Q*cert

A Platform for Specifying and Verifying Query Compilers

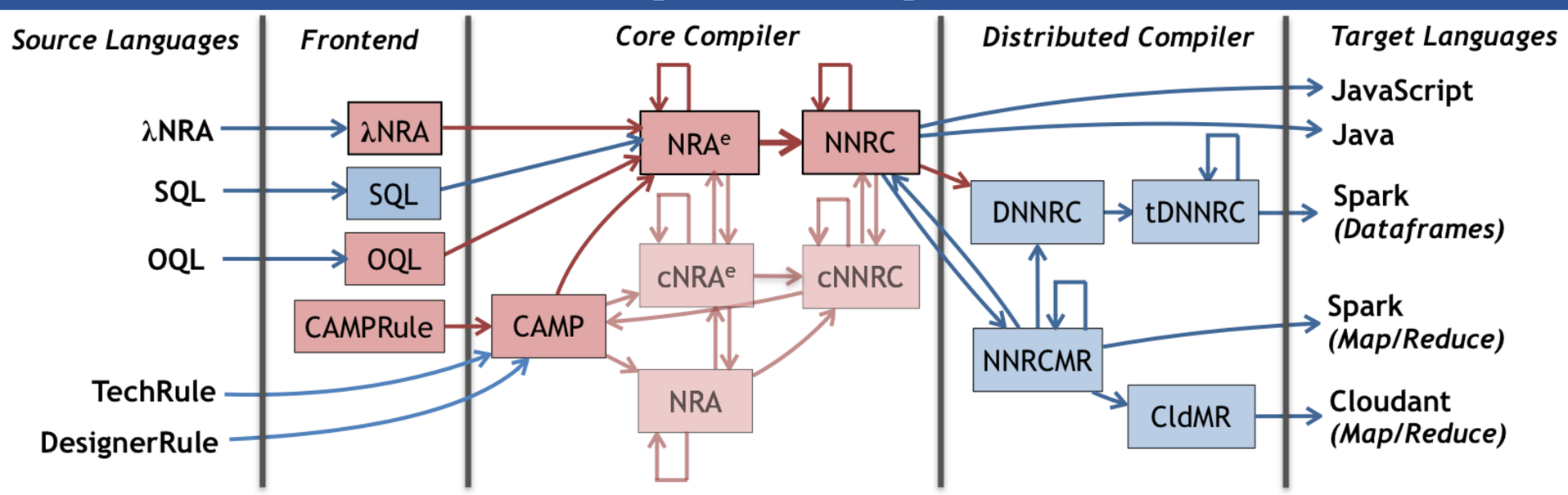
J. Auerbach, M. Hirzel, L. Mandel, A. Shinnar, J. Siméon
IBM Research

Goals:
Verified Query Compiler
Multiple Sources & Targets
Extensibility

Applications:
New Languages (e.g., DSLs)
Designing New Optimizations
Education Platform

Approach:
Written in Coq (Proof Assistant)
Mechanized Correctness Proof
Compiler Extraction

Compilation Pipeline



SQL: Structured Query Language
OQL: Object Query Language
λNRA: NRA with Lambdas
CAMPRule: Rule Macros for CAMP
TechRule: ODM technical rules
DesignerRule: ODM designer rules

NRA: Nested Relational Algebra
NRA^e: NRA with Environments
cNRA^e: Core NRA^e
NNRC: Named Nested Relational Calculus
cNNRC: Core NNRC
CAMP: Calculus of Aggregating Matching Patterns

DNNRC: Distributed NNRC
tDNNRC: Typed DNNRC
NNRCMR: NNRC + Map/Reduce
CldMR: NNRC + Cloudant Map/Reduce

with proof
without proof

Algebraic Equivalence

```
Lemma select_union_distr q0 q1 q2 :
    σ⟨ q0 ⟩(q1 ∪ q2) ≡ σ⟨ q0 ⟩(q1) ∪ σ⟨ q0 ⟩(q2).
Proof.
  … (* proof omitted *)
Qed.
```

Functional Rewrite

```
Definition select_union_distr_fun q :=
  match q with
  | NRAEnvSelect q0 (NRAEnvBinop AUnion q1 q2) =>
      NRAEnvBinop AUnion
        (NRAEnvSelect q0 q1) (NRAEnvSelect q0 q2)
  | _ => q
  end.
```

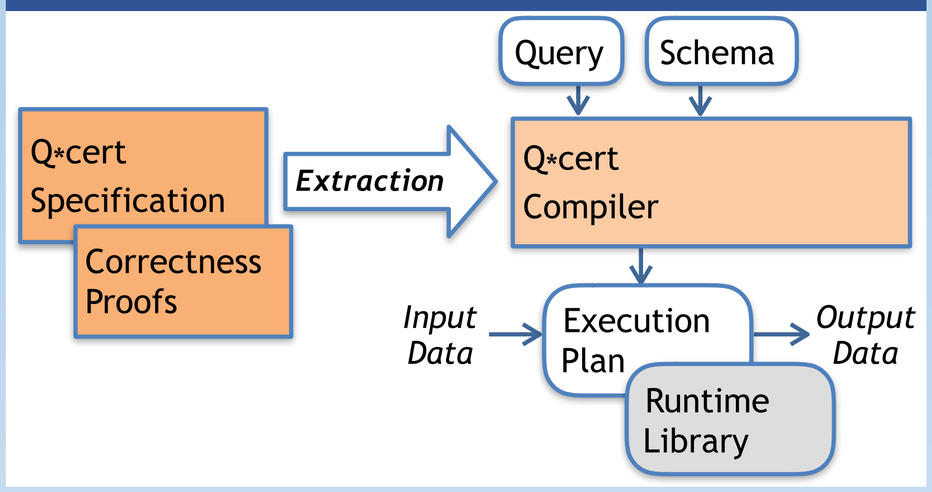
Correctness Proof

```
Property select_union_distr_fun_correctness q0 q1 q2 :
 select_union_distr_fun q ≡ q.
Proof.
 Hint Rewrite select_union_distr : envmap_eqs.
 prove_correctness q.
Qed.
```

Other Features

Configurable Optimizer
Nested Data Model with Objects
JSON Support
Aggregate Queries (supports TPC-H)
Type Checking
Foreign Types and Operations

Compiler Extraction



The Coq Proof Assistant

https://querycert.github.io