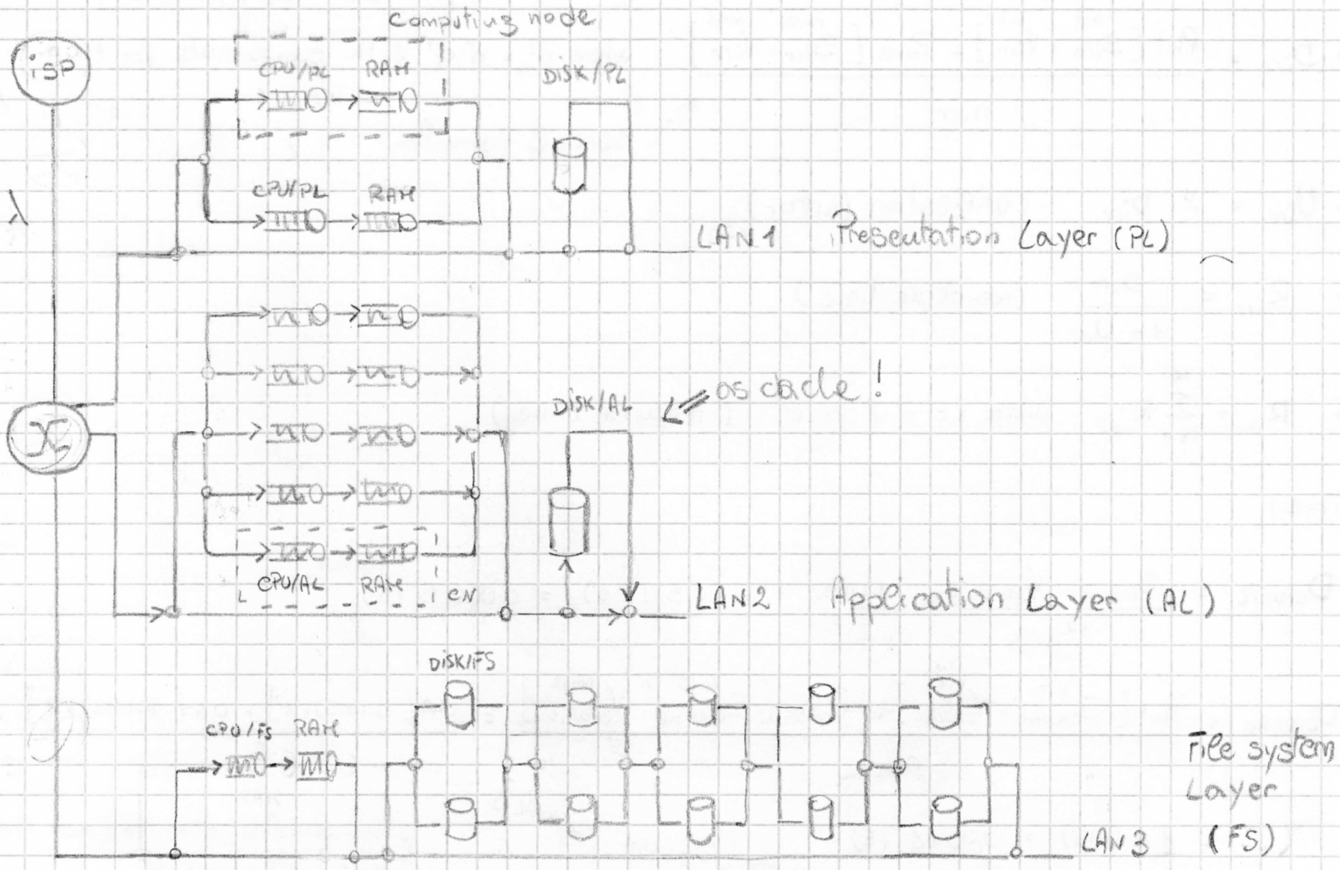
Exam 02.07.2013



Data:

- $\lambda = 100 \; (\frac{req}{sec})$;
- $B_{isp} = 1$ Gbps ; (full duplex ⇒ ⇄)
- $L = 100$ (μsec/pkt) { router latency }
- $F = 100$ Kbytes = ~~100 · 1024~~ 10^3 bytes = 100'000 bytes { file dimension }
- $R = 300$ bytes { Http request }
- $S_{CPU/PL} = 1$ (msec) ;
- $P_{hit} = 60\%$; $P_{miss} = 40\%$

  $S^{hit}_{CPU/AL} = 5$ (ms) ; $S^{miss}_{CPU/AL} = 10$ (ms)

  $S_{DISK/AL} = 5$ (msec · 10 Kbytes) ⇒ To serve the 100 kbytes (file dimension) ⇒ $V_{DISK/AL} = 10$
- $S_{CPU/FS} = 10$ (msec)

  $S_{DISK/FS} = 3$ (msec · 10 Kbytes) ⇒ To serve the 100 Kbytes (file dimension) ⇒ $V_{DISK/FS} = 10$

- $B_{LAN1} = 1$ GBs
- $B_{LAN2} = 1$ GBs
- $B_{LAN3} = 0,5$ GBs

Being a multiclass open QNs, I use these formulas: $\begin{cases} i = \text{queue identification} \\ r = \text{layer identification} \end{cases}$

- $D_{i/r} = \dfrac{P_{hit}\left[S_{i/r}^{hit} \cdot V_{i/r}^{hit}\right] + P_{miss}\left[S_{i/r}^{miss} \cdot V_{i/r}^{miss}\right]}{n_{i/r}}$ where $n_{i/r}$ = n° of i-th components for the r-th layer

general formula

- $U_{i/r} = \lambda \cdot D_{i/r}$ (utilization factor)

- $R'_{i/r} = \dfrac{D_{i/r}}{1 - U_{i/r}}$ (residence time)

- $R_r = \sum_{i=1}^{K} R_{i/r}$ where k = n° of queues (response time)

Thus:

- $D_{CPU/PL} = \dfrac{S_{CPU/PL} \cdot V_{CPU/PL}}{n_{CPU/PL}} = \dfrac{1\,(ms) \cdot 1}{2} = 0{,}5\,(ms) = 0{,}0005\,(s)$

- $D_{CPU/AC} = \dfrac{P_{hit}\left[S_{CPU/AC}^{hit} \cdot V_{CPU/AC}^{hit}\right] + P_{miss}\left[S_{CPU/AC}^{miss} \cdot V_{CPU/AC}^{miss}\right]}{n_{CPU/AC}} = \dfrac{0{,}6\left[5\,(ms) \cdot 1\right] + 0{,}4\left[10\,(ms) \cdot 1\right]}{5} =$

$= 1{,}4\,(ms) = 0{,}0014\,(s)$

is indipendent from the fact that I have a hit or a miss.

$D_{DISK/AC} = \dfrac{P_{hit}\left[S_{DISK/AC} \cdot V_{DISK/AC}\right] + \overset{P_{miss}\left[S_{DISK/AC} \cdot V_{DISK/AC}\right]}{\cancel{\emptyset}}}{n_{DISK/AC}} = \dfrac{\overset{0{,}6+0{,}4=1}{\cancel{0{,}6}}\left[5\,(ms) \cdot 10\right]}{1} = \underset{5}{\cancel{30}}\,(ms) = 0{,}0\underset{5}{\cancel{3}}\,(s)$

- $D_{CPU/FS} = \dfrac{\emptyset + P_{miss}\left[S_{CPU/FS} \cdot V_{CPU/FS}\right]}{n_{CPU/FS}} = \dfrac{0{,}4\left[10\,(ms) \cdot 1\right]}{1} = 4\,(ms) = 0{,}004\,(s)$

$D_{DISK/FS} = \dfrac{\emptyset + P_{miss}\left[S_{DISK/FS} \cdot V_{DISK/FS}\right]}{n_{DISK/FS}} = \dfrac{0{,}4\left[3\,(ms) \cdot 10\right]}{\cancel{5}\ 10} = \overset{1,2}{\cancel{2{,}4}}\,(ms) = 0{,}00\overset{12}{\cancel{24}}\,(s)$

- $U_{CPU/PL} = \lambda \cdot D_{CPU/PL} = 100\,\left(\frac{req}{sec}\right) \cdot 0{,}0005\,(sec) = 0{,}05$
- $U_{CPU/AC} = \lambda \cdot D_{CPU/AC} = 100 \cdot 0{,}0014 = 0{,}14$

$U_{DISK/AC} = \lambda \cdot D_{DISK/AC} = 100 \cdot 0{,}0\overset{5}{\cancel{3}} = \cancel{3}5$?

- $U_{CPU/FS} = \lambda \cdot D_{CPU/FS} = 100 \cdot 0{,}004 = 0{,}4$

$U_{DISK/FS} = \lambda \cdot D_{DISK/FS} = 100 \cdot 0{,}00\underset{12}{\cancel{24}} = 0{,}\cancel{24}\,12$

- $R'_{CPU/PL} = \dfrac{D_{CPU/PL}}{1 - U_{CPU/PL}} = \dfrac{0{,}5\,(ms)}{0{,}95} = 0{,}53\,(ms)$
- $R'_{CPU/AC} = \dfrac{D_{CPU/AC}}{1 - U_{CPU/AC}} =$

$R'_{DISK/AC} = \dfrac{D_{DISK/AC}}{1 - U_{DISK/AC}} =$

- $R'_{CPU/FS} = \dfrac{D_{CPU/FS}}{1 - U_{CPU/FS}} =$

$R'_{DISK/FS} = \dfrac{D_{DISK/FS}}{1 - U_{DISK/FS}} =$

Regarding the networks: Hypothesis: TCP doesn't know the local MTU of the LAN.

$\boxed{\text{REQUEST}}$

$$\#segment = \left\lceil \frac{R}{MTU_{TCP}} \right\rceil = \left\lceil \frac{300}{65'515} \right\rceil = 1;$$

$$\#datagram = \left\lceil \frac{R + \#seg(TCP_{ovhd})}{MSS} \right\rceil = \left\lceil \frac{300 + (1 \cdot 20)}{1480} \right\rceil = \#frame \quad \{MSS = MTU_{ETH} - IP_{ovhd} = 1500 - 20\}$$

$$Overhead = \#segment(TCP_{ovhd}) + \#datagram(IP_{ovhd} + Eth_{ovhd}) = (1 \cdot 20) + 1(20+18) = 58 \text{ bytes}$$

$$S^{req}_{ISP} = \frac{8(R + Overhead)}{B_{ISP}} = \frac{2864 \text{ (bit)}}{1 \cdot 10^9 \left(\frac{bit}{sec}\right)} = 2864 \cdot 10^{-9} \text{ (sec)} = 0{,}0029 \text{ (ms)}$$

$$S^{req}_{rout} = \#datagram \cdot L = 1 \text{ (pkt)} \cdot 100 \left(\frac{\mu sec}{pkt}\right) = 100 \text{ } (\mu sec) = 0{,}1 \text{ (ms)}$$

$$S^{req}_{LAN1} = \frac{R + Overhead}{B_{LAN1}} = \frac{358 \text{ (byte)}}{1 \cdot 10^9 \left(\frac{byte}{sec}\right)} = 358 \cdot 10^{-9} \text{ (sec)} = 0{,}00036 \text{ (ms)} = S^{req}_{LAN2}$$

$$S^{req}_{LAN3} = P_{miss}\left[\frac{R + Overhead}{B_{LAN3}}\right] = 0{,}4\left[\frac{358}{0{,}5 \cdot 10^9}\right] = [716 \cdot 10^{-9} \text{ (s)}] \cdot 0{,}4 = 286{,}4 \cdot 10^{-9} \text{ (s)} = 0{,}0002864 \text{ (ms)}$$

$\boxed{\text{REPLY}}$

$$\#segment = \left\lceil \frac{F}{MTU_{TCP}} \right\rceil = \left\lceil \frac{100'000}{65'515} \right\rceil = 2;$$

$$\#datagram = \#frame = \left\lceil \frac{F + \#segment(TCP_{ovhd})}{MSS} \right\rceil = \left\lceil \frac{100'000 + (2 \cdot 20)}{1480} \right\rceil = 68$$

$$Overhead = \#segment(TCP_{ovhd}) + \#datagram(IP_{ovhd} + ETh_{ovhd}) = (2 \cdot 20) + 68(20+18) = 2624 \text{ (byte)}$$

$$S^{rep}_{LAN3} = \left[\frac{F + Overhead}{B_{LAN3}}\right] \cdot 40\% = \left[\frac{102'624}{0{,}5 \cdot 10^9} \text{ (sec)}\right] \cdot 0{,}4 = [205'248 \cdot 10^{-9} \text{ (sec)}] \cdot 0{,}4 = [0{,}205 \text{ (ms)}] \cdot 0{,}4 = 0{,}082 \text{ (msec)}$$

$$S^{rep}_{LAN2} = \left[\frac{F + Overhead}{B_{LAN2}}\right] \cdot 60\% = \left[\frac{102'624}{1 \cdot 10^9}\right] 0{,}6 = 61'574{,}4 \cdot 10^{-9} \text{ (s)} = 0{,}0615744 \text{ (ms)}$$

=

$$S^{rep}_{LAN1} = \frac{F + Overhead}{B_{LAN1}} = \frac{102'624}{1 \cdot 10^9} = 102'624 \cdot 10^{-9} \text{ (sec)} = 0{,}10 \text{ (msec)}.$$

$$S^{rep}_{rout} = \#datagram \cdot L = 68 \cdot 100 = 6'800 \text{ } (\mu sec) = 6{,}8 \text{ (ms)}$$

$$S^{rep}_{ISP} = \frac{(F + Overhead)\,8}{B_{ISP}} = \frac{820'992 \text{ (bit)}}{1 \cdot 10^9 \left(\frac{bit}{sec}\right)} = 820'992 \cdot 10^{-9} \text{ (sec)} = 0{,}82 \text{ (ms)}$$

HANDSHAKING

$\# segment = \# datagram = \# frame = 6$

$Overhead = \# frame\,(TCP_{ovhd} + IP_{ovhd} + Eth_{ovhd}) = 6\,(20+20+18) = 348\ (byte)$

$$S^{hs}_{isp} = \frac{(\emptyset + Overhead)\,8}{B_{isp}} = \frac{2784}{1\cdot 10^9}\ (sec) = 2784\cdot 10^{-9}\ (sec) = 0{,}0028\ (msec)$$

$$S^{hs}_{rout} = \# datagram \cdot L = 6\cdot 100 = 600\ (\mu sec) = 0{,}6\ (ms)$$

$$S^{hs}_{LAN1} = \frac{0 + Overhead}{B_{LAN1}} = \frac{348}{1\cdot 10^9}\ (sec) = 348\cdot 10^{-9}\ (sec) = 0{,}00035\ (ms)$$

S^{hs}_{LAN2} ?

S^{hs}_{LAN3} ?

Then:

$U_i = \lambda \cdot D_i$

$R'_i = \frac{D_i}{1-U_i}$

$$D_{LAN1} = (S^{hs}_{LAN1} + S^{req}_{LAN1} + S^{resp}_{LAN1})\cdot 1 =$$

$$D_{isp} = (S^{hs}_{isp} + S^{req}_{isp} + S^{resp}_{isp})\cdot 1 =$$

$$D_{rout} = (S^{hs}_{rout} + S^{req}_{rout} + S^{resp}_{rout})\cdot 1 =$$

$$D_{LAN2} = S^{req}_{LAN2} + \{[S^{resp}_{LAN2}\cdot 1]\,0{,}6 + 0{,}4\,[S^{resp}_{LAN}\cdot 2]\} + S^{hs}_{LAN2}$$

$$D_{LAN3} = (S^{req}_{LAN3} + S^{resp}_{LAN3} + S^{hs}_{LAN3})\,0{,}4$$

Finally:

1) $R_{PL} = R'_{CPU/PL}$

$R_{AL} = R'_{CPU/AL} + R'_{DISK/AL}$

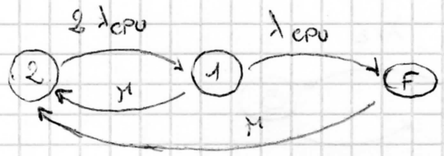
$R_{FS} = R'_{CPU/FS} + R'_{DISK/FS}$

$R_{tot} = R_{PL} + R_{AL} + R_{FS} + R_{isp} + R_{rout} + R_{LAN1} + R_{LAN2} + R_{LAN3}$ (average response time)

The bottleneck is the maximum of the response time

2) $A_i = \frac{MTTF}{MTTF + MTTR}$ where $MTTF = \frac{1}{\lambda_i}$ & $MTTR = \frac{1}{\mu}$

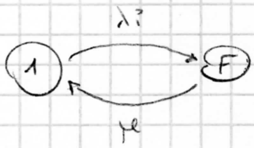
- CPU/PL



$$\begin{cases} P_2'(t) = \emptyset = -P_2(2\lambda_{CPU}) + \mu(P_1 + P_F) \\ P_1'(t) = \emptyset = -P_1(\mu + \lambda_{CPU}) + P_2(2\lambda_{CPU}) \\ P_F'(t) = \emptyset = -P_F(\mu) + P_1(\lambda_{CPU}) \\ P_2(t) + P_1(t) + P_F(t) = 1 \; ; \quad P_2(\emptyset) = 1 \end{cases}$$

$A_{CPU/PL}(t) = 1 - P_F(t)$

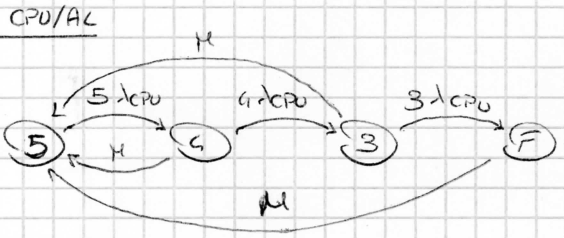
- DISK/PL, ROUTER, LAN (1/2/3), DISK/AC, CPU/FS



$A_i = 1 - P_F(t)$ where $i = DISK^{/PL}, NET, ROU$

- $A_{PL}(t) = A_{CPU/PL}(t) \cdot A_{DISK/PL}(t) \cdot A_{LAN(1)}(t)$

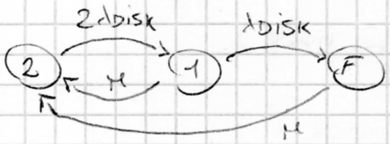
- CPU/AL



$A_{CPU/AL} = 1 - P_F(t)$

- $A_{AC}(t) = A_{CPU/AC}(t) \cdot A_{DISK/AC}(t) \cdot A_{LAN2}(t)$

- RAID (single pair)



$A_{pair} = 1 - P_F(t) \Rightarrow A_{RAID} = [A_{pair}]^5$

- $A_{FS}(t) = A_{CPU/FS}(t) \cdot A_{RAID}(t) \cdot A_{LAN3}(t)$

Finally: $A_{system}(t) = A_{router}(t) \cdot A_{PL}(t) \cdot A_{AC}(t) \cdot A_{FS}(t)$

3) About the presentation layer, we have a requests queue of dimension 5.

I hypothesize that the requests load is equally distributes on the CPU/PCS (that are 2)

So, the STD (State transaction Diagram) is:

$$\begin{tikzcd}
0 \arrow[r, bend left, "\lambda/2"] & 1 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda/2"] & 2 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda/2"] & 3 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda/2"] & 4 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda/2"] & 5 \arrow[l, bend left, "\mu"]
\end{tikzcd}$$

Where $\mu = \frac{1}{S_{CPU/PC}} = \frac{1}{0{,}001 \text{ (sec)}} = 1'000 \left(\frac{\text{req}}{\text{sec}}\right)$

Flow-in = Flow-out principle:

$$\begin{cases} P_0 \cdot \frac{\lambda}{2} = P_1 \cdot \mu \\ P_1 \cdot \frac{\lambda}{2} = P_2 \cdot \mu \\ P_2 \cdot \frac{\lambda}{2} = P_3 \cdot \mu \\ P_3 \cdot \frac{\lambda}{2} = P_4 \cdot \mu \\ P_4 \cdot \frac{\lambda}{2} = P_5 \cdot \mu \\ \sum_{i=0}^{5} p_i = 1 \end{cases}$$

About the single node:

$X = \sum_{i=1}^{5} p_i \cdot \mu$; (average throughput of a single node)

$N = \sum_{i=0}^{5} p_i \cdot i$; (average n° of requests in a single node).

$R = \frac{N}{X}$; (average response time for a single node).

About both nodes:

$\bar{R} = R$; (average response time for the 2 computing nodes)

Fraction of requests loss $= \frac{\lambda - \bar{X}}{\lambda}$ where $\bar{X} = 2 \cdot X$ (average throughput for the 2 computing nodes)