

新技术创造新世界

梁雅莉

英特尔行业解决方案集团中国区总经理





我们生活的世界日新月异...

50%

到2020年,50%的全球2000强企业将依托自己的能力创建数字化的增强的产品、服务和体验¹

80%

到2020年,大型组织中,数字化收入的预期增长将是80%²

50%

到2020年,50%的从业人员将是千禧一代³

变化前所未有的,这仅是冰山一角!

商业匠心·技术创新

1. 资料来源: IDC, <http://www.idc.com/getdoc.jsp?containerId=prUS41888916>

2. 资料来源: Gartner Group, <http://www.gartner.com/newsroom/id/3142917>

3. 资料来源: PWC, <https://www.pwc.com/gx/en/managing-tomorrows-people/future-of-work/assets/reshaping-the-workplace.pdf>



数据洪流

截至2020年

平均每互联网用户

1.5 GB

每天的流量

智能工厂

1 PB

每天的数据

云视频提供商

750 PB

每天的视频

商业匠心·技术创新



数据驱动，创新增长

数据形态正在发生革命性变化



在端到端的数据处理中实现价值创新





从云计算到人工智能



大型主机



工业标准
服务器



云计算

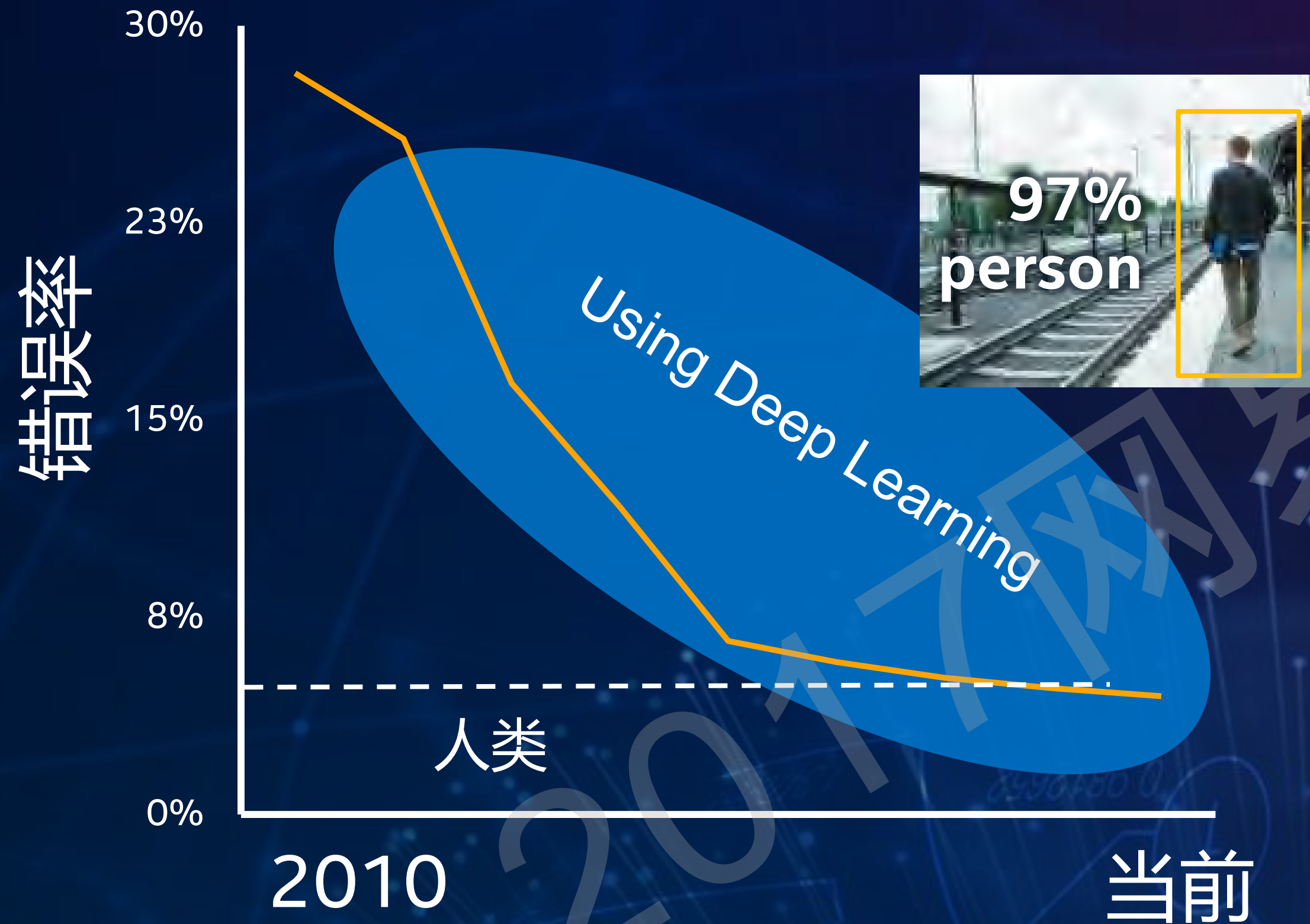
- ☑ 数据泛滥
- ☑ 计算突破
- ☑ 创新激增

人工智能

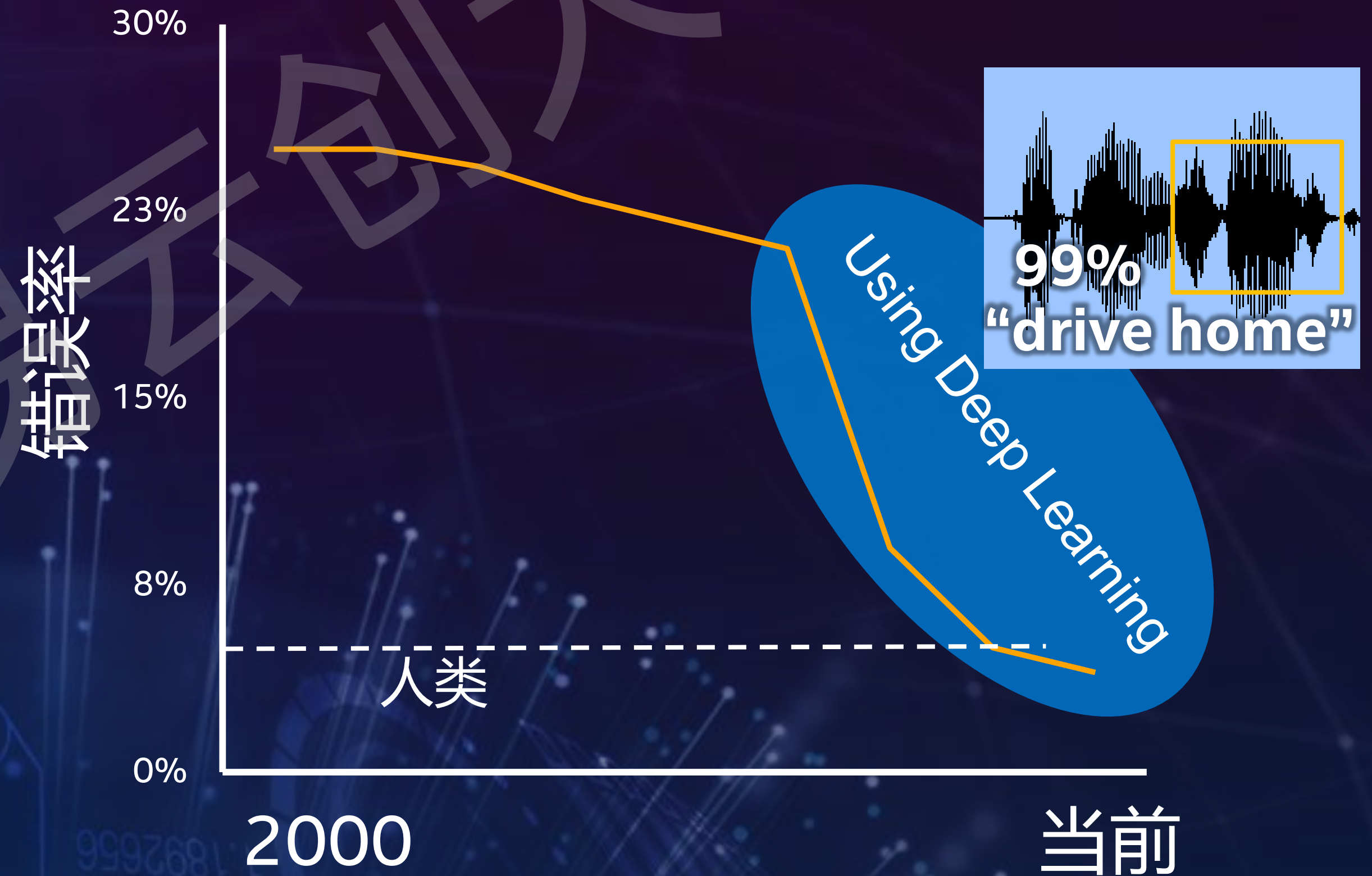
到2020年，人工智能对服务器的需求将增长12倍

深度学习已取得重大突破

图像识别



语音识别



在某些情境下，采用深度学习算法的图像/语音识别准确率已经优于人类，这不但能显著改善已有的相关应用，还将激发全新的应用。

释放数据价值，促成良性循环



云、人工智能、网络

存储、FPGA、5G

产生海量数据的物和设备

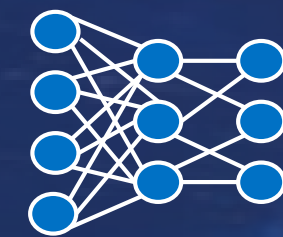


英特尔®机器学习/深度学习的技术体系

人工智能解决方案



工具



Intel® Deep Learning SDK

TAP

框架



Spark MLib
python™
Intel Distribution

库

Intel® Nervana™ Graph Compiler
Intel® MKL Intel® MKL-DNN

Intel® DAAL

硬件



Compute

Memory/Storage

Fabric

充分释放 技术潜能



英特尔® 至强® 可扩展平台



业界十年来
在平台技术上的最大进步



英特尔® 至强® 可扩展平台

高性能

高达

1.65X

每一代产品平均提升¹

业务持续性

高达

2X

每一代数据保护性能提升²

旨在实现

99.999%

服务器正常运行时间

敏捷性

4.2X

与四年前产品相比
虚拟机提升³

65%

与四年前产品相比
总拥有成本降低⁴

云 2017 人工智能

5G

1. 基于服务器虚拟化整合工作负载，支持的虚拟机可增多 4.28 倍；基于英特尔® 内部估计数据，1 个节点，2 个英特尔® 至强® 处理器 E5-2690，Romley-EP 平台，内存总量为 256 GB，运行 VMware ESXi® 6.0 GA，使用 Guest OS RHEL6.4、glassfish3.1.2.2、postgresql9.2。数据来源：请求号：1718；基准测试：服务器虚拟化整合；得分：377.6，21 个虚拟机，1 个节点，2 个英特尔® 至强® 铂金 8180 处理器，Wolf Pass SKX 平台，内存总量为 768 GB，运行 VMware ESXi6.0 U3 GA，使用 Guest OS RHEL 6 64 位。数据来源：请求号：2563；基准测试：服务器虚拟化整合；得分：1580，90 个虚拟机。数值越高越好。
2. 4 年总体拥有成本估计降低 65%，示例基于同等性能，使用 VMware ESXi® 虚拟化整合工作负载，对比：20 台已装 2 socket 服务器，英特尔至强处理器 E5-2690（以前称为“Sandy Bridge-EP”），运行 VMware ESXi® 6.0 GA，使用 Guest OS RHEL6.4，总成本为 919,362 美元；5 个新型英特尔® 至强® 铂金 8180 (Skylake) 处理器，运行 VMware ESXi6.0 U3 GA，使用 Guest OS RHEL 6 64 位，总成本为 320,879 美元，包括基本采购成本。服务器价格假设基于原始设备制造商针对以下配置发布的最新零售价格：2 socket 服务器，搭载基于 Broadwell 架构的英特尔至强处理器系统。根据所提供系统的实际价格，配置可能有所变化。



最为敏捷的可扩展人工智能平台

最大化投资 回报率

强大性能

已量产

训练时间
由数日缩短至数小时

提升高达113倍

与上一代相比¹，使用优化软件后性能

“实时的**推理性能**，
基于至强”

提升高达

2.2倍

与上一代相比²，
深度学习的性能

1. 平台：25 英特尔® 至强® 铂金 8180 CPU，2.50GHz（28 核），禁用 HT，禁用加速模式，通过 intel_pstate 驱动程序将 scaling governor 设置为 performance，384GB DDR4-2666 ECC RAM，CentOS Linux 7.3.1611 版（内核），Linux kernel 3.10.0-514.10.2.el7.x86_64，固态硬盘：英特尔® 固态硬盘 DC S3700 系列（800GB，2.5 英寸 SATA 6Gb/s，25nm，MLC）。性能测量：环境变量：KMP_AFFINITY=granularity=fine,compact，OMP_NUM_THREADS=56，CPU 频率设为 cpupower frequency-set -d 2.5G -u 3.8G -g 性能。对照平台：25 英特尔® 至强® CPU E5-2699 v3，2.30GHz（18 核），启用 HT，禁用加速模式，通过 intel_pstate 驱动程序将 scaling governor 设置为 performance，256GB DDR4-2133 ECC RAM，CentOS Linux 7.3.1611 版（内核），Linux kernel 3.10.0-514.el7.x86_64，操作系统驱动器：希捷® Enterprise ST2000NX0253 2 TB 2.5" 内置硬盘。性能测量：环境变量：KMP_AFFINITY=granularity=fine,compact,1,0，OMP_NUM_THREADS=36，CPU 频率设为 cpupower frequency-set -d 2.3G -u 2.3G -g 性能。英特尔 Caffe：(http://github.com/intel/caffe/)，修订版 b0ef3236528a2c7d2988f249d347d5fdae831236。推理能力的测量使用了“caffe time --forward_only”命令。训练能力的测量使用了“caffe time”命令。对于 ConvNet 拓扑，使用了虚构的数据集。对于其他拓扑，数据保存在本地存储上，并在训练前缓存在内存中。拓扑规格源于 https://github.com/intel/caffe/tree/master/models/intel_optimized_models (GoogLeNet、AlexNet 和 ResNet50)，GCC 4.8.5，MKLML 版本 2017.0.2.20170110.BVLC-Caffe：https://github.com/BVLC/caffe。

2. 平台：25 英特尔® 至强® 铂金 8180 CPU，2.50GHz（28 核），禁用 HT，禁用加速模式，通过 intel_pstate 驱动程序将 scaling governor 设置为 performance，384GB DDR4-2666 ECC RAM，CentOS Linux 7.3.1611 版（内核），Linux kernel 3.10.0-514.10.2.el7.x86_64，固态硬盘：英特尔® 固态硬盘 DC S3700 系列（800GB，2.5 英寸 SATA 6Gb/s，25nm，MLC）。性能测量：环境变量：KMP_AFFINITY=granularity=fine,compact，OMP_NUM_THREADS=56，CPU 频率设为 cpupower frequency-set -d 2.5G -u 3.8G -g 性能。对照平台：25 英特尔® 至强® CPU E5-2699 v4，2.20GHz（22 核），启用 HT，禁用加速模式，通过 acpi-cpufreq 驱动程序将 scaling governor 设置为 performance，256GB DDR4-2133 ECC RAM，CentOS Linux 7.3.1611 版（内核），Linux kernel 3.10.0-514.10.2.el7.x86_64，固态硬盘：英特尔® 固态硬盘 DC S3500 系列（480GB，2.5 英寸 SATA 6Gb/s，20nm，MLC）。性能测量：环境变量：KMP_AFFINITY=granularity=fine,compact,1,0，OMP_NUM_THREADS=44，CPU 频率设为 cpupower frequency-set -d 2.2G -u 2.2G -g 性能。Neon / ZP/MKL_CHWN 分之提交 ID：52bd02acb947a2adabb8a227166a7da5d9123b6d。使用了虚构的数据。在 mkl 模式中，基准测试使用了 main.py 脚本。使用的 ICC 版本：17.0.3 20170404，英特尔 MKL 小型库版本 2018.0.20170425；推理和训练的吞吐量使用 FP32 指令。

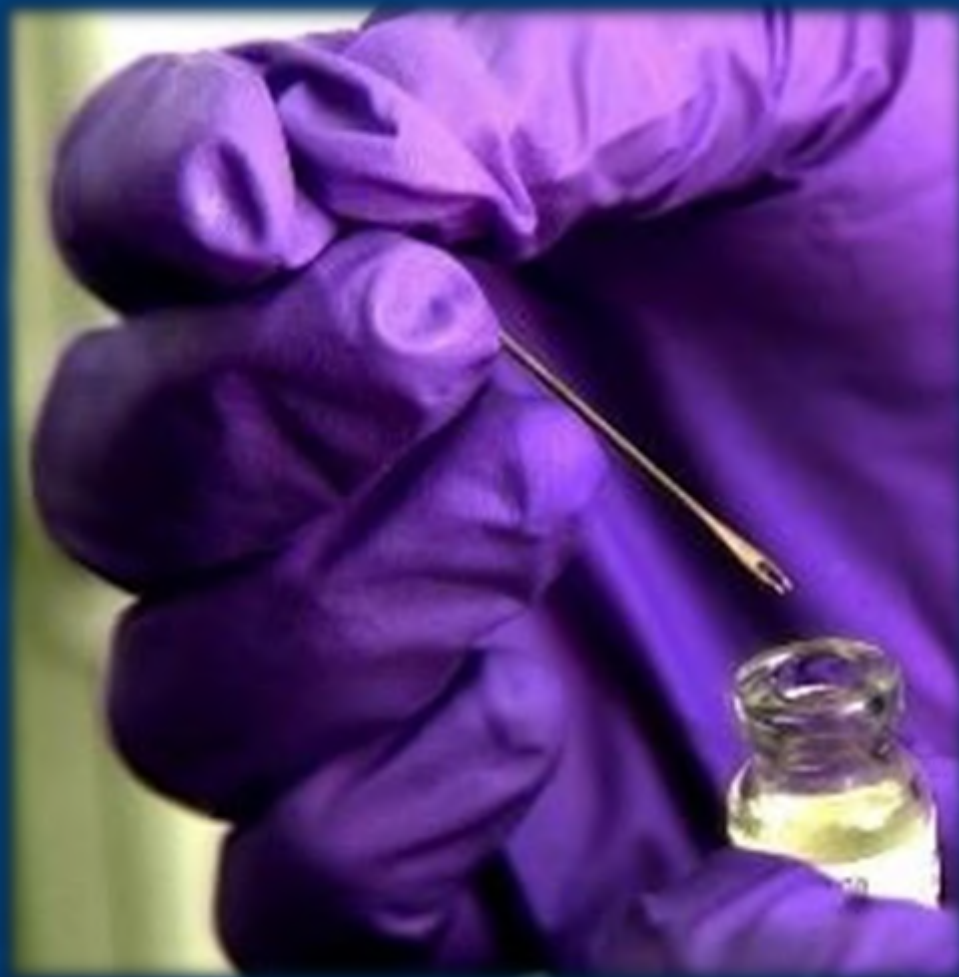


创新推动变革，让世界更美好

人工智能的产业规模和社会价值，堪比农业革命、工业革命和数字革命

加速

开发大规模的解决方案



疾病治疗
排除欺诈
解锁暗数据

释放

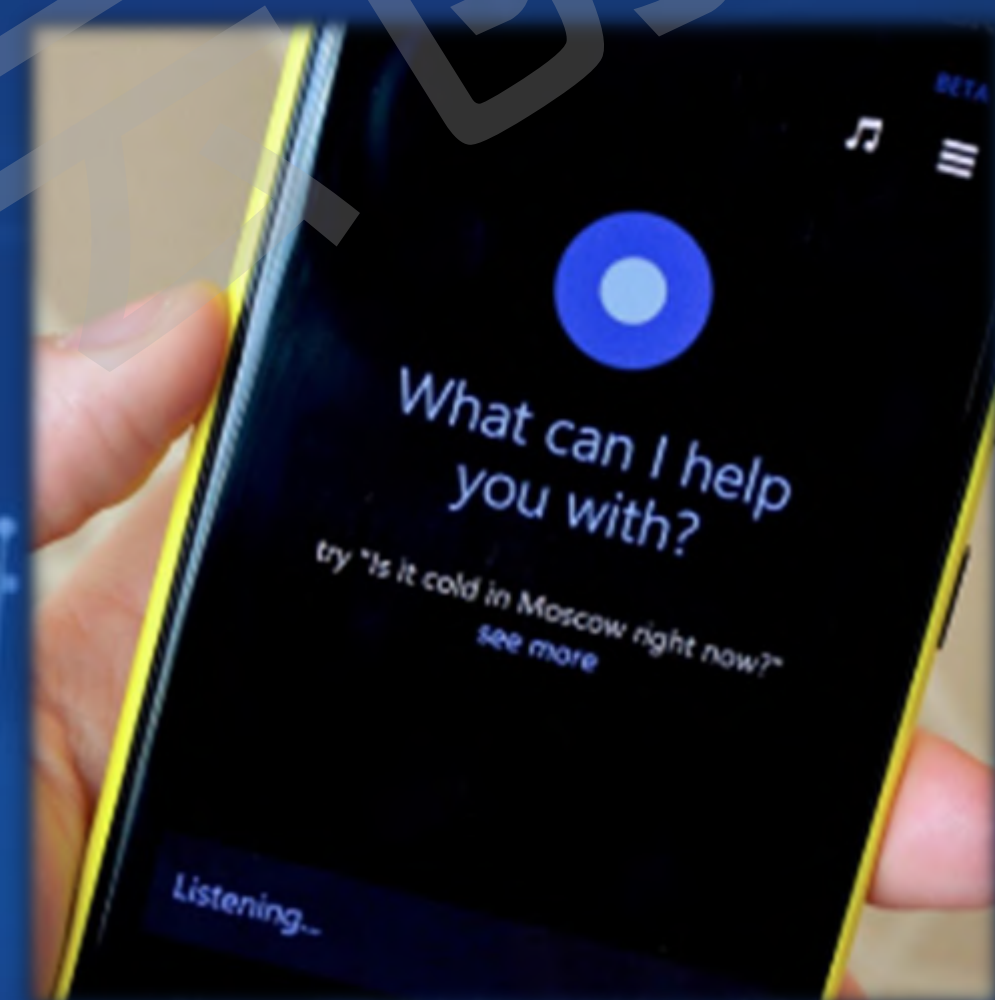
推动科学发现



深海/深空探测
粒子物理研究
解读大脑奥妙

增强

扩展人的能力



个人向导
增强决策
预防犯罪

自动

减少有风险和枯燥的任务



自动驾驶
灾难搜救
摆脱琐事

携手创新，共建未来



数据中心计算

- 新一代Purley平台服务器提升ROI, 降低TCO
- 合作开展FPGA应用实现
- 利用Intel QuickAssist技术加速计算与安全传输



数据中心存储

- 基于Intel 3D NAND与Optane SSD、SPDK技术提升存储性能
- 合作优化NetEase分布式存储系统
- 合作优化Ceph分布式存储系统



数据中心网络

- 合作开展DPDK应用优化
- 利用25G、40G网络加速云网络性能



软件优化

- 合作开展人工智能合作, 建立人工智能云平台, 优化人工智能算法
- 计算虚拟化与容器优化
- MySQL, 大数据平台, CaaS平台优化

2 0 1 7

网易云创大会

商业匠心 · 技术创新

2 0 1 7 网 易 云 创 大 会

