Re: Measuring IOPS and Raid penalty

Source:
http://www.tech−archive.net/Archive/Exchange/microsoft.public.exchange.admin/2006−03/msg00556.html

- *From*: "John Fullbright" <fullbrij@xxxxxxxxxxx>
- *Date*: Fri, 3 Mar 2006 13:55:23 −0800

750*.73 = 547.5

Add 20%

547.5*1.2 = 657 IOPS for Database

Now calculate the logs. Next, calculate your read/write ratio and apply the write penalty for your RAID type.

For Example:

Database disk requirement with a measured read/write ratio of 2:1.

Reads =657*.66 = 433.62
Writes = 657*.33 = 216.81

The write penalty for RAID 10 is 2:

Required IOPS capacity = 433.62+(216.81*2) = 867.14

That's about 10 spindles at 85 IOPS/spindle or 8 spindles at 120 IOPS/spindle

The write penalty for RAID 5 is 4:

Required IOPS capacity = 433.61+(216.81*4) = 1300.86

For raid 5, first we figure out the number of data spindles −1300.86/85 = 15.3 for 10K SCSI or 1300.86/120 = 10.845 for 15K SCSI. Then you need to add in parity spindles. 1/6 on HP or up to 1/11 on EMC.

RAID 5 on HP with 10K spindles = 19
RAID 10 on HP with 10K spindles = 10
RAID 5 on HP with 15K spindles = 13
RAID 10 on HP with 15K spindles = 8
RAID 5 on EMC with 10K spindles 18
RAID 10 on EMC with 10K spindles = 10
RAID 5 on EMC with 15 K spindles = 13

RAID 10 on EMC with 15K spindles = 8 spindles

Of course for NetApp ther is no write penalty:

RAID 4 10K = 9 spindles
RAID 4 15K = 7 spindles
RAID DP 10K = 10 spindles
RAID DP 15K = 8 spindles

John

=

"SW" <SW@xxxxxxxxxxxxxxxxxxxxxxxxx> wrote in message
news:A3D7EDF6−E941−4E8C−AA14−CE0E0A26E5FC@xxxxxxxxxxxxxxx

> If I use our peak then that is 0.730:
>
> IOPS/mailbox = (average disk transfer/sec) ÷ (number of mailboxes)
>
> our average disk transfer/sec peak was 0.730 and the number of mailboxes
> are
> 750
>
> Would this equal 730/750 = 097 or 0.730/750 = 9.7
>
> "John Fullbright" wrote:
>
>> "IOPS/mailbox = (average disk transfer/sec) ÷ (number of mailboxes)"
>>
>> According to "Optimizing Storage Performance for Exchange Server 2003"
>> you
>> should be using peak, not average. Think about it. If you use the
>> average,
>> you'll be undersized and performing poorly 50% of the time. If my minmum
>> is
>> 1, my average is 5, and my peak is 10, then if I design for average ...
>> This is a common sizing mistake.
>>
>> From the paper
>>
>> "8.
>> Identify the ex2003base server that experienced the highest load.

Use
the data collected from the server with the highest load as your
server/processor/storage baseline. Use the following best practices:

. Always design your system to allow 20 percent more
utilization
than you expect for peaks. This allows the storage and processors to
handle
spikes during peak periods.

. Megacycles per mailbox and IOPS per mailbox change as the
server configuration changes. The following list includes potential
factors
that can change the given megacycles per mailbox and IOPS per mailbox.

. Mailbox sizes are changed significantly

. Max message size is changed significantly

. Third party applications are added or removed

. Exchange features are added or removed.

. Average concurrency of the users changes (more or less
users
are online using the system at any given time).

9.
After the spreadsheet (included with the download of the guide
Optimizing Storage for Exchange Server 2003) is populated and your
mailbox
profiles are determined, you can design your storage solution. For
example,
if your analysis indicates that your standard mailbox profile translates
to
..75 IOPS per mailbox and 1.25 megacycles per mailbox, you can determine
the
following requirements for a 4,000 mailbox server:

. Mailbox Count: 4,000

. Peak DB IOPS: (4,000 × .70) = 3,000

. Peak Log IOPS: (DB IOPS/10) = 300

. Peak megacycles: (4,000 × 1.25) = 5,000 megacycles

To handle spikes, you should add a 20 percent buffer to your
processor

and storage design. With the addition of this buffer, the minimum hardware
requirements for this example are:

. Mailbox count: 4,000

. Peak DB IOPS: (3,000 + 20%) = 3,600

. Peak Log IOPS: (300 + 20%) = 360

. Peak Megacycles: (5,000 + 20%) = 6,000 megacycles

"

The standard in the same referenced paper is average write latency less than
20ms with no peaks lasting more than a few seconds over 50ms. You should size for a 20ms IO response time. A lot of people just sum the average sekk
time and rotational latency of a disk and divide one second by that to determine how many IOs a spindle can sustain. This method does not take into account IO response time. For example, lets take a 15K RPM SCSI spindle. For the following IOPS/spindle numbers, here are the response times you can expect:

10ms 87 IOPS
20ms 125 IOPS
50ms 200 IOPS

Not taking into account required response times is probably the second most
common sizing mistake.

"SW" <SW@xxxxxxxxxxxxxxxxxxxxxxxxx> wrote in message
news:FA50BED9−8B11−4118−AD1E−348D6BC0E8E0@xxxxxxxxxxxxxxx

> We have Exchange 2003 with a RAID 5 volume which has 6
> x 10,000 SCSI
> disks.
>
> We are trying to work out if we move to RAID 10 have
> 15,000 would

benefit
us.

To measure IOPS/mailbox

We Used the System Monitor tool to monitor Physical Disk\Disk
Transfers/sec
counter over the peak 2 hours of server activity.

To calculate our current IOPS/mailbox we used the following formula:

IOPS/mailbox = (average disk transfer/sec) ÷ (number of mailboxes)

our average disk transfer/sec was 0.193 and the number of mailboxes are
750

Ours

193.531/750= 0.258041 What does 0.258041 mean? We need to then go
further
and compare our current RAID 5 speed etc with the RAID 10 configuration
we
want to put in.

PLEASE can you explain you calculations to us to so we understand all
of
it?

Many thanks in advance!

.