>THIS IS THE WAY

Lightpaths for Protocol Performance

Kim Roberts
Optical Systems

>THIS IS NORTEL

Motivation

> The purpose of protocols for fast long distance networks is to provide quality connections for “real scientists”, studios, hospitals, etc.

> There is a two to five year time lag from starting your research to widespread user adoption of your solutions.

> To be relevant, you need to be trying to solve the problems that users will have two or more years from now.

> Will these be the same old problems, with just more steroids?

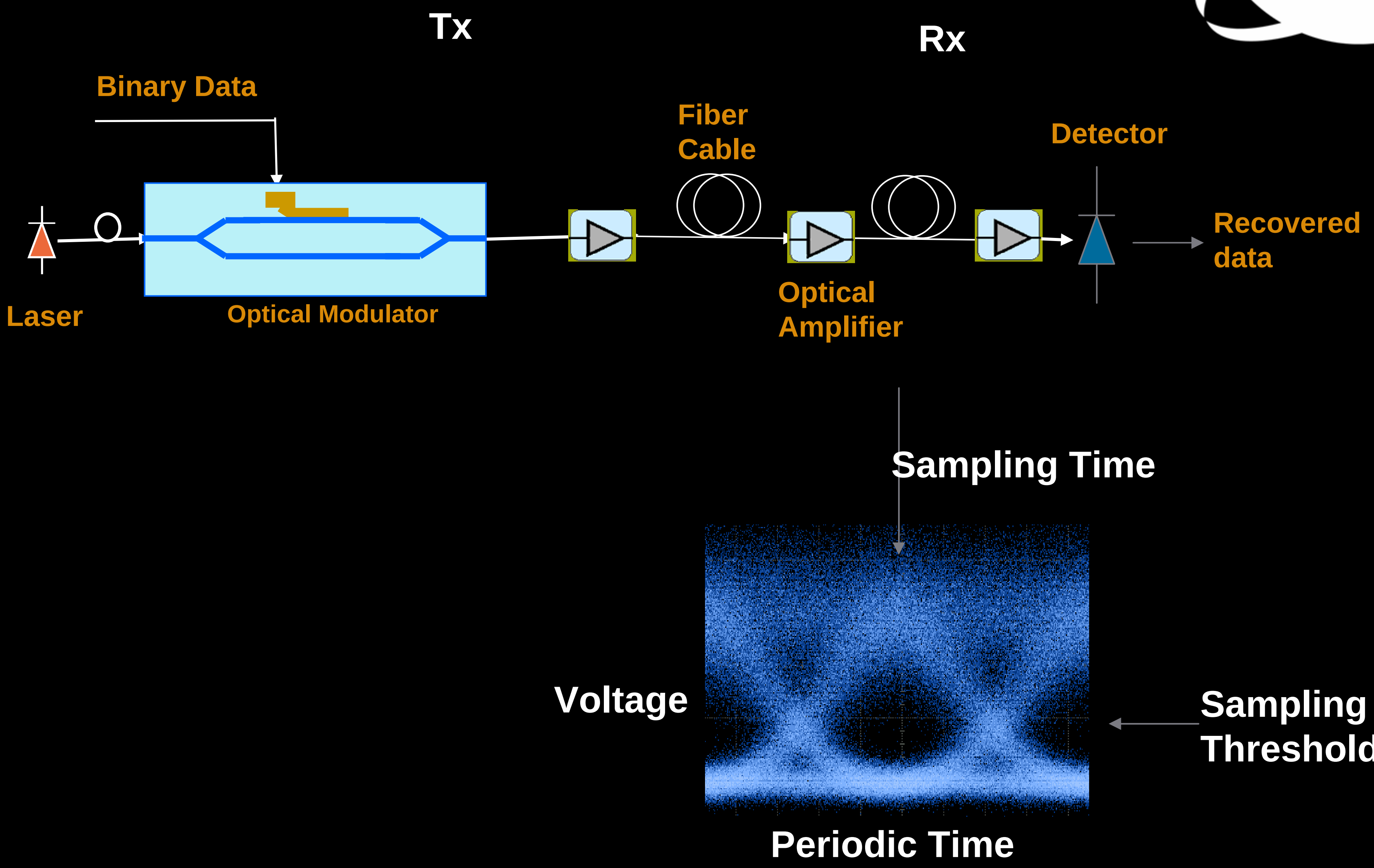
A LightPath

> A network connection that provides guaranteed capacity, constant latency, and high reliability.

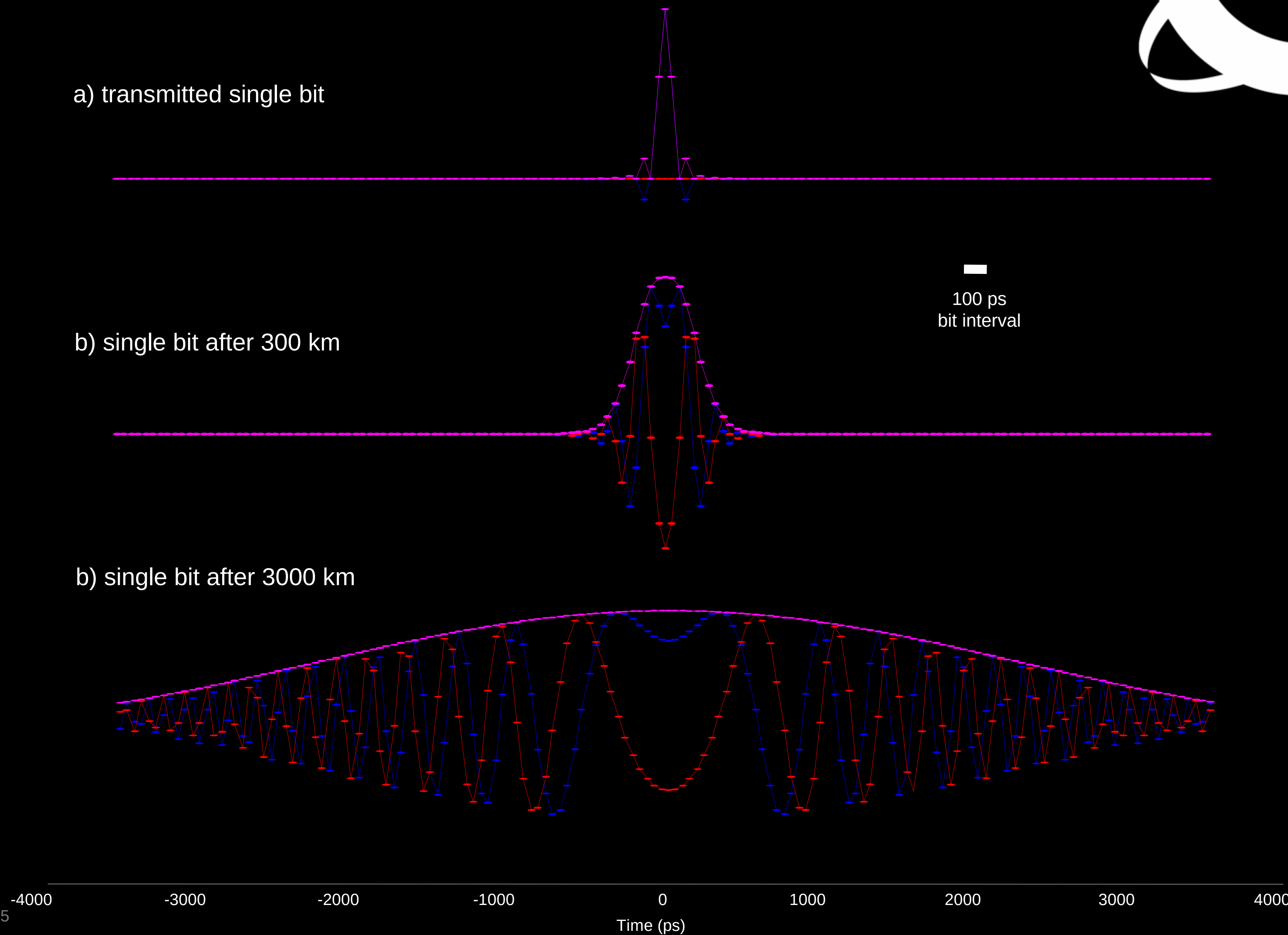
- Provisioned by a control plane.
- Megabits to Terabits per second
- The optimum implementation can be L0, L1, or L2, depending on the capacity/quality needed and facilities available
- L3 should be used for the many small bandwidth demands that do not need these guarantees

Layer 3	Routing	IP
Layer 2	Packet	Ethernet
Layer 1	Transport	OC-192
Layer 0	Photonic	Photons

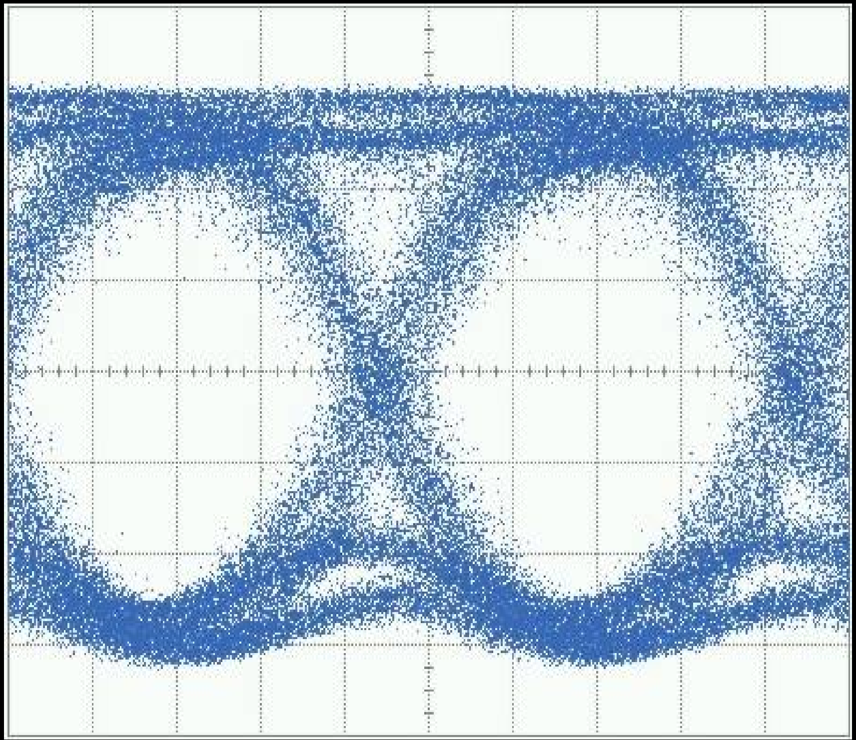
Traditional 10G Optical Link



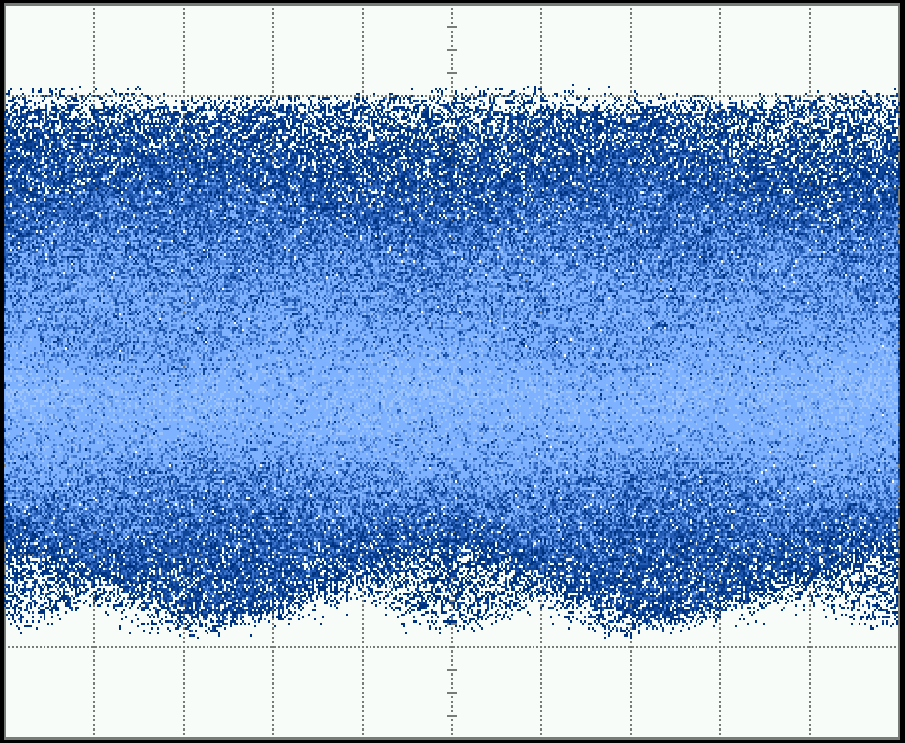
Chromatic Dispersion in Optical Fiber



Effect of Dispersion



Transmitted Eye



Eye after 320 km of Fiber

Dispersion Compensation Modules (DCM)

- Coils of 1 to 20 km of special fiber
- 5 μs to 100 μs of added delay per coil
- $3k to $10k per module
- Generally needed after every 80 km of cable

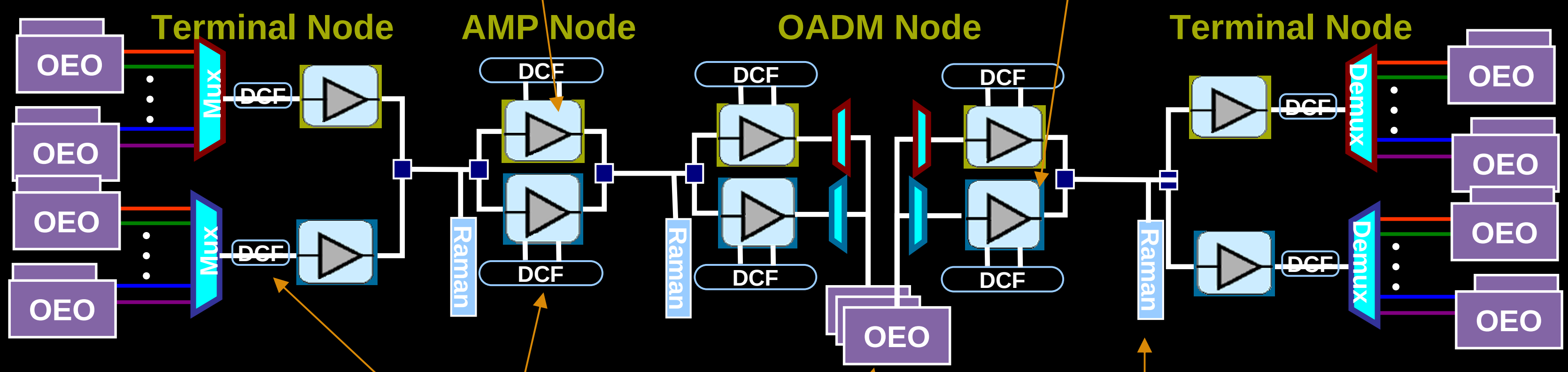
Details of a High Capacity Optical Line

$ C-BAND AND L-BAND EDFAS TO GET >800 Gb/s

$

DOUBLE EDFAS FOR MID-STAGE ACCESS

$



$

DISPERSION COMPENSATING FIBERS

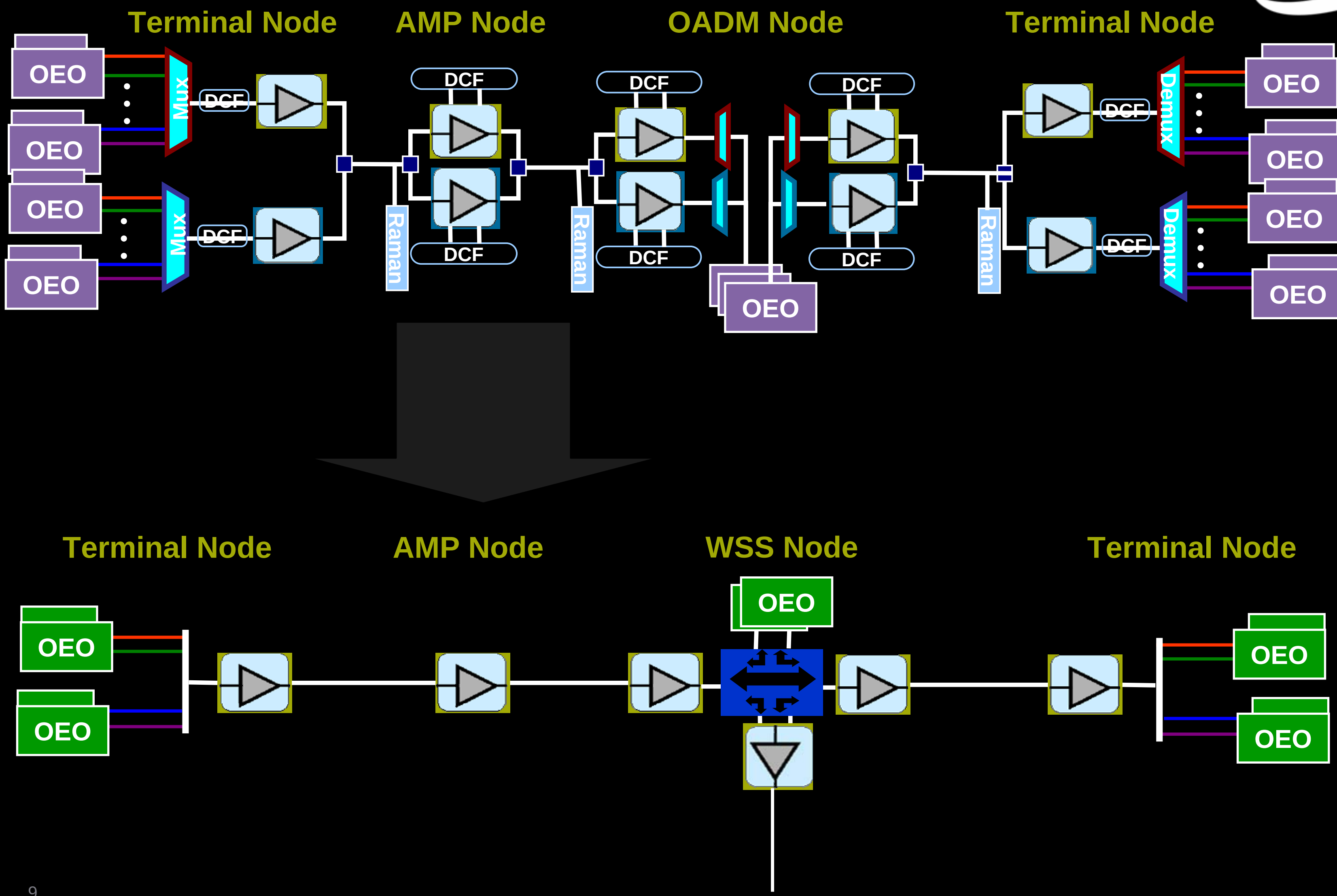
2 W RAMAN PUMPS

$

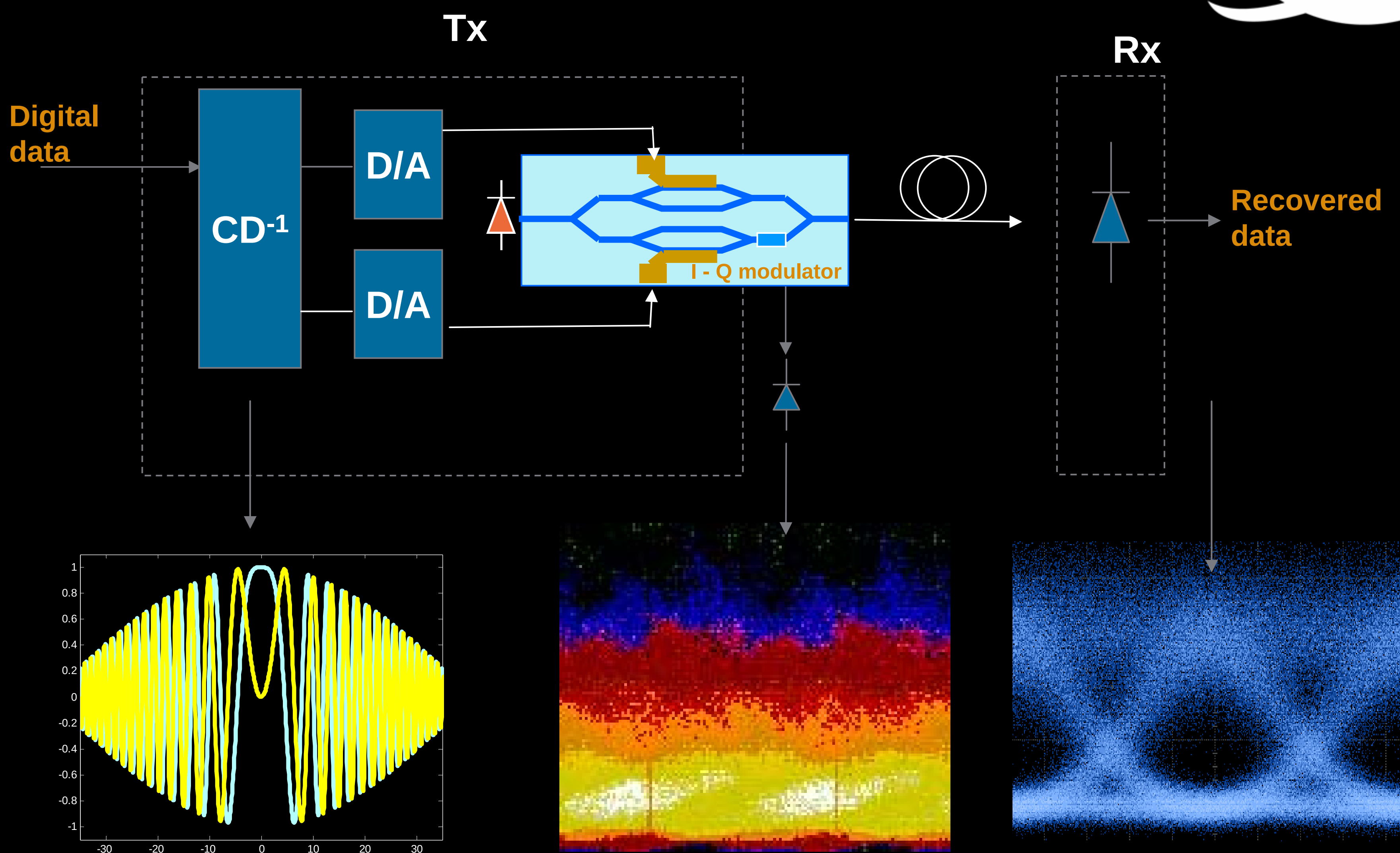
$

FIXED ADD-DROP

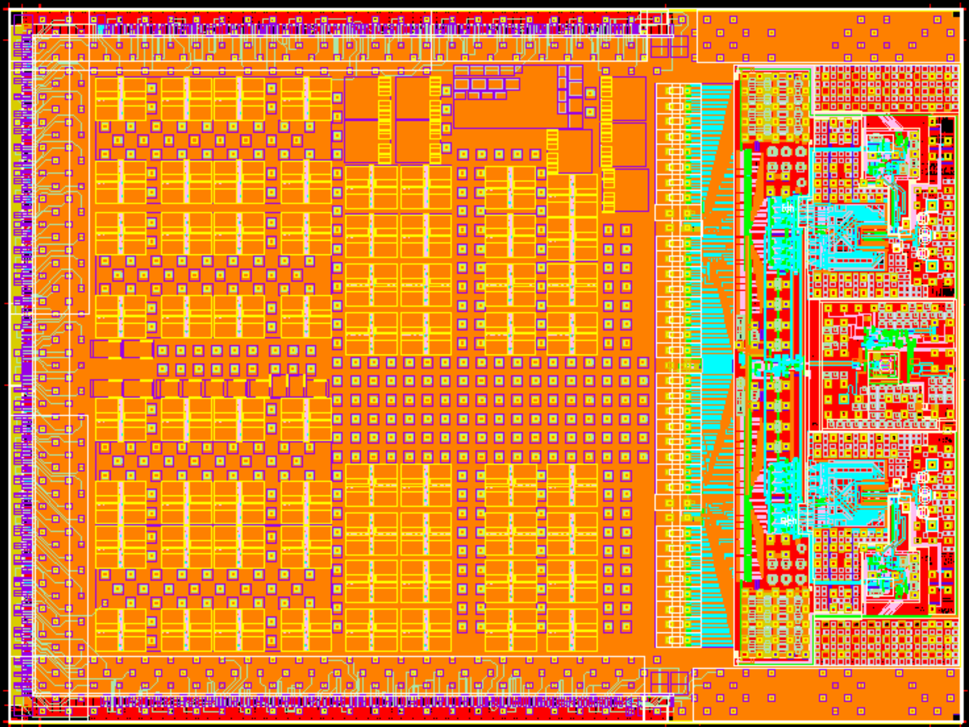
Simplified Optical Connectivity



Transmitter Pre-compensation



Dispersion Eliminator: WARP ASIC

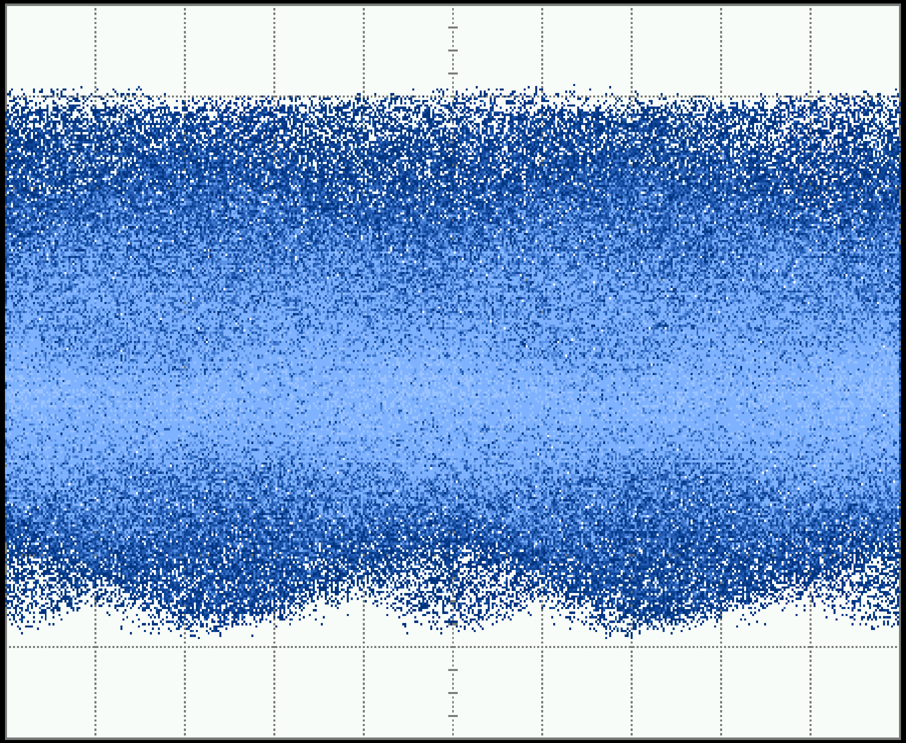


- 0.13 μm BiCMOS
- 2.5 Mb High Speed Memory
- 2.0 M Gates
- 6 T Ops per second
- Two 20 Gs/s 6 bit DAC
- 676 HTCE BGA - 14 layer

Linear and nonlinear precompensation of 10 Gb/s.

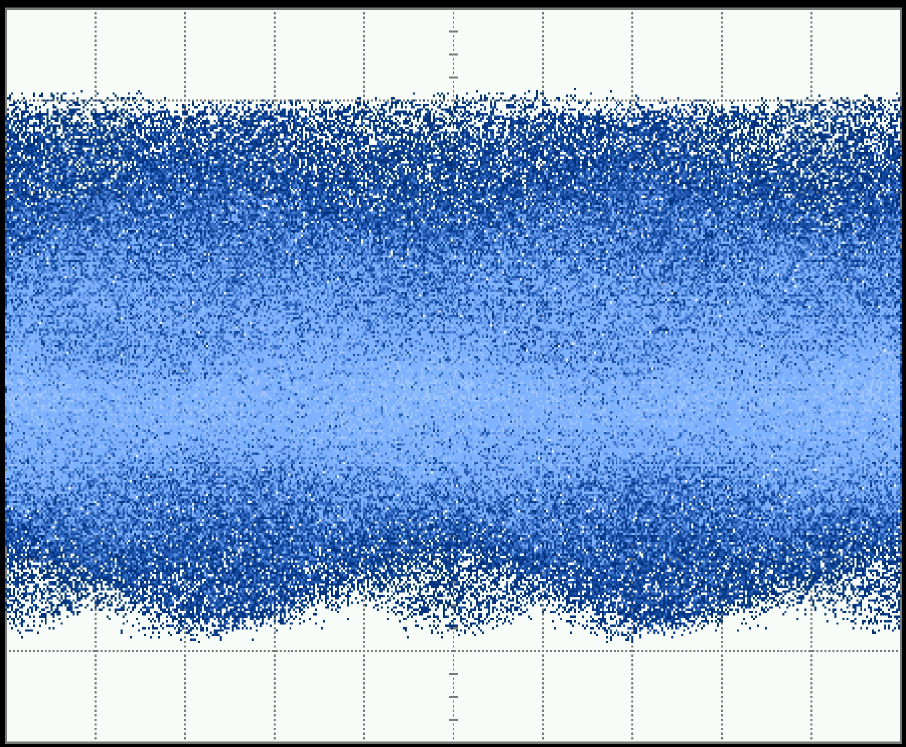
Silicon eliminates the optical channel degradations.

Electronic Dispersion Compensation
320 km of G.652 fiber, 2.5 dBm launched

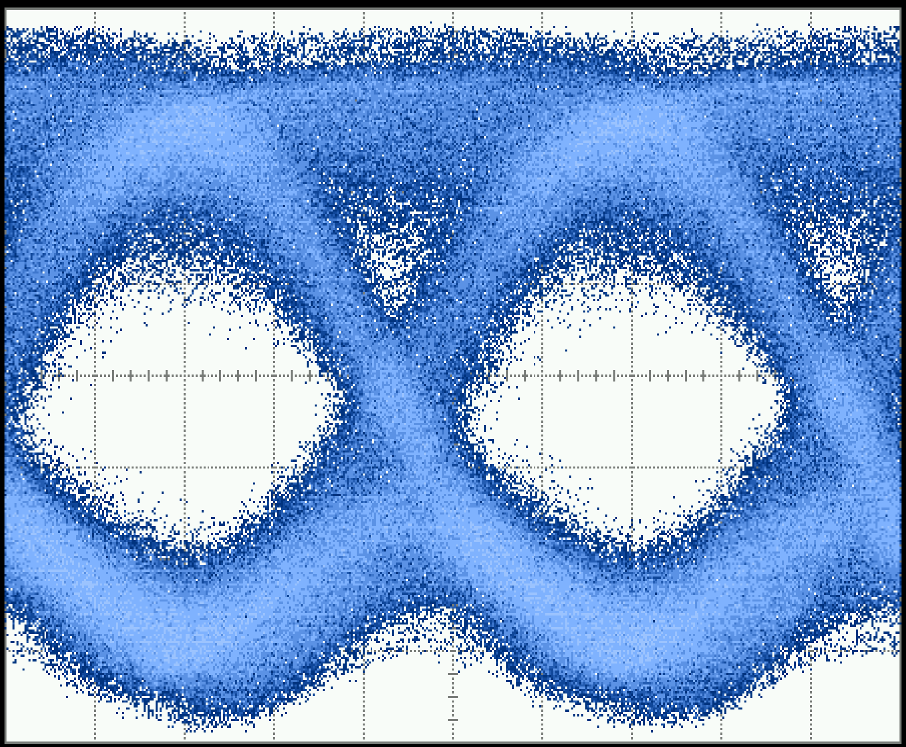


No Compensation

Received Eye Diagram
320 km of G.652 fiber, 2.5 dBm launched

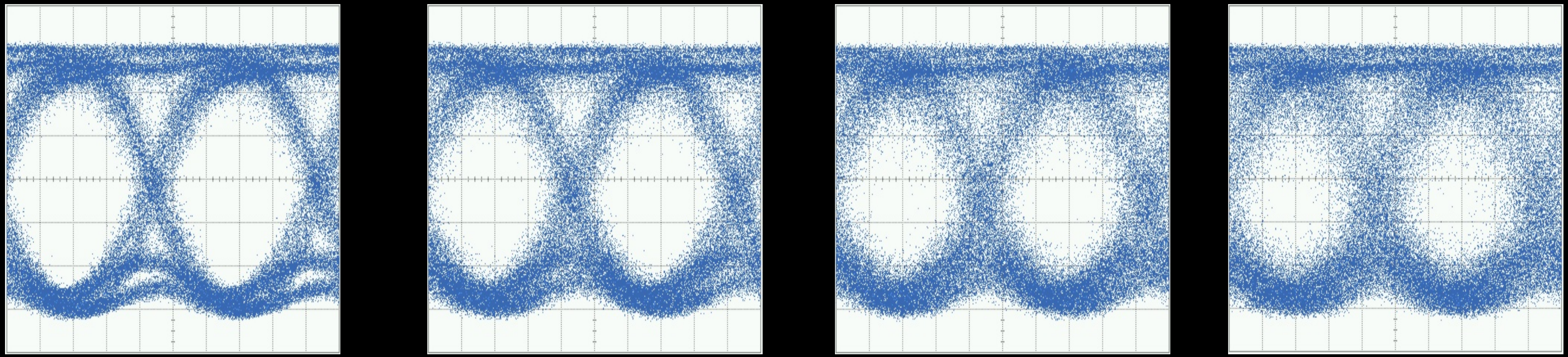


No Compensation



WARP Compensation
Overcomes Dispersion

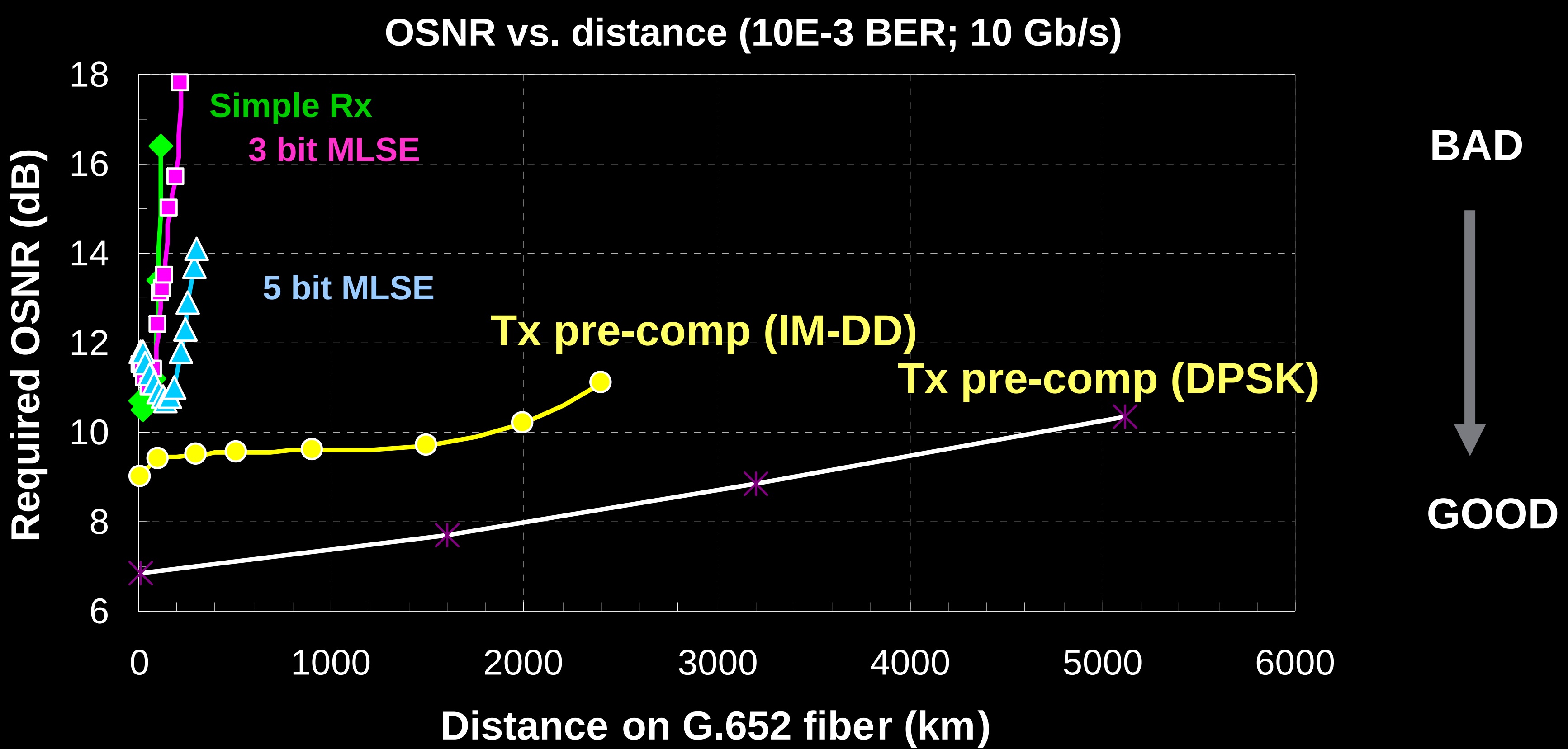
10 Gb/s
with no Traditional Dispersion Compensation



0 Km | 1600 km | 3200 km | 5120 km

Eye diagrams after transmission over standard G.652 fiber with Nortel WARP processing.

Electronic Dispersion Precompensation



Optical Dispersion Compensation is Obsolete

Forward Error Correction

Block Code

Information sequence: k bits

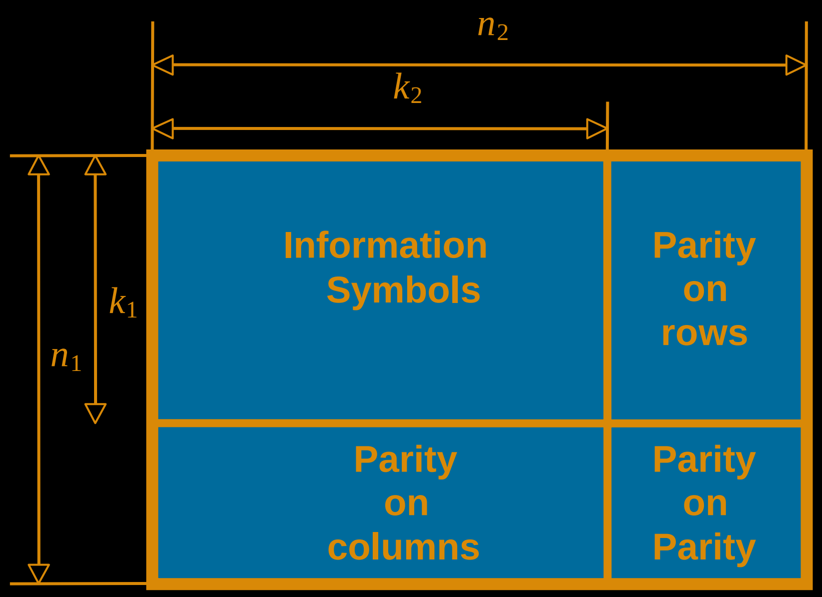
Codeword: n bits

Rate: k / n

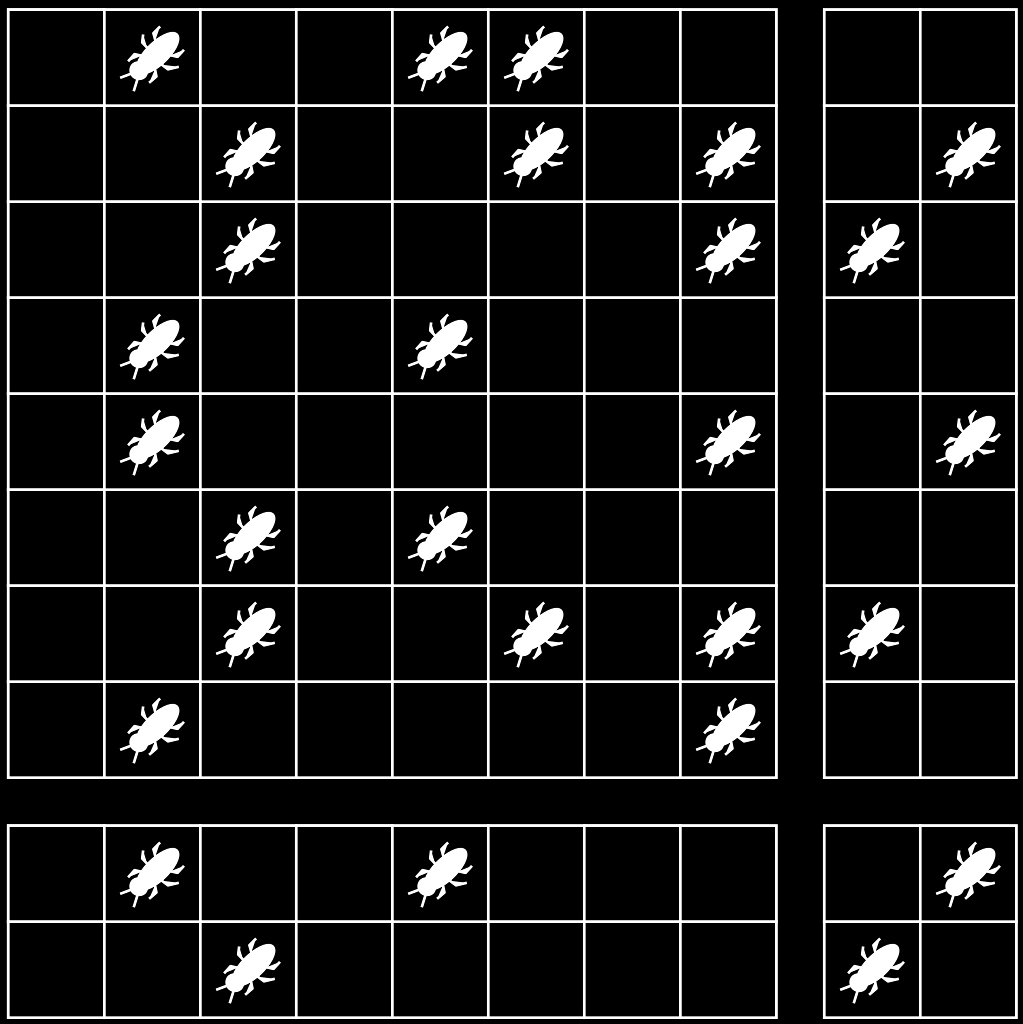
Bose-Chaudhuri-Hocquenghem (BCH) Codes

- For any positive integers m and $t < n/2$, there is a binary BCH code of length $n = 2^m-1$ which corrects all combinations of t or fewer errors and has no more than mt parity check symbols.

Product Code

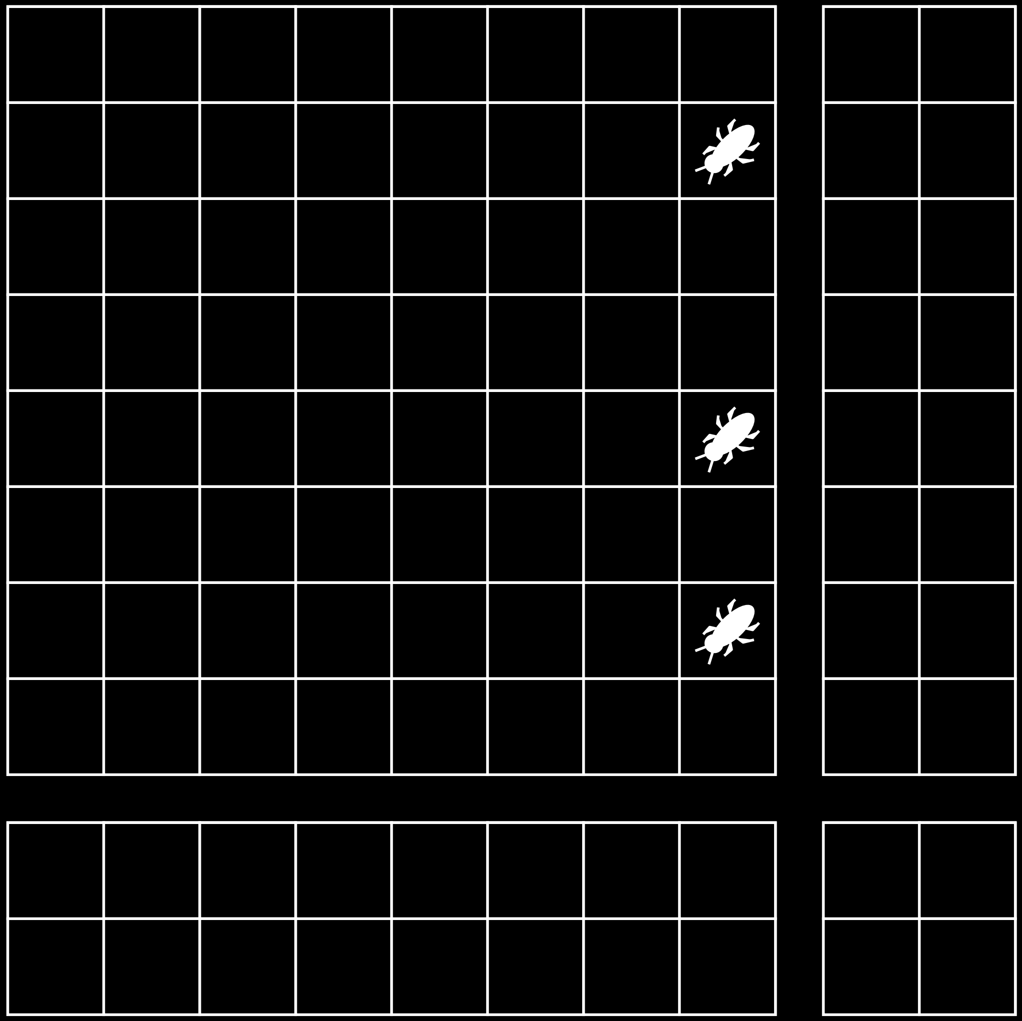


Two-dimensional Iterative decoding

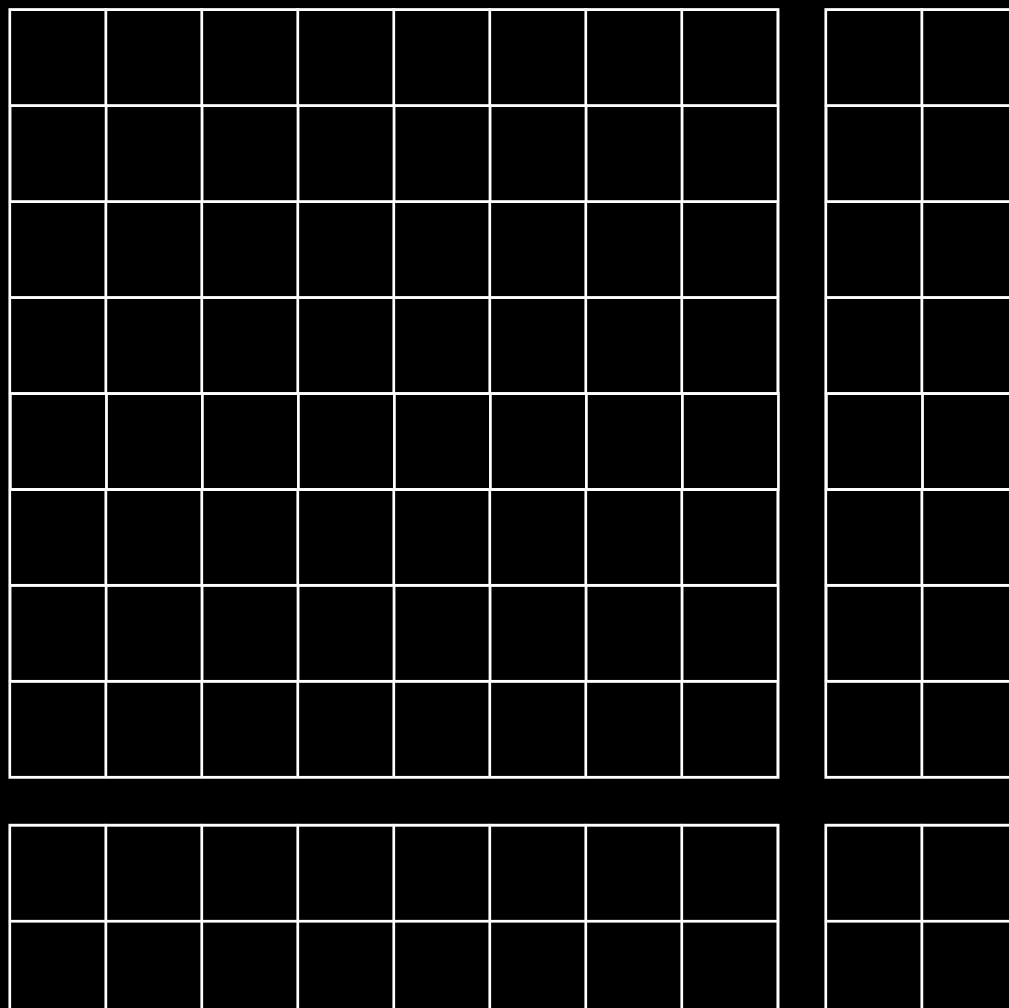


> Animation deleted

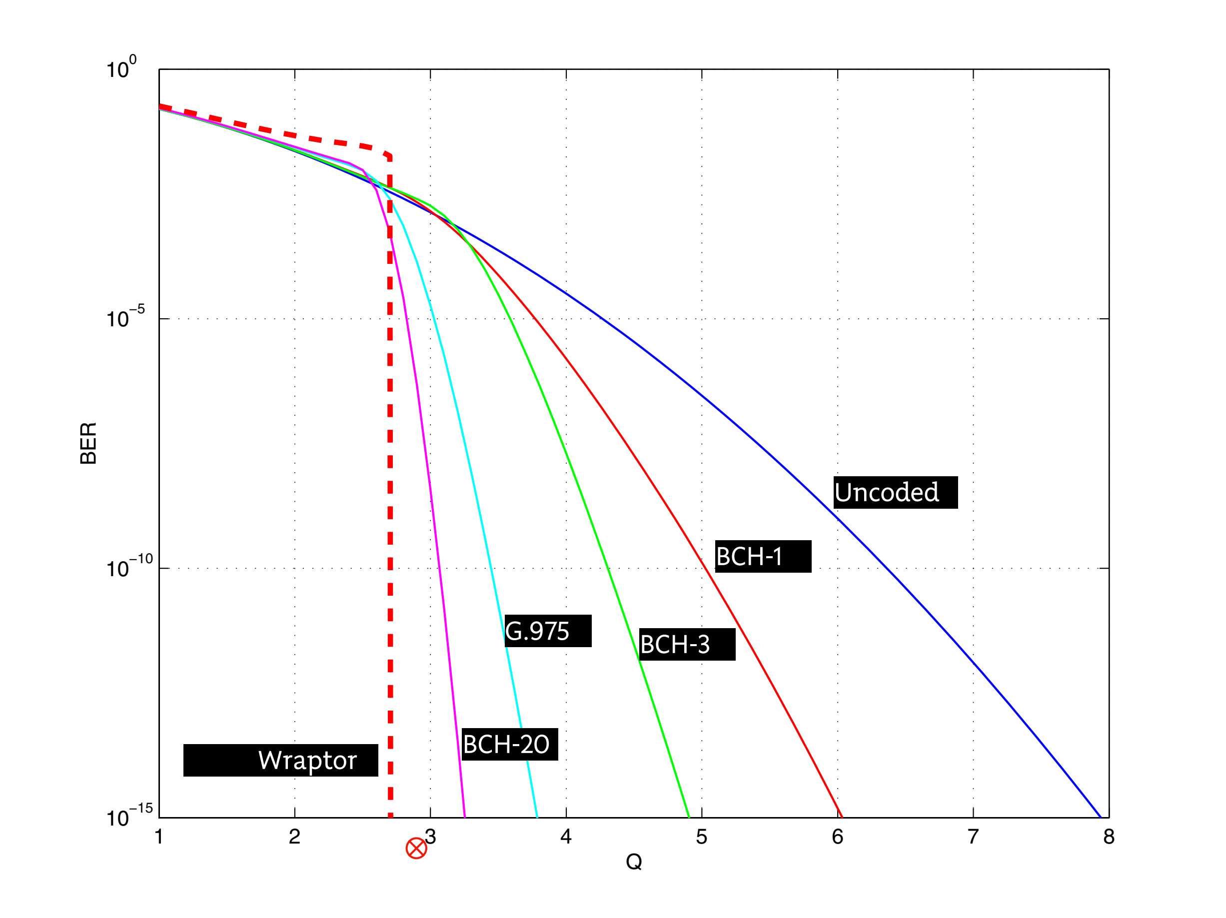
Two-dimensional Iterative decoding



Two-dimensional Iterative decoding



FEC Evolution in 10G Product



Forward Error Correction

- If systematic redundancy is added to the data, one can correct errors that later occur.
- Does not require retransmission, so minimal delay at full line rate.
- Correcting bit errors allows clean operation despite large amounts of optical noise, producing system gain.
- Can expect random packet error rates ~ 10^{-14}

Rate of Packet Loss

> FEC in the line and FEC between each chip eliminates random errors.
  - ~10^{-14} packet loss is once per (mathematical) millennium at 10G
  - We measured zero errors in a month under constant worst case optical noise conditions

> Layer 1 lightpath eliminates packet loss or reordering due to routing or packet switching.

> Packets can still be lost due to equipment failure or fiber cuts.
  - HDX has an availability of 99.9999%
  - Typical fiber route has 2 cuts per year per 1000 km of buried cable, with an 14 hour mean time to repair
  - Equipment can be designed for hitless protection switching, or for 50 ms protection.

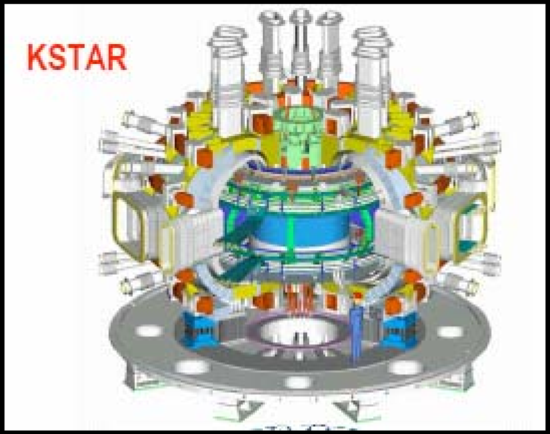
Latency and Jitter

> Minimize the physical path length
  - no satellite connections, use a direct fiber route.
  - no dispersion compensating fiber coils, at 50 μs per coil.

> Minimize the Buffering
  - Layer 1 lightpath, with no packet queues.
  - FEC rather than ARQ, so no ARQ buffer delays

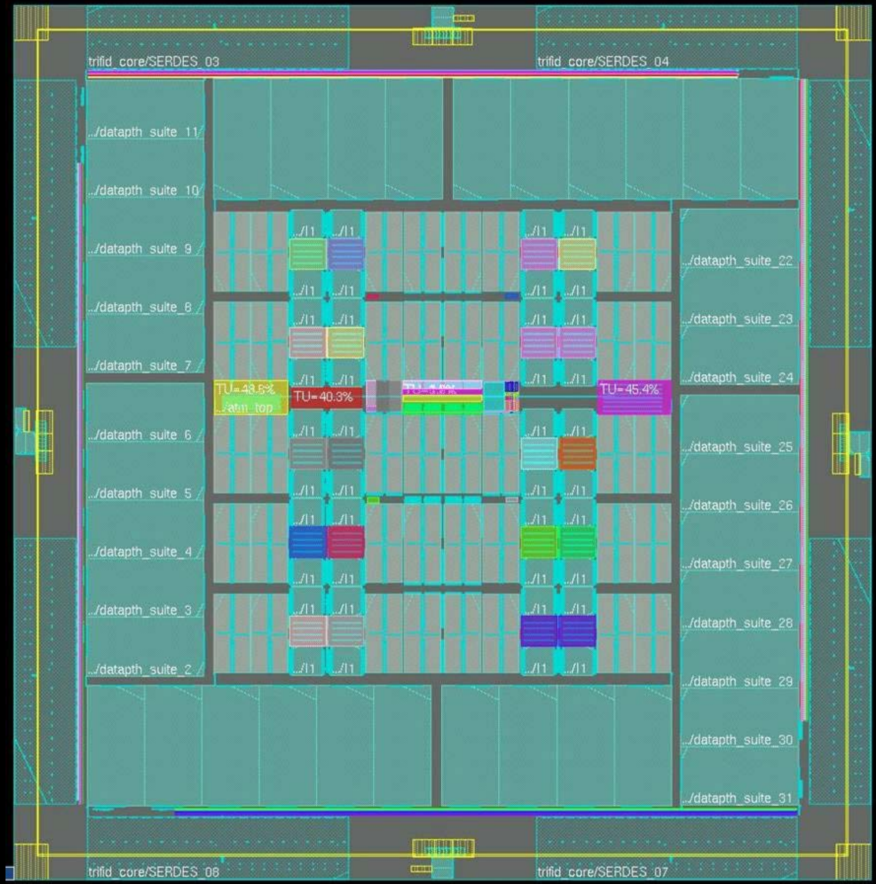
> Minimize the Jitter
  - Physical layer jitter is minimal
  - No queuing jitter

> How can you exploit a lightpath with minimum, constant, latency?



How do I send a large data set to several destinations?

400 Gb/s Layer 1 switching ASIC



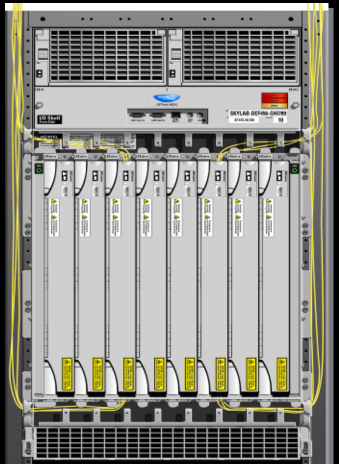
- One byte from each of 8000 inputs written into high speed 9 port RAM, every 160 ns.
- For each output, software has chosen which RAM location to constantly read.
- Any number of outputs can have chosen the same input.
- 50 Mb/s granularity
- Parallel connections build high capacity, e.g. 192 = OC-192
- Constant latency, no errors
  - Jitter < 1 ps
  - FEC on all inputs and outputs

These chips allow Layer 1 Multicast

One layer 1 switch:
- 50 Mb/s 70,000 ways,
- 240Gb/s streams 15 ways

HDX
Up to 1.28 Tb/s single shelf
3.84 Tb/s as a hub

HDXc
Up to 640 Gb/s



OME
Up to 160 Gb/s

What can you do with arbitrary Multicast?

For example: source from CERN, multicast in Amsterdam
One of those links across the Atlantic, multicast in Chicago
One of those links to Seattle, multicast to the Pacific.

- > **Multicast 10G or 40G one direction and 50M point-to-point return.**
- > **Multicast both directions.**
- > **Multicast one way without high speed return path.**

What about the Ends of the Lightpath?

> Mapping packets into the Lightpath without losing quality
  - Layer 0 Lightpath is optical end to end,
    - 10GE WAN or LAN.
  - Layer 1 Lightpath can traverse carrier networks such as transoceanic, and SONET/SDH connections.
    - GFP in STM-64/OC-192, or 10GE WAN

> What if the user chooses to share the Lightpath?
  - Time of day windows
    - optical or electrical switching of the full lightpath
  - Subdividing Layer 1 capacity
    - GFP mapping into STS-n
  - Layer 2 packet switching
    - provisioned pt-pt VLAN

GFP

> Generic Framing Protocol

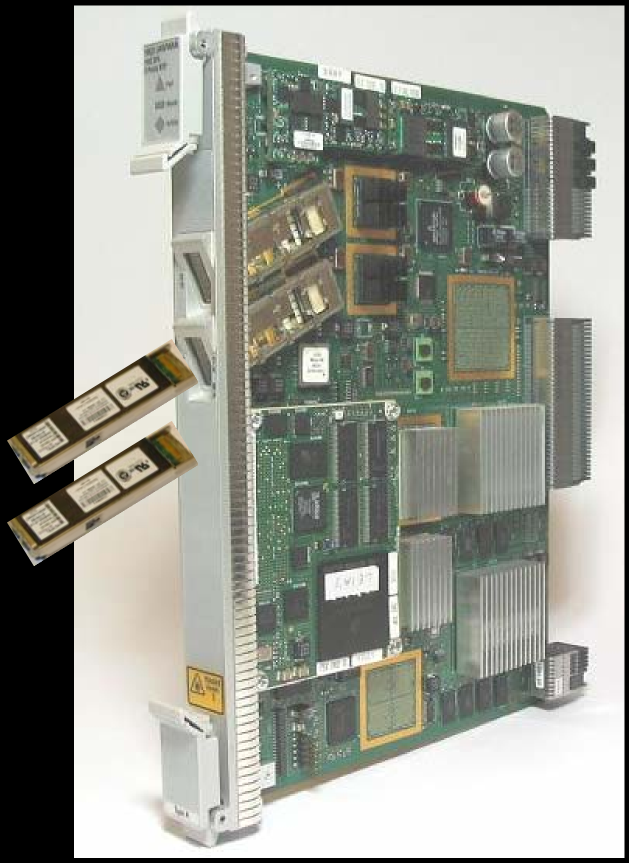
> Efficiently maps packets into and out of layer 1.

> Does not degrade quality of service.

> Do the mapping at the user site.

- allocate the share of the bandwidth
- then use layer 1 switching,
- demap at the destination site.

GFP mapping
2 x 10G into Layer 1 Lightpaths.



> Map GigE or 10GE into SONET/SDH using GFP.

> Alternative to 10GE WAN
- allows LAN interface prices
- better throughput than WAN
- not quite full LAN throughput

Frame Size	WAN	GFP
64	101.54%*	105.54%*
128	97.11%	101.32%*
256	94.61%	99.71%
512	93.27%	97.31%
1024	92.57%	96.59%
1280	92.53%	96.44%
1518	92.35%	96.35%
9216	91.94%	95.93%

10G Throughput

Bigger Pipes: 40G

> A successful 40 Gbps solution:
  - Completely compatible with a 10 Gbps line design
    - Use same deployed amplifier system as10Gbps
    - Same reach
    - Have same OADM and Branching capability
    - Have PMD/PDL specifications at least as tolerant as 10Gbps
    - Equal or better OAM&P metrics to those of 10 Gbps
    - Use same terminal platform as 10Gbps
  - Cheaper than 4 X 10 Gbps

> What will you do with 40G pipes?

Lightpath Channel Model

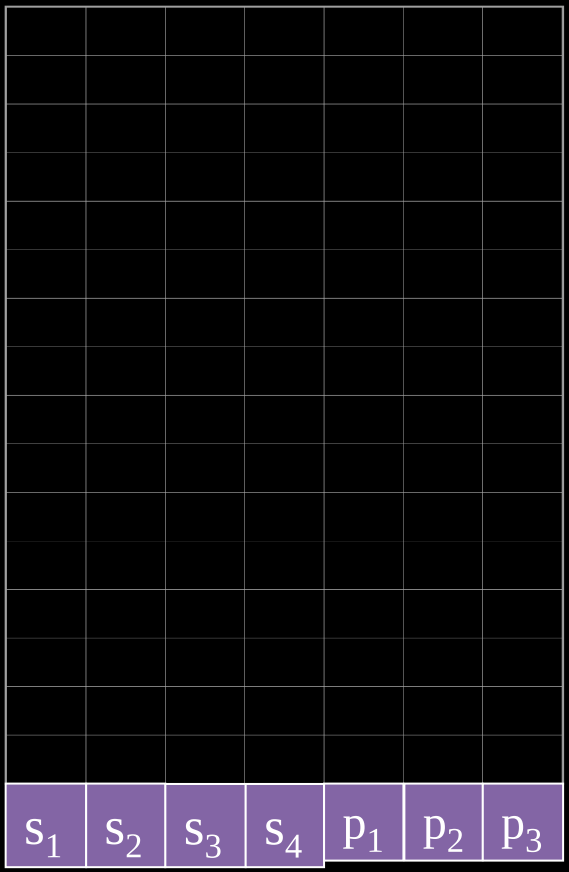
- Latency almost solely due to the speed of light in the shortest possible fiber connection and effectively zero packet jitter.
- No congestion, no packet reordering, guaranteed capacity.
- $\sim 10^{-14}$ probability of packet corruption from bit errors.
- The required level of availability.
- An engineered bound on the duration of interruptions.
- 10Gb/s to 40Gb/s of capacity per channel.
- Multicast connectivity to every required destination.

Can you fully exploit these Lightpaths?

Backup

Linear Block Code

Hamming Single Error Correction (7,4,3) code



Code(n,k,d)

$t = \lfloor d/2 \rfloor$

<u>Generated Parity bits</u>

$P_1 = S_1 + S_2 + S_3$

$P_2 = S_2 + S_3 + S_4$

$P_3 = S_1 + S_2 + S_4$

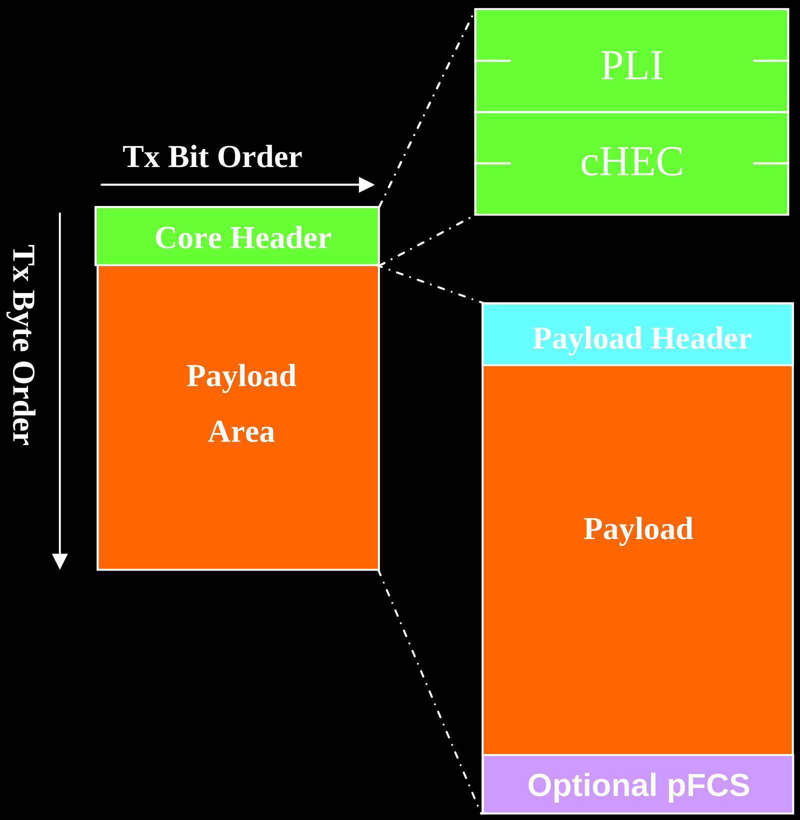
<u>Hamming distance</u>

is

number of places in which

any two code words differ

GFP Overview



- Simple & Scalable Protocol
  - Operates 1.5M to 10G
  - Scalable beyond 40G
- Unifying Transport mechanism
  - Currently supports
  - 802.3 Ethernet MAC
  - 8B/10B encoded clients
    - Ethernet, Fiber Channel, HDTV
    - Extendable for scalability
  - IP/PPP in HDLC like Frames
  - Easily extended for other protocols
- No Interference with QOS
- More efficient than legacy ATM, POS schemes
- In built OAM channel for Telemetry
- Standards Based
- Rollout on every Transport vendor's Silicon
- Works with Virtual Concatenation, LCAS

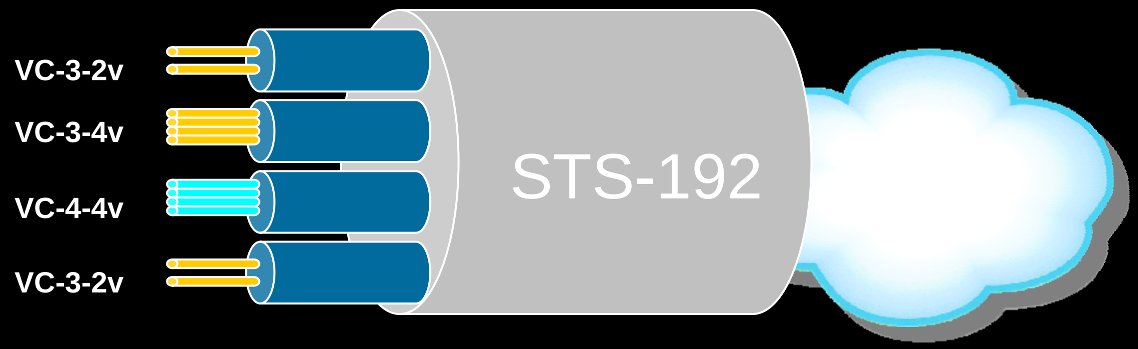
Virtual Concatenation (VCAT)

> "Right-sizes" the provisioned SONET path for the client signal
- Enables mapping into an arbitrary number of standard STSs
- Transport capacity decoupled from service bandwidth – less stranded bandwidth

> VCAT can be diversely routed through SDH/SONET network

> Recombined to contiguous payloads at end point of transmission
- Need to handle differential delays at egress due to diverse routing
- Do this using internal buffers – 5us/km of differential fibre

> Interworks with all existing SDH/SONET equipment
- Only source & sink terminals need to support VCAT



LCAS – Link Capacity Adjustment Scheme

> Efficiently offer dynamically-allocated bandwidth

> Hitless increase or decrease in capacity of a VCAT Link (or Virtual Concatenated Group, VCG) by adding/removing VCs

> Provides a soft protection & load-sharing mechanism

- Automatically decreases the link capacity if a VC path experiences a failure
- Automatically increases the link capacity when the network fault is repaired

> Provides an extra level of network and service resiliency

- Facilitates support of Service Level Agreements, especially in the area of network and service restoration
- Hitless bandwidth expansion and contraction reduces service interruptions in the event of network failure
- Eases network operations and maintenance actions