

超高集積・超広帯域光ネットワーク

研究開発法人産業技術総合研究所
電子光技術研究部門／データフォトンクスプロジェクトユニット
並木 周

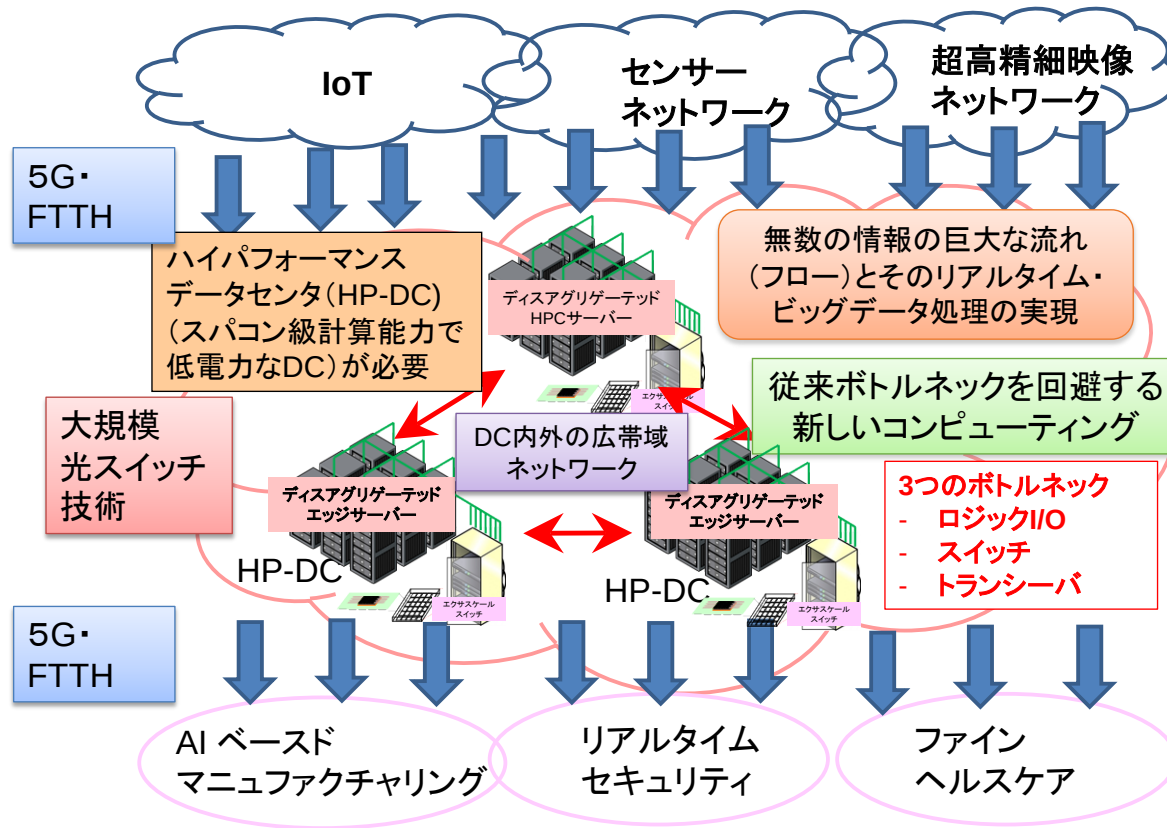
本日の講演内容

- ポストムーア時代に目指すべき光ネットワーク
 - データセンターの進展とディスアグリゲーション
 - エクサスケール光ネットワーク
- 基礎検討状況
 - 光スイッチ
 - トランシーバ
 - 波長セレクタ
 - ネットワーク制御方式
- 研究開発の方向性
 - シリコンフォトニクス・ファンドリについて
 - エコシステム構築に向けた取り組み

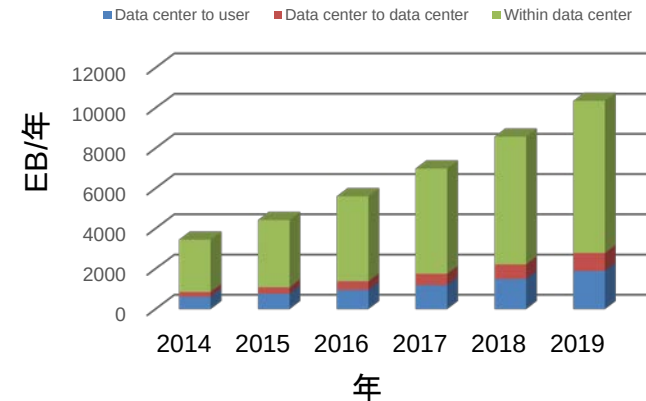
研究体制：データフォトニクスプロジェクトユニット

- 電子光技術研究部門と他部門を跨った垂直融合のための投影組織
- 電子光技術研究部門
 - 光ネットワーク技術G(井上崇GL、他6名)
 - 光伝送技術：設計、評価
 - 光ネットワーク技術：アーキテクチャ、実装
 - 光パスプロセッサG(池田和浩GL、他4名)
 - シリコンフォトニクス光スイッチ
 - シリコンフォトニクスG(山田浩治GL、他5名)
 - シリコンフォトニクス・トランシーバ技術
 - インタコネク・フォトニクスG(山本宗継GL、他4名)
 - オンボード・フォトニクス技術、集積型光コム発生
 - 3次元フォトニクスG(榊原陽一GL、他3名)
 - 3次元加工プロセス、機能デバイスへの応用
- 情報技術研究部門
 - サイバーフィジカルクラウド研究G(高野了成GL、他1名)
- 人工知能研究部門
 - 人工知能クラウド研究チーム(1名)

IoTビッグデータ:情報化から知識化社会へ



データセンターが扱う情報量

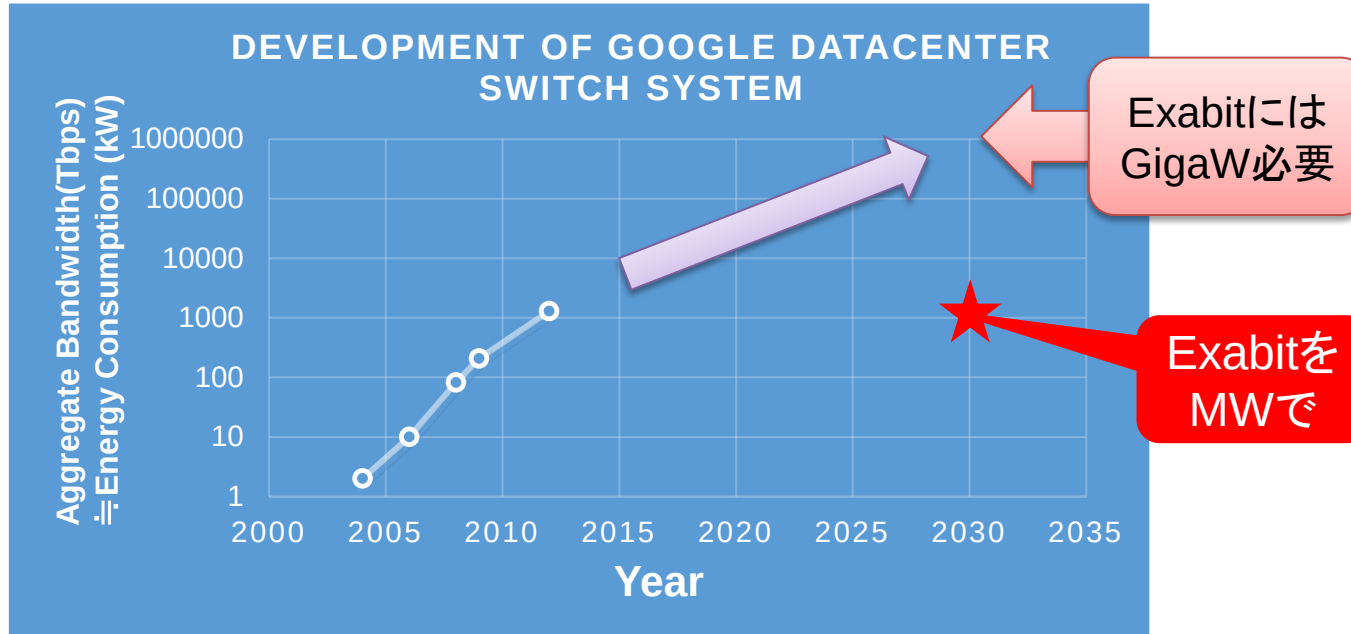


Source: Cisco Global Cloud Index, 2013–2018

データセンターを中心とした、情報通信インフラの再編が進む

従来スイッチ技術のさらなるスケールは困難

- Googleデータセンタースwitch技術の推移: 2030年には、エクサスケールへ
 - 既存技術の延長では、消費エネルギーは、ギガワット(GW)程度となってしまう
 - >>260MW: 全世界のGoogle DCs(10~14か所)

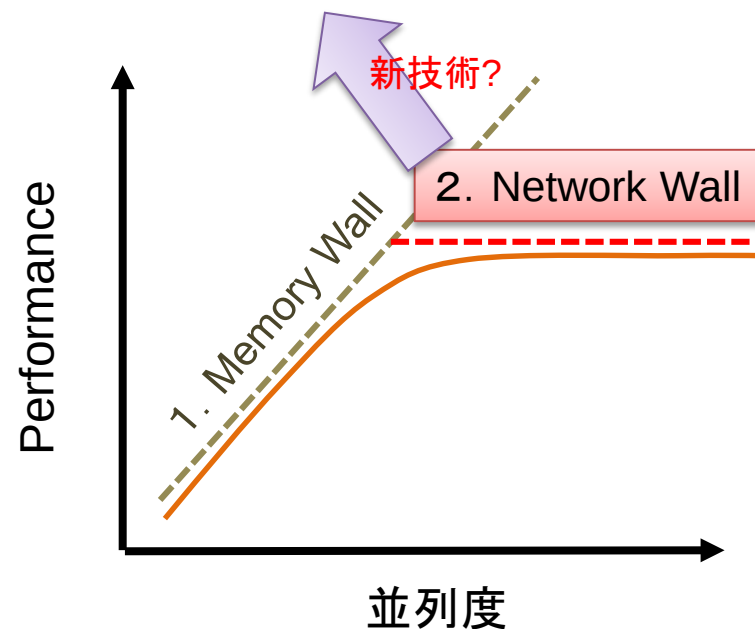
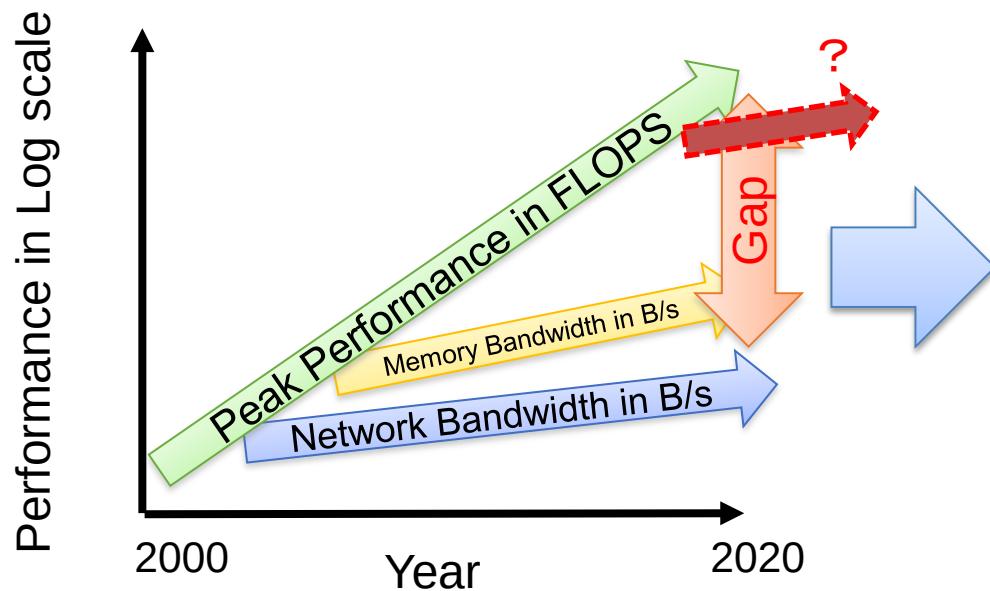


IP Router ~ CMOS
性能あたりの消費電力
▲10~12%/yr.

C. Lange, et. al., IEEE JSTQE, vol. 17, no.2, pp. 285-295, March/April, 2011

計算機のさらなるスケールが困難に

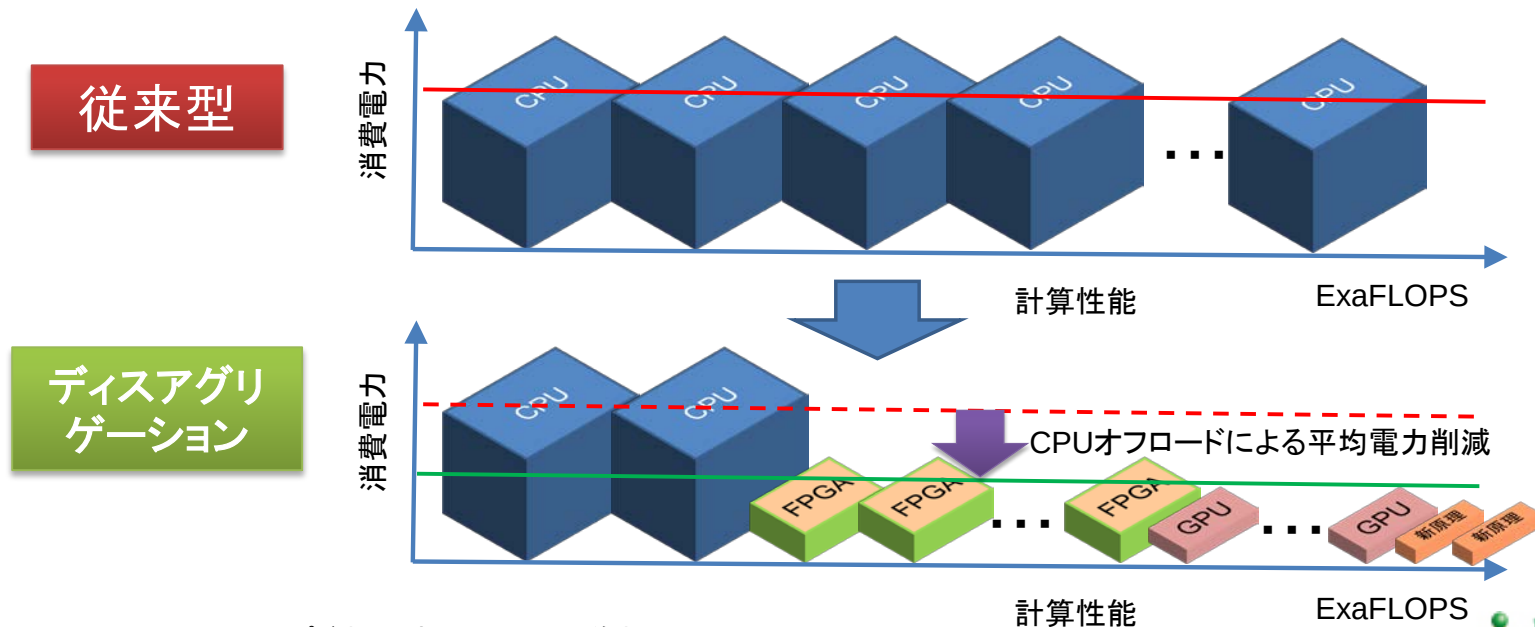
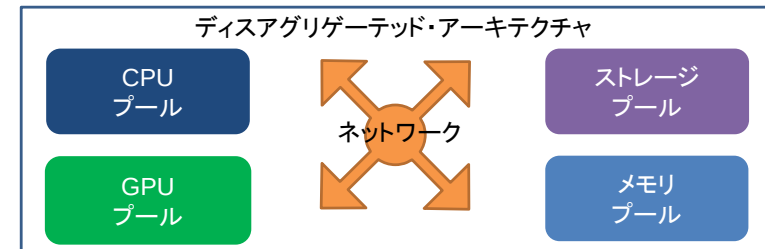
- Top500に見る傾向 → 並列化の限界



- メモリーウォールと、、
- ネットワークがボトルネックに

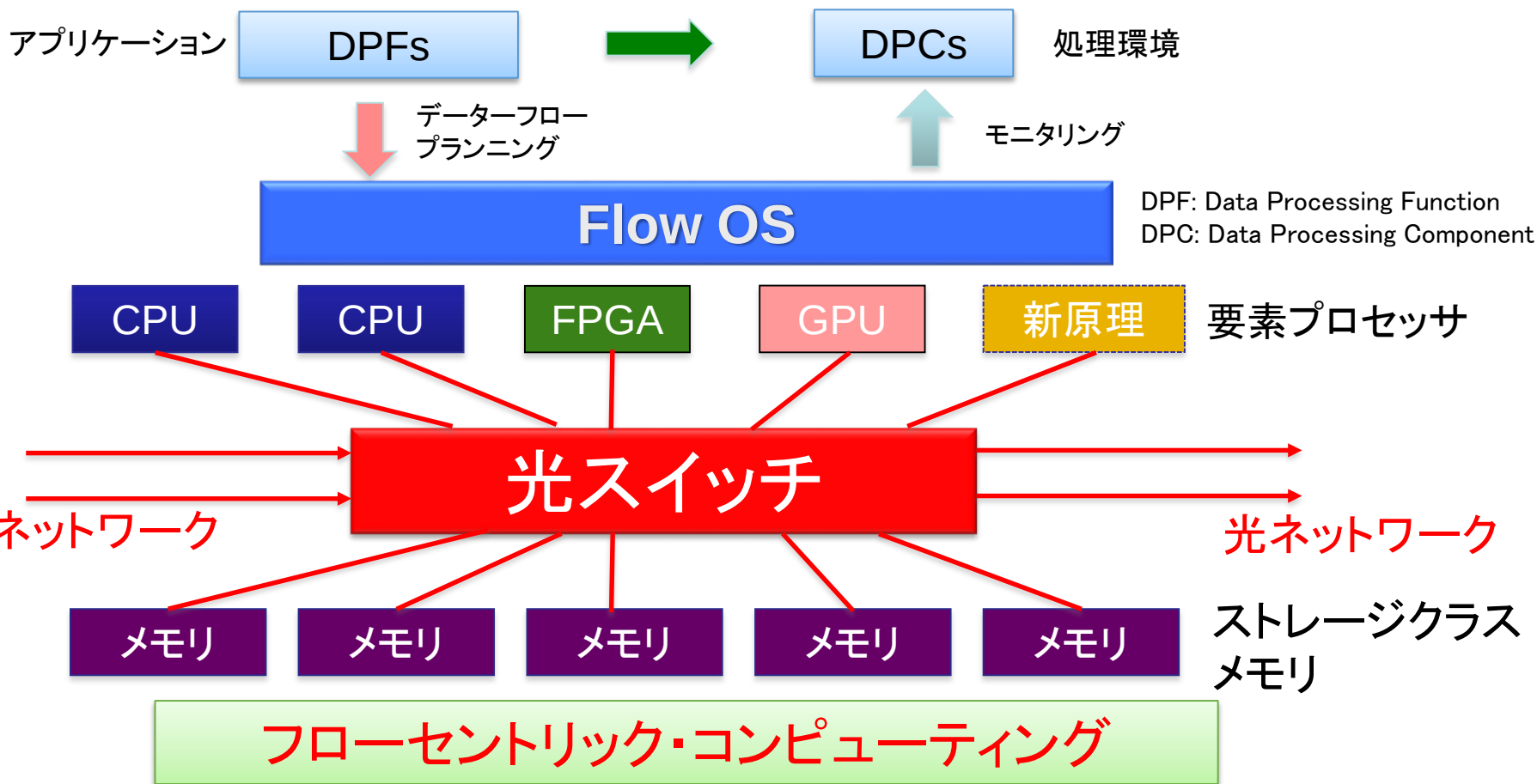
ディスタグリゲーションへの期待

- 従来型サーバー・アーキテクチャ
 - “All-in-one” CPU Blade Server → 無駄が多く、スケールしない。
- ディスタグリゲートッド・アーキテクチャ
 - 各機能ブロックを分離し、広帯域光インターコネクトで接続。
 - 無駄のないパワフルなシステム。
 - ボード間・ラック間の通信容量が巨大化
 - 光スイッチの導入は必至
 - 問題は、どうやって最適運用するか？



ディスアグリゲーションを司る、新しいコンピューティング

- データの流れを優先したアーキテクチャにより、実時間データ処理を性能保証
- ディスアグリゲーションに基づき、超高速光ネットワークと大容量ストレージクラスメモリを活用
- データセンタ全体を一つのオペレーティングシステム(OS)で効率的に運用・管理



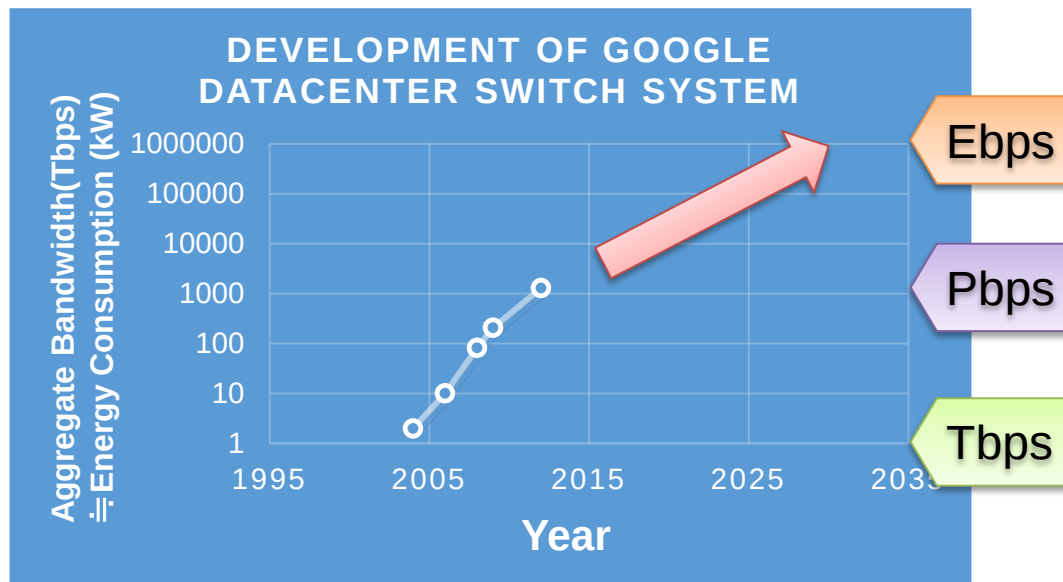
Switches: photonics vs. electronics

	Optical switches	Electrical Switches
Type of switching	Circuit switching	Packet switching
Granularity	Coarse	Fine and flexible
Bandwidth upgradability	Transparent > THz	Transceivers, LSI ~ 100G
Physical I/F	Optical connectors	Transceivers
Energy scaling	< W/port	0.2 nJ/bit (Infiniband) 6 nJ/bit (IP router)
<i>Energy consumption for 100,000 ports x 10Tbps</i>	<i>~ 1 MW</i>	<i>> 0.2 GW ($\propto$Infiniband)</i>
Scalability issue	Number of ports	Moore's law
QoS	Physically guaranteed	Depends
Control plane	Detached from data plane	Embedded in data plane
Latency (Start up)	~ sub μ s - s	< 500 ns per hop
Latency (Transmission)	~ 0 (speed of light)	< 500 ns per hop

Hybrid use is inevitable!

Let's talk about Exa Scale..

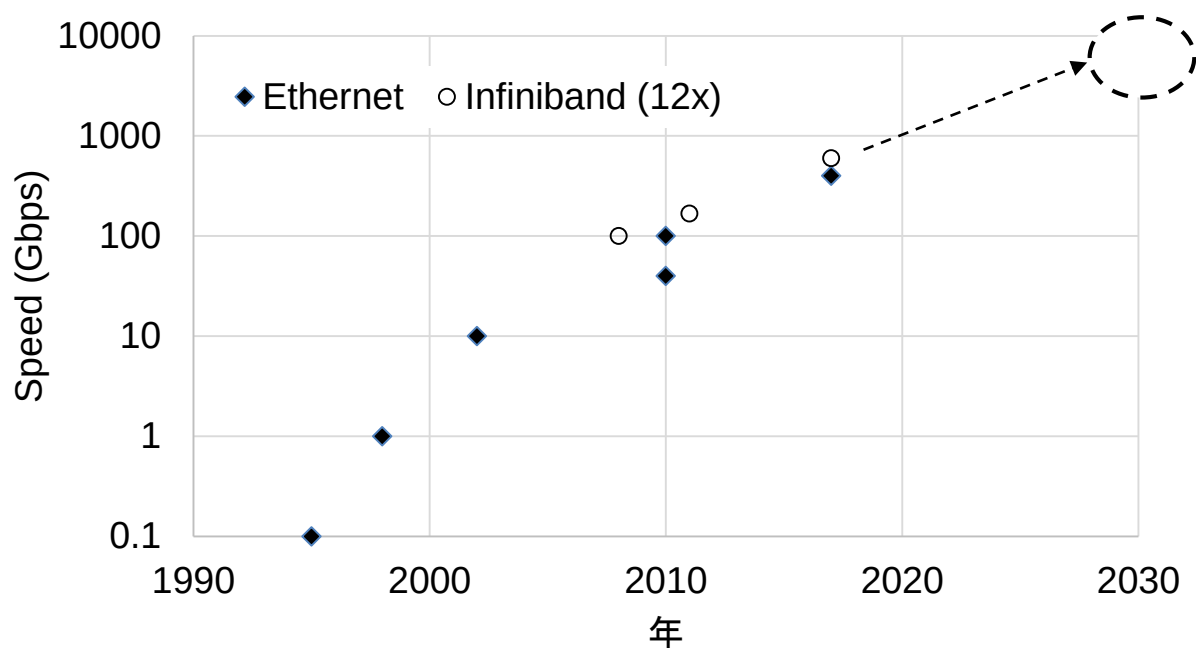
- One of Amdahl's laws says:
 - You need ~ 1 bit/s per FLOPS
 - Then, we need
 $\sim$ Exa bit/s for Exa FLOPS
 - ※TOP500では、 < 0.08 b/F
 - ※Post-Petaでは、さらに小さくなる
- Bandwidth vs. Radix
 - 100Gbps, 10M nodes?
 - 1Tbps, 1M nodes?
 - 10Tbps, 100k nodes?
 - 100Tbps, 10k nodes?



<http://www.theplatform.net/2015/06/19/inside-a-decade-of-google-homegrown-datacenter-networks/>

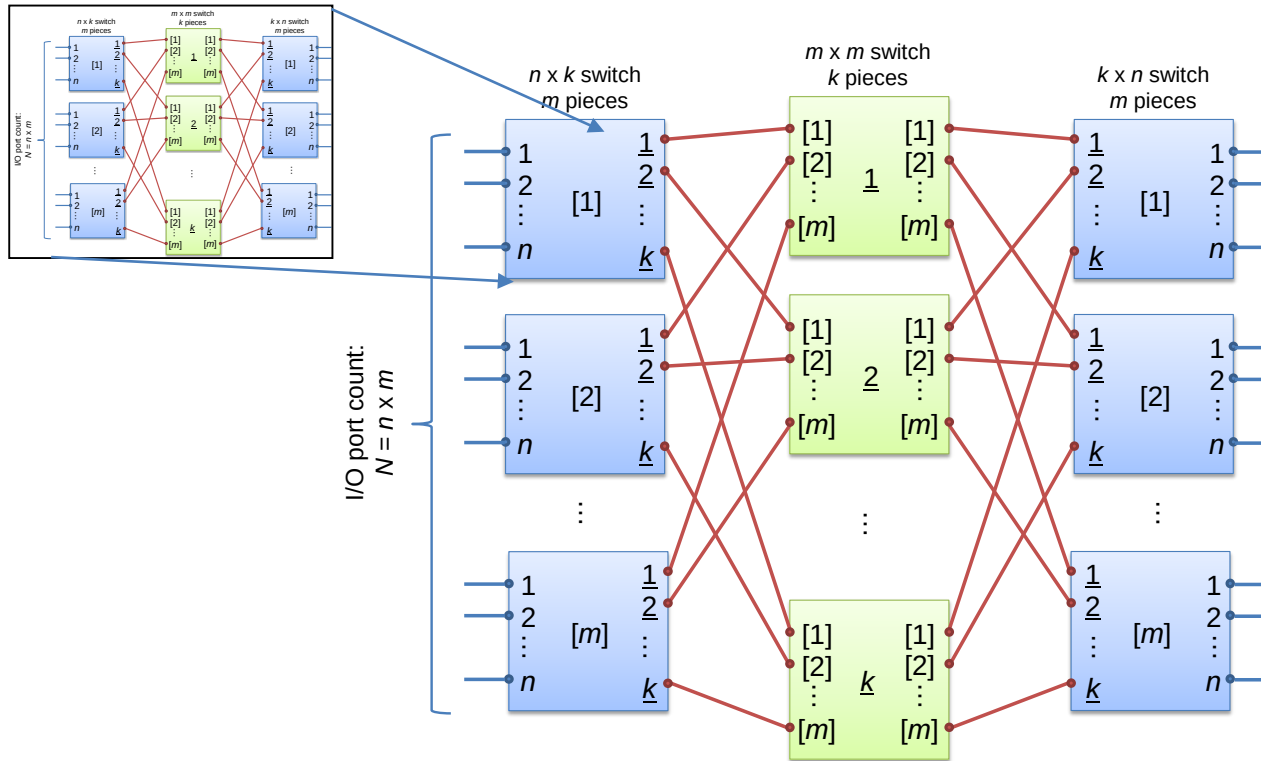
光トランシーバ、インタコネク速度の推移

- 2030年頃には、~10Tbpsへ



http://www.infinibandta.org/content/pages.php?pg=technology_overview
<http://www.ethernetalliance.org/subcommittees/roadmap-subcommittee/>

10万ポートのための、多段Clos Networks



Hierarchy order	# of hops	3D MEMS	Si-Photonics	PLC
0	1	512	32	8
1	3	131,072	512	32
2	9		131,072	512

$\hookrightarrow n^2/2$
 $\hookrightarrow n^2/2$

産総研のシリコンフォトニクス光スイッチ開発状況

小型、低ロス、低クロストークのための設計指針

- ✓ シリコン細線導波路(ただしTMモード)
- ✓ 全て方向性結合器で構成
- ✓ 熱光学効果による位相シフト
- ✓ PILOSSトポロジ

産総研「VICTORIES拠点」で
要素技術開発を実施

FY2014

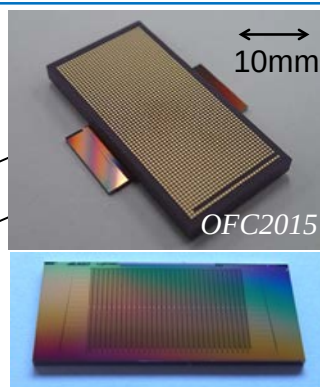
Record-large port
count 32x32

FY2013

Record-small 8x8



In-house process
(e-beam litho.)



FY 2015

Design
optimization
& test;
Control
circuit &
program

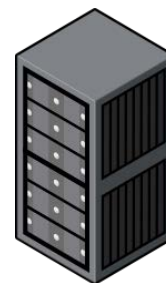
FY 2016

8 x 8 &
32 x 32
SW card



FY2017

Standard rack;
Clos network SW



<http://www.aist-victories.org/>

<これまで>

32x32までのポート数拡張を実証(1st sample)
光SWチップの作製技術・実装技術にメド

<今後>

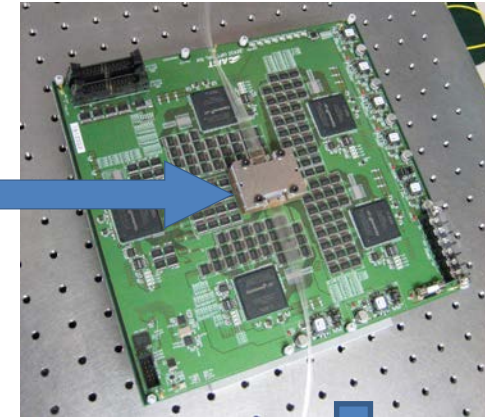
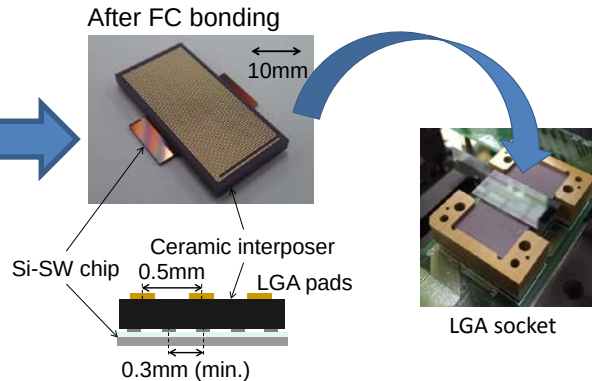
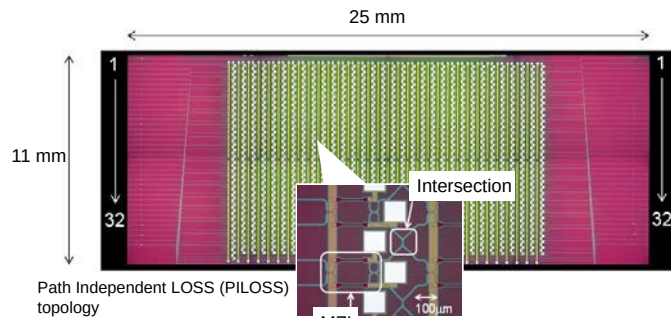
製品化を目指して「使える」スイッチを目指す
光学設計の最適化、制御回路との一体化

45-nm CMOS foundry process (SCR)

32 × 32 シリコンフォトニクス・スイッチ

K. Tanizawa et al.,
Opt. Express 23, 17599 (2015).

■ 32 x 32 silicon photonics switch integrated on control electronics



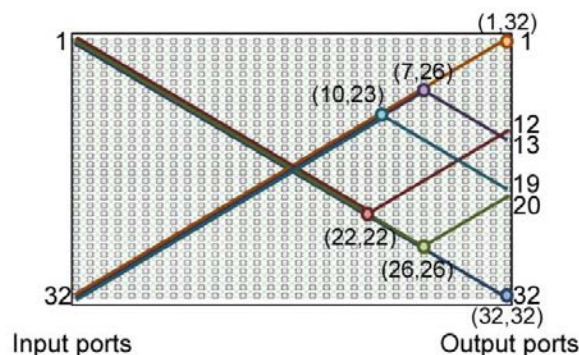
世界最大規模のシリコンフォトニクス素子
マツハツェンダスイッチ**1024個**、マイクロヒータ**2048個**、導波路交差**961個**を単一チップ上に集積
チップサイズは石英PLCの**1/40以下**

- Flip-chip package with LGA interposer
- Driven by pulse width modulation with FPGAs
- Polarization insensitivity, lower crosstalk and loss are now under development.

<Current optical performance>

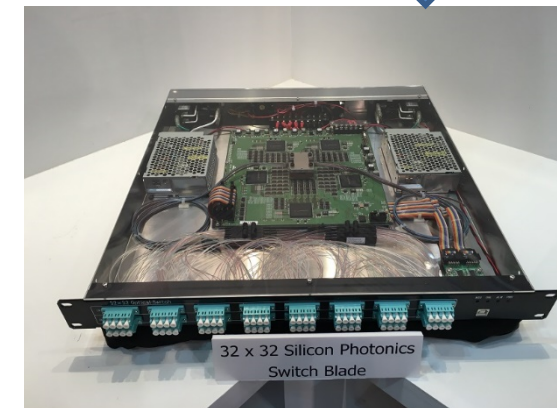
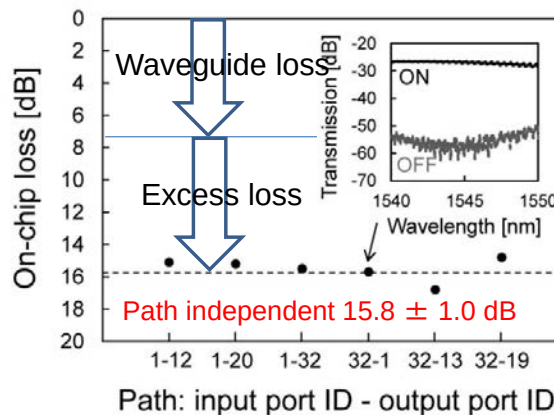
–On-chip loss: 15.8 ± 1.0 dB

–Crosstalk: < -20 dB in the operating bandwidth of 1.8 nm



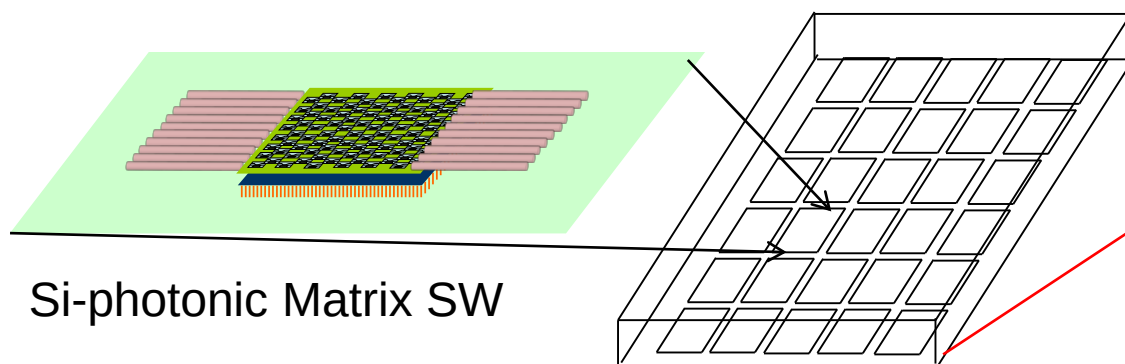
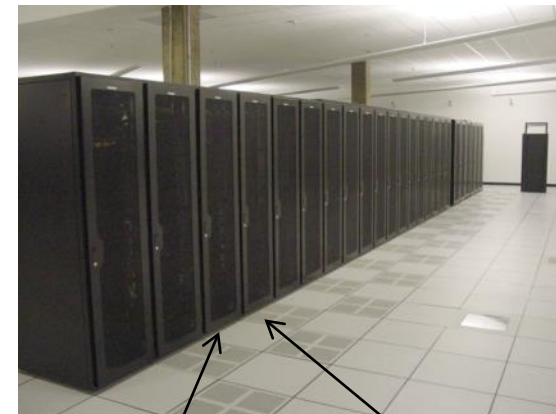
1 bar-state SW

31 cross-state SWs (Trimming)

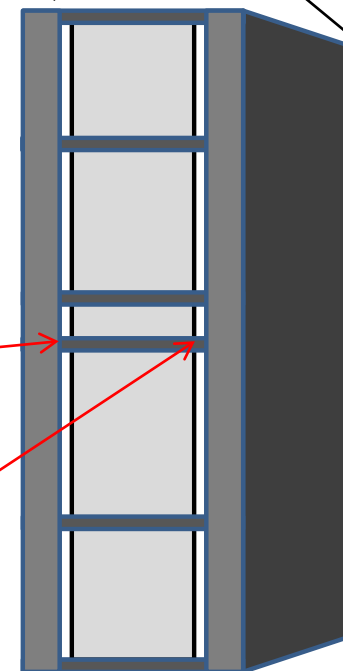


10万ポート光スイッチシステムの規模感

- 19' rack
 - 1U = W:48cm x H:5cm x D: 90cm
 - 42 U (H: ~2m)
- 32 x 32 Si Photonics SW Footprint
 - 9cm x 10cm → 5 x 9 pcs/1U
 - 45 x 42 = 1890 pcs/rack
 - ~ 12 kW per rack
- 100,000 x 100,000 fabric
 - 147,456 pcs of 32x32 SWs
 - $147,456 / 1890 = 78$ racks (~ 1 MW)
 - Approximately 78 racks are required
 - ... Less than 10% of all racks in mega DC (~1000 racks)



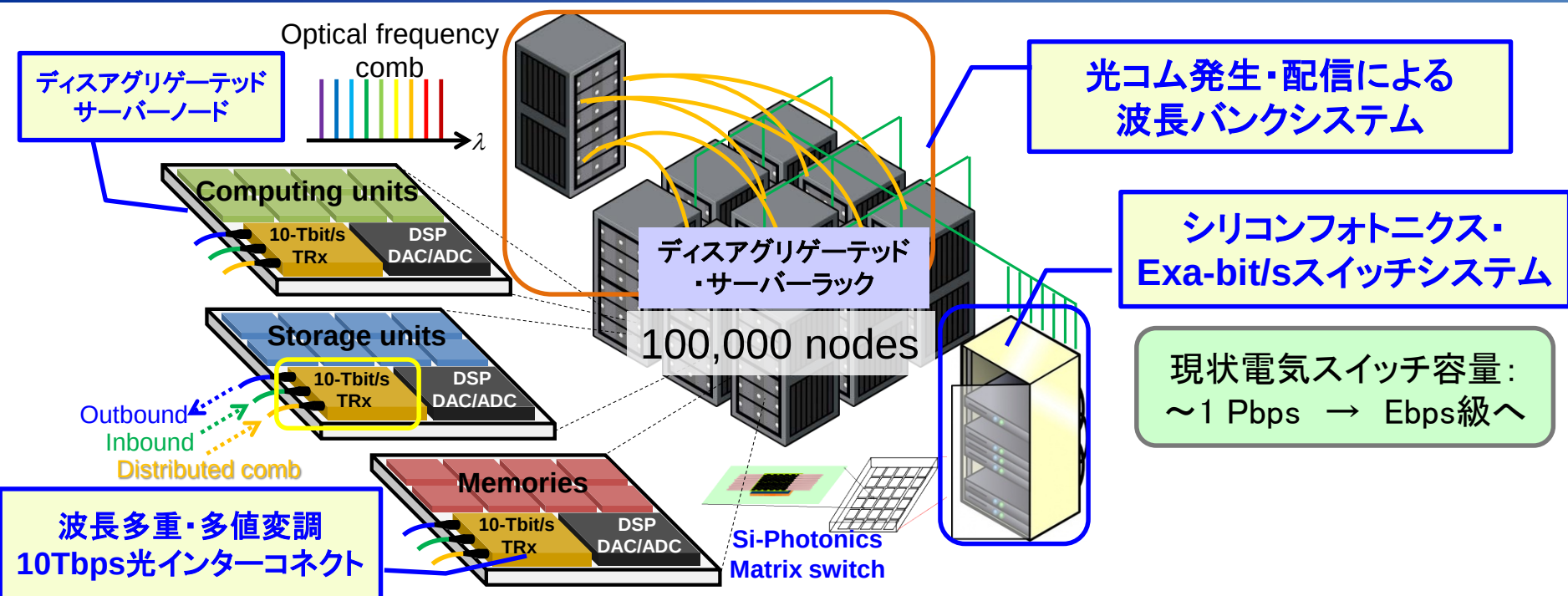
Si-photonic Matrix SW



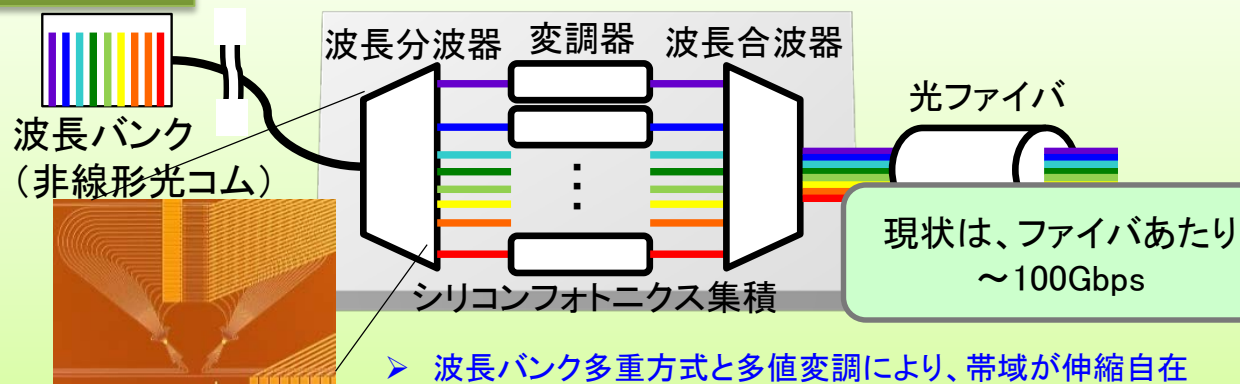
光スイッチ導入のインパクト

- エクサスケールで、1,000分の1の超低電力
 - 電気スイッチだと、1GW級の電力
 - 光スイッチだと、わずか1~2MW程度で実現
- サーキットスイッチ
 - 全体を司る上位レイヤー(OS)との連携が必要
- 新しい光トランシーバ技術が必要
 - 光スイッチと互換: 単芯シングルモードファイバ
 - WDMと多値変調は、必至！！
 - 従来光インターコネクト比
 - 10,000倍以上の伝送距離(～ 10m → > 100km)
 - 既存広域用光トランシーバ比
 - 100倍の信号帯域 (100G → 10T)
 - 数十倍の電力効率 (400pJ/bit → 10pJ/bit)
 - 100分の1のフットプリント(100Gと同サイズ)

エクサスケール光ネットワーク技術への挑戦 一例



波長バンク方式



メリット:

- ・温調と気密封止が不要
- ・高い波長精度

10Tbpsのための伝送方式検討

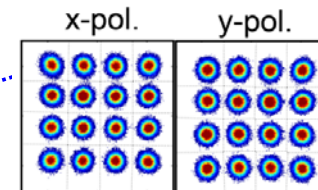
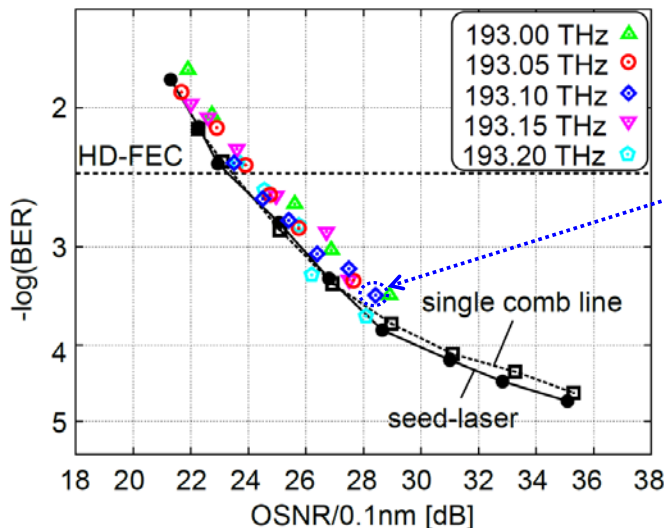
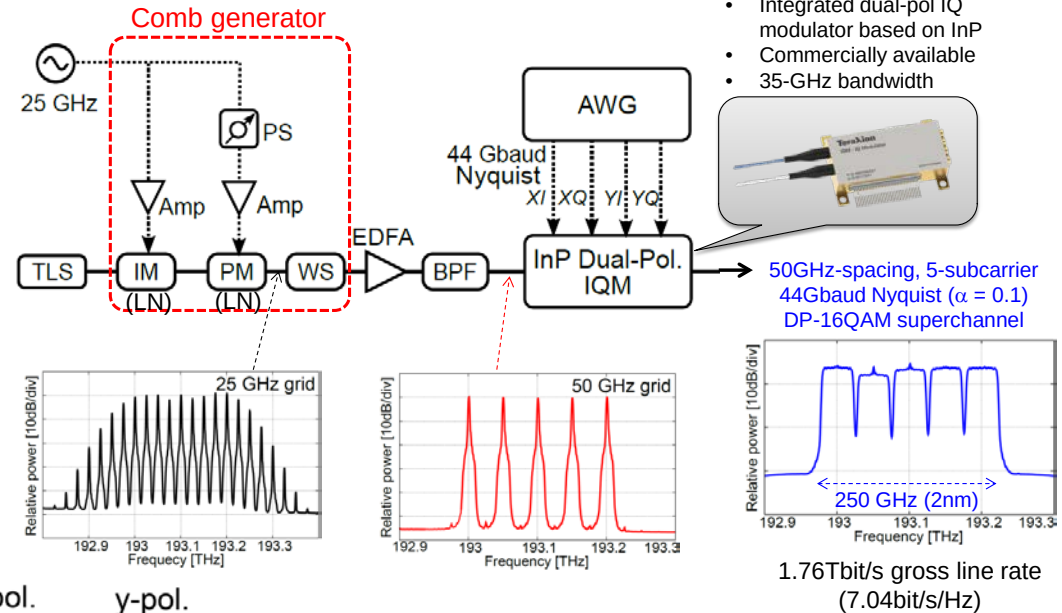
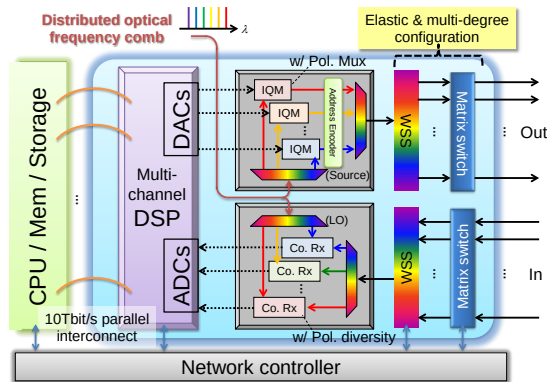
Baud rate (net)	Degree of multiplexing	Channel number	Channel spacing	Total Bandwidth	Required SNR/Q
28 (25)	2 (PAM4)	200	>50GHz	> 80nm	20dB
	4 (DP-QPSK)	100	30-50GHz	24~40nm	15dB
	8 (DP-16QAM)	50	30-50GHz	12~20nm	22dB
	12 (DP-64QAM)	33	30-50GHz	6~10nm	28dB
43 (40)	2 (PAM4)	125	>80GHz	> 80nm	22dB
	4 (DP-QPSK)	63	50GHz	25nm	17dB
	8 (DP-16QAM)	32	50GHz	13nm	24dB
	12 (DP-64QAM)	21	50GHz	9nm	30dB
56 (50)	2 (PAM4)	100	>100GHz	~40nm	23dB
	4 (DP-QPSK)	50	75GHz	30nm	18dB
	8 (DP-16QAM)	25	75GHz	15nm	25dB
	12 (DP-64QAM)	17	75GHz	10nm	31dB

Smaller is better

Should be < 15nm for Si-SW passband
Should be < 30dB for SNR margin

10Tbpsトランシーバに向けた基礎評価実験

1. Optical comb source based on cascaded LN modulators (IM + PM)
2. Integrated IQ modulator based on InP
3. 44Gbaud DP-16QAM, Nyquist of $\alpha = 0.1$
4. Five Carriers



Capable of transmission up to 320-km SSMF link

世界最高レベルのスペクトル密度達成

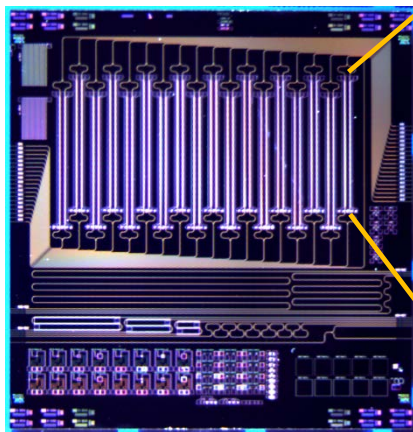
Carrier #	Total capacity
5	1.6Tbit/s
32	10.24Tbit/s

K. Solis-Trapala et al.,
OECC2016.

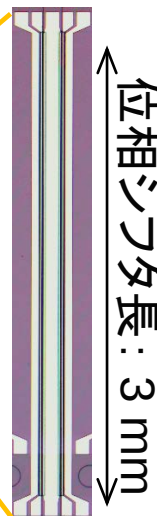
シリコン変調器の進捗

- ・PN構造付Si導波路、進行波型電極の設計完了し、光変調器の設計スキームを確立
- ・高速光電子融合デバイス測定系の構築を完了し、変調器評価に活躍
- ・2"ウェハーによる高効率型変調器の先行試作を完了し、高速高効率動作を確認
 - 静特性: $V_{\pi L} \sim 1V \text{ cm}$,
 - 動特性: $f_{3dB} > 20GHz$,
 - 標準的なシリコン変調器の約2倍の高効率
 - 32 Gbps程度的高速変調が可能
 - 変調動作は12.4 Gbpsまで確認(測定系の限界)
- ・SCRでの標準型変調器の本格検討試作が進行中。6月末完成予定

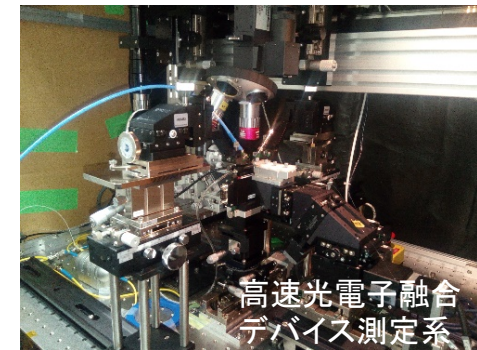
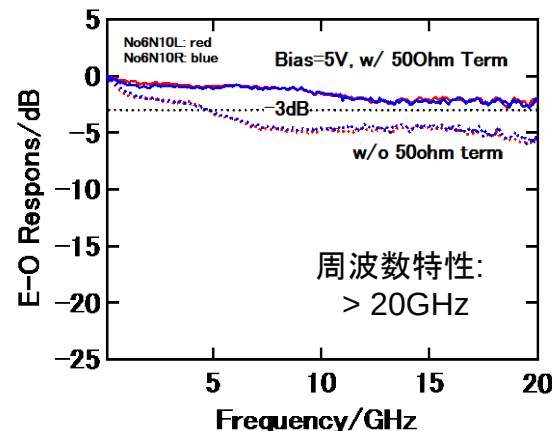
2"先行試作の概要



チップ構成



変調器

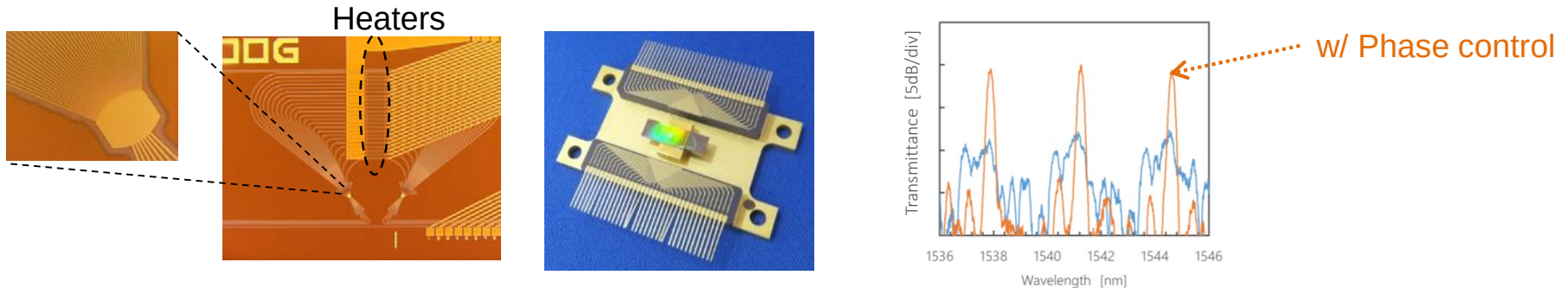


高速光電子融合
デバイス測定系

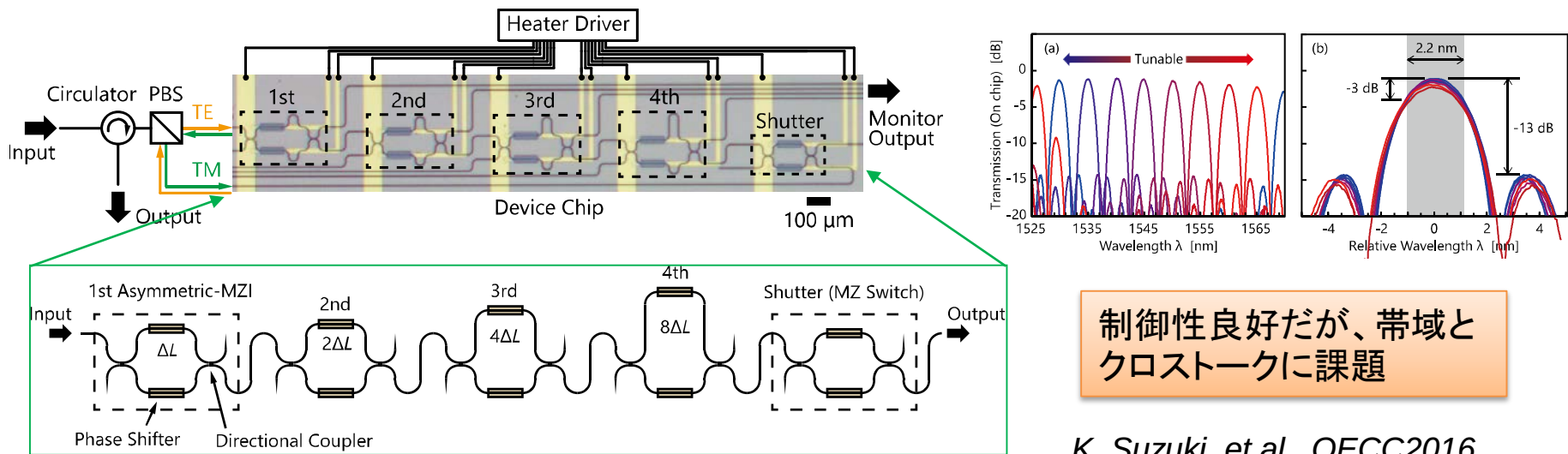
SCRによる波長セレクタに関する検討

- AWG基礎検討

- ナノワイヤ構造では、極限的位相調整が必要。



- 非対称マッハツエンダー干渉による可変波長フィルタ



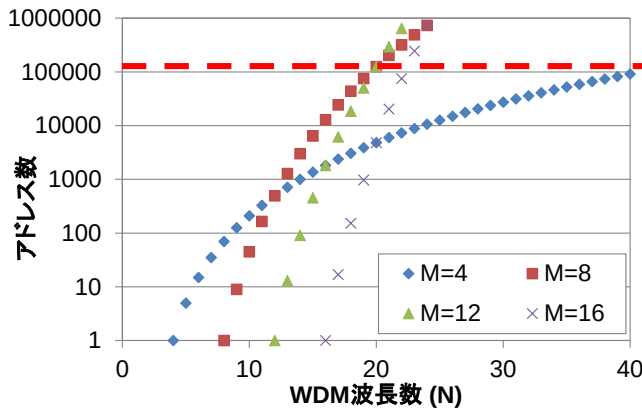
K. Suzuki, et al., OECC2016.

光スイッチ制御の簡素化に関する検討

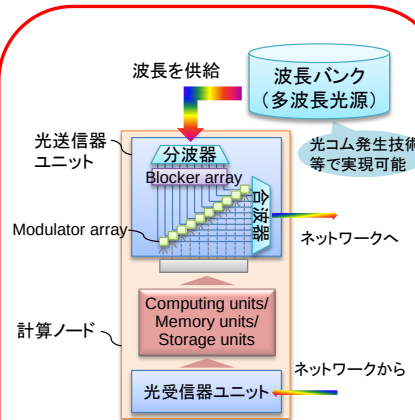
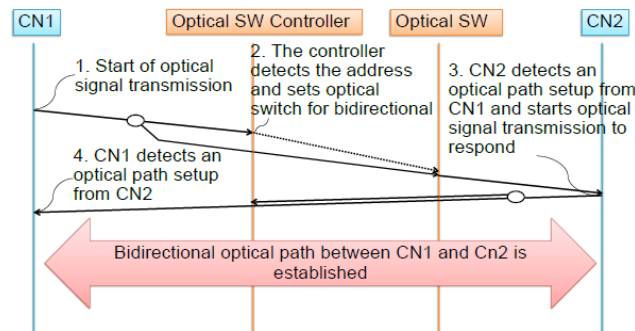
- 送信側がアドレスを波長配置にマッピング
- 各光スイッチがローカルに行先を判別
 - 制御プレーンによる同期制御が不要
 - 25~40波で、10万以上のアドレス空間

4波長中2波長にマークした場合

計算ノード	波長アドレス
1	λ_1 λ_2
2	λ_1 λ_3
3	λ_1 λ_4
4	λ_2 λ_3
5	λ_2 λ_4
6	λ_3 λ_4

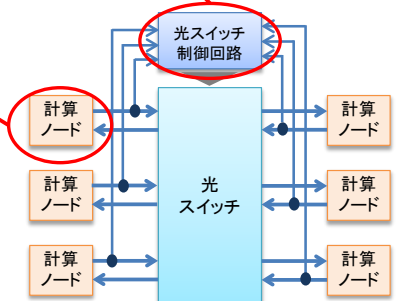
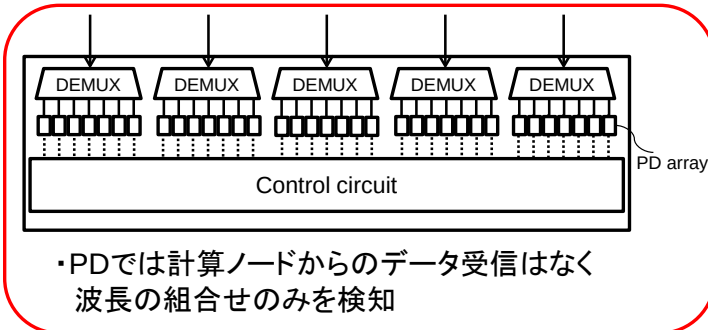


石井他、信学会総合大会2016



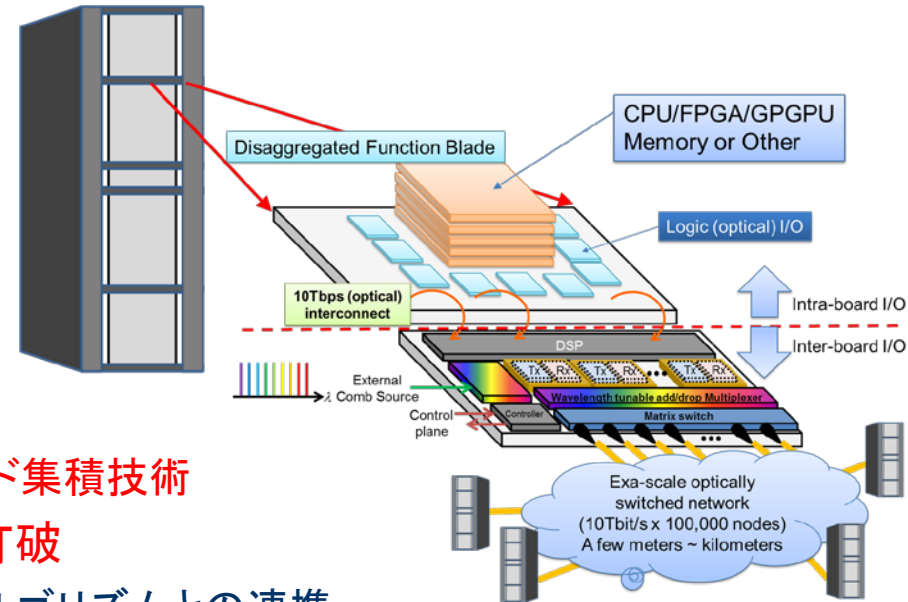
<多波長光源の配信>

- ☺ 光送信器ユニットの簡素化
- ☺ 高速波長変更が容易
- ☹ 偏波?



ポストムーア時代を拓くNW技術を創出するには

- エクサスケールNW全体の電力消費 **～10MW**
 - ネットワーク: $1\text{Ebps} = 10\text{Tbps} \times 100,000\text{ポート}$
 - スイッチシステム: 光で **<2MW** を目標
 - トランシーバは、**10pJ/bit** を目標 → 10MW
- **10pJ/bit以下** を実現する光トランシーバ
 - 新規変調器技術？
 - 新規ドライバ技術？
 - TEC、ハーメチック・フリー？
 - **DSP** を大幅に改善する技術
 - コスト、電力、レイテンシー、他
 - **さまざまな技術を組み合わせるハイブリッド集積技術**
- **個別技術の限界は、技術の垂直融合で打破**
 - フローセントリックコンピューティングやアルゴリズムとの連携
 - 電子光融合アセンブリ技術
 - **ロジックボードと光トランシーバの一体設計**



- 製造技術に関するエコシステムの構築

持続的発展を可能とするエコシステム造り

- 研究開発力を集約

- 水平分業化が進み、弱体化する企業の国際競争力の増強

- 垂直統合で事業化支援

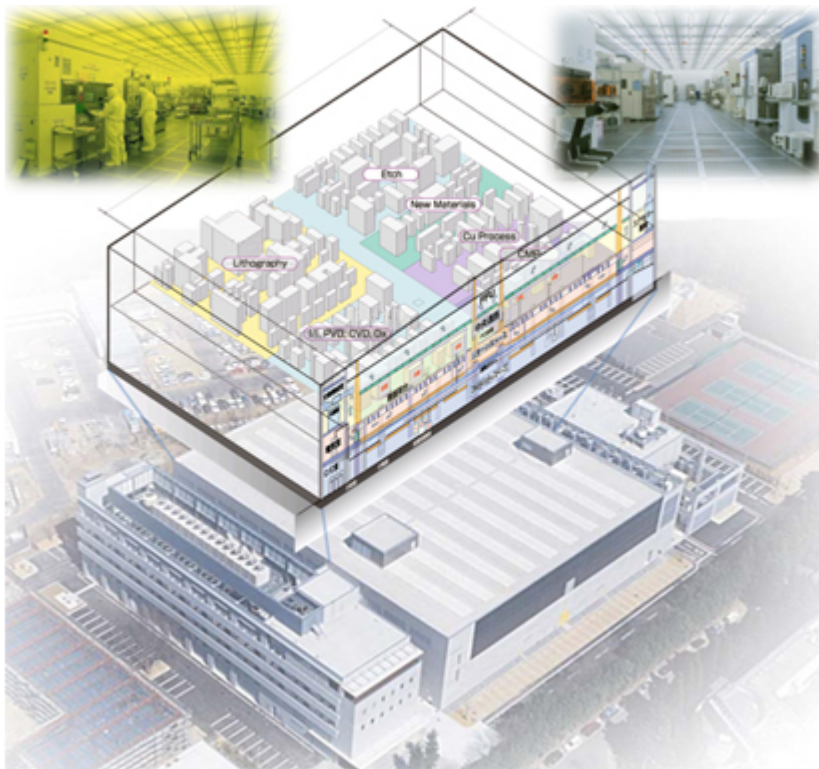
- 研究試作から量産まで
- デバイスからアプリケーションまで



テストベッドを
活用した
アプリケーション
とのコラボ



AIST-Super Clean Room (SCR)



- ✓ 3,000 m² floor area
- ✓ Cleanliness: Class 3 ($< 10^3/\text{m}^3$)
- ✓ 12-inch SOI wafer
- ✓ 45-nm CMOS technology utilizing ArF immersion lithography



Available for outsiders. <http://unit.aist.go.jp/tiaco/orp/scr>

産総研 2014年「共用施設利用制度」制定

産総研 共用施設利用制度

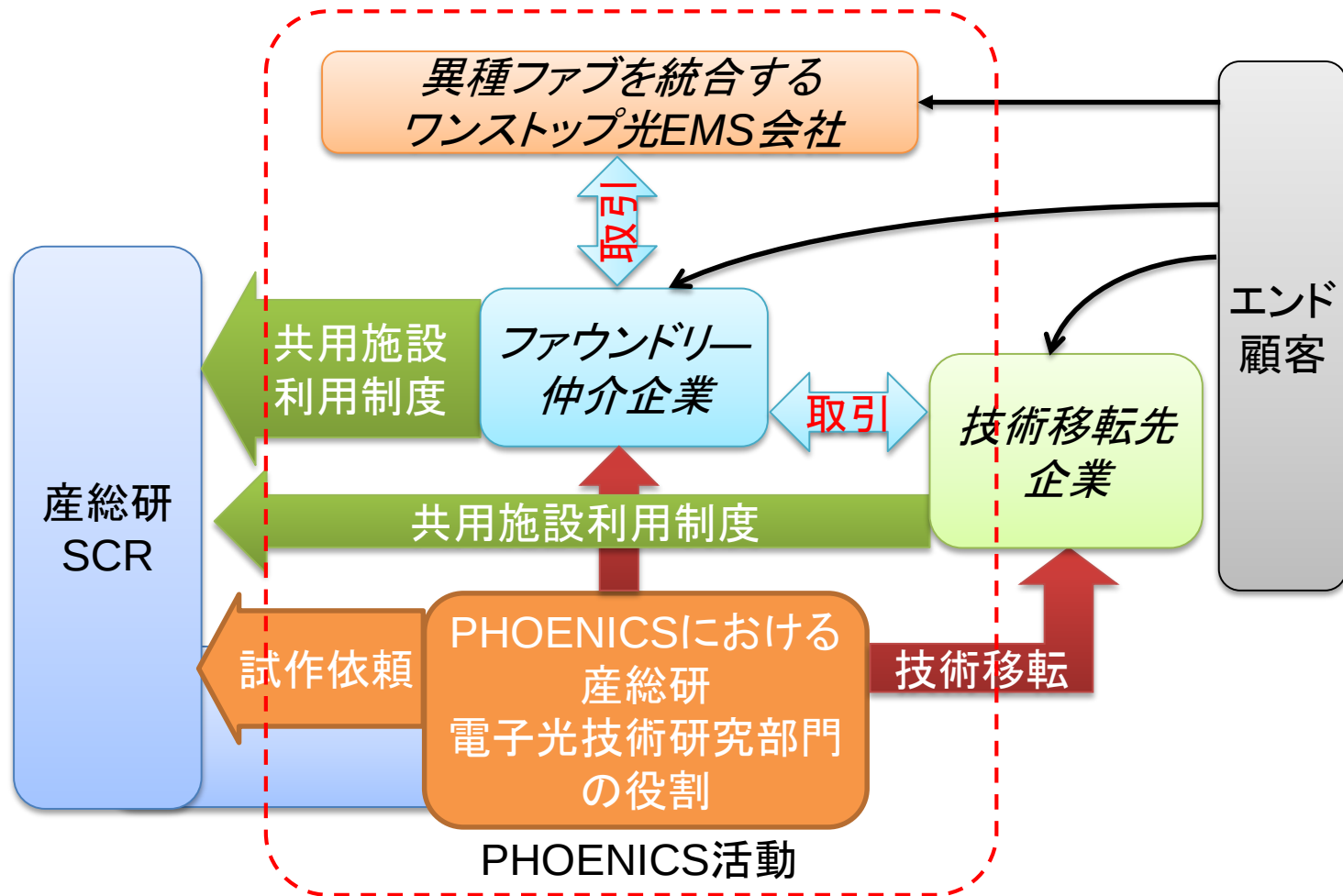
- 狙い:
 - 企業などにおける研究開発に関連した事業活動の活性化、産総研のオープンイノベーションハブ機能の強化
 - つくばイノベーションアリーナナノテクノロジー拠点(TIA-nano)事業強化の一環
 - SCRの利用促進により、半導体関連企業などへの研究開発支援の充実と促進
- 利用対象:
 - LSI量産工場に匹敵し、国内公的研究機関では最大規模のスーパークリーンルーム(SCR)
 - SCRに設置されている約100台の半導体製造装置
 - 今後、産総研の研究成果が活用されている他の大型研究施設、特殊装置などにも適用の範囲を広げていく予定
- 特長:
 - 利用約款に基づいた申請・回答ベースの簡便な利用手続き
 - 利用により発生した知的財産権は原則として利用者に帰属
 - 明瞭な秘密情報管理ルールへの遵守
 - 単価表ベースによる利用料金で、利用コストの見通しが容易

<https://unit.aist.go.jp/tia-co/orp/index.html>



シリコンフォトニクス・ファンドリ・サービス構想

- エコシステム構築へのステップ



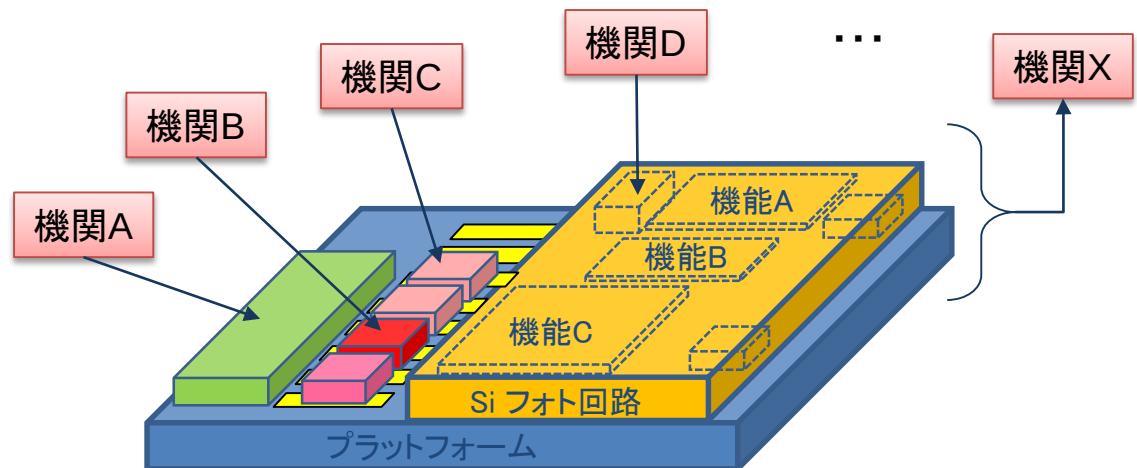
コンソーシアムによる、国内光デバイス11社の協業

光デバイス基盤技術イノベーション研究会(産総研コンソーシアム)

PHOENICS (PHOtonics ENgineering Innovation ConSortium)

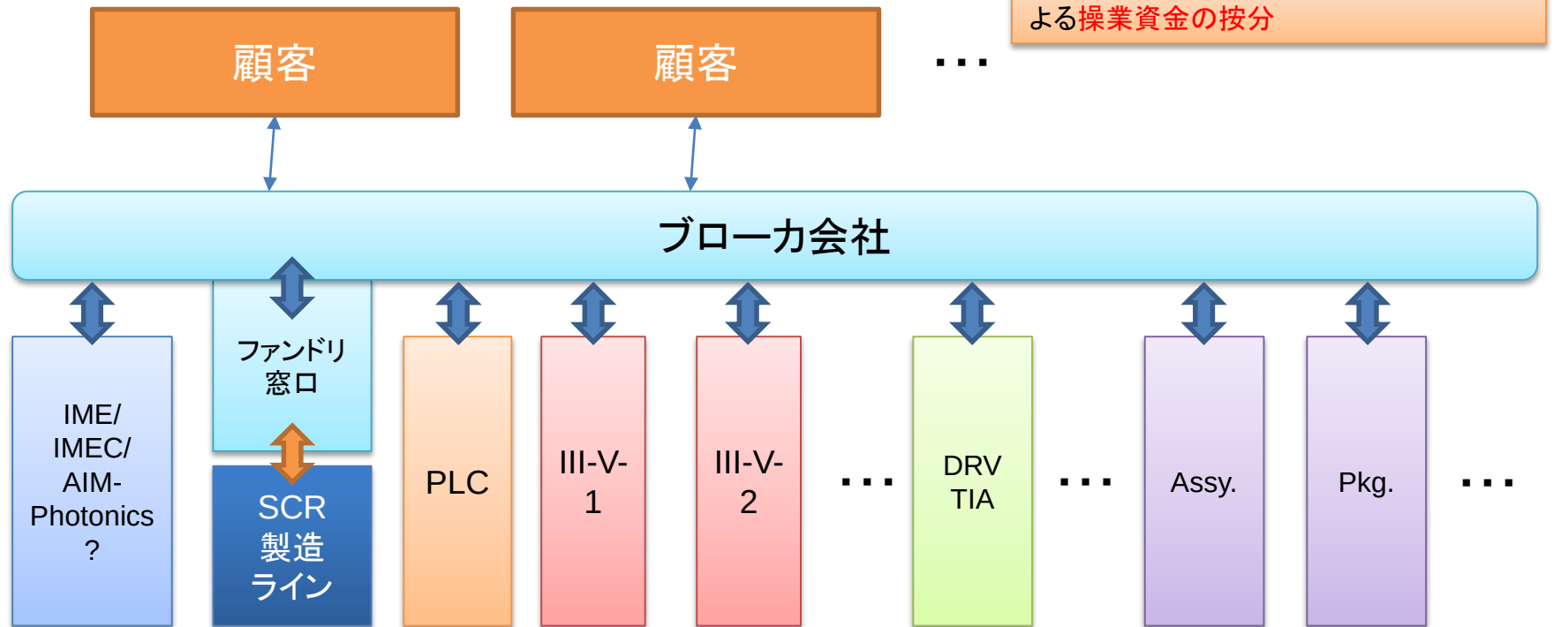
- 目的
 - 日本の光デバイス産業の国際競争力を強化するため、**散在する基盤技術を結集**する最適な方法を検討
 - 光デバイス設計・実装共通基盤の整備により、将来の**光デバイスのファブ**を、バーチャルに実現
 - 国プロなどを経て、光デバイスの**ワンストップEMS会社**の設立へ
- 期間
 - 2015年4月1日から3年間(コンソーシアム設置期間)
 - 並行して事業化へ展開

事業化・起業へ向けた、
参加者持ち出しからの出発



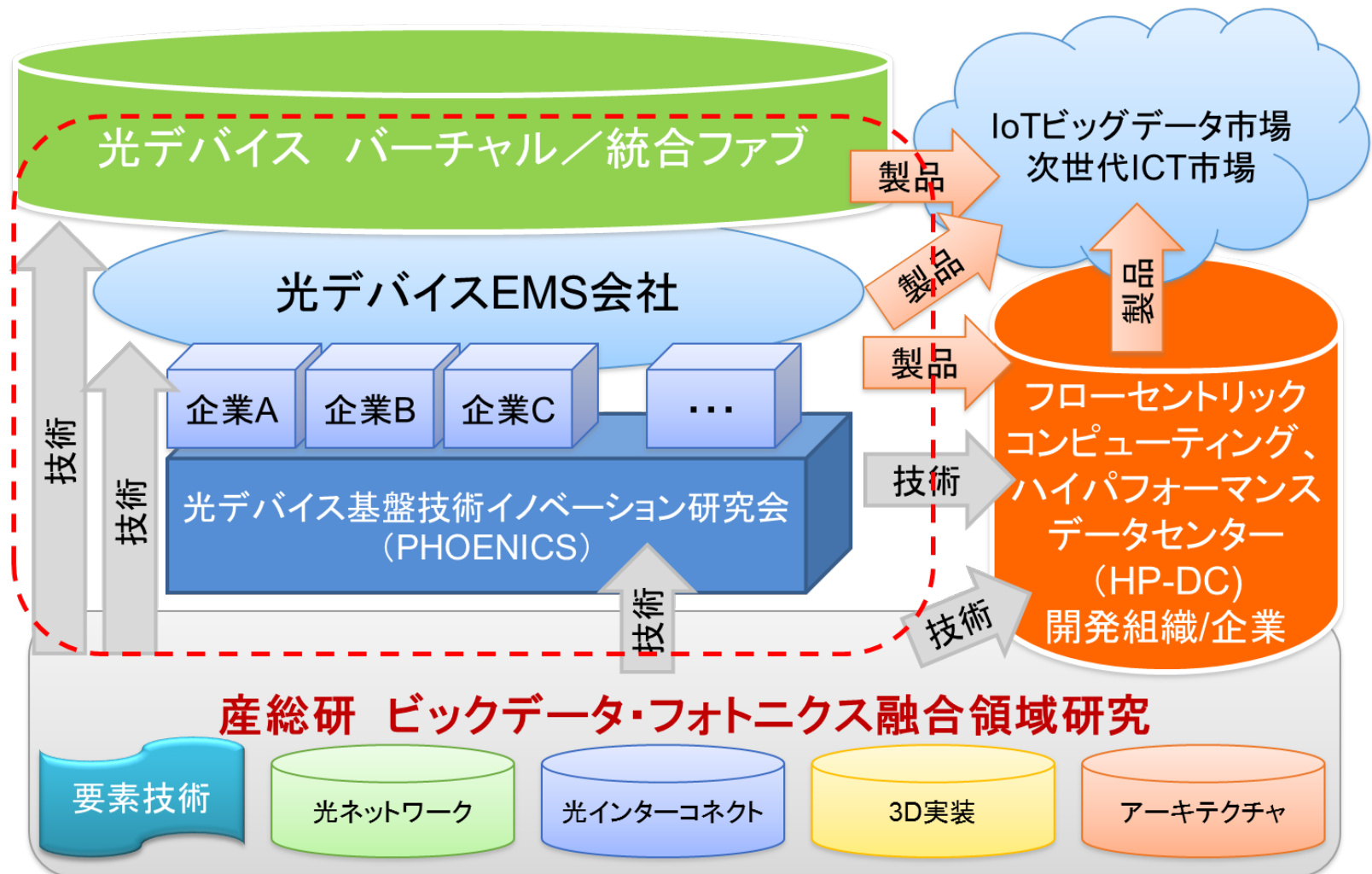
ハイブリッド集積デバイス エコシステム・イメージ

- バーチャルファブ、ブローカ会社
 - プラットフォーム標準化
 - ハイブリッド実装時の設計・品証
 - 責任分界点の設計
 - C-POTインターフェース規格



目指すべき持続発展可能なエコシステム(～2030)

- 垂直融合体制において、研究試作から量産まで、シームレスなファブ環境
– 日本だけでなく、グローバルな規模も視野



まとめ:ハードウェアの回帰と千載一遇の機会

- ポストムーア時代 = Game Changerの台頭
 - ネットワークとコンピュータが融合する新しい情報産業が形成される
 - アーキテクチャからデバイスまで含めたホリスティックなシステムの最適化が必要
 - 光スイッチによる、光ネットワークの飛躍的大容量化が鍵
- ハードウェア時代の再来:
 - シリコンフォトニクスなどによる、高度なハイブリッド実装技術と量産効果による低コスト化の両立が鍵である
 - ファブによる製造の集約やエコシステム造りが世界規模で必要
 - 産総研コンソーシアム「PHOENICS」では、散在する光デバイス技術を再編し発展させ、バーチャルファブを構築することによって、日本の高い光デバイス技術を結集し、世界的競争力を確保していく検討を実施中
- 世界の新しい情報産業にとって不可欠となる光デバイス基盤技術を日本が担っていく