VI Workshop PISATEL 2004

ERICSSON

A flexible measurement system for supporting Traffic Engineering in MPLS/DiffServ networks

Stefano Giordano

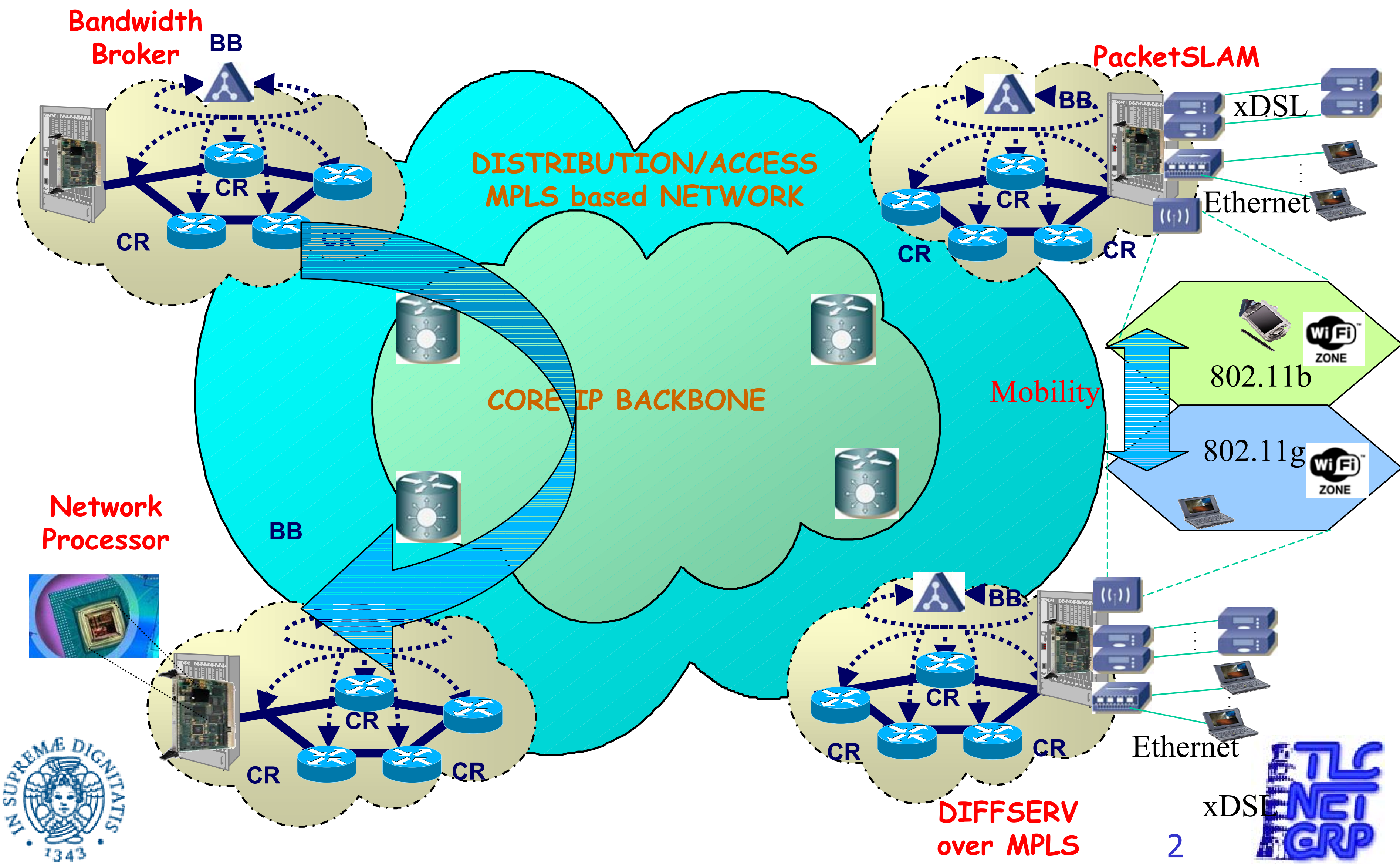
Università di Pisa

Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni

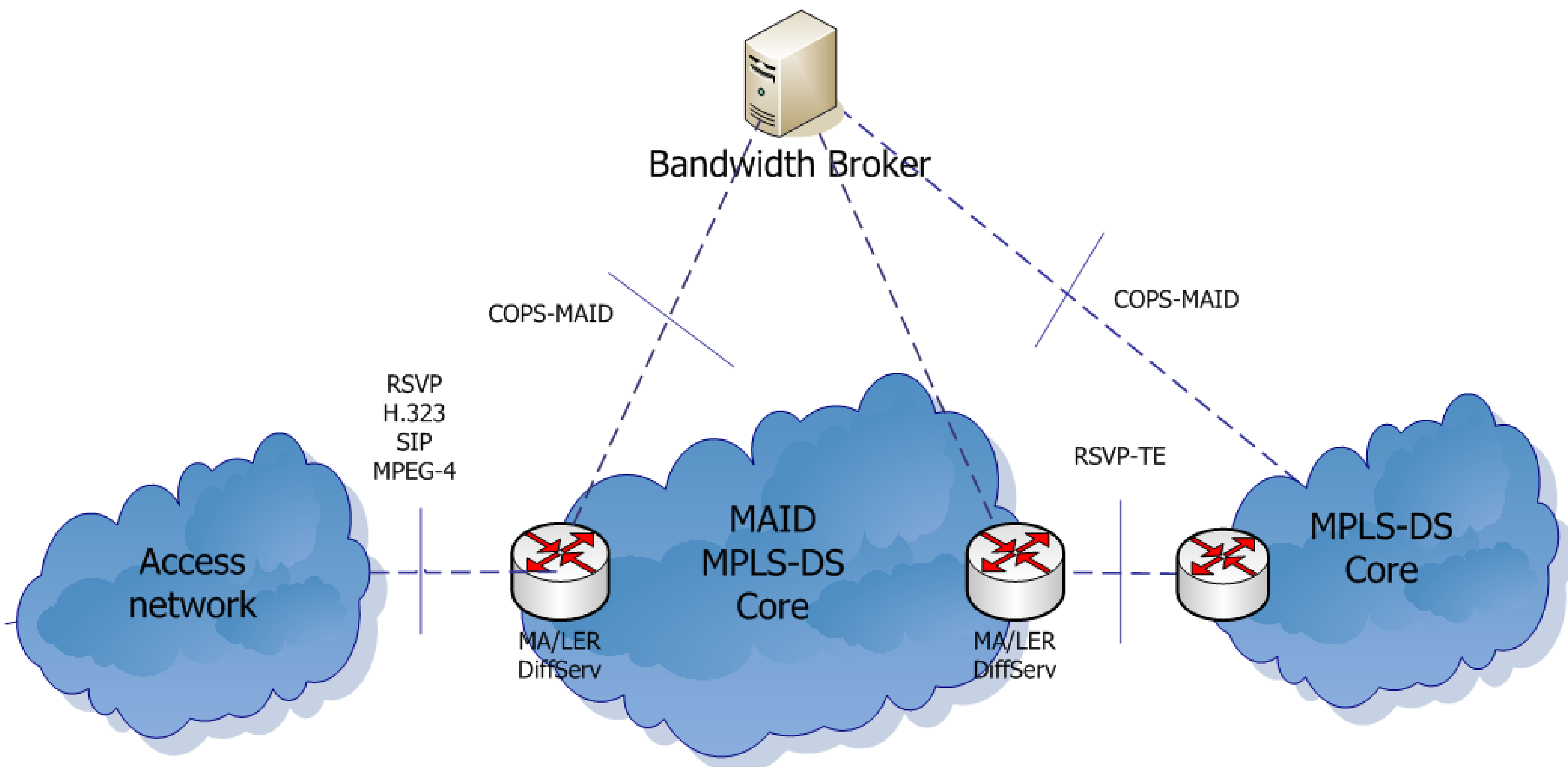
s.giordano@iet.unipi.it

netgroup.iet.unipi.it

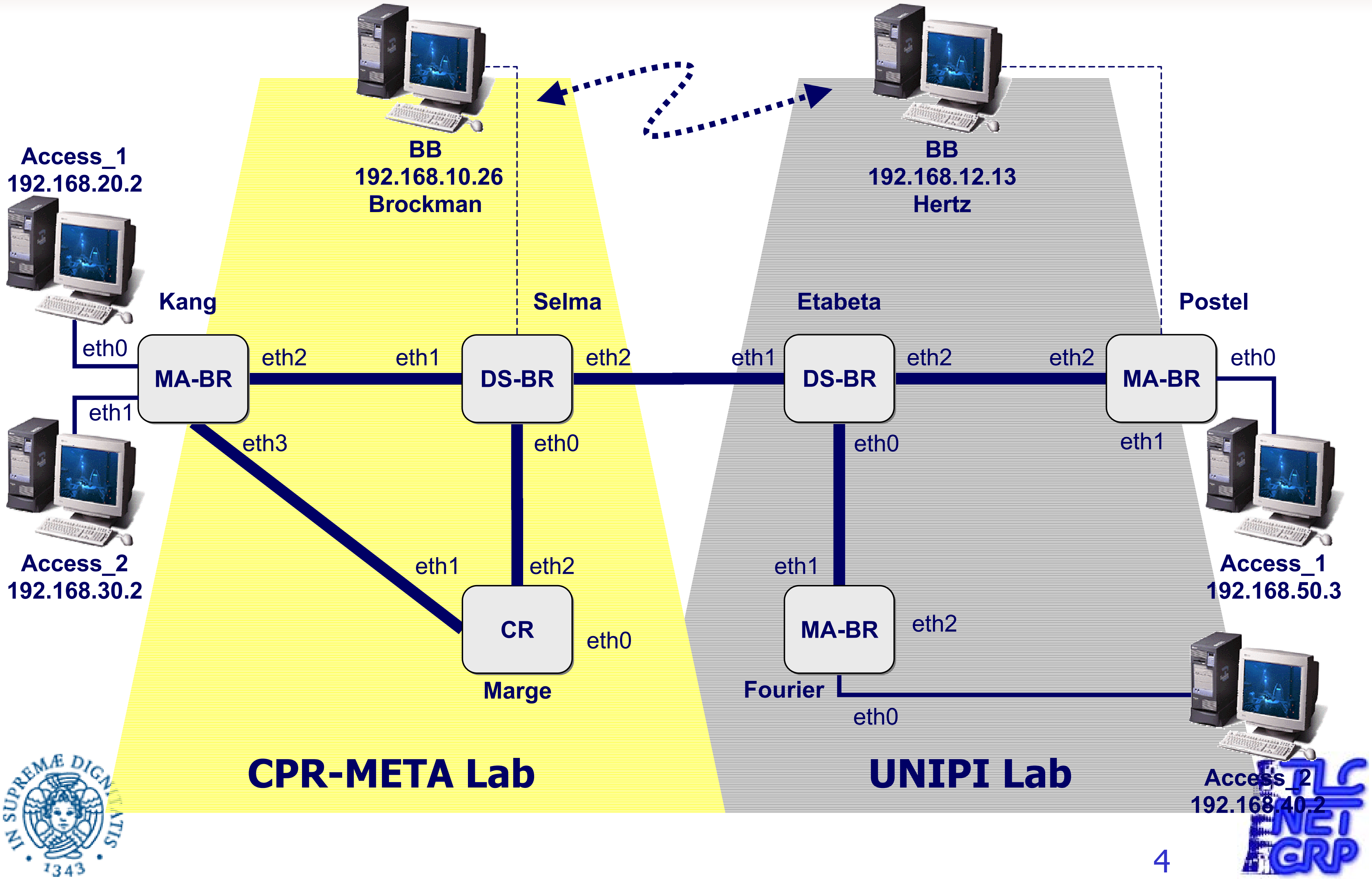
A reference scenario



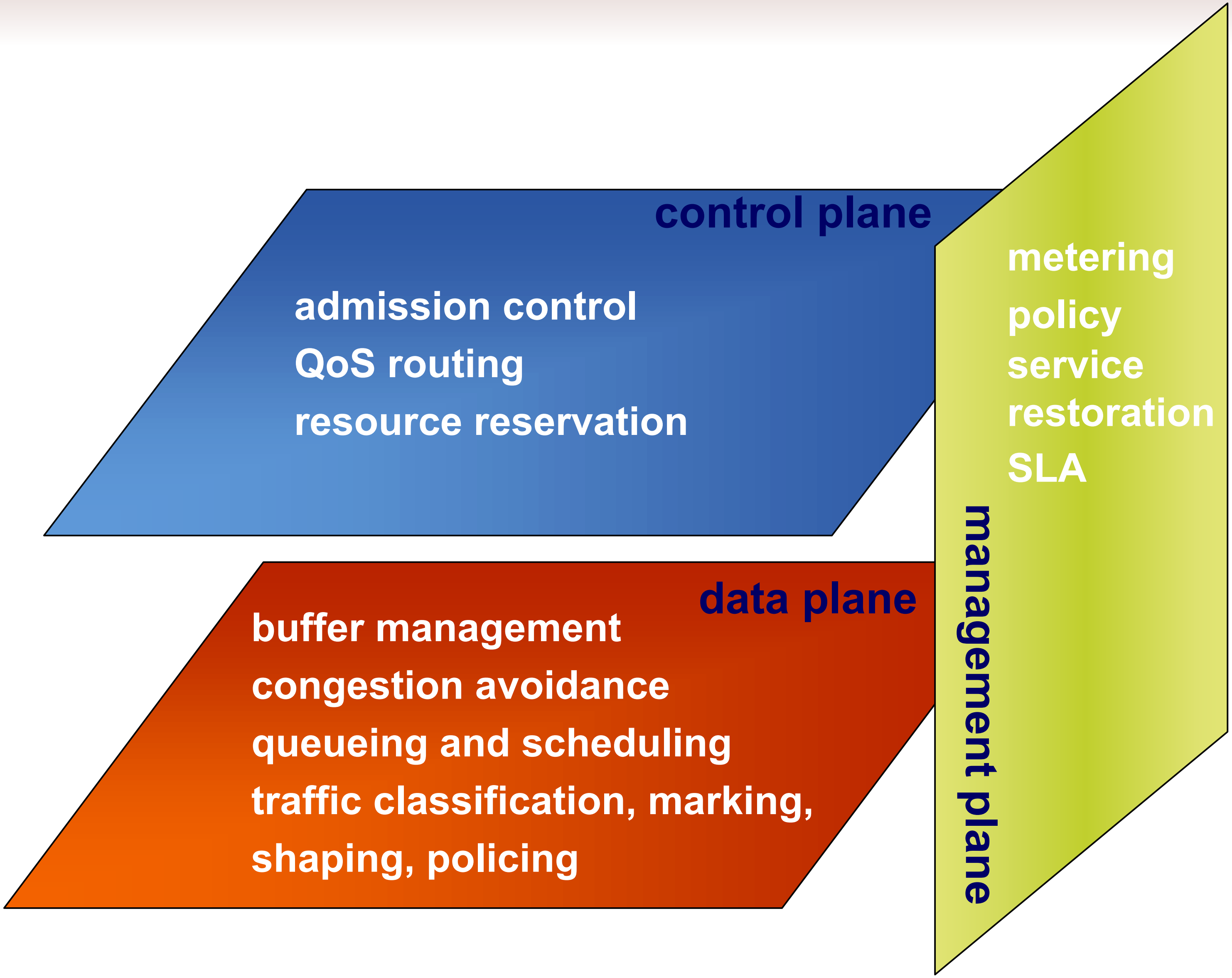
MAID (Multi-Access Inter-Domain) Architecture



Test-bed topology



QoS building blocks



QoS evolution

June 1994: rfc1633

Integrated Services (IntServ)

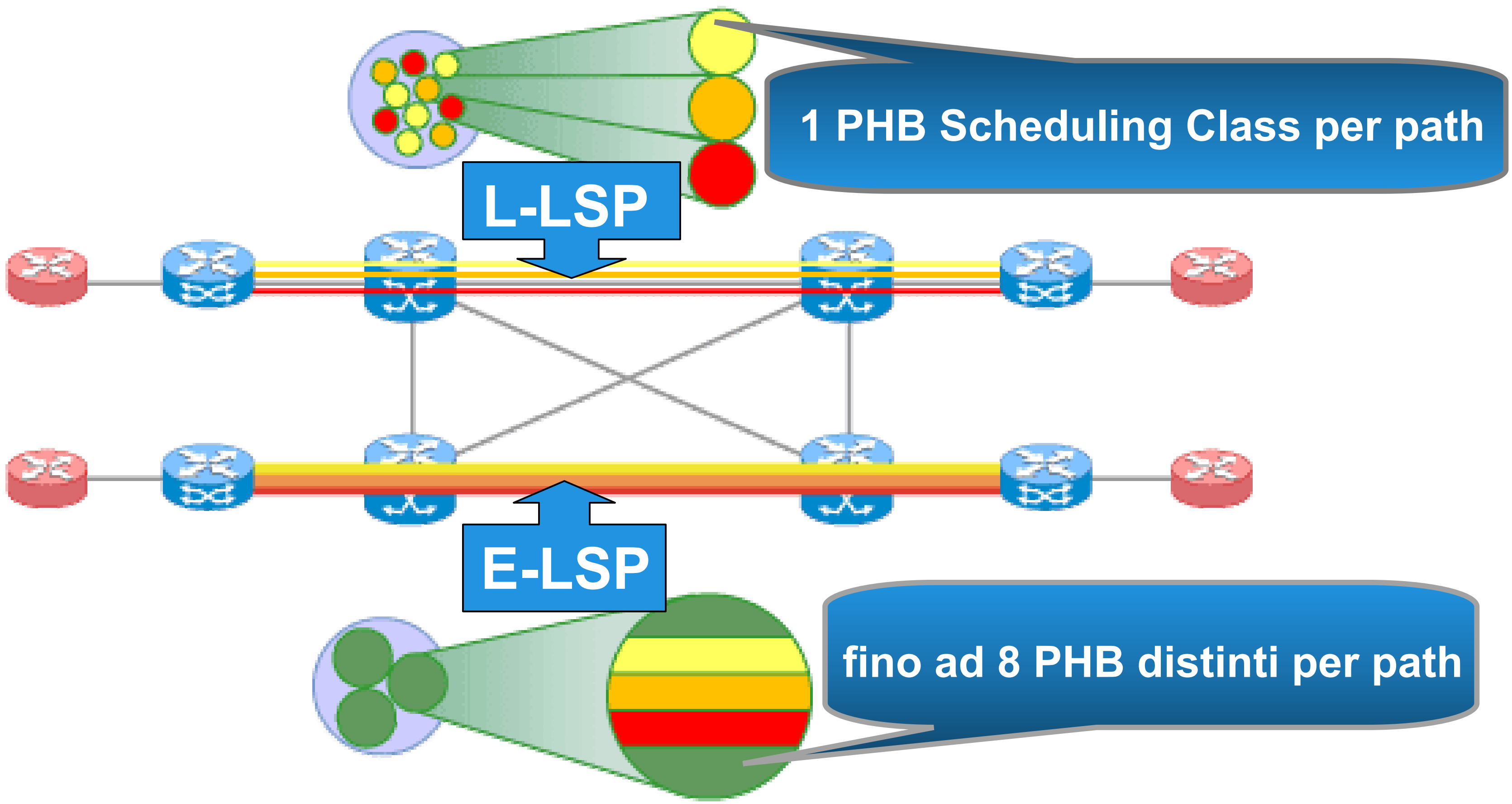
September 1997: rfc2205

Resource ReSerVation Protocol (RSVP)

December 1998: rfc2475

Differentiated Services (DiffServ)

MPLS Support of Differentiated Services
rfc3270, May 2002



Label Switched Path (LSP)

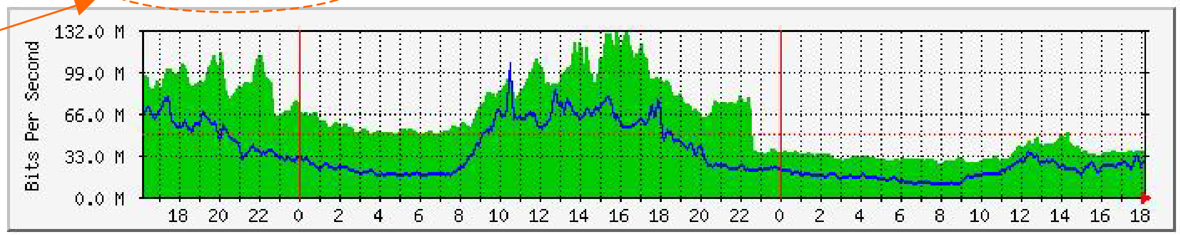
Universita' di Pisa

System:	m20pi-RE0 in
Interface:	ge-0/2/1
IP:	unipi-rtg.pi.garr.net (193.206.136.13)
Max Speed:	200 MBits/s (IP)
BGA:	50.0 MBits/s (IP)

The statistics were last updated **Wednesday, 8 December 2004 at 18:13,**
at which time **'RT-PI1-RE0'** had been up for **141 days, 4:23:33**.

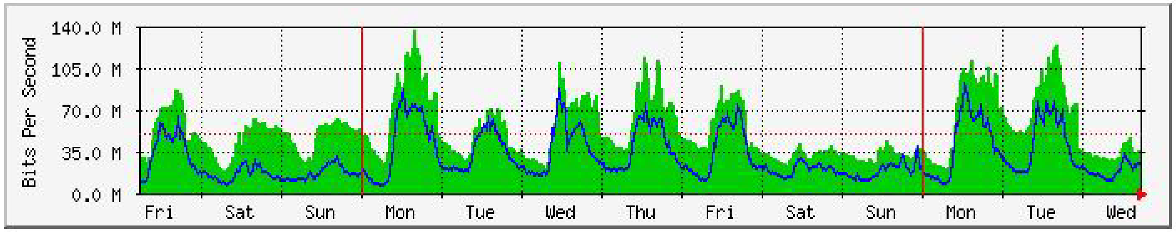
Typical Sampling Interval (and time window)

`Daily' Graph (5 Minute Average)



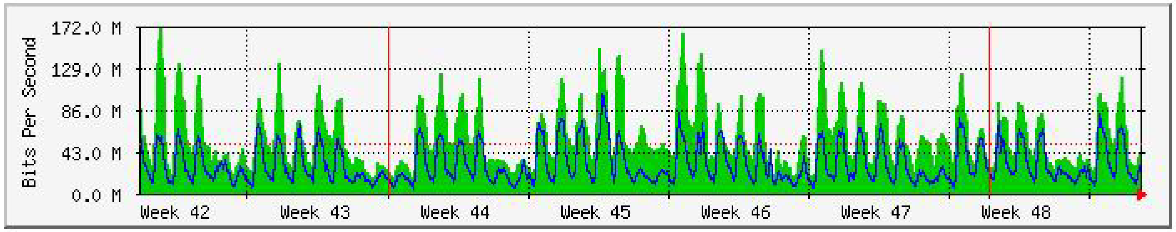
Max Bits/s: 131.1 Mb/s (262.1%) Average Bits/s: 63.0 Mb/s (126.1%) Current Bits/s: 38.2 Mb/s (76.4%)
Max Bits/s: 106.6 Mb/s (213.1%) Average Bits/s: 34.7 Mb/s (69.5%) Current Bits/s: 27.3 Mb/s (54.6%)

`Weekly' Graph (30 Minute Average)



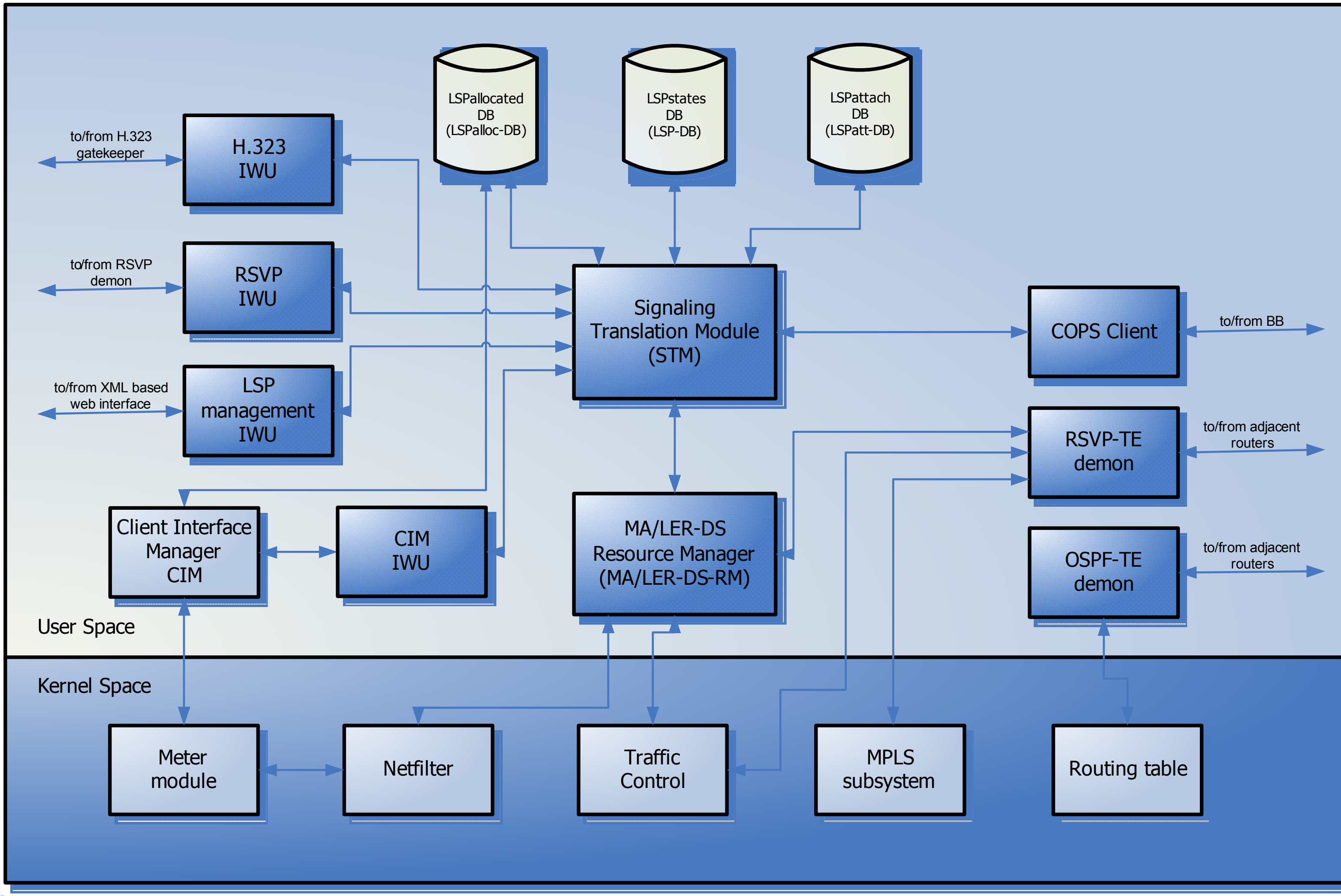
Max Bits/s: 137.8 Mb/s (275.7%) Average Bits/s: 53.9 Mb/s (107.9%) Current Bits/s: 36.7 Mb/s (73.5%)
Max Bits/s: 93.0 Mb/s (186.0%) Average Bits/s: 31.3 Mb/s (62.6%) Current Bits/s: 25.7 Mb/s (51.4%)

`Monthly' Graph (2 Hour Average)

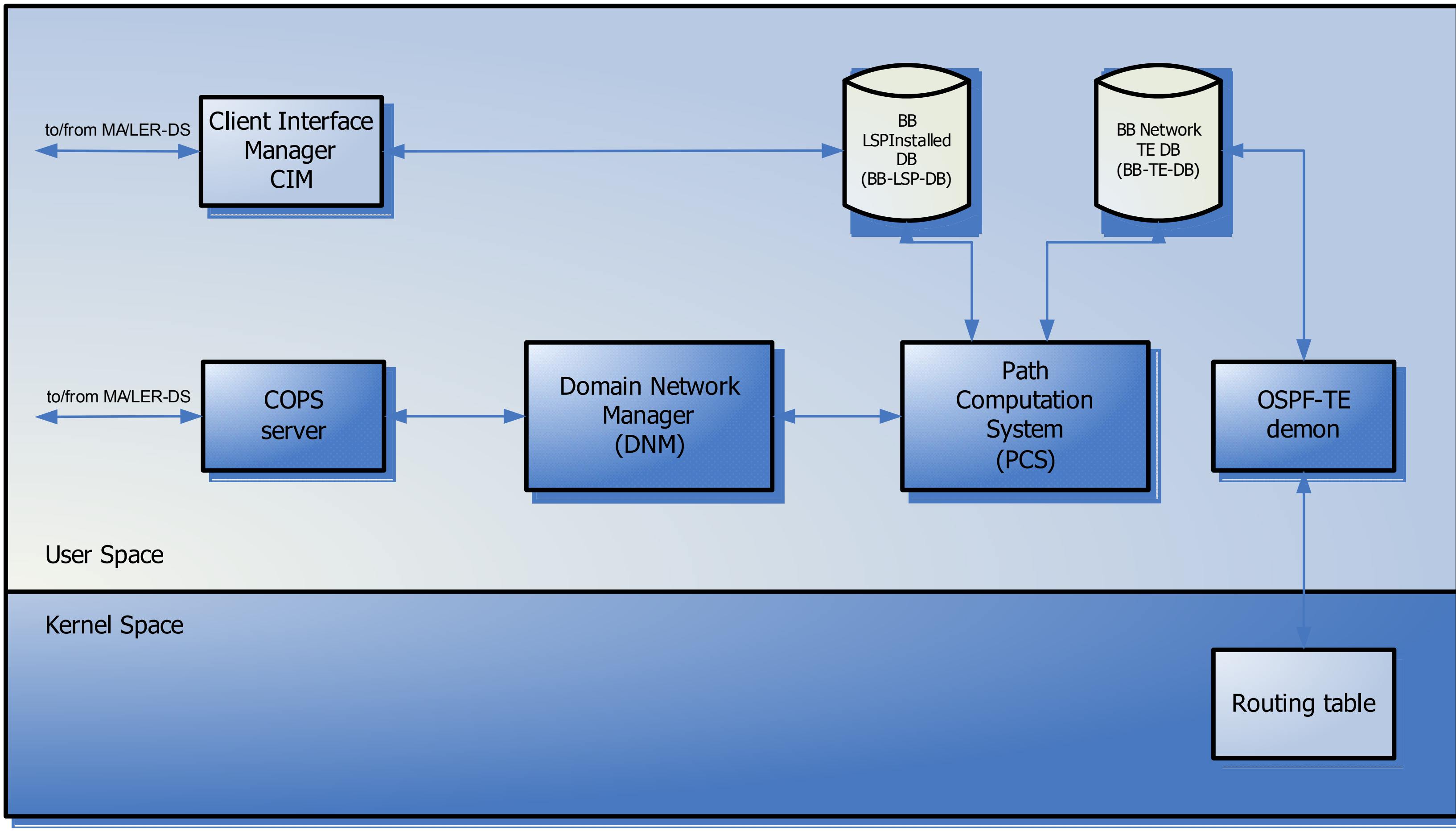


Max Bits/s: 171.3 Mb/s (342.7%) Average Bits/s: 58.5 Mb/s (117.0%) Current Bits/s: 43.6 Mb/s (87.2%)
Max Bits/s: 102.5 Mb/s (204.9%) Average Bits/s: 31.4 Mb/s (62.7%) Current Bits/s: 24.9 Mb/s (49.7%)

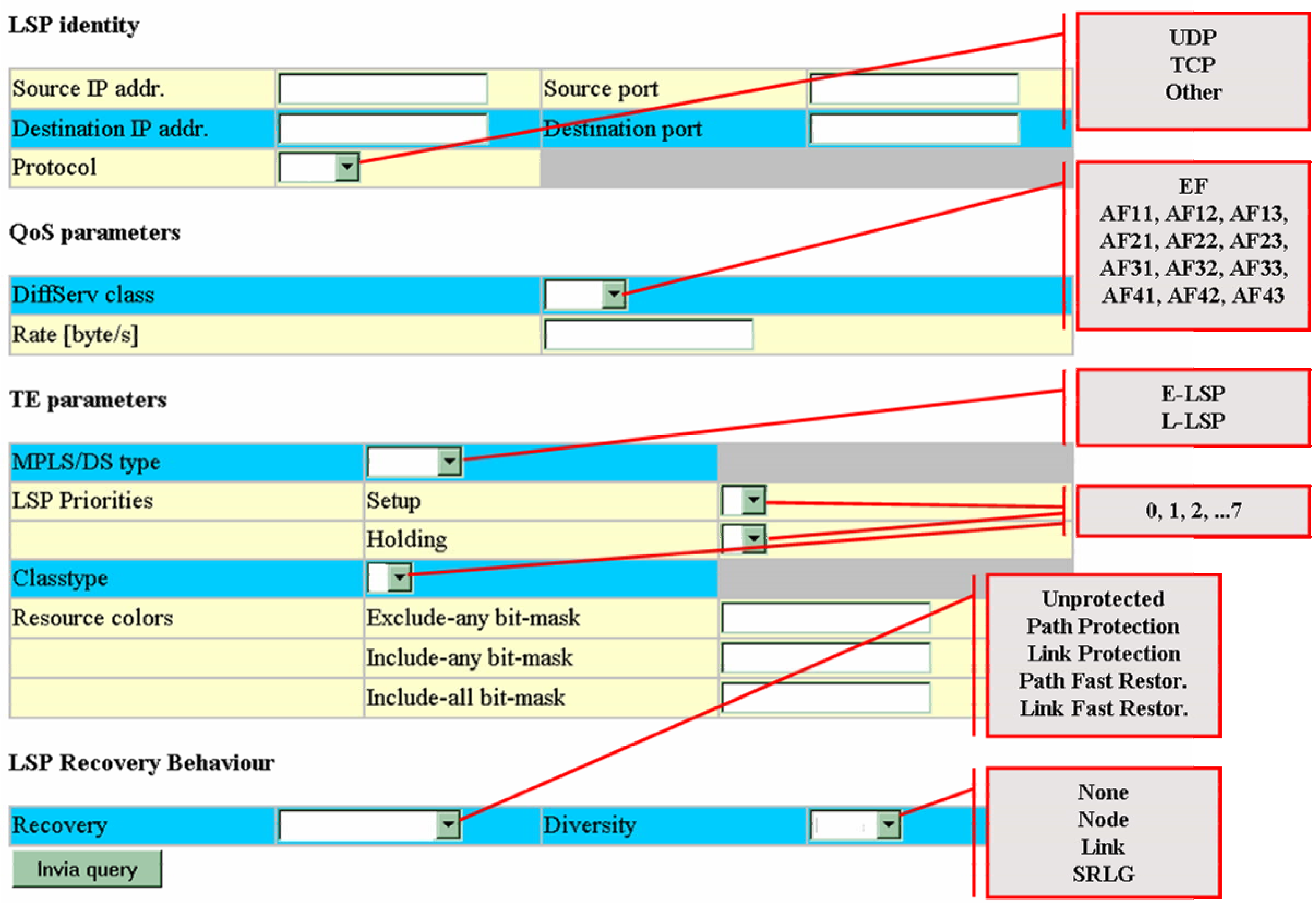
Control plane functionalities (Multiple Access Signaling – Label Edge Router)



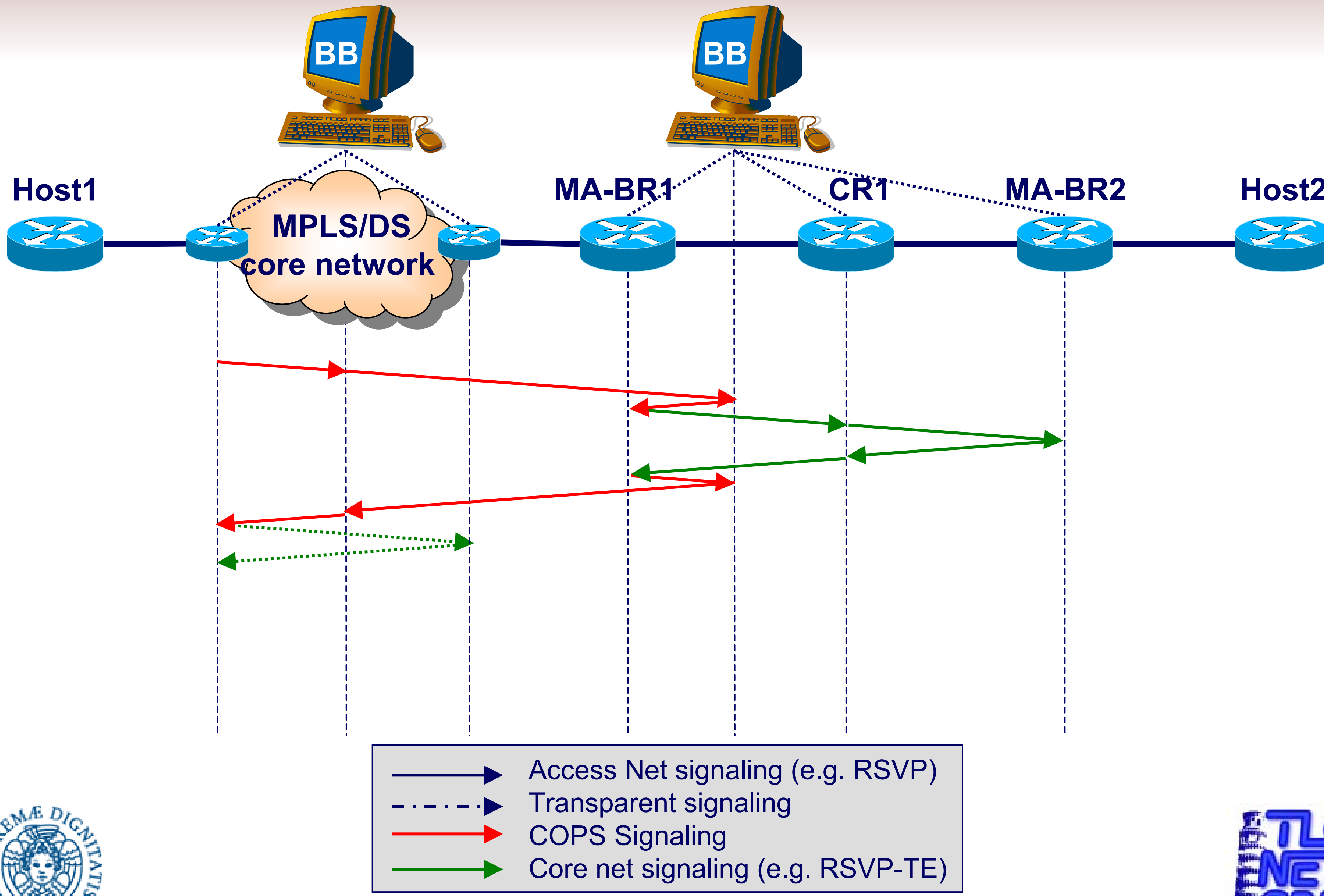
Control plane functionalities (Bandwidth Broker)



LSP management interface



Inter-domain control plane mechanisms via Web I/F



MPLS/DiffServ measurement system

- Develop a measurement system to be integrated into a MPLS/DiffServ network for:
  - management purposes (monitoring, statistics, etc…)
  - integration into the control plane to automate Resource Allocation, Admission Control (measurement and prediction based) and Traffic Engineering
- Requirements:
  - flexibility
  - configurability
  - modularity
  - minimize the impact on regular traffic dynamics

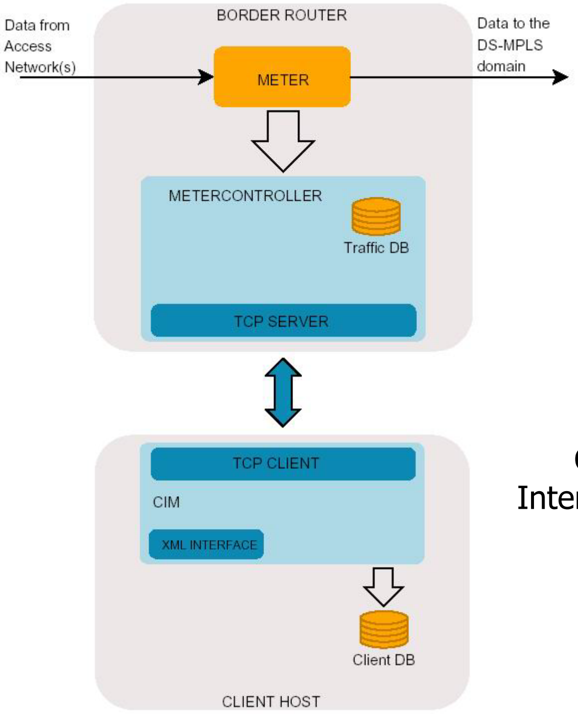
Requirements for a client server measurement system

Server-side requirements:

- **Measurements** of the traffic offered to edge routers of a DiffServ/MPLS domain taken over arbitrary and configurable time windows in the past.
- **Low latency** introduced by the system itself to reduce the unavoidable impact of measurements onto network performance.
- **Sampling** of network traffic with accurate timing with low sampling jitter, with no use of busy wait.
- **Timestamping** of traffic samples in order to enable time series processing, including prediction.
- **Data storage** of per-flow traffic time series in a proper database.
- **Remotization** of the measurement system in order to enable remote monitoring and management.
- **Multiple Client Support** in order to enable simultaneous operations, such as traffic monitoring, estimation and prediction.
- **Quick** delivery of information to client/s to prevent that high processing delays impact the relevance of measurements.
- **Interworking** with other router subsystems such as *routing*, *MPLS*, *RSVP-TE* daemon, etc.

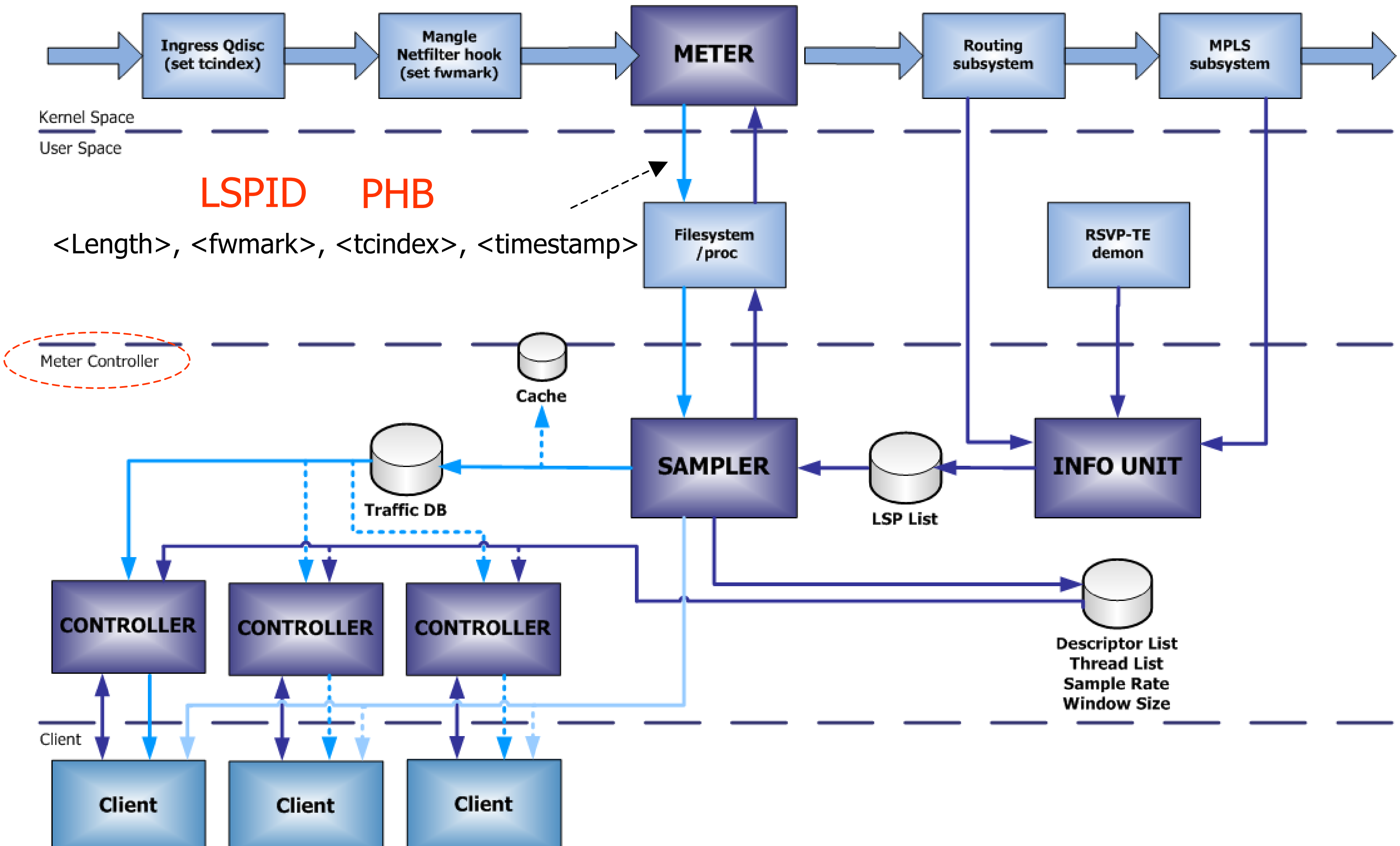
Client-side requirements:

- **Estimation**. In order to avoid unnecessary CPU load on the border router, the computation of the estimates must be performed in the client side.
- **Threshold crossing detection**. When the traffic load of a flow overcomes a preconfigured level the network control plane must react accordingly. The occurrence of a threshold crossing event must be notified as soon as possible to the control plane.

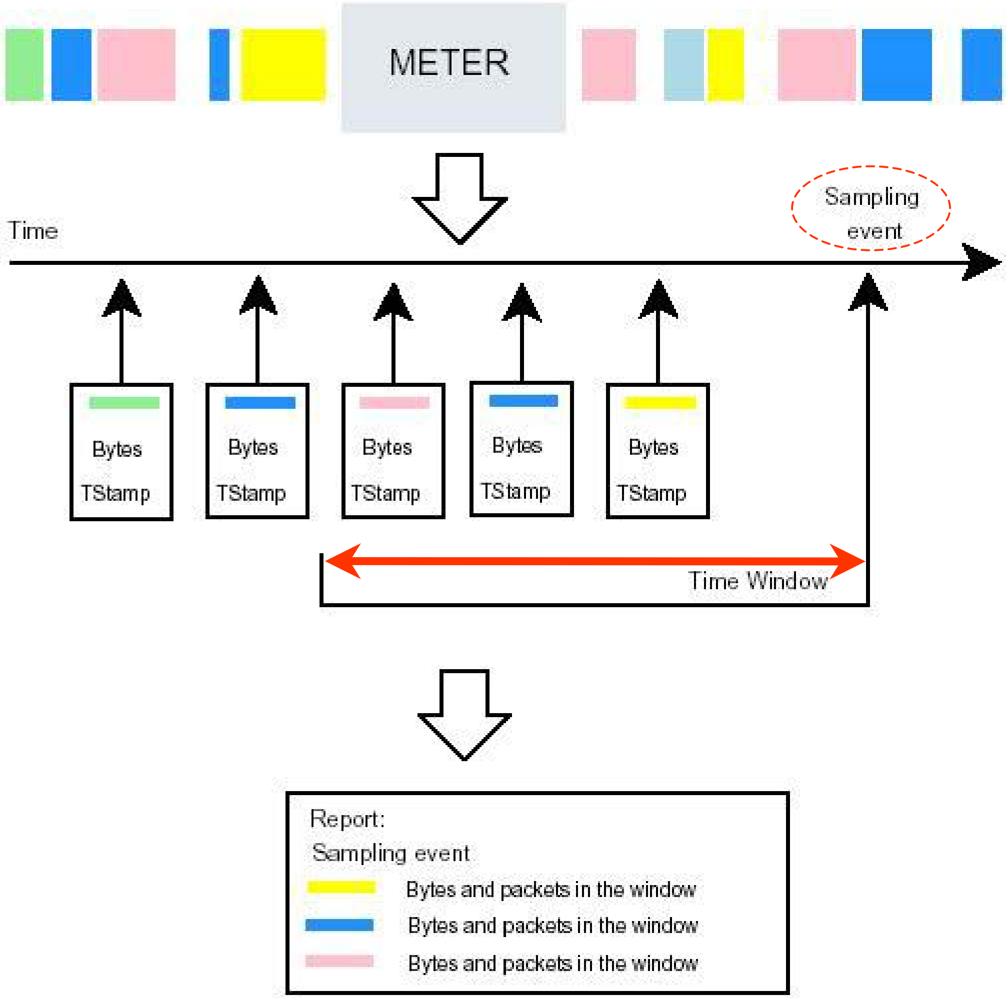


CIM (Client Interface Manager)

At a glance...

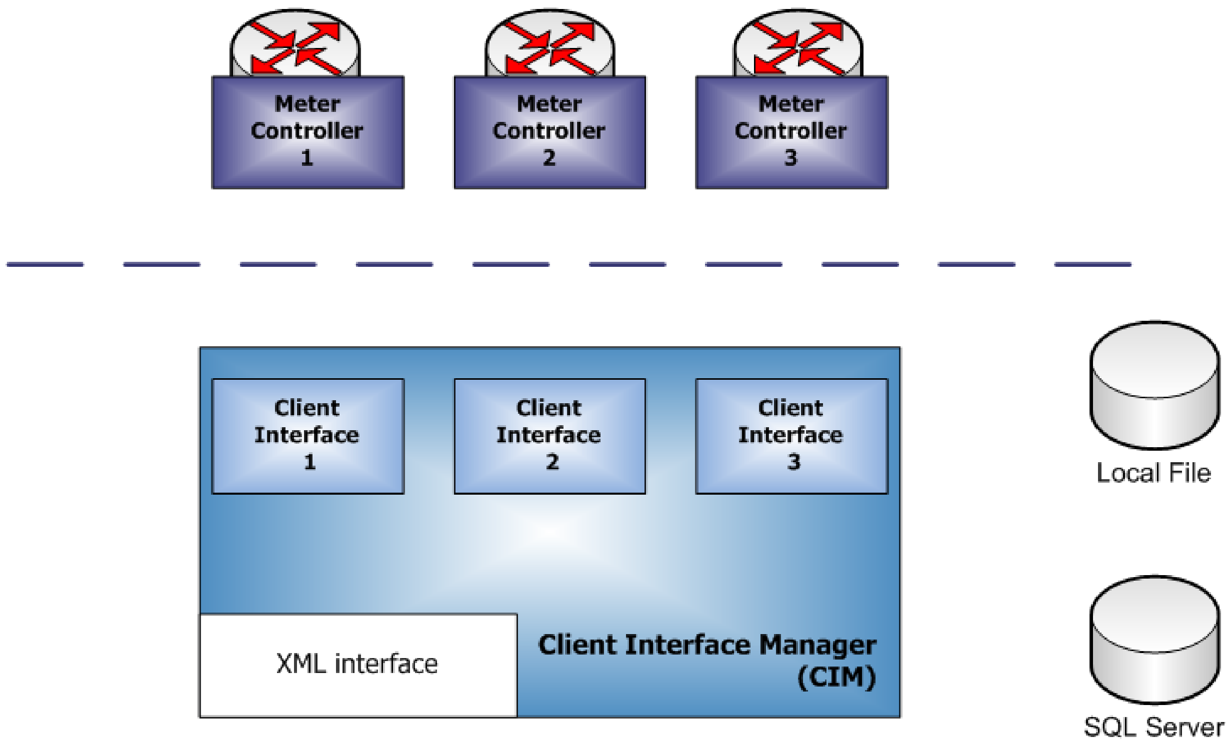

Ingress Qdisc (set tcindex)
Mangle Netfilter hook (set fwmark)
METER
Routing subsystem
MPLS subsystem
Kernel Space
User Space
LSPID
PHB
<Length>, <fwmark>, <tcindex>, <timestamp>
Filesystem /proc
RSVP-TE demon
Meter Controller
Cache
Traffic DB
SAMPLER
LSP List
INFO UNIT
CONTROLLER
CONTROLLER
CONTROLLER
Descriptor List
Thread List
Sample Rate
Window Size
Client
Client
Client
Client

Operational mechanism

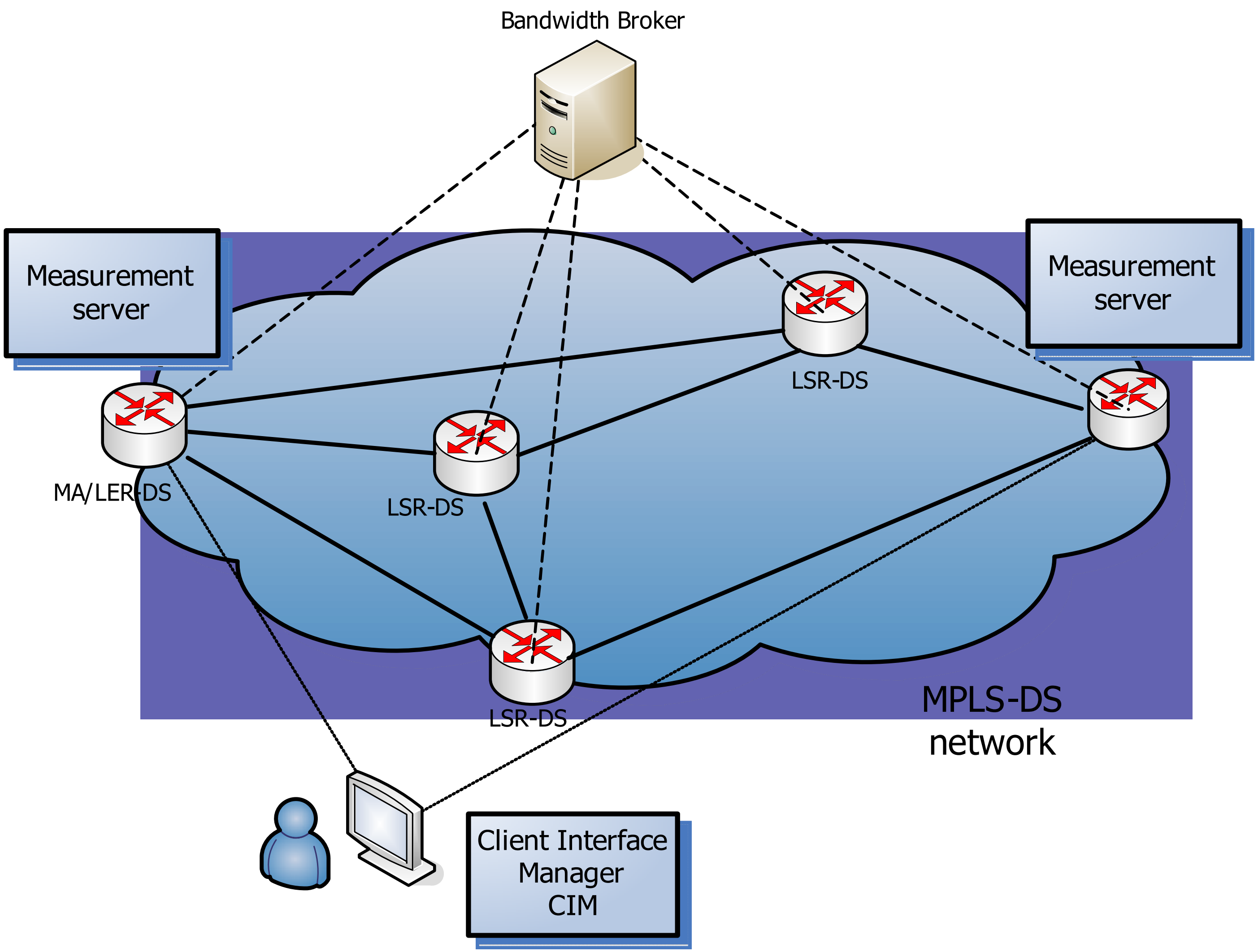


Client Interface Manager

- Client Interface Manager (CIM)
  - application that manages the Client Interface connected to the associated domain Edge Routers (coordinates modules, mechanisms and events in the client side)

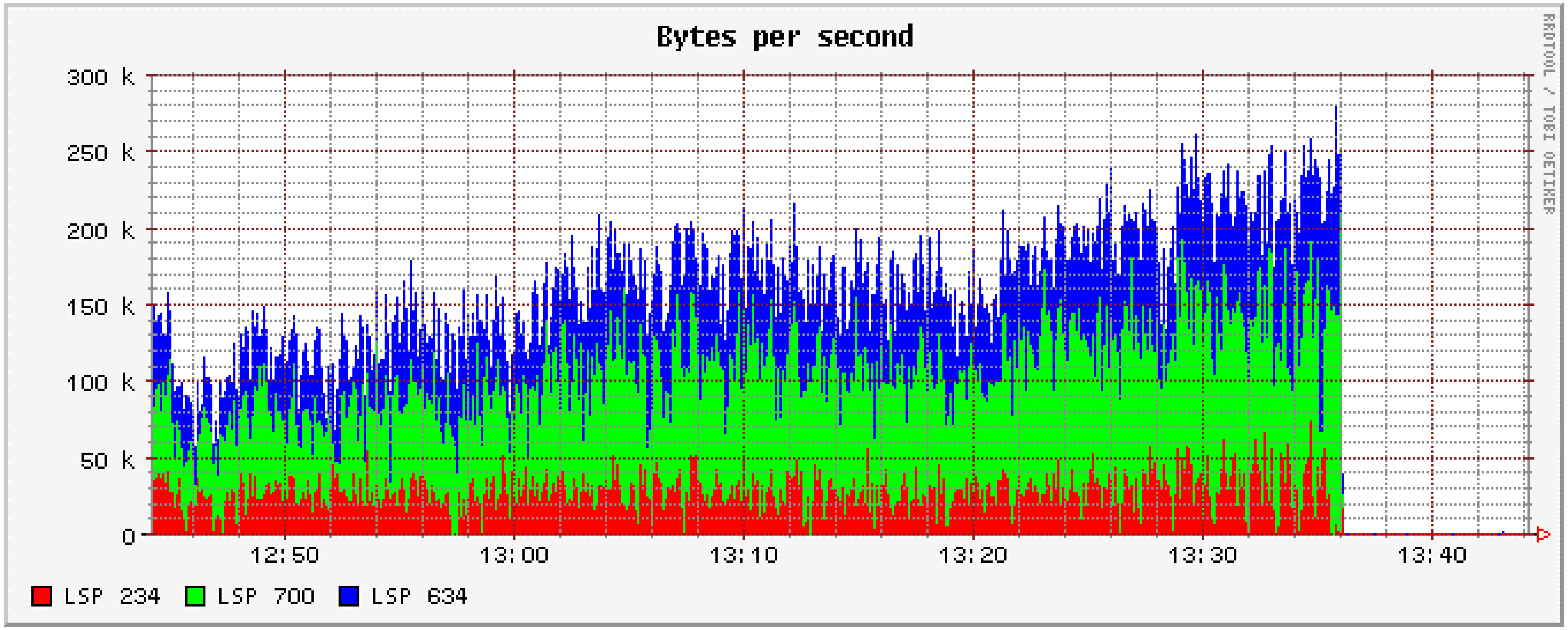


Application: monitoring/management



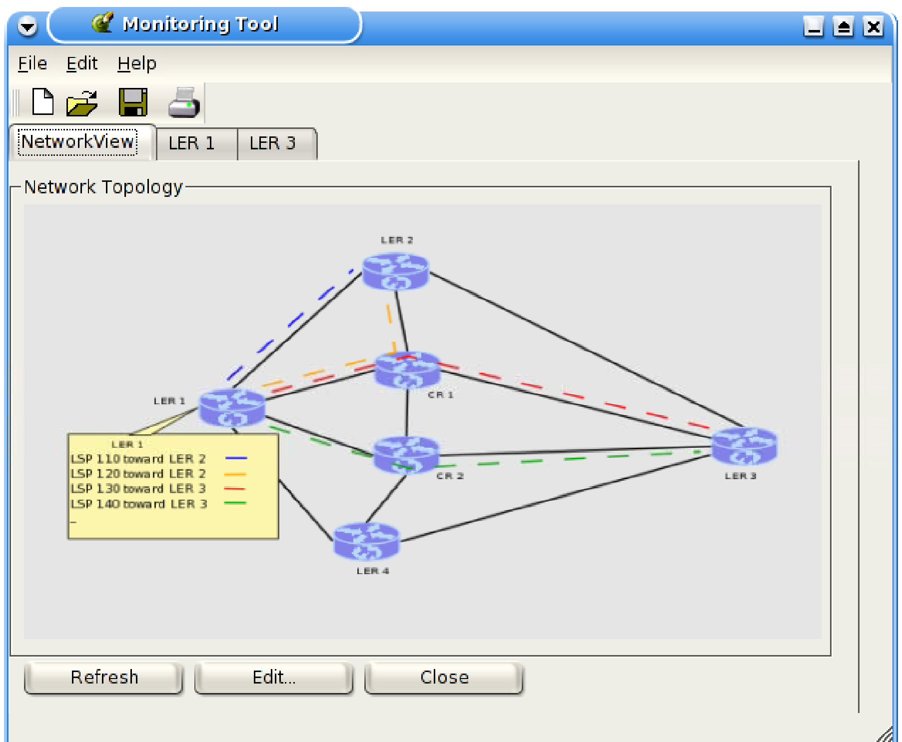
Application: monitoring/management

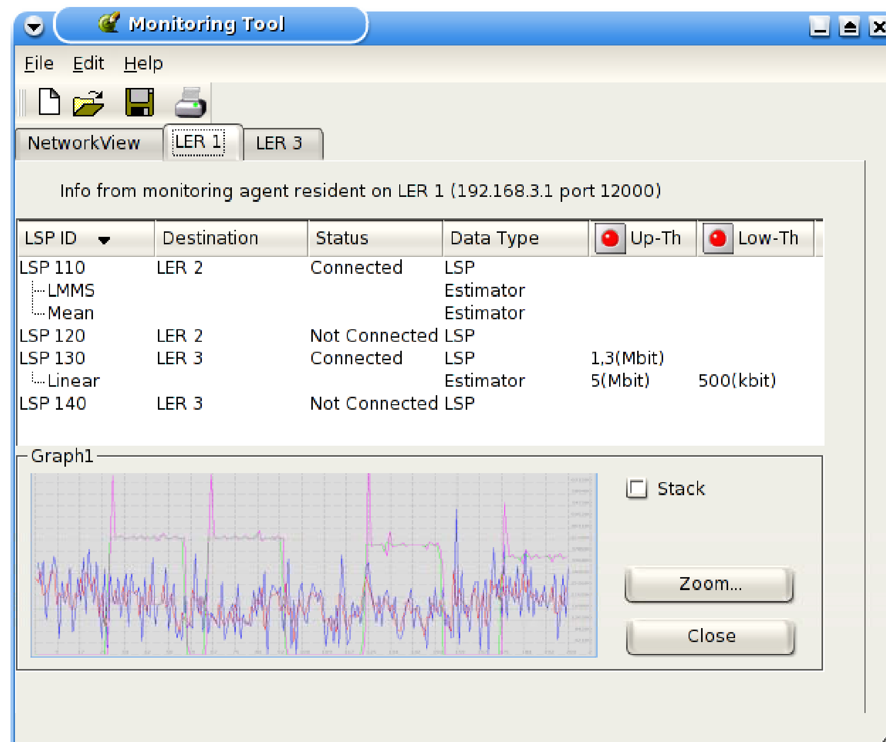
- RRDTool - produced graphs



Application: monitoring/management

- Tool developed by the Networks Group of the University of Pisa for a MPLS/DiffServ network traffic monitoring
  - Graphs for each LSP/PHB pair are available (measured and predicted values)

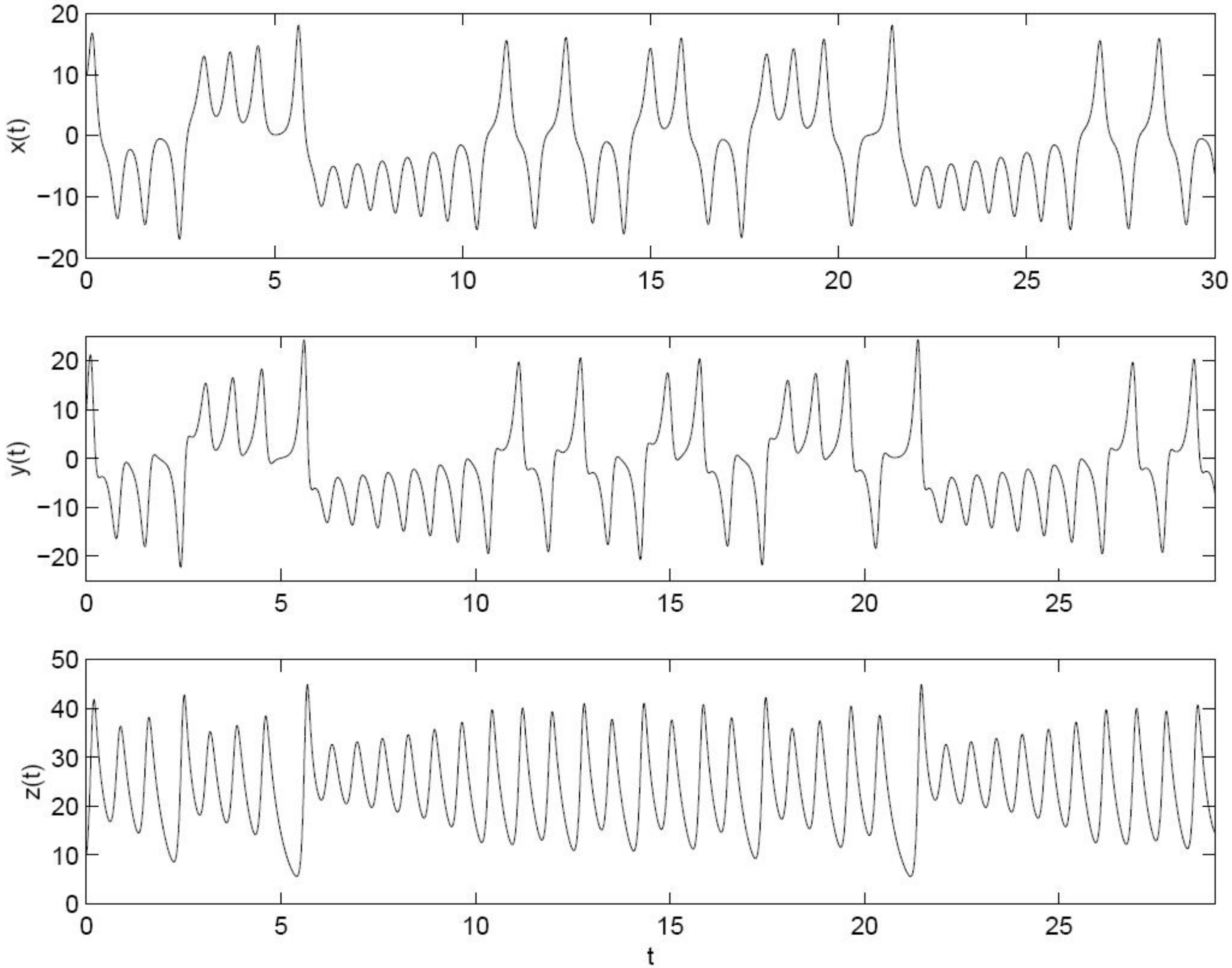




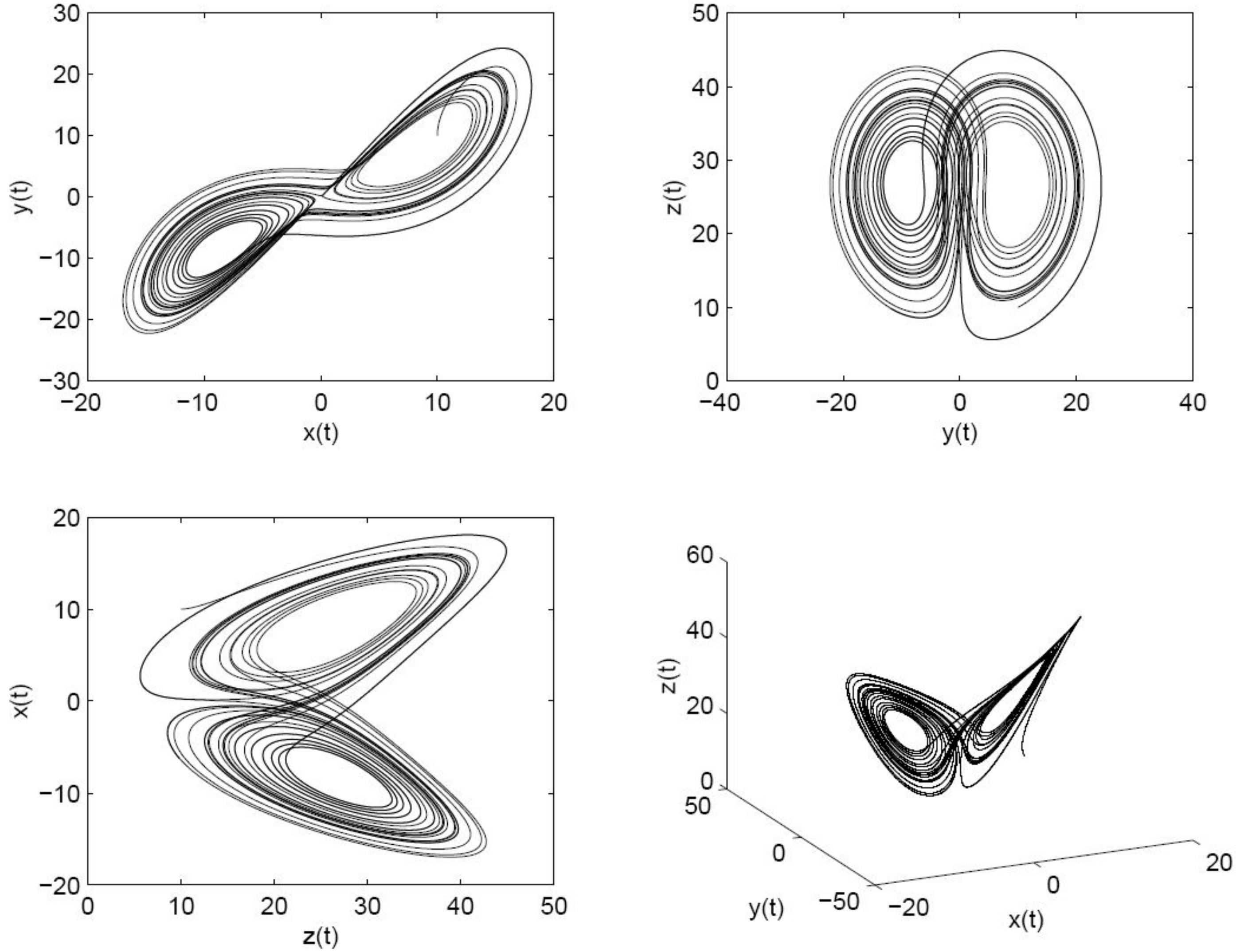
Lorenz Equations

$$\begin{cases} \dot{x} = 10(y - x) \\ \dot{y} = 28x - y - xz \\ \dot{z} = xy - \frac{8}{3}z \end{cases}$$

$$\begin{cases} x(0) = x_0 = 10 \\ y(0) = y_0 = 10 \\ z(0) = z_0 = 10 \end{cases}$$

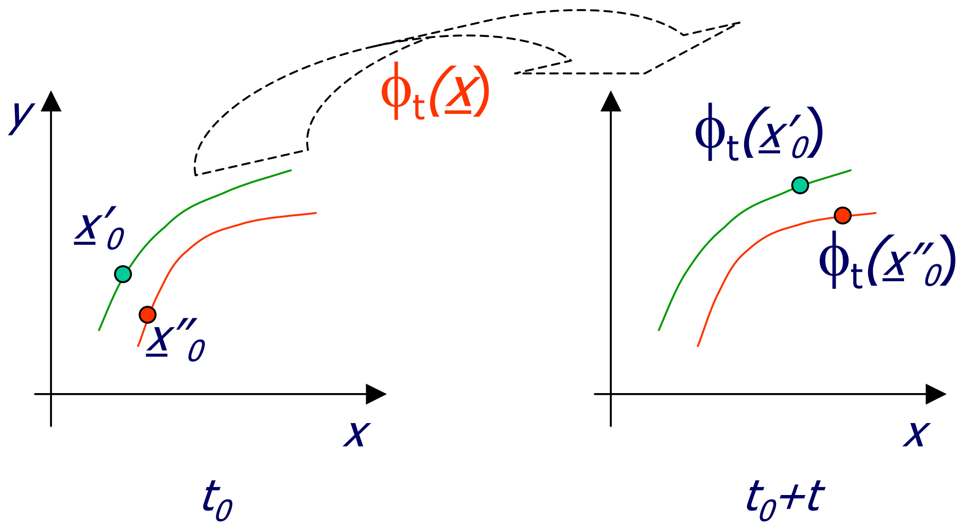


Lorentz attractor

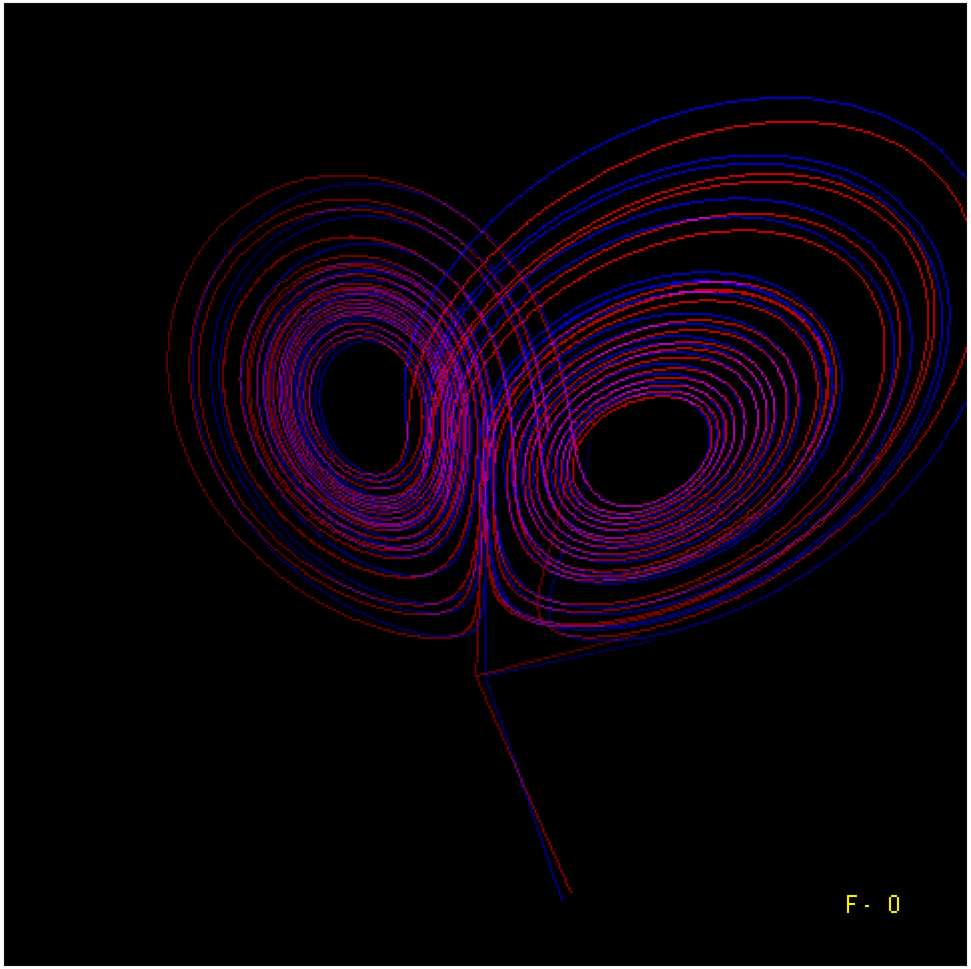


Attracting sets

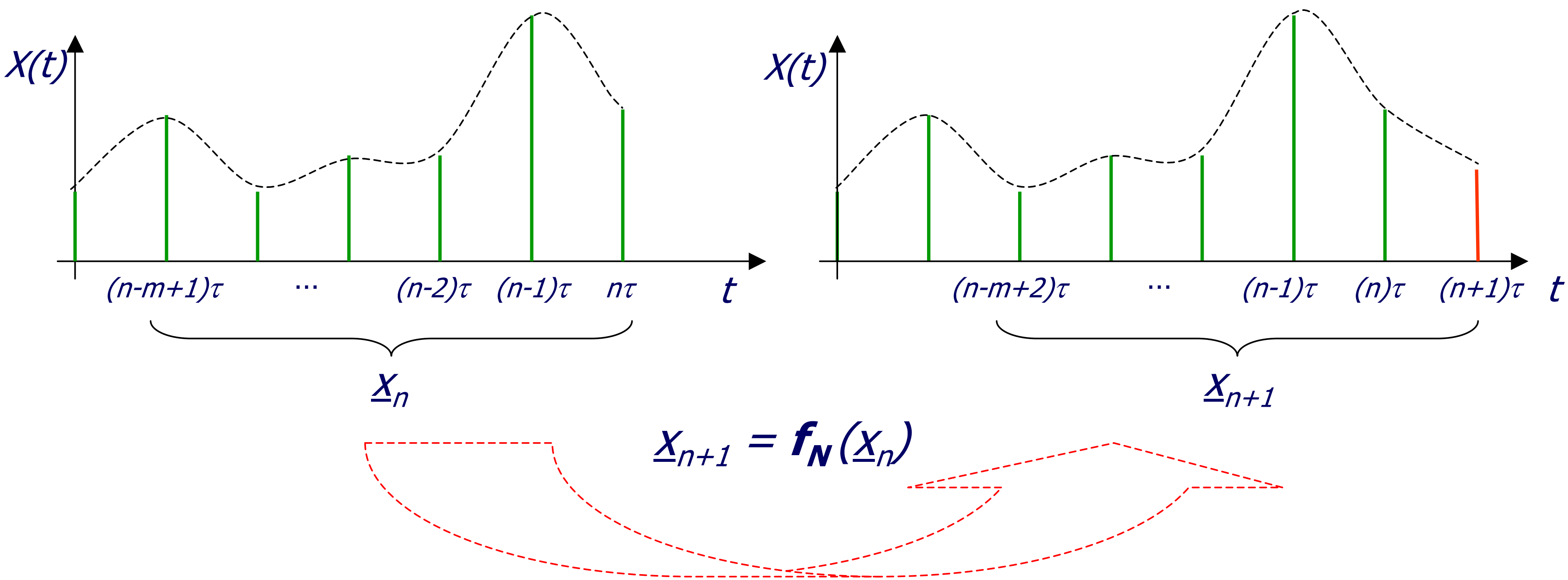
- For every $\underline{x}(t_0)=\underline{x}_0$ it is possible to find the solution $\phi_t(\underline{x}_0)$



The map $\phi_t(\underline{x}): R^n \rightarrow R^n$ *is called **flow** of the system*


F- 0

- The basic idea in adopting a chaotic predictor is that a windowed portion of the samples of a stochastic process changes due to an inner choatic behavior represented by the nonlinear map:



$$\underline{X}_{n+1} = \boldsymbol{f_N}(\underline{X}_n)$$

*The **dimension** of the state space is **m***

Steps to construct the prediction algorithm

- Consider the time series as obtained by sampling a continuous time function as

$$x_n = x(n\tau)$$

- The state vector at time n is then:

$$\underline{x}_n = \begin{pmatrix} x(n\tau) \\ x((n-1)\tau) \\ \vdots \\ x((n-m+1)\tau) \end{pmatrix}$$

Dimension m

- The parameter τ is also called **delay time**

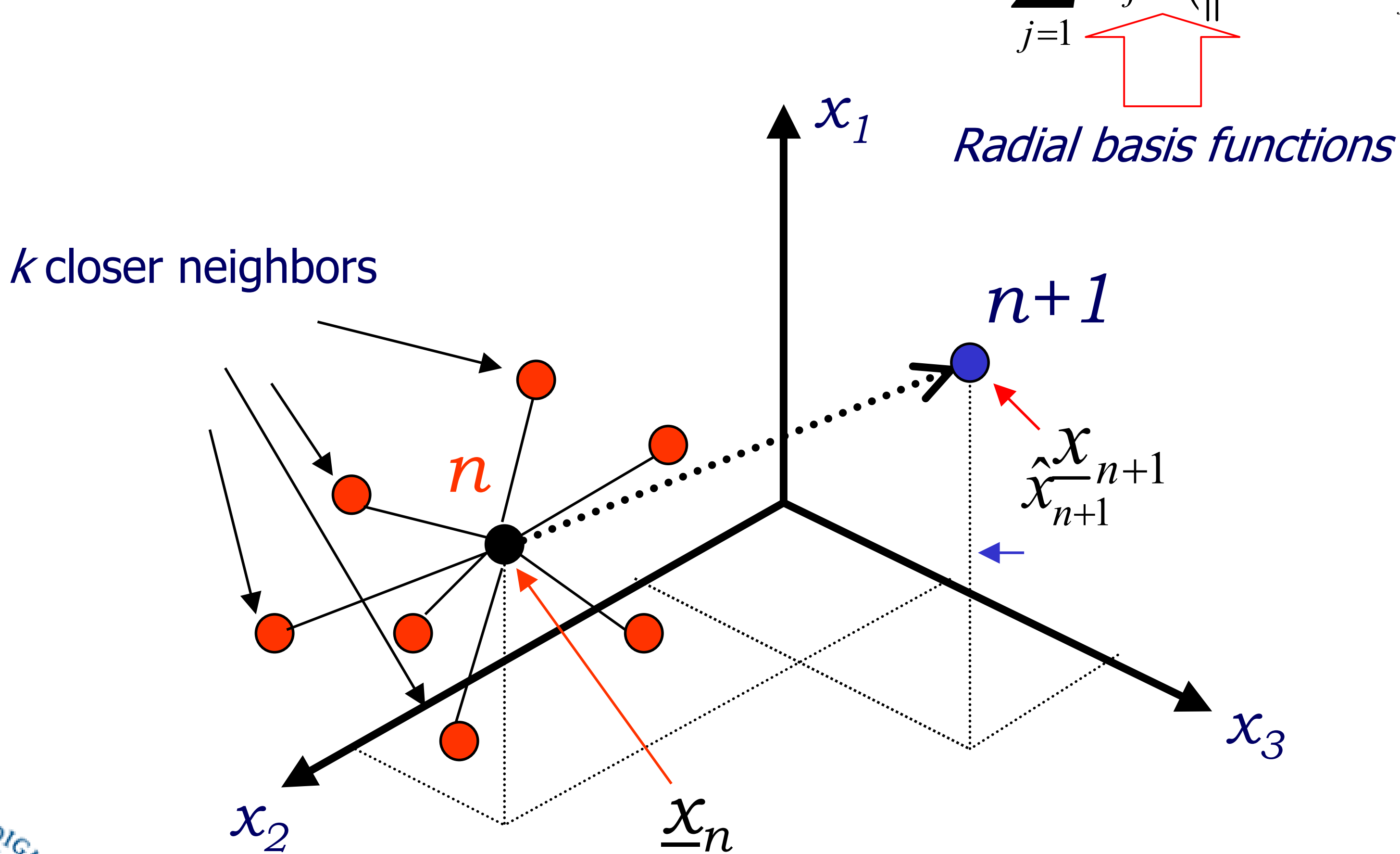
- Select a proper interpolation function f_N such that $\underline{x}_{n+1} = f_N(\underline{x}_n)$

- Notice that, with this choice of state space, the future state $\underline{x}_{n+1}$ has only the first component unknown!

RBFP - Radial Basis Functions Predictor

- Prediction function:

$$\hat{x}_{n+1} = \sum_{j=1}^{k} \lambda_j \phi\left(\left\| \underline{x}_n - \underline{x}_{n_j} \right\|\right) + \mu$$



- The form of the *radial basic functions* $\phi: R^+ \rightarrow R$ is the following:

$$\phi(r) = \left(r^2 + c^2\right)^{-\beta}, \qquad \beta > -1 \quad e \quad \beta \neq -1$$

c and β are costants

- Coefficients λ_i are determined through the knowledge of the past state evolution by imposing:

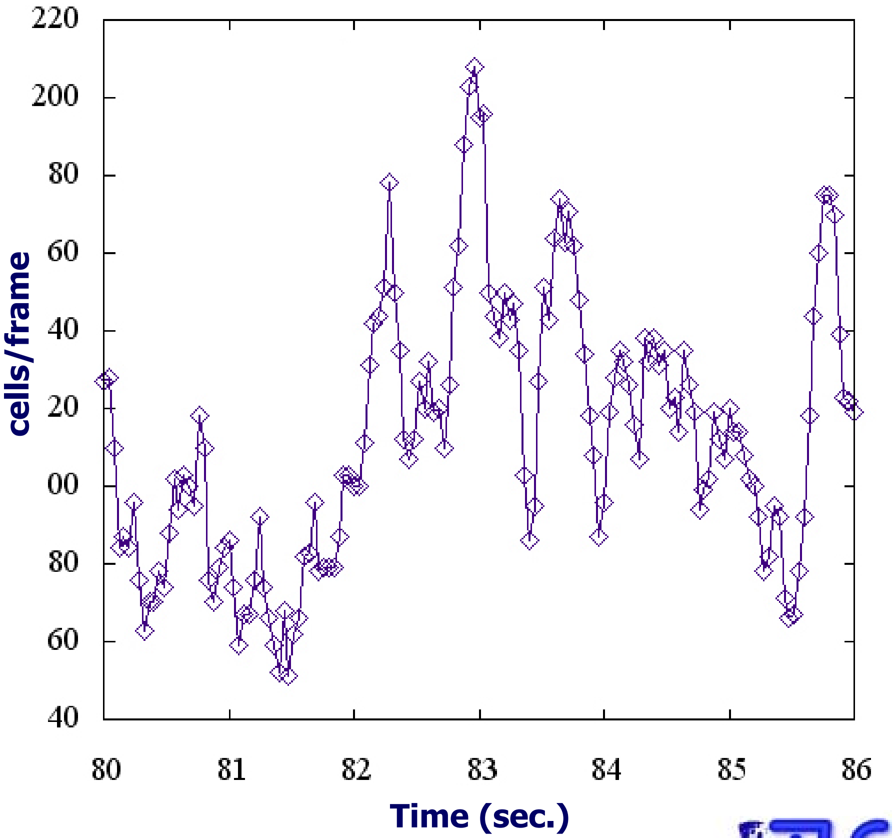
$$\hat{x}_{n_i+1} = \sum_{j=1}^{k} \lambda_j \underbrace{\phi\left(\left\| \underline{x}_{n_i} - \underline{x}_{n_j} \right\|\right)}_{\phi_{ij}} + \mu \qquad \text{and} \qquad \sum_{j=1}^{k} \lambda_j = 0$$

which leads to the linear system of *(k+1)* equations (to be solved by numerical techniques):

$$\begin{pmatrix} \phi_{11} & \cdots & \phi_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \phi_{k1} & \cdots & \phi_{kk} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ \mu \end{pmatrix} = \begin{pmatrix} x_{n_1+1} \\ \vdots \\ x_{n_k+1} \\ 0 \end{pmatrix}$$

Application to Network Traffic

- Videoconference traffic (prototypal codec)
- Time series: number of ATM cells transmitted during a frame period (1/25 sec.)
  - Average bit rate 1.668 Mbps
  - Peak-to-mean ratio: 4.8
  - Frame per second: 25
  - Number of frames: 48496

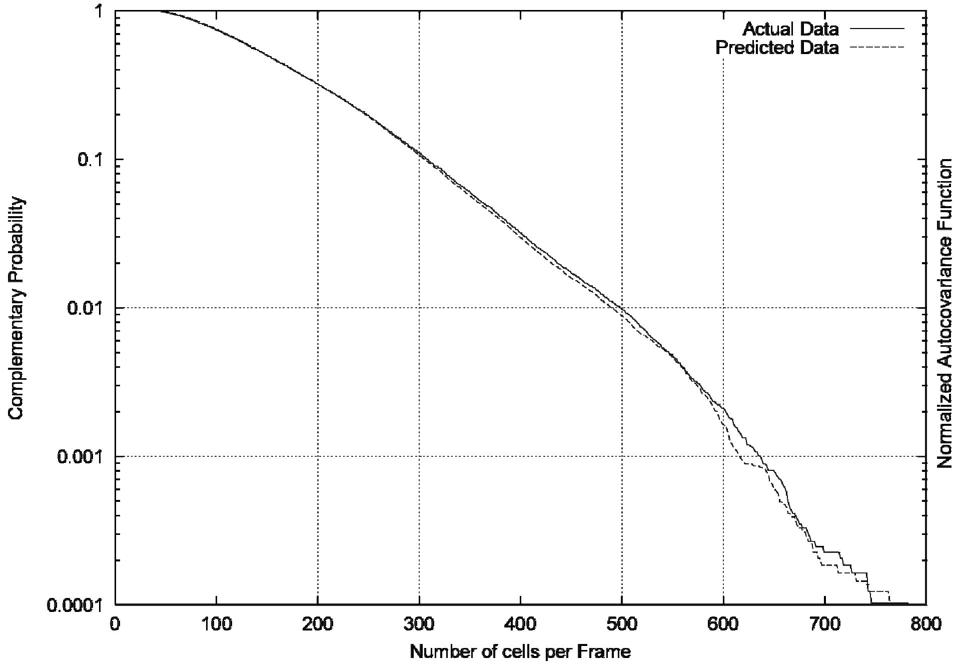


Automated mechanism of RBFP produces:

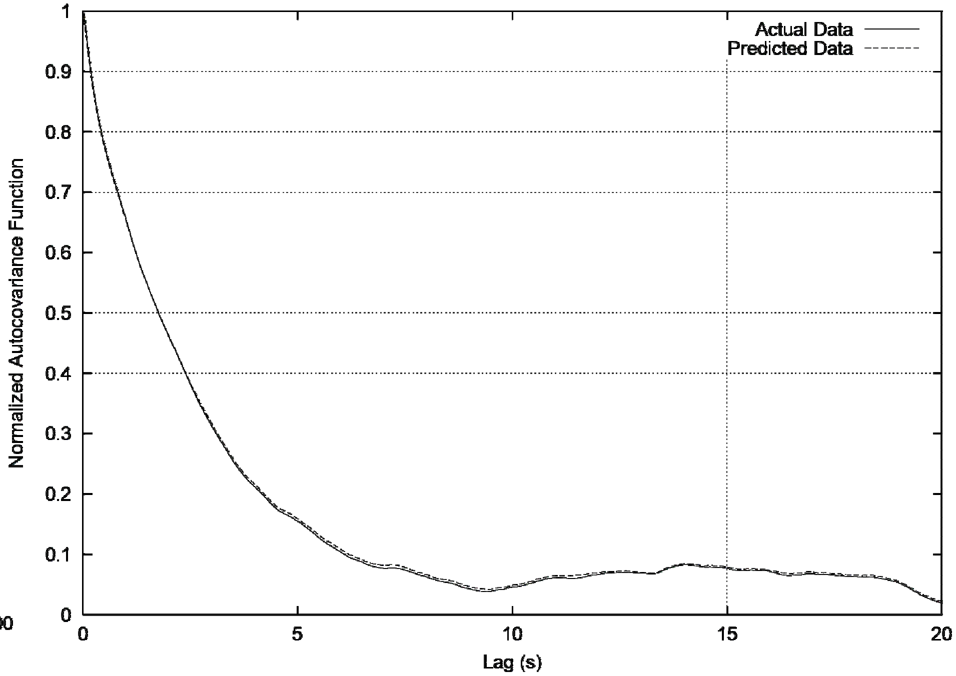
- Embedding dimension *m**=3
- Delay-time = 40 msec
- Number of neighbors *k = m*+1 = 4*
- Value of *c=0.28*
- Number of samples *N=3000*

RBFP Predictor performance: Video Sequence

Sample Autocovariance

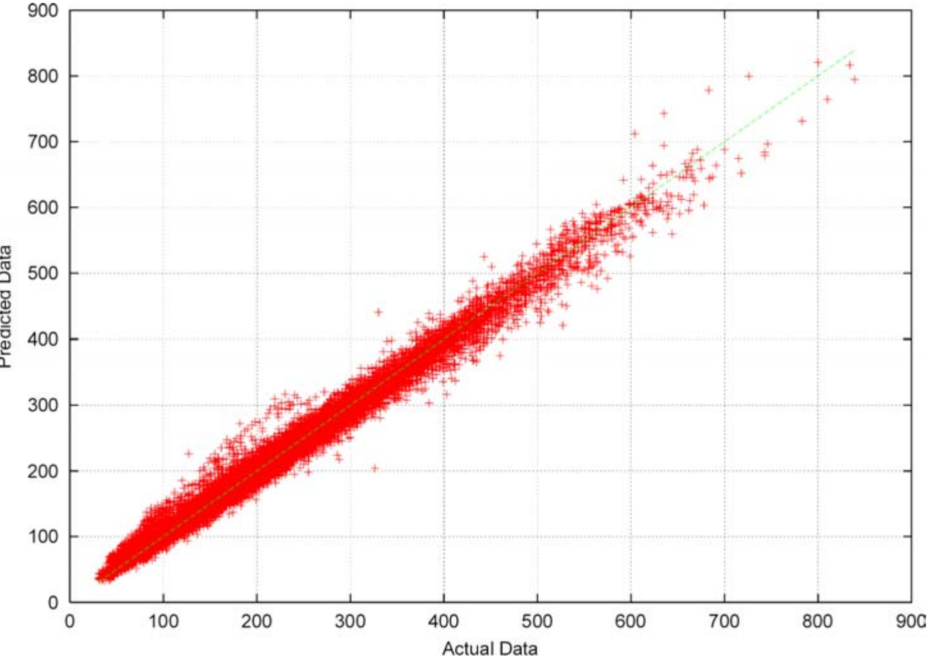


Complementary probability

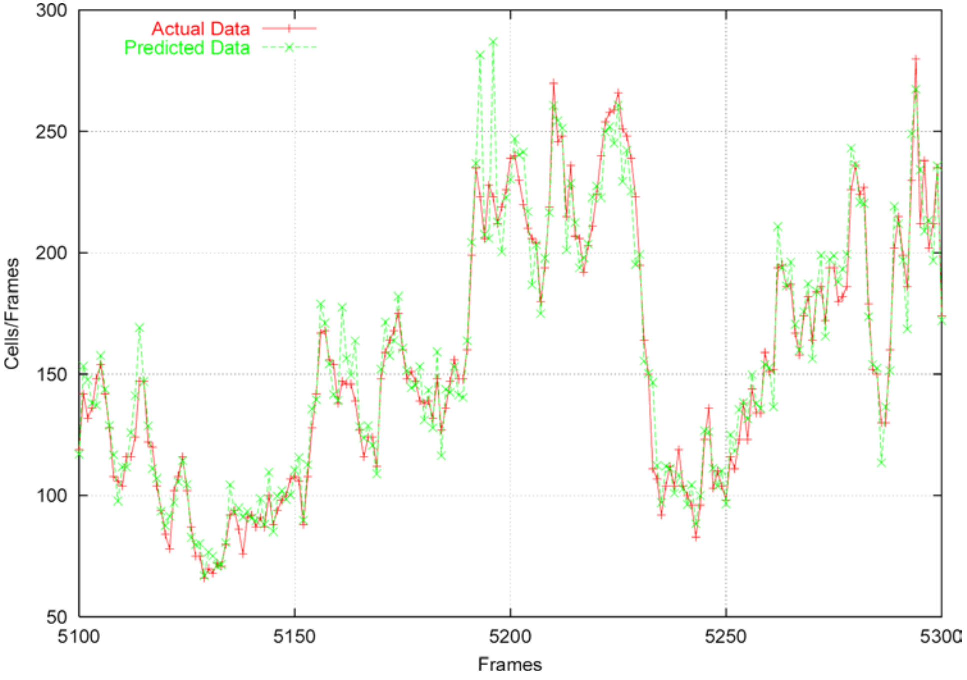


RBFP Predictor performance: Video Sequence

Actual/Predicted (APP) - Plot

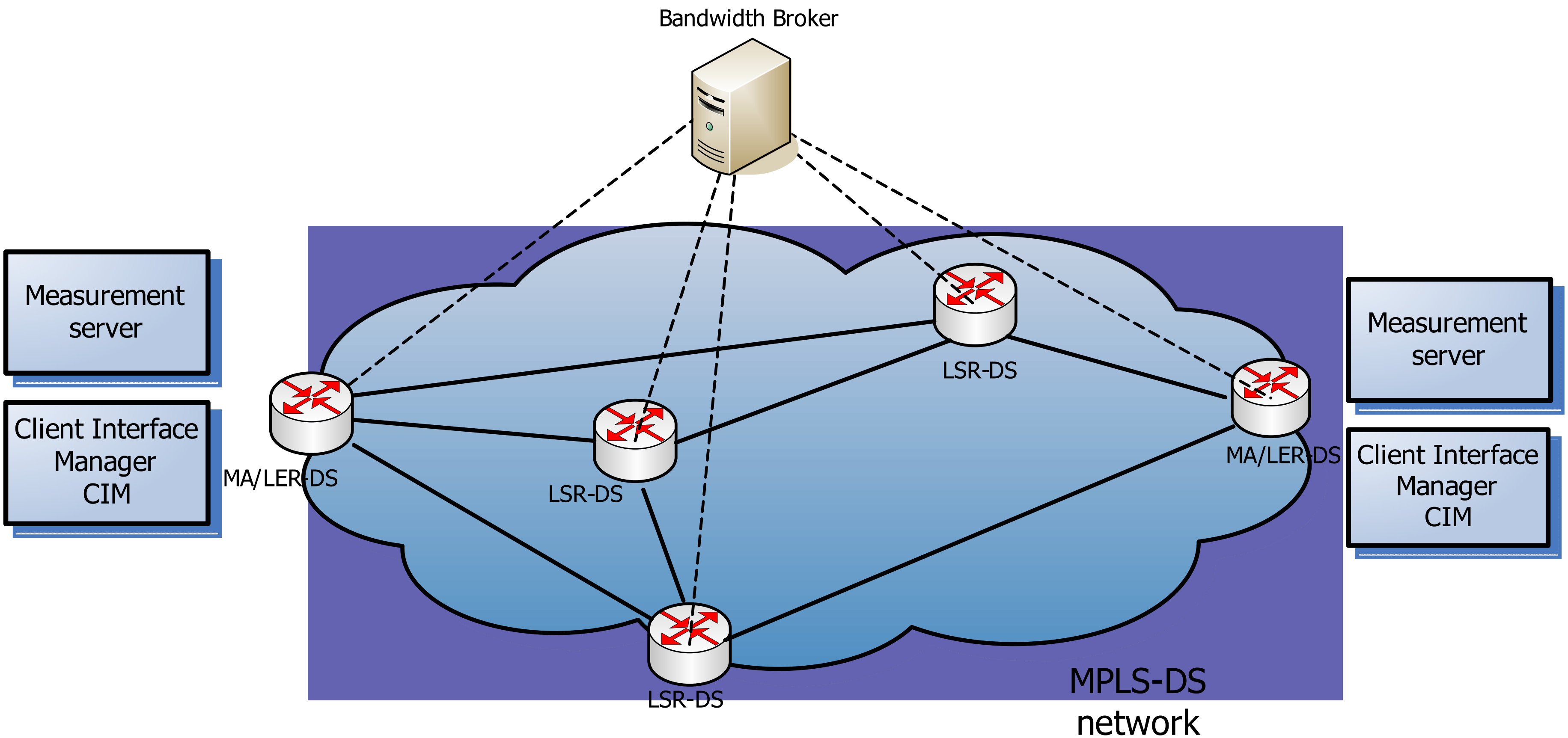


Actual/Predicted patterns

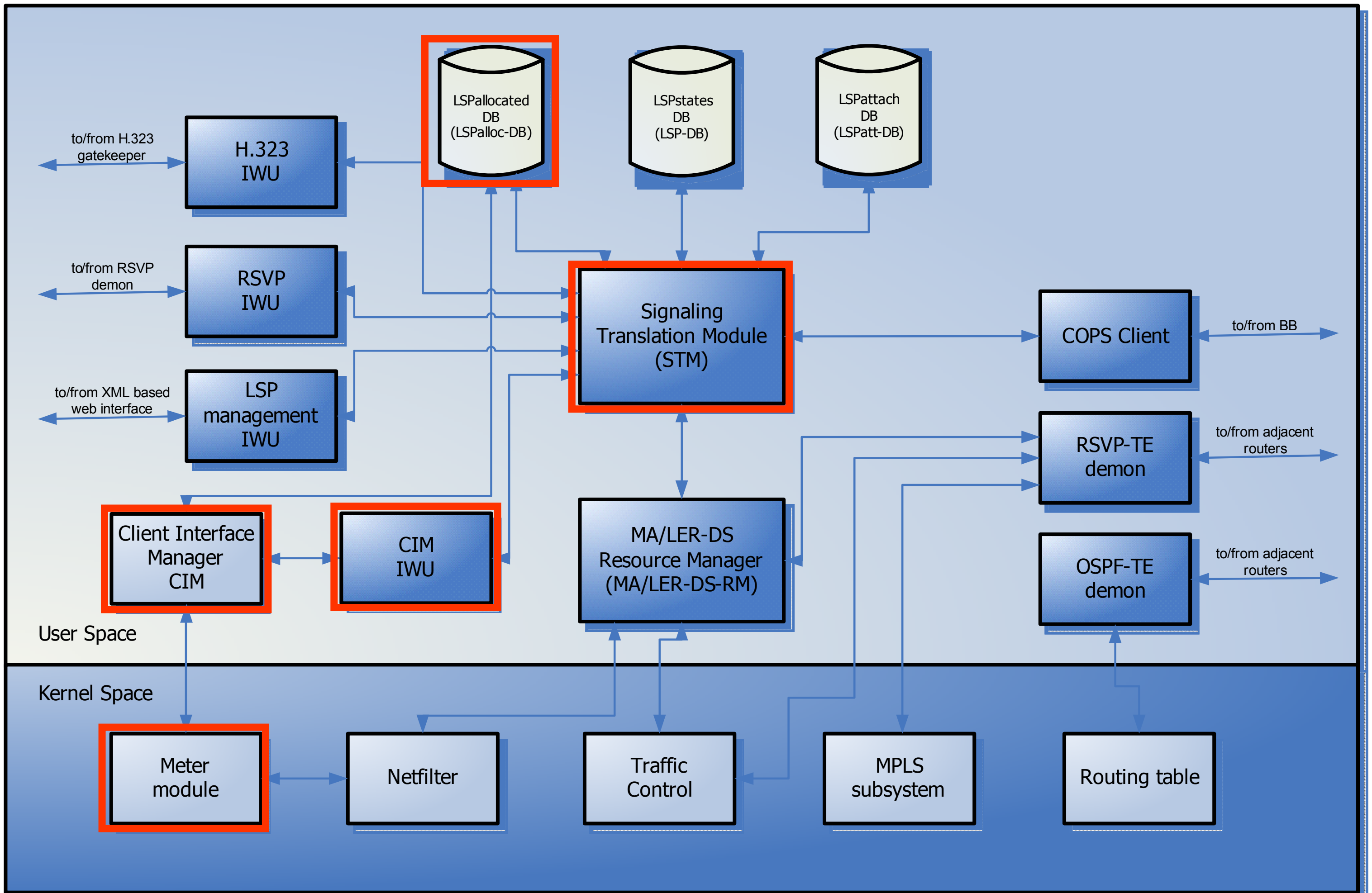


A simple case of linear prediction

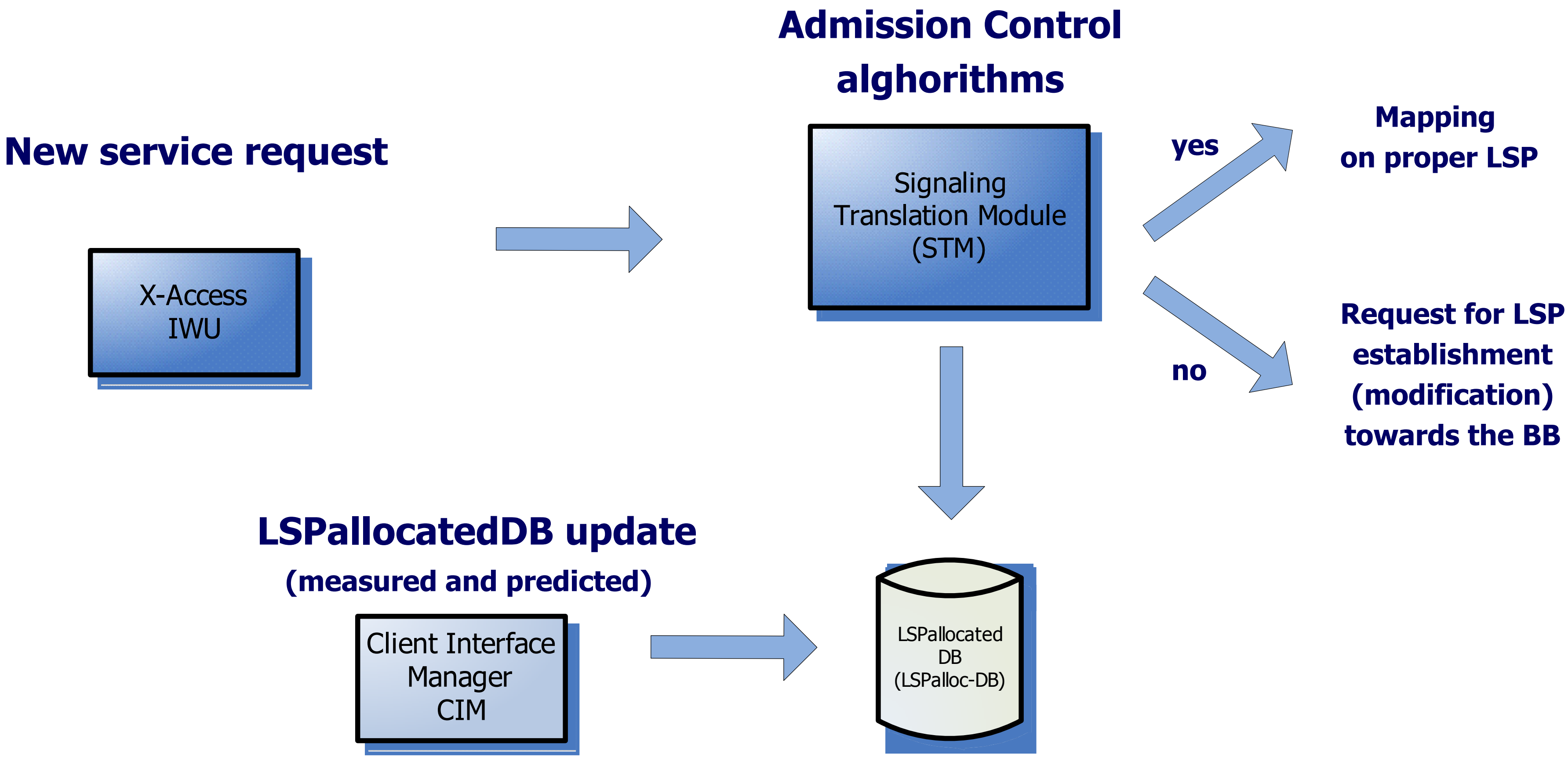
Application: Admission Control/Resource Allocation



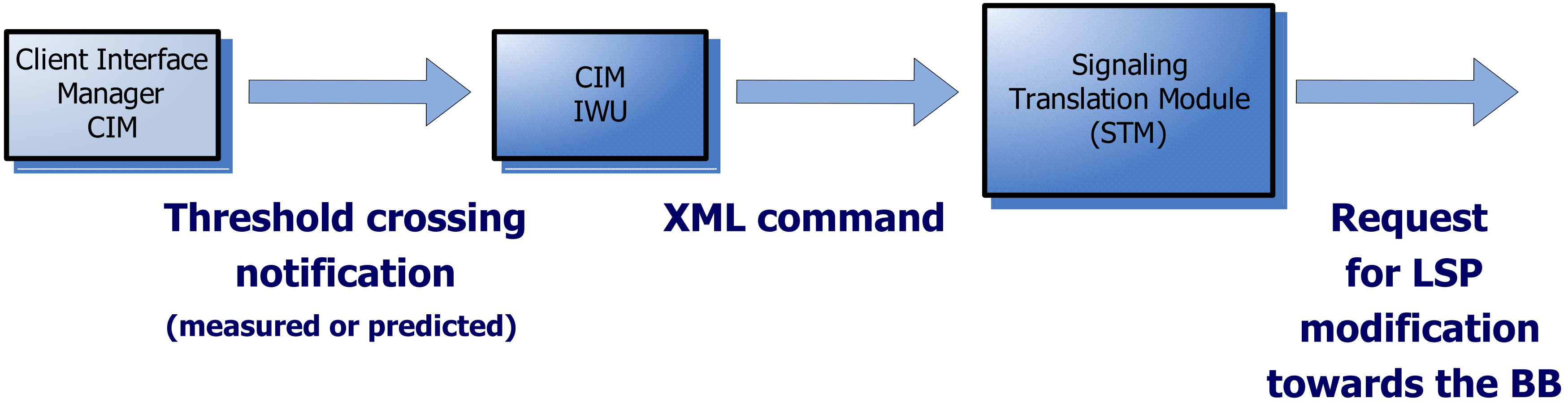
Application: Admission Control/Resource Allocation



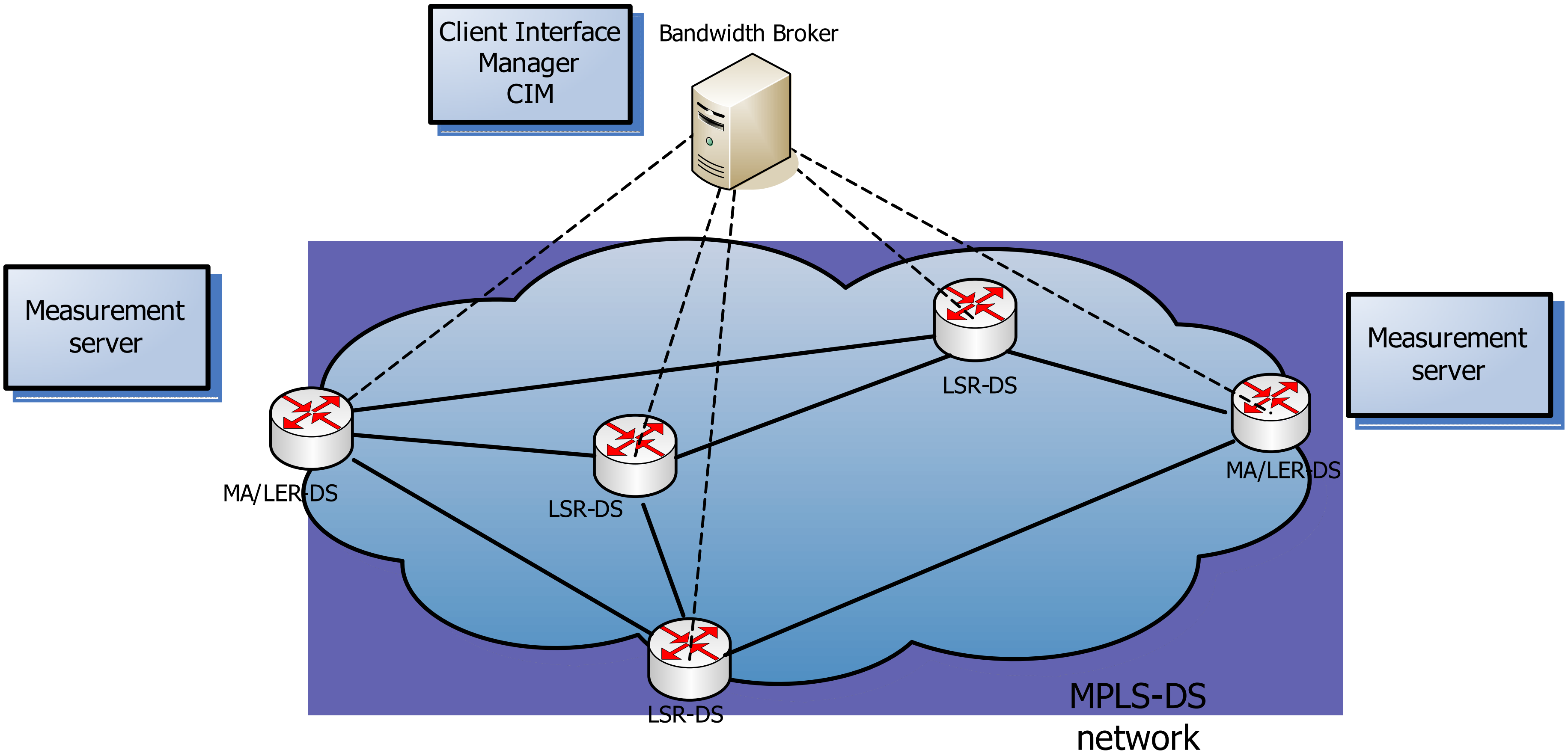
Application: Admission Control/Resource Allocation



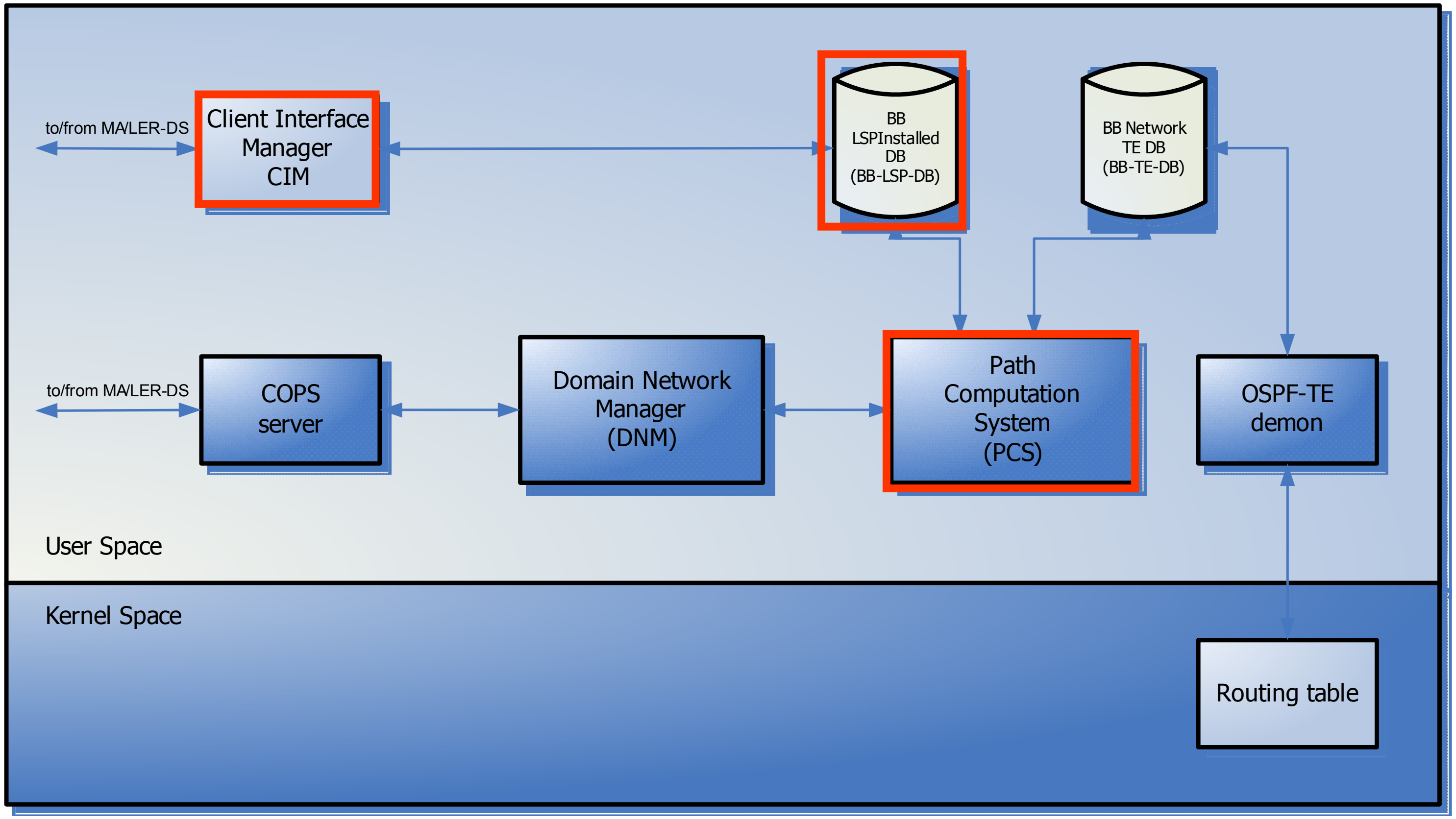
Application: Admission Control/Resource Allocation



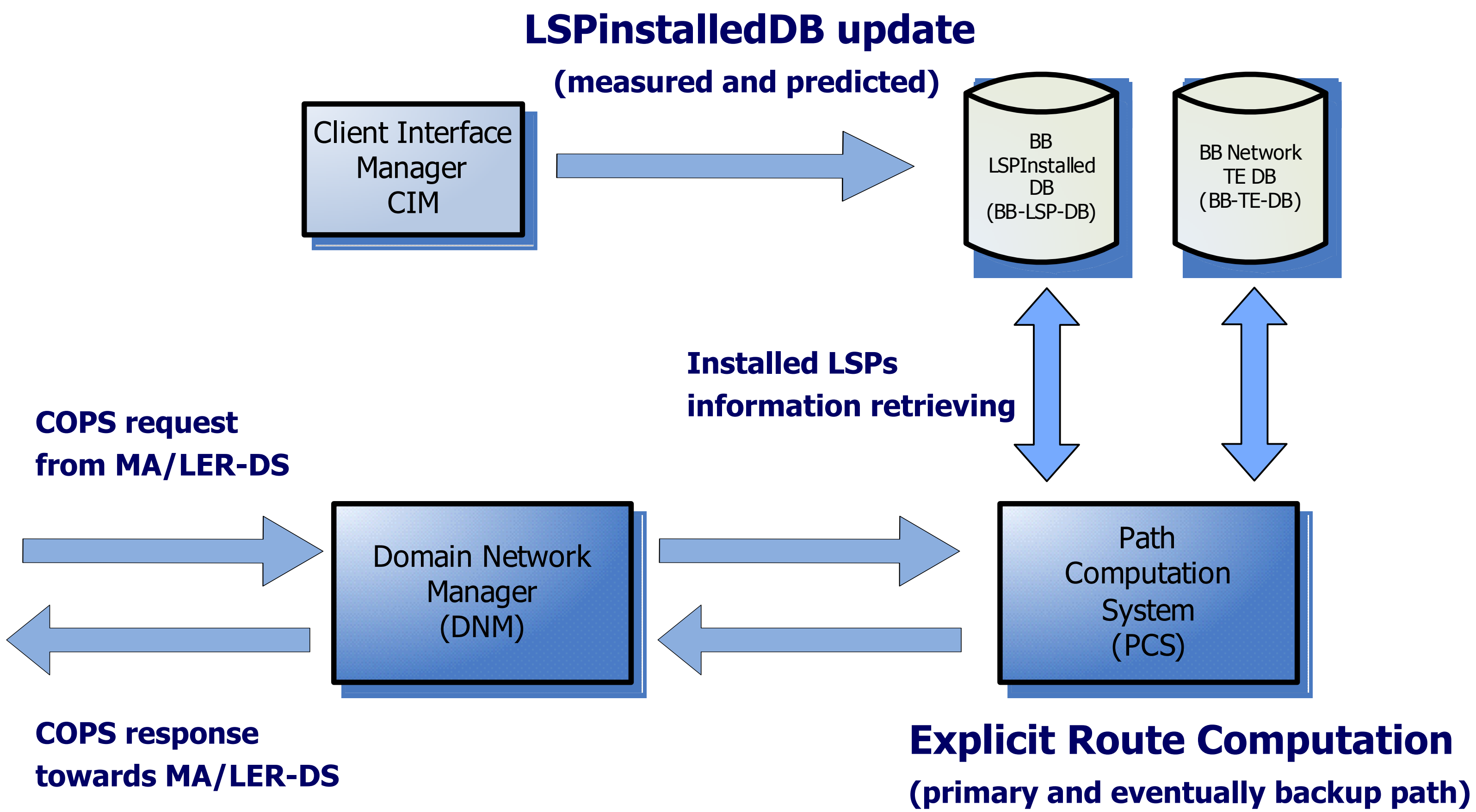
Application: Traffic Engineering



Application: Traffic Engineering

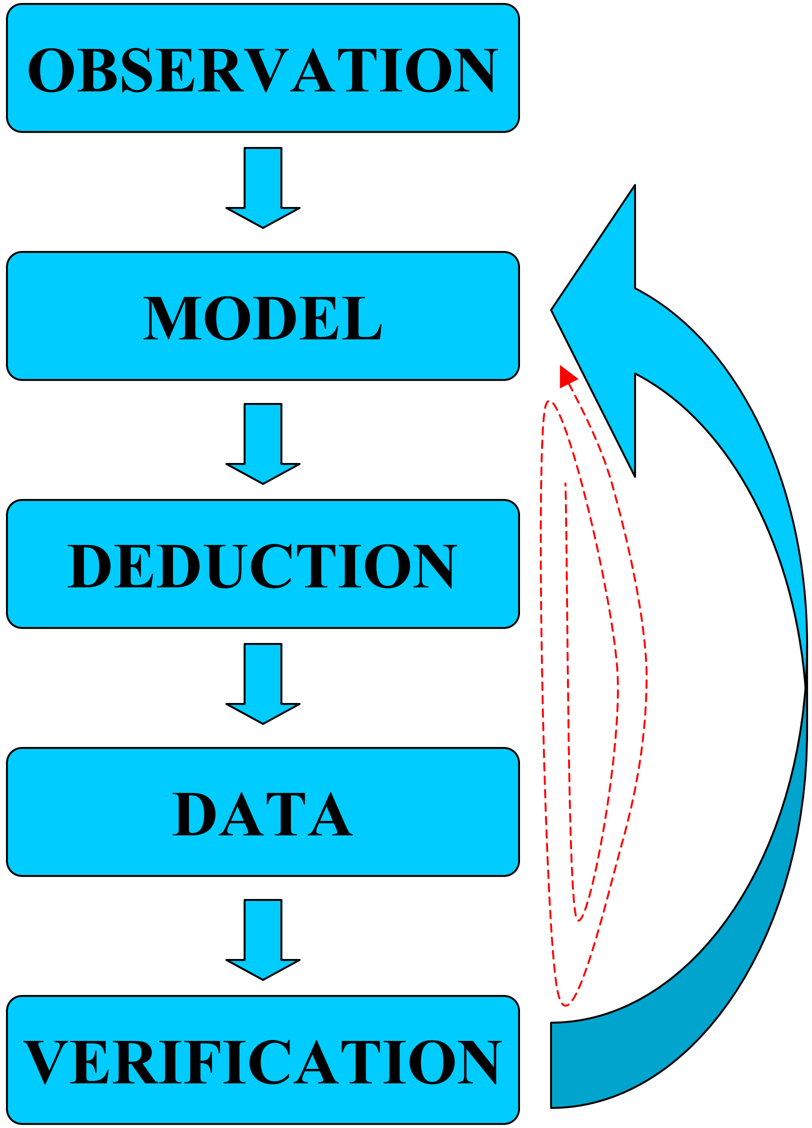


Application: Traffic Engineering



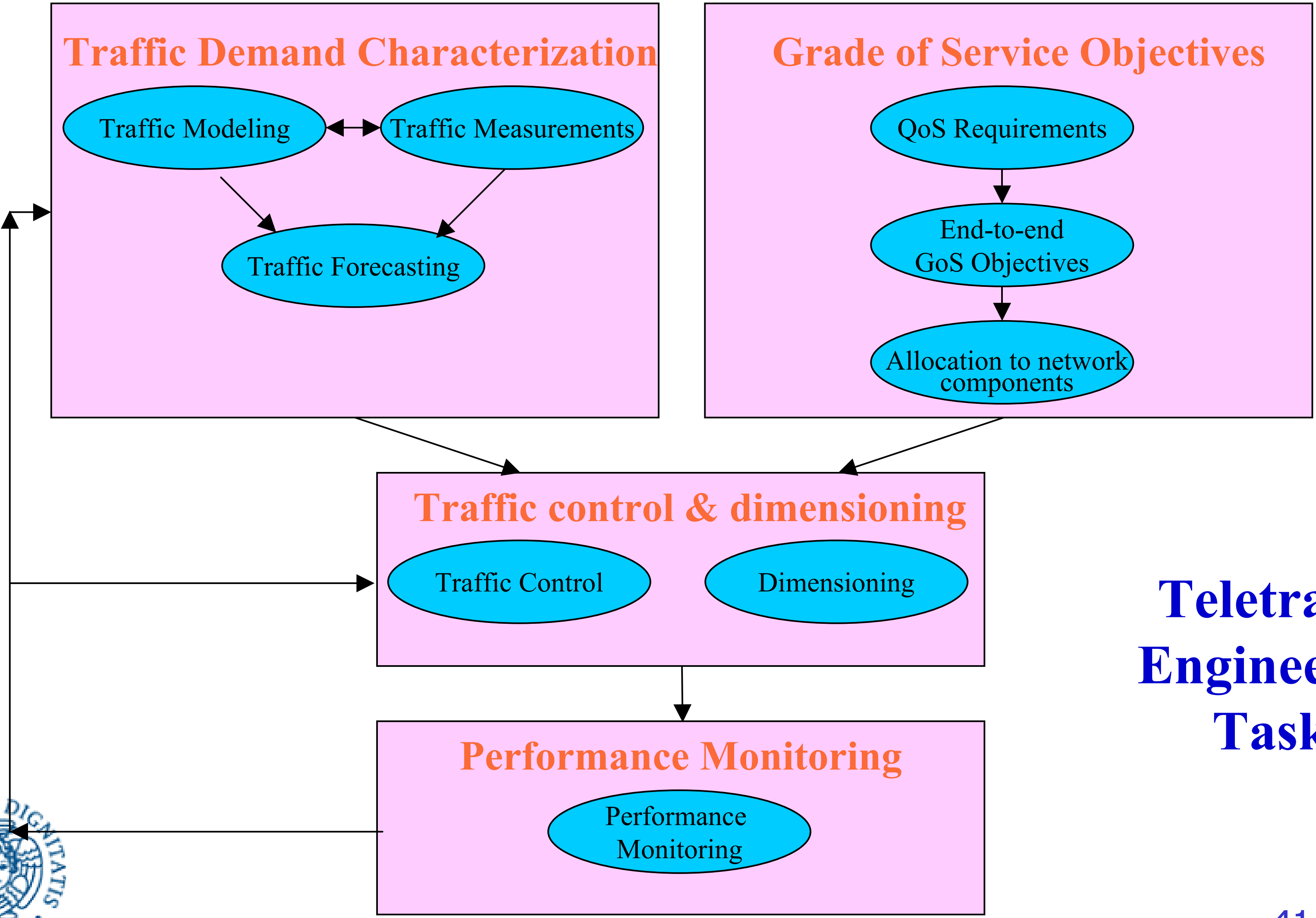
From ITU-D SG2 Question 16/2
“Teletraffic Engineering Handbook” – Chap.1

This scientific way of working is called the research spiral.



Teletraffic theory is an inductive discipline. From the observation of real systemswe establish theoretical models from which we derive parameters which can be compared with the corresponding observation of real systems If there is agreement he model has been validated. If not, then we have to elaborate the model further.

From ITU-D SG2 Question 16/2
"Teletraffic Engineering Handbook" – Chap.1

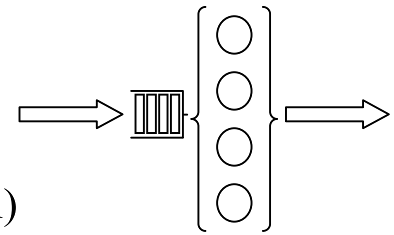


Teletraffic Engineering Tasks

From ITU-D SG2 Question 16/2
“Teletraffic Engineering Handbook” – Chap.1

- **Mathematical Model**

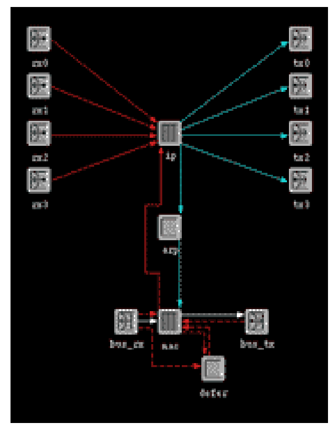
(inexpensive; no time consuming, general)



Theory of discrete state Stochastic Processes

- **Simulation Model**

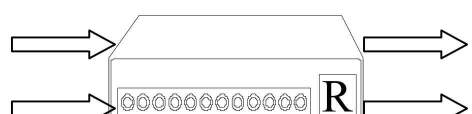
(expensive; simulation model is not general: every individual case must be simulated; <u>confidence intervals</u>; etc)

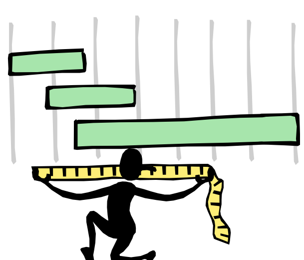


Discrete-event simulation: Trace Driven (inputs from collected data) or Artificial Inputs from statistical distributions

- **Physical Model**

(even more expensive, specific, time and resource consuming)





Prototypes: field-trials, real and artificial generators, analyzers, etc