

CAPACITY PLANNING

15 July 105

Student's name and surname

Exercise 1

A web site is connected to a 2 Gbps Ethernet, which is connected to an ISP through a router. The router has a latency of 100μsec/packet and connects the Ethernet to the ISP through a line of 1 Gbps full duplex. The web site is realized with 8 computing node (a CPU with a main memory and a local disk, used only for the OS's purpose) and with a file system, realized with 4 CPU and with a RAID 2 with 8 (4+4) disks (the RAID is connected to the 4 CPU with a 1 Gigabit FDDI). The incoming rate is 160 requests/sec and each HTTP request for a file is of 400 bytes, the dimension for the requested file is 100 Kbytes, the load is equally distributed among the server nodes.

The CPU service demand for each request is 10 msec in case of hit (file stored in the local main memory) and 20 msec in case of miss, the CPU-file-system service demand for each request is 5 msec, instead the service time of the each single disk for 10 Kbytes is 10 msec.

Compute the average response time in case all the components are fault-free, hypothesizing that the probability of hit is 80%.

Evaluate the SAFETY of the system, hypothesizing the following parameters:

- MTTF (MTTR) for the CPU with its main memory: 6 years (1 week)
- MTTF (MTTR) for the Ethernet and FDDI: 20 years (2 weeks)
- MTTF (MTTR) for each disk: 3 year (3 weeks)
- MTTF (MTTR) for the router: 20 year (3 weeks)
- Fault detection coverage factor equal for all the components and equal to 99%
- Fault recovery coverage factor equal for all the components and equal to 98%

Finally show the methodology to evaluate the average response time given the presence of faults and reparations in case the fault coverage factors are equal to 1.

Exercise 2

A file server receives requests from 8 workstations (users), that have a thinking time equal to 10 second. Every request needs 50 ms of CPU and 10 I/O requests to a disk, the time for

each I/O is 5 ms. The server can manage at most 4 users at the same time. Evaluate the average service time, the average throughput and the percentage of refused requests. In case of refuse the client send another request with the same thinking time of 10 sec.