QUANTUM ERROR CORRECTION ON FPGA

FILIPPO CORNA AND DAVIDE SALONICO
{filippo.corna, davide.salonico}@mail.polimi.it

POLITECNICO
MILANO 1863

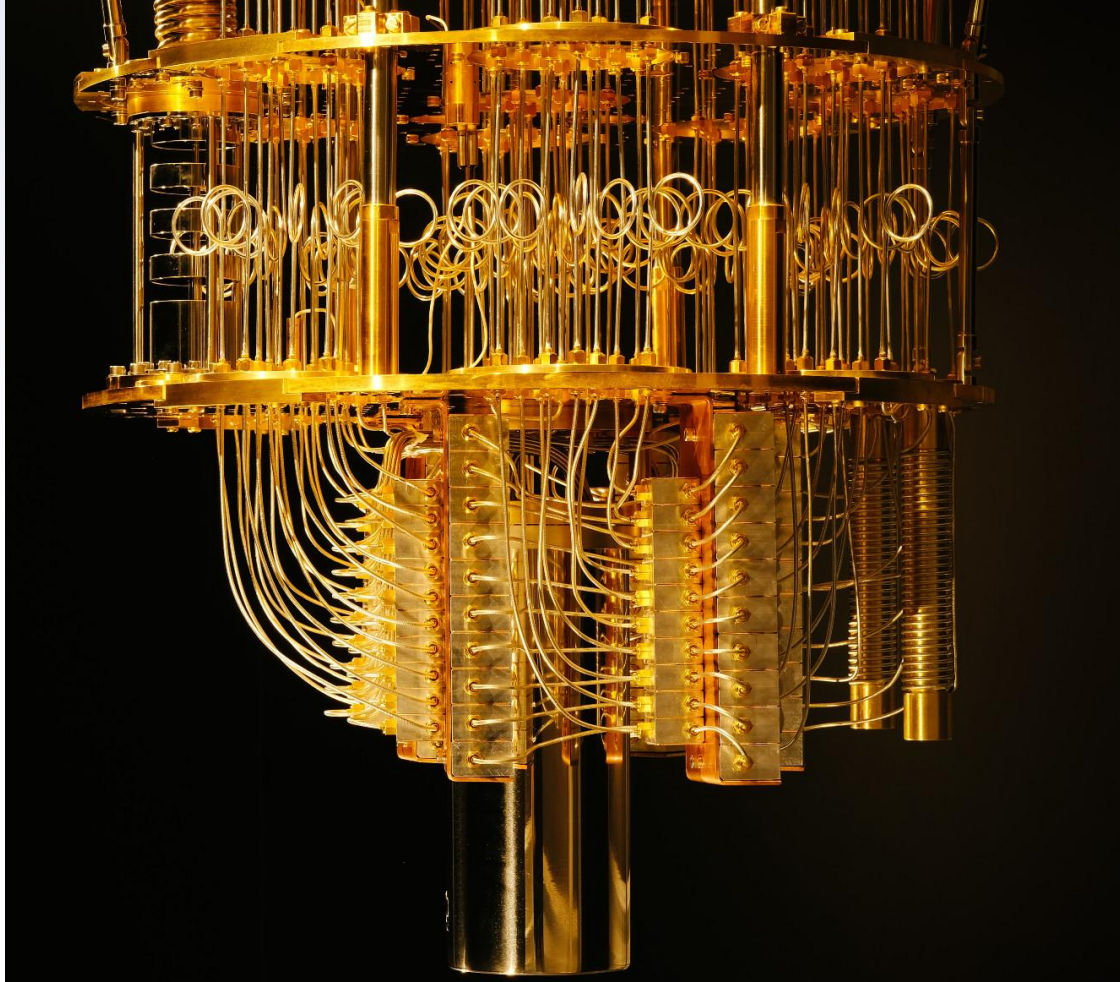
POLITECNICO MILANO 1863
NECST laboratory

18/07/2024 @ DEIB - NECSTLab Meeting Room

Outline

- Introduction to Quantum Computing
- Quantum Error Correction
- Sparse Blossom Algorithm
- Hardware Acceleration
- Experimental Evaluation
- Conclusion

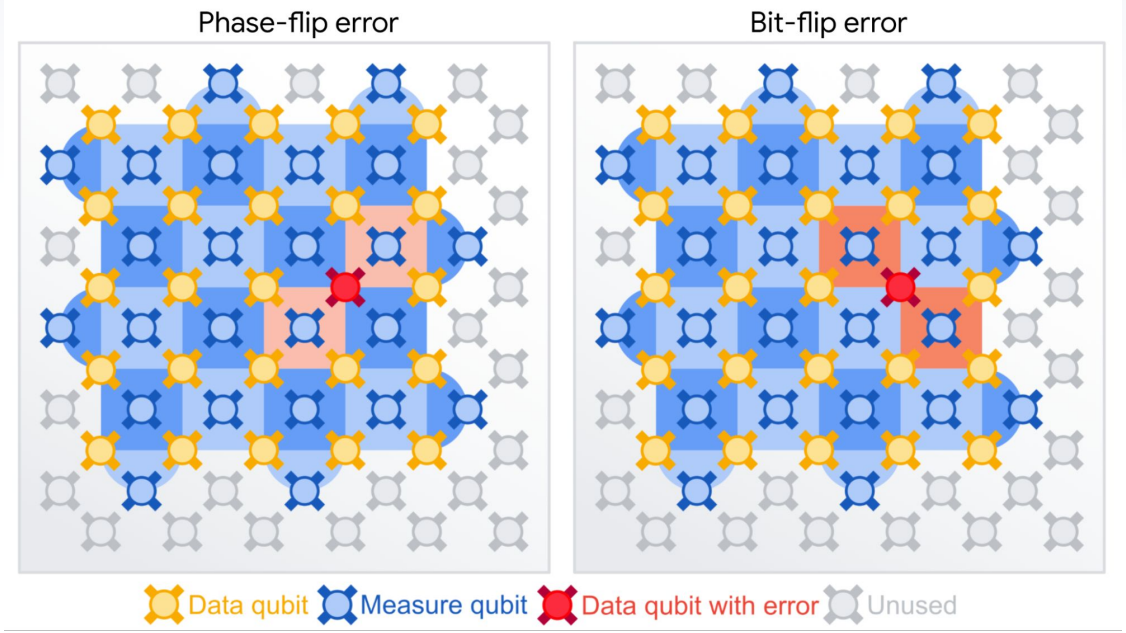
Quantum Computing & its limits



Source: https://www.dtu.dk/english/newsarchive/2023/05/the-quantum-computer-exists-but-is-not-all-that-powerful

Quantum Error Correction

- Encode information in a redundant way
- Detect errors
- Correct errors



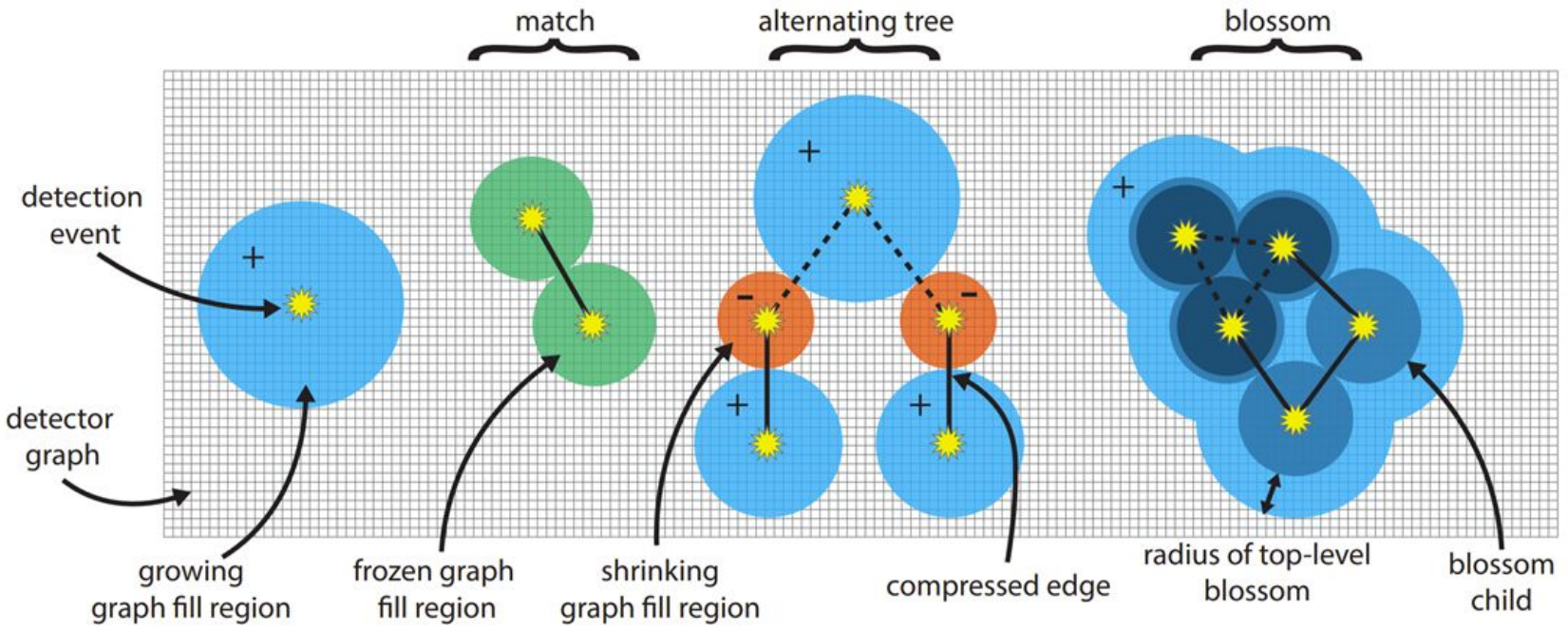
Source: https://physicsworld.com/a/breakthrough-in-quantum-error-correction-could-lead-to-large-scale-quantum-computers/

Fault tolerance

Errors are generated and start to accumulate very quickly

The only way to protect quantum information is constantly correct the newly occurring errors

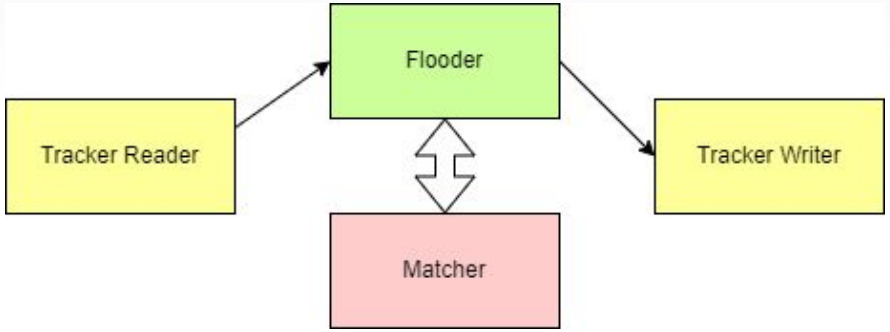
Sparse Blossom Algorithm (SBA)



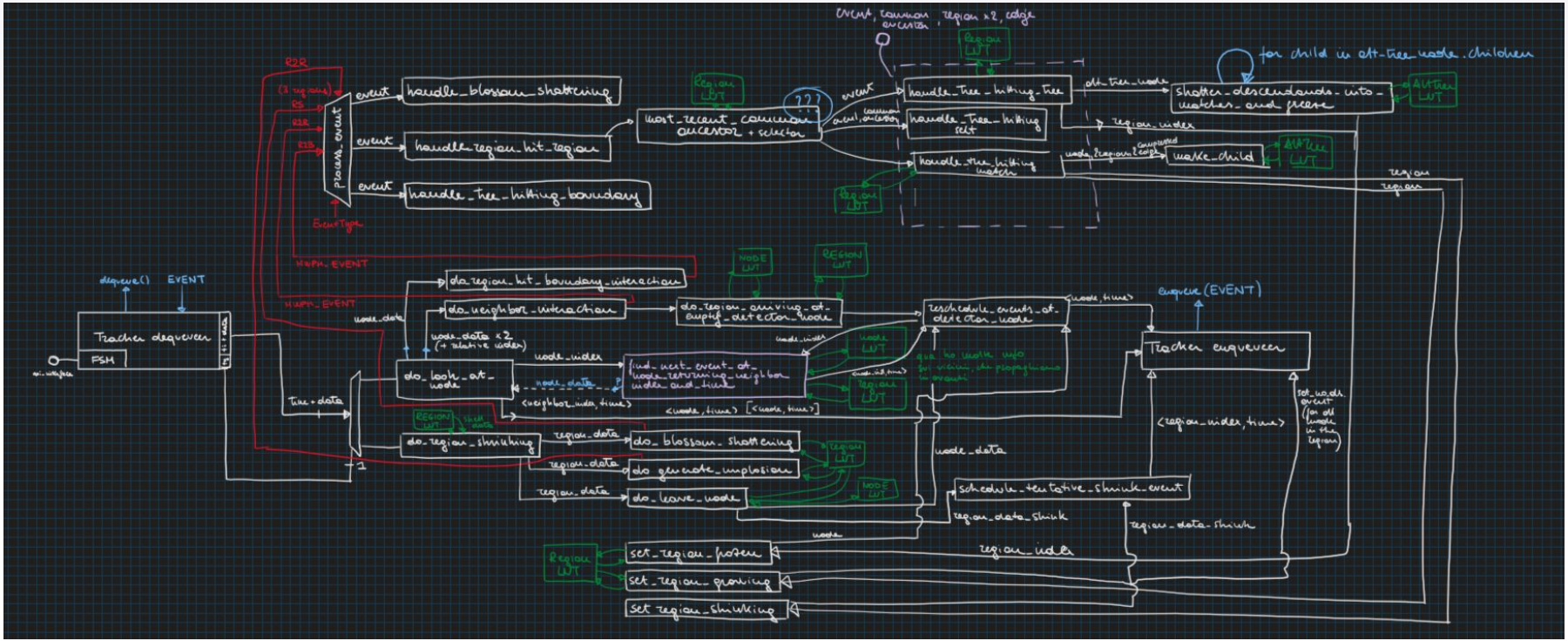
Source: Sparse Blossom: correcting a million errors per core second with minimum-weight matching (Higgott, Gidney)

SBA components...

- **Flooder**: manages nodes-regions interactions
- **Matcher**: manages regions/blossoms interactions
- **Tracker**: keeps track of the events to be processed and/or invalidated



… and a more detailed view

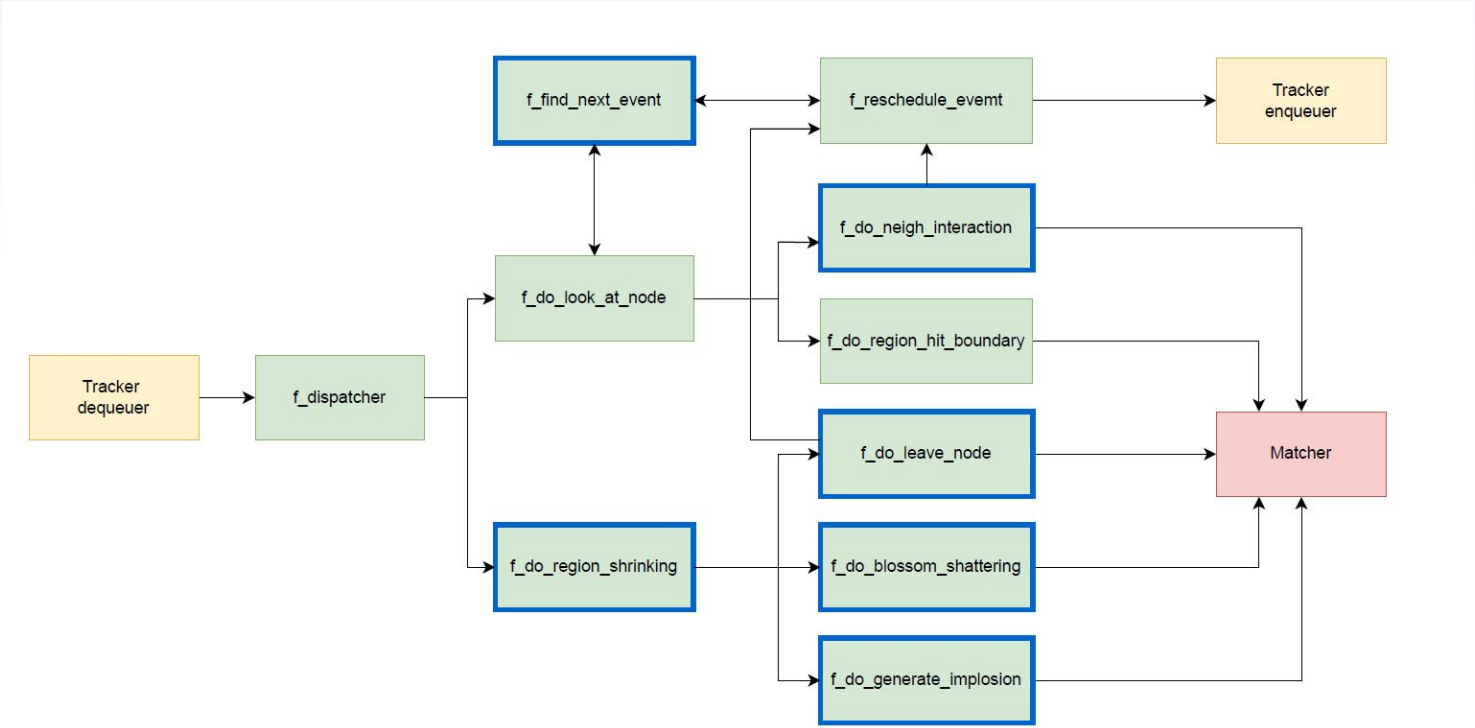


Methodology

Goal: reduce execution time of the Sparse Blossom Algorithm by implementing an hardware FPGA-based version

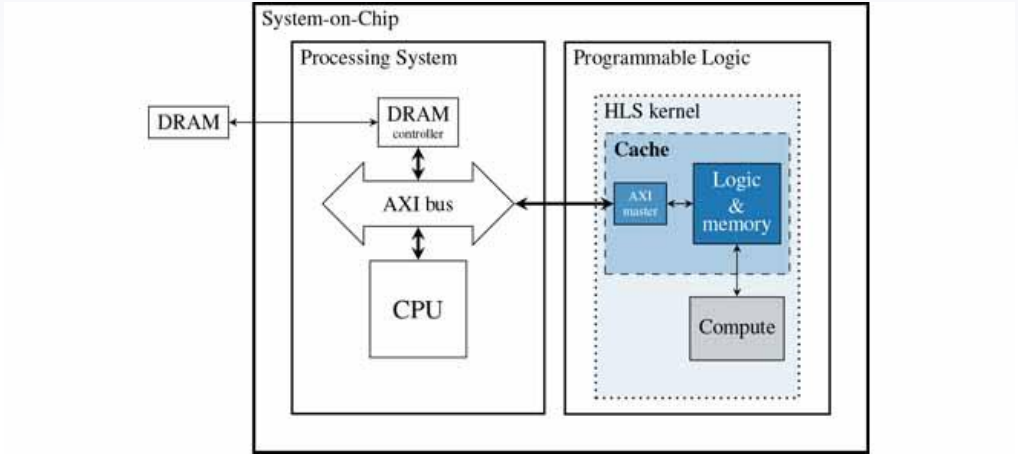
- Estimation of bounds for data structures
  - Hybrid of theoretical and empirical estimations
- Optimization of data type
- Pipelining and Unrolling for subroutines
- Bottom-Up approach

Flooder analysis



DaCH Cache

- HLS level cache
- Dataflow (emulating LCS pattern)
- Two levels
- Multiport
- Highly customizable
- Profiling Functions
- *Problems during synthesis*



Source: Array-Specific Dataflow Caches for High-Level Synthesis of Memory-Intensive Algorithms on FPGAs (Brignone and others)

Experimental Setup

Vitis HLS and Vitis Unified IDE

Intel® Xeon® CPU E5-2680 v2

- @ 2.80GHz
- 396 GB RAM

Alveo U55C High Performance Compute Card

- 16 GB HBM with 460 GB/s bandwidth
- 1,304K LUTs
- 2,607 Registers
- DSP Slices 9,024

Source: https://www.xilinx.com/products/boards-and-kits/alveo/u55c.html

Experimental Results

Correctness

Manually tested on a small graph:

1. Modify PyMatching source code
2. Generate a log of the events happening in the algorithm
3. Checking correctness of results w.r.t. our hardware version

```
(pym_venv) davide.salonico@nags21:~/PyMatching$ /home/users/davide.salonico/pym_venv/bin/python /home/users/davide.
salonico/PyMatching/src/test.py
SET_DESIRED_EVENT NODE
TentativeEvent{.time=27677302, .type=LOOK_AT_NODE, .node=0x1c42598}
MATCHER DO_RHB_INTERACTION
MwpmEvent{.type=REGION_HIT_BOUNDARY, .dat={.region=0x2106840, .edge=CompressedEdge{.obs_mask=0, .loc_from=0x1c42598
, .loc_to=0}}}
TentativeEvent{.time=27677302, .type=LOOK_AT_NODE, .node=0x1c42598}
MATCHER NO_EVENT
SET_DESIRED_EVENT NODE
SET_DESIRED_EVENT NODE
TentativeEvent{.time=16059418, .type=LOOK_AT_NODE, .node=0x1c426a8}
MATCHER DO_NEIGH_INTERACTION
MwpmEvent{.type=REGION_HIT_REGION, .dat={.region1=0x2197630, .region2=0x2106840, .edge=CompressedEdge{.obs_mask=0,
.loc_from=0x1c426a8, .loc_to=0x1c42598}}}
TentativeEvent{.time=16059418, .type=LOOK_AT_NODE, .node=0x1c426a8}
MATCHER NO_EVENT
TentativeEvent{.time=27677302, .type=LOOK_AT_NODE, .node=0x1c42598}
MATCHER NO_EVENT
SET_DESIRED_EVENT NODE
SET_DESIRED_EVENT NODE
TentativeEvent{.time=16059418, .type=LOOK_AT_NODE, .node=0x1c42620}
MATCHER DO_NEIGH_INTERACTION
MwpmEvent{.type=REGION_HIT_REGION, .dat={.region1=0x2197630, .region2=0x2106840, .edge=CompressedEdge{.obs_mask=0,
.loc_from=0x1c42620, .loc_to=0x1c42510}}}
TentativeEvent{.time=16059418, .type=LOOK_AT_NODE, .node=0x1c42620}
MATCHER NO_EVENT
TentativeEvent{.time=32118836, .type=LOOK_AT_NODE, .node=0x1c42510}
MATCHER NO_EVENT
SET_DESIRED_EVENT NODE
```

Conclusions & future works

A meaningful speedup evaluation is not feasible at this stage of development

- Completing the acceleration & speedup evaluation
- Cache configuration space exploration
- RTL cache

THANK YOU FOR YOUR ATTENTION

FILIPPO CORNA AND DAVIDE SALONICO
{filippo.corna, davide.salonico}@mail.polimi.it

POLITECNICO
MILANO 1863

POLITECNICO MILANO 1863
NECST laboratory

18/07/2024 @ DEIB - NECSTLab Meeting Room