Chair III: Database Systems
Chair XXV: Data Science and Engineering
Department of Informatics
Technical University of Munich

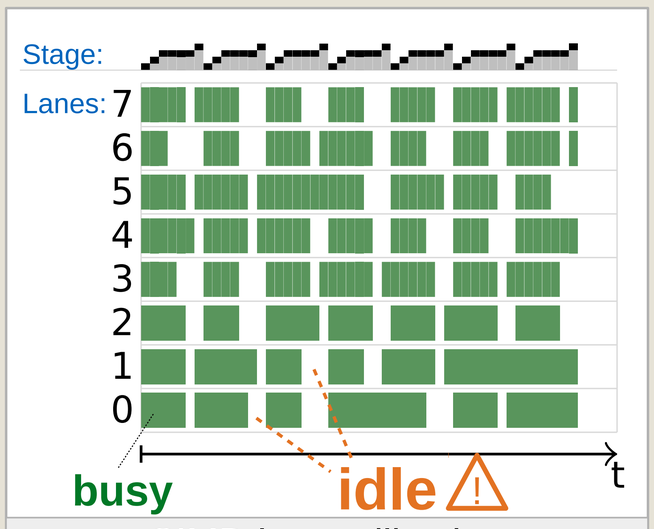
Make the Most out of Your SIMD Investments: Counter Control Flow Divergence in Compiled Query Pipelines

Harald Lang, Andreas Kipf, Linnea Passing, Peter Boncz*, Thomas Neumann, Alfons Kemper

Technical University of Munich
Munich, Germany
first.last@in.tum.de

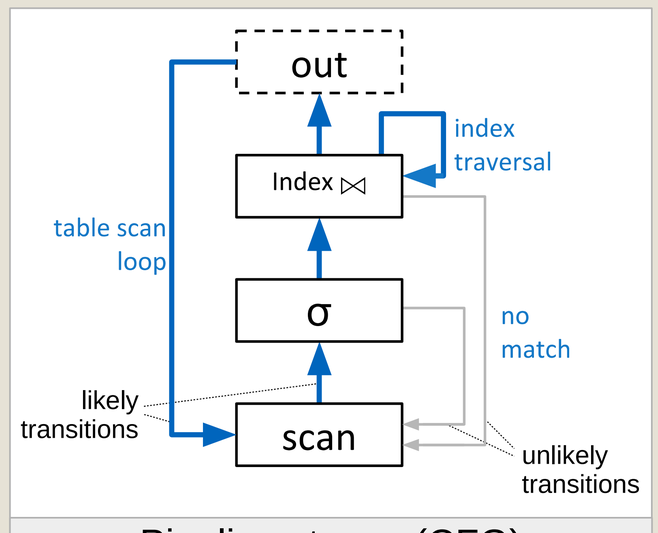
* Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
boncz@cwi.nl

Challenge



SIMD lane utilization

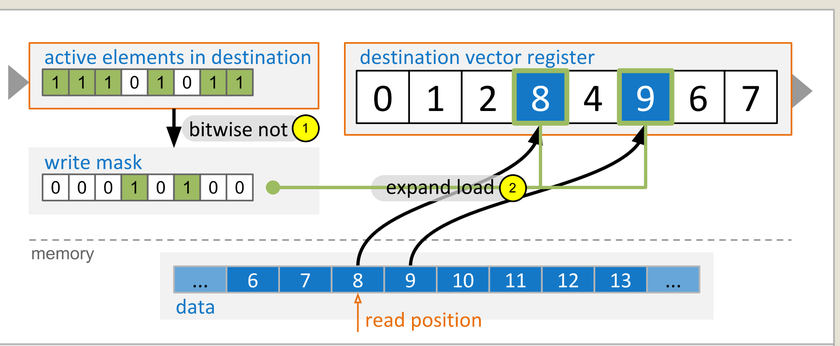
- SIMD lanes may (temporarily) become inactive due to control flows divergence.
- The resulting underutilization of vector-processing units causes performance degradations.
- We propose efficient algorithms and strategies to fill these gaps.



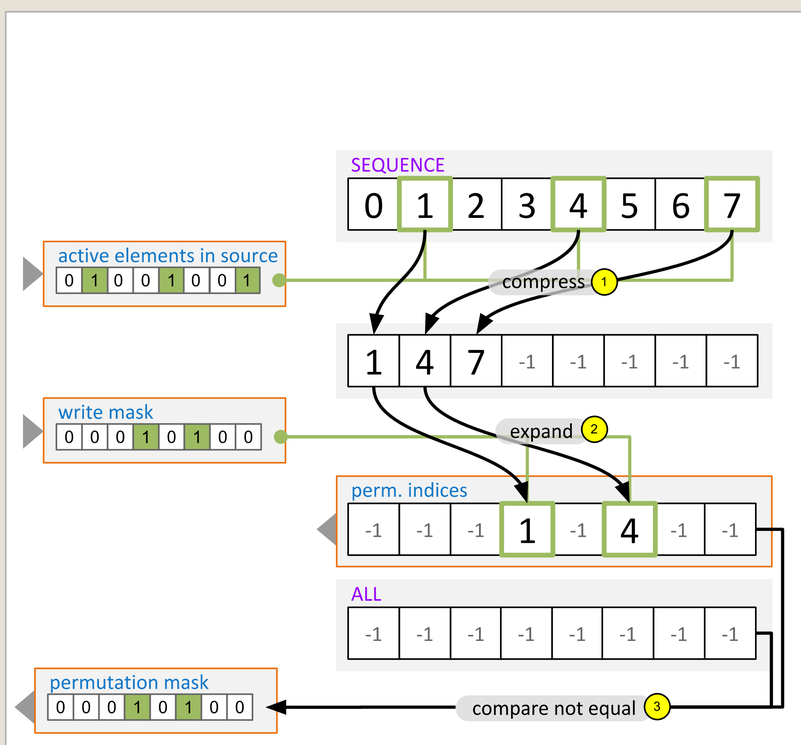
Pipeline stages (CFG)

Refill algorithms

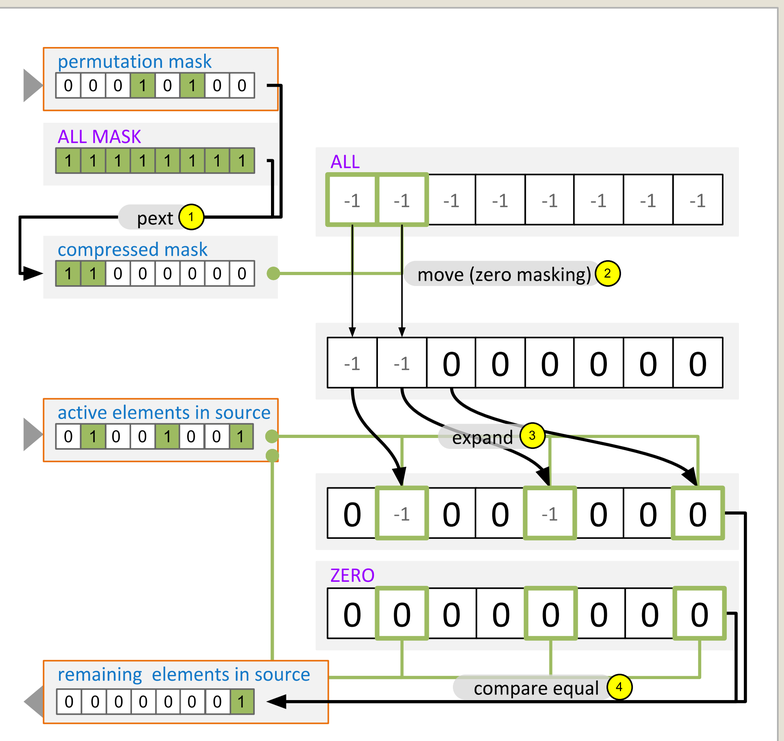
- Copy new elements to idle SIMD lanes in a destination register, without modifying active lanes.
- Basic building block to counter underutilization.
- Many different flavors, e.g.
  - copy from memory to vector registers
  - copy between vector registers.
- Enabled by AVX-512 instructions (i.e., `compress`, `expand`)



Moving elements from memory to register is natively supported in AVX-512.



(a)

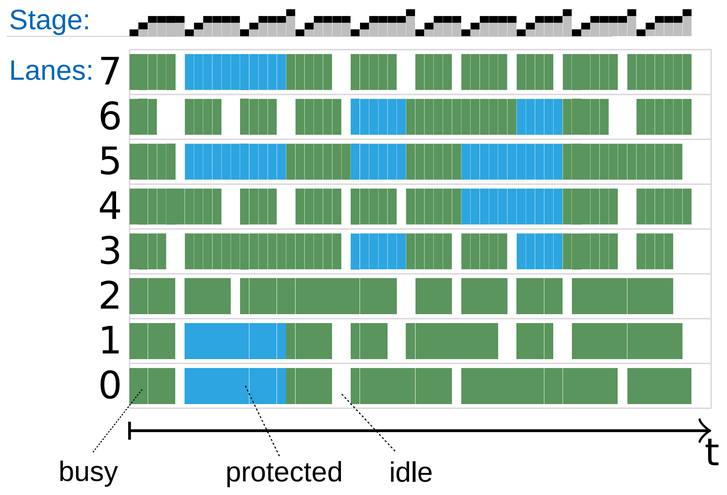


(b)

Selectively moving elements between registers is done using the **`permute`** instruction. The tricky part is to determine the permutation indices (a) and to mask the moved elements in the source register accordingly (b).

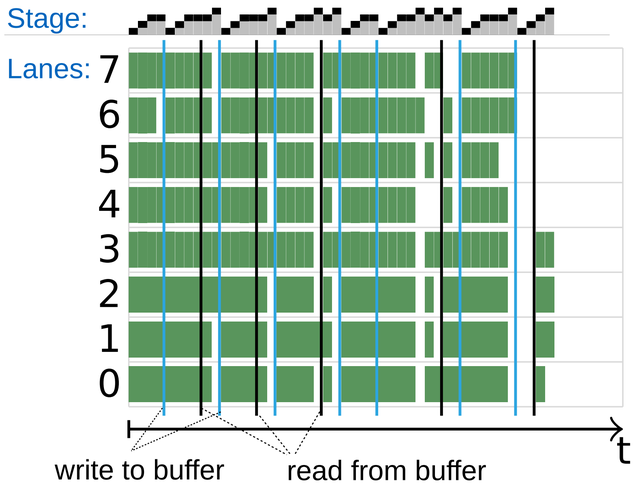
Refill strategies

Integration of refill algorithms with compiled query pipelines:



Partial Consume

- Refill at pipeline source (e.g., table scan)
- In case of underutilization, operators return the control flow
- Lanes with active elements are protected with very little overhead

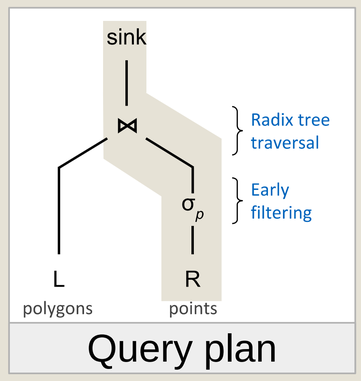


Consume Everything

- Operators allocate additional registers which are used as buffers.
- In case of underutilization, operators either refill idle lanes with buffered elements, or flush active lanes to the buffer and return the control flow.

Results

Approximate geospatial join



Query plan

System	Performance **Baseline** (AVX-512)
Knights Landing Phi 7210	**3559 Mtps**
Skylake-X i9-7900X	**910 Mtps**

(Manhattan polygons, 15m precision)

Performance improvements

System	Improvement
Knights Landing	**+ 20 %**
Skylake-X	**+ 35 %**

Acknowledgements:
This work has been partially supported by the German Federal Ministry of Education and Research (BMBF) grant 01IS12057 (FASTDATA and MIRIN), and the DFG projects NE1677/1-2 and KE401/22. This work is further part of the TUM Living Lab Connected Mobility (TUM LLCM) project and has been funded by the Bavarian Ministry of Economic Affairs, Energy and Technology (StMWi) through the Center Digitisation.Bavaria, an initiative of the Bavarian State Government.