

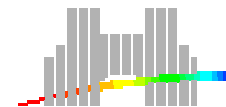
Achieving performances in active networks : a mandatory step to provide dynamic network services for Grid middleware and applications

GridBus workshop

J.P. Gelas – L. Lefèvre

Equipe INRIA RESO / LIP– Lyon, France

Jpgelas@ens-lyon.fr, laurent.lefevre@inria.fr



The GRID

- Remote tasks
- Wide area distributed heterogeneous applications

With Middleware environments (Globus, Legion, Condor, Netsolve, SGE, Nimrod...)

Providing applications services : data management, machines enrollment, security, APIs...

- On long distance networks (IP, TCP)

Grid targets

- Multi cluster architecture
 - Set of clusters / parallel machines
 - Few sites
 - Power process sharing
- Large scale computing
 - Thousand of machines
 - Cycles stealing

For Grid designers : « networks are simple and do not support Grid applications and middleware »

Networks and Grids

- Grid designers do not care about networks
 - ->TCP / IP for all
- But Grid applications need new services for transporting their data :
 - Reliable multicast, QoS, security, streams adaptation...
- Some of these services will have difficulties to be standardized (IETF..) or will never be available in network equipments...
- Need solutions to rapidly test, deploy and experiment new services in the network
- Active networks can help reaching this goal

Contents

- Active networking
- Mixing Grid and dynamicity in the network : the Active Grid architecture
- High performances in active networks : the Tamanoir approach
 - Execution environment layer
 - Inside kernel
 - Distributed resources

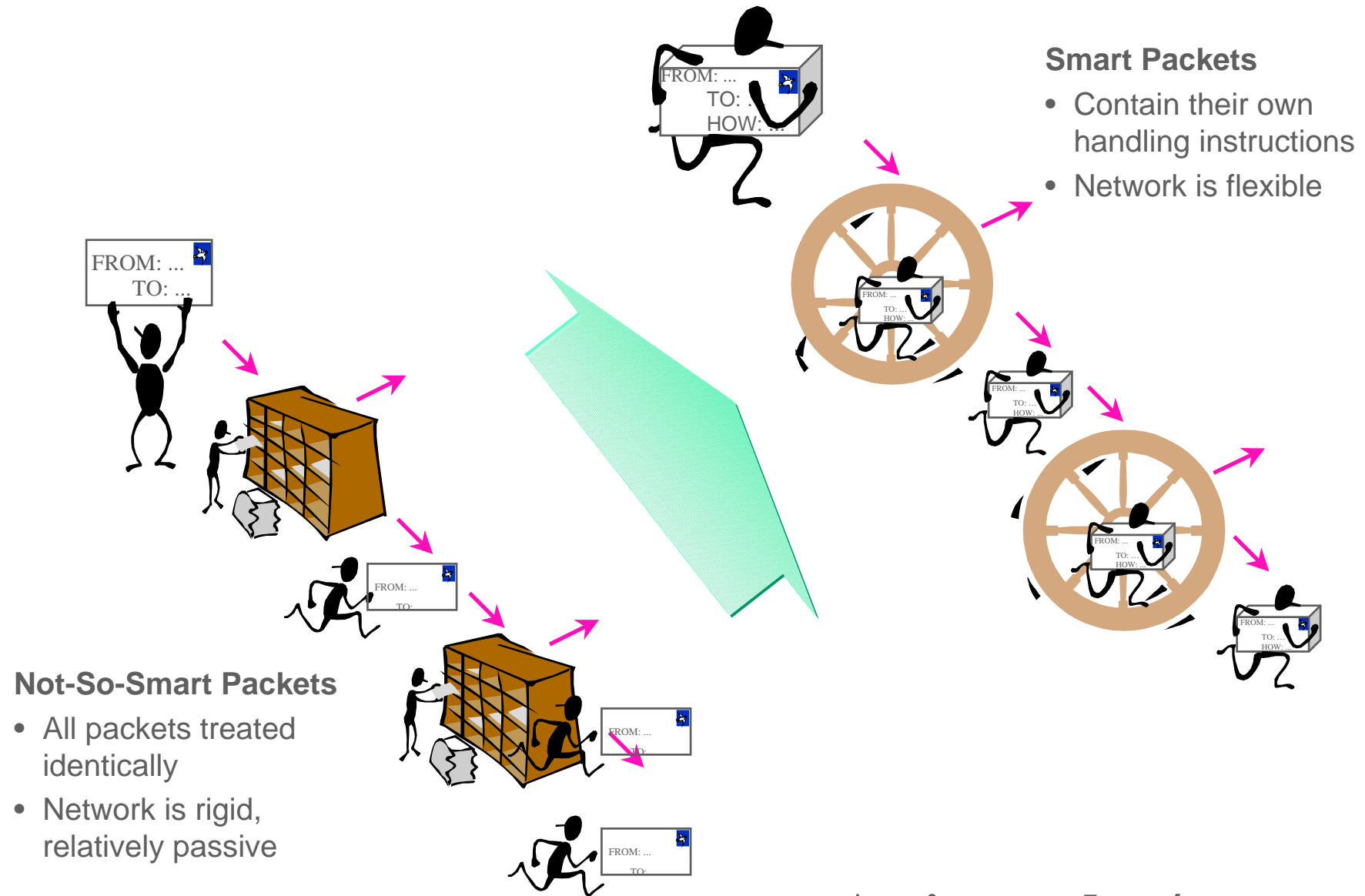
Active networks

- « *Provide intelligence and processing power inside networks* » to
 - Improve usage of network resources
 - Dynamically deploy new protocols
 - Manage equipments heterogeneity
 - Support data streams heterogeneity
 - Propose new network services to operators and applications

The Network is considered "**active**":

- 1** User code can be injected into intermediate systems to customize network services
- 2** The intermediate nodes can build the payload of packets, not only the header.

Active networks



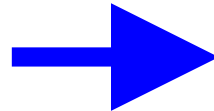
- D. Tennehouse slides

=> **Active nodes / routers**
GridBus workshop, 7th June 2003, Melbourne

Different Approaches

("out-of-band" code injection)

Discrete Approach



Configurable Node

Active Services
ANN

CANES

Active Bridging

ANTS

ANCORS

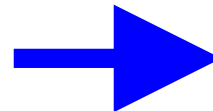
PLAN

Messenger

Smart Packets



Integrated Approach

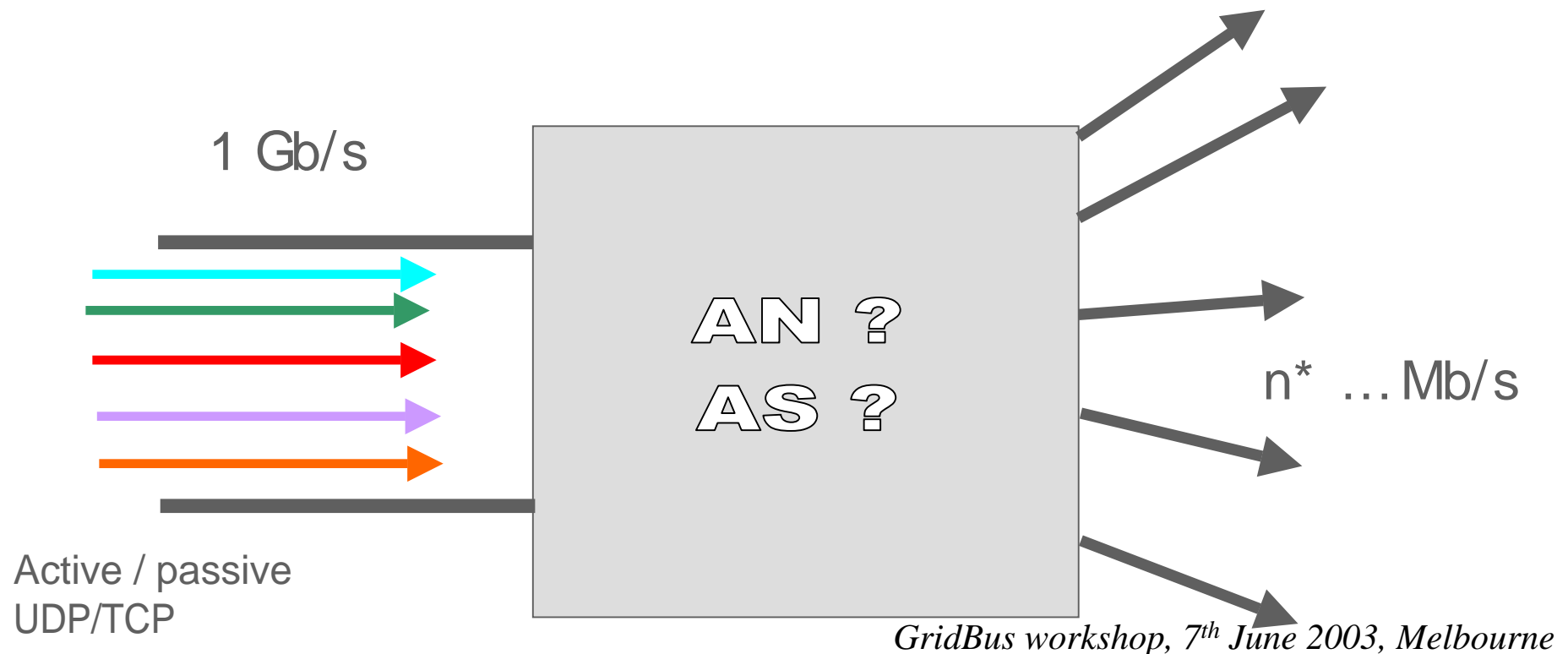


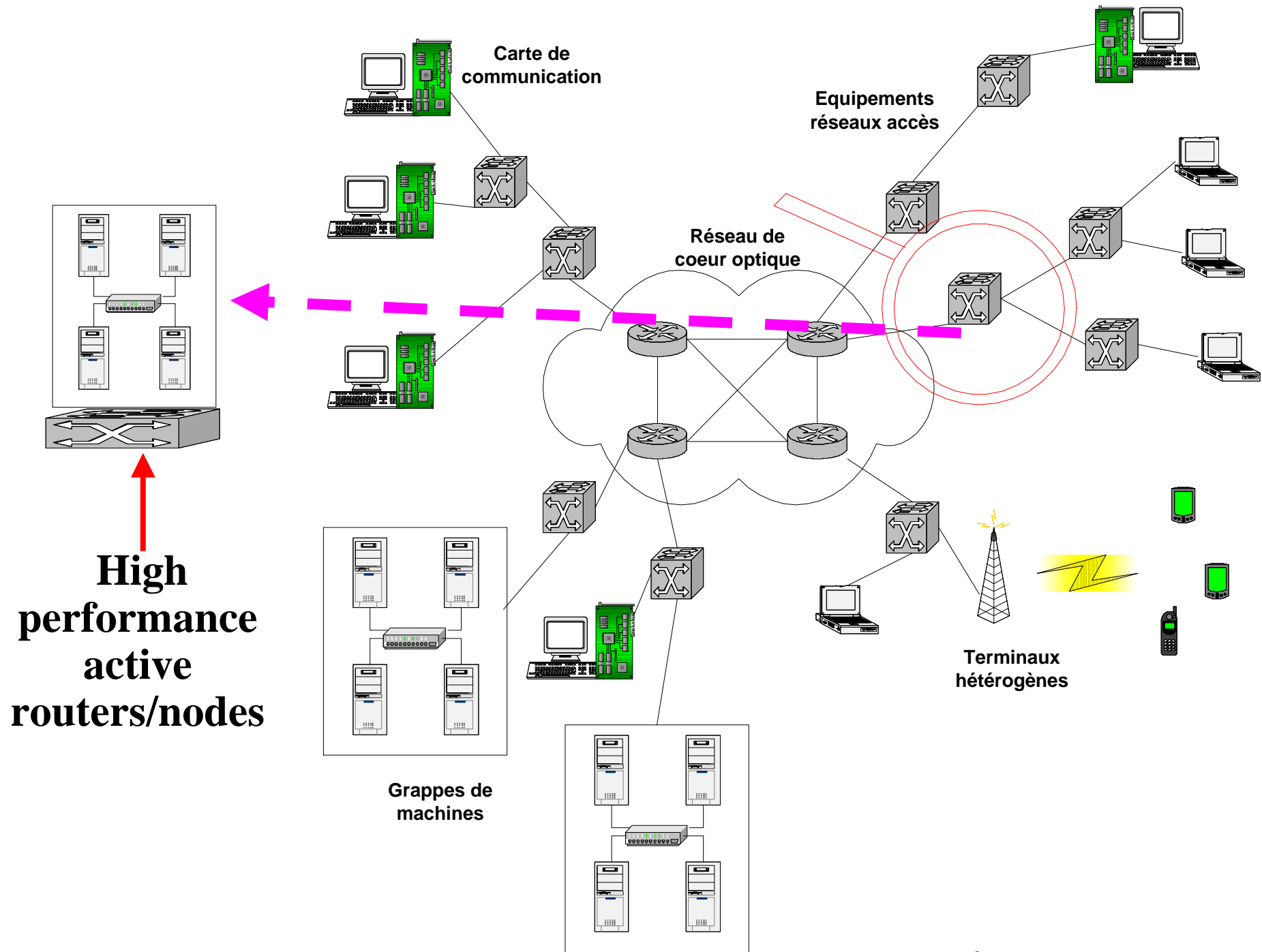
Packet Programming

("in-band" code injection)

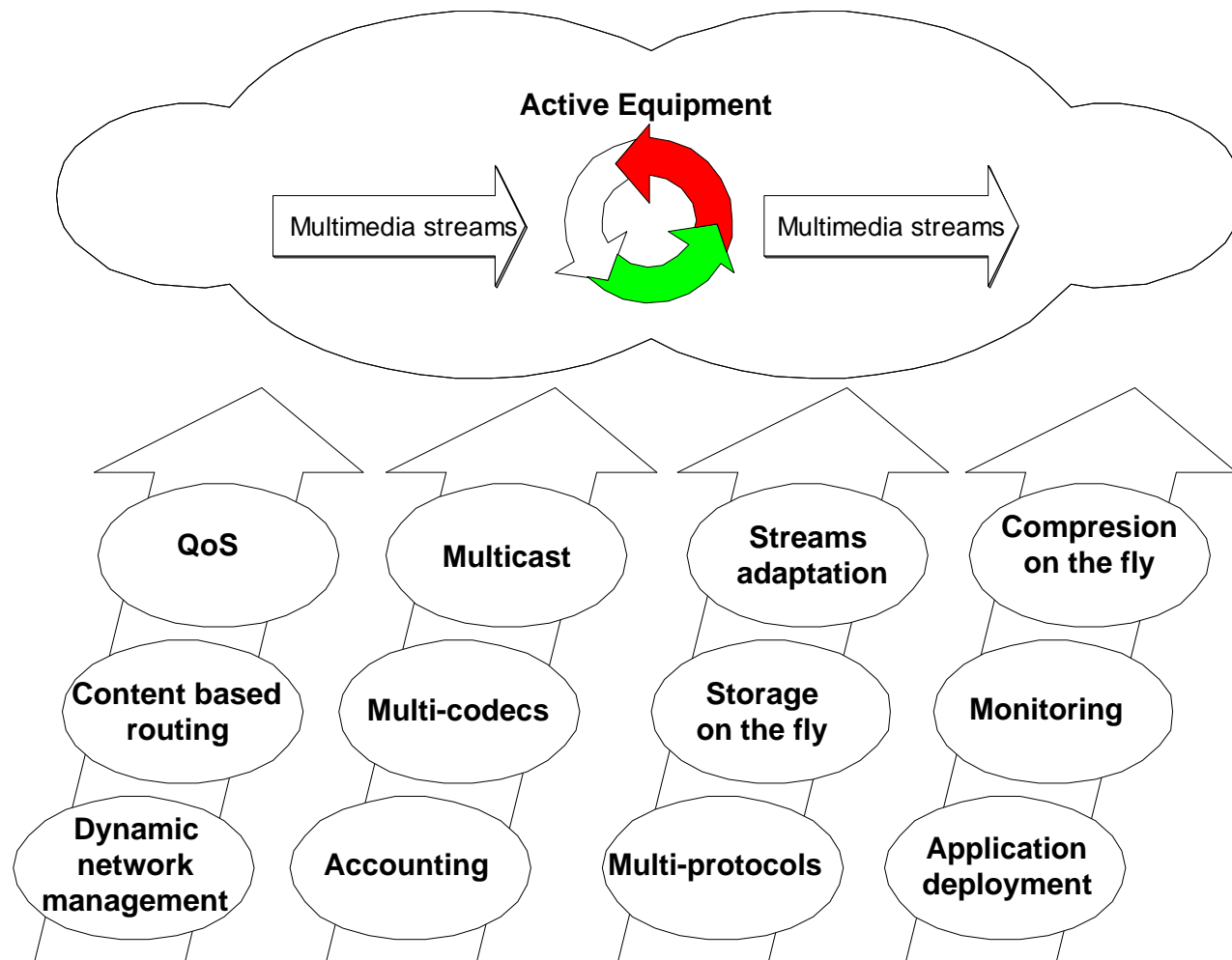
What do we want ? AN and AS

- High performance active node
- Passive and active packets
- Dedicated services : QoS, reliable multicast, cache...



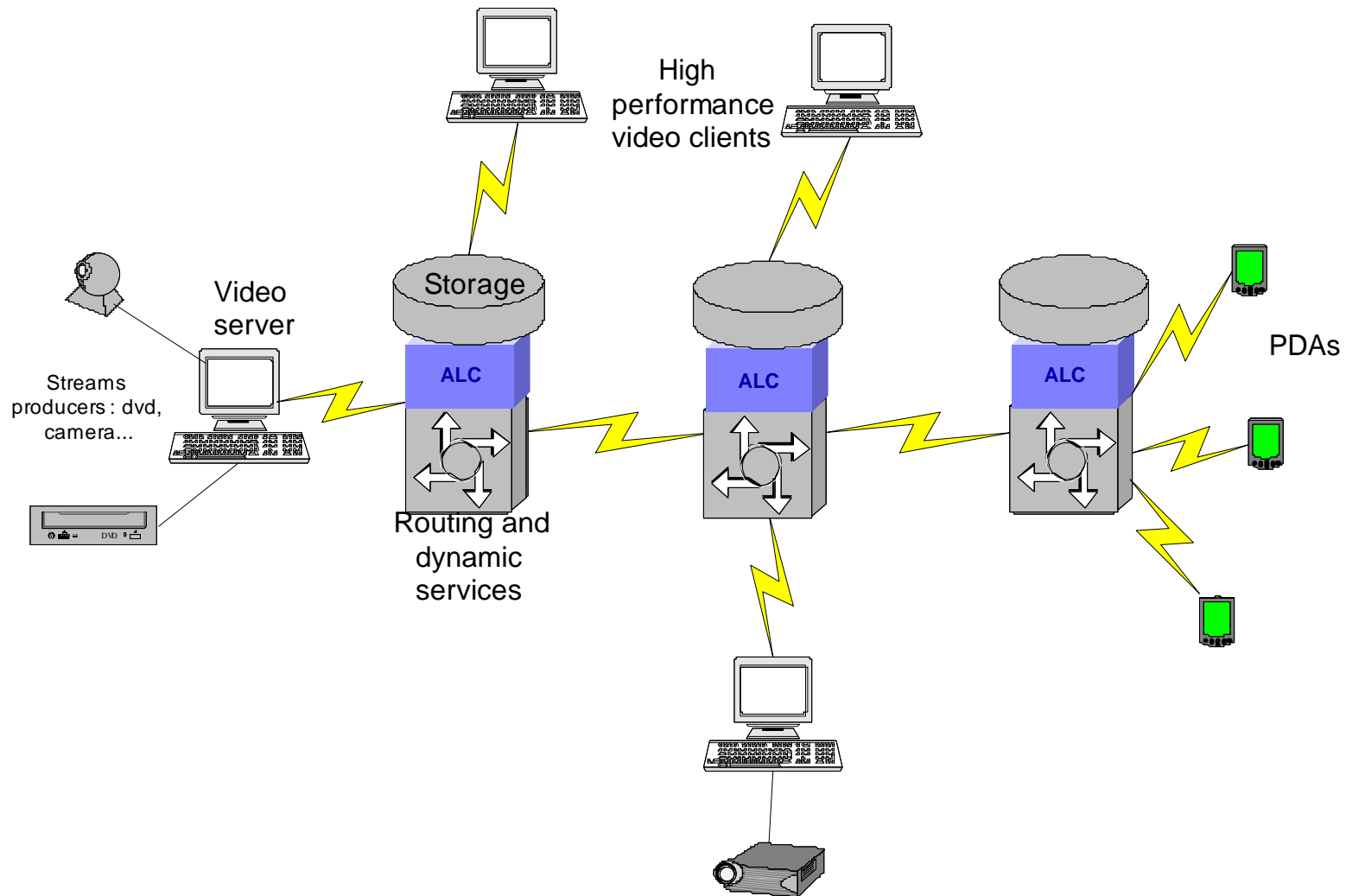


Active services tool-box for multimedia and Grid applications



Examples of active services

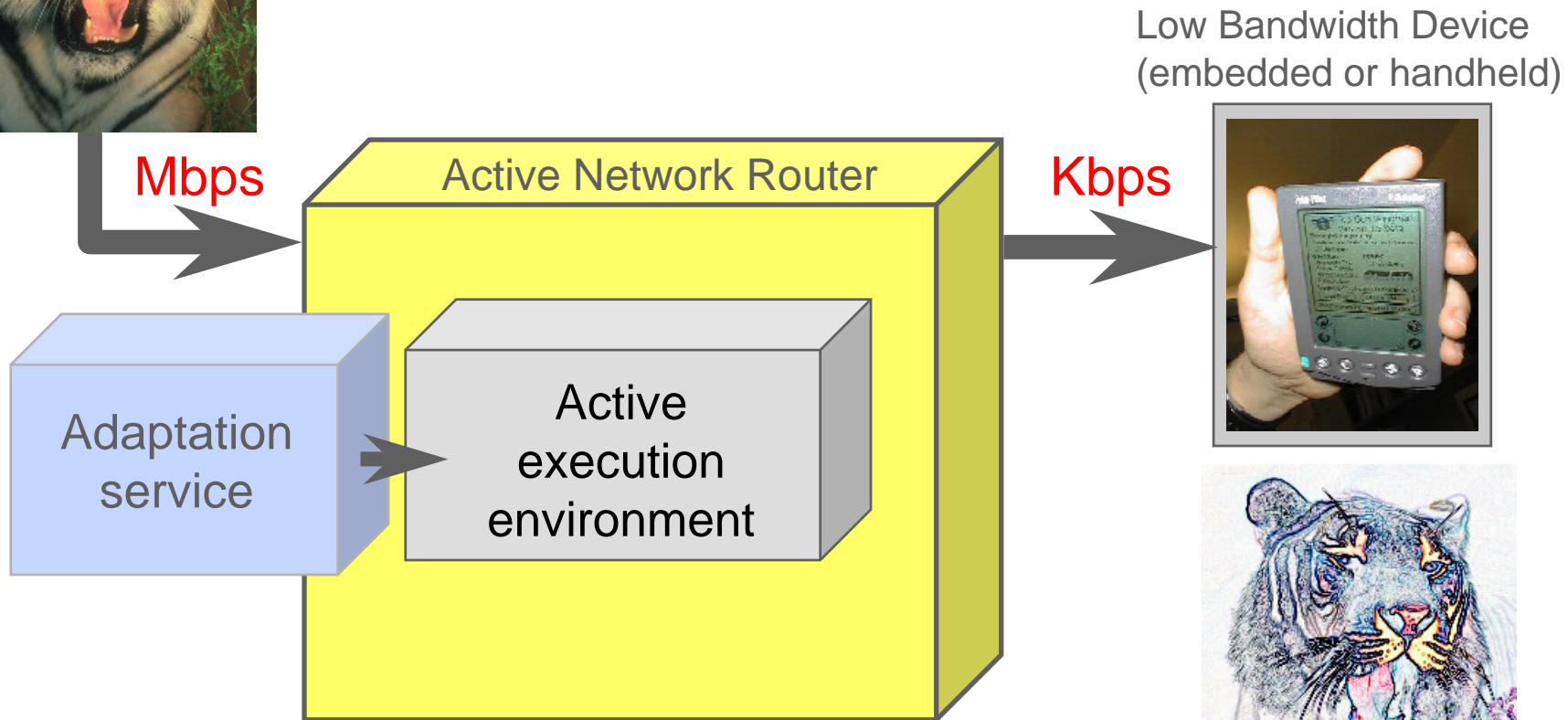
Active Video adaptation



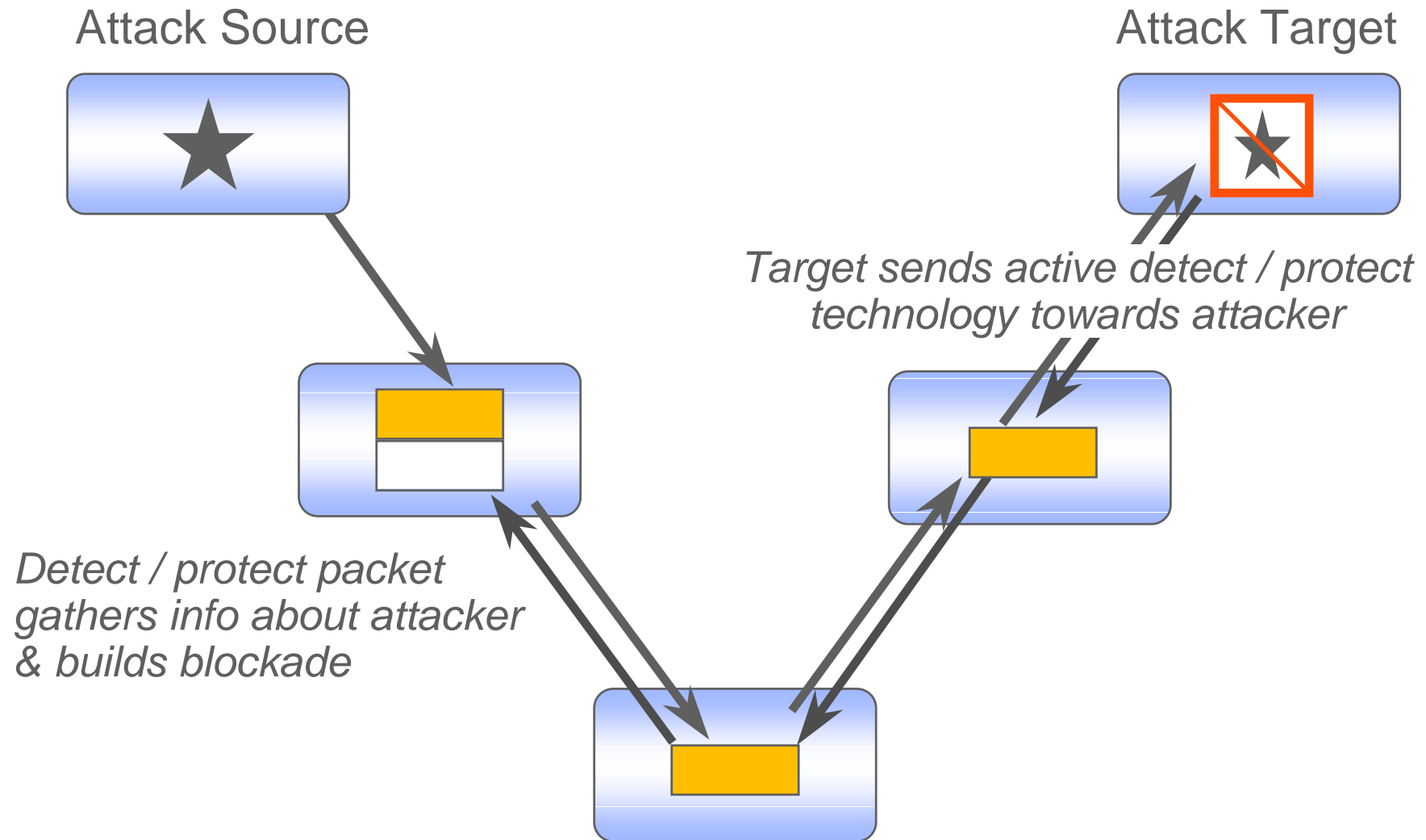
Adaptation multimedia stream



Large Data Stream Reduced to Small One
at Nearest Capable Network Point



Network attack traceback against DDos



Active Grid

Active Grid

« A more intelligent and dynamic network to support Grid middlewares and applications with adapted services. »

Grid control/management streams

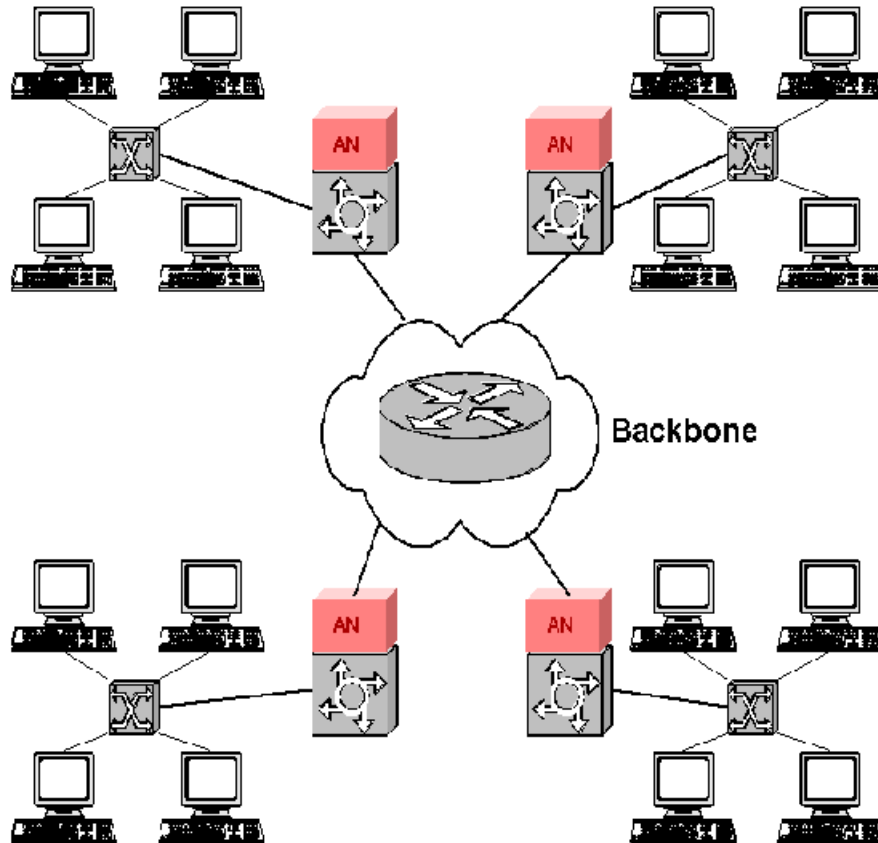
- Middleware environment deployment
 - Heterogeneity, dynamic topology
 - Dynamic enrollment
 - Machine subscribing
- Application deployment
 - Collective communications : multicast, gather
 - Source deployment, results gathering
 - Fault tolerance
- Grid support
 - Monitoring, network sensors

Grid applications streams

- Grid application input / output
 - Parameters
 - data
- WAN parallel processing
 - Communications between tasks
 - Point2point, global
 - QoS

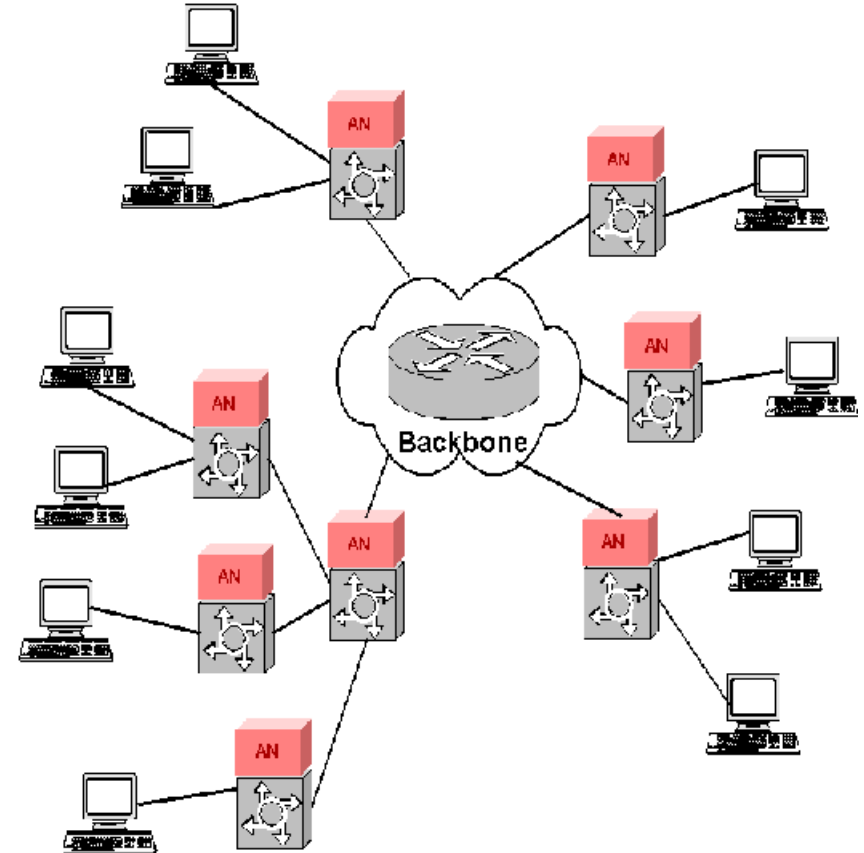
Active Grid

Multi-cluster computing



- AN : cluster head
- Manage data streams entering and leaving
- Local communications protocols
- Aggregate output streams

Large-scale computing



- AN associated with set of nodes
- AN hierarchies on different networks

Requirements for Grid Network Services

Meta cluster computing

Set of parallel machines or clusters linked together.

- *Grid environment deployment* : OS heterogeneity support, dynamic topology reconfiguration, fault tolerance.
- *Grid application management* : multi- and gathercast communication for binaries deployment, parameters and collection of results of distributed tasks.
- *Grid support* : collection of data control, nodes synchro, node workload info.

- ## Large scale computing

Thousand of connected machines

- *Grid environment deployment* : dynamic enrollment of unused machines.
- *Grid application deployment* : fault tolerance, check-pointing protocols.
- *Grid Support* : workload information of subscribed machines.

AGrid

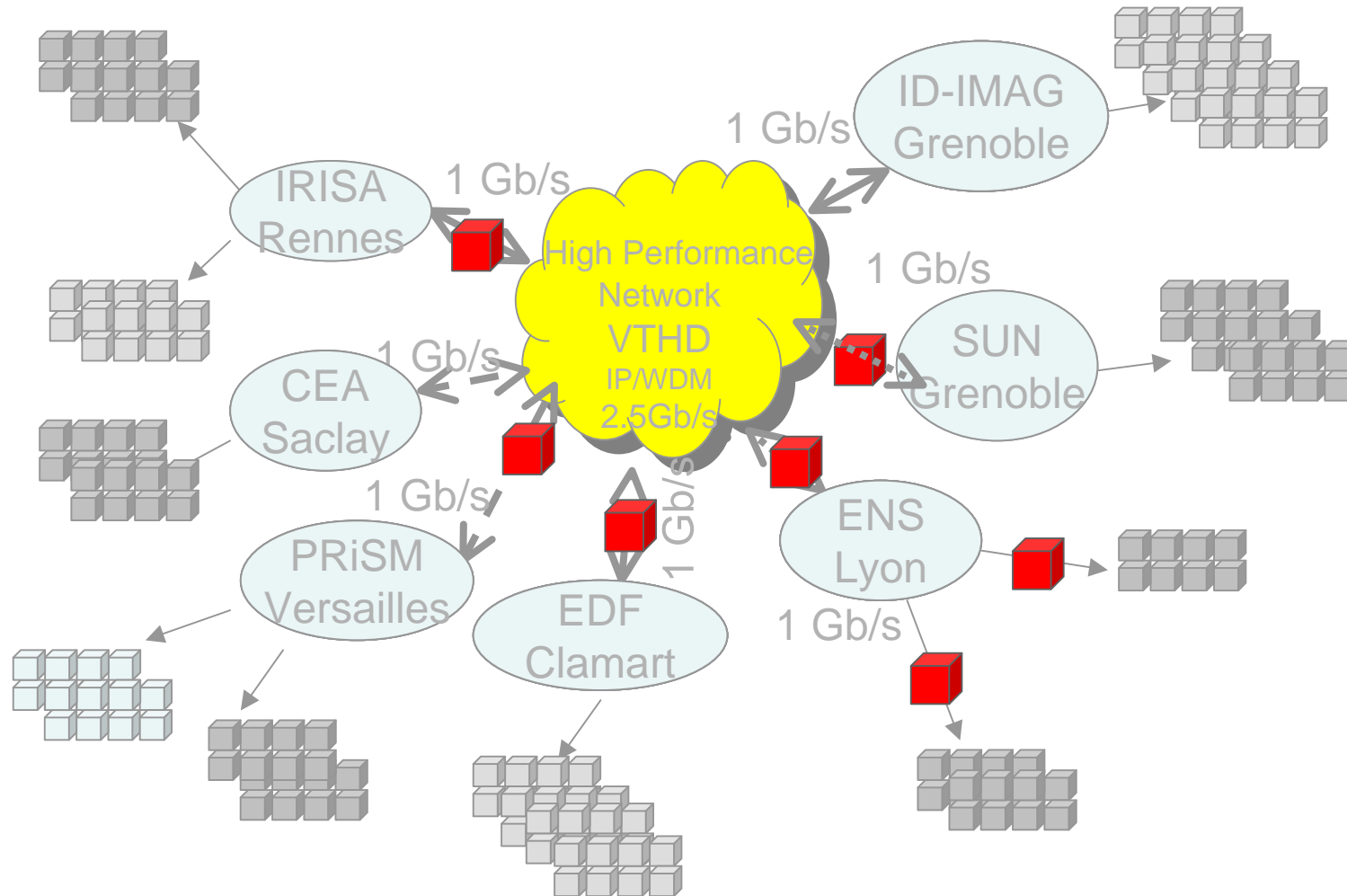
- A-Grid can improve Grid applications :
 - Application deployment
 - ARM : source code, binaries
 - Active cache : gathering, loss recovery
 - QoS : data transfer
 - Grid support
 - A-Grid provides information to middleware
 - Distant tasks management
 - WAN process
 - Active QoS for streams
 - Efficient transport

Tamanoir testbed around HP backbone

- RNRT VTHD project (Very high speed network) (<http://www.vthd.org>)
- Deploying high performance active node around wide area backbone (2.5 Gbit/S, 1 GEth)
- Supporting wide area applications (grid and multimedia applications)
- Development of :
 - wide area visualization tools for active node management,
 - Distributed and P2P active traffic generator



Tamanoir Active Grid support



« Active Grid : an intelligent and dynamic network which supports Grid middlewares and applications with adapted services »

GridBus workshop, 7th June 2003, Melbourne

Active Network and Grid Management

Policy

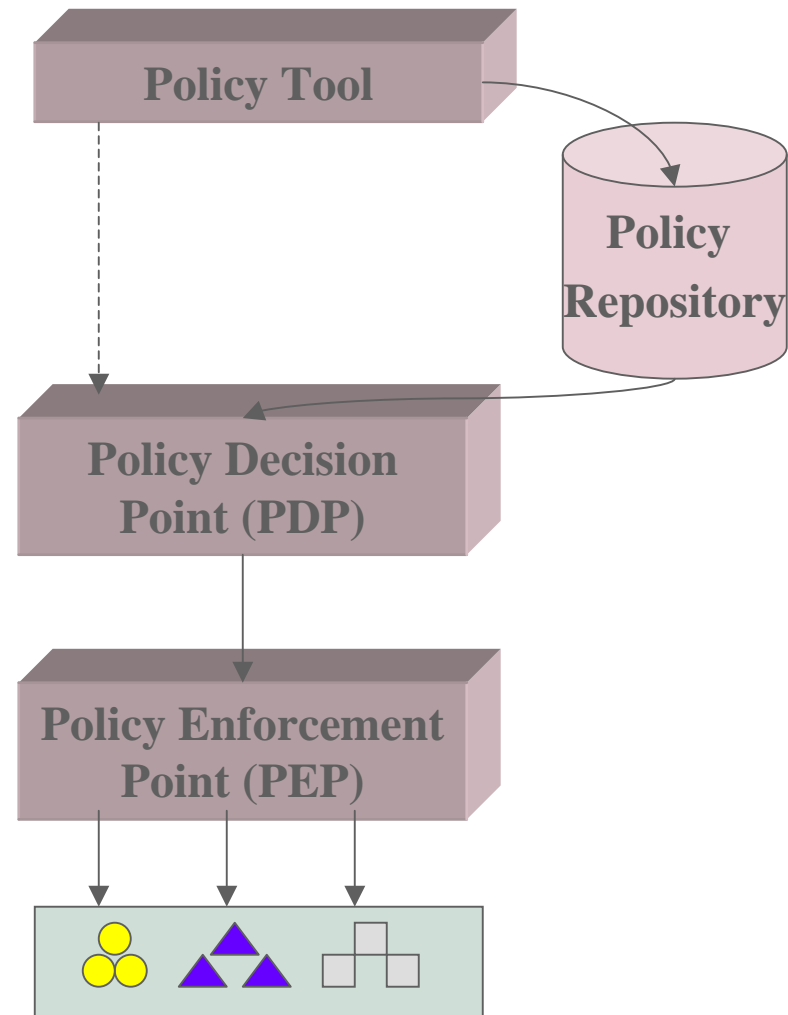
- A way to guide a behavior of the network through high level declarative directives
- Ex: IF (sourceHost == host1) AND (destHost == host2) THEN ProvideGoldService();

PBM (Policy Based Management)

- Allows control elements to be configured or scheduled on the fly
- Applies integrated management
 - System management
 - Network management
 - Service management
- \Rightarrow To cooperate in Grid computing

Why Policy-Based “Grid” Management?

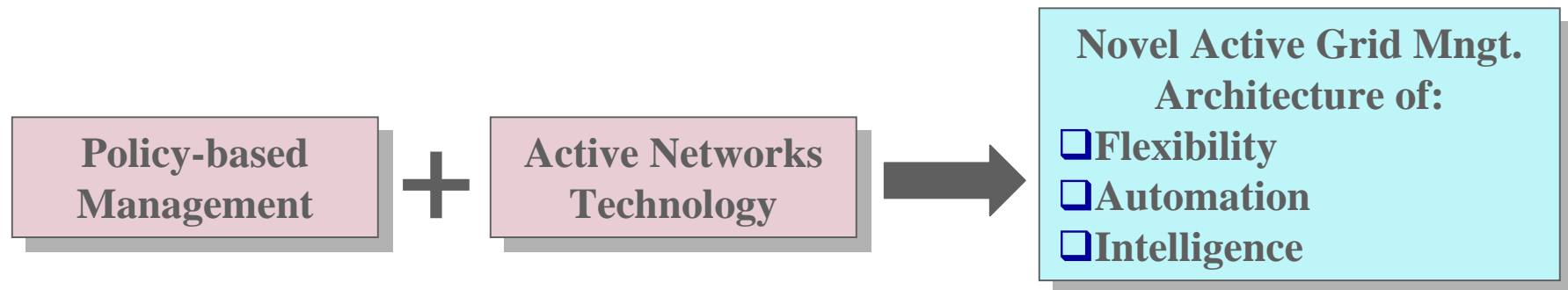
- Relieving network & service administrator from the burden of configuring every single Grid resource manually
- More flexible: administrator re-configures Grid by giving or changing policies
- More concerned about end2end management of Grid services



GridBus workshop, 7th June 2003, Melbourne

Why PBM+AN?

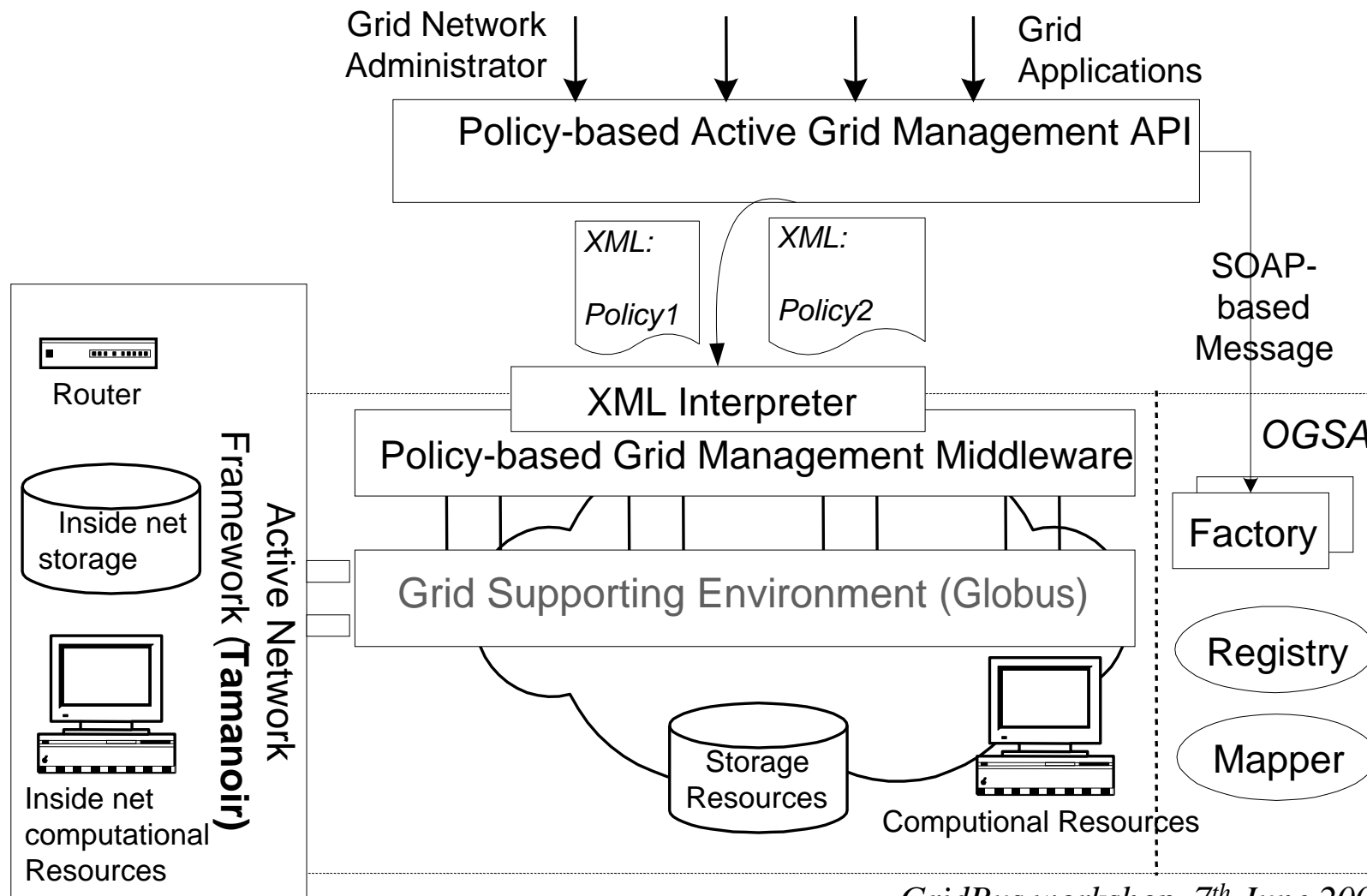
- PBM (Policy Based Management) is more suitable for the flexible and scalable management of Grid.
- AN (Active Network) speeds up the rapid creation and deployment of Network Grid services by introducing intelligence inside the network and other Grid resources.
- On one hand, *active networks* is a kind of enabling technology for
 - policy transit
 - policy downloading
 - Policy enforcement
- on the other hand, *PBM* also provides the management of AN themselves.



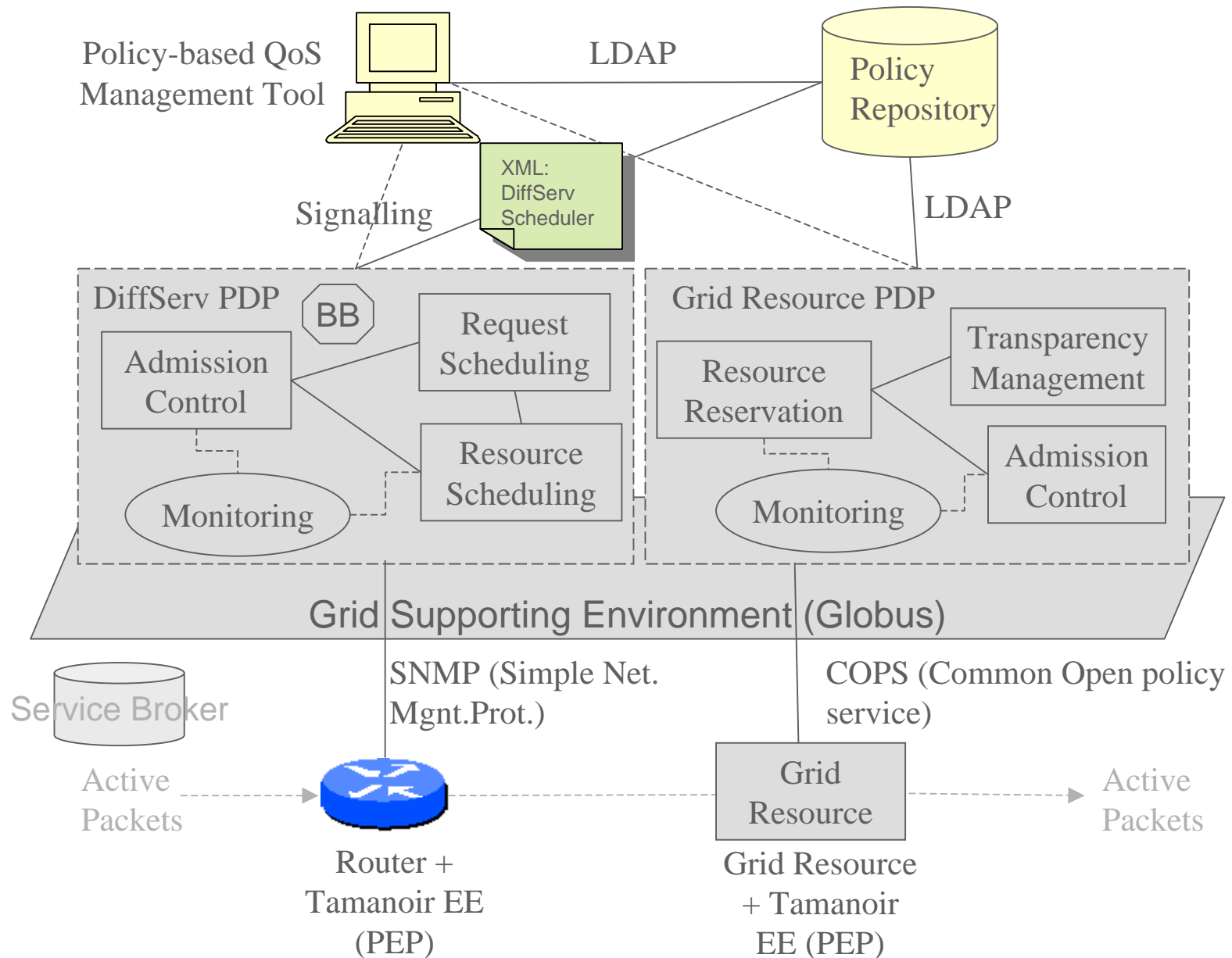
Architecture

- OGSA (Open Grid Service Architecture): standard mechanism for creating, naming and discovering grid services.
- PBM + AN must fit in OGSA proposal.
- Active Grid architecture provides :
 - Mechanisms to dynamically adapt Grid networks elements to Grid services requirements
 - Grid resources management

Overall Active Grid Architecture



Middleware architecture

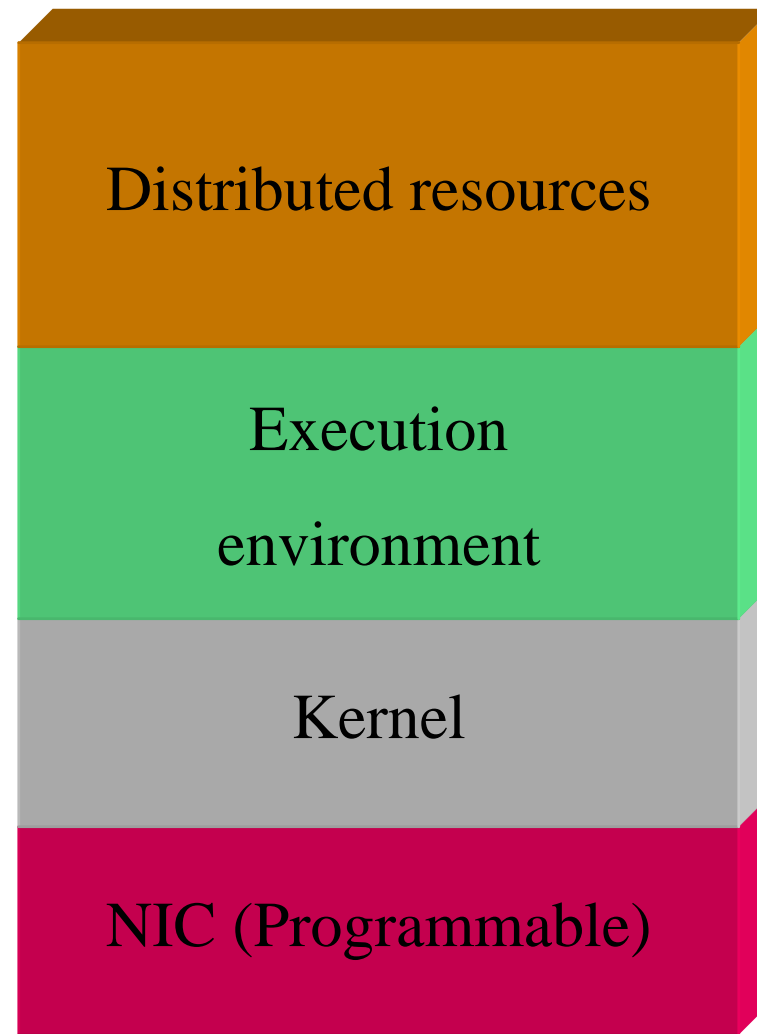


High performances in active networks

- For the moment : a few Mbit/s
- What do we need ? Performances for data transport / packets processing
 - Optimized execution environment
 - Compilation / Portability
 - Heterogeneity (OS / Networks)
 - Multi-services / scalability
 - Distributed / upgradeable architecture



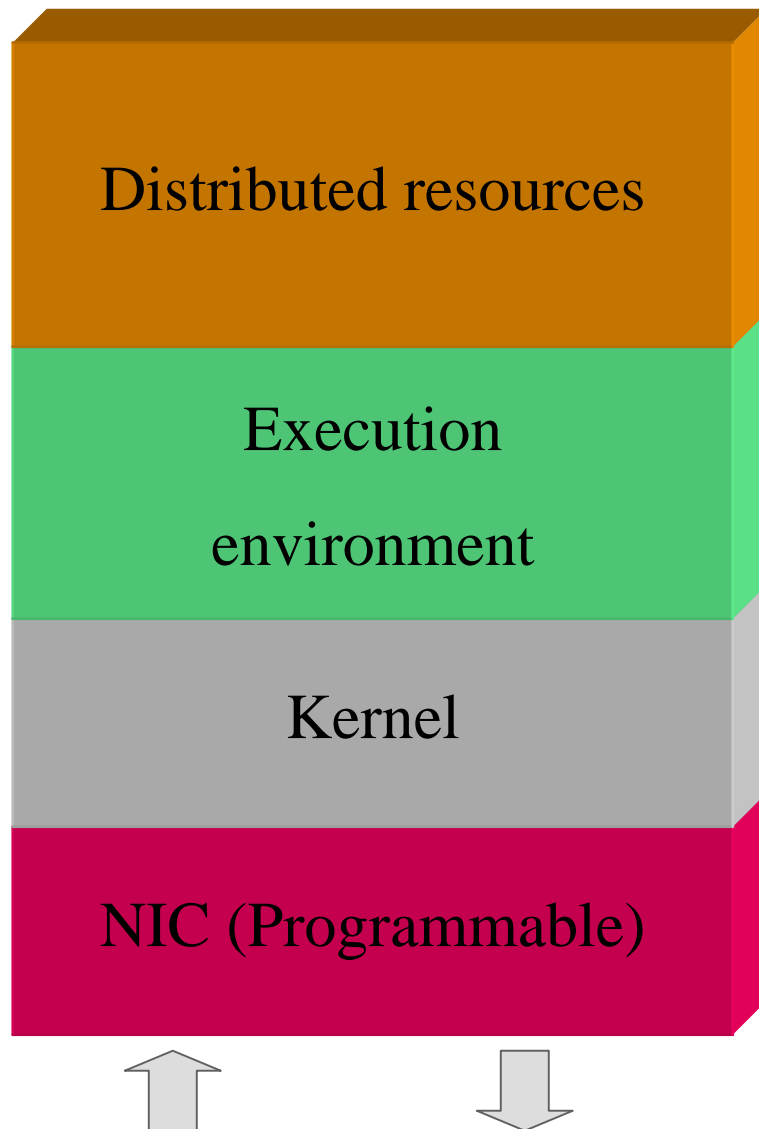
Tamanoir Architecture



GridBus workshop, 7th June 2003, Melbourne



Tamanoir Architecture : adapted for heterogeneous services



- Resources consuming services : distributed storage, streams transcoding, on the fly compression, cryptography...
- Services deployment / linked with middleware : reliable multicast...
- Middle services : content based routing, QoS...
- Light network services : packet marking, QoS...

High Performance Execution Environment

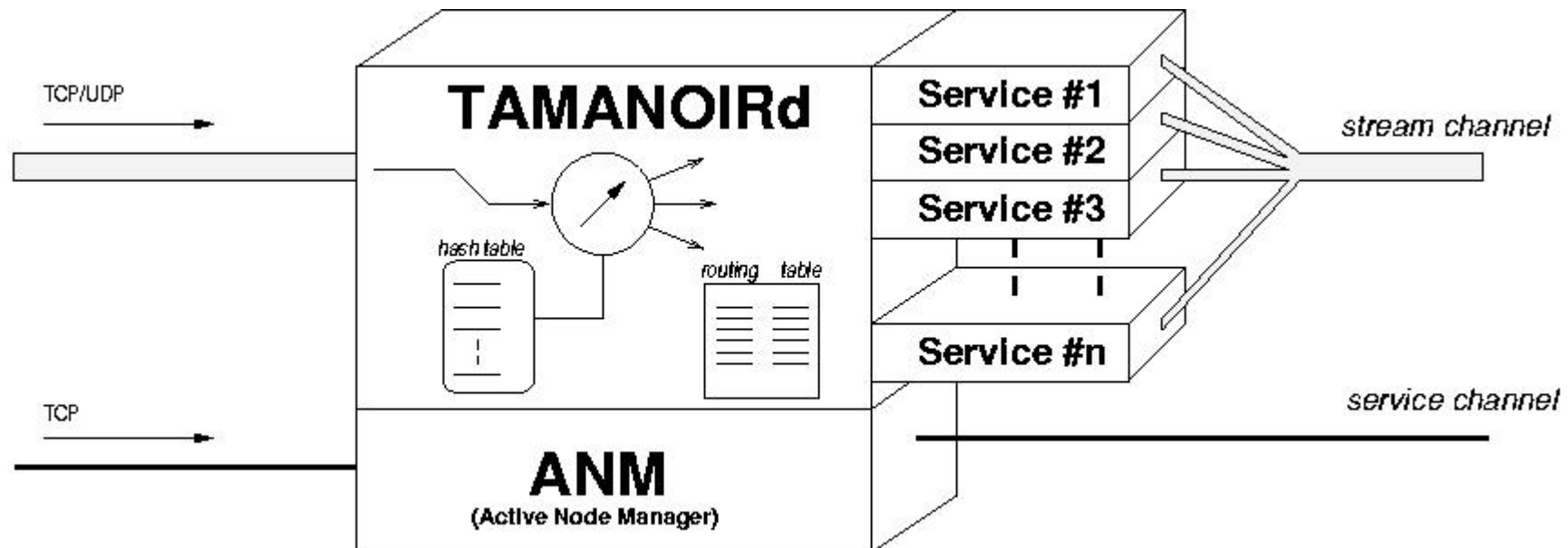
High performance AN support



- Execution Environment based on Java
- Distributed multi-threaded architecture
- Active packets : ANEP / service number
- Dynamic deployment of services
- Streams : UDP and TCP
- Compiled optimization (GCJ)

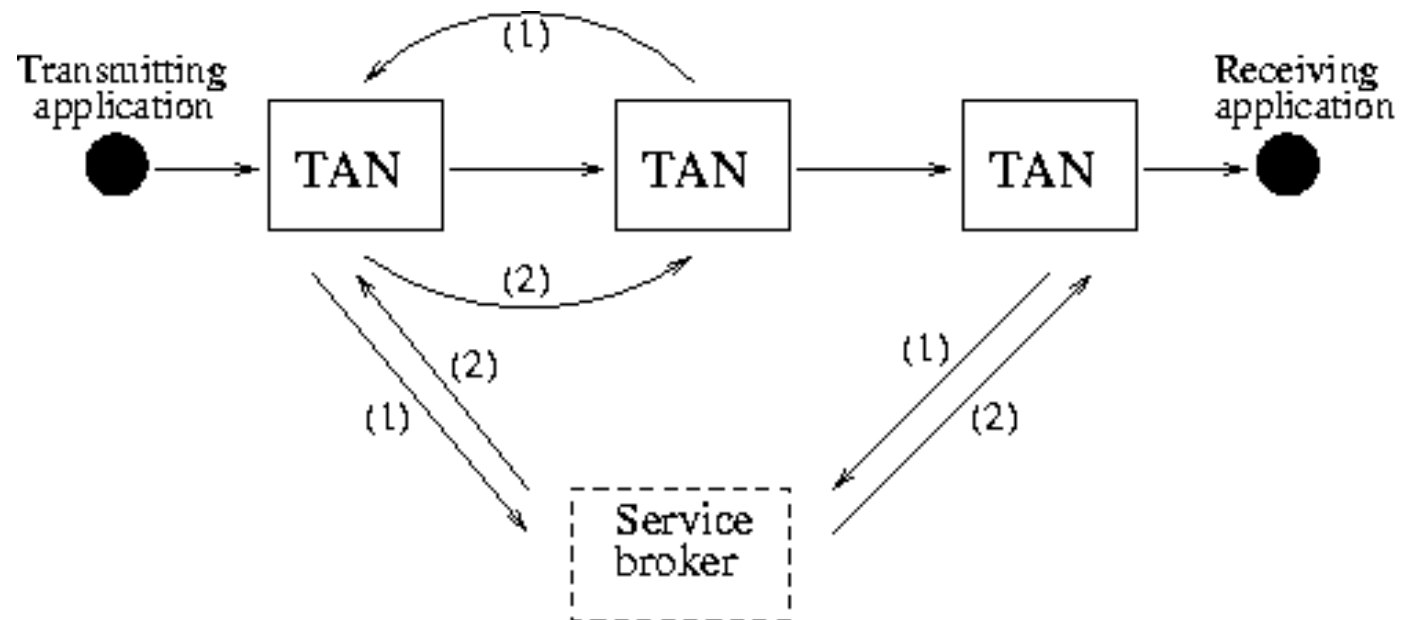
Tamanoir Active Node (TAN)

Execution environment



Active service deployment

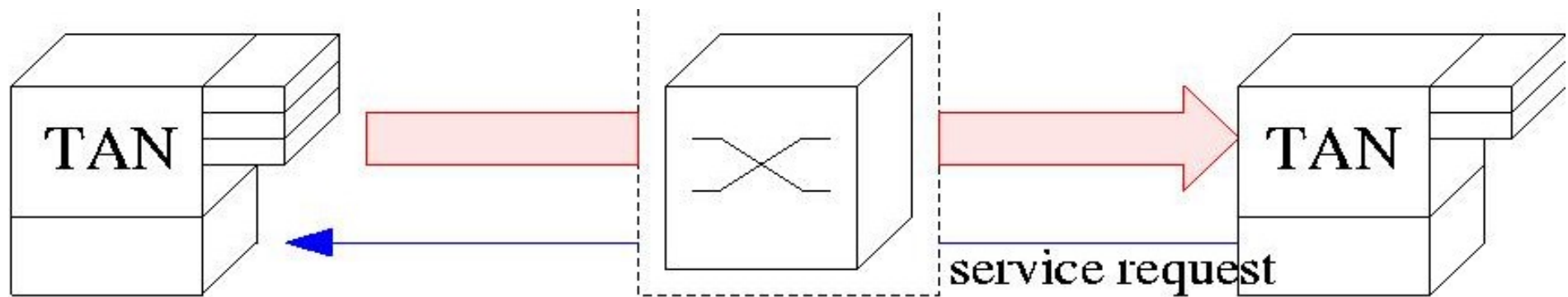
- From application / middleware
- From TAN
- From a Service Broker



service demand loading (1)

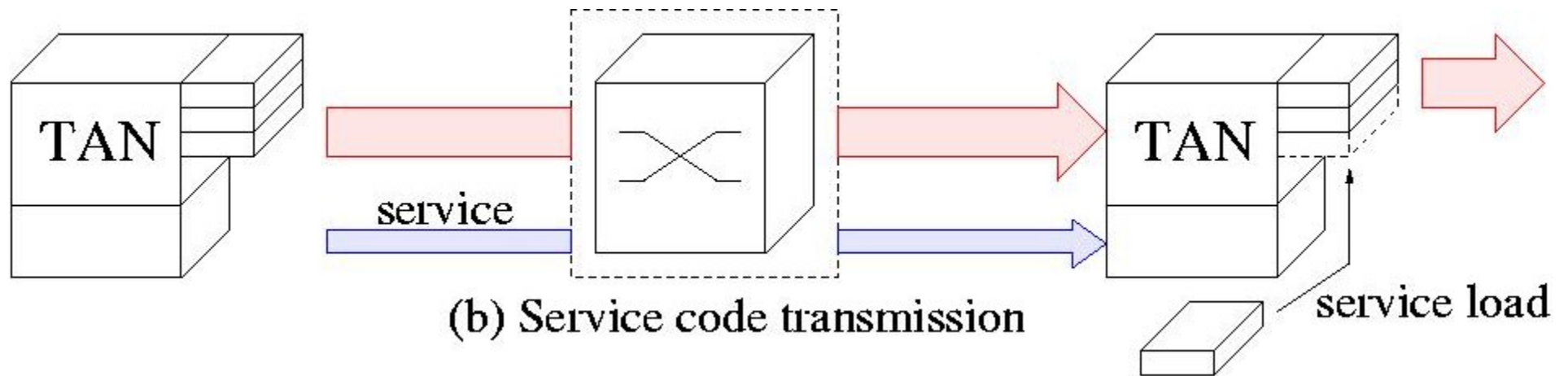
service transmission (2)

Service deployment



(a) Request of new service

Service deployment



Problem : What about data packets while the service is deployed ?

Active service deployment

- Service deployment in active nodes
 - UDP : loss of packets
 - TCP : slow down

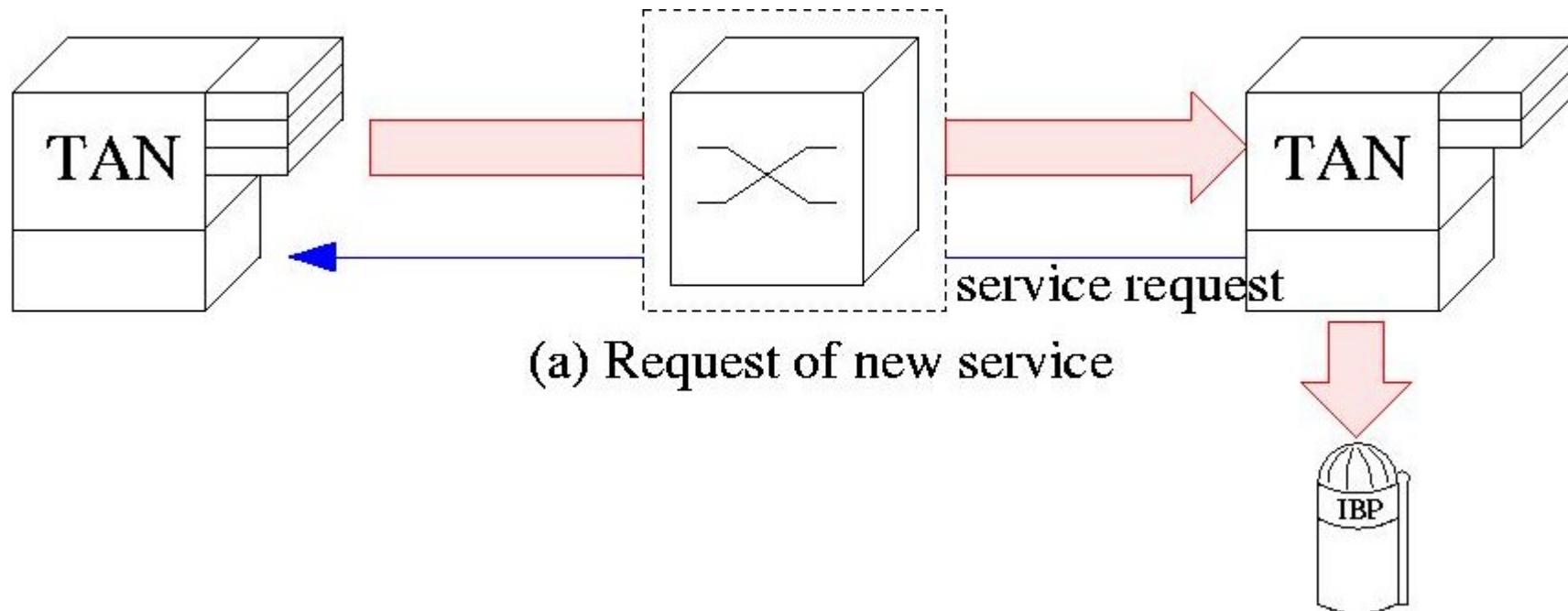
⇒ On the fly storage of packets during service install (IBP – LOCI / UTK)

⇒ <http://www.loci.cs.utk.edu>

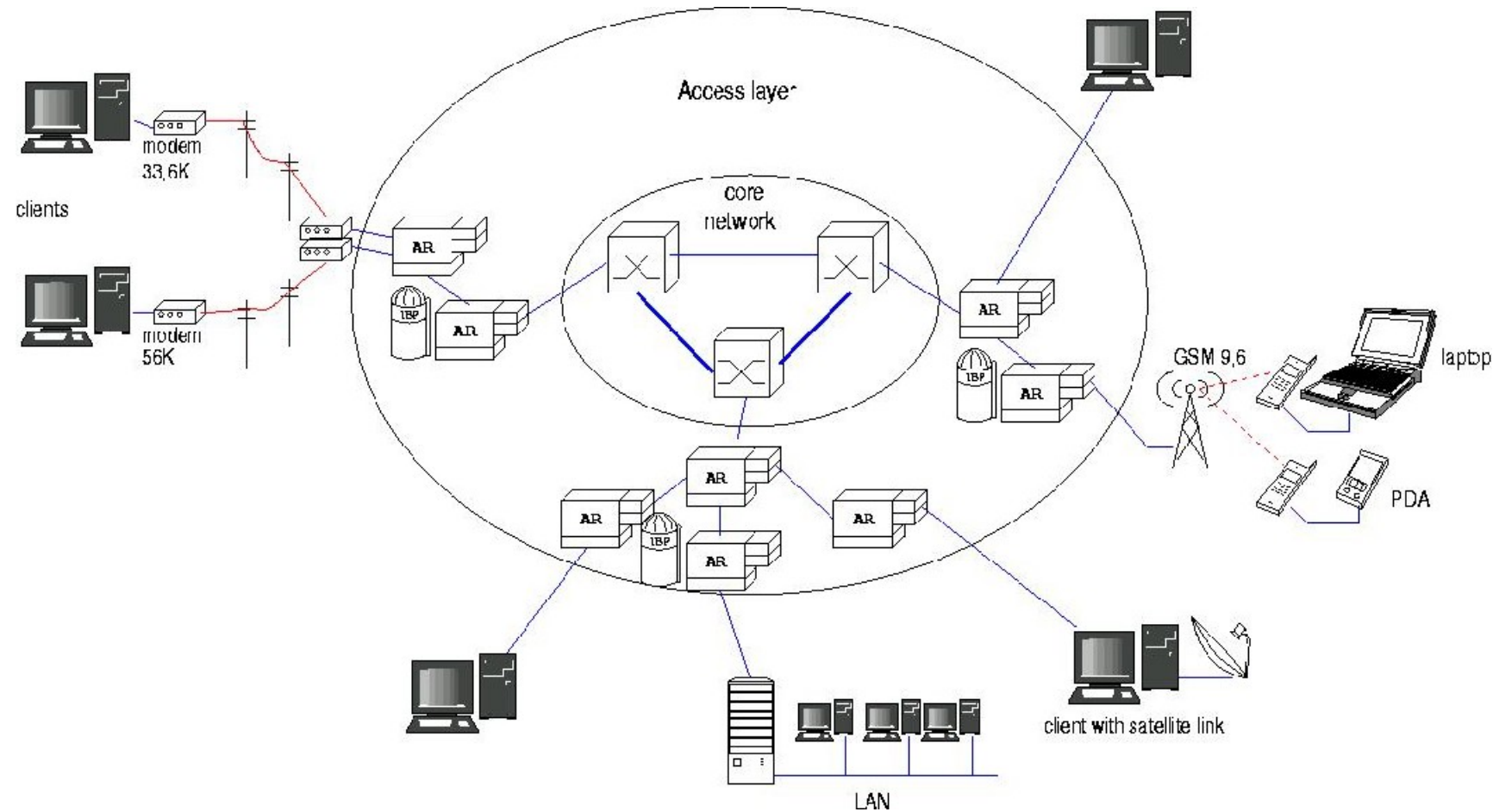
IBP logistical support for TAN

-Service Deployment -

- Data caching
 - IBPService provides caching for data reaching a TAN while the appropriate service has to be installed

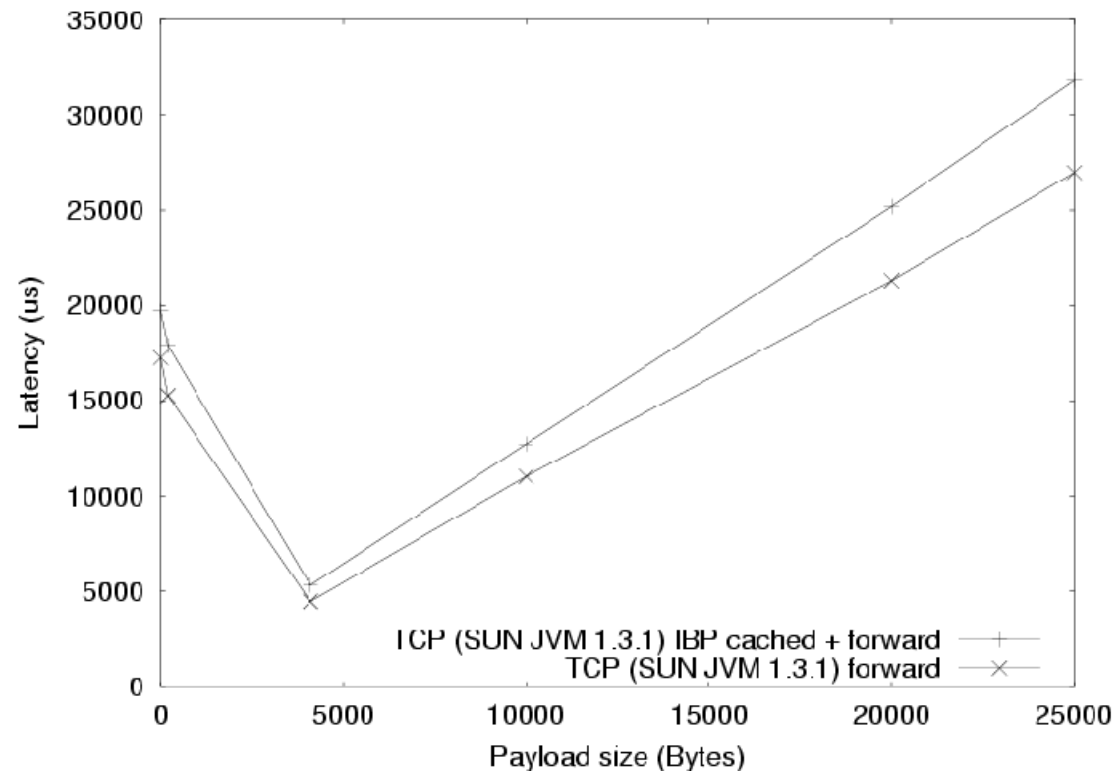


TAN /IBP locations



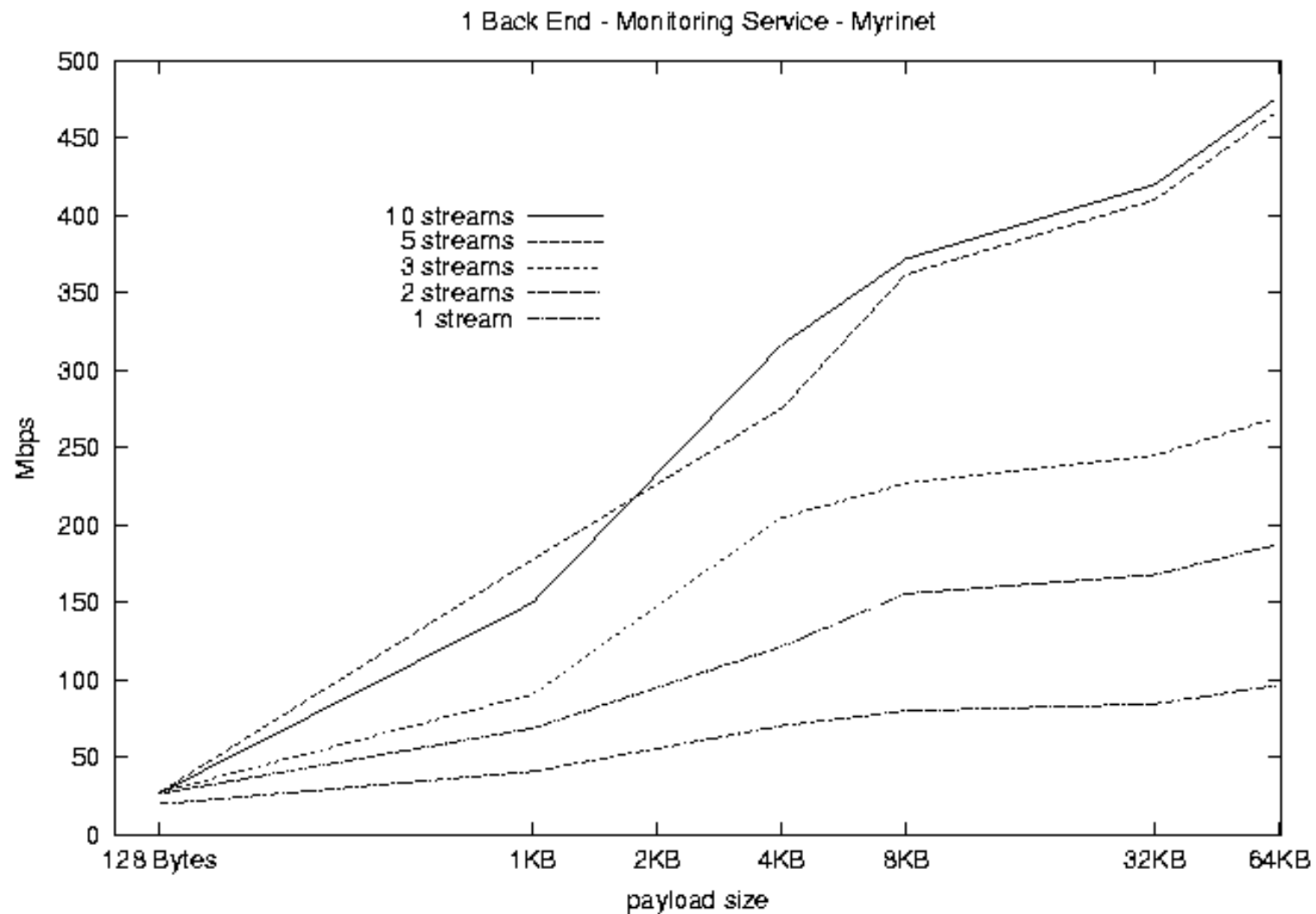
IBP services for Tamanoir

- Interaction between TAM and IBP
 - Independent processes
 - Socket-based communication
 - IBP services creates IBP Capabilities (pointers to IBP allocations)



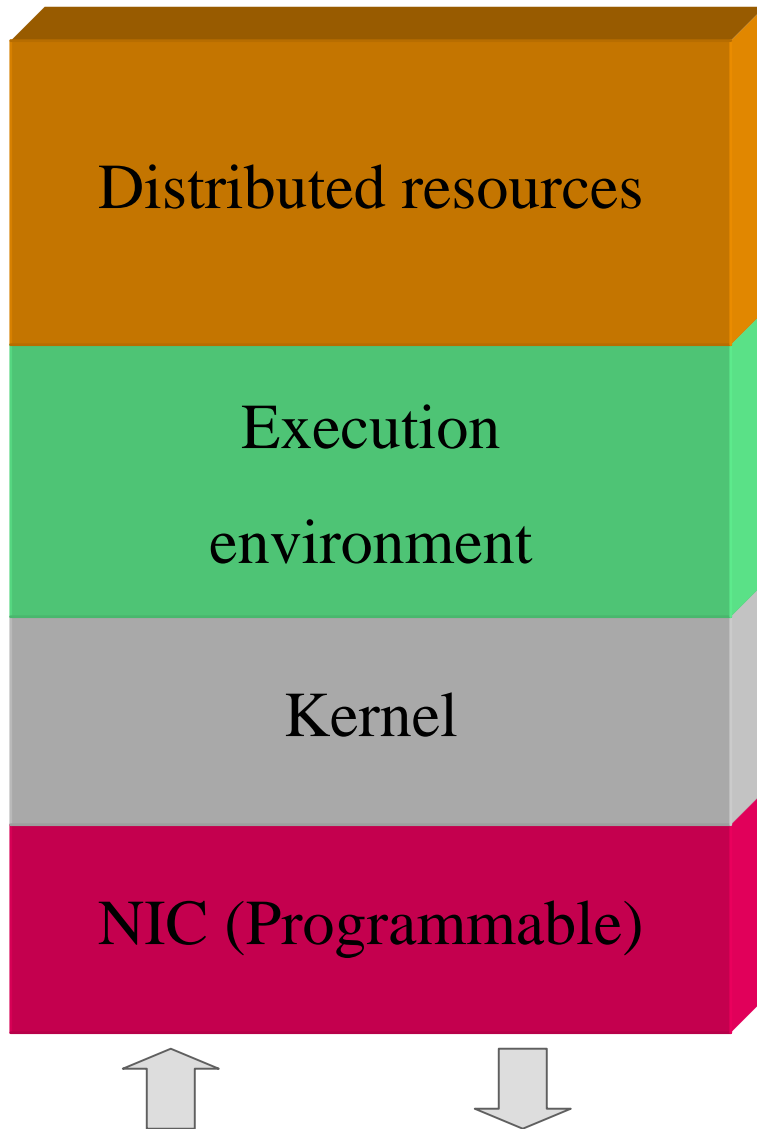
June 2003, Melbourne

Performance evaluation of user space EE



Solution / Problem

- Simple Java service in user space can support up to 450 Mbits ! ☺ (*we eat the ants !*)
- A stand-alone active node is not enough to fully support and apply active service to Gbit streams ☹
- Need to put lightweight services close to the network
- Explore kernel services



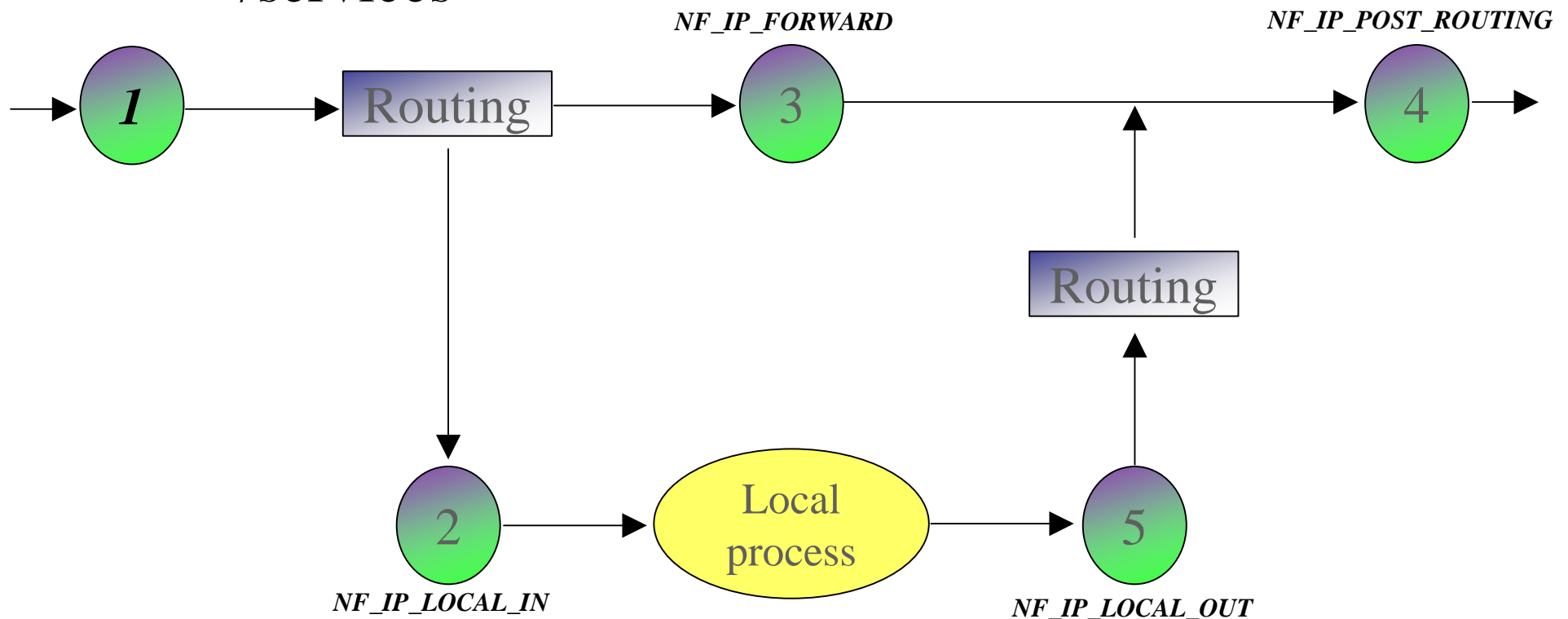
Kernel support
for active node

Kernel Support

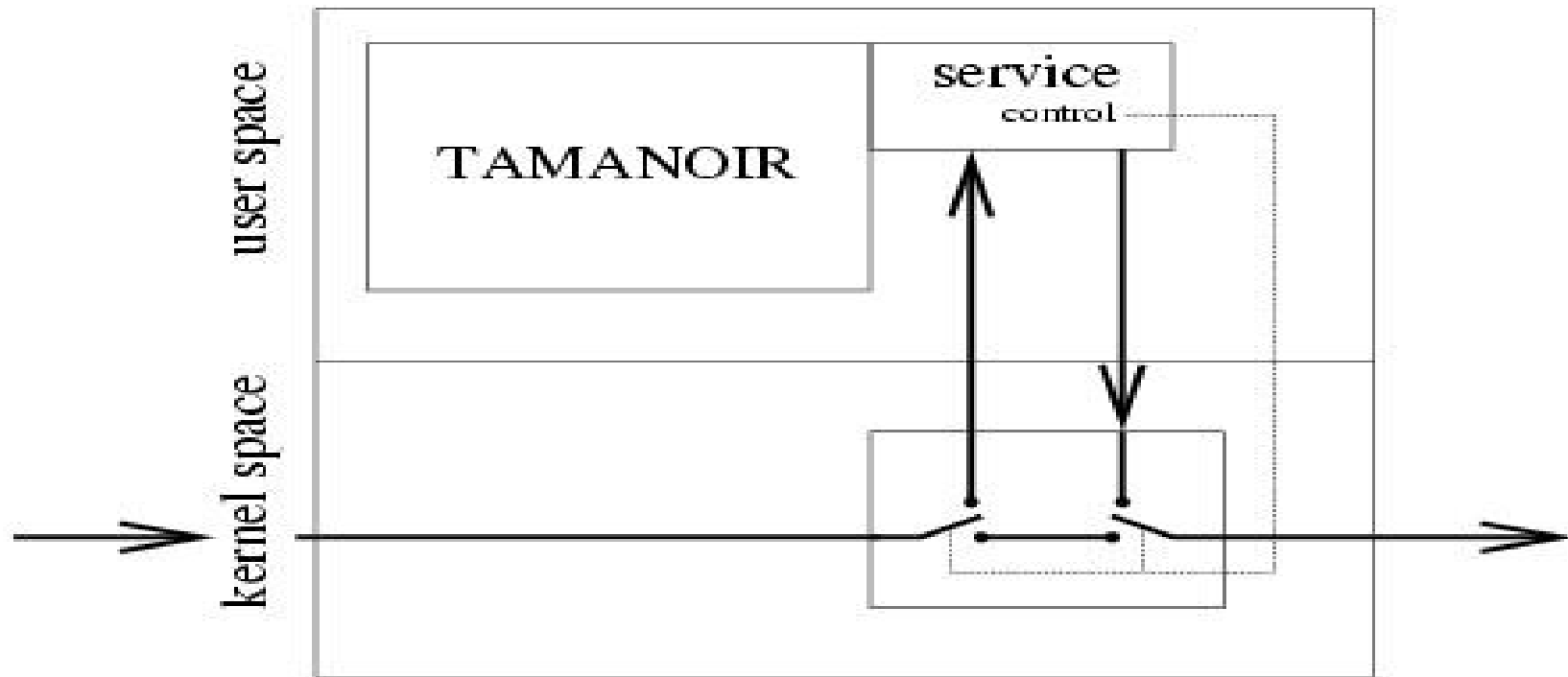
- Module inside Linux kernel : using netfilter
- Filtering ANEP packets
- Allowing packet to cross the active node through the kernel space
- Efficiently put ANEP packets inside the service in user space

Netfilter

- Protocols define hooks on the packet way inside IP stack
- Associate hooks and personalized applications /services



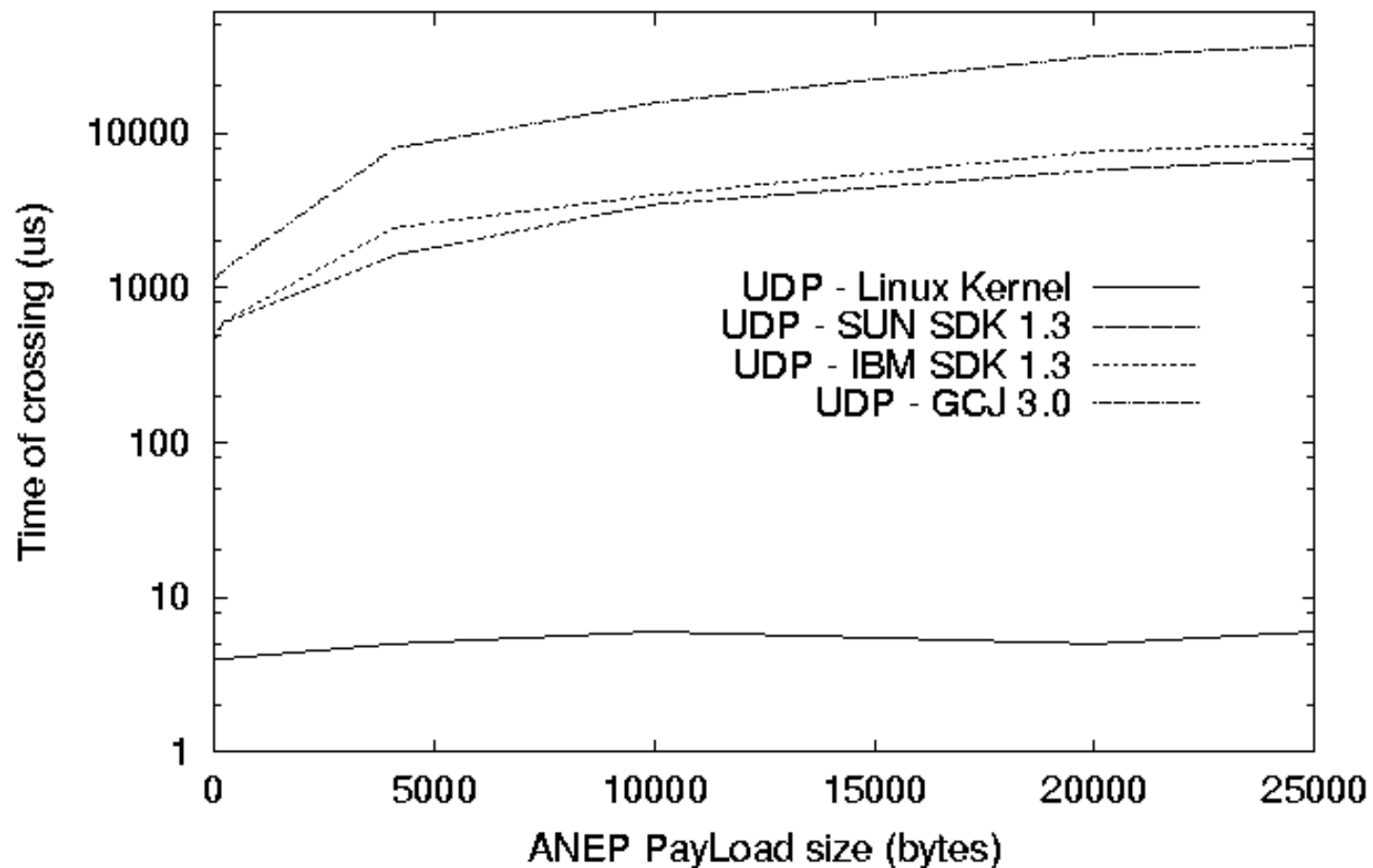
Communication between active service and OS module



- Message control / parameters

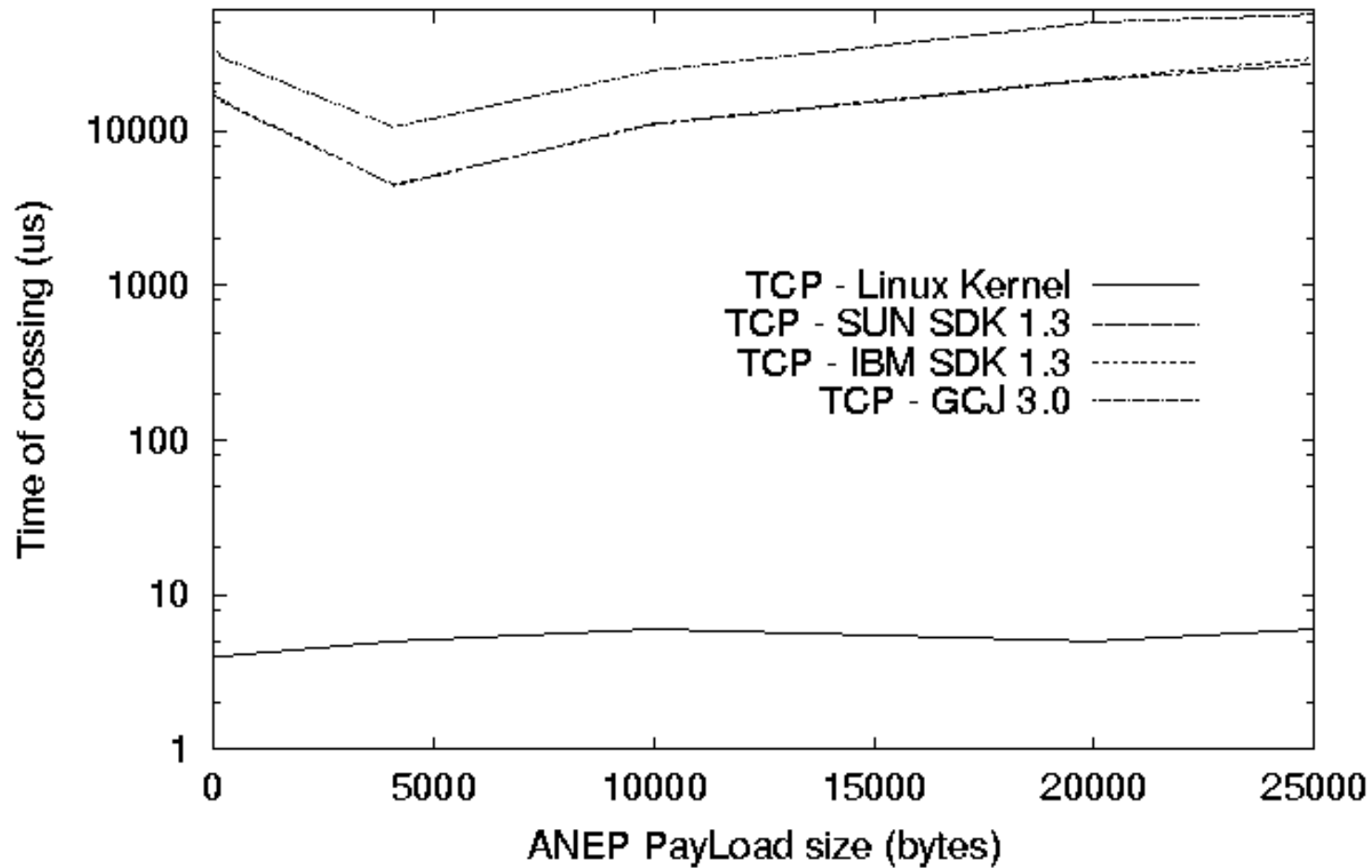
Performance evaluation

Time of crossing of ANEP packet according to the PayLoad packet



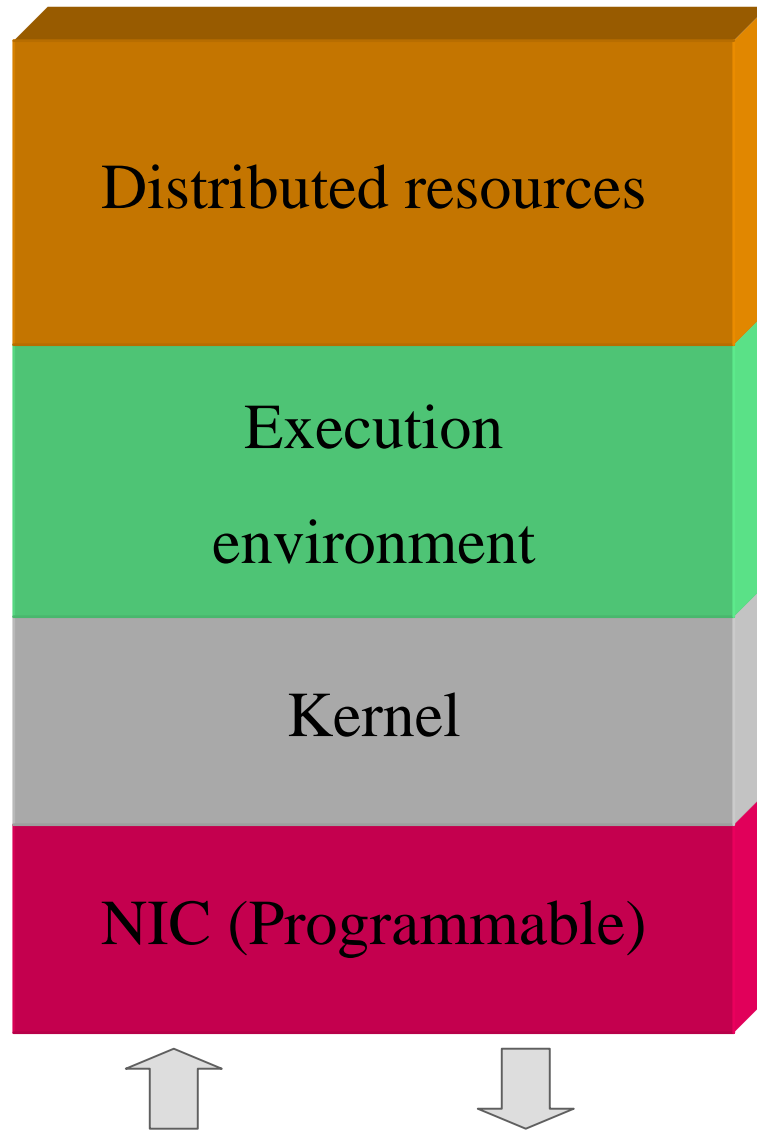
Performance evaluation

Time of crossing of ANEP packet according to the PayLoad packet



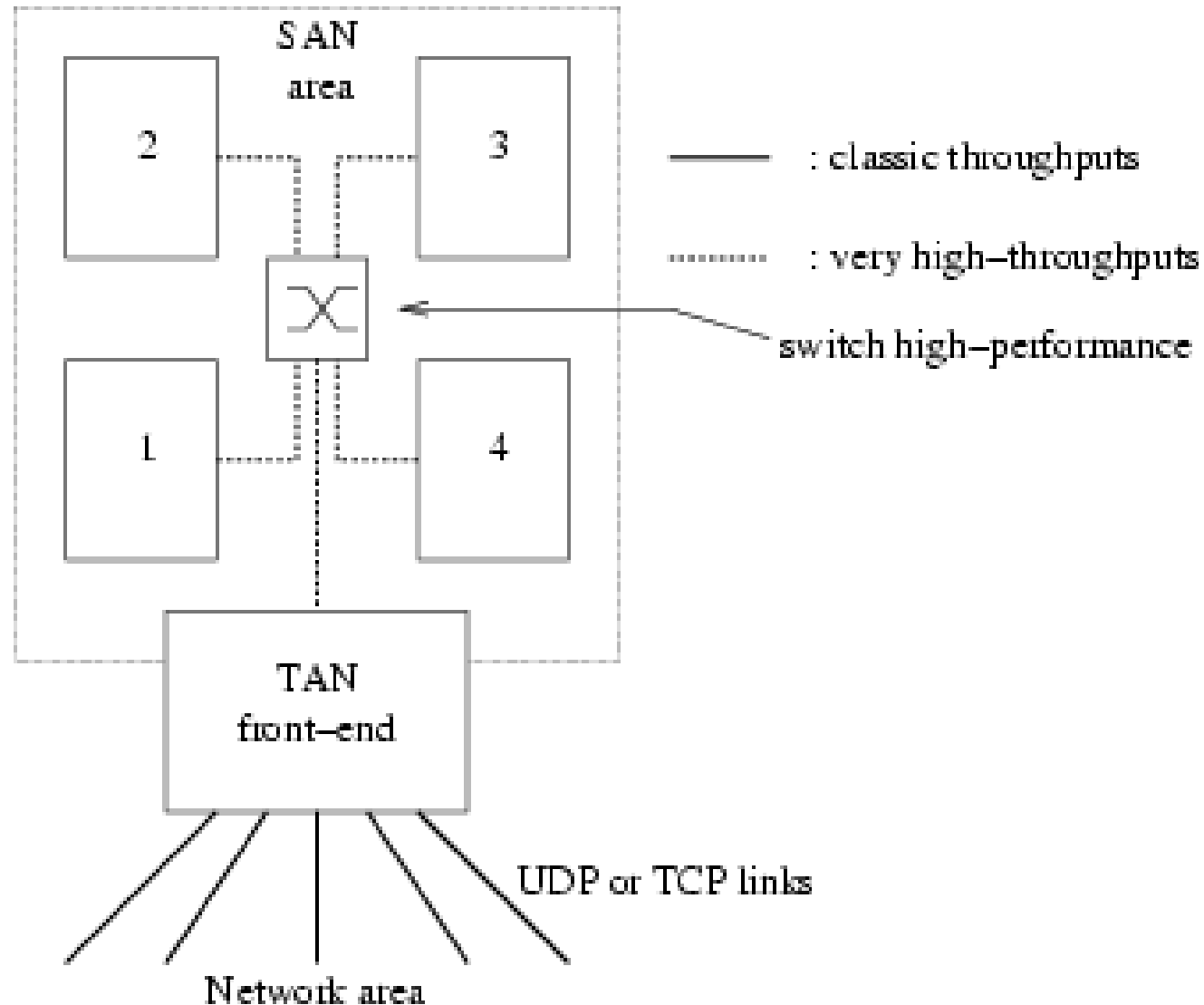
Solution / Problem

- Kernel active services greatly reduce latency for transient packets ☺
- Only applicable to lightweight services (state, CPU consumption..) ☹
- Reduce portability of active services
- What about CPU consuming services ?



Cluster in
active node

Tamanoir architecture : on a cluster

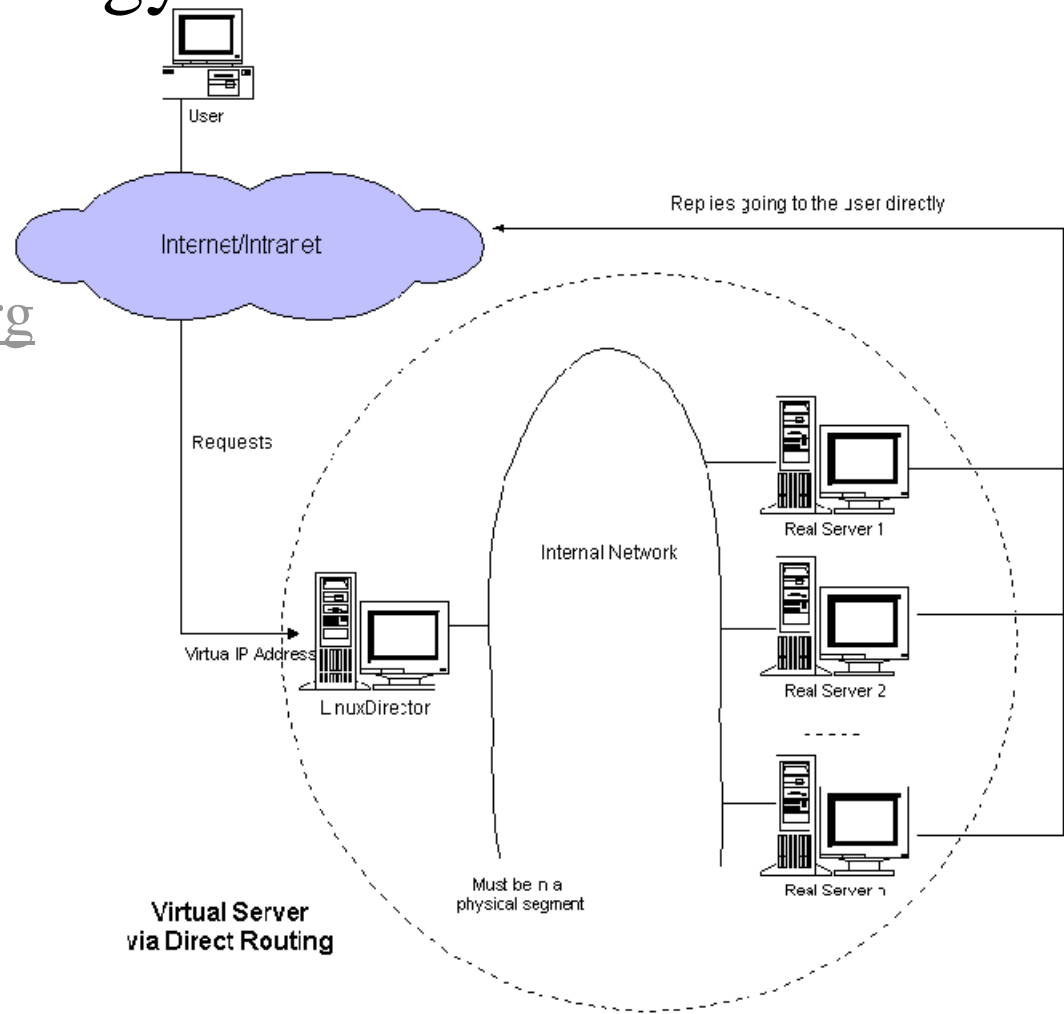


Pushing performances

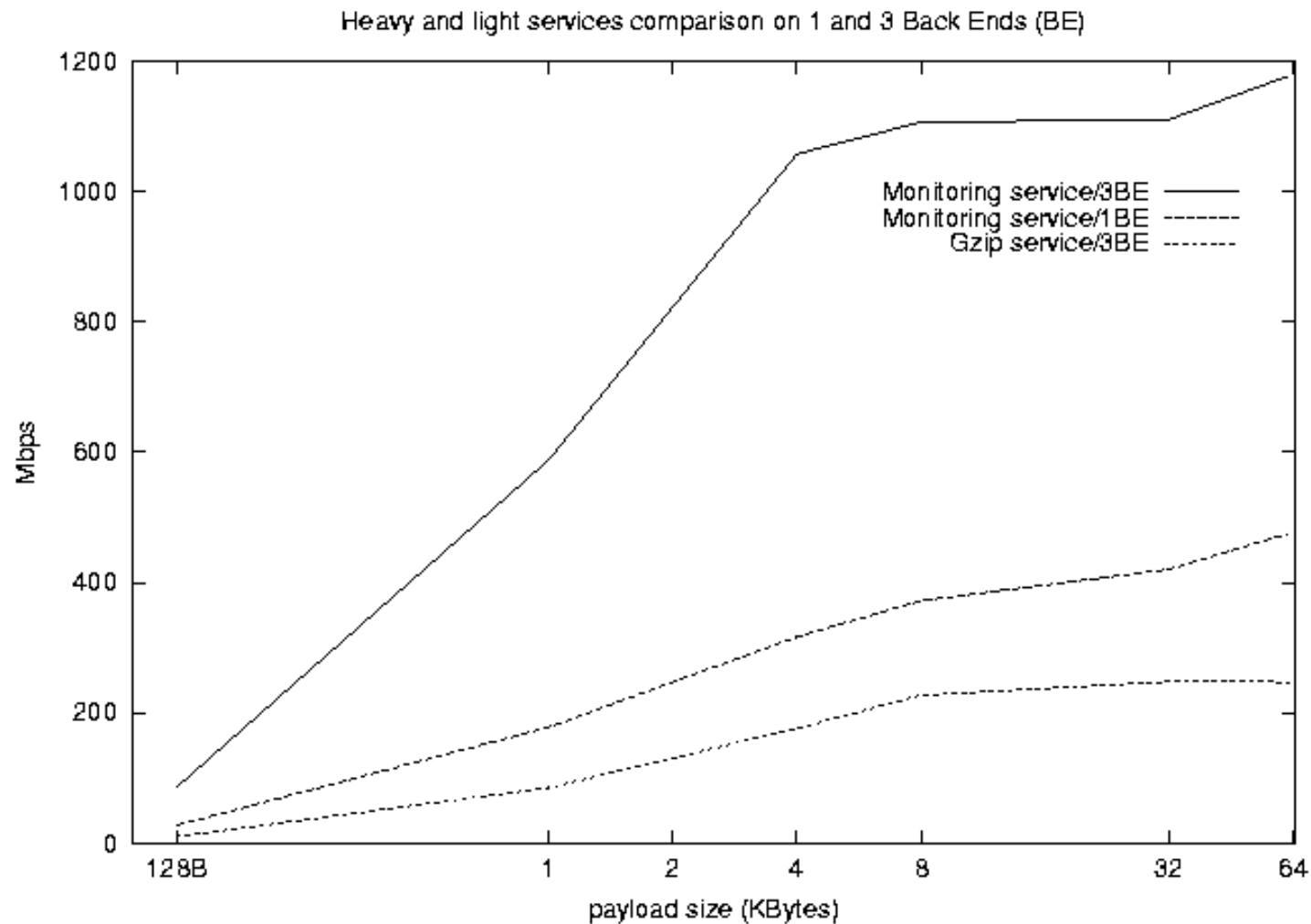
- Improving performances of active nodes : using clustering technology
- Linux Virtual Server

<http://www.linuxvirtualserver.org>

- NAT
- Direct routing
- Tunneling



Performance evaluation



Solution / Problem

- We support a full Gbit streams with a small cluster based Tamanoir active node !
😊
- We need more resources for CPU consuming active services
- How to balance workload between internal nodes of a TAN without knowing length of streams and needed services ?
(work in progress..)

Conclusion

- New level of performances : first software-based active node implementation able to support Gbits network : *world record* ! ☺
- Where can we find AN ?
 - For the moment : nowhere (except labs /Abone)
- Where could we find AN ?
 - Inside networks : around backbones, access networks
 - At home : box of network services linked with your ISP
 - On the move : active network technology for mobile / ad-hoc networks
- In search for killer application for active and programmable networks technology : Grid is the perfect candidate !

Future works

Next step :

- Active networks on Gbits networks (VTHD++)
- Deport active services in NIC
- Interaction between AN / AS and middleware (Etoile)
- Distributed active packets generator
- Deploying Tamanoir in dedicated equipments

More information :

<http://www.ens-lyon.fr/LIP/RESO/Tamanoir>



AGrid

- More information : Papers :
 - « Active networking support for the Grid », L. Lefèvre, CD Pham, P. Primet, B. Tourancheau, B. Gaidioz, J.P. Gelas, M. Maimour – International Workshop on Active Networks (IWAN 01), Philadelphia, Oct. 2001
 - « Towards the design of an active Grid », J.P. Gelas and L. Lefèvre, Workshop on Network Support and Services for Computational Grids ,International Conference on Computational Science (ICCS02), Amsterdam, April 2002

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.