Challenge 1 – Back to Earth

concept

Stanislav Majzlik
28-06-2016

Summary

Hooray, Zombies gone! Our planet is safe again. Thank you Steven, you are the best. So let's take our planet again, and first start with the new infrastructure to support re-colonization. Big thanks to Bill, Warren, Mark or whoever gave us the money, lets spend it on some great stuff.

We are going to build multi site environment fro the world's new infrastructure. The most critical application is the HumanityLink software suite. Performance of this software is paramount.

Overview

Earth

According to Wikipedia, Earth (otherwise known as the world) is the third planet from the Sun, the densest planet in the Solar System, the largest of the Solar System's four terrestrial planets, and the only astronomical object known to harbor life.

Moon

According to Wikipedia, Moon is Earth's only permanent natural satellite. It is one of the largest natural satellites in the Solar System, and the largest among planetary satellites relative to the size of the planet that it orbits (its primary).

Earth – Moon

The Moon is in synchronous rotation with Earth, always showing the same face with its near side marked by dark volcanic maria that fill between the bright ancient crustal highlands and the prominent impact craters. The semi-major distance Between Earth and Moon is 384399 km.

Conceptual design

Data centers

The requirement was to build primary site on Earth and secondary site on Moon. Assuming that all of the humans alive are still evacuated on Moon this approach would not be logical. As a lot of preparations would need to be done and the HumanityLink is the core software I would leave the primary site on the Moon till we will be ready to send the moonwalkers back to earth. This decision will provide the HumanityLink software to the users on the Moon with best responses of the interface so we will avoid the usual end users complains about the software "not working" and made the end users support happy. In another words, I would

recommend to use primary site as active site and use the secondary site as a DR and failback destination.

Once the humans will have their tickets to SpaceX 1st. regular flight from Moon to Earth sitting in the Airport departure hall, behind the security check at least 2 hours before actual departure we can start with the data migration from Moon to Earth and switching the Primary and Secondary Sites roles between the sites.

Current fastest manned mission from Moon to Earth took 70 hours so this time will provide us downtime big enough to migrate all the data to Earth. Also we can use the flight to send offline copy of the data to use in case of failure during the migration.

After couple of discussions about best physical location of the Datacenters I have adopted idea of placing the datacenter in Slovakia. Most of the humans are replacing Slovakia with Slovenia and vice versa which is enough to confuse Zombie-terrorist as the distance between the center of the countries is 452 km and the countries are not adjoining countries (thanks Katarina for the idea). The second thing is that most of US and EU countries are considering Slovakia as a Eastern Europe country (which Slovakia never was) so no one really cares what is happening there (4 from 5 security officers recommend).

The best position on the Moon will be somewhere around Ptolemaeus crater as the area is visible from the Earth so with right equipment we can check if the HW engineer is really heading to datacenter with the correct spare part to fulfill our 4 hours fix time Gold SLA on HW maintenance contract.

To keep the complexity and configuration as simple as possible both datacenters will be configured exactly the same way. This will also provide some type of standardization. Cloud software suite will consist from products by VMware. As a hypervisor we will use VMware vSphere ESXi 6.0 Update 2. For management we will use VMware vCenter Server 6.0 U2, Site Recovery Manager 6.1 with vSphere replication will be used to migrate the data between the sites and use the secondary site as DR, failback site.

Network

Network connectivity between Earth and Moon is a tough challenge itself which most possibly would need couple of years and couple hundred thousand pages of documentation. As mentioned before, the semi-major distance Between Earth and Moon is 384399 km which give us average latency of 1,3s. In 2013 NASA presented results of successful mission Lunar Laser Communications Demonstration (LLCD). The main goal of LLCD is to prove the fundamental concepts of laser communications and transfer data at a rate of 622 megabits per second (Mbps), which is about five times the current state-of-the-art from lunar distances. NASA used devices created by MIT Lincon; Laboratory Lunar Laser Ground Terminal (LLGT) and Lunar Laser Space Terminal (LLST) which were using laser as a data carrier. To make the long story short NASA was able to test up to 20Mbits uplink / 622Mbits downlink on Earth to Moon distance.

So now all MIT and NASA engineers who were working on that mission are anyway locked out on Moon and have plenty of free time, I would ask them to modify the already tested technology to provide us synchronous connection and sort out also the issues around Moon not being in fixed position to Earth. I would recommend to find multiple points on Earth which will together with networking protocols like BGP give us uninterrupted network connection with low latencies between Earth and Moon. As a failback solution we can still use well tested RF connection but with limited throughput.

From now on I would assume that these intelligent people spent their time wisely and were able to sort all related issues so our network traffic has reasonable throughput and short latencies.

Storage

Storage will be provided by VMware Virtual SAN 6.2 with all flash drives. Standalone storage devices are not going to be used as this will add more complexity. SSDs will provide us best performance and fast response for the software running on top of it.

Failures to Tolerate (FTT) will be set to 1 to retain maximum usable capacity while still providing resiliency.

Physical hosts

HW, CPU and memory capacity were calculated according to assumption that the VM configuration will be following:

	number of VMs	vCPU per VM	Memory per VM (GB)	Storage per VM (GB)
Domain controller	2	1	8	60
PSC	2	2	2	30
VCSA	1	4	18	100
SRM	2	2	10	100
HumanityLink Web Server	3	4	8	100
HumanityLink DB Server	1	8	16	500
HumanityLink App Server	2	4	8	250
Web Servers	25	4	8	100
DB Servers	5	8	16	500
App Servers	10	4	8	250

Total allocated capacity is following:

Number of VMs: 53
Number of vCPUs: 222
Number of Memory: 474 GB
Storage size: 9280 GB

The minimum number of host to provide required capacity will be 4. Taking into account the standard N+2 I would recommend 5 host at a minimum and 6 as optimal.

The server that will be used is a 2U node HP Proliant DL380 G9 Server.

Processor	2 x 16-Core Intel® Xeon® Processor E5-2683 v4 (40M Cache, 2.10 GHz)
Memory	8 x HPE 16GB 1Rx4 PC4-2400T-R Kit
Storage controller	Smart Array P440ar/2GB FBWC
Hard Drive	3 x HP 1.6TB 6G SATA Mixed Use-2 SFF 2.5-in SC 3yr Wty Solid State Drive 1 x HP 480GB 6G SATA Mixed Use-2 SFF 2.5-in SC 3yr Wty Solid State Drive
Network Card	2 x HP InfiniBand FDR/Ethernet 10Gb/40Gb 2-port 544+FLR-QSFP Adapter

Operating system

All physical servers will be installed with VMware vSphere 6.0 U2 Enterprise Plus. All physical servers will boot from USB device.

Compute Capacity

Selected HW will provide use following compute capacity per host:

CPU sockets	2
CPU cores (physical)	32
CPU cores (logical with HT enabled)	64
RAM	128 GB

VMware vCenter Server

vCenter Server Virtual Appliance will be used on each site to manage local physical hosts. Each site will have 2 Platform Services Controllers and 2 Domain Controllers. Single AD domain will be used between the sites. Domain controllers will be also used to provide DNS and NTP services.

Each site will be configured with 1 virtual data center with 1 virtual cluster containing 6 ESXi hosts. Cluster will be enabled for vSAN, HA and vMotion.

While best practice would prefer to configure a separated management cluster I would not recommend it in this environment as the number the overall number of VMs is not high. This approach will also help us to save on power, space and cooling. Also this configuration is simple so it would not be difficult to manage, maintain and document.

Network

Each physical host is equipped with a HP InfiniBand FDR/Ethernet 10Gb/40Gb 2-port 544+FLR-QSFP Adapter which will provide us 4 uplink ports.

Following setup will be used:

Two virtual distributed switches with 2 uplinks will be created. First vDS will be dedicated for Management, vMotion and VM portgroups, second vDS will be dedicated for vSAN portgroup. Portgroups will be configured with two uplink ports, one from each Network Adapter. Each uplink port will be connected to separate switch. This will provide us with failover capacity in case of singe port failure as well in case of single Network adapter failure. For first vDS, VLANs will be used to segregate Management, vMotion and VM traffic.

Replication

For failover and DR we will use VMware Site recovery manager using vSphere replication. This will provide us central management for building automated recovery plans in advance, testing the recovery plans and actually automate the execution of the recovery plans.

New recovery plan will be created which will cover HumanityLink software suite VMs. vCenter server, PSCs and domain controllers wont be covered as both sites have their own vCenter server, PSCs and domain controllers. Protected site will be on the Moon, recovery site will be on the Earth. Once most of the humans will be on Earth (or on the way to Earth), we can then use the Planned migration feature of SRM to migrate the HumanityLink suite to Earth. After that we can reconfigure SRM to set the Earth site as protected site and Moon site as recovery site.