Wrapping entire Kubernetes clusters into a confidential-computing envelope with Constellation
What does cloud even mean?

Cloud Provider
at remote location

You

VM
VM
VM

Hypervisor
CPU
Memory
Devices
Threats

- Data in memory is plaintext
  - Physical access to memory
  - Insider attacks
  - Compromised cloud provider
  - Cross-tenants attacks
Confidential Computing to the rescue!

- Hardware-based Trusted Execution Environment (TEE)
- Protects data confidentiality and integrity in use
- Protects code integrity in use
- Based on memory encryption
AMD Secure Encrypted Virtualization (SEV)

- Confidential Virtual Machine (CVM)
- CVM and CPU in Trusted Computing Base (TCB)
- Transparent encryption of memory per CVM
AMD SEV: Memory encryption

- Separate key per CVM
- Inaccessible by software
- Pages AES encrypted in RAM
- Tracking page ownership
Problem solved?

- Is TEE hardware used?
- Is the right code running?
- Is the firmware up to date?

➢ Remote attestation needed!
Remote ATtestation procedureS: RFC 9334 (RATS)
Remote ATtestation procedureS: RFC 9334 (RATS)
Measuring what’s there

- **Launch measurement** during Confidential VM boot
- Measure guest Pages

\[ L_n = \text{Hash}( L_{n-1} \| \text{Page} \| \text{Metadata} ) \]
Measuring what’s there

- **Platform measurement** = Security Version Numbers (SVN)
  - Secure Processor Bootloader SVN
  - Secure Processor OS SVN
  - SNP firmware SVN
  - Processor microcode SVN
Attestation flow

- Secure Processor
- Confidential VM
- Verifier

Attestation flow process:
1. Secure Processor requests Attestation Report from Confidential VM.
2. Confidential VM responds with Attestation Report.
3. Secure Processor then requests Attestation Report from Verifier.
type snpAttestationReport struct {
    ReportedTCB        tcbVersion
    LaunchMeasurement  [48]byte
    ReportData         [64]byte
    IDKeyDigest        [48]byte
    ChipID             [64]byte
    ... more fields
    Signature          [512]byte
}

type tcbVersion struct {
    Bootloader    uint8
    TEE           uint8
    SNP           uint8
    Microcode     uint8
}

Evidence

Identifies code running in CVM

Data from attestation request

Platform measurement

Signed by VCEK
Versioned Chip Endorsement Key (VCEK)
Attestation report

type snpAttestationReport struct {
  ReportedTCB        tcbVersion
  LaunchMeasurement  [48]byte
  ReportData         [64]byte
  IDKeyDigest        [48]byte
  ChipID             [64]byte
  ... more fields
  Signature          [512]byte
}

type tcbVersion struct {
  Bootloader  uint8
  TEE         uint8
  SNP         uint8
  Microcode   uint8
}

Evidence

Identifies code running in CVM

Data from attestation request

Connects report to Chip Unique Secret and VCEK

Signed by VCEK

Platform measurement
Versioned Chip Endorsement Key (VCEK)

Secure Processor
- Chip Unique Secret
- Key Derivation Function
  - VCEK\textsubscript{priv}
  - Report

Hypervisor
- SVN

CVM
- VCEK\textsubscript{pub}
- Report

Verifier
- VCEK\textsubscript{pub}
- Report

AMD
- Root CA
- VCEK\textsubscript{pub}

Endorsement

Sign

Sign

Verify
constellation config generate <cloud>
constellation create
constellation init

kubectl [scale anything!]
How do we get there?
Measured Boot 101 - TPM

<table>
<thead>
<tr>
<th>PCR[0]</th>
<th>2c26b46b68fffc68ff99b453c1d30413413422d706483bfa0f98a5e886266e7ae</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR[14]</td>
<td>...</td>
</tr>
<tr>
<td>PCR[15]</td>
<td>0000000000000000000000000000000000000000000000000000000000000000</td>
</tr>
</tbody>
</table>

...
Measured Boot 101 – Extend PCR

\[ PCR[i]' = hash(PCR[i] || data) \]
Measured Boot Chain

TPM

- hash(OS)
- hash(Bootloader)
- Firmware
Confidential VM Boot

- Hypervisor

- Guest

  - initial memory
Launching a Guest

Hypervisor

Guest

initial memory
Launch measurement

Hypervisor

Guest

initial memory

ID Block

Launch Digest

Policy

Image Metadata

Signature
Launch measurement

Hypervisor

Guest

initial memory

Attestation report

Measurement

...
Protection rings

- ring 1: Userspace
- ring 0: Kernel
- ring -1: Hypervisor
Measured boot chain – SVSM

VMPL1

...  

VMPL0

SVSM

vTPM
Measured boot chain

- Root FS (dm-verity)
- Unified Kernel Image (UKI)
- Firmware (OVMF)
- SVSM
- vTPM

VMPL1

VMPL0
Unified Kernel Image

roothash=a5ee4b42f70ae1f46a08a7c92c2e0a20672ad2f514792730f5d49d7606ab8fdf
dm-verity
Constellation Nodes

Unified Kernel Image

Bootloader

Firmware

k8s

Pod

Bootstrapper

State Disk

rootFS

...
How to get from here...
... to a Confidential Cluster?
Cluster initialization

DevOps engineer

First node

aTLS handshake

Init call

Kubeconfig, ...
aTLS handshake

- ClientHello(nonce)
- ServerCertificate(AttestationStatement), ServerHelloDone
- ClientKeyExchange
- ChangeCipherSpec, Finished
- Verify Attestation
aTLS handshake - Attestation

VCEK

SEV Attestation report
- Report Data
- Signature

Attestation Key

vTPM Quote
- PCRs
- Nonce
- Signature
Autonomous Join

New node

Join Service

aTLS handshake

Request to Join

JoinToken, ...

New node
mutual aTLS handshake

Client

ClientHello(nonce)

ServerCertificate(AttestationStatement), AcceptableCAs(nonce), ServerHelloDone

ClientKeyExchange, ClientCertificate(AttestationStatement)

ChangeCipherSpec, Finished

Server

Verify Attestation

Client

Verify Attestation

ChangeCipherSpec, Finished

Server

Client
… bringing it all together
What are you waiting for?

```
constellation config generate <cloud>
constellation create
constellation init

kubect1 [scale anything!]
```
Thanks!

- Check it out on GitHub: github.com/edgelesssys/constellation
- Ask us for a demo!

Malte Poll
@malte@chaos.social
@malt3
github.com/malt3

Paul Meyer
@katexochen@infosec.exchange
@katexochen
github.com/katexochen
Learn more

- Constellation documentation
- Confidential Computing whitepaper
- Constellation cluster attestation
- Edgeless Systems blog