
MAAS - Datacenter

Tutor de la pràctica : Antonio Bueno

Marc Sànchez Pifarré, GEINF (UDG-EPS)

4 de Desembre de 2019

Contents

MAAS (Metal as a service)	3
Context	3
Possibilitats en funció del hardware.	3
Com funciona MAAS?	4
Hardware	6
Utilització del hardware	6
Configuració prèvia del hardware	7
Connectivitat	9
Instal·lació de MAAS	9
Primera configuració	10
Accés	10
Muntar el cluster mitjançant UI	11
Virtualització	12
Referències	13

MAAS (Metal as a service)

Context

Una empresa ISP de girona (l'anomenarem Fibrix) vol entrar al mercat de l'emmagatzemament web i facilitar als seus clients la possibilitat de poder allotjar-hi les seves pròpies màquines virtuals. De fet aquesta empresa té l'objectiu de unificar el domini de una connexió bona de dades, no només venen el servei de connexió sinó també afegint els serveis de datacenter, concretament la possibilitat que una altra empresa tecnològica del sector pugui contractar amb el mateix ISP un servei de cloud.

Fibrix ha aconseguit accés a un seguit de màquines físiques provinents de una altra empresa cloud del sector. Aquesta altra empresa ha de mantenir un volum molt gran de feina i ha decidit invertir en Fibrix regalant-li un hardware usat i una mica anticuat.

El hardware regalat és un hardware limitat en capacitats. Concretament ha rebut servidors DELL POWEREDGE 860 [1].

Els servidors que ha aconseguit Fibrix son servidors professionals, de marca propietària i cara però en quan a tecnologia es podrien anomenar vellots. Tenen la possibilitat de tenir 1 processador XEON doble nucli (dual core) amb freqüència de gir màxima de 2.66GHz i memòria RAM de fins a un màxim de 8Gb.

Però se n'adonen que aquests servidors tenen quelcom molt bo, Compatibilitat estàndard de BMC [2] amb IPMI [3] 1.5. Que ve a ser, gestió i control remot de la màquina sense necessitat de sistema operatiu.

Possibilitats en funció del hardware.

Opció 1

La primera possibilitat, també la més vanal, seria pensar en muntar una màquina física per cada client. Però....

- De seguida ens quedariem sense màquines.
- Capacitat d'ampliació limitada per client.
- El cost de la inversió en noves màquines s'hauria de repercutir als clients.

Mal negoci.

Opció 2

Dockerització (IAAS - Infrastructure as a service). Muntar una sèrie de servidors docker i que cada client pugui desplegar les seves aplicacions web, ja siguin html/css o qualsevol altre llenguatge i tecnologia mitjançant una interfície web d'administració.

Aventatges :

- Docker permet crear instàncies virtuals sobre màquines físiques.
- Docker permet crear xarxes internes i overlays entre altres dockers.
- Docker ocupa només el que s'ha desplegat i és escalable.
- Molts...

Docker éns aportaria escalabilitat i fins i tot ens podríem plantejar la alta disponibilitat realitzant bé la feina.

Inconvenients :

Docker està pensat per què els seus nuclis de treball suportin càrregues suficients com per mantenir l'estructura de les aplicacions + les rèpliques dels demés. En el nostre cas les màquines no son gaire grosses i per tant el marge d'escalabilitat és estret i no seria possible implementar una bona qualitat de servei.

Tanmateix el producte que es vol vendre als clients son màquines senceres, poden ser windows Server, Ubuntu, CentOS. **Docker no està pensat per suportar la integració de Sistemes operatius complets, està pensat per ser instal·lat sobre un kernel i emular-lo.**

Opció 3

El muntatge d'un cluster, capaç de tenir la flexibilitat d'afegir i treure màquines físiques per poder-lo fer escalable. Ens ha de permetre la creació de màquines virtuals mitjançant imatges i el desplegament de les mateixes. Tanmateix ens ha de subministrar eines per a la comprovació de l'estat del hardware de les màquines físiques.

Aquí entra MAAS [4].

Una plataforma per crear datacenters basada en virtualització i feta exclusivament pensant en grans infraestructures que volen tenir la possibilitat d'escalar. La clau de l'èxit és la manera que té maas d'aprofitar la màquina física utilitzant BMC i IPMI. les màquines no requereixen de sistema operatiu i posen a disposició del hardware tots els recursos dels que disposen.

L'empresa Fibrix decideix que aquesta és la solució més viable al seu objectiu.

Com funciona MAAS?

Maas defineix uns conceptes importants per funcionar.

- Region.
- Controller.
- Fabric.
- Pool.

L'esquema seria quelcom similar a:

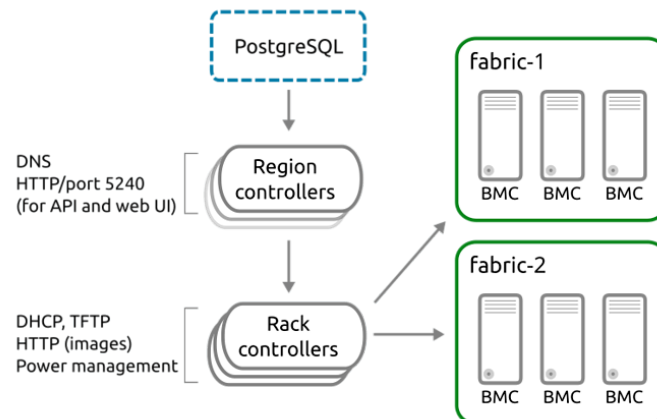


Figure 1: Rack and region scheme.

Region

Una regió és definida com un espai físic on s'hi situen un seguit de màquines, una empresa TIER 3, per aconseguir tenir el TIER 3 està obligat a tenir rèpliques de les màquines a 3 regions diferents. És part de la alta disponibilitat.

L'empresa Fibrix, degut a què és una start up i que encara no ha pogut desenvolupar el sistema està definida directament com una regió on hi podrà haver 1 o més controladors.

El servidor de regió ha de ser el que gestiona les peticions d'operació. És a dir, ha de ser capaç de gestionar el DNS per tot el que es consideri regi, en definitiva és el punt de gestió del datacenter.

Controller

Els controladors son els racks on hi poden haver diferents màquines físiques. Proveeix serveis d'alta capacitat de xarxa a tots els servidors del rack. En essència cobreix les necessitat que hi pugui haver en un rack de servidors físic dins d'un datacenter.

Se n'ocupa de serveis com per exemple :

- El wake on lan de les seves màquines.
- L'arrencada per PXE.

- Gestiona els serveis BMC amb IPMI o Virish.
- Gestiona la virtualització entre les diferents màquines.
- Se n'encarrega del dhcp de la xarxa i manté la subxarxa.

En general cada rack controller ha de tenir la seva subxarxa, ja sigui mitjançant el pas d'una NAT o sense.

Fabric

Totes les màquines físiques dels rack controllers es poden agrupar per generar el que s'anomenen Fabric. Cada Fabric consta de una sèrie de màquines físiques que per simplicitat direm que es troben al mateix rack, però en general son les que estan dins d'una VLAN.

En aquesta capa encara estem treballant amb màquines físiques i el concepte de virtualització encara no ha aparegut.

Pool

Aquí apareix el concepte de virtualització, i és que, totes les màquines físiques que gestiona maas acaben component el que s'anomena una piscina de recursos. Dins d'aquesta piscina és on es poden generar màquines virtuals.

Hardware

Com a hardware tenim un seguit de servidors DELL POWEREDGE 860 amb un dual Core i Majoritàriament 4Gb de RAMM.

El hardware del que disposem està preparat per treballar en mode cluster, definirem el mode cluster com els pc's que disposen de :

- WAKE ON LAN
- BMC Support
- Boot per PXE
- Virtualization support

Utilització del hardware

MAAS necessita els fabric's, és a dir, les màquines físiques, únicament com a màquines que presten el seu hardware. No cal que tinguin cap sistema operatiu. Per fer-ho, primer de tot utilitza el wake on lan per encendre-les. Un cop s'han encés, s'ha de configurar el BMC (Baseboard Management Controllers) [6] de la màquina, en aquest cas el sistema de Baseboard Management Controllers és

IPMI versió 1.5, això ho estipula el fabricant. BMC es configura per aconseguir una ip en xarxa, per tant s'ha de permetre que la màquina arrenqui des de PXE des de la BIOS. Curiosament aquests servidors estan programats per què si hi ha hardware que està habilitat però que no està posat (un disc sense endollar per actiu a la bios), que directament no arrenqui.

Configuració prèvia del hardware

A l'hora d'afegir màquines al hardware s'han de prendre diferents mesures per poder garantir que el node controller té control total sobre elles. Per tant s'han de realitzar els següents passos.

Configurar l'arrencada en xarxa des de la bios, MAAS ha de poder apagar i encendre les màquines quan ho requereixi per poder tenir un escalat més flexible.

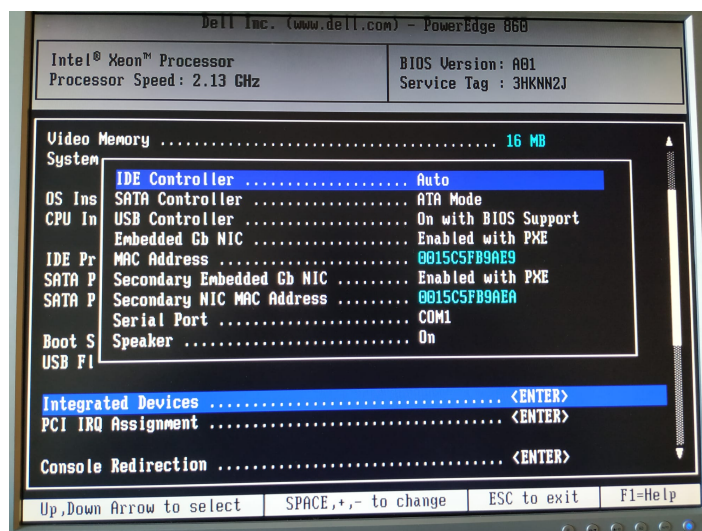


Figure 2: PXE Configuration

Configurar el Wake on lan mitjançant Ctr + S, per poder arrencar-les mitjançant la interfície de xarxa es fa servir el protocol :

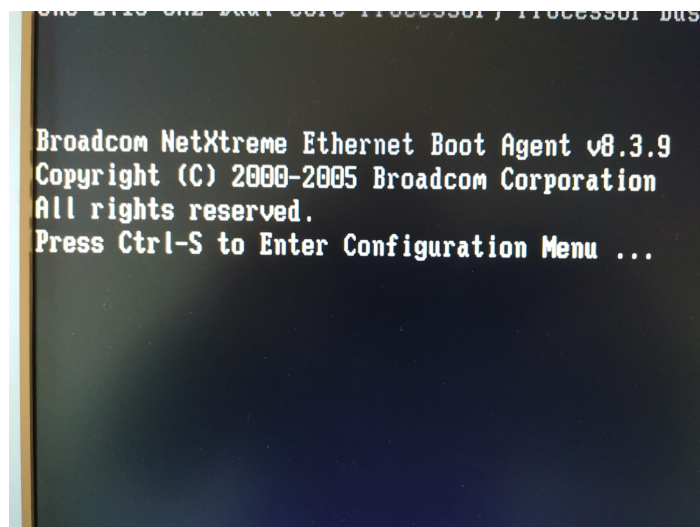


Figure 3: Accés WOL

Activant-lo n'hi ha prou.



Figure 4: WOL

Finalment s'ha de configurar el BMC (IPMI) accedint al seu panell amb Ctr + E. Volem que MAAS (el region controller) utilitzi les altres màquines únicament com a recursos hardware, sense fixar-nos en el sistema operatiu que porten. Degut a que les màquines ho permeten utilitzem IPMI però es podrien configurar mitjançant VIRISH [7], un sistema de cluster sobre virtualització. Aquest sistema sí que requereix que hi hagi un sistema operatiu a les màquines worker amb el sistema de virtualització QEMU [7] instal·lat.

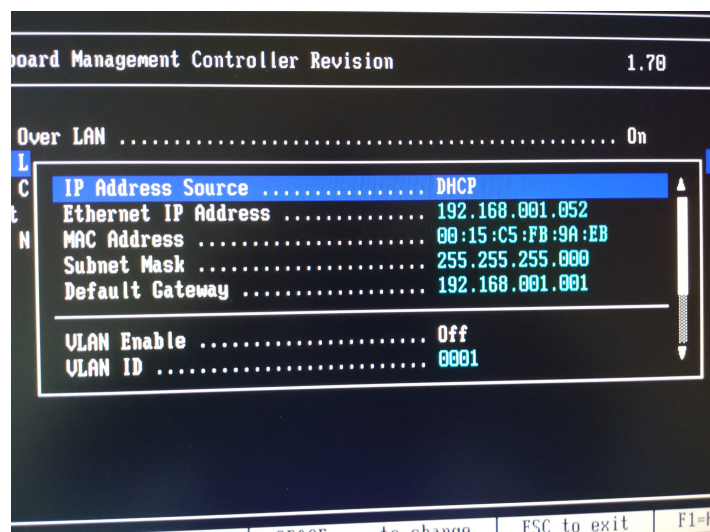


Figure 5: BMC

Connectivitat

El sistema a plantejar és la màquina Rack / Region controller, que té dues interfícies de xarxa, estigui connectada amb la interfície 1 a internet (amb IP pública). El controlador de regió conté el DNS local per poder resoldre les màquines físiques i virtuals de la mateixa regió mentre que el controlador de rack conté el DHCP per poder assignar ip's a les màquines que el componen.

Per tant des de l'exterior, cap a la xarxa del datacenter podem afirmar que hi ha un NAT. Internament, MAAS gestiona les IP's sense problemes generant subxarxes per cada Rack Controller. Aquestes subxarxes son configurables des de la interfície gràfica.

En el nostre cas, utilitzem el rang d'adreces que conté la xarxa 192.168.1.0/24.

La connectivitat entre el node rack i la resta de workers és mitjançant un switch ethernet.

Instal·lació de MAAS

Instal·lar Maas es pot fer directament des del paquets apt. La idea no és fer un tutorial de instal·lació sinó veure quins son els diferents problemes que apareixen a l'hora d'intentar muntar un datacenter.

Afegim el repositori

```
1 sudo apt-add-repository -yu ppa:maas/stable
```

i Instal·lem.

```
1 sudo apt update
2 sudo apt install maas
```

En el nostre cas, per simplicitat, muntarem tant el Region com el Rack en una sola màquina. De fet és com aconsellen que es faci si no requereixes duplicar la infraestructura en diferents llocs geogràfics diferents. Si es requerís es podria muntar un Region controller en una màquina física i un Rack controller en una altre màquina física.

Més informació a [5].

Primera configuració

El primer que s'ha de fer un cop s'ha instal·lat MAAS és la inicialització. En el nostre cas li he de dir que volem que es comporti tant de region com de rack controller, això passa per fer-ho amb la següent instrucció.

```
1 sudo maas init -mode all
```

Finalment, abans d'entrar a la interfície gràfica d'administració, hem de crear l'usuari root que hi tingui accés.

```
1 sudo maas createadmin --username=$PROFILE --email=$EMAIL_ADDRESS
```

on :

- \$PROFILE : És el nom d'usuari
- \$EMAIL_ADDRESS : és el correu electrònic de l'usuari.

Podem inicialitzar el password mitjançant aquesta mateixa instrucció passant-li `--password=$PASSWORD`, o podem directament fer que el demani sense posar-lo.

Accés

MAAS dóna accés a la configuració mitjançant la url `http://${API_HOST}:5240/MAAS` on `API_HOST` és l'adreça de la màquina. Un cop en aquesta url iniciem sessió amb l'usuari i el password que s'han inicialitzat anteriorment.

Muntar el cluster mitjançant UI

MAAS realitza una exploració de les màquines que pugui haver-hi a les xarxes en les que ell pertany, té les interfícies en mode promiscu per poder detectar màquines noves i per poder resoldre-les. De fet a la finestra de network discovery apareixen aquestes màquines.

NAME	MAC ADDRESS	IP	RACK	LAST SEEN	ACTIONS
unknown	00:1c:23:a1:d9:9b	192.168.1.50	maas-controller	Mon, 02 Dec 2019 09:04:02	
unknown	00:15:c5:f9:a2:e9	192.168.1.53	maas-controller	Tue, 03 Dec 2019 07:48:24	
unknown	00:1c:23:a1:a3:6a	192.168.1.55	maas-controller	Tue, 03 Dec 2019 07:56:43	

Figure 6: Network discovery

S'han configurat 3 màquines utilitzant la descripció dels passos de hardware que s'han comentat anteriorment.

Hem realitzat un filtre per què no surtin totes les de la xarxa externa (totes les màquines de la xarxa 84.88.154.0/23). Si despleguem una d'aquestes màquines mitjançant la fletxa descendent podem veure la seva adreça mac, que necessitarem per afegir-les al cluster com a machines.

De moment no en tenim cap.

FQDN	IP	POWER	STATUS	OWNER TAGS	POOL	NOTE	ZONE	SPACES	FABRIC	VLAN	CORES	ARCH	RAM	DISKS	STORAGE
0 machines available															

Figure 7: MAAS Empty Machines

Al botó add hardware de la part superior dreta de la pantalla, si el despleguem podem afegir una machine, la diferència entre machine i chassis es pot trobar a [8], Add Nodes. Afegim un node :

FQDN	IP	POWER	STATUS	OWNER TAGS	POOL	NOTE	ZONE	SPACES	FABRIC	VLAN	CORES	ARCH	RAM	DISKS	STORAGE
1 machine available															
1 Machine 1 Resource pool															
Commissioning															
glad-oryx-maas	192.168.1.50 (PXE)	On	Commissioning	bcds	Pool bcds	default			fabric-1	Default VLAN	amd64	0	0 GB	0	0 GB

Figure 8: Commissioning

I tot seguit ens salta a la finestra machines i ens mostra l'estat en el que es troba aquesta màquina.

En aquest moment està en estat Commissioning. És a dir està comprovant els recursos de què disposa aquesta màquina worker, tot seguit quan acabi el commission, passarà els testos de hardware, principalment comprova ramm i disc.

Veiem el llistat de coses que revisa...

NAME	TAGS	RUNTIME	DATE	RESULT	ACTIONS
00-maas-00-support-info	commissioning	0:00:01	Tue, 03 Dec. 2019 09:07:27	Passed View log	...
00-maas-01-cpuinfo	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:28	Passed View log	...
00-maas-01-hw	commissioning	0:00:02	Tue, 03 Dec. 2019 09:07:30	Passed View log	...
00-maas-02-virtuality	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:30	Passed View log	...
00-maas-03-install-ldpd	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:37	Passed View log	...
00-maas-04-list-modalities	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:37	Passed View log	...
00-maas-05-dhcp-unconfigured-ifaces	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:38	Passed View log	...
00-maas-06-get-fuel-api-data	commissioning	0:00:05	Tue, 03 Dec. 2019 09:07:44	Passed View log	...
00-maas-07-block-devices	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:44	Passed View log	...
00-maas-08-serial-ports	commissioning	0:00:00	Tue, 03 Dec. 2019 09:07:45	Passed View log	...
99-maas-02-capture-ldpd	commissioning	0:00:03	Tue, 03 Dec. 2019 09:07:45	Running	...
99-maas-03-network-interfaces	commissioning		Tue, 03 Dec. 2019 09:04:00	Pending	...
99-maas-04-network-interfaces-with-sriov	commissioning		Tue, 03 Dec. 2019 09:04:00	Pending	...

Figure 9: Commissioning single image

Un cop s'ha enllaçat la màquina, ha passat els testos i el commission llavors passa a l'estat de Ready. A partir d'aquí afegirem més màquines al cluster. Utilitzem aquest punt com un checkpoint on a partir d'ara pot ser que les màquines canviïn.

L'estat actual en el que es deixen les màquines és el següent :

Virtualització

Un cop maas coneix les màquines, té accés a elles i pot manegar-les, hem de fer una instal·lació neta de Ubuntu a cada una d'elles per poder tenir un sistema de virtualització. És necessari que per poder virtualitzar cada un dels nodes tingui la capacitat de fer-ho i per tant ens requereix que s'instal·li un Ubuntu. Es pot instal·lar ubuntu de manera automàtica a totes les màquines de cop mitjançant la opció deploy.

Machines 3 machines available		3 selected		Add hardware	Take action
3 Machines 1 Resource pool					
Filters		In (Selected)		Group by status	
<input checked="" type="checkbox"/> FQDN ~ IP	POWER	STATUS	OWNER TAGS	POOL NOTE	ZONE SPACES
Deploying 3 machines selected					
<input checked="" type="checkbox"/> gentle-sheep.maas	On ipmi	Deploying 18.04 LTS	bcds	Pool-bcds	default
<input checked="" type="checkbox"/> glad-onyx.maas	Off ipmi	Deploying 18.04 LTS	bcds	Pool-bcds	default
<input checked="" type="checkbox"/> secure-perch.maas	On ipmi	Deploying 18.04 LTS	bcds	Pool-bcds	default
Show all deploying...					

Figure 10: Add Nodes

Referències

- [1] - Dell powerededge 860
- [2] - BMC
- [3] - IPMI
- [4] - MAAS
- [5] - MAAS Installation
- [6] - BMC - IMPMI 1.5
- [7] - Virish - QEMU
- [0] - Add Nodes