Result Provenance in Named Function Networking

Claudio Marxer · Christian Tschudin

University of Basel, Switzerland

ACM ICN 2020, Montréal / virtual format · September 30, 2020

A Primer on Named Function Networking

– ICN/NDN with *named data* and *named functions*

```
/data/alice /data/bob /func/wordCount /func/maximum
```

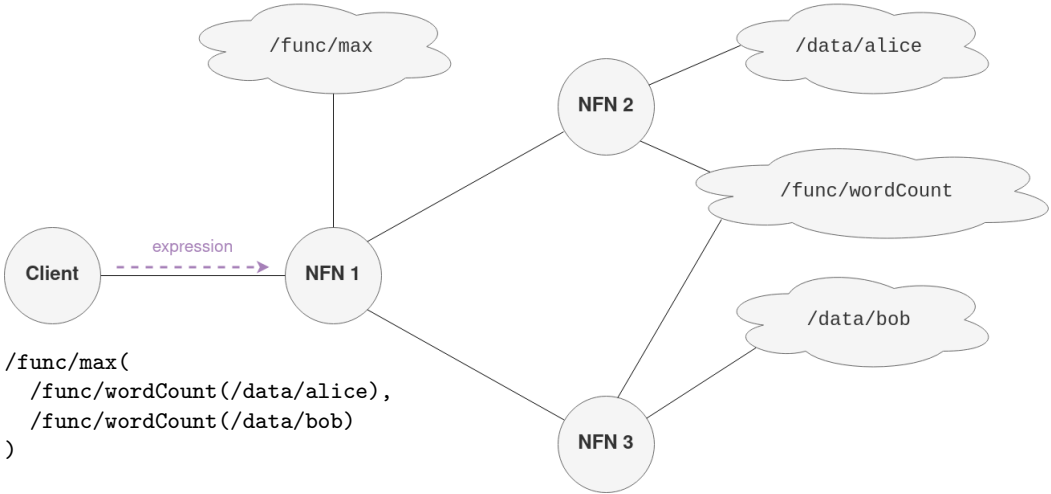
– Computation expressions: applications of named functions on named data

```
/func/maximum( /func/wordCount(/data/alice), /func/wordCount(/data/bob) )
```

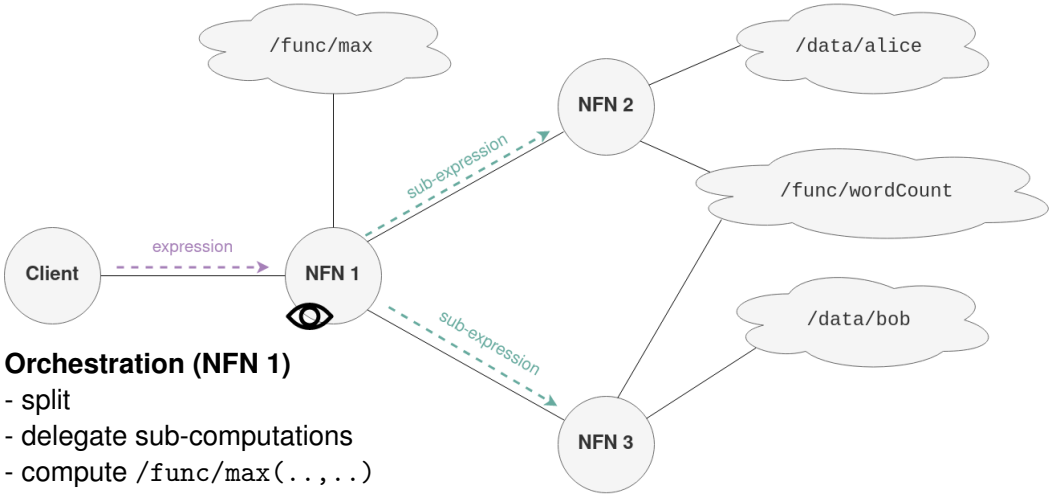
– In-network expression reduction (NFN-capable nodes)

Evaluation: computing result of function applications
Orchestration: where to place which (sub-) computation?

Example: NFN Orchestration + Evaluation



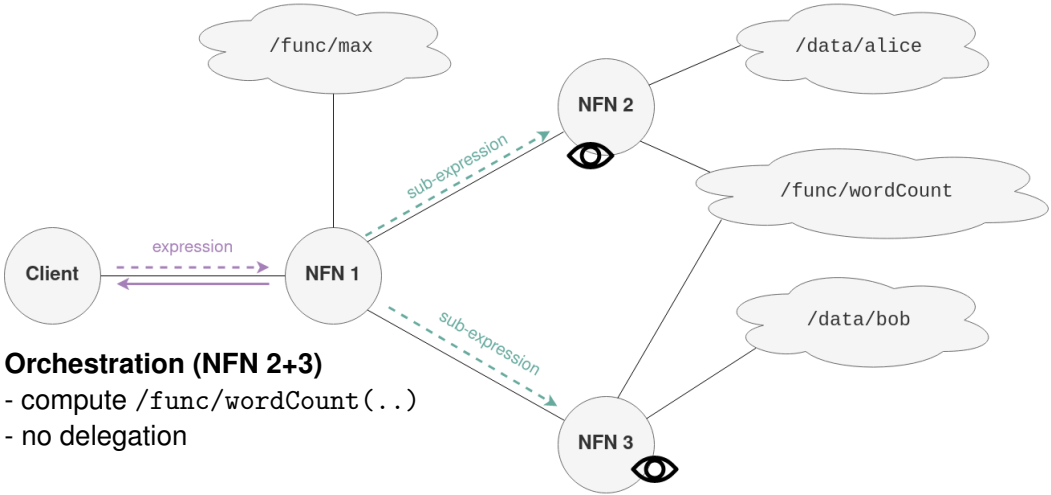
Example: NFN Orchestration + Evaluation



Orchestration (NFN 1)

- split
- delegate sub-computations
- compute `/func/max(..,..)`

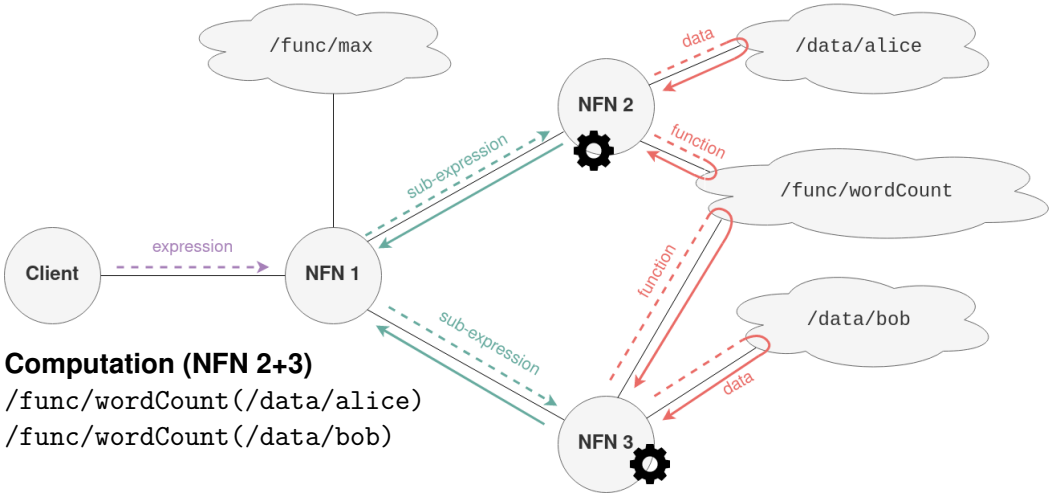
Example: NFN Orchestration + Evaluation



Orchestration (NFN 2+3)

- compute `/func/wordCount(..)`
- no delegation

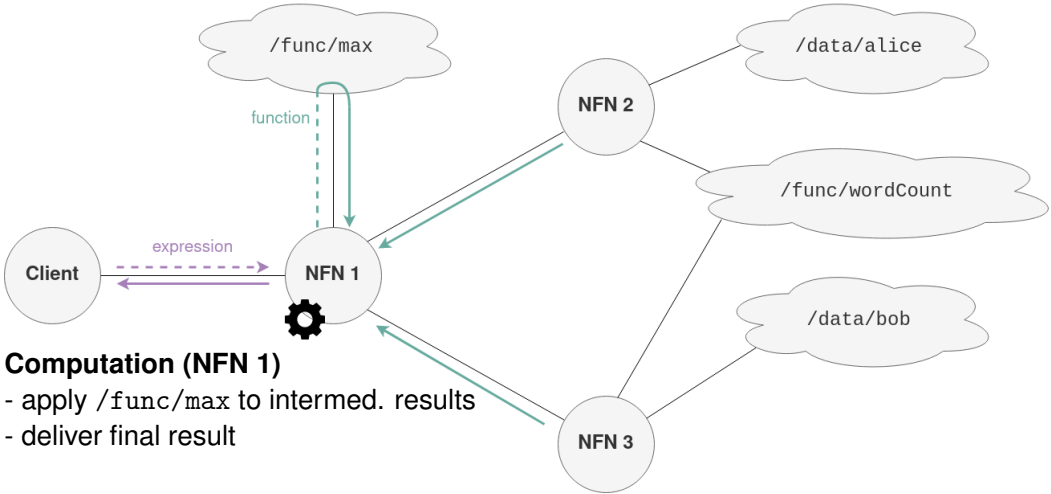
Example: NFN Orchestration + Evaluation



Computation (NFN 2+3)

`/func/wordCount(/data/alice)`
`/func/wordCount(/data/bob)`

Example: NFN Orchestration + Evaluation



Computation (NFN 1)

- apply `/func/max` to intermed. results
- deliver final result

- NFN mindset
- Security Challenge: Result Correctness
- Approach: Provenance Transparency
- Meta-Data: Provenance Records
- Provenance-Based Result Verification
- Ongoing and Future Work
- Conclusion

Result Correctness

– Good news: Convenient computation service for applications

– Bad news: NFN as a whole must be trusted that...
  - Evaluation *rules* are followed
  - Evaluation based on specified *data*

  $\rightarrow$ NFN result correctness is subject to extensive trust

– Goal of this work: Relaxed trust assumptions

– Approach: Log “genesis” of results in *provenance records*
  - Make involved computing entities (CE) traceable
  - Clients assess their trustworthiness

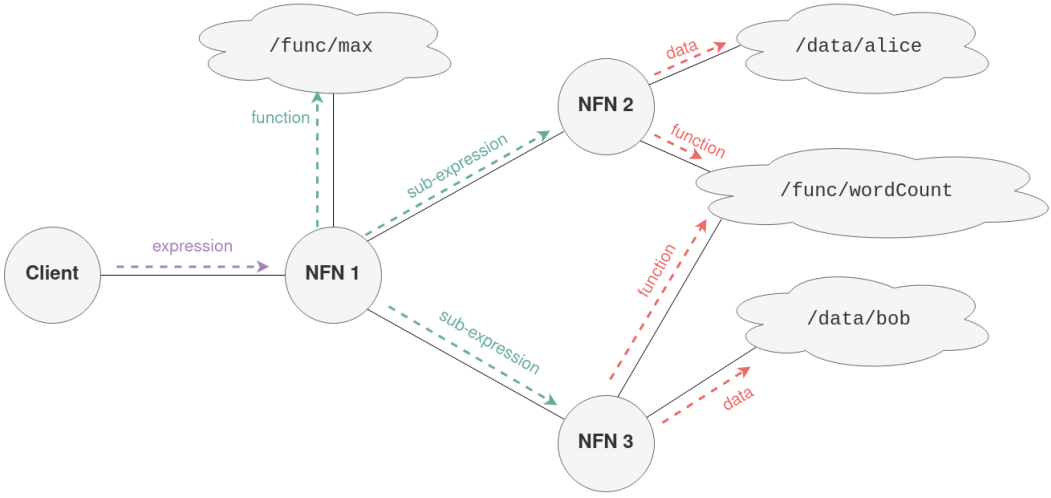
Provenance of Results in NFN

- Provenance meta-data in general[1]: DAG capturing a) *involved elements* (data, processes, hw/sw environment,..), and b) their *relationships*.

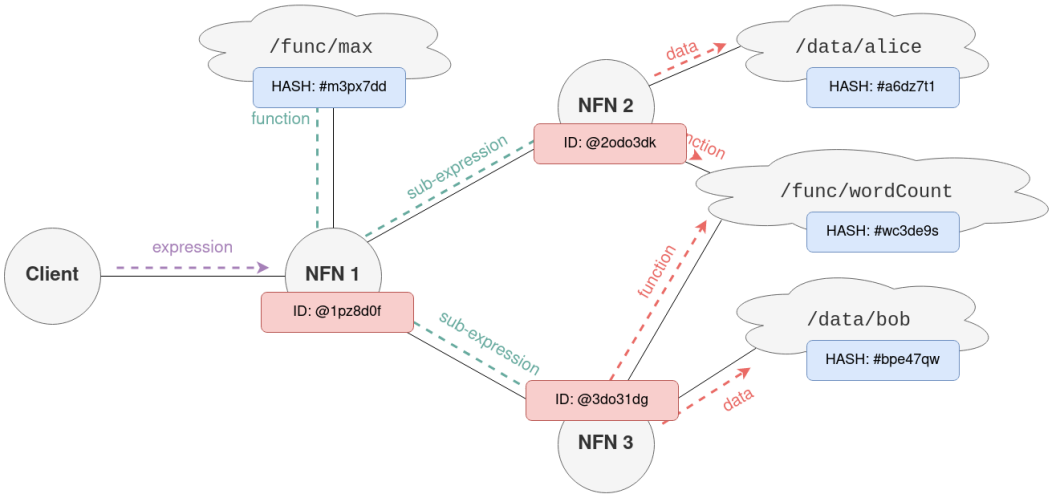
- *Provenance Records (PR)* in NFN: Capture for each computation step:
  - *a)* Identity of CE (public key)
  - *b)* Signatures and PRs of all inputs (data+function)
  - *c)* `hmac( result )`
  - *d)* `hmac( a + b + c + expression )`

[1] L. Carata et al. 2014. *A primer on provenance*. Communications of the ACM 57.

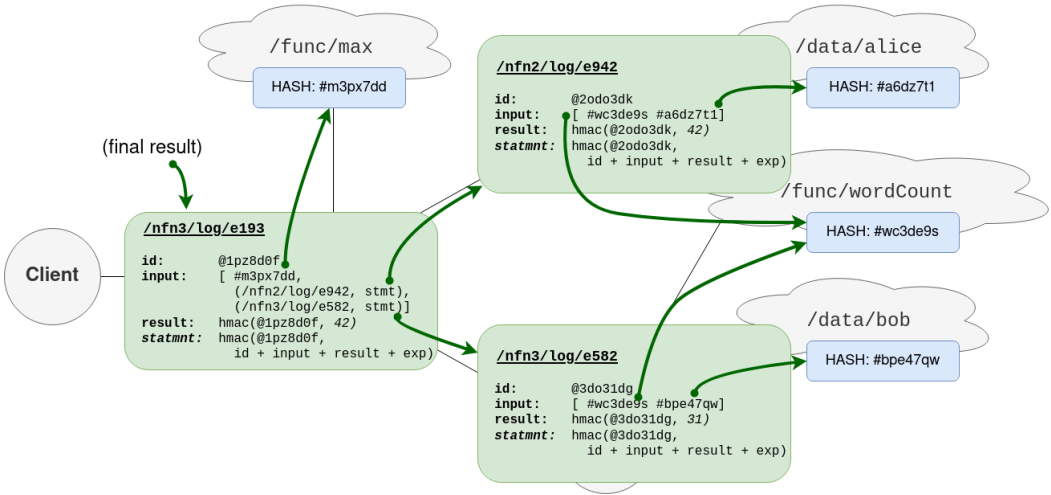
Example Revisited



Example Revisited



Example Revisited



Provenance-Based Result Verification

Input:

- PRs of all (sub-) computations
- list of trusted CEs

Steps:

- All involved CEs are trusted? *(False → result untrusted)*
- All `statement`-hmacs in all PRs are correct? *(False → forged or tampered PR)*
- `result`-hmac of final result correct? *(False → forged or tampered result)*

If successful: Final result is ...

- assumed to be correct under given trust assumptions
- authentic
- of integrity

Next Steps

Establishment of trust in CEs

- State: Predefined list of trusted CEs
- Ongoing: Reputation system
  - Clients exchange CE's reputation
  - Re-evaluation at random
  - PRs as not deniable proofs
  - Related: semantic web & dweb
- Future: Third-party certification

User-Constrained Orchestration

- Issue: Client has no further options if network delivers an untrusted result
- Ongoing: Clients proactively constrain NFN's orchestration (i.e. exclude untrusted CEs)

Next Steps

Availability of Provenance Records

- Implementation State:
  - PR in NDN's signature field
  - Tampering-resistant append-only log (by CEs)
- Future:
  - Issue: Incentive to not deliver disadvantageous logs
  - Replication (e.g. clients, TTP)

Faulty Primary Data

- Issue: Results derived from faulty primary data are faulty as well
- Future (NDN): Convention to flag authentic but faulty data (e.g. due to broken sensor)
- Future (NFN): Consideration in NFN result verification

Conclusion

- **Context:** Services in (recursive) *read-process-republish mode* (e.g. NFN)
- **Challenge:** Result correctness & relaxation of trust
- **Approach:** Transparent provenance & provenance-based result verification
- **Future:** Trust in CEs, User-Constrained Orchestration, Availability of PRs