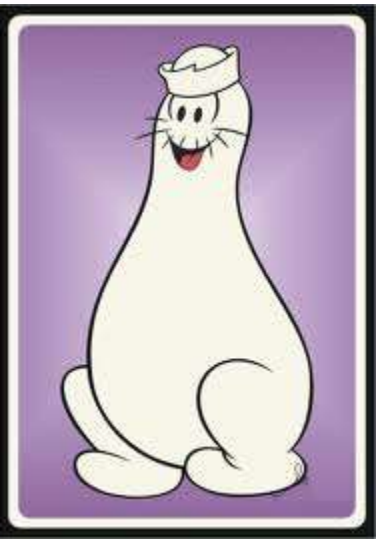


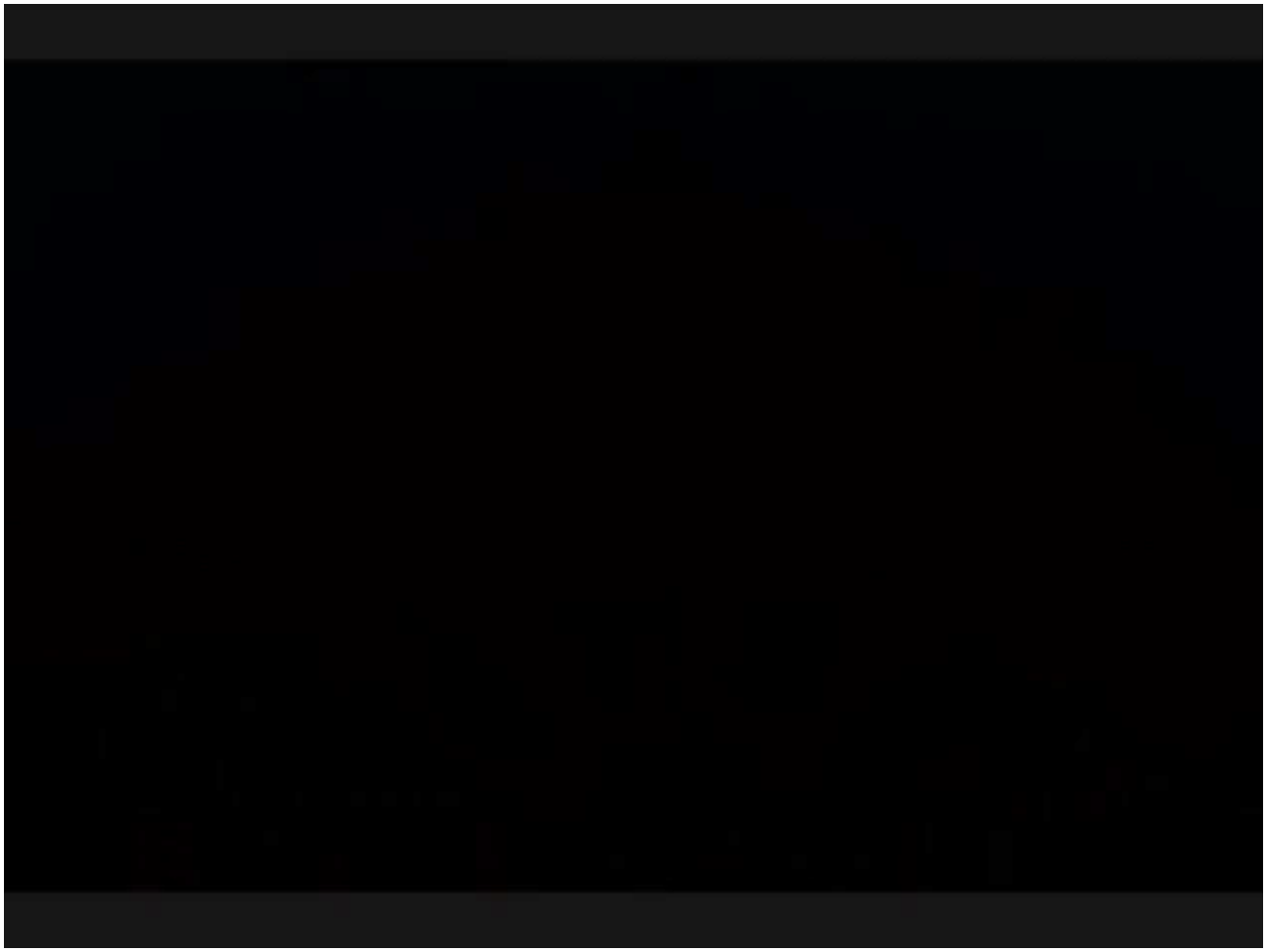
QCON – 2015
Containers and more Containers

UOLHOST

Marcus Vinícius Soares

mvc_msoares@uolinc.com

Complexidade Atual



Complexidade atual (e crescendo)


NO-SQL


TEMPLATE-ENGINE


WEB-SEVER


CACHE-SERVER


APPLICATION1


APPLICATION2


APPLICATION3


QUEUE-SERVER

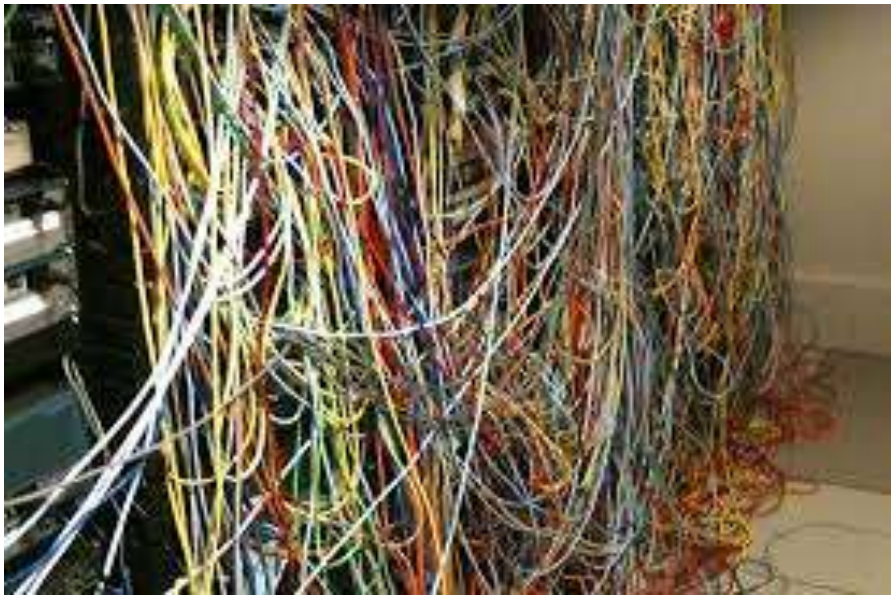

LOG-SERVER


STORAGE


DATABASE-READONLY


DATABASE-TRANSACTIONAL

História



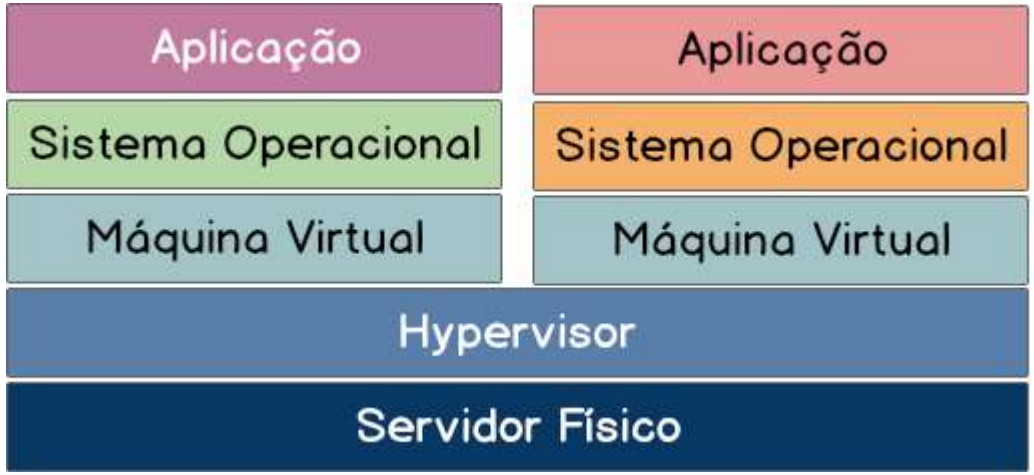
- Virtualização
- Servidores
- Rede
- Storage
- Aplicações – BINGO!!!
  - FreeBSD - Jail
  - Solaris - Zones
  - HP-UX - nPartitions, Vpars, IVms
  - IBM-Aix - PowevVM
  - Linux - OpenVZ

Virtualização

- Benefícios amplamente conhecidos
- Mas....
  - Tempo de provisionamento
  - Quantidade de VM's
  - Provisionamento Dinâmico
  - HPC performance
  - Burocracia interna

LXC – LinuX Containers

Modelo Tradicional com Máquina Virtual



LXC Style

LXC

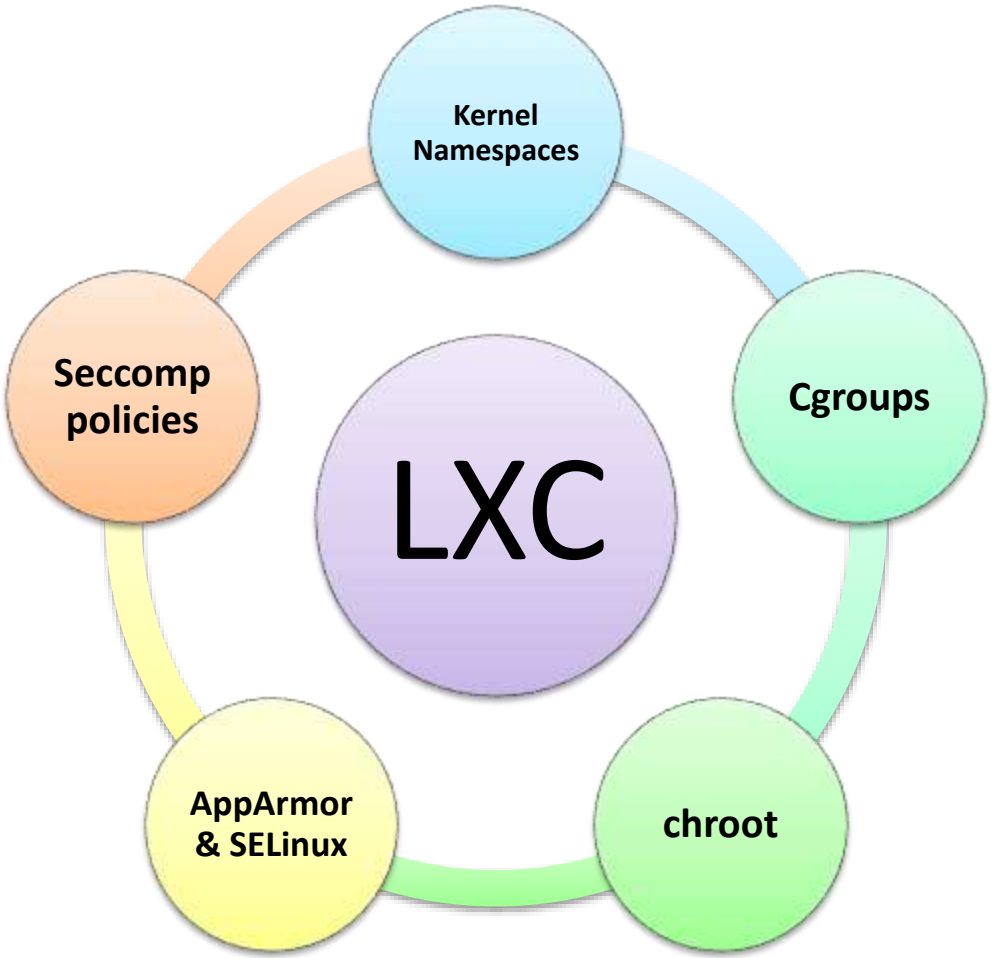
- Ligthweight

	Entrega	Deploy Manual	Deploy Automatico	Boot Time
Bare Metal	Dias	Horas	Minutos	Minutos
Virtualization	Minutos	Minutos	Segundos	Menos de um minuto
Lightweight Virtualization	Segundos	Minutos	Segundos	Segundos

LXC

- Agilidade e Flexibilidade
- Compartimentação
- Open source / Custo
- Larga adoção
- HPC ready
- Cloud stlye
- Kernel Version

LXC


Kernel Namespaces
Cgroups
chroot
AppArmor & SELinux
Seccomp policies
LXC

LXC

Sistema	Tunnable
blkio	- Weighted proportional block I/O access. Group wide or per device. - Per device hard limits on block I/O read/write specified as bytes per second or IOPS per second.
cpu	- Time period (microseconds per second) a group should have CPU access. - Group wide upper limit on CPU time per second. - Weighted proportional value of relative CPU time for a group.
cpuset	- CPUs (cores) the group can access. - Memory nodes the group can access and migrate ability. - Memory hardwall, pressure, spread, etc.
devices	- Define which devices and access type a group can use.
freezer	- Suspend/resume group tasks.
memory	- Max memory limits for the group (in bytes). - Memory swappiness, OOM control, hierarchy, etc..
hugetlb	- Limit HugeTLB size usage. - Per cgroup HugeTLB metrics.
net_cls	- Tag network packets with a class ID. - Use tc to prioritize tagged packets.
net_prio	- Weighted proportional priority on egress traffic (per interface).

LXC – Colocando para Funcionar

- Como fazer um server funcionar?
  - Instalar os pacotes
  - Config (rede / fs / memória / processador)
  - RootFS (images)
  - LXC Tools

LXC – Config

- Configurações dos recursos utilizados pelo container
  - Rede
  - Rootfs
  - Limitações
  - Mount binds
  - Segurança

LXC – Images

- Sua “ISO“
- Read-only para o container
- Shared
- Como gerar ➔ Docker

```
dclient = docker.Client(base_url=docker_url,version='1.16',timeout=120)
container = dclient.create_container(base, "/usr/bin/python -u
/mnt/build/docker-template-setup.py", volumes=["/mnt/build",
"/var/cache/yum"], name="build-%s" % pkg)
...
resp = dclient.export(container)
tar = subprocess.Popen(["/bin/tar", "-x", "-C", imagedir],
stdin=subprocess.PIPE)
...
.SPEC para gerar o RPM
```

LXC – Images

```
 /opt/lxc/templates/nodejs/current# ll
total 16
drwxr-xr-x  2 root root 4096 Aug  6 09:52 inittab
drwxr-xr-x 16 root root 4096 Aug  6 10:57 rootfs
drwxr-xr-x  2 root root 4096 Aug  6 10:57 scripts
-rw-r--r--  1 root root  650 Aug  6 09:48 template.ini
 /opt/lxc/templates/nodejs/current#
```

```
 /opt/lxc/templates/nodejs/current/rootfs/usr/include/node# ll
total 732
-rw-r--r-- 1 root root   1853 Jul  9 19:41 android-ifaddrs.h
-rw-r--r-- 1 root root  21866 Jul  9 19:41 ares.h
-rw-r--r-- 1 root root    652 Jul  9 19:41 ares_version.h
-rw-r--r-- 1 root root   9149 Jul  9 19:41 common.gypi
-rw-r--r-- 1 root root   1955 Aug  5 09:03 config.gypi
drwxr-xr-x 2 root root   4096 Aug  6 10:57 libplatform
-rw-r--r-- 1 root root   8437 Jul  9 19:41 nameser.h
-rw-r--r-- 1 root root   5271 Jul  9 19:41 node_buffer.h
-rw-r--r-- 1 root root  17655 Jul  9 19:41 node.h
-rw-r--r-- 1 root root   8463 Jul  9 19:41 node_internals.h
-rw-r--r-- 1 root root   4057 Jul  9 19:41 node_object_wrap.h
-rw-r--r-- 1 root root   2634 Jul  9 19:41 node_version.h
drwxr-xr-x 2 root root   4096 Aug  6 10:57 openssl
-rw-r--r-- 1 root root   2929 Jul  9 19:41 pthread-fixes.h
-rw-r--r-- 1 root root   5682 Jul  9 19:41 smalloc.h
-rw-r--r-- 1 root root   7728 Jul  9 19:41 stdint-msvc2008.h
-rw-r--r-- 1 root root  52889 Jul  9 19:41 tree.h
-rw-r--r-- 1 root root   1615 Jul  9 19:41 uv-aix.h
-rw-r--r-- 1 root root   1641 Jul  9 19:41 uv-bsd.h
-rw-r--r-- 1 root root   3213 Jul  9 19:41 uv-darwin.h
-rw-r--r-- 1 root root   9399 Jul  9 19:41 uv-errno.h
-rw-r--r-- 1 root root  52883 Jul  9 19:41 uv.h
-rw-r--r-- 1 root root   1781 Jul  9 19:41 uv-linux.h
-rw-r--r-- 1 root root   1985 Jul  9 19:41 uv-sunos.h
-rw-r--r-- 1 root root   1497 Jul  9 19:41 uv-threadpool.h
-rw-r--r-- 1 root root  16166 Jul  9 19:41 uv-unix.h
-rw-r--r-- 1 root root   1687 Jul  9 19:41 uv-version.h
-rw-r--r-- 1 root root  31145 Jul  9 19:41 uv-win.h
-rw-r--r-- 1 root root  15939 Jul  9 19:41 v8config.h
-rw-r--r-- 1 root root   8807 Jul  9 19:41 v8-debug.h
-rw-r--r-- 1 root root 210110 Jul  9 19:41 v8.h
-rw-r--r-- 1 root root   1721 Jul  9 19:41 v8-platform.h
-rw-r--r-- 1 root root  19460 Jul  9 19:41 v8-profiler.h
-rw-r--r-- 1 root root    729 Jul  9 19:41 v8stdint.h
-rw-r--r-- 1 root root   1033 Jul  9 19:41 v8-testing.h
-rw-r--r-- 1 root root  14655 Jul  9 19:41 v8-util.h
-rw-r--r-- 1 root root  15508 Jul  9 19:41 zconf.h
-rw-r--r-- 1 root root  87883 Jul  9 19:41 zlib.h
 /opt/lxc/templates/nodejs/current/rootfs/usr/include/node#
```

LXC – Tools

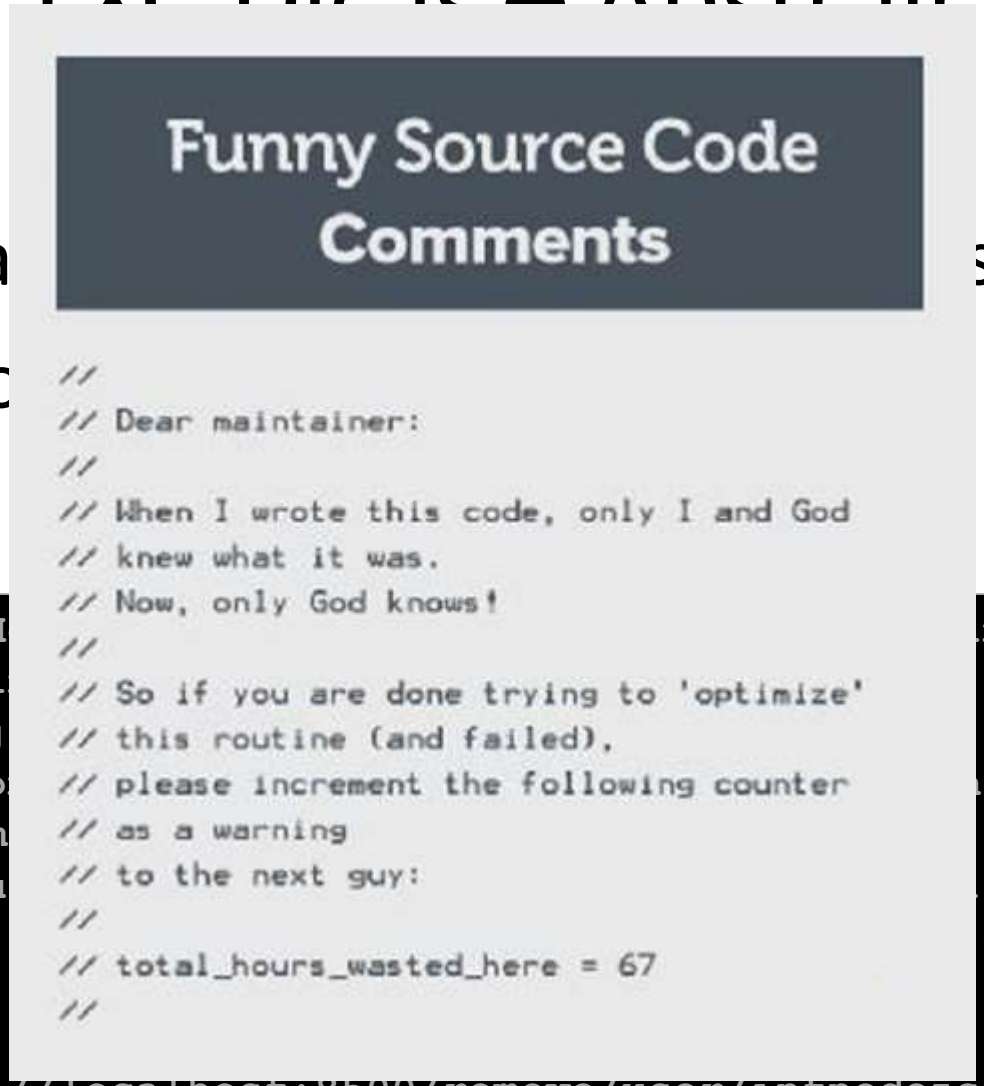
```
 ~# lxc-ls
b-u-intnodejs02-1  b-u-intnodejs03-1  b-u-intnodejs04-1  b-u-intnodejs05-1
 ~# lxc-info -n b-u-intnodejs02-1
Name:           b-u-intnodejs02-1
State:          RUNNING
PID:            64243
IP:             10.1.0.3
CPU use:        1.49 seconds
BlkIO use:      0 bytes
Memory use:     99.38 MiB
KMem use:       0 bytes
Link:           veth2
 TX bytes:      380.17 KiB
 RX bytes:      9.24 KiB
 Total bytes:   389.41 KiB
```

```
 ~# lxc-attach -n b-u-intnodejs02-1
bash-4.1# ps aux
PID   USER     TIME   COMMAND
    1 root       0:00 /sbin/init
    5 root       0:00 [route]
    6 root       0:00 [sh]
    8 root       0:00 /bin/bash /usr/sbin/monitor.sh
   49 nodejs     0:00 /usr/bin/node /opt/web/webapps//server.js
   50 nodejs     0:00 /usr/bin/python /usr/sbin/logit /opt/web/logs/nodejs.log
   94 root       0:00 /bin/bash
   97 root       0:00 sleep 3
   98 root       0:00 ps aux
bash-4.1#
```

```
 ~# lxc-
lxc-attach       lxc-clone        lxc-destroy      lxc-ls           lxc-start        lxc-unshare
lxc-autostart    lxc-config       lxc-execute      lxc-monitor      lxc-stop         lxc-usernsexec
lxc-cgroup       lxc-console      lxc-freeze       lxc-ps           lxc-top          lxc-wait
lxc-checkconfig  lxc-create       lxc-info         lxc-snapshot     lxc-unfreeze
```

LXC Dicas – Abstrair

- Esconda … salsicha
- Exemplo …




```
# curl -si -H "H                                    ntnodejs01" \
-H "X-User-Domai                                     pool03/pool03-
fs03/ac/intnodej
-H "X-Applicatio                                    stance: 1" \
-H "X-Memory-Lim
-H "X-Idle-Timou                                    " \
-H "X-Realhost:                                     84/


# curl -v "http://localhost:8500/remove/user/intnodejs01"
```

LXC Dicas – Geral

- Tenha seu kernel em debug mode em prod
- Kdump habilitado até ter certeza
- Testing ... Testing ... More Testing
- Overcommit

Avisos e Perguntas

Marcus Vinícius Soares

mvc_msoares@uolinc.com

http://www.uolhost.com.br/