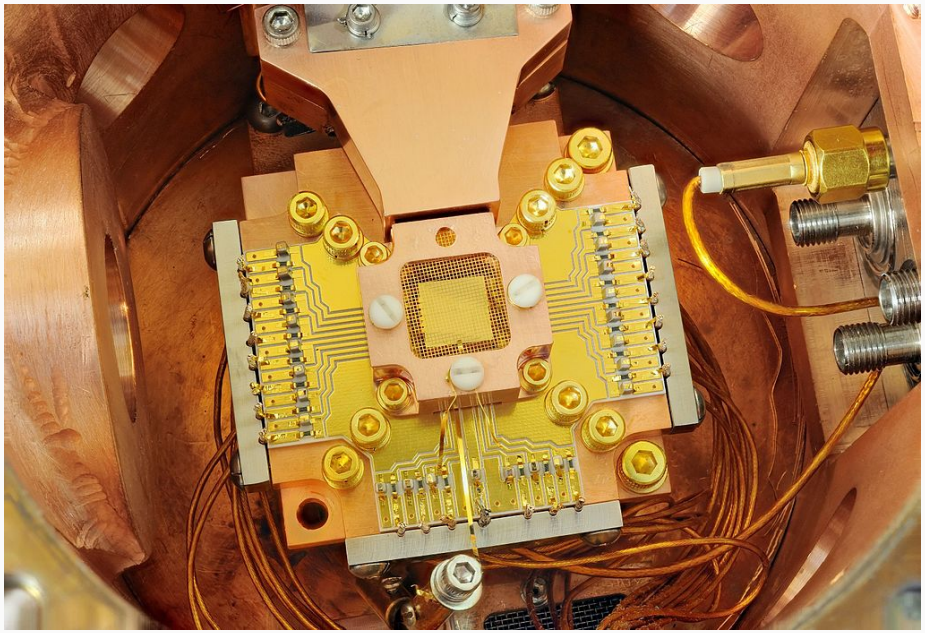
# Quantum Computer Architecture

## Scalable and Reliable Quantum Computers

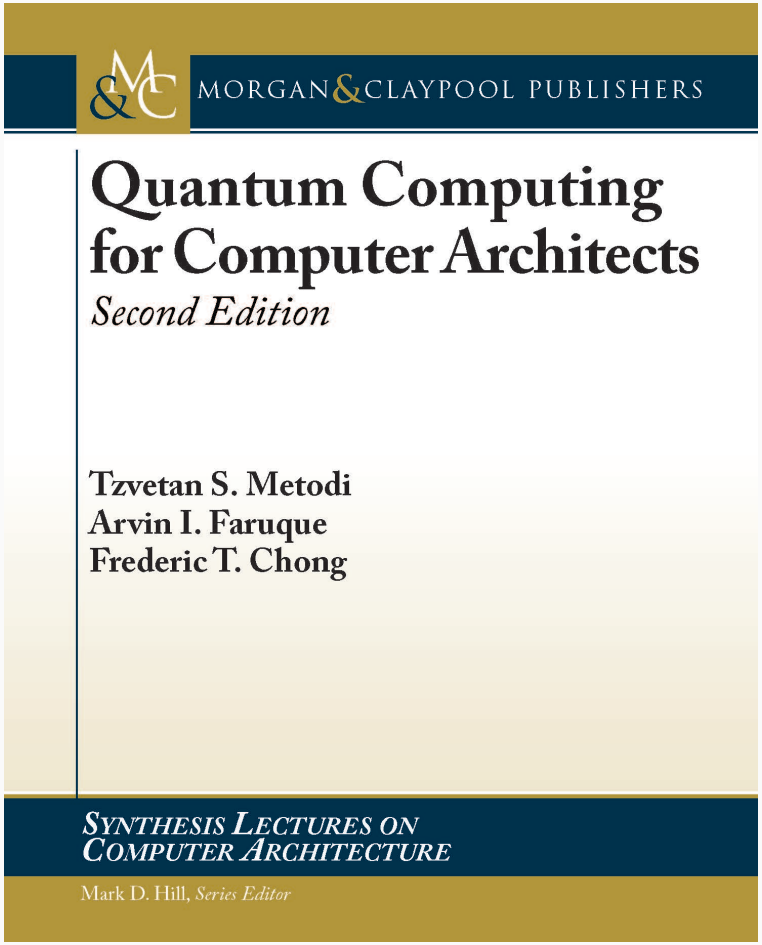
Greg Byrd (ECE)

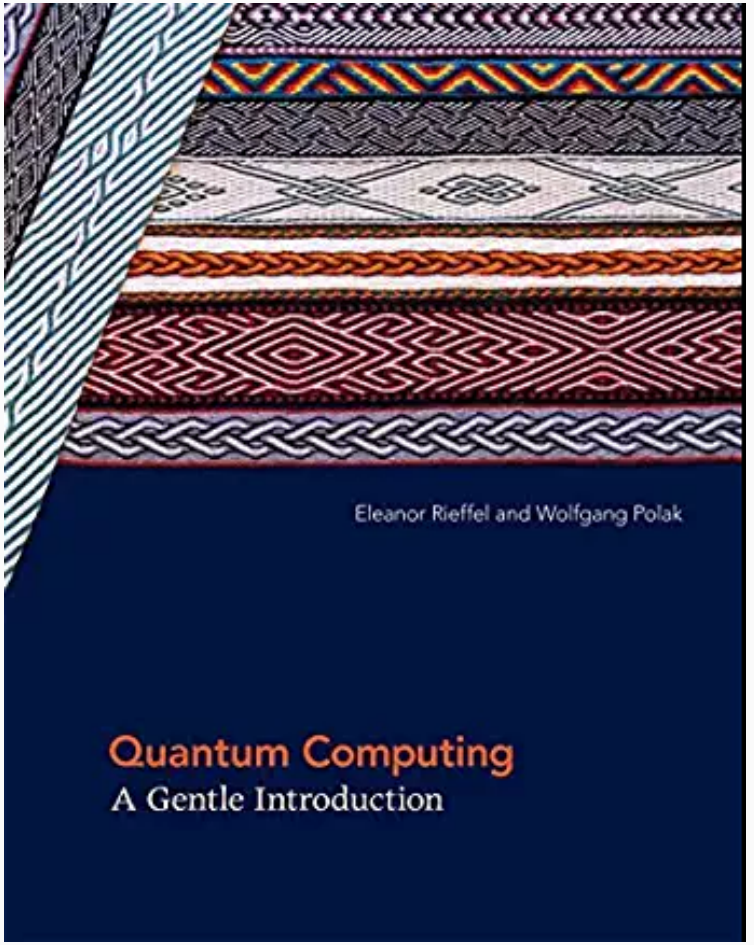
CSC 801 - Feb 13, 2018



# Overview

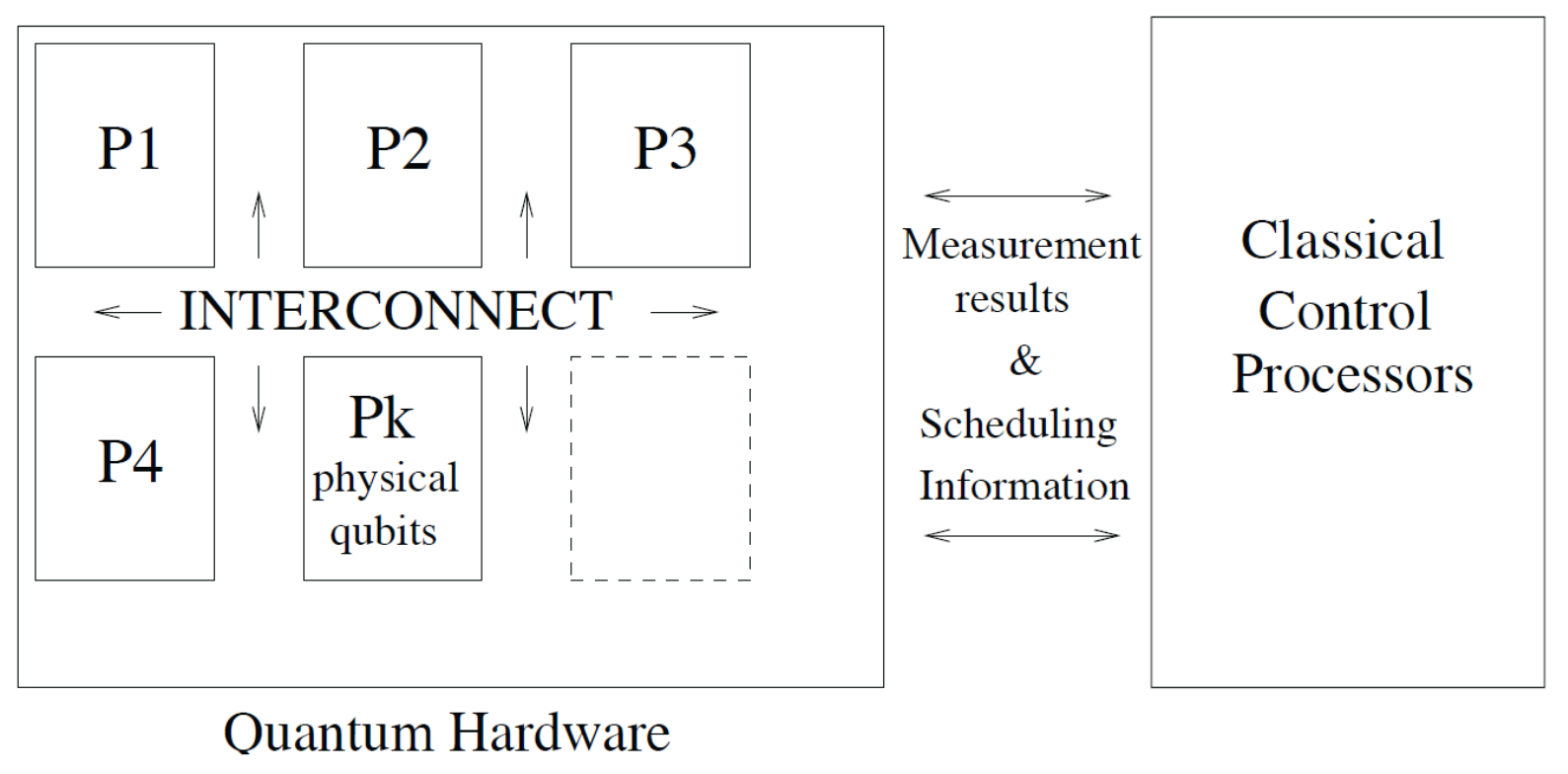
# Sources


MORGAN & CLAYPOOL PUBLISHERS
Quantum Computing for Computer Architects
Second Edition
Tzvetan S. Metodi
Arvin I. Faruque
Frederic T. Chong
SYNTHESIS LECTURES ON COMPUTER ARCHITECTURE
Mark D. Hill, Series Editor


Eleanor Rieffel and Wolfgang Polak
Quantum Computing
A Gentle Introduction

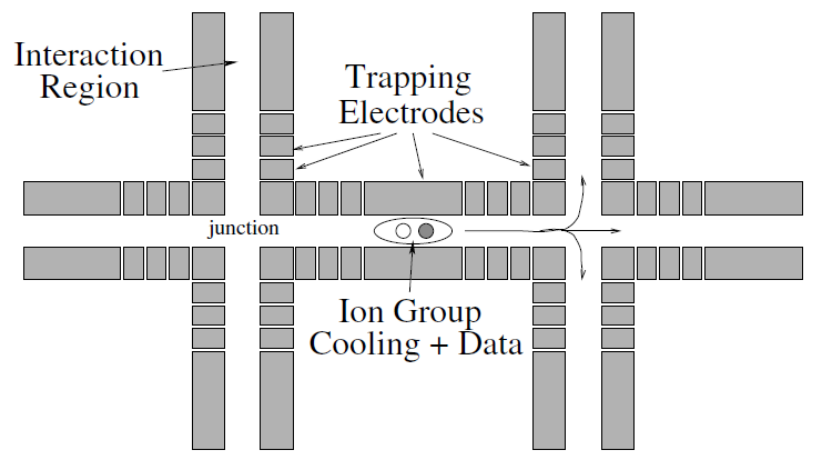
# Key Concepts

# Quantum Computer

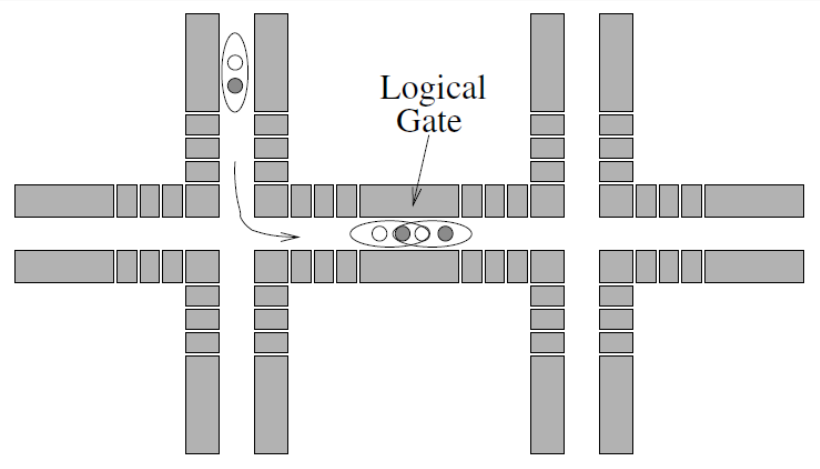


# Outline

# Ion Trap Operation

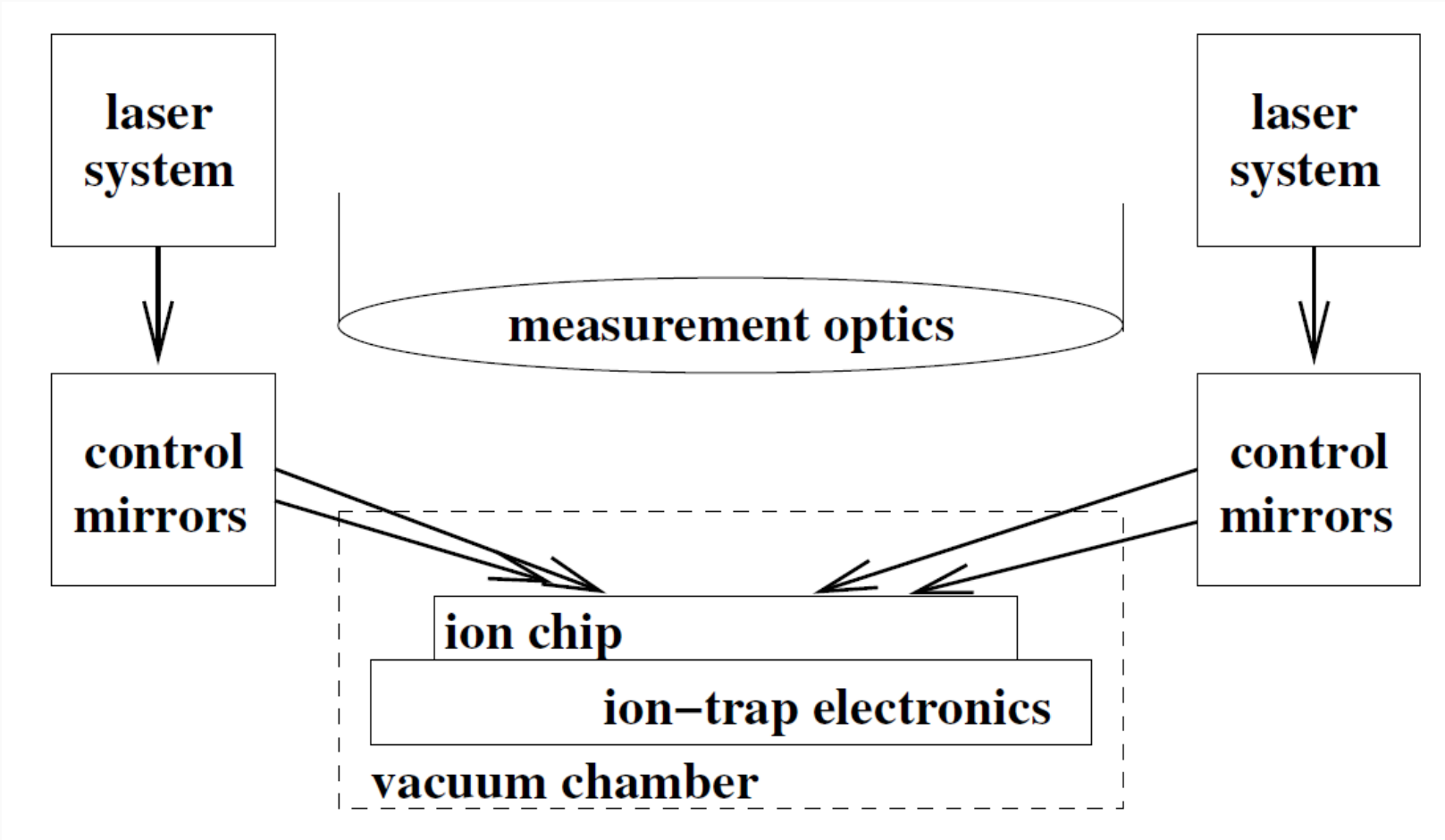



The ion can move into any of the six adjacent trapping regions.




Move two qubits together to perform gate operation.

# Ion Trap System



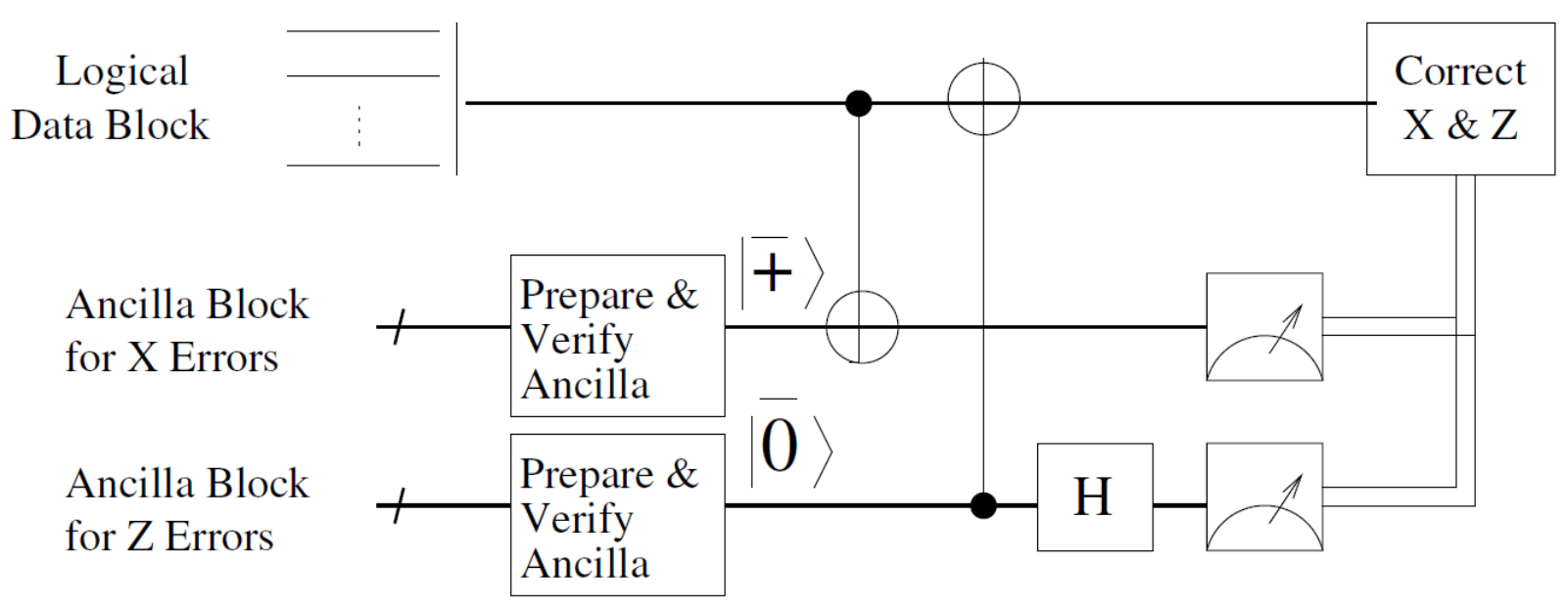
# Outline

# Error Model

"...errors are so frequent that the **speed of the error correction method** used is critical for the application run-time."

- Quantum states are continuous and therefore errors are continuous. In principle, it could take infinite resources to determine exactly what the error is.
- Measurement destroys the superposition of quantum data, but the only way to extract the error syndrome bits is to measure. Therefore, we can only measure indirectly.
- Quantum data is fundamentally more faulty than classical data. (E.g., 1 error in $10^8$ operations.)
- Quantum states are entangled, affecting how errors propagate.

# Steane Method for Error Correction




Two blocks of ancillary qubits, same size as encoded logical qubit.

1. CNOT propagates bit flips (X errors) to $|\overline{+}\rangle$. Measure and correct.
2. Reverse CNOT propagates phase flips (Z errors) to $|\overline{0}\rangle$. Measure and correct.

Can use one set of ancillary qubits, if correction in series.

# Steane [[7,1,3]] Code

One logical qubit is encoded using seven physical qubits.

$$\begin{aligned} |\bar{0}\rangle \quad = \quad & |0000000\rangle + |1111000\rangle + |1100110\rangle + |1010101\rangle + \\ & |0011110\rangle + |0101101\rangle + |0110011\rangle + |1001011\rangle \end{aligned}$$

Smallest code that allows *transversal* logical gates for:

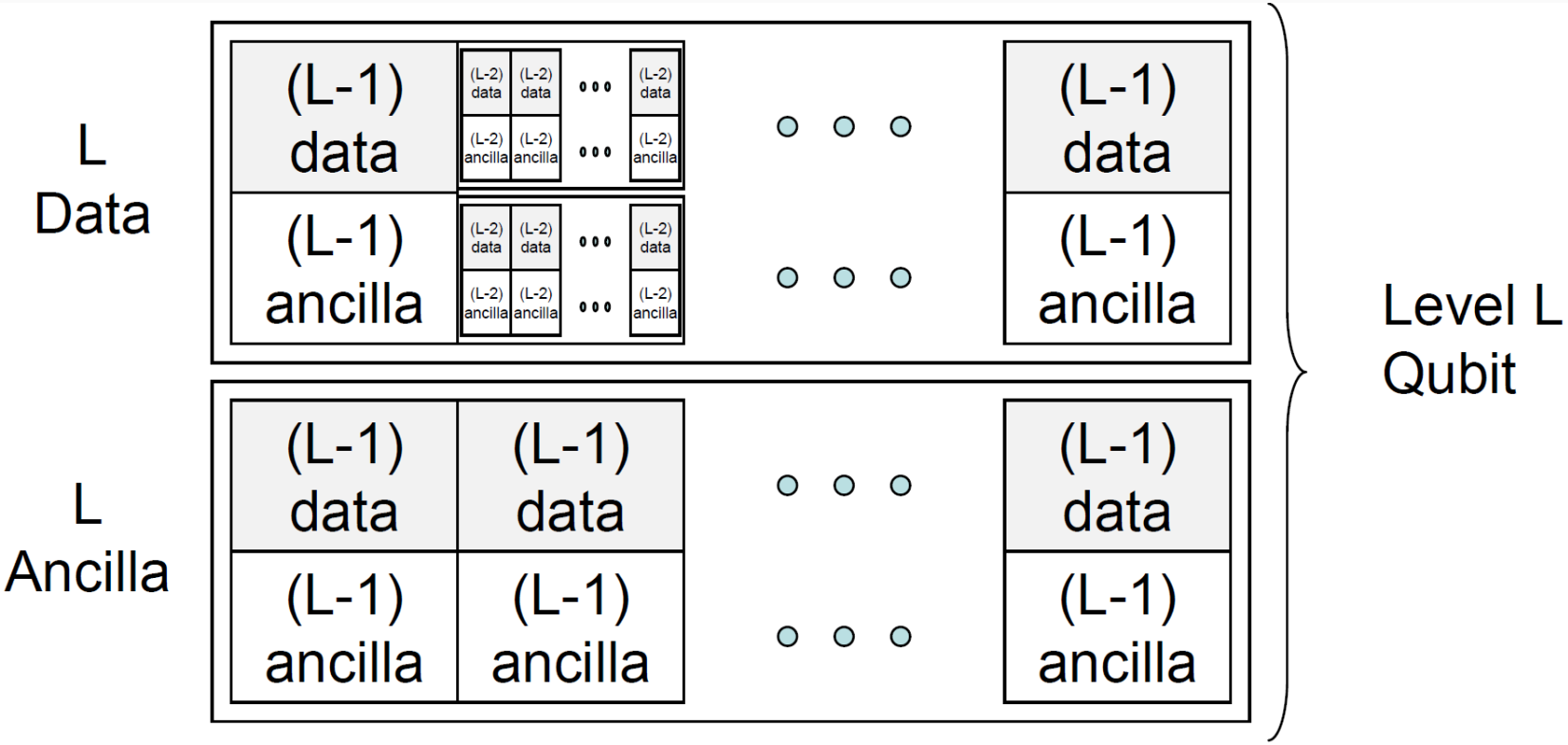
$$H, \mathrm{CNOT}, X, Z, Y, S$$

Transversal means that the gate is applied in parallel to all component qubits. The T gate is more complicated.

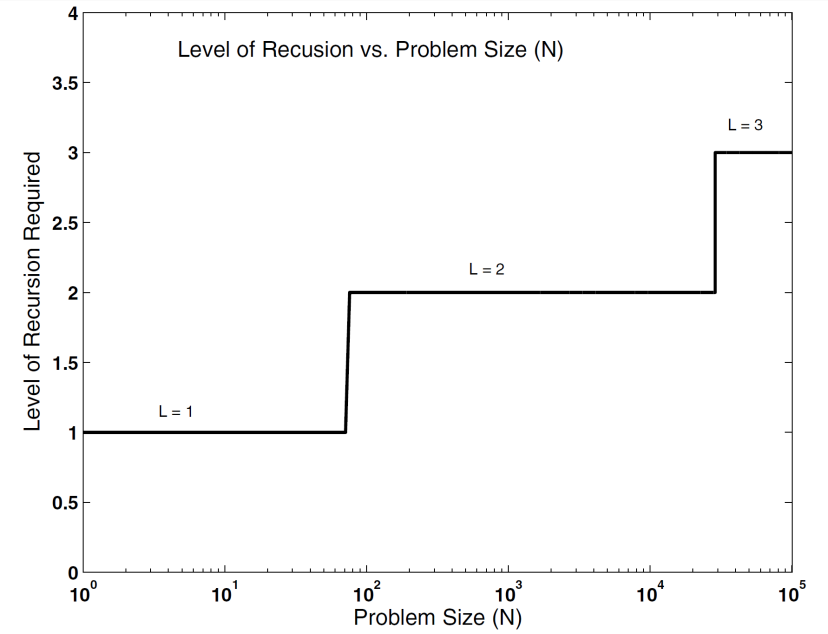
# Recursive Error Correction Code

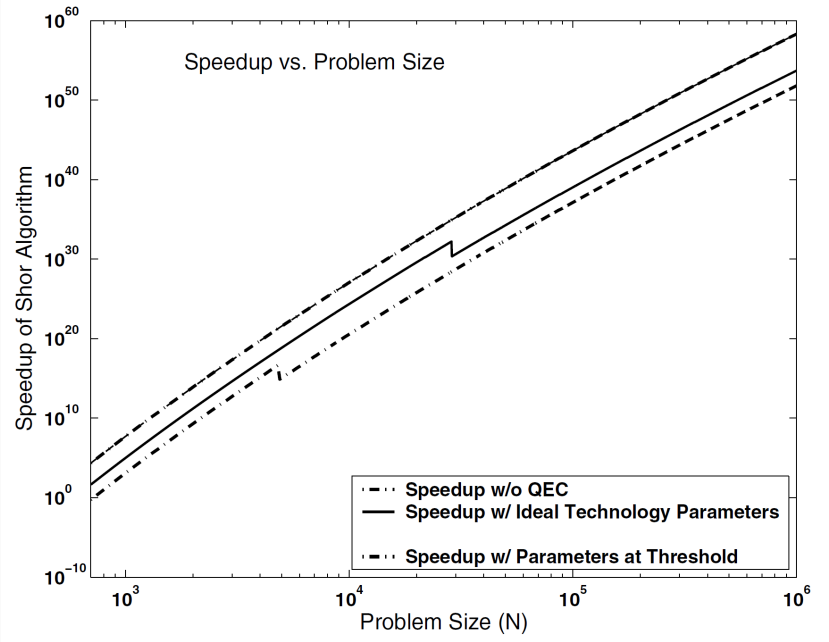
Recursive error coding allows error rate to be scaled according to the problem size.

- As problem scales (in time and qubits), error rate per logical operation (gate or move) must decrease.



# Shor Factorization Speedup



# Outline

# Movement of Qubits

For most physical implementation technologies (e.g., ion trap), qubits must be near each other to interact in a 2-qubit gate.

Movement is challenging:

- Time to physically move
- Decoherence (error) introduced by movement ops
- Many physical qubits must be moved (e.g., 49 for L2)
- Higher recursion in QEC $\Rightarrow$ qubits are further apart

In large applications, long-range qubit movement is needed.

# No-Cloning Rule

**Proof.**

Suppose a cloning operator $U$, such that $U|a0\rangle = |aa\rangle$ for all $|a\rangle$.

Let $|a\rangle$ and $|b\rangle$ be orthogonal quantum states.

$$U|a0\rangle = |aa\rangle \,, \text{and} U|b0\rangle = |bb\rangle$$

Consider $|c\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$. By linearity,

$$U|c0\rangle = \frac{1}{\sqrt{2}}(U|a0\rangle + U|b0\rangle) = \frac{1}{\sqrt{2}}(|aa\rangle + |bb\rangle)$$

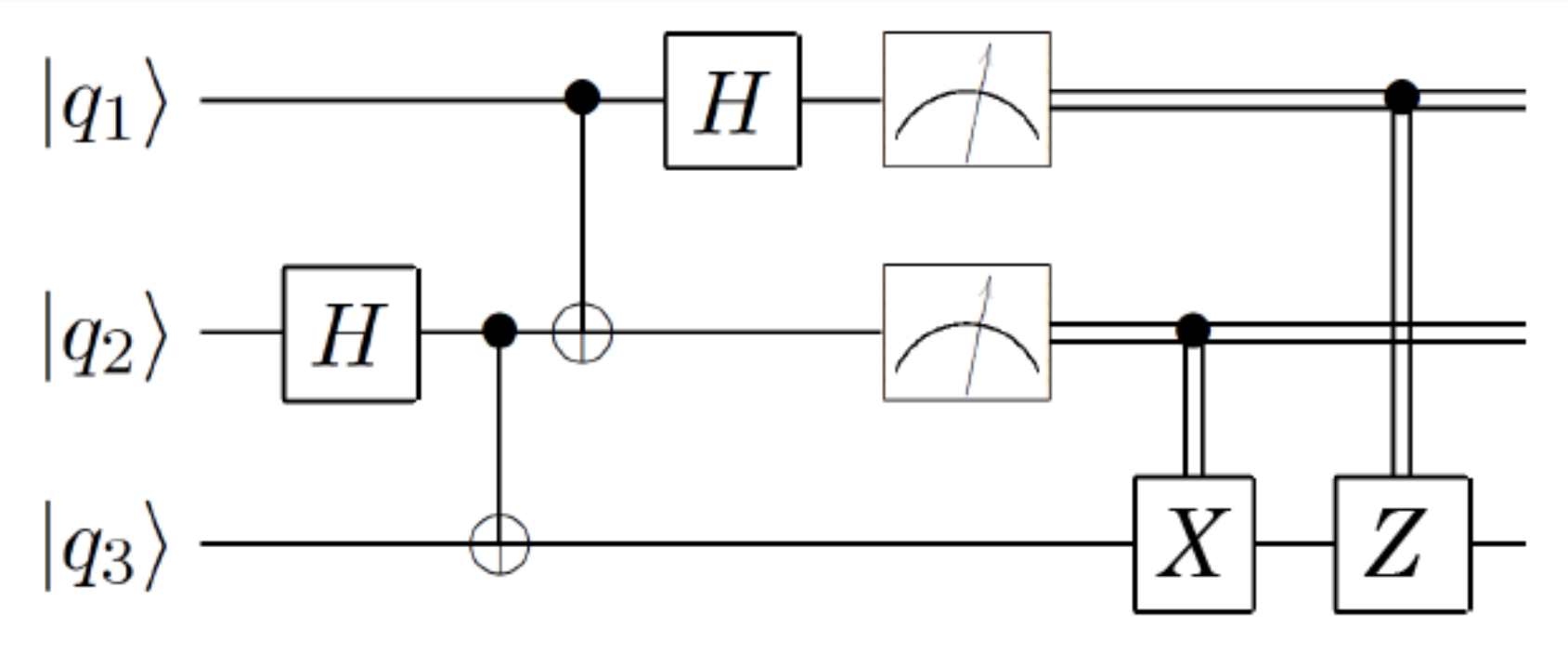
However, if U is a cloning operator, then

$$U|c0\rangle = |cc\rangle = \frac{1}{2}(|aa\rangle + |bb\rangle + |ab\rangle + |ba\rangle) \neq \frac{1}{\sqrt{2}}(|aa\rangle + |bb\rangle)$$

□

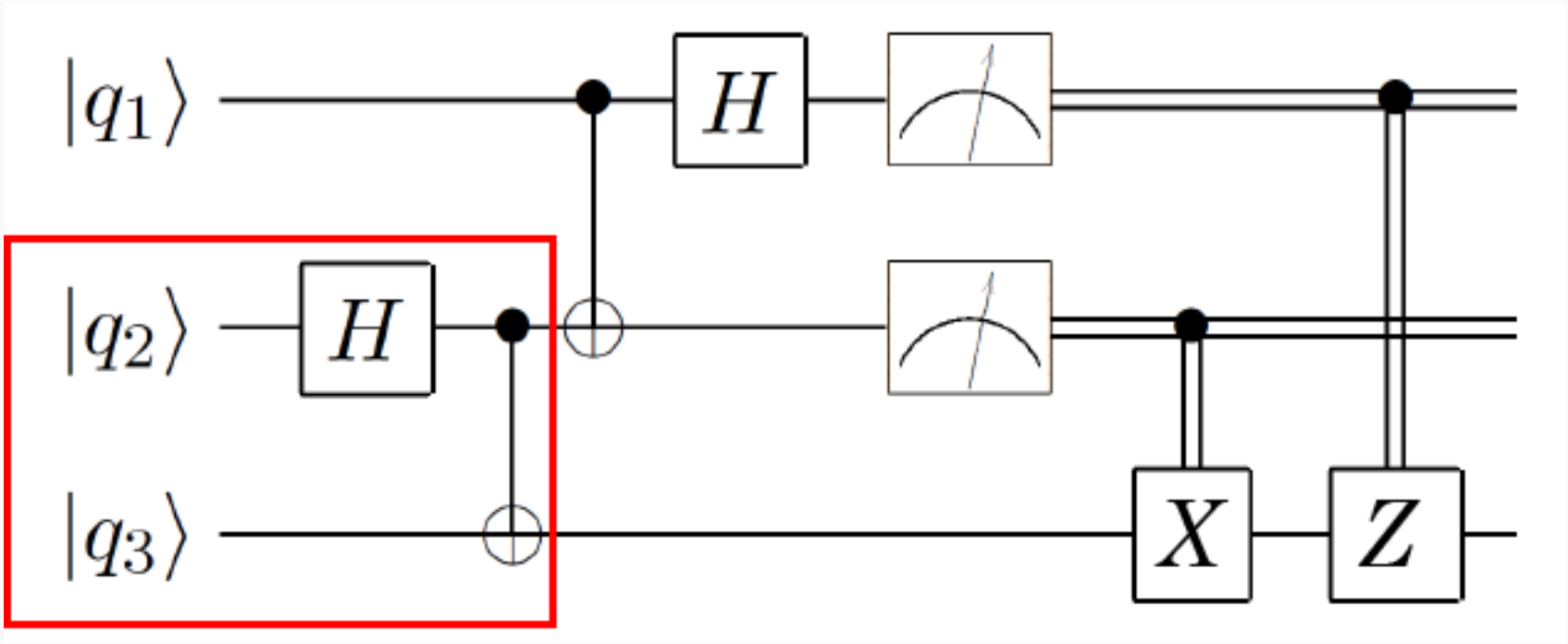
# Quantum Teleportation

However, we can use entangled EPR pair to achieve teleportation. We can recreate a state $|q_1\rangle$, even at a distance, but we have to measure the original qubit (and thereby destroy its state).



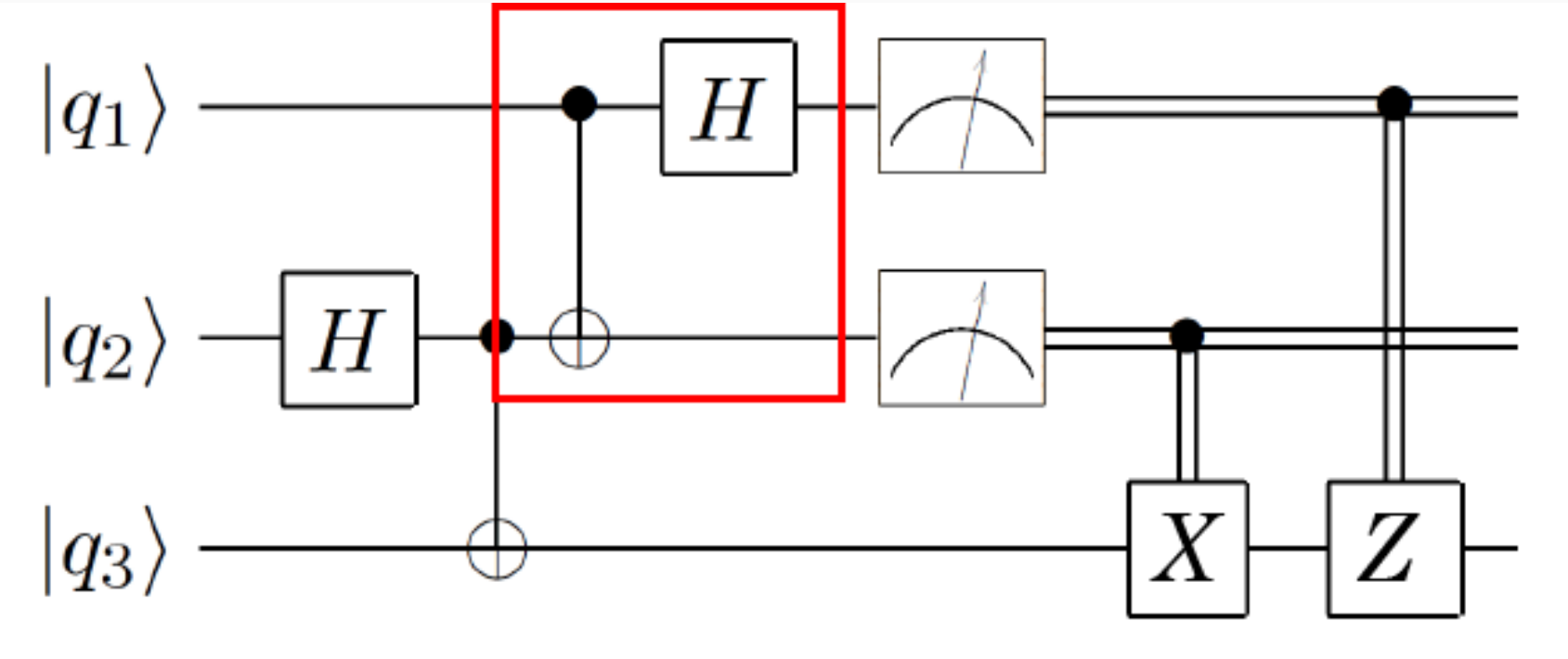
# Teleportation: Step 1

Alice has qubit $|q_1\rangle$ in unknown state $a|0\rangle + b|1\rangle$.




Create EPR pair $|q_2 q_3\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.
Alice keeps $|q_2\rangle$ and sends $|q_3\rangle$ to Bob.

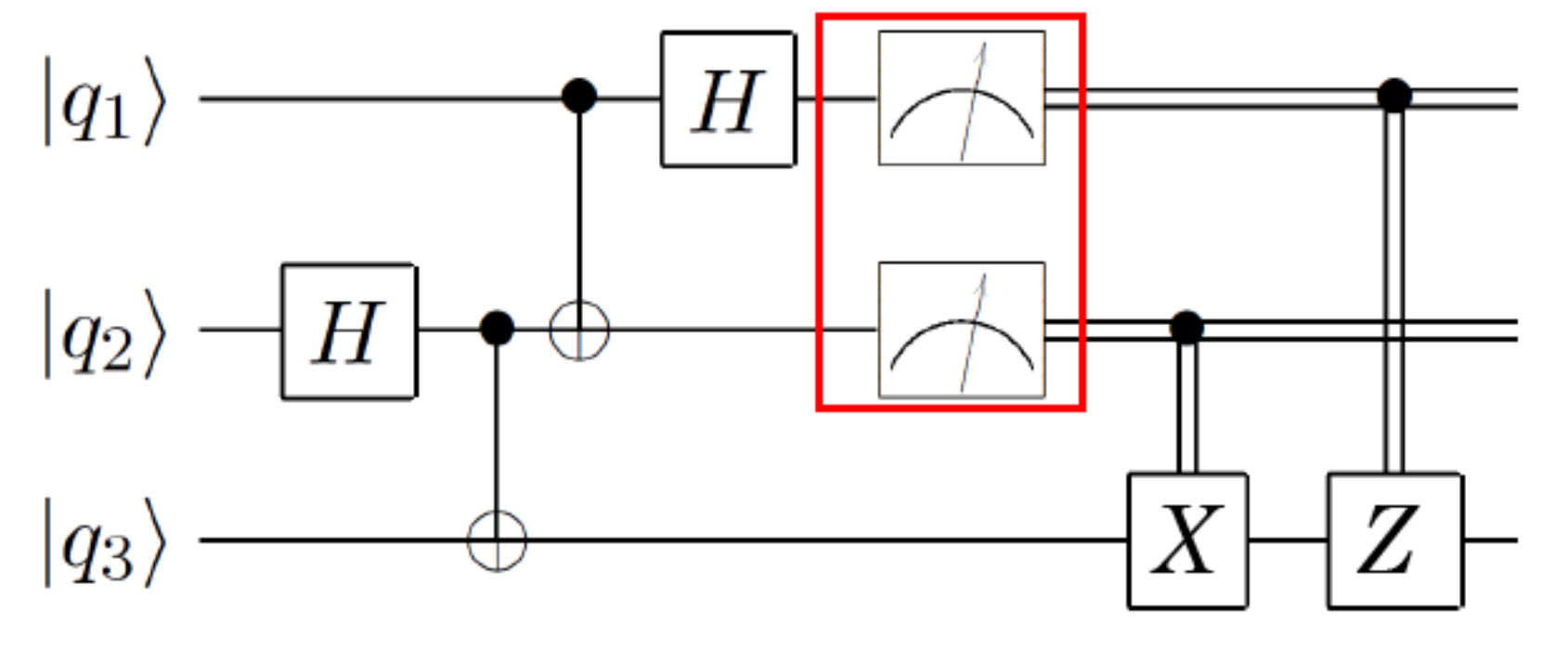
# Teleportation: Step 2




After CNOT and H gate:

$$
\begin{aligned}
|q_1 q_2 q_3\rangle = \frac{1}{2}( \quad & |00\rangle\,(a\,|0\rangle + b\,|1\rangle) \\
& + |01\rangle\,(a\,|1\rangle + b\,|0\rangle) \\
& + |10\rangle\,(a\,|0\rangle - b\,|1\rangle) \\
& + |11\rangle\,(a\,|1\rangle - b\,|0\rangle))
\end{aligned}
$$

# Teleportation: Step 3

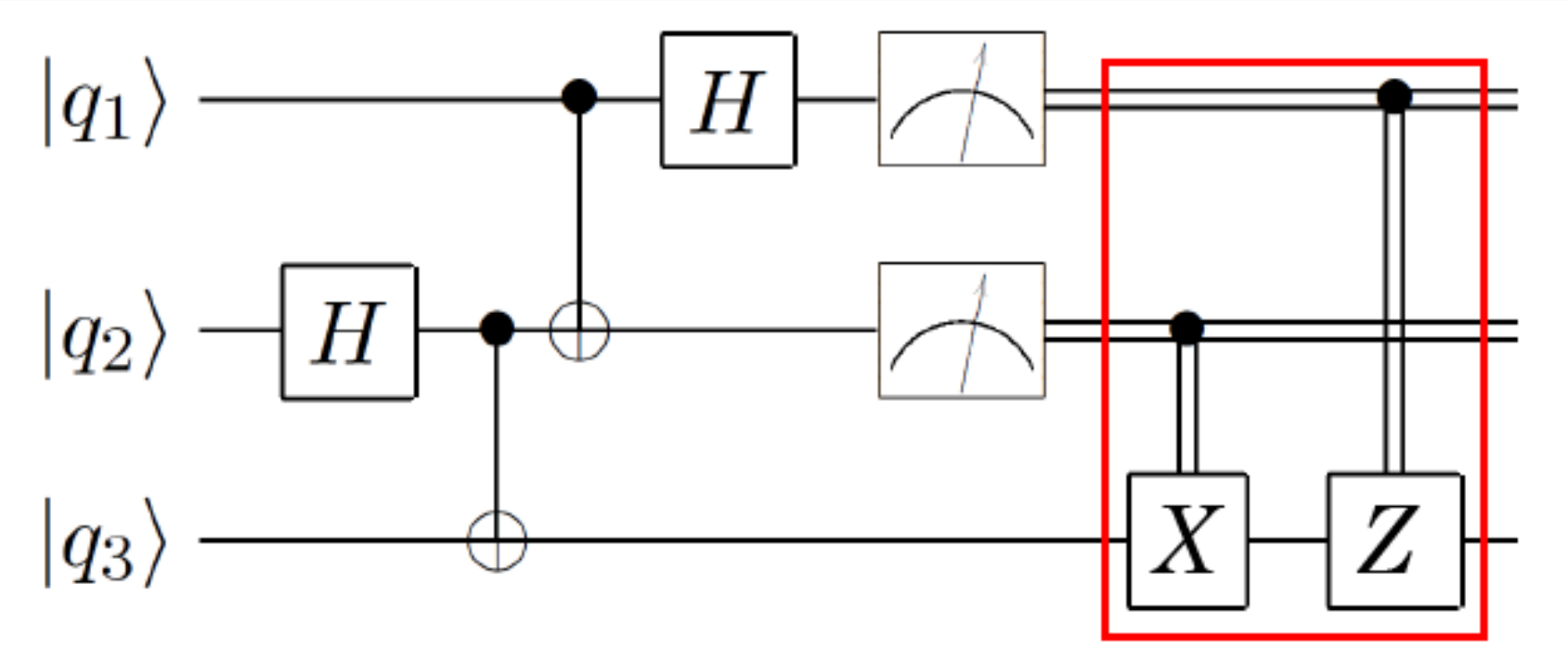



When Alice measures $|q_1 q_2\rangle$, we know the state of Bob's qubit $|q_3\rangle$, and what transform is needed to restore it to the original state. (See next slide.)

# Post-measurement Transform

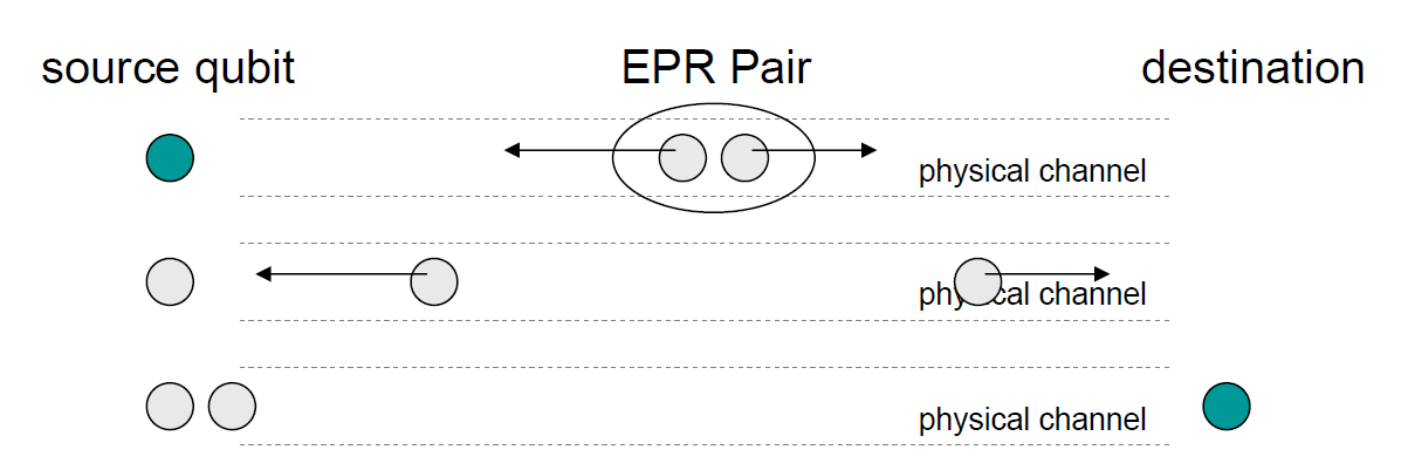
| $\lvert q_1 q_2 \rangle$ | $\lvert q_3 \rangle$ | Transform | Result |
|---|---|---|---|
| $\lvert 00 \rangle$ | $a\lvert 0 \rangle + b\lvert 1 \rangle$ | $I$ | $a\lvert 0 \rangle + b\lvert 1 \rangle$ |
| $\lvert 01 \rangle$ | $a\lvert 1 \rangle + b\lvert 0 \rangle$ | $X$ | $a\lvert 0 \rangle + b\lvert 1 \rangle$ |
| $\lvert 10 \rangle$ | $a\lvert 0 \rangle - b\lvert 1 \rangle$ | $Z$ | $a\lvert 0 \rangle + b\lvert 1 \rangle$ |
| $\lvert 11 \rangle$ | $a\lvert 1 \rangle - b\lvert 0 \rangle$ | $ZX$ | $a\lvert 0 \rangle + b\lvert 1 \rangle$ |

# Teleportation: Step 4



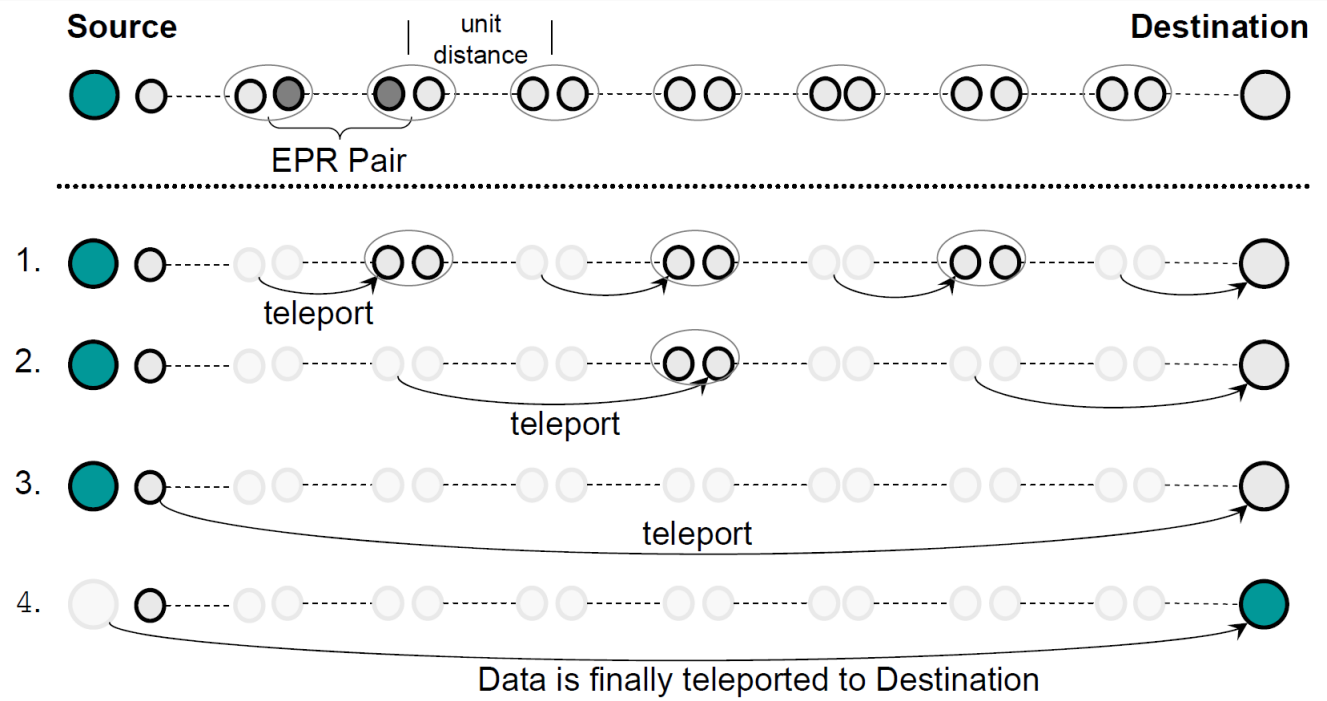

Alice sends two cbits (result of measurements) to Bob, who then performs the needed transforms.

# Communication through Teleportation




NOTE: Still moving the EPR qubit. Degrades over time/distance. But EPR pairs can be replaced/repaired (EPR *purification*).

# Quantum Repeaters




EPR entanglement survives teleportation. Can stage multiple pairs so that each qubit need not travel far.

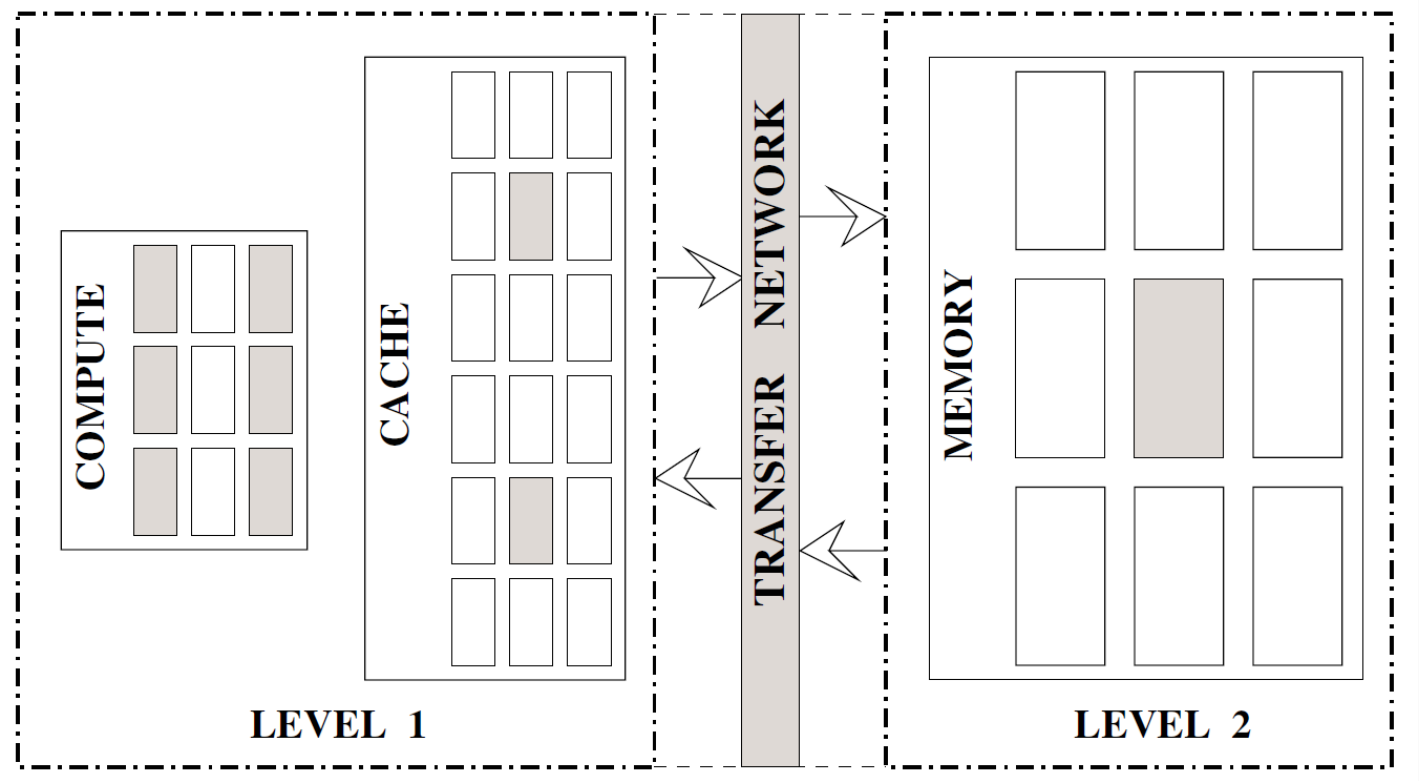
Picture for third phase is wrong. Should be teleporting EPR qubit in the middle to the right.

# Architectural Elements

# Two Fundamental Components

- Logical qubit
  - Apply 1- and 2-qubit gates
  - Error correction
- Communications channels
  - Teleportation-based
  - Quantum repeaters
  - EPR pair creation and purification

# High-Level Architecture



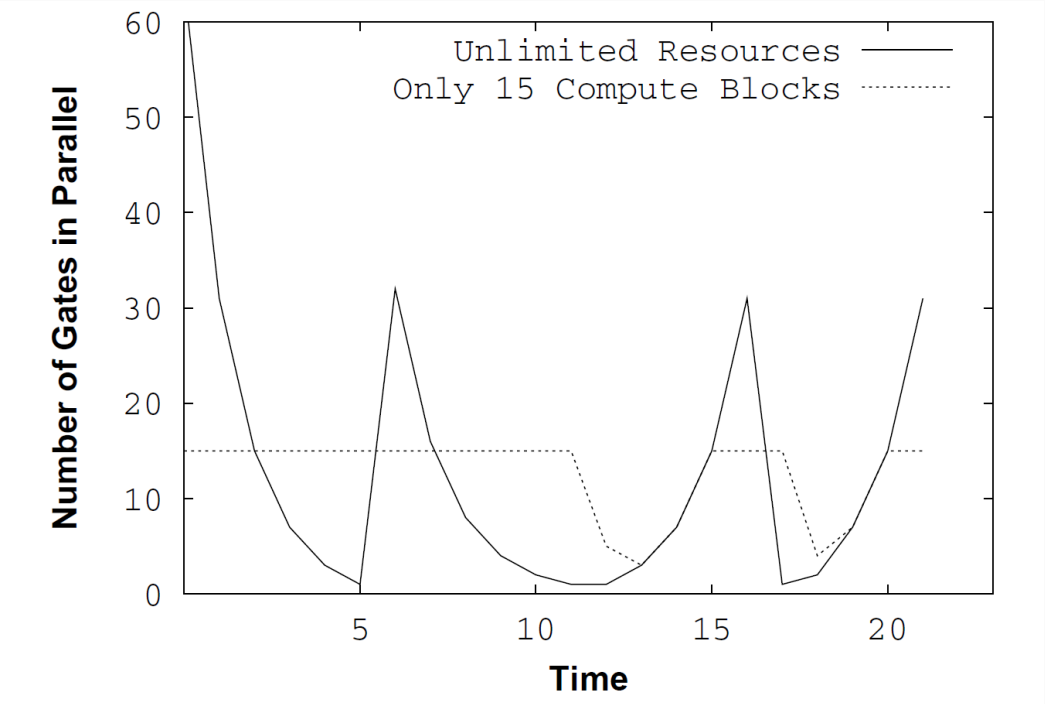

Blank tiles: logical data qubits

Shaded tiles: logical ancilla tiles

# Compute vs. Memory Tiles

Reasons to specialize tiles for compute vs. storage

- Different encoding schemes
- Limited qubit-level parallelism

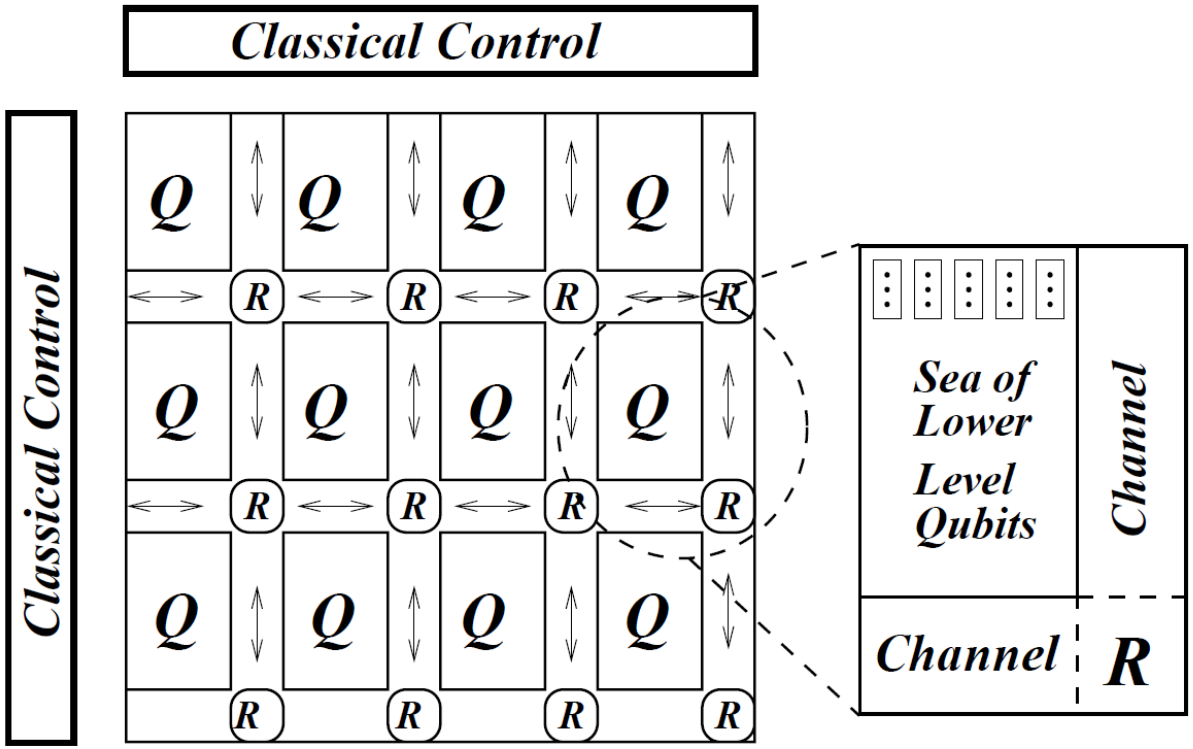



Parallelism vs. time for 64-bit adder

# QLA: Quantum Logic Array

# Quantum Logic Array

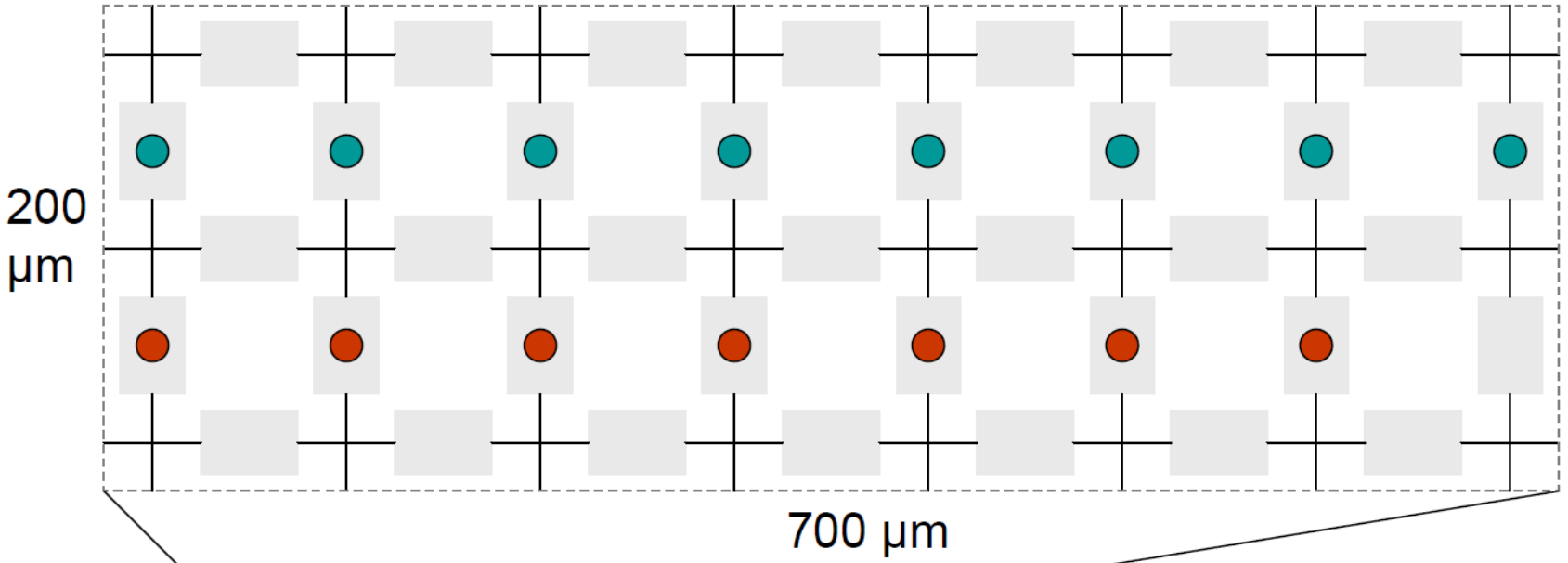
A detailed proposed architecture based on ion trap technology.



# Parameters

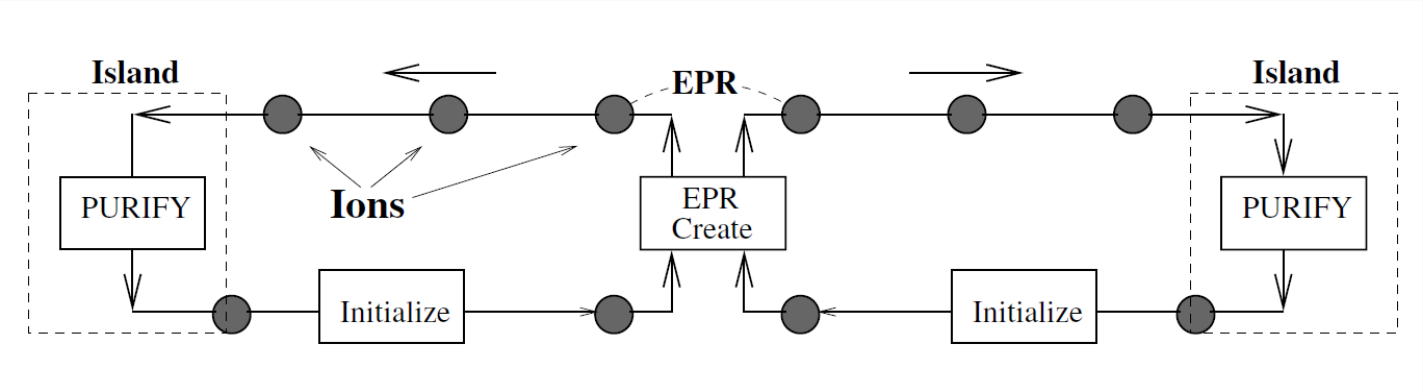
| Operation | Time $\mu$s now(future) | Failure Rate now(future) |
|---|---|---|
| Single Gate | 1 (1) | $10^{-4}$ ($10^{-8}$) |
| Double Gate | 10 (10) | 0.03 ($10^{-7}$) |
| Measure | 200 (10) | 0.01 ($10^{-8}$) |
| Movement | 20 (10) | 0.005 ($5 \times 10^{-8}$)/$\mu$m |
| Split | 200 (0.1) | |
| Cooling | 200 (0.1) | |
| Memory time | 10 to 100 sec | |
| Trap Size | $\sim$ 200 (1 – 5) $\mu$m | |

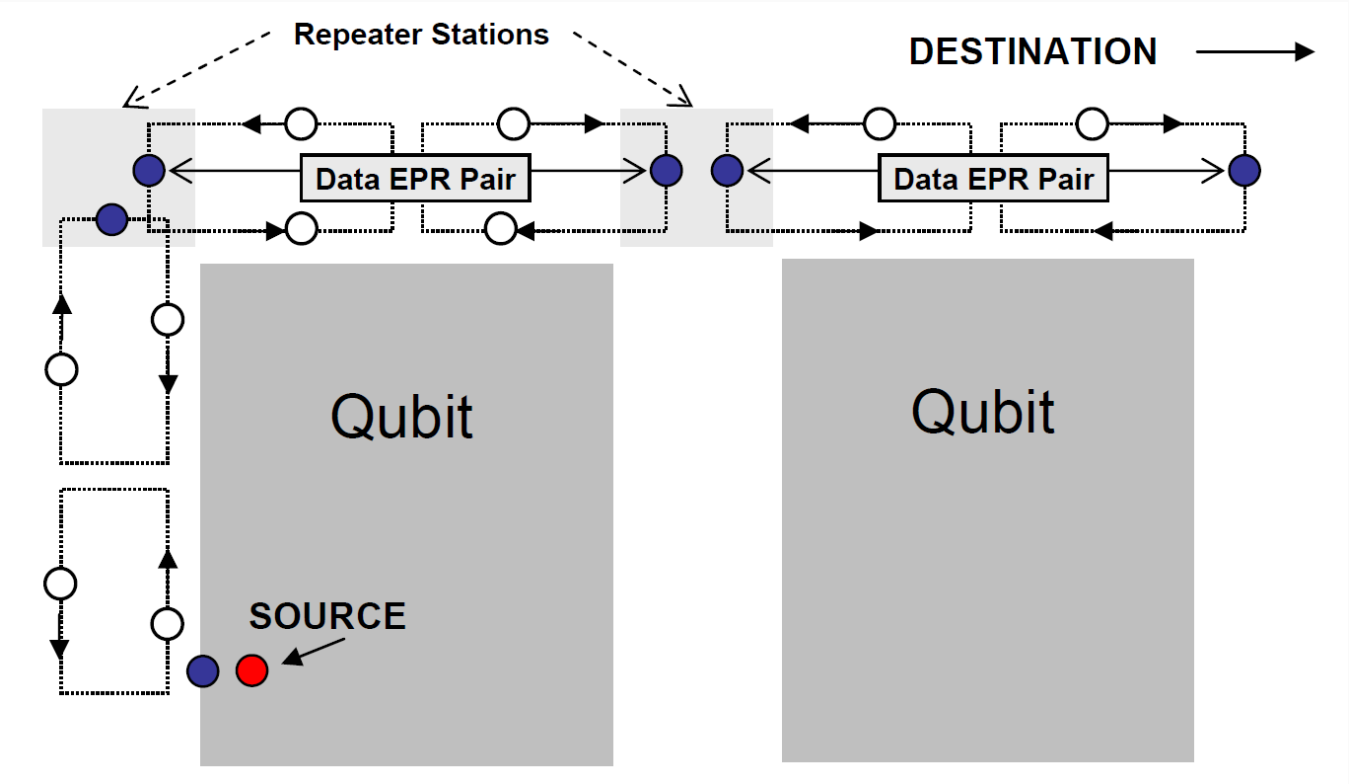
# Logical (L2) Qubit



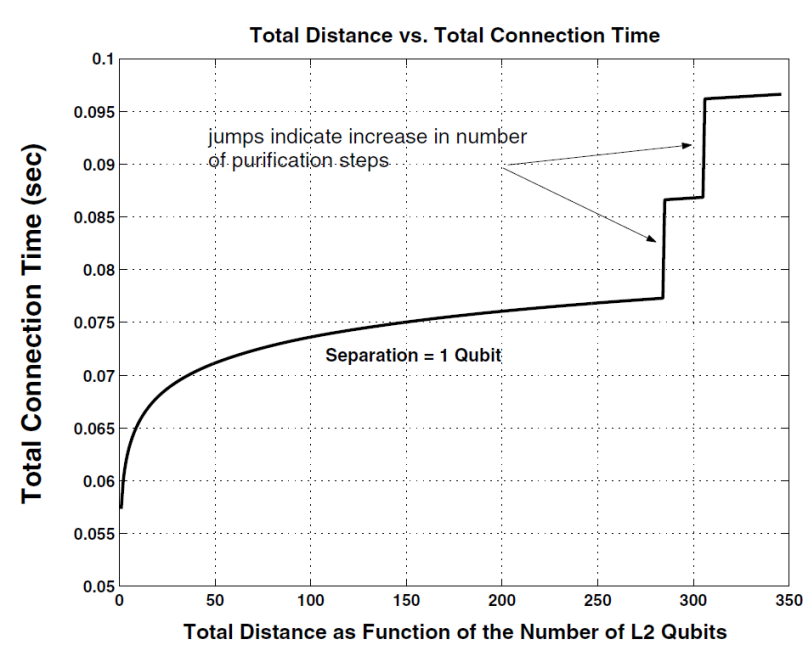

| | | | | | | |
|---|---|---|---|---|---|---|
| A1 | | A3 | A4 | A5 | A6 | A7 |
| D1 | D2 | D3 | D4 | D5 | D6 | D7 |
| A1 | A2 | A3 | A4 | A5 | A6 | A7 |

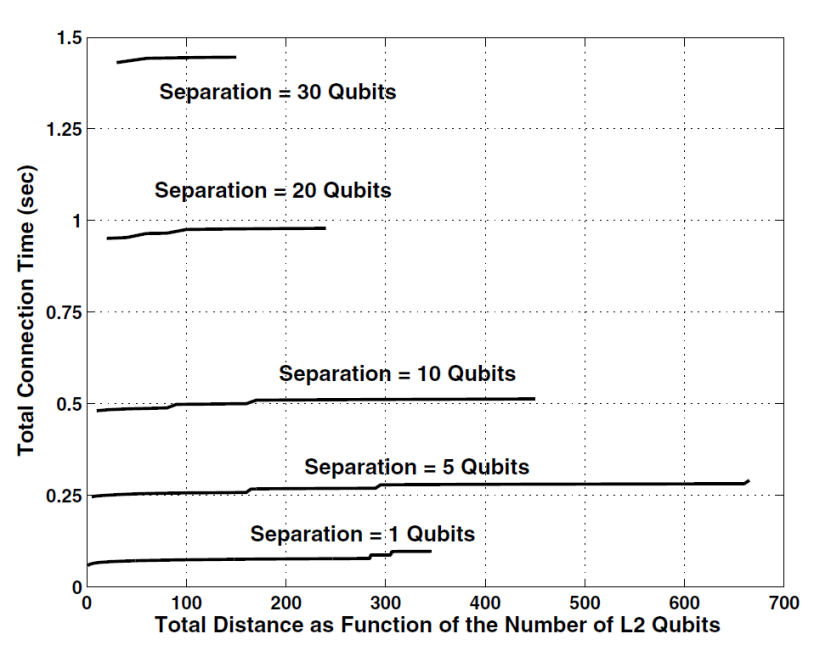
# Communication Channel



# Communication Latency




Time to transmit an L2 qubit (49 physical qubits).

# Shor Factorization: Est. Time and Resources

| | N=128 | N=512 | N=1024 | N=2048 |
|---|---|---|---|---|
| Logical Qubits | 37,971 | 150,771 | 301,251 | 602,259 |
| Toffoli Gates | 63,729 | 397,910 | 964,919 | 2,301,767 |
| Total Gates | 115,033 | 1,016,295 | 3,270,582 | 11,148,214 |
| Area($m^2$) | 0.11 | 0.45 | 0.90 | 1.80 |
| **Time(days)** | 0.9 | 5.5 | 13.4 | 32.1 |

Additional gains in area and performance

- Reduced recursion to decrease area
- Introduce memory hierarchy