# HTTP-FUSE CLOOP with Software RAID and DNS-Balance for Embedded Linux

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

Jun Kanai[1], **Kuniyasu Suzaki** [2],
Toshiki Yagi[2], Mitaro Namiki[1]
1) Tokyo University of Agriculture and Technology
2) National Institute of Advanced Industrial Science and Technology

# Outline

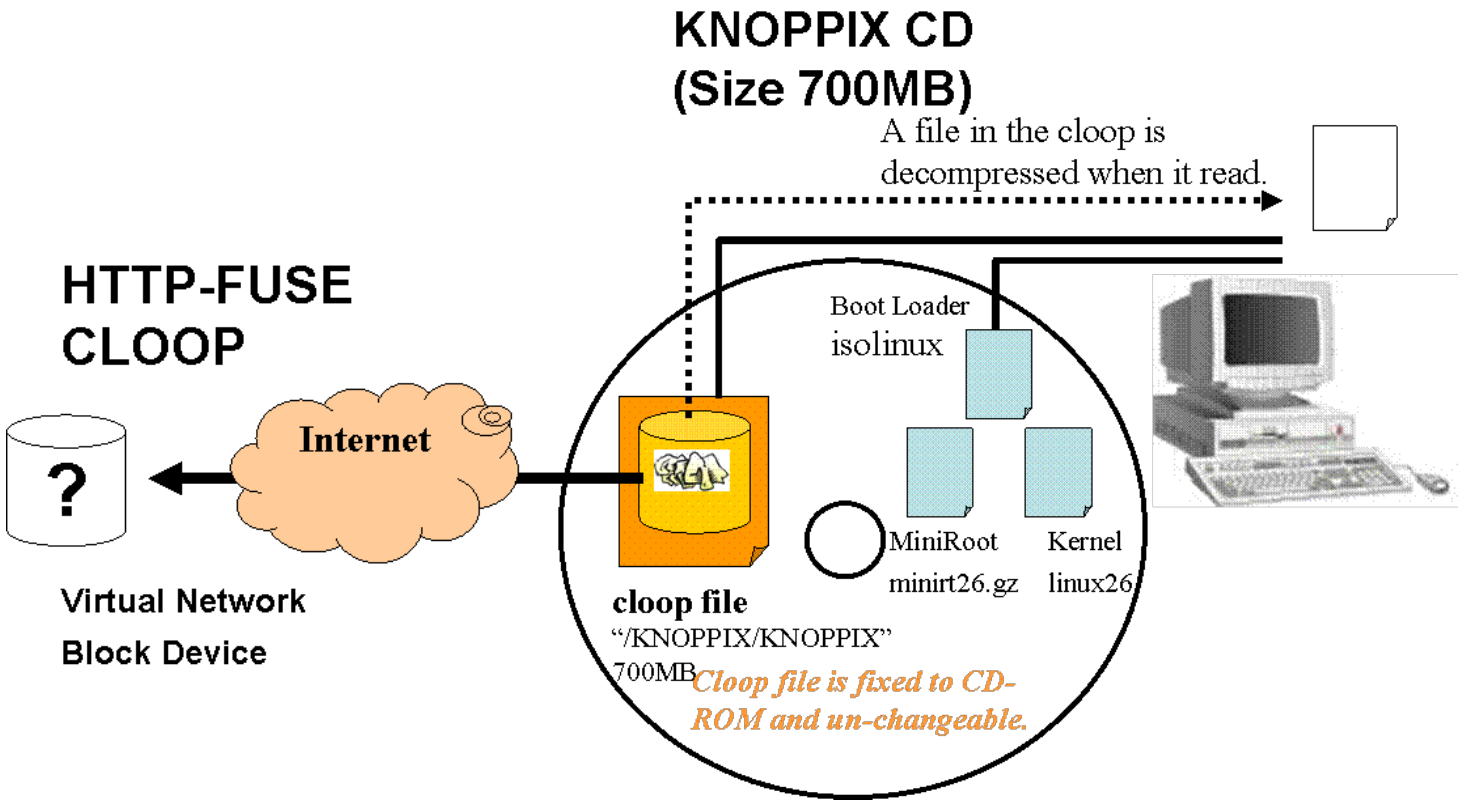
- What is HTTP-FUSE CLOOP?
- New Optimization
  - Software RAID
  - DNS Balance
- Customization for embedded Linux
  - Reduce memory copy
  - Replace de-compress algorithm
  - Performance on SH4, ARM9
- Conclusions
- Demo

# What is HTTP-FUSE CLOOP?

- HTTP-FUSE CLOOP is network block device which is designed for Internet Thin Client.[OLS'06, LCA'07 MinConf Virtualization]
  - Based on CLOOP (Compressed Loopback Device) which is used on 1CD Linux “KNOPPIX”

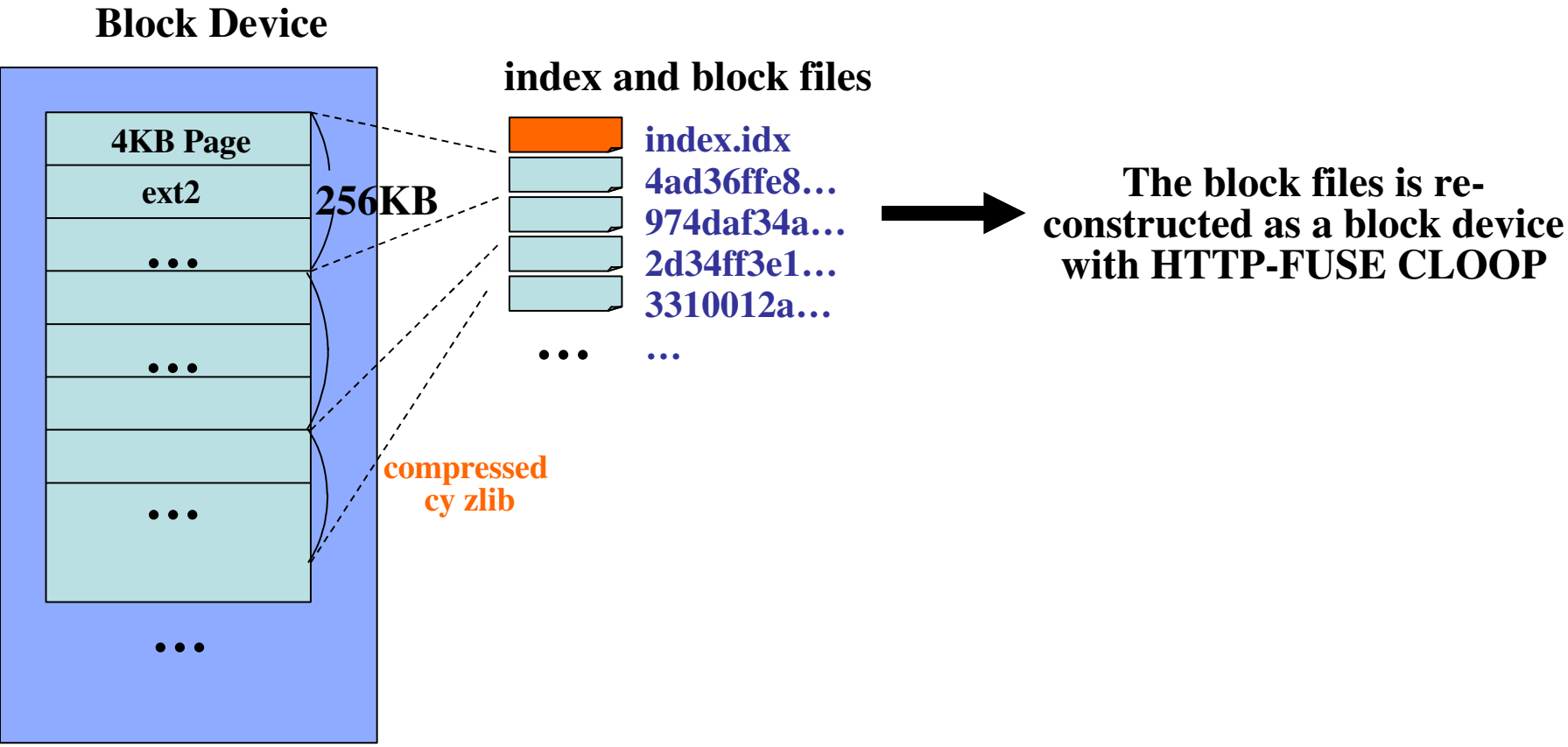
# CLOOP

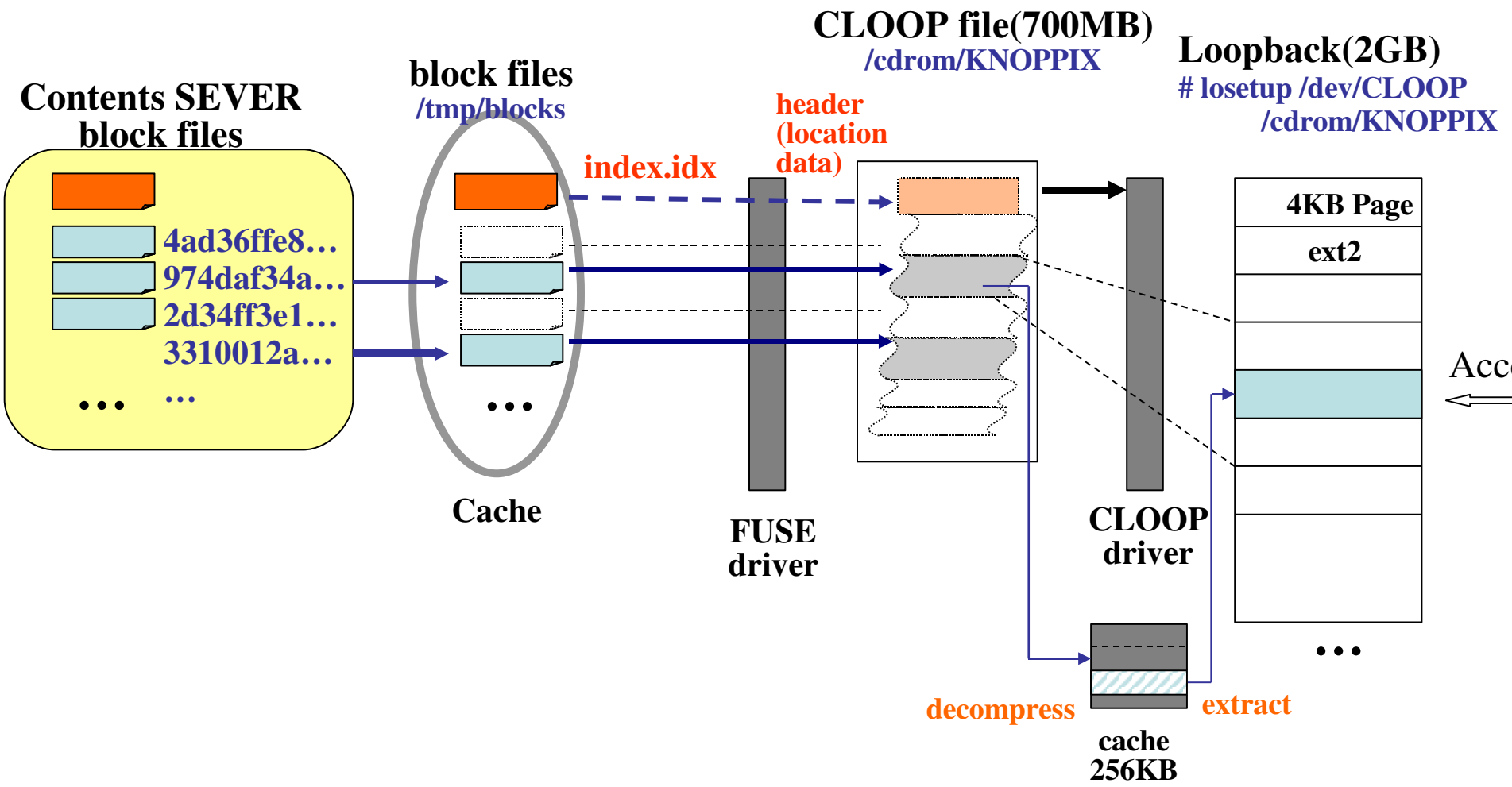
- CLOOP enables to pack 2.0GB contents (Root File System) to 700MB CD-ROM.
  - Each 64KB block is compressed by zlib and save a loop file.
- HTTP-FUSE CLOOP gets rid of COOP file form CD-ROM and exposed it to Internet.



# HTTP-FUSE CLOOP

- HTTP-FUSE CLOOP is made from existing block device.The block device is split, compressed. Each split block is saved to each block file.
  - Current split size is 256KB.
- The block files are managed by index file, which includes location information.
  - index file works as a header of CLOOP file.
- Each block file name is a MD5 value of its contents.
  - If there is a same contests 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


Block Device
4KB Page
ext2
256KB
index and block files
index.idx
4ad36ffe8…
974daf34a…
2d34ff3e1…
3310012a…
…
compressed
cy zlib
The block files is re-constructed as a block device with HTTP-FUSE CLOOP


Contents SEVER
block files
4ad36ffe8…
974daf34a…
2d34ff3e1…
3310012a…
block files
/tmp/blocks
Cache
index.idx
header
(location
data)
FUSE
driver
CLOOP file(700MB)
/cdrom/KNOPPIX
CLOOP
driver
Loopback(2GB)
# losetup /dev/CLOOP
/cdrom/KNOPPIX
4KB Page
ext2
decompress
extract
cache
256KB

# Box of HTTP-FUSE CLOOP

- The request for server is HTTP only.
- Small Linux Box can be the server of HTTP-FUSE CLOOP.

- These Linux Boxes are not so powerful. But they are bound up by software RAID.

USL-5P
HTTP-FUSE-KNOPPIX-BOX
（SH4-266MHz/64MBMem/CF/
100MbpsLAN/150g）

# Weak point of HTTP-FUSE CLOOP

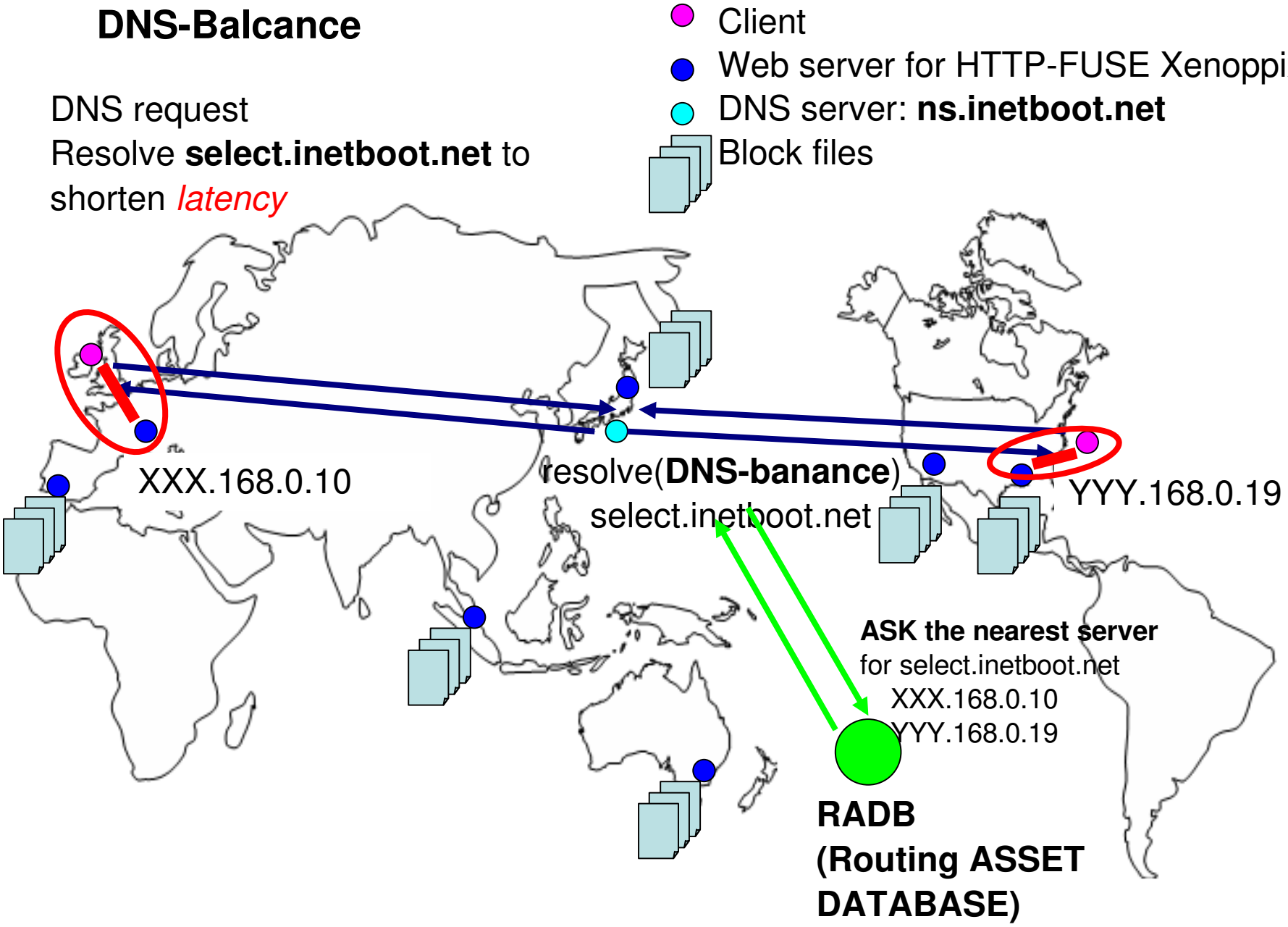
- Vulnerable for Network Latency
  - HTTP-FUSE CLOOP have to download small block files on demand (sequentially). It takes network latency severely.
- New Solution
  - DNS Balance
    - Find good server for a client PC.
  - Software RAID
    - Widen bandwidth with multiple access.

# DNS Balcance

- Name resolver (Load balancer) developed by Yokota [18th **IEEE** Int. Conf. On Advanced Information Networking and Application]
  - http://openlab.jp/dns_balance/dns_balance.html
    - Written by Ruby
  - On Internet
    - Find near server for Client.
    - Rrouting information is offered by RADB.net
  - On LAN
    - DNS Balance is used for Load balancing.
    - HTTP-FUSE CLOOP uses stateless HTTP. So it enables to replace server dynamically.

# DNS-Balcance

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

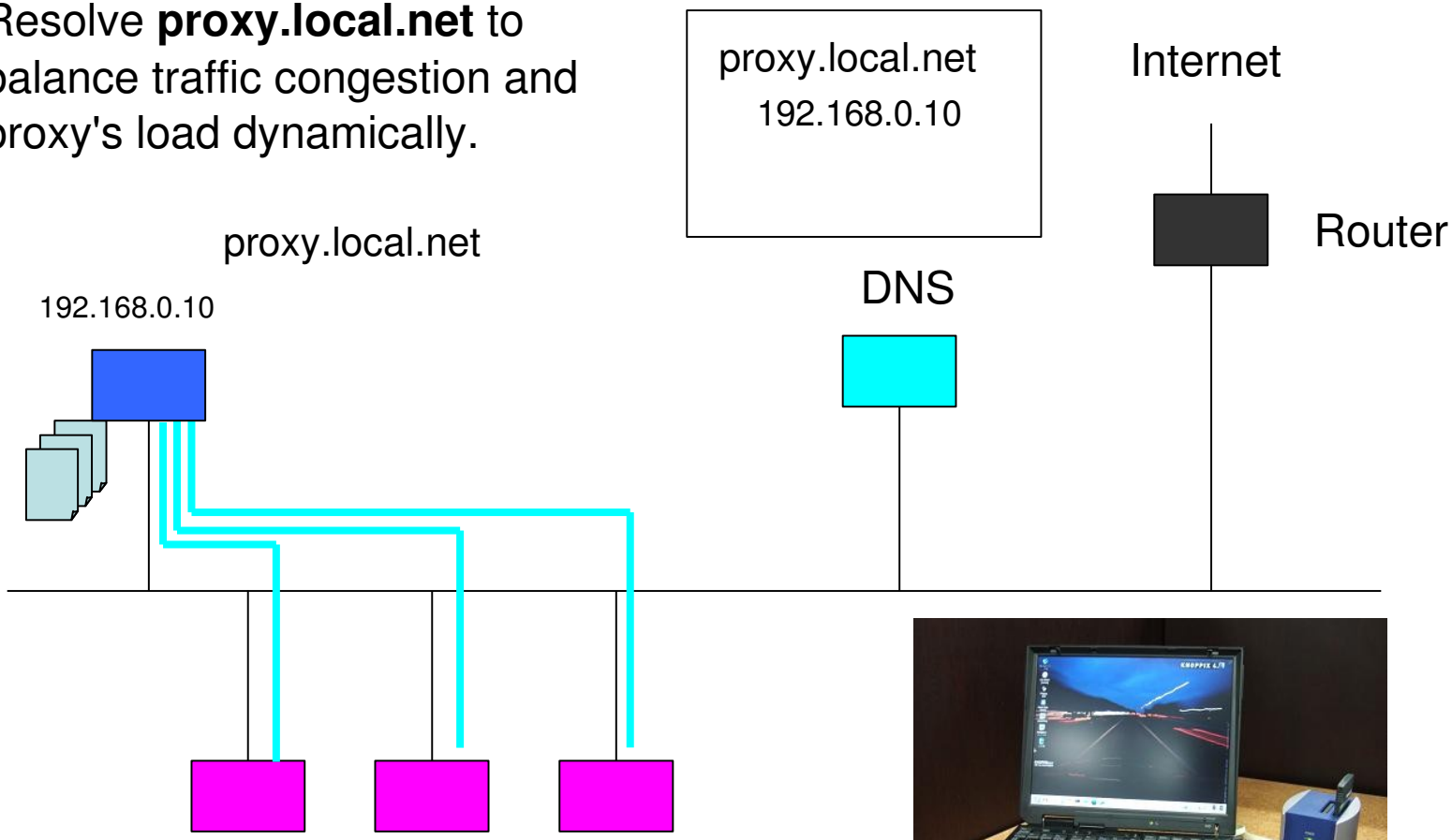


# Current HTTP sites

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

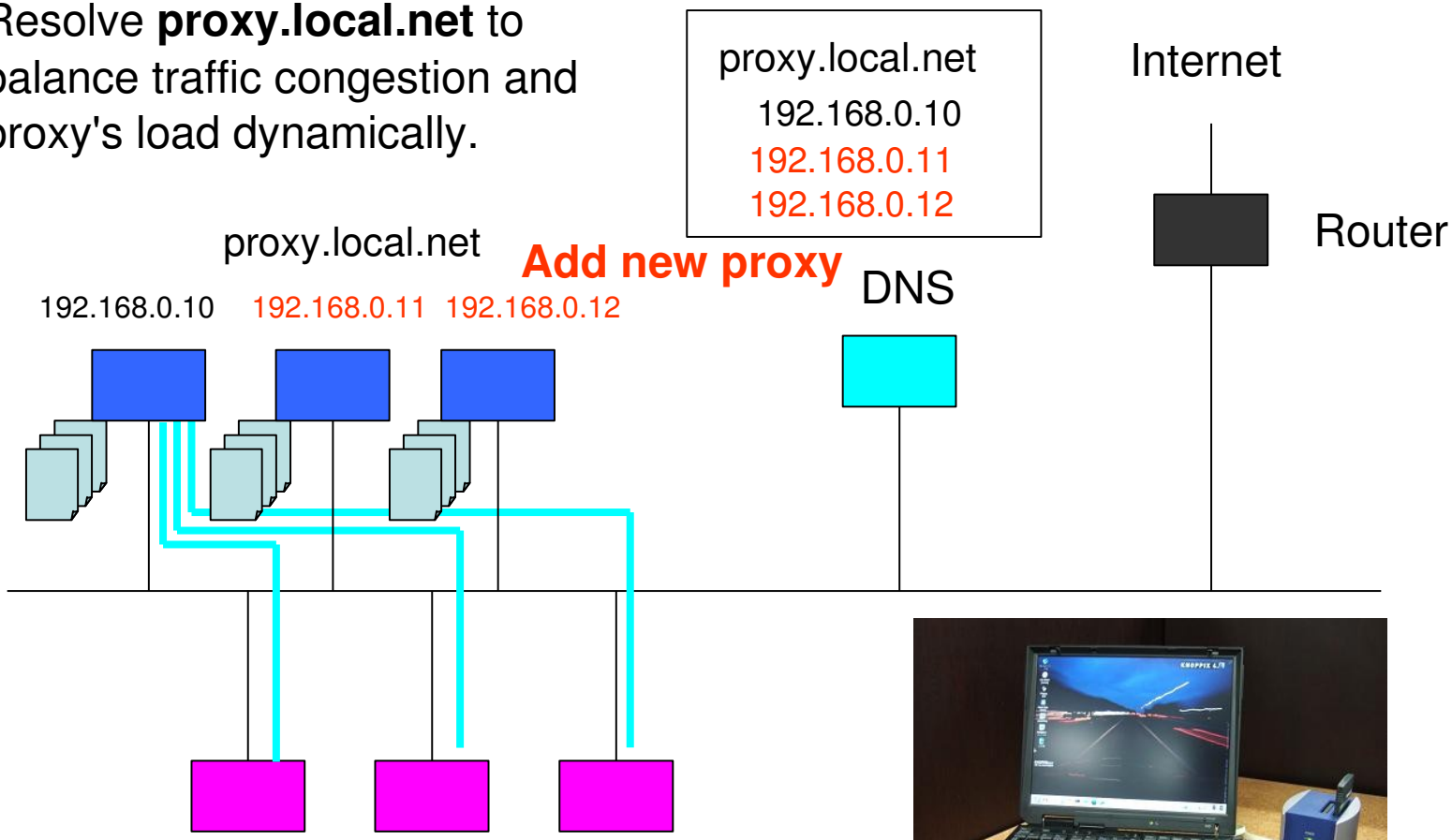
# LAN Level

DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



# LAN Level

DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



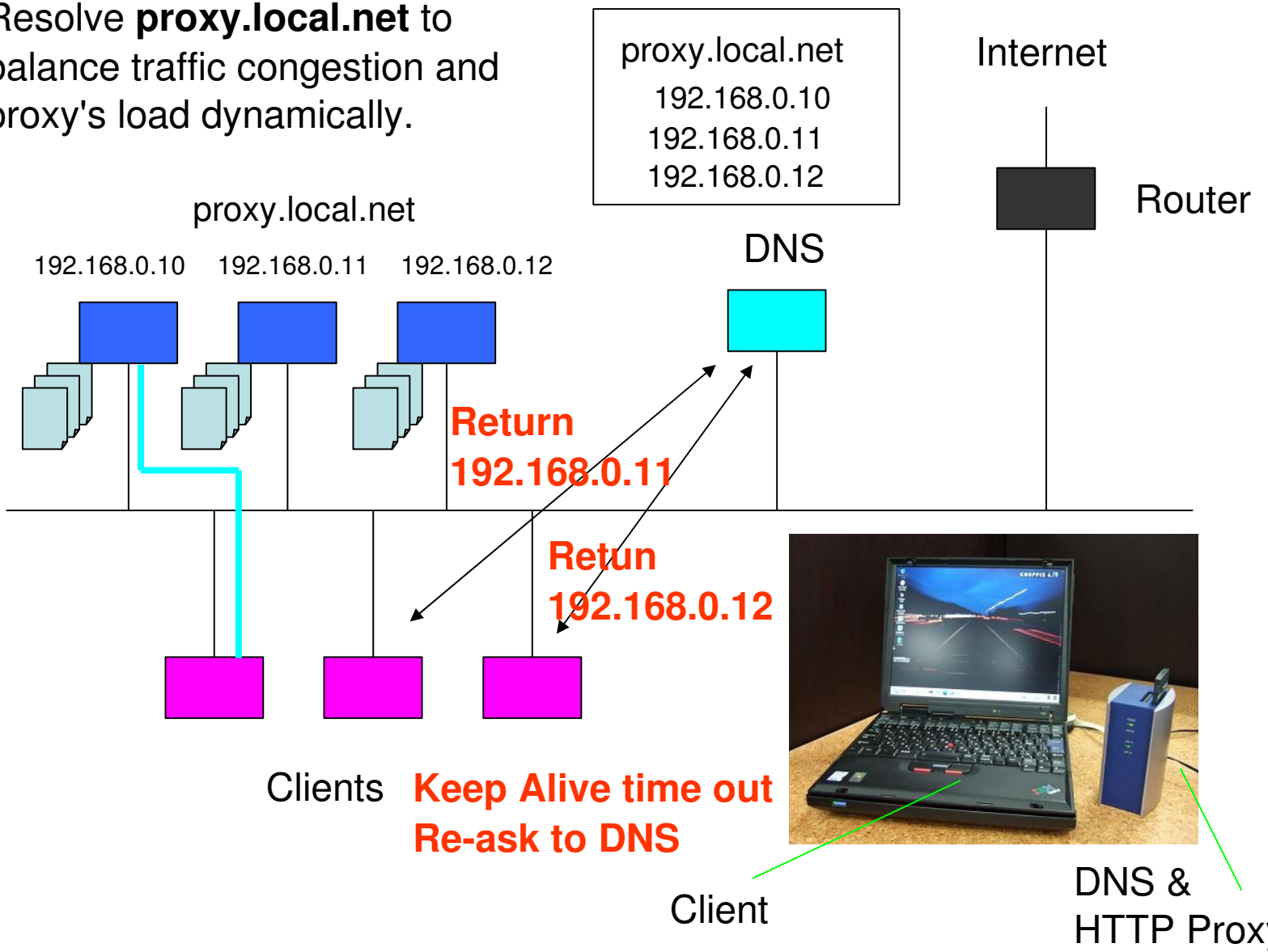



Client

DNS &
HTTP Proxy Cache

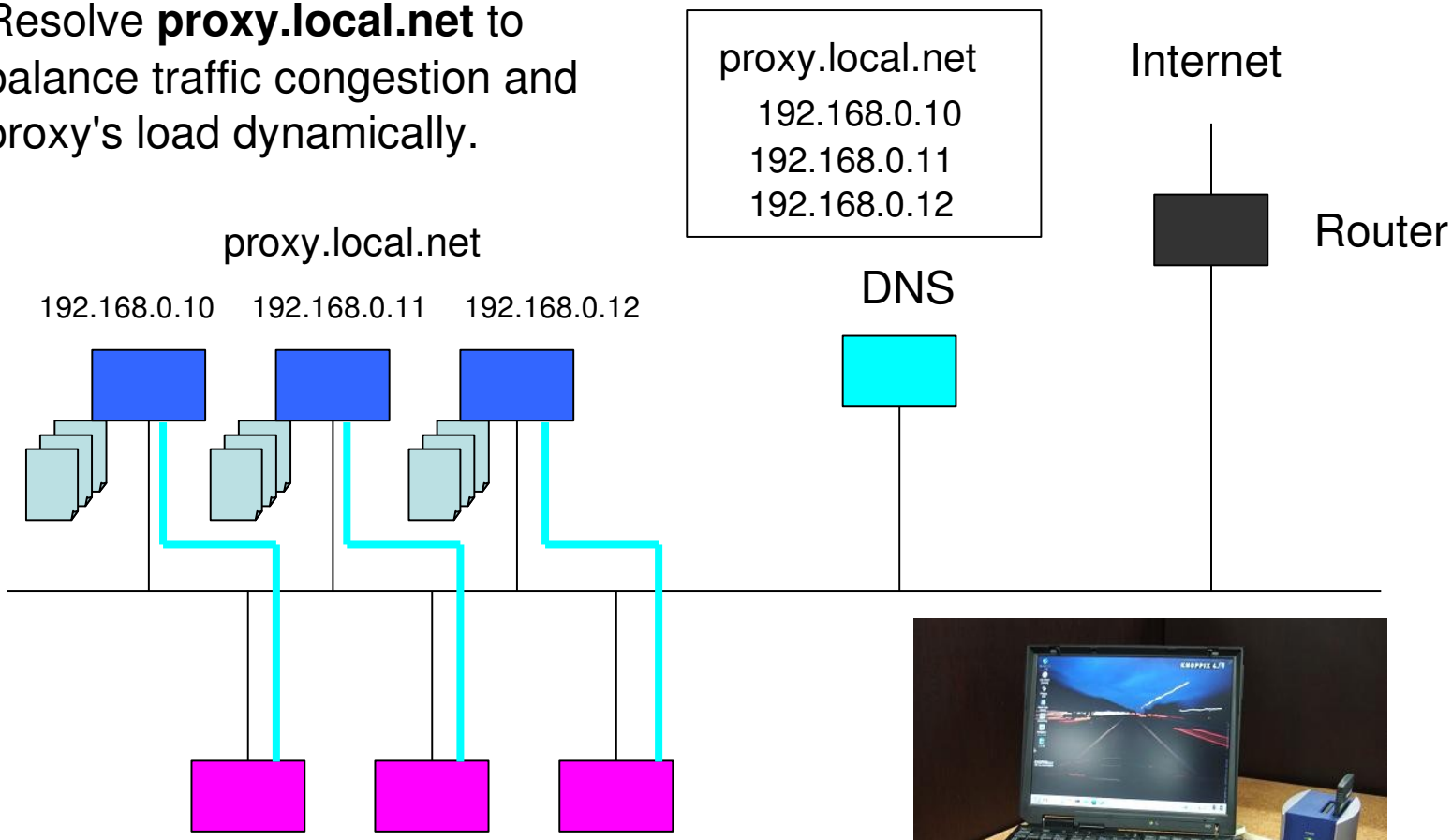
# LAN Level

DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



# LAN Level

DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



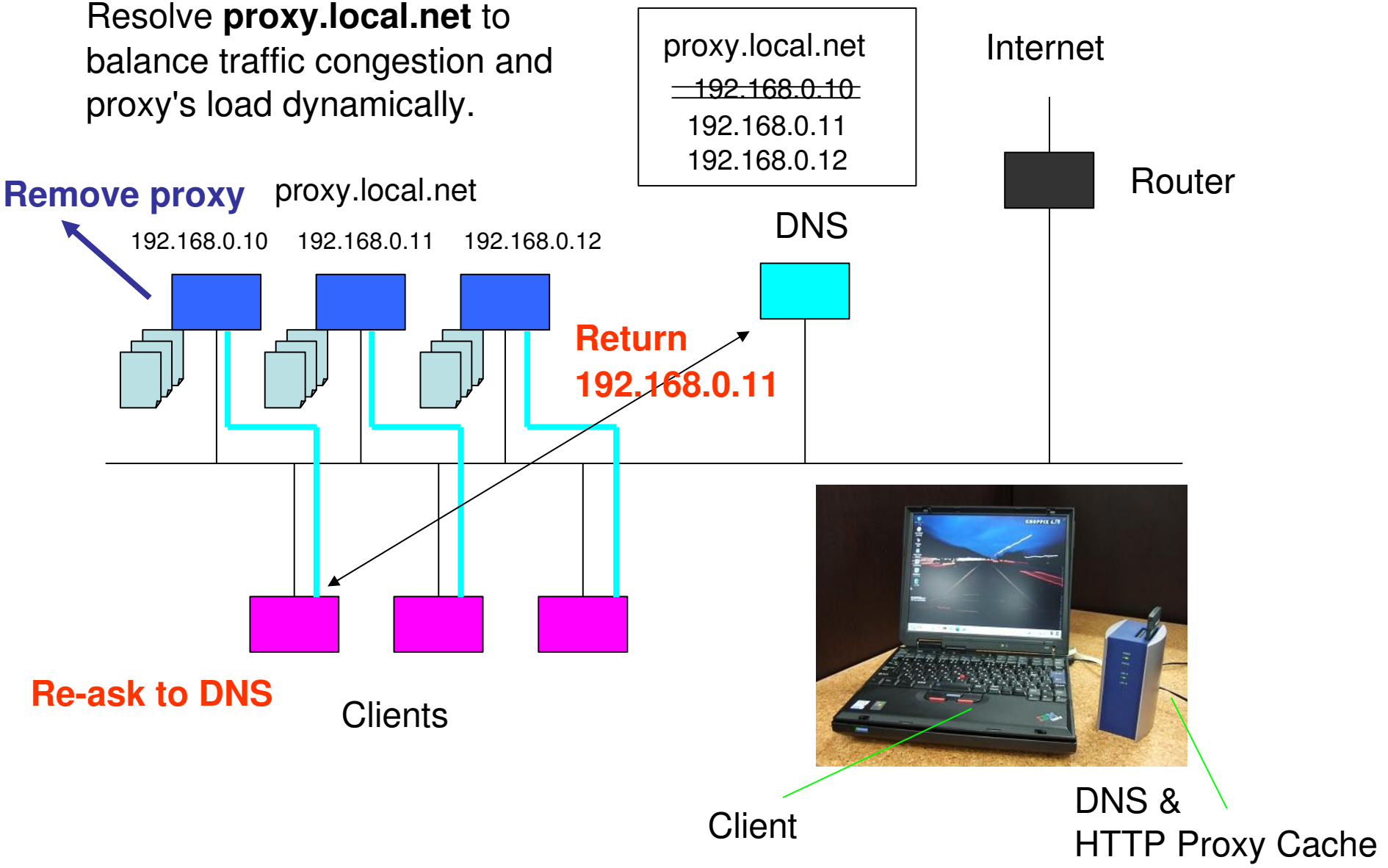



Client

DNS &
HTTP Proxy Cache

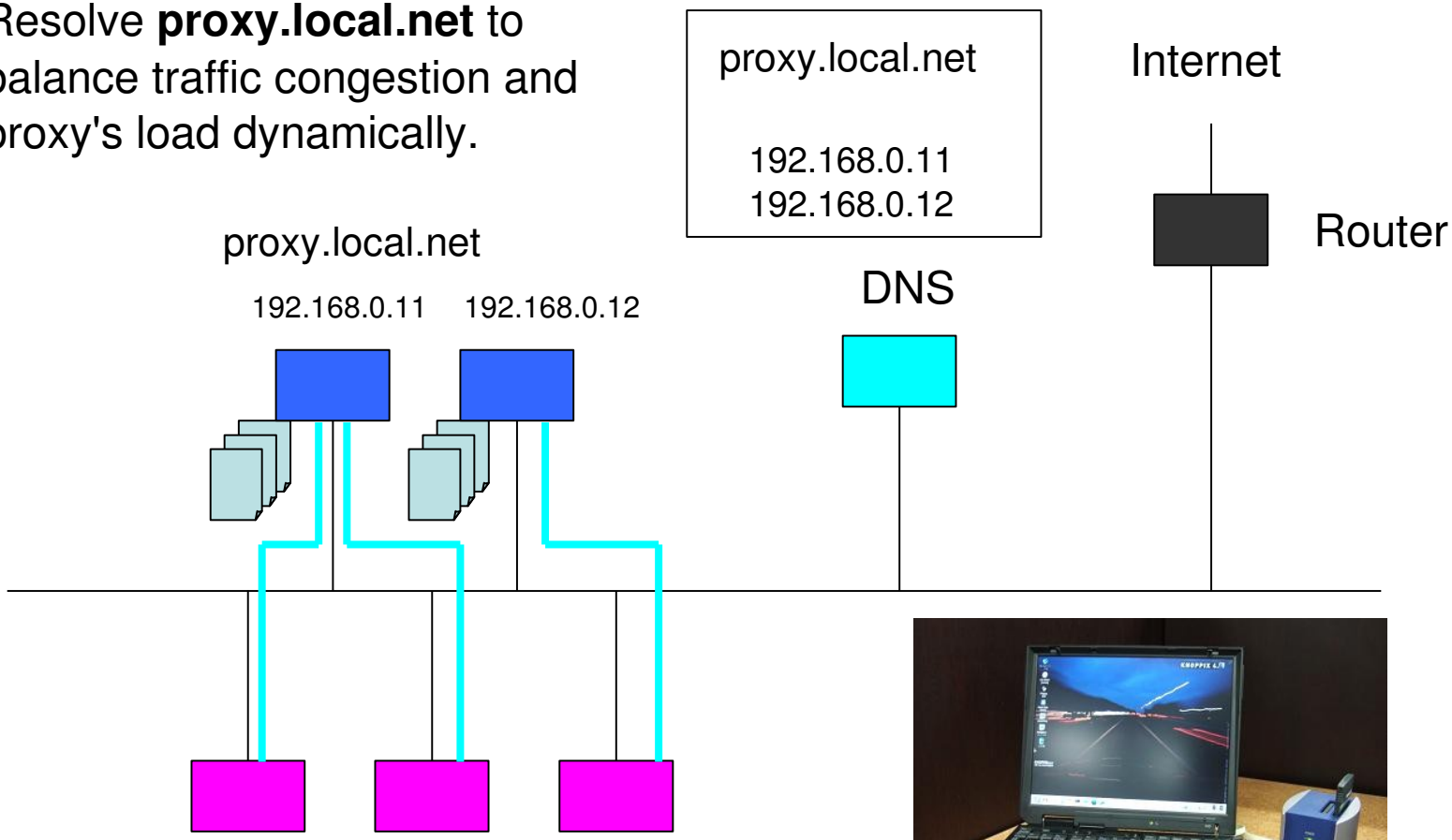
# LAN Level

DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



# LAN Level

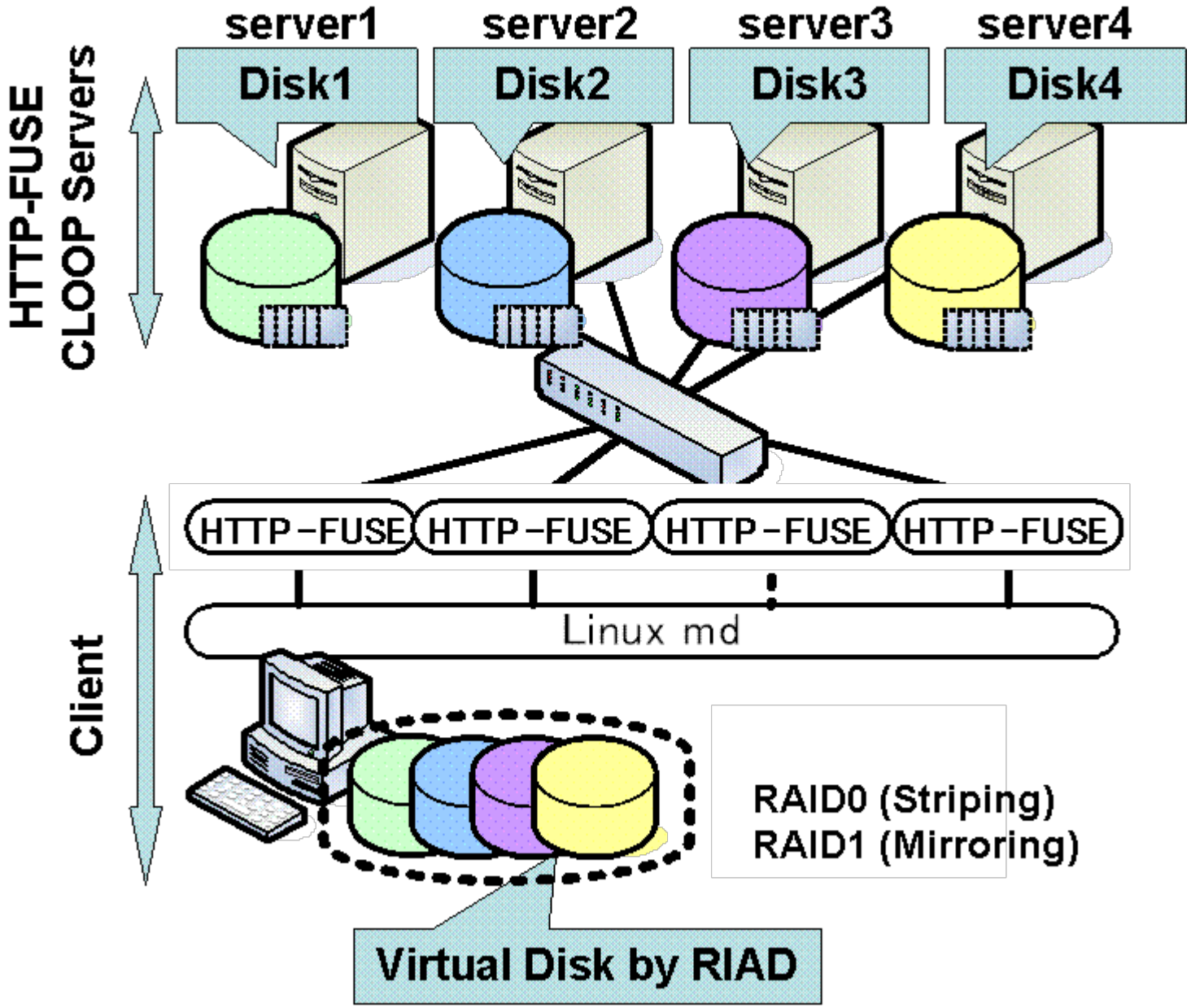
DNS request
Resolve **proxy.local.net** to balance traffic congestion and proxy's load dynamically.



# Software RAID

- Weakpoint of HTTP-FUSE CLOOP
  - accepts only one request
  - doesn't have recovery procedure
  - because it is a software device and doesn't assume any troubles.

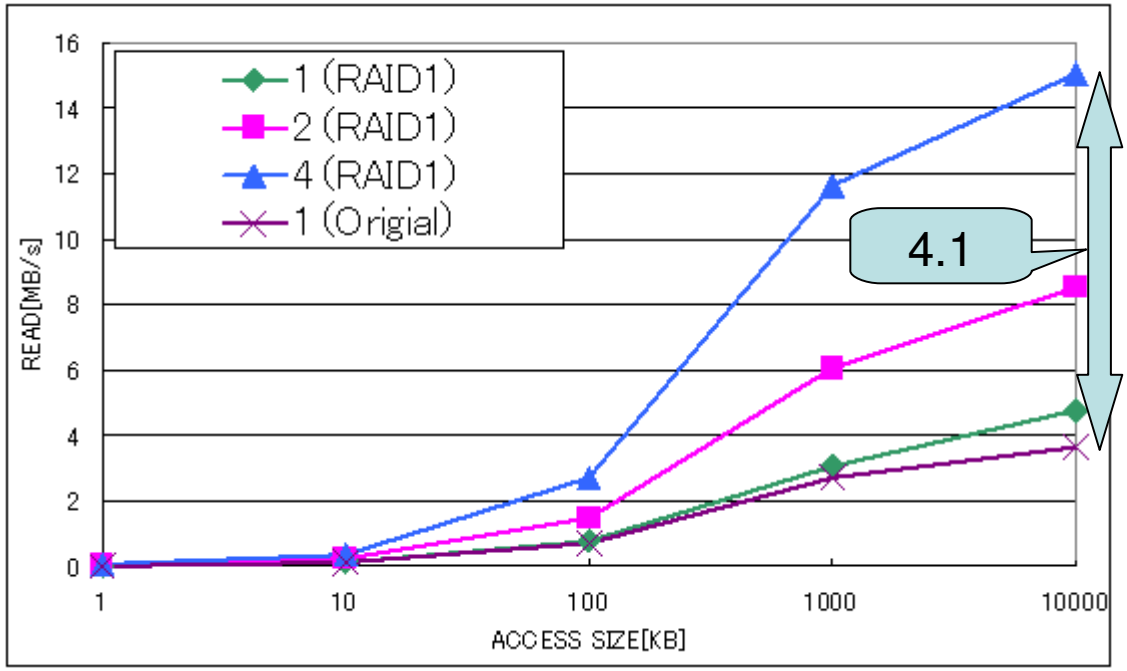
- MD (Multiple Disks) driver aggregate some block disks to a virtual disk as a RAID.
  - We applied MD to HTTP-FUSE CLOOP and enable to accept multiple requests.


server1
server2
server3
server4
Disk1
Disk2
Disk3
Disk4
HTTP-FUSE
CLOOP Servers
HTTP-FUSE
HTTP-FUSE
HTTP-FUSE
HTTP-FUSE
Linux md
Client
RAID0 (Striping)
RAID1 (Mirroring)
Virtual Disk by RIAD

# Performance of Software RAID

The bandwidth becomes
4.14 times wider than original.

Super linear is caused by
access overlap of MD



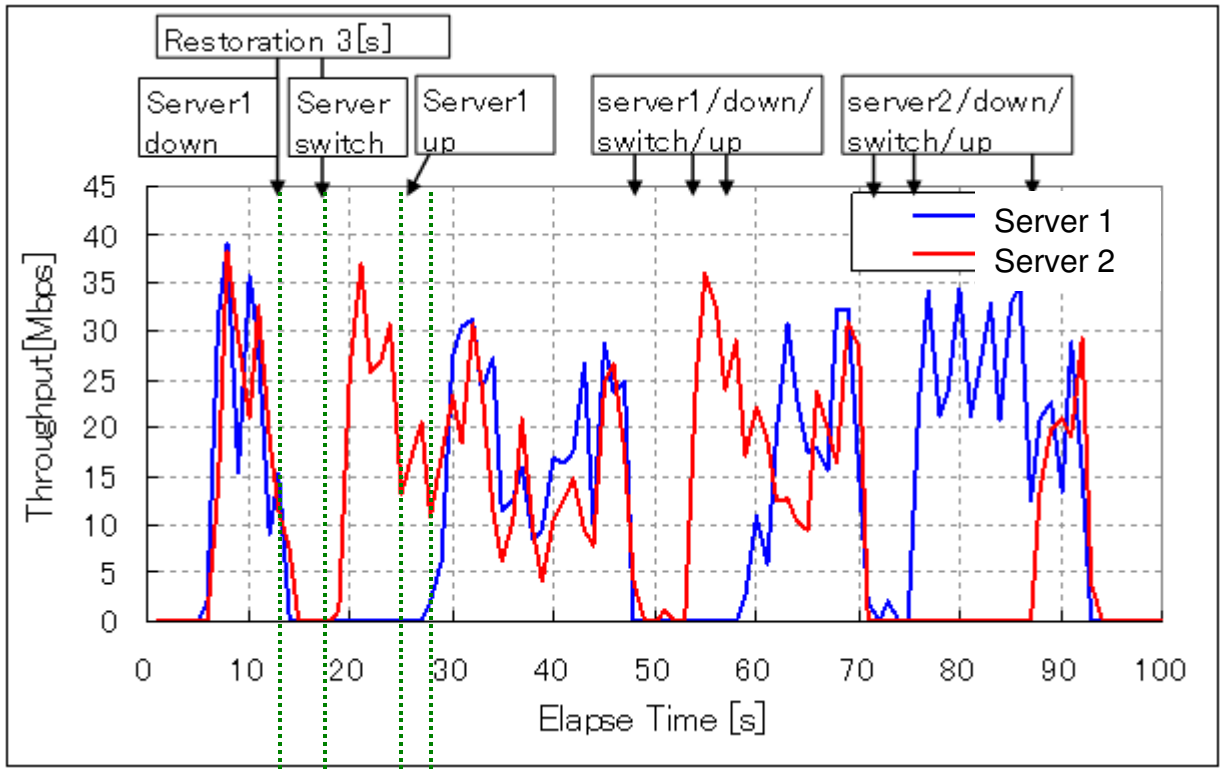
# Fault-tolerant

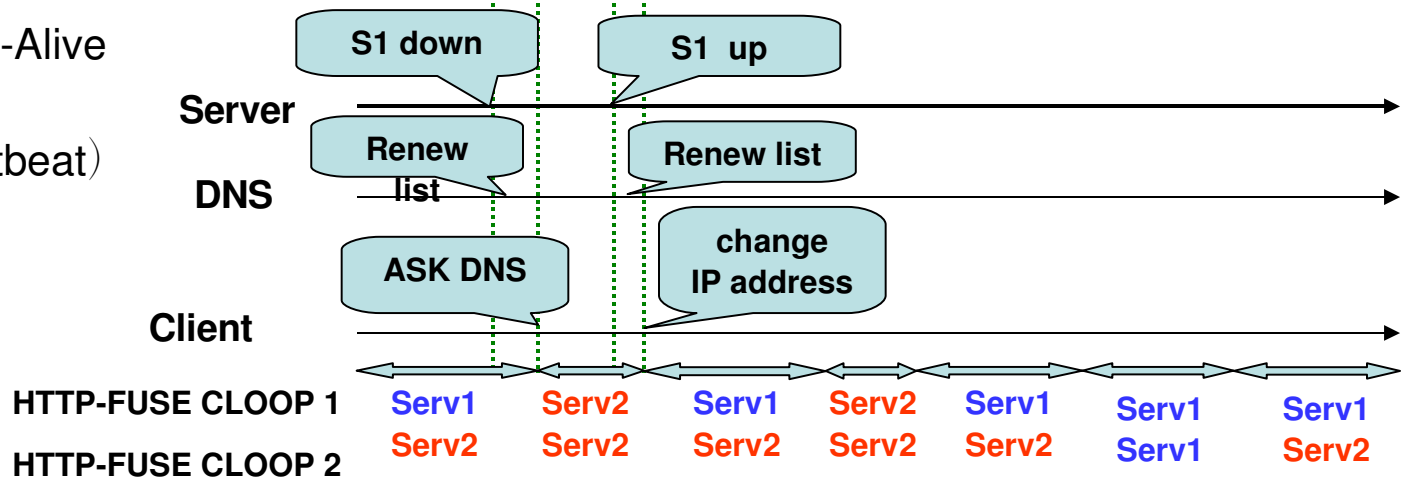
- The combination of MD and DNS Balance achieves fault-tolerance because the server of HTTP-FUSE CLOOP is replaced by DNS Balance dynamically.

- The DNS Balance checks the live servers.
  - When server is down, DNS Balance find the down server by ***heart-beat***. The client asks DNS to replace the dead server and DNS answers a live server.

# Fault-tolerant

- Fault Tolerant
  - Application don't recognize the server down.
- Restoration time is determined by
  - DNSTTL
  - HTTP-FUSE timeout
  - HTTP-Keep-Alive
  - List Update time（Heartbeat）



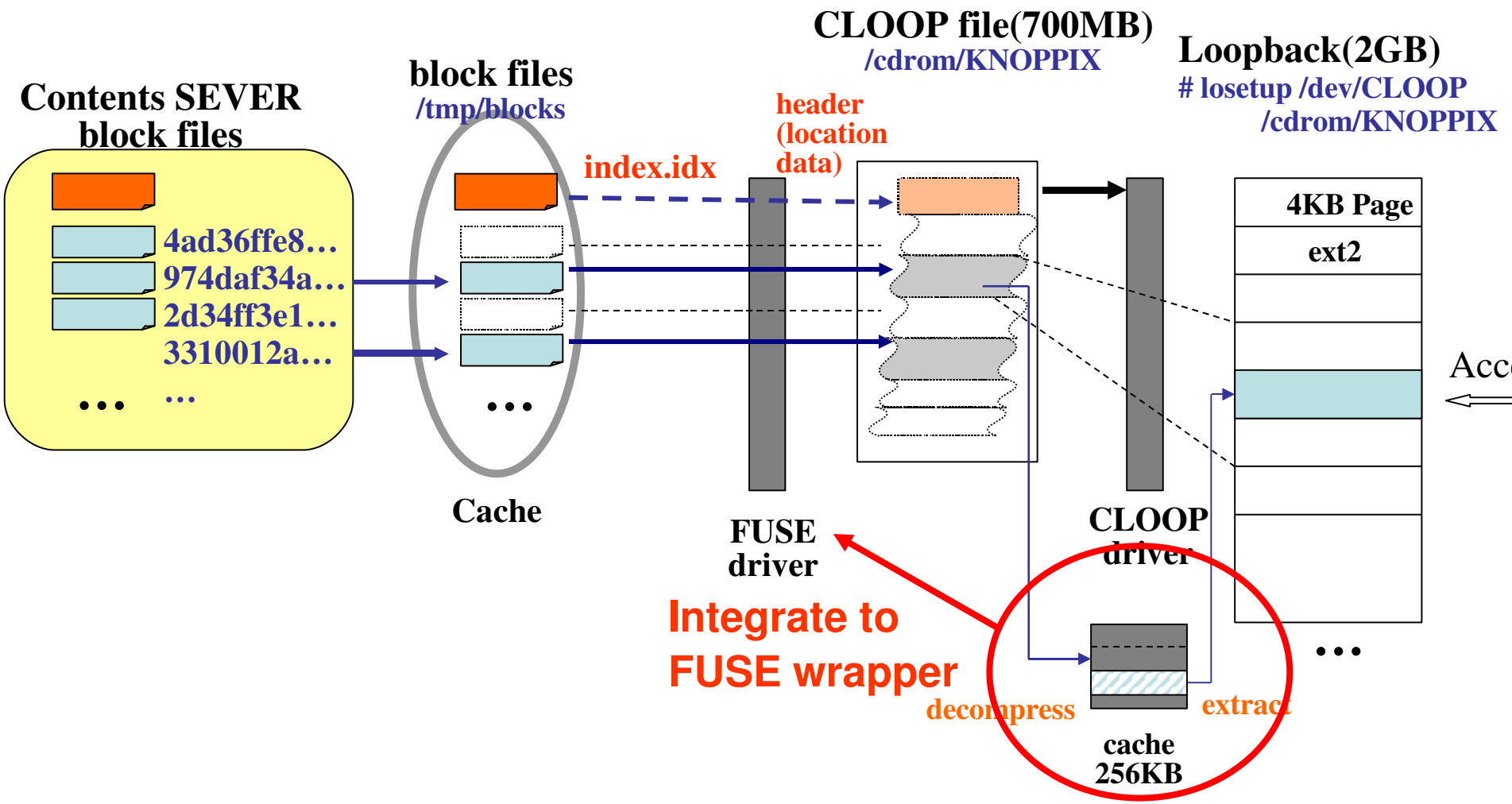
# For Embedded Linux

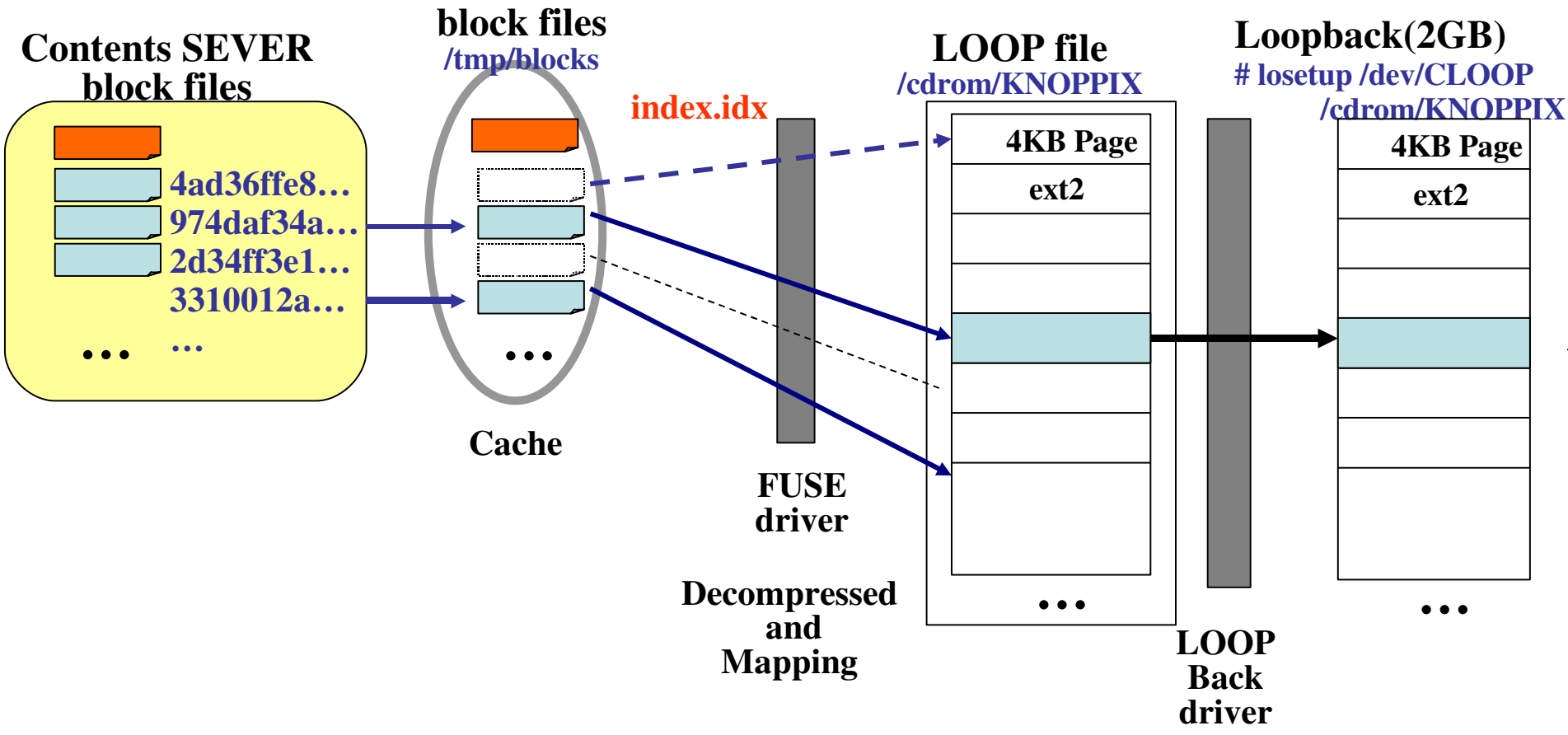
- Embedded Linux runs on low power machine
  - Low power CPU, Small Memory & Storage
  - Network is 100Mbps
    - Network is rich resource. HTTP-FUSE CLOOP compensates the small storage.

- HTTP-FUSE CLOOP is designed for powerful PC.

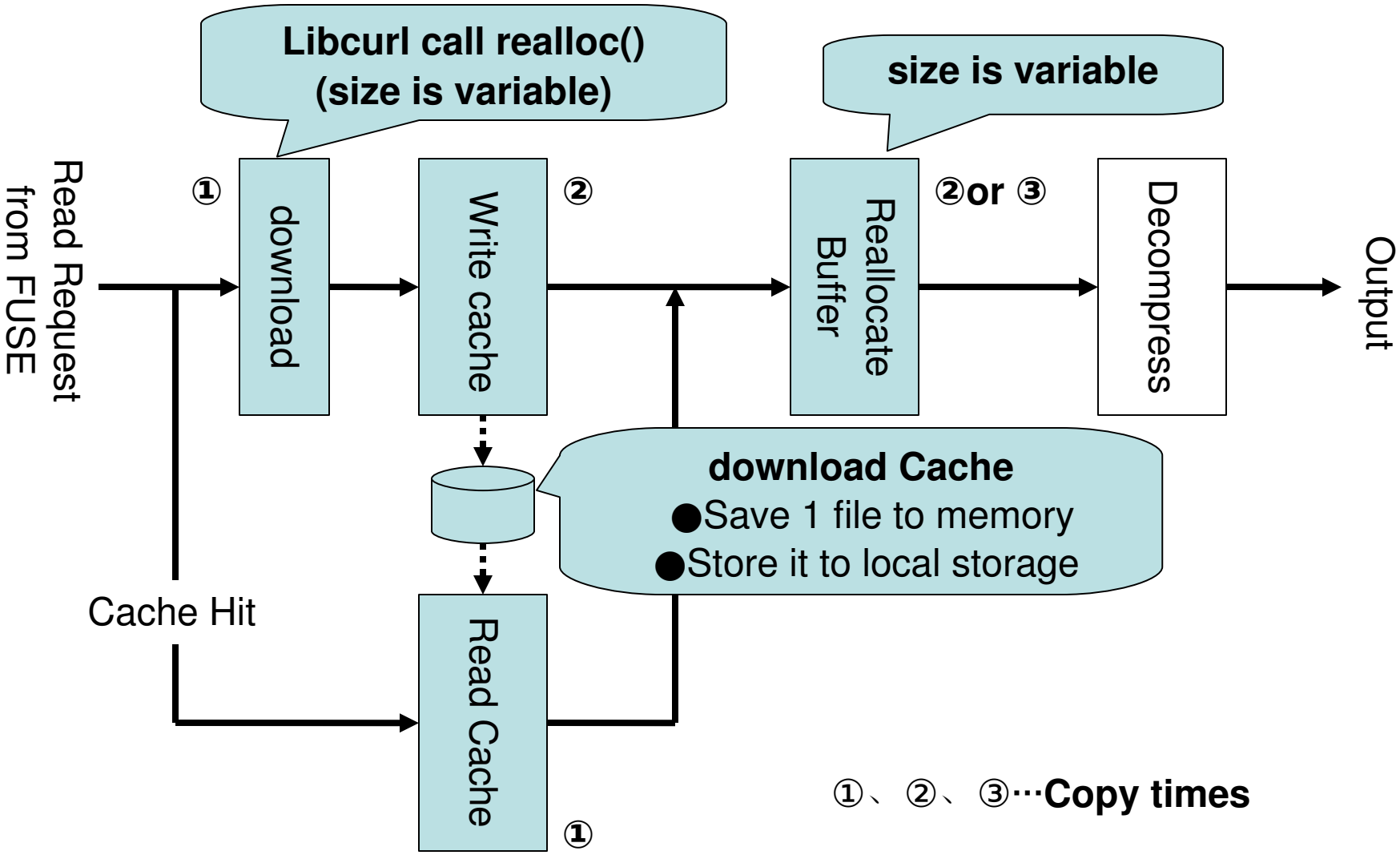
# Reduce memory copy & Replace light decompress

- HTTP-FUSE CLOOP isn't tuned well.
  - HTTP-FUSE CLOOP builds up existing drivers; FUSE, CLOOP driver.
    - Replace CLOOP driver with normal LOOP driver. The decompression is done at FUSE Wrapper.
  - Customize HTTP-FUSE CLOOP to reduce memory copy.
  - Libz is heavy for embedded system
    - Decompress is replaced

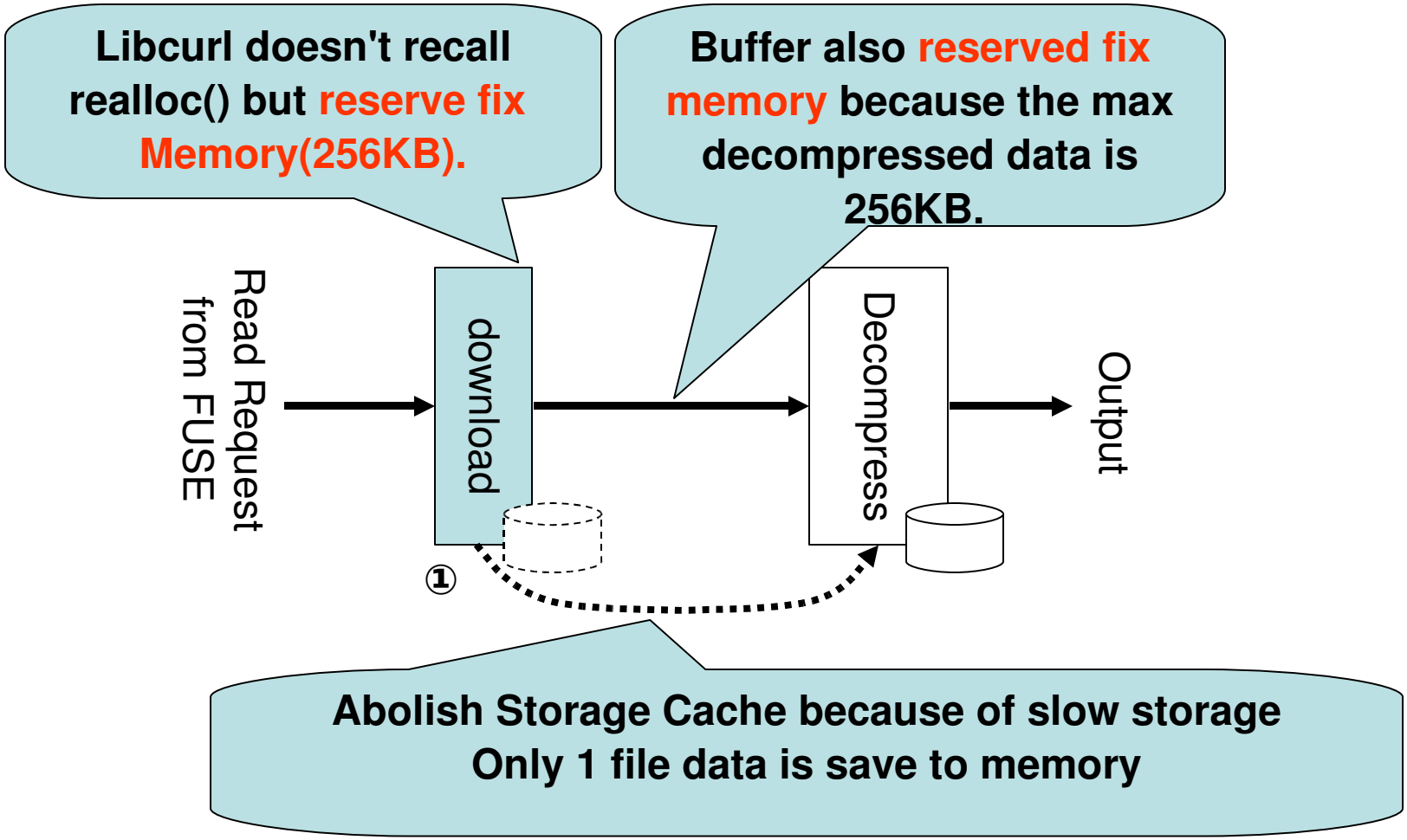

Contents SEVER
block files
4ad36ffe8…
974daf34a…
2d34ff3e1…
3310012a…
block files
/tmp/blocks
index.idx
Cache
CLOOP file(700MB)
/cdrom/KNOPPIX
header
(location
data)
FUSE
driver
Loopback(2GB)
# losetup /dev/CLOOP
/cdrom/KNOPPIX
4KB Page
ext2
CLOOP
driver
Integrate to
FUSE wrapper
decompress
extract
cache
256KB


Contents SEVER
block files
4ad36ffe8…
974daf34a…
2d34ff3e1…
3310012a…
block files
/tmp/blocks
Cache
index.idx
FUSE
driver
Decompressed
and
Mapping
LOOP file
/cdrom/KNOPPIX
4KB Page
ext2
LOOP
Back
driver
Loopback(2GB)
# losetup /dev/CLOOP
/cdrom/KNOPPIX
4KB Page
ext2

# Cache Copy of HTTP-FUSE-loop (Original)



# Cache Copy of HTTP-FUSE-loop (Revised)

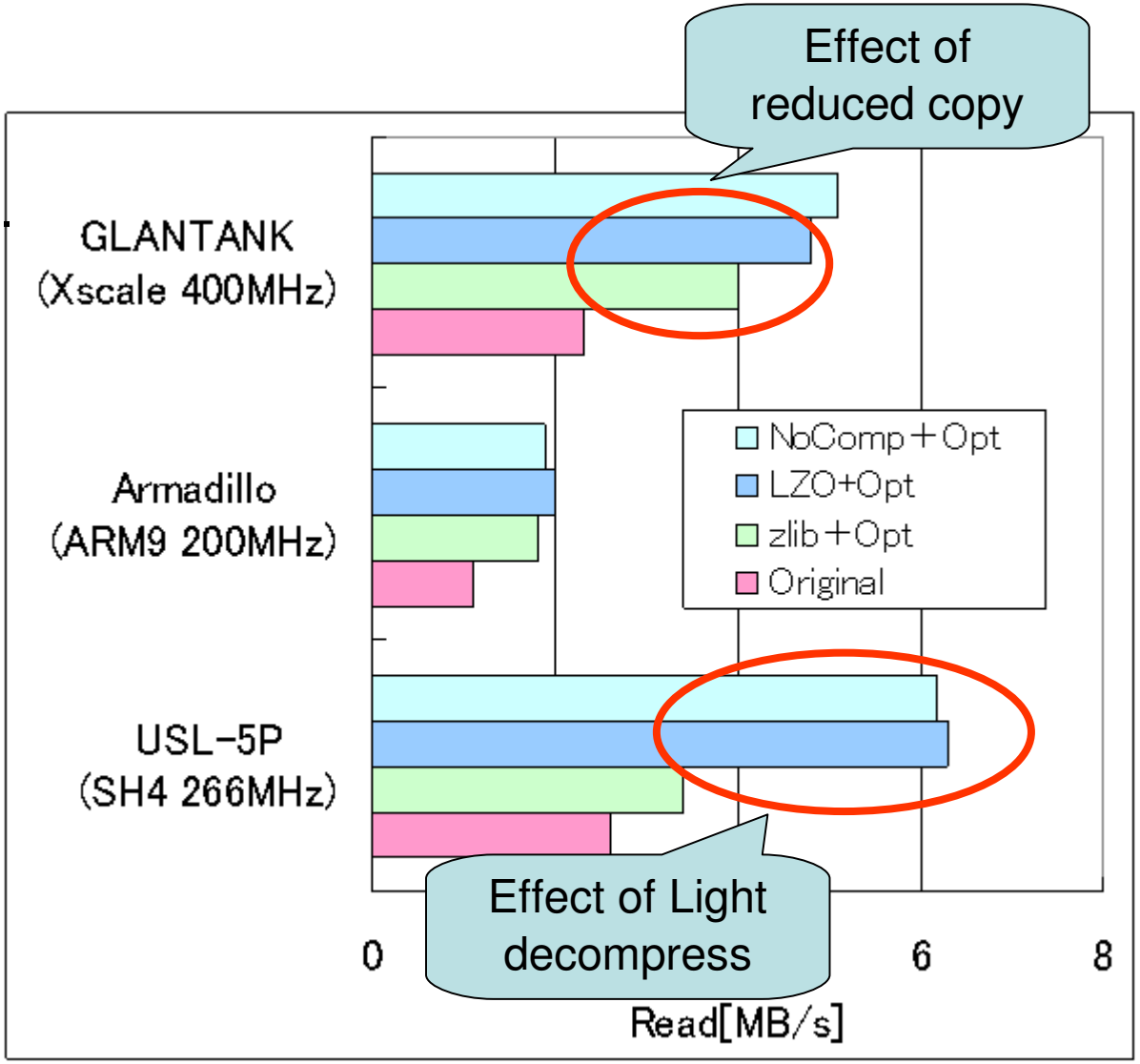


# Decompression

- Zlib is heavy on embedded system.
  - We replace Zlib with LZO
  - Or, no-compression because NIC is rich device on embedded system.

# Performance

- Memory optimization is effective on ARM9.
- Light decompress is effective on SH4
- Performance is 1.8～2.4 up.



# Conclusions

- We propose new optimization for HTTP-FUSE CLOOP.
  - Software RIAD makes wide bandwidth and DNS-Balance finds appropriate server.
    - Software RAID makes 4.14 time wider than original
  - The combination of Software RAID and DNS Balance makes Fault Tolerance.
- We customize HTTP-FUSE CLOOP for Embedded Linux.
  - Reduced Memory Copy and replace decompress with light one.
- The combination of Software RAID, DNS-Balance, and Embedded Linux is under construction now.