3.4 World Wide Web

もともとは、完全な分散システムだった。
成功が始まると、集中化してビジネス化された。
再び、分散化が起こりつつある。

WWWの生い立ち

- 1989 CERN (European Center for Nuclear Research) で文書管理のために考案
  - 他の文書へのリンクを含むテキスト（HyperText）
  - HTML
- 1993 Mosaic
  - クリックするだけでリンク先にアクセスすることのできるグラフカルユーザインタフェース
- 1994 World Wide Web Consortium
- Netscape Navigator vs. Internet Explorer

URL (Uniform Resource Locator)

スキーム名://ドメイン名:ポート番号/ファイル名

例：http://www.aida.k.u-tokyo.ac.jp:8080/index.html

- 「:ポート番号」、「/ファイル名」は省略可能
- 「:ポート番号」を省略した場合にはスキームごとの既定値が使用される

スキームの例

スキーム名	プロトコル	ポート番号
http	HTTP	80
https	暗号化されたHTTP	443
file	そのコンピュータ上のファイル	-
ftp	ファイル転送プロトコル	21
mailto	電子メール	25

クライアント

サーバ

80番ポートにコネクションを設定 →

```
GET /~aida/koujichu.html HTTP/1.0
```

←コマンドの終わりを表す空行

←

```
HTTP/1.1 200 OK
Date: Thu, 21 Apr 2005 02:41:27 GMT
Server: Apache/1.3.27 (Unix)  (Red-Hat/Linux) mod_ssl/2.8.12
OpenSSL/0.9.6b DAV/1.0.3 PHP/4.1.2 mod_perl/1.26
Last-Modified: Thu, 21 Apr 2005 02:17:43 GMT
ETag: "beedd-9d-42670d47"
Accept-Ranges: bytes
Content-Length: 157
Connection: close
Content-Type: text/html

<HTML>
  <HEAD>
    <TITLE>工事中</TITLE>
  </HEAD>
  <BODY>
    <IMG SRC="koujichu.gif" ALT="工事中">申し訳ありませんがただいま工事中です
  </BODY>
</HTML>
```

←ステータス

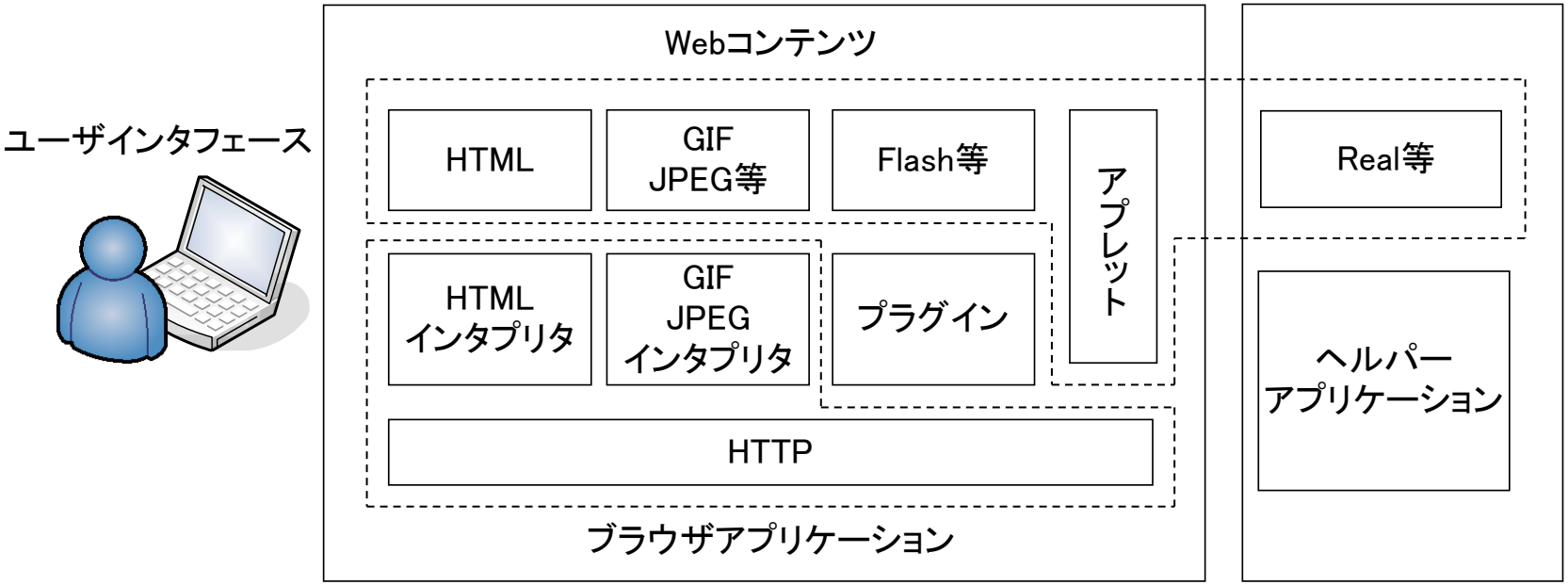
←ヘッダと本文を区切る空行

コネクションを切断 ←

ブラウザの構造

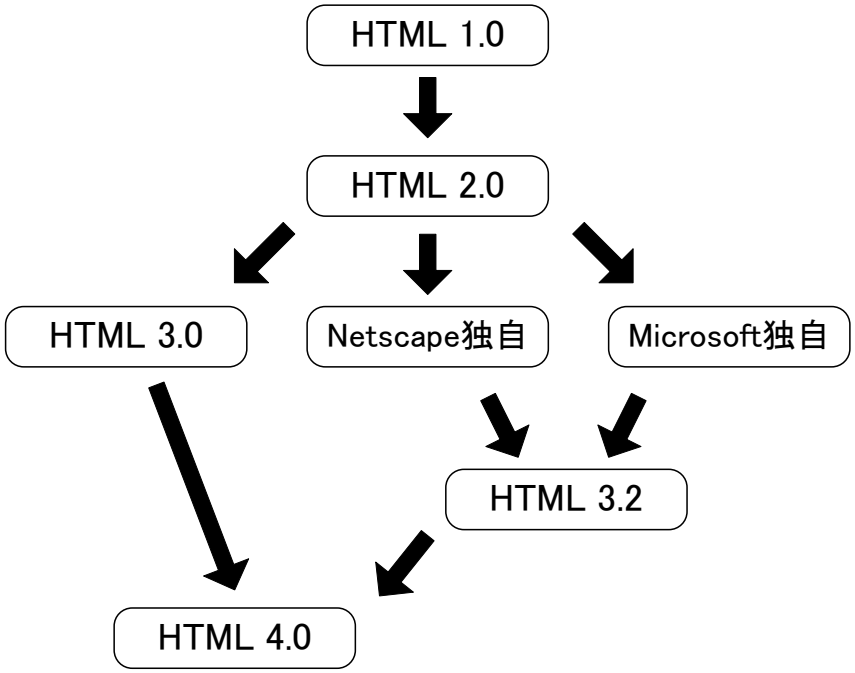
- 第1世代：全てのコンテンツをブラウザ自身が処理
- ヘルパーアプリケーション
  - 別ウィンドウで動作
- プラグイン
  - ブラウザウィンドウ内で動作
- アプレット
  - 実行可能なプログラムをコンテンツとしてダウンロード

ブラウザの構造


ユーザインタフェース
Webコンテンツ
HTML
GIF
JPEG等
Flash等
アプレット
Real等
HTML
インタプリタ
GIF
JPEG
インタプリタ
プラグイン
ヘルパー
アプリケーション
HTTP
ブラウザアプリケーション

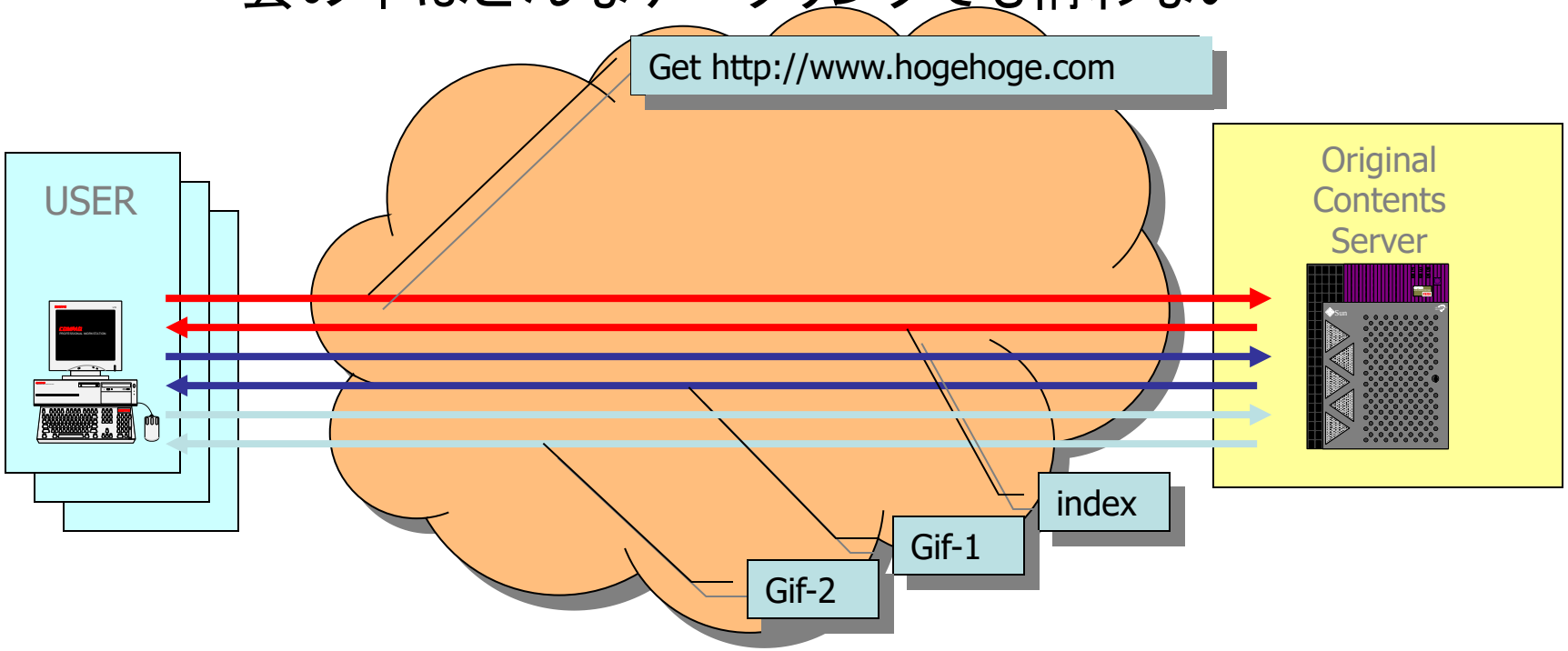
HTML (HyperText Markup Language)

- 文書構造を記述するためのSGML (Standard Generalized Markup Language)をハイパーテキスト記述に応用
- HTML5 へ



通常のWEBコンテンツの流れ

“雲の中はどんなデータリンクでも構わない”



一度のhttpリクエストに際し、ファイル形状ごとにそれぞれTCPセッションの確立から始めなければならないため、アクセス数の増加に対しサーバの負荷が指数関数的に増加する

要は、大変たくさんの TCP コネクション

タグ

<タグ名 属性名=“属性値”>文字列</タグ名>

開始タグ　　終了タグ

- タグ名、属性名の大文字／小文字は区別しない
- 終了タグは省略されることもある

基本的なタグ

タグ名	意味	属性など
H1～H6	見出し	数字の小さい方が大きな見出し
P	段落	align=“right”など
BR	改行	終了タグなし
I	イタリック	
B	ボールド	
FONT	フォント	size, face, colorなど
IMG	画像埋め込み	src=“ファイル名” alt=“代替文字列” width, height、終了タグなし

記述例

<H1>レベル1の見出し</H1>
<H2>レベル2の見出し</H2>
<H3>レベル3の見出し</H3>
入力本文中で改行しても画面上では改行されないので
画面上で改行したい場合には<P>や
を
使う必要があります。
<P>
<B>と</B>で囲まれた部分は<B>ボールド</B>で表示されます。

<I>と</I>で囲まれた部分は<I>イタリック</I>で表示されます。

<U>と</U>で囲まれた部分は<U>下線付き</U>で表示されます。

C:¥sample1.htm - Microsoft Internet Explorer

ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り リンク

アドレス(D) C:¥sample1.htm 移動

レベル1の見出し

レベル2の見出し

レベル3の見出し

入力本文中で改行しても画面上では改行されないので画面上で改行したい場合には<P>や
を使う必要があります。

<B>と</B>で囲まれた部分は**ボールド**で表示されます。
<I>と</I>で囲まれた部分は*イタリック*で表示されます。
<U>と</U>で囲まれた部分は下線付きで表示されます。

ページが表示されました マイ コンピュータ

ハイパーリンク（アンカー）

- <A href=“*URL*”>文字列</A>
- <A href=“ファイル名”>文字列</A>
- 多くのブラウザでは文字列の部分の色を変えてアンダーラインを引くことでハイパーリンクの存在を意識させる

HTML文書の全体構造

```
<HTML>
  <HEAD>
    <TITLE>タイトル</TITLE>
  </HEAD>
  <BODY>
    文書本体
  </BODY>
</HTML>
```

リスト、表

タグ名	意味	属性など
UL	Unnumbered List	type=“disc”/”circle”/”square”
OL	Orderd List	type, startなど
LI	List Item	リストの項目
TABLE	Table	borderなど
TR	Table Row	rowspan(行をまたがる)など
TD	Table Data	align, valign, rowspan, colspan, nowrapなど
TH	Table Header	ボールドでセンタリング

<UL>
<LI>
<B>と</B>で囲まれた部分は<B>ボールド</B>で表示されます。
<LI>
<I>と</I>で囲まれた部分は<I>イタリック</I>で表示されます。
<LI>
<U>と</U>で囲まれた部分は<U>下線付き</U>で表示されます。
</UL>
<TABLE BORDER>
<TR><TH>タグ名<TH>意味<TH>属性など
<TR><TD>TABLE<TD>Table<TD>borderなど
<TR><TD>TR<TD>Table Row<TD>rowspan(行をまたがる)など
<TR><TD>TD<TD>Table Data<TD>align, valign, rowspan, colspan, nowrapなど
<TR><TD>TH<TD>Table Header<TD>ボールドでセンタリング
</TABLE>

C:¥sample2.htm - Microsoft Internet Explorer

ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り リンク

アドレス(D) C:¥sample2.htm 移動

- <B>と</B>で囲まれた部分は**ボールド**で表示されます。
- <I>と</I>で囲まれた部分は*イタリック*で表示されます。
- <U>と</U>で囲まれた部分は下線付きで表示されます。

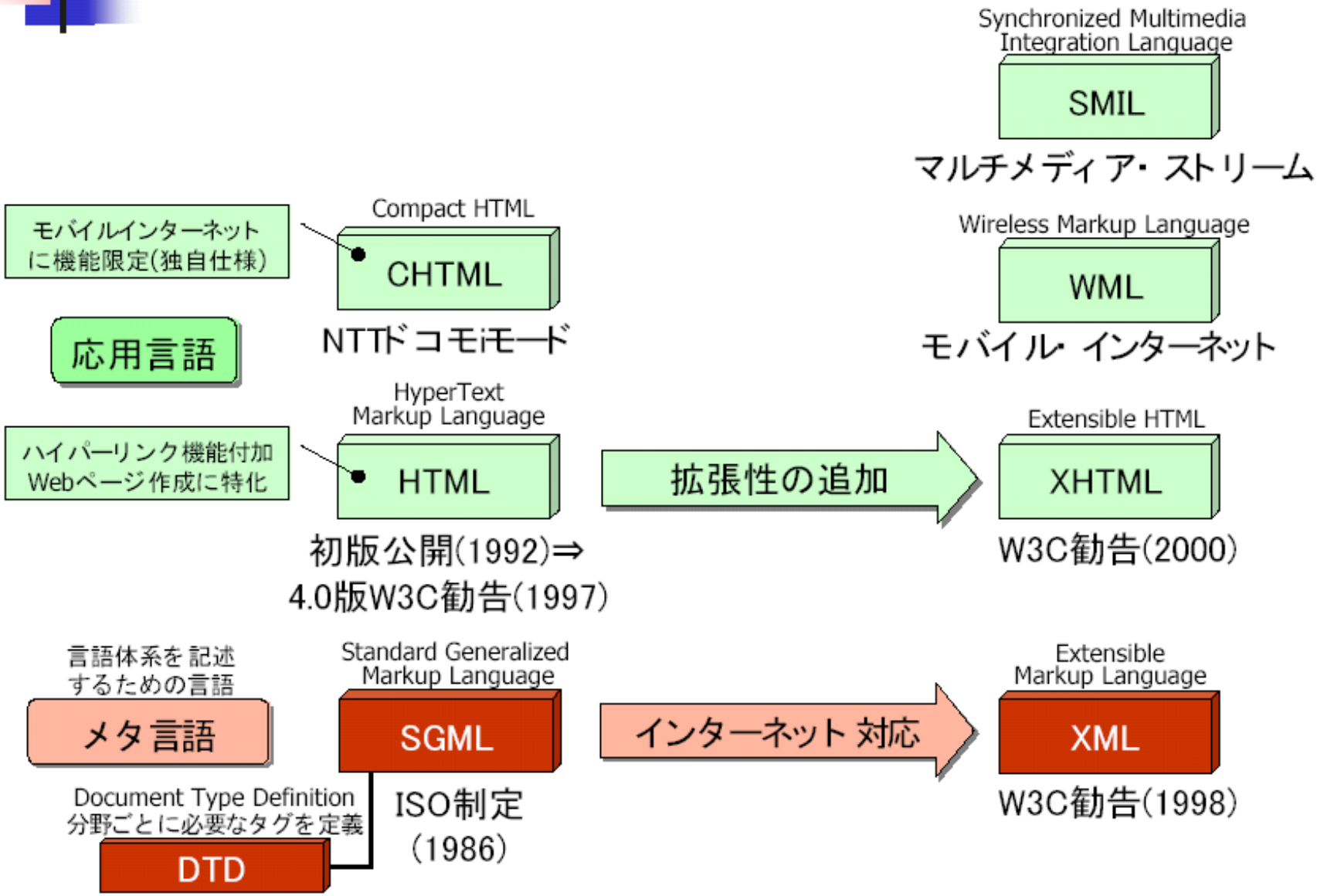
タグ名	意味	属性など
TABLE	Table	borderなど
TR	Table Row	rowspan(行をまたがる)など
TD	Table Data	align, valign, rowspan, colspan, nowrapなど
TH	Table Header	ボールドでセンタリング

ページが表示されました マイ コンピュータ

XML (eXtensible Markup Language)

- 見かけではなく項目の意味づけを記述
  - 利用者が新たなタグを定義
  - 企業間でのデータ交換など多目的に利用可
- ブラウザに表示するためには見かけを XSL (eXtensible Stylesheet Language) によるスタイルシートで記述

マークアップ言語の発展



XMLによるコンテンツ記述例

```
<?xml version="1.0" ?>
<?xml-stylesheet type="text/xsl" href="book_list.xsl"?>
<book_list>
<book> <title> Computer Networks, 4/e </title> <author> Andrew S.
    Tanenbaum </author> <year> 2003 </year> </book>
<book> <title> Modern Operating Systems, 2/e </title> <author> Andrew
    S. Tanenbaum </author> <year> 2001 </year> </book>
<book> <title> Structured Computer Organization, 4/e </title> <author>
    Andrew S. Tanenbaum </author> <year> 1999 </year> </book>
</book_list>
```

XSLによるスタイルシート

```
<?xml version='1.0'?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
    version="1.0">
<xsl:template match="/">
<html>
<body>
<table border="2">
  <tr>
     <th> Title</th>
     <th> Author</th>
     <th> Year </th>
  </tr>
```

XSLによるスタイルシート(続き)

```
<xsl:for-each select="book_list/book">
  <tr>
      <td> <xsl:value-of select="title"/> </td>
      <td> <xsl:value-of select="author"/> </td>
      <td> <xsl:value-of select="year"/> </td>
  </tr>
  </xsl:for-each>
</table>
</body>
</html>
</xsl:template>
</xsl:stylesheet>
```

C:¥book_list.xml - Microsoft Internet Explorer

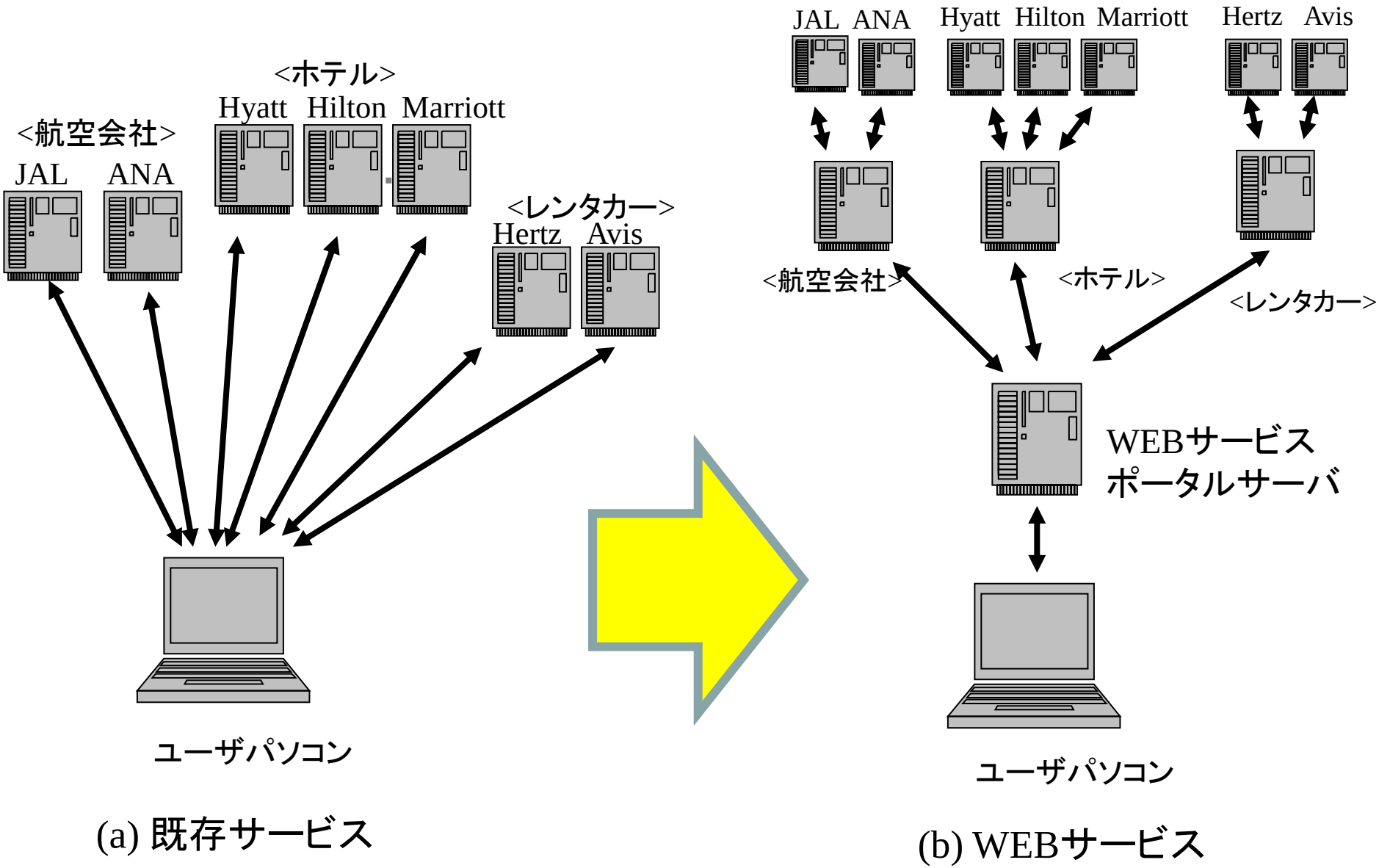
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り リンク

アドレス(D) C:¥book_list.xml 移動

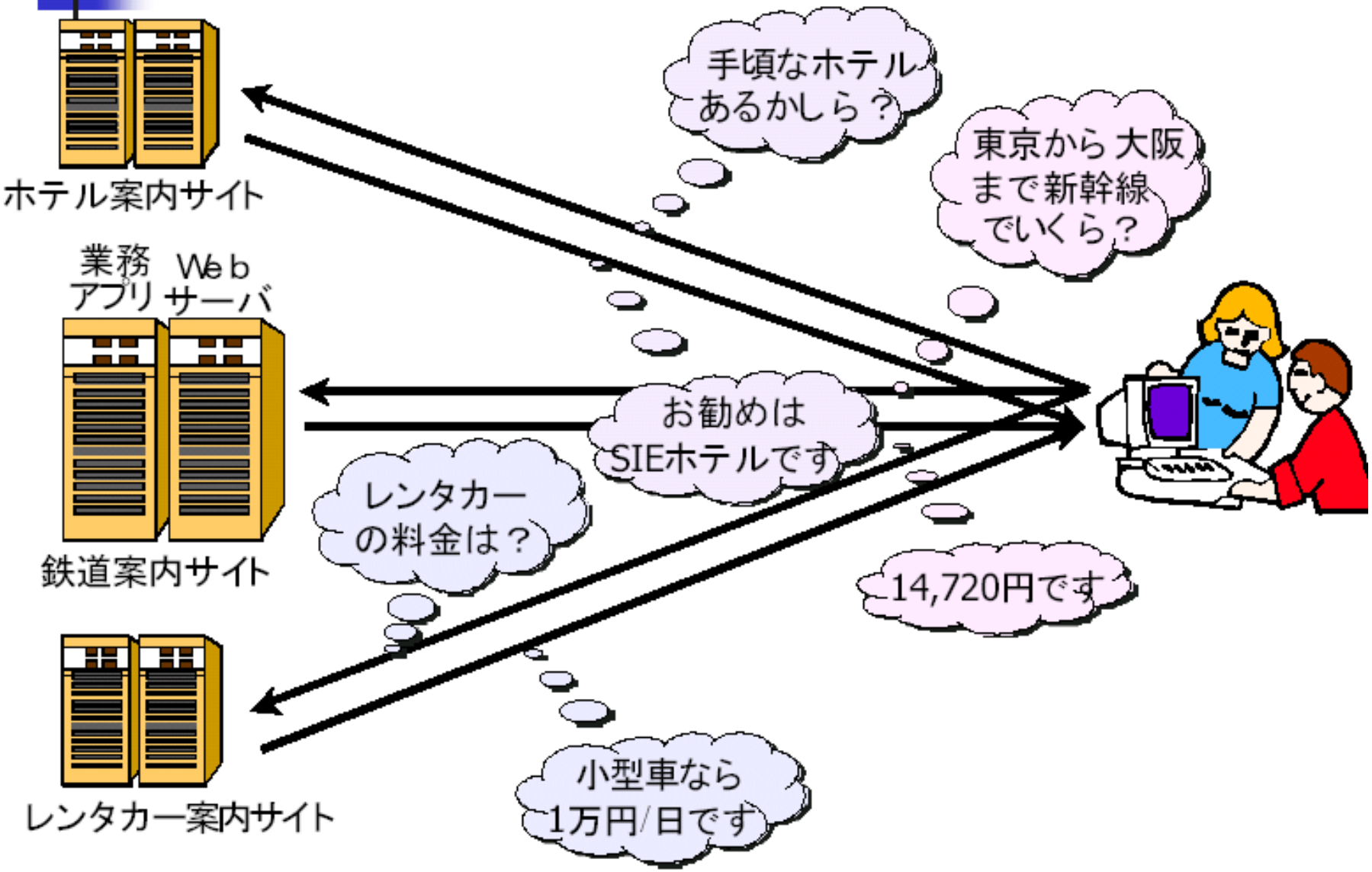
Title	Author	Year
Computer Networks, 4/e	Andrew S. Tanenbaum	2003
Modern Operating Systems, 2/e	Andrew S. Tanenbaum	2001
Structured Computer Organization, 4/e	Andrew S. Tanenbaum	1999

ページが表示されました マイ コンピュータ

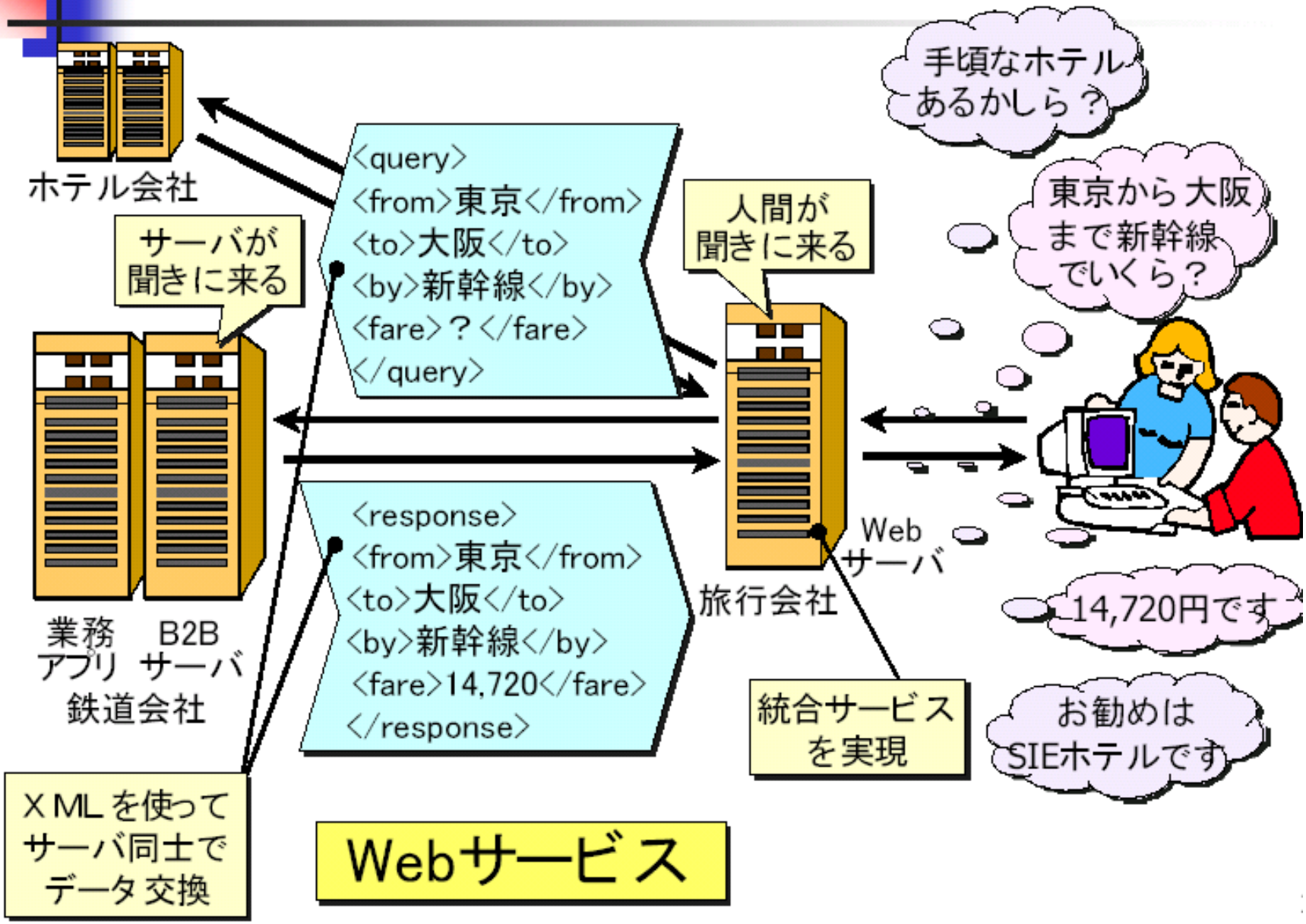


WEBサービスの基本概念(2000年初頭)

従来のWebサイト へのアクセス



Webサービスの概念


ホテル会社
サーバが聞きに来る
業務アプリ
B2Bサーバ
鉄道会社
<query>
<from>東京</from>
<to>大阪</to>
<by>新幹線</by>
<fare>？</fare>
</query>
<response>
<from>東京</from>
<to>大阪</to>
<by>新幹線</by>
<fare>14,720</fare>
</response>
人間が聞きに来る
Webサーバ
旅行会社
手頃なホテルあるかしら？
東京から大阪まで新幹線でいくら？
14,720円です
お勧めはSIEホテルです
統合サービスを実現
XMLを使ってサーバ同士でデータ交換
Webサービス

大規模（WEB）サーバ

最近の 傾向･方法

1. マルチプロセッサ型の計算機アーキテクチャの導入

   分散処理(機能分散、地理的分散)

2. キャッシュ技術の導入

   CDN、P2Pなど

3. DMA(Direct Memory Access)技術の導入

   P2P技術

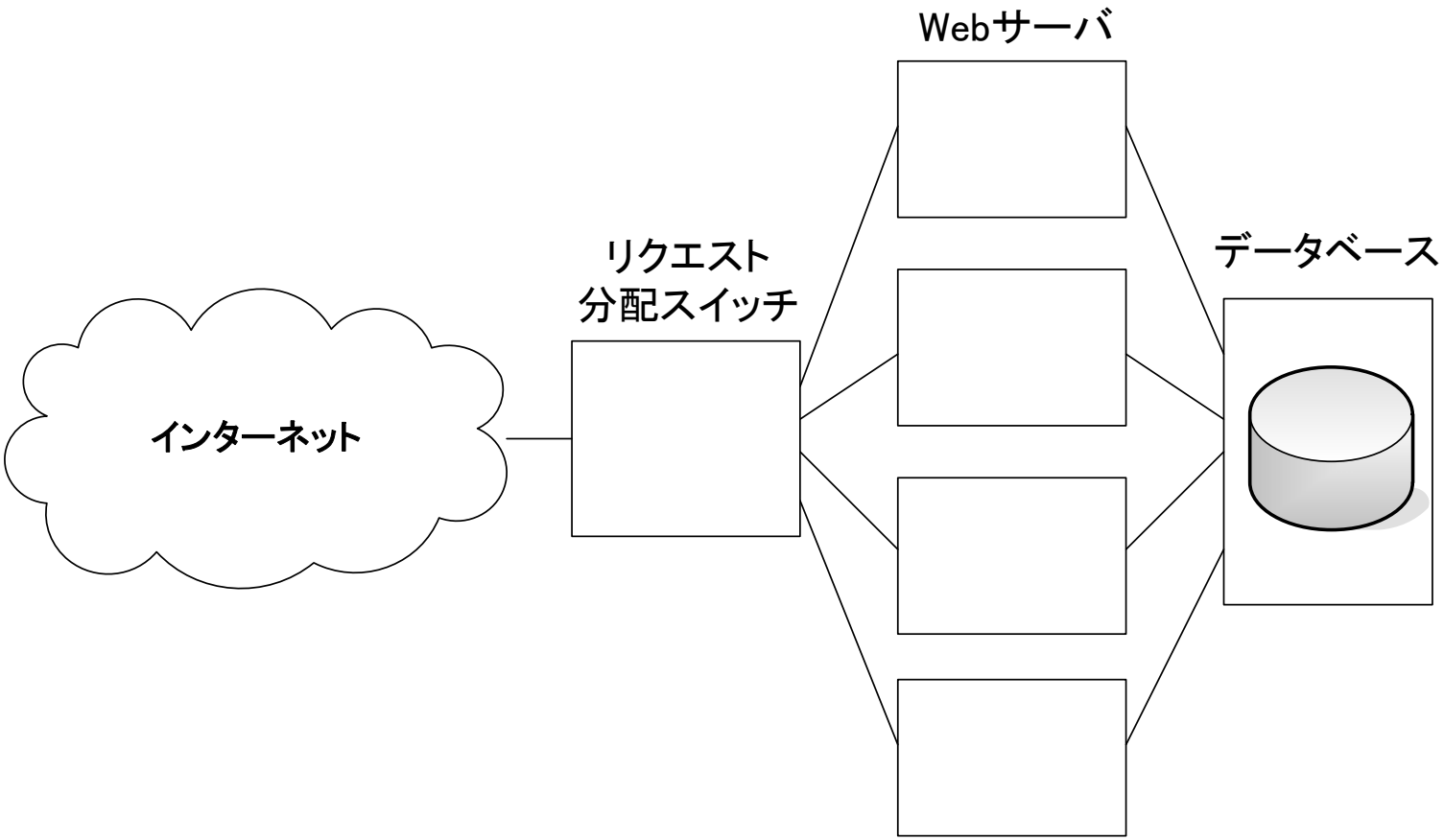
4. 仮想化技術の導入

   Virtualization、Overlayネットワーク

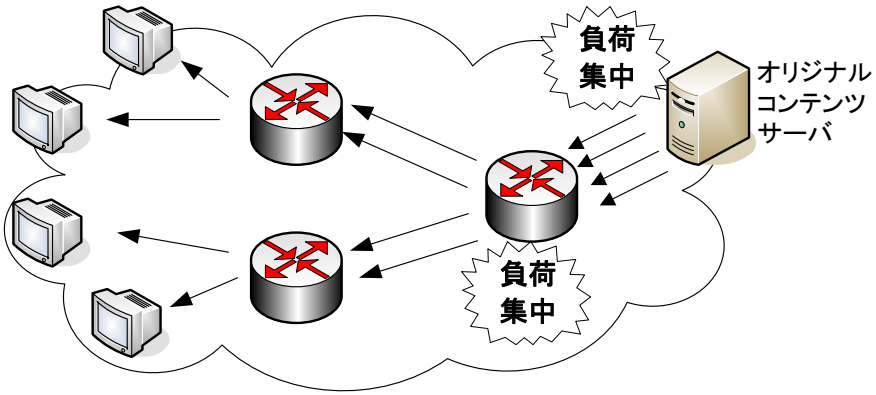
処理負荷の分散手法

1. 水平分散
   - 複数のサーバに並列的に処理を実行させる。
   - 特定の機能ごとに、専門化したサーバを用意し、適切に Dispatch する。
2. 垂直分散
   - Proxy機能や Cache機能を提供し、クライアントに近い仮想ノード/分身ノードが、データ処理を分担する。
   - データの変更時の Consistency の維持が課題となる。

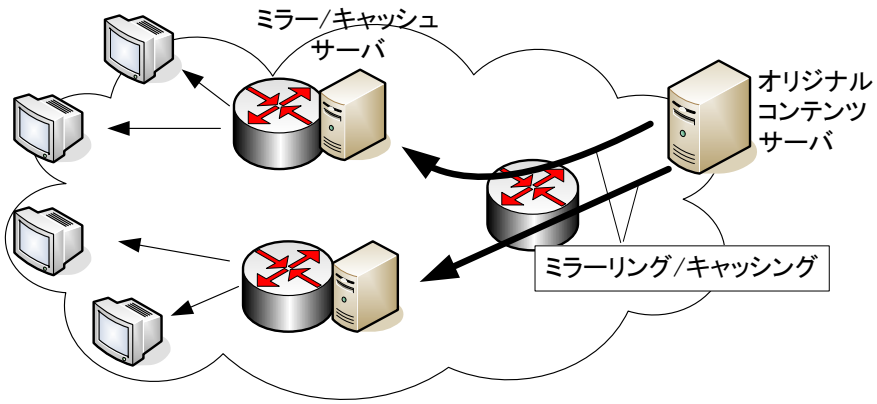
負荷分散



Content Delivery Network



従来のWebサーバシステム

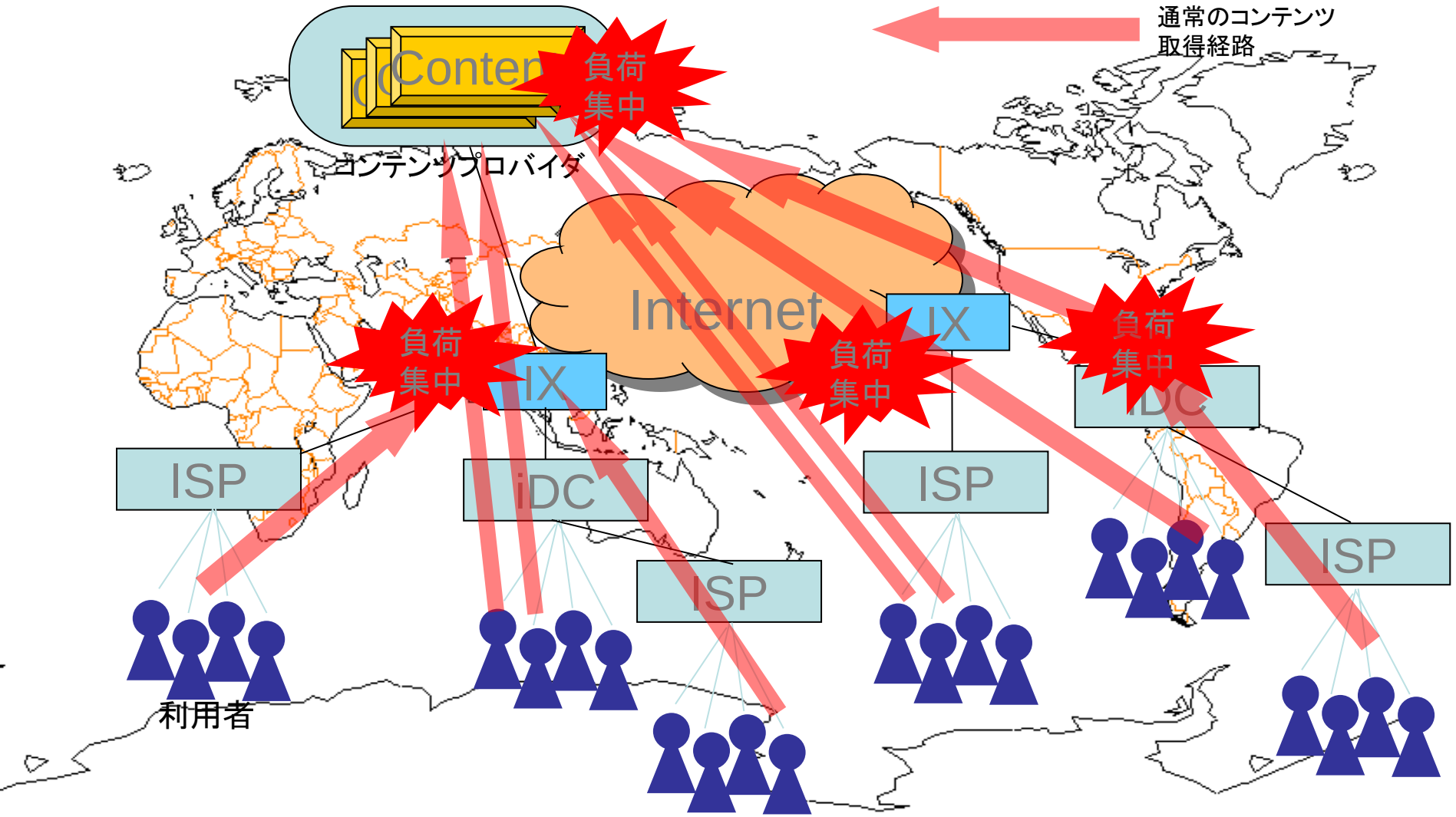


コンテンツ配信システム

CDS/CDNとは？

- CDS
  - Content(s) Delivery Service/System
  - Content(s) Distribution Service/System
  - Web/Streamingなどのrich contentsに対してキャッシュやミラーを積極的に用いた負荷分散システム(サービス)
    - ・リバースキャッシュ技術
    - ・キャッシュ管理技術
    - ・リクエストナビゲーション技術
- CDN
  - Content(s) Distribution Network
    - ・CDSの基盤ネットワーク
  - CDS間のコンテンツピアリングを行う単位
    - ・例：　A社のαというコンテンツをB社のCDSへ提供する

Webサービスの構造上の問題（負荷の集中）



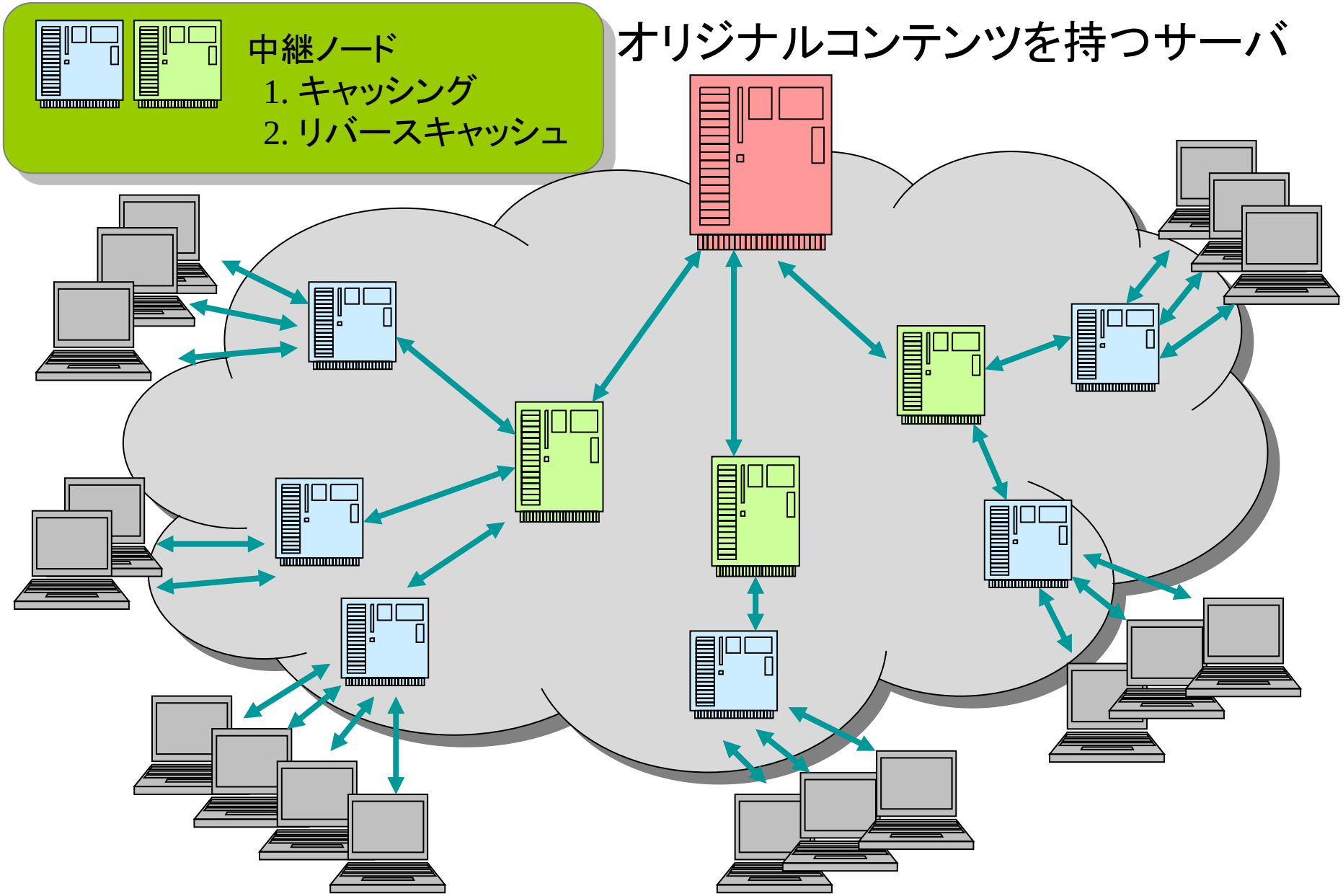
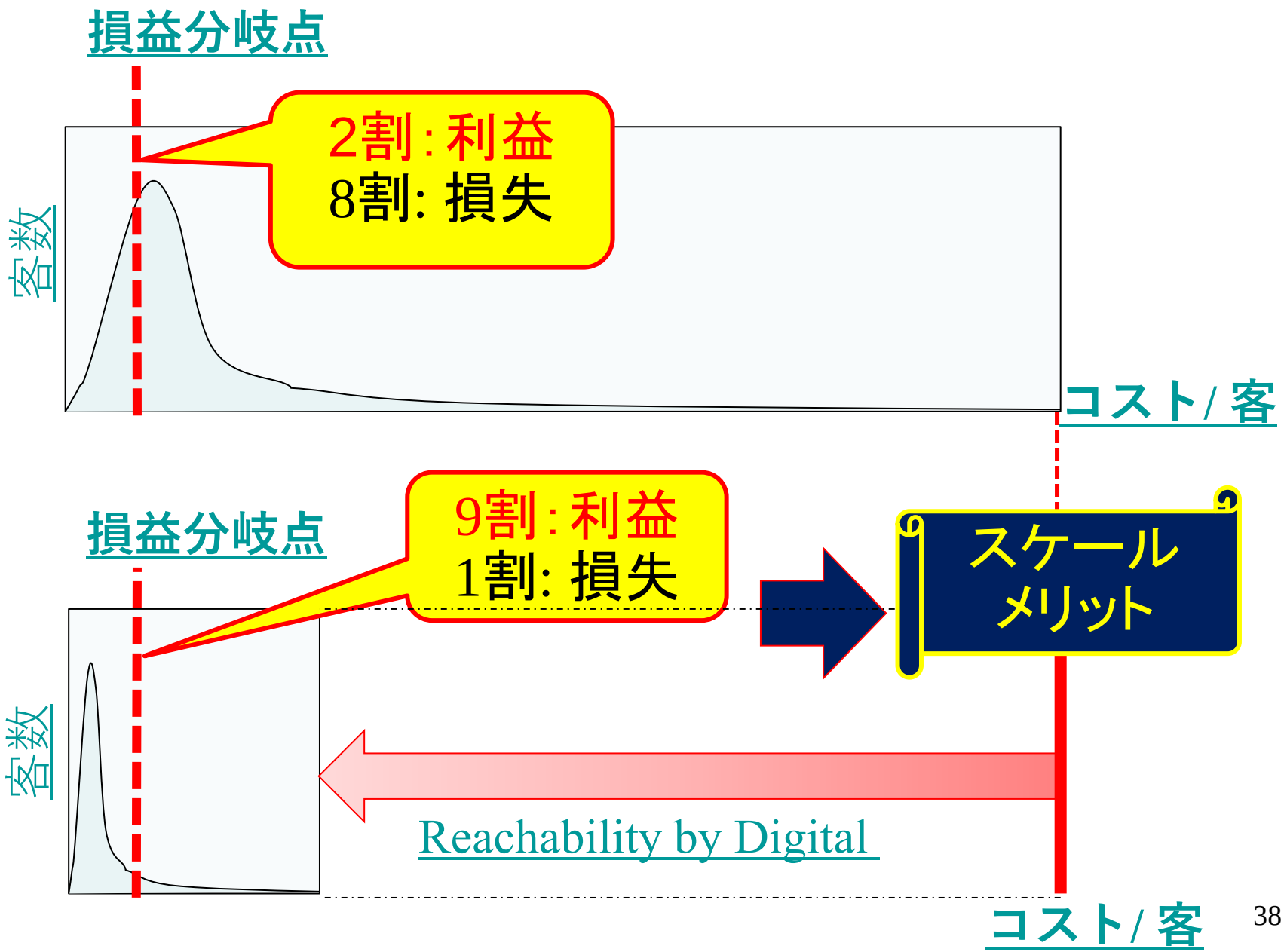


図8-12 CDNの全体構造概念図

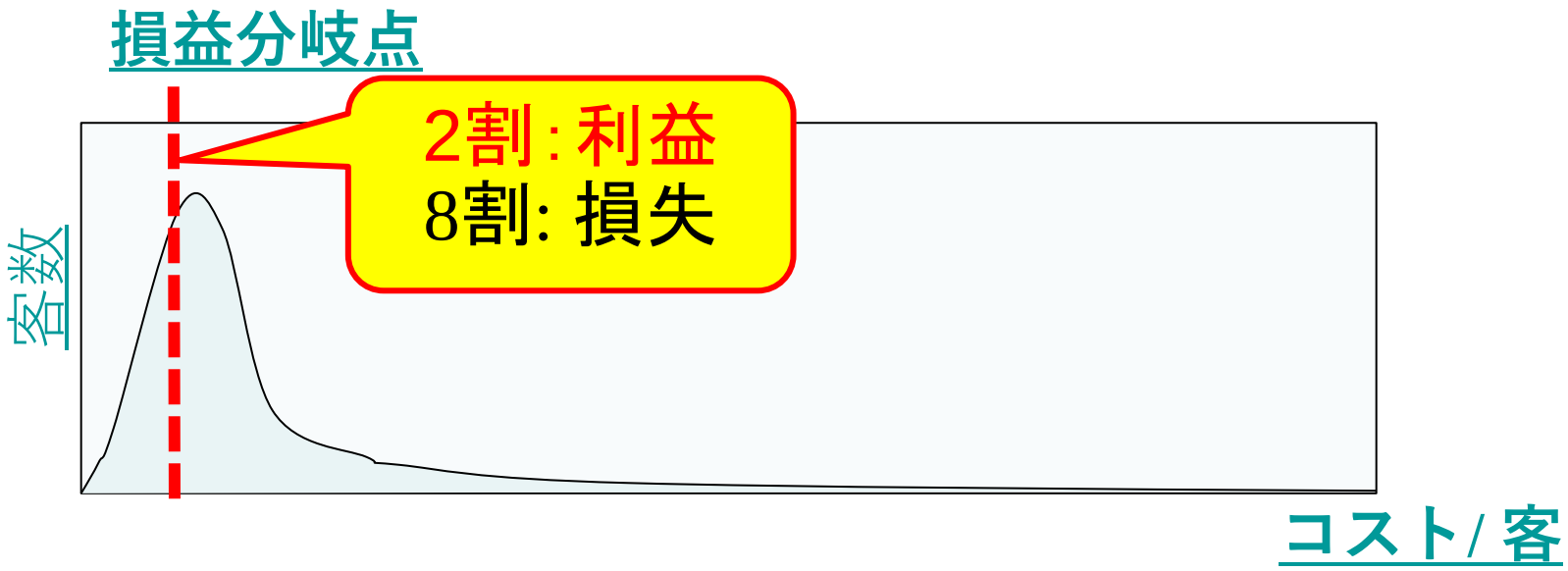
ロングテール

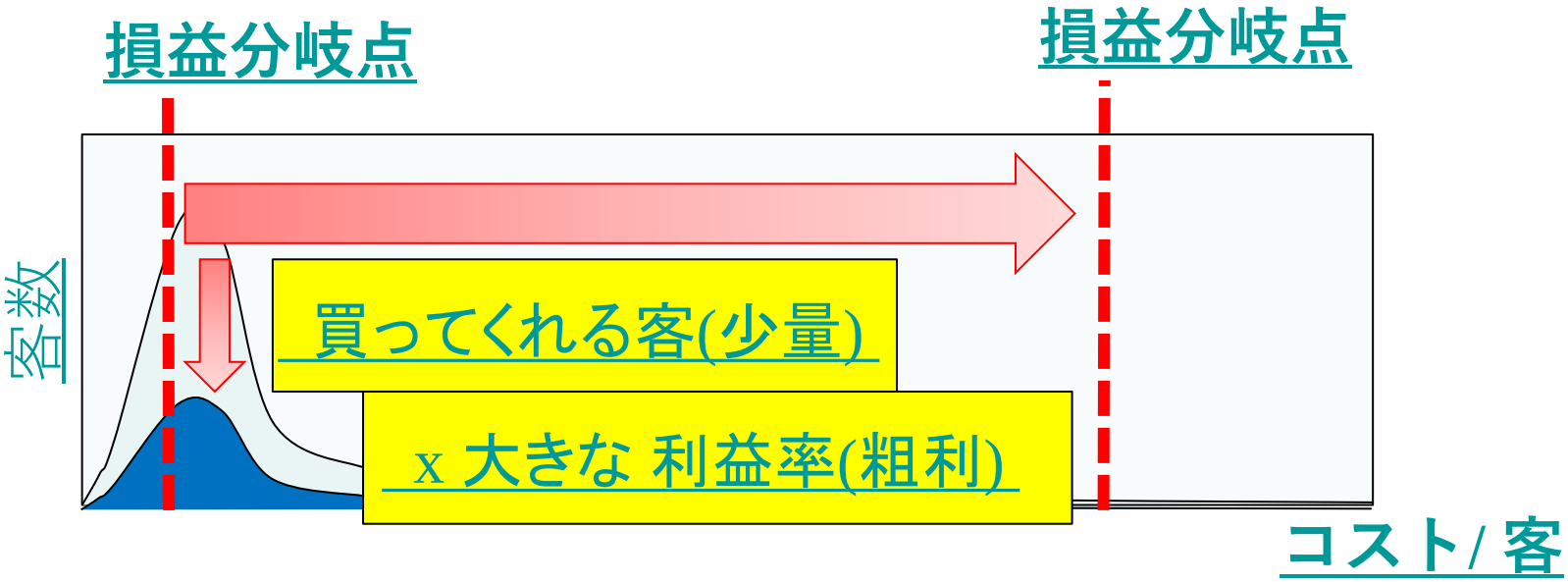
- 20:80の法則 (パレートの法則)
  - 20%の商品から80%の売り上げ
  - 20%の社員から80%の成果 etc.
- 実店舗では20%に注力
- アマゾン・ドット・コムでは230万品目中13万位以下の売り上げが57%

パレートの法則 vs Long-Tail Business



パレートの法則 vs Long-Tail Business





Peer-to-Peer System

What is P2P

- Apply the technical frameworks, that is applied to inside a computer so as to improve the processing performance, are applied to the distributed computer network environment.
- In particular, the followings are key components for P2P system;
  1. Cache (and mirroring)
  2. Redirection of data flow, e.g., DMA
  3. Abstraction of objects, to handle by unique and single method.
- P2P does not include computation

Something new in the Internet

1. Multi-processor computer architecture for distributed processing, aka functional distribution, geographical distribution
2. Cache and Proxy

   CDN(Contents Delivery Network)
3. Virtualization

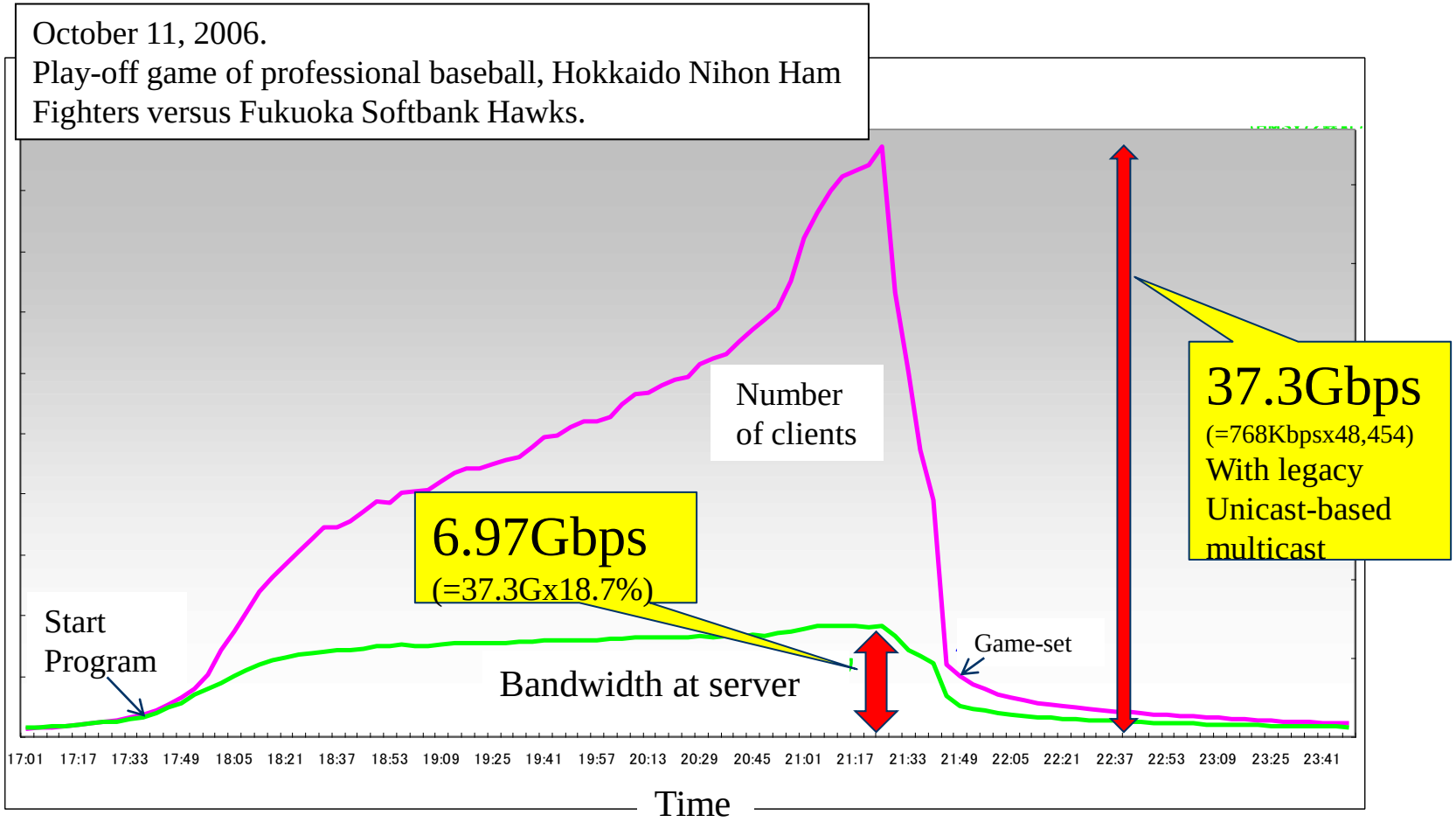
   Overlay networking

Background of P2P; Unawared technical evolution ?

- Two of drastic cost down in ICT
  - Copying
  - Recording and saving

- However, logistics of information may not be aware of this evolution……
  - Have BitTorrent or Joost been awared ?
  - BB-TV! by SoftBank YBB has operated, silently…

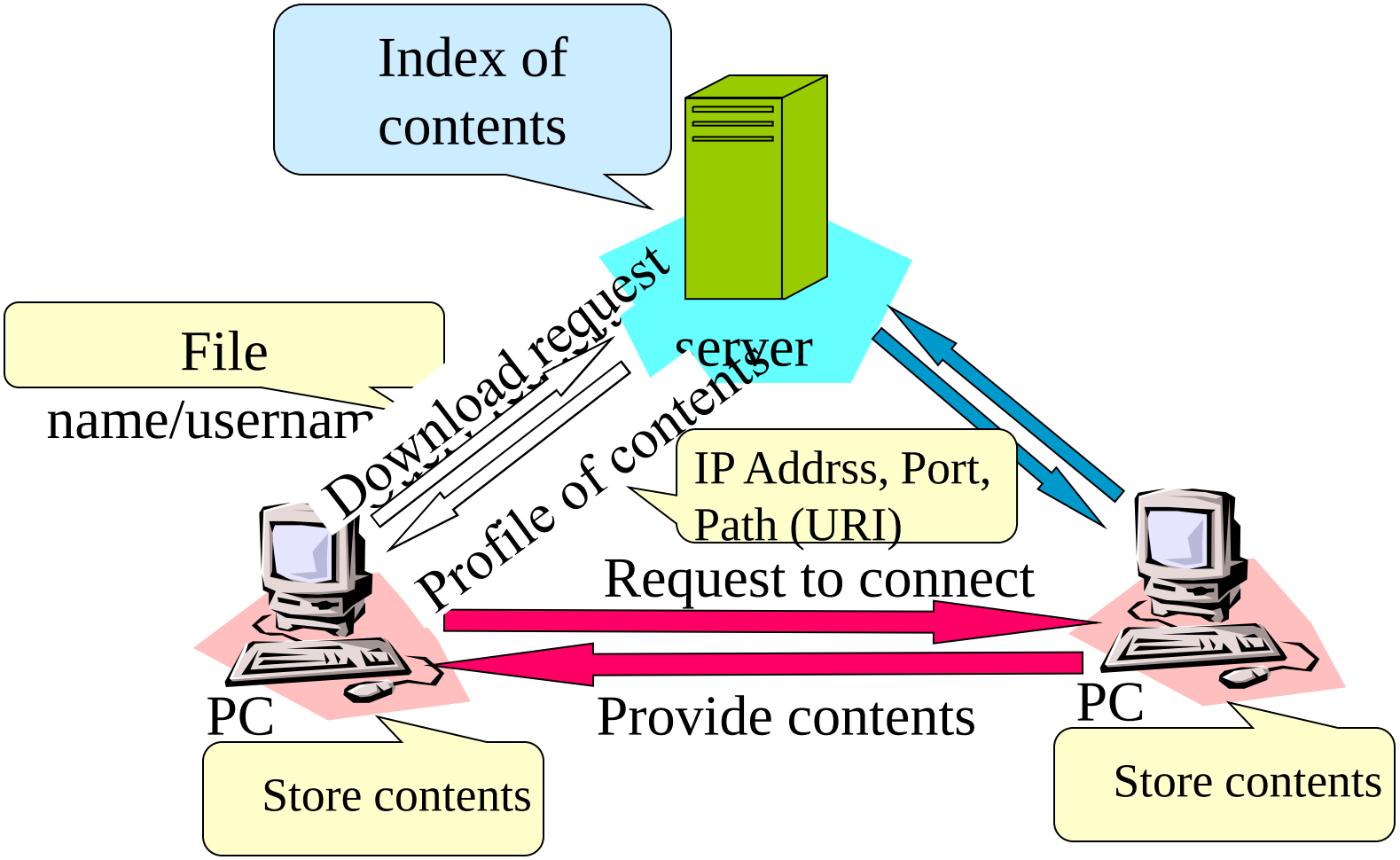
Peer-to-Peer overlay multicasting service by professional ISP, BB-TV! by SoftBank over ADSL network



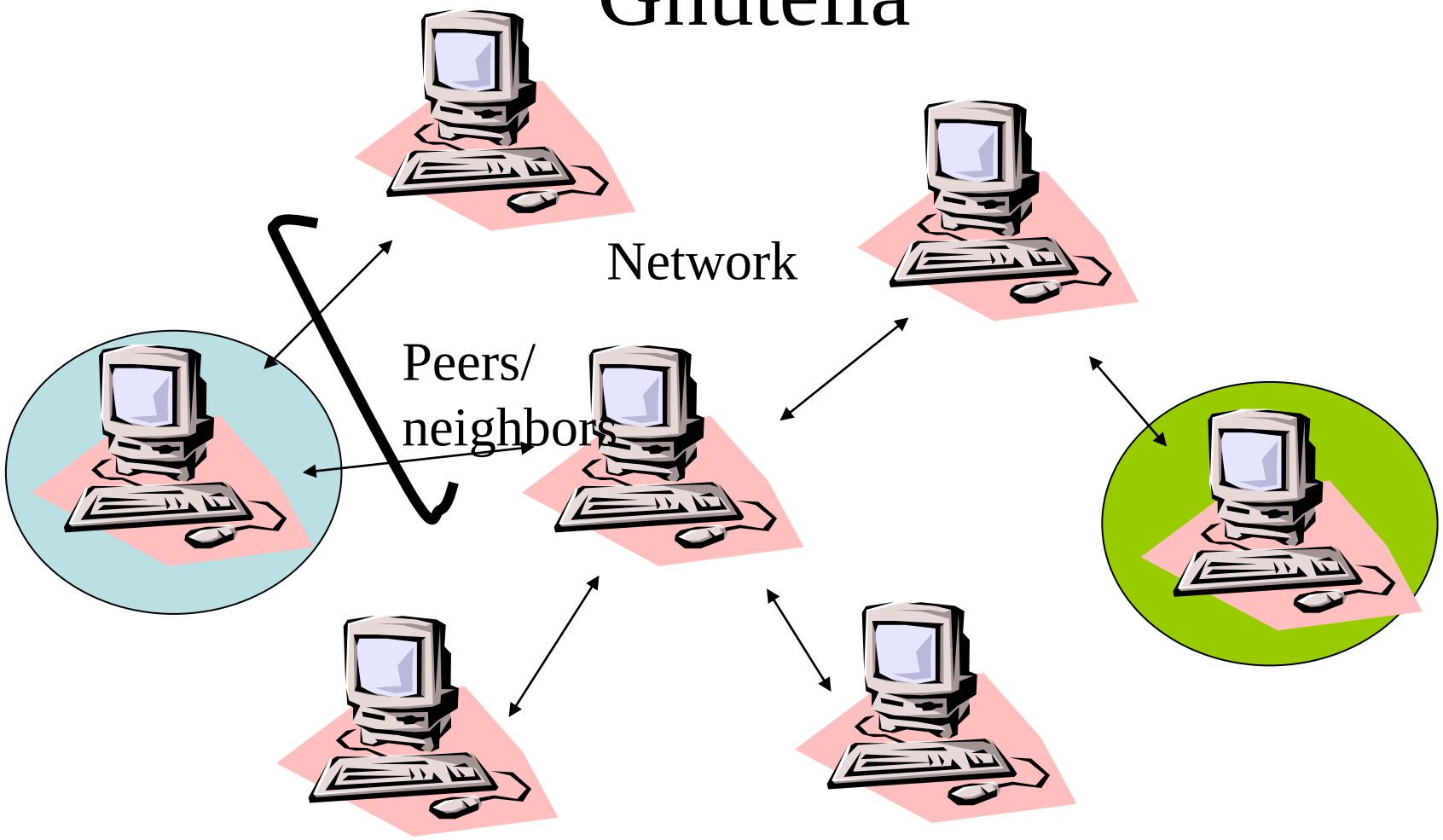
Peer-to-Peer Overlay Networking

- First generation
  - Napstar, WinMX
    - ➔directory server + Pee-to-Peer connection
      (*) similar to SIP and NGN
- Second generation
  - Gnutella
    - ➔ Server-less pure peer-to-peer
      (*) have a scaling problem…..
- Third generation
  - Freenet, Winny
    - ➔ introduction of network cache for scalability

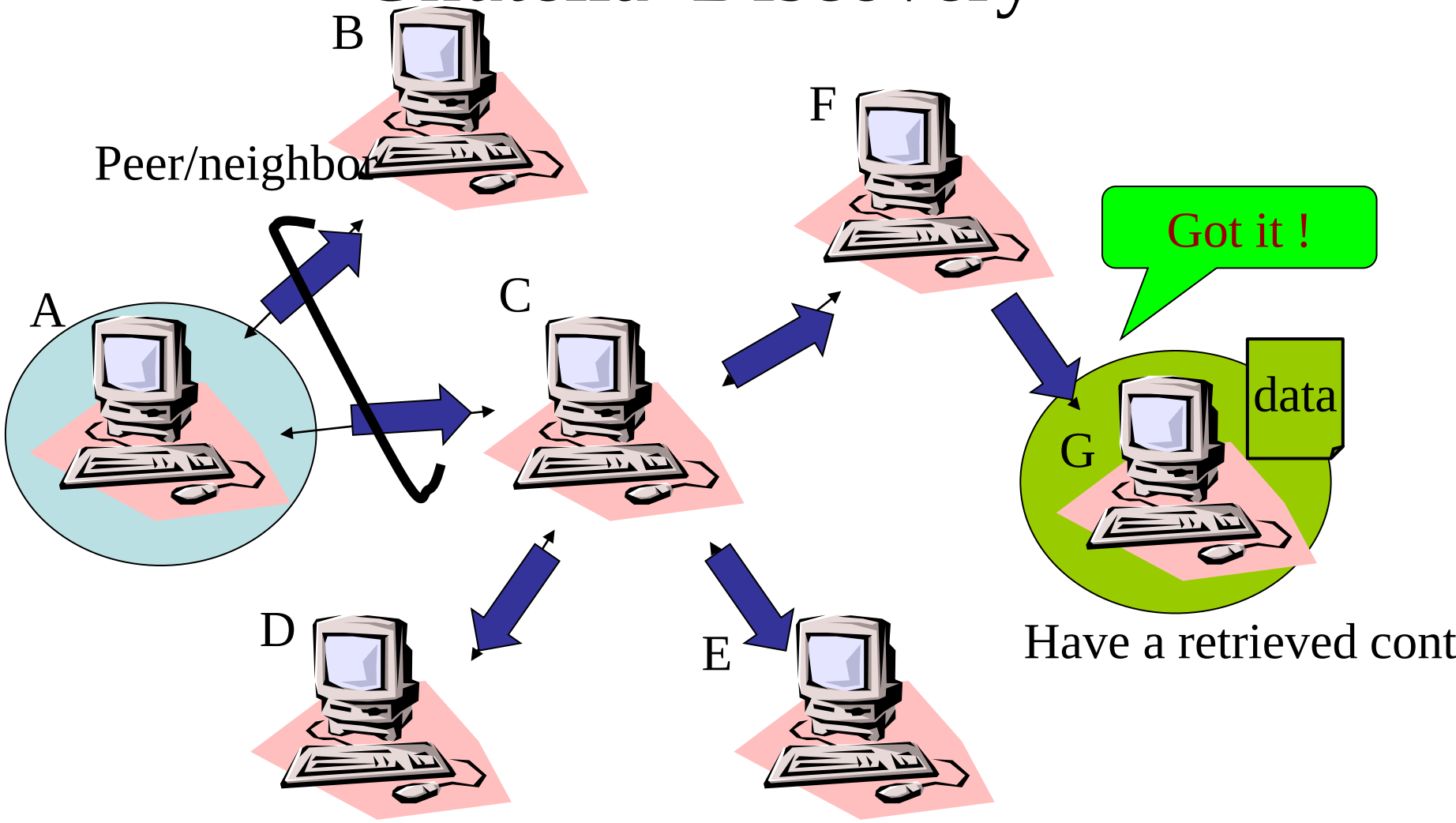
Napster


Index of contents
server
File name/username
Download request
Profile of contents
IP Addrss, Port, Path (URI)
Request to connect
Provide contents
PC
PC
Store contents
Store contents

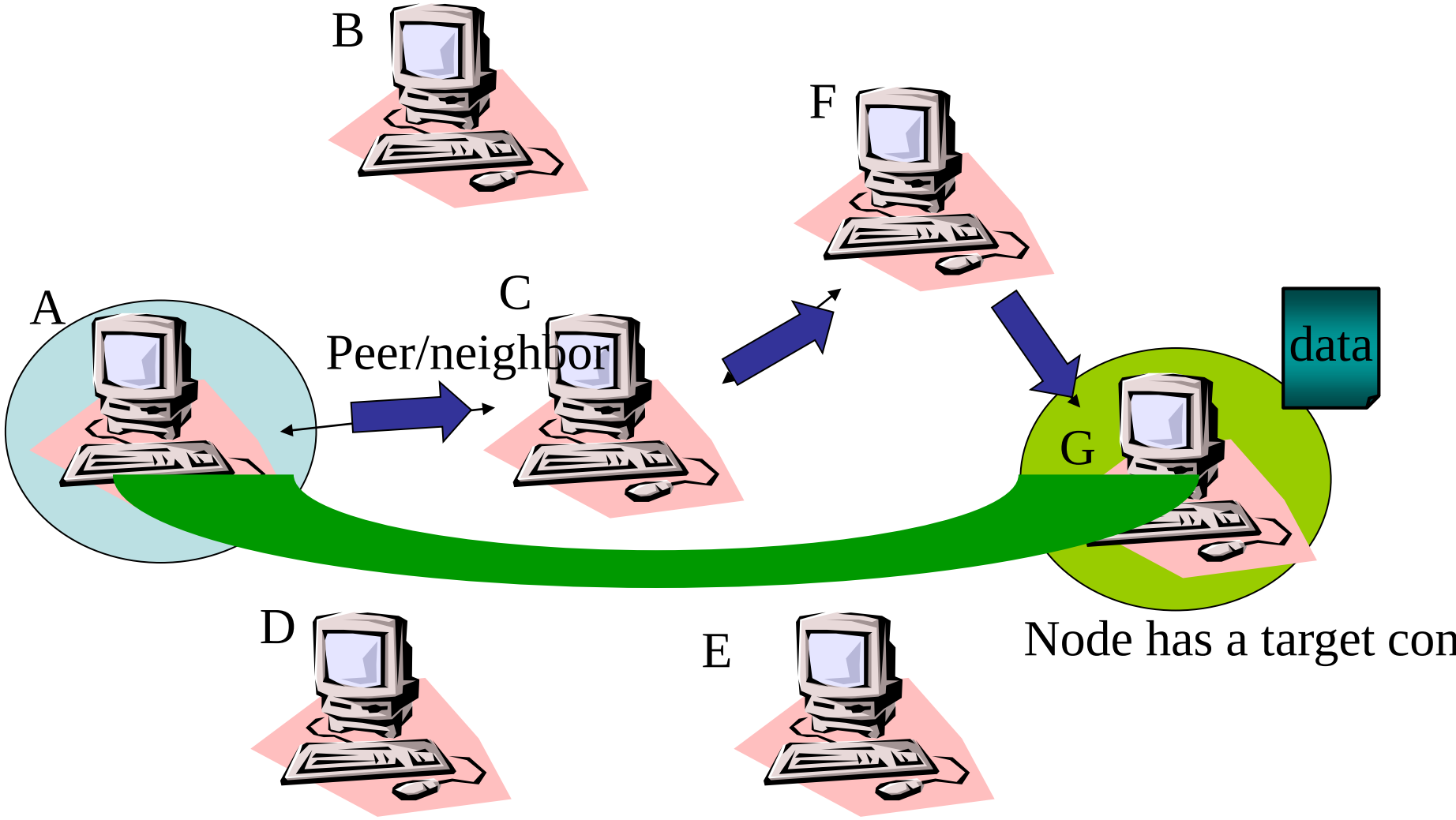
Gnutella



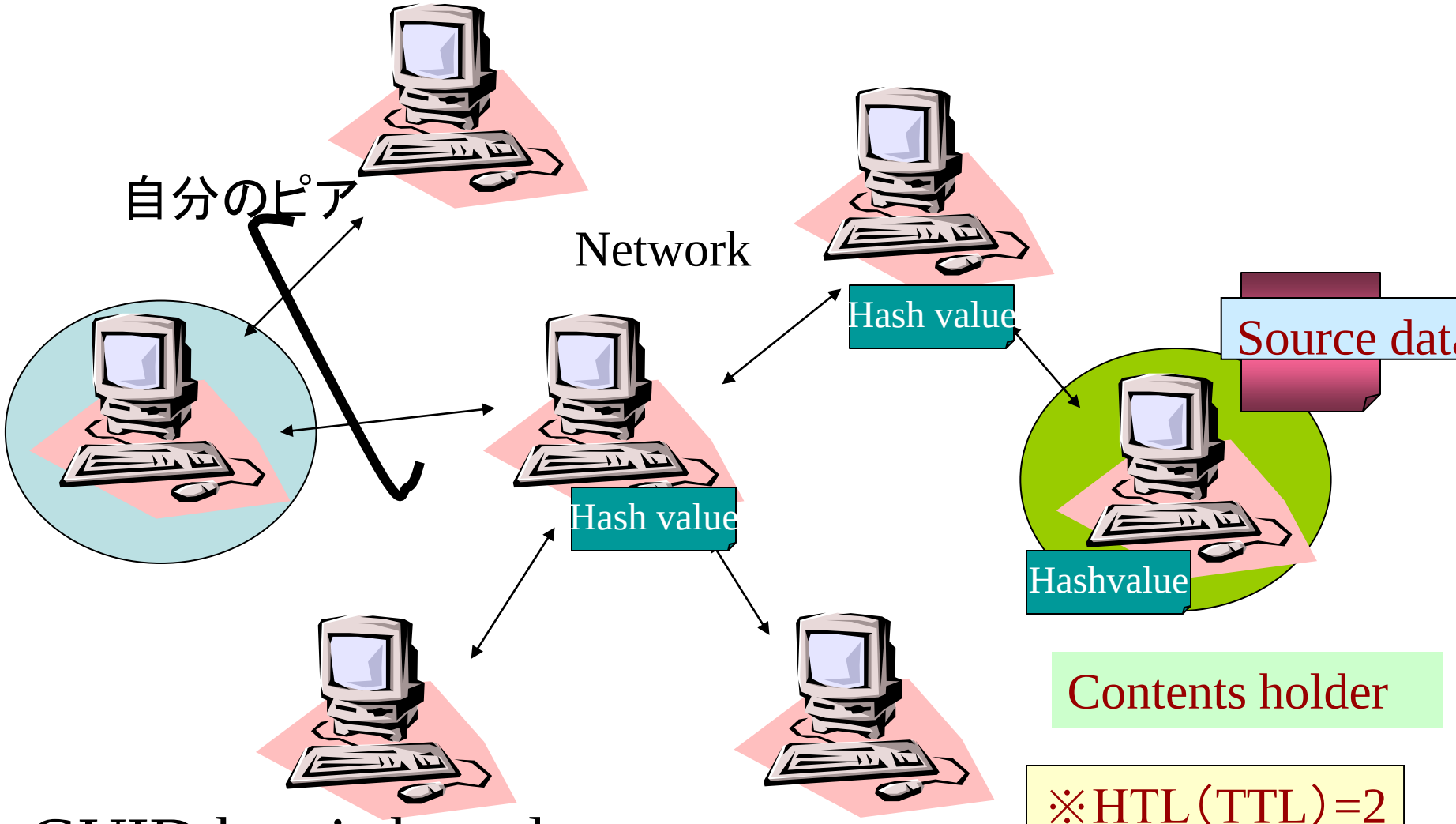
Gnutella : Discovery



Gnutella: Transmission

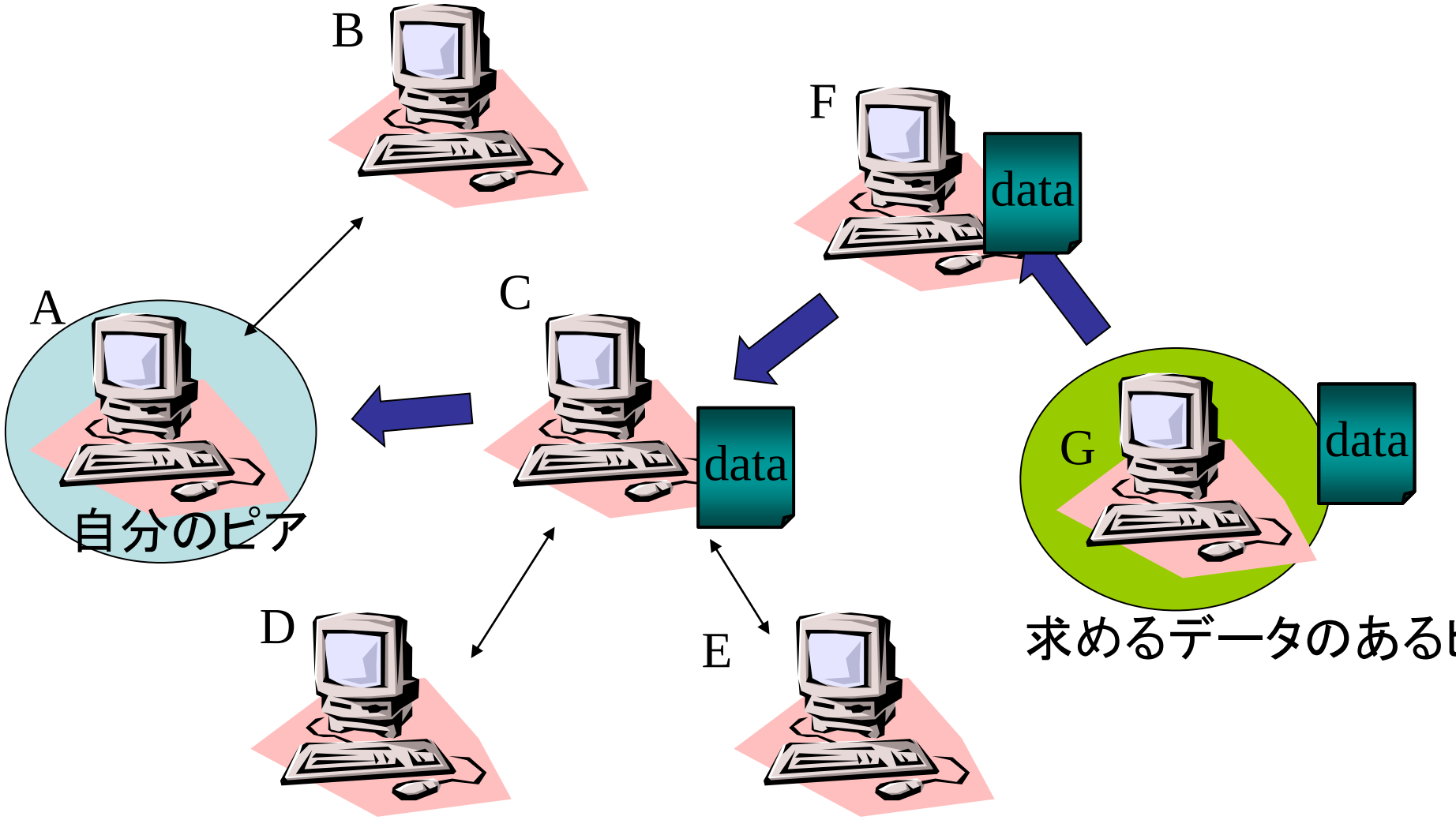


Freenet：put a file



- GUID key is broadcast

Freenet; file transfer



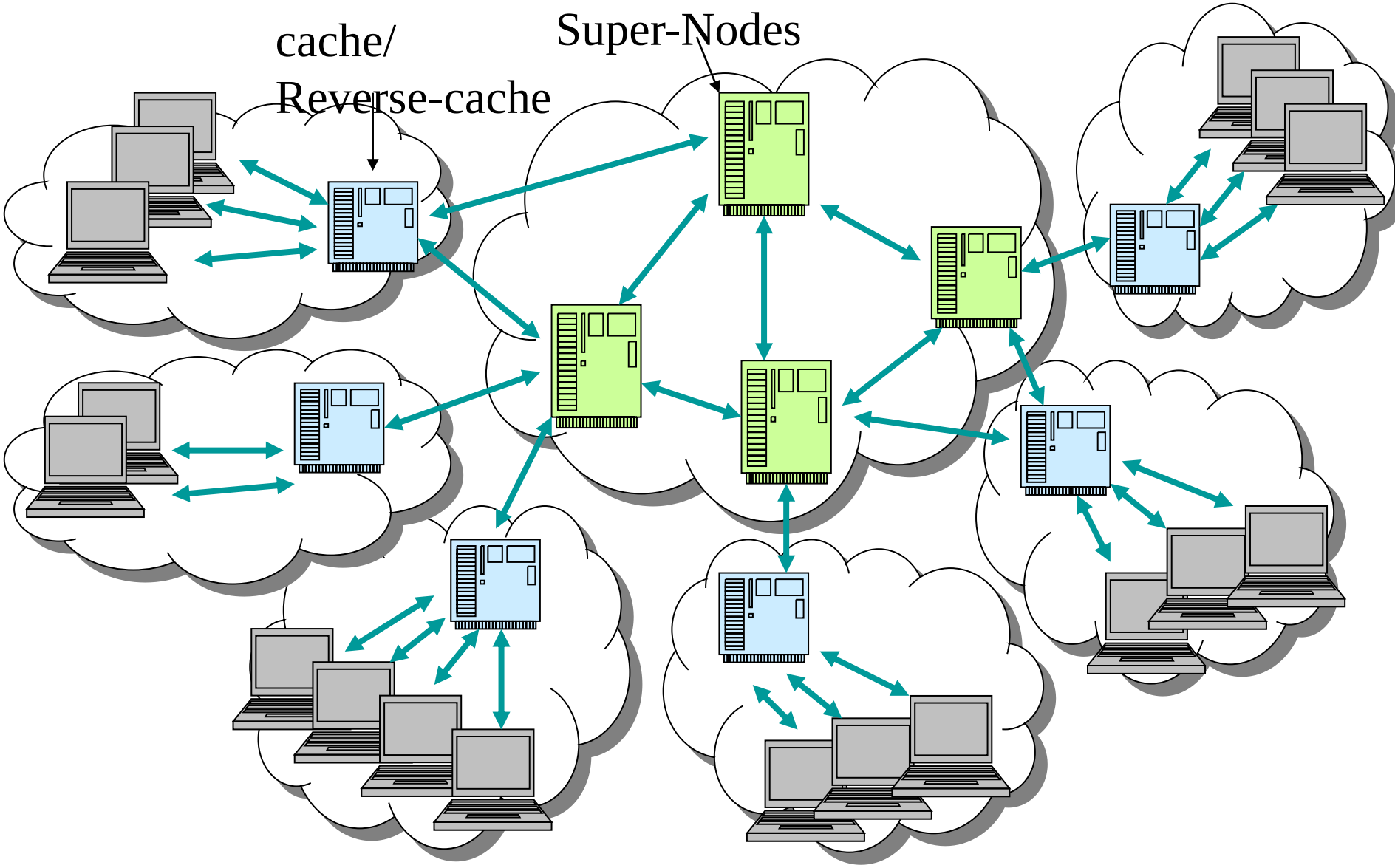


図9-5 Winny/SKYPEにおける階層的トポロジー構造の概念図

Distributed file storage in P2P system

- Napster
  - Separate Index and Storage
    - DMA operation by the introduction of out-band signaling
    - Reducing the copy overhead using the pointer
- Freenet/Winny
  - Separate file name and it’s location information
    - Operation like virtual memory system
  - Anonymity of file profile
  - Introduction of network cache
    - Paging, cache, reverse-cache, etc.,

What has Peer-to-Peer introduced ?

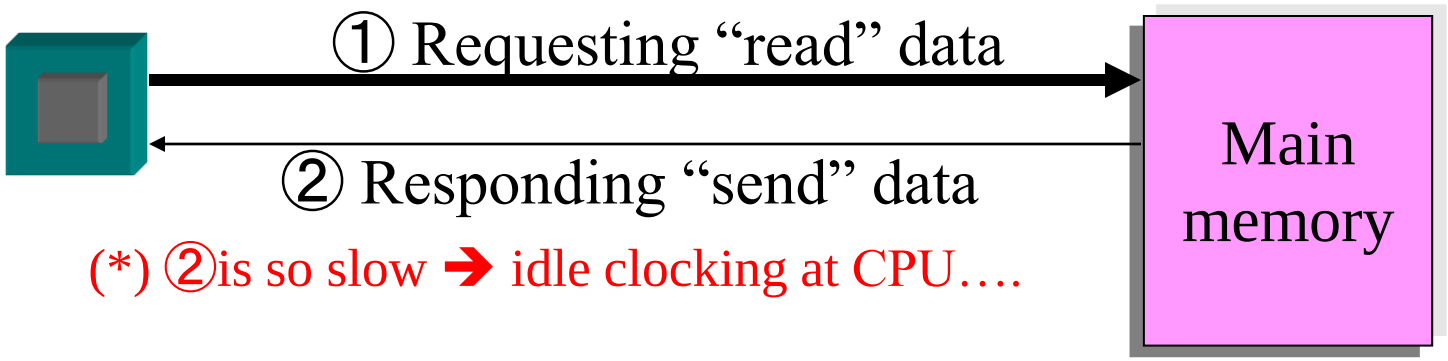
1. {networked} Cache and Proxy
2. {networked} DMA (Direct Memory Access)
3. {networked} Virtual memory system (by DHT)
   - Separate contents handler(identifier) and real storeing address
   - Access heterogeneous and multiple device with single {virtual} interface
4. Abstraction of contents by number (by DHT)

   concealing {file name, file descriptor, etc.}

Cache memory in computer

High speed processing

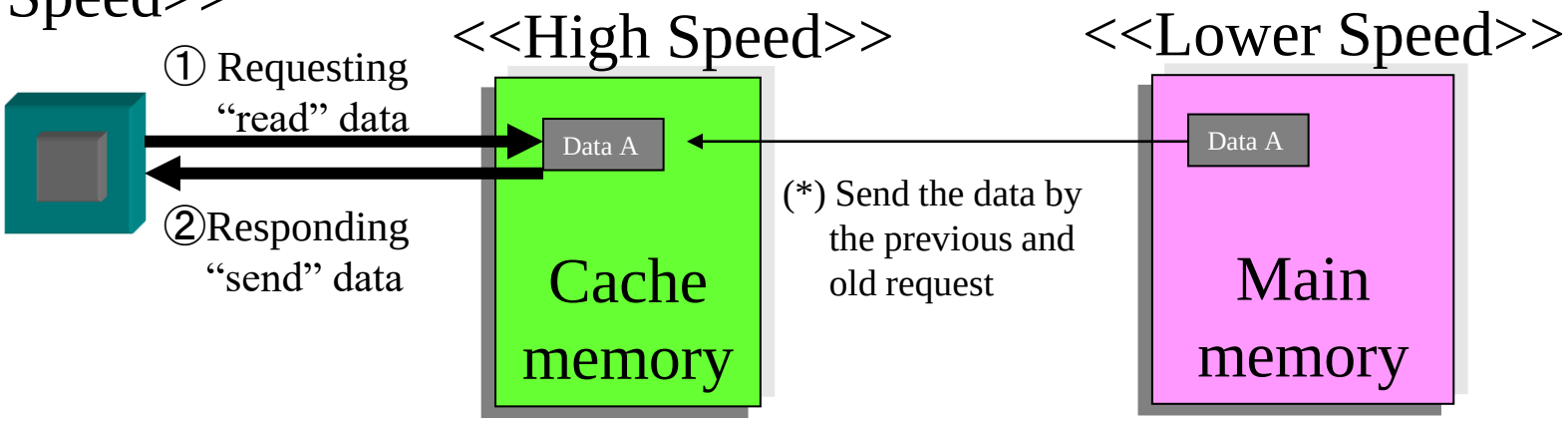
Lower speed processing



(*) ②is so slow ➔ idle clocking at CPU….

Introduction of cache memory system

<<High Speed>>

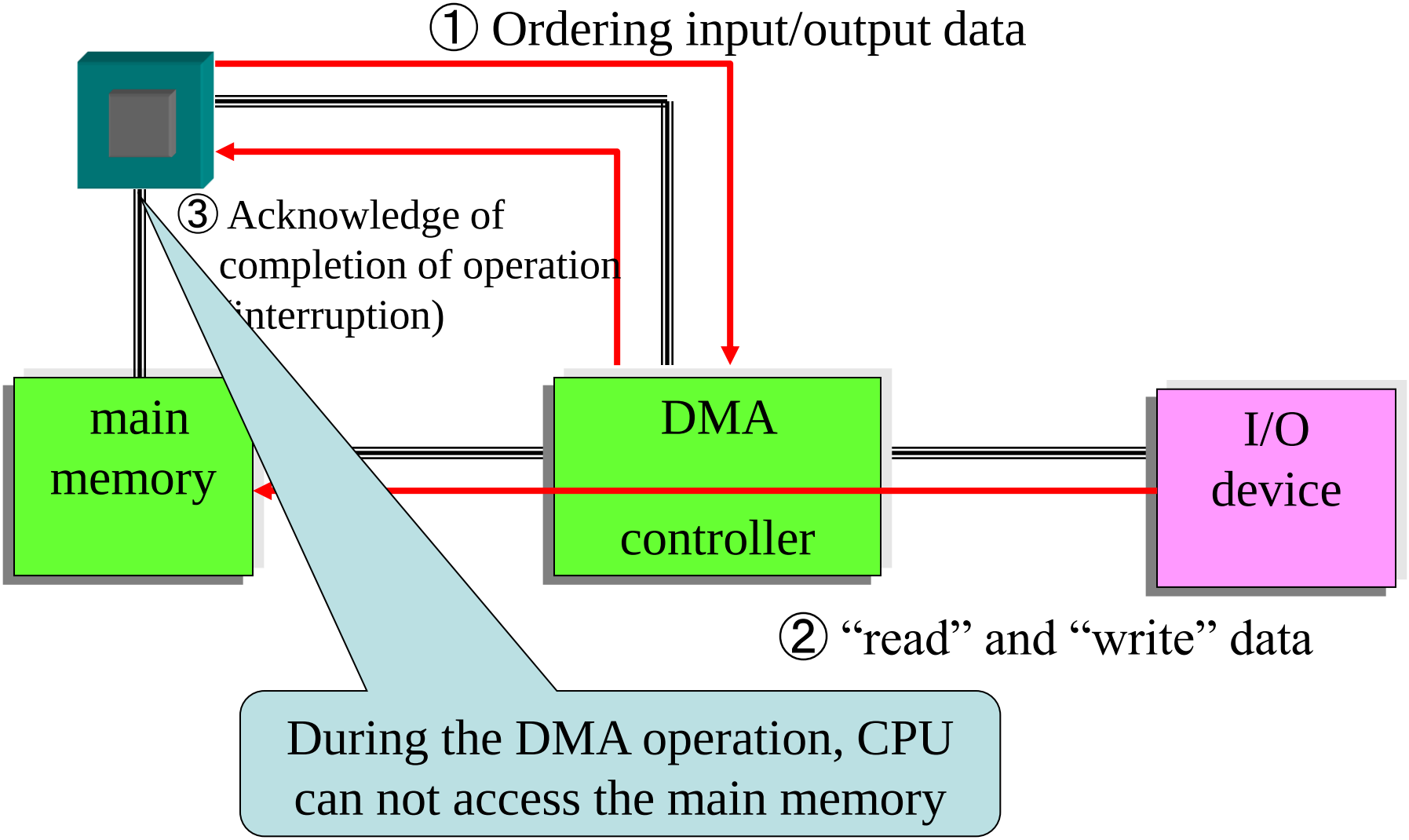


What has Peer-to-Peer introduced ?

1. {networked} Cache and Proxy
2. <u>{networked} DMA (Direct Memory Access)</u>
3. {networked} Virtual memory system (by DHT)
   - Separate contents handler(identifier) and real storeing address
   - Access heterogeneous and multiple device with single {virtual} interface
4. Abstraction of contents by number (by DHT)

   concealing {file name, file descriptor, etc.}

DMA; Direct Memory Access



Signaling system

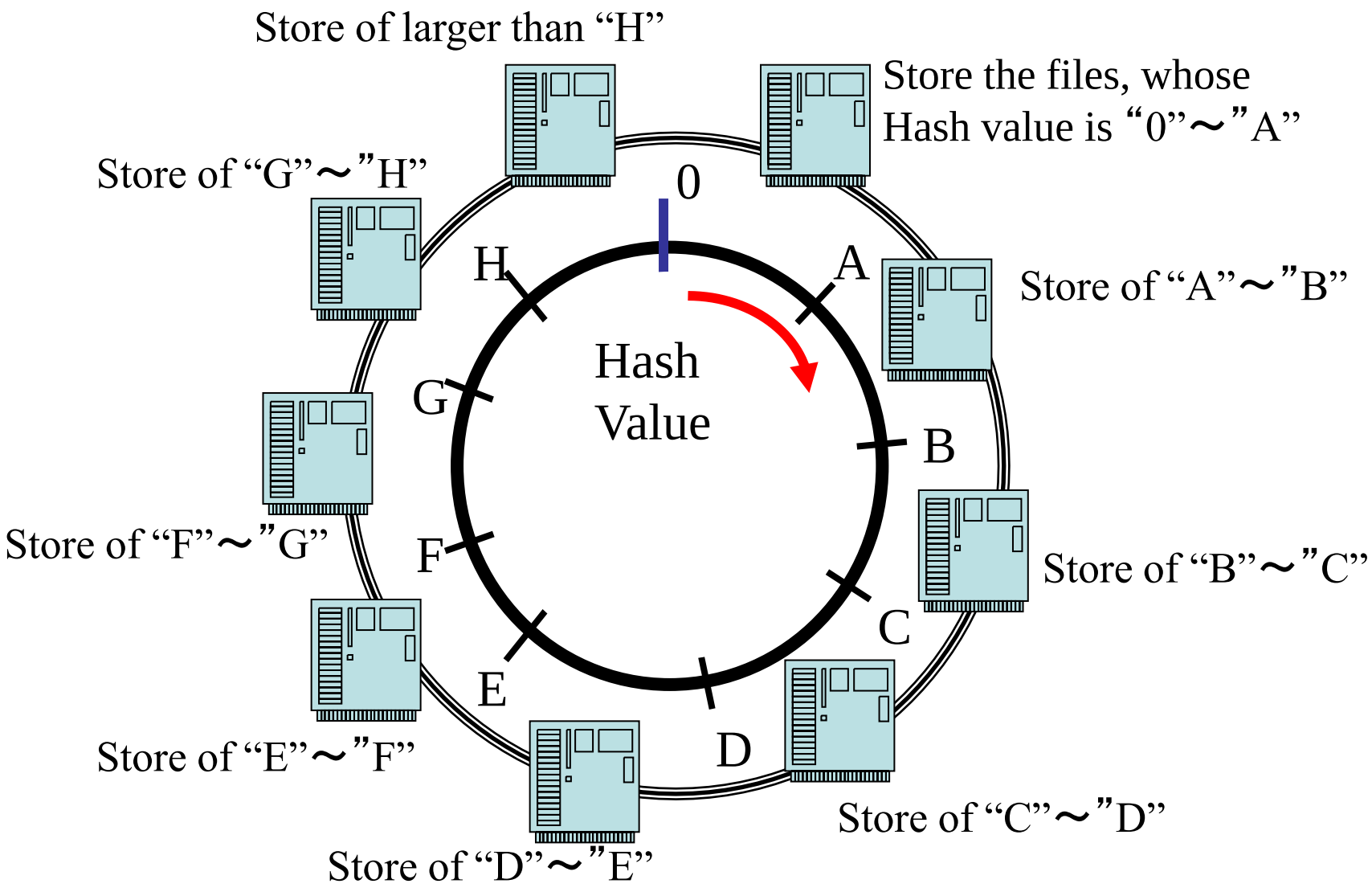
- Out-band signaling
  - Separate the Data-Plane and Control-Plane)
  - Example; Telephone, MP3 player, TV broadcasting

- In-band signaling
  - Share and use the same resource by Data-Plane and by Control-Plane
  - Example; internet

What has Peer-to-Peer introduced ?

1. {networked} Cache and Proxy
2. {networked} DMA (Direct Memory Access)
3. {networked} Virtual memory system (by DHT)
   - Separate contents handler(identifier) and real storeing address
   - Access heterogeneous and multiple device with single {virtual} interface
4. Abstraction of contents by number (by DHT)

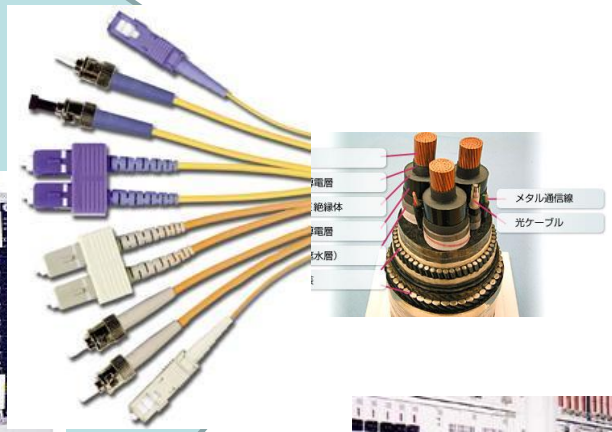
   concealing {file name, file descriptor, etc.}



Distributed file storage image in DHT system

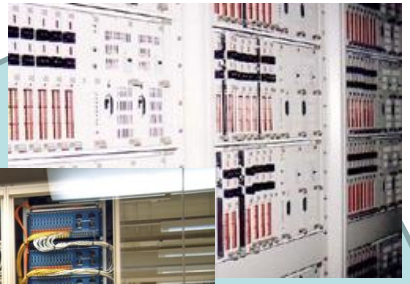
データセンター

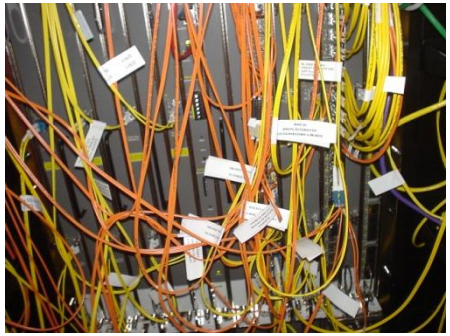
電話局

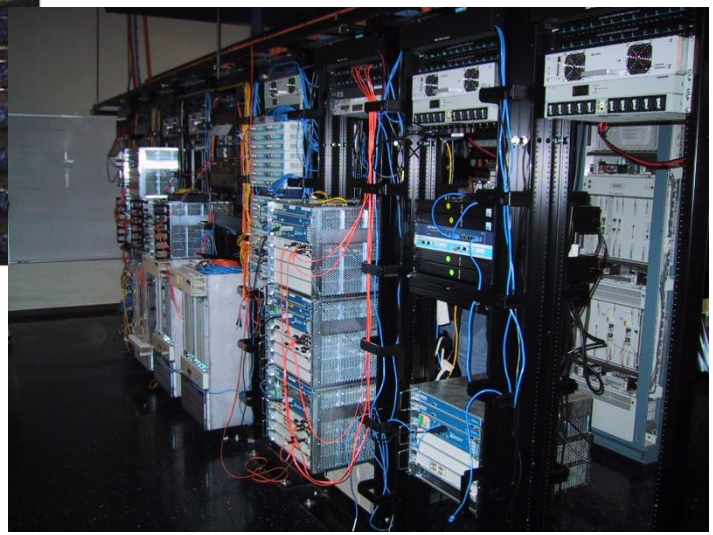


自宅

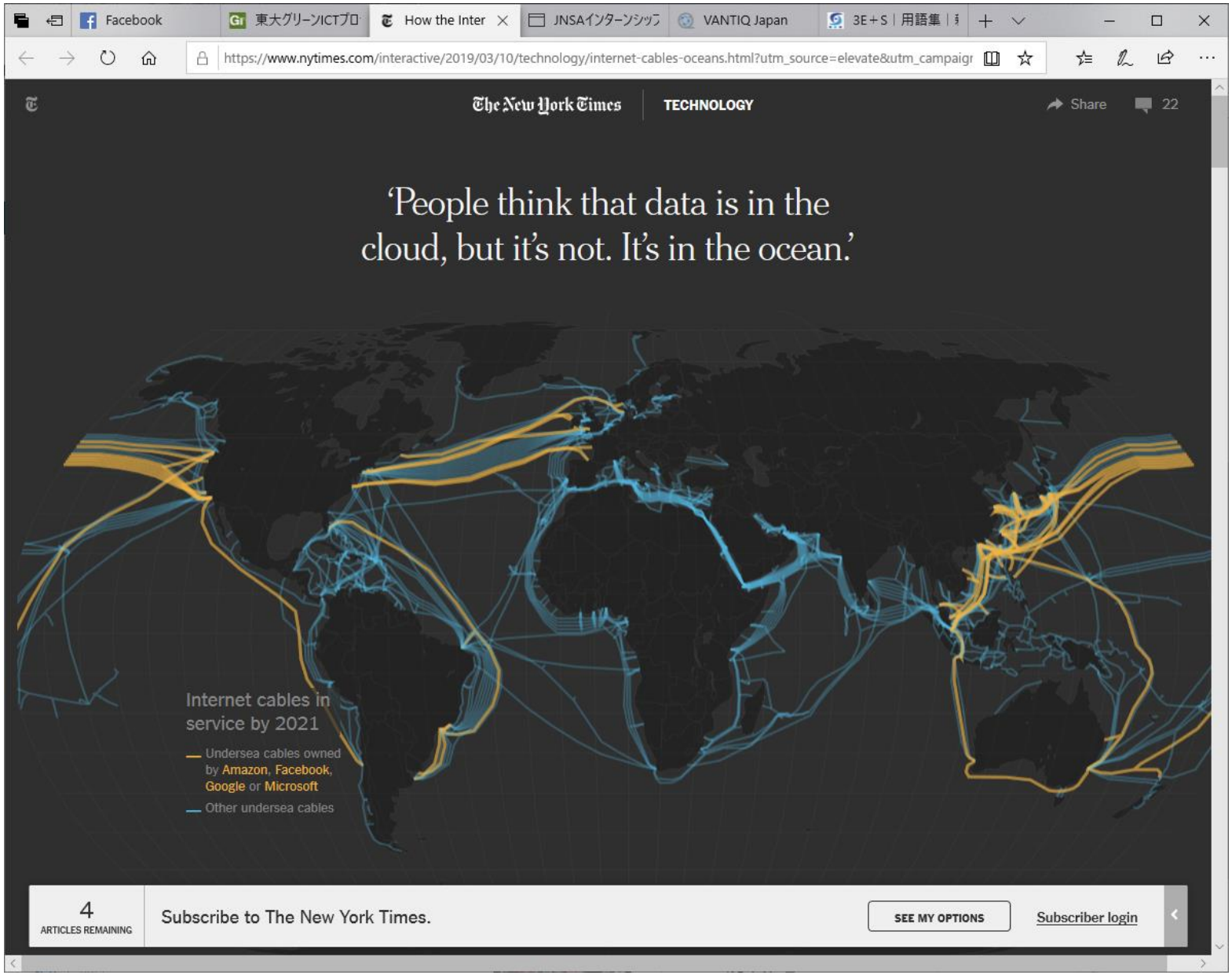
データセンタ



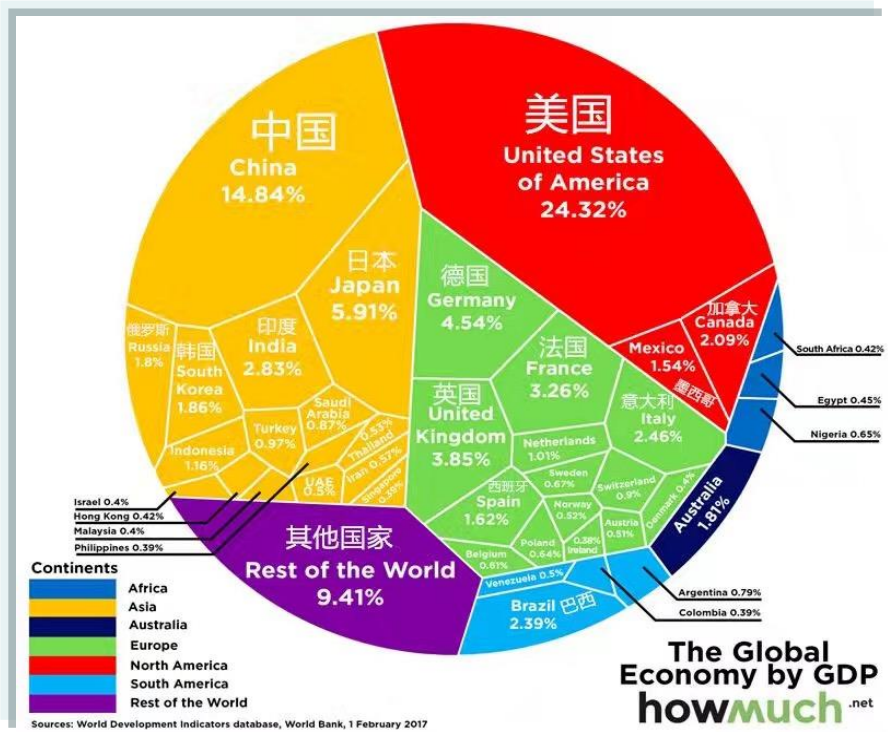




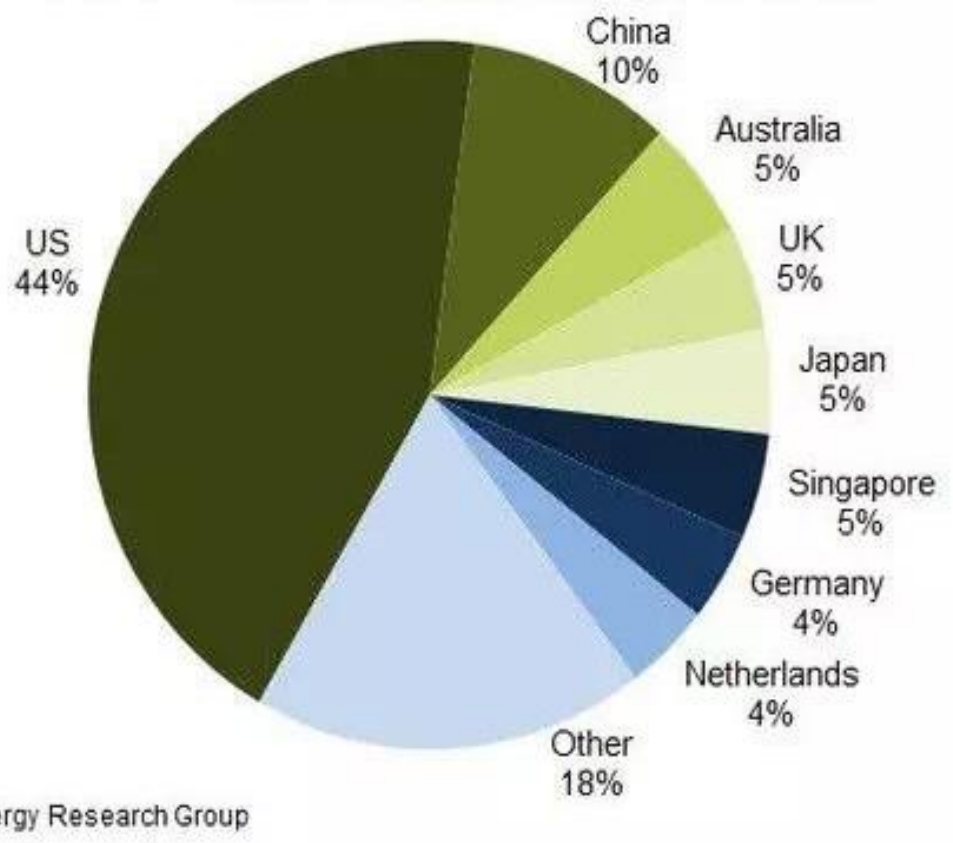
https://www.nytimes.com/interactive/2019/03/10/technology/internet-cables-oceans.html?utm_source=elevate&utm_campaign=LinkedIn&utm_medium=social&utm_term=0_597bc2b578-0e31f613ae-56581769&_lrsc=e6ed9c97-dc44-49f1-a6fc-6888803fb85b&fbclid=IwAR3tk1YDKjonsPctO6nM-u6TMX0Z5Vnw1qmEVg1rS4rqEYQAT7YQBn5fMbE



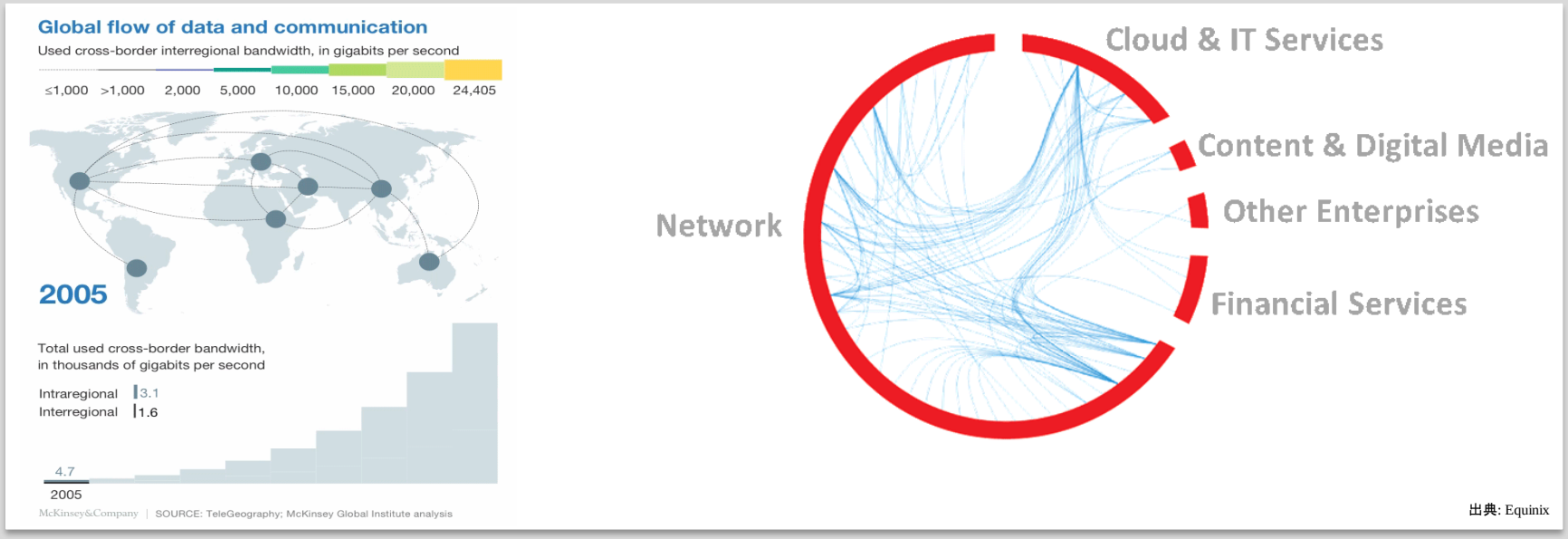
2019年頃　米国:中国:日本 ＝ 9：3：1



2016年世界各国GDP占比



グローバルな通信トラフィック vs デジタル密度の中心



出典: Equinix

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014

資料: EQUINIX社 古田敬 社長

データセンターで 何が起こっている？

1. 大規模化が加速
   - ✓ MW以下、数十MW、数百MW : **“異なる解“**
2. **オープン化** と **“PULL”** ビジネス構造化
3. 持続継続・加速する質量増加の中での **『爆縮』** 現象
   - ✓ 継続: ムーアの法則
     - □ 増加: 密度(**回路**・電力・熱) & **速度**
     - ■ 不変: **伝搬速度**、(冷却・搬送)媒体密度

Innovations in Data Center architecture

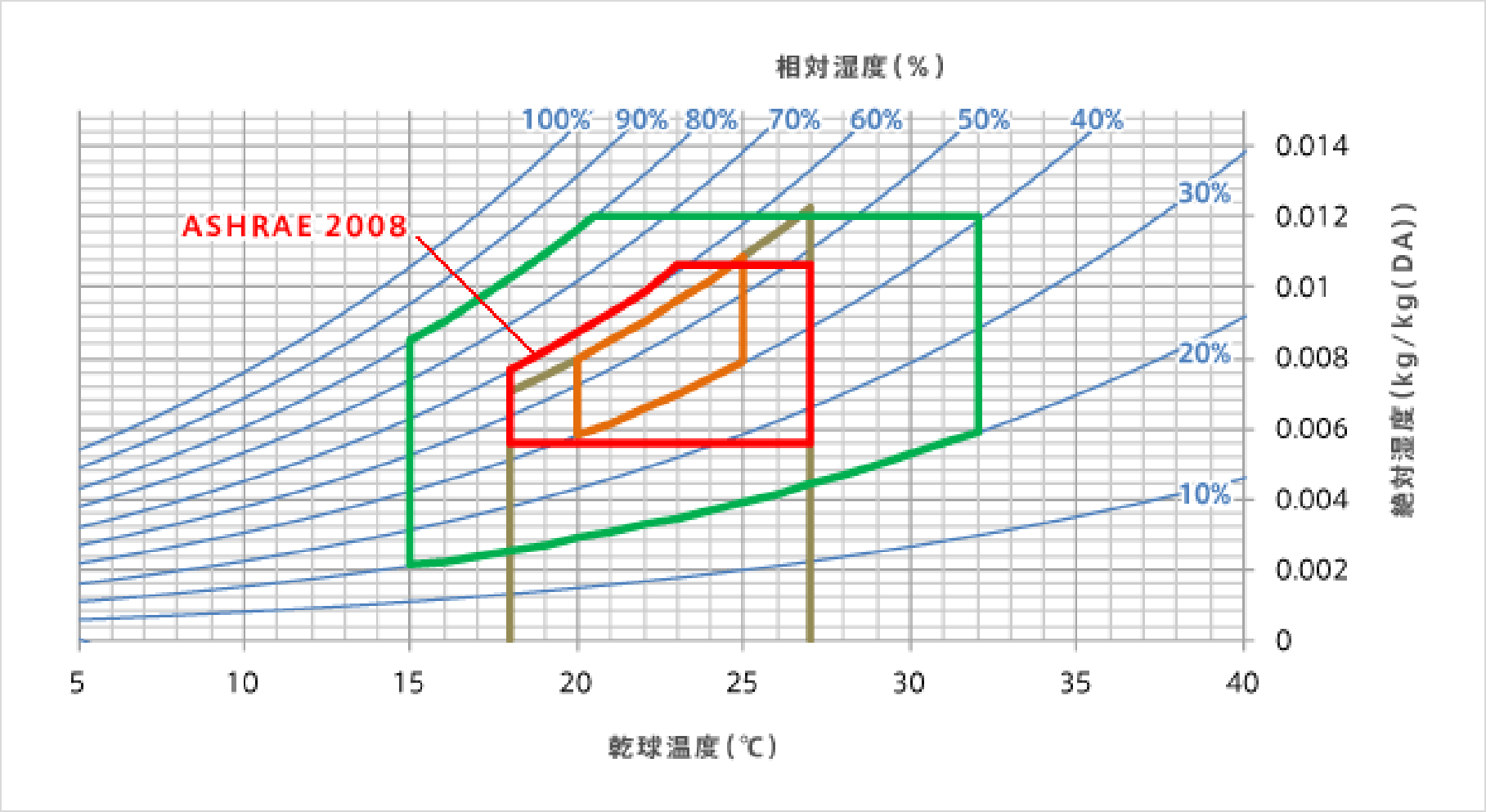
1. **Open & Transparent , i.e., white xx**
   - ✓ HW : Chip, board, server, switch, router, Electric power, HVAC
   - ✓ SW: Operating system, Middleware , Application

Innovations in Data Center architecture

1. **Open & Transparent , i.e., white xx**
   - ✓ HW : Chip, board, server, switch, router, Electric power, HVAC
   - ✓ SW: Operating system, Middleware , Application

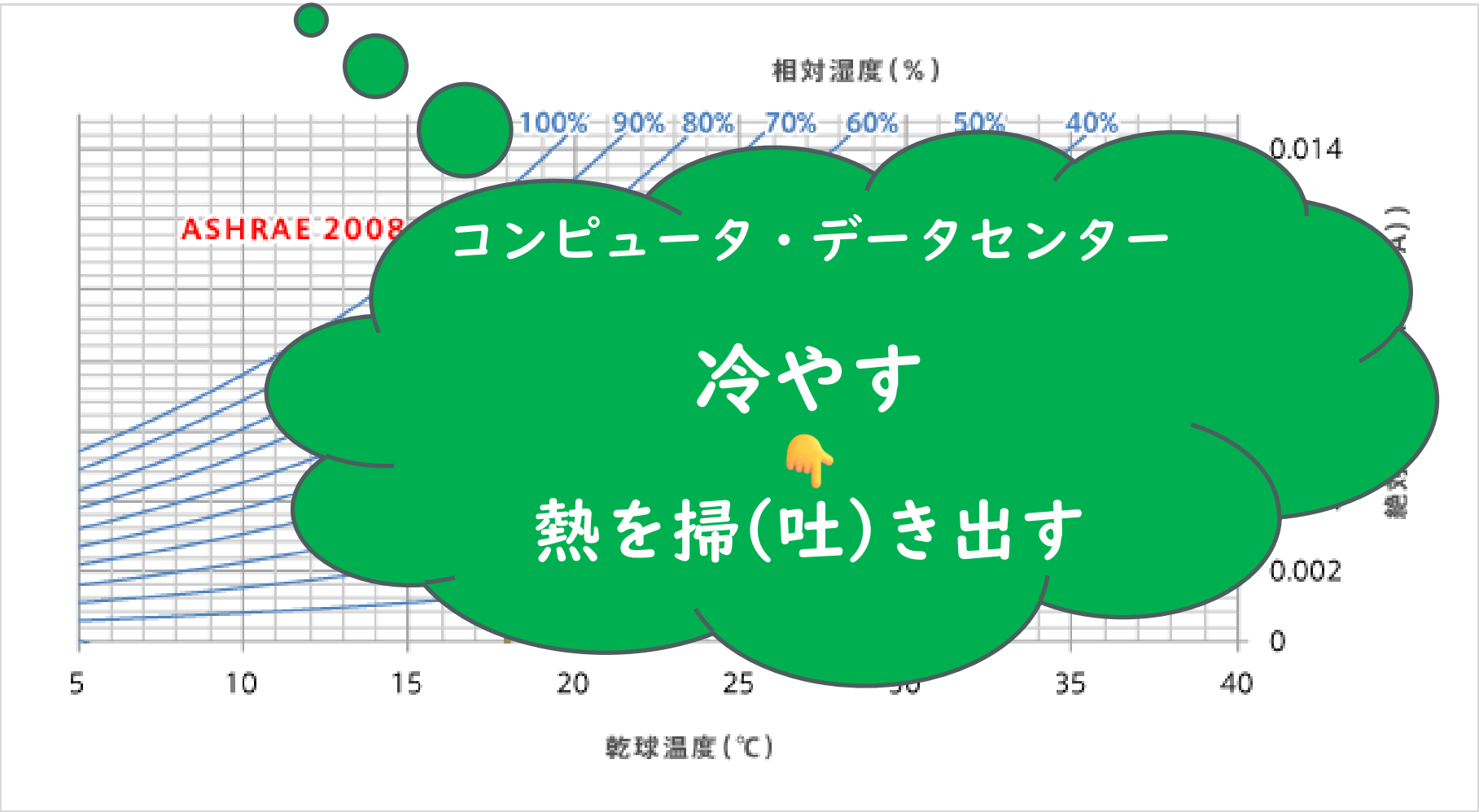
Open & Transparent
Not only for
server/switch/router,
but also for
“facilities”, with DCIM

動作条件の緩和(by ASHRAE)



ASHRAE：American Society of Heating, Refrigerating and Air-Conditioning Engineers（米国暖房冷凍空

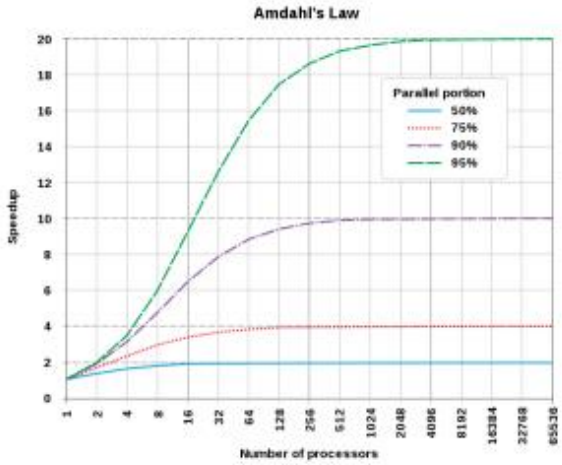
動作条件の緩和(by ASHRAE)

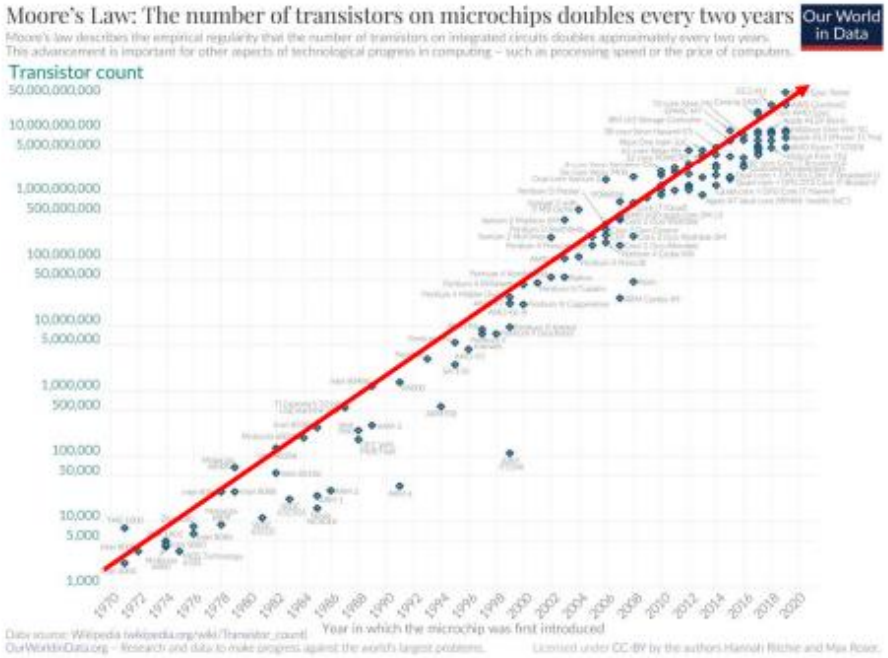


ASHRAE：American Society of Heating, Refrigerating and Air-Conditioning Engineers（米国暖房冷凍空

Moore's Law, What's Driving IT Power and Heat, and What We Can Do About It

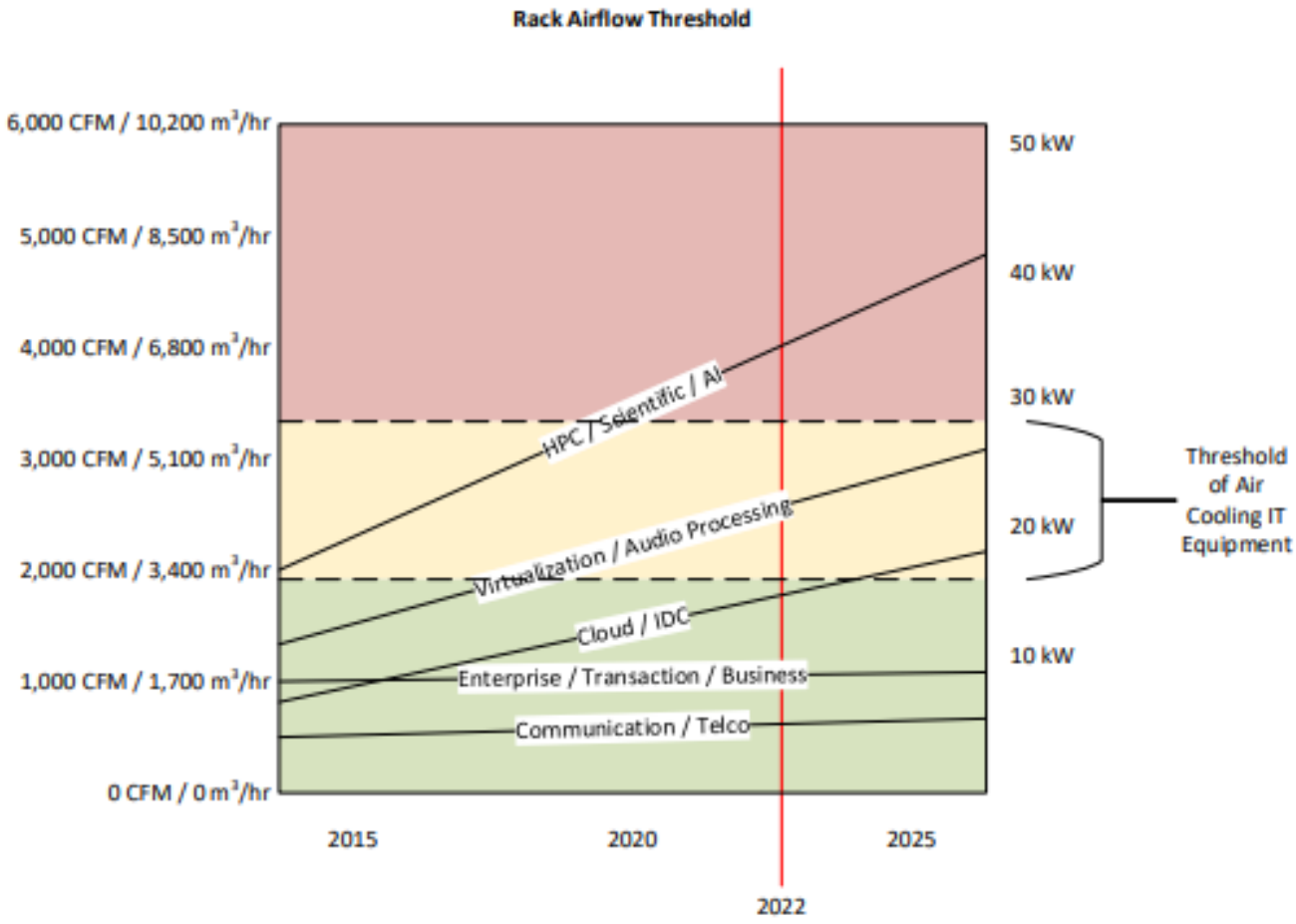
- Moore's Law shows that the number of transistors in a chip continues to grow as expected
- Increases in transistor count does increase power consumed and heat generation
- Increasing number of cores per chip alleviates some growth in power but is subject to Amdahl's Law concerning parallelism
- Without the ability to continue increasing chip core count, chip manufacturers can only increase chip clock frequency to continue performance gains

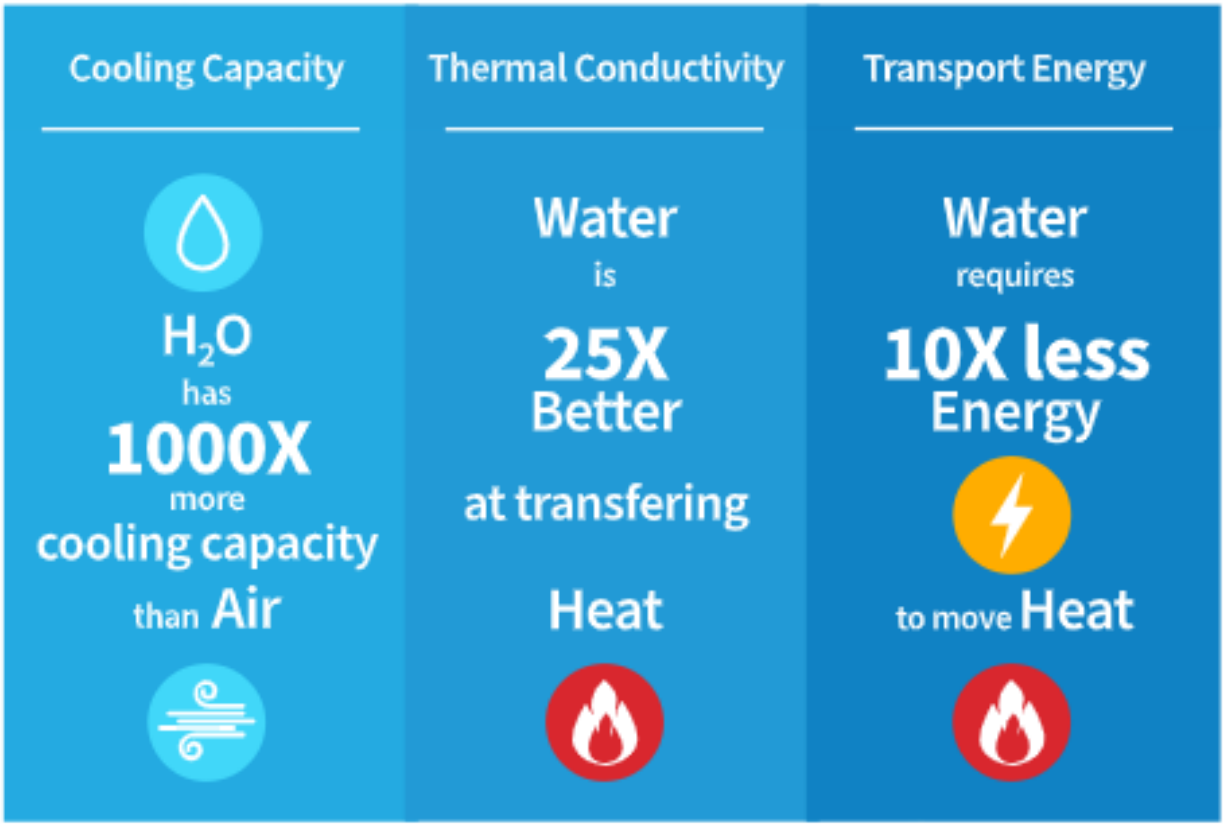




Confidential – For Internal Use Only

nVent


Rack Airflow Threshold
6,000 CFM / 10,200 m³/hr
5,000 CFM / 8,500 m³/hr
4,000 CFM / 6,800 m³/hr
3,000 CFM / 5,100 m³/hr
2,000 CFM / 3,400 m³/hr
1,000 CFM / 1,700 m³/hr
0 CFM / 0 m³/hr
50 kW
40 kW
30 kW
20 kW
10 kW
HPC / Scientific / AI
Virtualization / Audio Processing
Cloud / IDC
Enterprise / Transaction / Business
Communication / Telco
Threshold of Air Cooling IT Equipment
2015
2020
2025
2022



https://www.supermicro.com/ja/solutions/liquid-cooling より引用

要は、、何が起こっているの？

- モジュール構成の 再定義・再設計
  - 昔からの課題
    - **パフォーマンス**(具体的には、Bandwidth & Latency)

      CPU >> メモリネットワーク >> ストレージ

  - 新しい(けれども昔からの)課題
    - 『物理的距離』
    - 『熱』
    - 『電力使用量・使用料金』

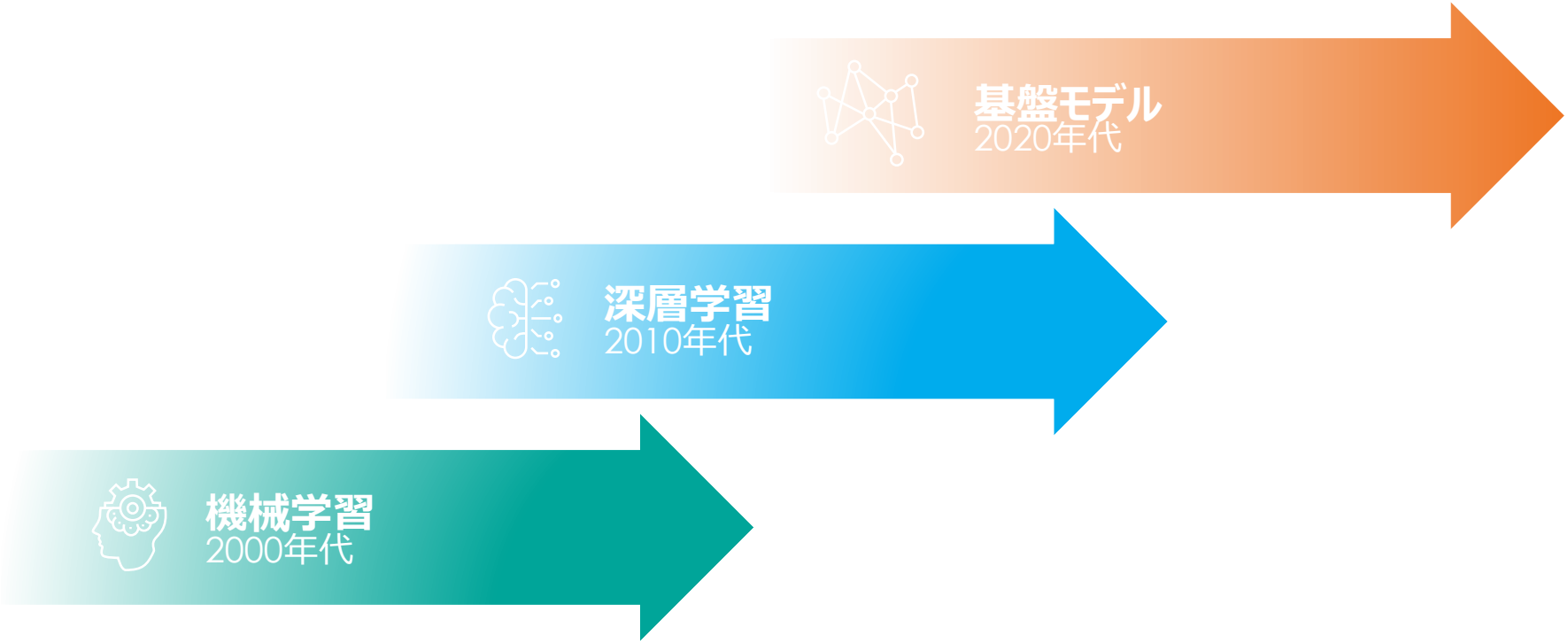
Innovations in Data Center architecture

1. **Open & Transparent , i.e., white xx**
   - ✓ HW : Chip, board, server, switch, router, Electric power, HVAC
   - ✓ SW: Operating system, Middleware , Application

2. **Data (Storage) Centric**
   - ✓ Big Data collection and analysis
     - ➢ Explosion of amount of data
     - ➢ Cross domain data integration
   - ✓ **Processor/CPU centric ➔ Data/Storage centric**
     - ➢ **Migration overhead: data >> processing image**

       (*) Contribution of

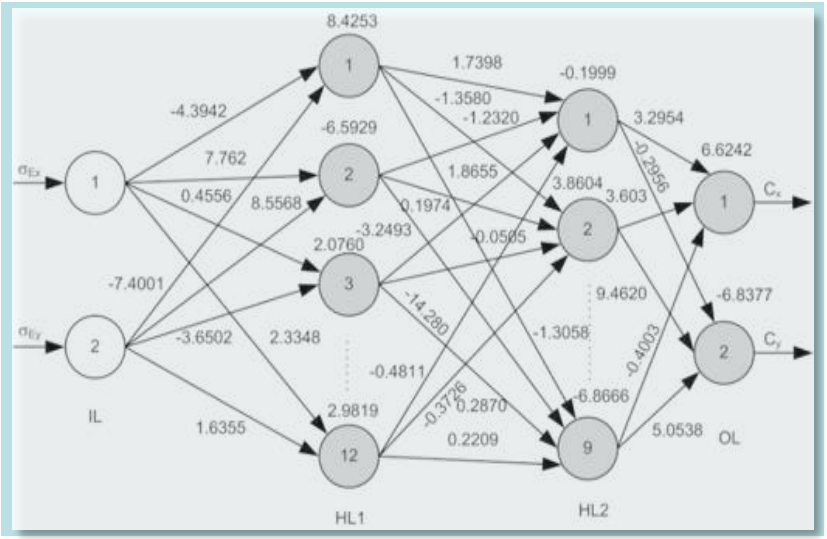
       1. **VM (Software-Defined North Bridge) technology**
       2. **Software-Defined Storage technology**



資料：SambaNova社

現代AIはデータフローの問題

```
37  #include <iostream>
38 using namespace std;
39
40 int _tmain (int argc, _TCHAR* argv[])
41 {
42
43      int iVal1 = 0, iVal2 = 0, iVal3 = 0;
44
45       printf("Enter three numbers:");
46       scanf("%d %d %d", &iVal1, &iVal2, &iVal3);
47
48      if (iVal1 >= iVal2)
49       {
50           if(iVal1 >= iVal3)
51             printf("Largest number = %.2d", iVal1);
52           else
53             printf("Largest number = %.2d", iVal3);
54       }
55       else
56       {
57           if(iVal2 >= iVal3)
58             printf("Largest number = %.2d", iVal2);
59           else
60             printf("Largest number = %.2d", iVal3);
61       }
62
63  getchar ();
64  return 0;
65 }
```



ソフトウェア 1.0

- コードで書かれている (C++, ...)
- ドメインの専門知識が必要

現代AIはデータフロー（ソフトウェア2.0）

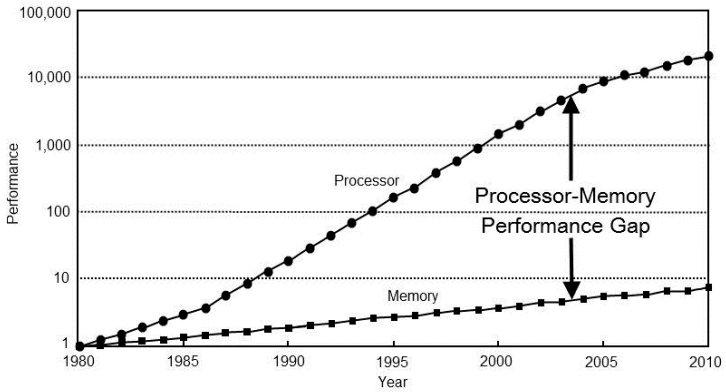
- コードではなくデータがモデルを鍛える
- ニューラルネットワークの重みで記述

資料：SambaNova社

AI半導体の概況

最近のAI演算におけるCPU/GPUの課題はスパース性への対応と消費電力

- 演算性能とメモリ性能の乖離
  - メモリの階層構造によりデータの遅延を隠蔽
  - キャッシュヒット率に依存
  - DENSEなGEMM（密行列計算）には最適な構造
  - スパース（疎行列）では性能ギャップが露呈
- 現代半導体においては、演算に必要な電力消費より、データの移動に必要な電力消費の方が遥かに大きい
  - **いかにデータを近くに配置し、データの移動距離を短くして演算出来るかが重要**



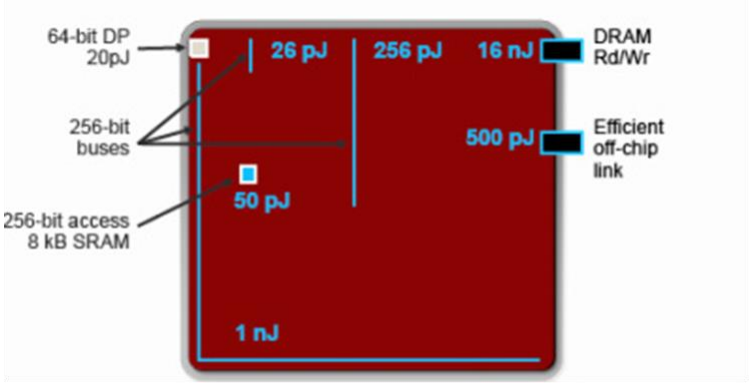


Figure: Computation cost is significantly lower than communication cost in 28nm NVIDIA chips (Source: Bill Dally)

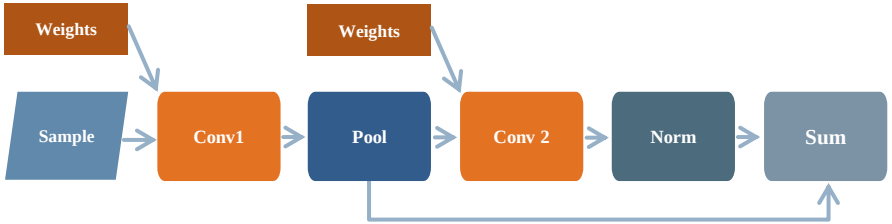
多くのスタートアップが従来方式に代わる「AI半導体」を提案している。

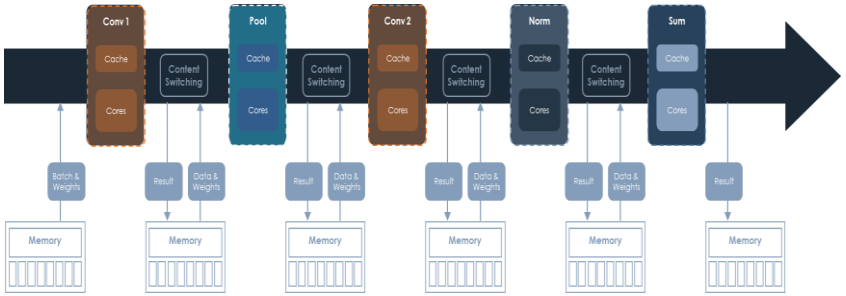
資料：SambaNova社

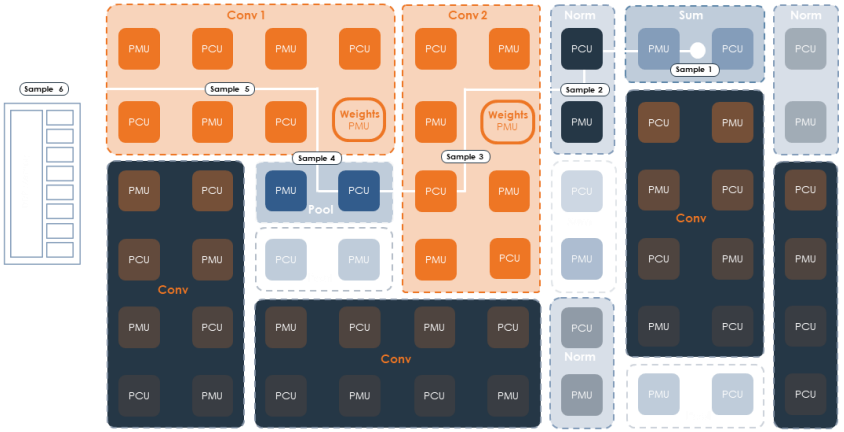
従来手法とデータフローの違い

データフローアーキテクチャは局所性と並列性を活用

畳み込みグラフ







計算カーネル毎にメモリアクセスが必要なため、高速メモリを演算器の近くに配置する必要
→ **高速メモリは容量が小さい**

外部メモリアクセスが最小化できるので、高速メモリが不要 → **メモリ大容量化可能**
データが移動しなければ演算が発生しない
→ **スパース性にも強い**

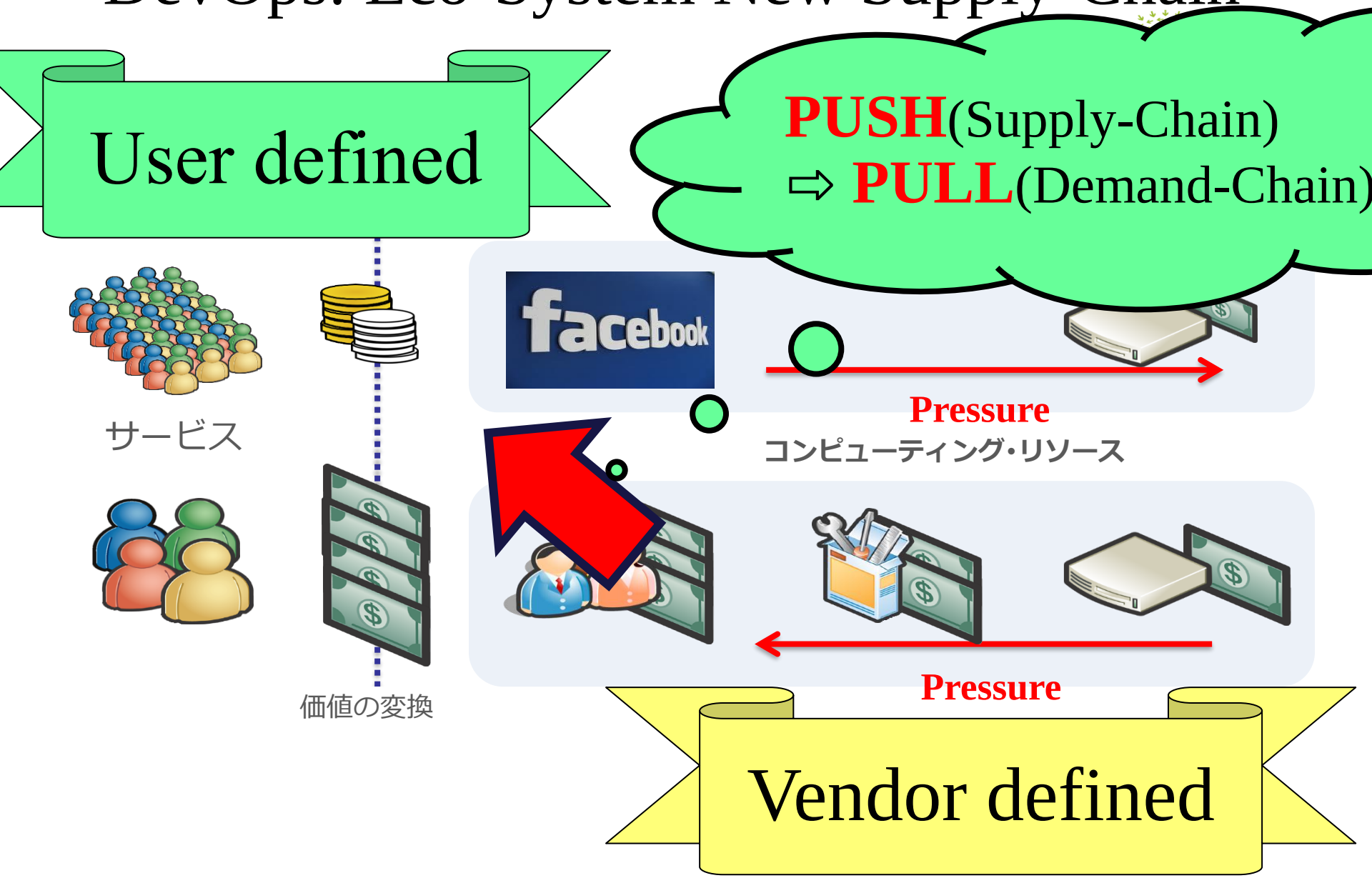
資料：SambaNova社

OCP (Open Compute Project)に代表される、DevOps型の 設計・実装

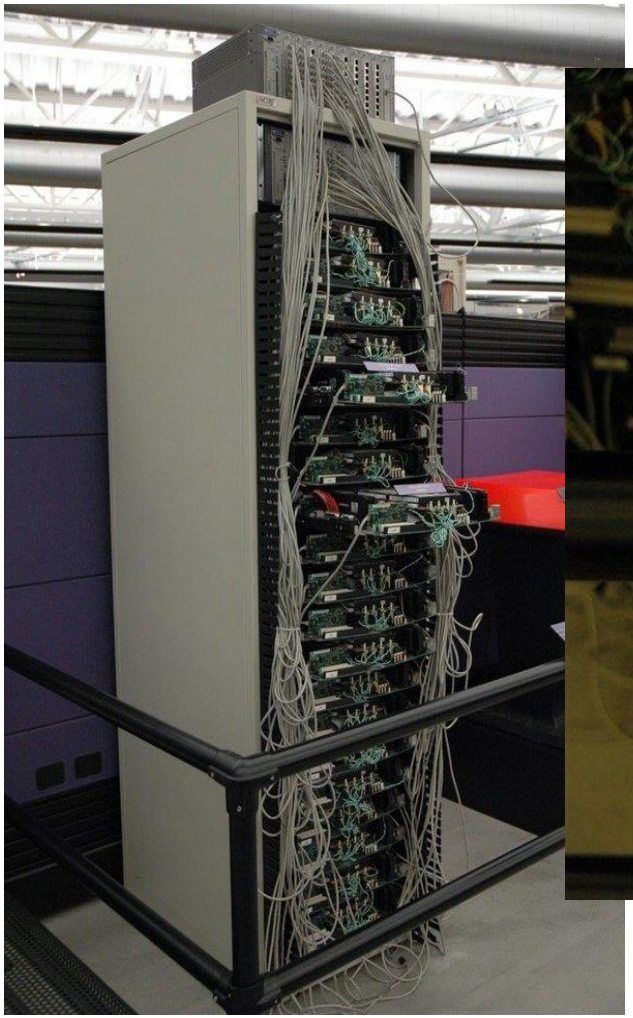
汎用品・技術による カスタマイズ化
(e.g., マイクロ・ファブ化)
Driven by GAFA+M, BAT

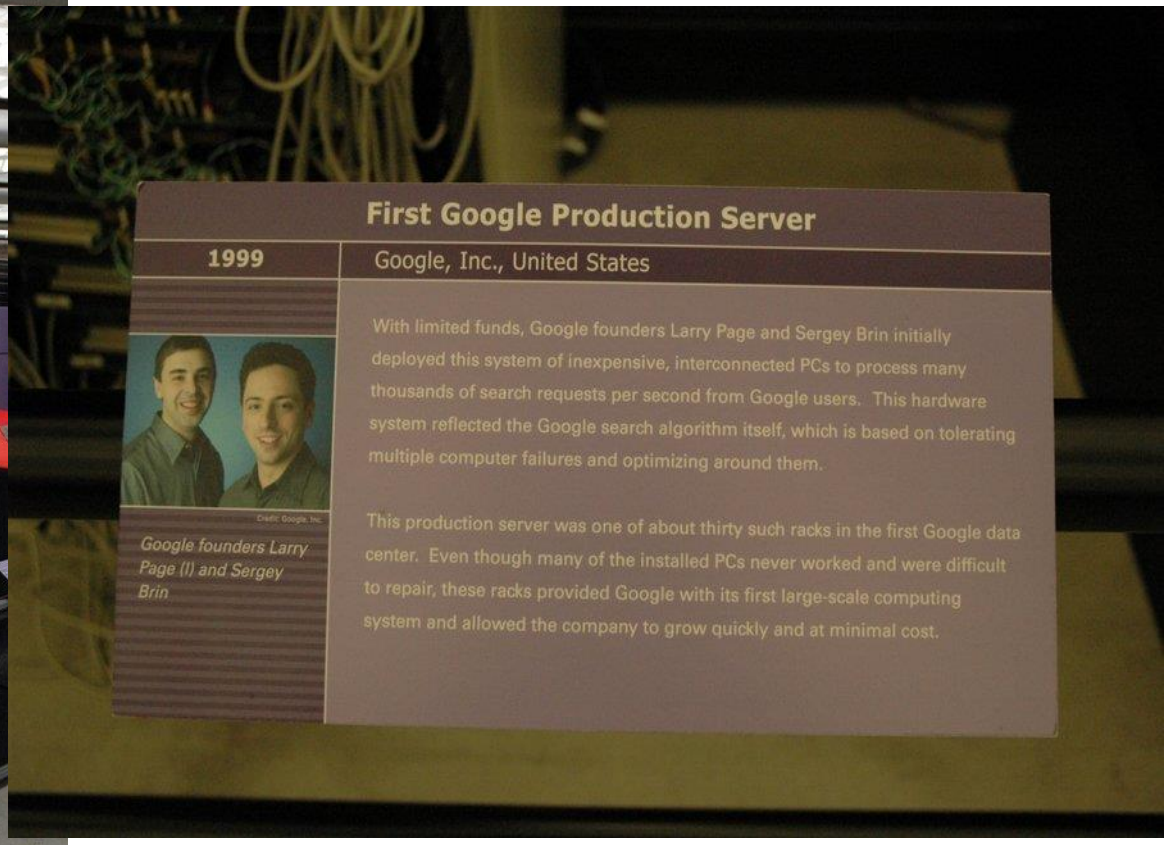
USA: Google, Amazon, Facebook, Apple, Microsoft
China: Baidu, Alibaba, Tencent(WeChat)

DevOps: Eco-System New Supply-Chain

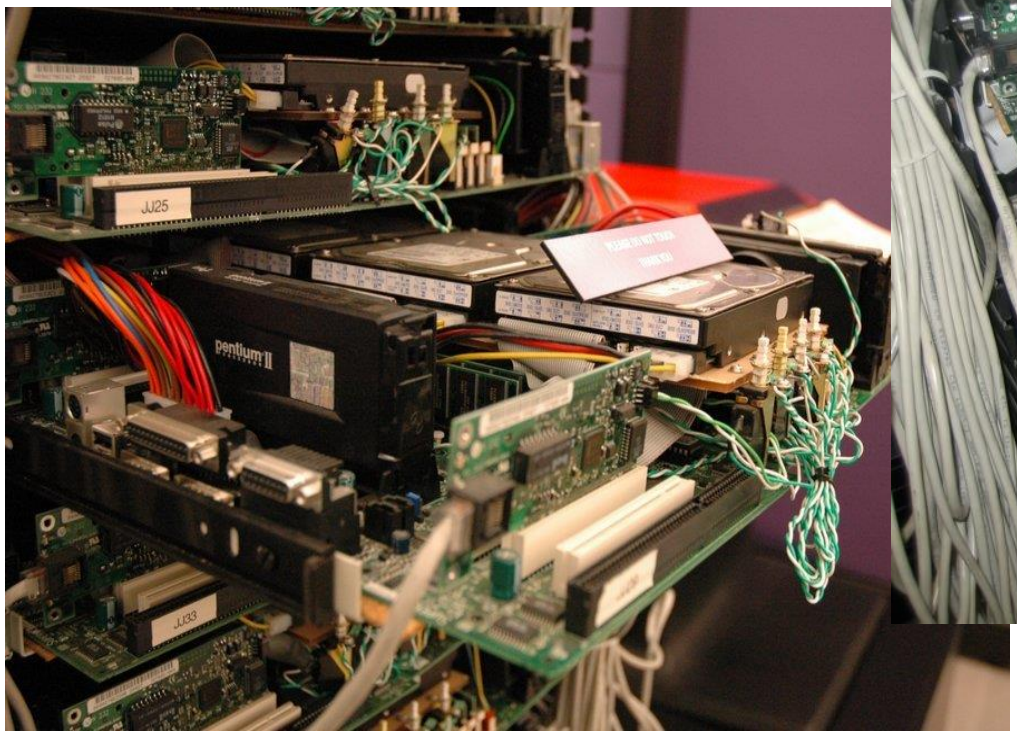


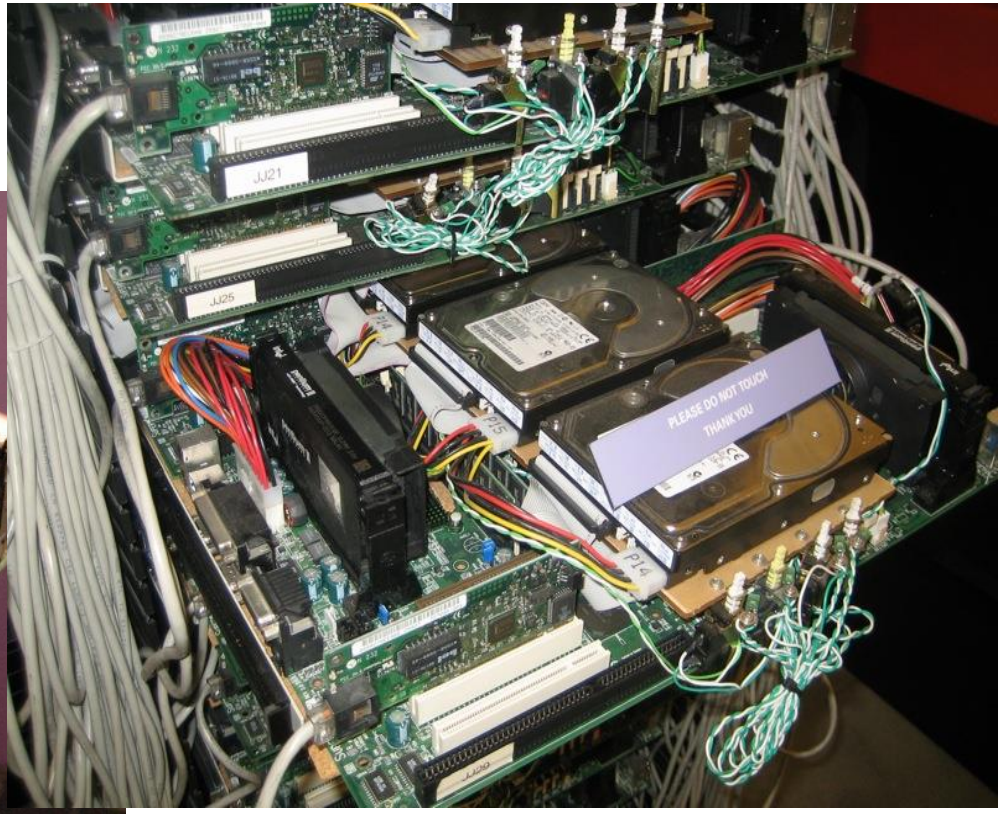
Google の 最初のサーバ (Computer Museum in California)

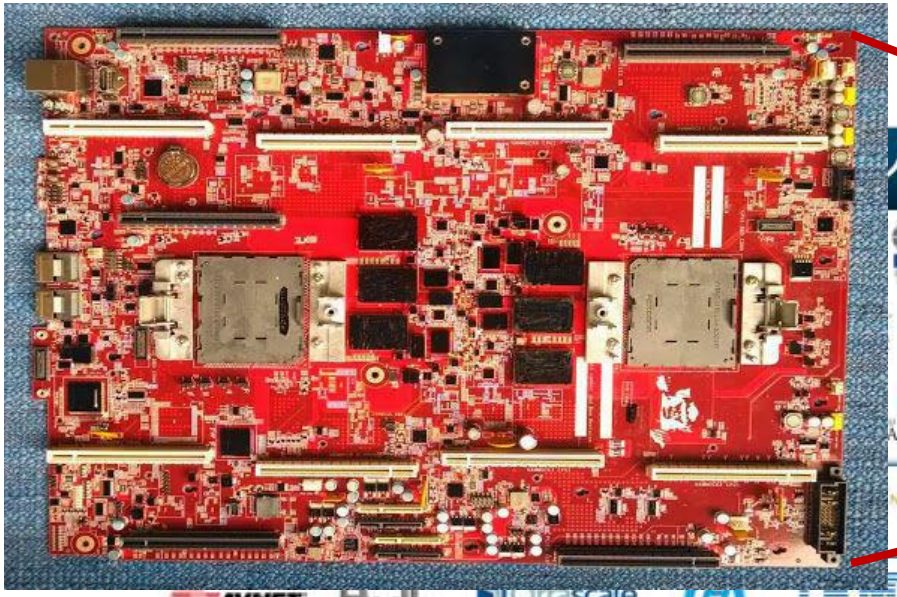


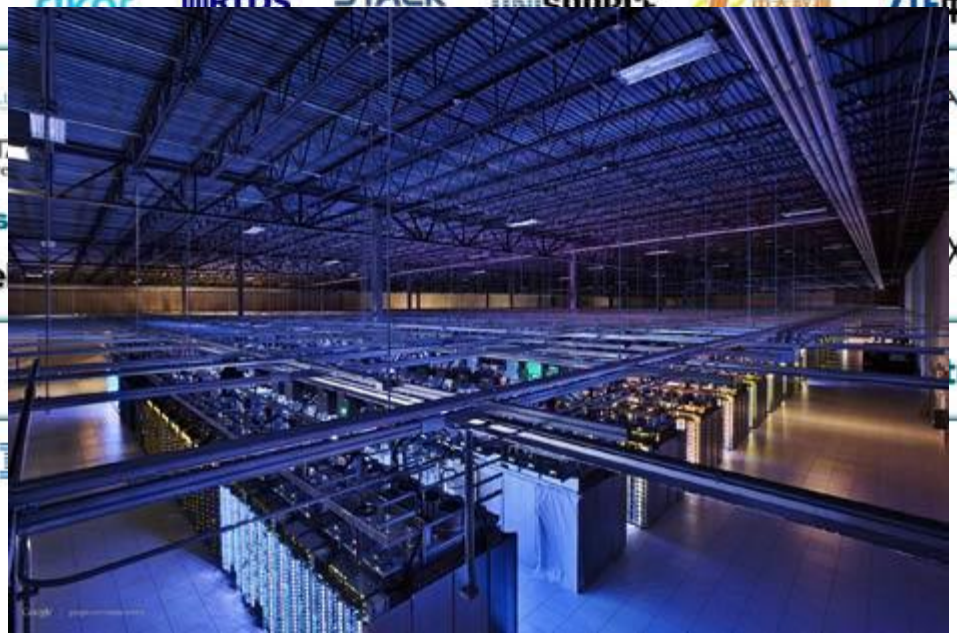


Google の 最初のサーバ
(Computer Museum in California)



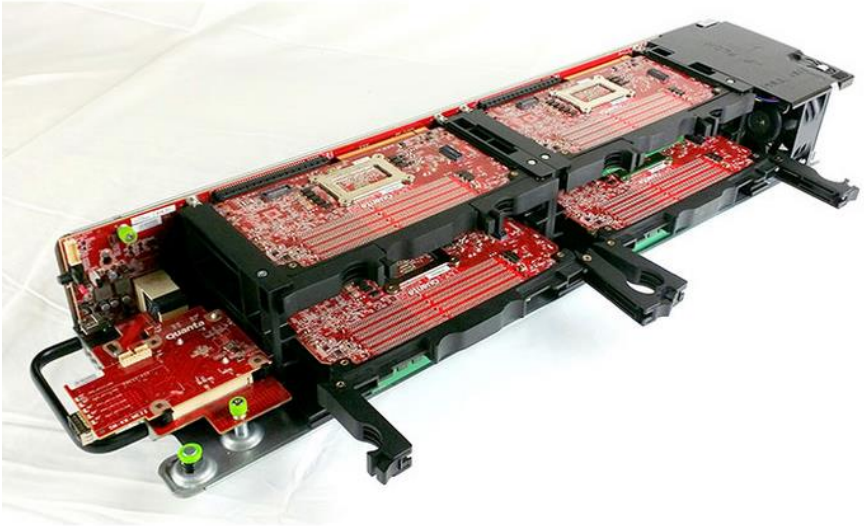






AVNET
inspur
OCF
ONE STOP SYSTEMS
STACK
ALTERA
HGST
numas
XILINX
Chip / SOC

Yosemite / 1S Server



Intel Xeon D-1500 SoC
HighPowered-SoC Micro Server 210×110mm M.2 SSD 10GbE 65W

OpenRackV2 192 SoC servers
PCI-Express x16 mechanical slots
X86 （ARM, Power）
40GbE Mellanox C-4 hybrid mezzanine card 400W

intel created with Xeon D processor and worked with Quanta to design the board and to get the microserver manufactured.
Facebook and Quanta designed the sideplane and the hybrid mezzanine card along with Mellanox.

出典: https://tech.fb.com/open-compute-project-announces-the-first-open-hardware-multi-node-rack-scale-high-performance-micro-servers-/

資料: OCP Japan 座長 藤田龍太郎 氏

Open COMPUTE SERVER

2U4ノード22インチ幅
Open Compute Project仕様
Quanta社製

2U3ノード 22インチ幅
Open Compute Project仕様
Quanta社製

12U24ノード 19インチ幅
Microsoft
Open Compute Project仕様
Gig

Microsoft OCP & Cloud Server

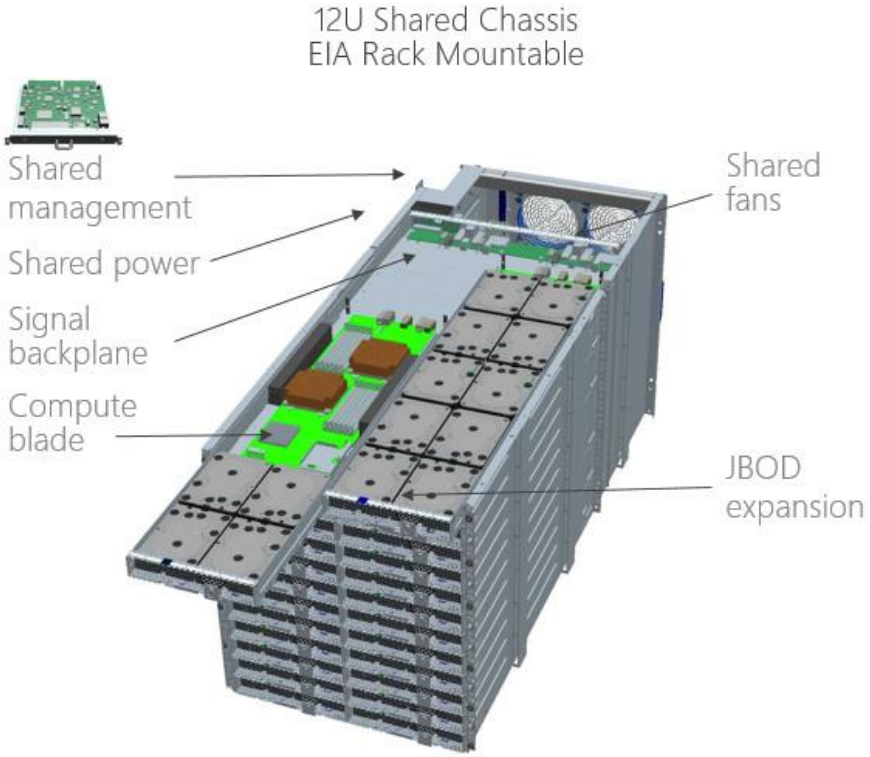


Diagram: Microsoft cloud server specification

Microsoft　OCPにコントリビュート

chassis v1.0
Blade v1.0
JBOD v1.0
Chassis Management v1.0
Network Mezzanine v1.0
SAS Mezzanine v1.0
Chassis Management Software source code

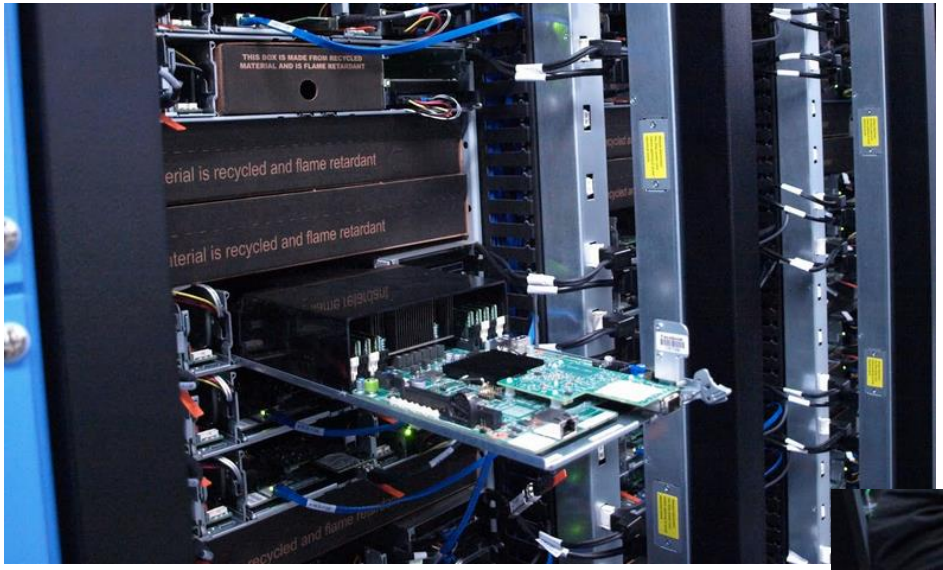
出典：https://gigaom.com/2014/01/27/microsoft-the-software-king-wants-to-tell-the-world-how-to-build-servers/

資料：OCP Japan 座長 藤田龍太郎 氏

Knox (Open Vault)

Cold Storage用サーバ

OCP のサーバーとラック



ネジを１本も使わずに
キッティング

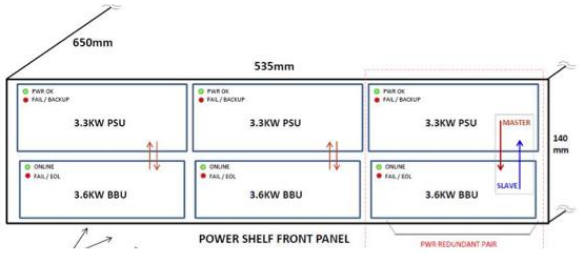
フタもなければ、
フロントパネルもない

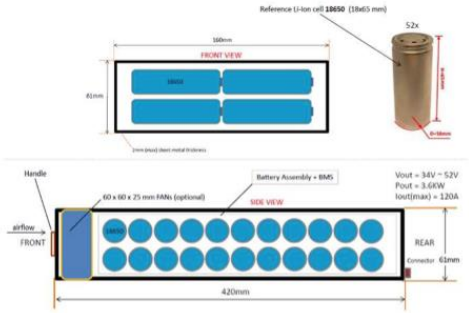
出典：http://wp.me/pwo1E-2Ku

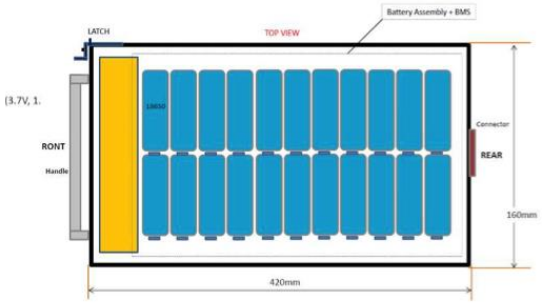
資料：OCP Japan 座長 藤田龍

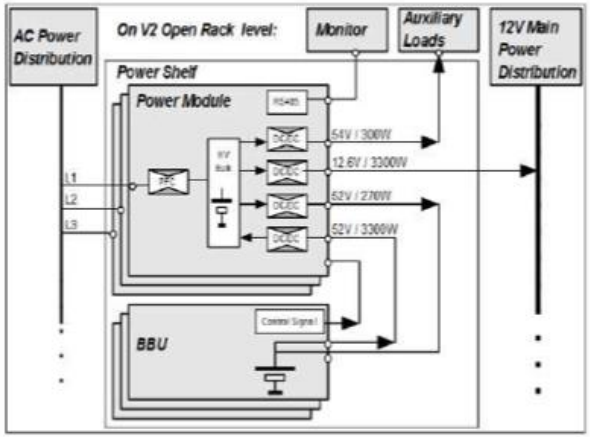
Power shelf

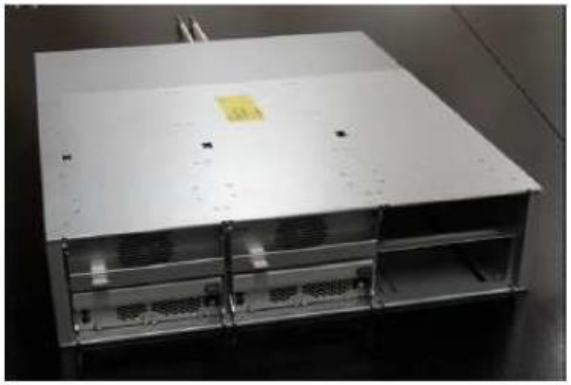
Power modules and Li-ON batteries in the same shelf
Single 12V Busbar output(535A)
Three-phase input
2+1 Redundancy + batteries
534mm x 612mm x 19mm

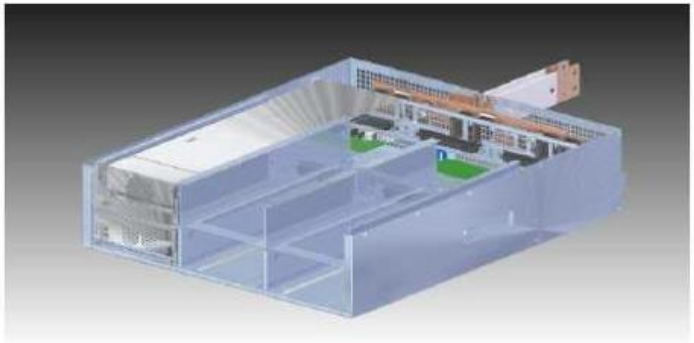






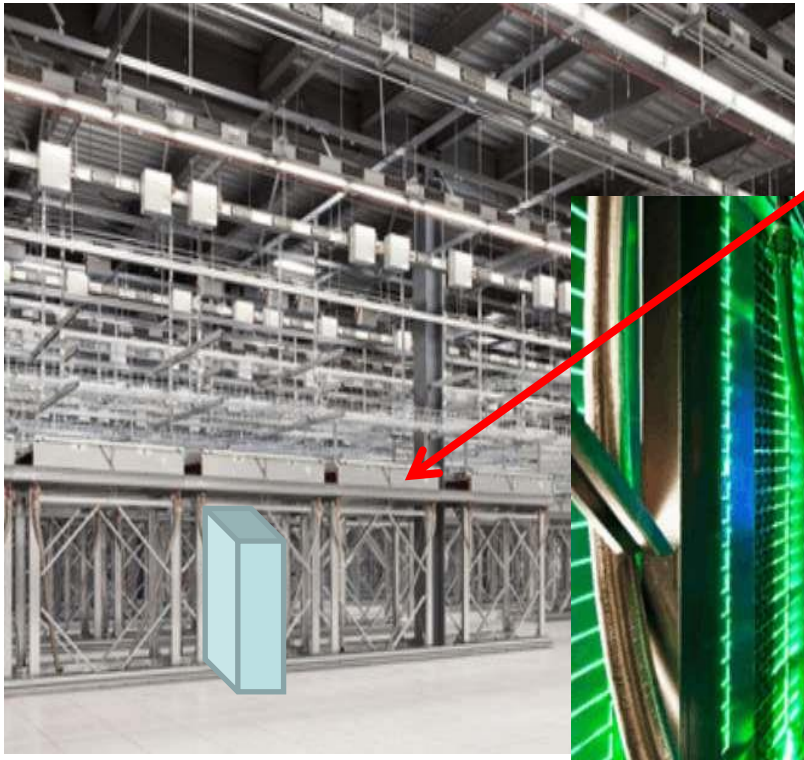






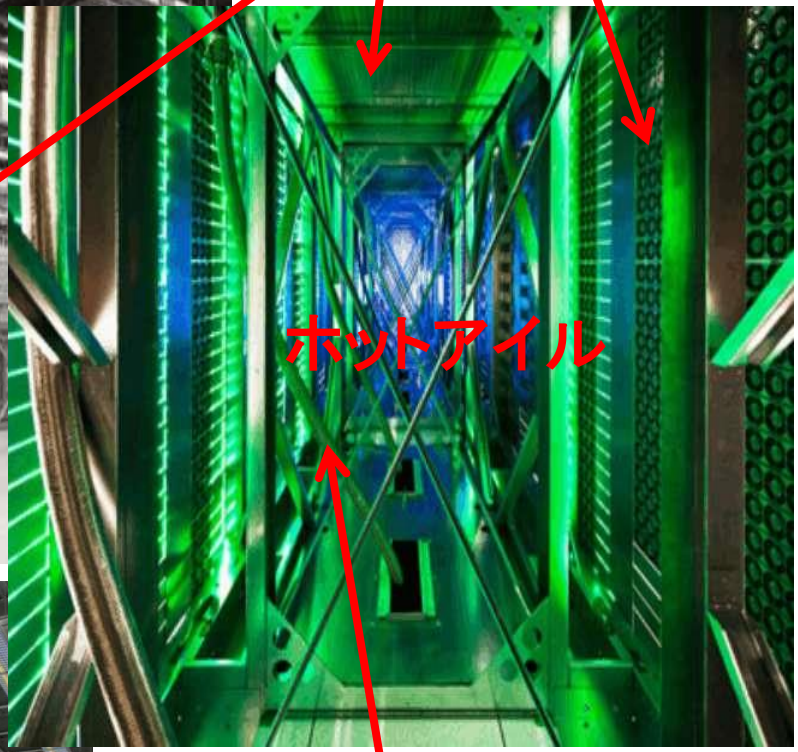
資料: OCP Japan 座長 藤田龍太郎 氏

Google

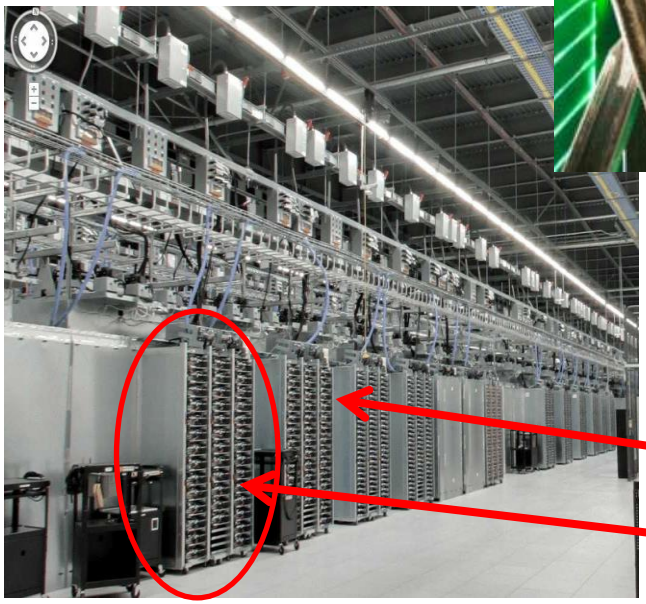


熱交換コイル

ラック背後のファン群



床下から冷却水



隙間が空いている

3ラックで1クラスタを構成？

Googleのデータセンター（Google）

GoogleのPUE低減方策　　負荷の移動

データセンターに冷却機を設置せず、暑い日には他のデータセンターに負荷を移動

- ベルギーのデータセンター
  冷却機は設置されていない。

- 世界各地にあるGoogleデータセンター
  ベルギーデータセンターで冷却機が必要になるような暑い日（年間7日程度発生）には、電源を落とし、他の地域のデータセンターに負荷を移動する。

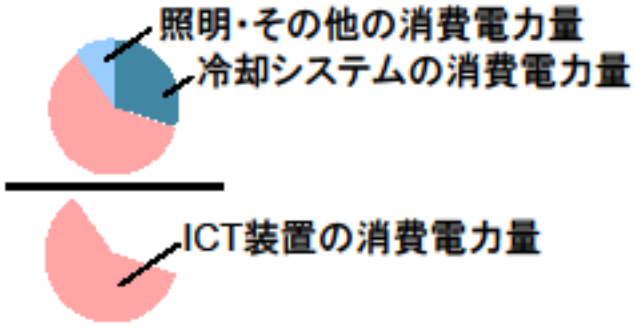
Follow the moom

世界各地でデータセンターを運営する事業者が、その負荷を、電力料金が安く気温が低い夜間の地域に移動させる方法アイデアは提唱されているが、実装できるのはGoogleなど超大手企業に限られる可能性が高い。

データセンタのエネルギー効率を示す指標PUEと向上策

● PUE (Power Usage Effectiveness 電力使用効率)

$$\mathrm{PUE} = \frac{\text{DC全体のエネルギー消費量}}{\text{ICT装置のエネルギー消費量}}$$



PUEが**小さい**ほど省エネルギー

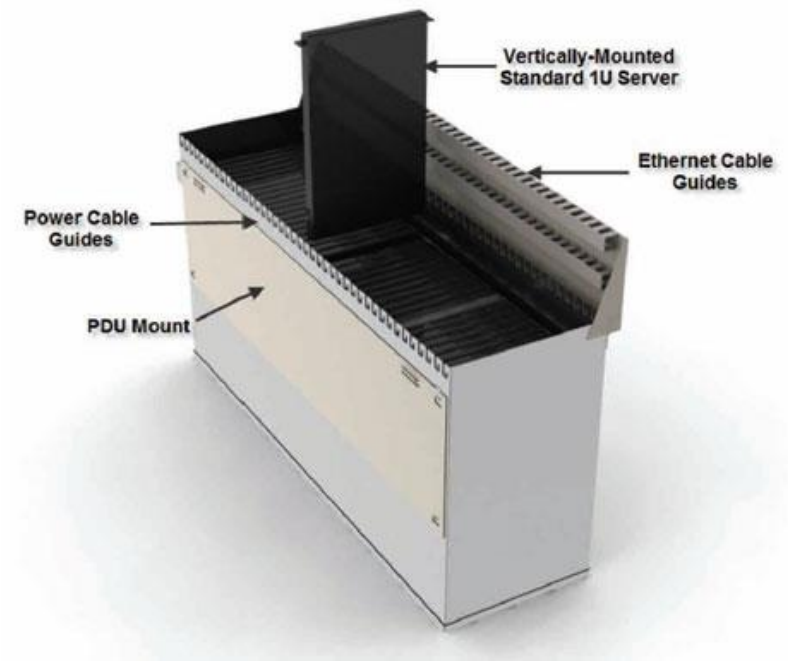
●PUEの向上策

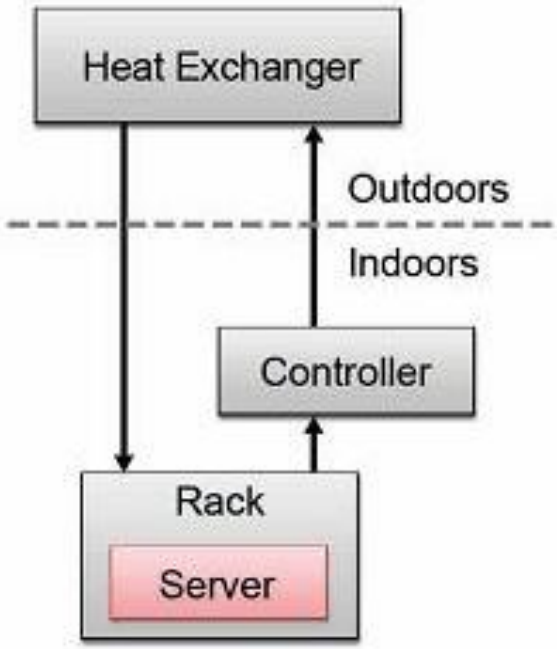
- IT装置の低消費電力化 → 効果大
- 空調システムの効率向上、容量適正化 → 効果大
- 電力供給システムの効率向上、容量適正化 → 効果中～大
- 照明や施設サポート部分の省エネルギー → 効果小

サーバをオイルに浸して冷却 (Green Revolution Cooling)

- サーバをオイルの冷却液(GreenDEF Coolant)に完全に浸すことで冷却するシステム。

データセンターの消費電力を45%削減。





サーバを冷却液に完全に浸す

サーバの消費電力を5～25%削減できる
サーバ内部のファンが不要になるため。

全体構成

ラックと屋外の熱交換機との間で冷却液を循環させる。
消費電力は、サーバの消費電力100Wあたり、5W。

山中のデータセンター（The Mountain Data Center）

- 米国ミズーリ州オザーク高原（Ozark Mountains）の地中、30万平米の施設 The Mountain Complexの中にあるデータセンター。

断熱性の高い岩盤が地熱冷却のための理想的なヒートシンクとなる

- オザーク高原の風景
- 岩盤をくりぬいたデータセンター

地下データセンター（Bahnhof）

冷涼な地下環境（15℃）を活用

▶改装前の核シェルター

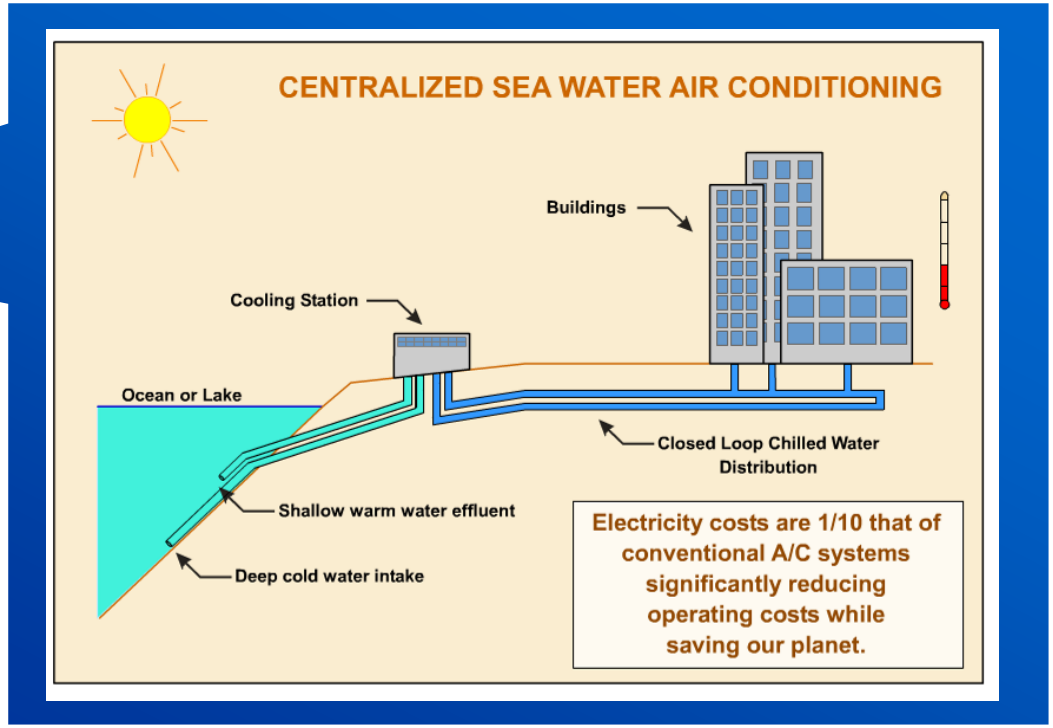
地下データセンター（Bahnhof）

▶データセンターの入り口

海水空調データセンター（モーリシャス）

- モーリシャスは世界初の、海水空調（SWAC: Sea Water Air Conditioning）をビルトインした1万平米のデータセンターを建設中 。2010年12月にオープン予定。
- 海水空調は米国やカナダ[illegible]用中。冷房コストを9割程度削減。

既に実用化済みの省エネ技術をデータセンターに適用。



モーリシャス

アフリカ、アジア、オートストラリアの中間に位置し、地震帯から離れており、政治的・社会的にも安定しており、ネットワーク接続環境もよいことから、アフリカ、アジア、中東の情報ハブとなることを目指している。

海水空調システム

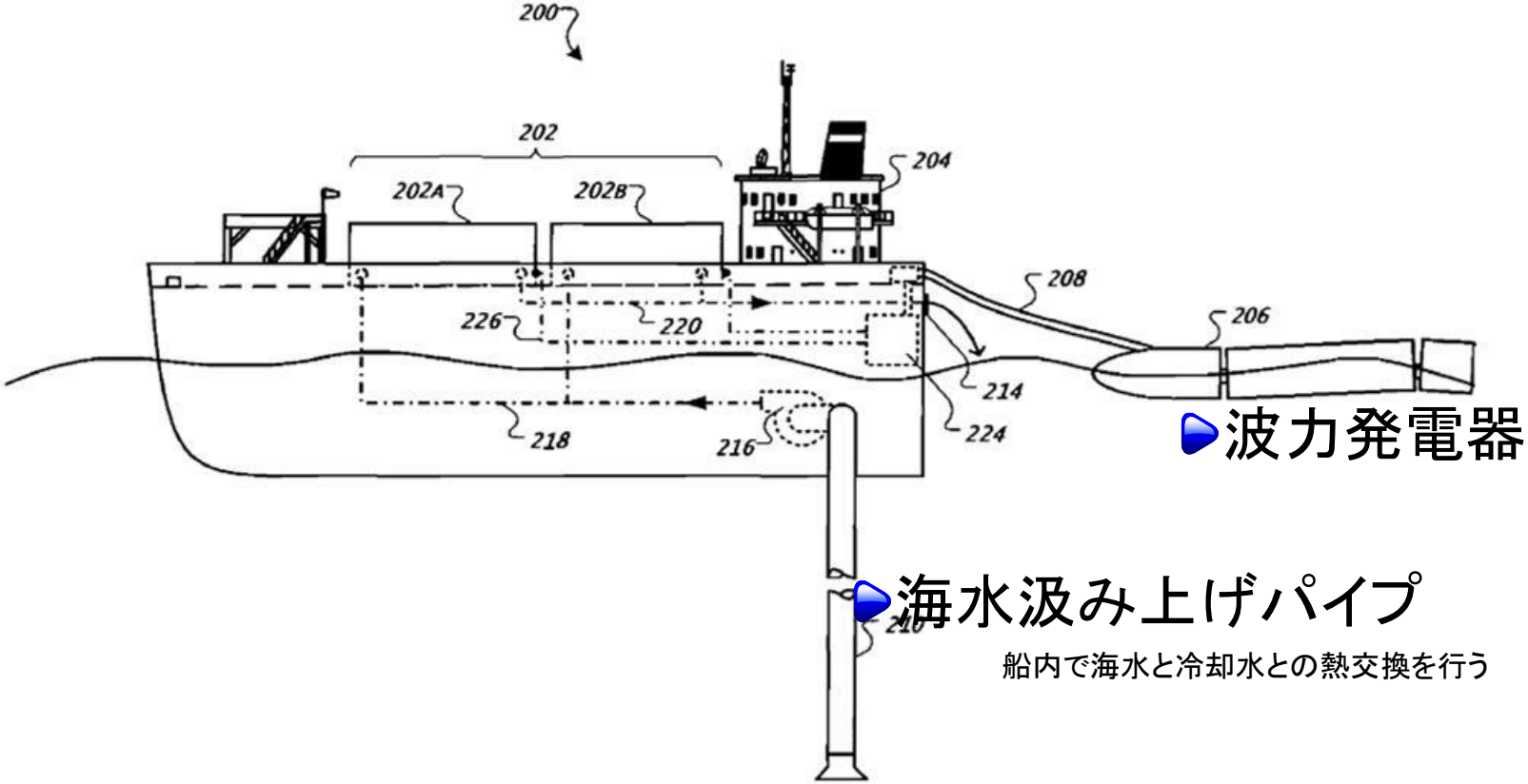
2マイル沖、水深1000mの深海から、5℃以下の海水を汲み上げ、データセンターの冷却水との熱交換を行う。

洋上のデータセンター（Google）

- Googleが2007年2月に特許を出願。2009年4月には特許が成立した。

土地不要で固定資産税がかからない。データセンターの消費電力は、波力発電で賄う。

海水をくみ上げて冷却に使用する。

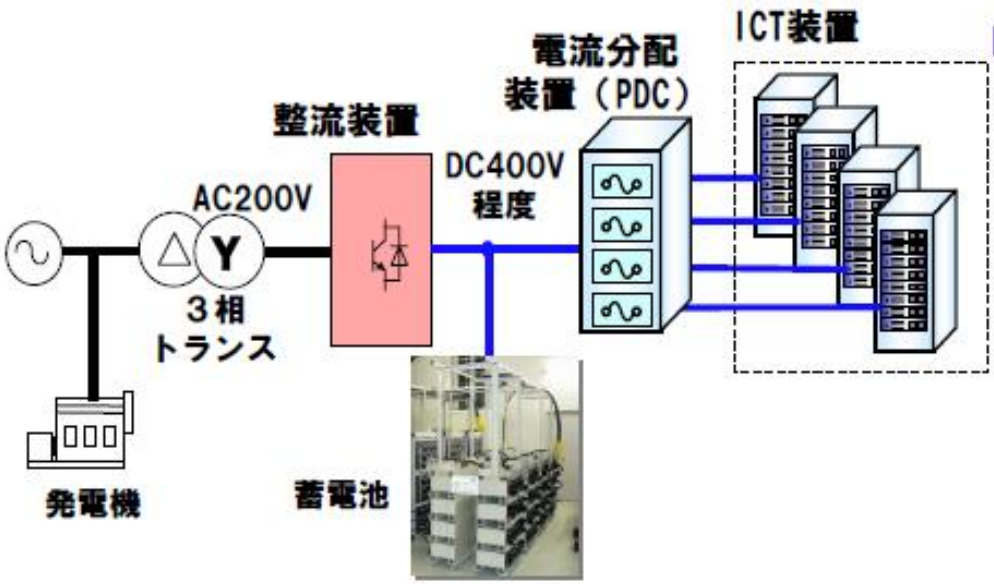


高電圧直流給電（NTTなど）

NTT

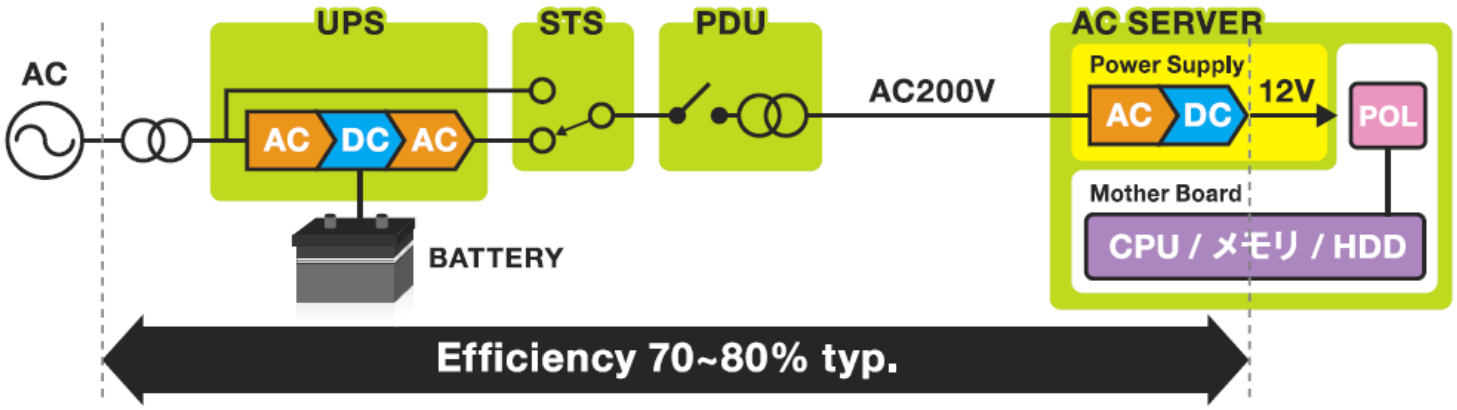
- 2008年より高効率高信頼な48V給電を高電圧化する高電圧直流給電の開発を実施。
- 2009年にNTTデータ三鷹ビルで実証試験を実施：グリーンITアワード 2009 経済産業大臣賞受賞

高電圧直流給電化で、AC UPSと比較して、**最大18％の省エネ**を達成。



A Typical AC(UPS) method

一般的な
交流(UPS)方式

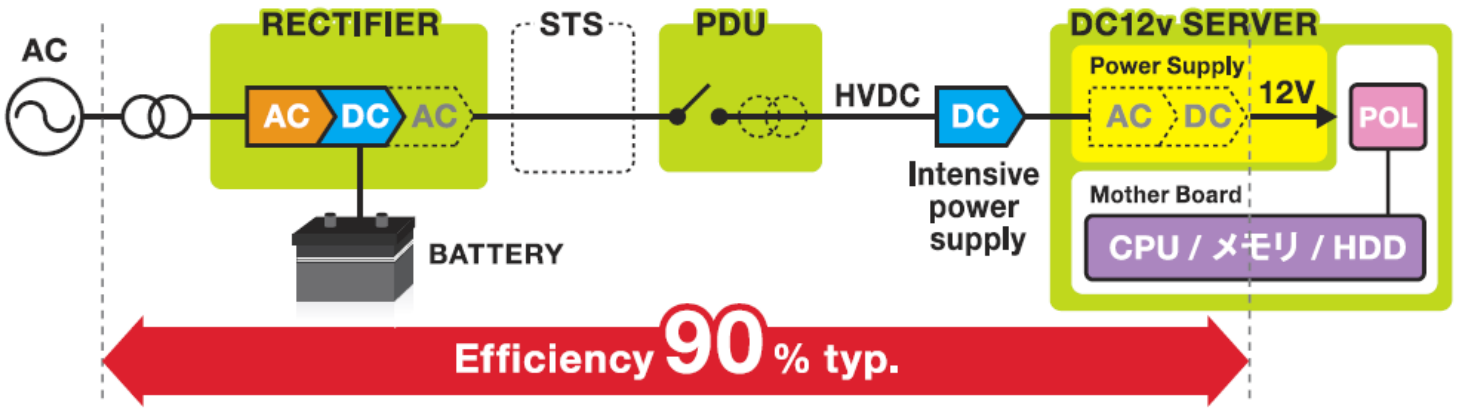


Approximately 10-20% electric energy reduction

約10～20%の電力量削減

HVDC

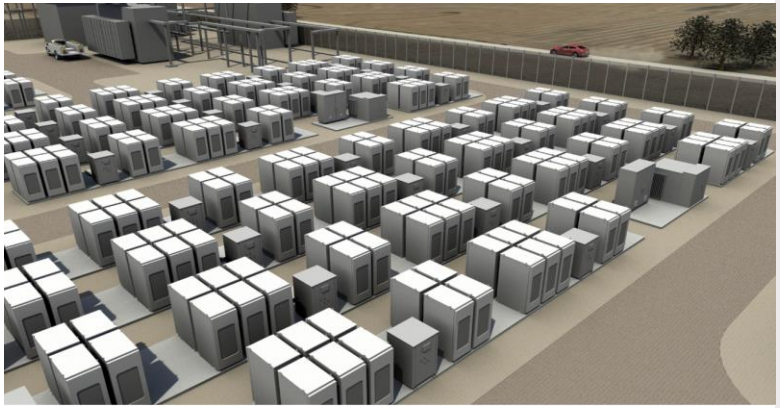
SERVER RACK SYSTEM
DC12V TYPE



さくらインターネット
高電圧直流給電＋太陽光発電

コンテナ型データセンター

Amzon Web Service向け　Tesla リチウムイオン蓄電池センター　4.8MWh



Above: Tesla's utility storage.

Image Credit: Tesla

テスラ家庭向け蓄電装置
2.5KWh10KWh
来年約40万円で発売予定

出所：Tesla Energy　http://www.teslamotors.com/presskit/teslaenergy
出所：　http://www.gizmodo.jp/2015/05/power_wallpower_packgigafactor.html

米国APPLEデータセンタは、再生可能エネルギーで稼働

メイデン、ノースカロライナ州のデータセンターの隣にあるアップル社の燃料電池ファーム

メイデン、ノースカロライナ州のデータセンターの隣にあるアップル社の太陽農場

燃料電池と太陽光発電(計２０MW)

2012年以降、すべてのAppleのデータセンターに供給される電力は、再生可能エネルギー資源で100パーセントまかなわれています。

ノースカロライナ州のデータセンターから15マイル程度、Appleの第二の太陽農場

出所：***DataCenterKnowledge***

APPLEと環境
http://www.apple.com/jp/environment/renewable-energy/

Google

デジタル化による炭素排出量大幅削減

- 企業や自治体がオンプレミスからGoogle Appsに移行することで最大85%の省エネ効果を達成できると、グーグルが試算している。
- グーグルは2021年6月27日、企業や自治体がオンプレミスからGoogle Appsに移行することで最大85%の省エネ効果を達成できるとする試算結果をブログで紹介した。「クラウドサービスの利用で作業効率の向上に加え、エネルギー消費量や二酸化炭素の排出量の減少とコスト削減につながる」としている。
- 試算によれば、企業や自治体がメールシステムをオンプレミスからGmailに移行することで、最大80倍のエネルギー効率が得られるという。さらにオフィスアプリケーションを含めてオンプレミスからGoogle Appsに移行すれば、省エネ効果は65～85%に達するとした。
- 同社によると、1万7000人の「Google Apps for Government」ユーザーがいる米国一般調達局（GSA）ではオンプレミスからの移行で二酸化炭素排出量が85%削減され、年間想定では28万5000ドルのコスト削減効果につながったとしている。

ITmedia エンタープライズ

日経XTECH
日経クロステック

Google

2021年4月23日

- 米グーグルは同社における「脱炭素」の進捗状況を20214月20日（米国時間）に発表した。2030年までにデータセンター（DC）やオフィスなどを二酸化炭素（CO2）を排出しないカーボン・フリー・エネルギーで24時間365日運営する目標を掲げており、**既に5カ所のDCでその目標をほぼ達成**したとする。
- 進捗状況はスンダー・ピチャイ最高経営責任者（CEO）名義のブログで発表した。グーグルは脱炭素の取り組みを3つのステップに分類している。

① 第1ステップは自社が排出するCO2に相当する**カーボンオフセット（CO2排出権）を購入**する「カーボンニュートラル」で、グーグルは07年になし遂げた。

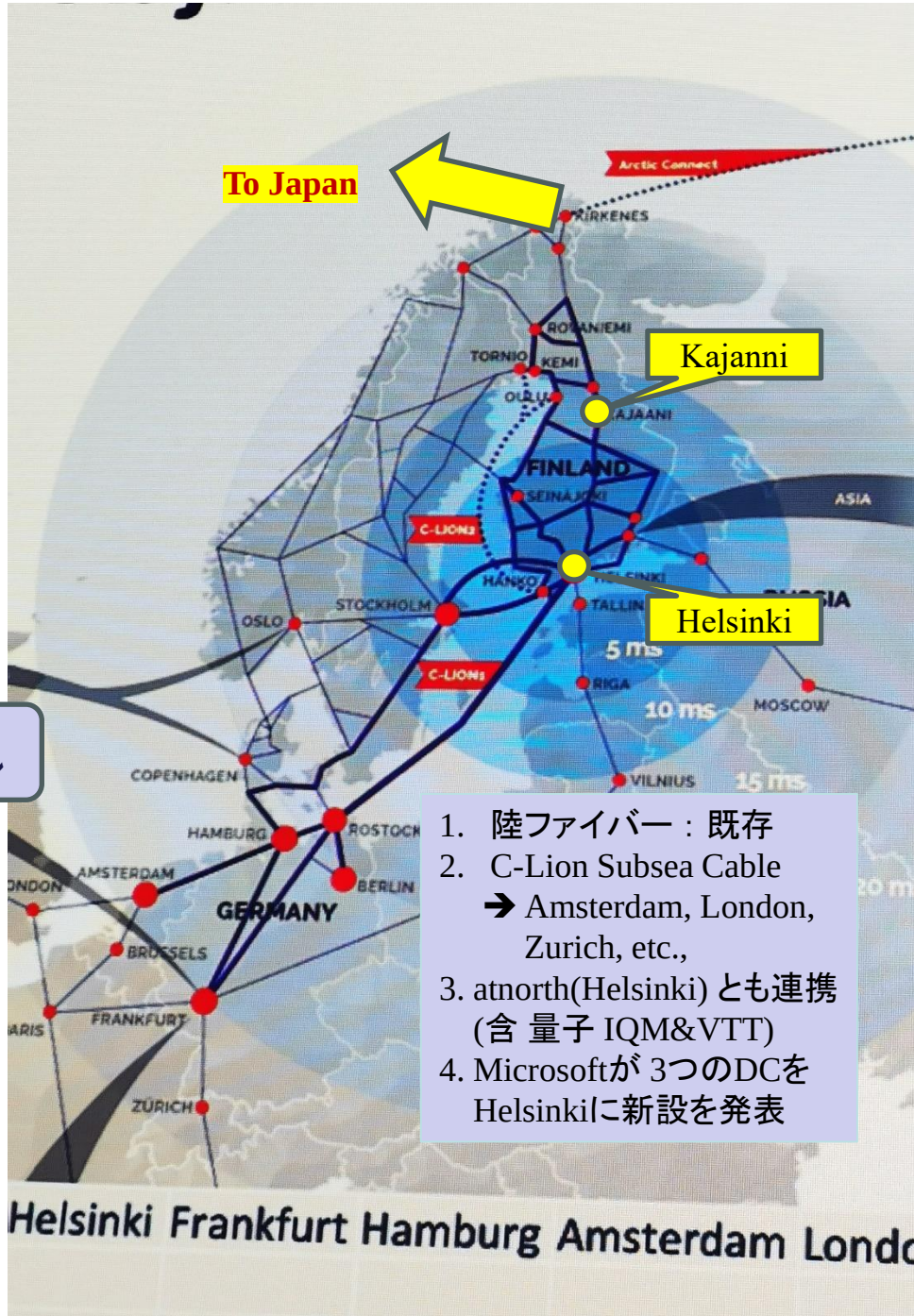
② 第2ステップは同社の年間電力使用量に相当する**再生可能エネルギーを購入**する「100%リニューアブル（再生可能エネルギー）」で17年に達成した。

【クラウド型データセンターが都会から疎開可能になった!!】

➔ DataCenter-Go-to-RenewalEnergyPowerPlant(REPP)

③ そして第3ステップとしてDCやオフィスが**消費する電力を常時クリーンエネルギーでまかなう**「24/7 カーボンフリー（24時間365日脱炭素）」を30年までに達成する。

1. 100% 水冷
2. 非常用自家発電機は無し
   (データバクアップは有り)
3. 場所・電力の拡張性 有り



ICT Solutions for Brilliant Minds

CSC

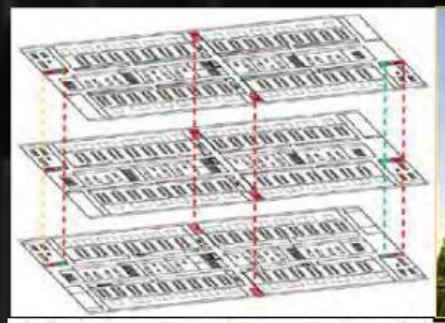
visited September 20(Wed), 2023

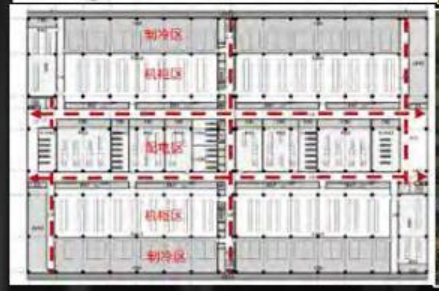
- 100%水冷 (🡸 HPCに特化)
- 簡易な施設構成(UPC,自家発)
- 電気代: **1/5 ～ 1/7** !!!
  - ✓ 30 € MWh (5-7円/kWh) from PPA
  - ✓ 日本 35円/kWh
  - ✓ Carbon **“Negative”** も発生
- **製紙工場の跡地を再利用**
  - ✓ 工場の躯体をそのまま利用
  - ✓ 250MWの(既存)電力引き込み
  - ✓ **100% 再生可能エネルギー電力**
- 地域冷暖房システムへの 熱供給
  - ✓ 夏季も販売可能(病院や宿泊施設等)
  - ✓ 大きな **税制優遇(22 € ➔ 0.6 € /MWh)**
  - ✓ 拡大予定

する

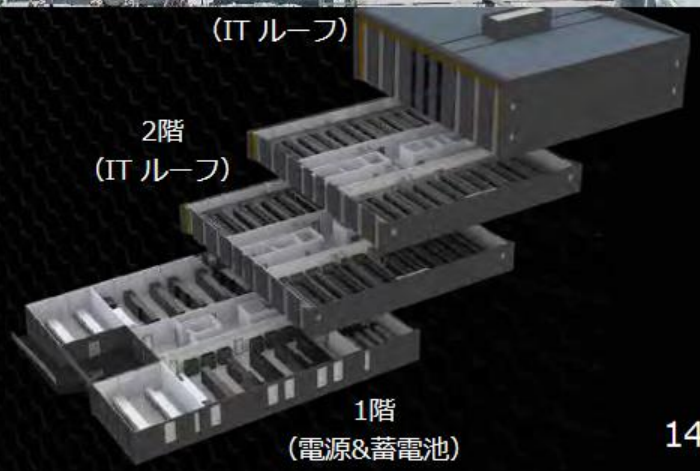
多層プレハブDC

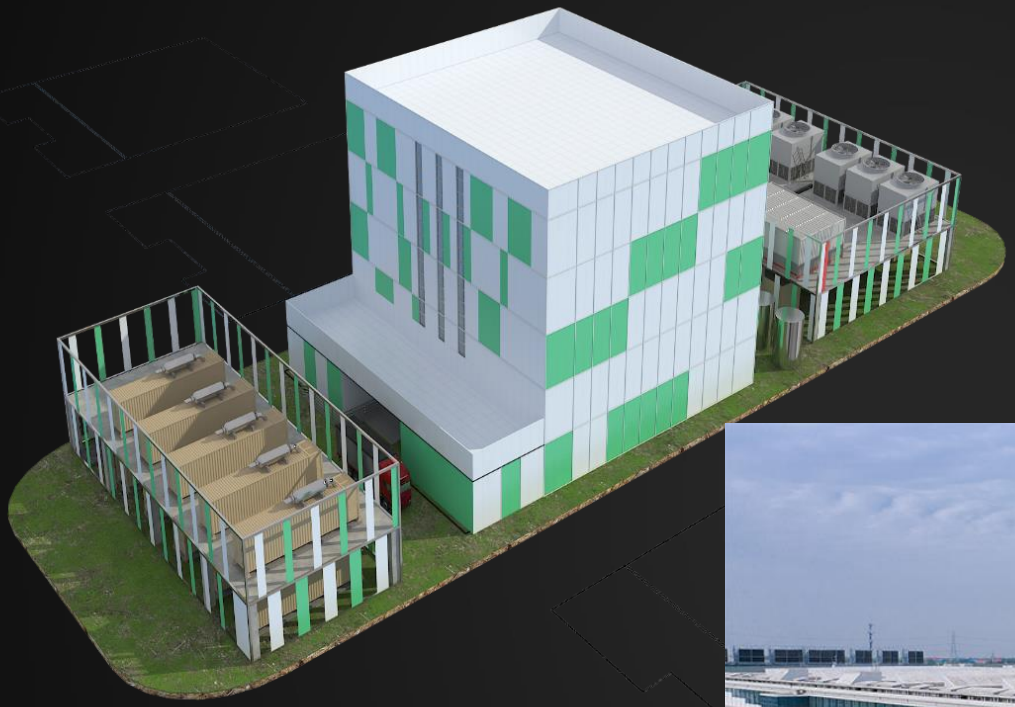
- アーキテクチャの標準化
  - 電気：N+Rデザイン
  - 空調：間接蒸発AHU+Fan Wall
- レイアウトと建物の標準化（モジュール、プレハブ）

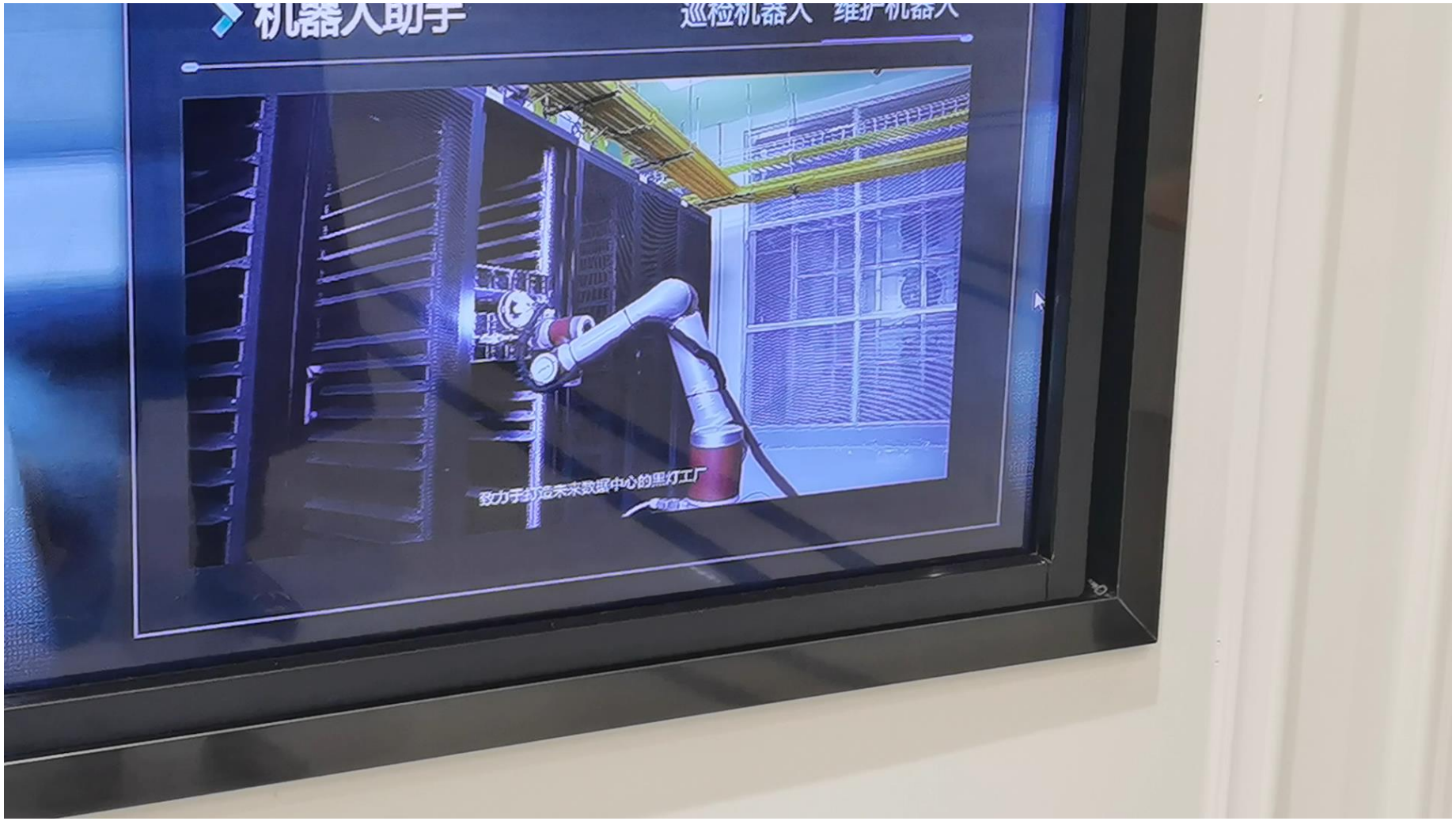




（IT ルーフ）






机器人助手
巡检机器人
维护机器人

過去&現在の Google の戦略

1. Follow the Moon
   - ✓ 夜は温度が下がるね
2. Go to North
   - ✓ 北は 温度が低いのよね
3. Follow the Sun with stable Wind, Water
   - ✓ 再生可能エネルギーの場所に移動しよう！
4. Generate, terminate and migrate over the globe
   - ✓ 電力需給 に 協力いたしましょうか?

すべては、仮想化技術の おかげ！

ビジネスケース

ドイツの会社（BMW)

	以前	現在		
	ドイツ	ドイツ	アイスランド	スウェーデン
	100% of IT Load Tier III	20% of IT Load Tier III	50% of IT Load New standard	30% of IT Load New standard
IT負荷	5 MW	1MW	2.5 MW	1.5 MW
価格	200万ユーロ	40 万€	50万€	20万€
総額	200万ユーロ		110万	

www.etixeverywhere.com

大事で速い仕事は近い場所の{dark-sideの}DC

Non-Criticalな仕事は、遠いけども、安い(Cheep) Green＆CleanなDC

Transfer DCs to energy clean sites (Iceland & Sweden)

✓ 100% Renewal Energy (Hydro & Geothermal)

1. Gentle & contribute to Earth
   - ✓ by use of renewal energy (RE-100)
   - ✓ by use of cool air (EP-100)

2. Cost 👇 & Productivity 👆 (EP-x000)

   Replacing power cable (copper)

   to

   tele-communication cable (glass)

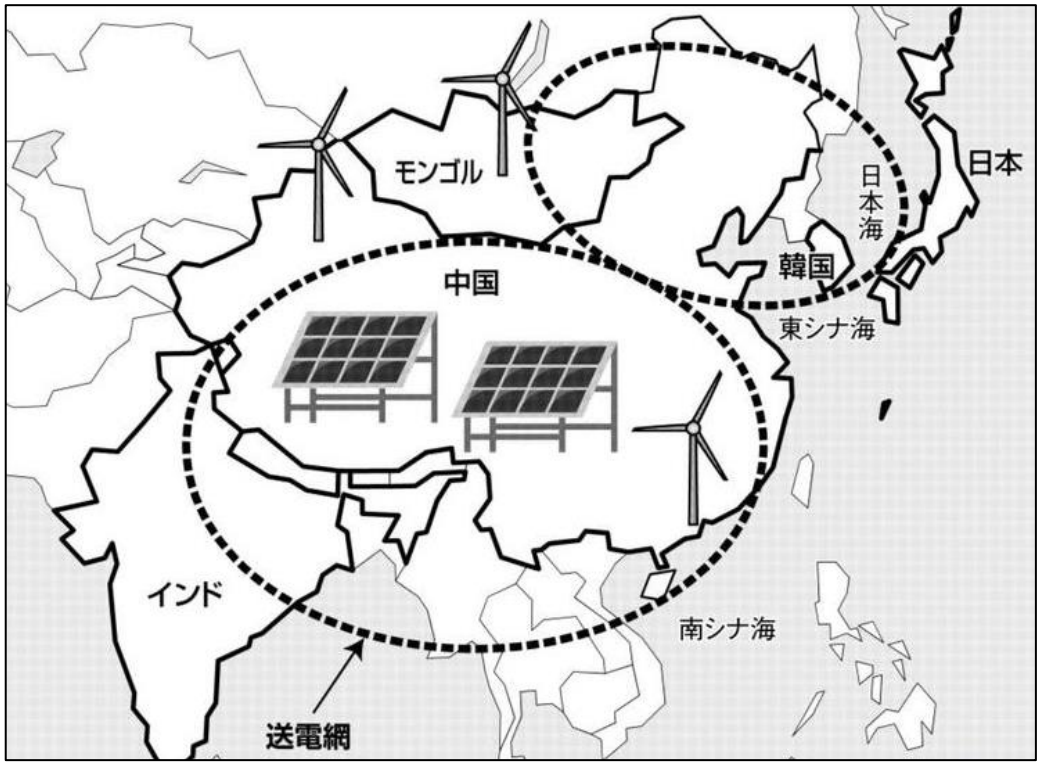
移動に必要なエネルギー=Energy Productivity

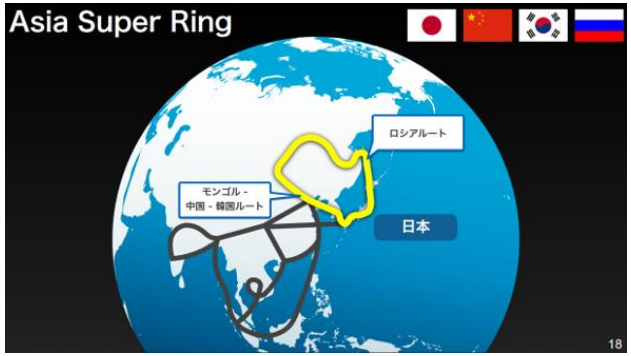
荷物	電力	通信
Material	Electricity (Power)	Digital bits (digital Function)
x0000	x00	1

	質量(g)	質量比
陽子/中性子	1.673×10^{-24}	1,840 ($=1.8 \times 10^{-3}$)
電子	9.109×10^{-28}	1
光子	$0 < m < 2 \times 10^{-51}$	$> 2.198 \times 10^{-24}$

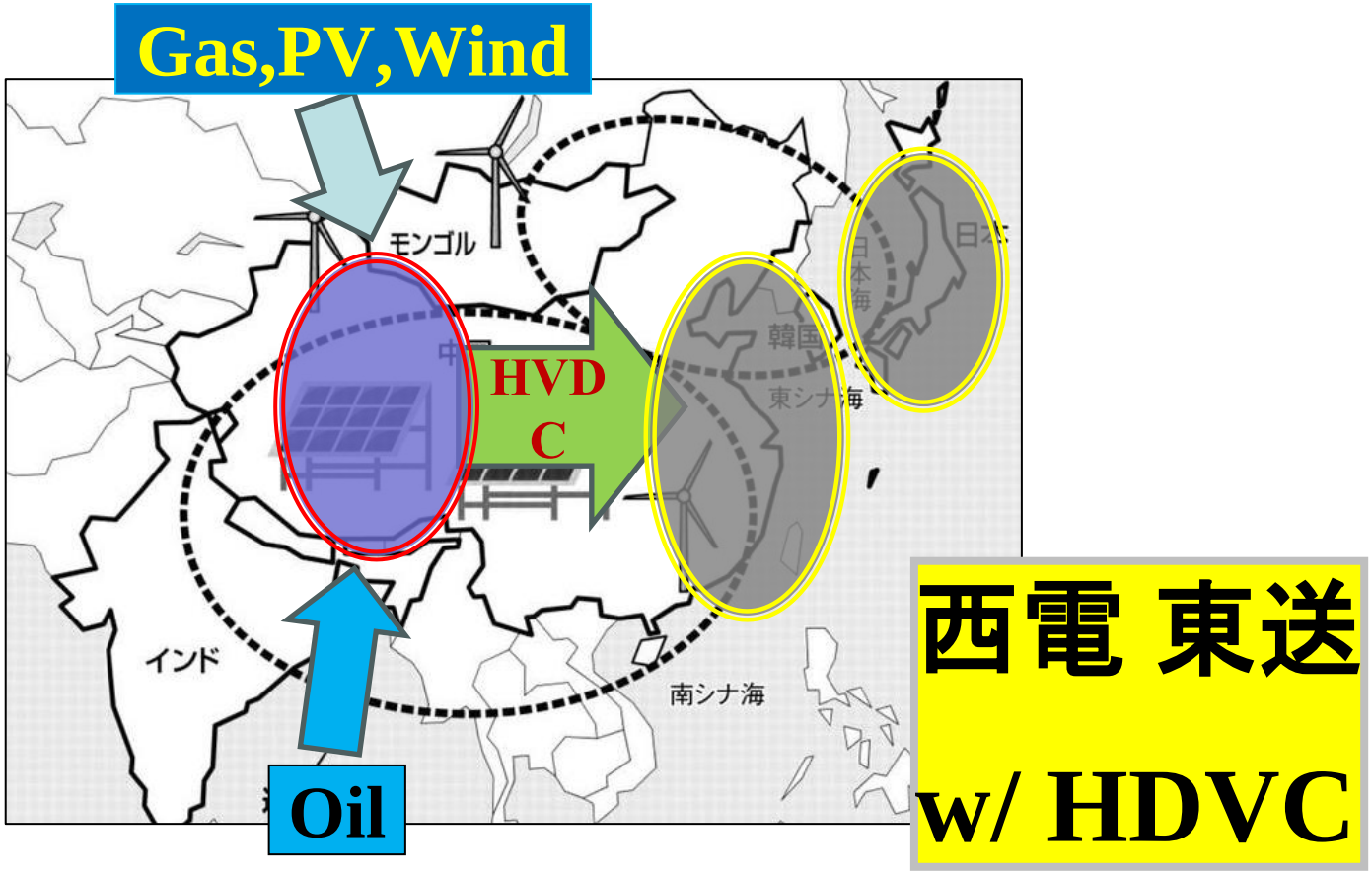
	速度 (Km/h)	速度 (m/sec)	速度比
モノ	36	10	1
電流	30万($=3 \times 10^{5}$)	3×10^{8}	3,000万
(電子)	0.0075 mm/sec	7.5×10^{-5}	1/750万
電磁波(光子)	30万($=3 \times 10^{5}$)	3×10^{8}	3,000万
(ガラス中)	20万($=2 \times 10^{5}$)	2×10^{8}	2,000万

Asian Super Power Grid Plan (2017)





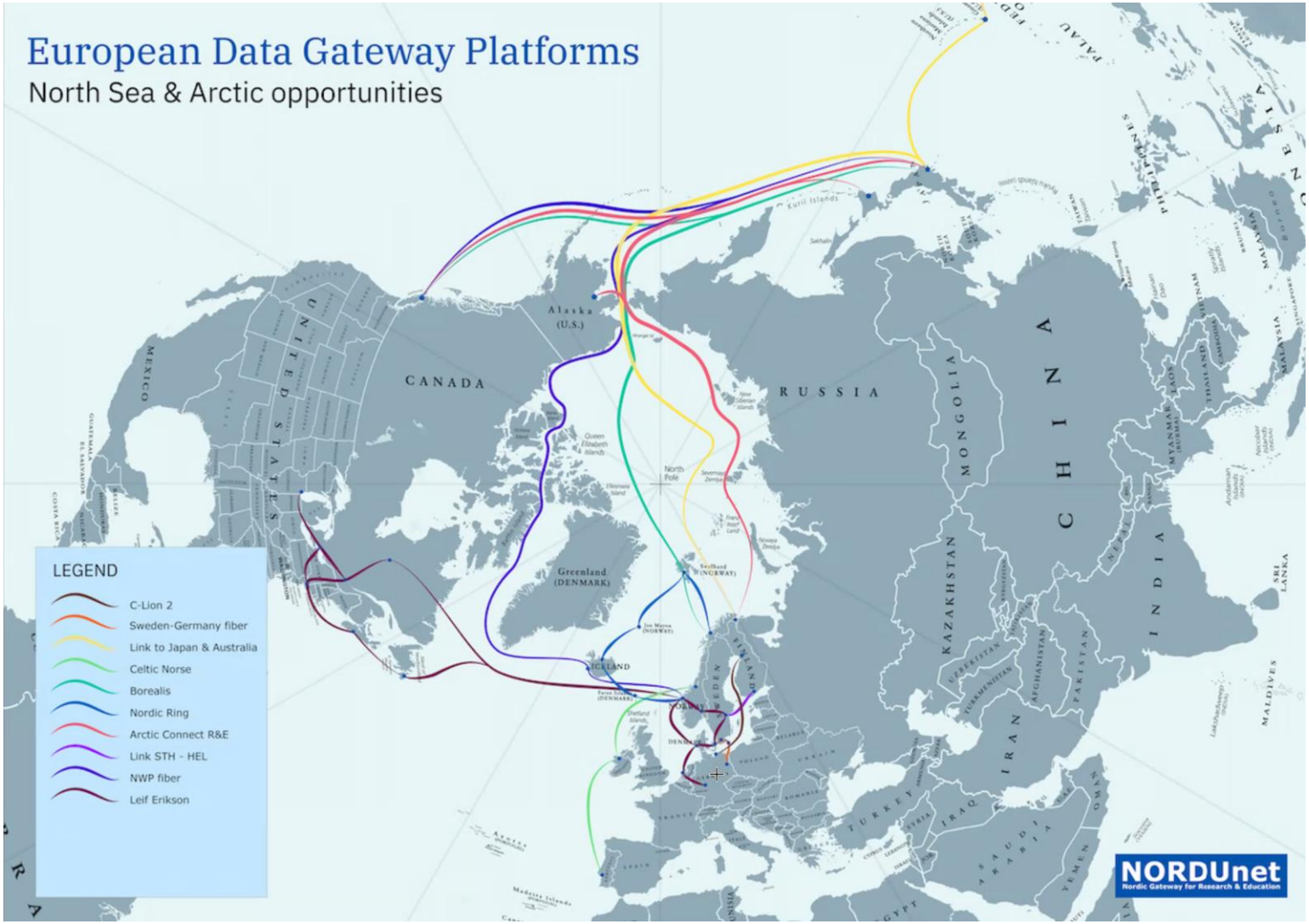
Asian Super Power Grid Plan (2017)



東数 西算(2021 第13期全人代)

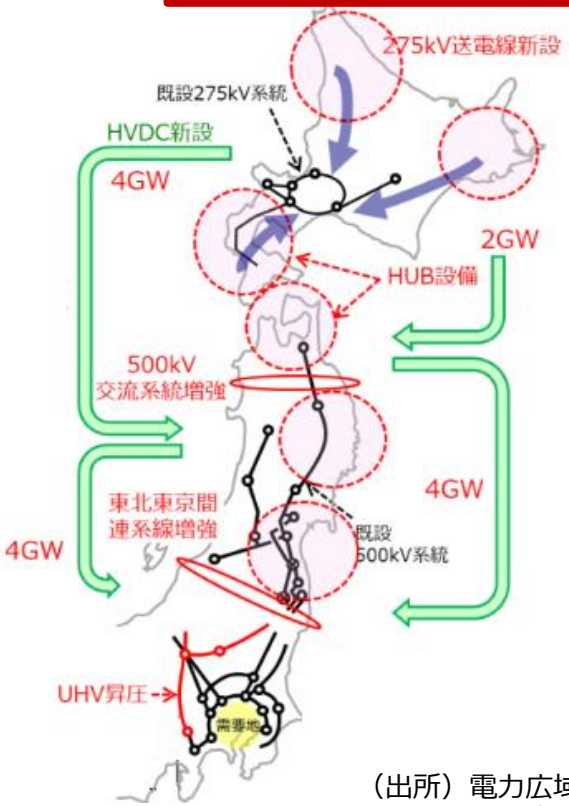
European Data Gateway Platforms

North Sea & Arctic opportunities

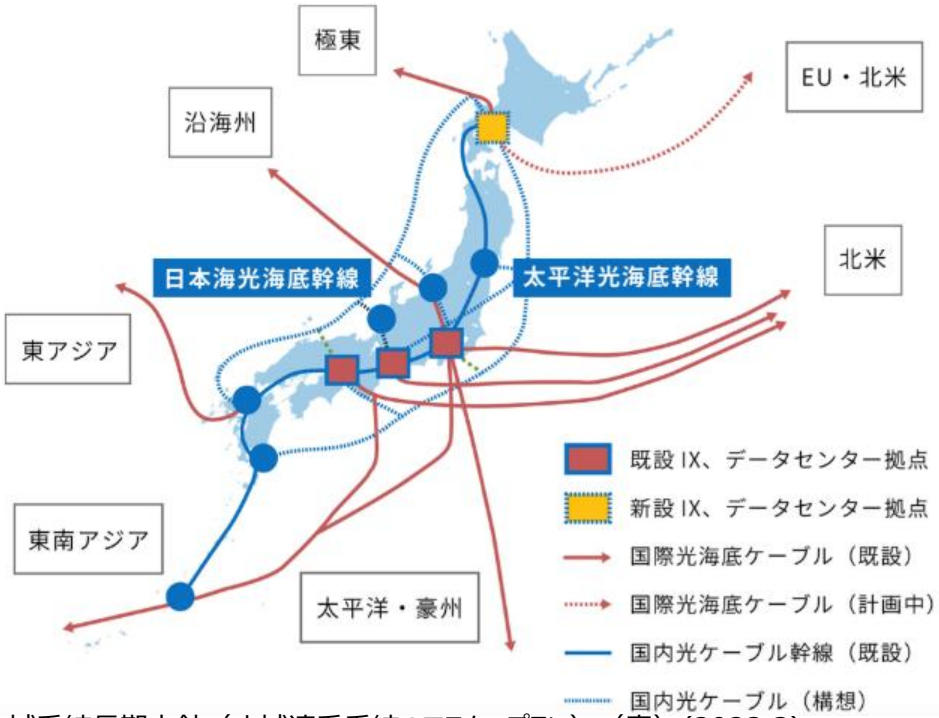


As a デジタル田園都市国家構想(Digital Garden City Initiative)

Electric Power



Information



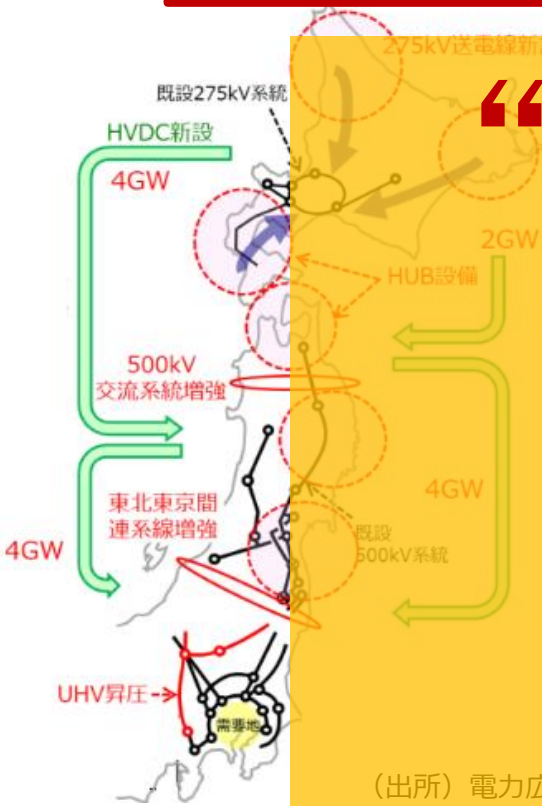
（出所）電力広域的運営推進機関：広域系統長期方針（広域連系系統のマスタープラン）（案）(2023.3)
北海道ニュートピアデータセンター研究会提言書：「北海道をデータセンターのパラダイスに」(2022.4)

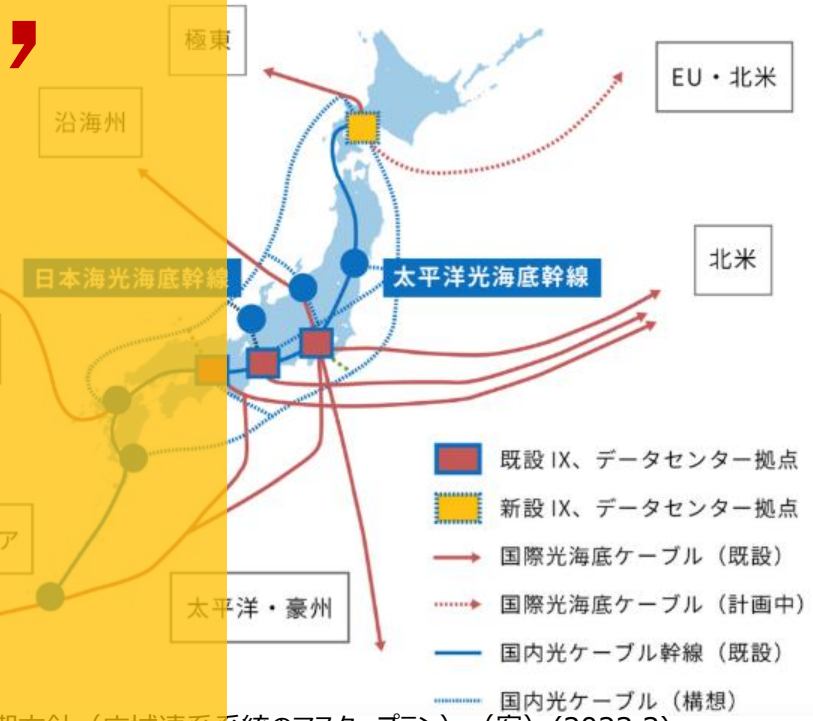
As a デジタル田園都市国家構想(Digital Garden City Initiative)

Electric Power

Information

“Cost”

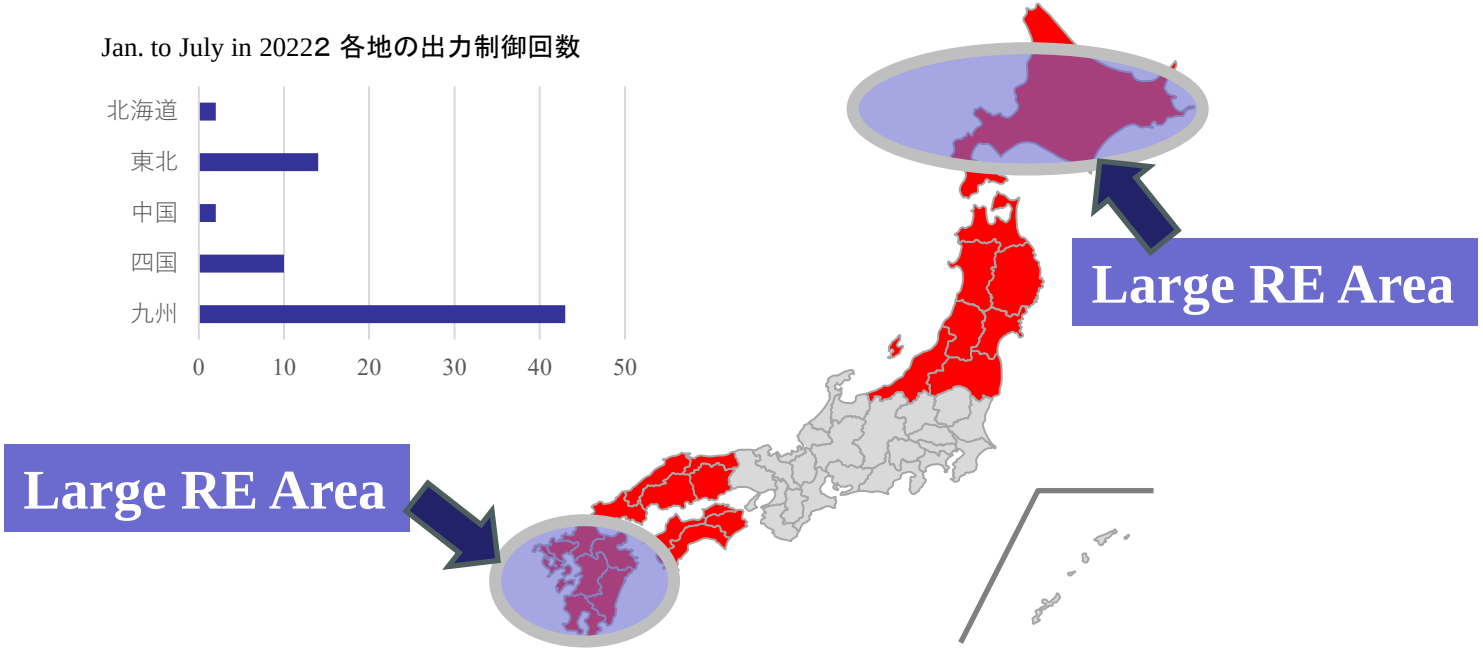




（出所）電力広域的運営推進機関：広域系統長期方針（広域連系系統のマスタープラン）（案）(2023.3)
北海道ニュートピアデータセンター研究会提言書：「北海道をデータセンターのパラダイスに」(2022.4)

Electrical Power shortage in spring and winter in Japan

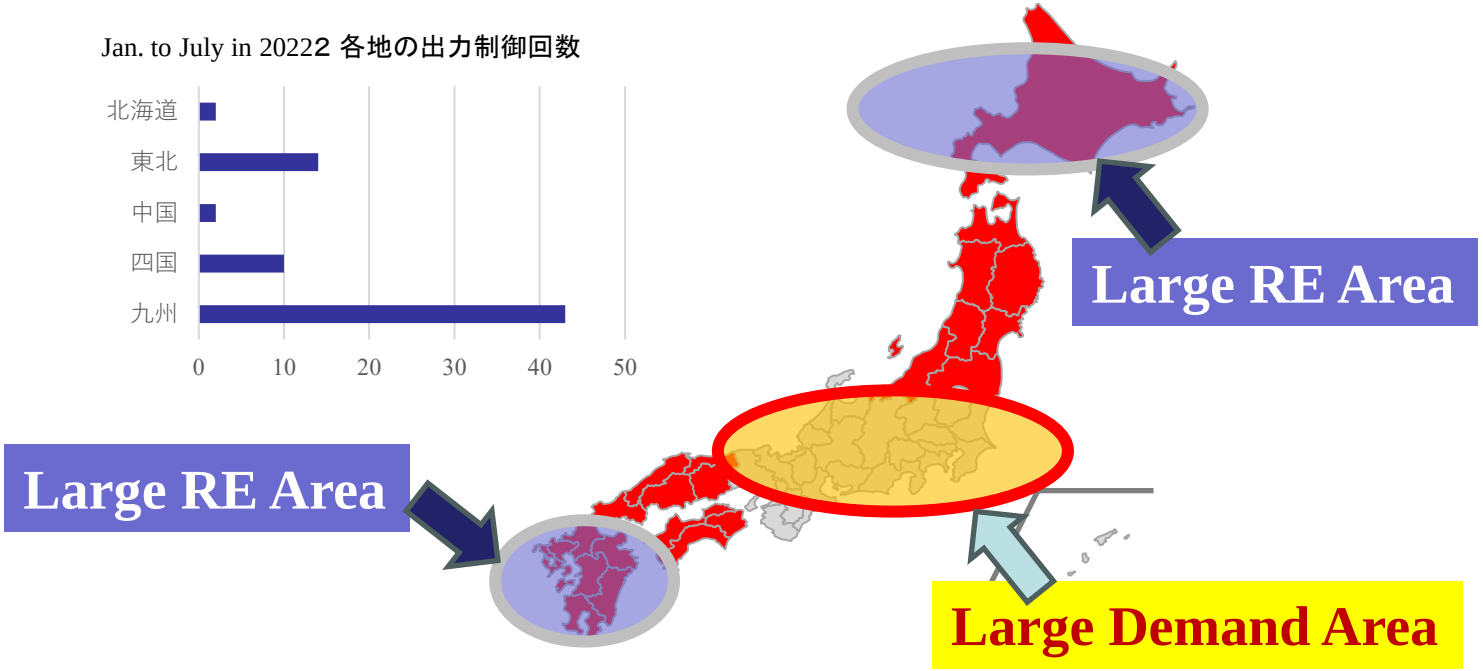
- ２０２２年再エネ出力制御は四国・中国・東北・北海道電力管内に拡大



【資料】 ビットメディア 代表取締役社長 高野 雅晴 氏
第４回イベント（2022年9月28日）
「データセンターの地域エネルギー貢献」

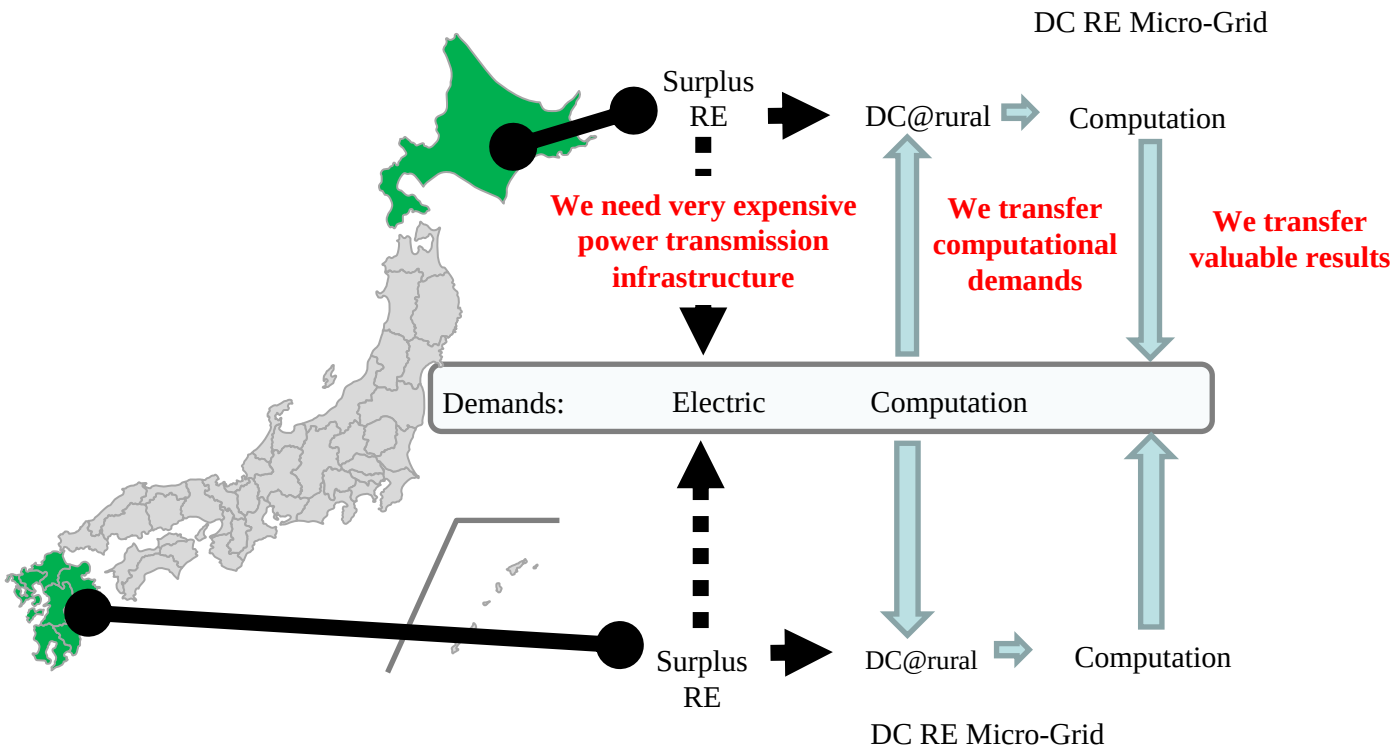
Electrical Power shortage in spring and winter in Japan

- ２０２２年再エネ出力制御は四国・中国・東北・北海道電力管内に拡大



【資料】 ビットメディア 代表取締役社長 高野 雅晴 氏
第４回イベント（2022年9月28日）
「データセンターの地域エネルギー貢献」

Data Center in rural area can help !!!

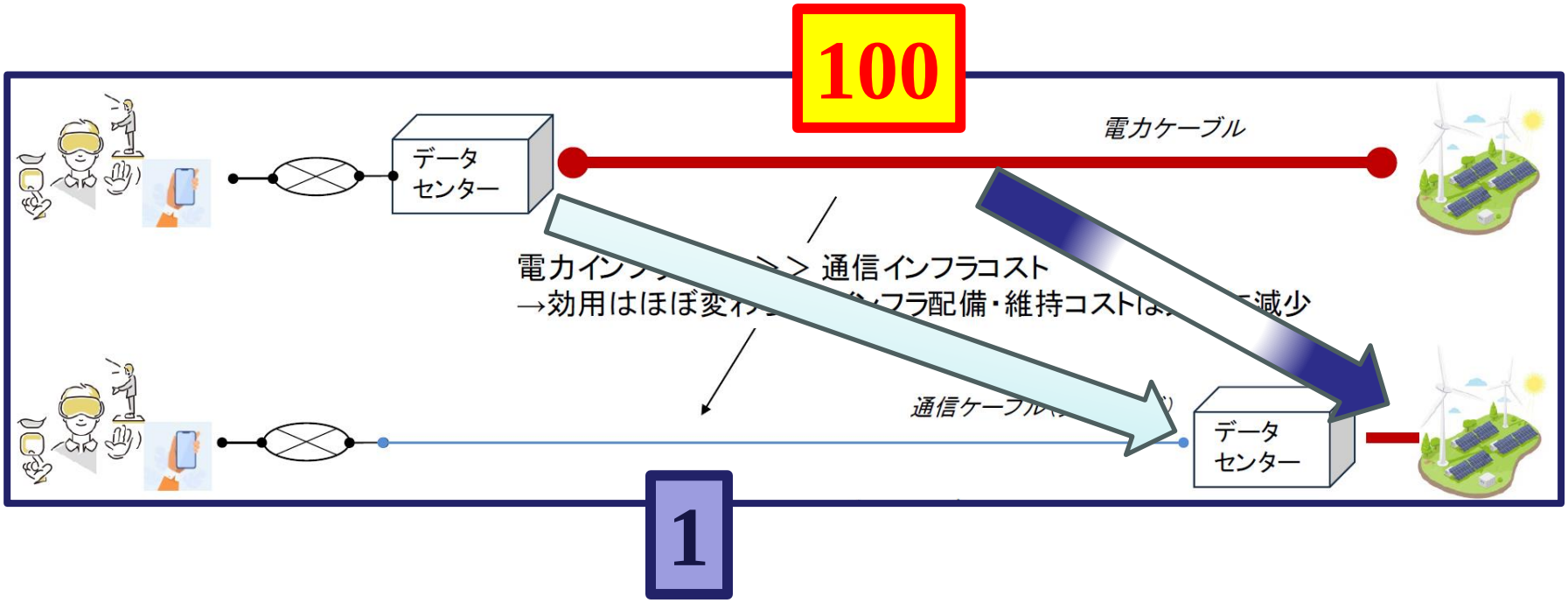


【資料】 ビットメディア 代表取締役社長 高野 雅晴 氏
第４回イベント（2022年9月28日）
「データセンターの地域エネルギー貢献」

https://highreso.jp/gpu/consulting

人工知能・ビッグデータ に 特化 すると、、、違う解も？

1. HCI/GPU 専用サービス
   - ➔ 緩い PKIs
     - ① 遅延特性
     - ② サービス停止頻度・時間
2. {北陸}電力会社との連携
   - ➔DR: Demand Response
     - ① ΔkW (＝緊急対応)
     - ② kWH (＝定常計画対応)


100
データ
センター
電力ケーブル
通信インフラコスト
データ
センター
1

Energy x Digital

EV as new component for new markets

【Operating rate】

- Private : 4～5%
- Commercial : 24%

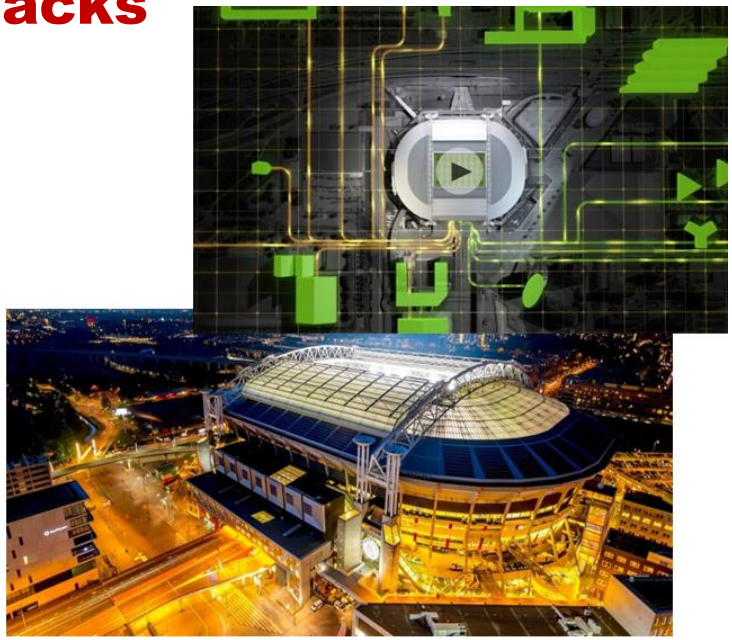
1. Battery
   a. Emergency
   b. DR(Demand Response)
2. Computation

Reuse of used EV battery at Amsterdam stadium

- **{Electric} Power distribution base**

  Capacity of EV's used battery packs

  - ✓ **3MWh only with 148 EVs**
  - ✓ **100KW in peak per EV**
  - ✓ **20KW in average per EV**



1. **Normal situation;**

   Power DR (Demand Response)

2. **Emergency case;**

   Local self power supply

体積当たりのエネルギー密度(

キャ
最大

400V & 12V DC System

100 KW in Peak
30 KW for continuous
↓
30 Houses
by 50cm x 50 cm x 50cm

Nuclear	Legacy	300～500 MW
	Now	1.0 ～ 1.5 GW
Hydrogen (Kurobe)		335 MW
Hydrogen		500MW ～ 1.0 GW

Nissan Leaf	Ave.	225 MW
	Peak	750 MW

HONDA FCV
30 kW (Ave.)
100 kW (Peak)

Macro

Parking Lots

Nissan Leaf： Ave. 30 KWh / car

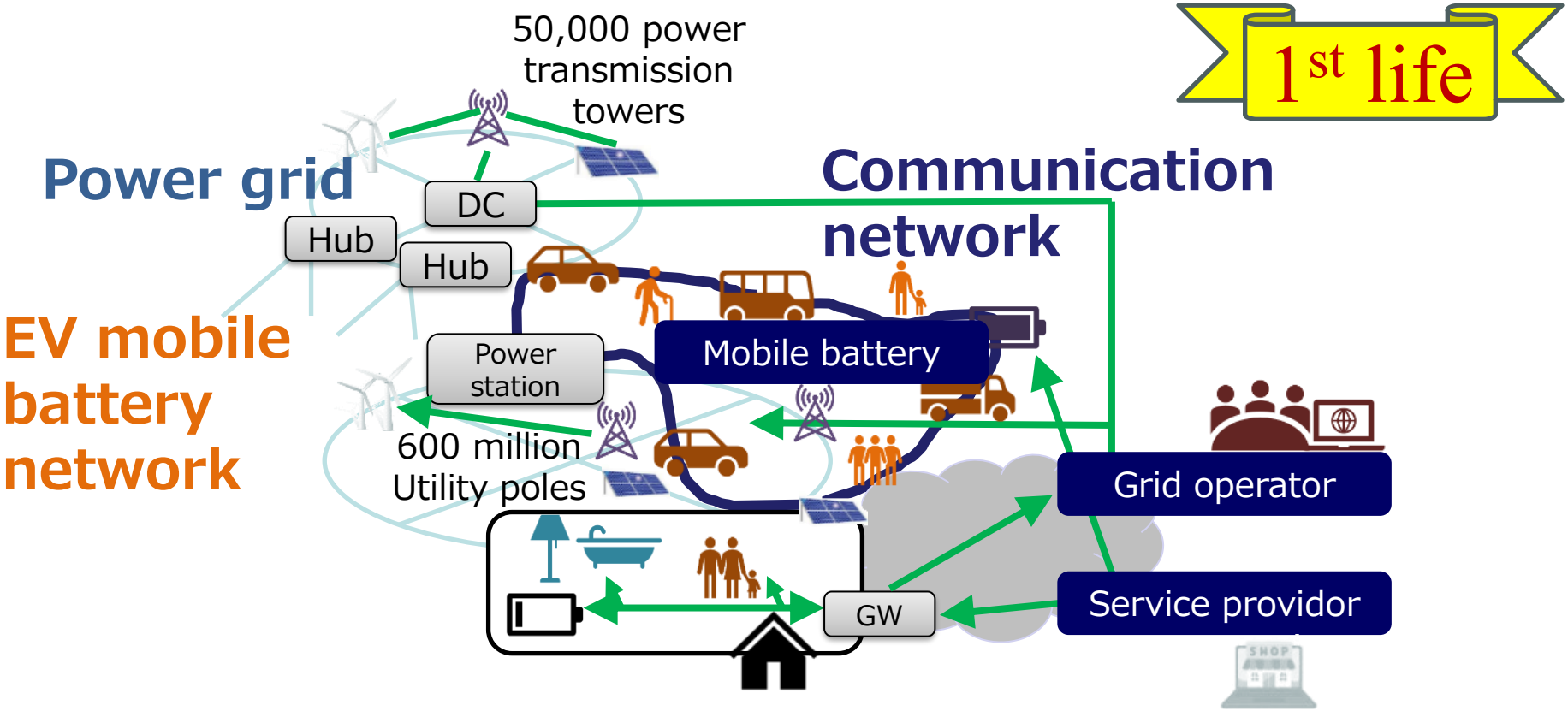
Total No. of Leafs： 350,000(Worldwide)

75,000 (in Japan)

(*) as of Sept.2016.

Nissan Leaf (100 cars)	Ave.	3 MW
	Peak	10 MW

①Battery x ②Computation x ③Mobility



（出所）岡本(2021)「GRIDで理解する電力システム」、電気新聞出版

①Battery x ②Computation x ③Mobility

1st life

◆ Nissan EV：30kW(Ave.), {100kW(Peak)}, 100kWh

1. 10^6(1 millions)：30GW(Ave.), 100 GWh(Peak)
   - Pumped storage power in TEPCO (=9GWh) **x 10**
   - Kurobe Dam@Japan (=335MW) **x 10**
2. 100 EVs ： 3MW(Ave.), {30MW(Peak}
   - Eng.No.2 Bldg (12F)： 700kW (≒35 EVs)
   - Hongo Campus(No.2 in Tokyo)：30MW (≒1 k EVs)

（出所）岡本(2021)「GRIDで理解する電力システム」、電気新聞出版

①Battery x ②Computation x ③Mobility

50,000 power

1st life

◆ Tesla(by AMD@CES2023) ： **10 TFLOPS (10^{13})**

https://car.watch.impress.co.jp/docs/news/1469629.html

◆ 富岳(No.3 Super Computer) ： **1,000 PFLOPS (10^{18})**

- **3 TFLOPS/node**
- 158,976 nodes (1.5×10^5)

https://www.fujitsu.com/jp/about/businesspolicy/tech/fugaku/specifications/

👇

100k($=10^5$) Tesla ≒ 富岳(No.3 Super Computer)

（出所）岡本(2021)「GRIDで理解する電力システム」、電気新聞出版

金銭コスト
温室効果ガス排出量

鉄道、道路 vs（列車）（自動車）	空港、港湾 vs（飛行機）（船舶）	{不要}
光ファイバ網 vs	携帯無線網 vs	衛星&WiFi網
同期型送配電網 vs（日本型）（EV電力網）	地域電力網 vs（米国型）	移動型電力源
【面】	【{固定}点】	【{移動}点】

敷設・維持・変更コストの減少
＝地球温暖化ガス排出量の削減

移動能力の変化/進化

【コンピュータ(計算)】

1950-1990	1990～	2000～
メインフレーム/WS	ノートブック	VM(Virtual Machine)
地面への恒久的拘束	地面からのHWの解放	SWのHWからの解放

【UPS/バッテリー/蓄電】

1800～	2000～	2025～
自家発	設置	自走
化石燃料との バンドル(拘束)	(化石燃料)発電設備 からの解放	地面への拘束 からの解放

High capacity mobile battery packages

High capacity mobile battery packages

Transfer cost … …
= Energy Productivity…

Electricity >> Digital bits

x00 : 1