

2009/3/38

第2回ハンズオンセミナー in 大阪(日本Androidの会)

PIAX on Android の挑戦

～飛翔しろ、Android。あの雲より高く～

~~キバヨシ~~ 吉田 幹

BBR, PIAX Inc. & 阪大

\$ BBR_



今日のお題 (1)

■ 理念・解説編

- 実は限界、インターネット
- ID/Locator分離
- ネットのトレンドは、Web2.0 から Cloudへ
- かたや、P2Pの方は...
- 構造化オーバーレイって知ってます？
- インターネットを救う P2P
- 社会を変える P2P
- Android が空を飛ぶ意味

今日のお題 (2)

- PIAX・Android編
 - PIAXとは
 - 世界一のおっ嬢様 Skip Graph
 - ID/Locator分離をオーバーレイで実現
 - Transport層のデザイン
 - NAT越えと認証
 - PIAX on Android に至るまで

理念・解説編

実は限界、インターネット

7年前は、

“IP Over Everything, Everything Over IP”

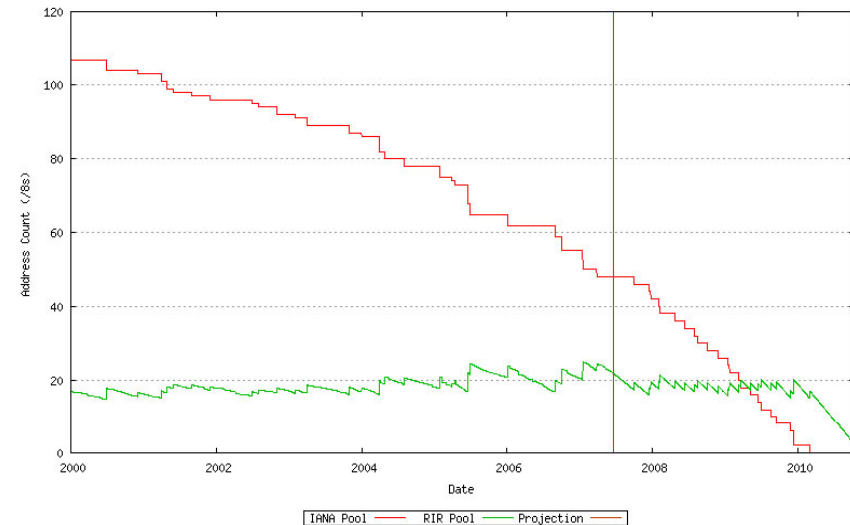
だったけど...

- これが間違ったスローガンであることが認識されつつあります

枯渇するアドレス

■ センサー、タグ類が大きく消費

- 2020年で100億ノード
- M2Mも
- Deep Network の時代へ
 - Smart Dust プロジェクト



Geoff Huston氏によるIPv4アドレス在庫枯渇予測グラフ

■ IPv6は現状打破しない

- <http://www.unixuser.org/~euske/doc/ipv6ex/>

■ 「JPNIC、近づくIPv4アドレスの在庫枯渇に対する取り組みを発表」(2007/6/19)

- <http://www.rbbtoday.com/news/20070619/42805.html>

失われた到達性

- 必要悪なNAT
 - アドレス枯渇と関連
 - end-to-end の透明性 (任意のIPアドレスへの到達性) が損なわれる
 - アドレス回収の際、まだ生きているトランスポート層が切断される。といった細かいけど面倒な話も
- もはや、Stupid でなくなった
 - NAT, Firewall, VPN... と複雑に
- 実質的には、Web中心のトポロジー
 - P2PでなくC/S型
 - 世の中はこれでよしとしているけど

困難なモバイル対応

- そもそも端末が移動するなんて思っていなかった
- 仕方なしに(と書くと失礼だけど)、
 - Mobile IP, Mobile IPv6
 - Home Addressを使って中継させる
 - LING (Location Independent Networking for IPv6)
 - 半分の 64bit をIDとして使用
 - NEMO (Network Mobility)
 - 移動するLAN
- 決め手がないまま、端末はどんどん移動する
 - 携帯モバイル機器(DSとか)。車載システムもいずれはネットに

大丈夫か？ ルータ

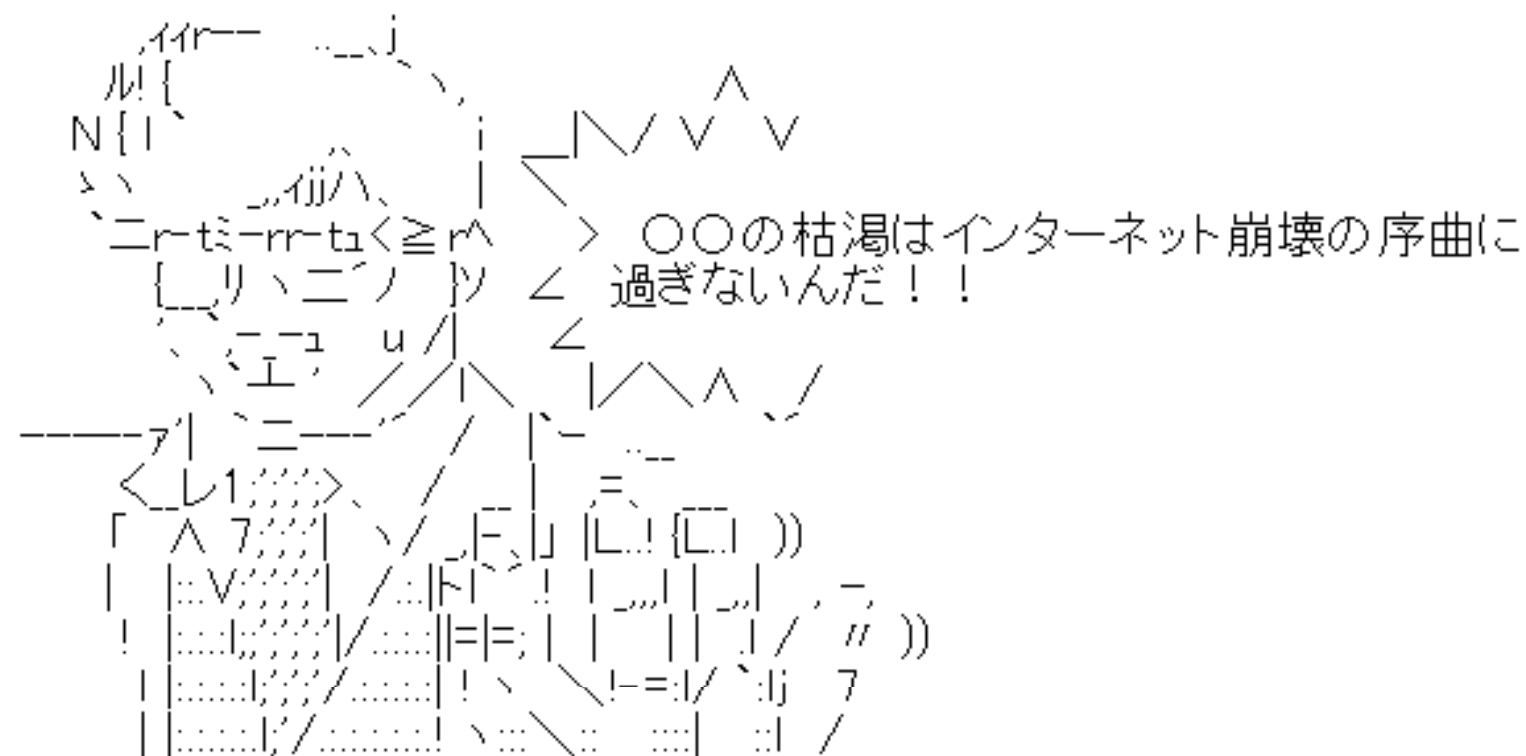
- 増え続ける帯域、増え続ける経路表(30万エントリーを越えつつある)
 - マルチホームが経路表増大の原因の1つに
 - 昨年5月のNTT東の大規模フレッツサービスダウン
 - 本当のところ、IPv6関係のバグが原因だったが、経路表の増大はやはり問題
- コアルータにおいては、小規模発電所レベルの電力を消費
 - 2020年には 48.7%に達するという試算も
 - 改善の措置がされない前提だとこうなる
 - <http://innovation.nikkeibp.co.jp/etb/20060417-00.html>

できないマルチキャスト

- 理論的には、
 - 複数のマルチキャストグループの経路統合は、一般に不可能
- プロトコルは山ほど (DVMRP, MOSPF, PIM-DM, CBT, PIM-SM, PIM-SSM, IGMP...) あっても、
 - まともに動かない orz
 - NATも問題の1つに
- 通信と放送の融合というけれど、
 - IntServ(RSVP), DiffServ とともに QoS保証は無理
 - メッシュ型P2P+ユニキャスト(OLM, ALM)が現実解?
 - BBブロードキャスト、ウタゴエ Ocean Grid

認証・セキュリティも

- 後追い
 - IPsec, SSL, TLS, PGP, PKI...
 - 認証はユーザにかけたいのか、IPにかけたいのか？
 - PKIは良いとして、ほとんどの人は公開鍵を持っていない
- モバイル環境では、IPsecは意味をなさない
 - SA(Security Association) が相手ノードのIPアドレスを使用するため
- 新たな原理原則が必要
 - アプリケーションを視野に置くことも



さて、問題の核心は

- つまるところ、アーキテクチャが古いため
 - 1970年代の技術が使われ続けていることが奇跡
- 狭いアドレス空間（いうまでもなく）
- **IPにおける identifier と locator の概念の混在**
 - IPアドレスの二元性(duality of the IP address)問題
 - 物理アドレスが論理アドレスの機能も有している
 - モビリティ、マルチホーム、IPsec に支障が...
- ルーティングプロトコル
 - 全ノードが同一の情報を集めて同じ動作をする**分散型集中制御**
 - 大域経路表の破綻(スケーラビリティに欠ける)
 - Dijkstra法(経路計算)の計算量はノード数 N として、 $O(N^2)$
 - 耐故障性、トポロジー変動への弱さ
- QoS, マルチキャスト経路制御
- 時刻同期

もっと大事な問題は

- 非持続的な進化発展性

- すでに社会インフラとなっているインターネットをそっくり変えることは不可能
- IPv6の失敗がよい教訓
 - 段階的に変えて行こうとすると、デザインがどうしてもきたなくなる。デザインを優先したのが、IPv6だった



- 悪い部分は変えていきたい。
けれど、それもできないジレンマ

- 新規技術に対する許容性が望まれる

- 科学的プロセスと技術的プロセスを繰り返し、技術をブラッシュアップさせていくための**フィールド**と、
- 実証された技術を取り入れることができる**柔軟なアーキテクチャ**

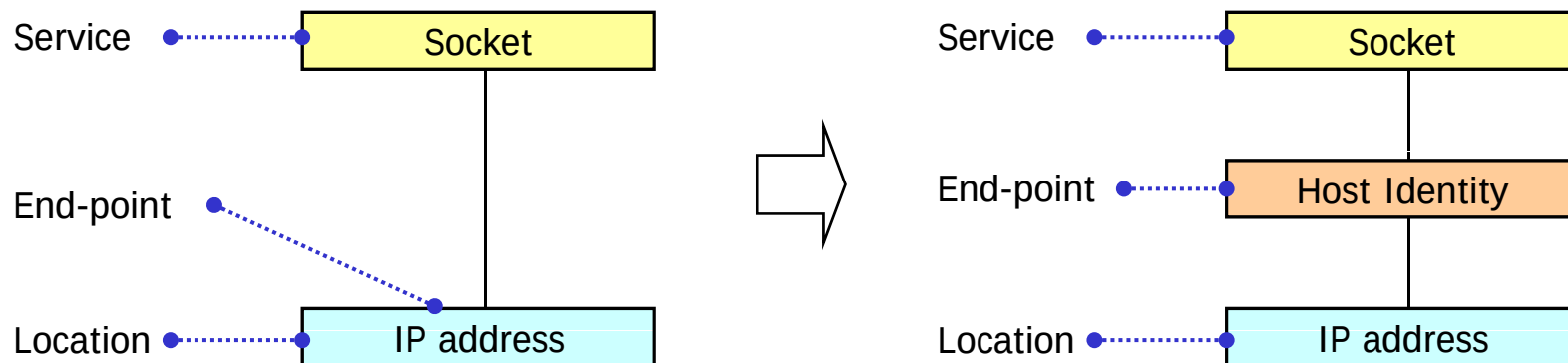
ネットワークのあるべき姿(私見)

- IDにおける end-to-end透明性
 - locatorが変更されても、シームレスにハンドオーバーする
- 現実結合原則
 - ネットワーク上のエンティティと現実社会の乖離を防ぐ
 - 双方向認証(IDベース)
 - 地理的位置情報のノード付与も重要
 - 無線アドホック系ではルーティング自体に関係する
- 自己xx的 ⇒ 持続性
 - 物理的なネットワークを選ばないことも含め
 - IP、IPv6、アドホック系、ZigBee...
- guarantee と best effort の区別
 - トラヒック計測も重要
 - この整理が出来た上でNGNの価値がでる

ID/Locator分離

ID/Locator分離

- IPアドレスの二元性問題 (duality problem of the IP address)
 - IPアドレス (or FDQN) が、ノード自身の識別子 (ID) と接続位置 (locator) の両方の意味を有している
 - モビリティ、マルチホーム、IPsec に支障を生じる
- ノードの個体 (identity) と接続位置 (location) の概念を分離し、ID と locator を別レイヤで扱う必要がある



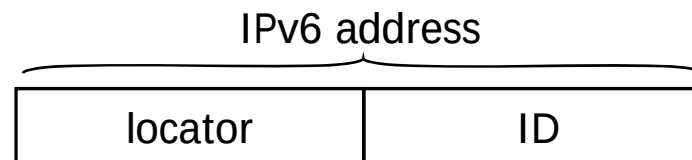
HIPにおけるID/Locator分離の例

IPにおけるID/Locator分離の概要

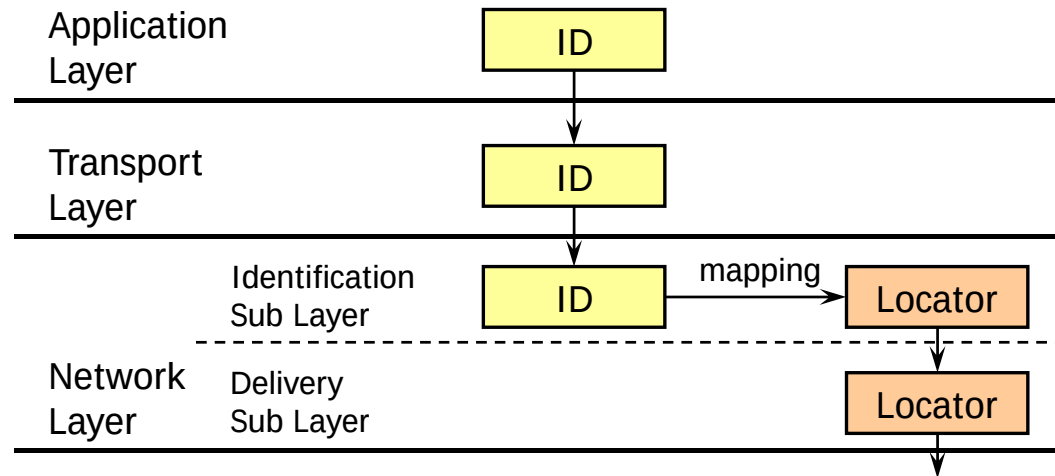
- IRTF, IETFにて検討が進められている
- 直近の問題として、マルチホーミングに伴う経路表の増大がある
- 提案方式
 - LINA(Location Independent Network Architecture), LIN6(Location Independent Networking for IPv6)
 - HIP(Host Identity Protocol)
 - LISP(Locator Identifier Separation Protocol) などなど
- いずれも、ID解決にネームサーバを用いる
 - ネームサーバは、(ID, locator)の組を管理する

LINA, LIN6

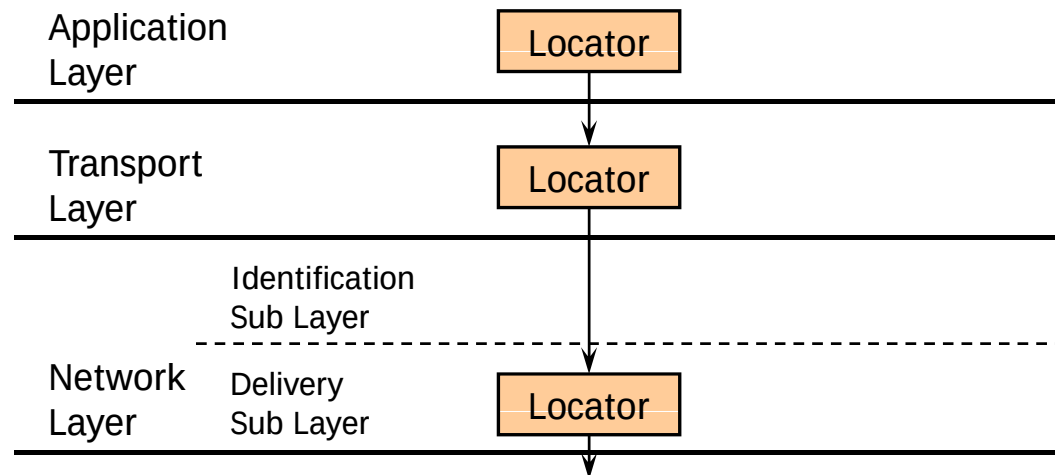
- アプリケーションは、送信先ノードの指定に、IDとlocatorの両方を利用可能
- Network層を2つの sub layerに分離
 - Identification sub layer と Delivery sub layer
 - Identification sub layer で、IDから locatorに変換 (ID解決)
- MA(Mapping Agent)がネームサーバとして機能
- MAが、Identification sub layerからの要求に応える形で、ID解決を実現
- LIN6 は、LINAのIPv6への適用
 - IDを40bitで表現し、24bitのprefixをつけて、下位64bitに埋め込む



LINAの処理フロー



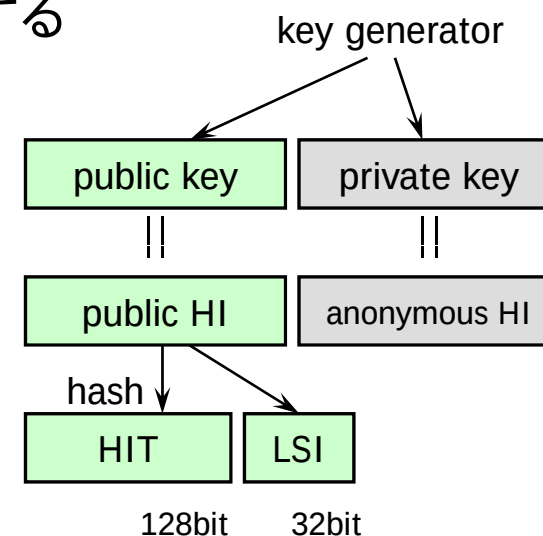
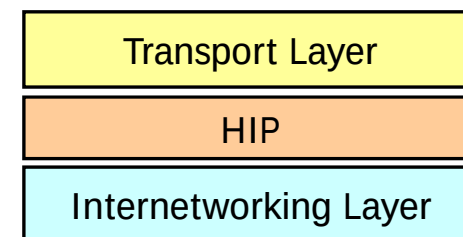
相手ノードがIDにより指定された場合



相手ノードがlocatorにより指定された場合

HIP Architecture

- TransportとNetwork層の間にID解決のための層を設ける
 - LINAと同じ
- 主な特徴
 - IDのために独立の空間を定義する
 - LIN6の場合は、IPv6アドレスに埋め込み
 - IDに認証と匿名性のための機能を付与する
 - このため、2種類のID(HI, HIT)を持つ
- ID解決のためのネームサーバとして、**DHT**(分散ハッシュ表)が検討されている
 - (ID, locator) の組をDHTにより管理
 - 認証機能により、**改竄**を防ぐ
 - secure-i3
 - i3(Internet Indirection Infrastructure)を拡張



既存方式の問題

■ 実現性

- LIN6においては、IPv6の普及が前提
- HIPでは、ホストにおけるプロトコルスタック、LISPでは、ルータの置き換えが必要
- 実現には時間とコストを要する

■ ID解決のオーバーヘッド

- ネームサーバを用いるため、オーバーヘッドを生じる
- ハンドオーバー時にID解決を要する場合、通信断が長くなる

■ セキュリティ強化の必要性

- 不正なエントリ参照や改竄を防ぐため、ノード認証の仕組みが必要

■ DHT自体の持つ技術課題

- 通信頻度の高いノードのID解決エントリを保持するDHTノードの参照負荷が高くなる ⇒ DHT内の負荷が均等でない
- churn状態(DHTノードの出入りが頻繁)と複製の持ち方について緻密に管理する必要がある
⇒ 検索成功の保証が困難

ネットのトレンドは、Web2.0 から Cloudへ

Web2.0

- 元は、ティム・オライリーの客寄せネタ

- 構成要素

1. The web as platform
2. Harnessing collective intelligence
3. Data is the next Intel inside
4. End of the software release
5. Lightweight programming models
6. Software above the level of a single device
7. Rich user experiences

- 砕くと

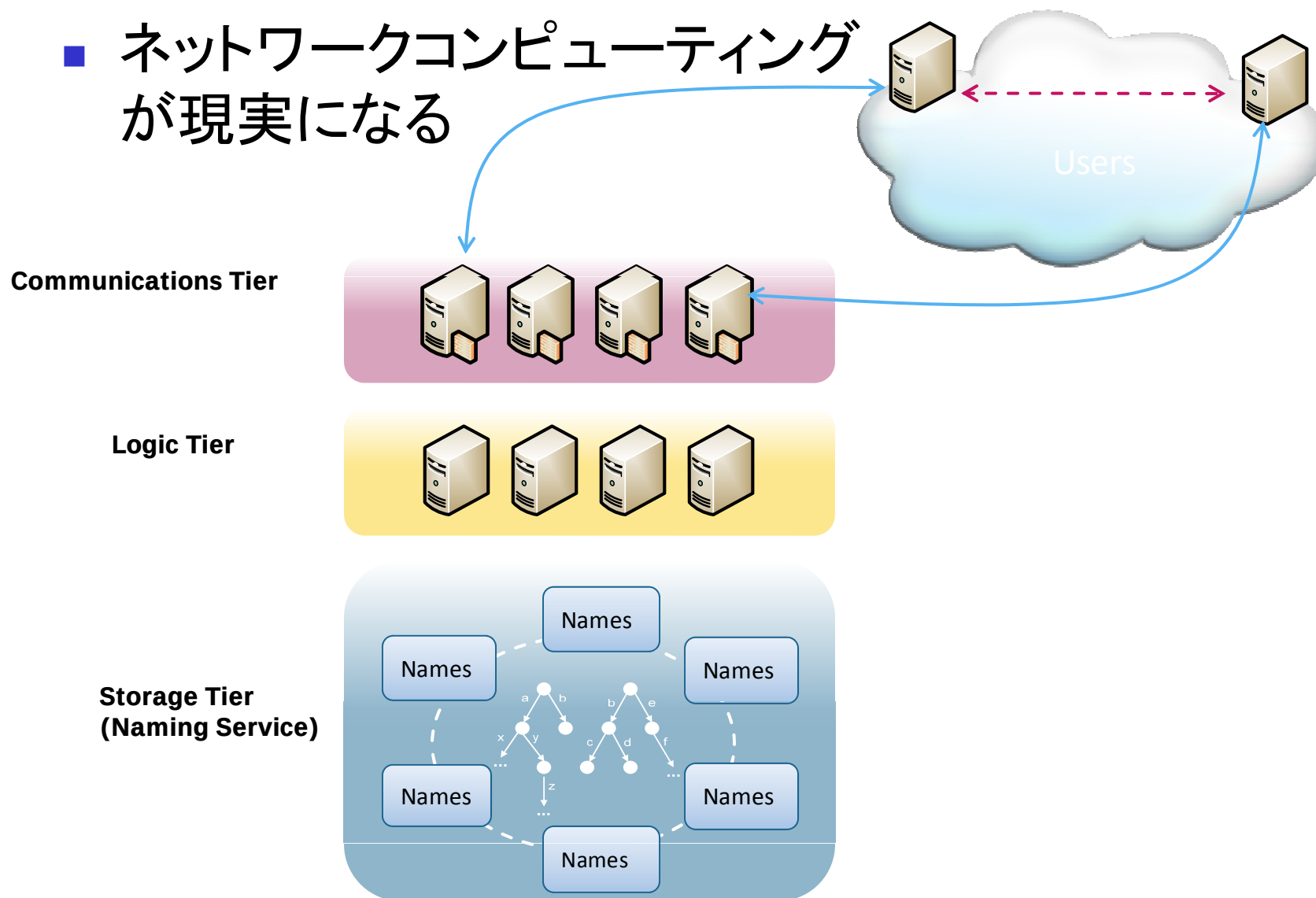
- ユーザの手による自由な情報整理
- ユーザ参加型、同時にユーザは貢献者
- Ajax, Web API (マッシュアップ), RSS 等の技術

Web2.0

- インターネットの民主化の流れと解釈
(やや金儲けネタになってるが...)
- 悔しいけど、的を得ている
- ポストWeb2.0として、「先進的Webサービス」といった検討も始まっている
 - Semantic Web Service, SaaS, SOA,...
- 要は、Googleバンザイ!?
 - 本当は、ちょっと違う
 - Googleは Cloudに転じた

Cloud の基本構造

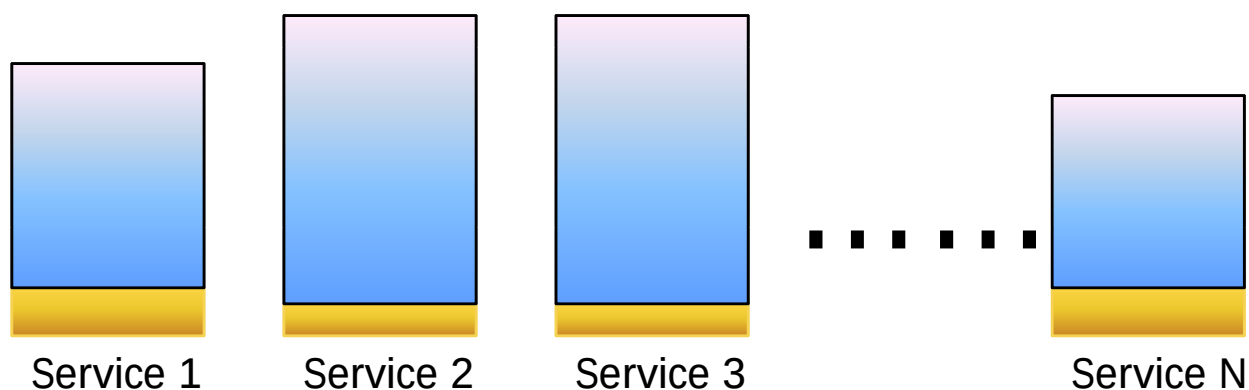
- ネットワークコンピューティングが現実になる



※MS: Windows Azure 資料より

Cloud OS まで登場してきた

- たくさんのサービスを24時間、365日稼働させるための、新たなOS



Windows® Azure™

※MS: Windows Azure 資料より

かたや、P2Pの方は...

P2Pを取り巻く変な前提 (1)

■ 帯域を不法に食う

- 帯域を特定の個人が使うのがなぜ悪いのか？
 - ネットにおける利用形態が多様化し、偏りは自然な話
 - タリフの方で対応するのが筋
- これとP2P形態のサービスをリンクさせて、P2Pが悪者であると言っている
- 例えば、複数のCDNを使って、大量のコンテンツをDLするのは構わないのか？

■ 著作権侵害、情報漏えい

- インターネット、さらには広く共通する問題
- むしろ、P2PはDRMとの相性が高く、さらにデリバリ能力が飛躍的に高い。コンテンツホルダには歓迎されるべきシステム形態であったはず

P2Pを取り巻く変な前提 (2)

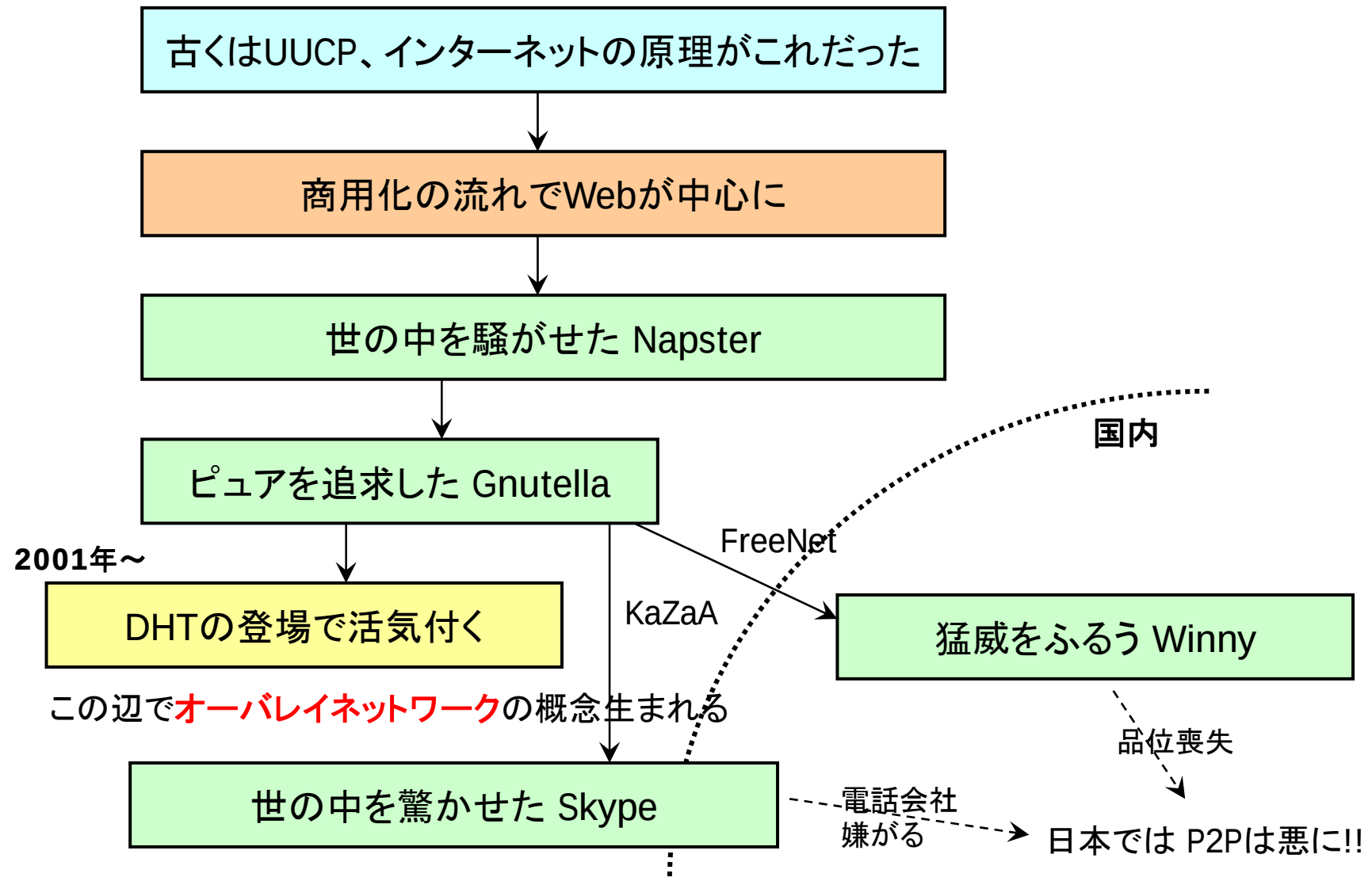
- 司法的に問題解決しないといけない
 - 法律の問題は、「キャッシュが違法コピー」という部分。
この影響範囲はP2Pにとどまらない
 - なぜ、P2Pだけを相手にする？
- 金子氏を擁護する、一部の公的団体
 - 「技術者に責任を負わせる現状の法律実務は誤り」とする。とあるが、
 - 戦う相手が違う。どこが問題なのか不明確
 - 法律が悪いとするなら、Winnyは正しいのか？
 - マスコミは喜んで取材するが、世の中の評価は逆
 - 特にベンチャー企業の痛手は深刻
 - みんな Web2.0へ

P2Pを取り巻く変な前提 (3)

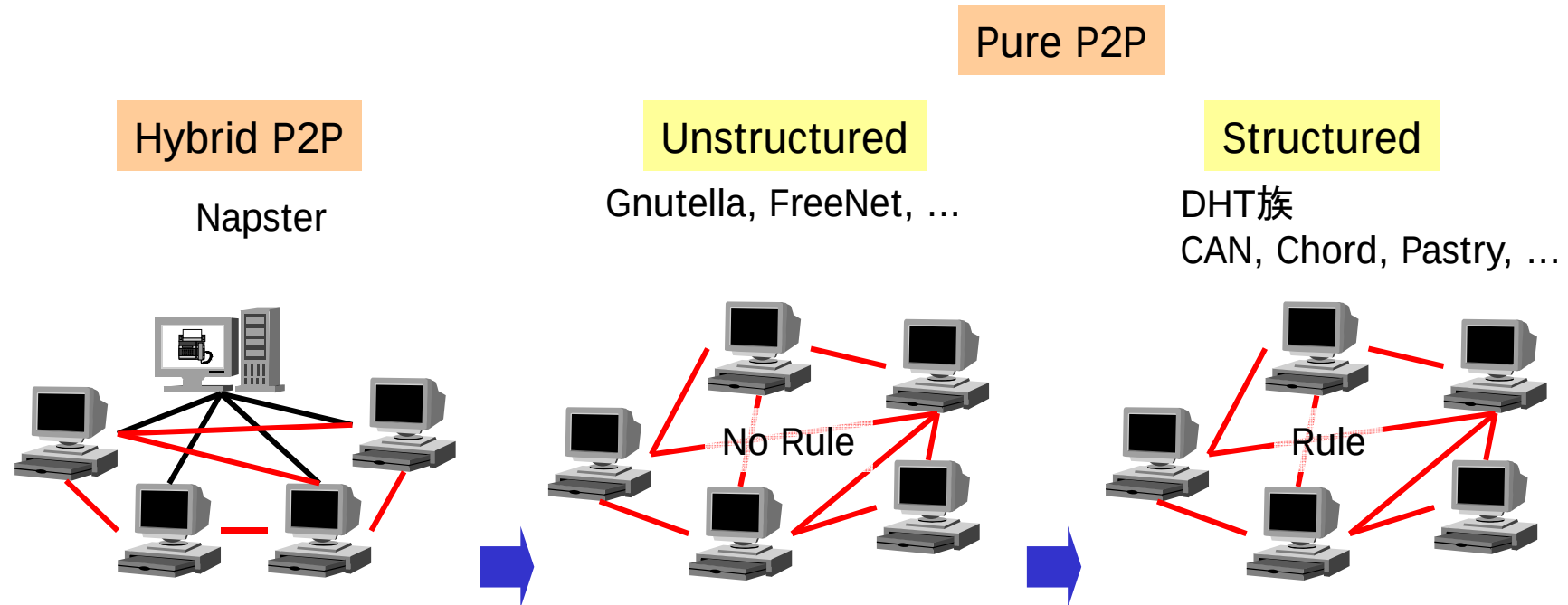
- P2Pは、ファイル共有システム？
 - Winnyとファイル共有システムとP2Pがごちゃまぜになっていないか
 - オーバレイネットワークという概念分離も見当たらない
 - 別形態のP2Pシステムを開発している立場から見て、いつもピンとこない前提

構造化オーバーレイって知ってます？

インターネットとP2P



P2Pシステムの進化



Overlay Network と呼ぶ
アプリケーション層でトポ
ロジを構成する論理的な
ネットワークのこと

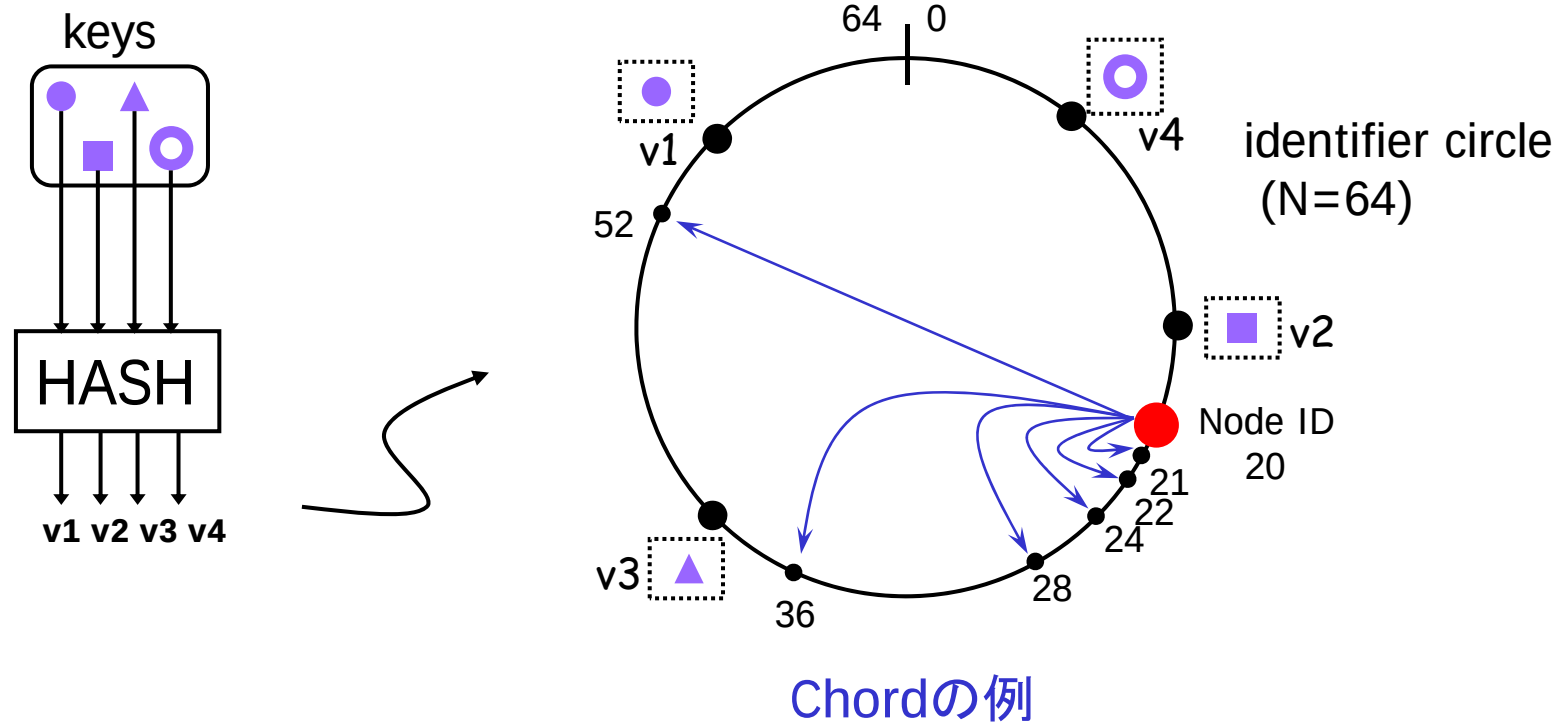
構造化オーバーレイ(Structured Overlay)

- ノード間のトポロジに一定の数学的規則がある
 - これにより、効率の良いリソース発見や経路制御が実現できる
 - 理論的解析やシミュレーションにより、計算量、通信量を導出できる
 - 広域、大規模なネットワークに対して、可用性とスケーラビリティが保証されている(高性能、高信頼性)
- 完全分散型のルーティングプロトコルと思えばよい
- 代表的なサービスモデルに DHT (Distributed Hash Table: 分散ハッシュ表) がある

	ノード数	1000	100万	10億
Chord	検索ホップ数	10	20	30
	テーブルサイズ	10	20	30
Pastry (d=4)	検索ホップ数	2.5	5	7.5
	テーブルサイズ	37.5	75	112.5

DHT

- (key, value) の組を $\text{hash}(\text{key})$ を近傍のIDを持つノードに管理させることにより、分散ハッシュ表を実現
- 指定されたIDを持つノードを高速に探し出す
- IDは通常128または160bit整数



DHTとその応用

■ 方式例

- CAN, Chord, Tapestry, Pastry, Kademlia,...
- ring方式、tree方式、hypercube方式、XOR方式などなど

■ 応用例

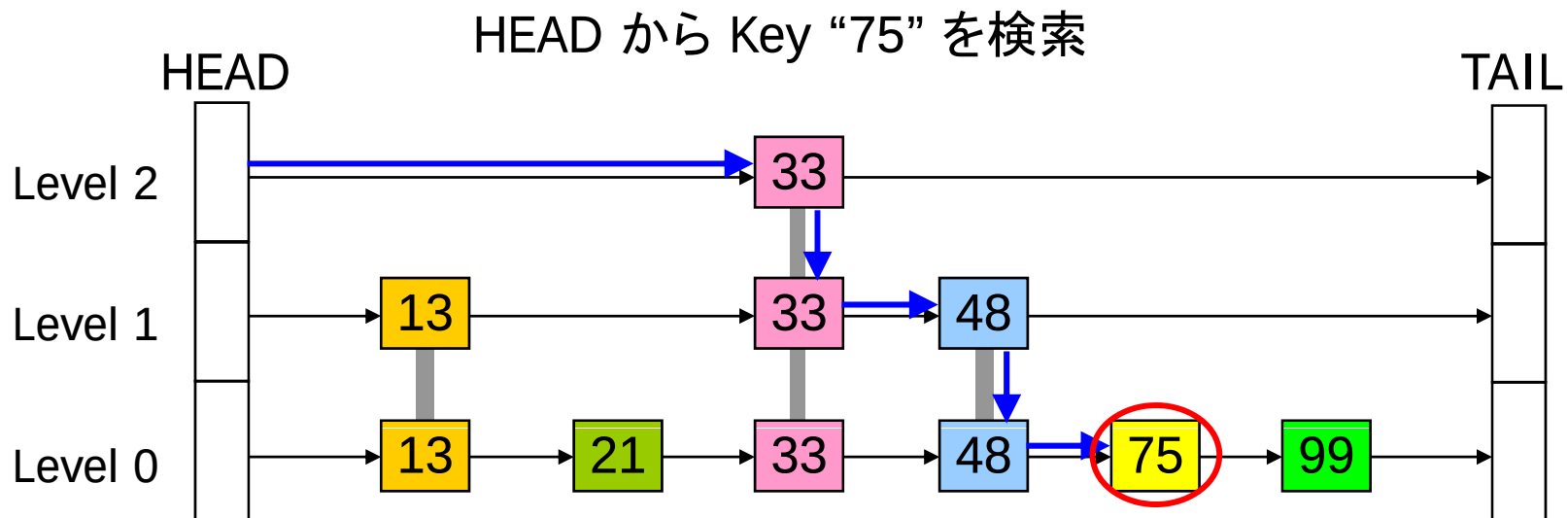
- 分散ファイルシステム
 - PAST, OceanStore, Shark, Glacier,...
 - 9.6倍のコピーを持つことで、60%のノードダウンに対して、99.9999%の可用性を確保できる (Glacier)
- ALM (アプリケーション層マルチキャスト)
 - Bayeux, Scribe, SplitStream, Overcast,...
 - 配信木を分散管理できる
- 評価のためDNSも
 - 現状、35%がDNS管理不備による無意味なメッセージとなっている
 - 管理不要、高い負荷分散効果が期待できる

■ 実験フィールド

- PlanetLab ... 全世界に800を超えるノード

Skip Graph

- 構造化オーバレイ界のニューウェーブ
- DHTよりさらにパワフル
 - 任意のkeyを持つオブジェクトをサーチ
 - 範囲探索
- 海外でもマイナー、国内で知っている人は皆無(だった)
- 基本とするデータ構造は skip list (W. Pugh, 1990) という新
手のバランスト



i3 (Internet Indirection Infrastructure)

- Chordを用いた間接参照のためのインフラ
- IDから物理アドレスへの変換だけでなく、multicast, anycast も実現
- 論文
 - Stoica, I., Adkins, D., Zhuage, S., Shenker, S., and Surana, S. Internet indirection infrastructure. In Proceedings of SIGCOMM (August 2002), ACM.
- IDと物理アドレスの組(trigger)を DHT によって管理することで、mobility, multicast, anycast に対処する
- この研究成果は、KBR(API, フレームワーク)の形にまとめなおされる

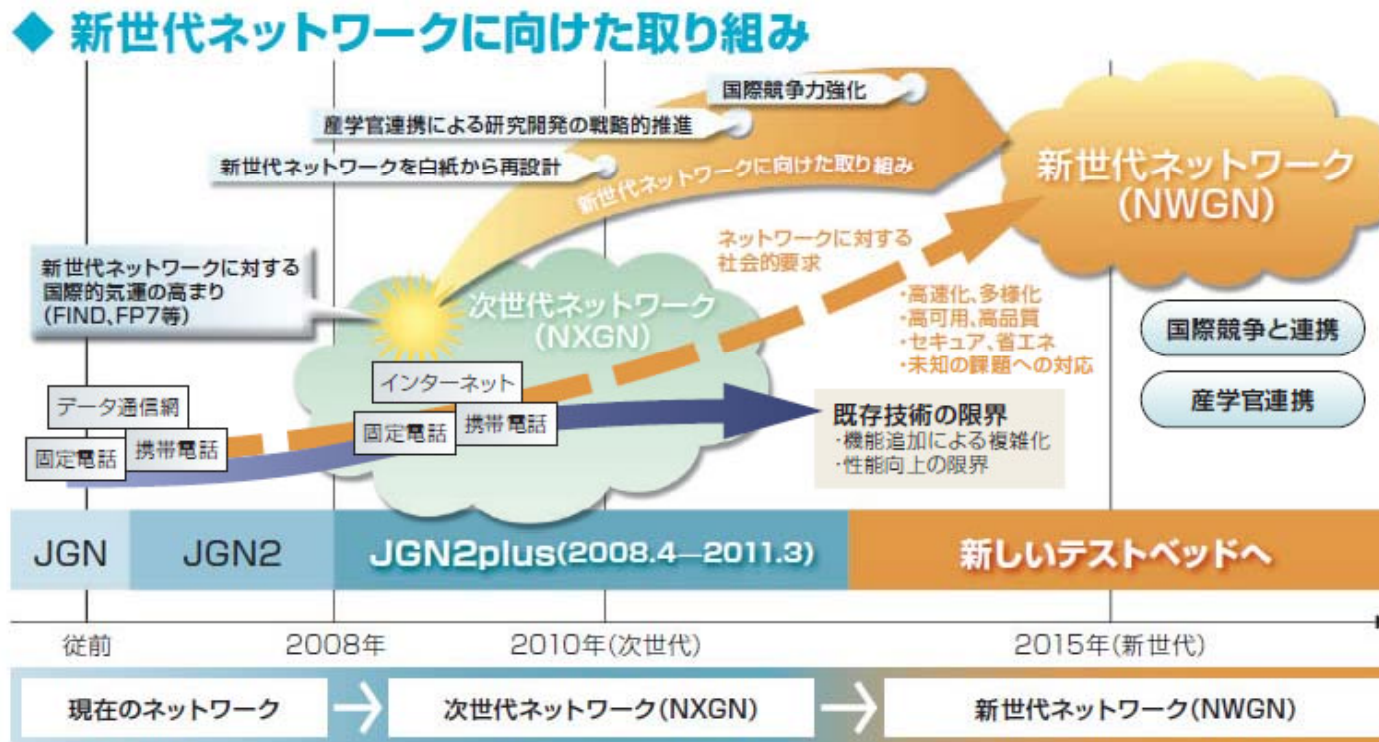
インターネットを救う P2P

- 新世代ネットワーク
- ID/Locator分離

新世代ネットワーク(NwGN)

新世代ネットワーク

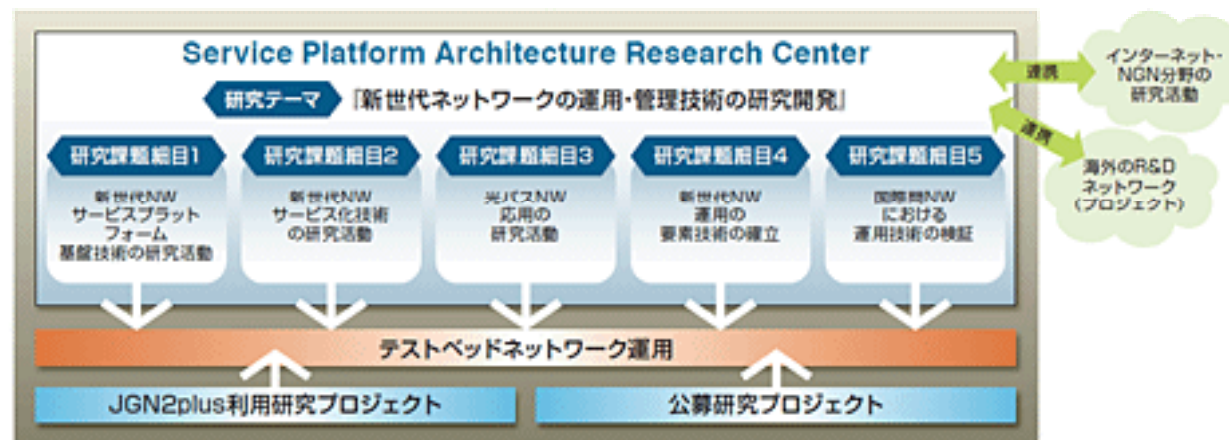
- 海外に先駆けて、将来に備え、新たなアーキテクチャを持つネットワークを実現する
 - アメリカ: GENI/FIND, 欧州: FP7/FI



※NICT, JGN2plus HPより

新世代ネットワーク

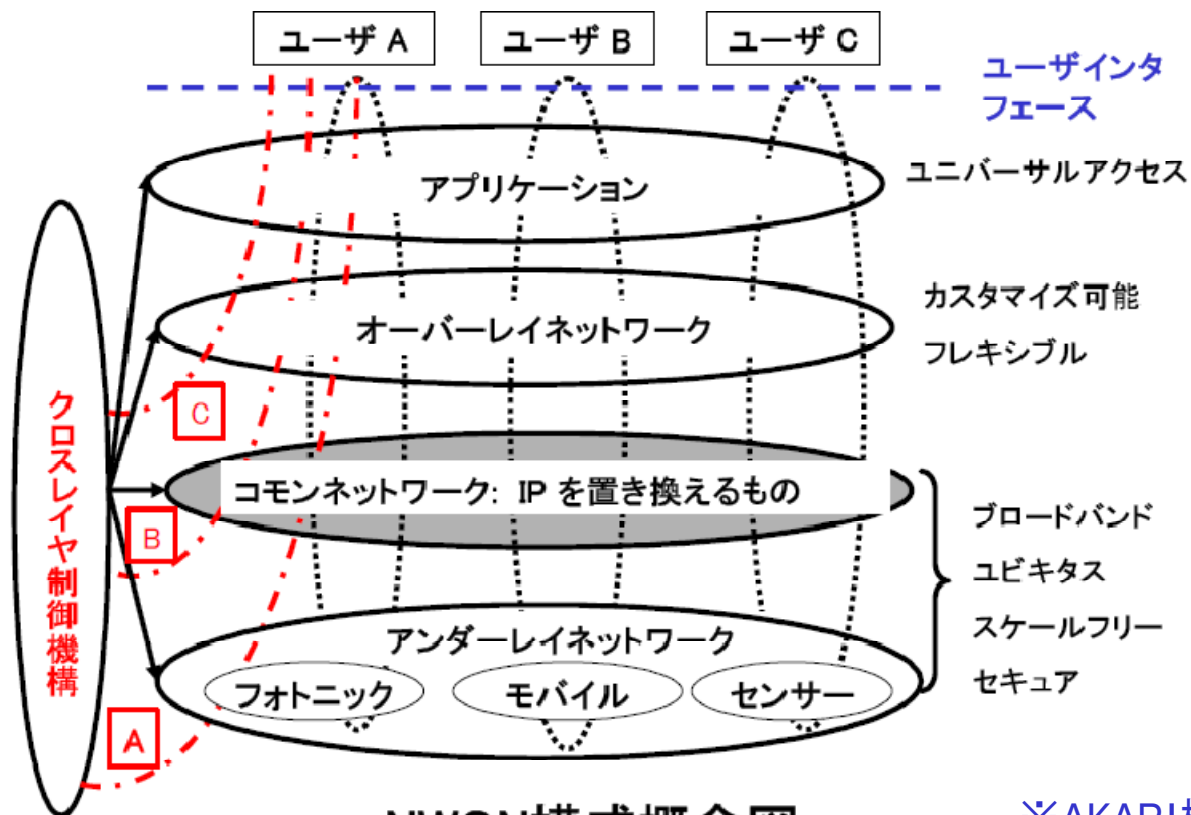
- NwGN (NeW Generation Network)
 - NGN (Next Generation Network) と全然違う
 - その心は、Network for Next Generations
 - 故平原さんの個人的見解
- NICT(情報通信研究機構)が推進を担当
 - AKARIプロジェクト
 - JGN2plus SPARC



※NICT, JGN2plus HPより

NwGNのキモ(私見)

- 薄い共通NW層, クロスレイヤ制御
- 持続的(Sustainable)
- Paths と Packets の統合

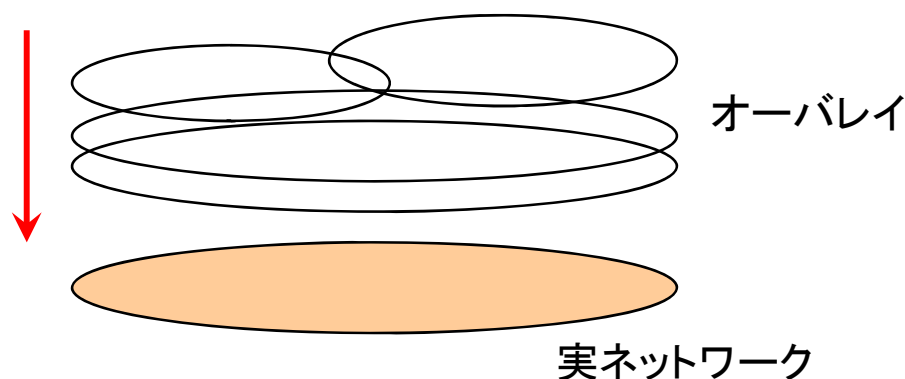


NWGN構成概念図

※AKARI概念設計書より

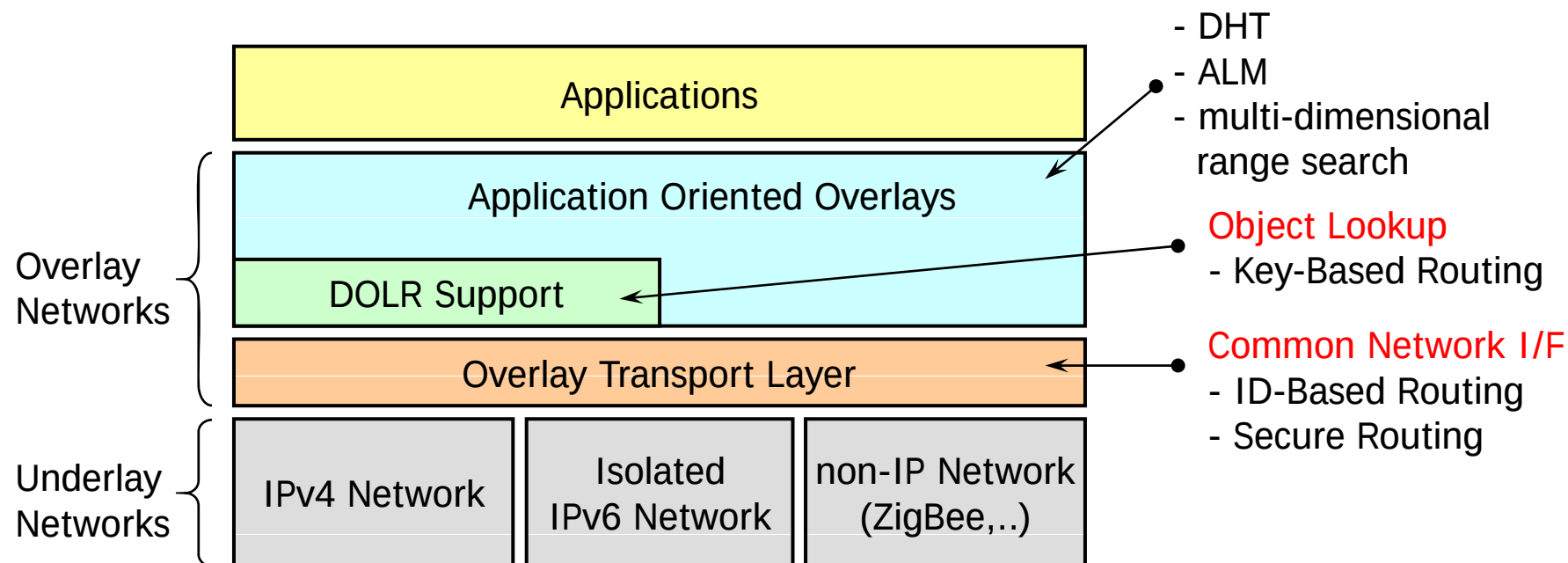
オーバーレイネットワークの意味

- NwGNで新しいアーキテクチャを考える鍵
 - オーバレイは最初から Clean Slate
 - 構造化オーバーレイでは、IDでルーティング
- 持続性は、新陳代謝により生まれる
 - その時代その時代で、もっとも良くできたオーバーレイ技術が実ネットワークになる
 - 進化の過程では、新たなオーバーレイが実ネットワーク上で開発される



オーバーレイが1つの解になる

- IDとlocatorの完全分離
 - 透過性も(NAT越え)
- IPv6, non-IPネットワークの統合
- 持続的進化
 - Overlay と Underlay の進化。ex. 光スイッチ技術



社会を変える P2P

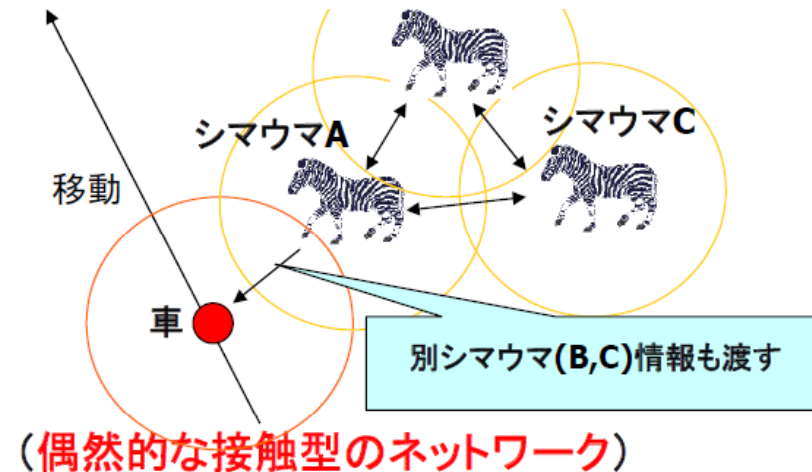
- センサーオーバレイ
- DTN
- Smart Grid

センサーオーバレイ

- 地球環境、人間社会のモニタリング
 - 気象観測、土壌汚染の観測、災害検知など
 - WIDEでは、Live E! プロジェクト
 - ゲリラ豪雨に効果の報告も（2008/9/19 Live E! シンポジウム）
- UCI (University of California, Irvine) で聞いた話
 - 約7年前
 - アメリカでは山火事が深刻な社会問題
 - ヘリコプターでの監視は発見が遅れる
 - 小さなセンサデバイスを開発して、山中に撒き、ボヤを早期発見できるようにできないか
 - 起こった研究
 - 塵のようなセンサ(センサダスト)をナノテクで開発する
 - 1000億を越えるノードにより自律構成される Deep Network

DTN

- Delay/Disruption/Disconnect Tolerant Network
または、Delay Tolerant Network
 - 劣通信環境でなんとか通信を成立させる
 - 惑星間通信, 深海通信, 戦場でのアドホック通信
- 事例
 - ZebraNet
 - 3,5000頭のシマウマの生態情報を収集する(ケニア)
 - 気球ネット
 - 災害時に気球を飛ばし、一時的なネットワークを構築
 - DUMBOプロジェクト
 - 象さん+Live E! (タイ)



DTN + OLPC

- OLPC (One Laptop per Child)
 - 開発途上国の子供たちにノートPCを
 - 教育ツールにより、貧困を解決する糸口を
 - けれど、ネットワーク環境は問題
 - OLPCは、無線LANのアドホックモードで通信できるけど
- OLPC に DTN を足すと何が起こるか
 - もう解りますよね :-)
 - 地理的経済的 Digital Divide の解消
 - ネットに未開も先進もない



太陽光発電とP2P

- 太陽電池は進化と普及の一途にある
 - 集中型化石燃料発電の時代から、分散型太陽光発電の時代に
 - 電力の生産と消費の細かな制御が必要
- SmartGrid
 - アメリカで起こっている電力系統制御技術
 - 持続的エネルギー消費社会に向かうための基礎技術
 - 「コンピュータと電話のような「進化」が電力網で起きる」
 - <http://business.nikkeibp.co.jp/article/pba/20080917/170808/>
 - P2Pが技術的に適していることは言うまでもない

Android が空を飛ぶ意味

これまでの梃子の原理

■ グローバル

- 今のネットビジネスを牽引する原理
- マーケットの拡大、ユーザ(会員)の獲得・拡大などなど
- 先駆者利益の構造、ビッグ3しか生き残れない

■ ネットは情報をグローバルに配信し、富を拡大させる梃子として活用されてきた

- なんとなく言えば、みんながみんな同じテレビ番組を見ているようなもの
 - 一部のタレントの知名度が上がり、富がそこに集中する
- 効率と利便性の追求がそのままの形で現れている
- 技術的には、画一的でちっとも高度じゃない

■ さて、ネットってそれだけの道具だろうか？

持続的社會と地産地消

- 世の中の経済はグローバルだけでは回っていない
 - 私たち個人個人は時間も空間も、現実にはbindされて生きている
 - エネルギー消費、情報交換、基本的な生活、趣味嗜好、何をとっても、基本はローカリティ
- ネットの基本は「つなぐ」という機能
- 特に、物流、エネルギー問題に着目すると、問題を解く鍵は、「持続的社會」と「地産地消」しかなく、ネットがその主要な役割を果たす
- こちらは人間性があって、高度な技術を要する

Cloud vs. P2P (1)

- Cloudで提供されるサービスの幅とモバイルデバイス同士が相互に提供し合うサービス(つまりP2P)の幅を考えた場合、ポテンシャル的にP2Pはとても大きい
- 例えば、今はやりの位置情報や地図連携のアプリ
 - ロケーションの取得や周辺サービスの検索、個人の嗜好とのマッチングなど
 - 最近は、セカイカメラのようなものも出てきた
 - 次の時代に必要なのは、**現実とバーチャルのbinding**
- 見てると、Cloud的な問題解決方法が取られているみたい。けれど、P2Pの方が圧倒的にやりやすくスケールする

Cloud vs. P2P (2)

- ネットの接続形態から来る制限
 - 携帯にせよ、インターネットにせよ、デバイスからの接続形態は、C/Sモデル
 - IPv4の枯渇、キャリアによるNATの多重化とこの傾向はますます進む
 - 「近くの友達より遠くのWebサーバの方がすんなりつながる」のが今のネット
- 携帯網は(ry まあいろいろあるとして、インターネット自体が end-to-endのメッセージ到達性を失っていて、C/Sモデルしかない
- その制限下で、Cloudはとても親和性がよくて、どんどん活用されていくだろう (...けど)

Cloud vs. P2P (3)

- 失われたメッセージ到達性を取り戻す働きをするのがP2P技術（正確にはオーバーレイ技術）
- こちらの技術が確立した時には、Cloudの役割は縮小する（なくなりはない）
- その主役となり、時代を牽引するのがAndroid
- Android が Cloudの端末というのは仮の姿。
Androidは、長期的には、Googleのビジネスモデルを破壊しかねない
- （なんと微妙な話でしょ）

がんばれ!! 職人たち

- ユーザはいつの世でも選択の自由を持っている
- また、自由な気概と気風はどんなでもないイノベーションを起こす(ここに資本力に対抗する鍵がある)
- デバイスの標準化とOSSは、固定したネットの世界を破る
- 問題解決のための技術を持っていることは強い
- 組み込み技術者のポテンシャルは大きい
 - 例えば、Android + OSSだと、プロトコルスタックをすげ替えるのは簡単な話