Reliability considerations for parallel 100G carrier-grade Ethernet transport

IEEE HSSG Meeting, Monterey, CA – January 17-19th, 2007

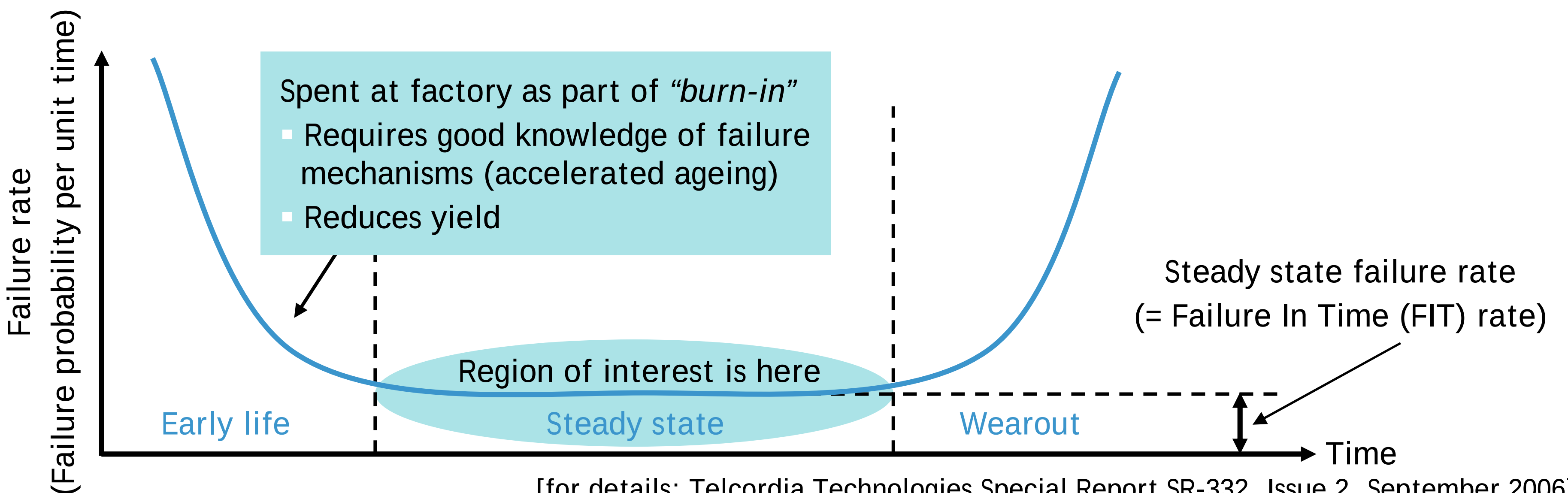
Peter J. Winzer, Marcus Duelk
Data Optical Networks Research / Bell Labs / Alcatel-Lucent
winzer@bell-labs.com

Outline

1. Reliability and its quantification
2. Wavelength redundancy for parallel PHYs in carrier-grade 100G Ethernet
3. Conclusion

What is reliability

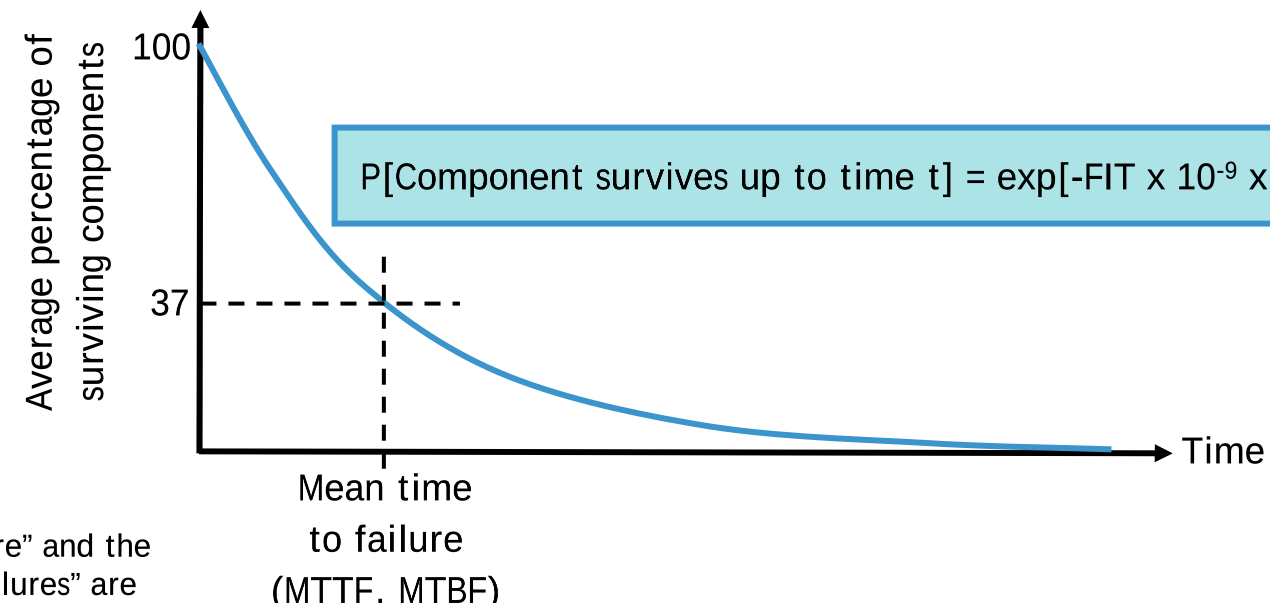
- Reliability quantifies, in a statistical sense, the failure behavior of devices, components, and systems as a function of time
- **"Reliability"** should not be confused with
  - **"Yield"**, which quantifies the fraction of satisfactory components at the time of manufacturing
  - **"Qualification"**, which states that a device, component, or system has passed a certain series of tests at a single (usually beginning-of-life) point in time.
- Typical failure behavior is described by a bathtub curve:



[for details: Telcordia Technologies Special Report SR-332, Issue 2, September 2006]

Steady state failure behavior

- **Failure in time (FIT) rate** describes steady state failure probability
  - FIT is normalized to 10^9 hours
    → 1 FIT = 1 failure in 10^9 hours
    → Probability that a component fails in any one steady state hour = FIT x 10^{-9}
- Constant FIT rate results in an **exponentially decaying** expected value of surviving components
- The **Mean Time to[1] Failure (MTBF, MTTF)**, expressed in hours, is given by 10^9/FIT.



[1] The "Time **to** failure" and the "Time **between** failures" are statistically identical.

[for details: Telcordia Technologies Special Report SR-332, Issue 2, September 2006]

Alcatel·Lucent

How to obtain FIT rates (I)

- FIT rates of **components** and **devices** are typically obtained through accelerated ageing
  - Requires good knowledge of the dominant failure mechanisms and how it can be accelerated (it's not always temperature or current!)

- FIT rates of **subsystems** are typically obtained by summing up the FIT rates of their components
  - If subsystem failure is caused by any one component failure and if
  - Component failures are statistically independent of each other

Example: Determination of the FIT rate of an optical transponder

Parts	Count	FIT Rate
Resistor	100	50
Capacitor	50	50
IC	1	30
…		
EEPROM	1	10
Laser driver	1	30
Laser	1	100
Diode+TIA	1	50
TRANSPONDER FIT RATE		**320**

Note: This kind of **FIT rate budgeting** is also done on a sub-system and system level (i.e., for full line cards and WDM systems), where a transponder is one item in the list

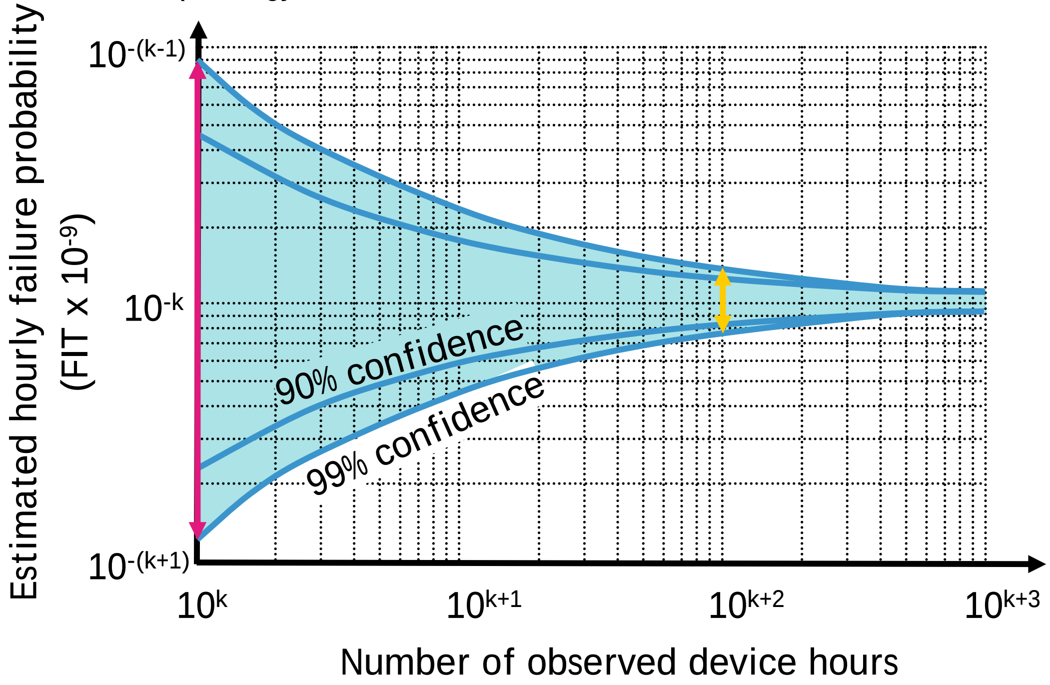
→ It is important that FIT rates of 100G transponders/transceivers are comparable to 10G counterparts !

How to obtain FIT rates (II)

- It is typically **not** enough to gather statistical data just from in-field failure observations
  - **In-field components** are not operating at their worst-case allowed limits
  - There is typically **not enough statistical information** available

Example:

If one tests components for 10 million device hours and finds no failure, one could claim a FIT rate better than $10^9/10^7 = 100$. However, such an estimate is highly unreliable, as quantified in the figure below (with k=7): An estimate based on a single failure (10^7 observed device hours at FIT x $10^{-9} = 10^{-7}$) is associated with an **uncertainty of almost an order of magnitude** at a 99% confidence level, as indicated by the pink double arrow. To get meaningful statistics, one needs to record at least 100 failures, as shown by the yellow double arrow. (Analogy to bit error measurements: One needs at least 100 errors to reliably measure a BER.)



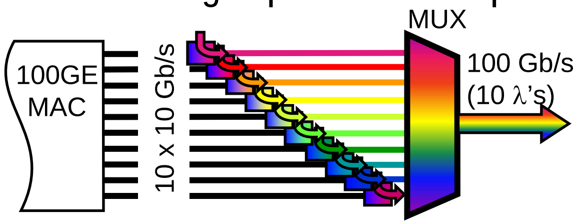
[Figure after Jeruchim et al, Simulation of Communication Systems, Kluwer Academic / Plenum Publishers, 2000]

Outline

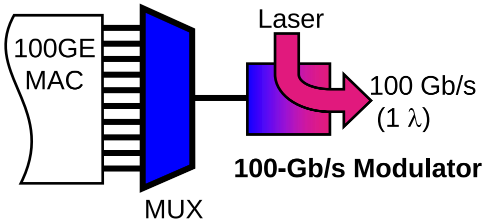
1. Reliability and its quantification
2. Wavelength redundancy for parallel PHYs in carrier-grade 100G Ethernet
3. Conclusion

100G parallel Ethernet transport

N-wavelength parallel transport



1-wavelength serial transport



10 x 10 Gb/s, 5 x 20 Gb/s, 4 x 25 Gb/s, 1 x 100 Gb/s

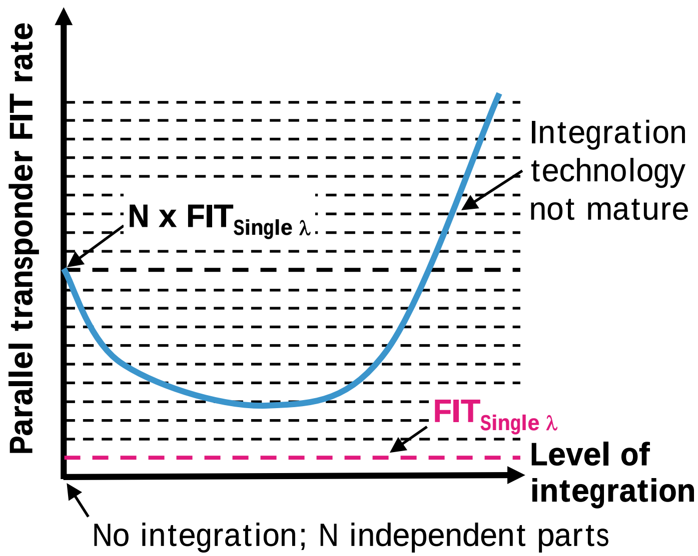
- To first order, the choice of bit rate per wavelength is driven by complexity and cost of high-speed opto-electronics
- Once high-speed components mature, serial transport has historically been lower cost

- An additional argument: **Reliability of parallel transport for carrier-grade Ethernet**
- **Full interruption** of 100G Ethernet link if only a **single wavelength** fails
  - Transmitter failure
  - Receiver failure
  - Wavelength multiplexer / demultiplexer failure
  - Link failure on any one wavelength (optical mis-routing of one wavelength)

Carrier-grade reliability is critical for 100G carrier-grade Ethernet, in both LAN and WAN

FIT rates of parallel transponders

- Failure rate of parallel transponders depends on **dominant failure mechanisms**
  - Failure of components that are common to all channels
  - Statistically independent failure of individual components
  - Statistical dependence between component failures
- If failures are statistically independent? **→ $FIT_{N\ wavelengths} \sim N \times FIT_{Single\ wavelength}$**
- Failure rate also depends on the **level of integration**
  - In *electronics*, integration typically reduces the FIT rate of the full system ($FIT_{Single\ component} < FIT_{N\ integrated\ components} < N \times FIT_{Single\ component}$)
  - Not necessarily same scaling rules for *opto-electronics* ($FIT_{N\ wavelengths} > N \times FIT_{Single\ wavelength}$ can be possible, if underlying integration technology not mature enough)
  - Until one understands and quantifies reliability and failure mechanisms, one cannot be sure of the FIT rate due to integration
    - Have to make a leap of faith
    - Risk management with limited input information
    - Can this be acceptable for carrier-grade Ethernet ?



Alcatel·Lucent

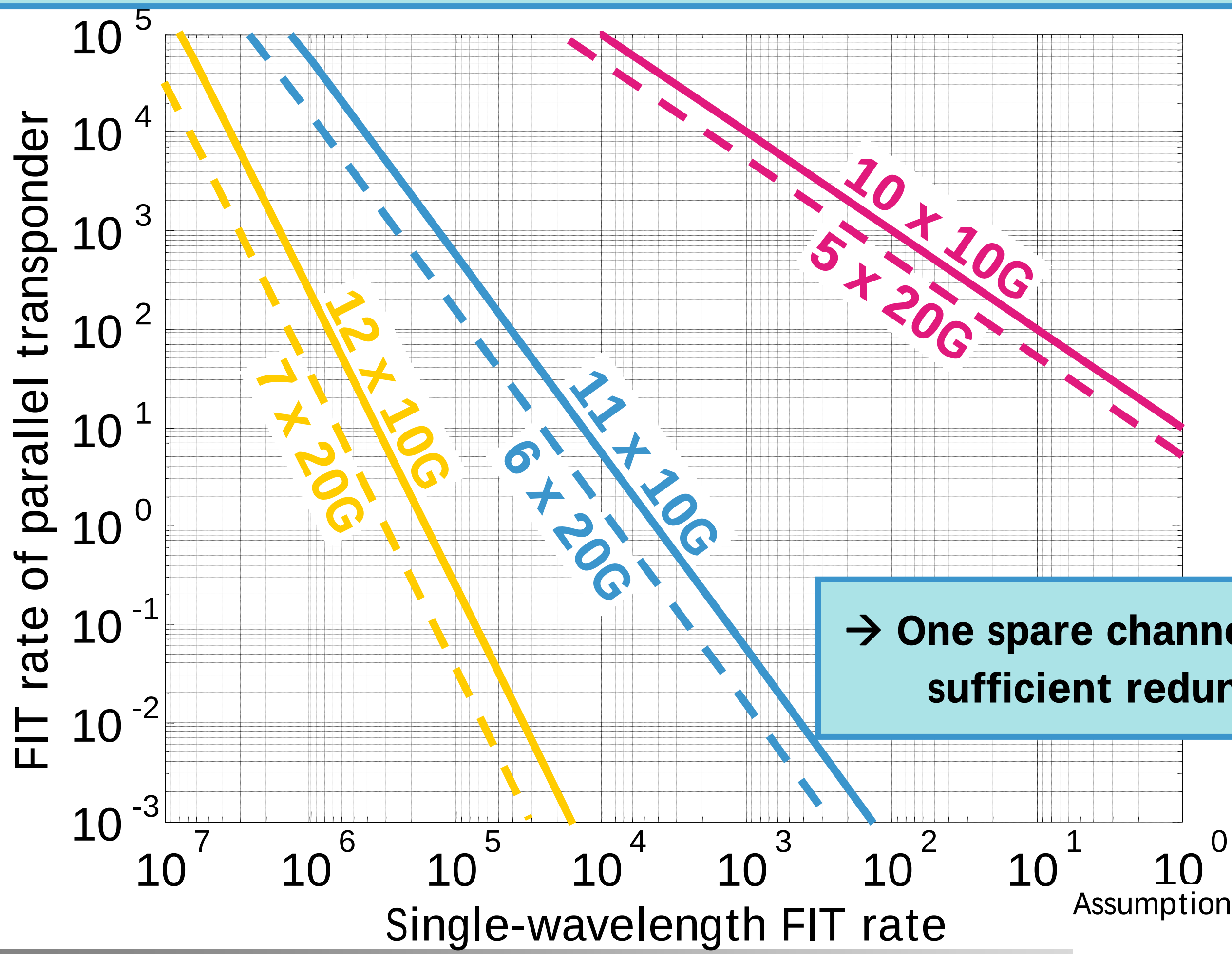
Reduction of FIT rate through redundancy

- Unprotected transponder operates at N x R Gb/s = 100G Ethernet rate
- Including k more wavelengths for redundancy: (N+k) x R Gb/s
- Mapping from N to N+k channels can be done in many ways, e,g.,
  - Within the MAC device on a packet basis
  - Within the PMD device on a fail-over switch basis
  - Through forward error control on a statically coded basis
  - …

Alcatel·Lucent

How much redundancy is needed

Reduce failure probability by including one or more channels for redundancy
(If any one channel fails, the spare channel(s) take over)



→ One spare channel provides sufficient redundancy

Assumption: Channels fail independently
(i.e., $FIT_{total} = N \times FIT_{single}$)

Alcatel·Lucent

Conclusion

- **Carrier-grade Ethernet needs carrier-grade reliability**

- Multi-channel transponders can have worse reliability than single-channel transponders
  - Parallel transponders must meet the FIT budget on a system level
  - A 100G linecard must have the same reliability as at 10G (or better)
  - FIT rates of parallel transponders depend on
    - Level of integration
    - Maturity of integration technology

  - Integration of various kinds may help, but **reliability needs to be proven**
    - Need to have good understanding of failure mechanisms
    - Need to perform accelerated ageing with the correct ageing factors
    - Not possible to deduce FIT rates from limited in-field observations

- **Introducing redundancy** eliminates reliability problems
  - One extra channel provides sufficient redundancy

Backup

Calculating the FIT rate of a parallel transponder

- Assuming *statistical independence* of individual channel failures

  N … Number of working channels
  k … Number of spare channels
  FIT … FIT rate of a single channel

$$P[\text{failure}] = \sum_{i=N+1}^{N+k} \binom{N+k}{i} \text{FIT}^i (1-\text{FIT})^{N+k-i} = 1 - \sum_{i=0}^{k} \binom{N+k}{i} \text{FIT}^i (1-\text{FIT})^{N+k-i}$$