

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Soley Masats, Jordi
Vidal Gonzalez, Albert**

Escenari 020

Data: 29/09/2023

1.-Descripció bàsica

Copy & Paste del full de càlcul "Resum"

AQUEST APARTAT NOMÉS PEL LLIURAMENT FINAL

IMPORTANT: Pels apartats següents no poseu només el número, justifiqueu el perquè. És més important el perquè que el número en si.

2.-Anàlisi de necessitats

2.1- Número de GB a emmagatzemar (en cru).

Ens arriben 100.000.000 de vídeos a emmagatzemar amb una mida de 5MB per vídeo.

$$100.000.000 \text{ videos} * \frac{5 \text{ MB}}{\text{video}} * \frac{1 \text{ GB}}{1000 \text{ MB}} = 500.000 \text{ GB}$$

Per tant, tindríem 500.000 GB a emmagatzemar.

Suposem que els 100MB de dades (per cada vídeo analitzat) que s'envien a un altre servidor no les hem de tenir en compte ja que es un servidor extern.

2.2- Velocitat requerida del sistema de disc (IOPS).

Cada dia fem:

- ➔ 50.000 input (vídeos que arriben)
- ➔ 1.000.000 output (vídeos que analitzem)

$$50.000 \text{ videos} * \frac{5 \text{ MB}}{\text{video}} + 1.000.000 \text{ videos} * \frac{100 \text{ MB} + 5 \text{ MB}}{\text{video}} = 105.250.000 \frac{\text{MB}}{\text{dia}}$$

Com aquestes dades son en un dia i IOPS son operacions per segon:

$$105.250.000 \frac{\text{MB}}{\text{dia}} * \frac{1 \text{ dia}}{86.400 \text{ segons}} = 1218,171 \text{ MBps} = 1218171 \text{ KBps}$$

$$\frac{1218171 \text{ KBps}}{4 \text{ KB per operació IO}} = 304.542,75 \text{ IOPS}$$

2.3- Tràfic amb el client (entre servers i de server a switch de connexió a xarxa):

Cada dia tenim:

- ➔ els 20GB d'informe diari enviat per l'empresa
- ➔ 1.000.000 vídeos analitzats per 100MB de dades per vídeo.
- ➔ 50.000 vídeos que arriben per 5MB per vídeo.

$$20.000 \text{ MB} + 1.000.000 \text{ videos} * 100 \text{ MB} + 50.000 \text{ videos} * 5 \text{ MB} = 100.270.000 \frac{\text{MB}}{\text{dia}}$$

$$100.270 \frac{\text{GB}}{\text{dia}} * \frac{1 \text{ dia}}{86400 \text{ segons}} * \frac{8 \text{ Gb}}{1 \text{ GB}} = 9,28 \text{ Gbps}$$

2.4- Tràfic amb el disc:

Cada dia tenim:

- ➔ 1.000.000 vídeos analitzats per 5MB per video.
- ➔ 50.000 vídeos que arriben per 5MB per vídeo.

$$1.000.000 \text{ videos} * 5 \text{ MB} + 50.000 \text{ videos} * 5 \text{ MB} = 5.250.000 \frac{\text{MB}}{\text{dia}}$$

$$5.250 \frac{\text{GB}}{\text{dia}} * \frac{1 \text{ dia}}{86400 \text{ segons}} * \frac{8 \text{ Gb}}{1 \text{ GB}} = 0,486 \text{ Gbps}$$

2.5- Pressió sobre la xarxa (ample de banda mínim necessito per servir el tràfic de client i disc). M'arriba?:

Per calcular al pressió sobre la xarxa, sumem el tràfic amb el client i el tràfic amb el disc.

$$9,28 \text{ Gbps} + 0,486 \text{ Gbps} = 9,766 \text{ Gbps}$$

Si que ens arriba ja que tenim una xarxa de 10 Gbps encara que aniríem justos ja que aquest valor es una mitjana i potser el pic supera els 10 Gbps.

3.-Decisions preses

3.1- Descripció dels elements d'emmagatzematge escollits, en funció de les necessitats. Quants tipus de cabines? (i perquè), RAID escollit a cadascuna d'elles. Nombre de cabines de cada tipus

3.2- Es justifica la necessitat d'un SAN? Si la resposta és si, raonar si el cost és assumible o no, i cas de no ser-ho calcular l'impacte sobre el rendiment del CPD

3.3.- Posem un *mirror*?

3.4- Empresa de *housing* escollida i perquè (relació entre el que ofereix, el que necessito i el que costa)

3.5- Posem monitorització?

3.6- Opció de backup?

3.7- Tràfic amb l'exterior afegit pel sistema de *backup/mirror* escollit. Quin *bandwidth* caldria?

4.-Recomanacions als inversors

4.1.- Anàlisi de Riscos (*Risk Analysis*)

Quines desgràcies poden passar i com les hem cobert?
Al menys s'han de cobrir els següents casos:

- Hi ha pèrdua d'un fitxer (per error o corrupció). De quan puc recuperar versions?

- Es trenca un disc (es perden dades? quan trigo en recuperar-me? el negoci s'ha d'aturar?)
- Puc tenir problemes de servei si falla algun disc?
- Cau la línia elèctrica. Què passa?
- Cau una línia de xarxa. Què passa?
- En cas de pèrdua o detecció de corrupció de dades no ens podem permetre seguir treballant fins que recuperem les dades correctes. Calculeu temps i costos de recuperació en cas de
 - Pèrdua/ corrupció d'un 1% de les dades
 - Pèrdua/ corrupció de la totalitat de les dades

4.2.- Anàlisi de l'impacte al negoci (*Business Impact Analysis*)

En funció de l'anàlisi de riscos anterior i del que costa estar amb la màquina aturada o no donar el servei complet, calcular quant perdo en diners per tenir-lo aturat i quan em costaria evitar aquesta situació.

Caiguda de la xarxa de dades:

Fallada de disc

4.3.- Creixement

Si creix el nombre de clients/ màquines/ dades (depèn de l'escenari), hem d'estar preparats. Quin creixement (en nombre de clients, etc...) podem assumir sense canviar el sistema (sobreprovisionament)? Quin és el recurs que s'esgota abans? Feu un informe de les implicacions que suposaria un increment d'un 20% en el volum de negoci (tot, clients, dades, ...)

4.4.- Inversions més urgents

Donat el CPD resultant és possible que no haguem escollit la millor opció per manca de diners. El CPD no és nostre, nosaltres només ho dissenyem, així que al final s'hauria de fer un informe als que posen els diners de en què valdria la pena invertir per millorar rendiment, seguretat o...