



HAL
open science

A Survey of Quantum Entanglement Routing Protocols-Challenges for Wide-Area Networks

Fabrice Dupuy, Claire Goursaud, Fabrice Guillemin

► **To cite this version:**

Fabrice Dupuy, Claire Goursaud, Fabrice Guillemin. A Survey of Quantum Entanglement Routing Protocols-Challenges for Wide-Area Networks. *Advanced Quantum Technologies*, 2023, 6 (5), 10.1002/qute.202200180 . hal-04244891

HAL Id: hal-04244891

<https://hal.science/hal-04244891>

Submitted on 18 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Survey of Quantum Entanglement Routing Protocols - Challenges for Wide-Area Networks

October 18, 2023

Fabrice Dupuy Claire Goursaud Fabrice Guillemin
Fabrice Dupuy*

Orange Innovation, NET/CASE, 46 Av. de la République, 92320 Châtillon,
France

Email Address: fabrice.dupuy@orange.com

Claire Goursaud

Université de Lyon, INSA Lyon, INRIA, *CITI EA 3720* Lyon, France

Fabrice Guillemin

Orange Innovation, NET/CASE, 2, Avenue Pierre Marzin, 22300 Lannion, France

Entanglement distribution, Entanglement routing, Swapping, Entanglement fidelity, Multi-commodities flow optimization

Abstract

In this paper, we analyse the key challenges (loss due to distance, entanglement routing, multi-commodities) for the coming Quantum Internet, relying on entanglement of quantum bits (for short, qubits) on top of an existing network. We present a unifying framework enabling to compare the various entanglement distribution, purification and routing protocols published so far. With regard to entanglement routing, the introduction of different time windows will be essential in order to cope efficiently with the main challenges like complex route calculation and fidelity estimation on the one hand, actual entanglement route selection and entangled photon generation on the other hand. For a roll-out on top of existing transmission networks, all the research publications for the last twenty years start to cover pretty well the global scheme. Nevertheless open questions remain, like the actual advantage of some task execution prior to the online quantum path selection, or the design of algorithms approximating the multi-commodities flow optimisation problem, or the issue of dealing with a processing time not much longer than the qubit life time.

1 Introduction

With the rapid development of quantum cryptography and quantum computing, it will become essential in the near future to transmit quantum information in the form of quantum bits (for short, qubits) over long distance links, given rise to the Quantum Internet [1]. In the past few decades, huge research effort has been devoted to the development of a quantum Internet as an interconnection of Wide Area Networks (WANs), each WAN enabling remote quantum devices to transmit quantum information. The universal resource for such quantum communication is an entangled pair of qubits (referred to as an ebit). In combination with classical information transmission (i.e., transmission of classical bits), the entangled pairs allow to implement a wide range of quantum protocols. The vision developed in [1] is that the quantum Internet will enable new major applications like secure communications, clock synchronization, raw information capture from telescopes or other quantum sensors, secure identification, and distributed quantum computing. The challenge is to move from the small scale of the current experiments to larger scale in terms of distance and source and destination pairs of qubits.

There are three major obstacles in the development of quantum communications over long distance (with no repeaters). The first one is due to the evanescence of qubits. The life time of a qubit is short and the buffering capacities are extremely limited. Today, the buffering time of a qubit does not exceed a few tens of milliseconds. This imposes to rapidly transmit qubits and very small buffers can be used in the network. As for the second obstacle, qubits (most of the time, photons) are subject to alteration when propagating along a fiber link. The quantum information carried by a qubit is altered by the interaction between the photons and the transmission medium. The third one is due to the non cloning principle of qubits. Contrary to classical bits, qubits cannot be copied. This prevents from retransmitting information in case of loss. Due to these obstacles, the transmission reliability decreases as the length increases.

To overcome the above obstacles for the transmission of qubits, the entanglement property of quantum states is an efficient tool in the development of the quantum Internet. The basic principle of entanglement is to couple to the quantum states of two or more particles so that they form a unique entity. By measuring the information carried by one of the particles, it is possible to obtain information on the other ones and this is valid whatever the distance between the particles, as long as they do not decohere. This is known as the Einstein-Podolsky-Rosen (EPR) paradox in quantum mechanics. Beyond loss of qubits, the coherence of two entangled particles is limited in time (a few tens of milliseconds), which in turn puts limits on the range of teleportation.

In spite of these limitations, it has been largely proven that quantum entanglement used to perform quantum teleportation is a promising technology to increase the transmission distance of qubits, with still a satisfying fidelity (i.e., purity of the quantum state so as it is possible to measure it with high probability). This approach relies on the generation, the distribution and the processing of quantum entangled pairs in addition to the transmission of classical data used

to recover information via quantum teleportation.

For a source to transmit a qubit to a destination (and eventually to other ones), one can rely on a series of contiguous segments connected two-by-two by elements referred to as repeaters. The protocol consists of distributing entangled qubits on each segment, and operating local Bell State Measurements (BSMs) at each repeater so as to set up end-to-end entanglement used to perform teleportation between the source and the destination. Entanglement can actually come in many forms and can involve more than two qubits; several papers (e.g., [2]) show how to deal with k -qubit in Grenberger-Horne-Zeilinger (GHZ) state and consequently with more complex quantum measurement modules, Bell States correspond to $k = 2$. In the rest of this paper, we shall focus on multiple communications involving two parties only (source and destination pairs) and therefore on Bell states and EPR pairs (namely, two entangled photons).

Experiments involving one source and destination pair have been successful over a few kilometers. Extending them to a WAN on one hand and to multiple source and destination pairs on the other hand is a challenging task which requires to solve some issues and to make design decisions:

1. Calculate the optimal number of repeaters to install along a path with adequate positions in order to perform quantum state swapping.
2. Estimate the quality of the quantum channels transmission.
3. Decide where to install sources capable of emitting entangled particles (for short, ebits) on demand and transmitting them to all nodes (repeaters or not), or deploy first generation quantum repeaters [3] supporting single-photon generation capabilities.
4. Estimate whether Quantum Error Correction (QEC) or entanglement purification protocols will be required, depending on the level of loss and noise in the quantum channels.
5. Choose among deterministic and probabilistic measurement schemes, when swapping/fusing multiple terminating qubit pairs on the repeaters along the paths (thus aiming at obtaining at least one end-to-end entanglement pair between each source and destination).
6. Elaborate heuristic methods of enabling such a global scheme to work fairly and efficiently with a large set of source and destination pairs.

The large number of publications in the field of quantum communications provides a very good basis for the survey proposed in this paper. Most, if not all, papers in the entanglement routing field introduce phases (active, external/internal, static/dynamic). In our understanding, a unified vision for these phases is missing in the technical literature. Our goal is to provide such a global framework listing the different tasks that could be involved in a quantum entanglement routing protocol, depending on the expected requirements, in terms of fidelity, latency or rate. The different types of classical/quantum operations

together with the classical communications can be categorized into four phases (and four time windows), enabling to check whether the challenges brought by scaling up to WANs are sufficiently covered. Such a framework does not provide restrictions but a range of tasks to select and to schedule when designing the protocol.

This paper is organized as follows: Section 2 highlights the challenges arising when shifting from the early point-to-point quantum communications to a long-distance Quantum Internet. Section 3 will review a chosen set of entanglement (configuration, purification, routing) protocols, and map them on the same framework. In Sec 4, we will list the open questions that could remain for such a WAN-scale entanglement network to come true.

2 Main challenges when scaling up to a WAN

Extending early point-to-point quantum communication experiments to a WAN is challenging because of the non cloning principle, the qubit evanescence and the transmission reliability. In this section, we will first introduce i) a quantum network model, then go through the issues to cope with, meaning: ii) the loss and the fidelity decrease of quantum information carried by qubits when crossing links over longer distance; iii) the different routing paths calculation within a short delay (to stay below the decoherence time τ of a quantum system), and iv) the concurrence on the entanglement resources between various source/destination pairs as entangled pairs are the resource for teleporting qubits.

2.1 Modelling a quantum network

The notation that will be used in the following is presented in Table 1.

2.1.1 Topology

A quantum network is modeled by means of an undirected graph $G = (V, E)$. A node $v \in V$ represents either a communicating party (the source or the destination of the qubits to communicate) or a repeater performing operations on qubits such as BSM. Vertices are labelled as v_0, v_1, \dots, v_n . An edge $\epsilon \in E$ is a quantum link connecting two nodes. The edge $\epsilon = (m, n)$ connects the nodes v_m and v_n . An edge ϵ is either an optical fiber or free space satellite-link and has a range L_ϵ (expressed in kilometers) and a power transmissivity η_ϵ .

In this paper, we consider an entanglement network, where entangled pairs of qubits are distributed on quantum links within time slots; the duration of a time slot will be discussed in the following as it has an impact on the quality of entangled pairs. The elementary resource is hence a pair of entangled qubits used to transfer data qubits from a source to a destination. A qubit of an entangled pair is distributed to an extremity of a link of graph G . A path through the network is a chain of adjacent links.

Table 1: Definitions and notation.

Notation	Definition	Equivalent in []
Node v_n	Any node of the network (Source, Destination, Repeater)	
Edge ϵ or $(v_m v_n)$	A physical link connecting the two nodes	A channel c in [4]
S_ϵ or Edge ϵ Width	Number of transmittable entangled pairs or number of modes per use of edge	$S(e)$ in [5]
L_ϵ	Edge Length	
η_ϵ	Transmissivity of edge ϵ	$\eta(e)$ in [5]
1 ebit or 1 EPR pair	One pair of entangled qubits	
Quantum Link	The successful distribution of an entangled pair (or ebit) over the edge, thus linking the two nodes	A Quantum Link too, in [4]
$p_d(\epsilon)$	The probability for the successful distribution of one entangled pair over edge	$p_0(e) = \eta(e)$ in [5]; p_c in [4]
$\omega_l^{(n)}$	The l^{th} path ω (a chain of Quantum Links or a virtual circuit) connecting the n^{th} pair of Source and Destination	a (W;h)-path, or a W-path in [4]
τ_g	Time between two generated entangled pairs; $1/\tau_g =$ entanglement generation rate or repetition rate (in Hz)	
τ_d	Distribution time and time unit	t in [5]
$T^* \tau_d$	protocol processing cycle or Latency	T in [5]; Δt in [6]
EXT_ϵ	the entanglement rate equals the amount of ebits generated per edge use	
EXT	the amount of end-to-end ebits successfully distributed over one path	$EXTE_t$ in [5]; R_g in [5]
THR	the throughput equals the amount of teleported qubits per second	
Fidelity F	the degree of similarity between a final quantum state and its original state	F_1 in [7]
parameter W	the parameter corresponds to the Werner State	The same notation everywhere
$d(v_m, v_n, F, \ell, \delta)$	the Connection Demand for a quantum link between source v_m and destination v_n , with a fidelity F , a latency ℓ and a throughput δ	$D_i = (s_i, d_i)$ in [8]
$ D ^{max}$	the maximum number of Connection Demands authorized in one processing time $T^* \tau_d$	
CDR	The Connected Demand Ratio = the number of demands that are actually connected via at least one end-to-end routing path over the total number of demands	the <i>entangled routing rate</i> in [8]

Each quantum link ϵ is characterized by a width S_ϵ , which represents the number of entangled qubit pairs or EPR pairs or ebits that can be distributed over the quantum link within a given time slot. The delay for any information to be transmitted over a link of length L is $\tau_d = L/c$, c being the light speed in the corresponding medium (e.g., a fiber).

2.1.2 Timing precision and synchronization

As mentioned above, time is slotted; entanglement distribution and routing protocols are based on accurate timestamps; each node clock should be synchronized. The time slot unit is denoted by τ_d . As in most articles [9, 10], we choose the time unit τ_d to be the distribution time over an edge whose length L_0 corresponds to a 3dB corresponding to a power loss of $1/2$ so that $\tau_d = L_0/c$.

The time slot unit τ_d should not be to be confused with the entanglement generation rate $1/\tau_g$. An entangled pair is generated over a quantum link every τ_g seconds. We will assume that all operations take either a negligible amount of time (e.g., quantum gate computation) or an integer multiple of τ_d (e.g., entanglement generation over a single hop, entanglement distillation or swapping, as described later). For practical reasons, the processing window time of the entanglement routing protocols will be a fixed time interval with length $T * \tau_d$ (denoted by Δt in [6]); it will also correspond to the latency routing protocol, i.e., the time needed to generate, distribute and swap **end-to-end entangled pairs**, thus enabling to teleport the **application-level qubits**.

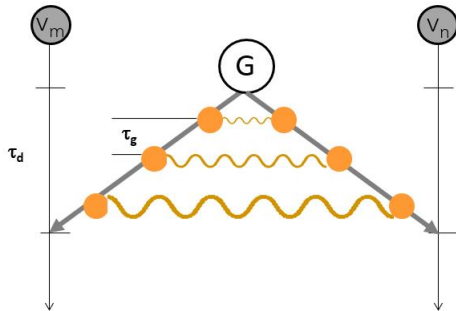


Figure 1: Three entangled pairs generated by (G) at rate $1/\tau_g$, whose transmission lasts τ_d seconds.

2.1.3 Connection Demands

The quantum network is requested to transport application-level qubits from sources to destinations. We thus define a demand $d(v_m, v_n, F, \ell, \delta)$ between a source and a destination as a five tuple composed of the source v_m , the destination v_n , a fidelity F , a latency ℓ and a throughput δ . Classic network protocols cannot be transposed here, due to the specificity of the quantum physical layer

(e.g. the non cloning principle) and the possible interdependence between the first 3 classic network layers (physical, data link and network).

To cope with qubit transmission demands, we need to deal with: 1) the distance between Source and Destination in a WAN is too high to authorize direct transmissions, intermediate quantum repeaters will be required; 2) routing entangled pairs through a quantum repeater/router requires mechanisms different from the ones in classical routing ; 3) the elementary quantum resource (here an entangled pair) cannot be shared by multiple end users (commodities), and is extremely fragile. The three challenges are described more precisely in the following three subsections.

2.2 The Distance Challenge

The propagation of a qubit (namely, a photon) across a channel alters its quantum state. As in any telecommunication network, transmission links are **lossy** and **noisy**. In the case of qubits, these causes of degradation are more critical because of the non replication and the limited storage capacity of qubits.

2.2.1 Loss

On edge ϵ , loss is characterized by the edge transmissivity η_ϵ , which is proportional to $e^{-\alpha L_\epsilon}$ in all cases (optical fibers or free space). In [11], it is shown that the maximum rate at which two remote parties can distribute quantum bits (qubits), entanglement bits (ebits), or secret bits over a lossy channel (e.g., an optical fiber) is equal to $C(\eta_\epsilon) = -\log_2(1 - \eta_\epsilon)$ per mode (the Pirandola-Laurenza-Ottaviani-Banchi or PLOB bound [11]). When $\eta_\epsilon \ll 1$, the maximum capacity of a single-mode edge ϵ (i.e. PLOB bound) can be approximated as $C(\eta_\epsilon) = 1,44\eta_\epsilon$.

When taking into account different multiplexing modes (Wavelength Multiplexing, Time Multiplexing, and Spatial Multiplexing), the edge width S can be seen as the product $S_{SM}.S_{TM}.S_{WM}$ of the widths, WM , TM and SM standing for wavelength, time, and space multiplexing respectively. For example, if the source generating the entanglement pairs over an edge has a generation rate of $1/\tau_g$ ($S_{TM} = 1/\tau_g$) and is based on S_{SM} spatially multiplexed photons, the edge has an entangled qubit capacity $C_{S_\epsilon}(\eta_\epsilon) = (S_{SM} * 1/\tau_g).C_1(\eta_\epsilon)$ (see [12], where $S_{SM} = M$ and $\tau_g = \tau$). Thus, if the edge ϵ supports S_ϵ modes, its capacity becomes $C_{S_\epsilon}(\eta_\epsilon) = S_\epsilon.C_1(\eta_\epsilon)$.

Photons and their quantum state are altered when crossing a fiber link. The probability $p_d(\epsilon)$ for a photon to reach the end node is also the wave loss ratio η_ϵ , that is, $p_d(\epsilon) = \eta_\epsilon = 10^{-\alpha.L_\epsilon/10}$ for some $\alpha > 0$; this quantity is sometimes multiplied by a factor η_c , the coupling efficiency within the optical element, and by η_d the detector efficiency. Following [4], the probability of having **exactly** i successful qubit transmission on the edge ϵ given that there S qubits are transmitted is

$$Q_\epsilon^i = \binom{S}{i} p_d(\epsilon)^i (1 - p_d(\epsilon))^{S-i}.$$

The probability to have **at least** 1 successful transmission for the width S_ϵ of edge ϵ is $1 - (1 - p_d(\epsilon))^{S_\epsilon}$.

Ebits are distributed to the extremities of the edge; a successful transmission of an ebit gives rise to a quantum link. The concatenation of quantum links between a source and a destination gives rise to a quantum path from the source to the destination. This quantum path is used to transmit an application-level qubit from the source to the destination. With these definitions, we can define two basic quantities:

- The throughput (THR), equal to the amount of **application-level qubits** actually teleported **per second**; it could be compared with the amount of secret key bits generated per second in the Quantum Key Distribution Protocols (also named *throughput* in [13]).
- The entanglement rate (EXT) of an edge or a path equals to the amount of **ebits** (entangled qubits) generated **per cycle**, over the physical edge or the virtual path; therefore the entanglement rate EXT_ϵ will be the mean number of entangled qubits successfully transmitted to the two extremities of the elementary edge ϵ , that is,

$$\text{EXT}_\epsilon = \sum_{i=1}^{S_\epsilon} i \cdot Q_\epsilon^i = S_\epsilon p_d(\epsilon). \quad (1)$$

Furthermore,

$$\text{THR} \leq \text{EXT}/(T \cdot \tau_d) \quad (2)$$

2.2.2 Noise

Beyond loss, edges are also **noisy** in the sense that there is a probability for the quantum system state to become less trusted and not exploitable for computation, transmission or storage. A noisy edge can transform the transmitted quantum system into a state a bit different from its original one. In an attempt to quantify the degree of similarity between the mixed density ρ_t in which the quantum system is after transmission, and its initial density $\rho_0 = |\psi\rangle\langle\psi|$, Schumacher [14] has defined the fidelity $F = \langle\psi|\rho_t|\psi\rangle$. In our case, the quantum system is an entangled pair (e.g., $|\psi\rangle$ can be the Bell state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$; then $\rho_0^+ = |\phi^+\rangle\langle\phi^+|$). The fidelity of the transmitted quantum system ρ_t^+ to its initial pure state ρ_0^+ is

$$F = \langle\phi^+|\rho_t^+|\phi^+\rangle.$$

To better illustrate the computation of fidelity, Reinhard Werner [15] proposed the following model, introducing the Werner state:

$$\rho_t^+ = W |\phi^+\rangle\langle\phi^+| + \frac{1}{4}(1 - W)\mathbb{I}_4,$$

where W is a weight coefficient, $\mathbb{I}_4 = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|$ is the identity matrix on the space spanned by the Bell states and $F = (3W + 1)/4$.

The density operator is a weighted sum on the Bell state with weight W , the probability of being in state $|\phi^+\rangle$ is F and equal to $\frac{1}{4}(1 - W)$ for the other Bell states. When W is very small, we see that we lose information about the initial state $|\phi^+\rangle$ and all Bell states have the same probability equal to $\frac{1}{4}$.

To overcome the noisy effects (due to the propagation across a transmission channel, the Bell-State Measurement by a repeater or even the quality of a quantum memory storage), purification methods are used and increase fidelity. First purification procedures were devised by Bennett [16], then improved by Deutsch [17], starting with 2 pairs distilled into one. The calculation formula for the resulting fidelity F_p after purifying two pairs of lower fidelity F can be found for example in [18] and is given by

$$F_p = \frac{F^2 + \frac{1}{9}(1 - F)^2}{F^2 + \frac{2}{3}F(1 - F) + \frac{5}{9}(1 - F)^2}.$$

In early stages, purification protocols, based on the quantum gate CNOT or similar logic gates, were very difficult to implement. More recent realizations of optimized purification circuits are less imperfect and can be used now for long-distance quantum communications. Several significant entanglement purification experiments using photons, atoms, and electron-spin qubits have been reported.

The value of the entangled pair fidelity can also decrease because in time. In [6, 19], models let the Werner weight W depends on time as $W(t) = W_0 e^{-\Delta t/T_{coh}}$, where T_{coh} is the expected coherence time of a qubit in the memory and W_0 is some initial weight. A long-time quantum memory capable of storing and measuring quantum information at the single-qubit level becomes essential. Nevertheless, since there have been remarkable progresses of increasing coherence time, it does not seem necessary in our comparative framework to integrate decoherence time and its dephasing effect.

2.2.3 Consequences of loss and noise

In view of the previous sections, we see that to transmit qubits over a **lossy and noisy long-distance edge** between a source and a destination (e.g., from the source node v_m to the destination node v_n in Figure 2 with $L > L_0$), it is necessary to install repeaters (represented as white circles in Figure 2). In the modelled network with quantum repeaters, all terrestrial edges are of length less than L_0 (L_0 corresponds to the length of an elementary edge for which the loss is 3dB); the length of free space physical links or edges will be a multiple of L_0 .

When successfully transmitted between two adjacent nodes, the entangled pair forms an elementary quantum link. After possible entanglement distillation (i.e., purification), these repeaters enable to perform BSMs, so that h elementary quantum links are fused into one end-to-end entangled pair between the source and the destination nodes. The latter corresponds also to a quantum path of h hops and length L (less than or equal to hL_0).

Like for classical networks, many quantum paths can connect a source v_m and a destination v_n for a given demand $d(v_m, v_n, F, \ell, \delta)$. Let us denote by

$\omega_l^{(d)}$ (or $\omega_l^{(d)}(F, \ell, \delta)$) the l th path of the d^{th} demand (or $d(v_m, v_n, F, \ell, \delta)$). For example in Figure 2, chaining three elementary links still between Source v_m and Destination v_n makes the path $\omega_1^{(v_m v_n)}$, and chaining six links makes the path $\omega_2^{(v_m v_n)}$.

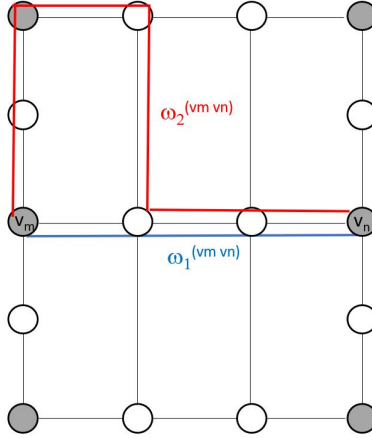


Figure 2: Two path $\omega_l^{(v_m v_n)}$ examples (one in blue, the second in red) to connect Source v_m and Destination v_n

The maximum capacity of the l th path $\omega_l^{(n)}$ composed of h edges ϵ_k , $k \in \{1, \dots, h\}$, is the minimum over k of all $S_{\epsilon_k} \cdot C_1(\eta_{\epsilon_k})$. The end-to-end entanglement rate over this path $\omega_l^{(n)}$ will be denoted by $EXT_l^{(n)}$ and calculated for the different routing protocols. As for the fidelity of the end-to-end entangled pair generated for $\omega_l^{(n)}$ (after swapping and fusing the h elementary pairs), it decreases also due to the quantum local operations performed by the repeaters along the path. If the fidelity decrease due the edge transmission only is around the same value F or W all along the path ($F < 1$; $F = (3W + 1)/4$), the additional fidelity decrease due to the $(h - 1)$ swapping operations on the intermediary nodes is $(1 + 3W^h)/4$ [20]. Therefore, depending on the requested end-to-end fidelity, some entanglement purification may be required along the path; we will see in section 3 how to integrate such operations in the protocols.

2.3 The Entanglement Routing Challenge

When extending point-to-point quantum communications to Quantum Internet/Quantum Wide-Area Networks, the routing path between each source-destination pair is not unique; a path calculation or selection is required. Routing paths in any conventional network can be constructed based on the topology of the network. Here, it is not the case, since the routing paths must be paths of entangled pairs; repeaters/routers have to deal with the conventional static network resources (i.e., nodes and links) and the dynamically-generated quan-

tum ones. In order to route or string correctly the entangled pairs, and thus generate a virtual path for the demanding Source-Destination (S-D) pair, the repeaters must perform BSM or decoding/encoding on correctly "selected" entangled qubits. To do so, they require:

1. routing information or a Forwarding Information Base (FIB), prior to the *on-line* local operations,
2. the set of connection demands like $demand(source : s_i; destination : d_i; fidelity : F; \dots)$ they have to satisfy during the given time window,
3. *dynamic traffic engineering* [8].

For some protocols, a preliminary (i.e., offline) calculation of the routes for each Source-Destination pair will suffice; others, and particularly those which can guarantee a fidelity threshold, will require an adaptive (i.e., online) calculation. Moreover, the corresponding routing algorithm is computed on a central server receiving from the repeaters the global information required for the routing path calculation, or an approximation one is executed on each repeater, with the available information.

Table 2: Single Demand routing algorithm categories: centrally/locally computed; adaptative/dynamic phase

Single Demand Routing Algorithms	offline computation / static phase	online computation / adaptative phase
Centrally computed	Algorithms calculating routing paths in the original network on a central server (results sent to each repeater node)	Algorithms calculating routing paths in the reduced network on a central server (results sent to each repeater node)
Locally computed	Algorithms calculating routing paths in the original network, on local repeater nodes (with less information)	Algorithms calculating routing paths in the reduced network, on local repeater nodes (with less information)

Most entanglement routing protocol studies start with the assumption that global information is available at each repeater node. However it requires quantum memories whose coherence time is compatible with the time necessary to locally compute the routes. Such protocols may prove unrealistic. Thus one comes to either

1. entanglement routing algorithms with less local information available at repeaters (i.e., heuristic approximation methods); the actual entanglement rates $EXT_l^{(n)}$ is inferior to the expected and ideal ones,
2. and/or some pre-calculation of the paths prior to the processing window time (i.e., "adaptative phase").

We will specify in the section 3 what the studied protocol is requiring, in order to route correctly.

2.4 The Multi-Demand or Multi-Commodity Challenge

In the Internet, a commodity is usually a demand consisting of a Source, a Destination, and potentially other requirements like the desired end-to-end transmission rate or the quality of service. Commodity = Demand = Entangled Connection Request.

Dealing with multiple demands $d_i = \text{demand}(\text{source: } s_i ; \text{ destination: } d_i ; \text{ fidelity: } F; \dots)$ is compared in [20] with the multicommodity flow optimization problem: the objective is to maximize in a network the total amount of flows (end-to-end entangled pairs connecting Sources and Destinations), given a set of connection demands (or commodities). Such optimization problems are NP-hard; it means that the routing algorithm, centrally or locally computed, offline or online (as in table 2) cannot result in polynomial time in at least one path proposal for each Demand. Therefore one has to design approximation algorithms which are based on relaxing some constraints and can therefore result in routing paths calculated in polynomial time. An interesting metric here is the number of demands that are actually connected in a given processing time window via at least one end-to-end routing path, over the total number of demands. Instead of "entanglement routing rate"[8], this metric will be denoted here "connected demand ratio" or CDR.

Table 3: Multi-Demand routing algorithms can be NP-hard

Multi-Demand Routing Algorithms	offline computation / static phase	online computation / adaptative phase
Centrally computed	(Approximate) algorithms calculating routing paths in the original network on a central server (results sent to each repeater node)	Approximate algorithms calculating routing paths in the reduced network on a central server (results sent to each repeater node)
Locally computed	Approximate algorithms calculating routing paths in the original network, on local repeater nodes (with less information)	Approximate algorithms calculating routing paths in the reduced network, on local repeater nodes (with less information)

In [20], the authors show i) that Linear Programming (LP) formulations can be used for relaxing such a problem and ii) that associating flows to a fraction enables a resolution of the Linear Programming in polynomial time (in the number of nodes).

The authors in [5] consider the case of non-intersecting commodity flows (between the different Source-Destination pairs), with the simple strategy for every single repeater node (including the nodes labeled as sources and destinations) to use the protocols described for one source-destination pair. The authors of [5] recommend a multi-flow spatial-division rule for routing multi-flow entanglement: the network is divided into spatial regions, each being 'allocated' to one

of the multiple source-destination pairs.

There are other tricks to consider like the one consisting in authorizing $|D|^{max}$ sets of demands during the processing time window. When starting a new time window, a first subgraph $G_{init}^{(n)}$ corresponding to the shortest paths (weighed by loss and pre-calculated fidelity) is calculated for each n^{th} demand for $n = 1, \dots, |D|^{max}$, the global information remaining in a classical information processing (central) server. Then after the entanglement distribution phase, every repeater sends an heralded signal to the central unit, which updates the first subgraphs $G_{init}^{(n)}$ with successfully entangled links (thus getting $G^{(n)}$), and sends back to the concerned repeaters the local routing information they will need in order to swap and eventually purify. The cost of this alternative is the two-way Classical Communications it requires, and therefore the additional time (a multiple of τ_d) it adds to the protocol running time.

One may note that some information or signal processing are to be done in (almost) real time, while some support pre-processing. We thus propose to distribute the tasks in a unified framework, depending on their timing constraints, and to review the related protocols.

3 A survey of entanglement routing protocols

In order to scale up to a WAN and to meet the main challenges (long distance, routing problem complexity, multiple demands), Section 2 highlighted that both classical and quantum information have to be processed, either to optimize the network configuration, to estimate the edge fidelity, to calculate the optimal paths for each Source-Destination pair, to improve the entanglement quality (fidelity), or to 'route' each ebit through the repeaters (swapping operation). We have also exhibited the fact that different time scales and types of algorithms are required. We propose here a global framework with a fixed number (namely, 4) of phases and associated time windows, in order to compare different entanglement routing protocols and provide a range of tasks to select and schedule when designing an entanglement routing protocol.

3.1 The global protocol design framework

Similarly to [4] and as illustrated in Figure 3, we introduce four phases to characterize the different tasks involved in a quantum entanglement routing protocol, and their associated time windows. The mean duration of each phase depends on the type of its operations: phase 4 Bell State Measurements are executed in seconds where the calculation of the repeater optimal positions can take a few minutes.

Phase 1 the network topology **validity phase**, requiring algorithms like the one in [21] aiming at optimally positioning the repeater nodes in an existing optical fiber infrastructure, or like the fidelity estimation protocol;

Relying on an existing Wide Area Network with its own topology (original constraint of the problem), the challenge is to place a minimal number of repeaters, in such a way that requirements on end-to-end entanglement-generation rate and fidelity are satisfied. This placement can be calculated using integer linear programming (ILP). While ILP problems are generally difficult to solve, the authors in [21] show that such problems can be solved well in practice for networks of up to 100 nodes.

As for the fidelity pre-estimation, [22] shows that the results of a set of experiments run during link initialization can be used to track expected behavior of the system and a mean fidelity value for each edge can be determined.

The Phase 1 time window is in months or years.

Phase 2 the information processing **static phase** (also named phase 1 in [4]), during which all possible paths for the demanding Sources and Destinations **can be** calculated, as well as the number of entanglement purification operations required to get the expected fidelity threshold; the information is sent to all repeater nodes; the corresponding time window (in minutes) relates to the processing time during which the $|D|^{max}$ Sources request entangled connections to Destinations, in order to teleport qubits.

Phase 3 the **entangled pair generation and distribution phase**, during which a large number of entangled pairs (ebits) is generated and distributed. This is the external phase in [5], the dynamic phase in [23]; of concern here are simple or multiplexing schemes [24] aimed at distributing one or more entangled pairs per edge; the associated time window, a multiple of τ_d , needs to be inferior to τ_{dch} (the qubit decoherence time in quantum memory). Since all nodes of one routing path need to receive the halves of the pairs entangled with its predecessor and successor at the same time, time synchronization is absolutely necessary. For simple routing protocols, this can be achieved by the current synchronization protocols using classical communications; for more complex ones, one will need nanosecond, even picosecond synchronization.

Phase 4 The actual **routing phase** (the internal phase in [5], adaptative phase in [25], or phase 4 in [4]) involves quantum computing (local operations) on each repeater node. The quantum algorithm aims at selecting the most optimal route, as viewed by the local classical information that repeaters are aware of. At the end of the fourth phase, most if not all sources can finalize the teleportation protocol by applying the teleportation-protocol on their qubit and the local half of the shared entangled pair, and then sending the two classical bits (i.e. the outcome of the measure) to the destination.

In the next subsections, we map some published entanglement purification and routing protocols with the above phases 2, 3 and 4. The denomination in

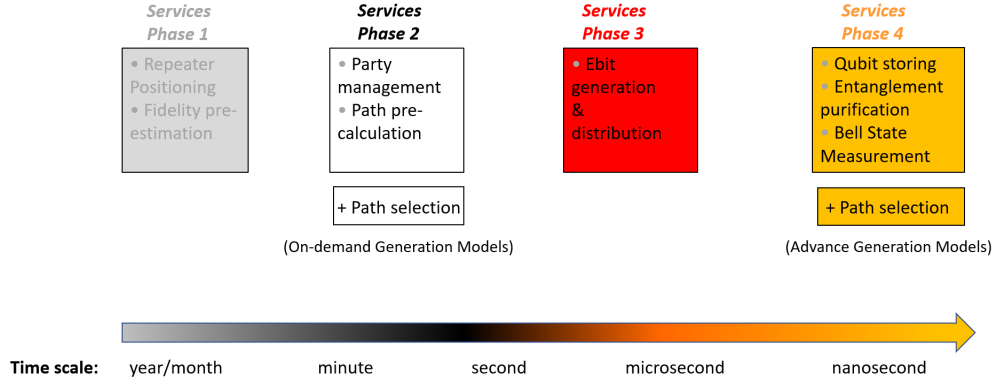


Figure 3: 4 main phases - 4 time windows

bold (e.g. **Step1**) will refer to the one used in the cited publication, the capital letter P in brackets (e.g. **P3**) will refer to the phase of our framework. We will also use the distinction introduced in [18] between:

1. the “advance generation” model in which the entanglement generation and distribution is carried out before the actual path selection; only resource allocation is considered in the routing design;
2. the “on-demand generation” model [4] [26] in which the entangled pairs are distributed on demand, along the selected path(s); potential failures of entanglement distribution are also considered in the routing design.

Thus, as depicted in Figure 3, all tasks that could be required during the design of an entanglement routing protocol can be surely associated to one of the four phases, except *path or route selection*, which can occur prior to the entanglement distribution (i.e. advance generation model) or after (i.e. on-demand generation model).

3.2 An ‘advance generation’ model based on Single-Link Single Path Routing

One first protocol, proposed by Briegel et al. [7] in 1998 and known as a first-generation quantum repeater, consists of three elements: entanglement distribution, entanglement swapping and nested purification protocols. Most authors [7] [27] refer to this scheme as the ‘purify and swap’ approach. Here, the number of pairs initially distributed will depend on both the number of nodes along the channel and the fidelity of the required final pair. The different steps of the protocol can be linked to the phases introduced in the previous section as follows:

Step1 (P3) The protocol begins by distributing many different copies of entangled pairs between the nodes.

Step 2 (P4) The initial distribution is followed by one round of purification, taking two entangled pairs to a single entangled pair of higher fidelity.

Step 3 (P4) A Bell-State Measurement is performed on the 1st, 3rd, 5th, etc. repeater nodes in order to generate ebits over 2¹ hops; the results of these measurements are sent via a classical communication channel. A second round of Bell-State Measurement is performed on the 2nd, 6th, 10th, etc. repeater nodes in order to generate ebits over 2² hops, the results being sent via a classical communication. And so forth, until all repeaters have swapped their local entangled qubits.

3.3 An 'advance generation' model based on Single-Link Multi-Path Routing

The Single-Link Multi-Path (SLMP) routing protocol [5] is based on parallel channels over each edge. Each cycle time T is divided into two phases: the "external" phase and the "internal" phase, which occur in the following order:

External phase (P4) each of the S pairs of memories across an edge attempts to establish a shared entangled (EPR) link.

Internal phase (P4) entanglement swaps (Bell State Measurements) are attempted locally at each repeater node between pairs of qubit memories. The authors of [5] consider these BSM attempts as links internal to the repeater nodes.

In [2], the authors evaluate the benefits of time-multiplexing in a two-dimensional quantum network. The paper extends the multi-path entanglement routing protocol presented in [5] by allowing quantum memories to hold onto entangled qubits for multiple time-steps before deciding on which qubit pairs to perform BSMs.

3.4 An 'on-demand generation' model, called Q-PASS / Q-CAST protocols

In the protocols proposed in [4], each time slot includes four phases as an extended model from [5] (or section above), the difference between Q-PASS and Q-CAST being that the former pre-computes potential 'good' paths between all possible sender-receiver pairs based on the network topology G (Q-PASS \subseteq P2). Then in each time slot (including P3 and P4), every node uses an online algorithm to make qubit-to-channel assignments based on the pre-computed paths of current S-D pairs and make local swapping decisions based on local link states. Q-PASS design includes both offline and online algorithms.

Q-PASS Phase 1 (P2) all nodes receive the information of current source-destination (A B) pairs that need to establish long-distance entanglements by the Classical Communications CC; *"this offline algorithm is run at a trusted server, with replica servers for robustness"*.

Q-PASS Phase 2 (P2 and P3) each node assigns its quantum memory qubits to quantum channels and attempts to generate quantum entanglements with neighboring nodes on the bound channels (P3)

Q-PASS Phase 3 (P4) each node knows its own link states via Classical Communications with its neighbors and shares its link states via the classical network. Since successful entanglements will quickly decay, each node can only receive the link states of a subset of other nodes. Phase 3 only includes classical information exchange.

Q-PASS Phase 4 (P4) also called the internal phase in [5], nodes perform quantum swapping to establish long-distance quantum entanglement using the successful quantum links.

3.5 An ‘advance generation’ model (Q-PATH), guaranteeing Fidelity

The authors of [18] propose Q-PATH, probably the first work that studies the fidelity-guaranteed entanglement routing problem for multiple sender-receiver pairs. Considering the high computational complexity of the Q-PATH algorithm, they further propose Q-LEAP, a Low-complExity routing Algorithm from the perspective of “multiPlicative” routing metric of the fidelity degradation. Based on these two routing design principles for a single source-destination pair, a greedy approach-based entanglement routing algorithm is then proposed for multiple source-destination pairs. After entanglement resource generation and distribution (P3), the Q-PATH/Q-LEAP algorithm consists of the following steps:

Initialization (P2): a “Purification CostTable” for each edge is computed. Since the fidelity constraint $F \geq F_{th}$ (for some fidelity lower bound F_{th}) must be satisfied for the routing path, if an edge cannot provide entangled pairs satisfying $F \geq F_{th}$ even after purification, this edge is deleted from graph G .

Path Selection Procedure (P2): To obtain the optimal routing path with minimum cost, the authors design an iterative method that utilizes the cost as the metric to continue each loop, which provides an indicator during the path searching. Here, k -shortest path algorithm is used to obtain multiple shortest paths with the same cost minimal cost.

Edge Cost Update (P4): Since the most important challenge is fidelity constraint, the routing design must check if the fidelity of entanglement connection is satisfied for each path.

Throughput Update (P4): The algorithm calculates the maximum achievable throughput purification decision.

3.6 An ‘advance generation’ model, using Forward Quantum Error-Correction

This protocol encodes logical qubits with small Calderbank-Shor-Steane (CSS) codes [28]; it then applies entanglement connection with the encoded qubits, and uses Quantum Error Correction to boost the fidelity of entanglement connection.

For example, each channel termination has 3 qubits in repeater memory and 3 auxiliary qubits. The algorithm precisely consists of the following steps:

Encoded generation step 3 memory qubits are prepared in the CSS encoded states (phase P2). Then, 3 physical Bell pairs are distributed (corresponding to phase P3) over the channel and used to implement 3 teleportation-based CNOT gates, thus entangling the 3 memory qubits of the local repeater with 3 others at the termination of the channel.

Encoded connection step (P4) encoded Bell measurements are simultaneously applied to all intermediate repeater stations, via pairwise CNOT gates between qubits in memory A and qubits in memory A'. The outcome is sent via 2 classical bits at each repeater forward.

Pauli Frame step (P4) according to the outcomes of intermediate encoded Bell measurements, the Pauli frame can be determined for qubits at the outermost station (namely destinations). Finally, one encoded end-to-end pair is created between source and destination nodes.

3.7 An ‘advance generation’ model based on the actual network state

The authors of [23] propose the following scheme: the entanglement resources are pre-distributed on all edges, the successful distribution constituting what the authors called the network state. This network state is distributed prior to requests. In this scheme, phases P2 and P3 are inverted. Therefore, no distribution times are needed upon incoming requests. The time that the multiple users have to wait for their end-to-end entangled pair simply amounts to the classical communication time. Each cycle, with the same duration, has three phases:

Dynamic phase (P3) the entanglement structure in the devices and across the network - called the network state - is generated, where the networking devices utilize the quantum channels to distribute the required entanglement. One way of doing this is to use the Quantum Network Configuration Protocol (QNCP).

Static phase (P3) a certain network state is stored on all nodes, enabling client requests to establish certain target (graph) states in the network and enabling during next phase the nodes to process this network state using only local operations and classical communication (LOCC).

Adaptive phase (partly P2 and fully P4) the clients request a graph state (i.e., in our early hypothesis, an end-to-end entangled pair) from the network, and the quantum network repeaters manipulate the states given during the static phase in such a way that the target state is established between requesting clients.

3.8 A last ‘advance generation’ model, using QNetworkRouting

One last example in the distribution-prior-to-routing category is QNetworkRouting [29]. After an initialization step (P1) involved also in distributing entangled pairs (P3), tasks of steps 2 to 5 (P4) consist of both physical steps (in which quantum operations are physically performed) and algorithmic steps (in which computations are executed).

The authors also show that entanglement purification can modify the network topology and affect the following routing design. As for general multi-request, multi-path, and multi-channel problems, the authors developed three algorithms that can allocate limited resource to efficiently generate entangled pairs between remote station pairs. Their stepwise routing scheme is based on one shared time window (the one corresponding to P4). It is expected that a queuing model might be constructed to tackle the network functioning beyond this processing time window. We think that such a queuing system is key in any of such schemes and we propose to use it in order to limit the number of commodities (or sender-receiver pairs) for a given time window related to (P2).

Let us analyse in the next section what such a framework is telling us: are the three challenges listed in Section 2 correctly addressed? If so, does the latter help to design an entanglement routing protocol fitting the end user requirements?

4 Analysis of the selected protocols - discussion

A few representative entanglement distribution and routing protocols have been mapped in Section 3 upon the same framework. This chosen set of protocols enabled us to list all the different tasks that could be required when designing an entanglement routing protocol: calculation of the repeater locations, (optional) fidelity pre-estimation, (optional) route pre-calculation, entanglement generation and distribution (with optional consideration of potential failures), qubit storage and retrieval, (optional) entanglement purification, path selection, Bell State measurements of selected pairs on each repeater along the paths. This makes it a universal framework in which one can select tasks and schedule them. Its phases 2, 3 and 4 can easily be mapped with the phases identified and described in the selected publications. Thanks to the larger view given by such a framework, we analyze in the present section whether the challenges stated in Section 2 (e.g. loss and noise due to long distance, dynamic or static routing,

and large number of demands) are sufficiently well covered and how to choose among the framework choice branches (e.g., on-demand or advance generation model).

4.1 Are the challenges correctly addressed ?

The tasks of phase 1 (repeater position calculation, fidelity pre-estimation) help to address the distance challenge ahead of the real-time routing concern. **Phase 2** introduces a time window during which a given number of quantum communications demands (i.e. multicommodity) can be received and scheduled; this enables a delineation between the tasks and workloads that can be done outside the constrained time slot related to real-time: for example path pre-calculation can be made and provided to the repeaters, with no consideration of the actual entangled resource distribution. Path selection occurs here in 'on-demand generation' models. This phase could clearly contribute to a less greedy routing algorithm ('routing challenge') and a more optimized resource allocation ('multi-commodity challenge').

Table 4: Are the three challenges correctly addressed ?

Distance Challenge	repeater position calculation (phase 1)
	entanglement distribution (phase 3)
	optical link fidelity pre-calculation (phase 1)
	optional purification (phase 4)
	local swappings to establish an E2E entangled pair (phase 4)
Entanglement Routing Challenge	path pre-calculation (phase 2)
	routing design (phase 1)
	(optional) local path selection (phase 4)
Multi-Commodity Challenge	multi-party management (phase 2)
	optional resource allocation strategy (executed during phase 4)

Phase 3 is at the heart of any entanglement network (LAN, MAN or WAN) and is common to all proposed entanglement-based protocols; it obviously deals with the distance challenge, but also with the multicommodities one (since one entangled resource cannot be shared by two connection demands). Hence, we may ask if phase 3 protocols should multiplex entanglement distribution in order to improve entanglement resource management.

Last, **Phase 4** concerns all Local Operations and Classical Communications (LOCC) to be performed by the repeaters during the actual routing time slot (routing here means measuring in the Bell basis); the actual list of tasks and their scheduling depend on whether one decides to design for example a 'forward' QEC protocol (with mainly one-way classical communication per cycle), a DLCZ-type protocol [27] or a SLMP-type protocol (with no encoded qubits).

Phase 4 contributes to the distance challenge (e.g. entanglement purification protocol), and to the multicommodities' one (how to allocate a set of entanglement resources to each partner communication), as shown in 4.

4.2 Different impacts on resource cost

Such a framework helps to highlight that the various entanglement routing protocols that could be designed to address the challenges above consume differently the underlying resources, such as:

- the number of distributed entangled pairs or ebits
- the number of classical communications in the control plane and therefore the overall process cycle time
- the routing tables stored on each node
- the (classical) computational resources required to deal with the routing algorithm complexity (see 3)

Concerning for example the cost in ebits:

- if the routes are calculated and selected prior to generating the entangled pairs (**Phase 3**), the corresponding '**on demand generation**' model will only have to distribute ebits (per edge) on the selected routes, the drawback being the size of the routing tables to store on each quantum repeater;
- if the ebits are generated and distributed (**Phase 3**) in advance to the actual route selection (**Phase 4**), the corresponding '**advance generation**' model will have to generate much more ebits (per edge), the drawback being the necessity to design a more dynamic routing protocol, able to cope only with local information.

Introducing a phase 2, prior to the entanglement generation and distribution phase (**Phase 3**), can bring the flexibility to choose one of the two models ('on demand generation' or 'advance generation') and therefore different ebit management schemes.

Concerning time, the main consumption could come from a too large number of classical communications that the designed routing protocol would require during **Phase 3** or **Phase 4**. For example, thanks to a not-heralded entanglement generation protocol and simultaneous entanglement connections, the process cycle time of $T * \tau_d$ (T being an integer) could just be the time for the synchronous entanglement generation on all edges (τ) plus the time corresponding to the Classical Communications between the $(h - 1)$ repeater stations and their Source-Destination; $T * \tau_d \sim (1 + h) * \tau_d$. Nevertheless, this is a lower bound, most of the referenced routing protocols requiring much more rounds of classical communications.

4.3 What are the choice branches given by such a framework? How to choose?

This framework opens up the possibilities of designing routing protocols that can best serve the needs of end users.

When one wishes to send one quantum information bit $|x\rangle = \alpha|0\rangle + \beta|1\rangle$ from Source S to Destination D, separated by $h - 1$ quantum repeaters: $(S - R_1)(R_1R_2)...(R_{h-1}D)$ the requirements on the quantum communication service provided by the WAN may differ depending on the type of the application: quantum key distribution, quantum sensing, distributed quantum computing, etc. To address this, [30] for example proposed to identify some Key Performance Indicators (KPI) and their possible value ranges, depending on the activity sectors:

1. one KPI is the end-to-end entanglement rate per Source-Destination pair, with or without guaranteed fidelity;
2. a second is the fidelity value requested by the end user applications;
3. the third KPI is the latency, or the process cycle time during which elementary entangled pairs are generated, distributed, distilled, swapped, providing the support for teleportation;
4. a fourth (related to the choice of a synchronization scheme [31]) is timing precision.

Let us take two examples: if the quantum bit $|x\rangle$ to transmit is one of a sequence constituting a 256-bit key, the complex numbers α , β and the latency are less important than the end-to-end rate (since there are at least 256 ones to transmit); if the quantum bit $|x\rangle$ to transmit is an intermediate result of a quantum computation, the complex numbers α , β and the latency are of paramount importance. We should soon be able to attach KPI range values to some choices that one would take within this framework, all based on the same network topology.

4.4 Discussion

Obviously, the literature starts to cover pretty well the list of tasks and problems that one will have when rolling out an entanglement network on top of an existing WAN infrastructure. As shown by Figure 4 and the referenced publications in Section 3, articulating correctly the four types of tasks (within phases 1 to 4) can enable to approach a solution to the problem of meeting the end-user requirements (e.g. communication latency, quantum fidelity or entanglement rate). Nevertheless, tackling perfectly the three challenges (distance, entanglement routing and multi-commodities) requires (1) to face an important issue in resource management (ebits, time, computational resources) and (2) to design a routing algorithm which maximizes the entangled routing rate of all demanding

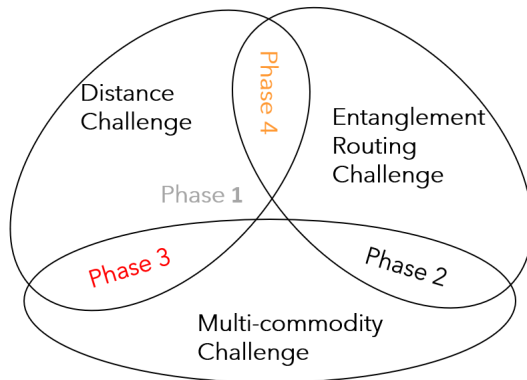


Figure 4: The proposed Framework Phases versus the Quantum Internet Challenges to meet

commodities (which is a NP-hard problem as recalled by [8] and represented by an empty intersection in the center of Figure 4).

From this framework, we need to embark in the use of a simulation tool, modelling the routing maximization problem, taking into account the technology constraints (e.g. quantum memory limitation, types of quantum repeaters [3]) and the key indicator values expected from the end users. Then, design choices like the selection between the on-demand or advance generation models would probably occur 'naturally'. More precisely, one needs to:

1. integrate the modelling of the main quantum entanglement network components,
2. study approximations of the entanglement routing problem, most multi-commodity routing problems being NP-hard (e.g. multi-commodity flow optimization [29]);
3. calculate the KPI patterns corresponding to some of the possible framework choices.

5 Conclusion

Quantum Internet, though still in its infancy, is a fast growing concept and a great challenge ahead. It could take the form of interconnected WANs, each WAN implementing an entanglement network on top of an existing (most probably optical fiber) infrastructure. To operate each WAN, one will need to calculate the optimal number of quantum repeater nodes and correctly position them, to work out heuristic methods to enable entanglement routing in a multi-commodity environment, and to combine several technologies (Single Photon Sources, Quantum Memories, repeaters of different - 1G to 3G - generations)

in a consistent way. The proposed framework, extracted from the survey of some referenced quantum routing protocols, is listing the tasks to carry out, highlighting the related cost in resources, and unifying the name of the protocol phases that can be found in most of the publications. As a next step, adding ranges of Key Performance Indicator values to some of quantum routing protocol patterns that such a framework will enable to design will complete the toolbox required for rolling out an entanglement network on top of an existing WAN infrastructure.

Acronyms

BSM Bell State Measurement.

EPR Einstein-Podolsky-Rosen.

QEC Quantum Error Correction.

WAN Wide Area Network.

References

- [1] Stephanie Wehner, David Elkouss, and Ronald Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412):eaam9288, 2018.
- [2] Ashlesha Patil, Joshua I Jacobson, Emily Van Milligen, Don Towsley, and Saikat Guha. Distance-independent entanglement generation in a quantum network using space-time multiplexed greenberger–horne–zeilinger (ghz) measurements. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 334–345. IEEE, 2021.
- [3] Sreram Muralidharan, Linshu Li, Jungsang Kim, Norbert Lütkenhaus, Mikhail D Lukin, and Liang Jiang. Optimal architectures for long distance quantum communication. *Scientific reports*, 6(1):1–10, 2016.
- [4] Shouqian Shi and Chen Qian. Modeling and designing routing protocols in quantum networks. *arXiv preprint arXiv:1909.09329*, 2019.
- [5] Mihir Pant, Hari Krovi, Don Towsley, Leandros Tassiulas, Liang Jiang, Prithwish Basu, Dirk Englund, and Saikat Guha. Routing entanglement in the quantum internet. *npj Quantum Information*, 5(1):1–9, 2019.
- [6] Paolo Fittipaldi, Anastasios Giovanidis, and Frédéric Grosshans. A linear algebraic framework for quantum internet dynamic scheduling. *arXiv preprint arXiv:2205.10000*, 2022.

- [7] H-J Briegel, Wolfgang Dür, Juan I Cirac, and Peter Zoller. Quantum repeaters: the role of imperfect local operations in quantum communication. Physical Review Letters, 81(26):5932, 1998.
- [8] Tu N Nguyen, Dung HP Nguyen, Dang H Pham, Bing-Hong Liu, and Hoa N Nguyen. Maximizing entanglement routing rate in quantum networks: Approximation algorithms. arXiv preprint arXiv:2207.11821, 2022.
- [9] Evgeny Shchukin, Ferdinand Schmidt, and Peter van Loock. Waiting time in quantum repeaters with probabilistic entanglement swapping. Physical Review A, 100(3):032322, 2019.
- [10] OA Collins, SD Jenkins, A Kuzmich, and TAB Kennedy. Multiplexed memory-insensitive quantum repeaters. Physical review letters, 98(6):060502, 2007.
- [11] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. Fundamental limits of repeaterless quantum communications. Nature communications, 8(1):1–15, 2017.
- [12] Prajit Dhara, Norbert M Linke, Edo Waks, Saikat Guha, and Kaushik P Shadreesan. Multiplexed quantum repeaters based on dual-species trapped-ion systems. Physical Review A, 105(2):022623, 2022.
- [13] KD Goodenough. Distributing entanglement in quantum networks. PhD thesis, Delft University of Technology, Netherlands, 2022.
- [14] Benjamin Schumacher. Sending entanglement through noisy quantum channels. Physical Review A, 54(4):2614, 1996.
- [15] Reinhard F Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. Physical Review A, 40(8):4277, 1989.
- [16] Charles H Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A Smolin, and William K Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. Physical review letters, 76(5):722, 1996.
- [17] David Deutsch, Artur Ekert, Richard Jozsa, Chiara Macchiavello, Sandu Popescu, and Anna Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. Physical review letters, 77(13):2818, 1996.
- [18] Jian Li, Mingjun Wang, Qidong Jia, and Kaiping Xue. Fidelity-guaranteed entanglement purification and routing in quantum networks. arXiv preprint arXiv:2111.07764, 2021.
- [19] Sebastiaan Brand, Tim Coopmans, and David Elkouss. Efficient computation of the waiting time and fidelity in quantum repeater chains. IEEE Journal on Selected Areas in Communications, 38(3):619–639, 2020.

- [20] Kaushik Chakraborty, David Elkouss, Bruno Rijsman, and Stephanie Wehner. Entanglement distribution in a quantum network: A multicommodity flow-based approach. IEEE Transactions on Quantum Engineering, 1:1–21, 2020.
- [21] Julian Rabbie, Kaushik Chakraborty, Guus Avis, and Stephanie Wehner. Designing quantum networks using preexisting infrastructure. npj Quantum Information, 8(1):1–12, 2022.
- [22] Rodney Van Meter, Takahiko Satoh, Thaddeus D Ladd, William J Munro, and Kae Nemoto. Path selection for quantum repeater networks. Networking Science, 3(1):82–95, 2013.
- [23] A Pirker, J Wallnöfer, and W Dür. Modular architectures for quantum networks. New Journal of Physics, 20(5):053054, 2018.
- [24] Evan Meyer-Scott, Christine Silberhorn, and Alan Migdall. Single-photon sources: Approaching the ideal through multiplexing. Review of Scientific Instruments, 91(4):041101, 2020.
- [25] Alexander Pirker and Wolfgang Dür. A quantum network stack and protocols for reliable entanglement-based networks. New Journal of Physics, 21(3):033003, 2019.
- [26] Yangming Zhao and Chunming Qiao. Redundant entanglement provisioning and selection for throughput maximization in quantum networks. In IEEE INFOCOM 2021-IEEE Conference on Computer Communications, pages 1–10. IEEE, 2021.
- [27] L-M Duan, Mikhail D Lukin, J Ignacio Cirac, and Peter Zoller. Long-distance quantum communication with atomic ensembles and linear optics. Nature, 414(6862):413–418, 2001.
- [28] Andrew M Steane. Enlargement of calderbank-shor-steane quantum codes. IEEE Transactions on Information Theory, 45(7):2492–2495, 1999.
- [29] Changhao Li, Tianyi Li, Yi-Xiang Liu, and Paola Cappellaro. Effective routing design for remote entanglement generation on quantum networks. npj Quantum Information, 7(1):1–12, 2021.
- [30] Laurent de Forges de Parny, Olivier Alibart, Julien Debaud, Sacha Gresani, Alek Lagarrigue, Anthony Martin, Alexandre Metrat, Matteo Schiavon, Tess Troisi, Eleni Diamanti, et al. Satellite-based quantum information networks: Use cases, architecture, and roadmap. arXiv preprint arXiv:2202.01817, 2022.
- [31] Virginia D’Auria, Bruno Fedrici, Lutfi Arif Ngah, Florian Kaiser, Laurent Labonté, Olivier Alibart, and Sébastien Tanzilli. A universal, plug-and-play synchronisation scheme for practical quantum networks. npj Quantum Information, 6(1):1–6, 2020.