

DRS Deepdive part II

October 22nd, 2009 | 6 Comments |

Yesterday I posted the DRS Deepdive. One of the questions still left open was how DRS decides which VM to move to create a balance cluster. After a lot of digging for non-NDA info I found this "procedure" in a VMworld presentation(TA16) amongst some other cool info.

The following procedure is used to form a set of recommendations to correct the imbalanced cluster:

```
While (load imbalance metric > threshold) {
  move = GetBestMove();
  If no good migration is found:
    stop;
  Else:
    Add move to the list of recommendations;
    Update cluster to the state after the move is added;
}
```

Step by step in plain English:

While the cluster is imbalanced (Current host load standard deviation > Target host load standard deviation) select a VM to migrate based on specific criteria and simulate a move and recompute the "Current host load standard deviation" and add to the migration recommendation list. If the cluster is still imbalanced(Current host load standard deviation > Target host load standard deviation) repeat procedure.

Now how does DRS select the best VM to move? DRS uses the following procedure:

```
GetBestMove() {
  For each VM v:
    For each host h that is not Source Host:
      If h is lightly loaded compared to Source Host:
        If Cost Benefit and Risk Analysis accepted
          simulate move v to h
          measure new cluster-wide load imbalance metric as g
    Return move v that gives least cluster-wide imbalance g.
}
```

Again in plain English:

For each VM check if a VMotion to each of the hosts which are less utilized than source host would result in a less imbalanced cluster and meets the Cost Benefit and Risk Analysis criteria. Compare the outcome of all tried combinations(VM<->Host) and return the VMotion that results in the least cluster imbalance.

This should result in a migration which gives the most improvement in terms of cluster balance, in other words: most bang for the buck! This is the reason why usually the larger VMs are moved as they will most likely decrease "Current host load standard deviation" the most. If it's not enough to balance the cluster within the given threshold the "GetBestMove" gets executed again by the procedure which is used to form a set of recommendations.

Now the next question would be what does "Cost Benefit" and "Risk Analysis" consist of and why are we doing this?

First of all we want to avoid a constant stream of VMotions and this will be done by weighing costs vs benefits vs risks. These consists of:

- Cost benefit
 - Cost: CPU reserved during migration on the target host
 - Cost: Memory consumed by shadow VM during VMotion on the target host
 - Cost: VM "downtime" during the VMotion
 - Benefit: More resources available on source host due to migration
 - Benefit: More resources for migrated VMs as it moves to a less utilized host
 - Benefit: Cluster Balance
- Risk Analysis
 - Stable vs unstable workload of the VM (historic info used)

Based on these considerations a cost-benefit-risk metric will be calculated and if this has an acceptable value the

ABOUT YELLOW-BRICKS.COM

Welcome to Yellow-Bricks.com a blog about virtualization, with a strong focus on VMware, HA and vCloud solutions.

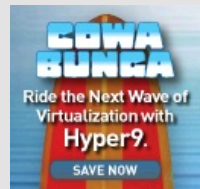
Some facts:

1000 Blog Articles

1971 Twitter Followers

4650 RSS Subscribers

SPONSORS



ADVERTISEMENTS

See the industry's best
I/O virtualization solutions Meet us at VMworld 2010, Booth 1413
www.virtensys.com

Monitor VM Performance
Improve VMware Performance Fast - Download Capacity Analyzer Now!
www.VKernel.com/LearnHow

Virtualization Costs
Allocate Costs and Manage TCO Decide What Services to Virtualize
www.digitalfuel.com/virtualization



RECOMMENDED



vSphere 4.0 Quick Start Guide

Shortcuts down the path of Virtualization

Buy it now just \$15.99.

ADVERTISEMENTS



Watch this video on www.youtube.com

SQLserver on VMware
Use Veloxum to improve performance immediately, Free Trial
youtube.com/ks cruzen

VM will be considered for migration.

I will consolidate both posts in a single blog page today to make it easier to find!

VMware I/O Performance

VM to LUN optimization for your business critical apps
www.virtualinstruments.com



Ads by Google

Posted in Management & Automation, PASS Syndication, Server | Tags: deepdive, drs, vcenter, vmotion, vSphere

« DRS Deepdive

DRS Deepdive Blog page »

You can skip to the end and leave a response. Pinging is currently not allowed.

6 Responses to “DRS Deepdive part II”



michael says:

Thursday, October 22, 2009 at 17:29

thanks for sharing! great info!



Brian Knudtson says:

Thursday, October 22, 2009 at 20:45

Duncan-

Can you dig into and explain why DRS assigns a Priority 1 to VMs that are being evacuated due to maintenance mode? Seems like it would make more sense to set these as Priority 5 so that with default DRS settings (Priority 3 and above) they would be applied automatically.

Thanks

brian



Cody Bunch says:

Friday, October 23, 2009 at 06:11

Wow. Good stuff. Thanks!



Andreas Berg says:

Friday, October 23, 2009 at 16:00

Nice. Very interesting reading. Thanks!



Arseny says:

Sunday, October 25, 2009 at 03:57

Brian, I believe think Priority1 is “higher” than 5, – and Priority1 would work even in Conservative DRS setting.

Duncan, thanks for great coverage! One thing I was trying to dive into recently was ‘Percent for each VM of entitled resources Delivered’, – can you cover this in ‘regular English’ afterwards? I believe it’s a math against memory pressure, CPU fairness and whatever...

Thanks again for this great coverage,

Have a great weekend,

Arseny



bitsorbytes says:

Tuesday, October 27, 2009 at 06:08

@Brian... If your putting a host into maintenance mode, then there is most likely an issue with it. I would rather have these VMs move off, then a cluster trying to balance itself.

Ads by Google

RECENT COMMENTS

- ▶ LucD on Standby vCenter Server for disaster recovery
- ▶ Forbes Guthrie on DRS Deepdive
- ▶ Duncan Epping on Standby NICs in an “IP-Hash” configuration
- ▶ Viktor on Standby NICs in an “IP-Hash” configuration
- ▶ Duncan Epping on Standby NICs in an “IP-Hash” configuration

TAGS

2.5 3.0.x 3.5 BC-DR Blogging Book

Bugs cloud design drs **ESX** esxi ha

Howto kb linux networking News patches

performance powershell Scripting

scripts security service console software srm

Storage Tools training u2 update

upgrade VCB vcdx vcenter vdi

VirtualCenter vmtn **VMware**

vmworld vSphere vstorage whitepaper

Workstation

DISCLAIMER

The views expressed anywhere on this site are strictly mine and not the opinions and views of VMware.

Leave a Reply

Name (required)

Mail (will not be published) (required)

Website

Submit Comment