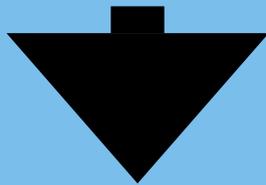


データセンタは今...!?

エスアールエス・さくらインターネット株式会社
田中邦裕

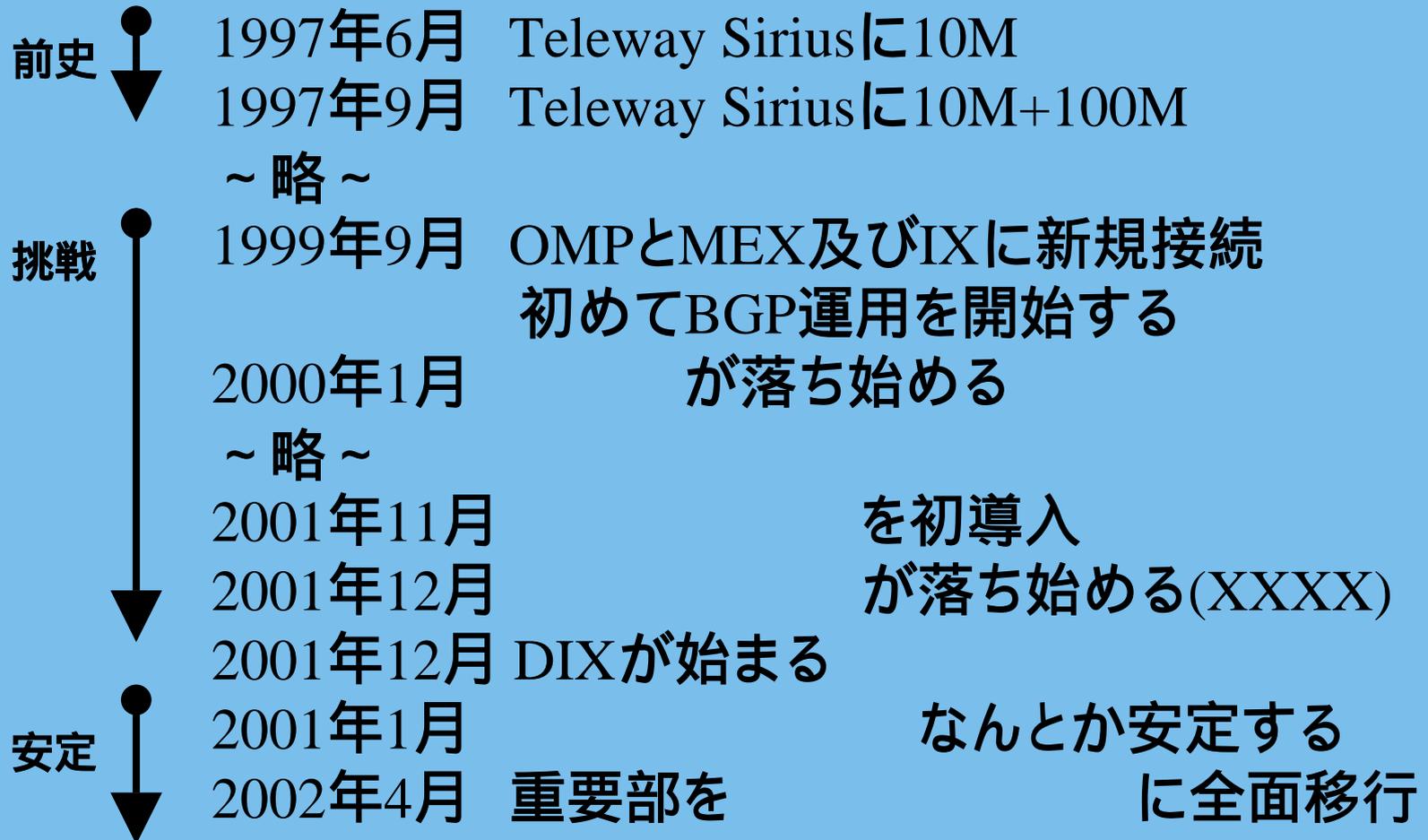
まずは手短かに自己紹介

- ・Aさん・・・マニア向けのウェブサイトがいっぱいあるよね
- ・Bさん・・・ の を使ってたらしいけど
- ・Cさん・・・営業をする気はあるのかしら
- ・D氏・・・さくら銀行系？
- ・E君・・・アダルトサイトをやるならさくらって聞いた



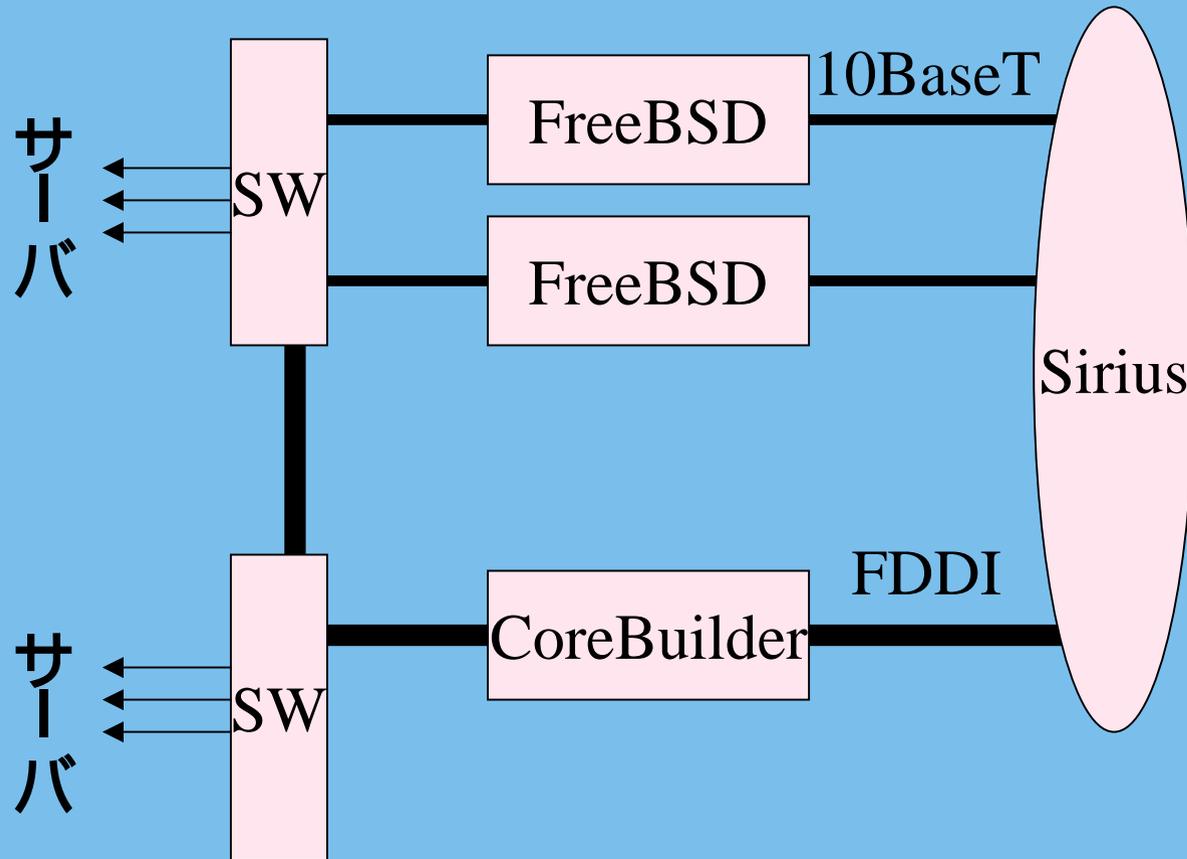
マニア向けのウェブサイト「も」多いんです
昨年12月に へ乗り換えました
実は、ちゃんと営業部だってあります
さくら銀行とは全く関係ないです
アダルトサイト「も」利用できるだけで・・・

戦歴



前史(1997-1999)

SRSとさくらインターネットが別々にTeleway Siriusと接続



末期のトラフィック: 全体で 90Mbps

挑戦期(1999-2001)

BGPの勉強から始めなければならない

…その時の技術員/東京1名・大阪2名(今の代表者2人のこと)

今までPCルータだったのでBGP用に新規購入が必要

…お金もないので必然的にCiscoということでも購入

大失敗

とりあえず、バックホーンはATMで構築

…90年代半ばはATMの全盛期・MegaAccessをたくさん買う

でもって、回線代はまけてもらう

…当時はあまり無かった非対称の形態で売れとしつこく迫る

ルータは高速で名をあげていた に決定

実績もあるとのことだったのだが…

シスコ以外はとては難しい

数ヵ月後・・・

のCPUが100%祭が開催される

→ あとで判ったこと

はフローを元にしたパケットフォワーディングを行っており一度フローが作成できていれば非常に高速である。しかし、フローが作成されていないければ、CPU処理となり、秒間に作成できるフローも限られている。

→ ということで

SrcのIP/PortとDestのIP/Portが異なると別々のフローが作成されるため、インターネット環境では性能が出ない。

とりあえず、flowmodeを変更して、DestのIPのみでフローを作成するように変更。小康状態。

でも、PortどころかSrc IPですらアクセスリストを書けなくなった

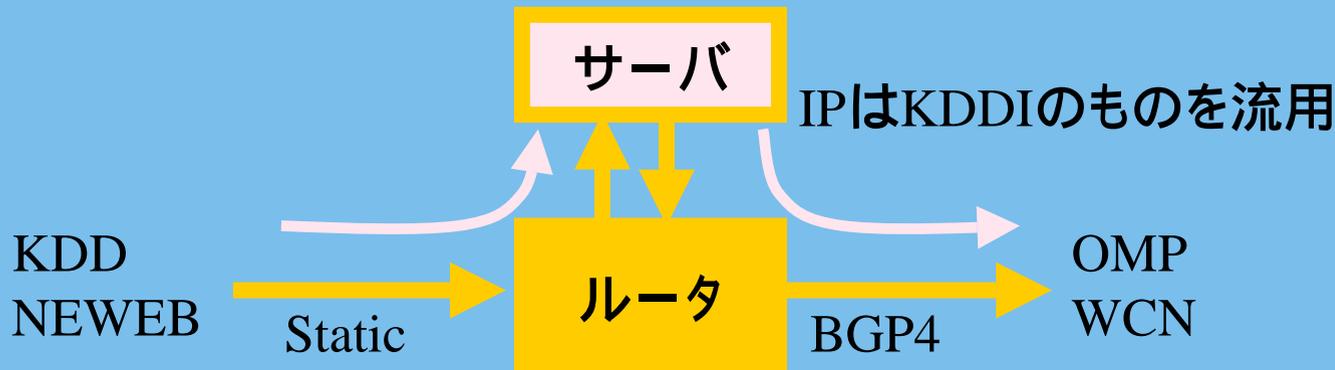
ルータだけじゃなく...

数ヵ月後(その2)...

Incomingをケチっていたので足りなくなる

→ とりあえず、StaticのIPを流用

KDDI(旧Sirius)の回線のうち外部接続ができるものがあったので、専用線でセンターまで引きこみ、Incomingだけ流用する。



→ その後回線を増設し、帯域不足は完全解消

その後

問題の多かった を撤去し に変更

→ まず、ボックス型のXXXXを1台導入(2001/12)

大きなバグが潜んでおり、断続的に障害が発生。
クリスマス休暇返上でファームウェアを交換させ、1ヵ月後に安定
今は問題なく稼動中。

ファームのソフトウェアとしての質はよいが、動作はおかしい・・・
某巨大C社のルータの場合、ソフトウェアとしてはヘン。しかし動作は定評

→ エッチ部を順次 に交換し、 も
導入(2002/3)

エッチに導入した がバグフィックス後3ヶ月程度問題が出ないことを
確認し、エッチルータすべてを に交換。
バックホーンルータも全て に交換。

ただ・・・

でできること

の というASICにて flow(C の NetFlow)を利用すれば、細かなアカウントिंगができるらしい。

がしかし

新作のルータにはよい事なんて何も無い

実際 に飛びついて、年が越せなくなりそうだった・・・。
何時間か通信をすると、ギガポートがブロックするトラブル。
(世界中でも御社しか問題がないよん、というお約束の回答、なんじゃそりゃ) 後でファームの不具合と判明

安定期(2001-)

ができなかったこと

その後、**のウリであるワイヤースピードの帯域制御を試すがこれがかからない。**

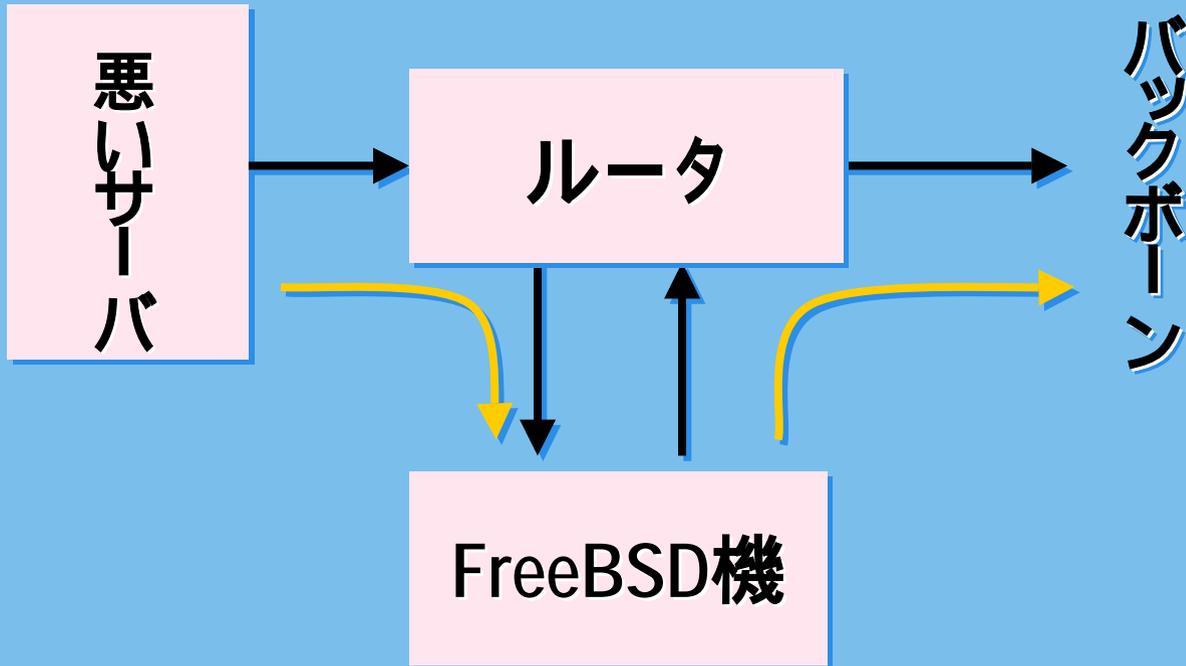
ちわー、
びょん
黒いハコ、初期不良交換で全部返すぞゴルァ!!
です。こないだのファームだめだった

結局、予備機のファームを入れ換えて1ヶ月様子を見て、本機の入れ換えをして1ヶ月様子を見て、やっと6月から帯域制限が可能に…(涙)
でも、現状では他のお客様に迷惑をかける等のことがなければ、営業的な対応以外で帯域制限はかけることは希になった。

そもそも、お客様、月9,800円の専用サーバで8Mbpsも使わんといてくれ…
それで、「さくらの専用サーバ」は遅いって言うのやめとくれ…

参考までに昔の帯域制限

帯域制御をしたいIPアドレスをポリシールーティングする



SSRの設定

```
acl 1000 permit ip 210.224.xxx.xxx any any any
```

```
acl 1000 permit ip 210.224.xxx.xxx any any any
```

```
ip-policy "RATE-LIMIT-1" permit acl 1000 next-hop-list "210.224.yyy.yyy" action policy-first sequence 10
```

```
ip-policy "RATE-LIMIT-1" apply interface IF-RENTAL
```

```
ip-policy "RATE-LIMIT-1" set pinger on
```

帯域制限の設定

カーネルに下の3行を追加して再構

```
options    IPFIREWALL
options    IPFILTER
options    DUMMYNET
```

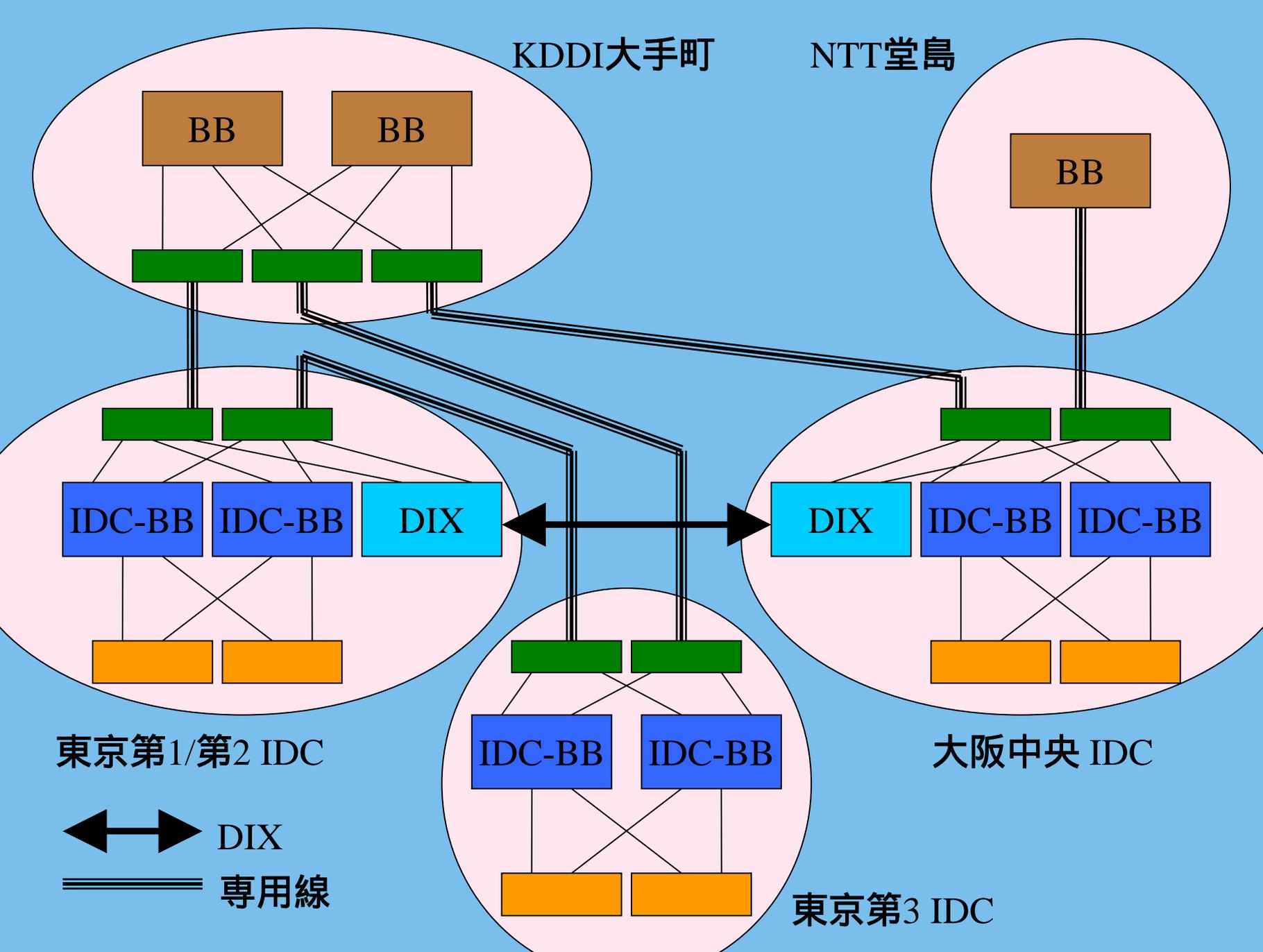
コマンドラインの設定 (rc等に記述しておく)

```
ipfw add pipe 1 ip from 210.224.xxx.yyx to any via fxp0
ipfw add pipe 2 ip from 210.224.xxx.yyy to any via fxp0
ipfw add pipe 3 ip from 210.224.xxx.yyz to any via fxp0
ipfw add pipe 4 ip from 210.224.xxx.yyq to any via fxp0

ipfw pipe 1 config bw 2Mbit/s
ipfw pipe 2 config bw 3Mbit/s
ipfw pipe 3 config bw 2Mbit/s
ipfw pipe 4 config bw 2Mbit/s
```

```
options BRIDGE
```

という便利な暗号;)もあるが、個人ユース以外では使わないほうが吉



KDDI大手町

NTT堂島

BB

BB

BB

IDC-BB

IDC-BB

DIX

DIX

IDC-BB

IDC-BB

東京第1/第2 IDC

大阪中央 IDC

IDC-BB

IDC-BB

東京第3 IDC

↔ DIX

≡≡≡ 専用線

IXにおける勝負

ピアリングしてください

イヤです

なんておっしゃらずに

でも、何だかんだ言いつつピアリング数も多くなりました。

NSPIXP2・・・46 peer

JPIX・・・・・・・80 peer

NSPIXP3・・・12 peer

いま、トラフィックの7割がIXで捌かれています

ま、ピアリングに重要なのは根回しとしつこさですよ

皆様のもとにも私からの熱いラブレターが届いているかもしれませんが・・・

文例集

私は、SRSさくらインターネットのたなかと申します。
この度は、御社にとてもメリットのあるピアリングのお話をさせて頂きたい
と思い、突然で失礼かとは存じますがメールをお送り致しました。
さて、弊社では国内有数のコンテンツを持っているデータセンタ事業者
でございます、ピアリングすることにより弊社だけでなく御社にも大きな
利益をご享受いただけるものと確信いたしております。

つきましては、御社がピアリング頂けるとの理解の上で弊社側の
設定を既に執り行っております。
あまり、downのままですとルータの負荷にもなりますので、手早くピア
リングを行っていただけるよう切にお願いいたします。

～まだまだつづく

さすがに、これは私でも送れない・・・

データセンターを作る

と、だんだん熱くなってきて、話が長くなって、白い目で見られるので、次にいきましょう。

戦歴

- 1999年9月 大阪及び東京にIDCを開設
大阪8ラック、東京12ラック
- 2000年4月 東西IDC第1期増設
55ラック
- 2000年10月 大阪IDC第2期増設
80ラック
- 2001年4月 東京第2-IDC増設
70ラック
- 2002年6月 東京第3-IDC増設
195ラック

センター開所時

❖ 東京第1 IDC ちょっとしたコンピュータ室



センターのグレードアップ

❖ 東京第2 IDC

センターと言えるようになった



大手IDCに並ぶスペックを

❖ 東京第3 IDC

お客様にキャリアのセンターと比較されなくなった

データセンター構築要員の採用

ゼネコン出身でネットワークをやりたいというやつが現れる
業者の手抜き工事を次々と発見。頭角を現す

業者のコンペによる低価格化

全てコンペにして、結局不況にあえぐ大阪市内の業者をまとめて
東京へ投入。工事をさせる

相手の決算期を逆読みして、値引き

半値8掛後、もうちょういたのむわ
お金はないけど、これほしいねん
これくらい、一緒にやってーなー

UPS

結局、UPSは定価の1 / 3に…
冗長化しても、価格を変えずにすむ

三菱がn + 1冗長でお願いします
というので、一緒の値段ならと
言い返すと…。

2週間ほど悩んだ揚げ句
「安くするから買ってくれ」
とのこと。
3月の決算期前だから…



これはUPS 2台分 背面に蓄電池 側面にもう一台

発電機

最初は三菱のものの予定が、
見積りを持ってこないので変更。
これまた、決算月だということで
搬入費をケチる



とりあえず、拡大。

空調

NTT-FのMACS IVと同等品らしい
空調

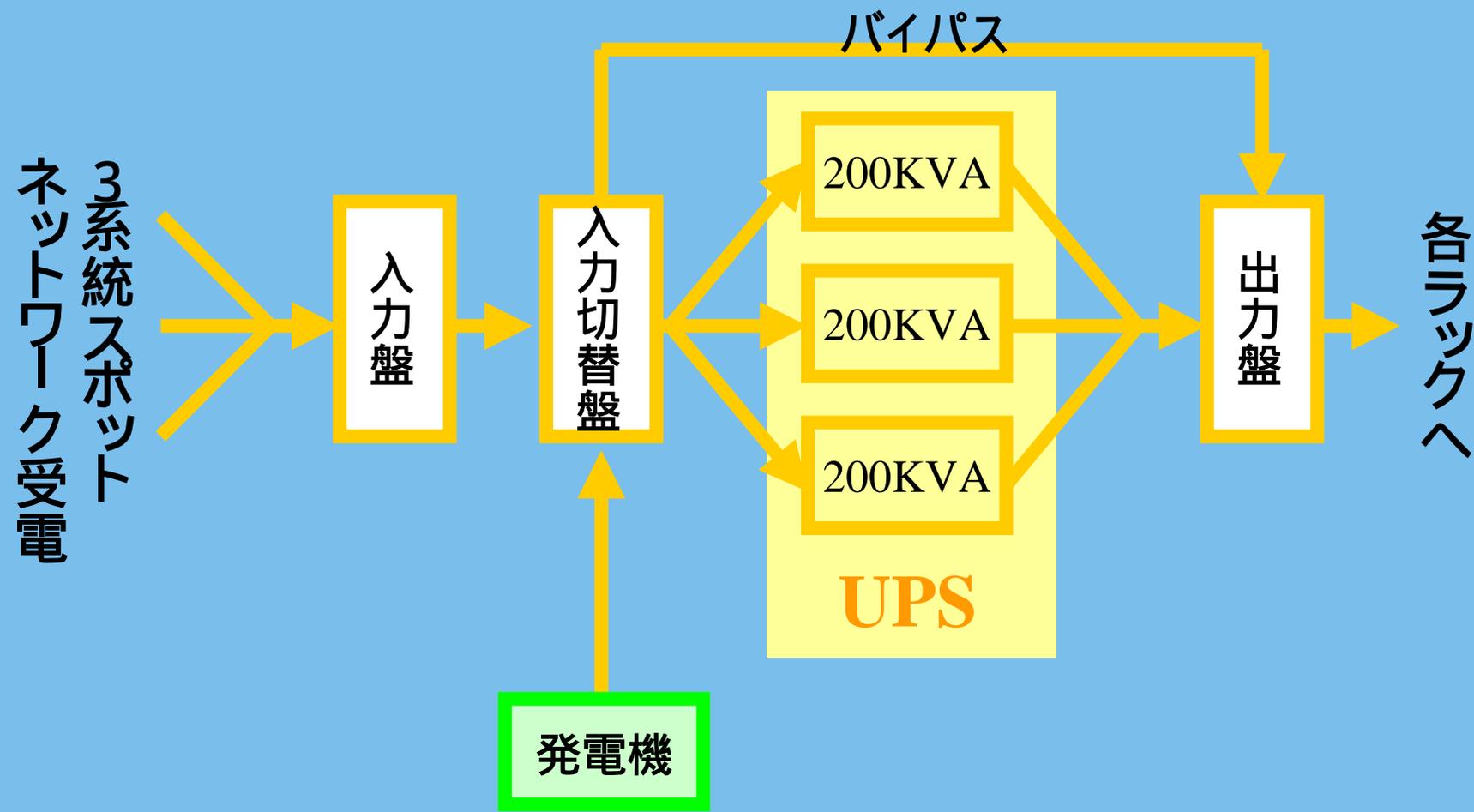
本来ならダイキンにしてコストを
浮かすところを、日立の大阪支店が
契約欲しさにダンピングした。
同じく、決算期前だから？

圧縮機は3系統で、1台での冗長性
をもつらしい
パスワード機能で、お客様が温度
設定を勝手に変えたりできないらしい

なんかいいらしい…。



電源系統図



リールラックには注意

- ❖ リールラックを195本も購入
- ❖ 棚板でもなんでも特注しますよ

ところがどっこい

- ❖ ドイツから取り寄せなんすよね
- ❖ あれ、はまりませんか？
- ❖ あ、それマウントレール外さないとだめです

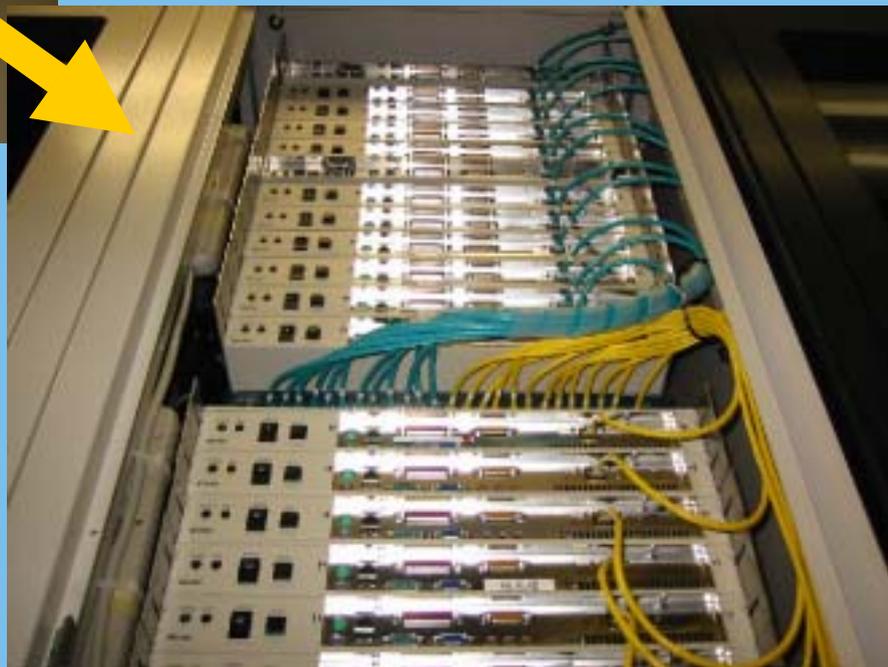
何だかんだ言いつつ...

- ❖ 全ての性能を向上させつつ、1ラックあたりの投資を2 / 3に圧縮
- ❖ データセンター作るなら、大阪の業者にやらせるとコストダウンで来ます(不況だから)
- ❖ でも、価格交渉には大阪人を登用してみたほうが吉です。

サーバの遍歴



当初はASUSTEKの作るNLXを導入していたが、台湾からの納品がまちまちで大変。
あと、マザーが限定される。
現在、600台程度稼働中



結局、ケースから全て作成
マザーボードも、現行の三代目で安定度向上。
現在、全ての自社サーバを自社製1Uに変更。
値切り散らかした結果、NLXより安価で高性能を達成
現在1,500台程度稼働中

で、ナニをやってる会社なの

❖ データセンター

- いわゆる、ラック貸し、ブロック貸し、1U貸し
- 部屋貸しもやります

❖ 専用サーバ

- 1台丸ごとサーバをレンタルするサービス
- 9,800円～75,000円で、使えます

❖ ウェブサービス

- ウェブスペースをレンタルするサービス (sakura.ne.jpってやつです)
- バーチャルドメインもやってます

❖ プロバイダ支援サービス

- DIXなどのIXサービスやニュースサーバ提供などのISP向けサービス
- DIXは既に20社様位のご契約をいただいて、4月以降安定運用中!!

なんとか、単月黒字になりました

で、ここからは泣き言を聞いてください

さくらウェブ 7/25の統計を見ました

順位	ユーザ名	転送量
1		9.44GB
2		8.41GB
3		7.57GB
4		6.81GB
5		4.54GB
6		4.08GB
7		4.01GB
8		3.99GB
9		3.95GB
10		3.92GB

順位	ユーザ名	ヒット数
1		1103128
2		1025367
3		1012721
4		636355
5		596516
6		562907
7		474043
8		472413
9		431921
10		429523

9.44GB = 926Kbps
10位までだけで、5.5Mbps平均
おまけに110万ヒットって……

ちなみに年間使用料1万円です

結局さくらウェブだけで一日9900万ヒット

専用サーバ

❖ 月9,800円で、8Mbps使う

9,800円のサービスの場合、10Mbpsを24分割している
でも、8Mbpsも使われると、10Mbpsで足りない
結局、100Mbpsを24分割することになった

❖ サーバをクラックされ、DoSの拠点になる

セキュリティホールが見つかり、1週間以内にクラック祭が
催される。某所で「クラックはさくらの自作自演」なんていわれた
ものの、実際は残業手当だけでも干されそうなんですぜ…。

❖ IPアドレスをクラスCとしてIPv6のgwにする

クラスCをJPNICから受けるのも大変なんだから…。

具体例

えいやっと止めたくなってしまう...

過去4時間はとめどなく
天上に張り付く

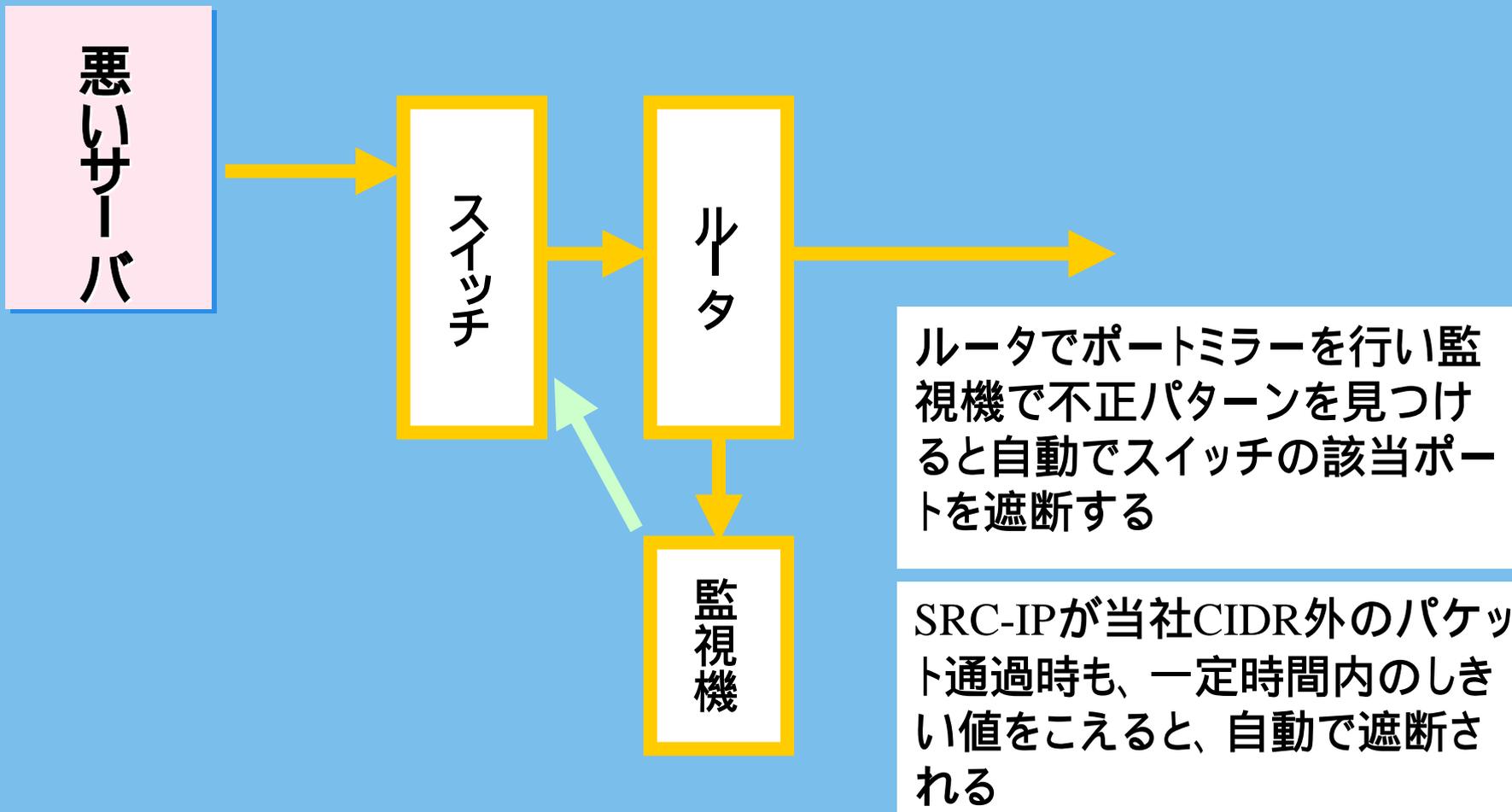
18時からずっと8Mbps
10Base-T Halfの限界



色の趣味が悪いのは、ペイントブラシで減色したせい

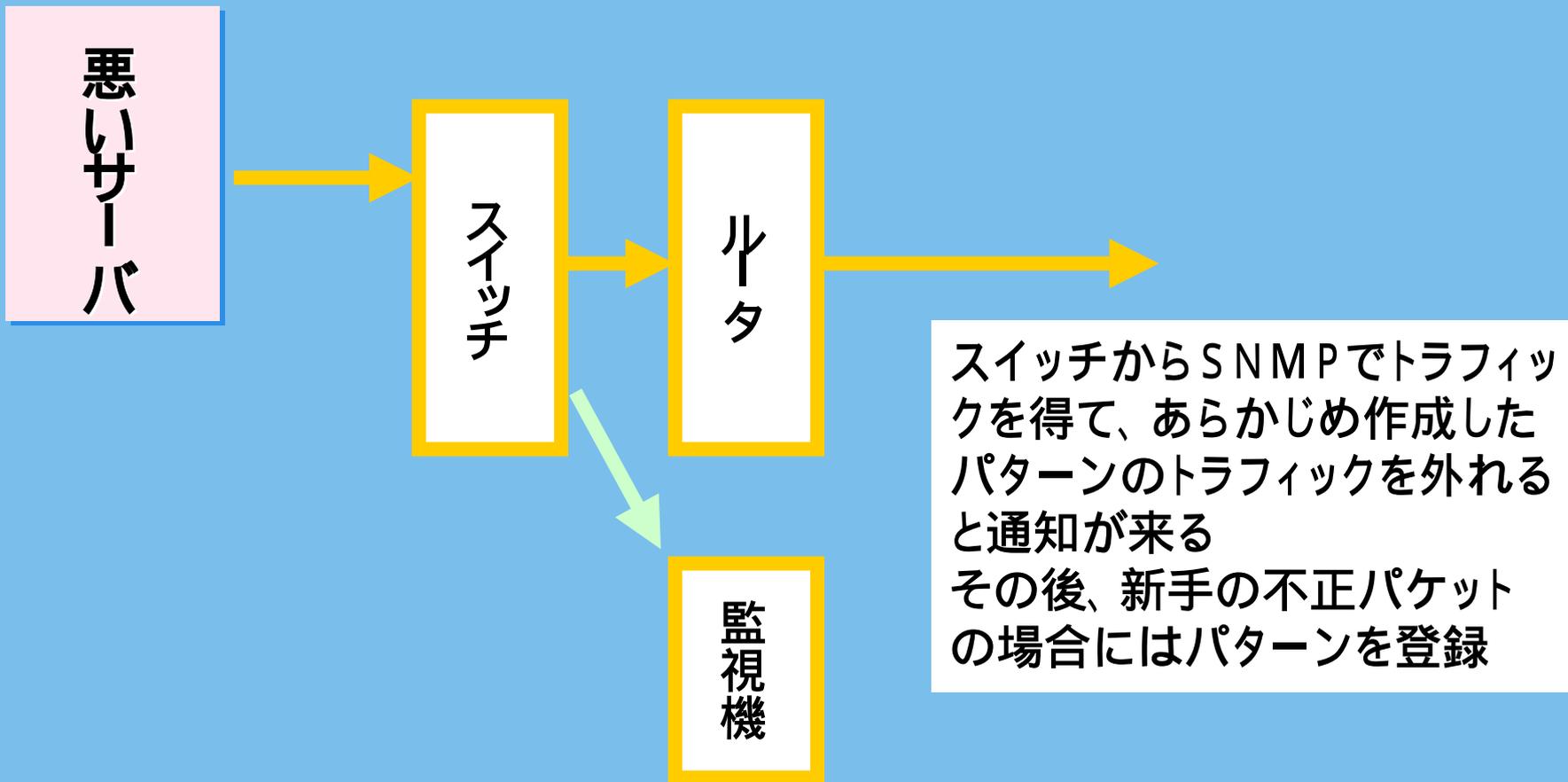
DoSプロテクタ

- ❖ 内側からDoSが始まった際に防げないのは困る



トラフィック急増検知

❖ トラフィック急増を検出して知らせる



不正使用予見

- ❖ あやしいコンテンツを事前に検出し通知
- ❖ type1・・・FTPを監視し、ファイルの拡張子と容量から事前に可能性を察知しておく
- ❖ type2・・・アクセスログを解析する際に、容量の大きなもののリストを作成しておく



お客様のコンテンツにクレームがついた際に、
「サイト難度」を図る参考にする
今後は、CGIのCPU使用時間も監視したいところ

難度・・・どれくらい面倒くさいかの独自指標

さいごに

ピアリングのラブレターが届いても
捨てないでくださいね

でも恋が成就しても「結婚」はしませんので…

お　　わ　　り