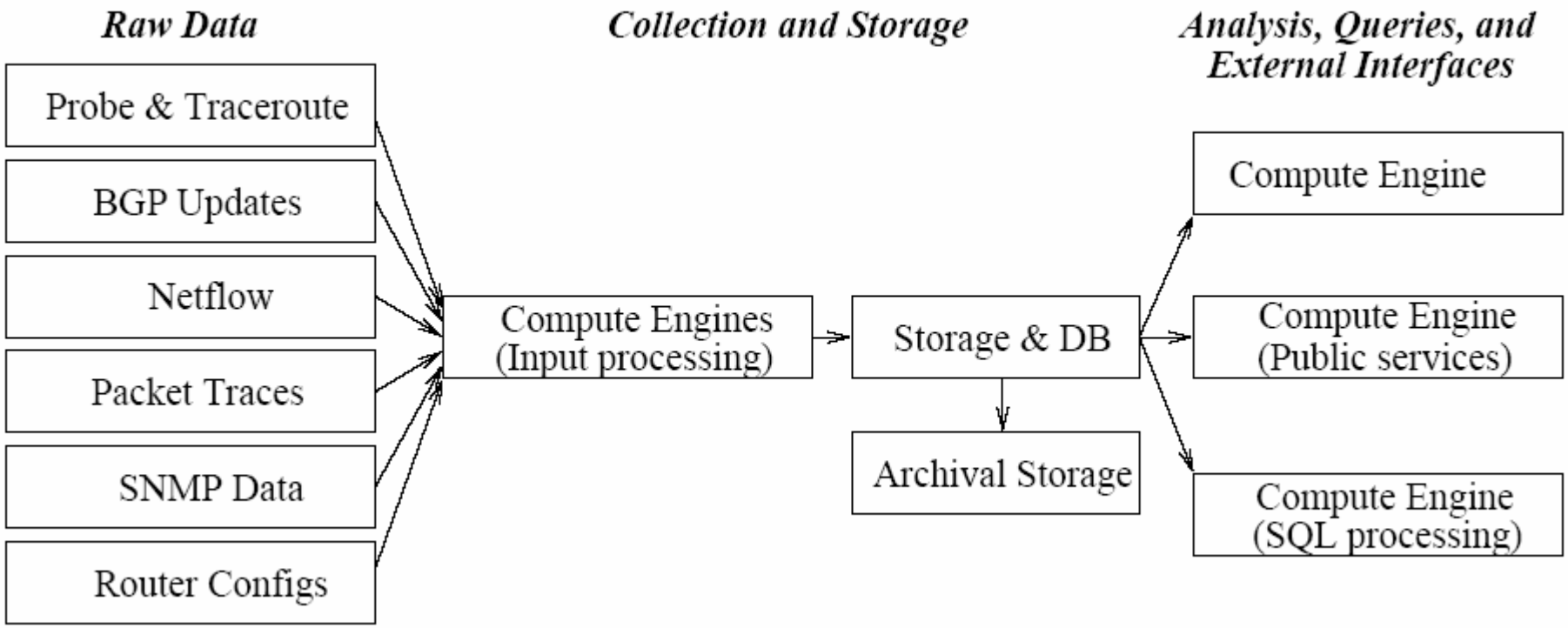
The Datapository

Dave Andersen, CMU
Nick Feamster, Georgia Tech

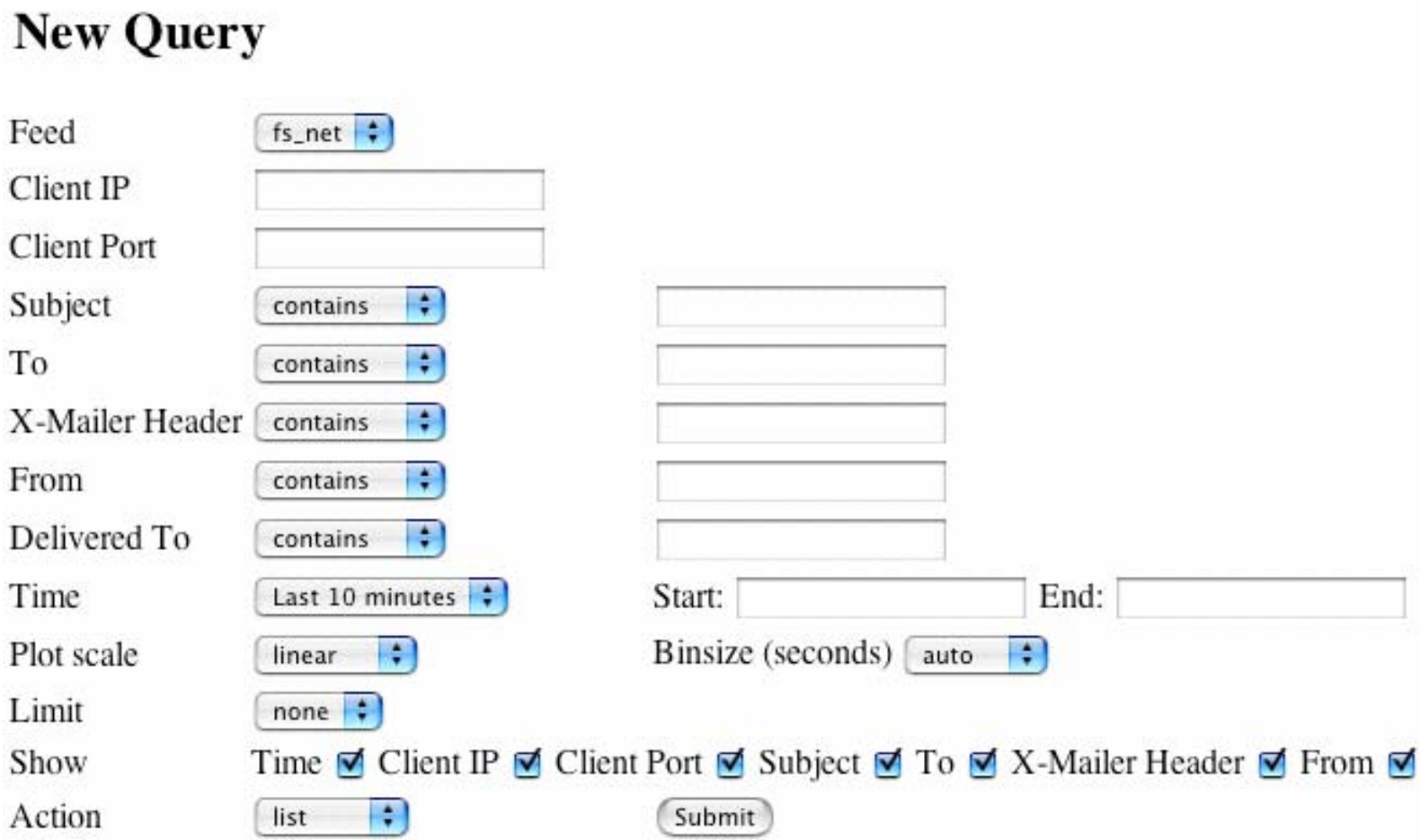
http://www.datapository.net/

Datapository Architecture



- **Separate:** collection, storage, analysis
- **Collection:** abstract type, format, and access method

A Handy Web Interface



Do Spammers Hijack BGP Routes?

Theory:

1. announce BGP route for mail server
2. Send lots of spam
3. Withdraw route, becoming invisible

Reality? Let’s check...

Selecting Spammers

```
SELECT * from spam
WHERE spam.time > ... AND spam.time < ...
    AND EXISTS
     (SELECT * from bgp
      WHERE bgp.time > spam.time - 20
          AND bgp.time < spam.time + 3600
          AND bgp.prefix = (spam.client_ip &
                    ((~0) << (32 - bgp.mask)))
```

Export Formats

- Web Interface
- XMLRPC
  - Text-based output
  - Programmatic interface
  - Output to Matlab
- (Brief demo)