

OS Circulation environment "Trusted HTTP-FUSE Xenoppix"

<http://unit.aist.go.jp/itri/knoppix/http-fuse/index-en.html>

<http://unit.aist.go.jp/itri/knoppix/xen/index-en.html>

Kuniyasu Suzaki⁽¹⁾, Toshiki Yagi⁽¹⁾, Kengo Iijima⁽¹⁾,
Junichi Tsukamoto⁽²⁾, Megumi Nakamura⁽³⁾, Seiji Munetoh⁽³⁾

- National Institute of Advanced Industrial Science and Technology
- University of Tokyo
- IBM Japan

Contents

- What is OS Circular?
 - Virtual Bootloader "Xenoppix"
 - Globalized Virtual Disk "Trusted HTTP-FUSE CLOOP"
- Current Status
- Conclusions

Purpose of OS Circular

- OS Migration System for anonymous Users
 - Boot OSes on anonymous PC without installation.
- FLOZ (Free Live OS Zoo) has same concept.
 - http://www.oszoo.org/wiki/index.php/Free_Live_OS_Zoo
 - "QEMU" runs OSes(Linux, Minix, Plan9, OpenSolaris, etc) on the server. The GUI is transferred by VNC on Client PC.
 - **Server Centric System**
 - Bad response because the server exits in Italy
 - Guest OS has no network service for security & resource

OS Circular

- OS Circular is **Client Centric System** and utilize virtualization technology.
 - Virtual Bootloader + Globalized Virtual Disks
 - Client PC boots with the Virtual Bootloader and get OS images via Globalized Virtual Disks
 - Current PC has powerful CPU, much Memory, and wide Network.
 - It prevents server bottleneck. The network is managed by the user.
 - OS images are maintained on the server and they are execute on a Client PC.
 - Same idea is implemented on Collective[NSDI'05] and Internet Suspend/Resume[WMCSA'02]
 - **What is different?**
 - *OS Circular utilizes existing, in-expensive and popular Infrastructure.*

Virtual BootLoader

- Virtual Bootloader is consisted of "Virtual Machine software" and host OS which runs Virtual Machine
- Virtual Machine offers an Abstraction Layer which offers a common PC environment on any PCs.
 - Full Virtualization enables us to use normal installer and security management for Guest OS.
 - The abstraction is used for OS migration.
 - SouIPAD [Mobisys'05], VAT of Collective[NSDI'05], Internet Suspend/Resume[WMCSA'02] use VMware.
- Host OS supports real device drivers
 - KNOPIX is used on SouIPAD and VAT of Collective, because KNOPIX automatically detects available devices and loads the appropriate Linux drivers.

Xenoppix as Virtual Bootloader

- Xenoppix (2005.09~) = Xen + KNOPPIX
 - Old Xenoppix uses Para-Virtualization and runs Plan9 & NetBSD. Current Xenoppix runs Para & Full Virtualization
 - Full Virtualization of Xen "HVM" prepares a common PC environment on anonymous PC.
 - HVM requires virtualization instructions; Intel VT or AMD-V.
 - KNOPPIX works on Domain0 of Xen (as the host OS).
 - Xen has no device drivers and utilize the drivers of the OS on Domain0. "Autoconfig" of KNOPPIX detects devices and setup drives.
 - So, Xen and KNOPPIX is the best marriage.

Globalized Virtual Disk

- Virtual Disk is Block Level Abstraction.
- The requirement for globalization.
 - Versioning
 - Partial update
 - Globalization
 - World Wide Deployment
 - Network/Storage Transparent for mobile computing
 - Handle network (dis/re)-connection
 - Security
 - OS itself should maintained by Security Software
 - Virtual disks have to keep validness of contents

Versioning

- 2 type of versioning
 - Non-persistent versioning
 - "undo" of operations during OS lives.
 - persistent versioning
 - "rollback" of OS image
- Trusted HTTP-FUSE CLOOP has same versioning of Venti[USENIX'02] of Plan9.
 - A SHA-1 hash of the data acts as the address of the data.
 - Data blocks are write once and aren't removed. They make ideal for permanent or backup storage.
 - Unfortunately Venti requires special protocol for file system and limits the scalability.
- Trusted HTTP-FUSE CLOOP saves data blocks to files.
 - The feature makes to utilize HTTP to distribute them.

Globalization (1/2)

- Virtual disk are desired to be shared on Internet.
- We must consider network latency and extensibility of server.
 - NBD, iSCSI and iFCP look fine, but they require special software on server. This feature prevents scalability.
- Policy of Trusted HTTP-FUSE CLOOP
 - Don't require special daemons and special ports on servers. The requirement is file distribution only.
 - For World Wide Deployment, utilize existing infrastructure.
 - Web hosting services, Mirror servers, Proxies

Globalization (2/2)

- We have to consider dis-connection of network for mobile computing.
 - StatelessLinux of RHEL5 takes snapshot on a local storage as cache.
- The blocks of Trusted HTTP-FUSE CLOOP are cached on a local storage.
 - While applications don't access extra blocks, it works fine.

Security

- Basically security management is independent of virtual disk, because security of OS should be managed by security software or package manager.
- The commitment of virtual disk is to keep uncontaminated contents.
- Trusted HTTP-FUSE CLOOP check the validness of contents when it mapped to a virtual disk.

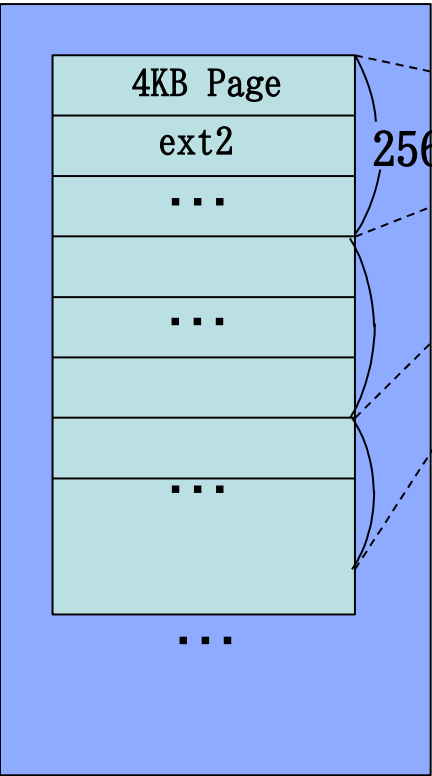
Trusted HTTP-FUSE CLOOP (1/2)

- Original block device is split by 256KB and compressed by zlib. Each data is saved to a block file.
- Block files are managed by "index" file which works as a header of CLOOP.
- Block file name is a SHA1 value of its contents.
 - If there is a same contents blocks, they are held together a same name file and **reduce total file space**.
 - The basic idea is resemble to "Venti of Plan9"
- Block files are reconstructed to a CLOOP file by FUSE wrapper.
 - FUSE is a User-land File System.
 - <http://fuse.sf.net>
- Each block file is measured with the SHA1 file name when it mapped to CLOOP.

Trusted HTTP-FUSE CLOOP (2/2)

- When a file is updated or created on an original block device, the relevant block files are newly created with new SHA1 name. The "index" file are also renewed.
 - Old block files are reusable.
- HTTP for file deliver
 - Most popular and well designed.
 - Web hosting is inexpensive.
 - 80 port is usually opened.
 - Other network block devices use special port which is usually closed.
- Block files are network/storage transparent.
 - **Block files are cached and reused on local storage.**
 - If necessary block files are stored in a local storage, network connection is not necessary.

Block Device

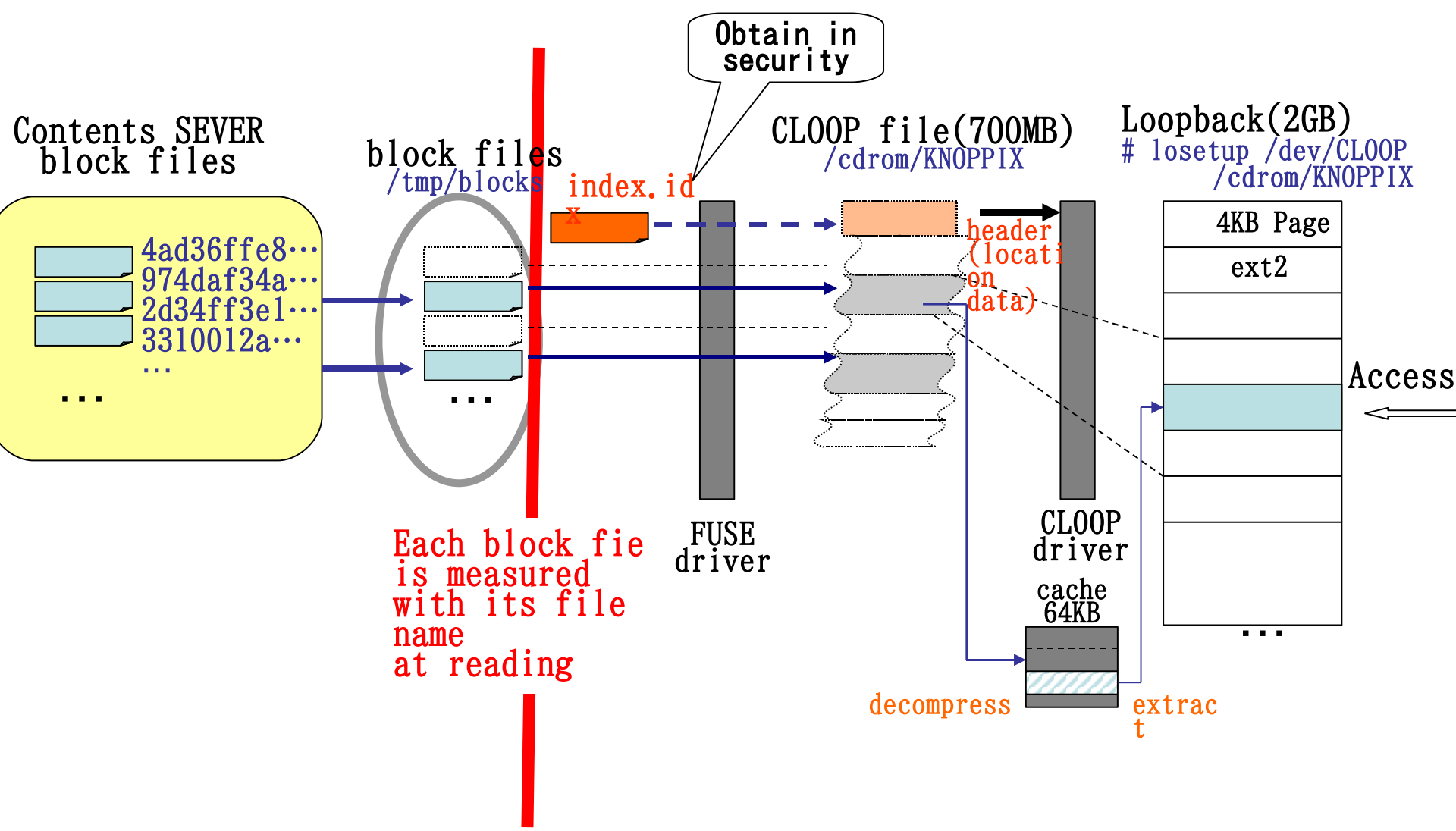


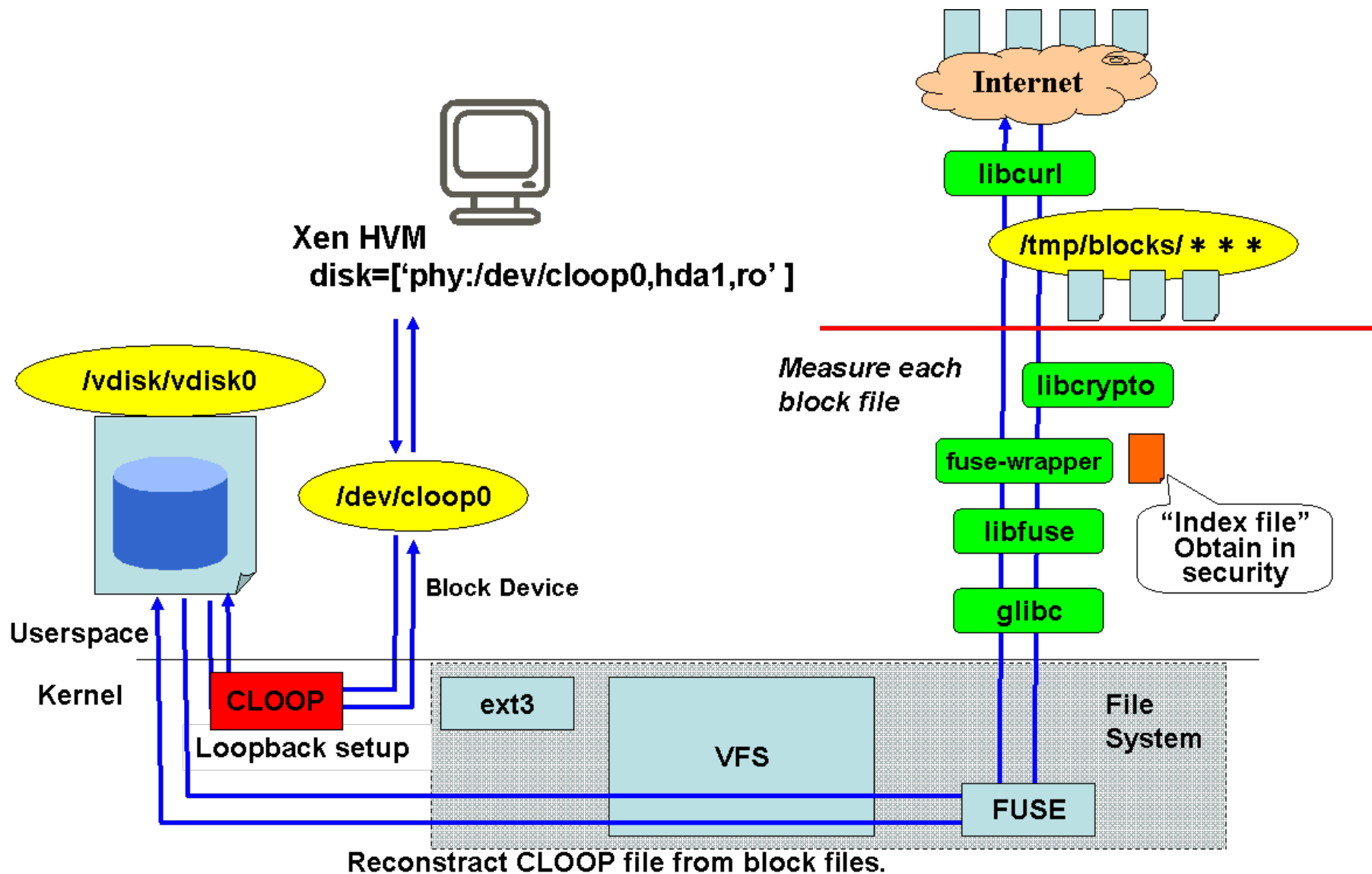
index and block files

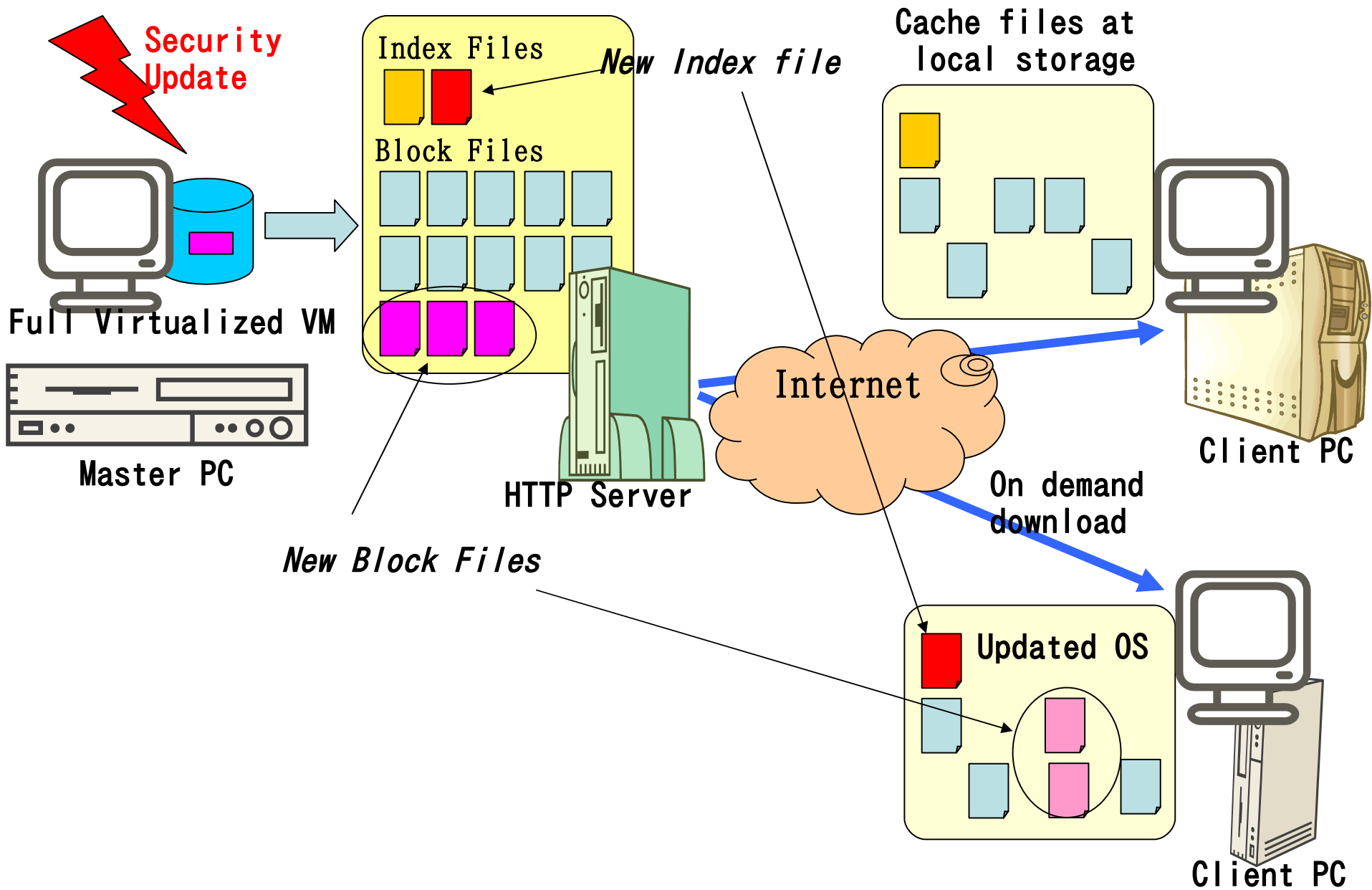


The block files are re-constructed as a virtual disk with HTTP-FUSE CLOOP

compressed by zlib







Weak point of Trusted HTTP-FUSE CLOOP

- Trusted HTTP-FUSE CLOOP is Vulnerable for Network Latency
 - On demand & small data request makes narrow bandwidth on long latency.
- Solutions
 - Shorten Latency
 - Find the nearest one among candidates download sites
 - Widen Bandwidth
 - Pre download necessary block files
 - Software RAID

Care for network latency (Find near download site)

– Client Side Solution: netselect

- netselect measures the latency of candidate servers and finds nearest one.

- <http://www.worldvisions.ca/~apenwarr/netselect/>

– Server Side Solution: DNS-Balance

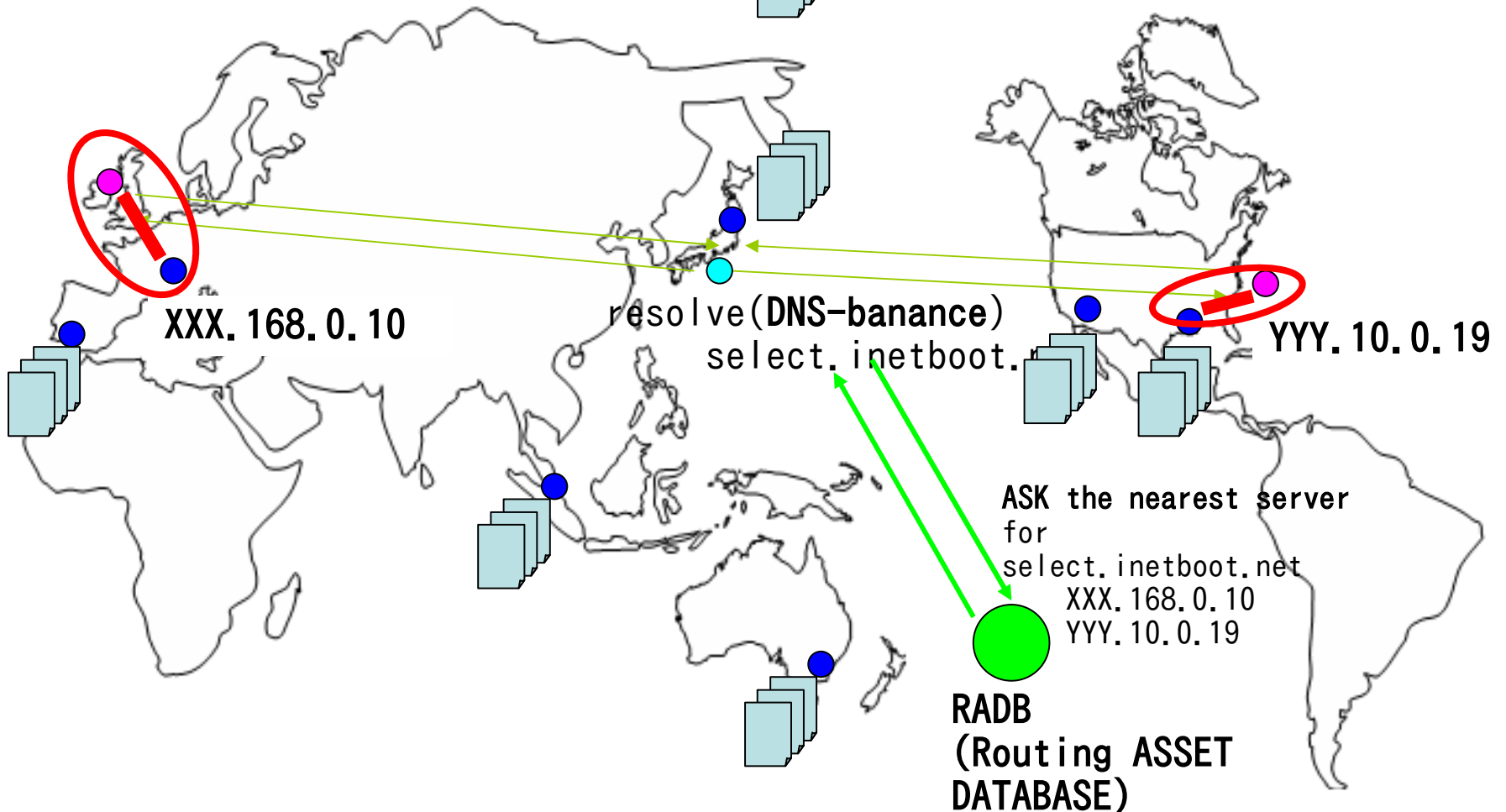
- DNS-Balance is a name resolver which suggests near candidate servers **with routing information offered by RADB.net**

- http://openlab.jp/dns_balance/dns_balance.html

DNS-Balance

DNS request
Resolve **select.inetboot.net**
to shorten *latency*

- Client
- Web server for HTTP-FUSE Xenoppix
- DNS server: **ns.inetboot.net**
- Block files



Care for network latency

(Multiple download connections)

- DLAHEAD (DownLoad AHEAD)
 - Preparation
 - Take a profile of downloaded block files at boot time.
 - At boot time
 - The block files are downloaded in advance with extra download connections. (Default is 4)
 - Effective and easy to implement
- Software RAID using MD
 - Mirroring(RAID1) virtual block device makes fault tolerant and broadband with multiple access.

Current Status of OS Circular

- Virtual Boot Loader
 - Xenoppix (Xen 3.0.3 + KNOPPIX 5.0.1)
 - Debian package
 - Trusted HTTP-FUSE CLOOP
 - Setup script for OS Circular
- OS Images is obtained by Trusted HTTP-FUSE CLOOP
 - Debian GNU/Linux
 - Secure updated with "apt-get" command
 - FreeBSD

Current Download sites

- Web Hosting Service is reasonable.
 - 5GB/ mount from \$10



Discussions

- Live OS update
 - Intel vPro, LUCOS[VEE' 06]
 - We have plan to integrate IDS and IPS on OS Circular.
- BlkTap of Xen [USENIX' 05]
 - Xen offers an interface of software device "BlkTap".
 - We want to make an offload engine of Trusted HTTP-FUSE CLOOP.

Conclusions

- OS Circular is Client Centric OS Migration System which is consisted of Virtual Bootloader "Xenoppix" and Global Virtual Disk "Trusted HTTP-FUSE CLOOP".
- The current target OSes are Debian GNU/Linux and FreeBSD.
 - Debian is updated by "apt-get" automatically.
- We hope OS Circular easy-to-try OS environment.