VC3: Trustworthy Data Analytics in the Cloud using SGX

Felix Schuster∗, Manuel Costa, C´edric Fournet, Christos Gkantsidis Marcus Peinado, Gloria Mainar-Ruiz, Mark Russinovich

Microsoft Research

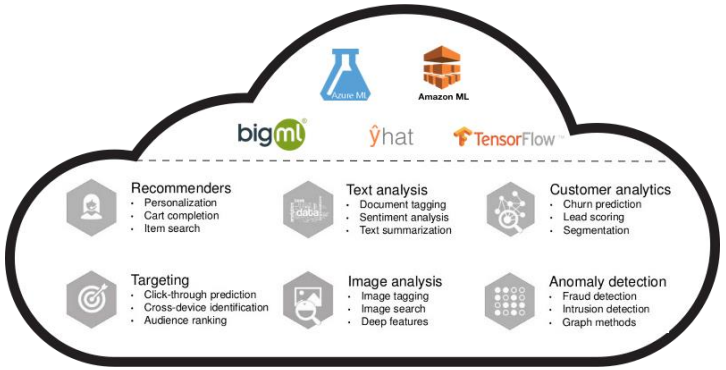
Presenter: Lun Wang

The New Role of Data

Cloud Service

Azure

amazon web services™



Alibaba Cloud

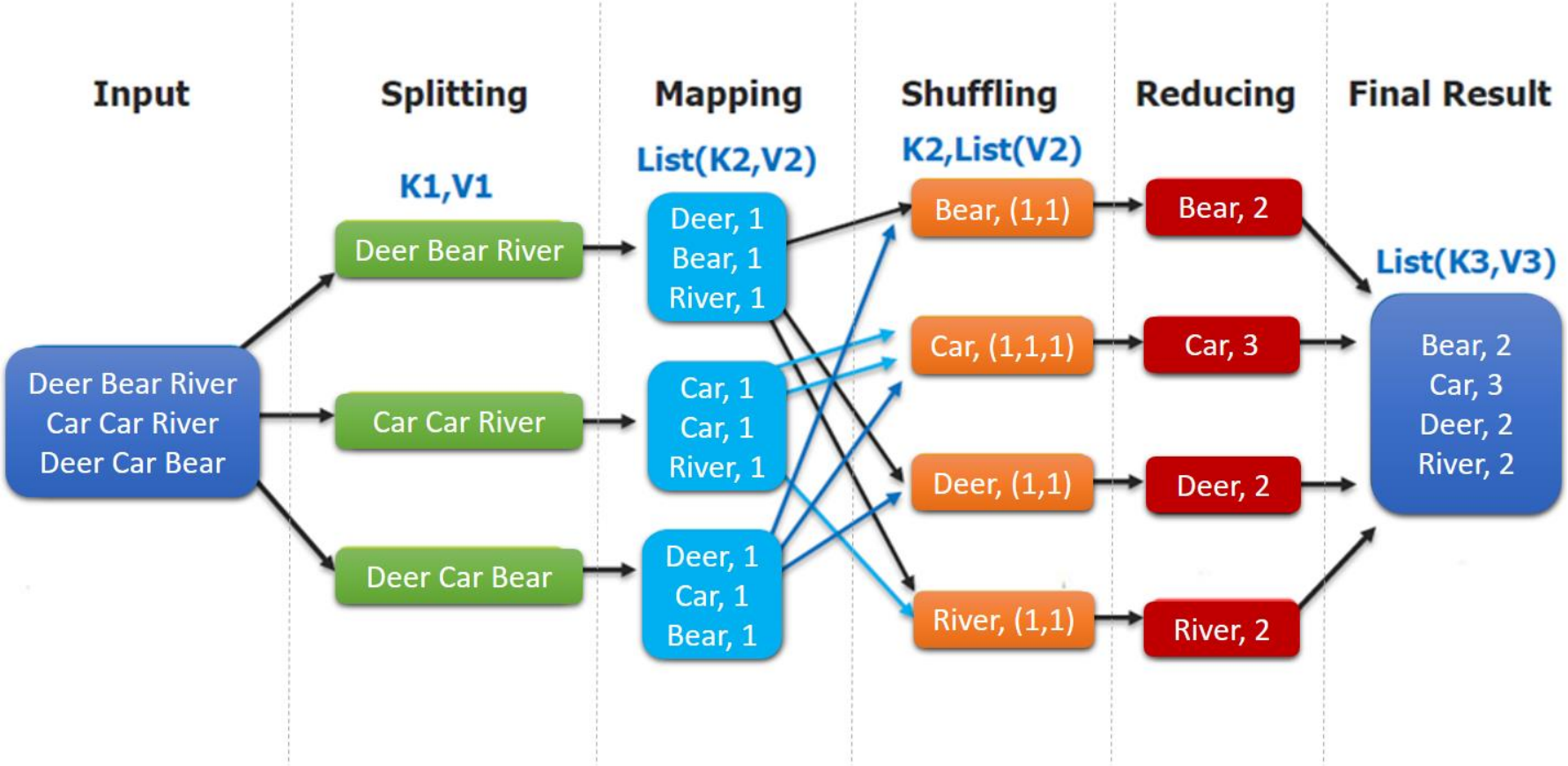
Google™

Trust the Service Provider?

- Facebook Data Privacy Scandal
- Uber Data Breach
- Equifax Data Breach
- Google Is Found to Collect Android Location Data

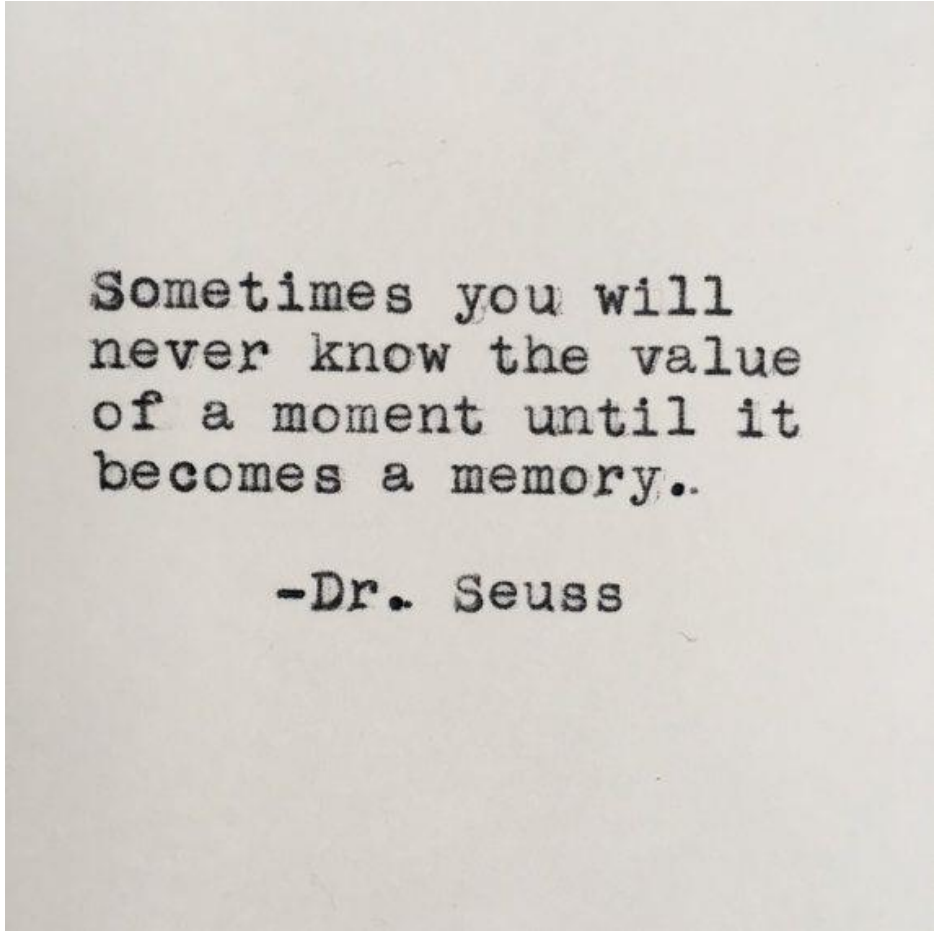
Cloud + TEE!

Map Reduce Model



MapReduce Word Count Process

Quote Enclave



- Quote:
  - Signature
  - SK can only be get by trusted member
  - PK is published by some corporation
- SGX Quote Enclave $\boxed{\mathsf{ESig}_{SGX}[C]\{text\}}$
  - PK published by Intel
  - “I am on a genuine SGX!”
- Cloud Quote Enclave $\boxed{\mathsf{ESig}_{Cloud}[C]\{text\}.}$
  - PK published by Cloud Provider
  - “I am on the cloud!”

Authenticated Encryption with Associated Data

$$\mathsf{Enc}_k[ad]\{text\}$$

Associated Data

Data

Integrity

Confidentiality

Authentication

Threat Model

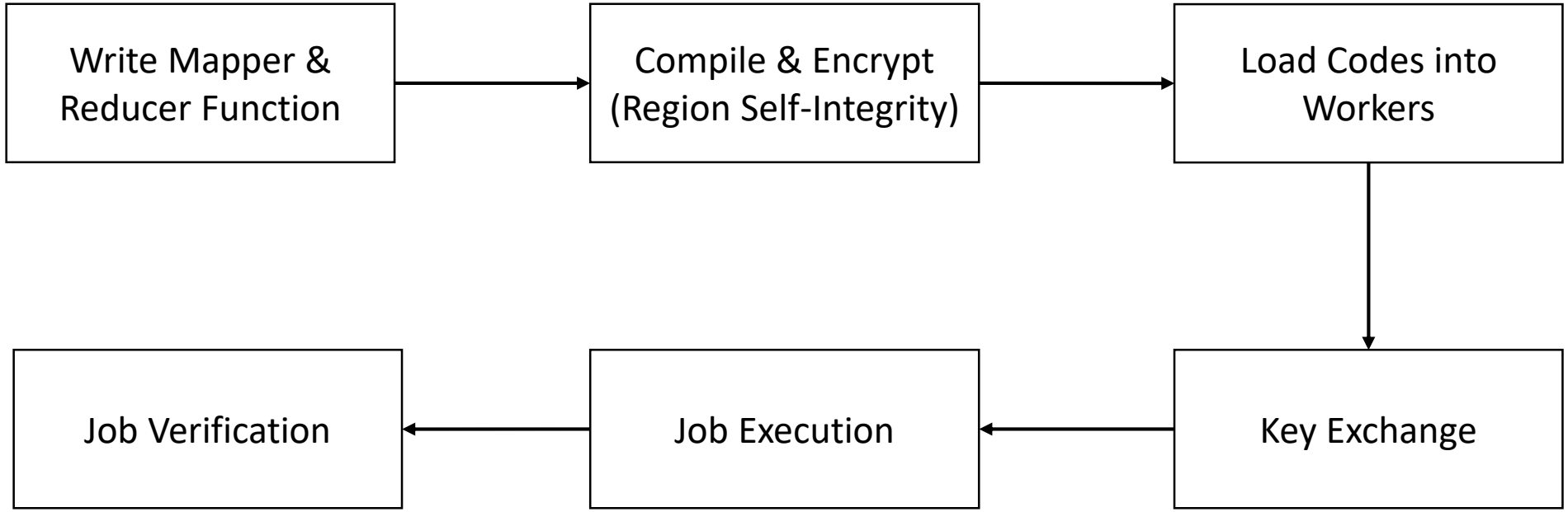
- Powerful adversaries:
  - hardware + software
  - except processor

- Denial-of-service
- Side-channels
- Traffic-analysis

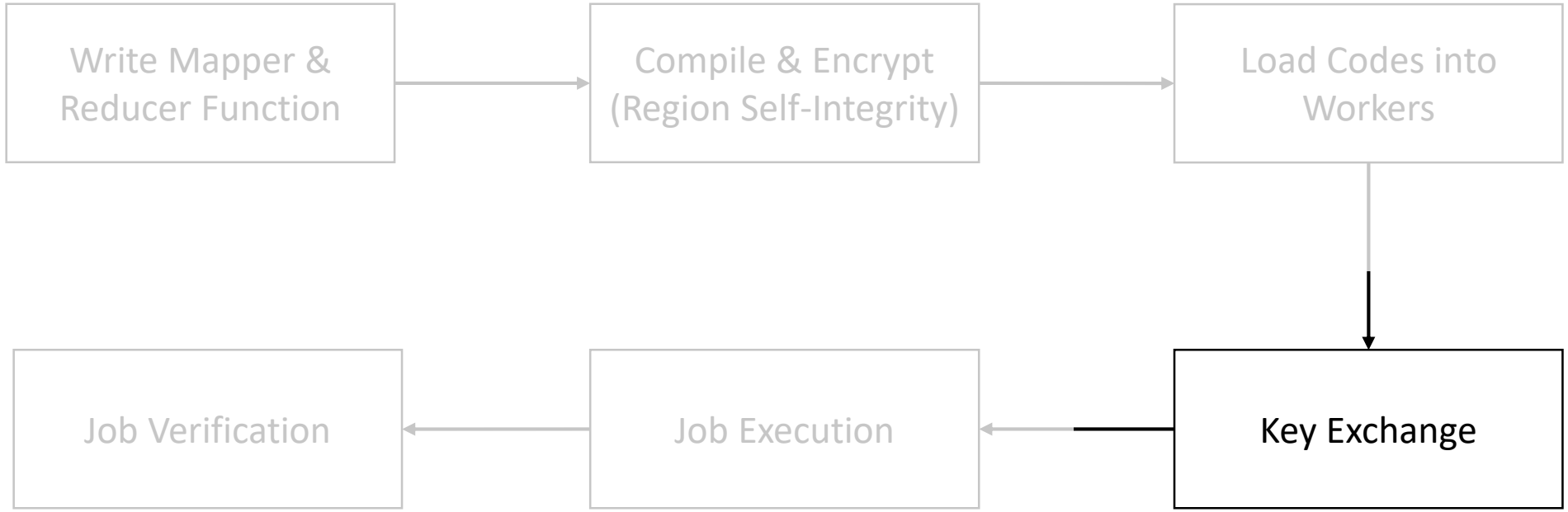
Security Guarantees

- Confidentiality
- Integrity
- Verifiability of execution

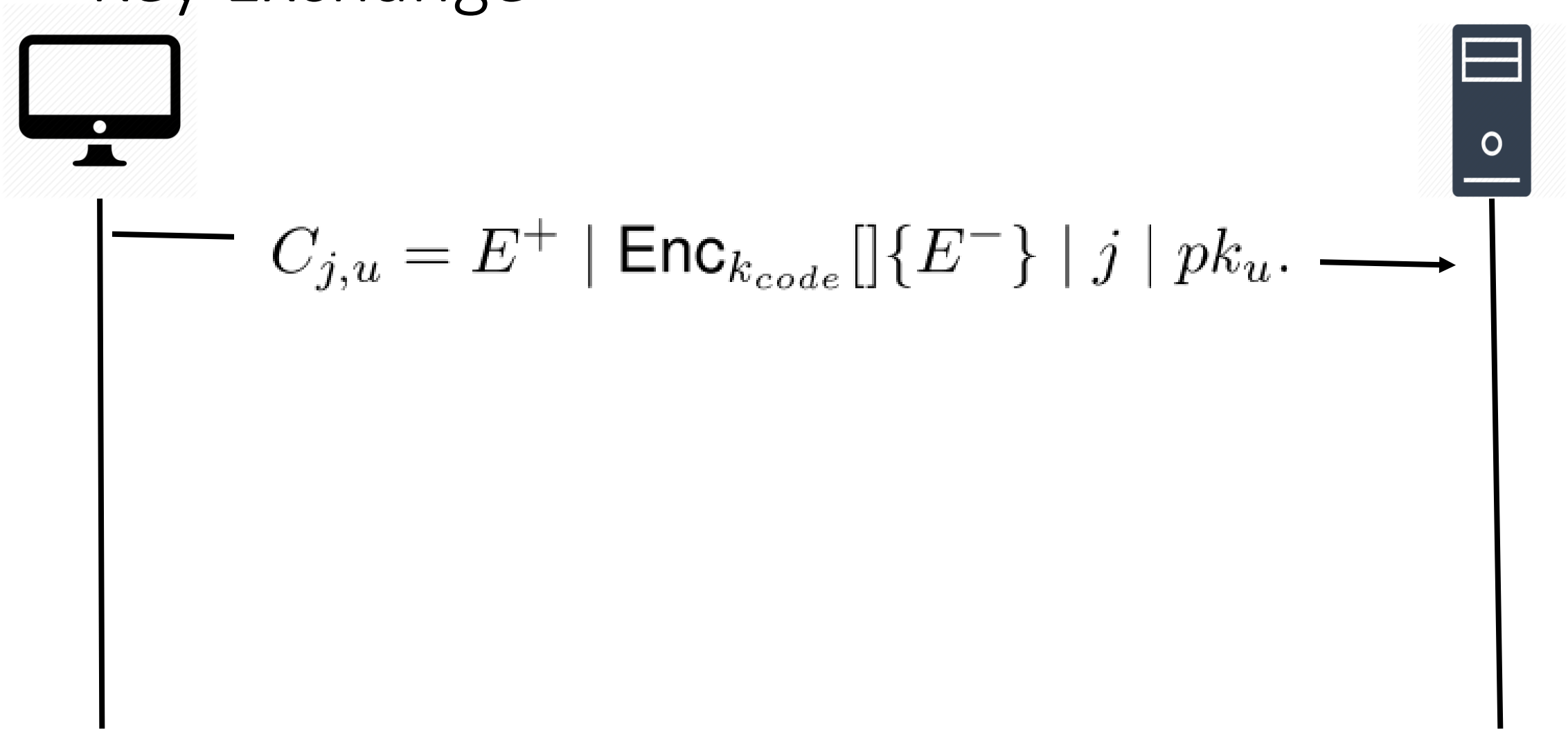
Design Overview



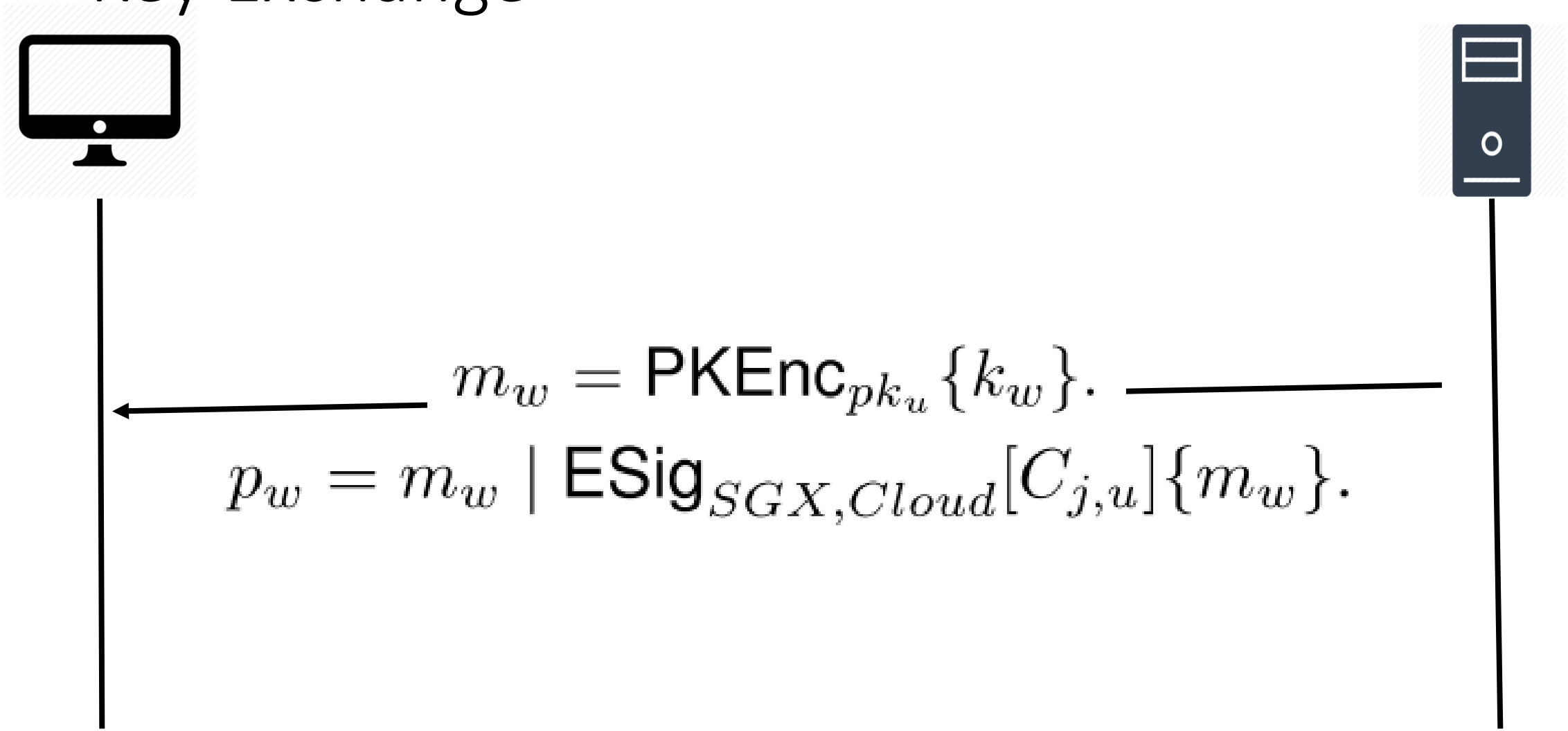
Design Overview



Key Exchange

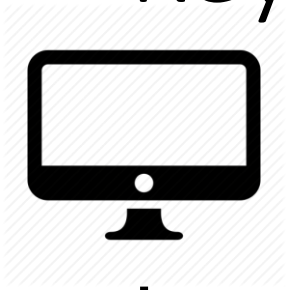


Key Exchange



Why they use so many keys?

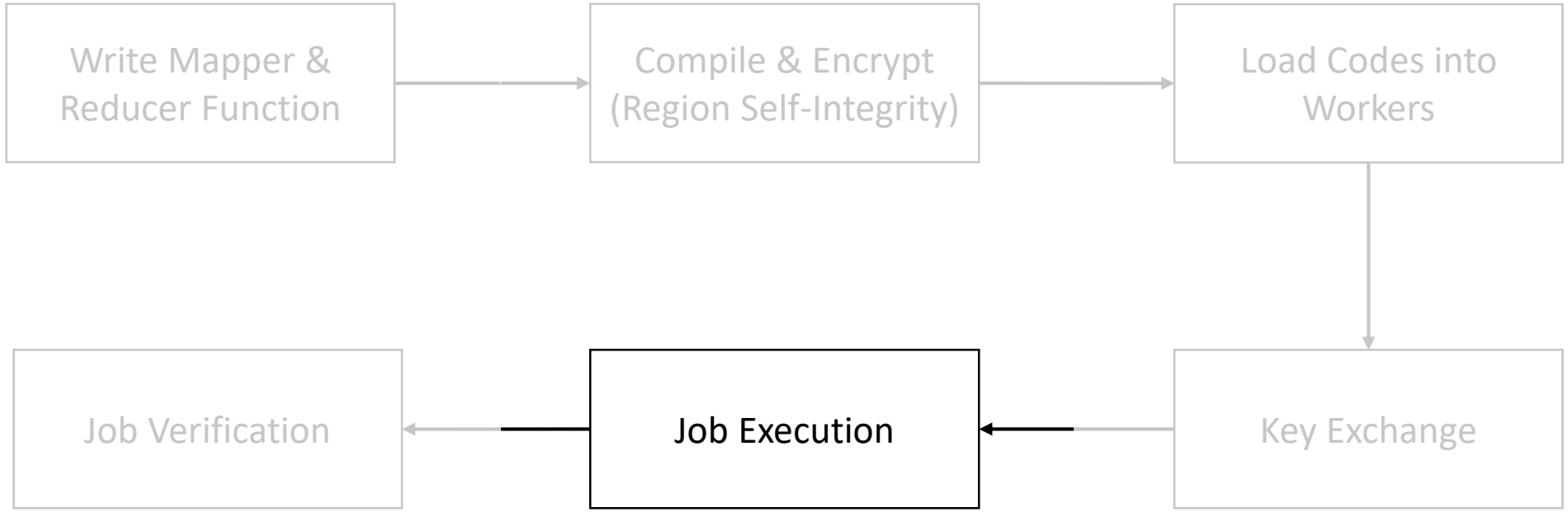
Key Exchange



$$JC_w = \mathsf{Enc}_{k_w}[]\{k_{code} \mid \mathbf{k}\}$$

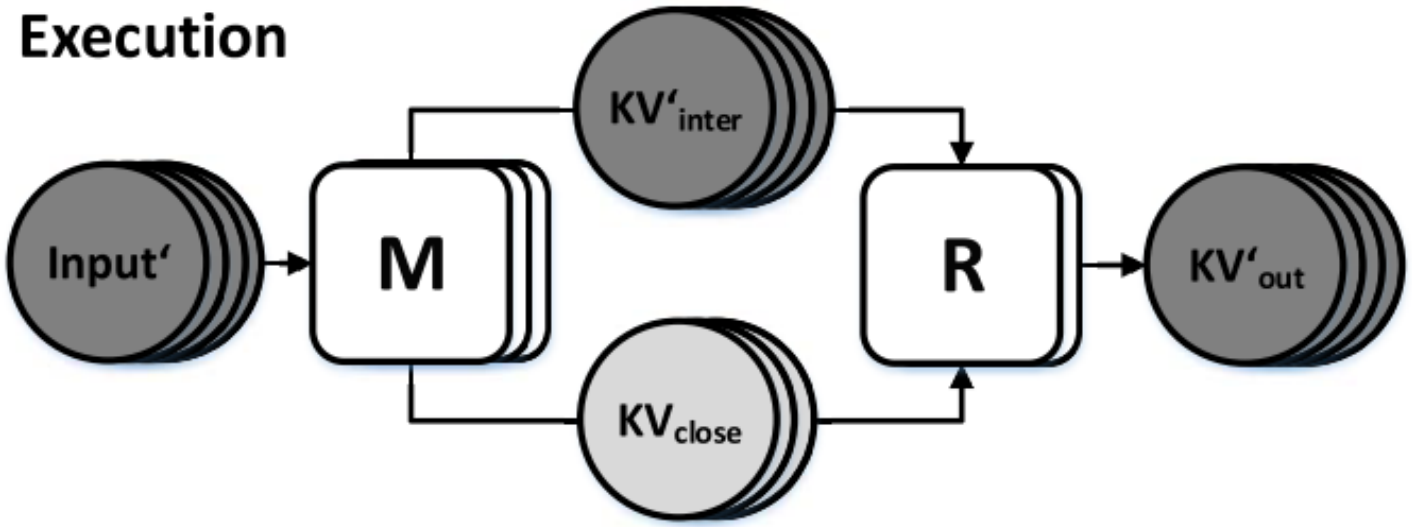
$$\mathbf{k} = k_{job} \mid k_{in} \mid k_{inter} \mid k_{out} \mid k_{prf}$$

Design Overview



Why we still need the r in the KVclose value?

Job Execution: Mapping



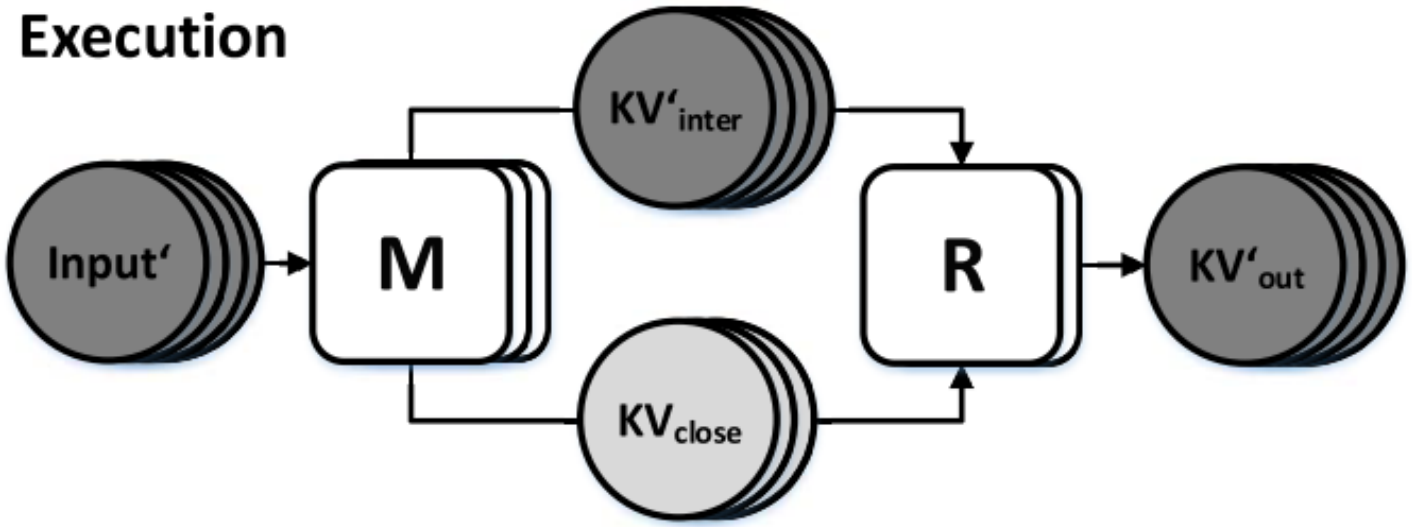
$$K'_{inter} = r \equiv \mathsf{PRF}_{k_{prf}}(K_{inter})\ mod\ R$$

$$V'_{inter} = \mathsf{Enc}_{k_{inter}}[j \mid \ell_m \mid r \mid i_{m,r}]\{\langle K_{inter} : V_{inter}\rangle\}$$

$$KV'_{inter} = \langle K'_{inter} : V'_{inter}\rangle$$

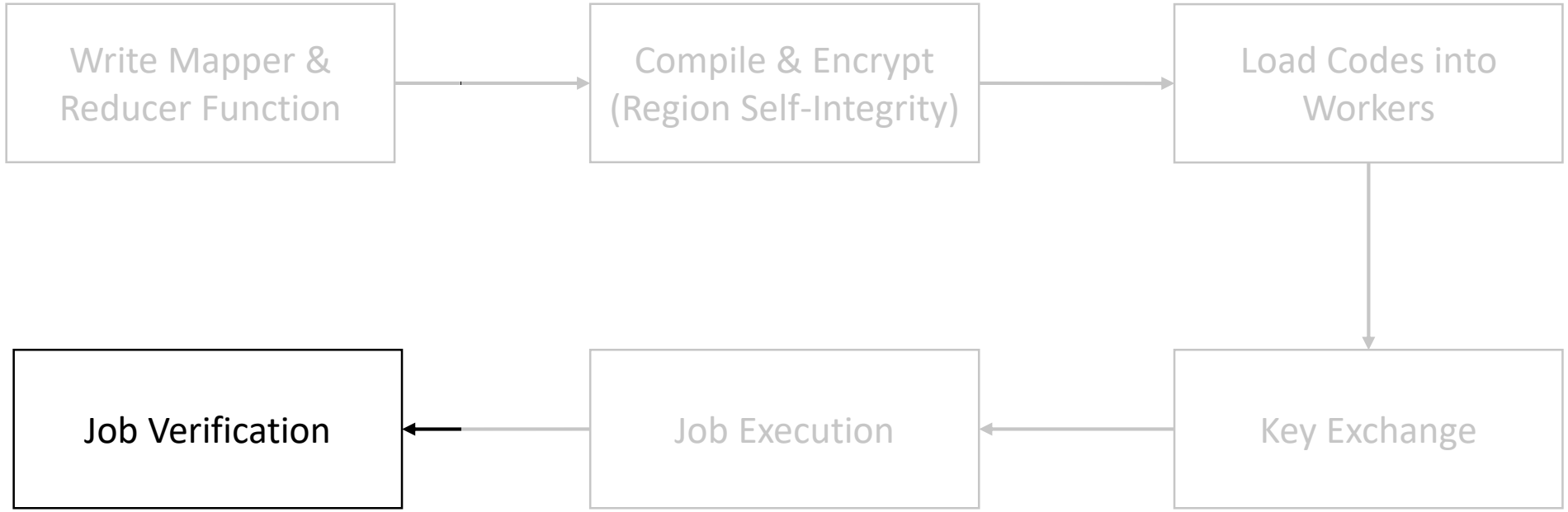
$$KV_{close} = \langle r : \mathsf{Enc}_{k_{inter}}[j \mid \ell_m \mid r \mid i_{m,r}]\{\}\rangle$$

Job Execution: Reducing



$$KV'_{out} = \langle \ell_{out} : \mathsf{Enc}_{k_{out}}[\ell_{out}]\{KV_{out}\}\rangle$$

Design Overview



Job Setup

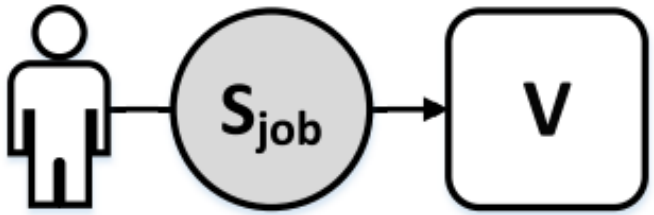
Key for Verification Message

$$j \mid k_{job} \mid \bar{R} \mid B_{in}$$

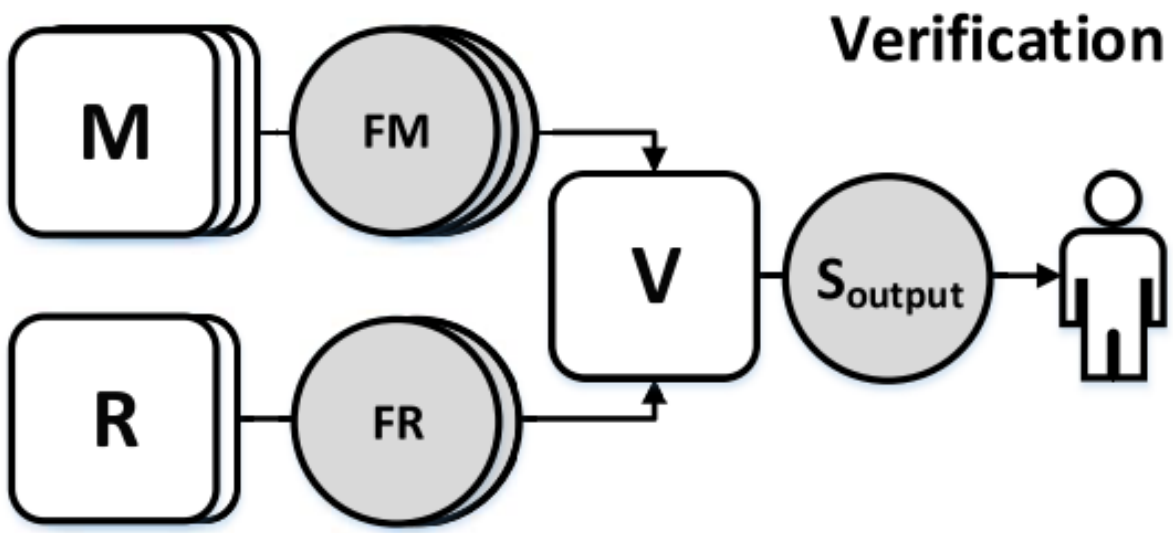
Job Identifier

Input Indexes

Number of Reducers

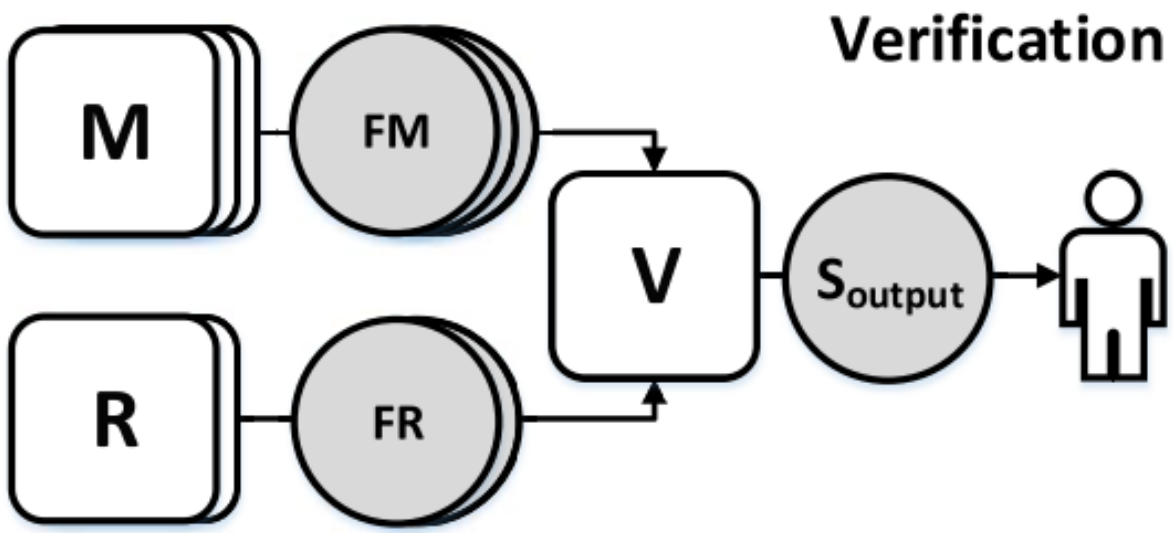


Job Verification: Mapper



$$FM = \mathsf{Enc}_{k_{job}}[j \mid \ell_m \mid B_{in,m}]\{\}$$

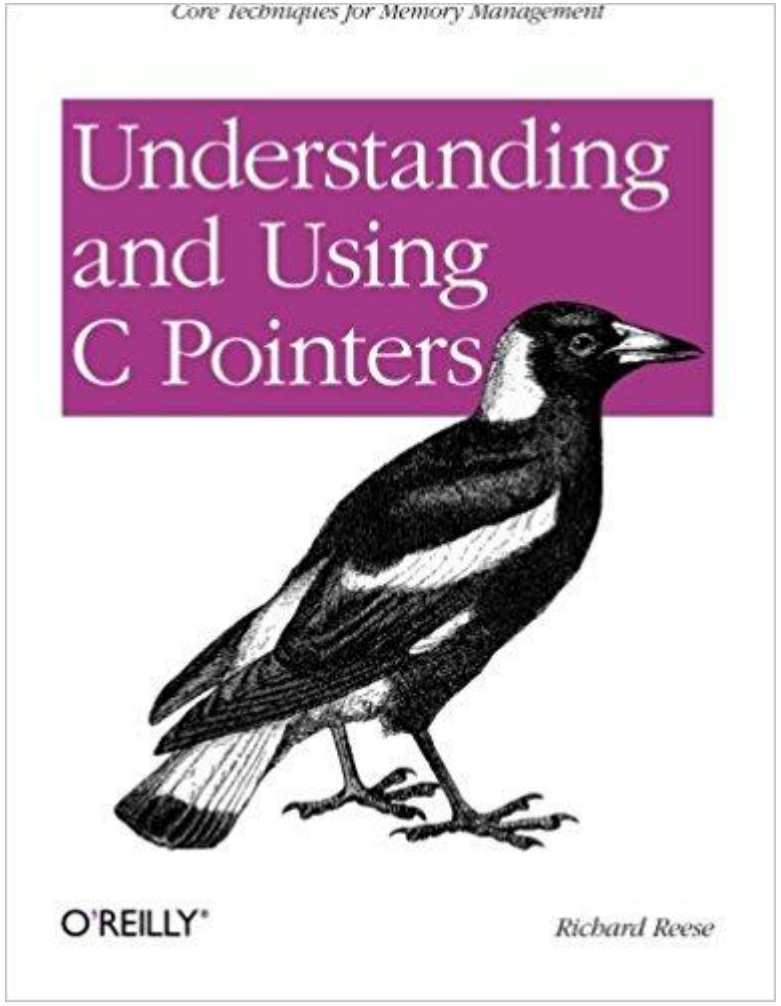
Job Verification: Reducer



$$FR = j \mid r \mid B_{out,r} \mid \mathsf{Enc}_k(j \mid r \mid B_{out,r} \mid P_r, \{\})$$

$$P_r \subseteq (\ell_m)_{m \in \mathbf{m}}$$

Are we good now?



- Defects
  - Write through pointers
  - Indirect call instructions
  - Read through pointers
- Self-Integrity Compiler!
  - Bitmap
  - Dynamic Check

Extension & Discussion

- How does the system guarantee confidentiality and integrity?
- How does system guarantees verifiability?
- Can we launch Meltdown or Spectre attack SGX?
- We need to trust intel to trust the quote scheme, so is this scheme just a transfer of trust? (Like many other schemes)