ARCHIVING SOFTWARE SYSTEMS

APPROACHES TO PRESERVE COMPUTATIONAL CAPABILITIES

Todd King

Why Archive Software

Developed software is a large investment and worthy of preservation.

But…
The long term preservation of software presents some challenges.

- Software often requires a specific technology stack to operate.
  - This can include software, operating systems and hardware dependencies.

Preservation of Capabilities (the old way)

On an archive horizon of 100 years we currently:

- Maintain ancient hardware long past its typical viability.
- Save the source code.
  - While this can preserve details of the implementation and algorithms, it may not be possible to reproduce the technology stack needed to compile and run the resulting applications.

Preservation of Capabilities (the new way)

Future Forward

Any method of design (and or creation) that is capable of being upgraded with relative ease.

Think in terms of

- Platforms, virtual machines, containers

Platforms

- The ubiquitous platforms of today will likely be supported in the future.
  - Intel 8086

- The host platforms may evolve in a variety of ways, but virtual machines can remain constant.
  - Intel 8086 -> 80286 -> 80386 -> Pentium -> Dual-Core -> i7

Virtual Machines

A virtual machine (VM) is a software implementation of a machine (for example, a computer) that executes programs like a physical machine.

Examples:

- Personal
  - VMware, KVM, VirtualBox, Parallels
- IaaS
  - OpenStack, Amazon (EC2,S3), Microsoft, Google

Containers

- Light-weight virtual machines.
- Make use of resource isolation (CPU, memory, block I/O, network, etc.) and separate namespaces, and to completely isolate the application's view of the operating system.

Example:

(www.docker.com)

Container Host Stack

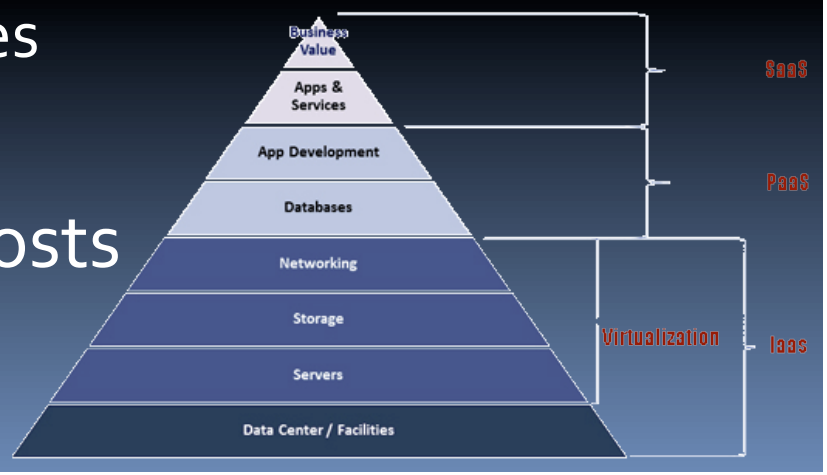
Allow

- External port mapping
- External file system mounting

Containers require a Linux Host

- Uses cgroups and namespaces

Can run on hardware or VM hosts



Cloning and Preservation

- Docker "images" are tar files
- Easily imported and exported.

PDS Archive

- Allow Containers – Policy change?
- Require x86 builds
- Define appropriate metadata (PDS Label)
  - Use Product_Software/Software_Binary
  - Define permissible values for "software_type" (Currently none are defined)
    - Docker_Image
  - Define permissible values for "software_format_type" (Currently none are defined)
    - "tar" or "Docker"

Thank You!