

# Intranet capacity planning: 3 case studies

15/01/08

# indice

- capacity planning di una intranet
  - classi di utenti, workload, parametrizzazione
- intranet con http proxy
  - influenza hit rate http proxy sulle prestazioni
- intranet 3-tier con due tipi di applicazioni
  - 3 livelli, influenza del mix di utenti sulle prestazioni

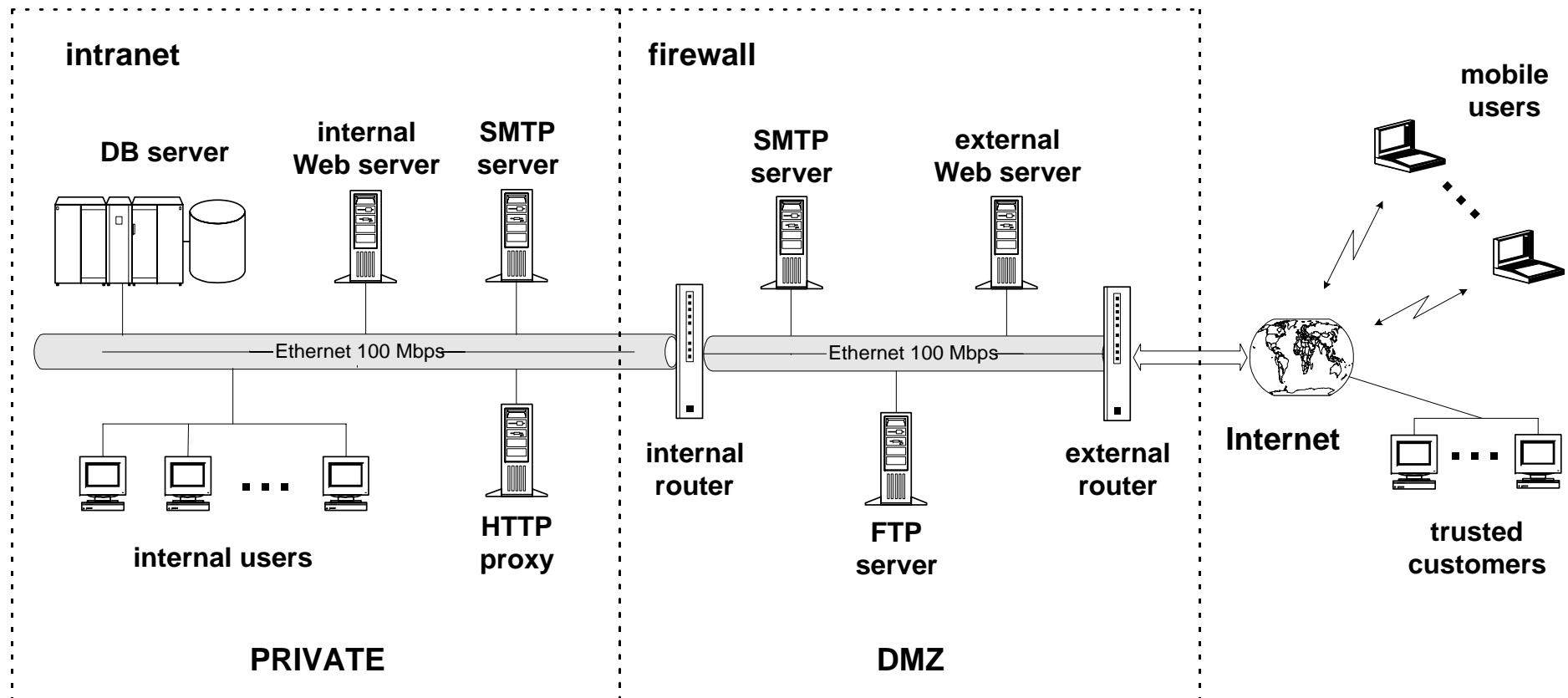
# capacity planning di una intranet

- struttura della intranet
- il workload (classi di utenti)
- parametrizzazione del modello
- prestazioni al variare del mix di utenti
- bottleneck

## caratteristiche generali

- intranet aziendale costituita da:
  - vari server (http, Web, DB, FTP, e-mail, proxy)
  - una Zona Demilitarizzata (2 firewall)
  - due lan (Ethernet 100 Mbps)
- due tipi di utenti
  - interni
  - esterni

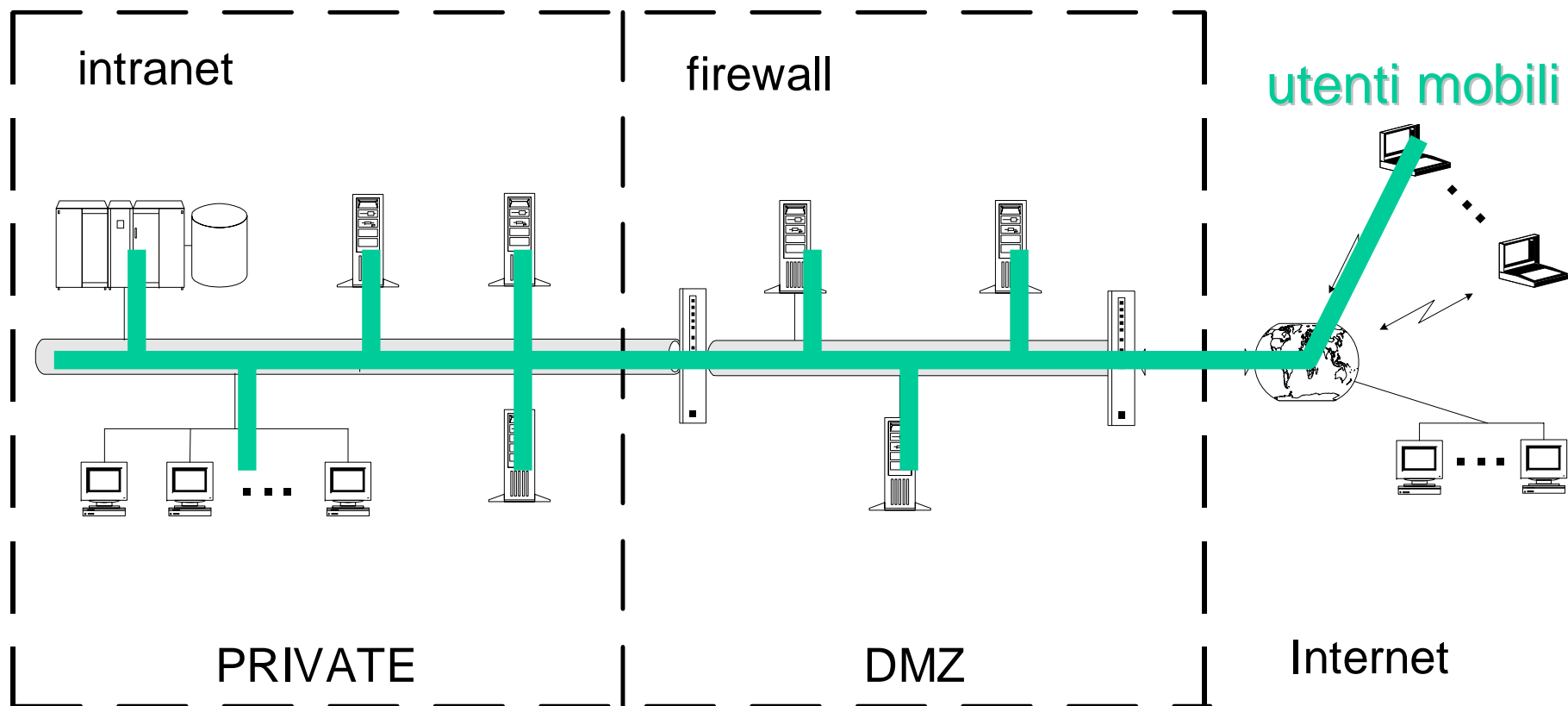
# elementi della intranet



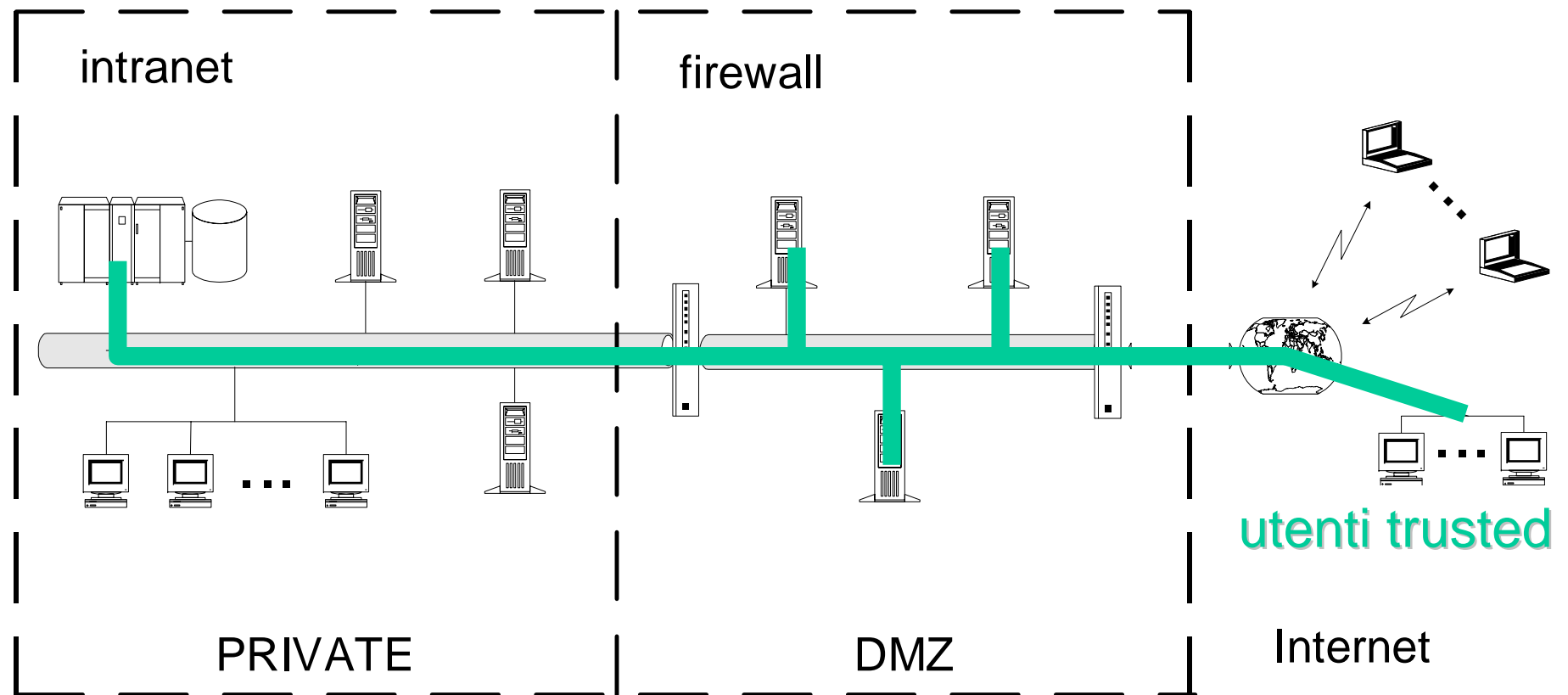
# workload

- utenti **esterni**
  - mobili (venditori, personale dell'azienda)
  - trusted (clienti registrati, http dinamiche)
- utenti **interni**
  - intranet
    - con accessi limitati ai server interni all'azienda (hit sul proxy)
  - Internet
    - con accessi ad internet (miss sul proxy)

# utenti esterni mobili: flusso degli accessi

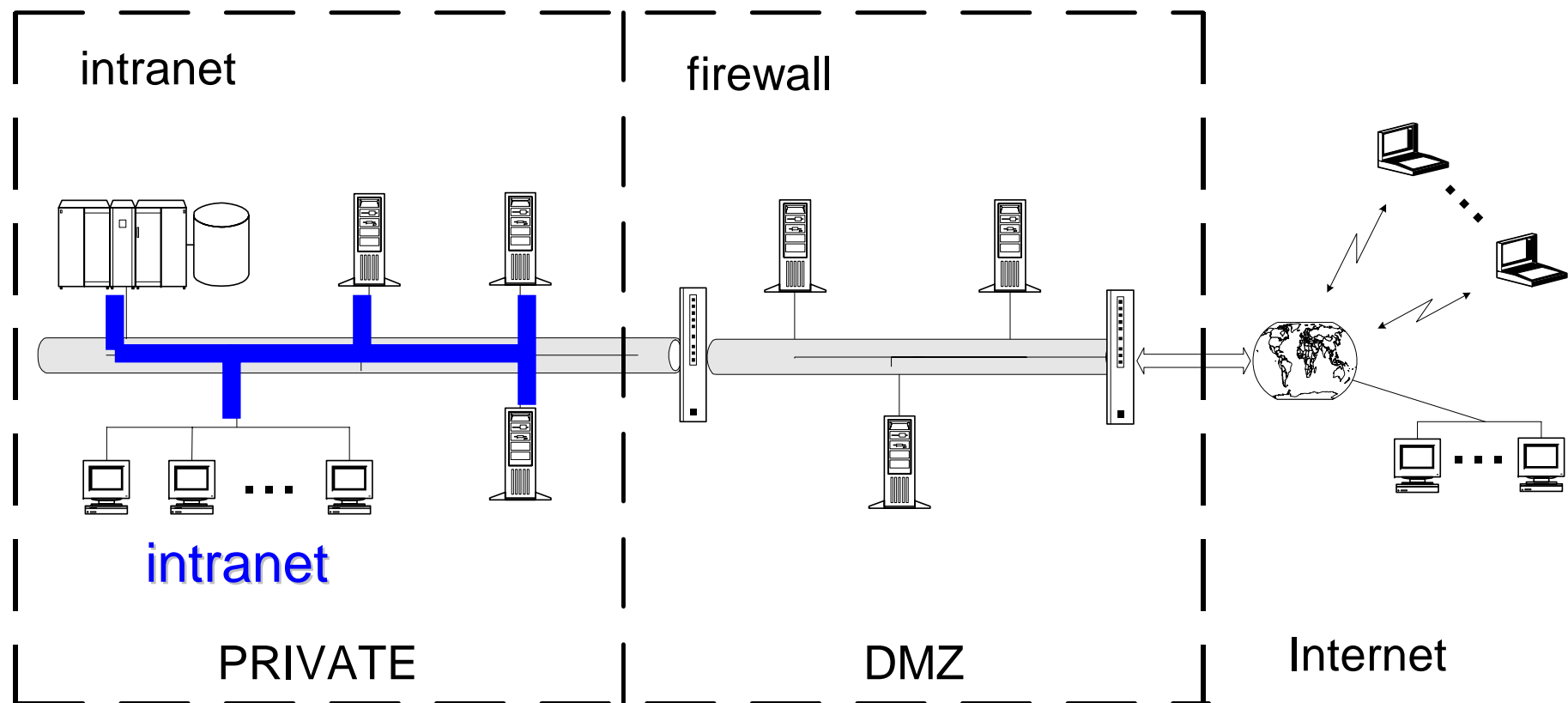


## utenti esterni registrati (trusted)

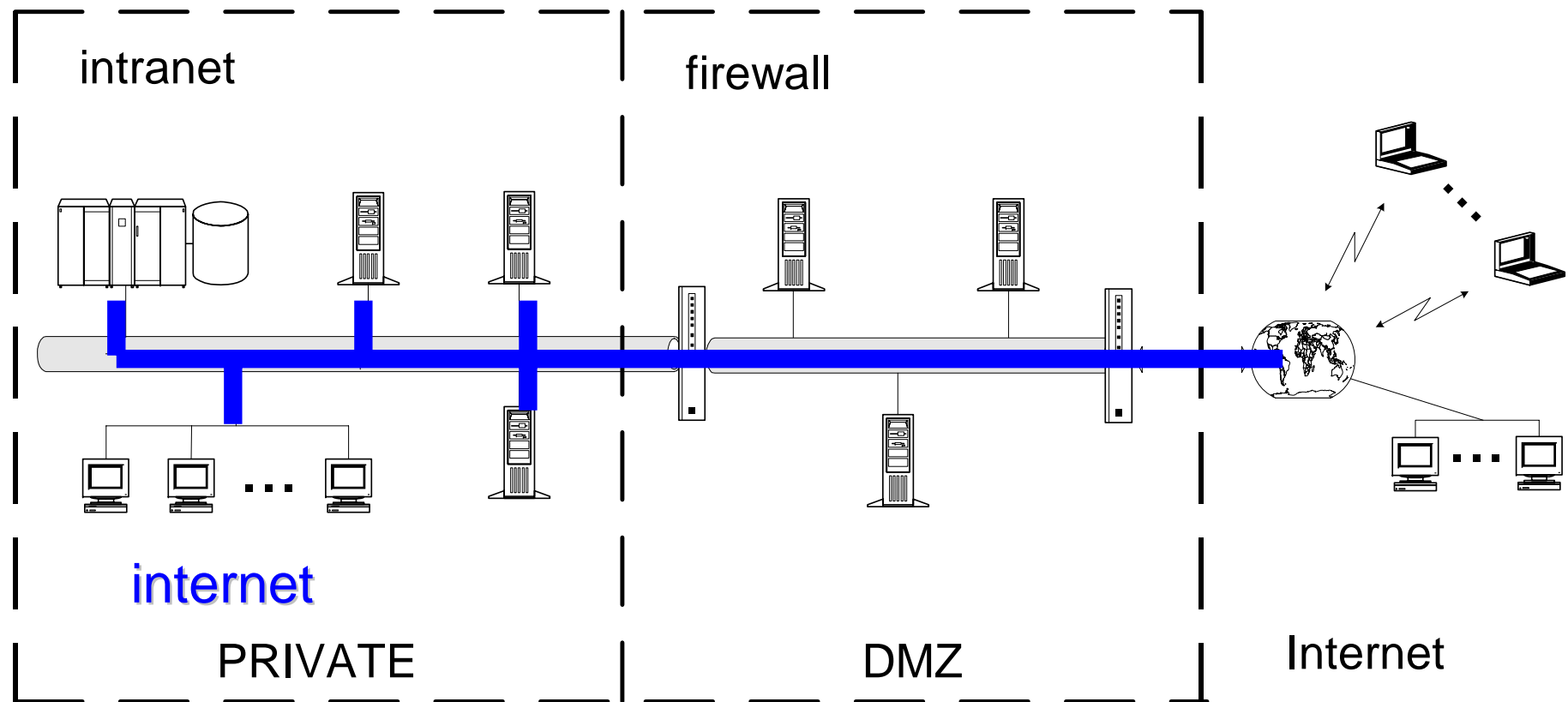




## utenti interni con accessi ai server interni



## utenti interni con accessi a Internet



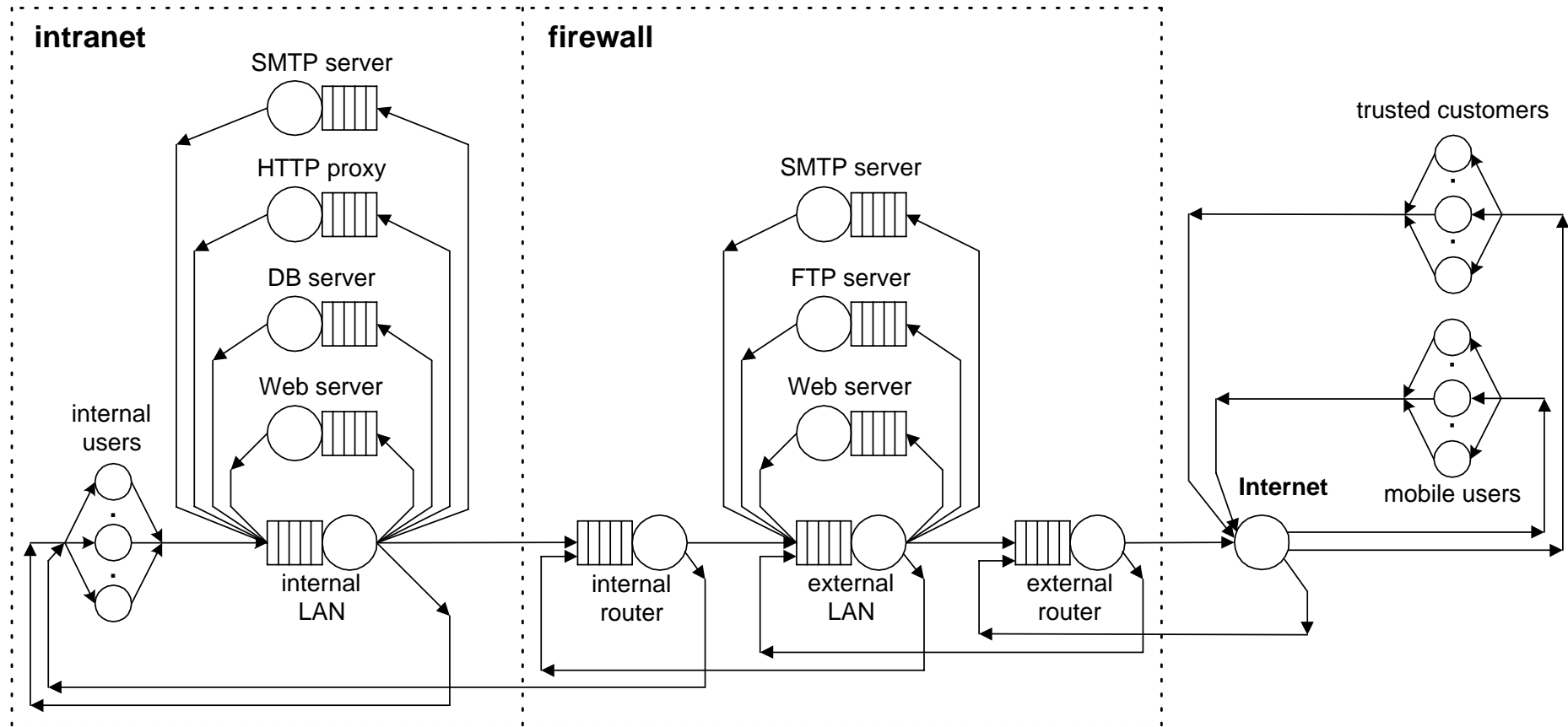
## distribuzione del carico sui componenti

	<b>intranet</b>	<b>proxy</b>	<b>Internet</b>	<b>mobile</b>	<b>trusted</b>
<b>int. LAN</b>	X	X	X	X	X
<b>int. Web serv.</b>	X	X	X	X	
<b>int. e-mail serv.</b>	X	X	X	X	X
<b>http proxy</b>		X	X		
<b>DB serv.</b>	X	X	X	X	X
<b>int. router</b>			X	X	X
<b>ext. LAN</b>			X	X	X
<b>ext. Web serv.</b>			X	X	X
<b>ext. FTP serv.</b>			X	X	X
<b>ext. e-mail serv.</b>			X	X	X
<b>ext. router</b>			X	X	X
<b>Internet</b>			X	X	X

## richieste di servizio

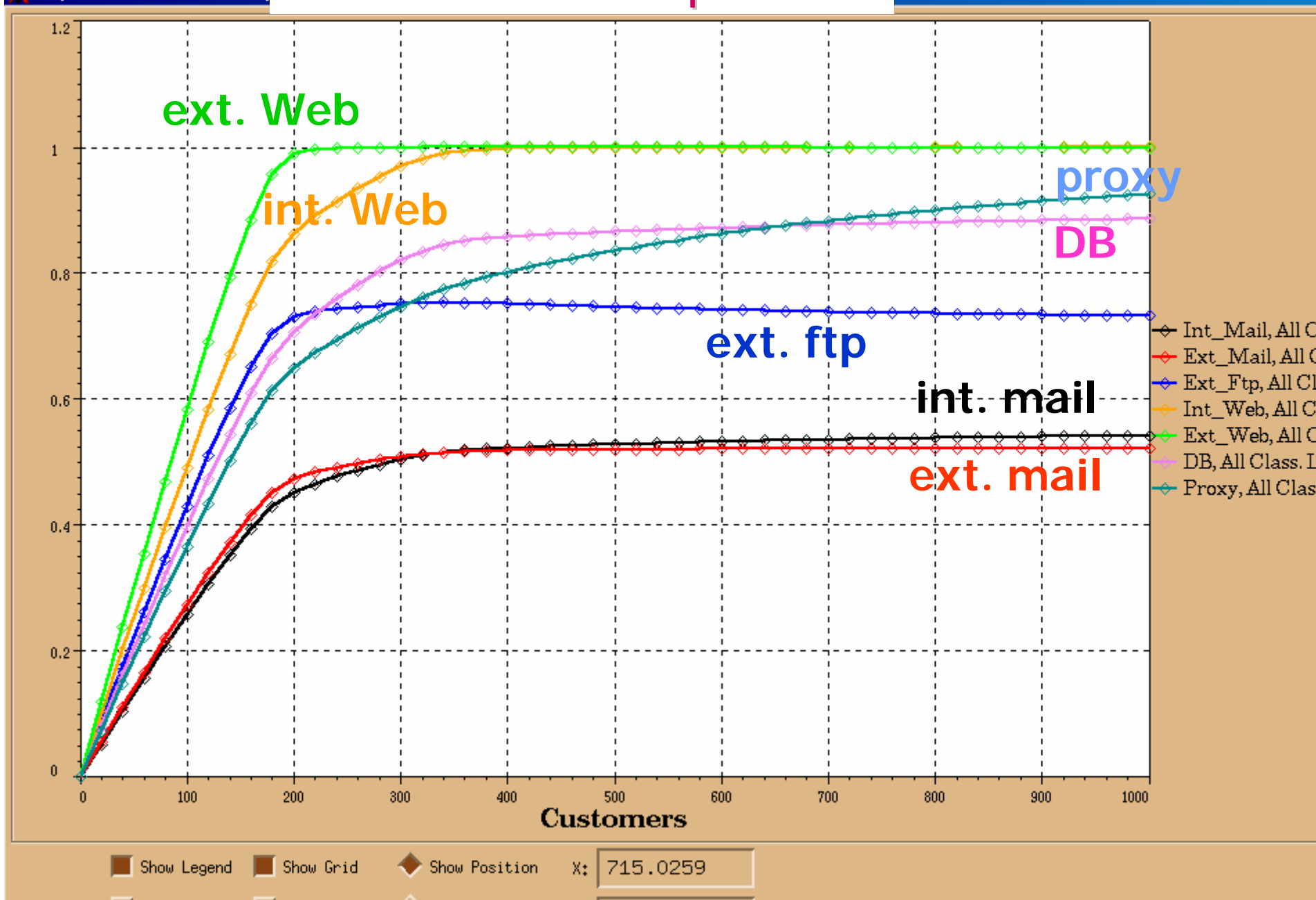
	<b>Intranet</b>	<b>Proxy</b>	<b>No Proxy</b>	<b>Mobile</b>	<b>Clients</b>
<b>Int. LAN</b>	0,01	0,01	0,01	0,01	0,01
<b>Ext. LAN</b>	0,00	0,00	0,01	0,01	0,01
<b>Int. FW</b>	0,00	0,00	0,02	0,12	0,12
<b>Ext .FW</b>	0,00	0,00	0,02	0,12	0,12
<b>Int. Mail</b>	0,08	0,08	0,08	0,08	0,08
<b>Ext. Mail</b>	0,04	0,04	0,04	0,12	0,08
<b>Ext. Ftp</b>	0,08	0,00	0,00	0,08	0,32
<b>Int. Web</b>	0,22	0,10	0,10	0,15	0,15
<b>Ext. Web</b>	0,00	0,00	0,30	0,15	0,35
<b>DB</b>	0,22	0,15	0,12	0,10	0,10
<b>Proxy</b>	0,05	0,30	0,15	0,10	0,10

# modello della intranet

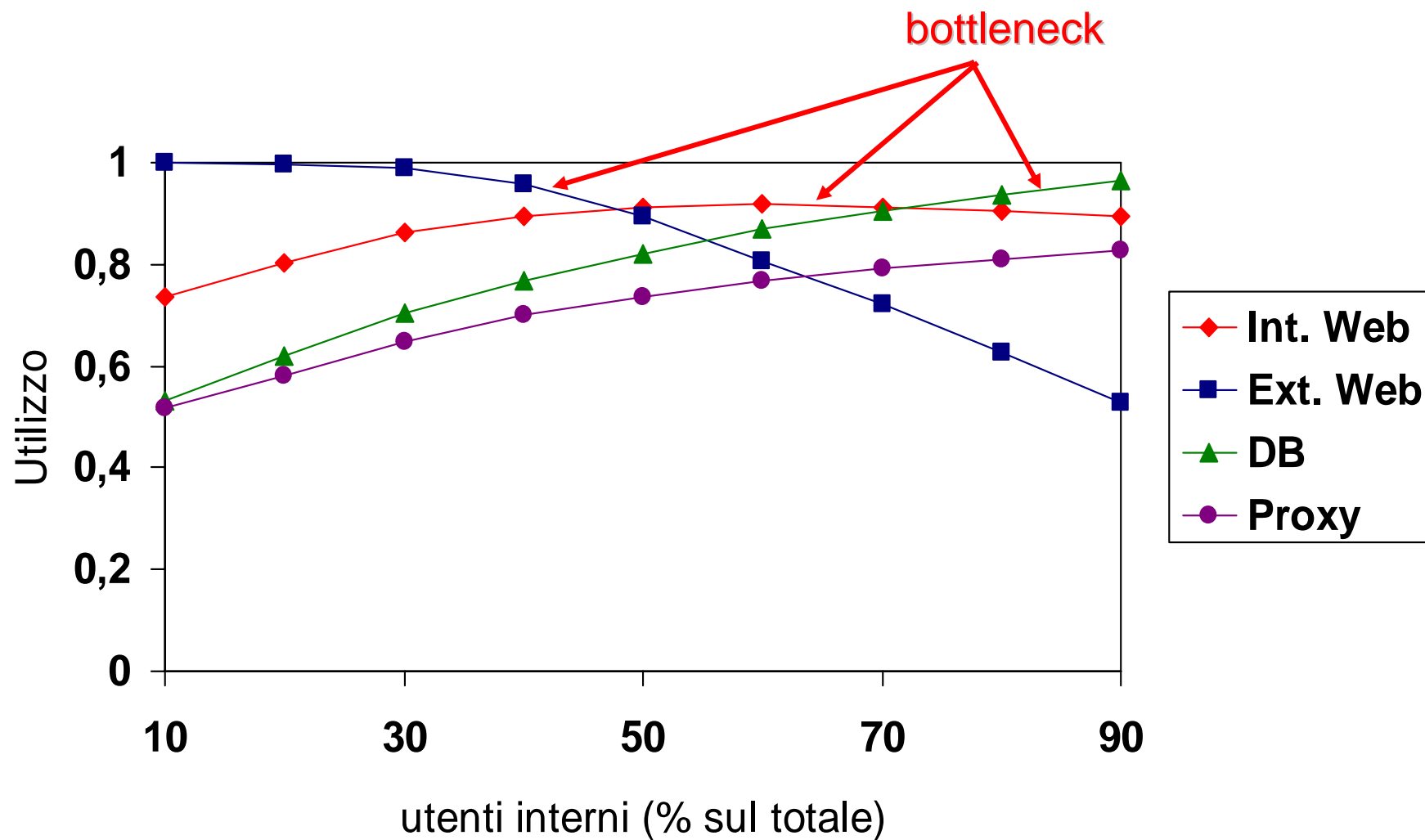


# utilizzo dei componenti

Graphics Performance Index:



## 200 utenti: variazione del "population mix"

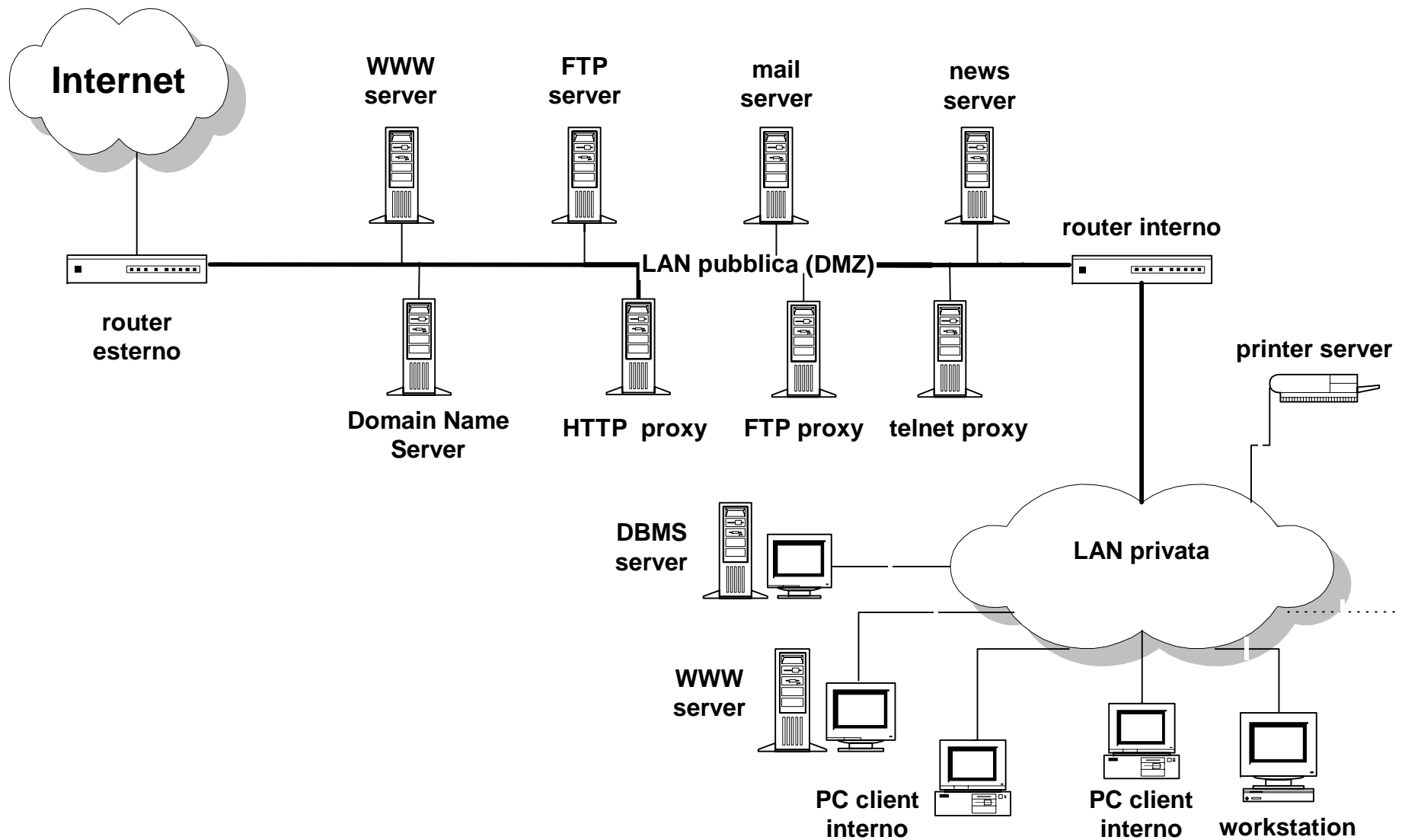


# intranet con http proxy

- influenza hit rate http proxy sulle prestazioni



# architettura



# parametri

Nome del componente	Descrizione
Internet	E' la grande rete, comprensiva di tutti i siti Web raggiungibili da un generico utente.
LAN Esterna	E' il backbone di rete che implementa la DMZ. Abbiamo scelto una Ethernet a 10 Mbps.
LAN Interna	E' il backbone di rete aziendale al quale sono connesse tutte le stazioni server e client dell'impresa. Anche qui, abbiamo scelto una Ethernet a 10 Mbps.
HTTP Proxy	E' il calcolatore che fa da application gateway per il protocollo HTTP. Qui vengono memorizzate le pagine HTML.
Web Interno	E' il sito Web dell'azienda accessibile solo dagli utenti interni all'azienda stessa.

Nome della classe	Nome simbolico	Descrizione
RICH	Richiesta HTTP di una pagina HTML	Una richiesta è un pacchetto di 1500 byte
PAGE	Pagina HTML risultato di una richiesta HTTP	Una pagina è un frame di 150 Kbyte
NEWPAGE	Pagina HTML acquisita dal sistema tramite un accesso ad Internet	Una pagina è un frame di 150 Kbyte

Done

Save

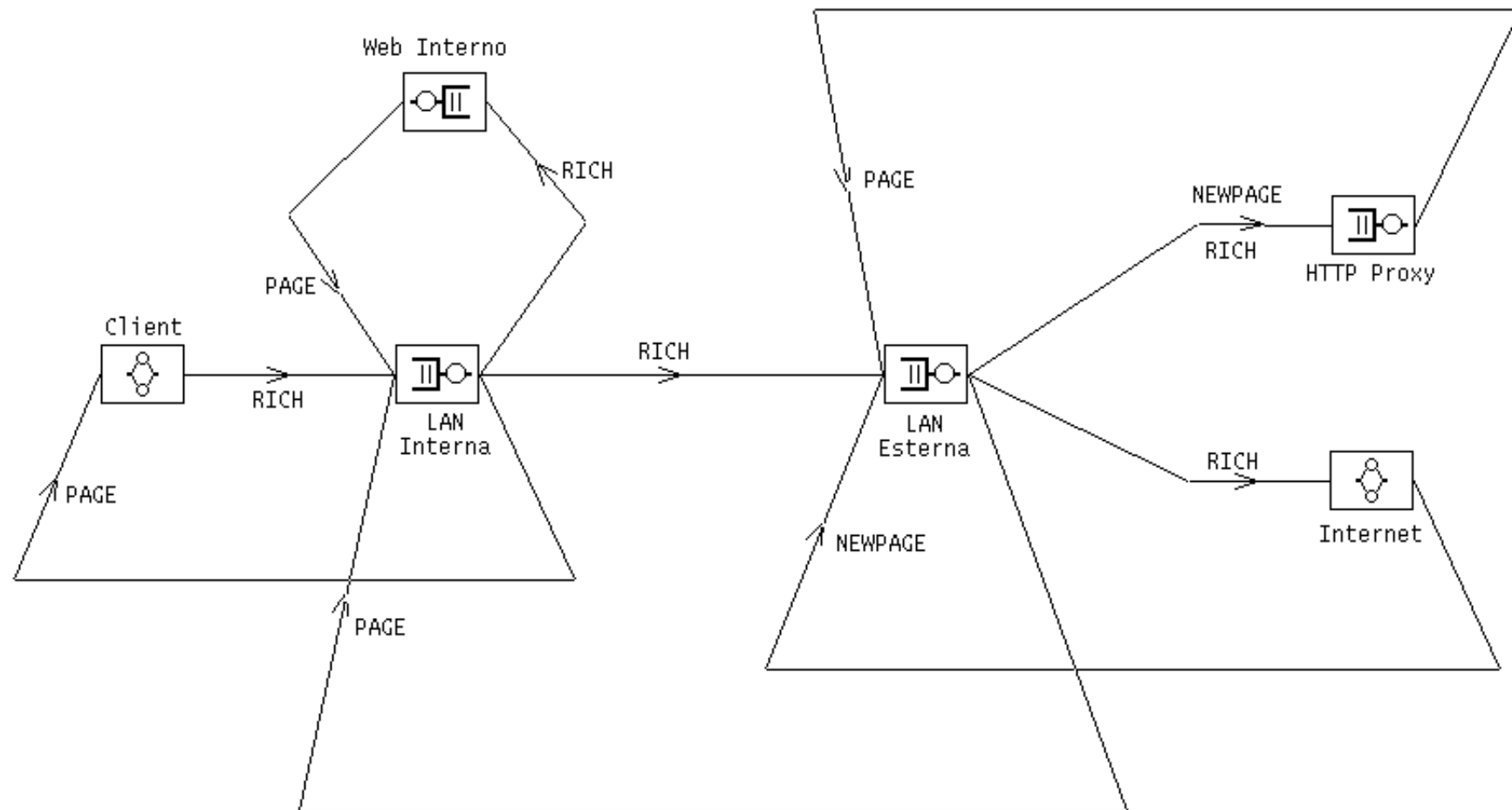
Abort

Left button - Move object

Middle button - Create node

Right button - Show operation menu

# modello del flusso del traffico HTTP



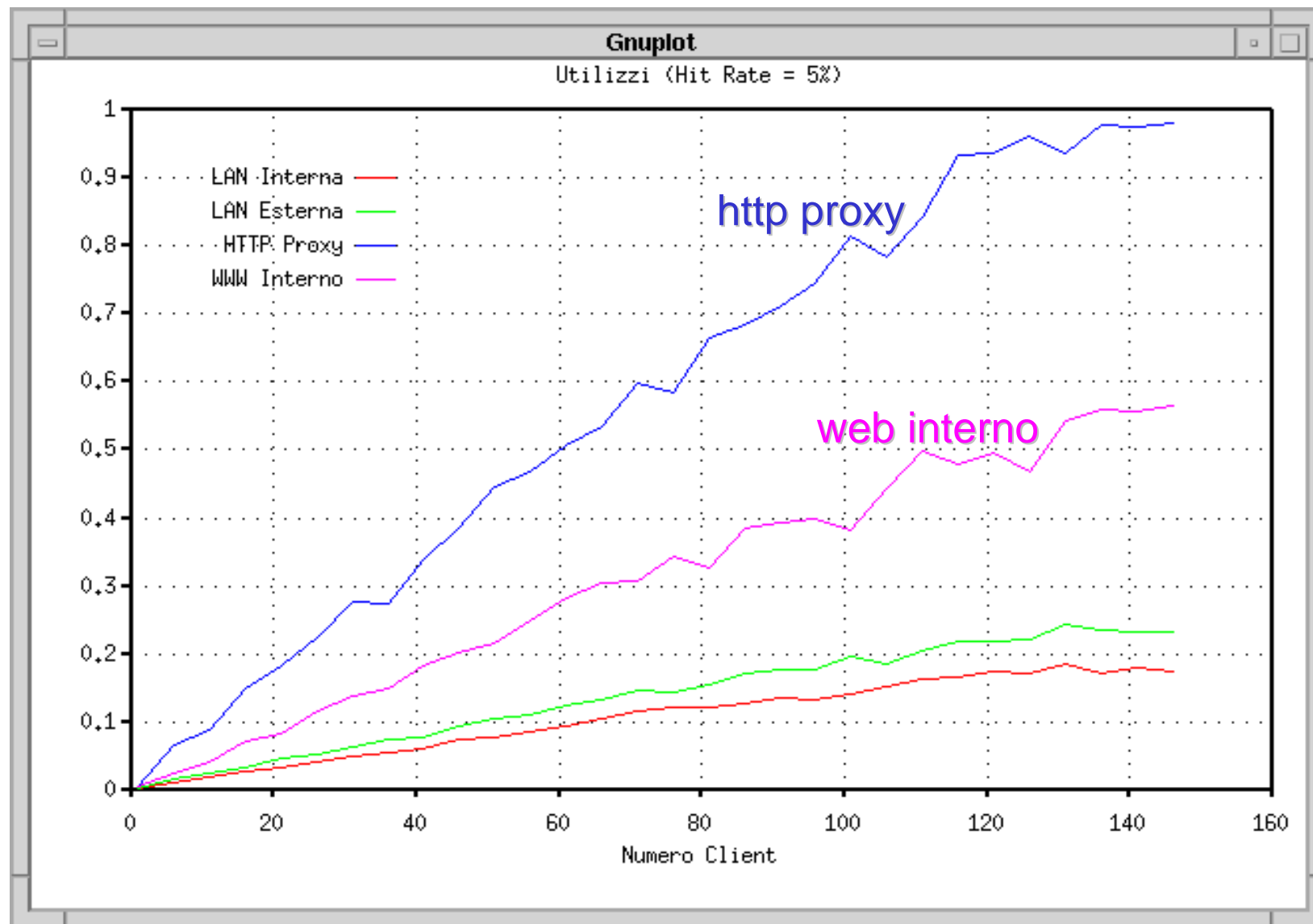
# parametri esperimento 1

Nome grandezza	Valori	Commenti
Numero di Client	<b>da 1 a 146 step 5</b>	Questo è l'unico parametro variabile nell'esperimento
Think time dei Client	50 sec.	Distribuzione esponenziale
Percentuale di accessi al Web Interno sul totale	33%	
Tempo di servizio del Web Interno	1 sec.	Il Web Interno elabora solo job di classe RICH
Hit Rate dell'HTTP Proxy	<b>5%, 50% ed 80%</b>	Ad ogni valore corrisponde una distinta esec. della simulazione
Tempo di servizio HTTP Proxy	1 sec.	Per job di classe RICH
	1 sec.	Per job di classe PAGE
Tempo di servizio della LAN Interna e della LAN Esterna	0.0012 sec.	Per job di classe RICH
	0.12 sec.	Per job di classe PAGE
Tempo di delay per Internet	40 sec. Coeff. Variaz. = 10	Distribuzione iperesponenziale

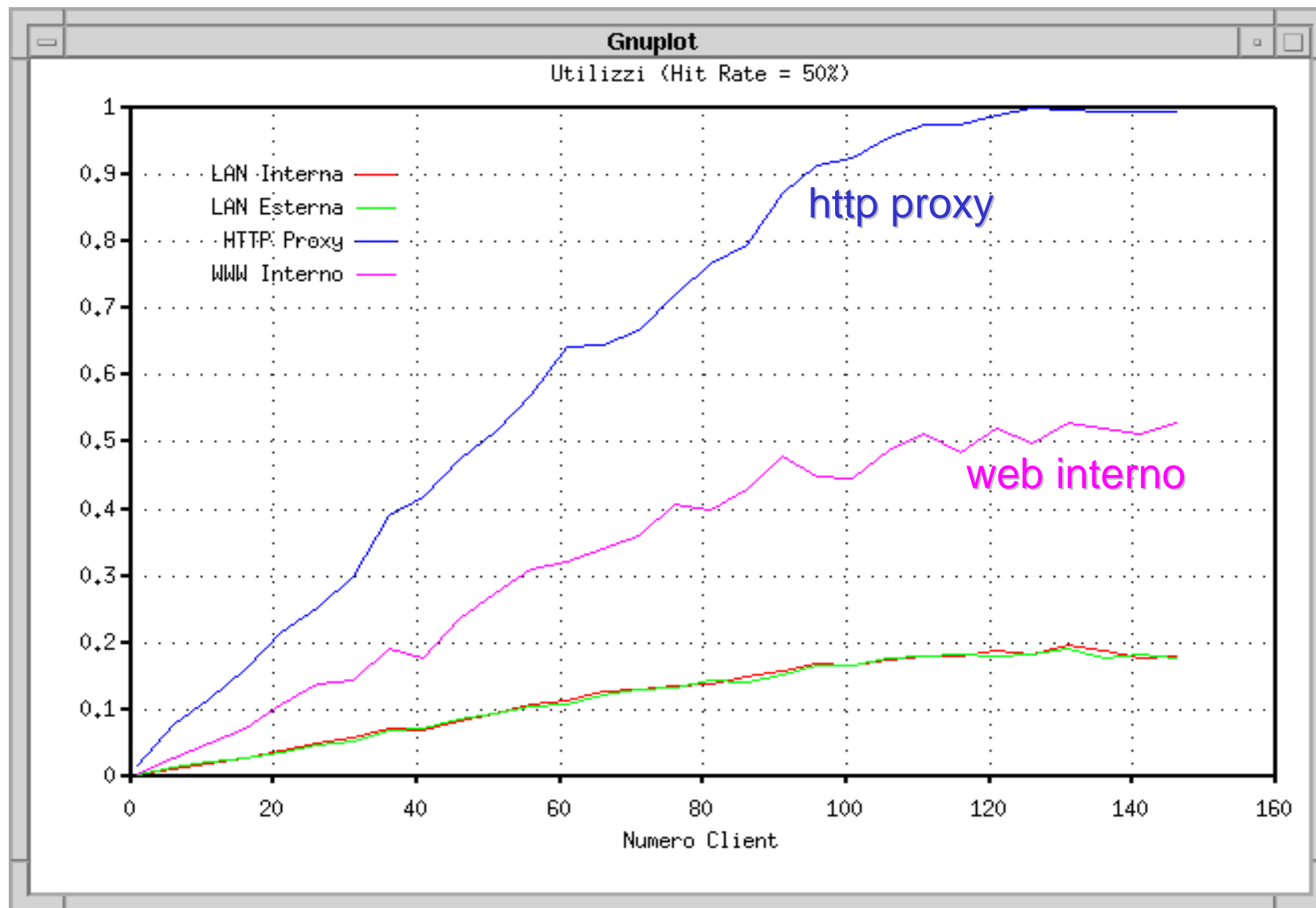
## parametri esperimento 2

Nome grandezza	Valori	Commenti
Numero di Client	<b>80, 120 e 145</b>	Ad ogni valore corrisponde una distinta esec. della simulazione
Think time dei Client	50 sec.	Distribuzione esponenziale
Percentuale di accessi al Web Interno sul totale	33%	
Tempo di servizio del Web Interno	1 sec.	Il Web Interno elabora solo job di classe RICH
Hit Rate dell'HTTP Proxy	<b>da 14% a 60%</b>	Questo è l'unico parametro variabile nel sistema
Tempo di servizio HTTP proxy	1 sec.	Per job di classe RICH
	1 sec.	Per job di classe PAGE
Tempo di servizio della LAN Interna e della LAN Esterna	0.0012 sec.	Per job di classe RICH
	0.12 sec.	Per job di classe PAGE
Tempo di delay per Internet	40 sec. Coeff. Variaz. = 10	Distribuzione iperesponenziale

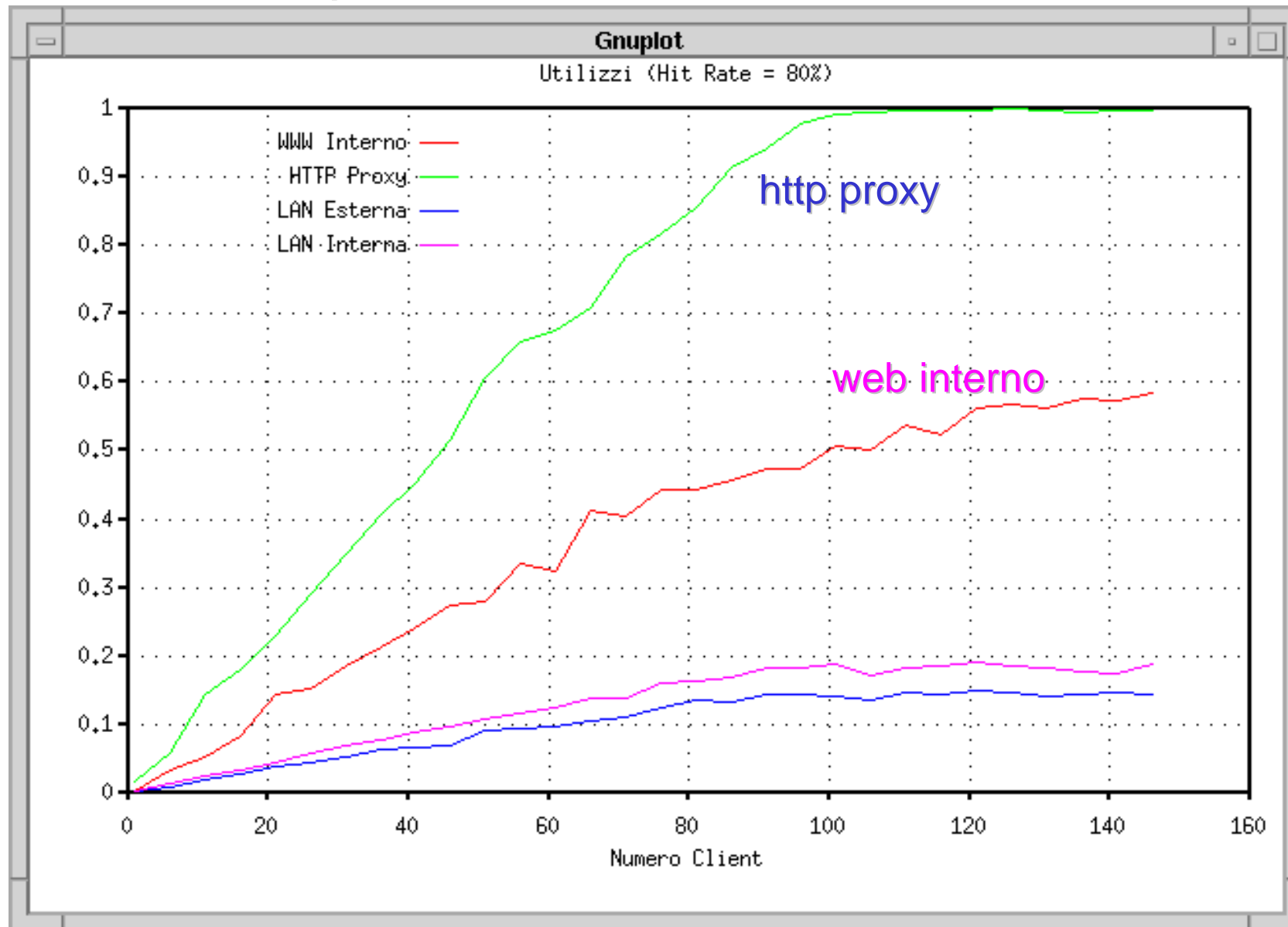
esp.1, utilizzi: hit rate 5%



esp.1, utilizzi: hit rate 50%

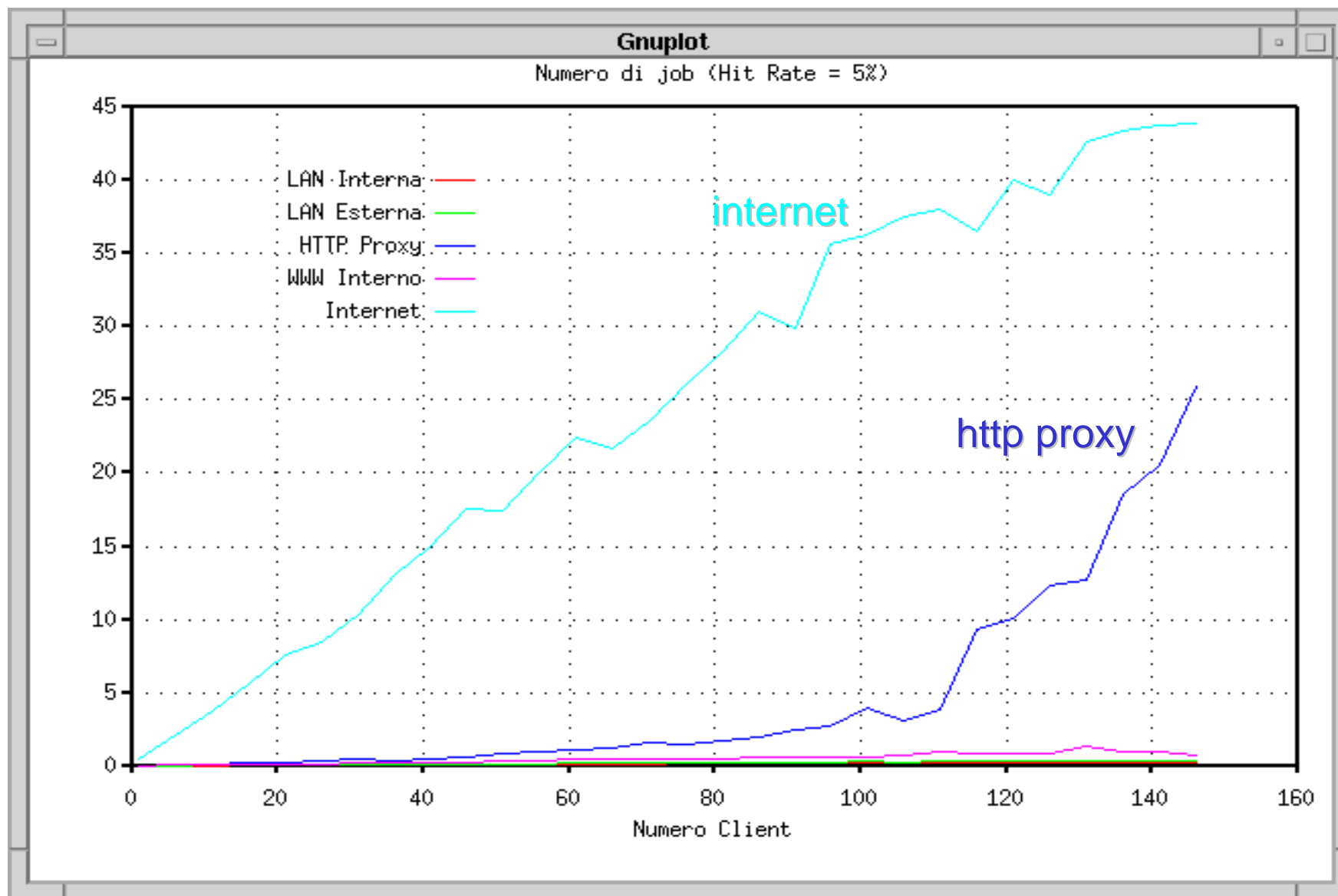


esp.1, utilizzi: hit rate 80%

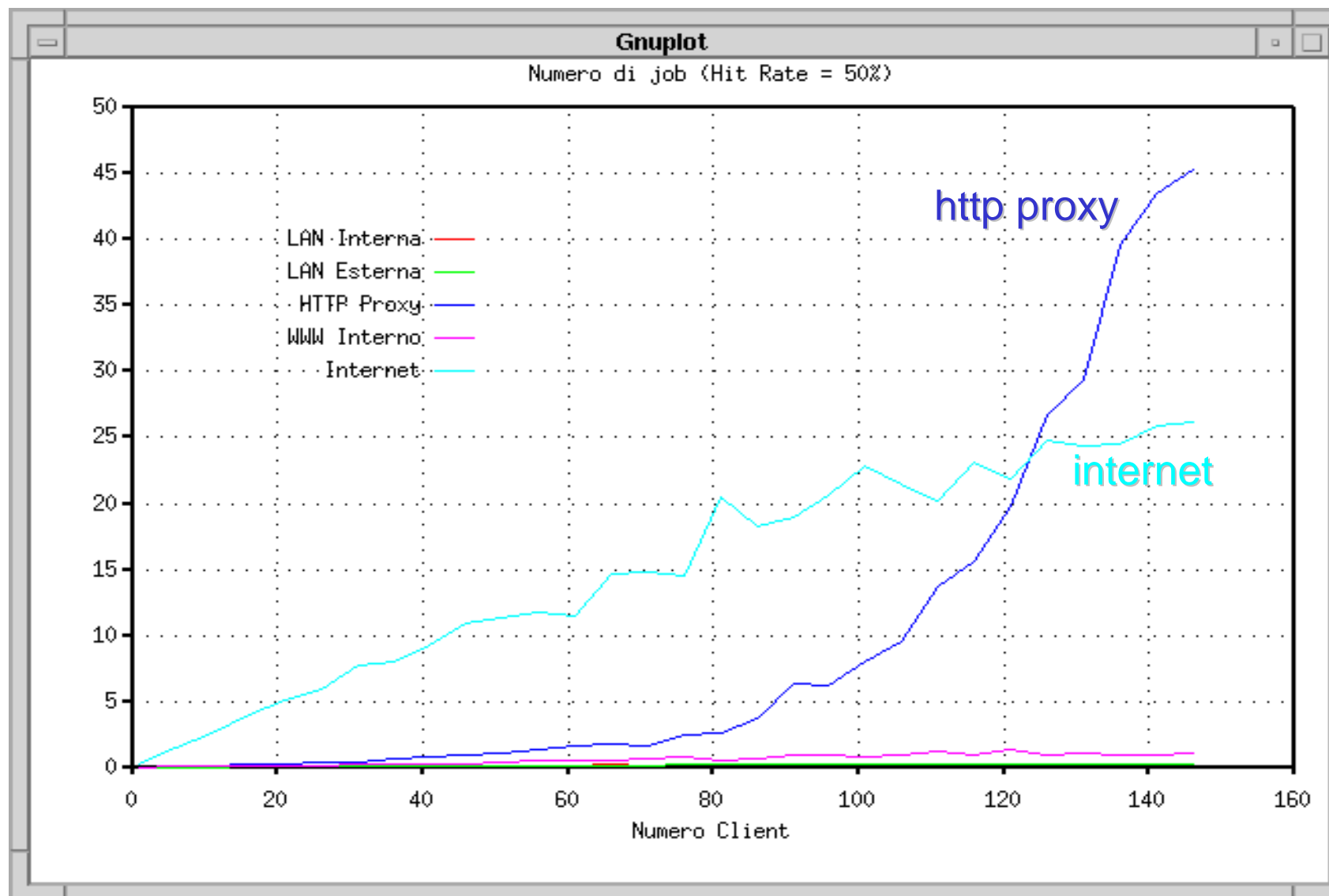




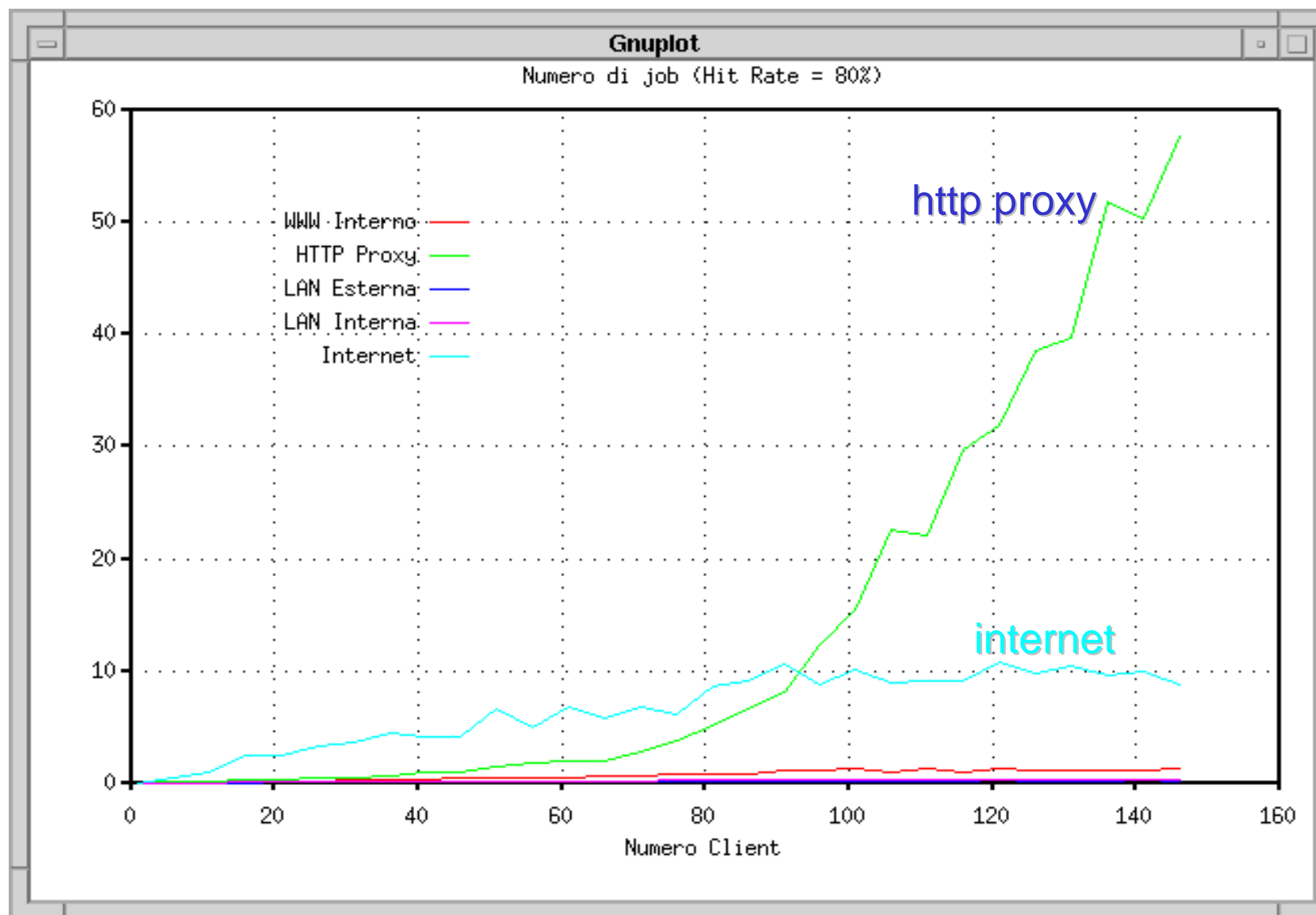
esp.1, num. client: hit rate 5%



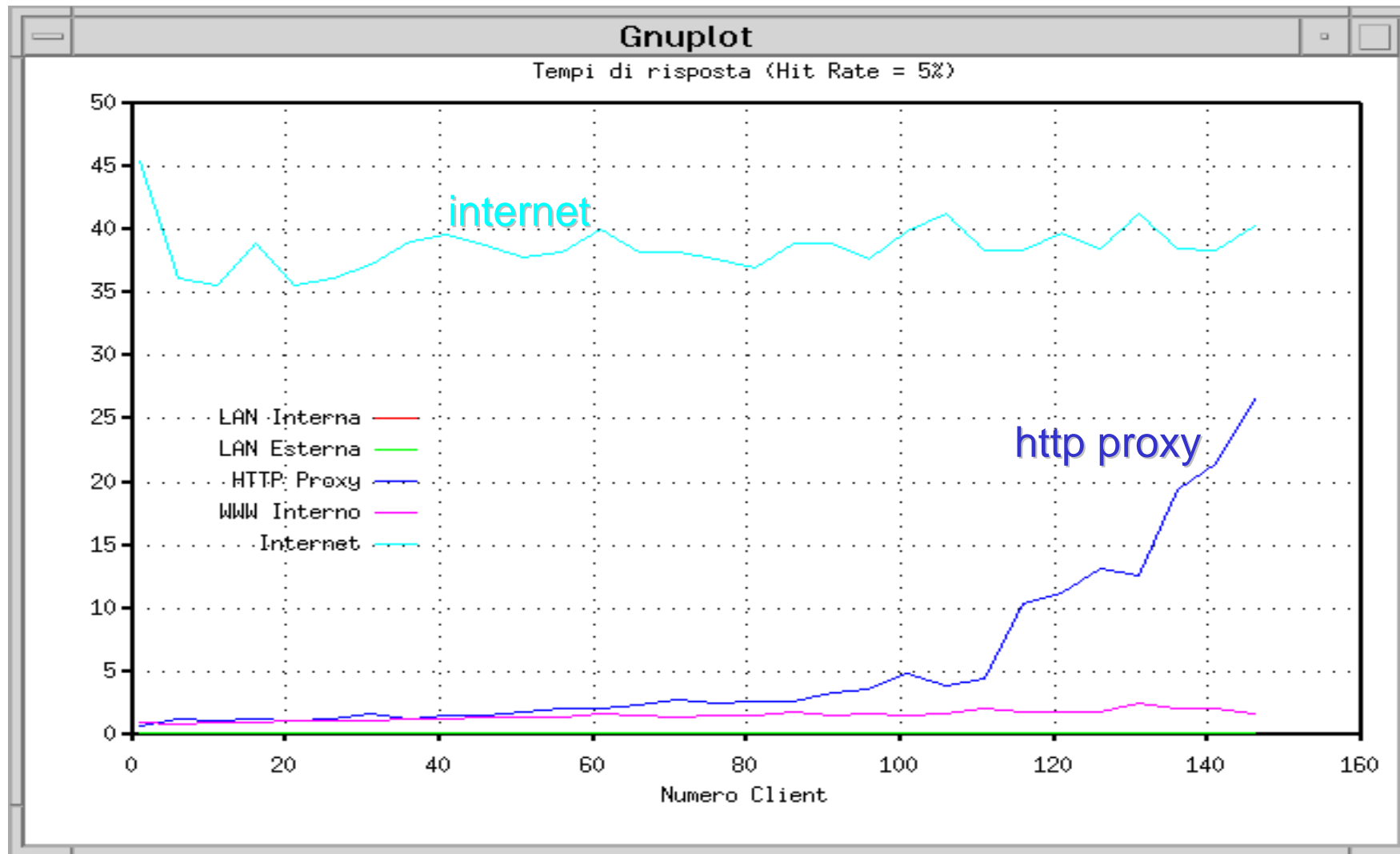
esp.1, num. client: hit rate 50%



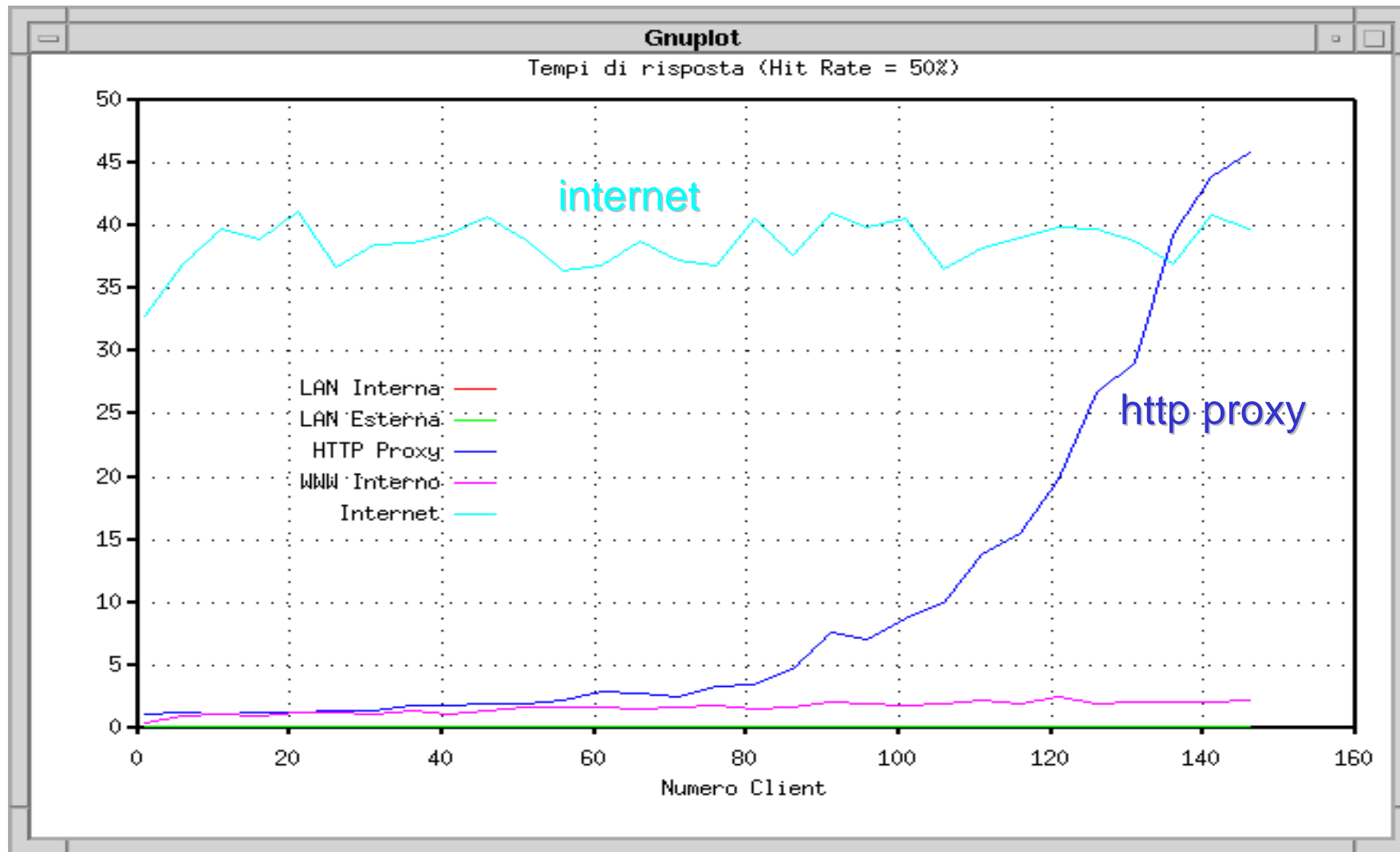
esp.1, num. client: hit rate 80%



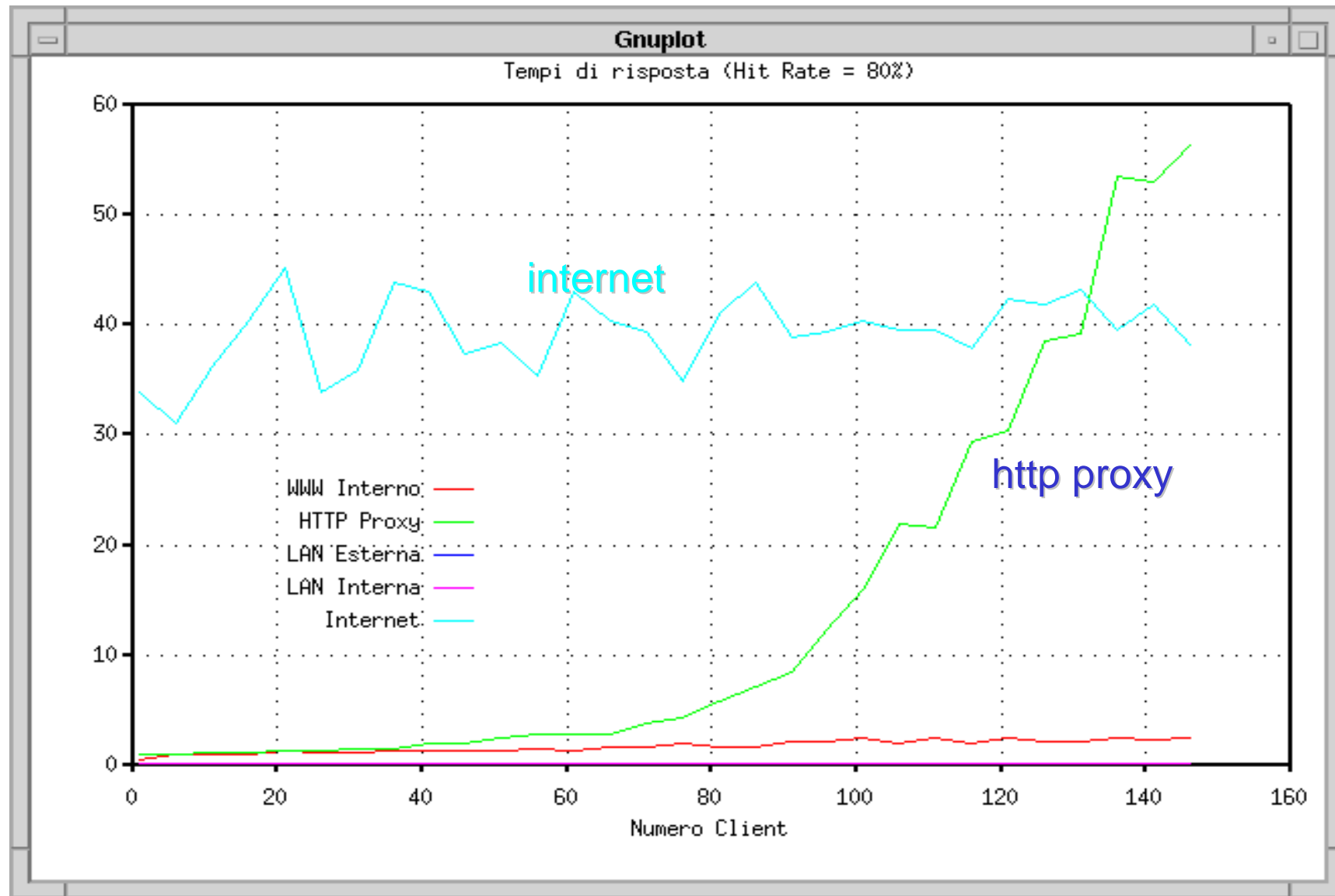
esp.1, tempi di risposta: hit rate 5%



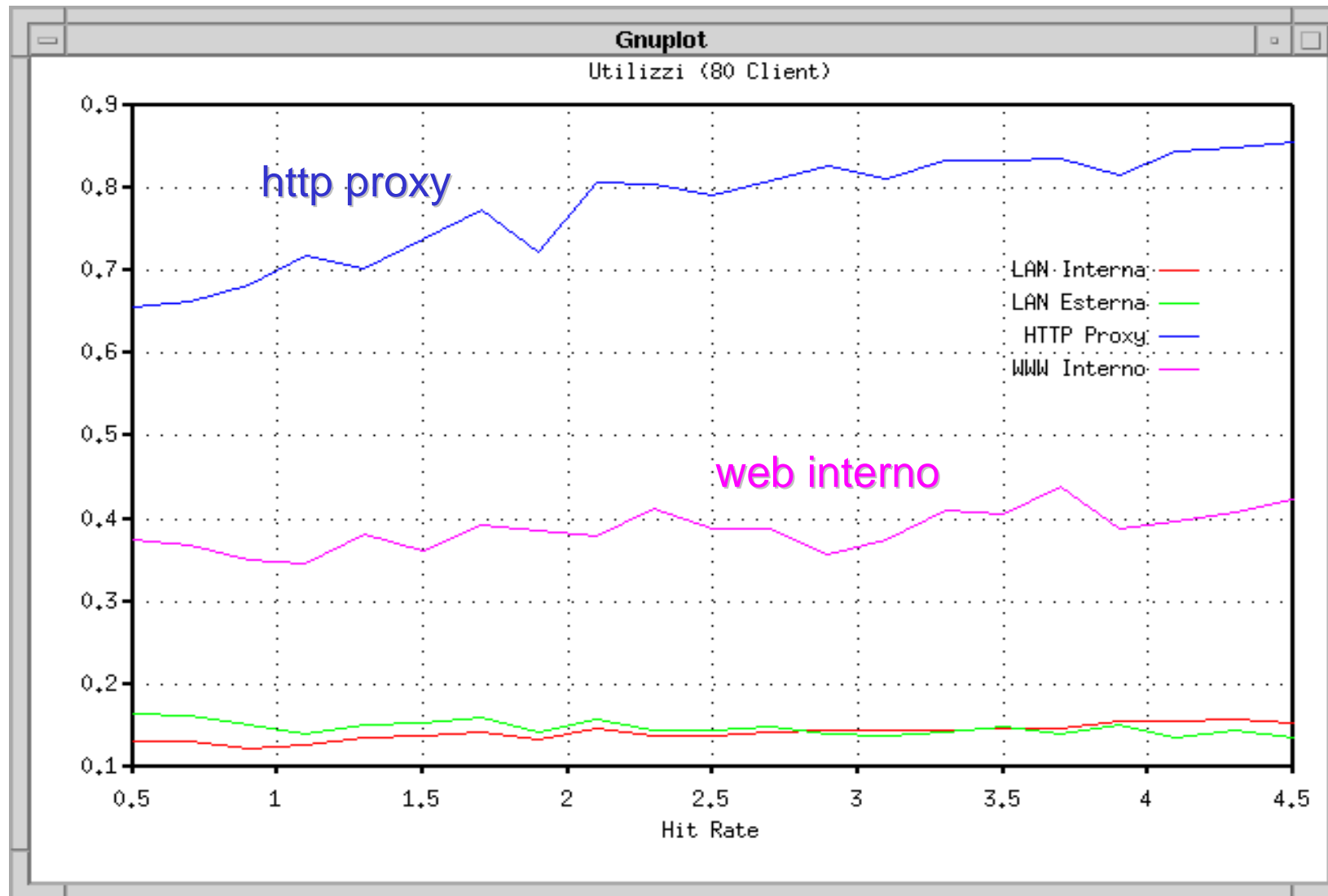
esp.1, tempi di risposta: hit rate 50%



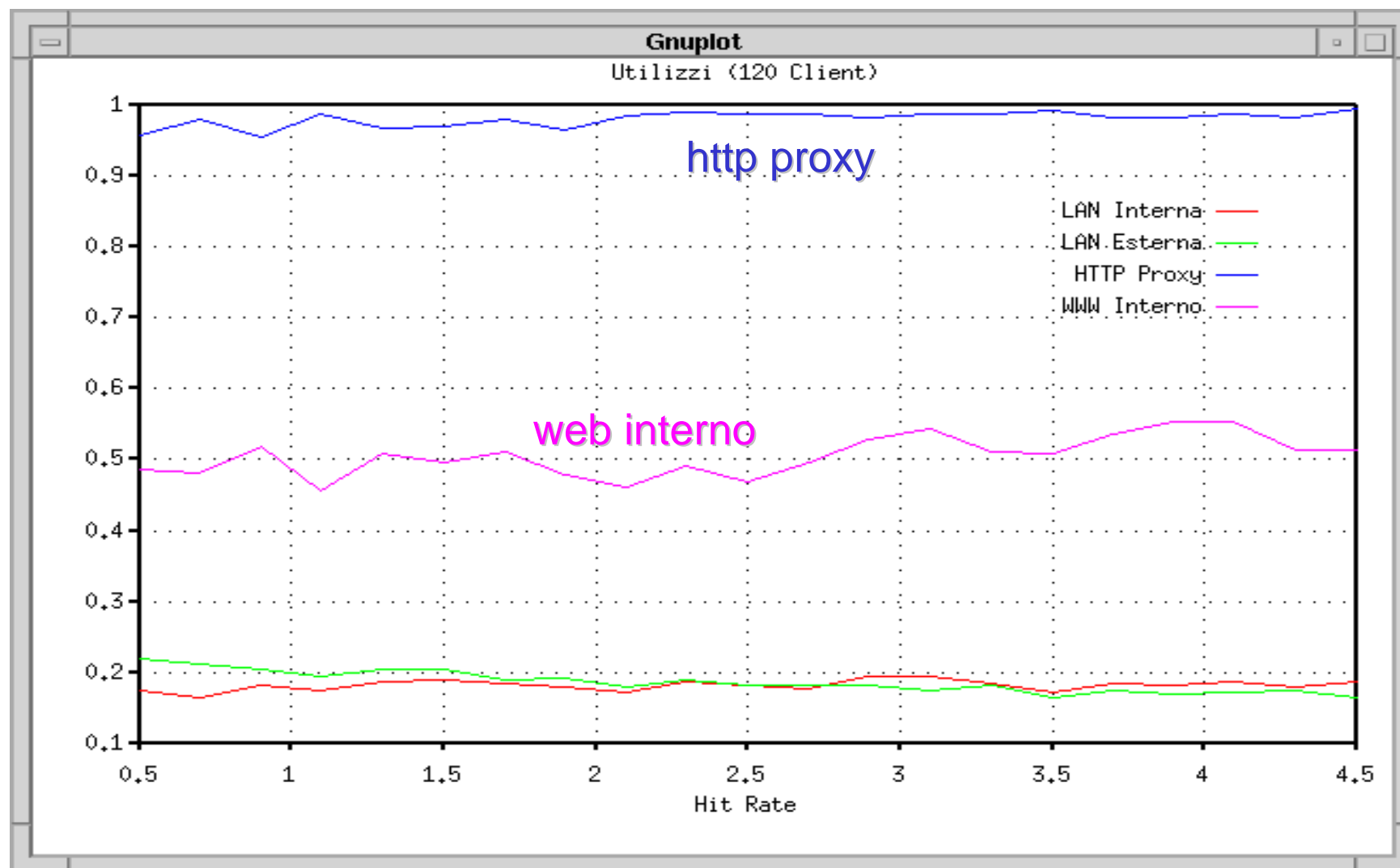
esp.1, tempi di risposta: hit rate 80%



esp.2, utilizzi: 80 client

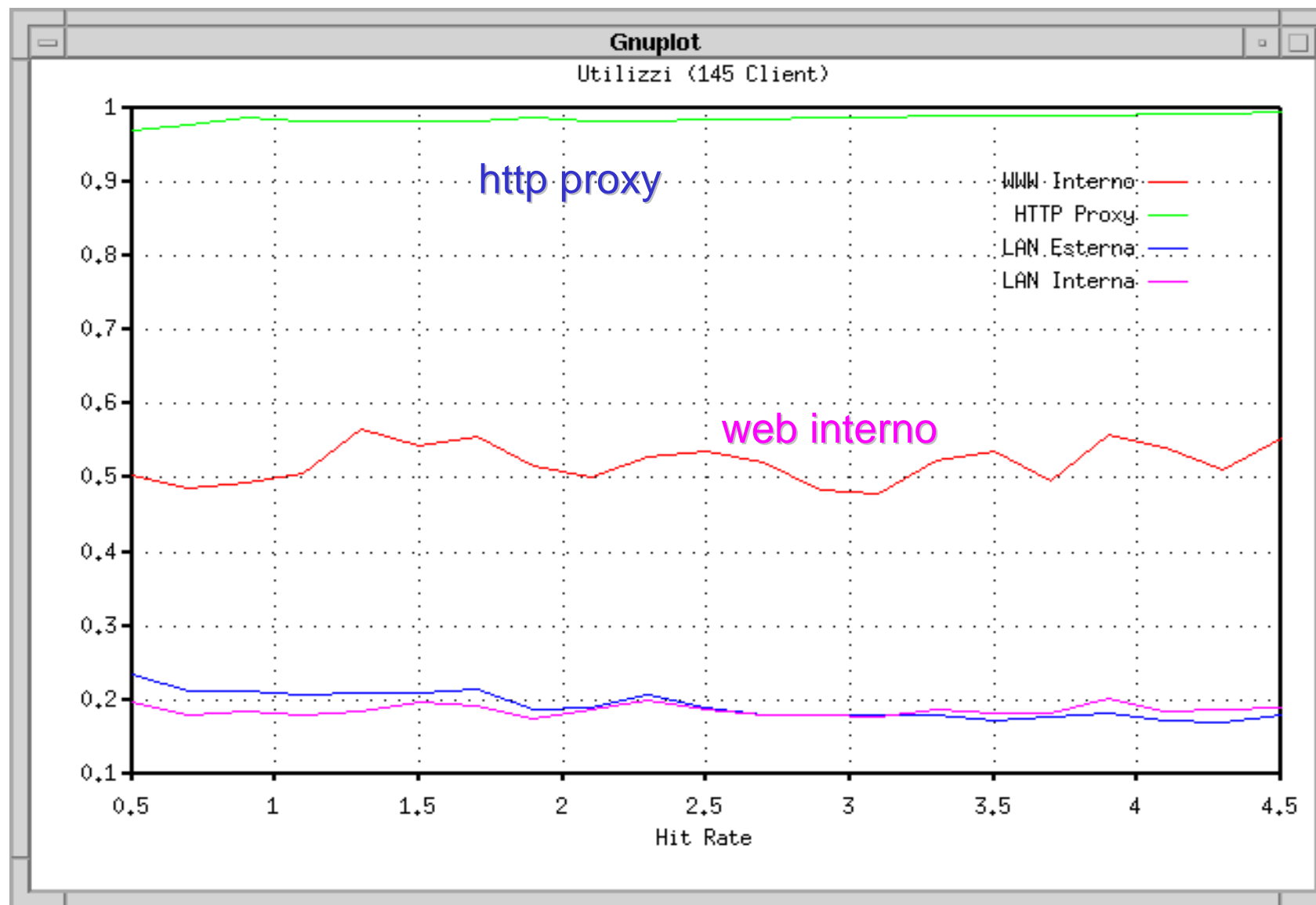


esp.2, utilizzi: 120 client

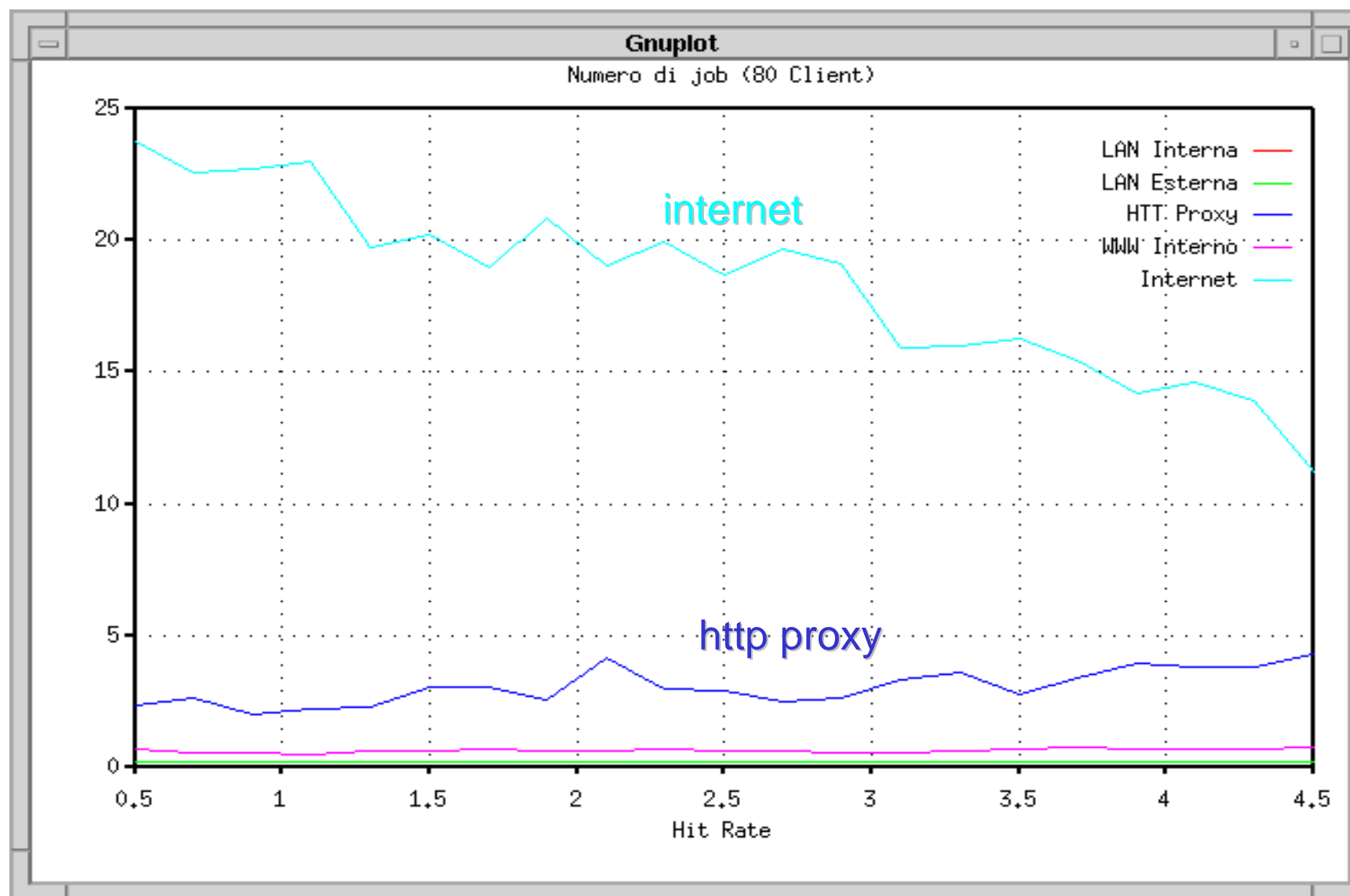




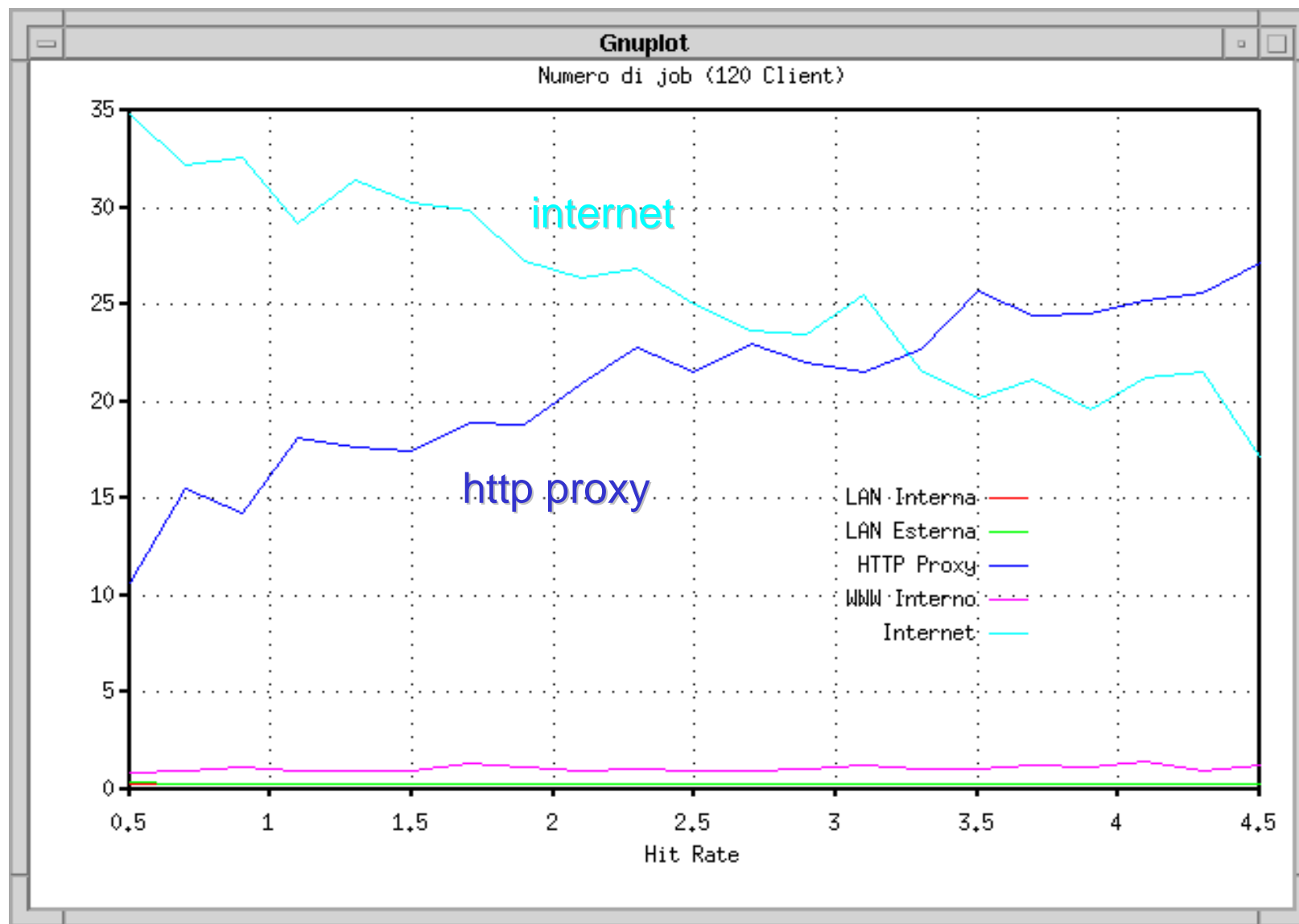
esp.2, utilizzi: 145 client



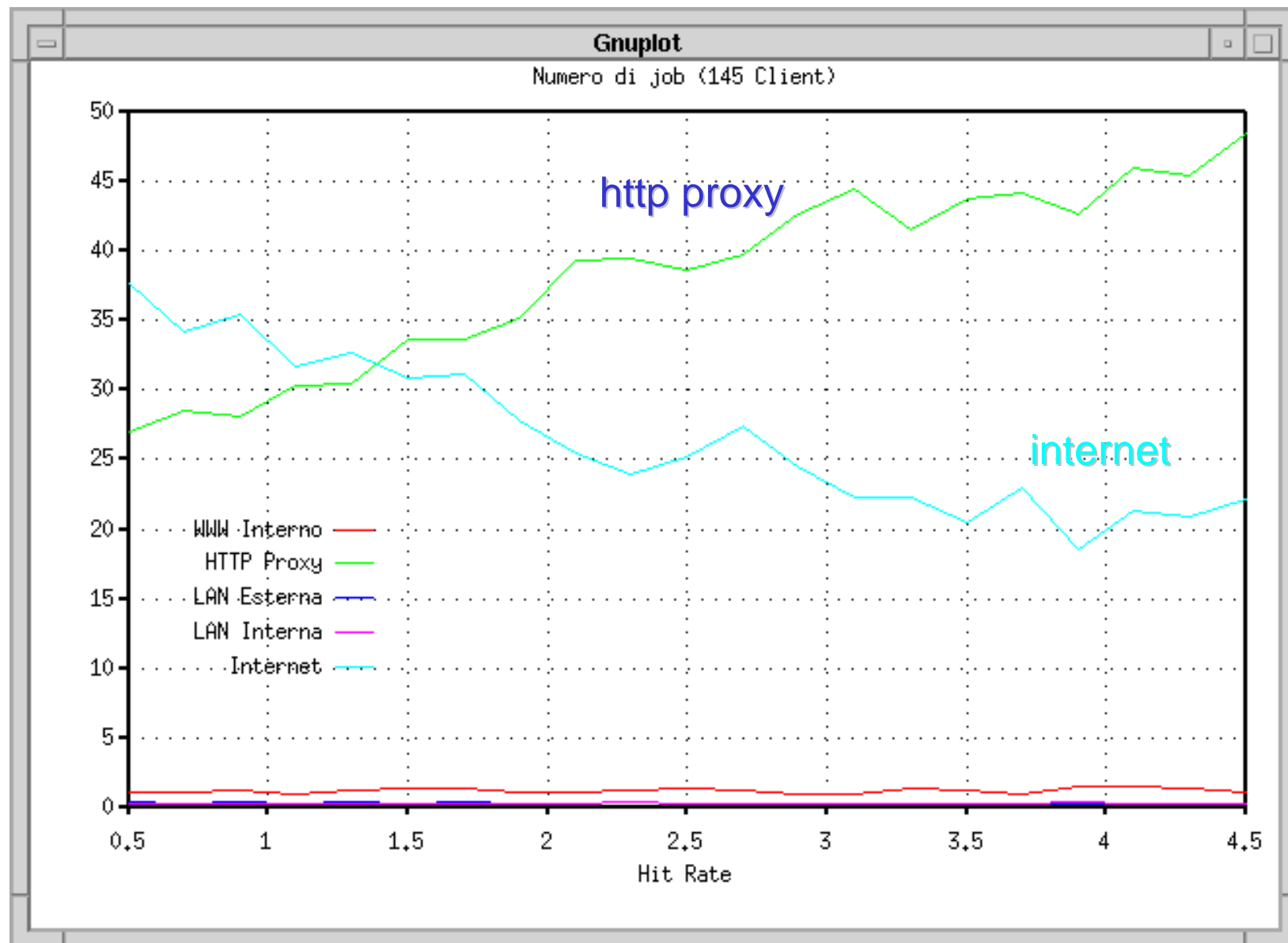
esp.2, num. client: 80 client



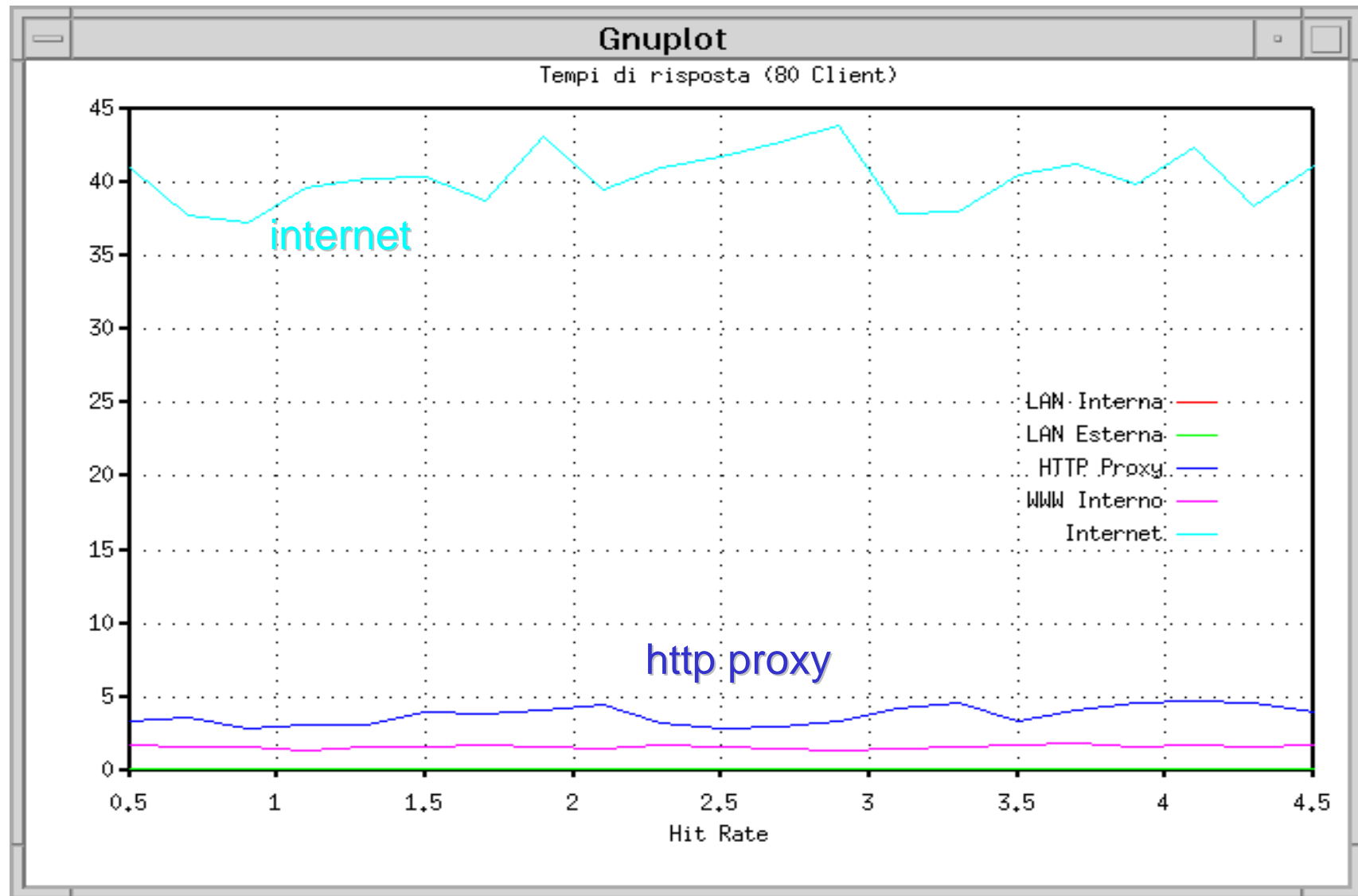
esp.2, num. client: 120 client



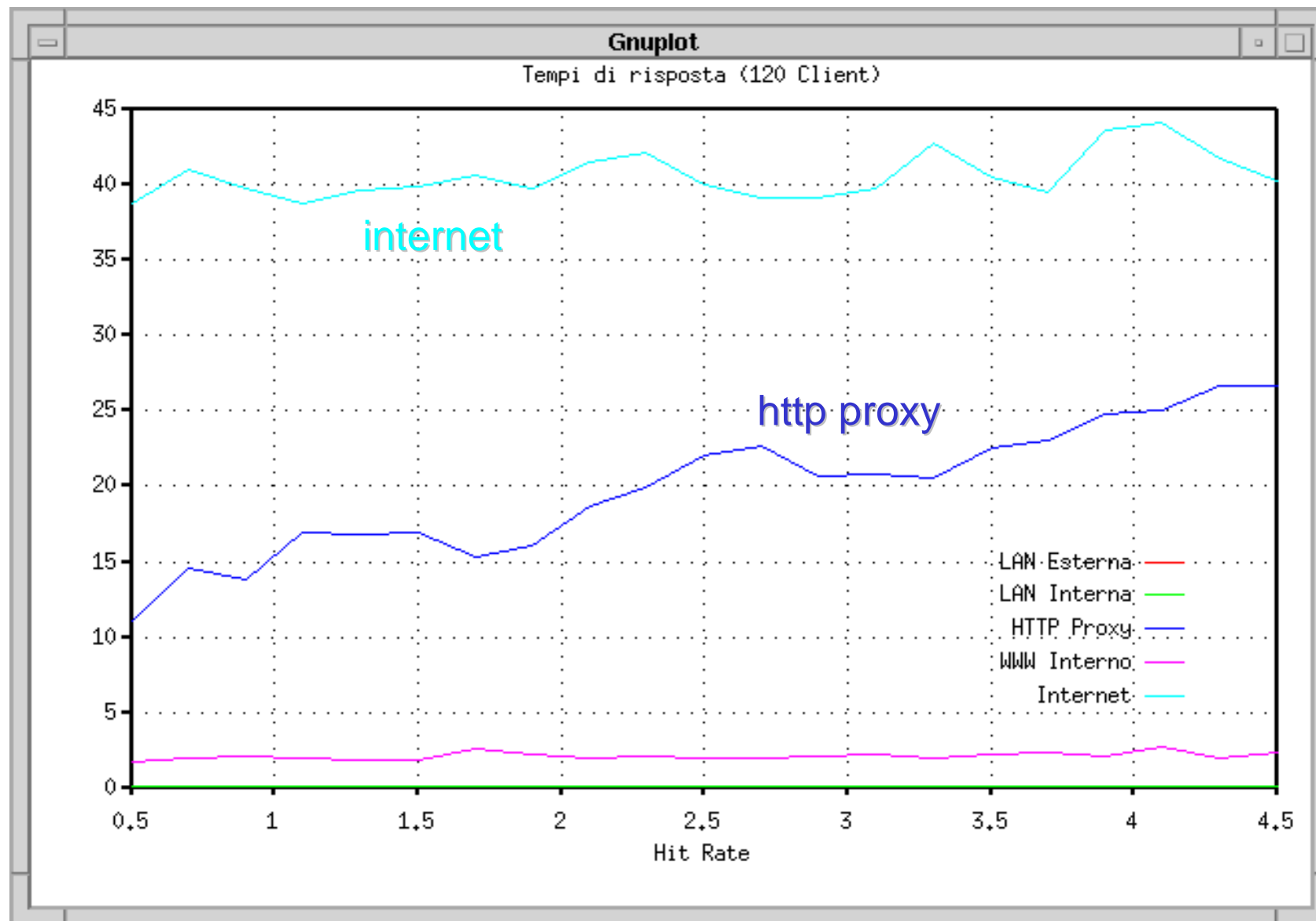
esp.2, num. client: 145 client



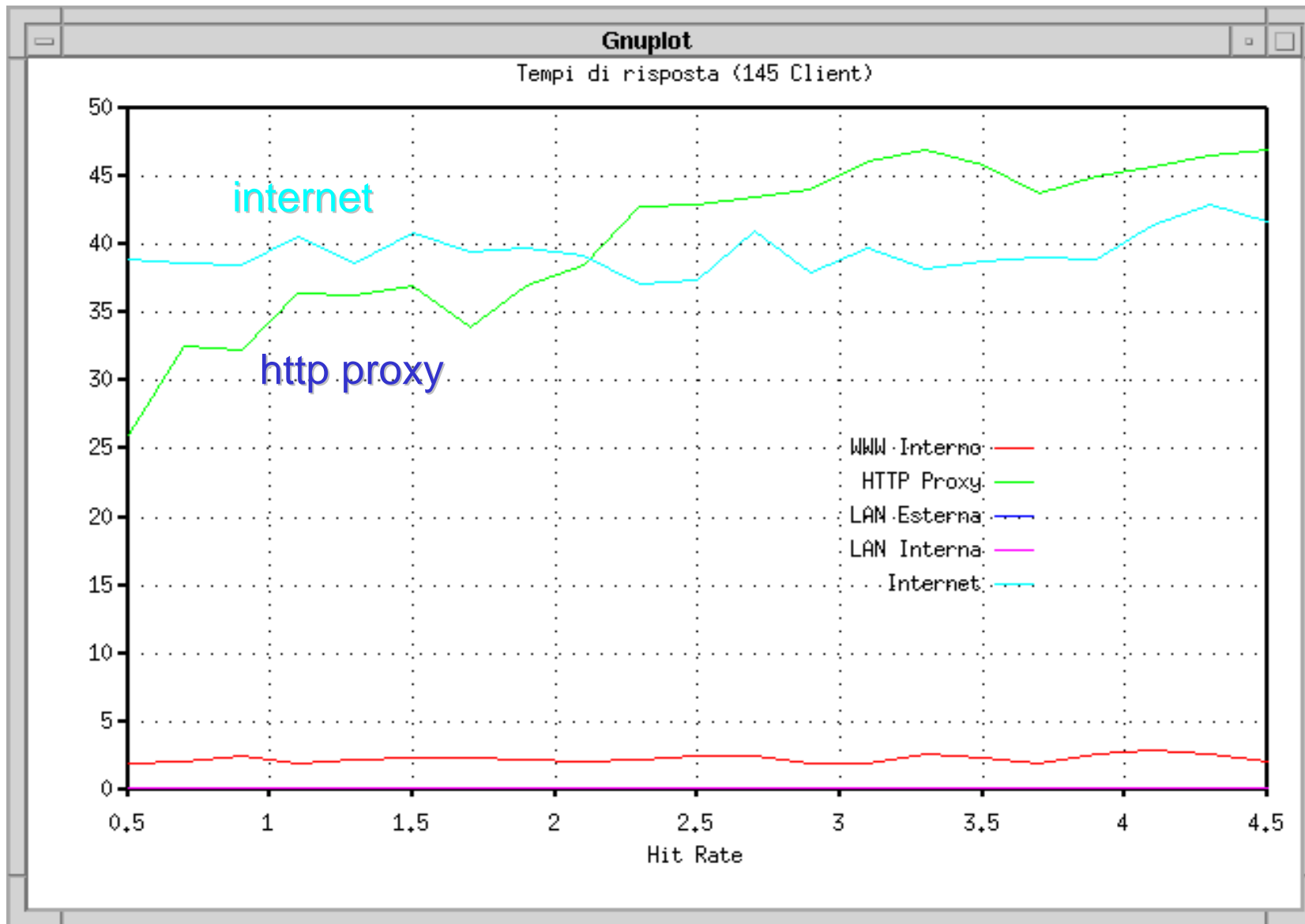
esp.2, tempi di risposta: 80 client



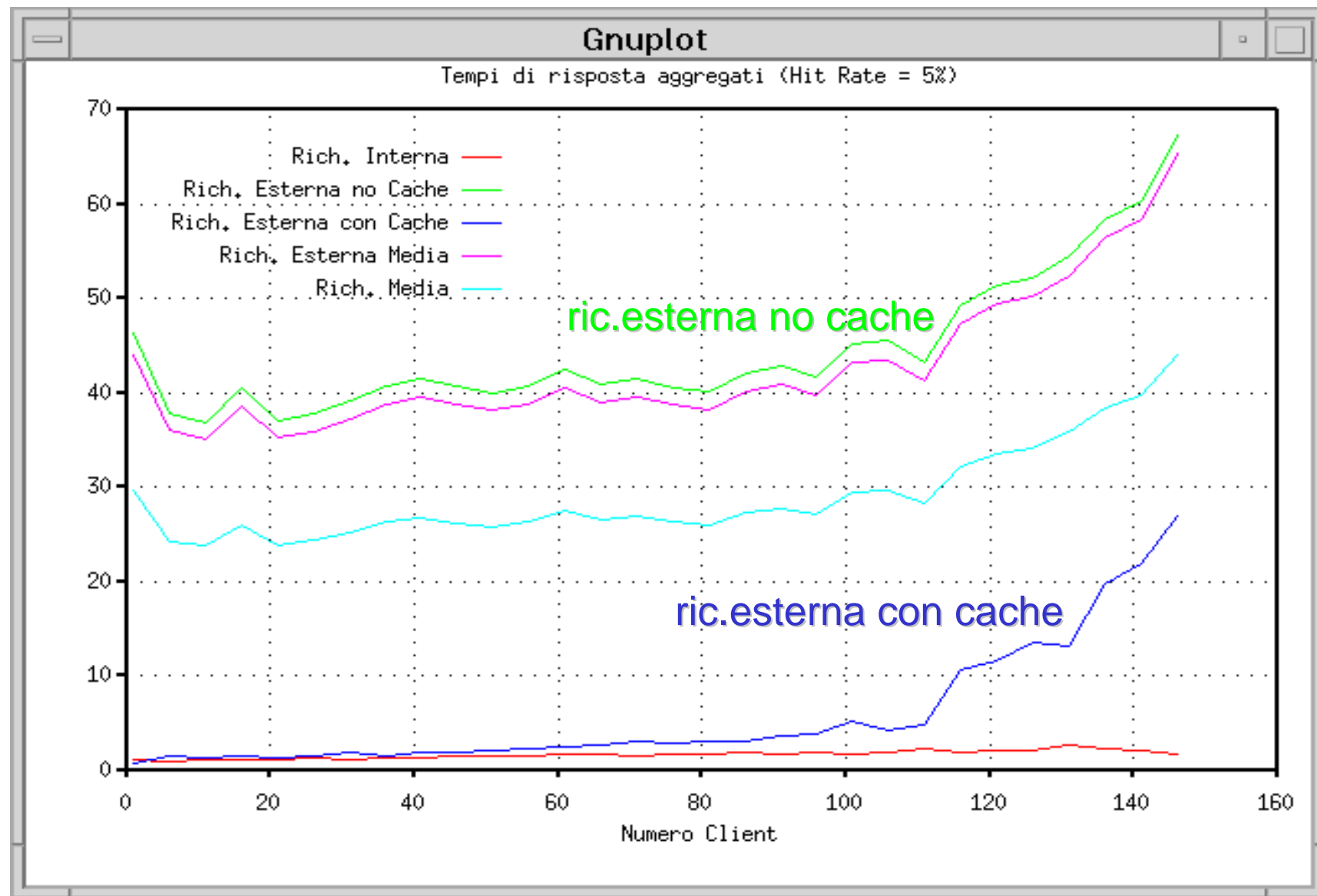
## esp.2, tempi di risposta: 120 client



## esp.2, tempi di risposta: 145 client

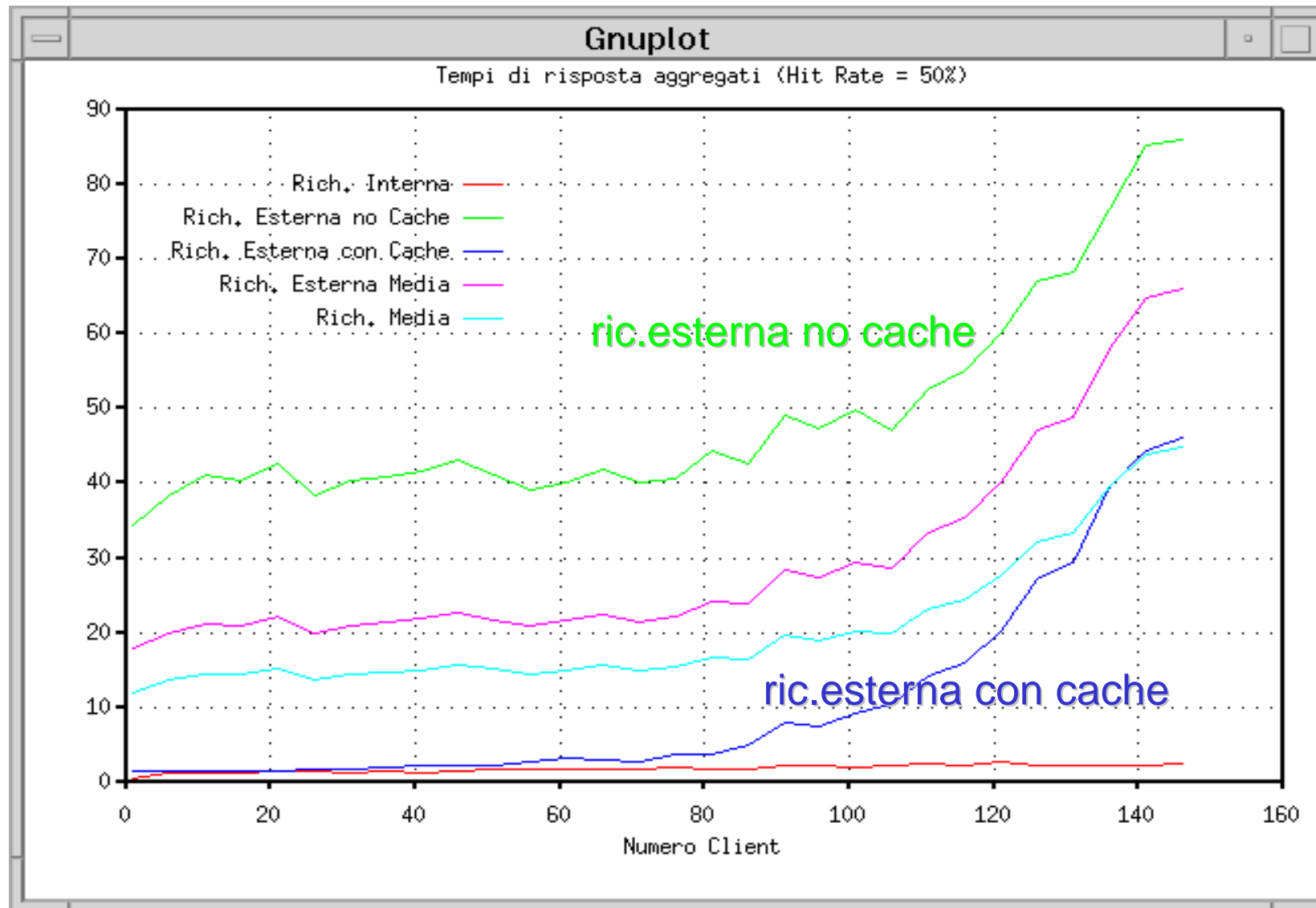


esp.1, tempi di risposta aggregati: hit rate 5%

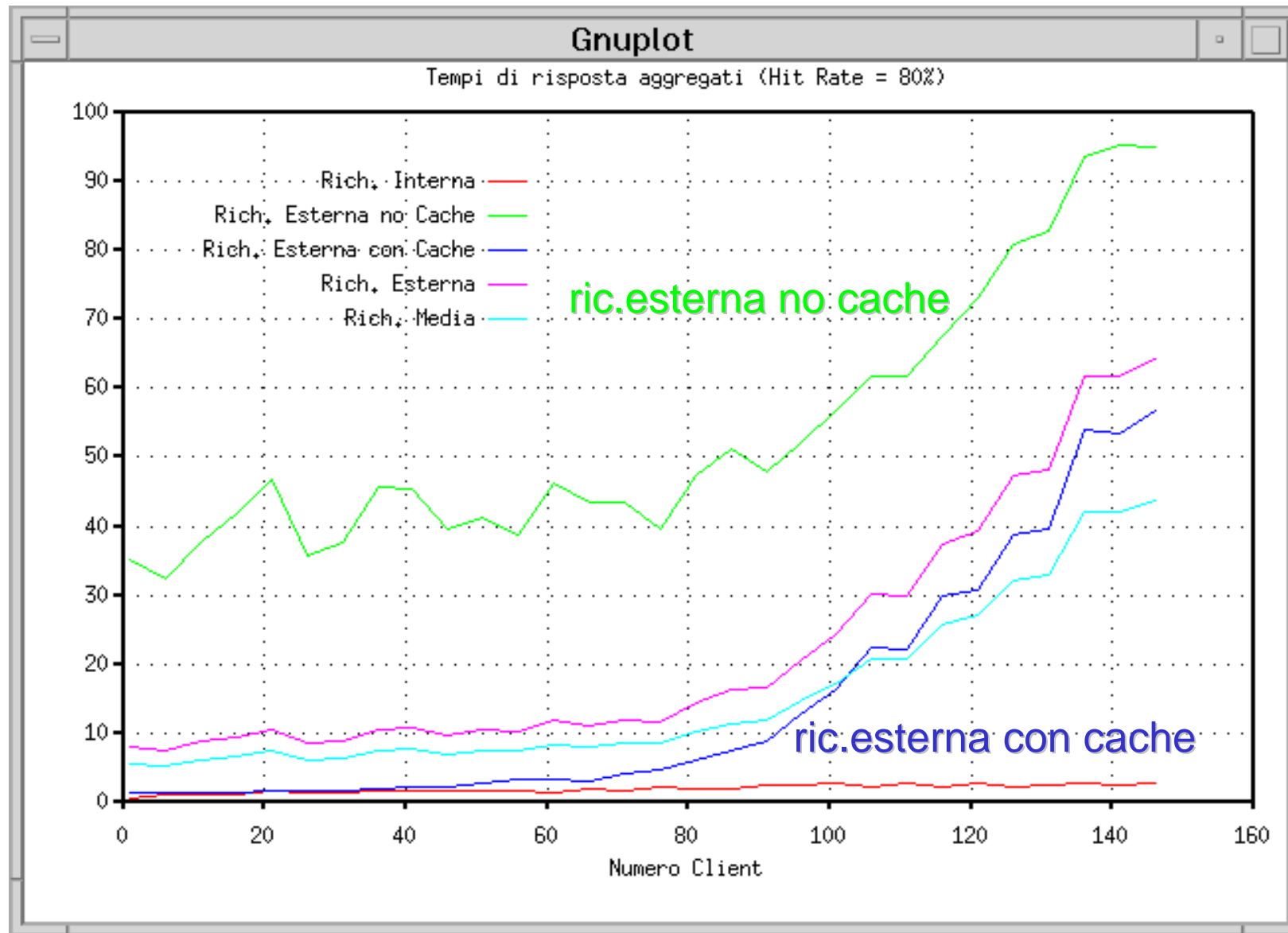




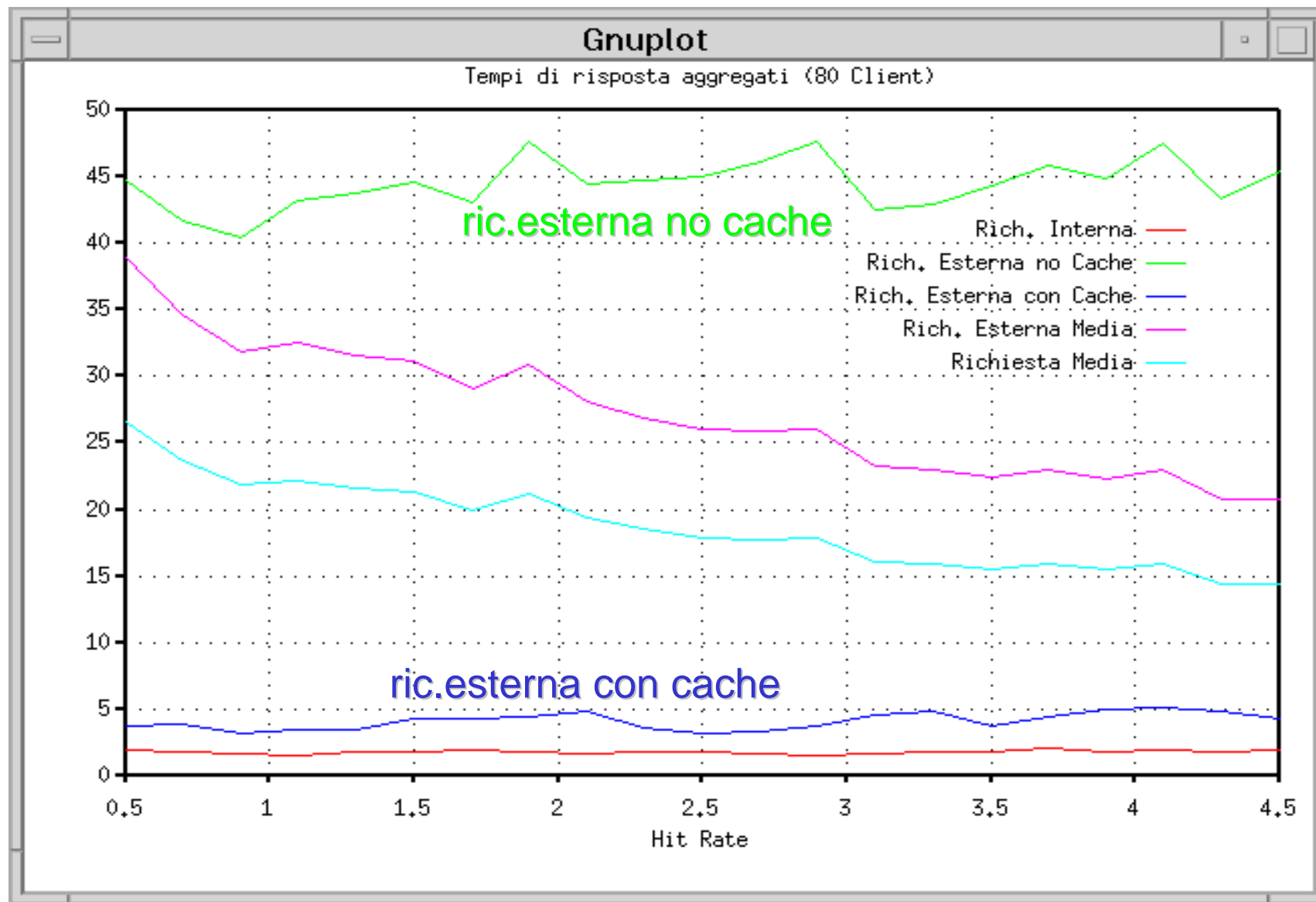
esp.1, tempi risposta aggregati: hit rate 50%



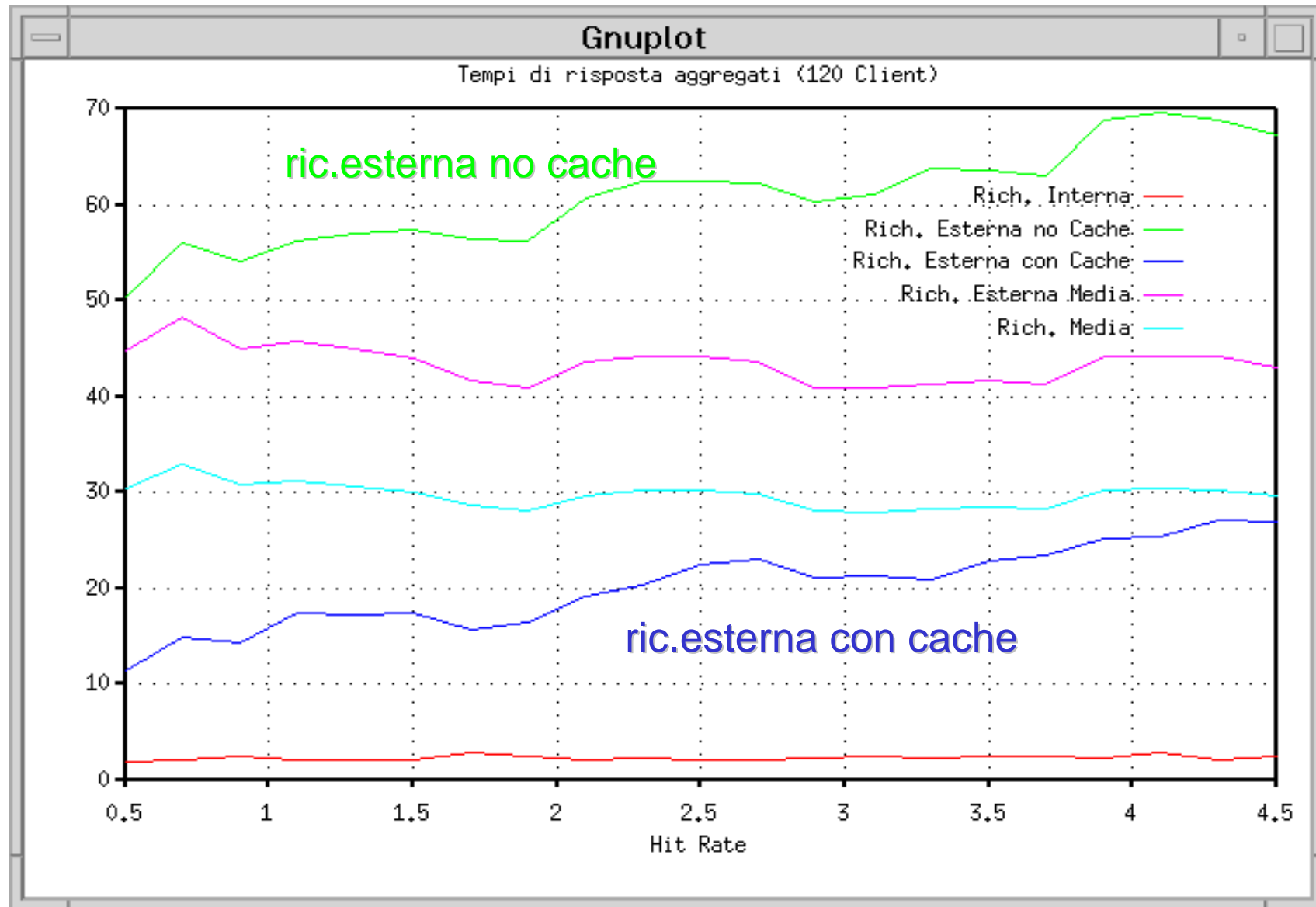
esp.1, tempi risposta aggregati: hit rate 80%



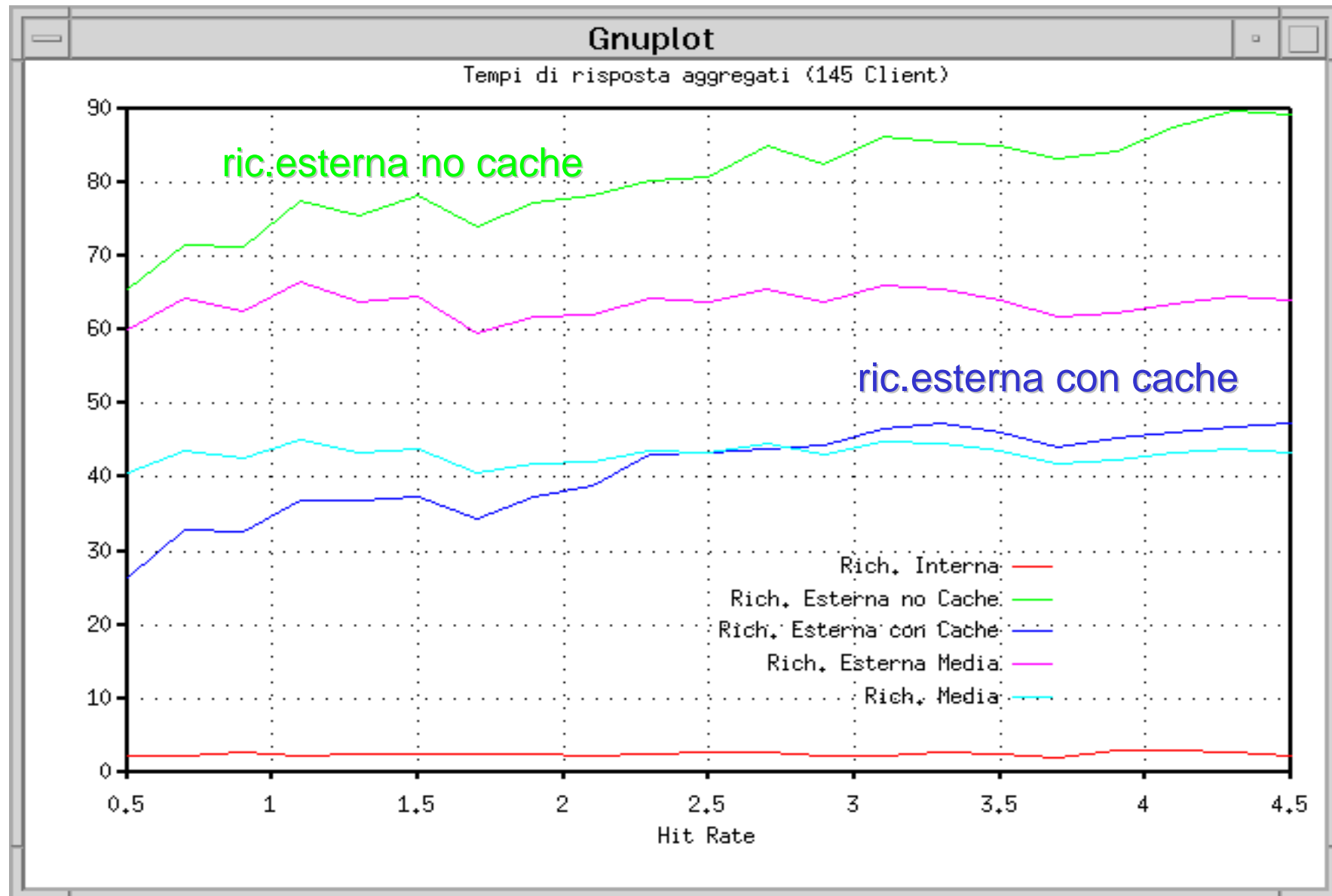
## esp.2, tempi risposta aggregati: 80 client



## esp.2, tempi risposta aggregati: 120 client



## esp.2, tempi risposta aggregati: 145 client

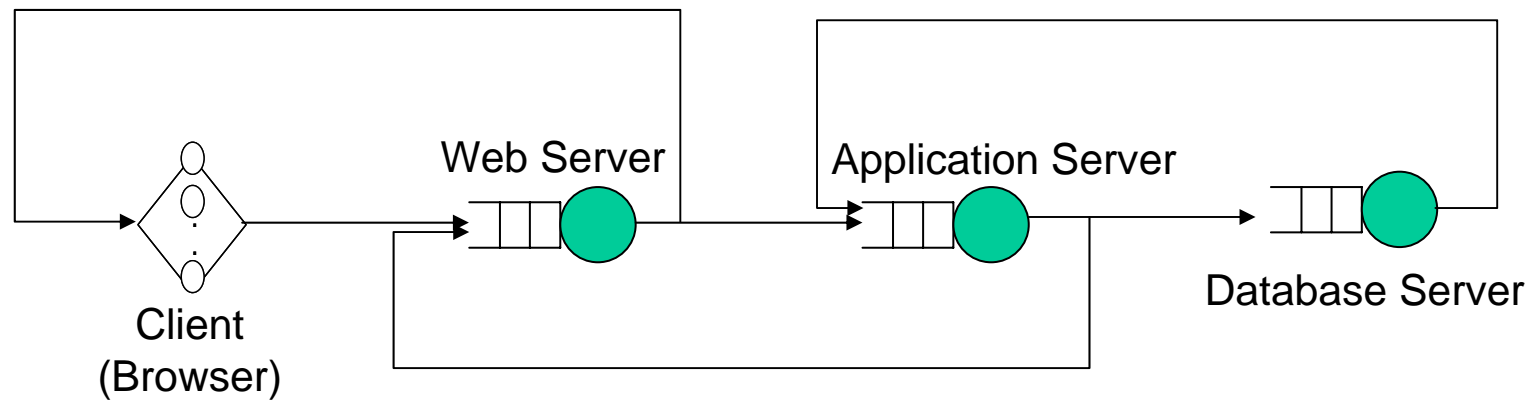


# modello di web application 3-tier

- struttura della intranet
- il workload (2 classi di utenti)
- prestazioni al variare del mix di utenti
- bottleneck

28/04/05 *R. Turrin*

# architettura dell'impianto



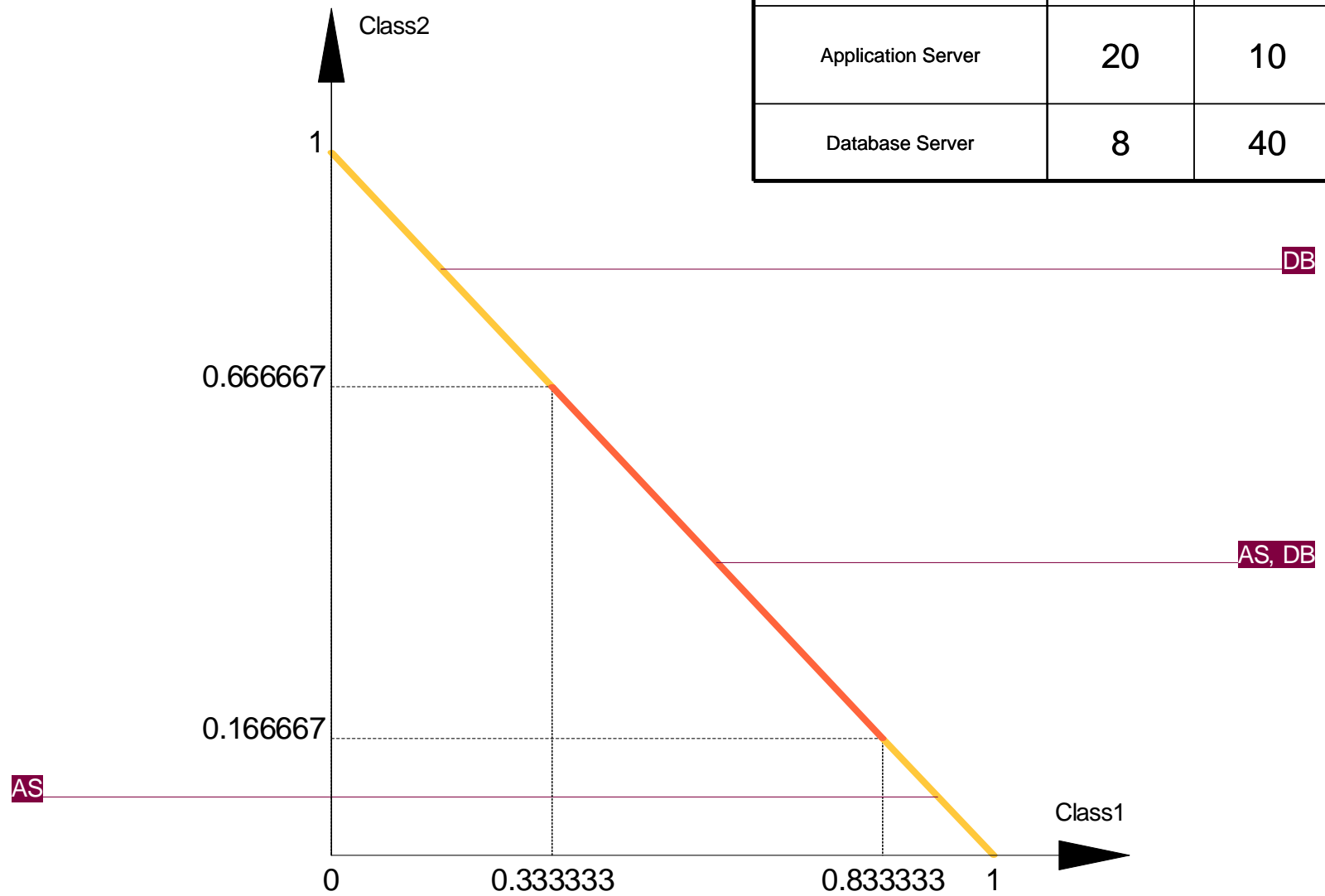
## workload - 1

Station	Service Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



# bottleneck switching

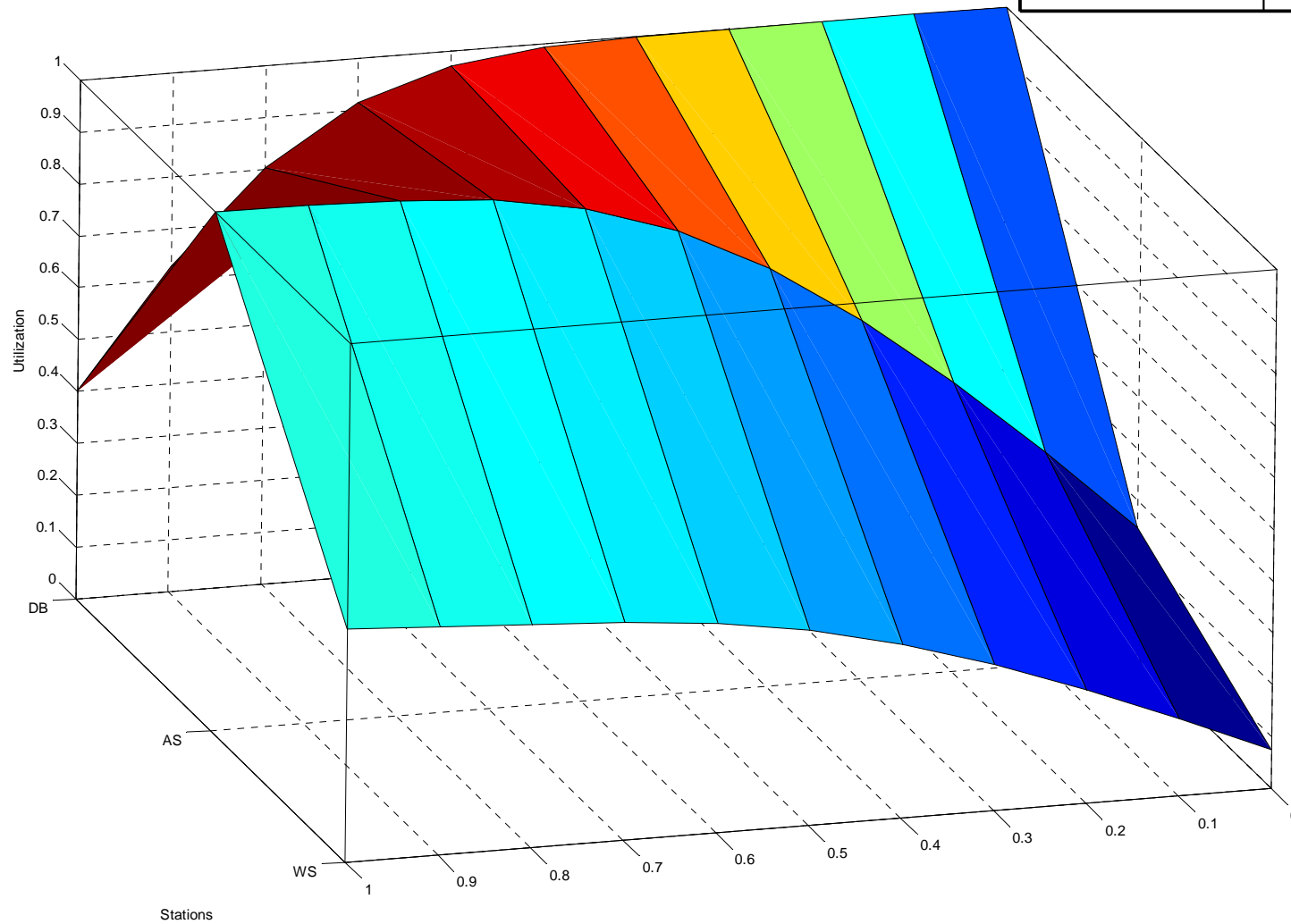
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



N=20

# Utilizzo

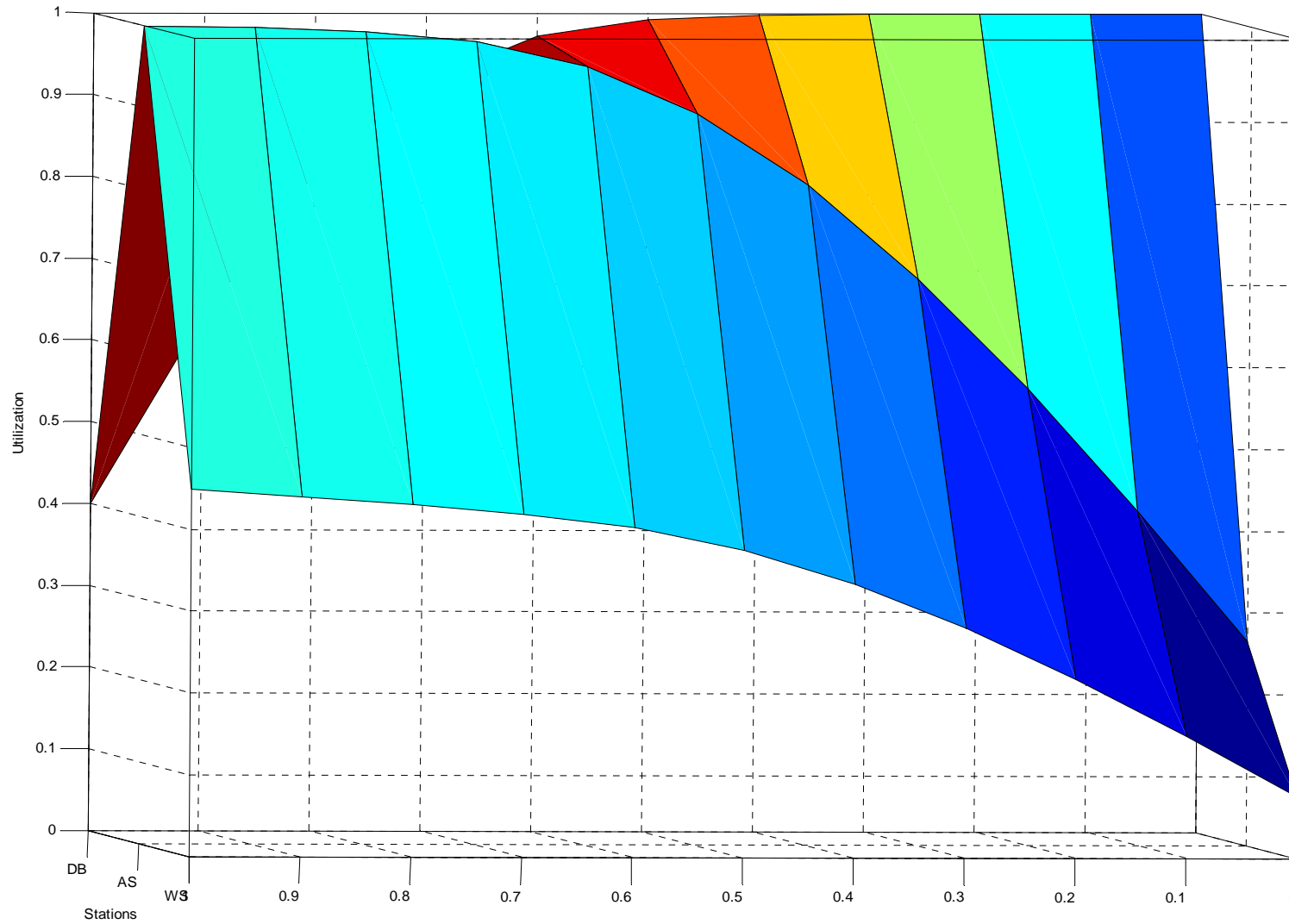
Station	System Demands (D) <sub>[sec]</sub>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



N=20

# Utilizzo

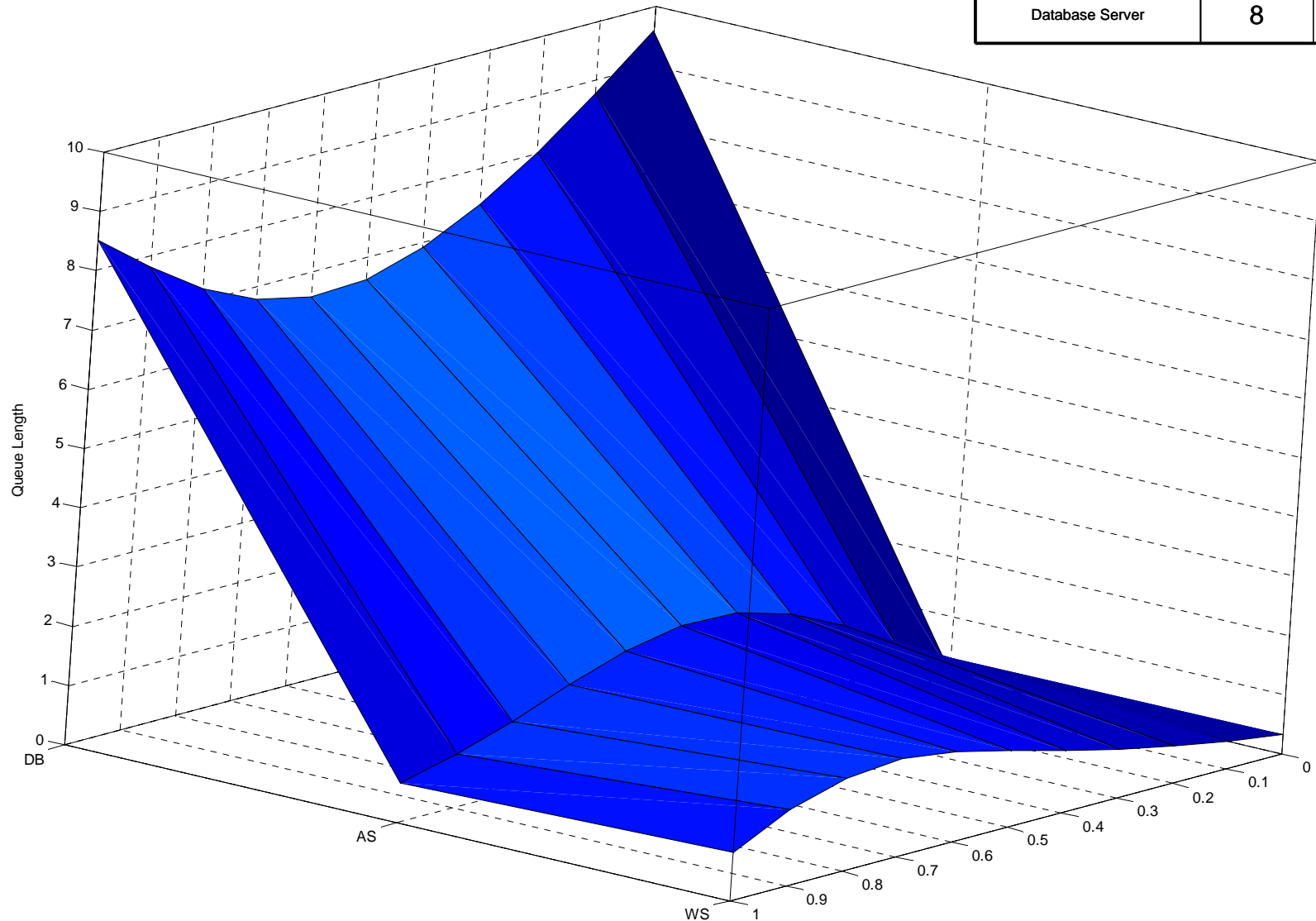
Station	System Demands (D) <sub>[sec]</sub>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



# lunghezza delle code

N=20

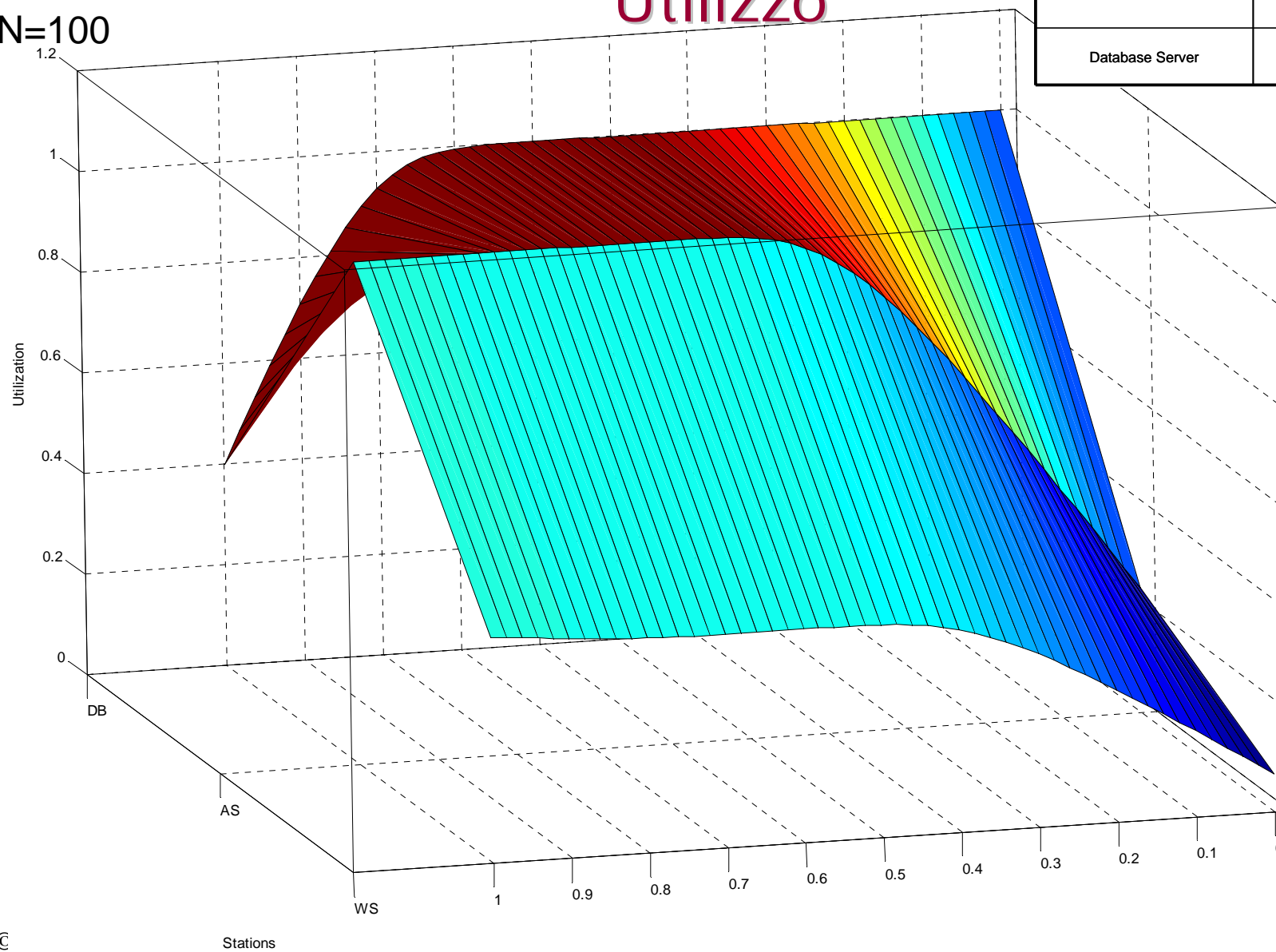
Station	System Demands (D) <sub>[sec]</sub>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



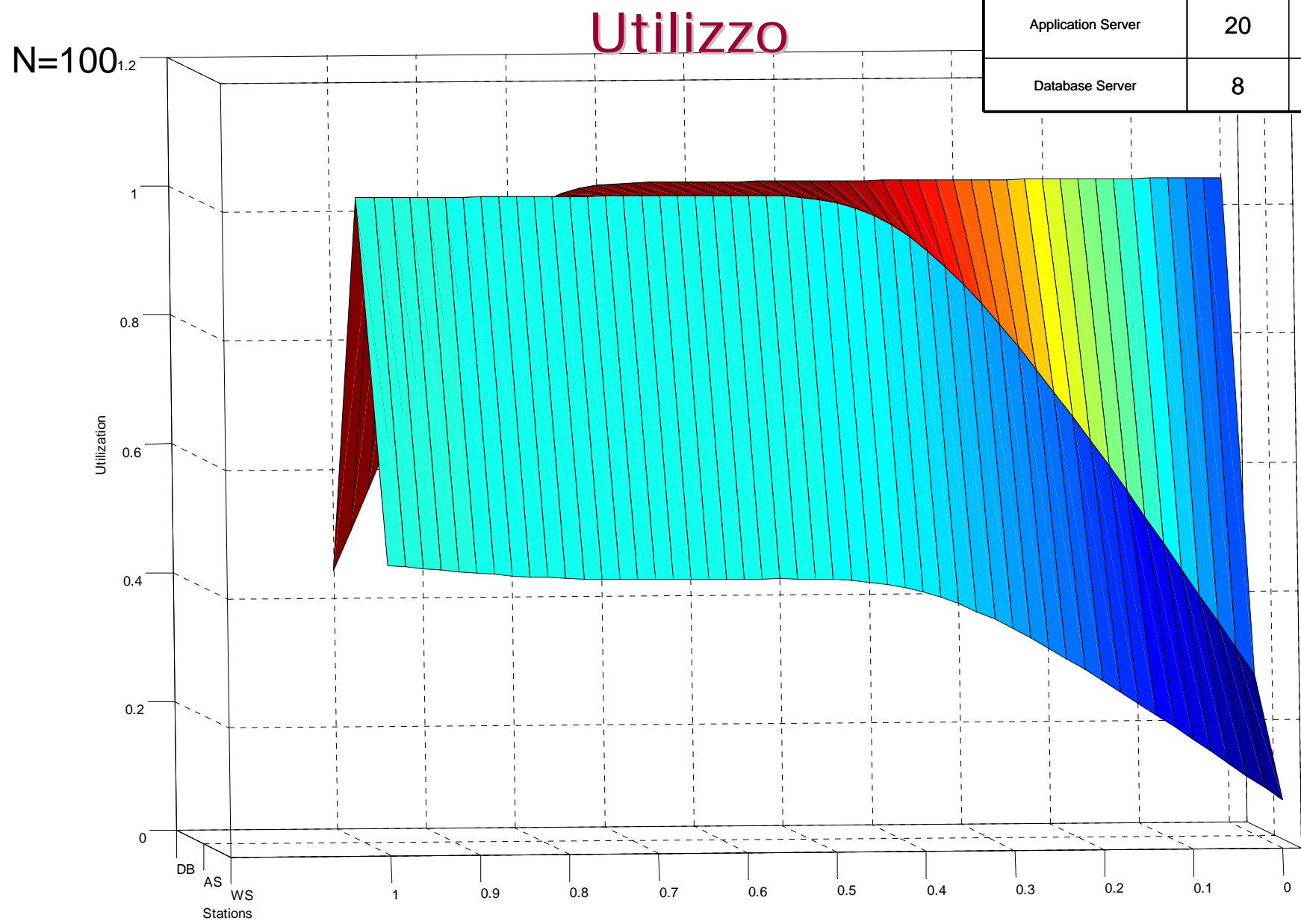
N=100

Utilizzo

Station	System Demands (D) <small>[sec]</small>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



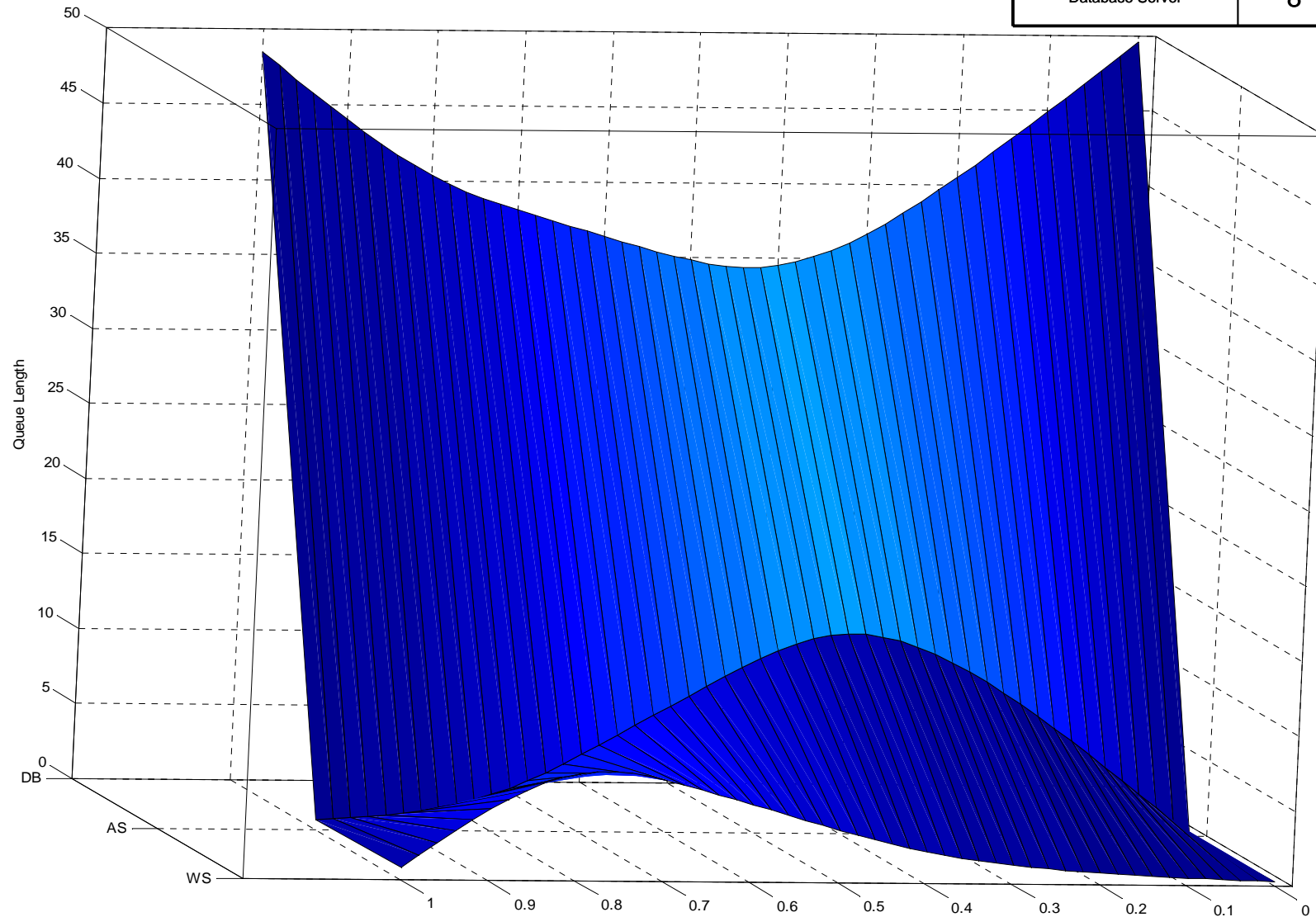
Station	System Demands (D) <sub>[sec]</sub>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



N=100

# Lunghezza di coda

Station	System Demands (D) <small>[sec]</small>	
	Class 1	Class 2
Web Server	9	3
Application Server	20	10
Database Server	8	40



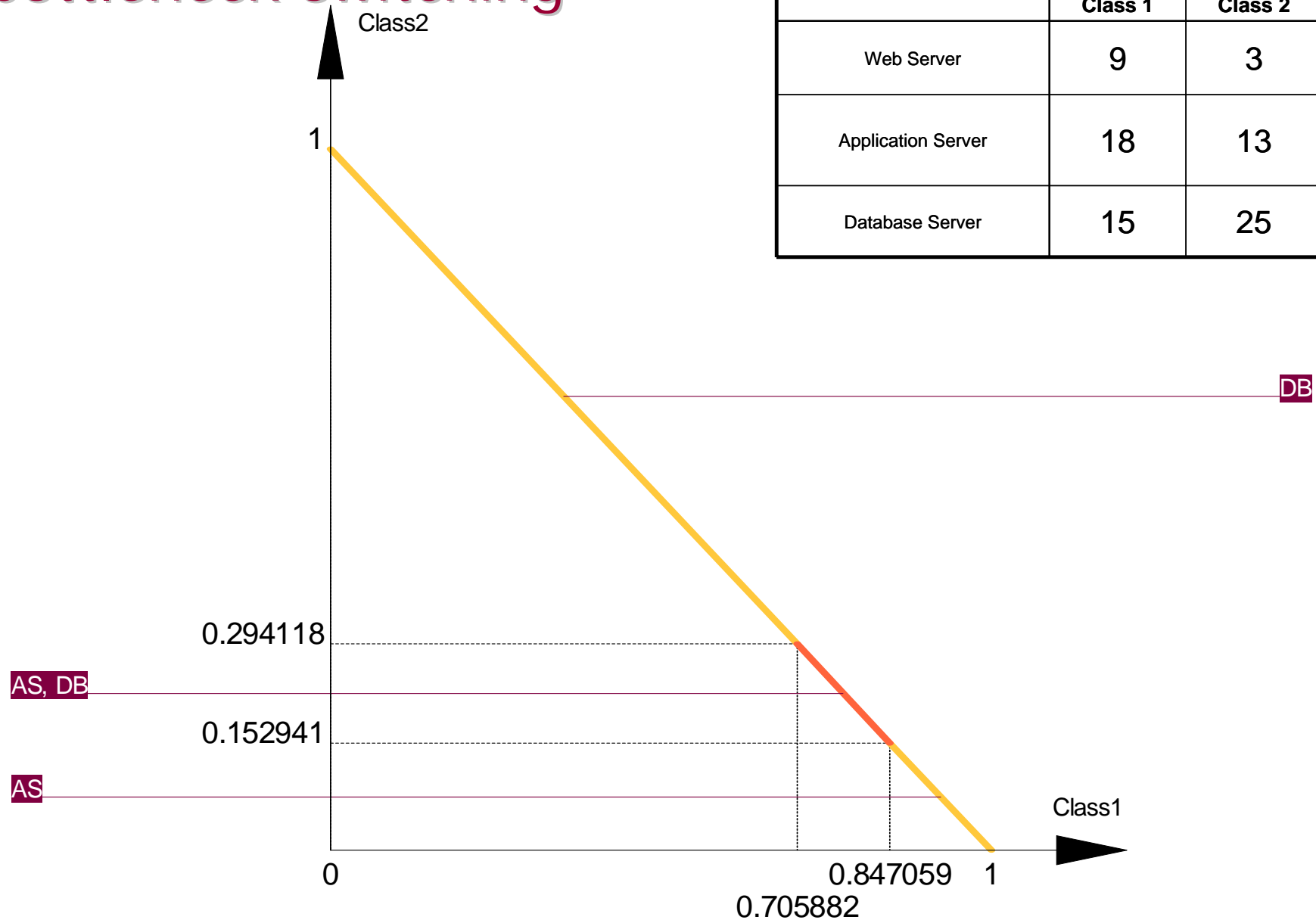
## Workload - 2

Station	Service Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



# bottleneck switching

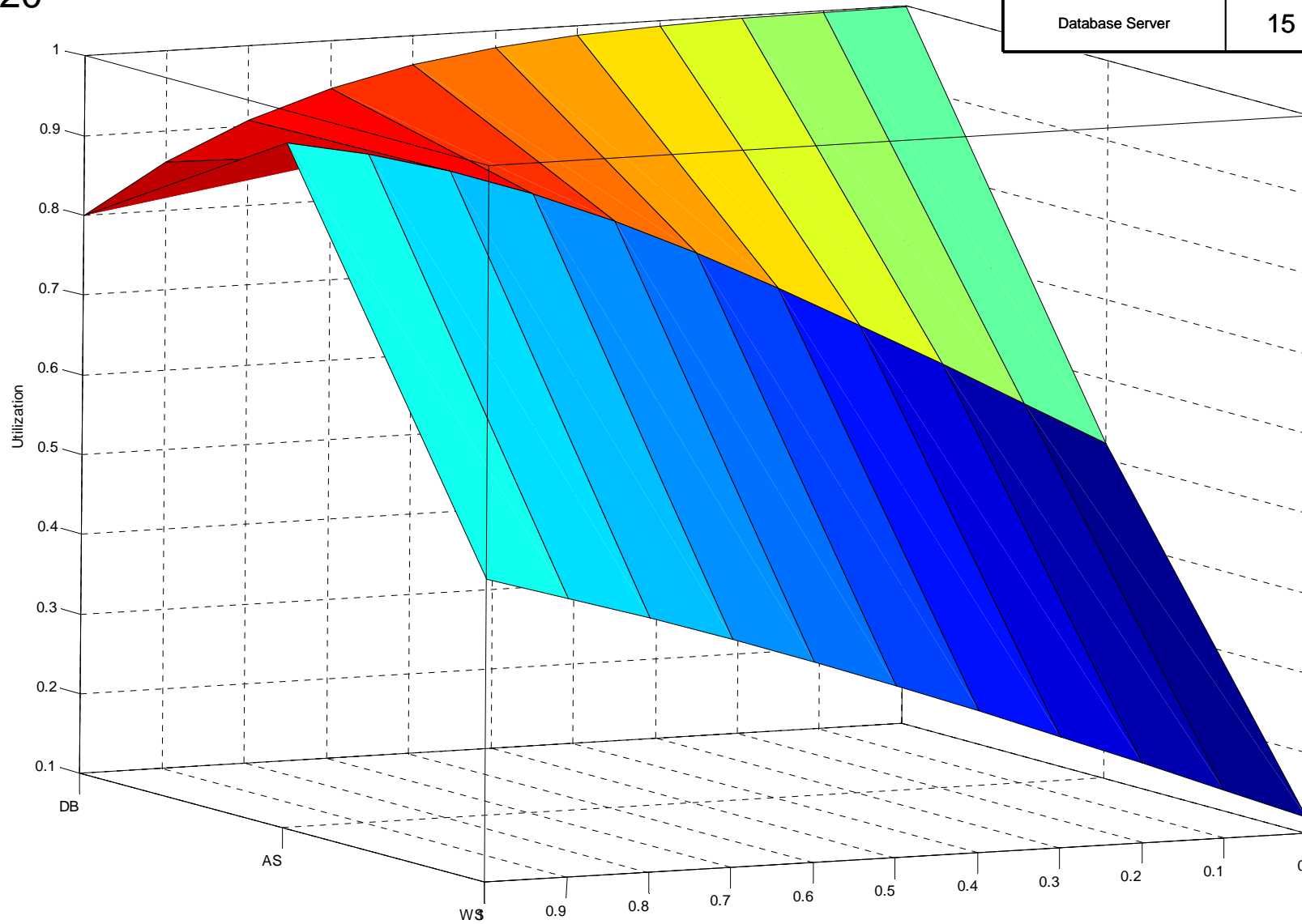
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



N=20

# Utilizzo

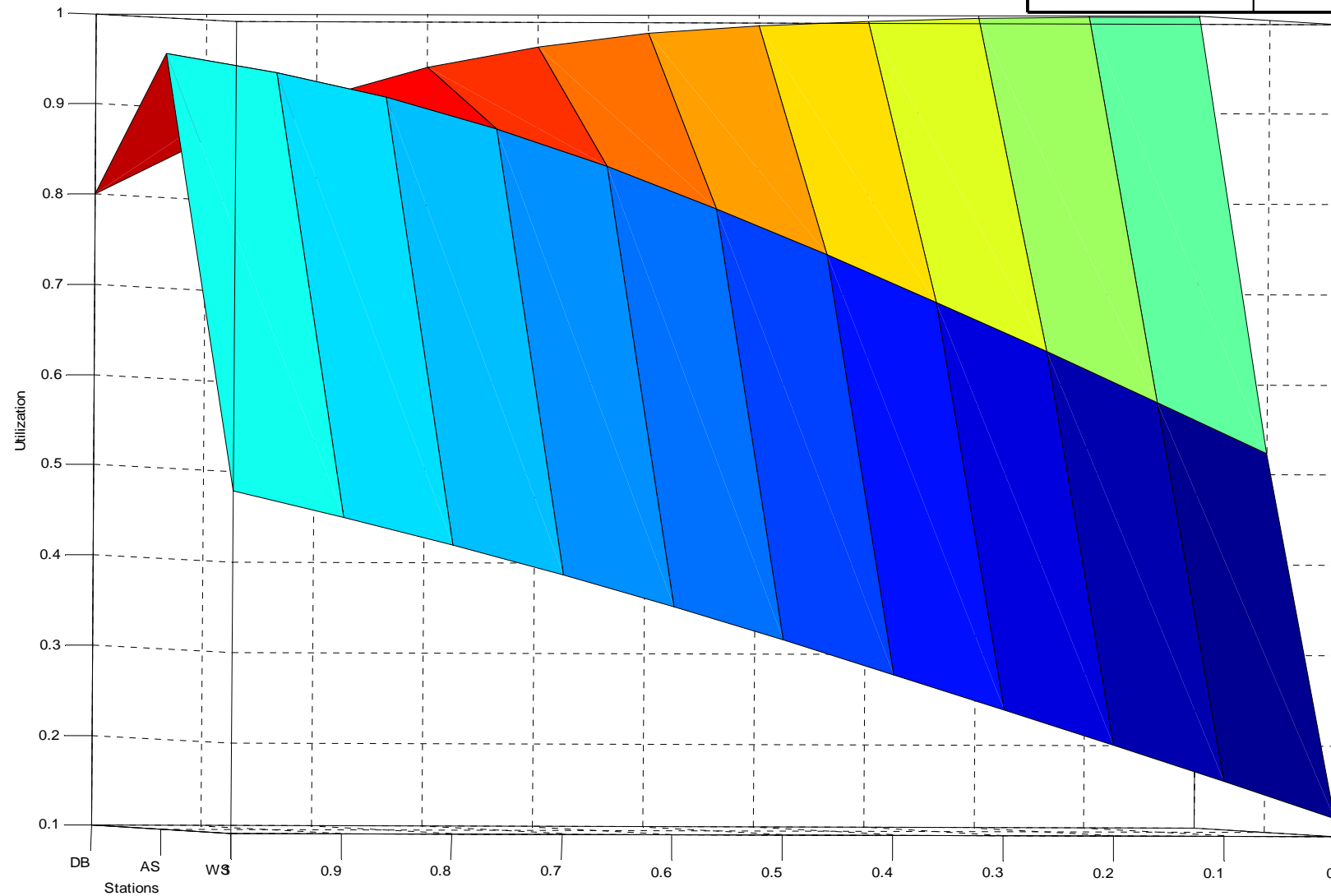
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



N=20

# Utilizzo

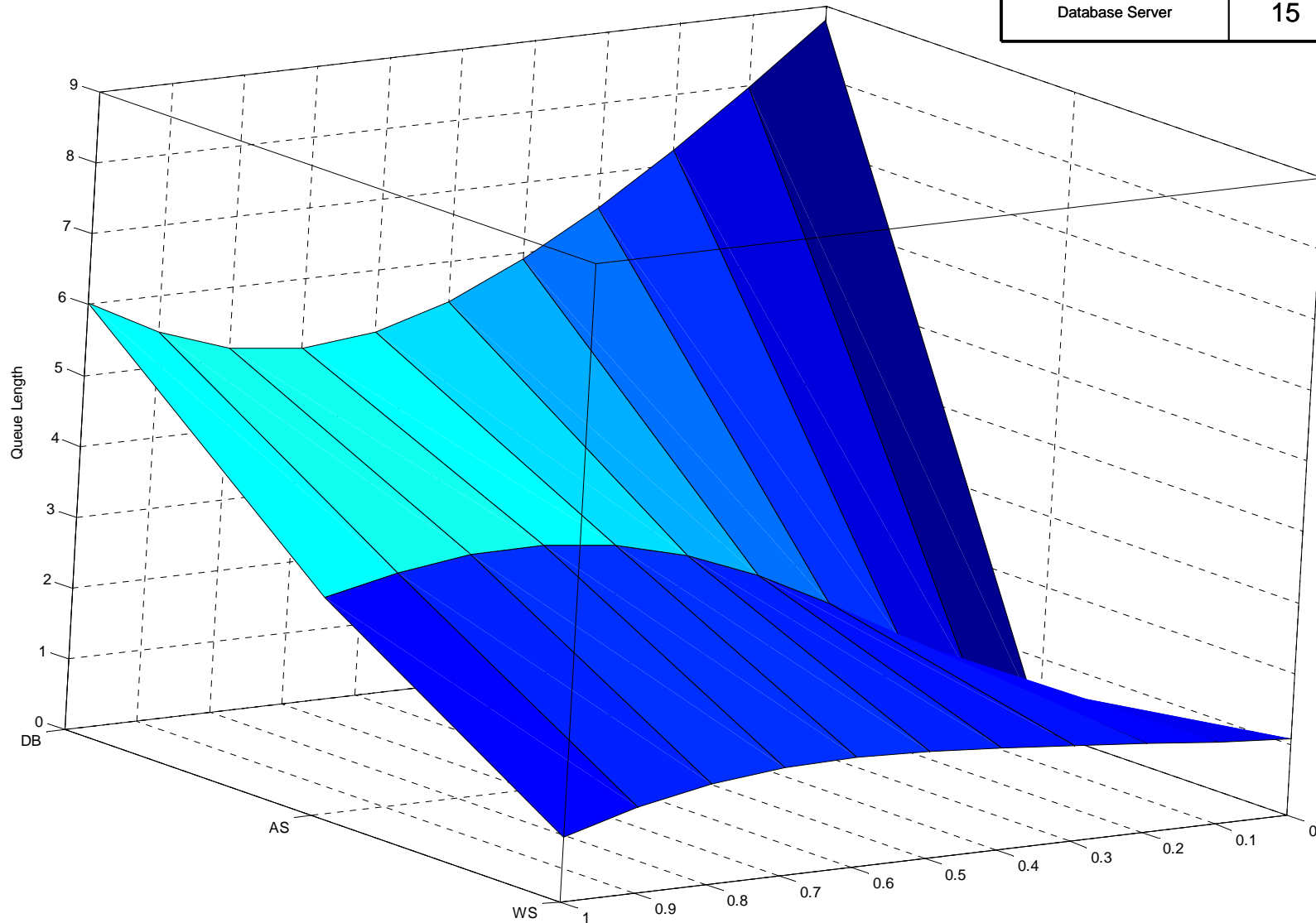
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



N=20

## Lunghezza di coda

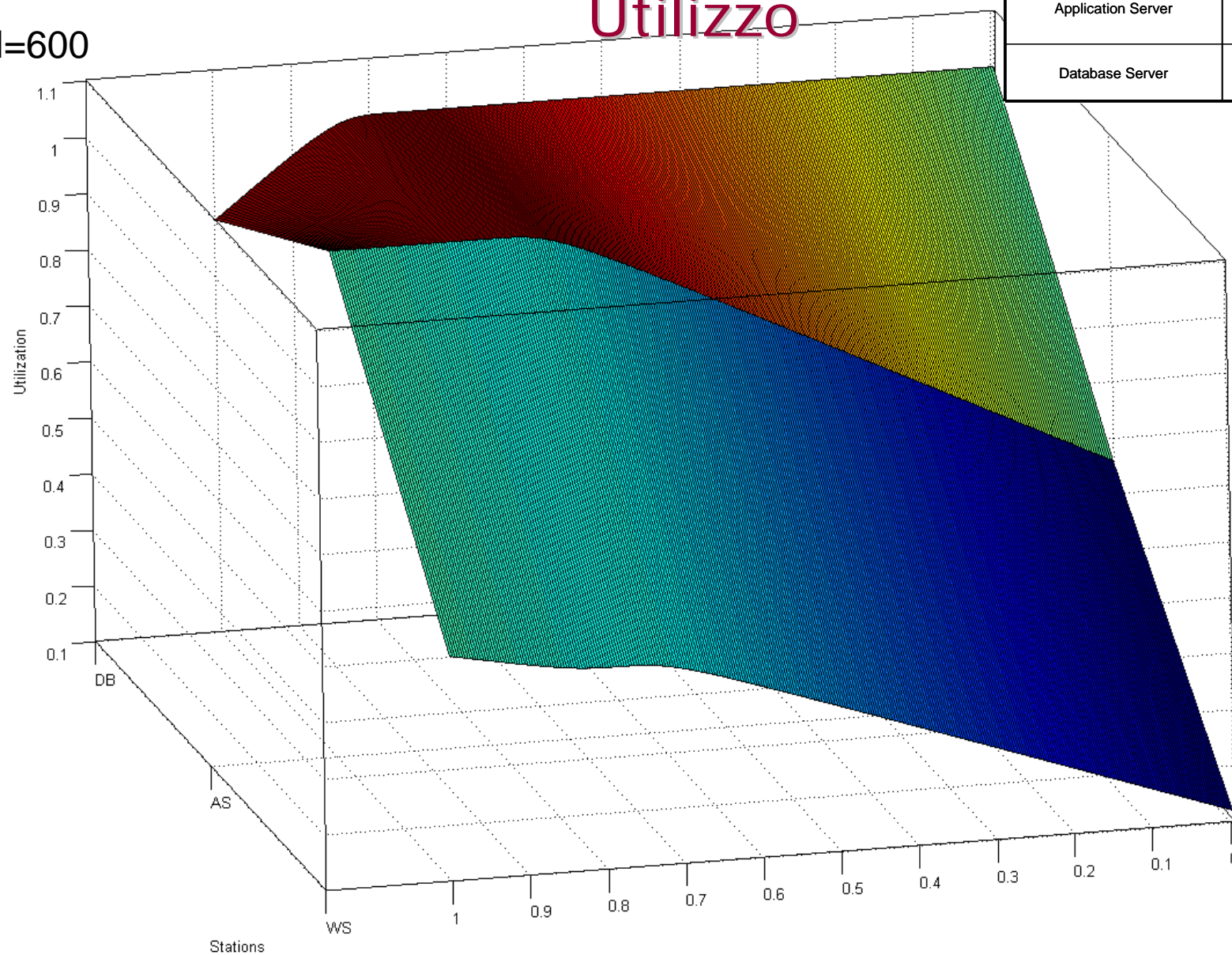
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



N=600

Utilizzo

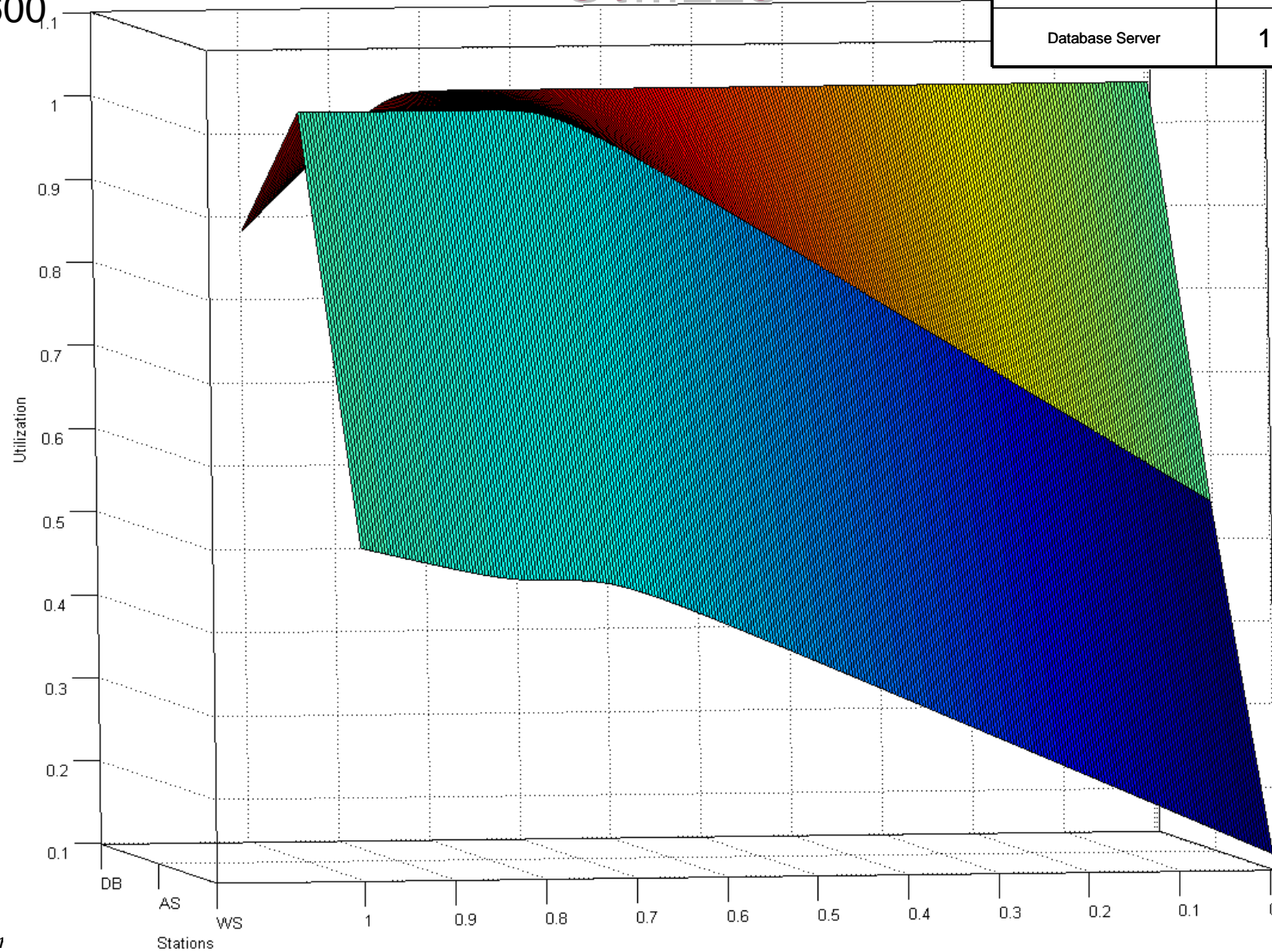
Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25



# Utilizzo

Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25

N=600





# lunghezza delle code

N=600

Station	System Demands (D) [sec]	
	Class 1	Class 2
Web Server	9	3
Application Server	18	13
Database Server	15	25

