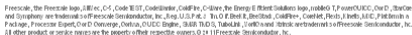


May 8, 2013



# Concepts

## Operating Systems

## Virtualization

## Linux Containers

## Usage

Schematics

Demo

Benchmarks

## Relevance

Spread

Use Cases

## QA



## IT trends

- ▶ More resources
  - ▶ Better hardware at lower costs
  - ▶ Higher standards for software quality
- ▶ More users
  - ▶ Contact with technology at an earlier age
  - ▶ Shared access to the same device
- ▶ Data consolidation
  - ▶ Data warehousing
  - ▶ Service unification
  - ▶ Differentiated access
- ▶ Increased flexibility
  - ▶ Versatile configuration
  - ▶ Focus on usability

# OS Recap

- ▶ Resources
  - ▶ CPU
  - ▶ Memory
  - ▶ Peripherals
- ▶ Structures
  - ▶ The scheduler
  - ▶ The pager
  - ▶ Filesystems
- ▶ The kernel
  - ▶ Handles hardware
  - ▶ Exposes capabilities
  - ▶ Manages resources

Figure: The Memory Pager

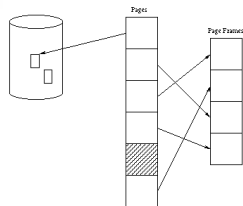
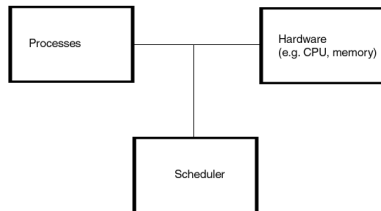


Figure: The Scheduler



# Virtualization

- ▶ Key aspects:
  - ▶ Simulation (of HW / SW)
  - ▶ Virtual machines
  - ▶ Autonomous computing
  - ▶ Utility computing
- ▶ Advantages:
  - ▶ Better resource usage
  - ▶ Lower running costs
  - ▶ Improved security
- ▶ Concerns:
  - ▶ Management
  - ▶ Isolation
  - ▶ Performance
  - ▶ Applicability



# OS-level Virtualization

- ▶ One host
- ▶ Multiple running OS instances
- ▶ Rootfs, system libs, binaries

OS instance = a process hierarchy

OS level virtualization = **partitioning** the process tree

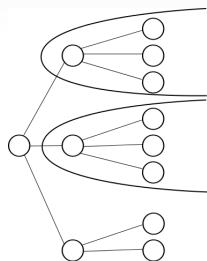
Advantage: **close to 0% performance overhead**

Flaw: **shared kernel**

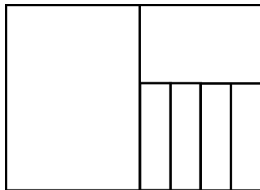


# Kernel Support

- ▶ Namespaces:
  - ▶ Abstract resources
  - ▶ Processes see the resource as their own
  - ▶ Isolation between namespaces



- ▶ Control Groups
  - ▶ Resource management among processes
  - ▶ Hierarchical support
  - ▶ Interaction with resource responsible structures:
    - ▶ Scheduler
    - ▶ Pager





# Sample Process Hierarchy

```
init(1)-+--dnsmasq(2162)
        |-klogd(2175)
        |-lxc-start(2964)---init(2966)----+--init(2972)
        |                                 |-sh(2971)
        |                                 '-syslogd(2969)
        |
        |
        |
        |-lxc-start(2974)---init(2976)----+--init(2982)
        |                                 |-sh(2981)
        |                                 '-syslogd(2979)
        |
        |
        |-netserver(2167)
        |-sh(2179)
        |-syslogd(2173)
        '-udevd(962)-+--udevd(1189)
                    '-udevd(1190)
```

# Process IDs

```
init(1)--dnsmasq(2162)
|-klogd(2175)
|-lxc-start(2964)---init(2966)(1)--init(2972)(7)
|
|                                     |-sh(2971)(6)
|                                     '-syslogd(2969)(4)
|
|
|
|
|-lxc-start(2974)---init(2976)(1)--init(2982)(7)
|
|                                     |-sh(2981)(6)
|                                     '-syslogd(2979)(4)
|
|
|
|-netserver(2167)
|-sh(2179)
|-syslogd(2173)
'-udevd(962)--udevd(1189)
        '-udevd(1190)
```

# Namespace Segregation

```
init(1)--dnsmasq(2162)
|-klogd(2175)
|-lxc-start(2964)--- init(2966)(1)--init(2972)(7)
|                                     |-sh(2971)(6)
|                                     '-syslogd(2969)(4)
|                                     PID Namespace 1
|
|-lxc-start(2974)--- init(2976)(1)--init(2982)(7)
|                                     |-sh(2981)(6)
|                                     '-syslogd(2979)(4)
|                                     PID Namespace 2
|
|-netserver(2167)
|-sh(2179)
|-syslogd(2173)
'-udevd(962)--udevd(1189)
                '-udevd(1190)
```



FreeScale, the FreeScale logo, 32/64, C-4, CodeWISIT, CodeLinker, ColdFire, C-Move, the Energy Efficient Solutions logo, iM8000 T, PowerQUICC, QorIQ, iZorQ and Symphony are trademarks of Freescale Semiconductor, Inc., Reg. U.S. Pat. & Tm. Off. BeeR, BeeStad, ColdFire, CoolNet, Hots, iJelly, iJED, P4tInn is a Package, Processor Expert, QorIQ Converge, QorIQ, QorIQ Engine, QorIQ TLDs, TuboLink, iV8000 and iV8000 are trademarks of Freescale Semiconductor, Inc. All other product or service names are the property of their respective owners. © 2014 Freescale Semiconductor, Inc.

# Filesystem Segregation

"chroot on steroids"

```
init(1)---dnsmasq(2162)
|-klogd(2175)      root: /var/lib/lxc/foo1/rootfs/
|-lxc-start(2964)--- init(2966)(1)++-init(2972)(7)
|
|                 |-sh(2971)(6)
|
|                 '-syslogd(2969)(4)
|
|                 PID Namespace 1
|
|                 root: /var/lib/lxc/foo2/rootfs/
|-lxc-start(2974)--- init(2976)(1)++-init(2982)(7)
|
|                 |-sh(2981)(6)
|
|                 '-syslogd(2979)(4)
|
|                 PID Namespace 2
|
|-netserver(2167)
|-sh(2179)
|-syslogd(2173)
'-udevd(962)++-udevd(1189)
      '-udevd(1190)
```

# CPU Partitioning

```

init(1)--dnsmasq(2162)
      |-klogd(2175)
,-----|-lxc-start(2964)-----
|      |      25%
|      |
|      |
|      |
|      |
1 core  |-lxc-start(2974)-----
|      |      75%
|      |
|      |
|      |
'-----|-----
        |-netserver(2167)
        |-sh(2179)
        |-syslogd(2173)
        '-udevd(962)--udevd(1189)
                '-udevd(1190)

```

# Demo

1. Start 2 containers
2. Check PIDs
3. Assign them a single core on the host
4. Balance CPU usage 25% - 75%

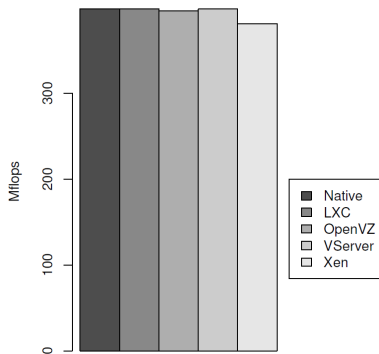


FreeScale, the FreeScale logo, ARM, Cx, CodeWIT, CodeLinker, ColdFire, C-More, the Energy Efficient Solutions logo, iM8000 T, PowerQUICC, QorIQ, iZaQ and Symphony are trademarks of FreeScale Semiconductor, Inc., Reg. U.S. Pat. & Tm. Off. Bell, Be!lStad, ColdFire+, CoolNet, Flex, iFlex, iMCC, P4tterns a Package, Processor Expert, QorIQ Converge, QorIQ, QVCC Engine, S3AR T4D5, TubuLink, iZaQ and iZaQ are trademarks of FreeScale Semiconductor, Inc. All other product or service names are the property of their respective owners. © 2011 Freescale Semiconductor, Inc.

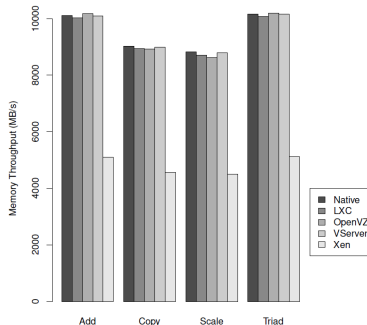
## System Performance

## CPU performance

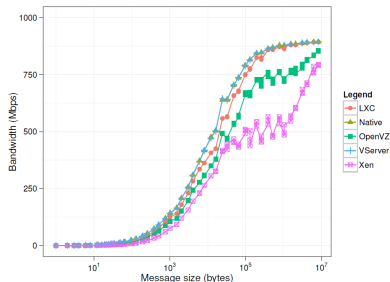
### Linpack



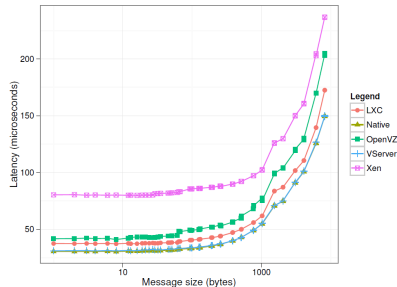
## Mem throughput



# Bandwidth NetPIPE



## Latency





# Isolation

PERFORMANCE ISOLATION FOR LU APPLICATION. THE RESULTS REPRESENT HOW MUCH THE APPLICATION PERFORMANCE IS IMPACTED BY DIFFERENT STRESS TESTS IN ANOTHER VM/CONTAINER. DNR MEANS THAT APPLICATION WAS NOT ABLE TO RUN.

	LXC	OpenVZ	VServer	Xen
CPU Stress	0	0	0	0
Memory	88.2%	89.3%	20.6%	0.9%
Disk Stress	9%	39%	48.8%	0
Fork Bomb	DNR	0	0	0
Network Receiver	2.2%	4.5%	13.6%	0.9%
Network Sender	10.3%	35.4%	8.2%	0.3%









Thank you!

Questions?



FreeScale, the FreeScale logo, i.MX, C-5, CodeWIT, CodeLinker, ColdFire, C-More, the Energy Efficient Solutions logo, iM800 T, PowerQUICC, QorD, iLibra and Symphony are trademarks of FreeScale Semiconductor, Inc., Reg. U.S. Pat. & Tm. Off. Bell, IL, Bedford, ColdFire+, ColdJet, Flex, iLibra, iM800, iM800M, iM800M-P, iM800M-P2, iM800M-P3, iM800M-P4, iM800M-P5, iM800M-P6, iM800M-P7, iM800M-P8, iM800M-P9, iM800M-P10, iM800M-P11, iM800M-P12, iM800M-P13, iM800M-P14, iM800M-P15, iM800M-P16, iM800M-P17, iM800M-P18, iM800M-P19, iM800M-P20, iM800M-P21, iM800M-P22, iM800M-P23, iM800M-P24, iM800M-P25, iM800M-P26, iM800M-P27, iM800M-P28, iM800M-P29, iM800M-P30, iM800M-P31, iM800M-P32, iM800M-P33, iM800M-P34, iM800M-P35, iM800M-P36, iM800M-P37, iM800M-P38, iM800M-P39, iM800M-P40, iM800M-P41, iM800M-P42, iM800M-P43, iM800M-P44, iM800M-P45, iM800M-P46, iM800M-P47, iM800M-P48, iM800M-P49, iM800M-P50, iM800M-P51, iM800M-P52, iM800M-P53, iM800M-P54, iM800M-P55, iM800M-P56, iM800M-P57, iM800M-P58, iM800M-P59, iM800M-P60, iM800M-P61, iM800M-P62, iM800M-P63, iM800M-P64, iM800M-P65, iM800M-P66, iM800M-P67, iM800M-P68, iM800M-P69, iM800M-P70, iM800M-P71, iM800M-P72, iM800M-P73, iM800M-P74, iM800M-P75, iM800M-P76, iM800M-P77, iM800M-P78, iM800M-P79, iM800M-P80, iM800M-P81, iM800M-P82, iM800M-P83, iM800M-P84, iM800M-P85, iM800M-P86, iM800M-P87, iM800M-P88, iM800M-P89, iM800M-P90, iM800M-P91, iM800M-P92, iM800M-P93, iM800M-P94, iM800M-P95, iM800M-P96, iM800M-P97, iM800M-P98, iM800M-P99, iM800M-P100, iM800M-P101, iM800M-P102, iM800M-P103, iM800M-P104, iM800M-P105, iM800M-P106, iM800M-P107, iM800M-P108, iM800M-P109, iM800M-P110, iM800M-P111, iM800M-P112, iM800M-P113, iM800M-P114, iM800M-P115, iM800M-P116, iM800M-P117, iM800M-P118, iM800M-P119, iM800M-P120, iM800M-P121, iM800M-P122, iM800M-P123, iM800M-P124, iM800M-P125, iM800M-P126, iM800M-P127, iM800M-P128, iM800M-P129, iM800M-P130, iM800M-P131, iM800M-P132, iM800M-P133, iM800M-P134, iM800M-P135, iM800M-P136, iM800M-P137, iM800M-P138, iM800M-P139, iM800M-P140, iM800M-P141, iM800M-P142, iM800M-P143, iM800M-P144, iM800M-P145, iM800M-P146, iM800M-P147, iM800M-P148, iM800M-P149, iM800M-P150, iM800M-P151, iM800M-P152, iM800M-P153, iM800M-P154, iM800M-P155, iM800M-P156, iM800M-P157, iM800M-P158, iM800M-P159, iM800M-P160, iM800M-P161, iM800M-P162, iM800M-P163, iM800M-P164, iM800M-P165, iM800M-P166, iM800M-P167, iM800M-P168, iM800M-P169, iM800M-P170, iM800M-P171, iM800M-P172, iM800M-P173, iM800M-P174, iM800M-P175, iM800M-P176, iM800M-P177, iM800M-P178, iM800M-P179, iM800M-P180, iM800M-P181, iM800M-P182, iM800M-P183, iM800M-P184, iM800M-P185, iM800M-P186, iM800M-P187, iM800M-P188, iM800M-P189, iM800M-P190, iM800M-P191, iM800M-P192, iM800M-P193, iM800M-P194, iM800M-P195, iM800M-P196, iM800M-P197, iM800M-P198, iM800M-P199, iM800M-P200, iM800M-P201, iM800M-P202, iM800M-P203, iM800M-P204, iM800M-P205, iM800M-P206, iM800M-P207, iM800M-P208, iM800M-P209, iM800M-P210, iM800M-P211, iM800M-P212, iM800M-P213, iM800M-P214, iM800M-P215, iM800M-P216, iM800M-P217, iM800M-P218, iM800M-P219, iM800M-P220, iM800M-P221, iM800M-P222, iM800M-P223, iM800M-P224, iM800M-P225, iM800M-P226, iM800M-P227, iM800M-P228, iM800M-P229, iM800M-P230, iM800M-P231, iM800M-P232, iM800M-P233, iM800M-P234, iM800M-P235, iM800M-P236, iM800M-P237, iM800M-P238, iM800M-P239, iM800M-P240, iM800M-P241, iM800M-P242, iM800M-P243, iM800M-P244, iM800M-P245, iM800M-P246, iM800M-P247, iM800M-P248, iM800M-P249, iM800M-P250, iM800M-P251, iM800M-P252, iM800M-P253, iM800M-P254, iM800M-P255, iM800M-P256, iM800M-P257, iM800M-P258, iM800M-P259, iM800M-P260, iM800M-P261, iM800M-P262, iM800M-P263, iM800M-P264, iM800M-P265, iM800M-P266, iM800M-P267, iM800M-P268, iM800M-P269, iM800M-P270, iM800M-P271, iM800M-P272, iM800M-P273, iM800M-P274, iM800M-P275, iM800M-P276, iM800M-P277, iM800M-P278, iM800M-P279, iM800M-P280, iM800M-P281, iM800M-P282, iM800M-P283, iM800M-P284, iM800M-P285, iM800M-P286, iM800M-P287, iM800M-P288, iM800M-P289, iM800M-P290, iM800M-P291, iM800M-P292, iM800M-P293, iM800M-P294, iM800M-P295, iM800M-P296, iM800M-P297, iM800M-P298, iM800M-P299, iM800M-P300, iM800M-P301, iM800M-P302, iM800M-P303, iM800M-P304, iM800M-P305, iM800M-P306, iM800M-P307, iM800M-P308, iM800M-P309, iM800M-P310, iM800M-P311, iM800M-P312, iM800M-P313, iM800M-P314, iM800M-P315, iM800M-P316, iM800M-P317, iM800M-P318, iM800M-P319, iM800M-P320, iM800M-P321, iM800M-P322, iM800M-P323, iM800M-P324, iM800M-P325, iM800M-P326, iM800M-P327, iM800M-P328, iM800M-P329, iM800M-P330, iM800M-P331, iM800M-P332, iM800M-P333, iM800M-P334, iM800M-P335, iM800M-P336, iM800M-P337, iM800M-P338, iM800M-P339, iM800M-P340, iM800M-P341, iM800M-P342, iM800M-P343, iM800M-P344, iM800M-P345, iM800M-P346, iM800M-P347, iM800M-P348, iM800M-P349, iM800M-P350, iM800M-P351, iM800M-P352, iM800M-P353, iM800M-P354, iM800M-P355, iM800M-P356, iM800M-P357, iM800M-P358, iM800M-P359, iM800M-P360, iM800M-P361, iM800M-P362, iM800M-P363, iM800M-P364, iM800M-P365, iM800M-P366, iM800M-P367, iM800M-P368, iM800M-P369, iM800M-P370, iM800M-P371, iM800M-P372