single GRMP message. These messages should meet following conditions:
-Are sent from the same CE to the same FE.
-Do not need explicit message responses.
Such messages include that like event report messages, FE or LFB action manipulate messages, attribute manipulate messages, etc.

2) Using list data format [3 Section 3.4.3]
The list data format is used in several GRMP messages so that these messages can set more than one commands (that have same command type) in one message body. Examples of these GRMP messages are like:
-FE attribute manipulate message [3 Section 4.1.6], which allow CE to manipulate several FE attributes at one blow.
-FE attribute query and response messages [3 Section 4.1.7], which allow CE to query several FE attributes at one blow.
-FE event report message [3 Section 4.1.8], which allow FE to report several FE events via one message.
-LFB management messages [3 Section 4.2]
-Datapath management messages [3 Section 4.3]

Protocol requirement compliance level: ( )

5.12.3 Netlink2

<Text for this section>

5.13 Protocol Requirement: Asynchronous Event Notification

The ForCES protocol MUST be able to asynchronously notify the CE of events on the FE such as failures or change in available resources or capabilities. (refer to section 5, requirement# 6)

5.13.1 FACT

Compliance = T
FACTÆs Event Notification message class includes the Asynchronous FE Event notification message used to report asynchronous FE events to the CE. Please see section 5.5 in [2] for more details on this.

5.13.2 GRMP

In GRMP, a FE asynchronously informs CEs of a monitored failure, resources and capabilities changes, and other asynchronous events via GRMP FE event report message [3 Section 4.1.8]. These events include all that are defined within GRMP scope, by ForCES FE model, and possibly by vendors. GRMP use an object class identifier [3 Section 3.4.5] to describe such inclusion. Current document of GRMP has defined following asynchronous events, which belong to GRMP event class:

Object Class = 0 (GRMP class)
FE Event Type
-FE status event such as FE up, down, etc.
-LFB status event such as LFB up, down, etc.
-FE heartbeat
-FE port change, such as port up, down, etc.
-FE memory change
-FE DoS attack alert

Protocol requirement compliance level: ( )

5.13.3 Netlink2

<Text for this section>

5.14 Protocol Requirement: Query Statistics

The ForCES protocol MUST provide a means for the CE to be able to query statistics (monitor performance) from the FE.

5.14.1 FACT

Compliance = T
FACTÆs Capabilities and Control message class includes the Query request and response messages which can be used by the CE for querying the FEÆs properties and statistics. Please see sections 5.2.7, 5.2.8 in [2] for more details on this.

5.14.2 GRMP

GRMP defines statistics regarding FE performance as FE or LFB attributes. The FE attributes of statistics are the statistics that

take whole FE as a statistic object, and the LFB attributes of

statistics usually take the individual LFBs as statistic objects. In GRMP, the statistics attributes include that defined in FE model, by vendors, and by GRMP protocol itself. GRMP uses FE attribute query and response messages [3 Section 4.1.7] and LFB attribute query and response messages [3 Section 4.2.4] to query the statistics.

GRMP can also support query of statistics defined by network management tools like SNMP by using MO get message [3 Section 4.5.1] and MO response message [3 Section 4.5.3].

Protocol requirement compliance level: ( )

5.14.3 Netlink2

<Text for this section>

5.15 Protocol Requirement: Protection Against Denial of Service Attacks

Systems utilizing the ForCES protocol can be attacked using denial of service attacks based on CPU overload or queue overflow. The ForCES protocol could be exploited by such attacks to cause the CE to become unable to control the FE or appropriately communicate with other routers and systems. The ForCES protocol MUST therefore provide mechanisms for controlling FE capabilities that can be used to protect against such attacks. FE capabilities that MUST be manipulated via ForCES include the ability to install classifiers and filters to detect and drop attack packets, as well as to be able to install rate limiters that limit the rate of packets which appear to be valid but may be part of an attack (e.g. bogus BGP packets).

5.15.1 FACT

Compliance = T
FACT uses separate control and data channels to provide robustness in the protocol against Denial of Service (DoS) attacks. Please see section 3.3 in [2] for more details on this. Also, the Configure Request and Response messages in FACT could be used to install filters on FEs which can be used for rate-limiting the malicious traffic.

5.15.2 GRMP

GRMP supports protection against DoS attacks by means of following defined mechanisms:

1) A model for GRMP slave module [3 Section 4.6.1]
In this model, all GRMP messages sending to CE are put to two different channels: the data channel, which is only for packet

redirection messages, and the control channel, which is for other

GRMP messages that are usually generated by GRMP slave itself. Note that before redirected packets enter GRMP slave, a filter LFB defined by FE model is usually applied to decide what kind of packets are allowed to be redirected to CE. Messages on the two channels pass through a packet scheduler, then are put on the link connecting to CE. The scheduler is managed by CE by setting some scheduling policies (disciplines) to it. This policy setting can be done either at the scheduler startup time or on the fly during its runtime. In this way, the CE can control the traffic over the two message transmission channels dynamically according to the monitored traffic status. This also provides a means for CE to protect control channel transmission and to defend against DoS attacks.

2) GRMP DoS protection policy [3 Section 4.6.2]
This is defined in GRMP as a FE attribute. In this policy attribute, scheduling priorities, channel bandwidths, and congestion control policies for the individual data channel and control channel can be assigned.

3) GRMP DoS attack alert policy [3 Section 4.6.3]
This is also defined as a FE attribute. This policy specifies in which state a DoS attack is considered happened. DoS attack monitoring is performed by monitoring the status and statistics of the scheduler in the GRMP slave model.

4) A DoS attack alert event report [3 Section 4.1.8]
This is an event report message sent from FE to CE to report that a DoS attack is monitored according to the preset DoS attack alert policy. The event report can also include some information concerning the attackers.

When CE has received the DoS attack alert event report, it may need to change DoS protection policy in some way to ensure the control channel transport. CE can also change attributes of the filter LFB put at the data channel entrance so that the packets that are doubted from DoS attackers can be filtered.

Protocol requirement compliance level: ( )

5.15.3 Netlink2

<Text for this section>

5.16 Protocol Requirement Summary Table

This section is a summary of the compliance levels claimed for each protocol above and is included as a convenience.

Protocol Requirement	FACT	GRMP	Netlink2
1. Configuration of Modeled Elements	T	?	?
2. Support for Secure Communication	T	?	?
3. Scalability	T	?	?
4. Multihop	T	?	?
5. Message Priority	T	?	?
6. Reliability	T	?	?
7. Interconnect Independence	T	?	?
8. CE Redundancy or CE Failover	T	?	?
9. Packet Redirection/Mirroring	T	?	?
10. Topology Exchange	T	?	?
11. Dynamic Association	T	?	?
12. Command Bundling	T	?	?
13. Asynchronous Event Notification	T	?	?
14. Query Statistics	T	?	?
15. Protection Against Denial of Service Attacks	T	?	?

Security Considerations

This document is a comparison between three protocols in order to help in the selection of the best approach to use as the ForCES protocol. Security considerations are addressed in each of the protocol proposals and MUST be included as part of the fitness evaluation for each proposal.

References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2 Audu, A. et al., "ForwArding and Control ElemenT protocol (FACT)", work in progress, September 2003, <draft-gopal-forces-fact-05.txt>

3 Wang, W. et al., "General Router Management Protocol (GRMP) Version 1ö, September 2003, <draft-wang-forces-grmp-00.txt>

4 Salim, J. H. et al., "Netlink2 as ForCES Protocol", work in progress, June 2003, <draft-jhsrha-forces-netlink2-01.txt>

5 Khosravi, H. et al., "Requirements for Separation of IP Control and Forwarding", work in progress, July 2003, <draft-ietf-forces-requirements-10.txt>

6 Yang, L. et al., "Forwarding and Control Element Separation (ForCES) Framework", work in progress, August 2003, <draft-ietf-forces-framework-08.txt>

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

8 Barnes, M., "Middlebox Communications (MIDCOM) Protocol Evaluation", work in progress, Nov 2002, <draft-ietf-midcom-protocol-eval-06.txt>

9 Yang, L. et al., "ForCES Forwarding Element Functional Model", work in progress, October 2003, <draft-ietf-forces-model-01.txt>

10 F. Baker, "Requirements for IP Version 4 Routers", RFC1812, June 1995.

11 S. Floyd, "Congestion Control Principles", RFC2914, September 2000.

Acknowledgments

Author's Addresses

Alex Audu
Alcatel R&I
1000 Coit Road
Plano, TX 75075
Phone: 1-972-477-7809
Email: alex.audu@alcatel.com

Hormuzd Khosravi
Intel
2111 NE 25th Avenue
Hillsboro, OR 97124
Phone: 1-503-264-0334
Email: hormuzd.m.khosravi@intel.com

David Putzolu
Intel
Mailstop JF3-206-H10
2111 NE 25th Avenue
Phone: 503-264-4510
Email: david.putzolu@intel.com

Weiming Wang
Department of Information and Electronic Engineering
Hangzhou University of Commerce
149 Jiaogong Road
Hangzhou, 310035, P.R.China
Phone: +86-571-88057712
Email: wangwm@hzcnc.com