http://people.allbsd.org/~hrs/FreeBSD/sato-FBSDW20150130-1.pdf

第4回FreeBSDワークショップ

(18:30から)

佐藤　広生 <hrs@FreeBSD.org>

東京工業大学/ FreeBSD Project

2015/1/30

開催背景

- **日本国内の＊BSD活動は2000年以降、縮小の一途です**
  - 少なくともユーザ数は大幅に減った
  - 海外では明るい話題がそれなりにあるのに...
- 盛り上げたいのはやまやまですが、何をするのが良いですか？

本ワークショップの進行

- 18:30〜19:00　自己紹介＋話題にしたいトピックの提示
- 19:00〜19:30　提示トピック
- 19:30〜19:45　休憩
- 19:45〜21:00　ライトニングトーク

意見は自由に発言ください！

オーガナイザの自己紹介

- 名前：佐藤 広生
  - FreeBSD コアチームメンバ、リリースエンジニア(2006-)
  - FreeBSD Foundation 理事(2008-)
  - その他の＊BSD／オープンソース関連の活動いろいろ
  - 東京工業大学助教(2009-)

自己紹介タイム

- 名前（所属）
- 開発者 or 利用者
- 興味がある／話題に
  したい内容

  をどうぞ

本ワークショップの進行

- 18:30〜19:00　自己紹介＋話題にしたいトピックの提示
- 19:00〜19:30　提示トピック
- 19:30〜19:45　休憩
- 19:45〜21:00　ライトニングトーク

意見は自由に発言ください！

http://people.allbsd.org/~hrs/FreeBSD/sato-FBSDW20150130-1.pdf

ライトニングトーク

http://people.allbsd.org/~hrs/FreeBSD/sato-FBSDW20150130-1.pdf

使ってみようVIMAGE Jail

佐藤　広生 <hrs@FreeBSD.org>

東京工業大学/ FreeBSD Project

2015/1/30

VIMAGE Jail って何だ

- **VIMAGE**
  - ＝　FreeBSDのネットワークスタックを仮想化する仕組み。
  - ＝　経路表、インタフェース等も含めて全部。
  - ＝　Jail に関連付けられている。

- **使うには？**
  - ＝　GENERICに入っていないので buildkernel しましょう

```
# echo "options VIMAGE" >> /usr/src/sys/amd64/conf/GENERIC
# cd /usr/src && make buildkernel && make installkernel
```

- **呼び方**
  - ＝　VIMAGE：カーネルのオプション名から
  - ＝　vnet：シンボル置き換えのためのフレームワークの名前
  - 開発者も統一感なく使っていたりする

VIMAGE Jail って何だ

- **技術的にはどうなってるのか**

  ＝　ネットワーク関連のグローバル変数を格納する
  　　メモリ領域を確保

  ＝　ソースの中の変数参照を、全部ポインタに置き換え

  ＝　シンボル置換マクロを用意して、
  　　元コードをなるべく変えないように工夫

```
#define V_if_indexlim           VNET(if_indexlim)
```

  - 実行コンテキストがどの vnet に属しているかは、struct thread の td_vnet に設定（基本的には継承）
  - 高々100kB 程度のオーバヘッドで、完全に独立したネットワークスタックがつくれる
  - スタックの独立性が高く、MP環境でのロック競合は基本的に発生しない

使い方

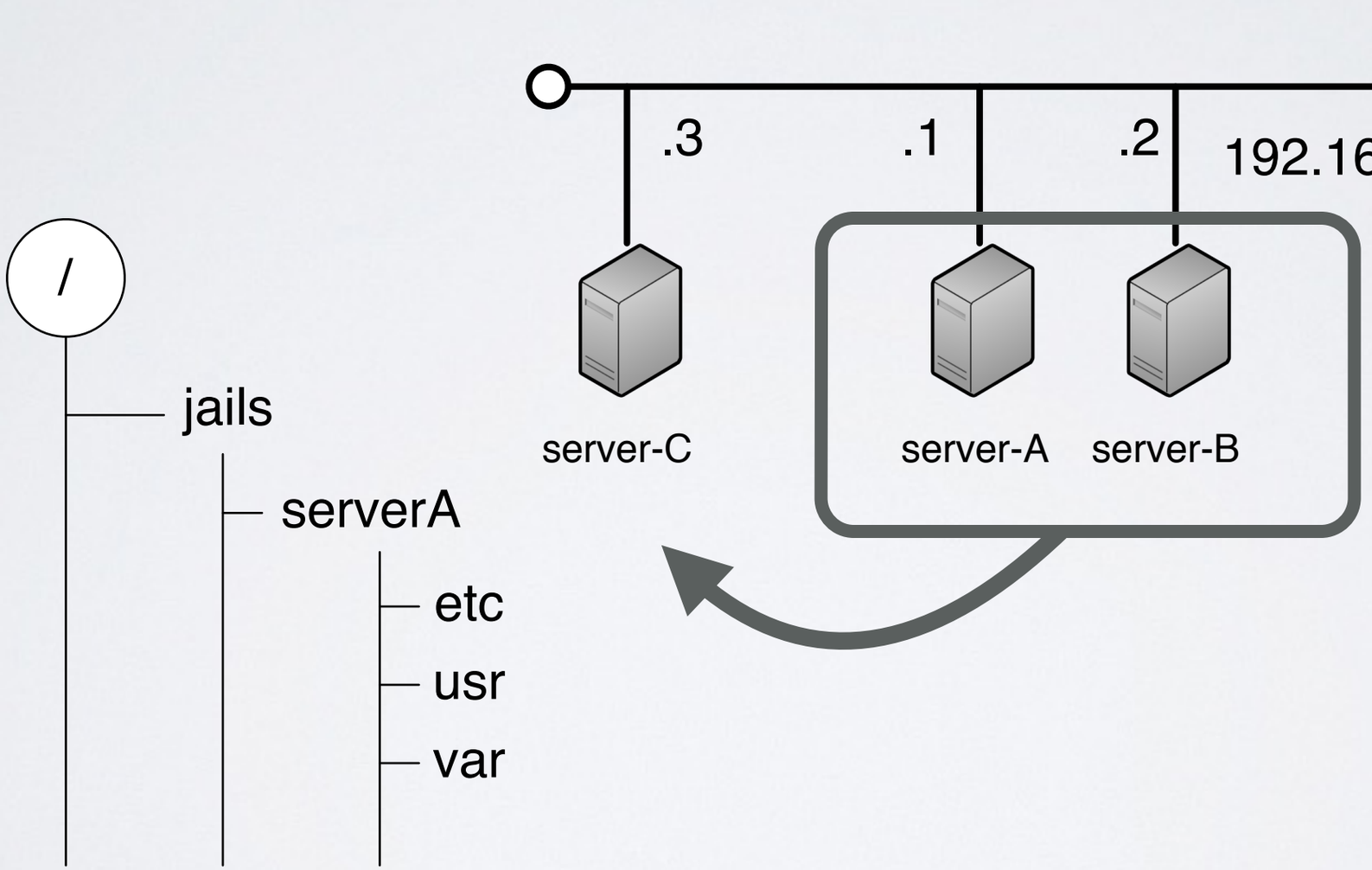
- **覚えておくべきこと**
  - ＝　Jailに関連付けてある。vnet パラメータを付けてつくるだけ
  - ＝　jail コマンド：chroot の高機能版だと思ってて良し

```
# jail -c name=hoge vnet persist
# jexec hoge /bin/sh
# ifconfig
lo0: flags=8008<LOOPBACK,MULTICAST> metric 0 mtu 16384
        options=600003<RXCSUM,TXCSUM,RXCSUM_IPV6,TXCSUM_IPV6>
        nd6 options=21<PERFORMNUD,AUTO_LINKLOCAL>
        groups: lo
#
```

- これだけで同じシステムで異なるネットワークスタックの環境完成。

実運用への応用

- 複数のFreeBSDシステムを1台に集約したい



サーバ集約

- VIMAGE Jail で独立したユーザランドをつくる
- ホスト環境が起動した時に、VIMAGE Jail を自動的に構成するように指定

```
serverC# mkdir -p /jails/serverA

serverC# cd /usr/src && make installworld \
                DESTDIR=/jails/serverA

serverC# mergemaster -U -i -d /jails/serverA
```

（もちろんすでに動いているマシンのファイルをまるまるコピーしてもOK）

- /etc/rc.conf @ serverC

```
jail_enable="YES"
jail_list="serverA"
```

サーバ集約

- /etc/jail.conf @ serverC

```
host.hostname = "${name}.allbsd.org";
path = "/jails/${name}";
exec.clean;
exec.system_user = "root";
exec.jail_user = "root";
exec.start += "/bin/sh /etc/rc";
exec.stop = "";
exec.consolelog = "/var/log/jail_${name}_console.log";
mount.devfs;
devfs_ruleset = "10";
mount.fdescfs;
mount +=  "procfs /jails/${name}/proc procfs rw 0 0";
allow.mount;
allow.set_hostname = 0;
allow.sysvipc;
allow.raw_sockets;


serverA { vnet; };
```

グローバル設定

（${name}などは展開される）

jail単位の設定

サーバ集約

- server-C 上で server-A のイメージを起動してみる

```
serverC# /etc/rc.d/jail start
Starting jails: serverA.
serverC# /etc/rc.d/jail console serverA
:
root@serverA:~ #
```

- **jail スクリプトがやっていること**
  - **start:** /etc/jail.conf の設定に従って jail を作成
  - **console:** jail の中で login(1)を実行

前掲のjail.confだと、作成後に /etc/rc が走るので、
ほぼホスト環境と同じような初期化処理が行われる
（コンソール出力は/var/log/jail_$name_console.log へ）

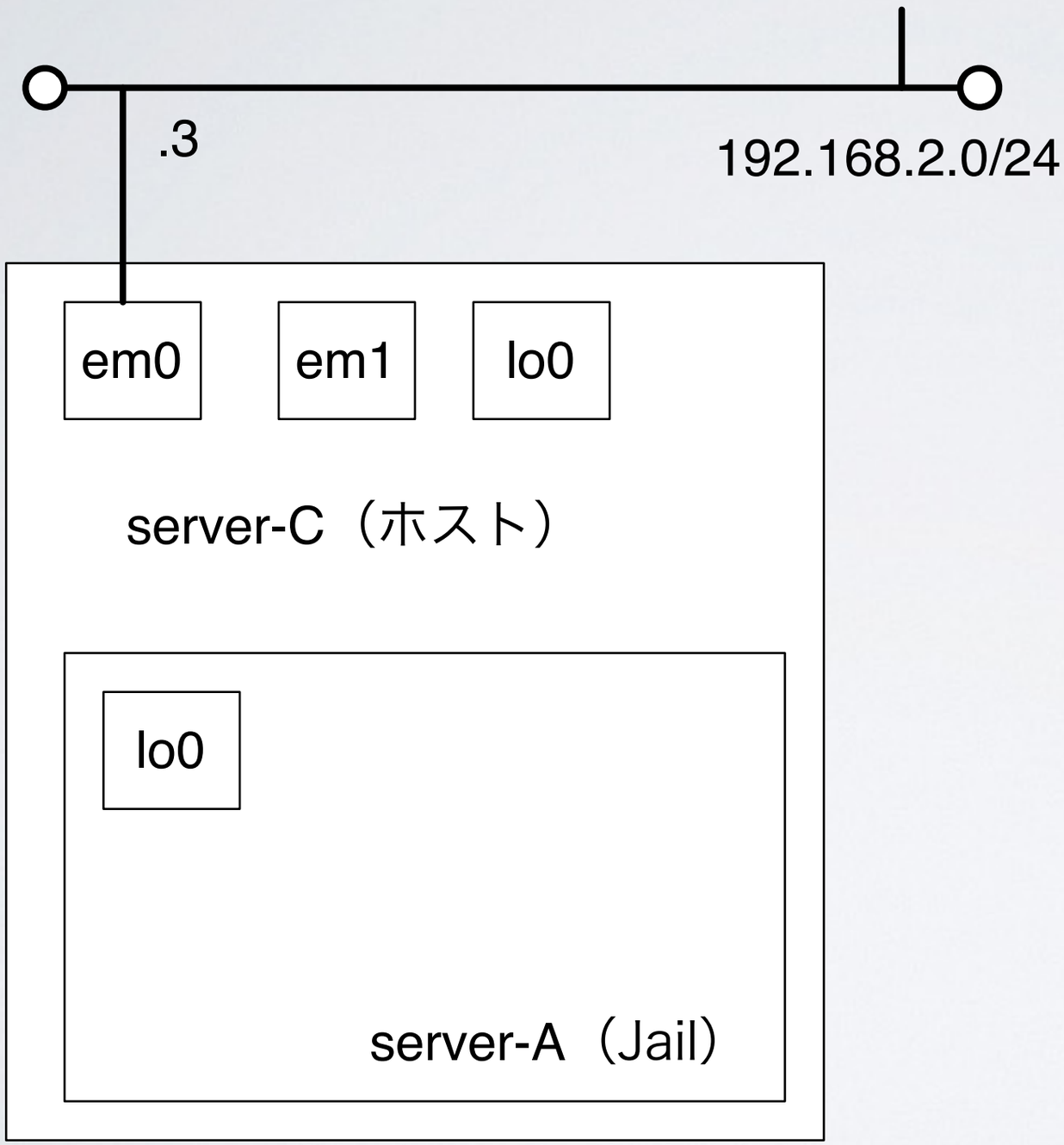
サーバ集約

- server-A には lo0 しかないぞ

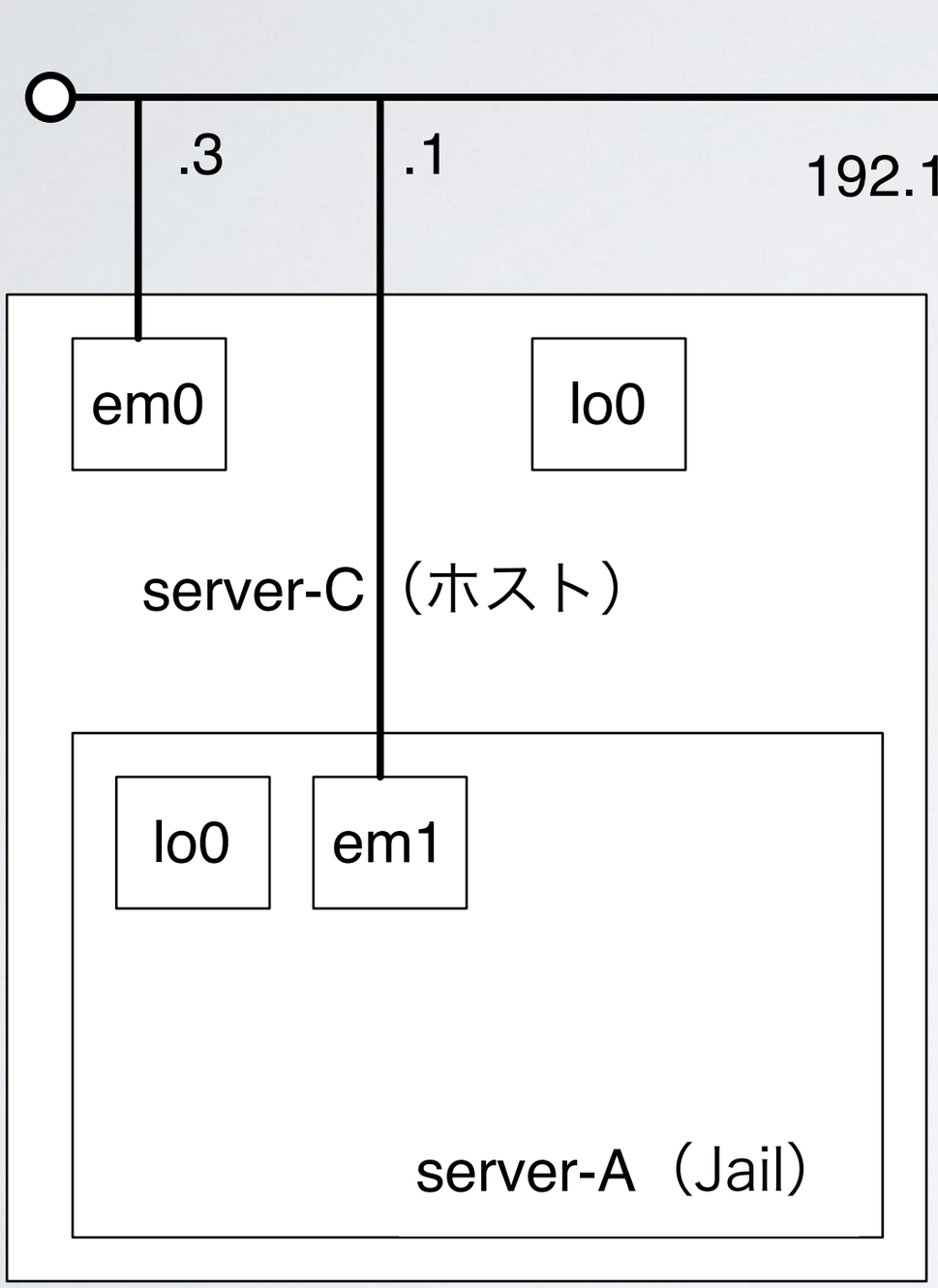
```
root@serverA:~ # ifconfig
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> metric 0 mtu 16384
        options=600003<RXCSUM,TXCSUM,RXCSUM_IPV6,TXCSUM_IPV6>
        inet6 ::1 prefixlen 128
        inet6 fe80::1%lo0 prefixlen 64 scopeid 0x1
        inet 127.0.0.1 netmask 0xff000000
        nd6 options=21<PERFORMNUD,AUTO_LINKLOCAL>
```

server-C （ホスト環境）を経由して外部と通信するような設定が必要

サーバ集約



サーバ集約



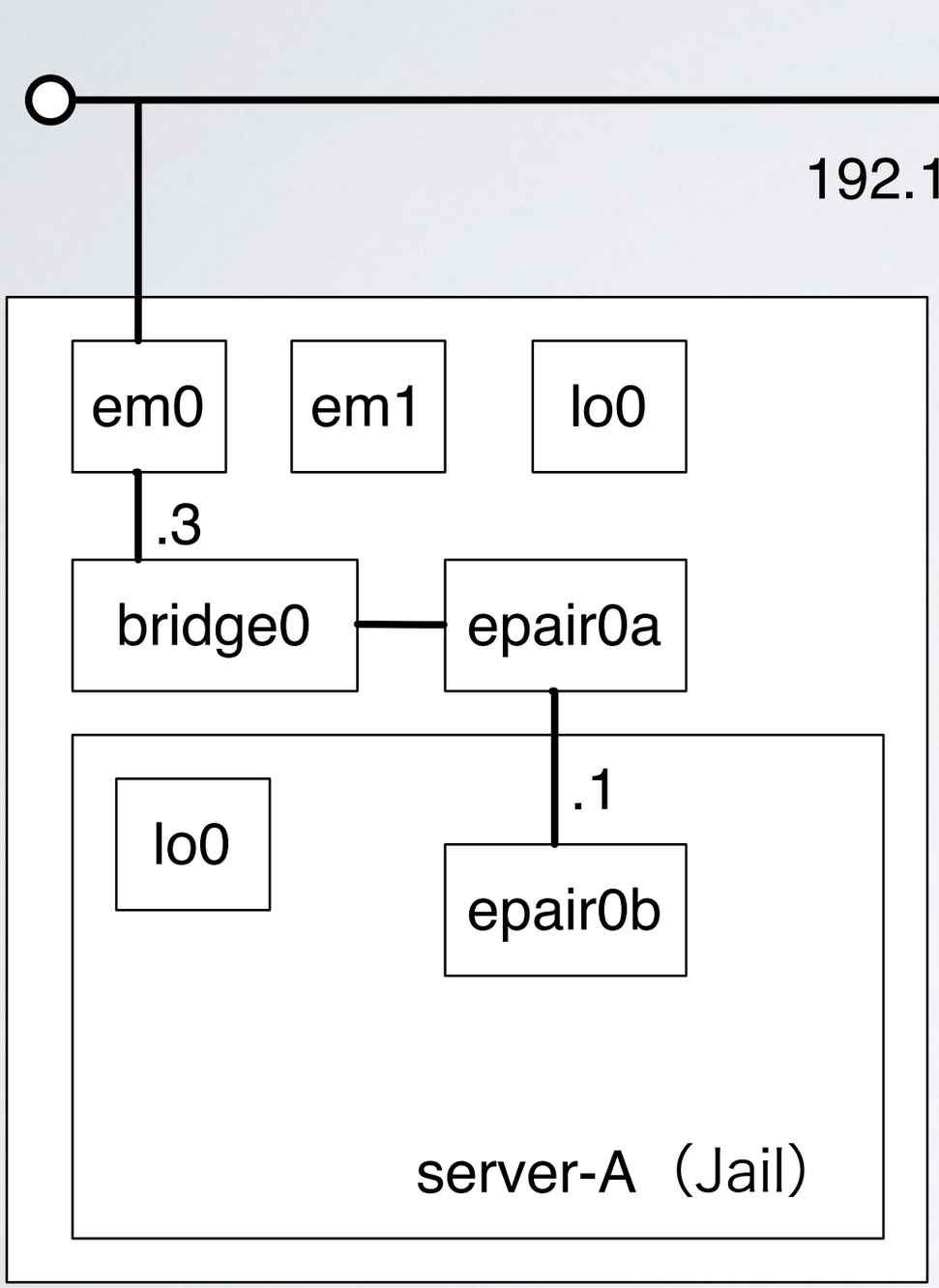
- 方法1
  - server-C の物理NICを server-Aに移動させる

```
serverC# ifconfig em1 vnet serverA
```

```
/etc/jail.conf
serverA {
    vnet;
    vnet.interface = "em1";
}
```

独立性が高い構成

サーバ集約

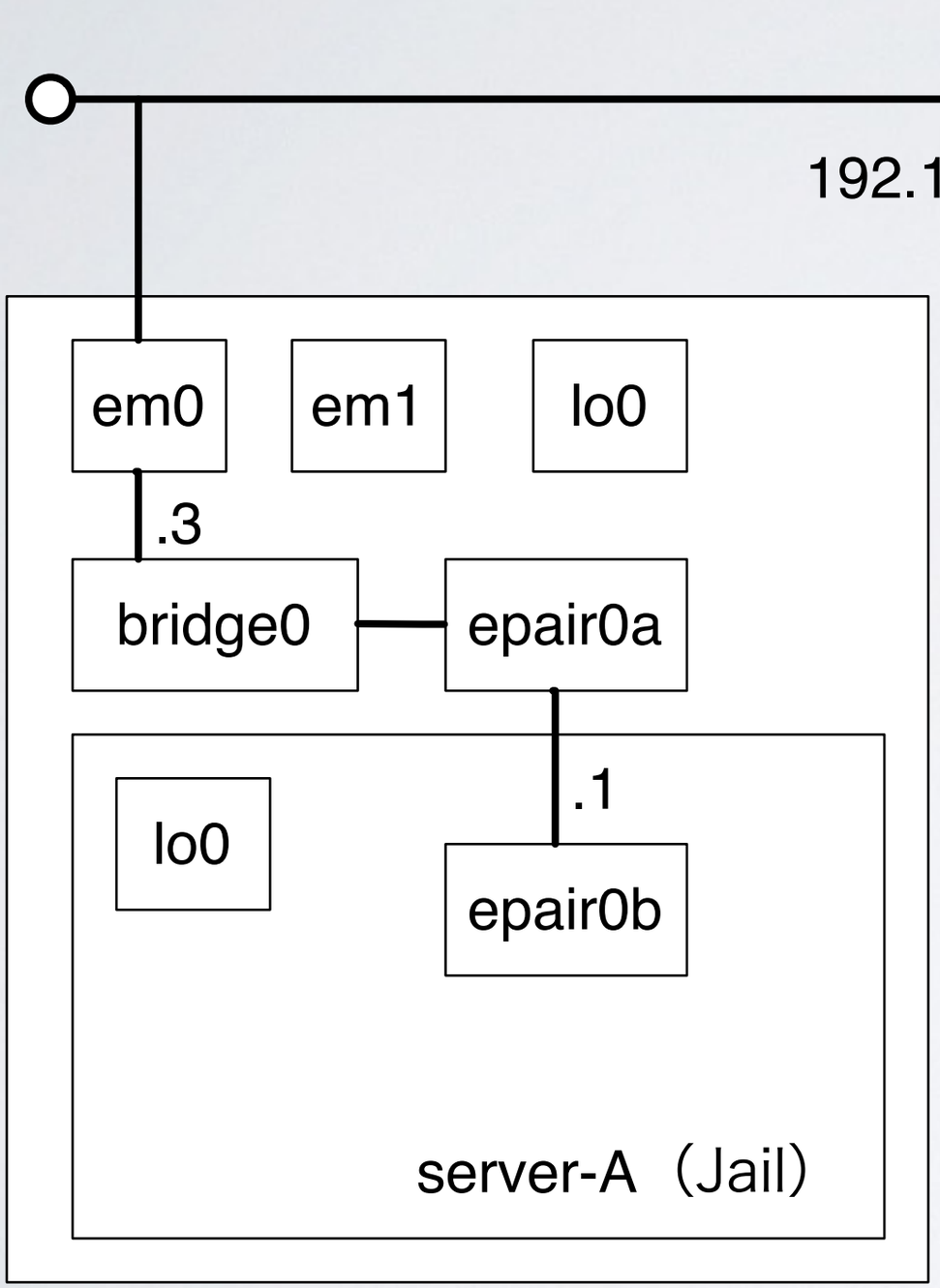


- 方法2
  - AとCをepair（仮想NIC）で接続
  - C側のepairをem0とブリッジ

```
serverC# ifconfig epair0 create
serverC# ifconfig epair0a up
serverC# ifconfig epair0b vnet serverA
serverC# ifconfig bridge0 create
serverC# ifconfig bridge0 addm em0
serverC# ifconfig bridge0 addm epair0a
serverC# ifconfig bridge0 inet \
                              192.168.2.3/24
```

経路設定などは不要

サーバ集約



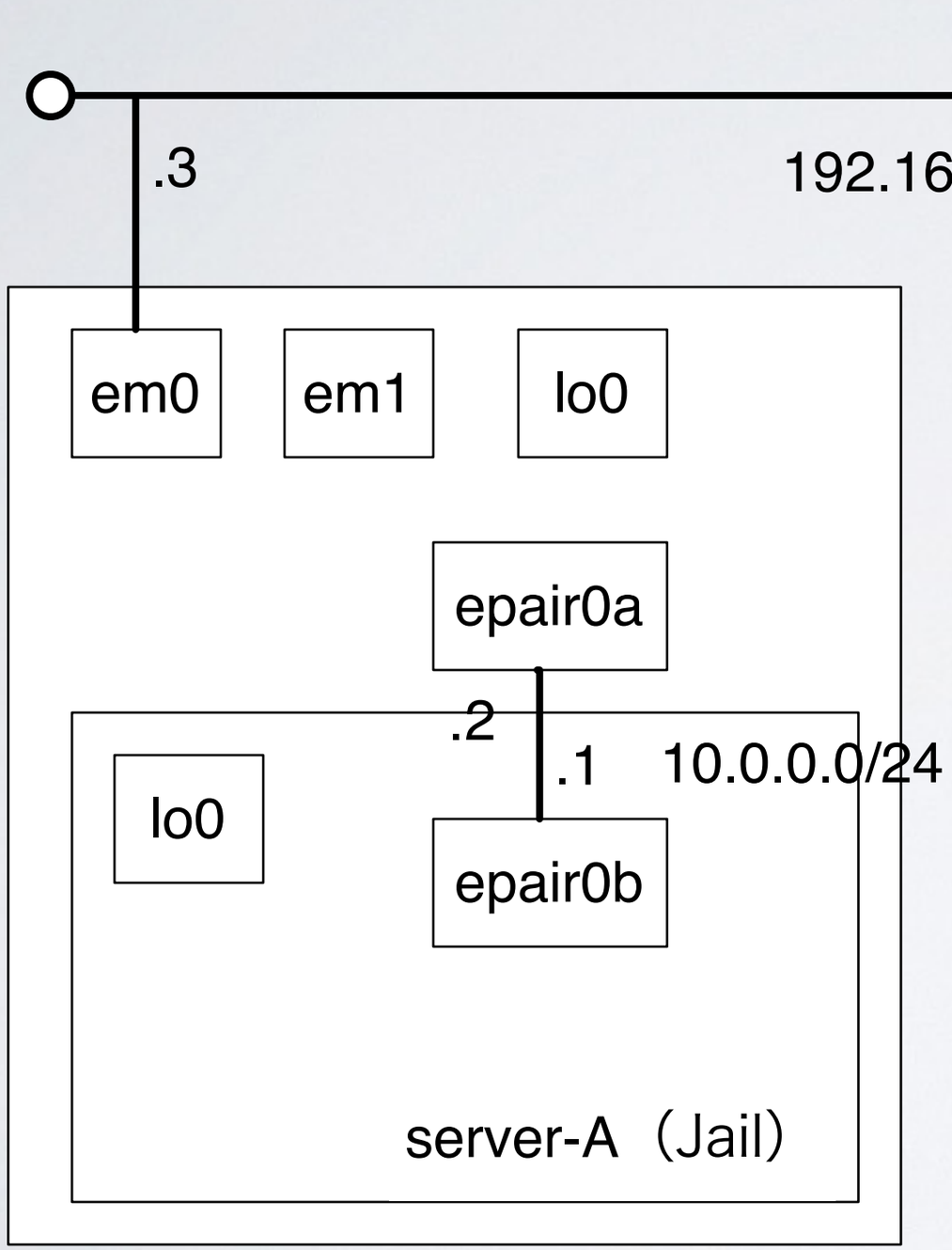
- 方法2
  - AとCをepair（仮想NIC）で接続
  - C側のepairをem0とブリッジ

/etc/rc.conf

```
cloned_interfaces="bridge0 epair0"
ifconfig_epair0a="up"
ifconfig_bridge0="addm epair0a \
                  addm em0"
ifconfig_bridge0_alias0=" \
                 inet 192.168.2.3/24"
```

経路設定などは不要

サーバ集約



- 方法3
  - AとCをepair（仮想NIC）で接続
  - Cをルータとして動かす

経路設定や
static NAT設定が必要だが、
外部のIPアドレス空間の
節約ができる

まとめ

- /jails/serverA の下にユーザランドを全部コピー
  （新規につくりたければmake installworldでOK）
- /etc/rc.d/jail start で起動
- /etc/rc.d/jail console でコンソールアクセス
  （もちろん sshd をあげて SSH でアクセスとかもできる）
- あとは実マシンと同じように、アプリケーションを入れて使える（jail はホスト名なども別々に設定可能）

- sysutils/ezjail ：もっと高機能な管理スクリプト

AsiaBSDCon


AsiaBSDCon2015
A Technical Conference for Users and Developers on BSD-based Systems
Redistribution and use in sou nd bi forms, with or hout modification, are
permitted provided that the met
1. Redistributions of source t notice, this list of
conditions and the follow
2. Redistributions in binary copyright notice, this list of
conditions and the fol nentation and/or other
materials provide it
THIS SOFTWA Y TH CONTRIBUTORS "AS IS" AND

AsiaBSDCon

- 2015は投稿がだいぶ増えたので、3並列セッションになりました
  （およそ30件程度）

- 3/12,13 がチュートリアル、14, 15 が論文発表です。

- コミュニティ支援のためにも、ぜひ参加ください。

告知

- AsiaBSDCon 2015（JR飯田橋駅付近）
  2015/3/12-15。

- FreeBSD勉強会（有楽町線・麹町駅付近）
  不定期（おおよそ月一回）

- FreeBSDワークショップ（JR飯田橋駅付近）
  月一回のしゃべる会。