Science

MAAAS

Filling the Light Pipe
David J. Richardson
Science **330**, 327 (2010);
DOI: 10.1126/science.1191708

more to the genetic variation between the two strains than the centers. At first glance, this result is paradoxical, because the centers have much higher gene densities than the arms; where there are more genes, one would expect to detect many more variant loci that regulate gene expression. This paradox, however, can be explained by another difference between the two regions: The arms also have much higher rates of genetic recombination (exchange of genetic material between the maternal and paternal genomes of an individual) than the centers. Investigators have found a similar pattern in other organisms, including *Drosophila*; in these organisms, genome regions with higher levels of genetic recombination also have greater DNA sequence variability (*3*).

The process producing this pattern is likely to be genetic hitchhiking, in which changes in allele frequencies at sites that are under selection affect the frequencies of variants at nearby sites on the chromosome. In other words, a variant present at one place in the genome can spread through a gene pool because it "catches a ride" with a closely linked DNA sequence variant that is under selection. Two forms of hitchhiking may be important: "selective sweeps" caused by the spread of variants favored by selection (*4*); and "background selection" (*5*), which involves the elimination of rare del-

eterious mutations from the population (see the figure). Other things being equal, genome regions with low levels of recombination are the most likely to experience hitchhiking effects, resulting in a reduction of variation within regions close to the sites under selection. Because of this effect of hitchhiking, the effective population size (N_e) in these regions (a measure of the number of individuals in the population that contribute to the next generation) is reduced. Thus, the combination of linkage and low recombination reduces both the variability of the region and its potential for adaptation, because levels of variability within a population and the efficacy of selection are both proportional to N_e (*6*).

Rockman *et al*. investigated whether background selection could explain the variation in levels of gene expression between the chromosome centers and arms. They used information on genomic parameters such as gene densities and recombination rates across the *C. elegans* genome and also estimated the values for two unknown variables: the intensity of selection against deleterious mutations and the level of inbreeding in the population (*7*). The inbreeding parameter is important for *C. elegans* because it is a hermaphrodite with a high frequency of reproduction by self-fertilization. The resulting lack of genetic differences between the paternal and maternal genomes of an individual greatly reduces the

effectiveness of recombination.

The results are consistent with a major role for background selection, but contributions from selective sweeps cannot be ruled out. A limitation of the study is that it is based on examining differences between only two strains, one of which has been in laboratory culture for many years.

Nonetheless, the analysis convincingly shows that hitchhiking effects are the most likely cause of the observed patterns of variation and provides the first evidence that hitchhiking can shape quantitative variability. It will be interesting to see whether similar patterns are observed in more representative samples from the population and for DNA sequence variants as well as quantitative variability. It is also important to seek for similar patterns with respect to quantitative trait variation in other systems, such as human populations.

References
1. J. Flint, T. F. C. Mackay, *Genome Res.* **19**, 723 (2009).
2. M. V. Rockman, S. S. Skrovanek, L. Kruglyak, *Science* **330**, 372 (2010).
3. W. Stephan, *Phil. Trans. R. Soc. B.* **365**, 1245 (2010).
4. J. M. Smith, J. Haigh, *Genet. Res.* **23**, 23 (1974).
5. B. Charlesworth, M. T. Morgan, D. Charlesworth, *Genetics* **134**, 1289 (1993).
6. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).
7. R. R. Hudson, N. L. Kaplan, *Genetics* **141**, 1605 (1995).

APPLIED PHYSICS

Filling the Light Pipe

David J. Richardson

As data transmission rates continue to increase, optoelectronic engineers are developing techniques to deal with the approaching "capacity crunch."

It has been a landmark year for the field of optical telecommunications, with the award of the 2009 Nobel Prize to Charles Kao for his insight in the mid-1960s that the future of communications lay in single-mode silica-based optical fibers (*1*) as well as the 50th anniversary of the first demonstration of the laser—both key technologies responsible for the development of the global-scale communications networks of today (*2*). Recently, however, a growing realization has emerged within the telecommunications industry that the end of the phenomenal growth in optical fiber communication capacity is within sight. At this year's Optical Fiber Communication Conference (OFC 2010), several groups

Optoelectronics Research Centre, University of Southampton, Southampton SO17 1BJ, UK. E-mail: djr@orc.soton.ac.uk

reported results within a factor of ~2 of the ultimate capacity limits of existing optical fiber technology. Without radical innovation in our physical network infrastructure—that is, improvements in the key physical properties of transmission fibers and the optical amplifiers that we rely on to transmit data over long distances—we face what has been widely referred to as a "capacity crunch" that could severely constrain future Internet growth, as well as having social and political ramifications.

Most of the information we exchange over the Internet is carried through optical fibers encoded on a beam of laser light. Data traffic on the world's networks is growing at around 40% year-on-year (*3*). This growth is driven primarily by social networking, cloud computing, and bandwidth-hungry video services such as YouTube. Video is responsible for most

of the increased demand, and with high definition firmly established in the marketplace and both 3D and ultrahigh-definition formats emerging, this trend is set to continue.

The data-carrying capacity of a single optical fiber is determined by the spectral bandwidth over which suitably low-loss signal transmission can be achieved and our technical ability to exploit this bandwidth through suitable data coding and decoding schemes. The bandwidth of current fiber systems is limited to around 11 THz by the optical amplifiers used to boost the power of flagging signals attenuated by fiber transmission loss. Until relatively recently this bandwidth was considered to be effectively infinite, and the primary issue was one of developing cost-effective means to keep pace with the increasing traffic demands. Key to this has been the development of a technique referred to as
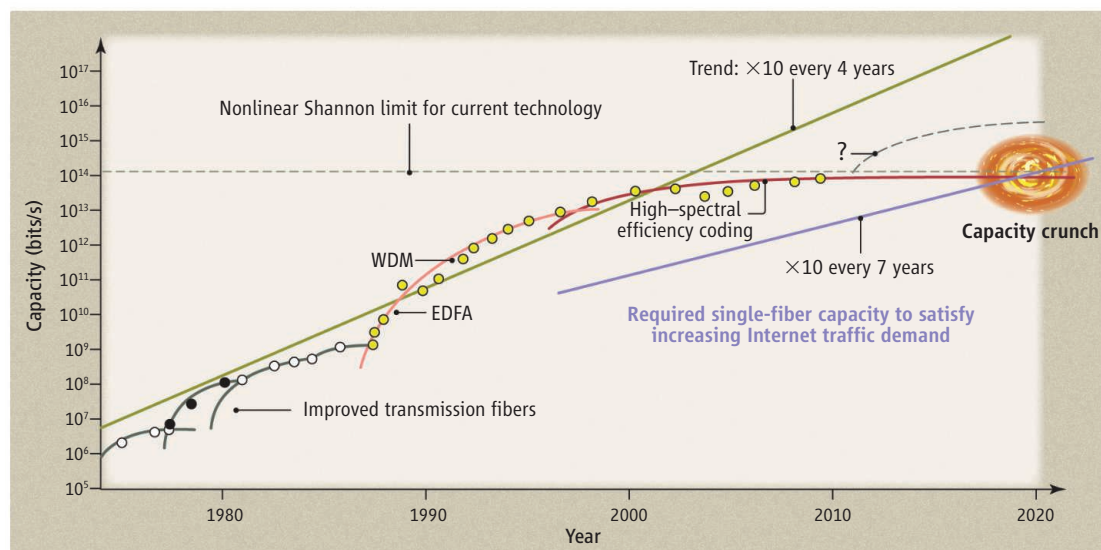
And on the data flows. The figure shows both the record data transmission capacity of fiber transmission systems (numerical values scaled to be representative for a ~1000-km system) as a function of year along with an estimate of the single-fiber capacity required to meet the 40% per annum in Internet traffic. The nonlinear Shannon limit for the existing fiber technology is rapidly being approached and a "capacity crunch" looms ahead without major innovation in the basic fiber infrastructure used within our networks, as illustrated by the question mark labeling the idealized future growth curve. Progress toward the capacity limit of current fiber technology has resulted from several key breakthroughs over the years, which include the development of low-loss, single-mode transmission fiber, the invention of the erbium-doped fiber amplifier (EDFA), the development of wavelength division multiplexing (WDM), and most recently the development of DSP-based coherent detection, which allows for high-SE signaling. Although a capacity crunch has been marked around 2020, this should not be taken too literally, as there is considerable uncertainty in this date given differing estimates of traffic growth for different countries and different future fiber deployment scenarios.

wavelength division multiplexing (WDM) in which the bandwidth is occupied by multiple, independent and spectrally distinct wavelength channels each carrying encoded data.

The maximum amount of information that can be transported per unit bandwidth for each WDM channel is quantified by a number referred to as the spectral efficiency (SE) as measured in terms of numbers of bits per second per Hz (b/s/Hz). An upper bound on the SE was established by Claude Shannon in 1948 on the basis of fundamental information theory principles and is given by $SE = \log_2(1 + SNR)$ where SNR is the signal-to-noise ratio (4). It was later shown that optical nonlinearity, which restricts the signal powers that can be usefully sent through the fiber for a given transmission distance, limits the SE values that can be achieved in practice (5). The current consensus is that the maximum practical value is likely to be ~10 b/s/Hz. Such values of SE should be possible using advanced coding schemes borrowed from radio communications that exploit the phase, amplitude, and polarization of the optical field (6).

Until recently, binary amplitude coding (turning the laser on and off to define "1" and "0" bits, respectively) was the preferred and only practical coding option—providing for SEs of less than 1 b/s/Hz and considerably much less in most practical implementations. However, as a result of steady technological

improvements, most notably in the speed and power of electronic digital signal processing (DSP), the use of (coherent) phase encoding has now become possible. Moreover, at the same time, through advances in laser and filtering technologies, the individual WDM wavelength channels can now be packed more closely together without intolerable interference, providing a further boost to overall bandwidth utilization. These developments have culminated in the flurry of impressive results at OFC 2010, which included demonstrations of a record 69.1 terabit/s data transmission over 240 km of fiber at a SE of 6.4 b/s/Hz (7), 64 terabit/s transmission over 320 km at a SE of 8.0 b/s/Hz (8), and finally, extending the distance, transmission of quadrature phase shift–keyed signals at 3 b/s/Hz over >10,600 km (9). These are staggering results and are within a factor of ~2 of the nonlinear Shannon limit for the current fiber technology, and it is therefore hard to envisage substantial further improvements. Although it will take much effort to transfer these high-SE results over to commercial systems, it will provide welcome headroom for several years to come. However, in terms of capacity scaling of existing fiber systems, the end is within sight. Once the capacity of conventional single-mode fiber-based systems is exhausted, the only option will be to add additional parallel systems, an option that is highly undesir-

able from a cost and, ultimately, a power sustainability perspective.

Further innovation and breakthroughs in the basic fiber infrastructure are urgently required. Potential ways forward include the use of spatial division multiplexing, where N optical cores rather than one are incorporated within the fiber to provide a factor of N increase in the data-carrying capacity; the development of broader-bandwidth amplifiers; reduction of fiber attenuation; and the mitigation of optical nonlinearity (either by reducing the nonlinearity of the transmission fiber itself, or by introducing active electronic or optical means to compensate it). However, it should be stressed that while many ideas exist, improving upon the current technology will be extremely challenging and network operators will also need to look increasingly at ways of reducing traffic—for example, by caching data locally throughout the network, because storing large amounts of data is far easier and cheaper than transmitting it over long distances. Perhaps adopting different charging models for Internet use could also substantially affect traffic demands, and we may all increasingly need to get used to the idea that bandwidth (just like water and energy) is a valuable commodity to be used wisely.

References

1. K. C. Kao, G. A. Hockham, *Proc. IEEE* **113**, 1151 (1966).
2. T. H. Maiman, *Nature* **187**, 493 (1960).
3. Cisco Visual Networking Index: Forecast and Methodology, 2009–2014 (www.ciscosecure.net/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html).
4. C. E. Shannon, *Bell System J.* **27**, 379 (1948).
5. P. P. Mitra, J. B. Stark, *Nature* **411**, 1027 (2001).
6. R.-J. Essiambre, G. J. Foschini, G. Kramer, P. J. Winzer, *Phys. Rev. Lett.* **101**, 163901 (2008).
7. A. Sano et al., in *Proceedings of the 2010 Optical Fiber Communication Conference* (Optical Society of America, Washington, DC, 2010), PDPB7.
8. X. Zhou et al., in *Proceedings of the 2010 Optical Fiber Communication Conference* (Optical Society of America, Washington, DC, 2010), PDPB9.
9. J. X. Cai et al., in *Proceedings of the 2010 Optical Fiber Communication Conference* (Optical Society of America, Washington, DC, 2010), PDPB10.