



Advanced architecture for INTER-domain quality of service MONitoring, modelling and visualisation



INTERMON-IST-2001-34123

INTER-DOMAIN COMMUNICATION AND DATA EXCHANGE

Work-package No. / Title	WP3 / Inter-domain communication and data exchange
Date	19-09-2003
Distribution	INTERMON Consortium (Project Officer, Evaluators, Peer Reviewers)
Document Identifier	
File name	
Version	V 1.0
Editor	E. Boschi
Authors	E. Boschi, S. D'Antonio
Contact Person(s)	E. Boschi boschi@fokus.fraunhofer.de , S. D'Antonio salvatore.dantonio@napoli.consorzio-cini.it

Change History

V1.0	Initial version

Table of Contents

SUMMARY	4
1 INTRODUCTION.....	5
2 AGREEMENT BETWEEN ISPS	5
3 STARTING A PROCEDURE BETWEEN DIFFERENT DOMAINS	6
4 END-TO-END SCENARIO	8
4.1 How BGP-4 HELPS WITH FINDING THE DOWNSTREAM ISP	10
4.2 INTER-DOMAIN ROUTE CHANGES	10
5 REFERENCES.....	12

Summary

This document describes the Inter-domain communication within the InterMon system. We will discuss the type of data that will be exchanged and the communication protocol between the different domains. We will specify which components of the InterMon architecture are involved and how (which actions do they perform) describing the communication and data exchange protocols.

1 Introduction

Single ISPs usually monitor their network. They are interested in information on resource usage, traffic accounting, fault diagnosis and troubleshooting. The way this monitoring is done (e.g. the measurement framework), the metrics used (which are many) and the network characteristics are often very ISP specific.

From the users point of view end-to-end (E2E) performance is interesting. The problem is that you don't know how many providers you're going back and forth through and the intra-domain information a single ISP has would not be enough. ISPs should work through making their metrics accessible to other ISPs. They should cooperate. This way they could provide E2E information (either with E2E measurement or concatenating the measurements).

In order to achieve the interoperability between management systems three main conditions have to be satisfied:

1. It needs to be decided how to share data among ISPs (Communication Protocol)
2. The different ISPs have to agree upon the measurement framework (Different domains have different measurement tools) and on common measurements
3. The different ISPs have to agree on how to compare the data (as already said, many characteristics are very ISP specific)

The InterMon system offers a way to do this. It provides the ISPs joining it a common platform where data could be collected and shared using a common data format (IPFIX). We start with the assumption that different domains have different tools. The system makes it possible to exchange raw data and aggregated data depending on the user and the purpose, by using the common data format. Using the InterMon system, ISPs can collect, store, visualize, analyse and exchange data. The InterMon platform provides the interface towards the outside world (to the user layer and other domains).

Particularly important in the inter-domain scenario is a mechanism providing Authorization, Authentication, and Accounting with different capabilities and levels. An AAA component in the InterMon architecture is used for this.

In the remainder of this document we will try to give an answer to the following open questions in inter-domain monitoring:

1. Do methodologies need to be the same in different domains?
2. Which type of aggregation can be done?
3. How concatenate measurements from different domains to provide a view across different domains?
4. The data storage within a domain should be aggregated?

Question 3 is particularly important to us. We will describe our solution in chapter 3 and 4.

We will also try to define the following issues:

1. A common set of parameters to suggest to the tool developers.
2. A way to start a procedure between different domains.

Issue number 2 is discussed in detail in this document. In particular, section 2 describes how a data request message can be exchanged between domains using the AAA system for secure communication.

2 Agreement between ISPs

An ISP interested in monitoring data, related to a network domain managed by another ISP, needs to cooperate with it in order to obtain such data. This implies a negotiation process that results in an agreement between the two parties. A customer and a service provider sign such an agreement. The customer is the ISP that requests the monitoring data, whereas the service provider is the ISP that

manages the intra-domain QoS monitoring and provides the related results. In case of inter-domain QoS monitoring, the negotiation process is recursively triggered between each pair of involved ISPs.

The agreement can be considered as formed by two different parts: one containing contact information of the parties, and the other consisting of a list of fields to be filled in by the provider requesting the monitoring service.

This second part is called Specification of Monitoring Service and represents a sort of technical annex containing a set of parameters to be handled when performing the QoS monitoring. In other terms, this document is intended to provide a standard way of specifying the monitoring service. The technical parameters composing the contents of such document are the following:

- *Scope*, that uniquely identifies the geographical/topological region where the monitoring activity has to be performed by indicating the boundaries of that region; in other terms, such parameter specifies the network ingress and egress points for the traffic flow.
- *Flow Identification* indicating which traffic flow has to be monitored; it provides rules useful for the classification of the packets belonging to the traffic flow (source/destination network addresses, ports, transport protocol, etc.).
- *Time Schedule* specifying the start and the end of the monitoring task.
- *Metrics*, indicating what has to be measured. A set of QoS parameters upon which the two parties have to agree is specified in this field. Such parameters can be either metrics specifically defined or standard metrics, such as throughput, packet loss, one-way delay and jitter. Furthermore, the following attributes are associated with each metric:
 - *Notification threshold*, indicating a threshold that triggers a notification to the customer if the threshold is reached.
 - *Report schedule*, specifying when the measurement reports have to be sent to the customer.
- *Reporting Document Type*, describing which kind of documents has to be sent to the customer.
- *Reporting Destination Address*, indicating where the ISP, which is in charge of making the measures, has to send the reports (email, postal, fax, etc.).
- *Authentication data*, needed to identify and authenticate the requesting entity.
- *Encryption service*, in order to preserve the confidentiality of the information exchanged between the peering entities.

The definition of a standard format for the technical document facilitates the negotiation phase since the parties can appropriately process the information in an automated fashion.

Therefore, the introduction of a template, from which the agreement is prepared, enables the dynamic negotiation and delivery of the monitoring service.

3 Starting a procedure between different domains

A data exchange between different domains wouldn't be imaginable in the real life without an appropriate system for authentication and authorization. The Intermon architecture uses a AAA component to mediate every interaction with other domains. Figure 1 shows the InterMon architecture, highlighting in light blue the components that will take part in the Inter-domain communication.

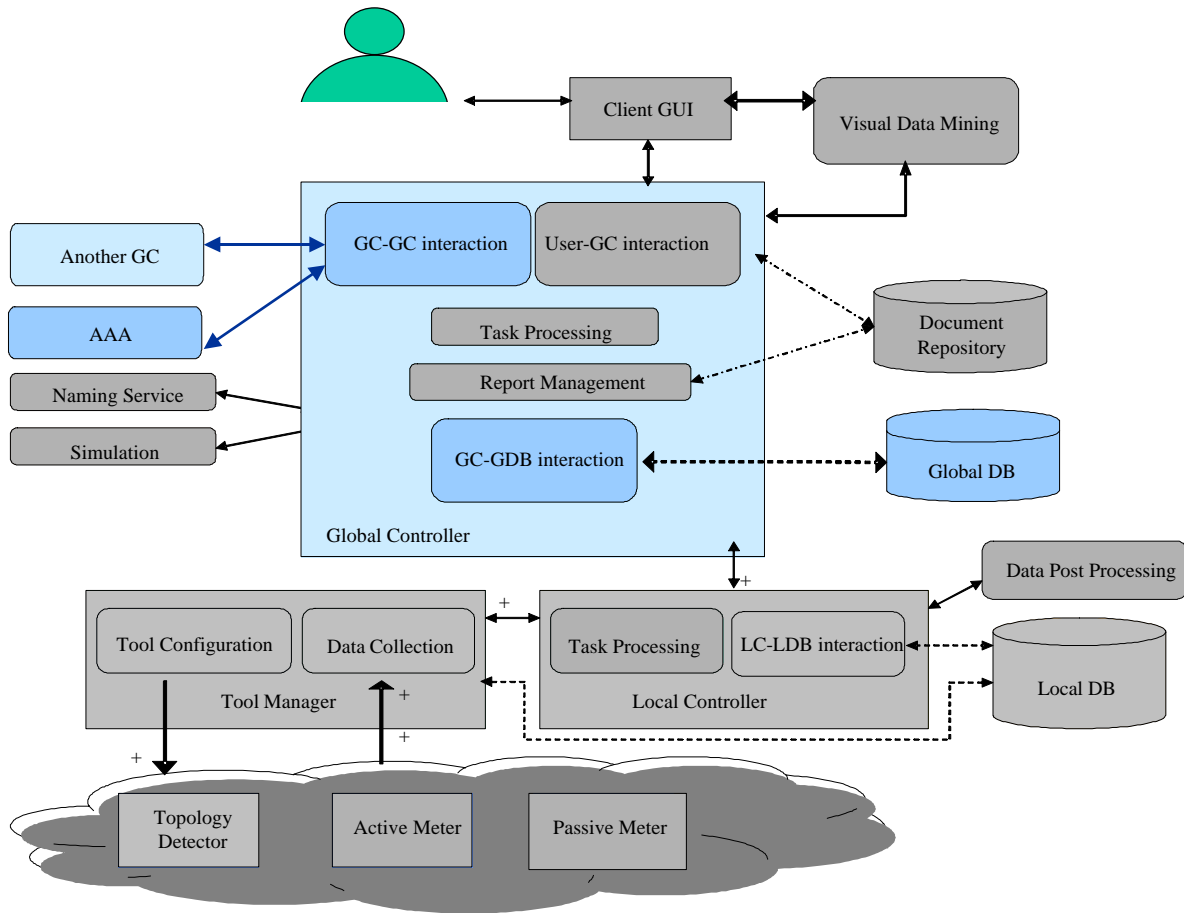


Figure 1. The InterMon architecture

The communication between the GC-GC Interaction components of two different ISPs goes through their local AAA servers. Each AAA server will store in a Table which ISP (or, more in detail, which user or group of that ISP) has the right to do/get what according to previously negotiated agreements. The possible “actions” are collection of data and configuration (e.g. ISP A wants ISP to measure something and configures a task using AAA).

Figure 2 shows how a procedure between different domains can be started. We refer here to the collection of measurement data. The ISP A requests data to the ISP B, sending a Data Request message where it is specified what kind of data are needed and the ISP A identifies itself (1). The Data Request message is initiated by the GC–GC interaction component of the ISP A. The request is sent from the local AAA component to the AAA component of ISP B. This components verifies together with the local GC-GC interaction component whether the ISP A is authorized to obtain the requested data and sends the answer (authorized, not authorized, error) back to the ISP A (2). This communication is done using the Diameter protocol. Only upon confirmation, the data are fetched from the global Database and sent, using an encrypted channel, to ISP A (3). The session key is provided by the AAA component after the transaction approval. Note that, in the Intermon architecture, the interface provided by the Database to the GC-GC Interaction component is the same one provided to the client GUI (Cf. Figure 1).

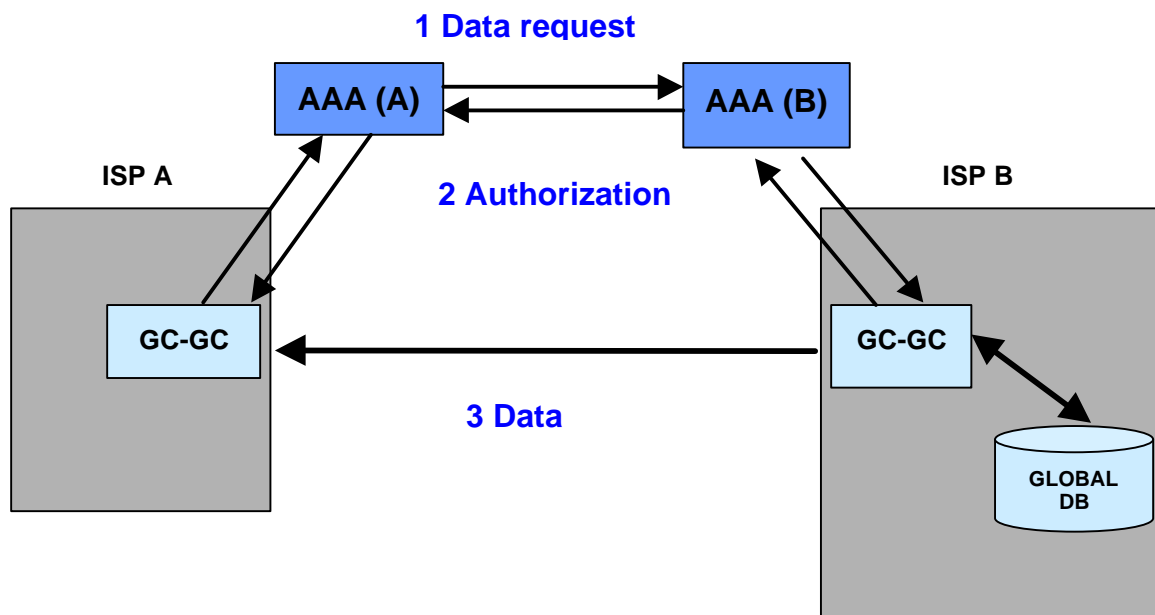


Figure 2. Data request and exchange

4 End-to-end scenario

In the INTERMON architecture the Global Controller (GC) is responsible for the management of monitoring activities within each Autonomous System (AS). In particular, upon reception of RMS (Request for Monitoring Service) document from the client GUI, the GC translates it in a standard document, called Specification of Monitoring Service (SMS). This document contains a set of

parameters specifying the QoS monitoring service. In case of SMS spanning over multiple domains, a certain number of GCs – those belonging to the crossed domains – are involved in fulfilling the monitoring request. More precisely, each GC in the chain is in charge of collecting the intra-domain measurement results and exporting them in order to monitor the QoS distribution in different networks. How is it possible to let multiple GCs cooperate in order to deploy such multi-domain service? An answer can be found if we think that all of the GCs are able to analyze SMS documents and nothing prevents a GC from being a network client of a neighboring GC. This consideration suggests the adoption of a cascade model where a multi-domain SMS can be recursively split in two parts: a single domain SMS plus a remaining part related to elsewhere, i.e. one or more downstream domains. In other terms, after receiving the SMS, the generic GC selects the subset of information related to its own AS and passes such a subset to the Local Controllers (LCs) that in turn translate it in appropriate configuration parameters to be forwarded to the Tool Manager. The rest of the SMS document is sent to the peer GC en-route towards the destination.

Once the measurement activities have been performed, in order to compute end-to-end performance each GC has to collect and concatenate data from both its LCs and the adjacent GC: data related to its own AS will be obtained by combining results provided by LCs, whereas those related to the remaining part of the path (if any) will be asked from the neighbouring GC.

Figure 3 shows an example of multi-domain SMS splitting.

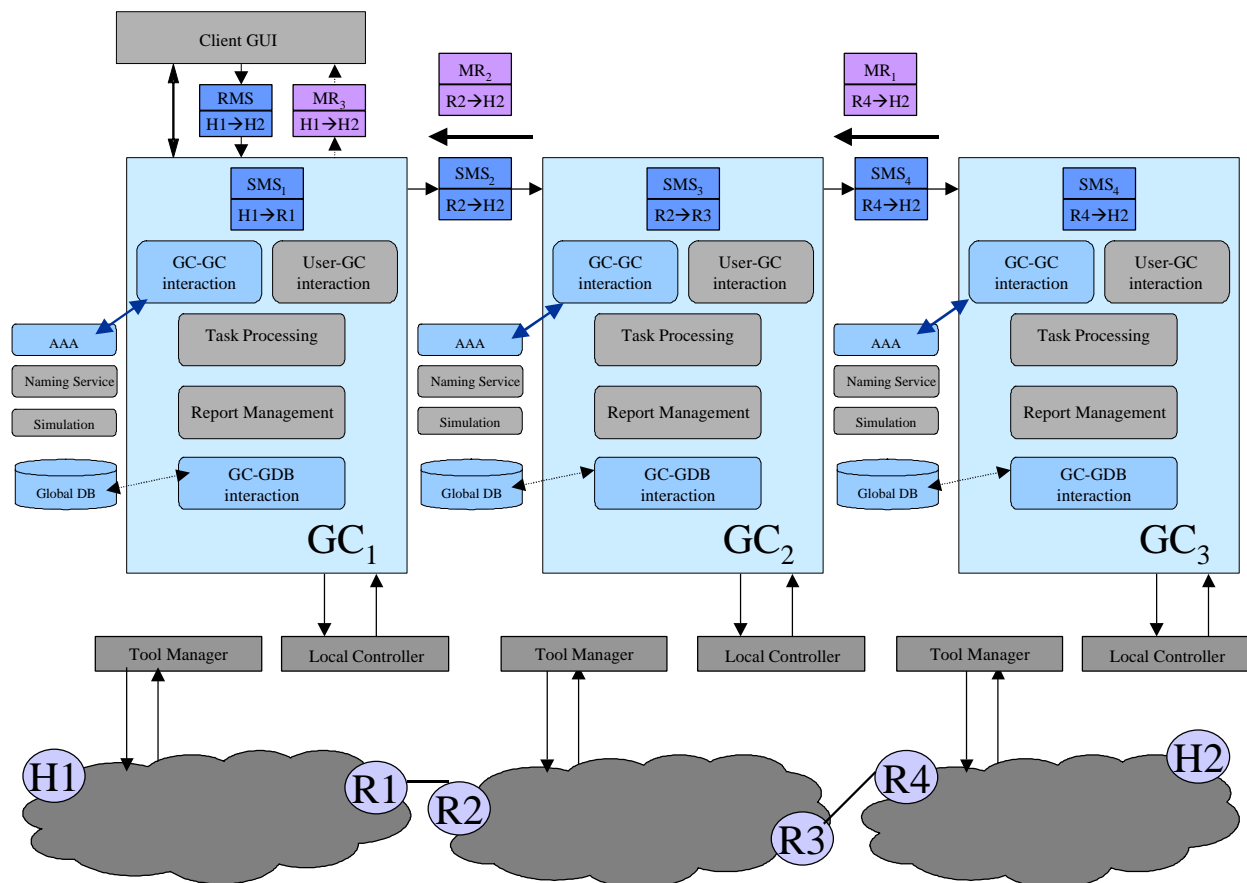


Figure 3. An example of multi-domain SMS splitting

Global Controller GC₁ receives a Request for Monitoring Service (RMS) from the Client GUI. In this example, RMS has as subject the traffic flow whose source and destination are, respectively, H1 and H2. GC₁ translates it in a SMS document and, afterwards, decomposes this document into two parts:

- The first one (i.e. SMS₁) refers to the Autonomous System managed by GC₁ itself;

- The second one (i.e. SMS₂) is passed to GC₂, which is the next Global Controller en route towards H2. SMS₂ specifies how to make measures within downstream ASs. Upon reception of SMS₂, GC₂ recursively executes the same actions as GC₁, by splitting the received request into two parts, called SMS₃ and SMS₄. Since the Scope field of SMS₄ refers to the last Autonomous System along the path, both SMS₃ and SMS₄ are intra-domain SMSs and they are handled by GC₂ and GC₃, respectively.

The process of exporting and composing monitoring results is triggered by GC₃ that produces an intra-domain monitoring report (i.e. MR₁) when the flow ends and forwards it to GC₂. Each Global Controller along the monitoring chain builds a report, by assembling local data with results received from the adjacent Global Controller. Finally, the Global Controller interacting with the Client GUI (i.e. GC₁) is in charge of composing the global report (i.e. MR₃) containing the end-to-end performance. Such a document is sent to the user.

4.1 How BGP-4 helps with finding the downstream ISP

BGP [1] is fundamentally a distance vector protocol but instead of maintaining just the cost to each destination, each BGP router keeps track of the exact path used and tells its neighbours the exact path it is using.

More precisely, A BGP route consists of the following:

- A destination, described as an IP address prefix.
- Information that describes the path to the destination, including the following:
- AS path, which is a list of numbers of the ASs that a route passes through to reach the local router. The first number in the path is that of the last AS in the path—the AS closest to the local router. The last number in the path is the AS farthest from the local router, which is generally the origin of the path.
- Path attributes, which contain additional information about the AS path that is used in routing policy.

In BGP, two systems form a transport protocol connection between one another. They exchange messages to open and confirm the connection parameters. The initial data flow is the entire BGP routing table. Incremental updates are sent as the routing tables change. BGP does not require periodic refresh of the entire BGP routing table. Therefore, a BGP speaker must retain the current version of the entire BGP routing tables of all of its peers for the duration of the connection. *KeepAlive* messages are sent periodically to ensure the liveness of the connection. Notification messages are sent in response to errors or special conditions. If a connection encounters an error condition, a notification message is sent and the connection is closed.

So, using BGP as inter domain routing protocol means having a way to deduce the path taken by the packets as the whole path is kept in the BGP routers routing tables. Those routers have the information necessary to identify the downstream ISP for each flow sent to a given destination.

4.2 Inter-domain route changes

Up to now we have assumed that all packets belonging to a single traffic flow cross the same Autonomous Systems for the entire flow duration. This implies that the path from source to destination taken by the packets does not change during the flow lifetime. GCs along this path are in charge of monitoring the traffic flow and managing the SMS splitting.

However, inter-domain routing changes are not rare events in the Internet, so it might happen that packets belonging to the same traffic flow take different paths towards the destination. In other terms,

after a routing change occurs, the traffic flow will cross a different set of ASs and, therefore, the flow monitoring will involve other Global Controllers. In this case, the inter-domain management of SMS documents and the composition of measurement results are much more complicated with respect to the case when no routing change occurs.

In order to better explain how SMSs splitting is performed in case of routing changes, let us consider the following example.

Source and destination of the flow to be monitored are, respectively, H1 and H2. Thanks to information provided by an inter-domain routing protocol (e.g. BGP), packets belonging to such flow cross the following Autonomous Systems from the source to the destination: AS1, AS2, AS3 and AS6 (see Figure 4). Therefore, Global Controllers responsible for monitoring activities inside such ASs participate to the SMS splitting process.

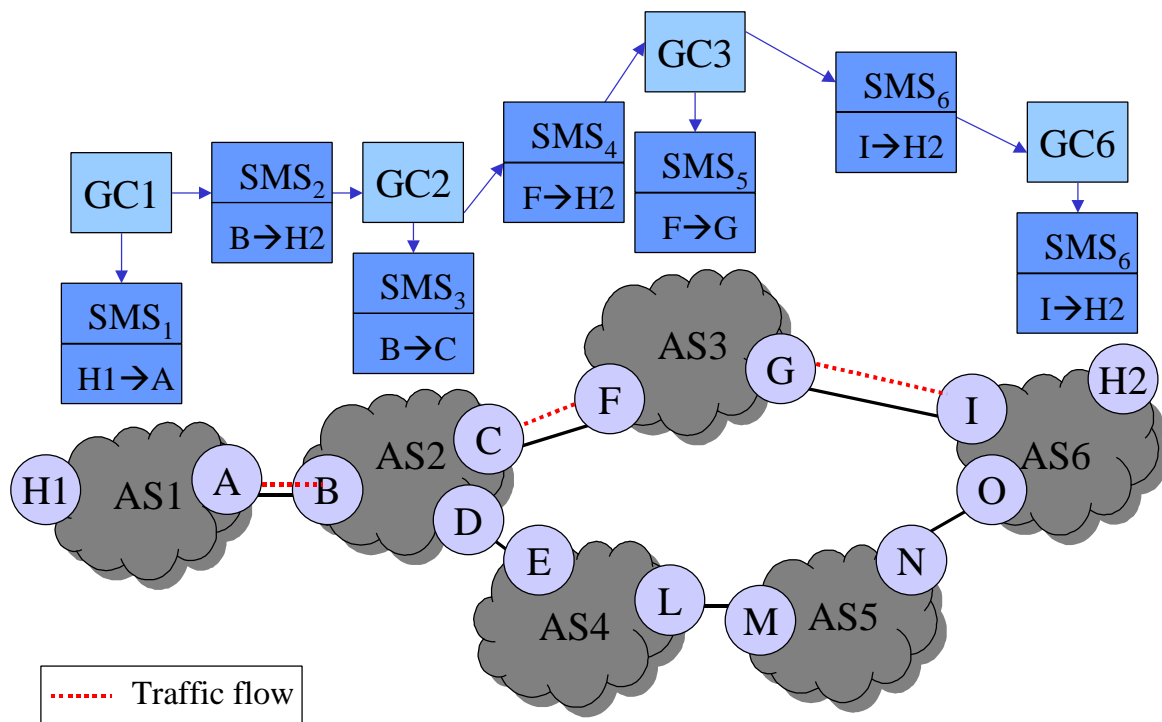


Figure 4. Inter-domain SMS splitting

Now, let us suppose that a link failure between AS3 and AS6 causes a routing change at time t_1 . After this event, Autonomous Systems crossed by the traffic flow are AS1, AS2, AS4, AS5, and AS6 (see Figure 5).

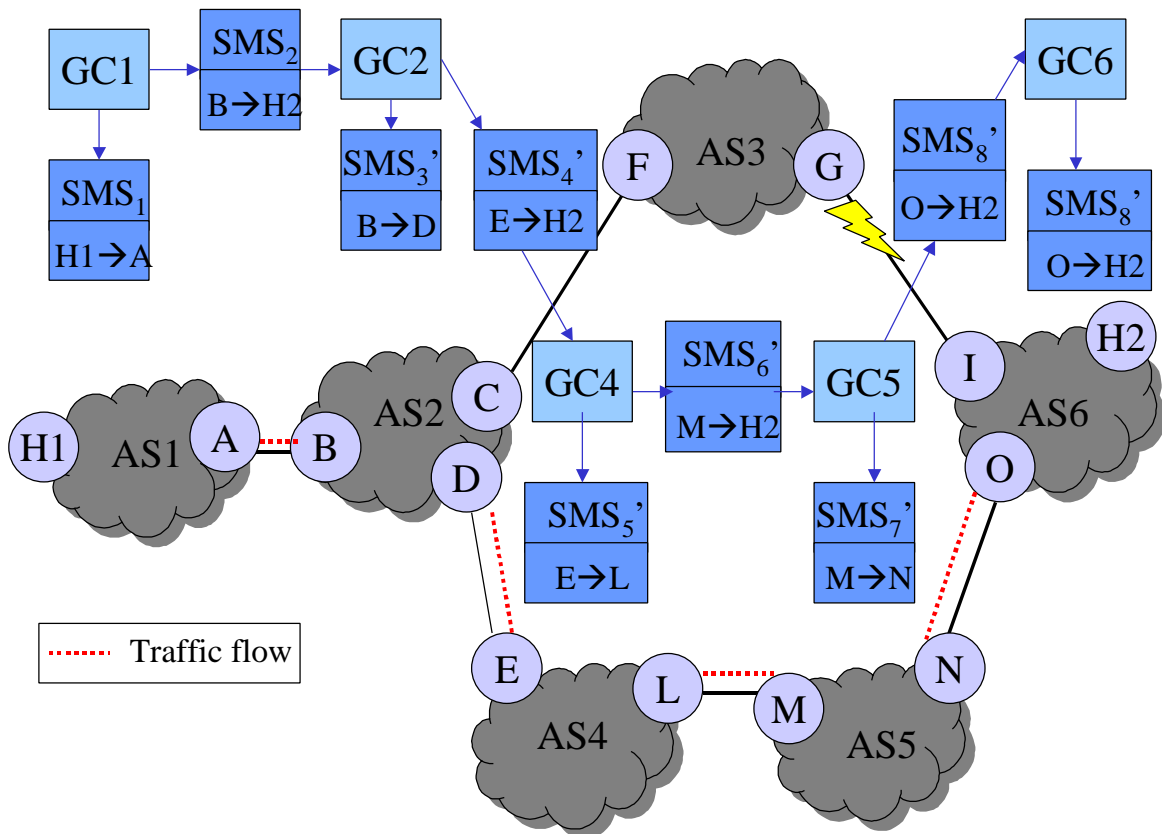


Figure 5. Effects on SMS splitting caused by a link failure

If t_0 and t_2 are, respectively, the start time and the end time of the flow, then we can distinguish two different scenarios:

1. before the routing change, the monitoring request spawns a number of SMS documents, whose time schedule is $[t_0, t_2]$
2. after the routing change, SMS related to AS1 is still valid. However, due to the link failure, packets arriving at AS2 are routed towards AS4, so GC2 is induced to split SMS₂ into two new SMS documents, called, respectively, SMS₃' and SMS₄'. SMS₃' refers to AS2 and its scope is B–D, whereas SMS₄' is sent to GC4 and its scope is E–H2. The time schedule of both documents is $[t_1, t_2]$.

Therefore, starting from t_1 on some modifications affect role and behaviour of GCs along the monitoring chain. More precisely,

- GC2 and GC6 keep on monitoring the traffic flow, for which the network ingress and egress point are changed;
- GC3 finishes monitoring activities related to SMS₃;
- GC4 and GC5 start performing intra-domain monitoring on the basis of specifications contained inside SMS₅' e SMS₇', respectively;

5 References

- [1] RFC 1654, A Border Gateway Protocol 4 (BGP-4)