未来网络发展大会
FUTURE NETWORK
DEVELOPMENT CONFERENCE
NANJING · CHINA / 中国 · 南京

未来网络技术发展系列白皮书（2025）

量子互联网与算网协同体系架构白皮书

第九届未来网络发展大会组委会

2025年8月

版权声明

编写说明

主要编写单位：

紫金山实验室

江苏省未来网络创新研究院

北京邮电大学

主要编写人员：

张浩、李媛、张晨、黄韬、刘韵洁

前　言

从量子这个概念的提出，到以半导体技术为基础的第一次量子革命，孕育出了现代计算机文明，给人们的社会生活带来了巨大的变化。其中极具代表性的应用场景之一就是计算机通信和互联网，其使得人与人之间的交流变得非常方便。近几十年来，以操控量子态为基础的第二次量子革命又带来了新的量子信息技术，比如量子通信、量子计算和量子精密测量。这类新技术都是以量子力学原理来进一步突破原有的技术路线。其中量子通信是利用量子不可克隆原理从物理上实现绝对安全通信；量子计算是利用量子态叠加原理实现并行运算，极大提高计算速度；而量子精密测量则是突破标准量子极限进一步提升测量精度。在实用化的过程中，随着用户和节点数目的增加，很自然地就形成了量子网络。当网络的覆盖面变得很大，类似于当今全球互联网时，就形成了量子互联网。所以在将量子信息实用化的过程中，对量子互联网进行深入的研究和发展是必然趋势。

目前量子互联网的发展还处在初期阶段。由于其和经典互联网的基本原理不同，很多经典互联网的发展模式和技术都无法直接借鉴过来。现阶段不论是底层的硬件技术，如量子门操作速度和保真度、量子纠错和量子存储时间等，还是上层的量子互联网体系架构，如运行模式和协议栈，都不成熟。这也导致在量子互联网的研究中还面临很多新的问题和挑战。

本白皮书首先简洁地介绍和梳理量子互联网相关的基本原理和

技术，包括部分量子信息基础知识和代表性协议等。随后介绍量子互联网的发展现状和代表性的体系架构方案。最后围绕量子互联网的基本技术路线提出构建未来量子互联网的运行模式，讨论和展望量子算网协同的研究内容和可能的发展方向。本白皮书旨在通过对量子互联网的介绍、梳理和展望，为量子互联网从基础理论研究朝工程和产业化发展提供一个架构和技术层面的参考。

目　　录

一、量子信息技术概述

1.1 量子信息基本概念

1.1.1 从经典力学到量子力学

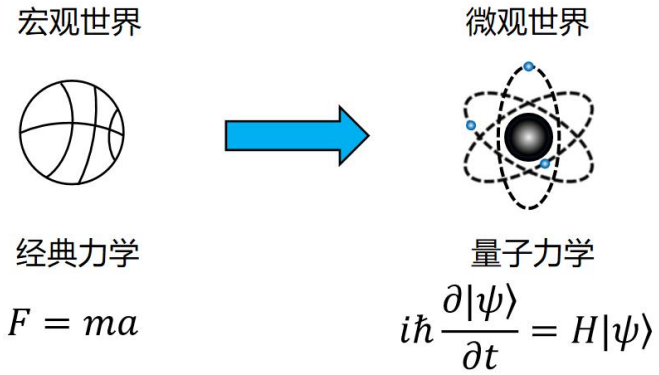

图 1. 从宏观尺度的篮球到微观尺度的原子。相应的物理理论从经典力学过渡到量子力学。

在日常生活中，我们肉眼所能见到的物体的运动行为都属于经典物理所研究的范畴。比如一块被水平扔出去的石头做抛物线运行，踩油门让车加速等。这些运动规律都可以被牛顿力学所描述。通过给定物体的质量和受力情况就可以通过 $F = ma$ 这个公式去计算物体的加速度，再结合运动学公式和初始状态计算该物体往后任意时刻的运动状态。然而牛顿力学可以计算的运动规律是有范围的，即低速宏观弱引力场情况。如图 1 所示，当我们研究的物体尺寸从日常生活中见到的宏观世界，如飞机、汽车和篮球，逐渐变小到了原子尺寸的微观世界时，情况大不相同。而描述这个微观世界粒子运动规律的理论就是量子力学。在量子力学中，微观粒子的运动状态由波函数 $|\psi\rangle$ 表达。

只要完全搞清楚物体的波函数$|\psi\rangle$随着时间如何变化就可以完全掌握物体的运动状态。此时经典力学中的$F=ma$这个公式已经无法使用，需要用薛定谔方程$i\hbar\frac{\partial|\psi\rangle}{\partial t}=H|\psi\rangle$来计算物体的波函数$|\psi\rangle$。其实基于量子力学的技术和产品早已被我们使用，比如电脑和手机中的半导体，其中的原理就用到了量子力学能带理论。

1.1.2 量子态及其演化

上面说到了微观世界粒子的运动规律需要用量子力学所描述。而在量子力学中，一个粒子的状态，也就是量子态，用波函数描述。在符号上，我们习惯用ψ来表示，或者用狄拉克符号$|\psi\rangle$表示。比如一个光子有水平和竖直两个偏振状态，此时我们就可以分别将其表示为$|\mathrm{H}\rangle$（水平 Horizontal 首字母）和$|\mathrm{V}\rangle$（竖直 Vertical 首字母）。一个原子有自旋向上和向下两种状态，可以表示为$|\uparrow\rangle$和$|\downarrow\rangle$。甚至一只猫的死和活的状态，都可以表达为$|\text{死}\rangle$和$|\text{活}\rangle$。量子力学中的概念有很多，为简单起见，我们只在这里介绍几个后续内容涉及到的重要概念。下面我们重点介绍两个非常重要的基本概念：叠加态和纠缠态。

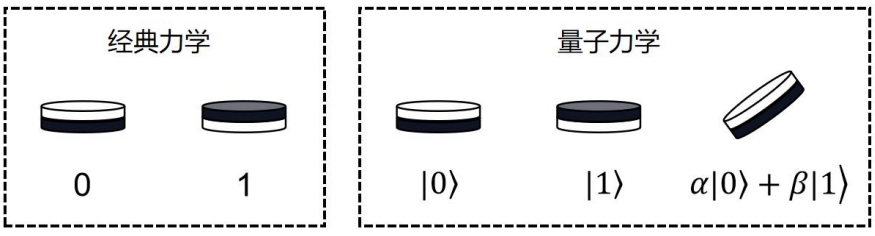

图 2. 经典物理的态和量子力学中的态。抛硬币为例，白色向上为 0，黑色向上为 1。量子力学区域最右侧为 0 和 1 的叠加态。

叠加态是指一个系统同时处于两种或多种量子态的状态。在数学

上表示为一个系统的几个量子态线性叠加。比如一个光子同时处于水平偏振和竖直偏振，表示为$|\psi\rangle=\frac{1}{\sqrt{2}}(|\mathrm{H}\rangle+|\mathrm{V}\rangle)$。表达式中的$\frac{1}{\sqrt{2}}$是归一化因子。一个原子同时处于自旋向上和向下$|\psi\rangle=\frac{1}{\sqrt{2}}(|\uparrow\rangle+|\downarrow\rangle)$。如果一个物体同时处于 n 种状态，就是$|\psi\rangle=\frac{1}{\sqrt{n}}(|\psi_1\rangle+|\psi_2\rangle+...+|\psi_n\rangle)$。这种情况也许会让没接触过量子力学的人感到很奇怪。因为在我们先前经验中见到的世界里，一个物体的状态只会出现某一种。以抛硬币为例来对比一下经典世界和量子世界，如图 2 所示。硬币白色朝上为 0，黑色朝上为 1。左侧虚线框内的经典世界要么出现 0 要么出现 1。右侧的虚线框内的量子世界就很奇特，既可以独立出现 0 和 1，也可以出现 0 和 1 的叠加态，也就是 0 和 1 同时存在。这种叠加态其实对于我们生活在宏观世界的人来说很难想象。因为我们从出生到现在见到的世界里的状态都是某一个确定的状态，比如光子要么处于水平偏振，要么就是竖直偏振。一只猫，要么是死的，要么就是活的。同时处于死和活的状态是一种什么样子，我们根本想象不出来。所以在量子力学建立的早期，许多量子物理学家都很难接受其他人甚至自己提出的一些理论带来的“奇怪”结论。其中对于叠加态的质疑就是著名的薛定谔的猫。物理学家们通过猫能否同时处于死和活的状态来质疑量子力学的正确性。因为按照量子理论，微观世界的粒子是有叠加态的。但是宏观世界，我们的生活经验告诉我们，这是不可能的。因为我们从未看到过一只猫既是死的又是活的。物理学家们通过将猫的死活状态和微观粒子的状态绑定在一起，想从猫不可能同时处于死和活这个宏观

世界的事实反过来质疑量子力学叠加态是不成立的。至于为什么宏观世界看不到薛定谔猫这种现象，开放系统理论认为宏观世界中的物体并不是一个孤立系统，周围有很多物体和其相互作用，这就导致了量子态的退相干，很难处于量子世界中的叠加态，也就解释了为什么我们无法在经典世界中看到这类现象。

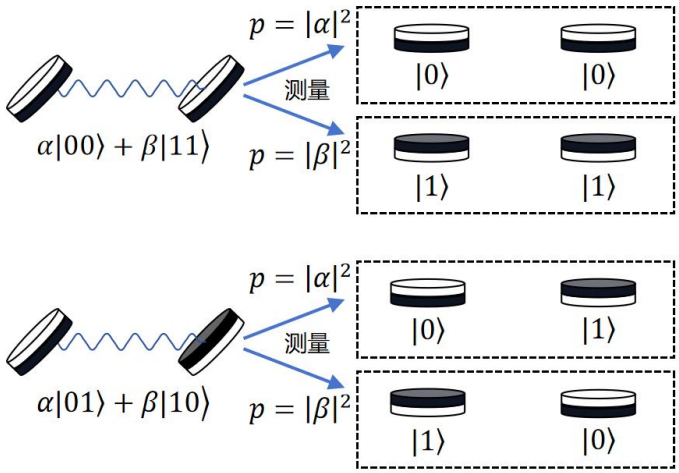

图 3. 纠缠态及其测量后的状态。p 为该测量结果出现的概率。

另一个重要的概念是纠缠态。纠缠态是两个或多个系统的状态不能表示为各个系统量子态直积形式的态。以两个物体为例，通常在没有任何关联的情况下，物体 1 的状态为$|\psi_1\rangle$，物体 2 的状态为$|\psi_2\rangle$。将他们看成一个复合体系来表达，其状态为$|\psi\rangle=|\psi_1\rangle\otimes|\psi_2\rangle$，其中$\otimes$为直积符号。有时为了简便，省略中间的直积符号，直接写为$|\psi\rangle=|\psi_1\rangle|\psi_2\rangle$。如果是纠缠态，其复合系统的状态就无法写成直积形式，也就是$|\psi\rangle\neq|\psi_1\rangle\otimes|\psi_2\rangle$。以光子为例，如果两个光子组成的复合系统的偏振状态表达式为$|\psi\rangle=\frac{1}{\sqrt{2}}\left(|\mathrm{H}_1\rangle|\mathrm{H}_2\rangle+|\mathrm{V}_1\rangle|\mathrm{V}_2\rangle\right)$，其无法写成两个光子各自状态的直积形式。此时两个光子之间是纠缠的，对其中一个光子进行 H/V

测量，如果测量结果为 H，则另外一个光子的状态也是 H。同样测量到 V，则另外一个光子的状态就变为 V。目前的量子理论认为这种关联是非局域的。当两个相距很远的光子处于纠缠态，这种关联依然成立。总体来看，两光子组成的复合系统其实就是一个叠加态，即同时处于量子态$|H_1\rangle|H_2\rangle$和$|V_1\rangle|V_2\rangle$。图 3 为两个纠缠态的例子，分别考虑两个纠缠态$\alpha|00\rangle+\beta|11\rangle$和$\alpha|01\rangle+\beta|10\rangle$，系数满足归一化条件$|\alpha|^2+|\beta|^2=1$。当纠缠态被测量时，其结果会出现关联。而这两种结果出现的概率由量子态中的系数决定的。例如对于纠缠态$\alpha|00\rangle+\beta|11\rangle$，测量后的结果总是两个白色朝上（00），或者两个黑色朝上（11）。而得到这两个结果的概率分别为$|\alpha|^2$和$|\beta|^2$。对于纠缠态$\alpha|01\rangle+\beta|10\rangle$，结果总是一黑一白，也就是 01 或者 10。

对于两比特的最大纠缠态，我们叫做 Bell 态（Bell state），其有四种情况，分别表示为：

$$|\Psi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle),$$

$$|\Psi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle-|11\rangle),$$

$$|\Phi^+\rangle=\frac{1}{\sqrt{2}}(|01\rangle+|10\rangle),$$

$$|\Phi^-\rangle=\frac{1}{\sqrt{2}}(|01\rangle-|10\rangle).$$

这四个 Bell 态是量子信息中经常用到的纠缠态，后续内容也会经常涉及。对于三比特的纠缠态，常见的有 GHZ（Greenberger-Horne-Zeilinger）态，也就是$|GHZ\rangle=\frac{1}{\sqrt{3}}(|000\rangle\pm|111\rangle)$。n 比特 GHZ 态的表达式为

$\frac{1}{\sqrt{n}}(|00...0\rangle \pm |11...1\rangle)$，上述狄拉克符号中有 n 个 0 和 n 个 1。

当一个物体的量子态随着时间进行演化的，其数学上的表达为 $|\psi(t)\rangle = U(t)|\psi(0)\rangle$，其中 $U(t)$ 为演化算符。在上式中给定时间 t 后，就可以计算出该时刻的量子态。而演化算符 $U(t)$ 由系统的哈密顿量 H 决定。以上的计算表达式本质都是计算薛定谔方程 $i\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle$ 来求解系统的波函数 $|\psi(t)\rangle$。

更详细的介绍可参考书籍[1-4]。

1.1.3 量子操作

对一个系统进行的各类量子操作都可以叫做量子逻辑门。当然这类量子操作最终都是作用在系统的量子态上，可以通过量子操作来改变系统的状态。比如最简单的比特翻转门 X，可以将 0 和 1 进行翻转，其表达为：$X|0\rangle = |1\rangle$ 和 $X|1\rangle = |0\rangle$。当基矢为 $|0\rangle=\begin{bmatrix}1\\0\end{bmatrix}$，$|1\rangle=\begin{bmatrix}0\\1\end{bmatrix}$ 时，X 门的矩阵形式为 $X=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$。再如 Hadamard 门，当基矢为 $|0\rangle=\begin{bmatrix}1\\0\end{bmatrix}$，$|1\rangle=\begin{bmatrix}0\\1\end{bmatrix}$ 时，其矩阵形式是 $H=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$。所以 Hadamard 门有 $H|0\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ 和 $H|1\rangle=\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$。如果是两比特系统，除了有每个量子位的单量子比特操作，还有两量子比特门。最常见的就是受控非门（CNOT），其作用就是当控制位比特为 0 时，靶位不做任何操作，当控制位为 1 时，靶位进行比特翻转。有时候也可以将控制位的态反过来控制，比如控制位为 1 时靶位翻转，控制位为 0 时不做操作。考虑控制位为 1 进行靶位的比特翻转情况，其 CNOT 门的态矢表达形式为

$\mathrm{CNOT}=|00\rangle\langle 00|+|01\rangle\langle 01|+|11\rangle\langle 10|+|10\rangle\langle 11|$。当基矢为$|00\rangle=[1\ 0\ 0\ 0]^T$，$|01\rangle=[0\ 1\ 0\ 0]^T$，$|10\rangle=[0\ 0\ 1\ 0]^T$和$|11\rangle=[0\ 0\ 0\ 1]^T$时（这里的 T 表示矩阵转置），其矩阵形式为

$$\mathrm{CNOT}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}.$$

两比特的量子门还有控制 Z 门、控制相位门和交换门。在量子计算中，通常会用量子线路的形式来表达门操作序列。图 4 所示为几个单量子比特门和两量子比特 CNOT 门的量子线路表示。

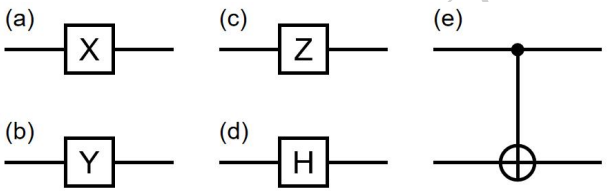

图 4. 量子门操作线路图。（a）泡利 X 门，也称为比特翻转操作。（b）泡利 Y 门。（c）泡利 Z 门，也称为相位翻转操作。（d）Hadamard 门。（e）两比特 CNOT 门，上侧带黑色点的量子比特为控制位，下侧带圆圈的量子比特为靶位。

更详细的介绍可参考书籍[5]。

1.1.4 量子测量

测量是量子力学中一个非常重要的概念和过程。因为量子力学中，系统的状态都是由波函数ψ来描述。但是波函数并不是一个可观测量，无法直接被观测到。所以要想实验上确切知道一个系统的状态，必须通过测量一个可观测量去获取状态信息。对于量子测量的基本假设是

其由一组满足完备性条件$\sum_k M_k^\dagger M_k=I$的测量算子$\{M_k\}$作用在被测量系统的状态空间上，测量后系统的状态以 $p(k)=\langle\psi|M_k^\dagger M_k|\psi\rangle$的可能性由$|\psi\rangle$变为$\frac{M_k|\psi\rangle}{\sqrt{\langle\psi|M_k^\dagger M_k|\psi\rangle}}$。这里要求算符满足完备性条件是因为测量得到的所有可能状态的概率之和为 1，即$\sum_k p(k)=\sum_k\langle\psi|M_k^\dagger M_k|\psi\rangle=1$。以单量子比特的二能级系统为例子，假设初始量子态为$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle$，系数已归一化。如果测量算子为$M_0=|0\rangle\langle 0|$和$M_1=|1\rangle\langle 1|$，则测量得到结果为 0 的概率为 $p(0)=|\alpha|^2$，测量后状态为$\frac{M_0}{|\alpha|}|\psi\rangle=\frac{\alpha}{|\alpha|}|0\rangle$；测量得到的结果为 1 的概率为 $p(1)=|\beta|^2$，测量后状态为$\frac{M_1}{|\beta|}|\psi\rangle=\frac{\beta}{|\beta|}|1\rangle$。以上的$\frac{\alpha}{|\alpha|}$和$\frac{\beta}{|\beta|}$可以忽略，所以有效的状态分别为$|0\rangle$和$|1\rangle$。

这里的假设为一般测量假设。在量子信息中有两个非常重要的特殊的测量分别是投影测量和半正定算子值测量（positive operator-valued measure，简称 POVM）测量。对于投影测量，其测量算子除了满足以上假设的完备性条件外，还要满足正交投影算子的厄米性条件。这里的测量算子是被观测系统状态空间上的一个可观测量厄米算子 $M=\sum_k kP_k$。其中 P_k 是到本征值 k 的本征态空间 M 上的投影。测量的可能结果为测量算子的本征值 k，其中对应的概率为 $p(k)=\langle\psi|P_k|\psi\rangle$，测量后的状态为$\frac{P_k|\psi\rangle}{\sqrt{p(k)}}$。对于 POVM 测量，主要用于那些不关注测量后系统的状态，而是关注测量后系统得到不同结果的概率的测量场景。这里不再对其做详细介绍，更多内容可参考书籍[5]。

1.2 典型量子应用

1.2.1 量子通信

量子通信包括量子密钥分发、量子隐形传态、量子安全直接通信、量子秘密共享和量子密集编码等，这里我们主要介绍前三种通信方案。更多关于量子通信内容可以参考书籍[4]和文献[6]。

（1）量子密钥分发

保密通信的思想是发送方先将信息加密成密文，然后将密文通过信道发送给接收方，接收方再用密钥解密。由于密文是被加密过的信息，即使窃听者将密文截获，也需要正确的密钥才可以得到准确的信息内容，否则就难以获取信息。所以只要通信双方事先可以共享绝对安全的密钥，那么就可以确保信息的传送是绝对安全的。经典通信中的加密是基于数学计算复杂度来实现的。一些好的加密算法通常是经典计算机无法在多项式时间内有效求解的，那么这类算法被认为是暂时安全的。由于经典通信的信息安全是基于数学计算复杂度的，其算法无法保证绝对的安全，所以有时候就会出现算法被破解造成信息不安全而需要更换新的加密算法的情况。如果拥有新的高效破解算法或者绝对计算优势的计算机，那么经典的信息安全就会受到严重威胁。相比于经典保密通信，量子通信是利用物理原理的绝对安全性来实现通信的绝对安全。其从物理原理上保证信息安全。第一个量子通信模型是 Bennett 和 Brassard 在 1984 年提出的 BB84 协议[7]。该协议本质上是量子密钥分发（QKD），就是实现通信双方共享绝对安全的量子

密钥，然后结合一次一密来实现绝对安全的通信。除了上述基于单光子的 BB84 协议，还有基于纠缠态的 QKD 协议，比如著名的 E91 协议[8]和 BBM92 协议[9]。这些早期的协议都是基于理想的物理实现，比如完美的单光子源和测量设备等。然而在现实应用中，这些理想的实验条件很难达到，使得实际运行的量子通信存在安全漏洞。随着研究的深入，人们不断发展出可以应用于实际非完美物理系统下的 QKD 协议来弥补安全漏洞和提高密钥率。比较典型的方案有诱骗态-QKD（Decoy state QKD）[10,11]、测量设备无关-QKD（MDI-QKD）[12]和双场-QKD（TF-QKD）[13]。其中诱骗态-QKD 是通过发送不同强度的诱骗信号来检测窃听者，从而抵御了因非完美单光子源导致的光子数分离攻击。MDI-QKD 通过在通信双方中间位置引入不可信的第三方进行测量来移除探测器漏洞。TF-QKD 将原有的基于双轨编码的 MDI-QKD 采用单轨编码，使得第三方测量由双光子干涉变为单光子干涉，理论上将原有的密钥分发距离提升了一倍。下面我们将简单介绍 BB84-QKD 协议的基本原理。

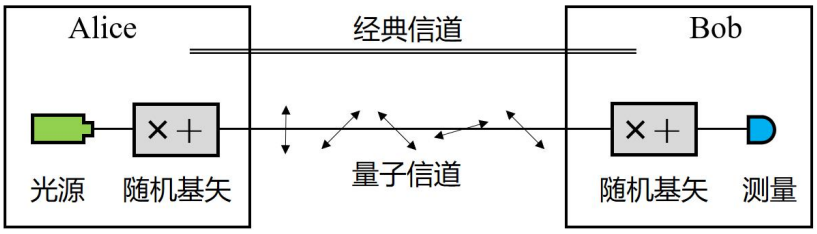

图 5. BB84 协议示意图。

BB84 协议具体的内容如下[7]：如图 5 所示发送方 Alice 随机地选择基矢＋（Z 基）或×（X 基）来制备单光子的偏振态。然后将光

子发送给 Bob。Bob 也随机地选择基矢＋或×来测量其接收到的光子偏振态，保留测量结果，并公布其所用的测量基。此时 Alice 也公布其所选择的基矢。Alice 和 Bob 通过经典通信共同比对双方公布的基矢，保留相同基矢所对应的结果。表格 1 给出了一个密钥协商的例子。可以看出，当 Alice 和 Bob 的基矢选择一致时，比特序列被保留。然后通过窃听检测、纠错和隐私放大生成最终的安全密钥。

Alice 比特序列	0	1	1	0	1	0	0	0	1	1
Alice 基矢选择	＋	＋	×	＋	×	×	×	＋	＋	×
光子偏振	H	V	+45°	H	+45°	-45°	-45°	H	V	+45°
Bob 测量基矢	×	＋	＋	＋	×	＋	×	＋	×	×
Bob 测量结果	45°	V	V	H	+45°	V	-45°	H	-45°	+45°
Bob 比特序列	1	1	1	0	1	1	0	0	0	1
匹配与否	否	是	否	是	是	否	是	是	否	是
密钥		1		0	1		0	0		1

表 1. BB84 协议密钥协商过程。

E91 协议为第一个基于纠缠的 QKD 协议，由 A. K. Ekert 在 1991 年提出[8]。和 BB84 不同的是，通信双方需要事先共享纠缠对，也就是 EPR 对。然后双方随机地从三个测量基矢中选择一个对各自持有的量子比特进行独立测量。随后双方将多次测量的测量基矢通过经典

信道进行比对，相同的基矢所对应的结果各自持有，不同的测量基对应的结果进行窃听检测。检测安全后，各自保留的结果将作为安全密钥。而 BBM92[9]协议是 BB84 协议的纠缠版本。其先分发纠缠然后用 BB84 一样的测量基去测量纠缠光子对，然后通过经典通道比对结果。

（2）量子隐形传态

量子隐形传态是通过纠缠信道直接传输未知量子态的一种途径[14]。如图 6 所示，通信双方 Alice 和 Bob 共享一对纠缠量子比特，也就是 Alice 的粒子 2 和 Bob 的粒子 3 纠缠。Alice 想把粒子 1 中的未知量子态传送给 Bob。其只要对粒子 1 和 2 做一个 Bell 态测量，然后将测量结果通过经典通信告知 Bob。然后 Bob 根据测量结果来对粒子 3 做相应的量子操作即可得到原本粒子 1 的量子态。注意此时粒子 1 的量子态已经发生改变。具体的数学表达如下：假设粒子 1 为一个量子比特，其未知量子态可以表达为$|\psi\rangle_1=\alpha|0\rangle+\beta|1\rangle$，其中系数满足$|\alpha|^2+|\beta|^2=1$。粒子 2 和 3 为纠缠态$|\psi\rangle_{23}=|\Phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$。所以这三个粒子一起组成的复合系统的量子态就可以写为$|\psi\rangle_{123}=|\psi\rangle_1\otimes|\psi\rangle_{23}$。将粒子 1 和 2 的量子态重新写成四个 Bell 态的形式，于是上述量子态改写为

$$|\psi\rangle_{123}=\frac{1}{2}[|\Phi^+\rangle_{12}(\alpha|0\rangle_3+\beta|1\rangle_3)+|\Phi^-\rangle_{12}(\alpha|0\rangle_3-\beta|1\rangle_3)+|\Psi^+\rangle_{12}(\alpha|1\rangle_3+\beta|0\rangle_3)+|\Psi^-\rangle_{12}(\alpha|1\rangle_3-\beta|0\rangle_3)]$$

。

分析上述表达式可以看出，对粒子 1 和 2 做 Bell 态测量，会使这两个

粒子随机塌缩到四个 Bell 态中的一个。此时粒子 3 也会变成对应的量子态。Alice 将测量结果告知 Bob，然后 Bob 根据结果来修正粒子 3 的量子态。如果 Alice 测量结果为$\left|\Phi^{+}\right\rangle$，则 Bob 不需要对粒子 3 做任何操作。如果结果是$\left|\Phi^{-}\right\rangle$，Bob 需要对粒子 3 做$\sigma_z$操作。而测量结果是$\left|\Psi^{+}\right\rangle$和$\left|\Psi^{-}\right\rangle$，则对应操作分别是$\sigma_x$和$\sigma_y$。

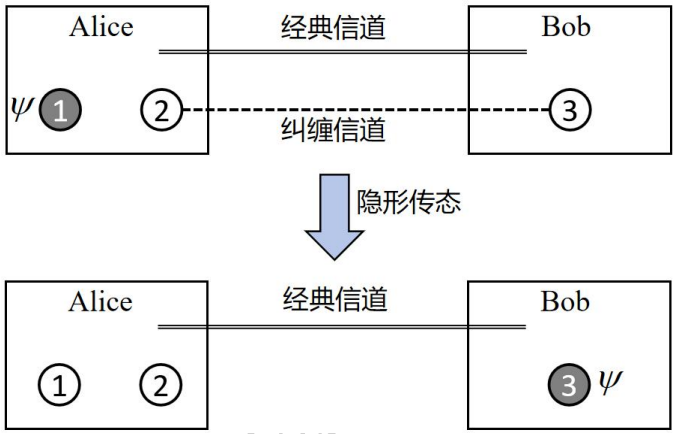

图 6. 量子隐形传态。ψ为需要从载体 1 传送到载体 3 的量子态。

量子隐形传态在量子通信和计算中有很多重要的应用。除了可以直接传递未知量子态，还有一个重要应用是纠缠交换。如果粒子 1 和另外一个粒子 4 处于纠缠态，那么通过隐形传态，也就是对粒子 1 和 2 做 Bell 态测量，可以实现粒子 3 和 4 之间的纠缠。该操作通常被应用于量子中继中扩展量子信道。

（3）量子安全直接通信

除了上述介绍的量子通信范式，还有一类通信模式叫量子安全直接通信（QSDC）[15]。这类方案无需信息加密、密钥协商和解密这些过程，而是利用量子信道来直接安全传输信息。QSDC 思想是由龙

桂鲁等人提出[16]。这里简单介绍一下基于纠缠的两步 QSDC 方案的思想[17]，大概的步骤如图 7 所示。

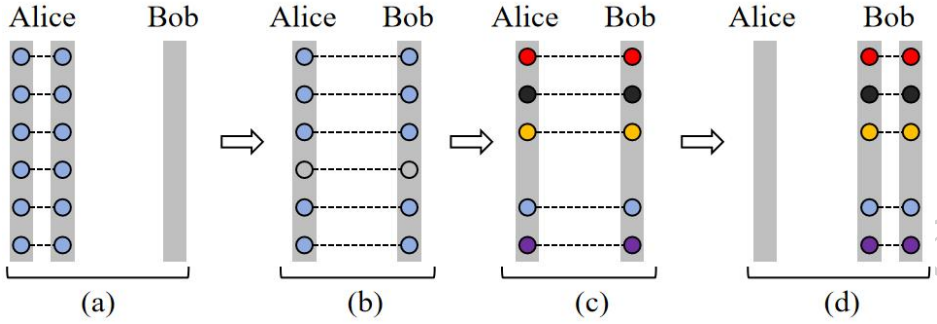

图 7. 基于纠缠的两步 QSDC 方案。（a）Alice 制备 Bell 态。（b）Alice 和 Bob 共享 Bell 态且双方进行窃听检测。白色圆圈为窃听检测的纠缠对。（c）Alice 编码信息。不同颜色圆圈代表不同的 Bell 态。（d）Alice 将持有的量子比特发送给 Bob，Bob 做 Bell 态分析获取信息。

通信双方 Alice 和 Bob 事先约定四个 Bell 态对应 2 比特信息，即 $|\Psi^-\rangle$，$|\Psi^+\rangle$，$|\Phi^-\rangle$和$|\Phi^+\rangle$分别对应 00，01，10 和 11。以上四个 Bell 态在图中分别标记为蓝色，红色，橙色和紫色。图 7（a）所示，Alice 先制备一系列某个 Bell 态，比如$|\Psi^-\rangle$。Alice 将这一系列纠缠对的另一半量子比特发送给 Bob，如图 7（b）。然后 Alice 和 Bob 对他们共享的纠缠对序列进行窃听检测，即抽取一部分纠缠对进行测量比对。如果窃听检测结果为不安全，则立即终止通信。如果结果为安全，Alice 将会对其持有的光子序列进行编码操作，如图 7（c）所示。即使用如下四个幺正操作：$U_{00}=I$，$U_{01}=\sigma_z$，$U_{10}=\sigma_x$和$U_{11}=i\sigma_y$分别会将共享的纠缠对$|\Psi^-\rangle$转变为$|\Psi^-\rangle$，$|\Psi^+\rangle$，$|\Phi^-\rangle$和$|\Phi^+\rangle$，对应编码信息 00，01，10 和 11。图 7（c）中黑色圆圈代表编码第二轮要执行误码

率检测的比特。随后 Alice 将编码后的光子再次传送给 Bob，如图 7（d）所示。Bob 对所持有纠缠对序列进行 Bell 态测量获取 Alice 编码的信息。此时双方再进行一次误码率检测，如果达标 Bob 就保留测量结果。如果检测不达标，就重传。通过以上过程 Alice 就可以将信息安全地传送给 Bob。

1.2.2 量子计算

量子计算主要是利用量子态的可叠加性来构造量子计算机。从发展历史的角度来看，量子计算机大概可以分为四个阶段。第一阶段（大概时间段为 1980 年-1985 年）是从 1980 年 Benioff 和 Manin 首次提出量子计算机的概念，到 1981 年 Feynman 提出量子模拟机，再到 1985 年 Deutsch 提出通用量子计算机，即量子图灵机。第二阶段（大概时间段为 1985 年-1994 年）为量子计算机的自由探索时期。第三阶段（大概时间段为 1994 年-2016 年）主要以 1994 年 Shor 提出大数因子分解法和 1996 年 Grover 提出量子搜索算法为代表的量子算法显示出了量子计算机的巨大社会实用价值，由此掀起了人们对量子计算机研究的第一波热潮。这一期间还有许多重要的进展，例如 1995 年 Cirac 和 Zoller 提出离子阱量子计算机；2000 年 Kitaev 提出拓扑量子计算；与此同时研究人员也在实验上探索了不同的量子计算平台，如超导量子比特等。第四阶段（大概为 2016 年至今）以 IBM 公司研制出 5 量子比特云平台开始，各种科研机构和公司纷纷研制实用化量子计算机。在这期间，2017 年 IBM 实现 50 量子比特的量子计算机；2019 年谷歌推出 Sycamore，实现量子优越性；2020 年中国科学技术大学推出

了“九章”玻色采样机；2021 年中国科学技术大学研制出“祖冲之”号超导量子计算机；2024 年谷歌发布了 105 个物理量子比特的 Willow 量子芯片。目前为止，“九章”已经发布了其四号版本，“祖冲之”已经发布了其三号版本。

量子计算机之所以有重要的研究价值，是因为其有望在某些计算任务中展现出远远超越经典计算机的强大算力。这一点主要是因为其利用量子力学中的态叠加原理。一个 n 比特的量子计算机，对其做一次操作，就可以实现对2^n个计算基矢同时进行操作。如果是两台量子计算机做并行运算，其计算能力的增加不是加法而是乘法。也正是这样，Feynman 觉得研究量子系统可以用量子计算机，而经典计算机对于体系稍大的量子系统就无法有效模拟[18]。这是基本原理上可以预见的能力，但是真正让量子计算机展现出其巨大的实用价值的当属 Shor 大数因子分解算法和 Grover 量子搜索算法的提出。Shor 算法给出的结果是对于经典计算机无法有效求解的大数因子分解问题，量子计算机可以有效求解[19]。这个算法的提出，严重威胁了现有的 RSA 公钥密码体系。而 Grover 算法展示了相比经典计算机，量子计算机在无序数据库中搜索样本任务中具有$\sqrt{N}$的加速[20]。这一算法的提出降低了 AES 对称密码体系的安全性。这两个算法都展现出了量子计算机在信息安全方面的重要影响。因此人们开始逐渐意识到量子计算机的重要性，进而投入更多的人力物力去研制量子计算机。

量子计算机计算任务需要按照特定的量子算法在量子比特上实现一系列的量子门操作来实现，如图 8 所示。其中一个可以实现通用

量子计算机的线路模型是组合任意的单量子比特门和两量子比特受控非门。这些量子门操作是按照人为设计的程序在量子比特上进行演化执行的。在具体的物理系统中，一般通过外部的光场脉冲或者磁场等来控制量子比特。

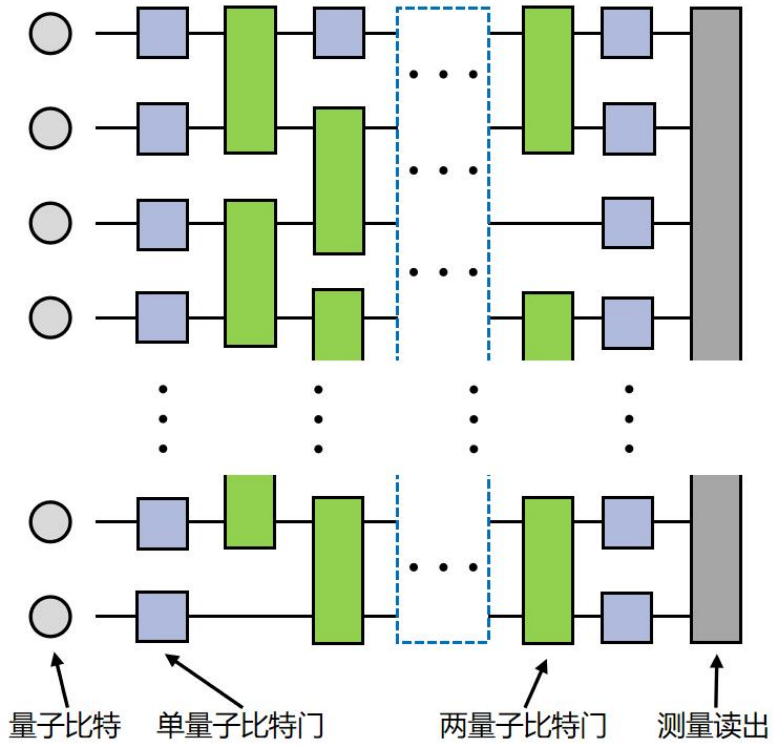

图 8. 通用量子计算机线路模型。

量子计算的并行性优势可以通过 Deutsch 算法来粗略体验一下。更一般的算法可以参考 Deutsch-Jozsa 算法[5]。如图 9 所示为 Deutsch 算法的量子线路图。两量子比特的输入态为$|\psi_{\text{in}}\rangle=|01\rangle$。经过上下两个 Hadamard 门，一个$U_f$（这里的$U_f$是映射$|x,y\rangle\to|x,y\oplus f(x)\rangle$）和上线路量子比特 Hadamard 门后，量子态变为$|\psi_{\text{out}}\rangle=\pm|f(0)+f(1)\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]$。然后通过测量第一个量子比特就可以确定$f(0)\oplus f(1)$的值，进而知道函数$f(x)$的全局性质。在经典设备中，计算$f(0)\oplus f(1)$至少需要两次，

而在量子设备中只需要一次就可以。从上面这个简单的例子就可以领略到量子计算在某些特殊任务中的优势。

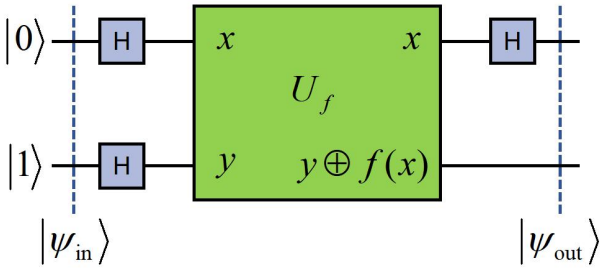

图 9. Deutsch 算法量子线路图。

量子计算在量子网络中的应用也是一个很重要的分支。考虑在初期的量子计算机，由于实验条件的限制，量子计算机无法像今天的电脑一样普及。可能只有少数几个实验室或者服务商才拥有量子计算机，而一般的用户作为顾客去远程使用服务商手中的计算资源。为了保证用户安全地将自己的计算任务委托给服务商手中的量子计算机来计算，盲量子计算模型应运而生[21,22]。借助单向量子计算机模型，用户通过和服务商手中的量子比特进行交互让计算量子比特执行一系列的操作来完成计算，而用户数据的安全性则由随机加密来保证。

除了以上的盲量子计算涉及到非局域地使用计算资源外，分布式量子计算机也是一个通过量子网络平台来进一步扩展计算能力的模型[23,24]。分布式量子计算机的思想是利用非局域的控制非门来协同多个小型量子处理器来进一步扩展成为一个更大的量子计算机。如果在量子网络中执行分布式量子计算，那么多个计算节点之间的非局域控制非门的执行质量非常关键，其会直接关系到整个分布式量子计算的计算效率。

目前由于受到比特数和门操作保真度等技术的限制，很难实现基于量子纠错码的大规模量子计算机。所以现阶段处于含噪音中等尺度的量子计算模型的研究，该阶段的目标主要是利用大约 100 量子比特规模的无量子纠错码的量子处理器来探究某些计算任务，例如量子化学、机器学习和组合优化问题等[25-27]。

1.2.3 量子精密测量和传感

在物理学中精密测量一个物理量是一项非常基础且重要的工作。通常情况下会采取将该物理量映射到相位上，然后通过精确测量相位来测量该物理量[28,29]。而相位的测量会经历态准备、相位编码、读出和估算这几个步骤。这个方案对像引力波探测、原子钟和陀螺仪等这类干涉传感器是通用的。所以这类测量方案的目标就是实现尽可能小的$\Delta\theta$来尽可能精确地测量相位θ。如果用有限个无关联的原子去测量相位，其相位的不确定度会受限于标准量子极限（standard quantum limit，简称 SQL），也就是$\Delta\theta_{\mathrm{SQL}}=1/\sqrt{N}$。所以量子精密测量就是研究如何通过量子资源去突破这个标准量子极限的界限，从而将物理量测量的更精确。研究发现通过引入压缩和纠缠可以突破标准量子极限。由于测量原子之间量子关联的引入，会使得相位的不确定满足$\Delta\theta<1/\sqrt{N}$。在没有噪音的理想情况，这类测量的极限就是海森堡极限（Heisenberg limit，简称 HL）$\Delta\theta_{\mathrm{HL}}=1/N$。

由于本书主要聚焦于量子互联网，所以简单介绍与量子网络相关的分布式精密测量方案。分布式量子传感主要是利用非局域量子关联来实现对空间分布参数的精密测量，具有超越经典测量极限的灵敏度。

较为典型的应用场景为多节点的量子相位估计[30-32]、全球的量子时钟网络[33]和长基线望远镜[34]等。更多关于量子精密测量和传感的内容可参考文献[28,29]。这里简单介绍两个应用场景。

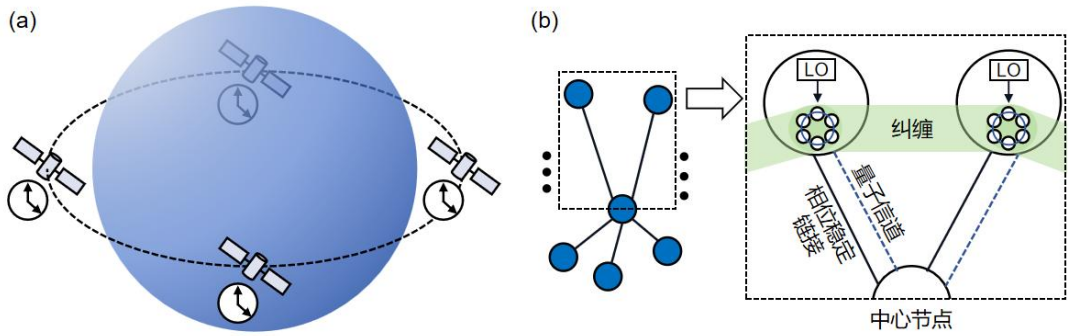

图 10. 全球量子时钟网络。

（1）量子时钟网络

量子时钟网络由哈佛大学 Lukin 组提出[33]。该研究组通过结合量子网络和精密测量提出一种用于地理上相距很遥远的光学原子钟网络的量子协同方案，进而实现安全的全球时钟。图 10 所示为基于多个卫星原子钟参与的网络。图 10（a）中是多个卫星原子钟围绕地球。图 10（b）为一个中间节点和几个其他节点连接在一起形成一个时钟网络。每个不同节点上的原子钟包含了大量的原子作为参考频率。每个时钟也拥有自己独立操控的本地振子。通过周期性地询问量子比特来维持时间，并且利用测量数据来稳定自己的本地振子频率在原子跃迁的参考频率上。每个节点分配一部分的量子比特去形成一个贯穿所有节点的纠缠态。通过该纠缠网络来获得一个每个节点都可以访问的超精确的钟信号。每个钟循环分为三个阶段：（1）初始化：制备钟原子态；（2）测量：本地振子的询问；（3）反馈：根据测量结果修正

激光频率。这种分布式的时钟结构可以让每个参与者在不丧失自主权和安全性的情况下都能通过集体数量的优势来提高自己本地时钟的稳定性。这种整体合作带来的性能增加会激励更多的节点加入，随着参与者的增多，同时又进一步增强了对信道中断的鲁棒性。整个网络的安全性则由量子通信来保证。

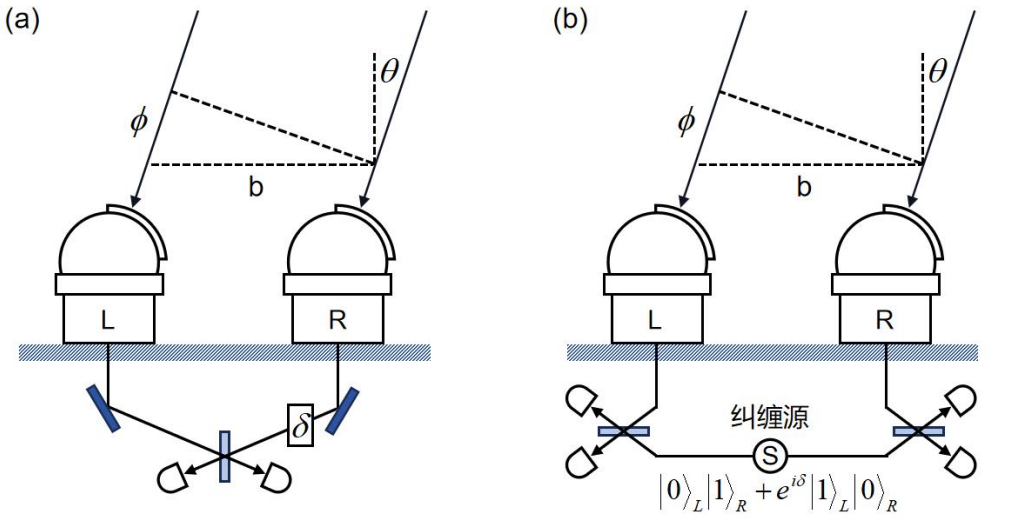

图 11. 长距离基线干涉望远镜。（a）传统远距离干涉望远镜。（b）基于量子中继器纠缠分发的远距离干涉望远镜。L 和 R 分别代表左和右。

（2）长基线望远镜

基于纠缠网络的长基线望远镜由 Gottesman 等人提出[34]。该方案是一个利用远距离纠缠分发来提高望远镜的探测能力。和前面利用量子关联来突破标准量子极限的方式不同，该方案是通过远距离的纠缠分发来提升干涉望远镜的基线距离从而提高望远镜的观测能力。对于直接探测干涉望远镜的原理如图 11 所示。被测物体发出的光照射到两个望远镜上。左侧望远镜接收到的光比右侧望远镜接收到的光要多传输了 $b\sin\theta$ 的路程。如果光的波长为 λ，则这段额外的路程会导

致左侧比右侧望远镜的光多出一个$\phi=(b\sin\theta)/\lambda$的相位。数学上表达为$|0\rangle_L|1\rangle_R+e^{i\phi}|1\rangle_L|0\rangle_R$，这里的 0 和 1 分别代表 0 和 1 个光子态。通过精确测量这里的相位ϕ就可以准确知道源的位置。由于ϕ和基线的距离成正比，所以延长两个望远镜的基线长度就可以获得更高的探测精度。

相比于一对固定基线的望远镜，望远镜阵列能获取更多的探测信息。根据 van Cittert-Zernike 定理可知，基线函数的能见度是源分布的傅里叶变换。所以如果我们可以测量到所有基线的能见度，就可以完全想象出源。通过一些离散数量的基线，可以很好地近似源亮度分布。

然而要实现如图 11（a）所示的长距离的基线望远镜有两个主要的困难。首先如果望远镜是建在地面的，则由于大气密度的震荡会影响望远镜之间的相对相位。另外要想克服光子丢失和相位错误来远距离传输单光子是非常困难的。针对以上两个困难，Gottesman 等人的方案提出利用量子中继实现远距离的纠缠分发来解决。原理如图 11（b）所示，在两个望远镜之间执行纠缠分发，将纠缠态的两个光子分别和望远镜接收到的探测光子汇聚到分束器进行干涉。随后通过测量后选择想要的结果。由于使用了量子中继，可以有效地延长基线的距离，从而提高望远镜的探测精度。除此之外还可以使用大规模的量子网络关联更多的望远镜阵列来获取更多的探测信息。

1.5 实验系统

各类量子信息任务的执行都需要一个具体的物理系统来实现。对

于量子通信，进行远距离的比特传输会选择光子作为量子比特载体。这是由于光子飞行速度快且相互作用弱，非常适合作为飞行比特进行量子比特传输。由于实际通信的需要，光量子比特可能会有不同波长的要求，比如可见光波段和通信波段等。在远程量子通信中，还需要一些相干时间长的存储量子比特用于量子中继。对于量子计算，目前研究较为广泛的物理系统有线性光子系统、超导量子比特、原子系统和自旋系统等。每个系统都有各自不同的结构和原理，因此也形成了自己的特点。对于量子计算的物理实现，DiVincenzo 提出了 5 个要求[35]：

（1）具有良好表征量子比特的可扩展物理系统（具有二能级系统）；

（2）具有初始化量子比特到一个简单基准态的能力（可以初始化）；

（3）具有很长相干时间，比门操作时间更长（相干时间长）；

（4）可以做普适的量子门操作（普适的 N 量子比特门）；

（5）具有对特定量子比特的测量能力（可以读出）。

满足以上五个要求，才可以作为实现通用量子计算机的物理系统。

这里简单介绍几个物理系统。图 12 给出了几个物理系统的示意图。更详细内容可以参考文献[36]。

（1）线性光学系统：利用光子作为量子比特，借助单光子源、线性光学器件和光子探测器来操控和测量光量子比特，从而实现量子计算。典型的线性光学量子计算方案有 KLM 方案、单向量子计算和随机行走量子计算。

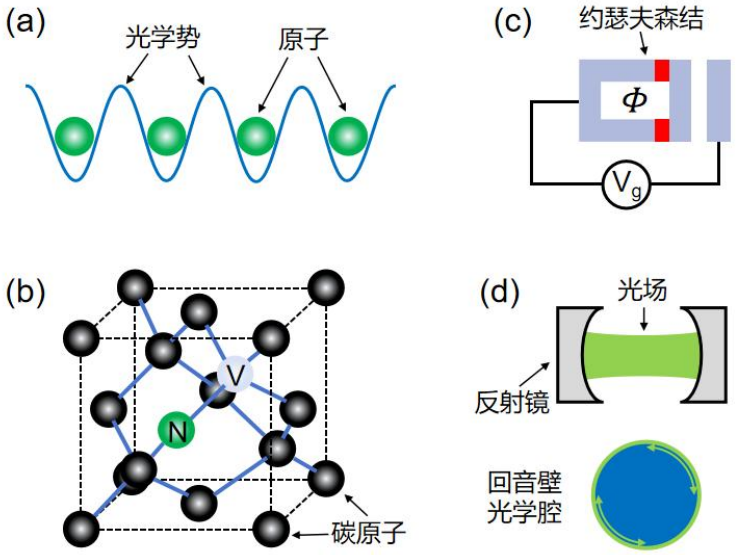

图 12. 各类物理系统。（a）一维光晶格原子系统。（b）氮-空穴金刚石色心。N 代表氮原子，V 代表空穴。（c）超导量子电路。（d）法布里-伯罗腔和回音壁模式微盘腔。

（2）原子系统：原子系统主要包括中性原子、极性分子和离子。对于中性原子，在超低温环境中，通过光晶格将原子捕获在光学势中形成阵列结构。图 12（a）给出了一维结构的光晶格原子。其二维结构有点类似于我们平时在商场里见到的鸡蛋放在蛋托中的场景，鸡蛋好比是被捕获的原子，蛋托是光学势。由于单纯的原子之间相互作用比较弱，很难扩展。所以目前都是将原子的一个电子激发到一个很高主量子数轨道，这样就可以形成一个很大的电偶极矩，从而具有很强的偶极相互作用。这种状态的原子就是里德堡原子。离子系统通常是用电场或者磁场将离子囚禁起来，利用激光对其进行精确操控。由于离子本身带电，所以离子之间通过库伦力相互作用。原子系统具有很好的相干时间，但是相比于其他系统，它们的初始化、操控和读出时

间也很长。

（3）固态自旋系统：自旋系统分为电子自旋和核自旋两类系统。目前样品主要是通过固态材料的参杂来实现大量的自旋阵列形成自旋比特。通常固态自旋系统主要分为参杂原子和量子点。比较典型的参杂原子系统有氮空穴金刚石色心。如图 12（b）所示，通过将金刚石结构中的一个碳原子用氮原子替代形成一个氮原子和邻位空穴的结构。其内部具有电子自旋和核自旋。在量子信息处理中可以操控这两种自旋。量子点是将电子囚禁在势阱中形成离散能级结构而被用来作为量子比特。自旋系统也具有很长的相干时间，但是其和外场相互作用比较弱，这也导致了其操控起来相对困难一些。

（4）超导量子比特：超导量子比特是基于超导 Josephson 结的电路系统，结构如图 12（c）。其工作的温度大约为 10mK。从尺度上看，超导量子比特属于宏观的，但是却表现出微观的量子特性。由于非线性约瑟夫森结的引入，使得超导量子比特的能级间隔变得不相等。这就可以让人们很好地利用其最低的两个能级去编码信息。根据电路拓扑和物理参数的不同，超导量子比特可以分为电荷量子比特、磁通量子比特和相位量子比特这三大类。超导量子比特之间或者与腔的耦合比较强，可以实现快速的门操作，但是这也导致了其对噪音比较敏感，从而使得相干时间比较短。

（5）腔系统：腔系统本身是将光子囚禁在一个束缚结构中的装置。最简单的腔模型是法布里-伯罗腔，其结构是两个平行的镜子，可以将光束缚在里面很长时间。除了这种腔，还有环形的回音壁腔、

蜂窝状的光子晶体腔和超导谐振腔等。图 12（d）给出了法布里-伯罗腔和回音壁模式微盘腔示意图。相比自由空间，腔中光场具有不一样的性质。腔中的光场和原子相互作用的研究形成了腔量子电动力学。在量子信息中，腔可以作为一种工具来更好地耦合光和原子等量子比特，从而实现对量子比特更好的操控，同时也可以作为多个量子比特之间耦合的数据总线。

二、量子互联网架构

2.1 量子互联网概述

量子互联网是由很多量子节点组成的一个巨大的网络系统，其构建的目的是运行经典互联网所不能实现的量子应用，比如量子通信和量子计算等。其中一个大家熟知的应用就是前面内容介绍的 QKD，可以利用量子互联网在相聚很远的用户之间实现安全量子密钥共享。目前量子互联网的发展面临很多自己独特的困难和挑战。这是因为一方面由于底层依赖的物理原理是量子理论，所以发展量子互联网并不能将已经很成熟的经典互联网的所有模式和技术照搬过来，需要大量新的探索。比如量子互联网中需要进行端到端的纠缠分发，这是经典互联网中所没有的。另一方面，由于底层量子技术的发展还处于初期，很多硬件的指标还无法满足实际应用的需求，这为量子互联网的实际部署带来很多困难。比如量子比特的相干时间不够长，量子门的操作

精度不够高等，会导致许多量子任务无法有效完成。当然量子互联网和经典互联网也有很多相似之处。比如都是网络系统，许多问题的处理都涉及图问题。还有在实际的运行上，都会面临路径选择和资源调度等。

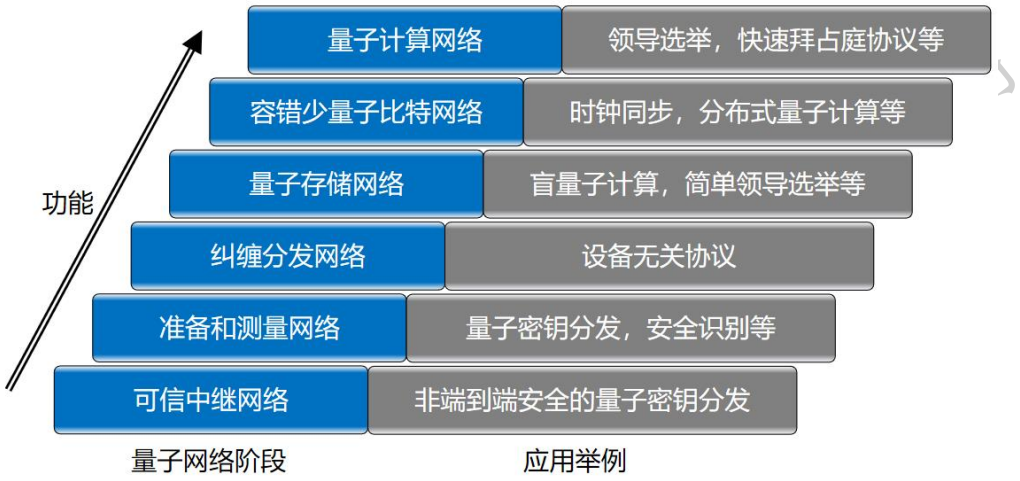

图 13. 量子互联网发展的几个阶段。

根据上层的量子应用功能和对底层技术的需求，Wehner 等人指出量子互联网的发展会经历如图 13 所示的几个关键阶段[37]。第一阶段为可信中继网络，其可以实现非端到端安全的量子密钥分发。第二阶段为准备和测量网络，其可以实现端到端的 QKD 和安全识别。该阶段可以实现端到端的量子应用功能。第三个阶段为纠缠分发网络，其可以实现端到端的纠缠分发。该阶段可以实现设备无关的量子应用协议。第四个阶段为量子存储网络，其允许端点处有局域的量子存储功能。由于量子存储的使用，使得量子互联网可以运行更多复杂的量子任务，比如纠缠纯化和一些分布式协议等。在这个阶段，可以执行量子隐形传态。该阶段典型的量子应用为盲量子计算和简单的领导选举等。第五阶段为容错少量子比特网络，其允许拥有局域的少量子比

特的容错操作能力。可实现的应用功能为更高精度的时钟同步和分布式量子计算等。第六阶段为可以任意交换量子通信的量子计算网络。典型的应用为领导人选举和快速拜占庭协议等。

在量子互联网中，底层技术的发展水平决定整个实际网络的应用部署。提高底层的硬件水平对量子网络的发展至关重要。除此之外，软件的发展也起着关键的协调作用。就像一台计算机，除了硬件水平过硬之外，还需要很流畅的操作系统协调所有的硬件来完成最终的任务。单纯只有好的硬件设备，例如性能很好的 CPU，没有流畅的操作系统，整个电脑也无法有效展现出其最好的算力。而对于互联网来说，协议栈就像是一个操作系统，协调着整个网络的运行。因此量子互联网协议栈的研究也必不可少。而协议栈的具体架构会受到各层协议和技术的影响，比如网络运行模式和量子中继的类型就是很重要的因素。本书主要关注量子互联网体系架构等方面的研究，所以会主要聚焦于网络运行模式和协议栈等方面的内容。接下来会介绍量子中继、协议栈和量子数据交换等内容。更多关于量子互联网内容可参考文献[37-39]。

最近二十年，世界上部分国家和地区推动和布署了外场量子网络，例如美国 DARPA [40]、欧洲 SECOQC[41]、瑞士 SwissQuantum[42]、日本 Tokyo[43]、中国的量子城域网[44-47]和基于量子科学实验卫星的天地网[48,49]等。由于目前量子中继技术在实验上实现的难度较大，以上这些布署的量子网络都是基于可信中继的量子通信网络。而最近中国科学技术大学团队在外场实现了基于量子中继的多节点纠缠分

发网络，最远节点距离达到 12.5 公里[50]。

2.2 量子中继及其分类

量子中继类型	原理和操作步骤
第一代	预报式纠缠分发 + 纠缠纯化 + 纠缠交换
第二代	预报式纠缠分发 + 量子纠错码 + 纠缠交换
第三代	量子纠错码
全光	簇态产生 + 纠缠交换

表 2. 四类量子中继的原理。

量子中继是实现远距离量子通信的关键部分。由于信号衰减和噪音的影响，光子丢失和量子态退相干严重阻碍着远程量子通信的实现。而量子中继器则是分而治之，通过将远距离的量子通信分割为多个短距离通信，从而实现远距离量子通信的目标。目前量子中继大致可以分为四类，分别为第一、二、三代和全光中继。各类量子中继的原理如表 2 所示。第一代中继的原理是首先在相邻节点之间实行预报式纠缠产生，然后纠缠纯化提高保真度，最后纠缠交换来延长纠缠信道的距离，最终建立起长距离的纠缠信道。预报式纠缠分发可以克服光子丢失错误，纠缠纯化可以弥补操作错误。第二代量子中继是采用纠错码的方式替代了纠缠纯化来弥补操作错误，对于光子丢失错误依然采用预报式纠缠分发。第三代量子中继都采用纠错码的方式来弥补光子丢失和操作错误。全光中继则是先在相邻节点之间产生图态，而后做纠缠交换和 Bell 态测量来延长纠缠信道。由于使用图态，可以降低对

量子存储的要求。对量子中继的梳理可以参考文献[51-53]。

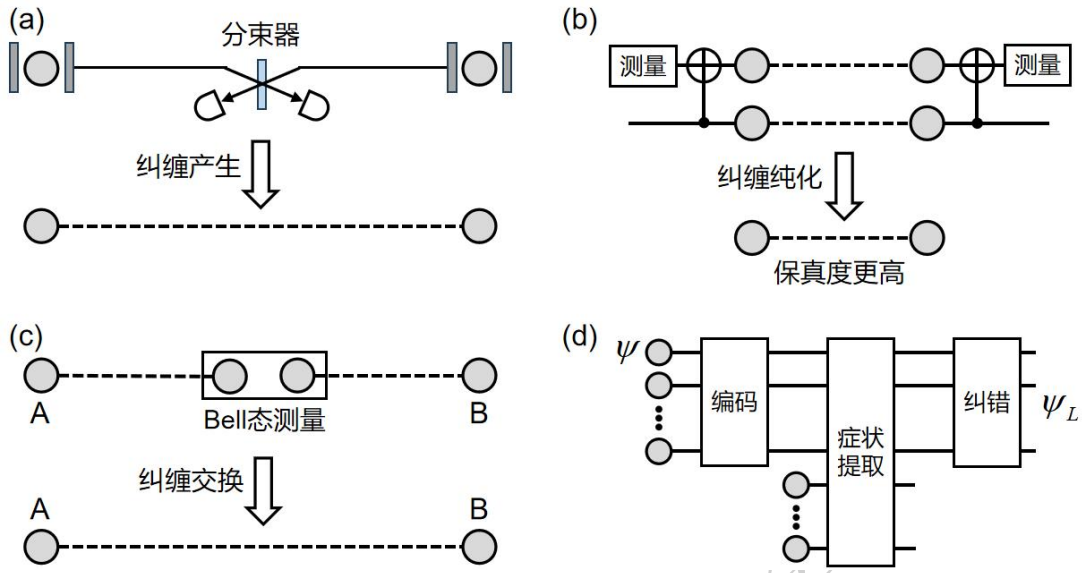

图 14. 量子中继的相关量子操作过程。（a）预报式纠缠产生（基于中间点测量的方式）。（b）非局域纠缠纯化。（c）纠缠交换。虚线代表纠缠关联。（d）量子纠错码。

首先介绍预报式纠缠分发、纠缠纯化、纠缠交换和量子纠错码。

（1）预报式纠缠产生是用测量后选择的方式预报式地在相邻节点之间产生纠缠态。和直接将制备好的纠缠态光子传输到两个节点相比，这种预报式的方式可以有效解决光子丢失错误带来的影响。在量子中继中，典型的预报式纠缠产生方案之一是如图 14（a）所示的中间点测量方案。两侧的节点各自将与存储量子比特纠缠的光子发送到中间点处进行 Bell 态测量，根据测量结果来预报式选择成功建立两侧节点纠缠的事件。

（2）纠缠纯化是通过局域操作和经典通信来提升纠缠态保真度的过程[54-58]。其可以用于解决量子中继中操作错误带来的纠缠态保

真度降低的问题（量子态保真度是衡量一个实际量子态和目标量子态相似程度的物理量）。纠缠纯化需要牺牲额外的纠缠资源。典型的非局域纠缠纯化方案是 Bennett 等人提出的对两对纠缠态做局域的双边 CNOT 门操作，然后测量靶位纠缠粒子来比对结果，保留源位保真度更高的情况[54]。Bennett 等人纠缠纯化方案如图 14（b）所示。

（3）纠缠交换是通过对两对纠缠粒子中的每对的其中一个进行联合的 Bell 态测量实现剩余未测量两个粒子之间的纠缠，如图 14（c）所示。在量子中继中，可以用纠缠交换来延长纠缠信道的距离。实际上纠缠交换是对纠缠态中的一个粒子进行量子隐形传态操作。

（4）量子纠错码是通过在大量的物理比特上编码逻辑比特来纠错的一种手段。由于系统上的冗余编码，可以通过测量辅助比特提取数据中的症状信息来恢复量子态[59-61]。量子纠错码的大概步骤如图 14（d）所示。典型的量子纠错码有 Shor 的 9 量子比特码、CSS 码、稳定子码和表面码等。量子纠错码通常被用于容错量子计算和长距离量子通信。

四类量子中继器操作原理大概如下：

（1）第一代量子中继[62-65]如图 15（a）所示，先将长距离的通信节点分为若干个短距离的节点。然后在两两相邻节点之间做预报式纠缠产生，待相邻节点之间的纠缠信道形成以后，根据需要来进行纠缠纯化以保证纠缠态的保真度。随后在中间节点之间做纠缠交换延长纠缠信道的距离，直到形成端到端的纠缠分发。

（2）第二代量子中继[66-68]如图 15（b）所示，首先在相邻节

点之间产生编码的逻辑 Bell 态。然后在中间节点里成对的物理量子比特之间做 CNOT 门和测量来实现编码的 Bell 态测量。通过以上步骤就可以实现端到端的编码 Bell 态。

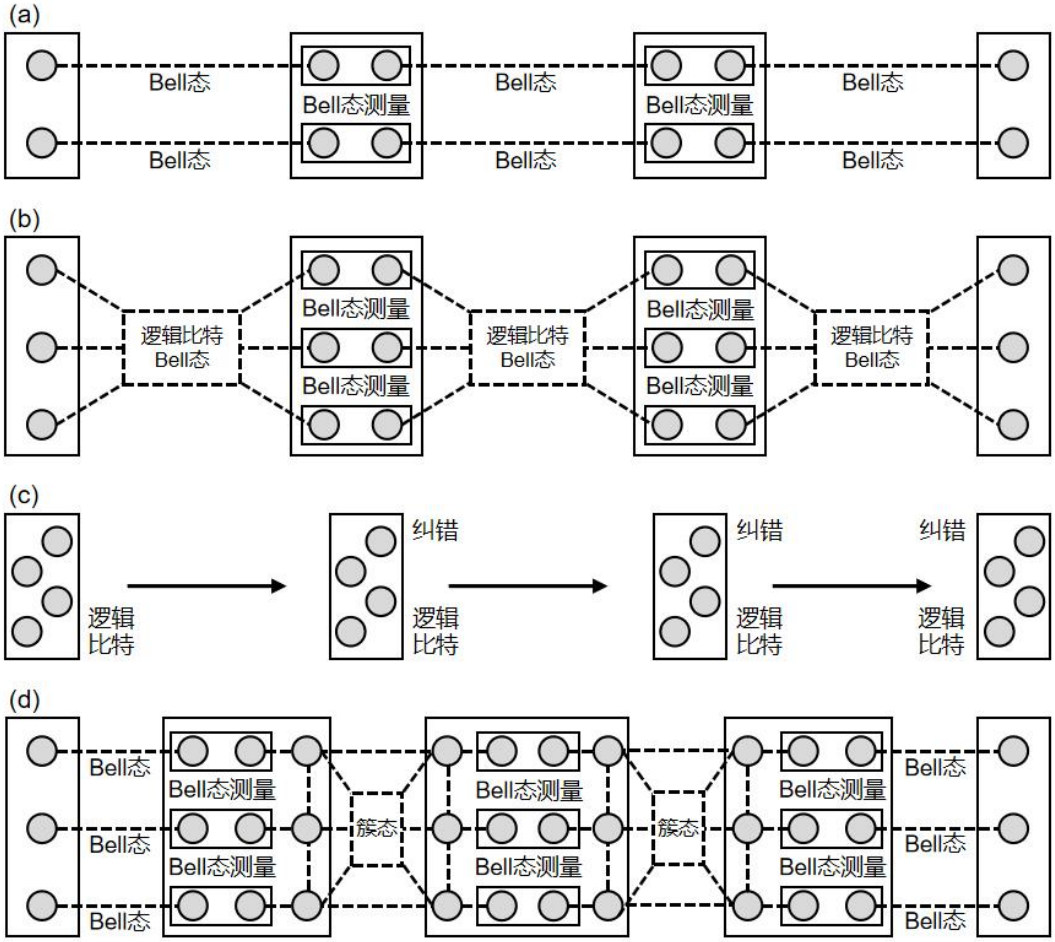

图 15. 四类量子中继模型。（a）第一代量子中继。（b）第二代量子中继。（c）第三代量子中继。（d）全光量子中继。虚线为纠缠关联。图中只画出大概的框架，涉及的预报式纠缠分发，纠缠纯化，编码和纠错等具体过程并未画出。

（3）第三代量子中继[69-71]如图 15（c）所示，其直接编码量子态，将编码后的量子态直接传送到下一个中继节点，在每个中继节点处进行纠错，然后继续传输到下一个中继节点，如此重复，直到量

子比特达到接收方。

（4）全光量子中继[72,73]如图 15（d）所示，首先在两个用户和其相邻节点之间产生纠缠信道，同时在除两个用户节点以外的相邻两个节点之间产生簇态。然后在所有中间节点的对应的量子比特之间做 Bell 态测量。如果 Bell 态测量成功，则继续在一级叶量子比特上进行 X 测量，同时在剩下的一级叶量子比特上进行 Z 测量。最后将所有的测量结果发送给用户双方，最终实现端到端的纠缠分发。

除了以上四类量子中继外，还有一类适用于近期量子互联网的安全经典中继[74]。利用 QSDC 和经典抗量子密码结合的方式来实现端到端的安全通信。其原理主要是在相邻节点之间用 QSDC 传输信息，这样就可以保证信息在信道的传输过程中是安全的。在节点处信息的安全性则依靠经典的后量子密码来维持。相比于纯量子中继，该中继模型的实现难度更低，可以在近期资源和技术受限的量子网络部署中起到一个过渡的作用。

2.3 量子互联网协议栈

量子互联网协议栈的提出在很大程度上借鉴了经典互联网协议栈架构的思想。所以在介绍量子互联网协议栈之前，我们先大概介绍一下经典互联网协议栈。

2.3.1 经典互联网协议栈模型

经典互联网协议栈目前广泛应用的模型有 OSI 七层模型和 TCP/IP 四层模型。如图 19 所示。实际应用中，通常将这两个模型合

成一个新的五层模型。将 OSI 模型中的会话层、表示层和应用层合并成一个应用层。而将 TCP/IP 模型中的网络接口层用 OSI 模型的物理层和数据链路层替代。新的五层模型从下到上分别为物理层、数据链路层、网络层、传输层和应用层。物理层负责将比特编码到物理载体上并通过物理介质进行传输。数据链路层负责在两个相连节点之间传输数据。网络层负责将数据从一个节点路由到另外一个节点，最终将数据传输到目的节点。传输层负责端到端的数据传输，应用层负责在用户层面执行各种应用协议。

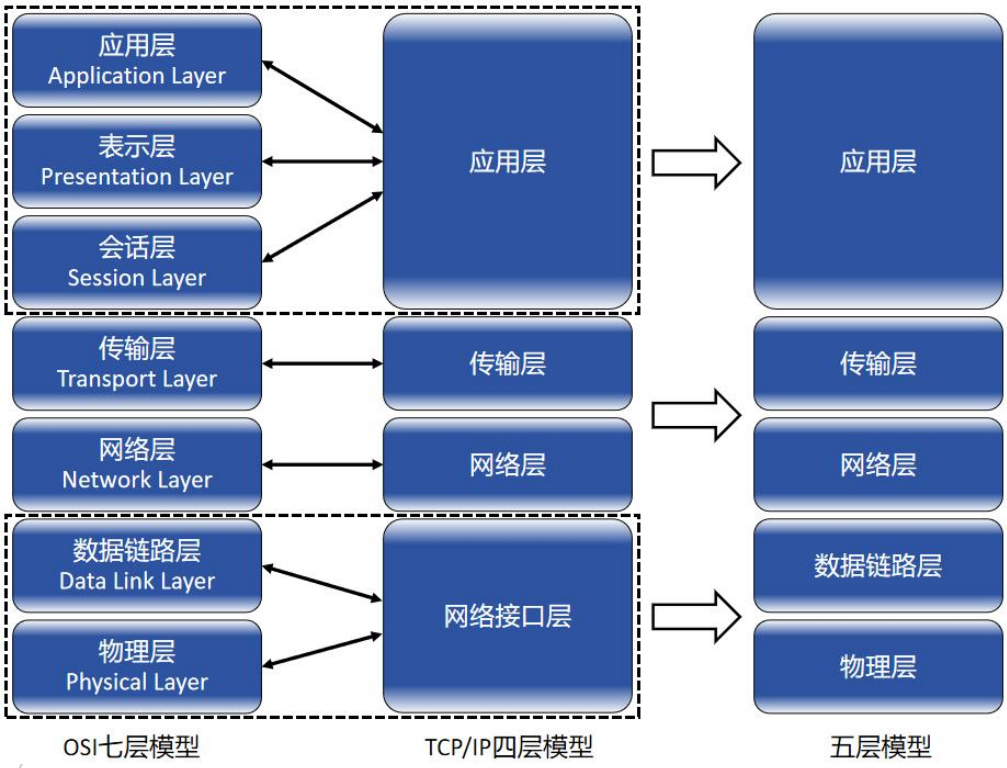

图 19. 经典互联网协议栈模型。

2.3.2 量子互联网协议栈方案

和经典互联网类似，要想顺利运行量子互联网中的任务，设计一个合理高效的网络协议栈具有很重要的现实意义。由于量子互联网底

层依赖的物理原理是量子力学，比如单光子和纠缠态的性质和操控等，和经典互联网完全不一样。所以量子互联网的许多架构的设计方案可能已经无法完全复制经典互联网模式。最近代尔夫特理工大学研究组开发并实验实现了平台无关的量子网络节点的操作系统，在这之前实验上可以演示的操作软件几乎都是根据具体的实验设备和量子网络应用协议和功能而专门设置的[75]。近期 Cisco 在理论上设计了可扩展量子网络数据中心的架构，该结构可以互联多个量子处理器，推动大规模量子计算的发展，并联合加州大学圣塔芭芭拉分校推出量子网络纠缠芯片[76]。如图 20 所示，这里简单介绍一下现存的几类协议栈方案。更多关于量子互联网协议栈的内容可以参考文献[77,78]。

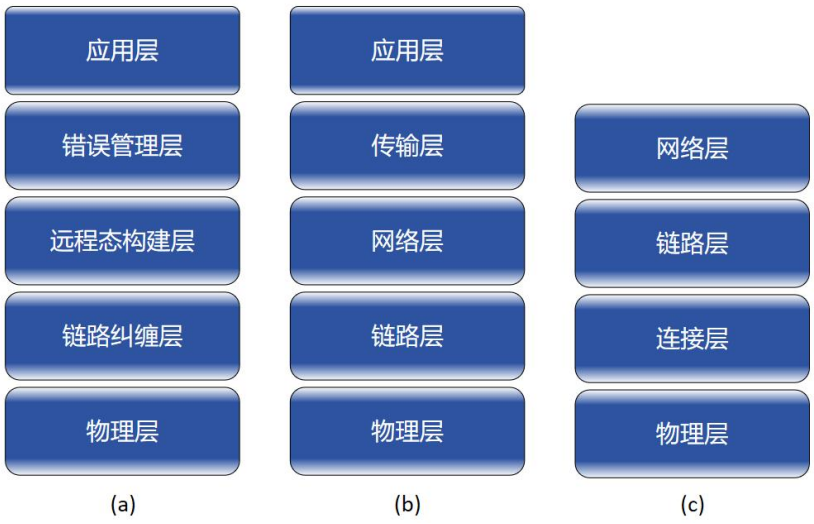

图 20. 量子互联网协议栈方案。（a）Van Meter 等人方案[79-82]；（b）Wehner 等人[83-86]、Li 等人[88]和 Bacciottini 等人[89]方案；（c）Dür 等人方案[87]。

（1）Van Meter 等人（日本庆应义塾大学）方案

Van Meter 等人针对第一代量子中继器的量子互联网提出了五层

协议栈[79-82]，如图 20（a）所示。其从下到上分别为物理层、链路纠缠层、远程态构建层、错误管理层和应用层。物理层主要包含关于量子比特的基本操作，例如发射和吸收光子、编码、相邻节点的纠缠产生和测量等。而链路纠缠层则包括控制物理层运行的所有经典控制信息。这一层主要负责对物理层量子比特操作的指令的管理，例如决定某时某量子比特之间的纠缠产生。远程态构建层负责实现端到端的纠缠分发，值得注意的是错误管理作为与远程态构建层的一个互嵌层。应用层则是各类量子应用的执行，例如量子通信和量子计算等。

（2）Wehner 等人（荷兰代尔夫特理工大学）方案

Wehner 等人提出的量子互联网协议栈和经典的五层协议栈在分层和名称上是一致的，但是每层所对应的功能却不同[83-86]，如图 20（b）所示。物理层主要负责实现相邻节点的纠缠产生。链路层则是实现更鲁棒的纠缠信道。网络层实现端到端的纠缠分发。传输层负责传输量子比特数据。应用层依然是各类量子应用。和 Van Meter 等人的方案不同的是，这里的方案并没有考虑将纠错单独作为协议栈的一个层。

（3）Dür 等人（奥地利因斯布鲁克大学）方案

Dür 等人主要考虑的是预先构建模式的量子互联网协议栈[87]，如图 20（c）所示。预先构建模式是指在请求之前提前构建纠缠信道。在纠缠资源方面考虑了两方和多方纠缠情况。该协议栈包括四层，从下到上分别为物理层、连接层、链路层和网络层。这里的物理层功能等同于上一个模型中的物理层和链路层，其主要负责产生短距离鲁棒

的纠缠态，物理上连接量子网络的器件，还负责量子存储和信号转换等。连接层主要负责构建点对点和点对多点的量子连接。链路层负责根据网络的需求来产生任意的图态。网络层则实现网络之间的图态的产生。该模型将端到端的纠缠互联作为最终目标来考虑整个协议栈的构建，没有提及更上层的量子应用。所以这里并没有应用层。

（4）Li 等人（中国科学技术大学）方案

Li 等人也针对两方纠缠的量子互联网提出了一个五层协议栈[88]。该方案和 Wehner 等人的模型类似，如图 20（b）所示。但是不同的是 Wehner 等人考虑的纠缠产生是按需构建，而 Li 等人考虑的是预先构建的。这种差异就导致了协议栈在链路层上的功能有差异。所以 Li 等人的协议栈链路层负责控制链路纠缠和处理网络层反馈的信号。

（5）Bacciottini 等人（美国马萨诸塞大学默斯特分校）方案

Bacciottini 等人借鉴经典互联网的模式提出基于分组交换的尽力而为的量子网络架构，并给出了量子网络协议栈[89]。该量子网络协议栈名称上和传统互联网的五层协议栈是一致的，与 Wehner 等人和 Li 等人的协议栈架构上也是相同的，如图 20（b）所示。其物理层包括量子硬件和节点中实现量子操作所需的经典控制硬件，如纠缠交换、内存量子比特操作和纠缠产生用到的中间步骤；链路层提供一个用于请求和获得链路预报式纠缠产生的接口；网络层通过纠缠交换消耗链路纠缠来产生端到端纠缠；传输层从端到端层面管理纠缠流；应用层在传输层之上消耗端到端纠缠来实现自定义逻辑。Bacciottini 等人明确指出该方案是基于分组交换网络模式构造的无连接量子网络。

三、量子互联网分组交换技术

3.1 基于量子封装网络的分组交换方案

为了让量子互联网可以像经典互联网那样运行分组交换模式[90]，美国加州大学戴维斯分校 Yoo 和美国西北大学 Kumar 在 2021 年首次提出利用经典-量子混合数据报实现量子封装网络的思想[91]。随后 Cisco 的 DiAdamo 等人在 2022 年提出了用经典-量子混合帧结构实现量子互联网分组交换的方案[92]。在 2024 年，美国加州大学戴维斯分校和美国西北大学研究组进一步解释了量子封装网络的概念[93]，并在实验上首次演示了对单光子数据报的交换功能[94]，随后进一步实现端到端的纠缠分发[95]。下面我们以 Cisco 的方案为例，详细阐述量子分组交换的基本原理。

在他们的方案中，实现分组交换量子网络的关键是经典-量子混合帧结构。其结构如图 16（a）所示。图（a）上侧的结构是经典互联网帧结构，其包括包头、负载和包尾。简单来说包头包含了路由和纠错等一些关键信息。包尾则是预示着整个帧的结束。负载是需要传输的被编码的信号。由于包头中携带的用于路由的信息，比如地址信息等，所以当帧到达某个节点以后，处理器会读出包头中的地址信息，然后分配下一个通道。由此将整个帧从发送方顺利传输到接收方。整个帧在互联网中的传输路径都是包头引导的。图（a）下侧是 Cisco 方案的经典-量子混合帧结构。结构和经典帧是一样的，只是将经典

帧中的经典负载换成了量子负载。所以在传输过程中，这里的负载将会是量子信号。图（b）是混合帧的产生机制。利用控制单元控制经典发射器发射经典包头信号和量子源发射量子负载信号，然后利用多路复用器将包头和负载结合，形成混合帧。多路复用可以是时分复用和波分复用。图（c）是混合帧的信号处理过程。当混合帧到达一个节点以后，通过解复用器将经典的包头和量子负载信号分开。随后用光开关将经典包头传送给经典处理器，量子信号则传输给量子存储。当经典包头处理结束，多路复用器又将其和量子负载组合后传输。

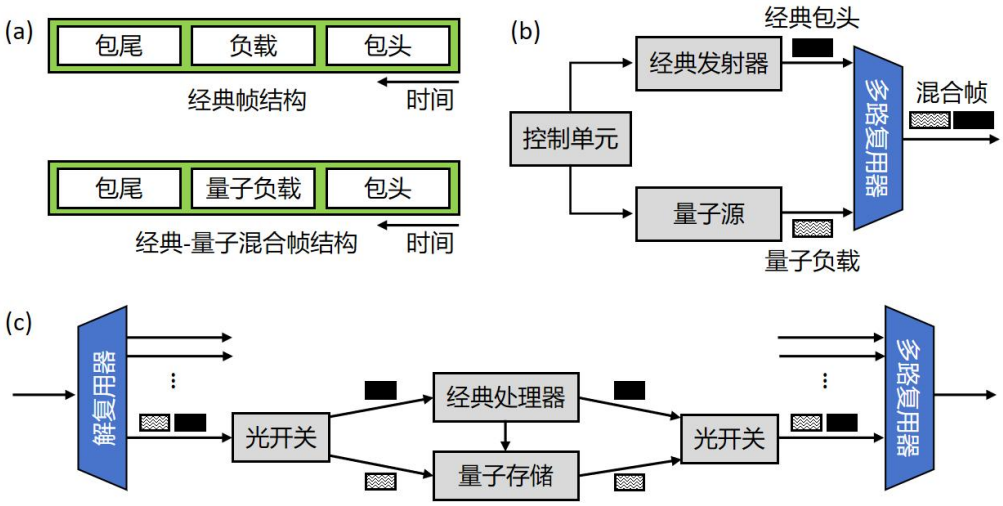

图 16. 量子互联网分组交换混合帧结构 （a）经典互联网帧结构和量子互联网混合帧结构。（b）经典-量子混合帧的产生。（c）混合帧的信号处理过程。

图 17 给出了一个利用混合帧结构在量子网络中实现分组交换模式的量子信号传输的例子。发送方 A 将需要传输的量子信号作为混合帧的负载部分。混合帧的包头编码接收方 B 的地址。当混合帧传输到节点 1 时，包头部分将会被经典处理器处理，比如读取包头的有效

信息。该节点会根据包头中携带的地址信息来决定出帧需要被发送的下一个节点。而量子负载则存储于节点 1 的量子存储中等待经典包头再次被发送。经过一段时间后，经典包头处理完成。其产生的包头再次和量子负载结合成帧被发送到选择好的下一个节点 2。混合帧到达节点 2 后，同节点 1 一样处理帧，随后将帧传输到节点 4。最后混合帧通过节点 4 的转发到达接收方 B。以上是考虑有量子存储情况，如果不考虑量子载荷的存储，那么可以采用即时交换方案，即发送方一开始估算好其与接收方之间的距离和大约需要经过的节点数。通过计算经典包头在所有中间节点的总处理时间，然后在发送混合帧的时候，在经典包头和量子负载之间预留出足够的间隔时间。这样就可以保证在每个节点量子负荷总是在经典包头处理结束以后才到达该节点。随着经过的节点数的增多，负载和包头之间的时间间隔也逐渐减小。当包头到达接收方，量子负载也几乎紧随其后到达。

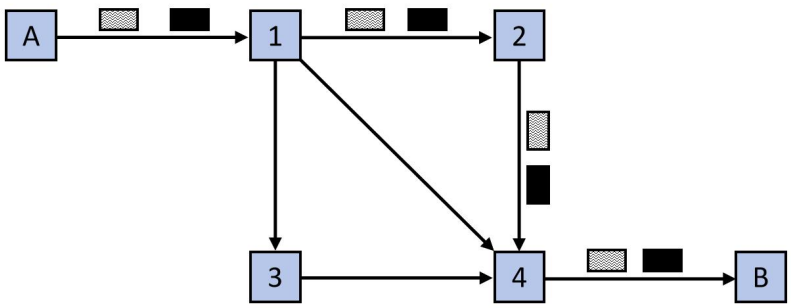

图 17. 量子互联网分组交换模式传输量子信号的一个例子。

根据应用层协议和物理层原理的不同，量子互联网一般分为单光子网络和纠缠网络。对于单光子网络，最直接的应用就是基于单光子的 QKD 协议。如果考虑对单光子网络应用以上分组交换模式，发送方可以直接将需要传输的单光子作为混合帧的量子负载传输到接收

方。对于纠缠网络，应用很广泛，除了基于纠缠的 QKD 协议，还有分布式量子计算等。如果考虑对纠缠网络运行该分组交换模式，我们一般考虑的都是端到端纠缠分发阶段。如图 18 所示，可以将纠缠分发分为两种情况。第一种叫中间点分发，另外一种叫发送方分发，分别如图 18（a）和（b）所示。（这里举例子为发送方发送帧，也可以采用接收方发送。因为很多时候纠缠分发的目的只是实现端到端的纠缠分发，无所谓是哪一方发送。但是因为通常情况下是发送方主动发送通信请求来联系接收方，所以采用发送方发送帧的方式更方便和直观。）这里拿一个纠缠对举例，对于中间点分发情况，可以将两个纠缠光子分别装载于两个混合帧当中作为量子负载。然后分别传输到发送方 A 和接收方 B。而对于发送方分发的情况，需要将其中一个光子存储在发送方 A 的量子存储中，而将另外一个纠缠光子装载于混合帧中发送到接收方 B。

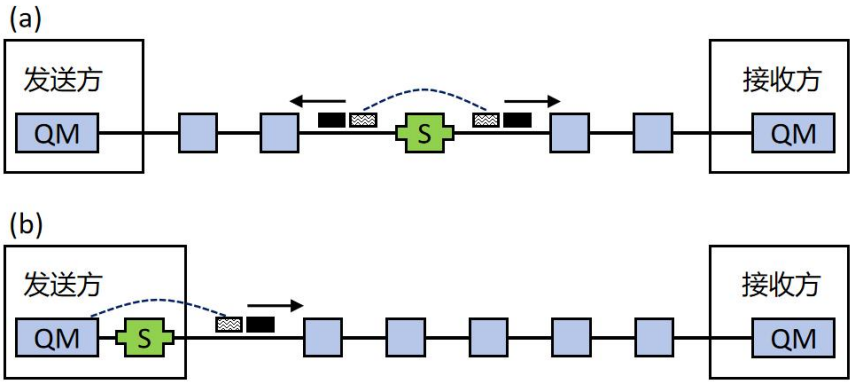

图 18. 纠缠分发网络的分组交换。（a）中间点分发方案。（b）发送方分发方案。S 为纠缠源，QM 为量子存储。浅蓝色方块为节点。黑色和条纹长方形分别为经典包头和量子负载。虚线代表纠缠关联。

3.2 经典帧辅助的混合分组交换方案

上面的分组交换方案理论上可以很好地应用于单光子网络和部分场景下的纠缠网络。对于纠缠网络的某些情况，比如第一代和第二代基于中间点的预报式纠缠产生量子中继网络，其无法有效实施。为了解决上述问题，扩展量子网络分组交换的适用范围，我们提出一个直接针对纠缠网络的经典帧辅助的混合分组交换方案[96]。方案整体思路为借助经典帧来决定路径，通过给帧的每一跳分配合适的纠缠信道，结合纠缠交换来扩展信道范围，最终实现端到端纠缠分发。在该过程中，只需要借助经典帧、纠缠信道分配和纠缠交换。由于使用到了经典帧辅助量子信道分发，所以我们称该方案为混合分组交换。同时由于该方案并不依赖于纠缠产生的方式而适用于所有的纠缠产生方案，也称作纠缠产生无关方案。具体方案如下。

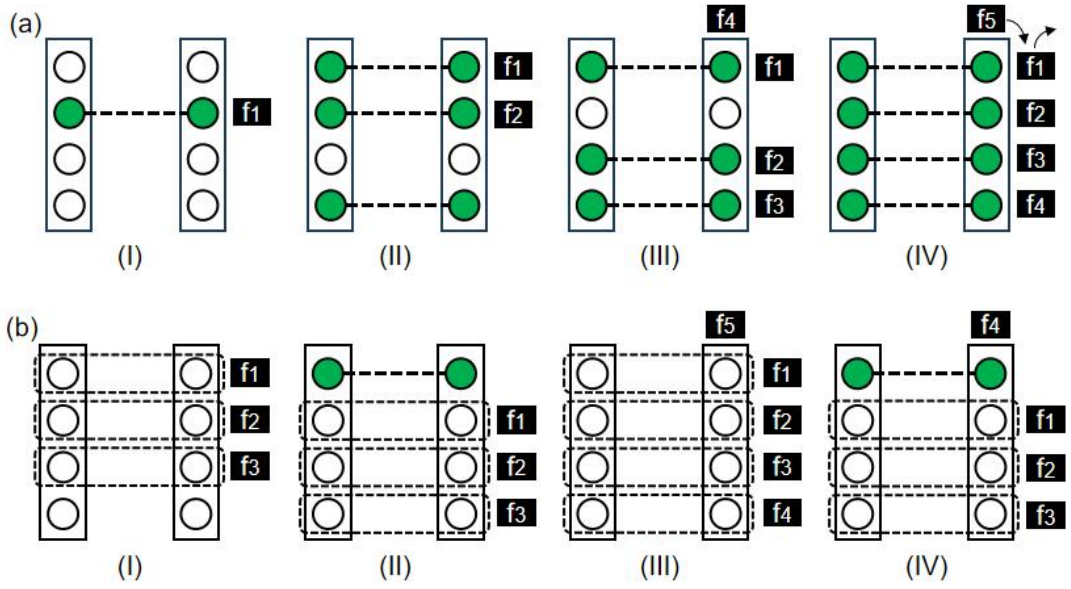

图 21. 相邻节点之间经典帧辅助的纠缠信道资源分配。（a）场景 1：无导向纠缠产生。（b）场景 2：经典帧导向的纠缠产生。

3.2.1 纠缠信道分配

首先考虑给经典帧分配相邻节点之间的纠缠信道资源。如图 21 所示，相邻节点之间的纠缠产生过程分为两种情况讨论：图（a）中的无导向纠缠产生和图（b）中的帧导向纠缠产生。对无导向纠缠产生，首先需要量子网络中所有相邻节点间一直重复进行纠缠产生。如果相邻节点间的量子比特之间纠缠产生成功，则暂时终止该两个比特之间的纠缠产生，直到这两个量子比特空闲时（无纠缠时），才继续执行纠缠产生。如果纠缠态存储时间太长，导致纠缠态保真度低，可以执行纠缠纯化或者纠错等来提高保真度。当然也可以设定一个阈值，当纠缠信道存储时间超过阈值，保真度太低，直接舍去该信道，重新产生一个高保真度的纠缠信道。在无导向情况中，网络中的相邻节点之间会保持一定数量的纠缠信道资源可供使用。当经典帧从上游节点到达下游节点时，控制单元会根据一定的调度算法分配一条纠缠信道给该帧（只考虑一条纠缠信道情况）。在该量子比特没有释放前，不允许将该信道及其量子比特用于其他用途。当多个经典帧从同一个上游节点到达同一个下游节点时，会有相应的排队机制给帧分配纠缠信道。比如按照时间先后顺序来排队，可以简单分为四种情况。第一种为图 21（a）中的（I），一个帧对应一个纠缠信道。这时候控制单元可以直接将该信道分配给该帧。第二种是图 21（a）中的（II），到达帧数量小于现存的纠缠信道资源，控制单元按照到达的先后顺序从上到下分配纠缠资源。比如 f1 和 f2 分别分配了从上至下第一个和第二个纠缠信道。第三种为图 21（a）中的（III），到达的经典帧数量超

过已有的纠缠信道。控制单元按照帧到达的顺序分配对应空间上的纠缠信道。未分到的帧则需要等待新的纠缠信道的产生。第四种是图21（a）中的（IV），所有量子比特都有纠缠信道，到达的经典帧数量多于比特数量。没有被分配到的帧需要等待被占据的量子比特被释放，然后新的空闲纠缠信道产生以后才可以被分配。

第二种场景为图21（b）中经典帧导向的纠缠产生，即帧被分配到量子比特后，控制单元启动纠缠产生程序来建立纠缠信道。相比无导向方案，这里纠缠信道的建立会比较慢，经典帧到达并被分配量子比特后方可为其产生纠缠，但是在网络业务量少的时候，更节省网络资源。对于帧导向的纠缠产生，我们也需要根据帧到达的数量来合理为其分配量子比特。简单分为四种情况讨论：第一种为图21（b）中的（I），所有的量子比特都是空闲状态，所到的经典帧数量少于量子比特数。节点按照帧到达的时间顺序分配空间的量子比特资源。图中虚线框表示框内的量子比特对已被划分给右侧相应的帧。第二种为图21（b）中的（II），部分量子比特因为被之前经过的帧占用，还未释放，到达的经典帧数量少于等于空闲量子比特数，此时控制单元将空闲的量子比特按照先后顺序合理分配给对应的帧。第三种为图21（b）中的（III），所有量子比特均为空闲状态，到达的帧数量多于比特数。此时按照时间顺序分配完量子比特，未被分配到的帧则需要等待再次释放的量子比特对。第四种为图21（b）中的（IV），有部分量子比特被占用，到达的帧数量多于空闲比特数。此时帧只能从空闲量子比特中分配，未被分配到帧需要等待新的被释放的量子比特。

3.2.2 纠缠交换和信道延长

当经典帧被分配纠缠信道以后，就需要进一步延长纠缠信道到更远的节点。这里有两个问题需要解决：（1）如何选择下一个节点；（2）如何延长纠缠信道到下一个节点。

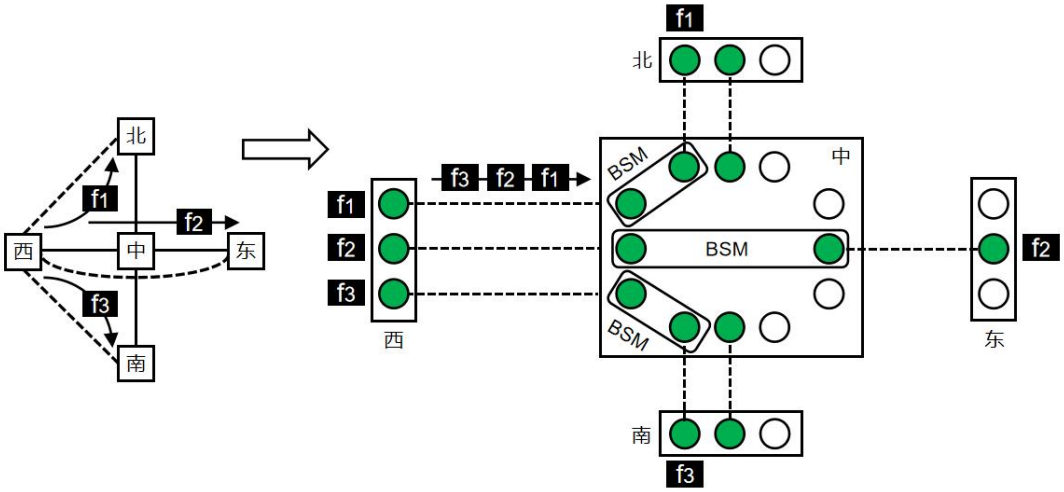

图 22. 经典帧辅助的量子纠缠信道延长方案。BSM 为 Bell 态测量。

图 22 给出了一个包含 5 个节点的十字型结构。假设我们需要建立西部节点和北、东、南节点的三个纠缠信道。由于帧 f1，f2 和 f3 的地址分别是北方、东部和南方节点。当这三个帧从西部节点传输到中间节点时，经典处理器会读取帧包头中的地址信息来决定下一个节点，随后将帧发送到下一个对应的节点。当节点在处理经典帧的同时，控制单元会按照上一节内容介绍的相邻节点纠缠信道分配方法给这三个帧分配纠缠信道。以无导向纠缠产生情况为例，当 f1，f2 和 f3 分别到达下一个目的节点后（对应图 22 中的北、东、南节点），控制单元会分配一个中间节点和目的节点之间的纠缠信道。然后中间节点会在每个帧所属的量子比特之间执行 Bell 态测量来完成纠缠交换。于是西部节点和各个帧的目的节点之间的纠缠信道建立完成。这里的

Bell 态测量，既可以是帧到达下一个节点后，当纠缠信道分配完成，由下一个节点发送消息到上一个节点让其执行，也可以是根据实际部署的节点长度等信息来提前商定一个时间来执行。例如根据帧选择的路径，提前计算帧到达下一个节点的时间以及纠缠信道分配的时间来设定一个时间 T 去执行，当帧离开该节点 T 时间后，自动执行 Bell 态测量。在这个结构中，通过上述方案，既完成了路径的选择，也完成了纠缠信道的延长。

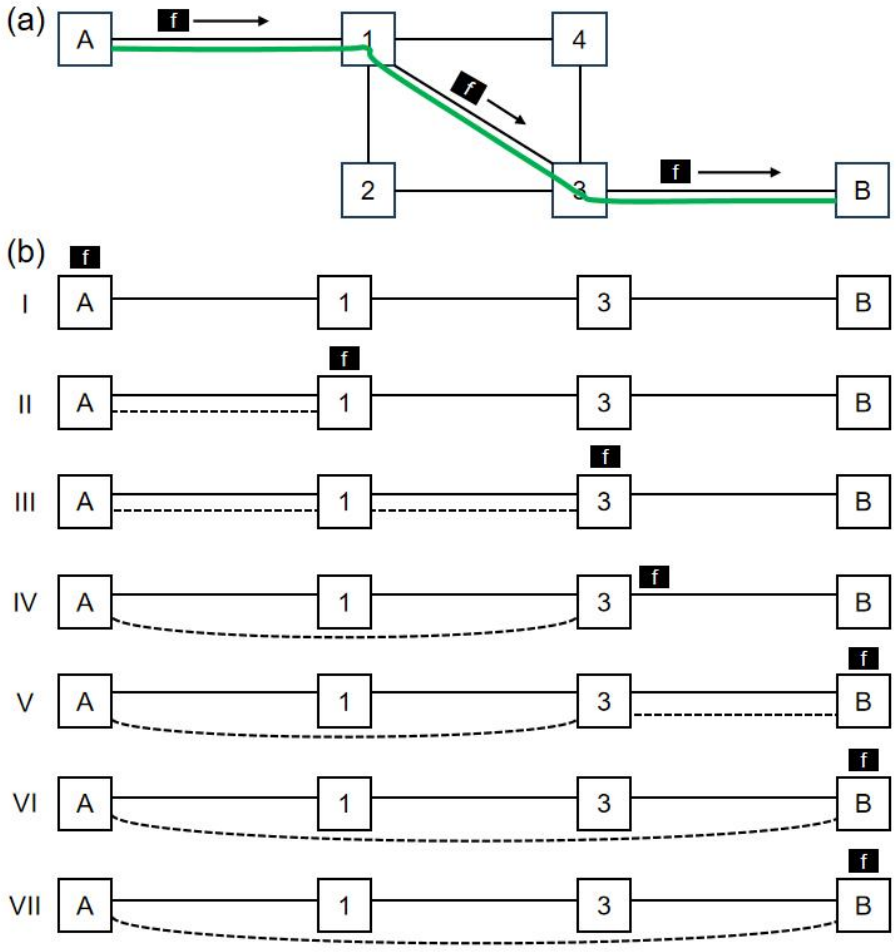

图 23. 经典帧辅助的混合分组交换端到端纠缠分发。标注 f 的黑色方框代表经典帧。虚线部分代表纠缠信道，黑色实线部分代表物理连接的通道。（a）网络拓扑和最终的线路。绿色线条代表路最终径。（b）

A-B 之间端到端纠缠分发的 7 个阶段。只给出 A-1-3-B 四个节点。

3.2.3 端到端的纠缠分发

以图 23（a）中的量子网络为例阐述端到端纠缠分发。发送方为 A，接收方为 B。目标是实现 A 到 B 的纠缠分发。过程如图 23（b）所示。（**I**）发送方 A 产生一个包含接收方 B 地址等信息的帧并将其发送到节点 1。考虑无导向的纠缠产生。（**II**）：帧到达节点 1，其包头被经典处理器分析。通过查找本地路由表和比对包头中的地址信息，选择节点 3 作为下一个目标节点。随后帧会被发送到节点 3。当帧被处理时，控制单元会给该帧分配一个节点 A 和 1 之间的纠缠信道。图中虚线部分代表纠缠信道。（**III**）帧到达节点 3 后被处理，包头信息被读取。节点 B 被选为下一个目标节点。同时该帧被分配一条纠缠信道。随后帧被发送到下一个通道。（**IV**）节点 1 中的经典处理器控制量子设备执行纠缠交换。测量结果通过经典网络发送给接收方 B。此时经典帧所对应的纠缠信道扩展为 A 和节点 3 之间。此时的帧还在 3 和 B 之间的通道传输（这里假设了 Bell 态测量时间小于帧在 3 和 B 通道传输时间）。（**V**）经典帧到达接收方 B，包头被处理。控制单元分配给帧一条节点 3 和 B 之间的纠缠信道。（**VI**）节点 3 执行纠缠交换，测量结果发送给 B。此时端到端的纠缠信道形成。（**VII**）节点 B 接收到所有的 Bell 态测量结果。根据结果节点 B 对纠缠信道做对应的量子操作来修正纠缠态。同时给 A 发送任务完成的确认信息。任务结束。

3.2.4 讨论

在上述方案中，我们需要通过运行经典互联网分组交换技术来辅助完成量子信道的纠缠分发。而整个过程中，被划分的量子信道逐跳延长和经典互联网的分组交换有类似之处。在以上这个例子中，我们只假设了一条纠缠信道，其实根据需要可以分发多条用于其他用途，比如单次分发多条端到端的纠缠信道用于量子任务执行，还有多条信道用于节点之间的纠缠纯化和纠错。除了以上一个帧的例子，还可以同时执行相同用户之间的多个帧请求，或者不同用户之间的多个帧请求。从例子中可以看出这里的混合分组交换和经典分组交换有所不同，当帧离开上一个节点以后，对应的量子资源不能立即释放给其他用途，需要等到 Bell 态测量结束后才可以释放。如果需要再次建立相邻节点之间的纠缠信道，需要等到两个节点的量子比特都被释放。和 DiAdamo 等人的方案不同的是，这里的混合分组交换方案直接使用经典帧，不需要合成经典-量子混合帧。在这一点上可以直接使用经典互联网的基础设备，省略了一些额外的操作。除此之外，由于帧的传输和纠缠信道的建立是两个独立过程，所以并不需要经典帧和量子信号在传输上有时间关联。整个网络的中间节点间可以采用不同的纠缠产生方式，在这一点上具有更高的灵活度，使得该方案对纠缠分发的过程具有很好的兼容性和鲁棒性。

四、量子互联网运行模式设计

在量子互联网发展的初期，由于物理层各类量子技术的限制，许

多器件和功能相对来说都不成熟。对于该阶段量子互联网的研究，需要考虑其实际的硬件限制。对此我们提出了一套初期少资源量子互联网运行模式来执行任意用户之间的任务请求[97]。

在量子设备资源少的情况下，如何使量子网络满足运行有一定需求的量子应用，比如最小保真度、最小吞吐量和低时延等，是一个很重要的问题。这里考虑量子互联网中可以兼容第一、二和三代量子中继器技术。

4.1 基本假设

首先假设研究的量子网络有以下几个特点：

（1）网络设备数量少，量子内存小；

（2）各类量子应用要求多，如低时延、最小保真度和最小吞吐量等；

（3）量子比特相干时间短，转移量子态会降低态的质量；

（4）传输量子数据可以有很多技术，比如利用量子纠错码逐跳转发和利用远程纠缠态进行量子隐形传态；

（5）量子纠错码和纠错方式需要根据业务的需求、路径的资源和质量等条件确定，远程纠缠态的构建需要路径上的量子路由器和中继器等进行协同操作。

4.2 量子网络设计整体要求

（1）量子网络布局：整个网络分为用户网络和主体网络两部分。

如图 24 所示，主体网络为中央虚线所包含的区域，也就是不包含用户节点的网络节点部分。用户网络则为用户和其最近邻的路由节点组成的网络区域，例如右下角虚线框内 C1，C2 和 3 这三个节点组成的局域用户网络。

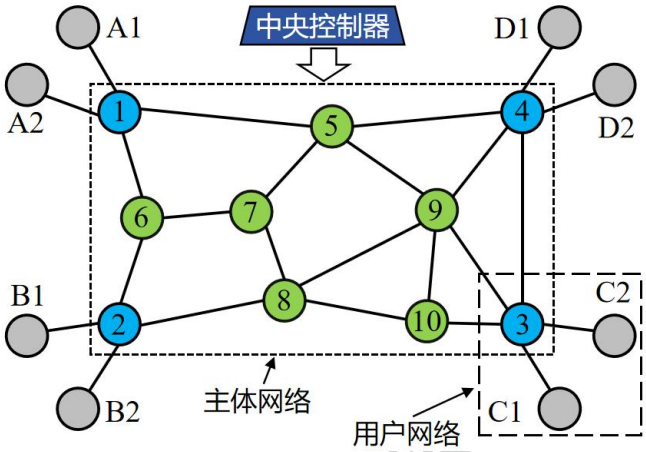

图 24. 初期少资源量子网络设计示意图。

（2）量子网络节点类型：用户、用户端相连的量子路由器、主体网络量子路由器和主体网络量子中继器。

（3）量子路由器部署：用户端相连的量子路由器使用第三代量子中继器技术且量子路由器具有一定的请求判别功能，主体网络可以兼容第一、二和三代量子中继器技术。

（4）量子网络调控模式：取消自治域模式，主体网络采用全网统一调控的集中式模式。由中央控制器向量子路由器和量子中继器下发规则集和转发表等。主体网络只负责目标用户端量子路由器之间的数据分发或纠缠分发。

（5）量子网络连接和资源分配模式：网络采用面向连接、固定路径和预留资源的方式运行，由中央控制器根据业务需求和网络资源

使用情况来计算定制化的连接和资源分配方案。

（6）帧结构设计：正式传输量子数据时，包头仅携带请求标识（ID）和路径 ID，不需要携带源地址、目的地址和端口号等。同一个请求的量子数据帧的大小相同。

（7）内存和网卡：用户端量子计算机内部不区分本地量子内存和量子网卡。

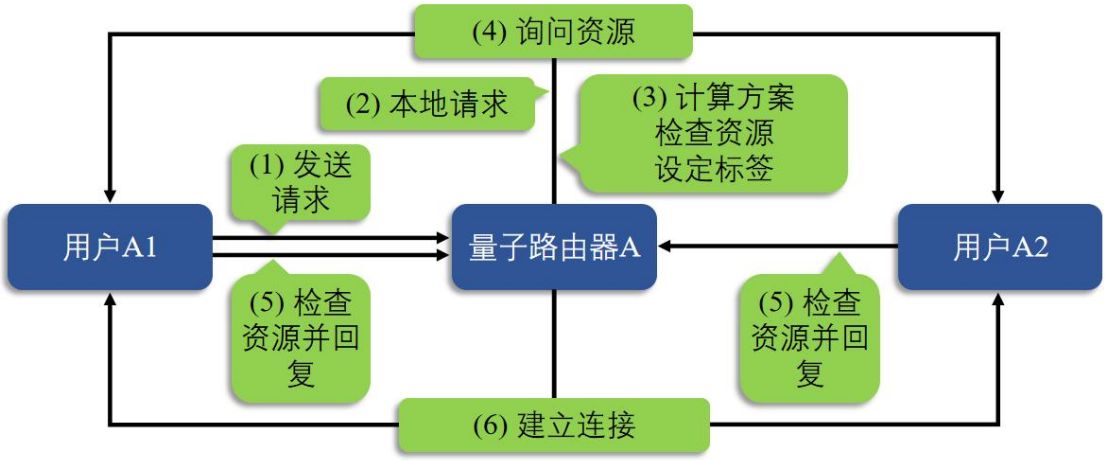

图 25. 本地请求建立连接的流程图。

4.3 量子请求运行方案

首先设定所有用户的请求都需要先通过与其连接最近的量子路由器来预先判别和处理，随后根据需要来决定是否将请求发送至中央处理器。如果是同一个量子路由器上的相邻用户节点之间的通信需求，则不需要请求中央处理器，直接由该量子路由器来完成。如果请求是不同量子路由器相连的用户之间的请求，则发送方量子路由器将请求发送至中央处理器来处理。

本地请求建立连接的流程如图 25 所示。

（1）用户 A1 发送请求到量子路由器 A。

（2）量子路由器 A 判断该请求为本地请求。

（3）量子路由器 A 为该请求计算方案、检查资源、设定标签。

（4）量子路由器 A 向用户 A1 和 A2 询问资源是否满足该请求。

（5）用户 A1 和 A2 检查资源并回复量子路由器 A。

（6）为该请求建立连接。

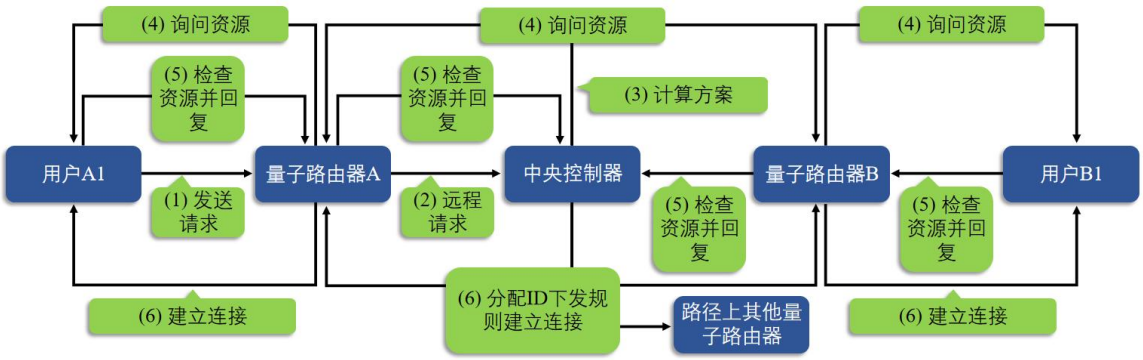

图 26. 远程请求建立连接的流程图。

远程请求建立连接的流程如图 26 所示。

（1）用户 A1 发送请求到量子路由器 A。

（2）量子路由器 A 判断该请求为远程请求，将其转发至中央控制器处理。

（3）中央控制器为该请求计算方案。

（4）中央控制器向量子路由器 A 和 B 询问资源是否满足完成该请求，A、B 向用户 A1、B1 询问资源是否满足完成该请求。

（5）用户 A1、B1 检查资源并回复量子路由器 A、B，量子路由器 A 和 B 检查资源并回复给中央控制器。

（6）中央控制器为该请求分配 ID，向选定路径上的量子路由器

下发规则，建立连接。

五、量子应用协议运行示例

基于纠缠的量子互联网执行量子信息任务时，首先需要建立端到端的纠缠信道。随后在该信道的基础上结合经典通信来完成一系列应用层任务，比如 QKD 和分布式量子计算等。所以端到端的纠缠分发是非常关键的一步。而实现端到端的纠缠分发可以有很多途径。如果类比经典互联网模式来执行量子互联网端到端纠缠分发，在网络层模式上可以选择面向连接和无连接，交换技术上可以选择电路交换和分组交换。电路交换只能适用于面向连接，而分组交换配合其他辅助控制系统，比如软件定义网络，可以适用于面向连接和无连接两种模式。和经典互联网相似，面向连接的电路交换方式，需要在网络中选择一条路径来预留资源。这种方式可以很好地保证通信的质量。但是由于需要预留资源，就会在包含大量用户业务的大规模网络中造成网络资源无法被充分利用。而无连接的分组交换就可以很好地解决这个问题，其不需要提前预留资源。如此一来，网络中的信道资源就不会被某一个或几个业务占用，资源得到充分利用。但是由于完全的无连接分组交换运行，也会面临一些自身的难题，比如可能会出现很多帧都被转发到某一条路径上，造成该路段的严重拥挤而导致很长的时延。此外帧在转发的过程中由于没有预留资源而造成丢失，也就是传统互联网中的丢包等，这些都会影响网络运行质量。为了改善以上的问题，可

以通过中央控制器的参与来宏观调控实际的分组交换运行。本章就是针对这个问题在量子互联网中设计一个中央控制器参与调控的分组交换实模式实现端到端纠缠分发，进而运行应用层的量子应用协议。主要思想是中央控制器为帧提前选定一条路径，但是不预留资源。在选定的路径上运行量子互联网分组交换。如此一来，遇到大量用户业务时，可以根据业务量来宏观调控帧的走向，一定程度上避免网络中某一条路径的严重交通拥堵。以下内容以 QKD 和分布式量子计算为例，展示完整的运行过程。

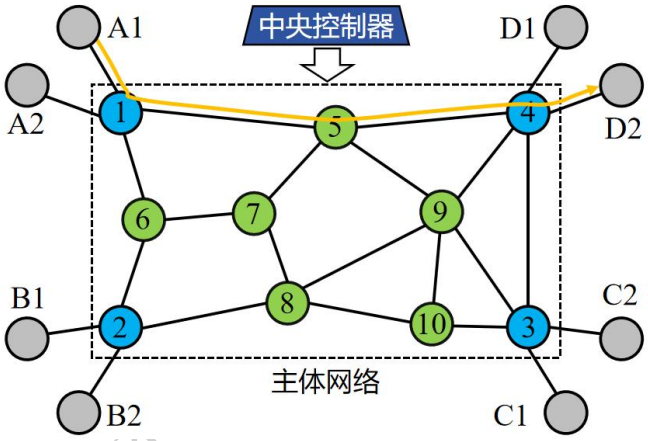

图 27 用户 A1 和 D2 之间执行 BBM92-QKD 应用协议。

5.1 量子密钥分发

本部分内容以量子互联网中执行 QKD 为例，结合混合分组交换技术以及上一章中的部分请求调度过程详细阐述整个流程。由于上一章中涉及的均为面向连接的预留资源的模式，我们这里仅采用其前期中央控制系统调度选取路径的部分，其他部分均在本章重新设计。

量子任务请求：如图 27 所示，用户 A1 需要和用户 D2 进行 QKD。

采用基于纠缠的 BBM92 协议。关于量子任务其他要求如下表格所示。

量子任务请求列表	
发送请求用户	A1
目标用户	D2
具体任务	QKD
协议类型	BBM92
网络类型	纠缠网络
中继类型	第二代量子中继
连接模式	面向连接但不预留资源
交换模式	混合分组交换
纠缠分发保真度	≥xxx
吞吐量	≥xxx
密钥率	≥xxx
…	…

表 3. 量子任务请求列表。

量子互联网对 QKD 任务执行过程如图 28（a）和（b）所示。左侧为过程图示，右侧为对应的文字讲解。

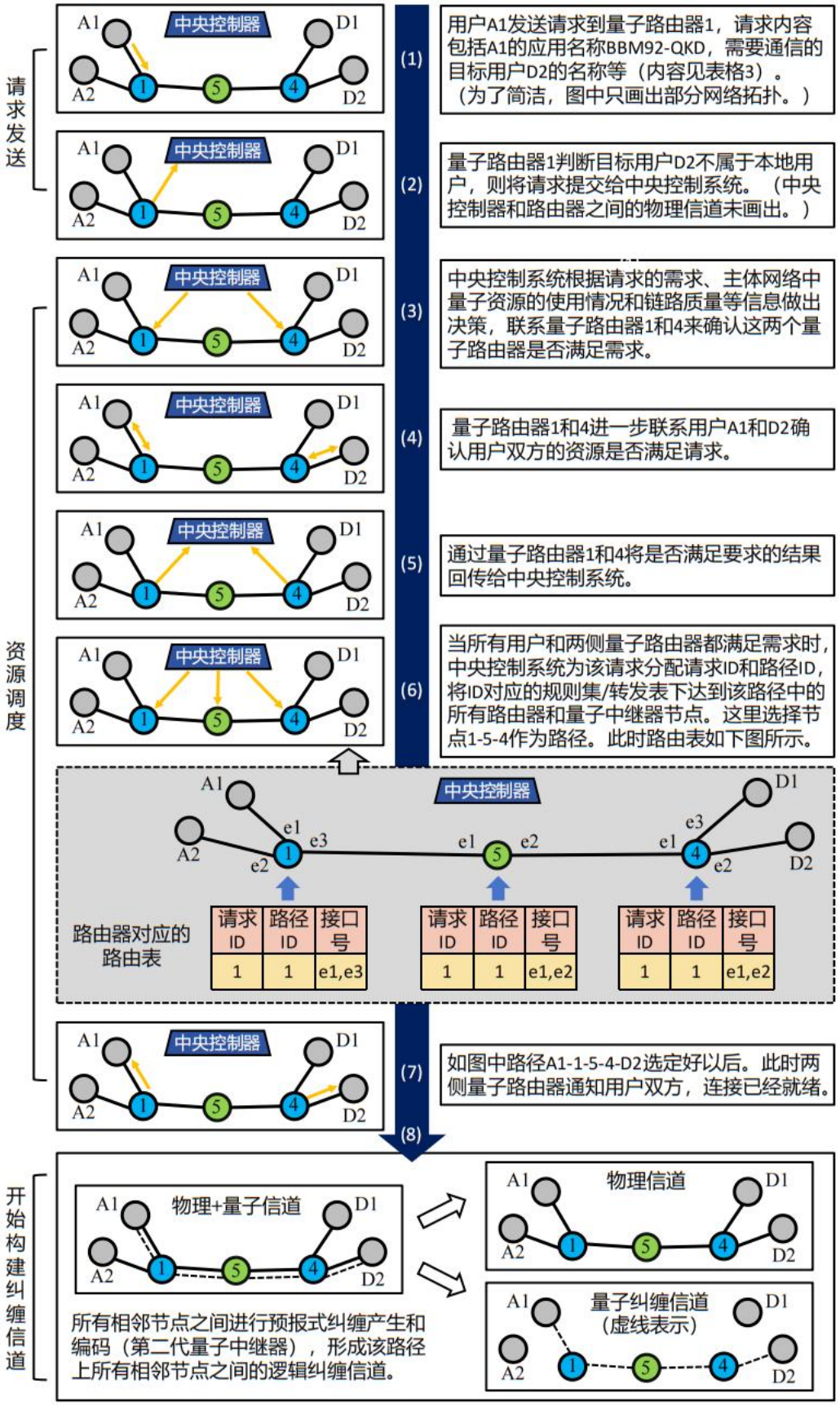

图 28（a）. 量子互联网用户 A1 和 D2 之间执行 BBM92-QKD 过程。

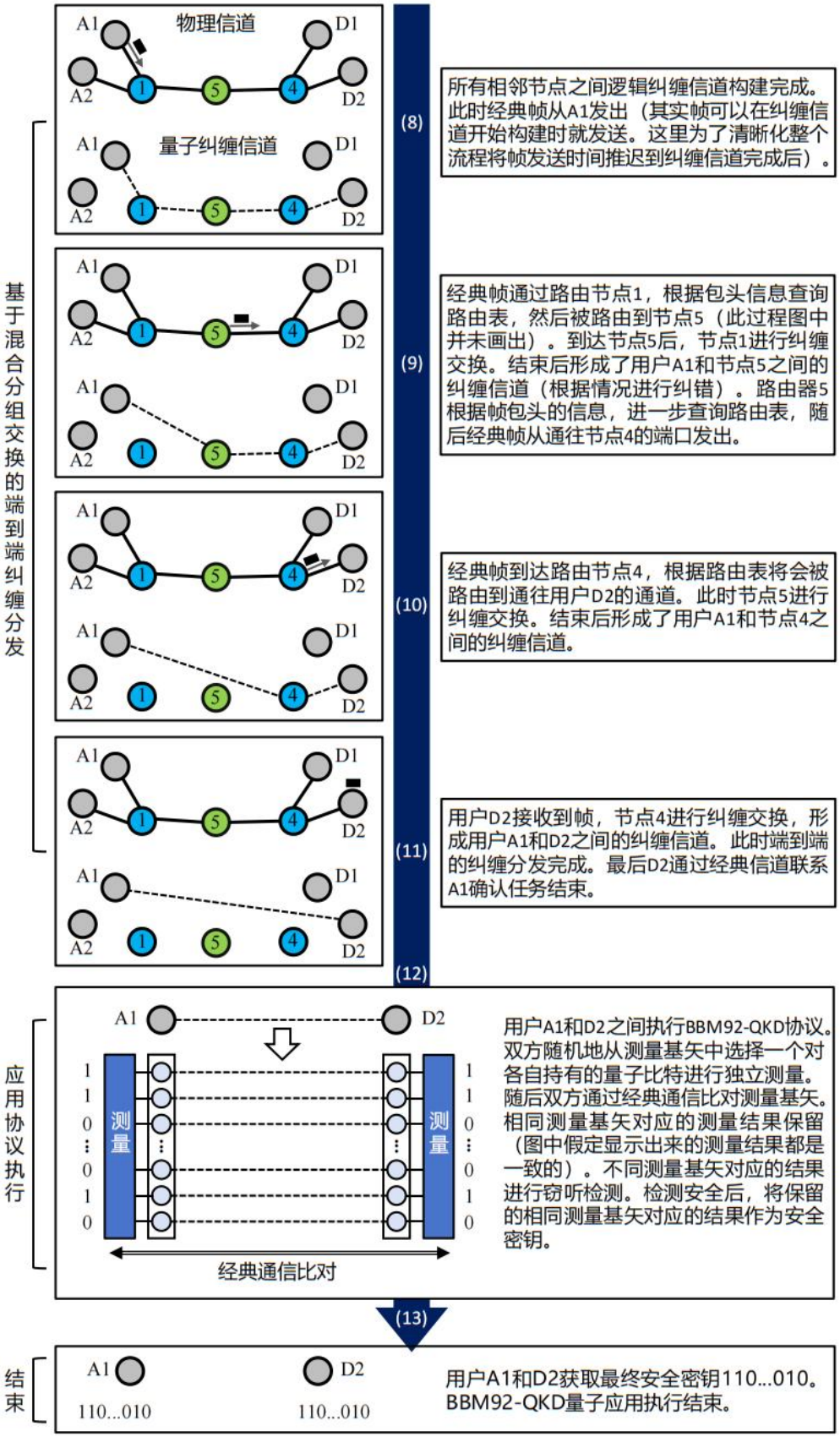

图 28（b）. 量子互联网用户 A1 和 D2 之间执行 BBM92-QKD 过程。

在整个过程中，图中(1)-(2)为请求发送，(3)-(7)为资源调度（主要是路径选择）。(8)-(11)为正式开始端到端纠缠分发。以上整个过程在中央控制系统调控下运行分组交换技术。这种模式可以在充分利用量子互联网资源的同时，也可以改善交通拥堵的情况。端到端的纠缠分发建立完成后，就可以执行 QKD 协议，即(12)和(13)。该过程需要多次使用经典互联网来传递信息，比对测量结果。

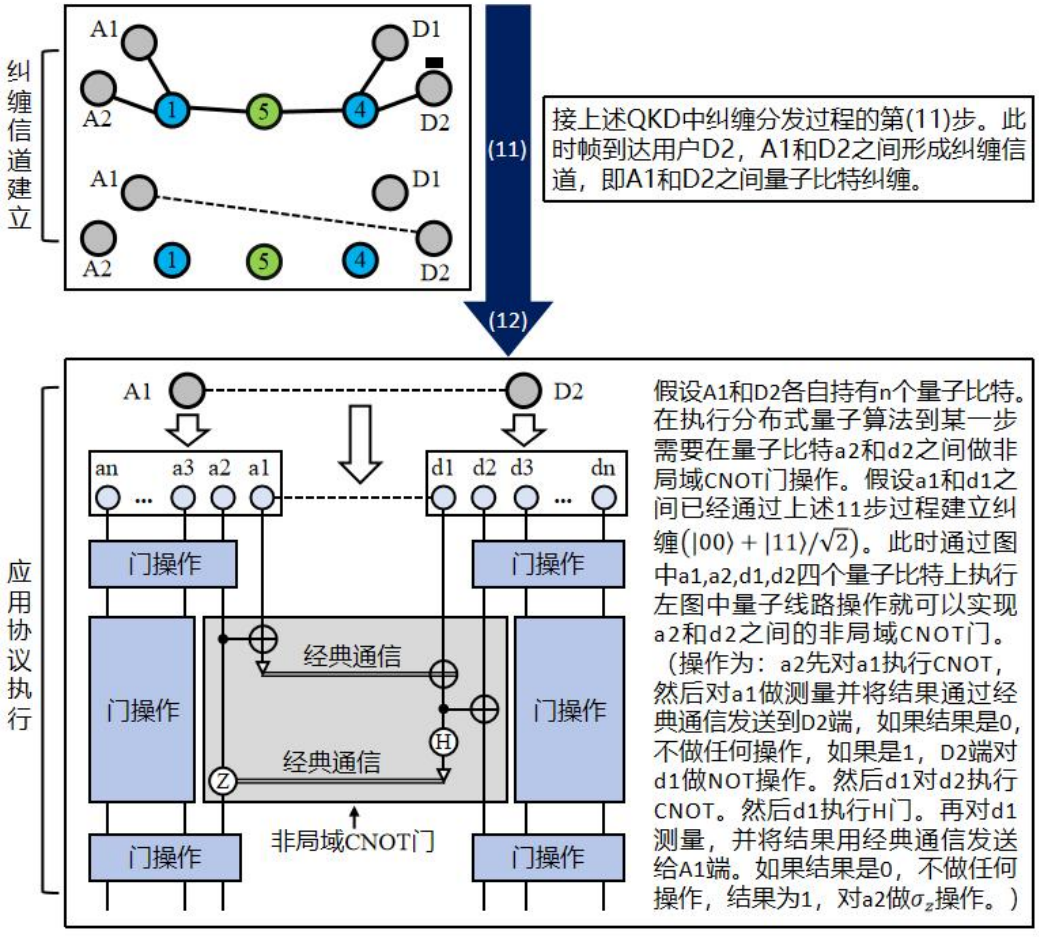

图 29. 量子互联网中用户 A1 和 D2 之间的分布式量子计算执行过程。

5.2 分布式量子计算

分布式量子计算是量子互联网的又一个重要量子应用任务。其利

用量子互联网将分布在不同端点的量子处理器连接起来共同计算一个大型计算任务。在分布式量子计算中，需要量子互联网来实现非局域的 CNOT 门，过程如图 29 所示。这里形成端到端的纠缠分发过程省略，直接参考上一节图 28 中(1)-(11)。

六、量子算网协同

6.1 量子计算协同化发展趋势

6.1.1 量子云计算

目前量子计算机的运行需要极其特殊的环境（如极低温、复杂的控制系统等）和高昂的运维成本，使得个人和企业用户难以在本地部署。量子云计算是一种新兴的计算架构，它允许用户通过云平台访问和使用量子计算资源而无需自己拥有和维护物理量子计算机[98]。

目前，绝大部分量子云计算平台仍依赖于经典架构进行管理和调度，这种架构通常被称为量子-经典混合云计算。在这种模式下，用户提交的量子线路任务首先由经典网络云服务接收，随后转发至量子处理器执行。计算完成后，结果会返回经典云，并最终传送给用户。

未来，量子云计算的一项关键创新是将其与量子网络进行深度整合，从而实现分布式量子云计算。

6.1.2 量子-超算融合计算

在高性能计算（HPC）系统的发展历程中，始终通过引入新的专

用加速器支持新的计算范式，例如，从早期处理器架构中的算术逻辑单元（ALU）和浮点运算单元（FPU），到后来的向量处理器和图形处理单元（GPU），专用加速器的演进不断推动高性能计算能力的提升。随着量子计算技术的发展，量子处理单元（QPU）被视为一种新型加速器，适用于那些在传统计算中资源需求随问题规模呈指数增长的任务，如整数分解、化学和物理中的电子结构计算，以及高能物理中的散射振幅计算等。

然而对于其他一些问题，目前还没有相关算法显示量子计算机具有计算优越性。例如对于预处理、后处理、输入输出（I/O）和可视化等辅助任务，目前经典计算资源仍然是更优选择。因此融合经典超级计算机和量子计算的“量子-超算融合计算”架构，有望带来大量能实现“量子优越性”的混合算法。其中最具代表性的例子之一是变分量子算法。这种算法类似深度学习，通过参数化的量子线路求解目标函数，再通过经典优化器迭代更新量子参数，逐步逼近最优结果。

要实现真正高效的量子-超算融合计算，还需要解决数据传输、资源管理和工作流管理等多方面的问题。

将 QPU 集成到 HPC 系统中主要有三种部署方式[99]：

远程访问：QPU 作为独立的运算单元，通过网络接口与 HPC 系统进行交互。这种方式部署灵活，但通信延迟可能会抵消部分量子计算的效率优势，而且需要确保传输数据的安全性和完整性。

本地集成：量子硬件位于经典计算基础设施的物理邻近位置，能有效降低延迟，提供了更高的性能和安全性，但部署和维护成本都相

对更高。

节点集成：将 QPU 集成到 HPC 节点内部，理论上可以实现最佳性能。但目前量子计算机的运行环境要求较为苛刻，如低温冷却系统、高频信号发生器和精确的环境控制等，节点集成面临显著的工程挑战。

6.1.3 分布式量子计算

当前，量子芯片所支持的量子比特数量仍较为有限，尚不足以支持规模较大的复杂量子算法。而对其规模的进一步扩展则受限于退相干、串扰、芯片拓扑以及控制电子学的复杂性。因此，突破单芯片能力限制，成为当前量子计算架构设计的重要方向。分布式量子计算被认为是克服这一瓶颈的可行路径之一。未来，分布式量子计算系统可以通过量子网络连接分布于不同空间的不同量子硬件，构建起异构、可扩展的量子计算系统。

根据在网络中的通信方式，分布式计算可以大致分为两类。

（1）量子节点之间仅进行经典通信，量子信息不在节点间直接传输。这一类又包括以下两个类型[100,101]。

线路分割（Circuit Cutting）：该方法将大规模的量子线路拆分成多个小规模的量子线路，分散在多个量子节点上执行，再通过经典后处理整合各个子线路的输出，得到最终的结果。但需要注意的是，该方法的复杂度会随着被切割的量子比特数量或门的数量呈指数增长，因此它仅适用于稀疏交互或可分解结构的量子线路。

尴尬并行（Embarrassingly Parallel）：这类任务指的是一些天然适合并行化的任务。它们可以在多个量子芯片上独立运行，仅需要最后

汇总结果。但前提是单个量子节点具备承载任务基本单元的能力。如果量子节点上的量子比特数量过少，则需要辅以线路切割或线路分布进行进一步的任务分解。

（2）量子节点之间不仅有经典通信，还有量子通信。这种模式也叫做线路分布，允许不同节点间直接传输量子信息，可实现多种量子机制。它将单个量子线路分成多个子线路，分别运行于不同的量子节点中，各节点通过远程量子门或量子隐形传态保持量子相关性。以量子隐形传态为例，传输一个量子比特需要同时传输 2 个经典比特的测量结果。原则上，该方法可以执行任意的量子算法，但也对量子网络提出了更高要求。

可以看到，量子计算的发展对计算和网络的共同协作提出了更高要求。

6.2 量子算网协同发展背景

在经典计算体系中，“算网协同”已被广泛认为是未来基础设施发展的重要趋势，用户无需关心计算资源的物理位置和网络位置，通过算力和网络的共同协作为业务提供端到端的服务质量保障。

在量子体系中，由于其独特性和复杂性，算力和网络的协作更为重要。具体来说，主要有以下原因：

6.2.1 量子应用对保真度的特殊要求

不同于经典应用，量子应用不仅要求资源充足，还必须满足量子态的保真度的要求。保真度用于衡量制备的量子态与理想量子态的接

近程度，反应了量子态的质量。保真度的取值范围是 0-1，1 表示理想状态，小于 0.5 则不再可用。部分量子应用还可能有更高要求。这意味着在量子网络中即使网络具备高吞吐量，如果不能保证量子态传输的保真度，也无法实现量子应用的正常运行。

6.2.2 量子算力特性对通信延迟的敏感性

量子计算中的计算量子比特需要在整个计算过程中保持量子状态，即整个计算时间需要小于量子比特的相干时间。这要求网络需要在严格的时间范围内内完成所需的量子通信和经典通信。

不同技术路线的相干时间差异较大：

超导体系：约百微秒；

中性原子体系：约秒级；

固态自旋体系：电子自旋约毫秒至秒级，核自旋秒至分钟级；

离子阱体系：超过 1 小时。

对于早期尚未完全成熟的量子网络，对网络的要求尤为严苛，算网协同需对此作出高效响应。但量子计算机进一步发展后，相干时间得以延长，这一要求会逐渐得到放宽。

6.2.3 计算量子比特与通信量子比特的资源分配权衡

量子算力资源中的量子比特可以分计算量子比特和通信量子比特。目前实验中两类量子比特通常采用不同的物理实现，如固态自旋体系的氮-空穴金刚石色心的量子计算方案采用电子自旋与核自旋分别承担通信与计算功能。

未来可能实现可以自由划分用途的通用量子比特，这面临算力资

源和网络通信资源的权衡问题。若将量子比特投入网络通信，可以并行执行的远程操作就越多，降低通信开销。但相对的计算量子比特会减少，降低算力。

6.2.4 早期量子网络资源受限

预计初期的量子网络中量子路由器上的量子比特数量、纠缠成功概率等都很低，能够提供的带宽小于 1000 qubits/s。该数值远小于目前经典网络的带宽。这意味着涉及到量子网络的业务更需要精细地调度网络资源，提高网络资源的利用效率。

6.2.5 涉及资源更复杂，协同需求更强

一个量子业务的完成可能需要多个量子计算节点、经典算力资源、量子网络以及经典网络。涉及到的资源将比经典体系的更复杂，因此更需要进行协同来完成。尤其对于第二类分布式量子计算（线路分布），多个量子节点间不仅在最后收集结果时需要交互，计算过程中同样会频繁进行阻塞式通信。

6.3 量子算网协同基础理论和研究方向

目前几乎没有量子算网协同的相关讨论，在这里，我们提出了一些未来的研究方向。类似经典体系中的算网操作系统[102]，量子算网系统的基础理论建立在对资源、业务以及调度的逻辑抽象模型上，量子算网协同也需要建立其自身的逻辑抽象模型。

6.3.1 资源抽象与建模

在量子算网协同中，既包括经典的算力资源和网络资源，也包括

量子算力资源和量子网络资源。其中经典资源部分与经典体系的算网协同中的抽象模型相同。下面我们重点讨论经典体系里没有包含的部分，即量子算力资源和量子网络资源。

（1）量子算力资源建模

对于量子算力资源，采用与经典算力资源统一的“资源量+供需关系+时空属性”三个维度的节点描述方法。其中“供需关系”和“时空属性”与经典算力资源的描述类似。而“资源量”则比经典的算力资源更为复杂。这主要源于目前多种体系的量子计算机并存，不同体系在多个维度上性能差异明显，且缺乏统一、权威的综合度量标准，使得资源量评估和对比变得极为复杂。

目前针对“资源量”的评估可以主要可以从三个视角展开[103]：

基础测控指标：包括量子比特数量、量子比特相干时间、量子比特连通性、量子门操作时间和保真度等，直接反应了量子计算机的底层能力。例如，在量子比特连通性较差的非全连接结构中，若需实现非相邻量子比特之间的量子门操作，必须插入大量 SWAP 门。这会增加量子线路深度，进而导致噪声和错误概率提高。

综合性能指标：包括量子体积、算法量子比特数、随机线路采样测试、镜像基准测试和每秒可靠的量子操作数等。这些指标衡量了量子计算机在运行量子线路时的整体性能。

应用性能指标：量子计算机在运行一些具体的量子算法（如量子傅里叶变换、Grover 搜索算法、相位和幅度估计算法、变分量子算法和量子近似优化算法等）时的实际表现。

如何准确建模量子算力资源，以便在量子业务到来时实现最适合的匹配，是未来的重要研究方向。

（2）量子网络资源建模

量子网络资源，同样可以采用“资源量+供需关系+时空属性”三个维度的对链路进行描述。其中“供需关系”和“时空属性”与经典网络资源的描述类似。对于“资源量”的描述，则相对复杂。

经典网络可以从网络资源提提供的“带宽、时延、抖动”服务能力来对网络资源进行度量。对于基于第三代量子中继器技术的量子网络，可以比较自然地沿用“带宽、时延、抖动”的概念。而对于基于第一和二代量子中继器技术的量子网络，量子比特并非由发送端逐跳传输到接收端，而是网络链路自身通过纠缠产生和交换等操作，建立端到端的纠缠态，来完成量子比特的传输。因此有关概念的定义需要重新调整。

类比带宽，目前一般使用“吞吐量”来衡量量子网络的性能，它定义为单位时间内形成的端到端纠缠态的数量。

由于纠缠态的建立具有概率性，且与经典网络中的数据包不同，量子网络中的端到端纠缠态是相同的量子态，因此难以直接复用时延和抖动的概念。通过记录每对端到端纠缠态的建立时间来进行性能评估，如首对纠缠态的建立时间和建立间隔时间等。

此外，不管采用哪种技术的量子网络，都应注意“保真度”这一重要指标。保真度和吞吐量可能存在权衡关系。纠缠纯化、量子纠错等技术可以提升保真度，但通常代价是吞吐量下降。

在量子网络发展早期，资源相对稀缺，如何准确地进行建模以及如何为量子业务分配量子网络中的量子资源是值得研究的课题。

6.3.2 量子业务建模

由于目前量子业务仍处于早期探索阶段，远未像经典体系那样形成成熟多样的服务形态，因此建模难度较高。它主要包括量子业务的自身特征、量子业务的运行条件和量子业务各个功能模块之间的交互拓扑关系。它们的核心都是业务对资源的要求。

具体来说，对于需要利用经典算力资源进行预处理和后处理的量子业务，需要对 CPU/GPU 和内存进行评估。对于量子云计算和量超融合计算中需要经典通信的部分，确定量子业务在经典网络中所需的带宽、时延和抖动等指标。对于量子计算部分，需要计算出对量子比特数量、量子比特连通性、量子比特相干时间和错误率等的要求。对于第二类分布式量子计算中需要量子通信的部分，确定量子业务在量子网络中所需的端到端纠缠态的数量、建立完成最小时间和保真度等指标。

准确抽象出量子业务对算力和网络的需求并建模，依赖于对量子算法的实现细节、量子计算机的硬件特性和量子网络机制有深入理解，是一项需要未来深入研究、持续推进的系统工程。

6.3.3 调度框架建模

调度建模的目的是将量子业务模型和量子资源模型相匹配，以最大化系统性能和效率。

对于不需要量子通信的量子业务来说，即上述介绍的量子云计算、

量超融合计算和第一类分布式量子计算，节点之间仅通过经典网络连接，调度涉及经典算力资源、量子算力资源和经典网络，与经典体系中最大的不同就是如何根据量子业务的要求来选择适合的量子算力资源，目前有多个不同架构体系的量子硬件平台，由于他们的性能和特性不同，会导致不同平台执行相同的应用需要不同的资源分配。

对于需要量子通信的第二类分布式量子计算就复杂的多。该调度框架需要包含经典算力资源、量子算力资源、经典网络资源和量子网络资源。除了需要根据量子业务的特点选择合适的量子算力资源外，还需要考虑一些额外的问题。一是需要寻找量子业务拆分到各个量子算力资源中的最优划分方案，使得对量子网络和量子算力资源的要求越小，且服务质量越高。二是选择的量子算力资源会影响对网络的要求。量子通信的时间需要小于量子算力资源的量子比特的相干时间，否则计算量子比特退相干将会导致计算失败。不同量子计算平台的相干时间各不相同，选择相干时间短的量子计算平台意味着对量子网络提出了更高的要求，这种情况更适配于短距离的量子通信。三是需要对量子算力资源中的计算量子比特和通信量子比特的分配进行调度决策，对算网的性能进行权衡。四是对于量子隐形传态、纠缠交换和纠缠纯化等操作，需要量子通信的同时，还伴有经典通信来传递测量结果，因此在选择网络资源时，需要协调经典网络和量子网络的时空属性，保证一致。不过虽然量子通信和经典通信的起点和终点需要一致，但他们可以各自选择各自不同的最优路径去实现通信。

七、总结与展望

在这个科技高速变革的信息时代，传统计算机互联网扮演着极其重要的角色。现代社会人与人之间的信息交换大多依赖于传统互联网平台。然而随着科技的发展，传统的信息安全和计算能力正在面临着以量子计算机为代表的新一代量子科技带来的巨大挑战，同时也是巨大的机遇。挑战来自于强大的量子计算机理论上可以攻破目前的公有密钥加密体系，严重威胁现代社会的信息安全。机遇是因为量子计算机的诞生可以大幅度提升算力，解决一些传统计算机无法有效计算的问题。幸运的是我们还可以发展量子通信来实现安全通信。而这一系列量子科技的交叉融合和大规模实用化就需要量子互联网。与传统互联网类似的是，量子互联网也是运行大量通信节点的网络平台。不同的是量子互联网平台运行的是量子通信、量子计算和量子传感等任务。目前量子互联网的发展还处于初期阶段。一方面这是由于量子互联网是许多量子信息应用的平台，通常是当上层的应用发展到一定的阶段才去研究它，导致这方面的研究起步相对较晚。另一方面是因为量子互联网基本物理原理和传统互联网差别很大，很多传统互联网的模式和技术无法直接复制到量子互联网，需要大量的新的探究。除此之外，底层量子技术还不成熟，限制了量子互联网的实验研究。所以目前量子互联网从底层的量子比特性能到上层的网络技术再到整体的运行架构都需要更多更深入的研究和发展。

技术上，对于量子互联网的发展，下一个需要解决的关键问题是

构建实用化的量子中继。目前基于预报式的量子中继模型最需要的是长相干时间的量子存储。量子存储不仅对长距离量子通信很关键，还能用于长时间存储量子数据。而基于量子纠错码的量子中继需要在纠错码技术上有所突破，这也是目前量子计算所面临的重要挑战。只有实验上真正实现量子纠错码技术，才能构建大规模的量子计算机。量子网络数据交换技术也很关键。成熟的量子中继结合数据交换技术才能进一步实现网络层的路由功能，形成真正的网络通信。在实际部署过程中，通过共用传统互联网的一部分基础设施，比如光纤和光开关等，来发展量子互联网是一条非常有潜力的途径，可以直接使用经典通信的同时节约很多资源。

业态方面，和传统互联网类似，当量子互联网技术成熟以后，结合上层量子应用协议，就会诞生出一些新的混合业态，比如量子算网协同。与单纯的某个量子应用协议不同的是，这些新型业态需要依赖量子互联网来实现。因此这些新型混合业态对量子互联网提出了更多的要求，需要对网络资源进行抽象和建模，开发更多的网络功能。

本白皮书首先全面梳理了与量子互联网相关的内容，包括量子信息基础知识、量子互联网的应用层协议（量子通信、量子计算和量子精密测量）、量子中继、量子互联网协议栈和量子分组交换。通过对基本架构和相关基本知识的介绍和梳理，让读者对量子互联网有一个清晰的认识。在此基础上，本白皮书主要介绍量子互联网的混合分组交换技术和初期少资源情况下的网络模式设计方案。进一步以应用层协议 BBM92-QKD 和分布式量子计算为例，详细阐述中央控制器调

度下面向连接但不预留资源的分组交换量子互联网运行过程，给出了建立端到端量子纠缠信道和执行应用层协议的详细步骤。然后重点介绍了一种新型混合业态，即量子算网协同。最后从发展背景、网络技术和业态上对量子互联网进行总结和展望。

附录 A：术语与缩略语

中文名称	英文缩写	英文全拼
贝尔态	/	Bell state
贝尔态测量	BSM	Bell state measurement
GHZ 态	GHZ state	Greenberger-Horne-Zeilinger state
受控非门	CNOT gate	Controlled-NOT gate
半正定算子值测量	POVM	Positive operator-valued measure
量子密钥分发	QKD	Quantum key distribution
诱骗态量子密钥分发	Decoy state-QKD	Decoy state quantum key distribution
测量设备无关量子密钥分发	MDI-QKD	Measurement-device-independent quantum key distribution
双场量子密钥分发	TF-QKD	Twin-field quantum key distribution
/	EPR	Einstein-Podolsky-Rosen
量子安全直接通信	QSDC	Quantum secure direct communication
/	CSS	Calderbank-Shor-Steane
/	RSA	Rivest-Shamir-Adleman
高级加密标准	AES	Advanced encryption standard
/	KLM	Knill-Laflamme-Milburn

中央处理器	CPU	Central processing unit
开放式系统互连	OSI	Open system interconnect
传输控制协议	TCP	Transmission control protocol
网际互连协议	IP	Internet protocol
标识	ID	Identification
高性能计算	HPC	High-performance computing
图形处理器	GPU	Graphics processing unit
量子处理单元	QPU	Quantum processing unit

参考文献

[1] 曾谨言，量子力学，科学出版社。

[2] 喀兴林，高等量子力学，高等教育出版社。

[3] 郭光灿和周详发，量子光学，科学出版社。

[4] 尹浩，韩阳等，量子通信原理与技术，电子工业出版社。

[5] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press.

[6] F. Xu, X. Ma, Q. Zhang, H. K. Lo, J. W. Pan, Secure quantum key distribution with realistic devices, Rev. Mod. Phys. 92, 025002 (2020).

[7] C. H. Bennett and G. Brassard, Quantum cryptography: public key distribution and coin tossing, in: Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing (IEEE, New York, 1984), p. 175-179.

[8] A. K. Ekert, Quantum cryptography based on Bell’s theorem, Phys. Rev. Lett. 67, 661 (1991).

[9] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum cryptography without Bell's theorem, Phys. Rev. Lett. 68, 557 (1992).

[10] H.-K. Lo, X. Ma, K. Chen, Decoy state quantum key distribution, Phys. Rev. Lett. 94, 230504 (2005).

[11] X.-B. Wang, Beating the photon-number-splitting attack in practical

quantum cryptography, Phys. Rev. Lett. 94, 230503 (2005).

[12] H.-K. Lo, M. Curty, and B. Qi, Measurement-device-independent quantum key distribution, Phys. Rev. Lett. 108. 130503 (2012).

[13] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, Overcoming the rate-distance limit of quantum key distribution without quantum repeaters, Nature (London) 557, 400-403 (2018).

[14] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Phys. Rev. Lett. 70, 1895 (1993).

[15] D. Pan, G. L. Long, L. Yin, Y. B. Sheng, D. Ruan, S. X. Ng, J. Lu, and L. Hanzo, The evolution of quantum secure direct communication: on the road to the Qinternet, IEEE Commun. Surv. Tutor. 26, 1819 (2024).

[16] G. L. Long and X. S. Liu, Theoretically efficient high capacity quantum key distribution scheme, Phys. Rev. A 65, 032302 (2002).

[17] F. G. Deng, G. L. Long, and X. S. Liu, Two-step quantum direct communication protocol using the Einstein-Podolsky-Rosen pair block, Phys. Rev. A 68, 042317 (2003).

[18] R. Feynman, Simulating physics with computers, International Journal of Theoretical Physics 21, 467 (1982).

[19] P. W. Shor, Algorithms for quantum computation: discrete logarithms

and factoring, in Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, p124, (1994).

[20] L. K. Grover, in Proceedings of the 28th Annual ACM Symposium on Theory if Computing, STOC'96 (ACM, New York, NY, USA, 1996), p212 (1996).

[21] J. F. Fitzsimons, Private quantum computation: an introduction to blind quantum computing and related protocols, npj Quantum Inf. 3, 23 (2017).

[22] Y.-C. Wei, P.-J. Stas, A. Suleymanzade, et al., Universal distributed blind quantum computing with solid-state qubits, Science 388, 509-513 (2025).

[23] X. Liu, X. M. Hu, T. X. Zhu, C. Zhang, Y. X. Xiao, J. L. Miao, Z. W. Ou, P. Y. Li, B. H. Liu, Z. Q. Zhou, C. F. Li, and G. C. Guo, Nonlocal photonic quantum gates over 7.0 km, Nat. Commun.15, 8529 (2024).

[24] D. Main, P. Drmota, D. P. Nadlinger, E.M. Ainley, A. Agrawal, B. C. Nichol, R. Srinivas, G. Araneda, and D. M. Lucas, Distributed quantum computing across an optical network link, Nature (London) 638, 383 (2025).

[25] J. Preskill, Quantum Computing in the NISQ era and beyond, Quantum 2, 79 (2018).

[26] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, et al., Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. 94, 015004

(2022).

[27] H.-L. Huang, X.-Y. Xu, C. Guo, G. Tian, S.-J. Wei, X. Sun, W.-S. Bao, and G.-L. Long, Near-term quantum computing techniques: Variational quantum algorithms, error mitigation, circuit compilation, benchmarking and classical simulation, Sci. China-Phys. Mech. Astron. 66, 250302 (2013).

[28] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. 89, 035002 (2017).

[29] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. 90, 035005 (2018).

[30] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multi-parameter estimation in networked quantum sensors, Phys. Rev. Lett. 120, 080501 (2018).

[31] D. H. Kim, S. Hong, Y. S. Kim, Y. Kim, S. W. Lee, R. C. Pooser, K. Oh, S. Y. Lee, C. Lee, and H. T. Lim, Distributed quantum sensing of multiple phases with fewer photons, Nat. Commun.15, 266 (2024).

[32] L. Z. Liu, Y. Z. Zhang, Z. D. Li, R. Zhang, X. F. Yin, Y. Y. Fei, L. Li, N. L. Liu, F. Xu, Y. A. Chen, and J. W. Pan, Distributed quantum phase estimation with entangled photons, Nat. Photonics 15, 137 (2021).

[33] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye,

and M. D. Lukin, A quantum network of clocks, Nat. Phys. 10, 582 (2014).

[34] D. Gottesman, T. Jennewein, and S. Croke, Longer-baseline telescopes using quantum repeaters, Phys. Rev. Lett. 109, 070503 (2012).

[35] D. P. DiVincenzo, The physical implementation of quantum computation, Fortschritte de Physik 48, 771 (2000).

[36] Z.-L. Xiang, S. Ashhab, J. Q. You, and F. Nori, Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems, Rev. Mod. Phys. 85, 623 (2013).

[37] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: a vision for the road ahead, Science 362, eaam9288 (2018).

[38] K. Fang, J. Zhao, X. Li, Y. Li, and R. Duan, Quantum NETwork: from theory to practice, Sci. China Inf. Sci. 66, 180509 (2023).

[39] Z. Li, K. Xue, J. Li, L. Chen, R. Li, Z. Wang, N. Yu, D. S. Wei, Q. Sun, and J. Lu, Entanglement-assisted quantum networks: Mechanics, enabling technologies, challenges, and research directions, IEEE Commun. Surv. Tutor. 25, 2133 (2023).

[40] C. Elliott, A. Colvin, D. Pearson, et al., Current status of the DARPA quantum network, In Quantum Information and Computation III Vol. 5815, 138–150 (International Society for Optics and Photonics, 2005).

[41] M. Peev, C. Pacher, R. Alléaume, et al., The SECOQC quantum key distribution network in Vienna. New J. Phys. 11, 075001 (2009).

[42] D. Stucki, M. Legré, F. Buntschu, et al., Long-term performance of the SwissQuantum quantum key distribution network in a field environment, New J. Phys. 13, 123001 (2011).

[43] M. Sasaki, M. Fujiwara, H. Ishizuka, et al., Field test of quantum key distribution in the Tokyo QKD Network, Opt. Express 19, 10387 (2011).

[44] T.-Y. Chen, H. Liang, Y. Liu, et al., Field test of a practical secure communication network with decoy-state quantum cryptography, Opt. Express 17, 6540 (2009).

[45] S. Wang, W. Chen, Z.-Q. Yin, et al., Field test of wavelength-saving quantum key distribution network, Opt. Lett. 35, 2454 (2010).

[46] T.-Y. Chen, J. Wang, H. Liang, et al., Metropolitan all-pass and inter-city quantum communication network, Opt. Express 18, 27217 (2010).

[47] T.-Y. Chen, X. Jiang, S.-B. Tang, et al., Implementation of a 46-node quantum metropolitan area network, npj Quantum Inf. 7, 134 (2021).

[48] S.-K. Liao, W.-Q. Cai, J. Handsteiner, et al., Satellite-relayed intercontinental quantum network, Phys. Rev. Lett. 120, 030501 (2018).

[49] Y.-A. Chen, Q. Zhang, T.-Y. Chen, et al. An integrated

space-to-ground quantum communication network over 4,600 kilometres, Nature 589, 214-219 (2021).

[50] J.-L. Liu, X.-Y. Luo, Y. Yu, et al., Creation of memory-memory entanglement in a metropolitan quantum network, Nature 629, 579-585 (2024).

[51] K. Azuma, S. E. Economou, D. Elkouss, P. Hilaire, L. Jiang, H. K. Lo, I. Tzitrin, Quantum repeaters: From quantum networks to the quantum internet, Rev. Mod. Phys. 95, 045006 (2023).

[52] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, Quantum repeaters based on atomic ensembles and linear optics, Rev. Mod. Phys. 83, 33 (2011).

[53] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, Optimal architectures for long distance quantum communication, Sci. Rep. 6, 20463 (2016).

[54] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, Phys. Rev. Lett. 76, 722 (1996).

[55] J. W. Pan, C. Simon, C. Brukner, and A. Zeilinger, Entanglement purification for quantum communication, Nature (London) 410, 1067 (2001).

[56] X. M. Hu, C. X. Huang, Y. B. Sheng, et al., Long-distance entanglement purification for quantum communication, Phys. Rev.

Lett. 126, 010503 (2021).

[57] Y. B. Sheng and F. G. Deng, Deterministic entanglement purification and complete nonlocal Bell-state analysis with hyperentanglement, Phys. Rev. A 81, 032307 (2010).

[58] H. Zhang, X. Xu, C. Zhang, M.-H. Yung, T. Huang, and Y. Liu, Variational quantum circuit learning of entanglement purification in multiple degrees of freedom, Phys. Rev. A 108, 042611 (2023).

[59] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A 52, R2493-R2496 (1995).

[60] D. Gottesman, Stabilizer codes and quantum error correction, arXiv: quant-ph/9705052 (1997).

[61] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. 87, 307 (2015).

[62] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Quantum repeaters: the role of imperfect local operations in quantum communication, Phys. Rev. Lett. 81, 5932 (1998).

[63] L. M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Long distance quantum communication with atomic ensembles and linear optics, Nature (London) 414, 413 (2001).

[64] B. Zhao, Z. B. Chen, Y. A. Chen, J. Schmiedmayer, and J. W Pan, Robust creation of entanglement between remote memory qubits, Phys. Rev. Lett. 98, 240502 (2007).

[65] T. J. Wang, S. Y. Song, and G. L. Long, Quantum repeater based on spatial entanglement of photons and quantum-dot spins in optical microcavities, Phys. Rev. A 85, 062311(2012).

[66] S. Perseguers, L. Jiang, N. Schuch, F. Verstraete, M. D. Lukin, J. I. Cirac, and K. G. H. Vollbrecht, One-shot entanglement generation over long distances in noisy quantum networks, Phys. Rev. A, 78, 062324(2008).

[67] L. Jiang, J. M. Taylor, K. Nemoto, W. J. Munro, R. Van Meter, and M. D. Lukin, Quantum repeater with encoding, Phys. Rev. A 79, 032325 (2009).

[68] W. J. Munro, K. A. Harrison, A. M. Stephens, S. J. Devitt and K. Nemoto, From quantum multiplexing to high-performance quantum networking, Nat. Photonics 4, 792 (2010).

[69] A. G. Fowler, D. S. Wang, C. D. Hill, T. D. Ladd, R. Van Meter, and L. C. L. Hollenberg, Surface code quantum communication, Phys. Rev. Lett. 104, 180503 (2010).

[70] W. J. Munro, A. M. Stephens, S. J. Devitt, K. A. Harrison, and K. Nemoto, Quantum communication without the necessity of quantum memories, Nat. Photonics 6, 777 (2012).

[71] S. Muralidharan, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, Ultrafast and fault-tolerant quantum communication across long distances, Phys. Rev. Lett. 112, 250501 (2014).

[72] K. Azuma, K. Tamaki and H. K. Lo, All-photonic quantum repeaters, Nat. Commun. 6, 6787 (2015).

[73] Z.-D. Li, R. Zhang, X.-F. Yin, et al., Experimental quantum repeater without quantum memory, Nat. Photonics 13, 644-648 (2019).

[74] G. L. Long, D. Pan, Y. Sheng, Q. Xue, J. Lu, and L. Hanzo, An evolutionary pathway for the quantum internet relying on secure classical repeaters, IEEE Netw. 36, 82-88 (2022).

[75] C. Delle Donne, M. Iuliano, B. van der Vecht, et al., An operating system for executing applications on quantum network nodes, Nature 639, 321–328 (2025).

[76] H. Shapourian, E. Kaur, T. Sewell, J. Zhao, M. Kilzer, R. Kompella, and R. Nejabati, Quantum Data Center Infrastructures: A Scalable Architectural Design Perspective, arXiv:2501.05598 (2025).

[77] Y. Li, H. Zhang, C. Zhang, T. Huang, and F. R. Yu, A survey of quantum internet protocols from a layered perspective, IEEE Commun. Surv. Tutor. 26, 1606-1634 (2024).

[78] J. Illiano, M. Calefff, A. Manzalini, and A. S. Cacciapuoti, Quantum internet protocol stack: A comprehensive survey, Comput. Netw. 213, 109092 (2022).

[79] R. Van Meter, T. D. Ladd, W. J. Munro, and K. Nemoto, System design for a long-line quantum repeater, IEEE/ACM Transactions on Networking 17, 1002 (2009).

[80] R. Van Meter and J. Touch, Designing quantum repeater networks, IEEE Commun. Mag. 51, 64 (2013).

[81] R. Van Meter, Quantum networking and Internetworking, IEEE Netw. 26, 59 (2012).

[82] R. Van Meter, J. Touch, and C. Horsman, Recursive quantum repeater networks, Progress Informatics 8, 65 (2011).

[83] W. Kozlowski and S. Wehner, Towards large-scale quantum networks, in Proceedings of the Sixth Annual ACM International Conference on Nanoscale Computing and Communication (NANOCOM' 19) (Association for Computing Machinery, New York, NY, 2019).

[84] A. Dahlberg, M. Skrzypczyk, T. Coopmans, et al., A link layer protocol for quantum networks. In Proc. ACM Special Interest Group on Data Communication, SIGCOMM'19, 159-173 (ACM, New York, NY, USA, 2019).

[85] W. Kozlowski, A. Dahlberg, and S. Wehner, Designing a quantum network protocol, In proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies (CoNEXT'20), 16 (ACM, 2020).

[86] M. Pompili, C. Delle Donne, I. te Raa, et al., Experimental demonstration of entanglement delivery using a quantum network stack, npj Quantum Inf. 8, 121 (2022).

[87] A. Pirker and W. Dür, A quantum network stack and protocols for reliable entanglement-based networks, New J. Phys. 21, 033003 (2019).

[88] Z. Li, K. Xue, J. Li, N. Yu, J. Liu, D. S. Wei, Q. Sun, and J. Lu, Building a large-scale and wide-area quantum internet based on an OSI-alike model, China Communications 18, 10 (2021).

[89] L. Bacciottini, M. G. De Andrade, S. Pouryousef, E. A. Van Milligen, A. Chandra, N. K. Panigrahy, N. S. V. Rao, G. Vardoyan, and D. Towsley, Leveraging Internet principle to build a quantum network, arXiv:2410.08980 (2025).

[90] L. G. Roberts, The evolution of packet switching, Proc. IEEE 66, 1307 (1978).

[91] S. J. B. Yoo and P. Kumar, Quantum wrapper networking, IEEE Photonics Conference, IPC 2021 - Proceedings, (2021).

[92] S. DiAdamo, B. Qi, G. Miller, R. Kompella, and A. Shabani, Packet switching in quantum networks: A path to the quantum Internet, Phys. Rev. Research 4, 043064 (2022).

[93] S. J. B. Yoo, S. K. Singh, M. B. On, G. Dul, G. S. Kanter, R. Proietti and P. Kumar, Quantum wrapper networking, IEEE Commun. Mag. 62, 76-81 (2024).

[94] M. B. On, R. Proietti, G. Gul, G. S. Kanter, S. K. Singh, P. Kumar, and S. J. B. Yoo, Experimental demonstration of datagram switching

with monitoring in quantum wrapper networks, J. Light. Technol. 42, 3504 (2024).

[95] M. B. On, R. Proietti, G. Gul, G. S. Kanter, S. K. Singh, P. Kumar, and S. J. B. Yoo, Entanglement distribution in packet-switched quantum wrapper network, 2024 Conference on Lasers and Electro-Optics (CLEO), Charlotte, NC, USA, (2024).

[96] H. Zhang, Y. Li, C. Zhang, and T. Huang, Hybrid packet switching assisted by classical frame for entanglement-based quantum networks, arXiv:2310.02770 (2023).

[97] Y. Li, C. Zhang, H. Zhang, T. Huang, and Y. Liu, A design framework for early quantum networks, arXiv:2508.04967 (2025).

[98] H. T. Nguyen, P. Krishnan, D. Krishnaswamy, et al., Quantum cloud computing: A review, open problems, and future direction, arXiv:2404.11420 (2024).

[99] M. Ruefenacht, B. G. Taketani, P. Lähteenmäki, et al., Bringing quantum acceleration to supercomputers, IQM/LRZ Technical Report (2022).

[100] D. Barral, F. J. Cardama, G. Diaz-Camacho, et al., Review of distributed quantum computing: from single QPU to high performance quantum computing, Comp. Sci. Rev. 57, 100747 (2025).

[101] 王升斌, 窦猛汉，吴玉椿，郭国平和郭光灿，分布式量子计算

研究进展，量子电子学报 41, 1-25 (2024)。

[102] 张晨，黄韬，周俊等，《算网操作系统白皮书》，第七届未来网络发展大会，2023 年 8 月。

[103] 量子科技产学研创新联盟，量子计算性能评估基准报告，2024 年 12 月。