

Network Monitoring Session Description

CoreGRID Workshop on Grid Middleware
Dresden
June 2007

Augusto Ciuffoletti – INFN – Italy

A. Papadogiannakis – Forth - Crete

M. Polychronakis – Forth – Crete

Summary

- **Describe a network monitoring framework**
- **Define a solution**
- **Introduce the schema describing network monitoring requests within such solution**

Narrowing the scope: GRID resources

- **A GRID is a collection of resources: among others, network paths**
- **A path is seen as an atomic resource: we do not distinguish links, routers etc.**
- **A path exists between any pair of non-network resources**
- **Non-network resources are partitioned into Domains**

Narrowing the scope: Network Monitoring

- Network Monitoring addresses network paths
- Network Monitoring is **performed** by specific resources, that we call Network Monitoring Elements
- Network Monitoring is **managed** by specific agents, that we call Network Monitoring Agents
- There is at least one Network Monitoring Agent for each Domain
- Each network path can be monitored by at least one Network Monitoring Element

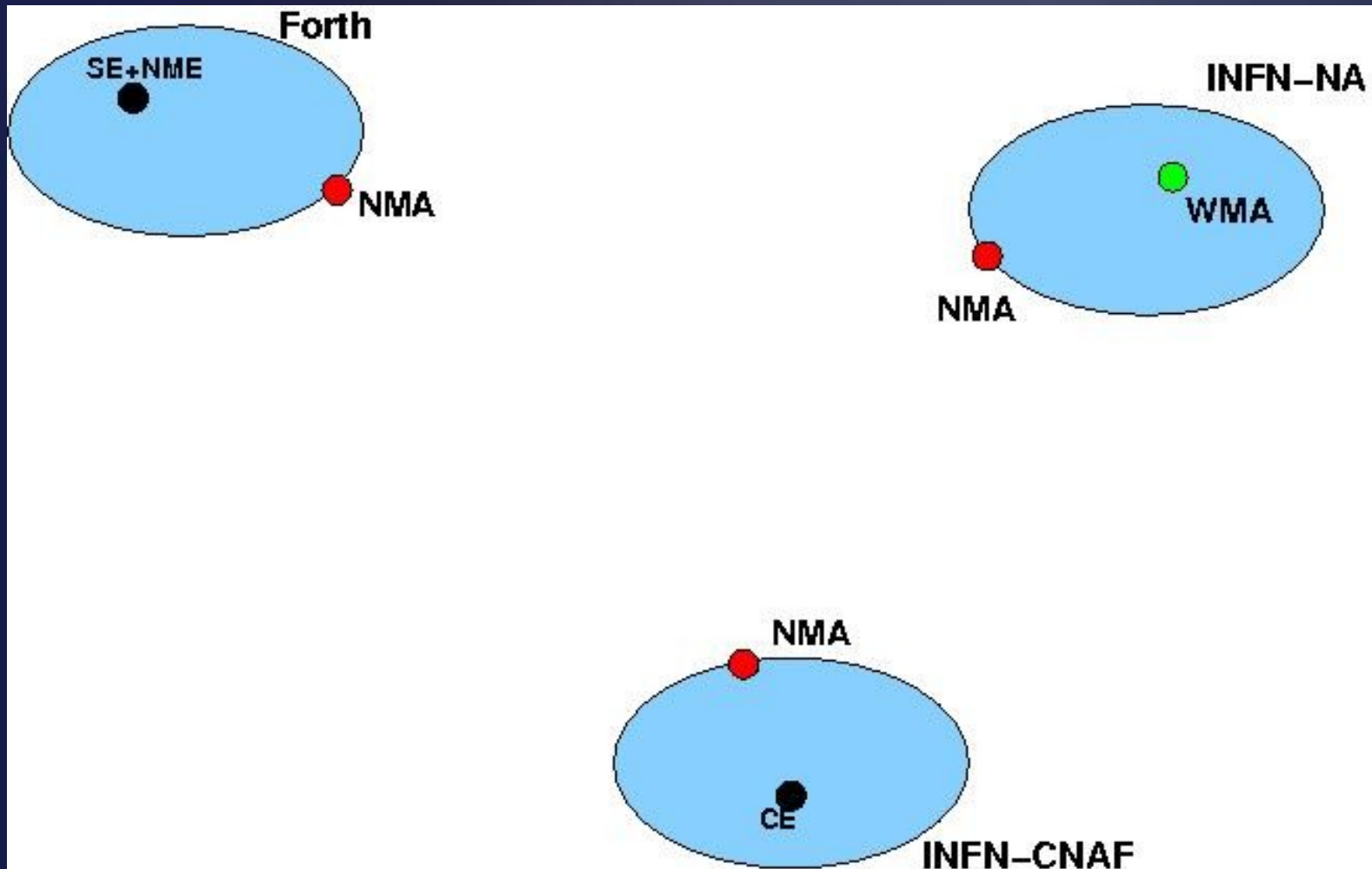
Narrowing the scope: NM management

- **Network Monitoring Agents form a community that cooperate peer-to-peer**
- **The scope of a Network Monitoring Agent is limited to a single Domain**
- **A Network Monitoring Agent has access to a directory describing the capabilities of the Network Monitoring Elements**
- **We envision the case of Network Monitoring Agents specialized in monitoring some remote domains**

Narrowing the scope: NM requests

- **Network Monitoring aims at the observation of the performance of the network resources allocated to a complex computational task, that we call Workflow**
- **NM requests are issued by the agent responsible for Workflow management**
- **NMA dynamically configure NMEs to operate measurements**
- **NMEs returns a stream of observations to the Workflow Management Agent**

The Big Picture (small example)



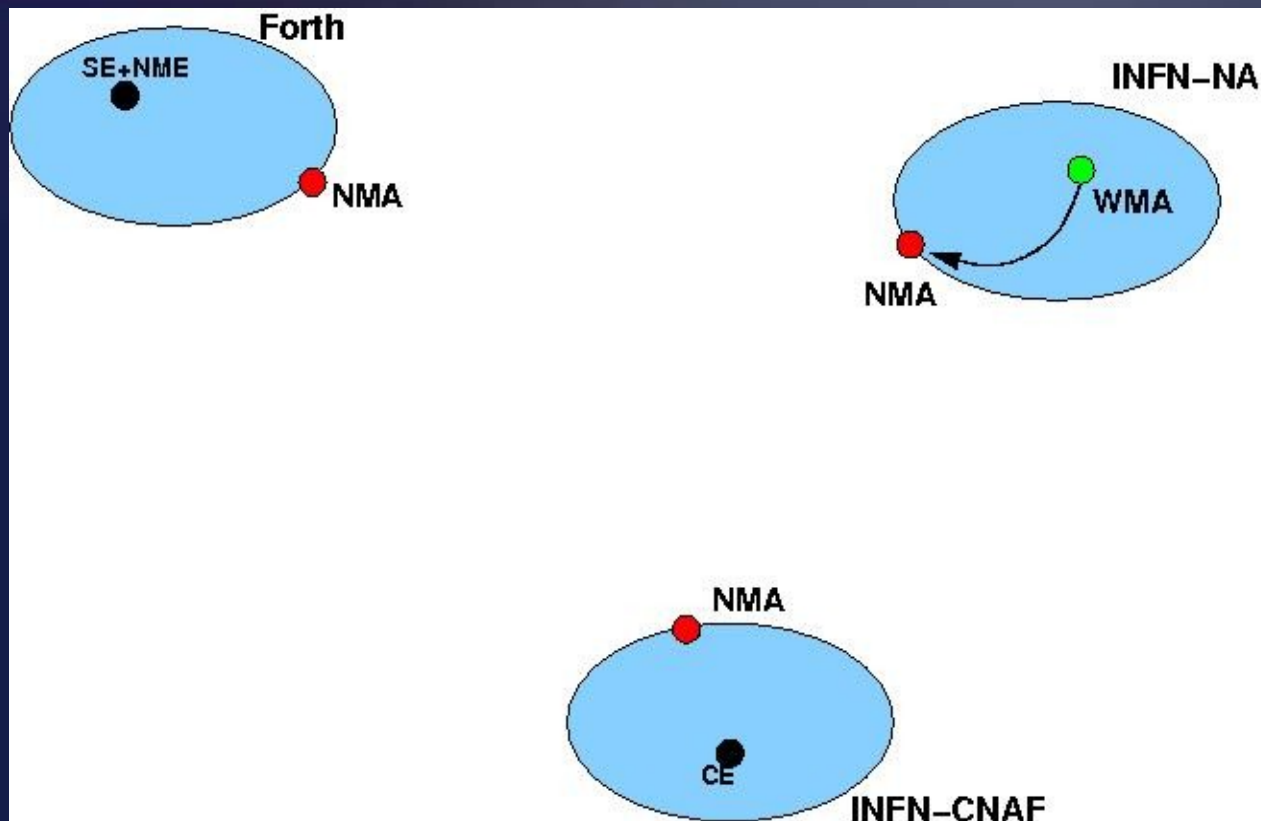
Observations: pro

- **No repositories needed for traces, since observations are produced and consumed on the fly**
- **Scalability is improved:**
 - **no n^2 structures in the system**
 - **no n^2 activities in the system**
 - **domains limit the scope of agents to a fraction of n**
- **Directories describe static capabilities of resources (which is what they are done for)**

Observations: con

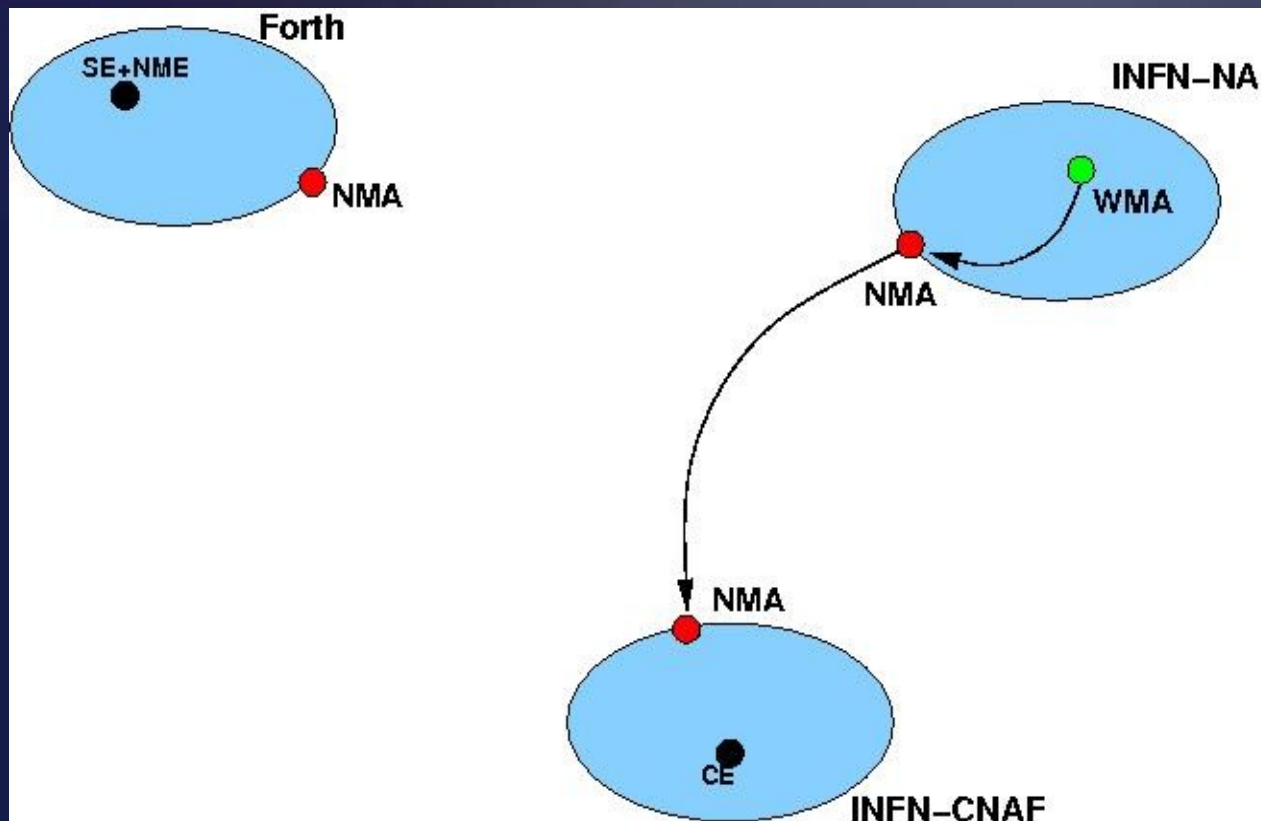
- **Design a coordination pattern for Network Monitoring Agents**
- **Network Monitoring Elements must be dynamically configurable (no ping.conf)**
- **A protocol to submit monitoring requests from Workflow Monitoring Agents to Network Monitoring Elements (via NMAs)**
- **Another to return observations backward, possibly as a multicast**

Request/Response protocol



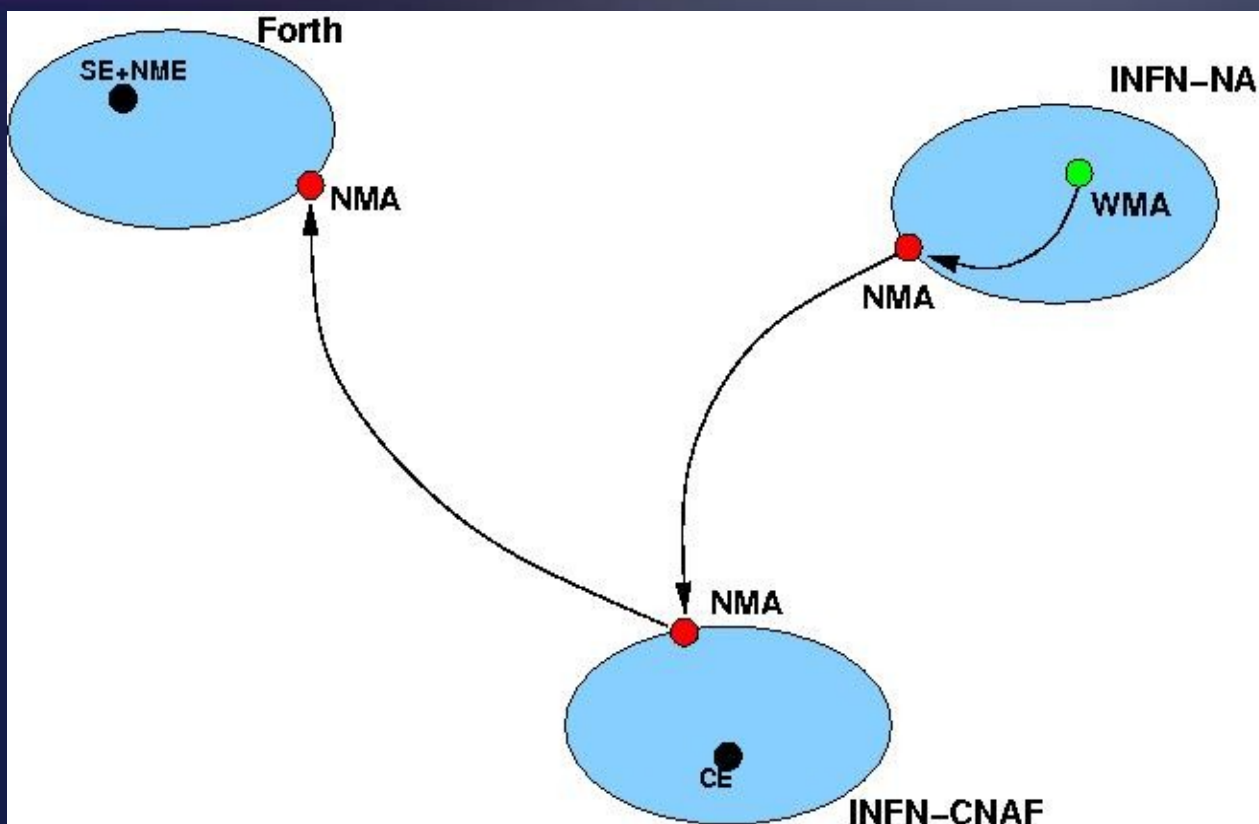
The Workflow Monitoring Agent issues a request to the local Network Monitoring Agent

Request/Response protocol



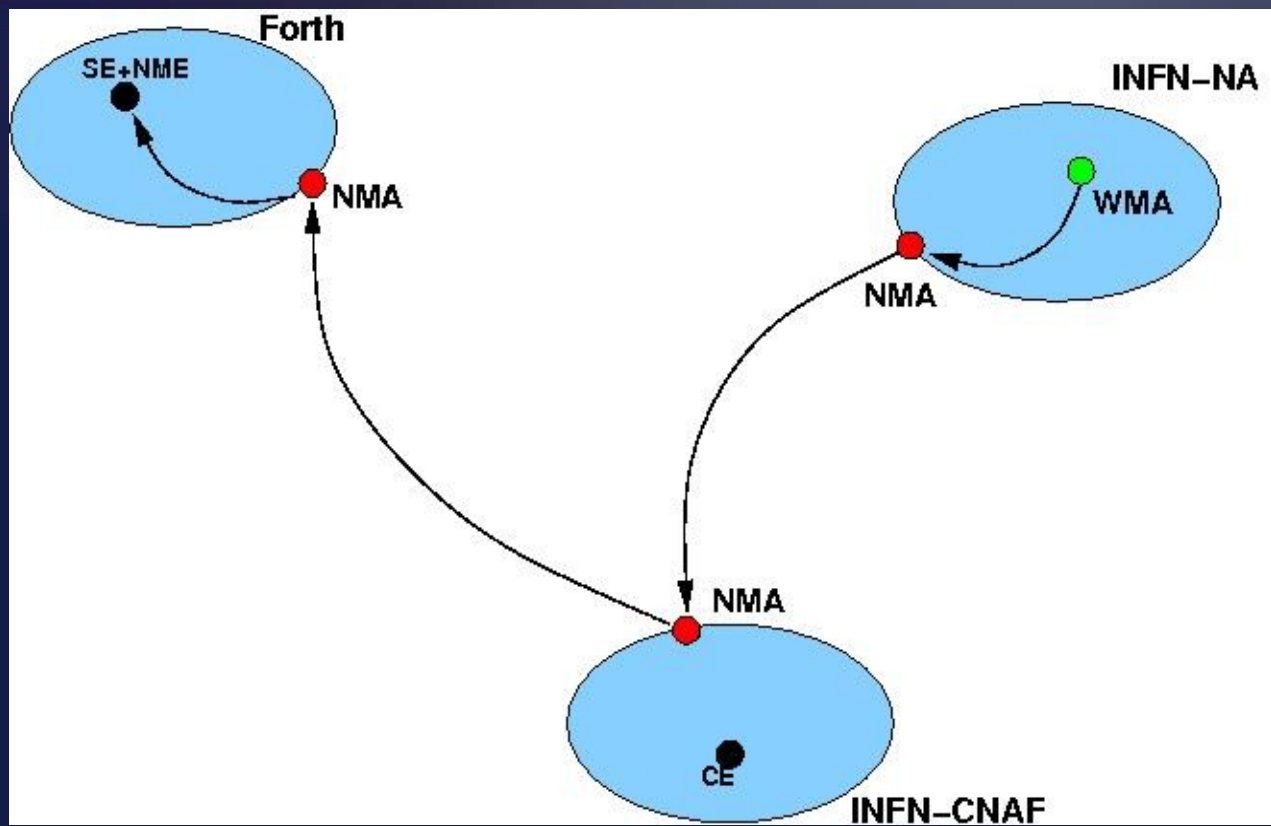
The Workflow Monitoring Agent routes the request to a peer (hierarchy?)

Request/Response protocol



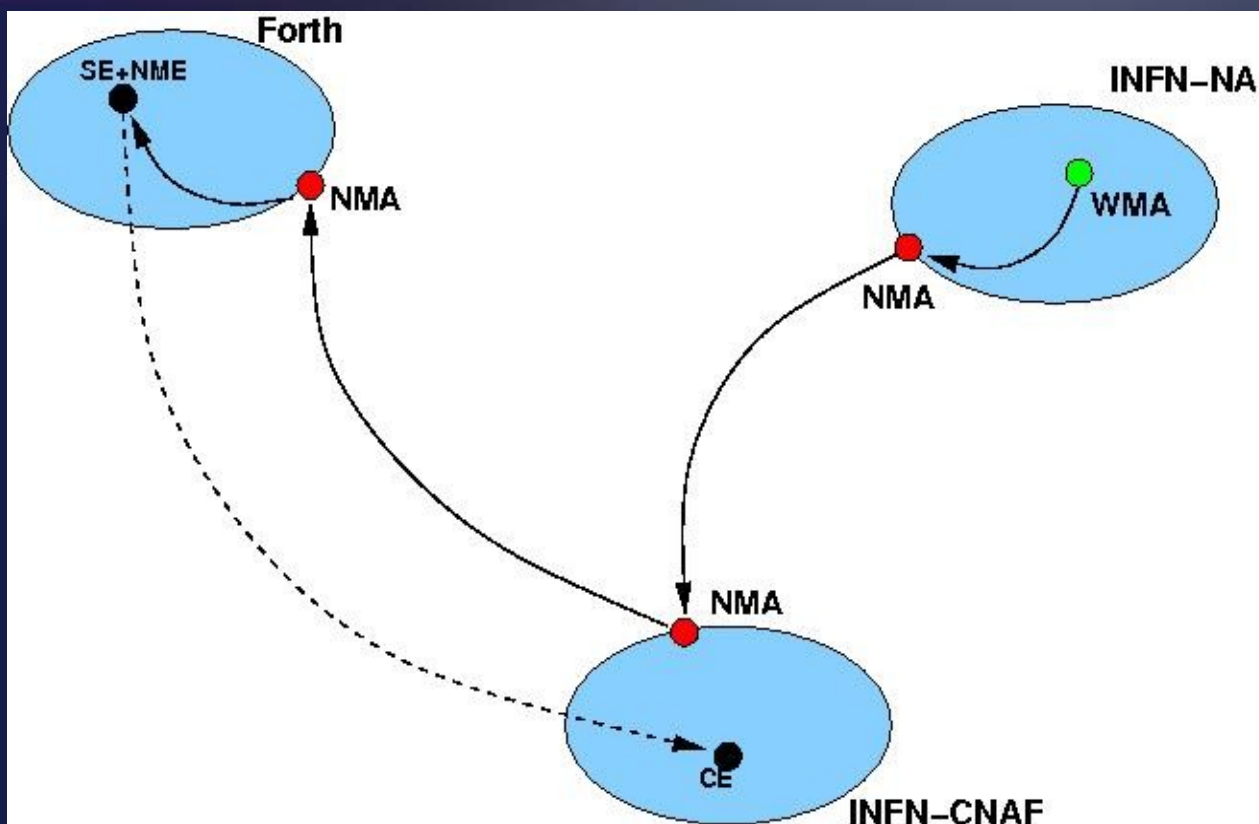
The Workflow Monitoring Agent routes the request to a peer (2nd step)

Request/Response protocol



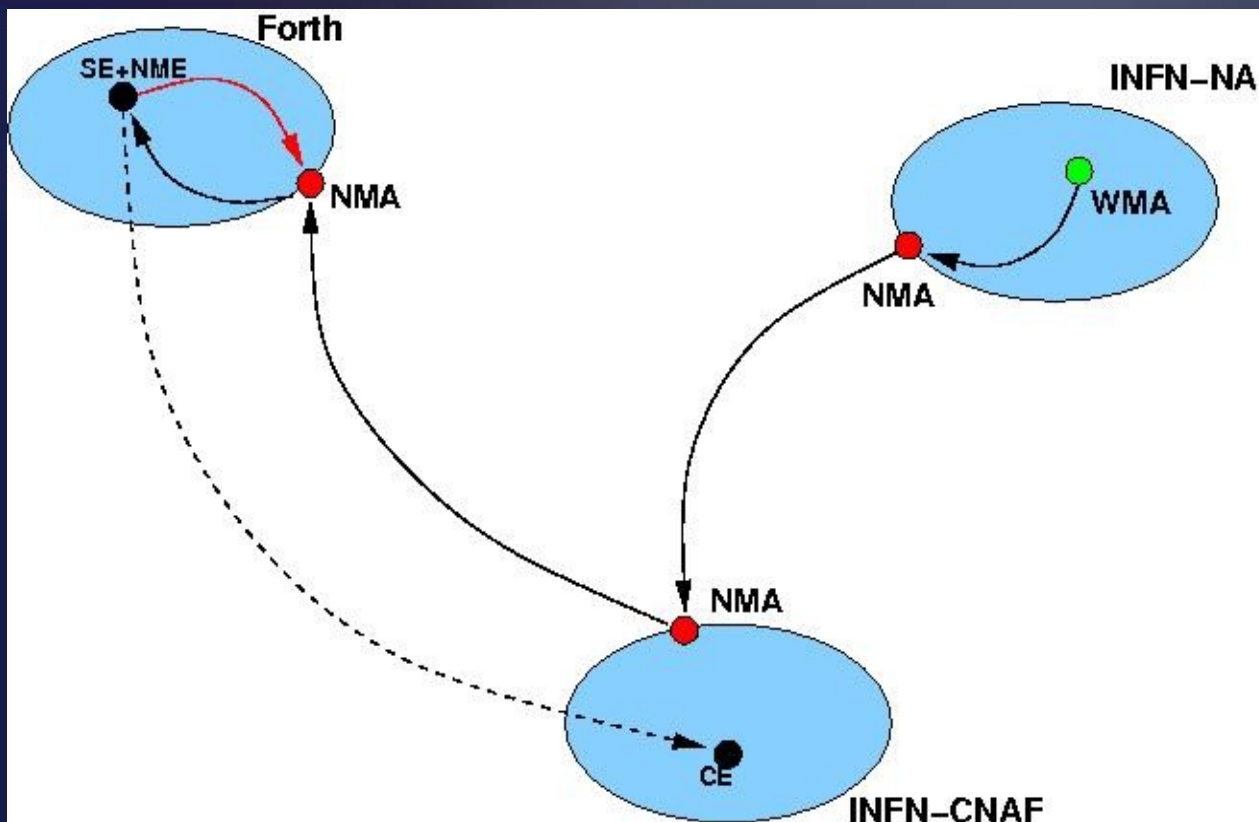
The Workflow Monitoring Agent delivers the request to the Network Monitoring Element

Request/Response protocol



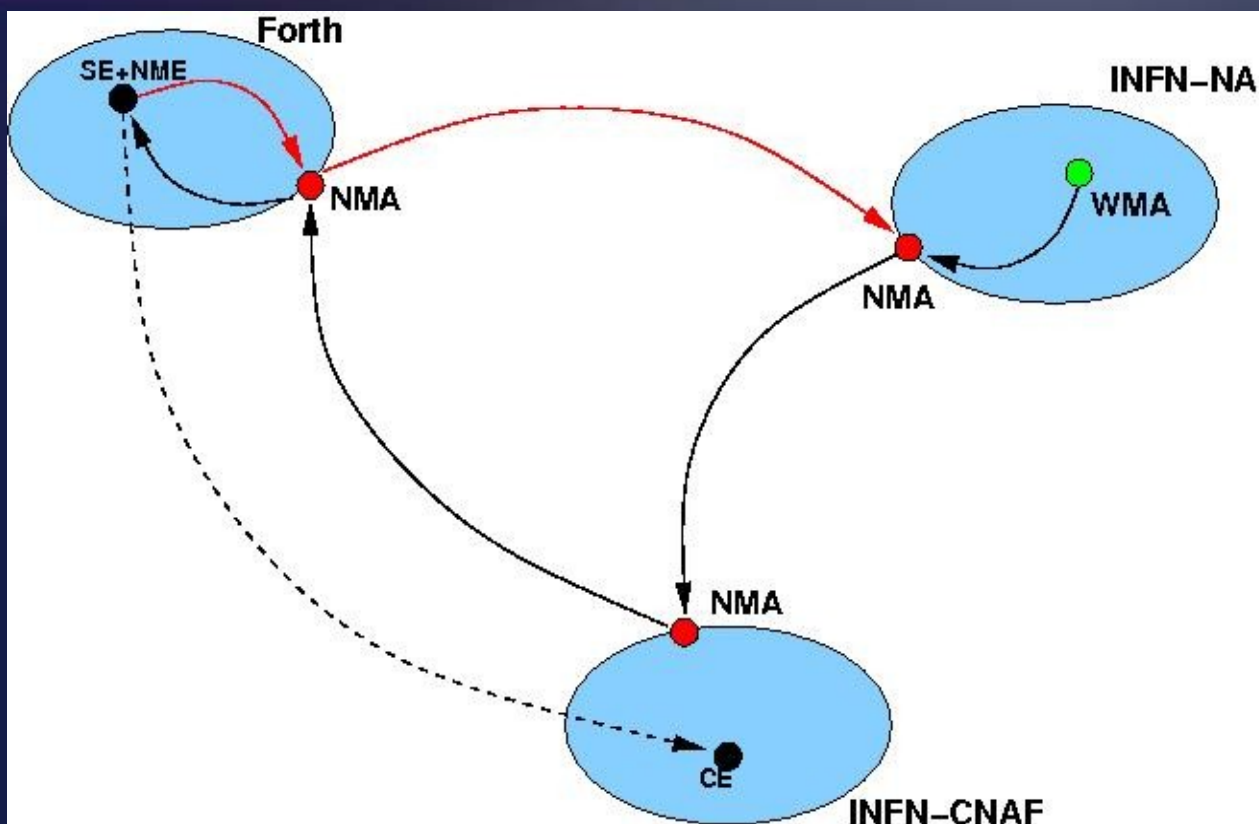
The Network Monitoring activity starts between the Storage and the Computing Element

Request/Response protocol



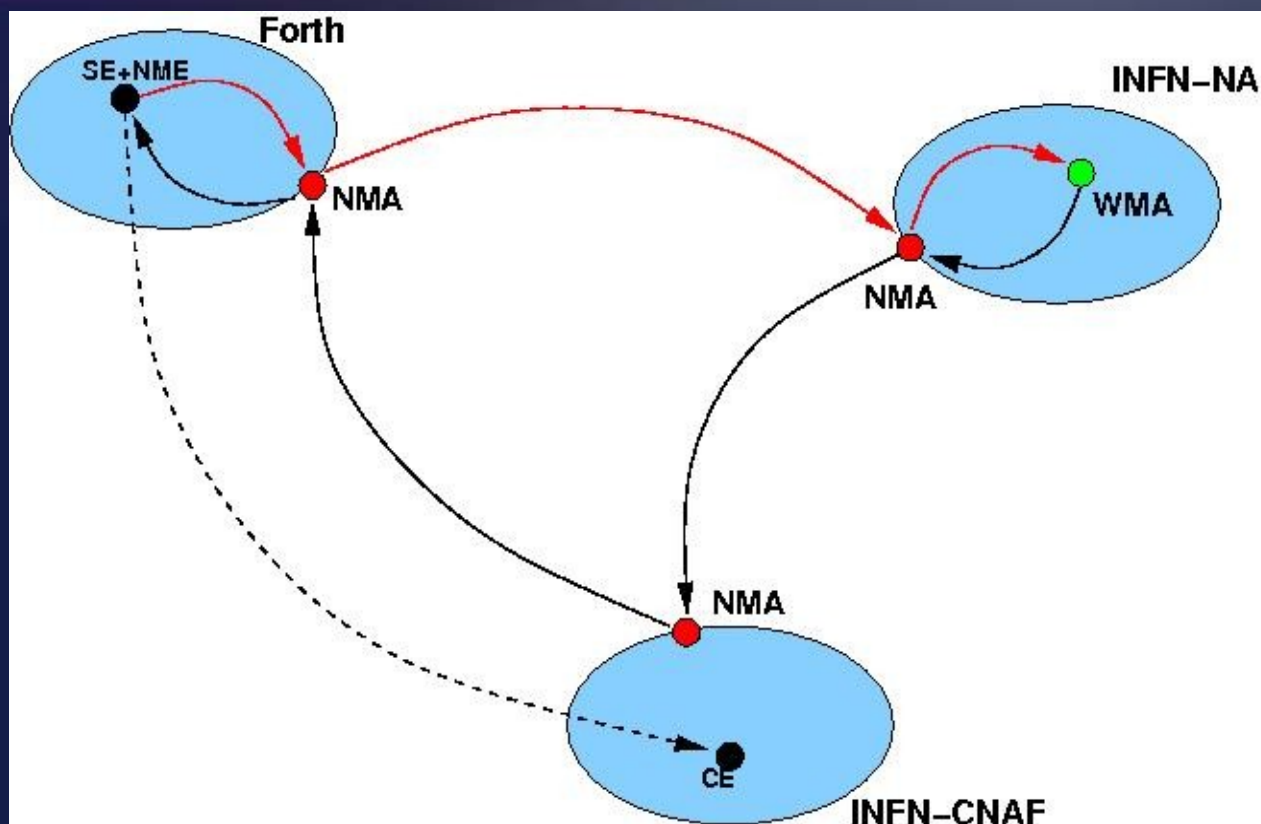
The stream carrying monitoring results is routed to the controlling Network Monitoring Agent

Request/Response protocol



The stream by-passes the intermediate Network Monitoring Agent

Request/Response protocol



The stream carrying the results reaches the requesting Workflow Monitoring Agent

Security issues

- **All communications are authenticated (except those produced by the monitoring tool inside the NME)**
- **Only communications between the Network Monitoring Agents require a global authentication support (a prototype exists)**
- **Other can be based on local policies**

The XSD schema

- The schemas have been written to fixate the ideas about the architecture.
- The use of XML addresses the portability of a possible implementation.
- The basic data item is the **NetworkMonitoringSession** type

Attributes of the Network Monitoring Session

Identification

```
<attribute name="SessionId" type="string"  
  use="required" />
```

Lifetime

```
<attribute name="StartAt" type="dateTime"  
  use="required" />
```

```
<attribute name="Duration" type="duration"  
  use="required" />
```

Resources for the result stream

```
<attribute name="BandwidthLimit"  
  type="nonNegativeInteger" default="0" />
```

```
<attribute name="Priority"  
  type="nonNegativeInteger" default="0" />
```

Elements of the Network Monitoring Session

Identification of the requesting WMA (for routing)

```
<element name="RequestFrom"  
type="nmsd:WorkflowMonitoringTaskType"  
maxOccurs="unbounded" />
```

Route stack (for reverse routing)

```
<element name="Route" type="nmsd:RouteStackType" />
```

Involved Domains (for NME selection)

```
<element name="NetworkElement"  
type="nmsd:NetworkElementType" />
```

Measurement descriptor (details in next slide)

```
<element name="MeasurementStream"  
type="nmsd:MeasurementStreamType" />
```

Definition of the Measurement Stream

A Measurement Stream contains several distinct data streams, with an identifying attribute:

```
<attribute name="CharacteristicStreamId"  
type="string" \>
```

each data stream is described by elements:

```
<element name="SamplePeriod" type="float"  
minOccurs="0" />
```

```
<element name="SourceIP" type="string"  
minOccurs="0" maxOccurs="unbounded" />
```

```
<element name="DestinationIP" type="string"  
minOccurs="0" maxOccurs="unbounded" />
```

...and by a choice of tool specific options...

Tool specific options

Each element contains options for one single tool:

```
<choice>
```

```
<element name="PingOptions"  
type="pt:PingOptionsType" />
```

```
<element name="AppmonOptions"  
type="am:AppmonOptionsType" />
```

```
</choice>
```

Each element is passed (in principle untouched) to a single NME.

Ping (active): packet size, data aggregation etc.

Appmon (passive): packet filter, data aggregation, characteristic, anonymization etc..

Separation of concern

- **NMA access all parts of the Network Monitoring Session description except the Measurement Stream descriptor (header?)**
- **WMA and NME produce/consume the sequence of Measurement Stream descriptors (payload?)**
- **The inter-NMA protocol acts as a transport for Network Monitoring Stream descriptors**

Tool oriented description

- **The WMA directly addresses a tool, and its capabilities**
- **Decouples design of network monitoring tools**
- **The protocol between the WMAs does not depend on the definition of the requested measurements**

Outline of the stream contents

- **One frame series for each element in the Measurement Stream descriptor**
- **Demultiplexing using Session and Measurement identifiers**
- **Also in this case, NMAs offer a transport service**
- **De-multiplexing is on the WMA, as well as frame unmarshalling.**

- We focus on monitoring activity finalized to workflow monitoring
- We introduce an architecture that includes consumers, producers and infrastructure
- We outline a request protocol, and design a schema that describes a monitoring session
- We outline a streaming protocol to deliver the results
- Experiments on the way...
- Full presentation and a paper on my WEB page
OGF20 link