

P_{hit}: fraction of requests that can be served from the proxy server's cache

Hit CPU time: CPU time, in seconds, needed to process the request at the cache proxy server in case of hit

Miss CPU time: CPU time, in seconds, needed to process a request at the cache proxy server in case of miss:

Miss:

- process the request
- ask the request to the WS
- store the document
- replace a document
- send the document to the client

Disk-time: disk time per kilobyte at the cache proxy server (in milliseconds)

$$D'_{LAN} = p_{hit} D_{LAN} + (1-p_{hit}) 2 D_{LAN}$$

$$D'^P_{router} = (1-p_{hit}) D_{router}$$

$$D'^P_{outL} = (1-p_{hit}) D_{outL}$$

$$D'^P_{internet} = (1-p_{hit}) D_{internet}$$

$$D'^L_{inL} = (1-p_{hit}) D_{inL}$$

$$D'^P_{CPU} = p_{hit} \text{ Hit CPU time} + (1-p_{hit}) \text{ Miss CPU time}$$

$$D'^P_{Disk} = \text{Disk time} \times \frac{\text{Document size}}{1000}$$

Esempio 10.5

stemi parametri di 10.4

+

CPU Time \times hit : 0.25 msec

CPU Time \times miss : 0.50 msec

Disk service time : 6 msec \times Kilo byte

P_{hit}	Throughput (rep/sec)	Response Time (sec)
0.20	0.364	37.82
0.30	0.416	32.72
0.40	0.483	27.61
0.50	0.581	22.50
0.0	0.292	48.0

Example 10.6

Same parameters as example 10.5 and 10.4

Link replaced by a T1 link (1.544 Mbps)

Cache hit ratio : 40%



new bottleneck: disk 45% util.

Response time 1.149 sec.

Throughput 3.347 rep/sec.