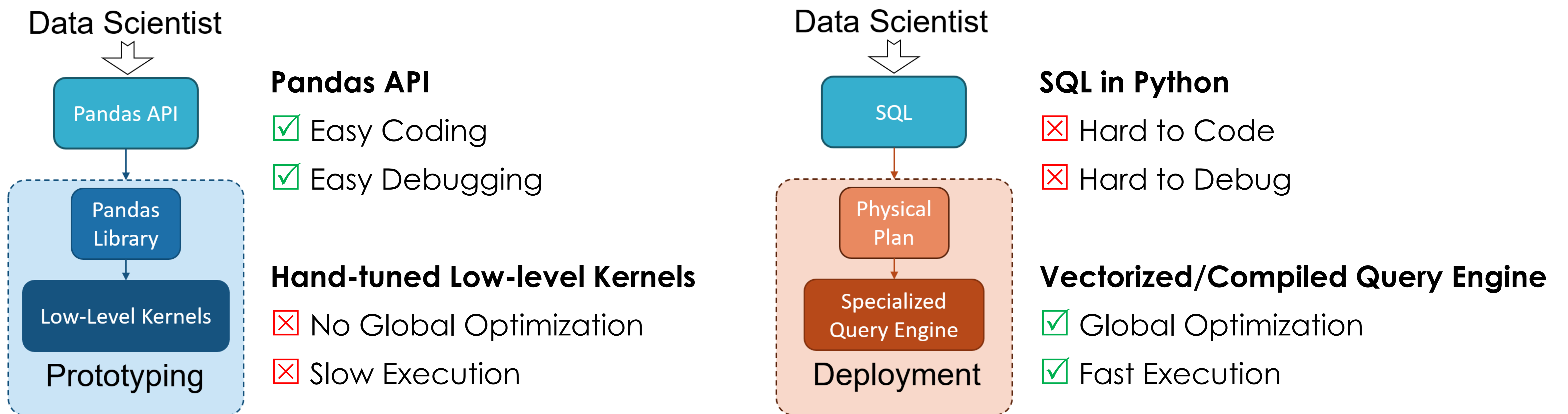


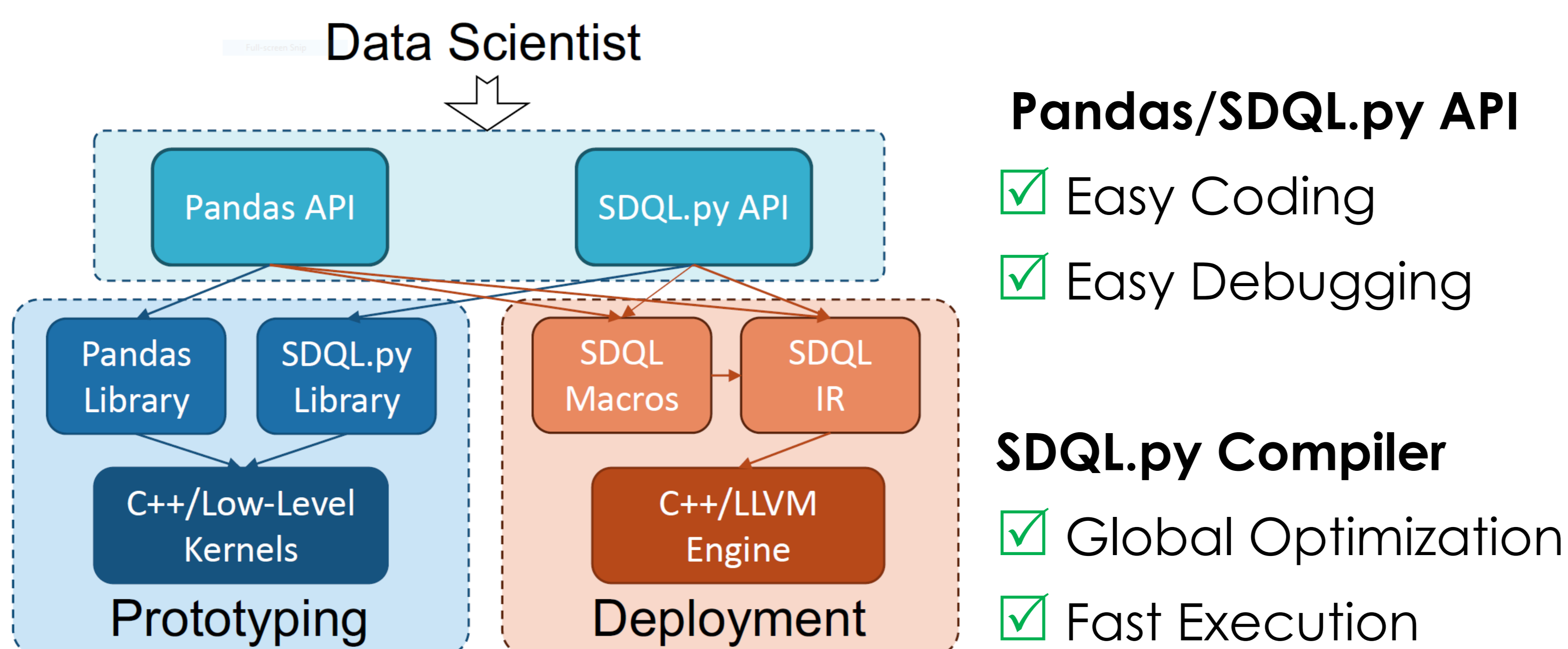
SDQL.py: An Efficient Query Engine in Python

Hesam Shahrokhi, Amir Shaikhha
School of Informatics, University of Edinburgh

Query Processing in Python: Current State of Affairs



Query Processing by SDQL.py



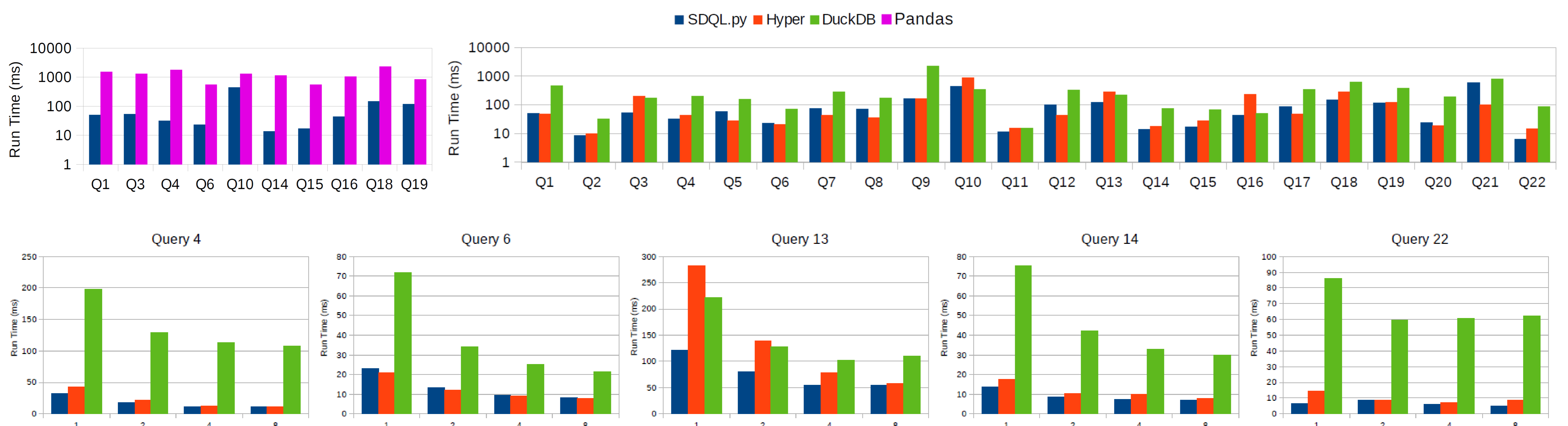
SDQL.py to SDQL IR Example

```
@sdql_compile({"R": R_type, "S": S_type}, exec_mode)
def query(R, S):
    R_indx = R.sum(lambda r: {r.k: r.c} if r.a=9 else {})
    return S.sum(lambda s:
        {record({"c": R_indx[s.k].c, "g": s.g}): True})
        if (s.e=="pass" and R_indx[s.k]) else {})
```

```
let R_indx = sum(r in R) if r.a=9 then {r.k -> r.c}
sum(s in S)
if s.e=="pass" and R_indx[s.k] then
    {<c: R_indx[s.k].c, g: s.g -> true}
```

Experimental Results

- Intel Core i5
- 16GB DDR4 RAM
- Ubuntu 20.04.3
- G++ 9.4.0, Python 3.8.10, and Pandas 1.4.1
- TPC-H Queries with Scaling Factor of 1



¹ Hesam Shahrokhi, Amir Shaikhha, "Building an Embedded Query Engine in Python", Under Review.

² Amir Shaikhha, Mathieu Huot, Jaclyn Smith, Dan Olteanu, "Functional collection programming with semi-ring dictionaries", OOPSLA'22.