sign	(1)	exponer	nt (8)	fraction (23)	
IEEE fp32					
sign	(1)	exponer	nt (5)	fraction (10)	
IEEE fp16					
sign	(1)	exponer	nt (8)	fraction (7)	
Brain float16					

FIGURE 3.27 Floating-point format for IEEE 754 single precision (fp32), IEEE 754 half-precision (fp16), and Brain Float 16. Google's TPUv3 hardware uses Brain Float 16 (see Section 6.11).

Copyright © 2021 Elsevier Inc. All rights reserved