Introduction to MLS

NALINI ELKINS

INDUSTRY NETWORK TECHNOLOGY COUNCIL

PRESIDENT@INDUSTRYNETCOUNCIL.ORG

A few words about me

- President: Industry Network Technology Council
- Founder & CEO: Inside Products, Inc. and Outside the Stacks, Inc.
- Advisory Board: India Internet Engineering Society
- RFCs: RFC8250 (Embedded performance and diagnostics for IPv6) and others
- Product developer (OEMed by IBM and others)
- Working with IPv6 for 15 years
- Working with network management, diagnostic, performance issues at large brick-and-mortar enterprises for over 30 years

Agenda

- Why Message Layer Security (MLS)?
- MLS charter
- Explain underpinnings of scalability
- Rachet / Merkle Trees
- MLS enterprise application

Charter: MLS

Need for group key establishment and message protection protocols with the following properties:

- Message Confidentiality
- Message Integrity and Authentication
- Membership Authentication
- Asynchronicity
- Forward secrecy
- Post-compromise security
- Scalability

https://datatracker.ietf.org/wg/mls/about/

Confidentiality

Message Confidentiality - Messages can only be read by members of the group

Implications / Questions:

- “Confidentiality” means encryption
- How will it be set up?
- Public key encryption? PKI?
- Certificates?

Integrity and Authentication

Message Integrity and Authentication - Each message has been sent by an authenticated sender, and has not been tampered with

Implications / Questions:

- Who will authenticate?
- Is there an authentication procedure?
- What does “tampered” mean? Is this more than checksum?
- What happens if something IS considered to be “tampered” with?

Membership Authentication

Membership Authentication - Each participant can verify the set of members in the group

Implications / Questions:

- Why is this necessary / desireable?
- How is this to be done?
- Is there a list of members somewhere?
- Who can add / remove members?
- Who has this list?
- How do we keep that secure?

Asynchronicity

Asynchronicity - Keys can be established without any two participants being online at the same time

Implications / Questions:

- How is this to be done?
- Is there a list of “keys” somewhere?
- How is this associated with participants?
- Who has this list?
- How do we keep that secure?

Forward Secrecy

Forward secrecy - Full compromise of a node at a point in time does not reveal past messages sent within the group

Implications / Questions:

- How is this to be done?
- Keys have to be rotated. How is this synchronized with participants?
- Who has this list?
- How do we keep that secure?

Post-Compromise Security

Post-compromise security - Full compromise of a node at a point in time does not reveal future messages sent within the group

Implications / Questions:

- How is this to be done?
- Does this mean that keys have to be rotated? How is this synchronized with participants?

Scalability

Scalability - Resource requirements have good scaling in the size of the group (preferably sub-linear)

Implications / Questions:

- How is this to be done?
- What is the highest size of the group?
- What is the scaling, if not sub-linear?
- What exactly IS sub-linear?

Linear Growth

- Starting at the age of 25, imagine if you could save $20 per week, every week, until you retire, how much money would you have stuffed under your mattress at age 65?
- To solve this problem, we could use a linear growth model.

From: https://math.libretexts.org/Bookshelves/Applied_Mathematics/Book%3A_College_Mathematics_for_Everyday_Life_(Inigo_et_al)/04%3A_Growth/4.01%3A_Linear_Growth#:~:text=Example%204.1.-,1%3A%20City%20Growth,grows%20by%20a%20constant%20amount.

Linear Growth

- Linear growth has the characteristic of growing by the same amount in each unit of time.

- In this example, there is an increase of $20 per week; a constant amount is placed under the mattress in the same unit of time.

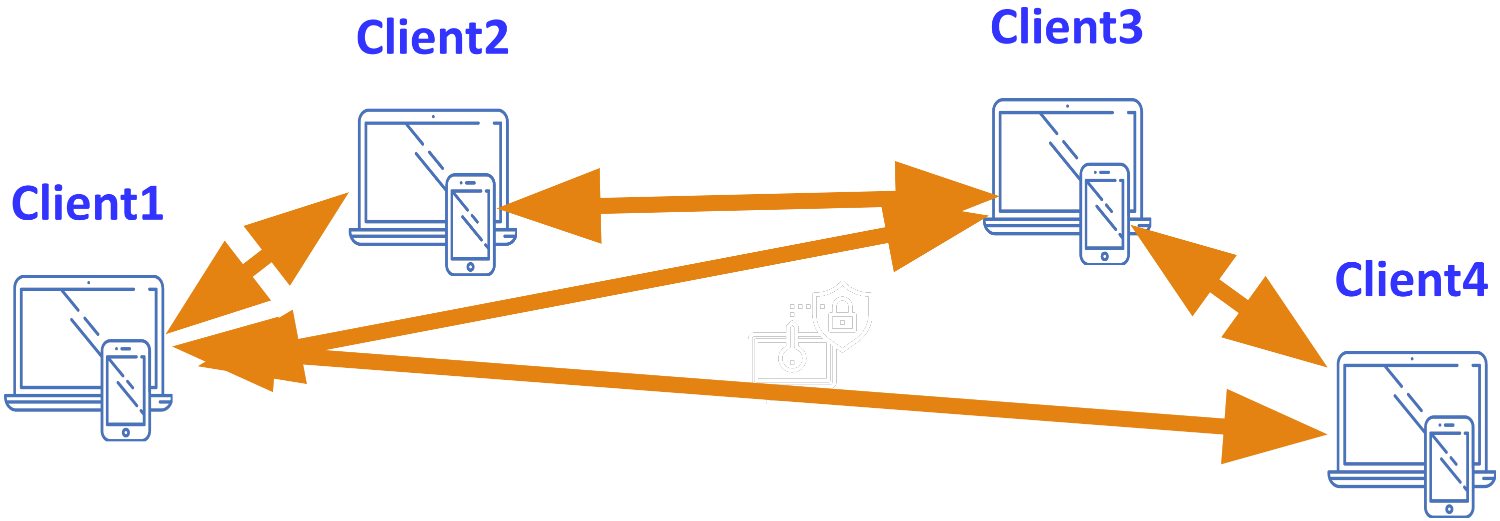
Linear Growth

- If we start with \$0 under the mattress, then at the end of the first year we would have $\$20 \cdot 52 = \$1040$.

- So, this means you could add \$1040 under your mattress every year. At the end of 40 years, you would have $\$1040 \cdot 40 = \$41{,}600$ for retirement.

- This is not the best way to save money, but we can see that it is calculated in a systematic way.

Linear Growth in Messaging Applications

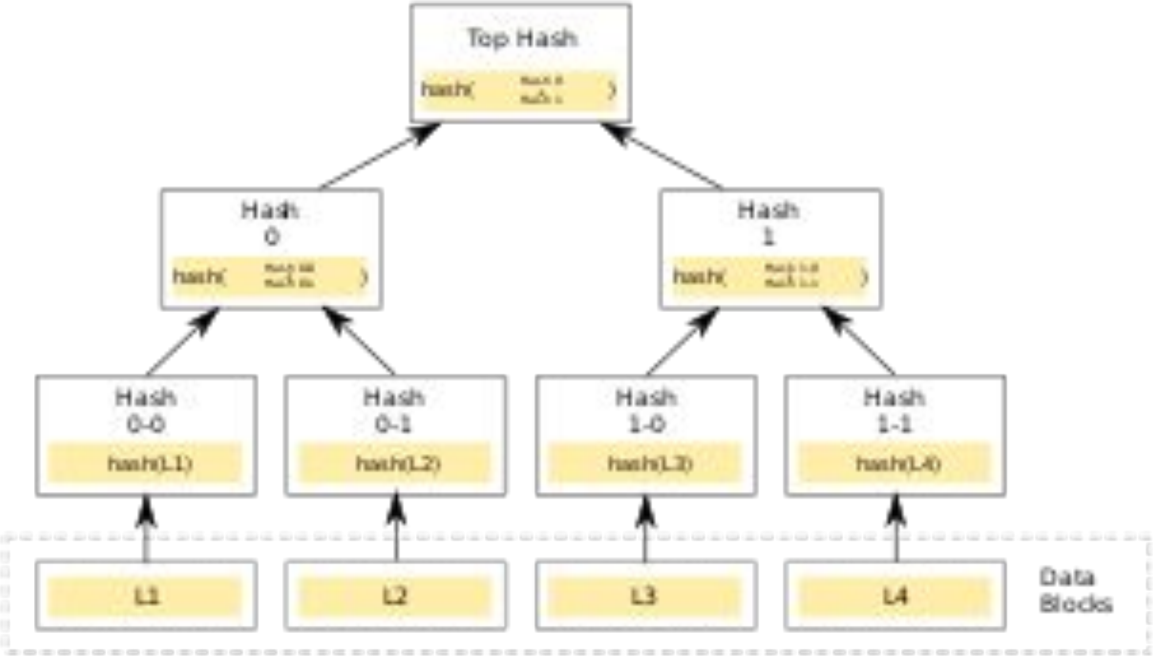


- Signal limits to 1,000 members
- WhatsApp limits the number of participants in a group to 1024 (though this can be circumvented client-side)
- MLS can be 2 to tens of thousands

Rachet Trees

- The way MLS achieves scalability is via “Rachet Trees”
- Rachet Trees are a type of Merkle Tree
- So, what exactly IS a Merkle Tree?

What is a Merkle Tree?



A hash or Merkle tree is a tree of hashes in which the leaves (i.e., leaf nodes, sometimes also called "leafs") are hashes of data blocks in, for instance, a file or set of files.

https://en.wikipedia.org/wiki/Merkle_tree

What is a Binary Search?

Binary search is an algorithm used to find a specific value in a sorted list of elements. It is an efficient algorithm that quickly locates the desired element by repeatedly dividing the search space in half.

1. **Initial setup**: Assume we have a sorted list of numbers. Let's say we want to find the number 7 in the list

[2, 3, 4, 5, 6, 7, 8, 9, 10].

Divide the List

2. **Divide the list**: We start by looking at the middle element of the list, which is 6. Since 7 is greater than 6, we know it must be in the second half of the list.

↓

[2, 3, 4, 5, 6, 7, 8, 9, 10]

3. **Divide again**: We now divide the second half of the list in half again and look at the middle element, which is 8. Since 7 is less than 8, we know it must be in the first half of the remaining elements.

↓

[2, 3, 4, 5, 6, 7, 8, 9, 10]

Find the answer

4. **Final step**: We repeat the process of dividing the search space until we find the desired element or determine that it doesn't exist. In this case, we find the number 7.

⇩

[2, 3, 4, 5, 6, 7, 8, 9, 10]

Binary sort eliminates half of the remaining elements in each iteration, making it highly efficient.

Please note that the list in the example was already sorted. If the list is not sorted, you would need to sort it first before applying binary sort.

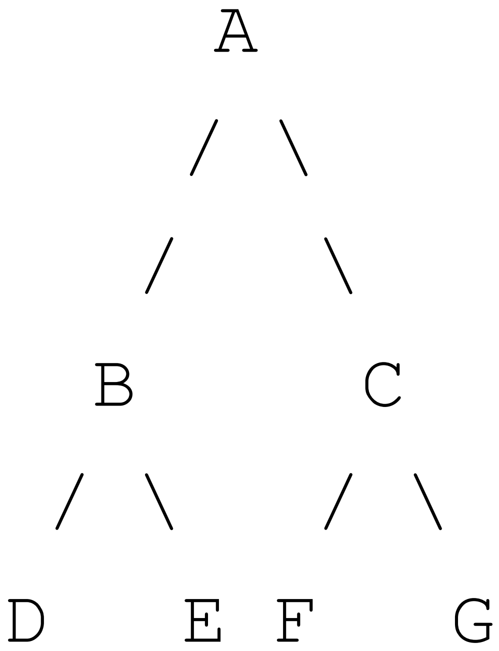
Why Binary Search?

- When you have to search a large number of items, it is better to use a binary search rather than sequential.

- Let’s now discuss binary trees and Merkle / rachet trees which are used in current cryptography (MLS / TLS)

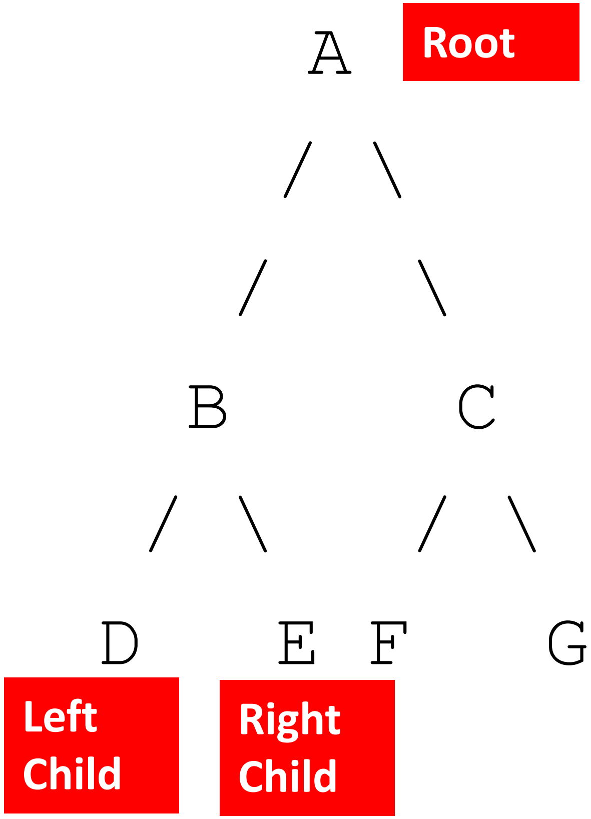
What is a Binary Tree?

- A binary tree is a type of data structure that consists of nodes connected in a hierarchical manner.
- Each node in a binary tree can have at most two children, referred to as the left child and the right child.
- The topmost node of the tree is called the root node.



What is a Binary Tree?

- A binary tree is a type of data structure that consists of nodes connected in a hierarchical manner.
- Each node in a binary tree can have at most two children, referred to as the left child and the right child.
- The topmost node of the tree is called the root node.



A

Root

B

C

D

E

F

G

Left Child

Right Child

Charter: Rationale

From charter:

“Several widely-deployed applications have developed their own protocols to meet these needs. While these protocols are similar, no two are close enough to interoperate.”

Example of messaging applications: Signal, Telegram, WhatsApp Messenger, WeChat, QQ Messenger, Viber, Line, and Snapchat.

Charter: Shared Code and Validation

From charter:

> "As a result, each application vendor has had to maintain their own protocol stack and independently build trust in the quality of the protocol. The primary goal of this working group is to develop a standard messaging security protocol for human-to-human(s) communication with the above security and deployment properties so that applications can share code, and so that there can be shared validation of the protocol (as there has been with TLS 1.3)."

Charter: Shared Code and Validation

From charter:

“It is not a goal of this group to enable interoperability/federation between messaging applications beyond the key establishment, authentication, and confidentiality services. Full interoperability would require alignment at many different layers beyond security, e.g., standard message transport and application semantics. The focus of this work is to develop a messaging security layer that different applications can adapt to their own needs.”

MIMI

The More Instant Messaging Interoperability (MIMI) working group will specify the minimal set of mechanisms required to make modern Internet messaging services interoperable.

https://datatracker.ietf.org/wg/mimi/about/

Existing Implementations

- MLSpp (C++) https://github.com/cisco/mlspp (Status: RFC)
- OpenMLS (Rust) https://github.com/openmls/openmls (Status: RFC)
- Wickr proprietary implementation (Rust) (Status: RFC)
- RingCentral proprietary implementation (C++) (Status: draft-11; RFC in progress)
- MLS* (F*) (Status: RFC in progress)
- BouncyCastle (Java) (Status: RFC in progress)
- go-mls (Go) (Status: RFC in progress)

https://github.com/mlswg/mls-implementations/blob/main/implementation_list.md

MLS General

- Clients (participants)
- Groups
- Groups: very small to very large.
- Users can have multiple MLS clients.

MLS Key Services

- Authentication Service (AS)
- Delivery Service (DS)

- Implementation options:
  - Centralized
  - Decentralized
  - Combined

MLS Authentication Service (AS)

- Role of the Authentication Service (AS)
- Responsibilities
  - Binding identifiers
  - Public key material
- Generation and validation of credentials (signing)

MLS Delivery Service (DS)

- Role of the Delivery Service (DS)
- Responsibilities
  - Message distribution
  - Storage
- Broadcaster in group messaging

MLS at an Enterprise: A one act Play with Prologue

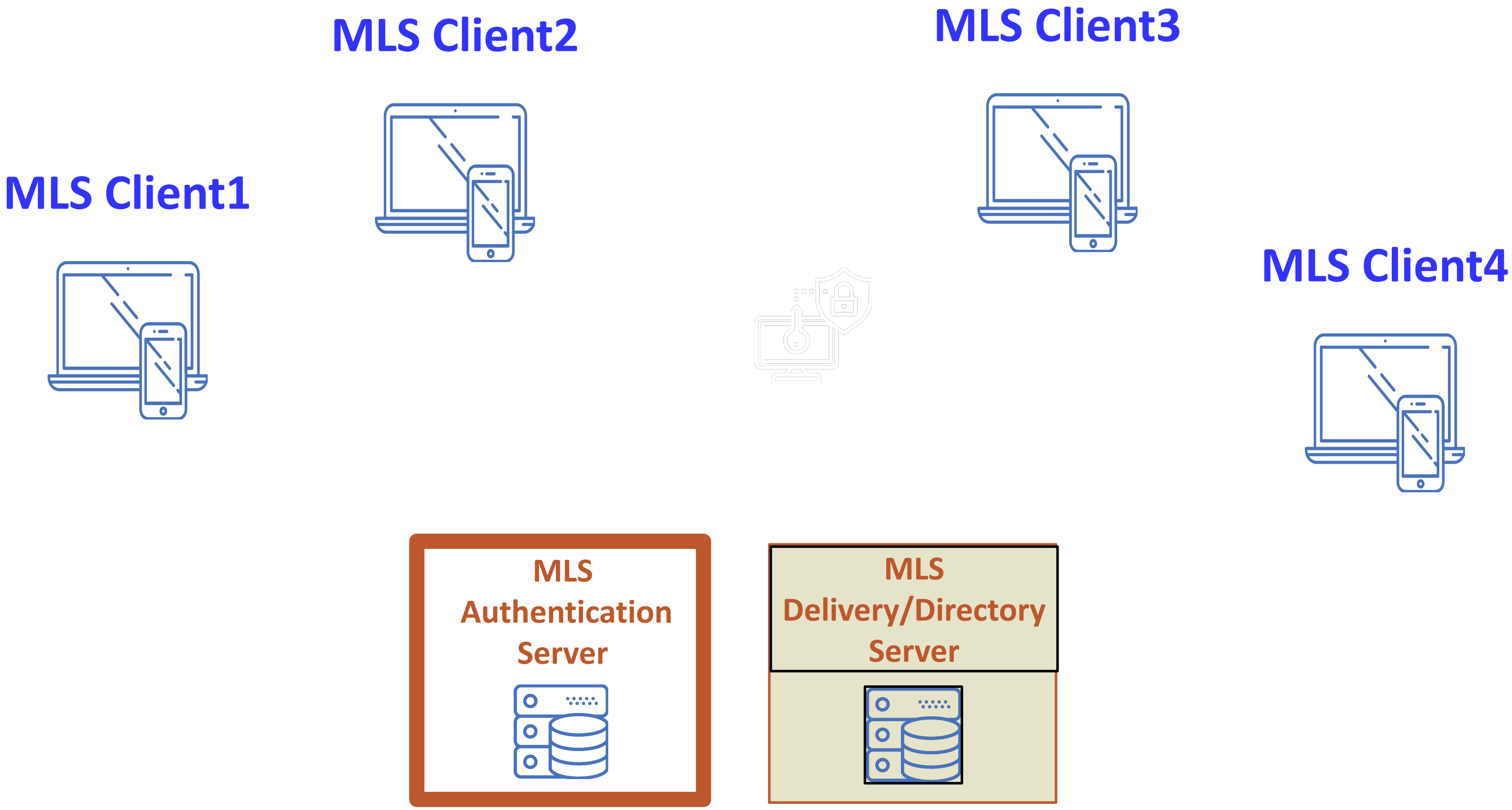
Dramatis Personae

Authentication Server (AS): appears only in Prologue for registration. The AS has signed the credential for the MLS Client (MC), so it need not appear in afterwards. The Client has received the AS signature public key.

Directory Server (DS): appears in Prologue (for storage of Key Package) and in Act 1: (for retrieval of Key Package). DS is required according to MLS spec.

MLS Client (MC): appears in Prologue (registration with AS and for storage of Key Package). It is the main actor in all subsequent drama.

Picture of Personae



MLS Client2

MLS Client3

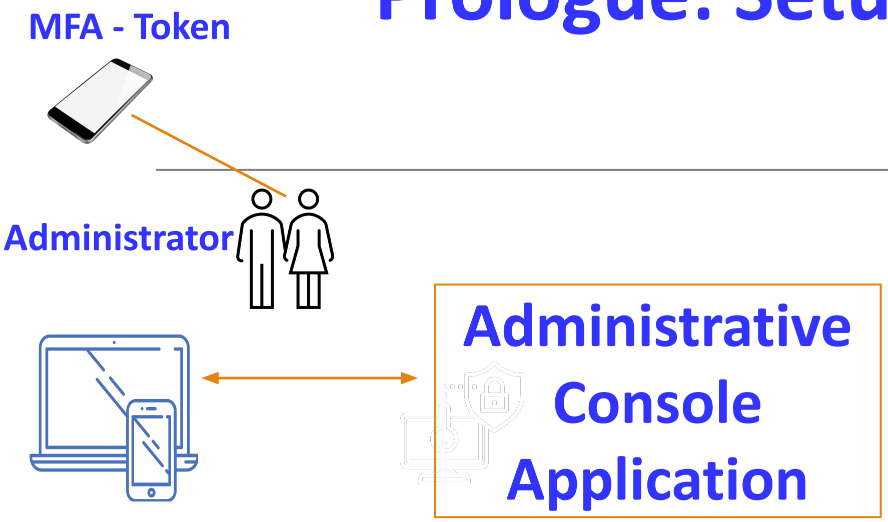
MLS Client1

MLS Client4

MLS Authentication Server

MLS Delivery/Directory Server

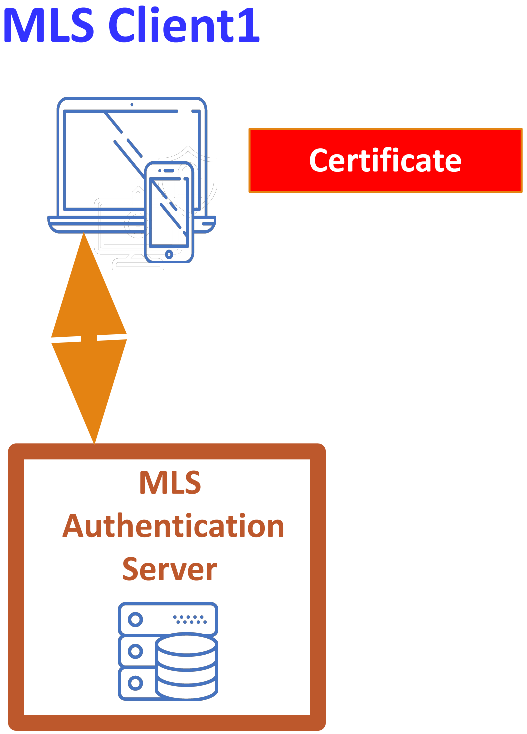
Prologue: Setup Administrators



Register Administrator

- The Administrator (human) logs on to an application.
- Often, Multi-factor Authentication (MFA) is done. Token is sent to cell phone to verify identity of administrator.
- The initial setup is done (authority, policy, etc.)

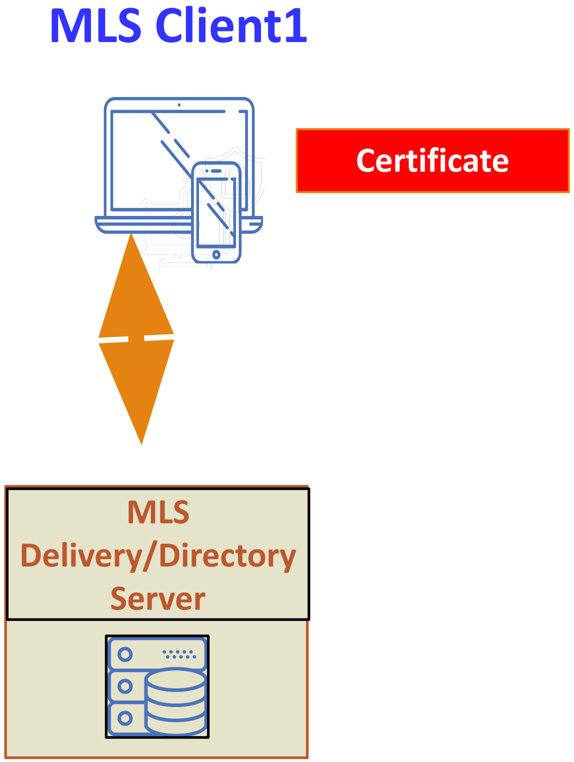
Prologue: Scene #1: MLS AS Enters



Register with MLS Authentication Server

- This is done any time an MLS Client wants to participate in group messaging.
- MLS code in Client creates the private keys, MLS identity, gets credential (certificate) signed by AS.
- Each client has to do this.

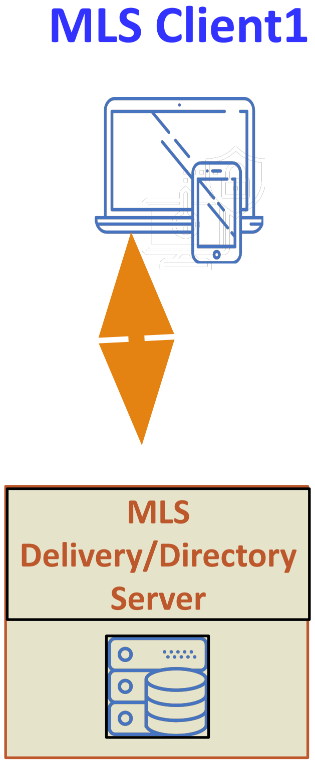
Prologue: Scene #2: MLS DS Enters



Register with MLS Delivery Server

- This is done following successful registration with the MLS AS in Scene #1.
- You need the Capabilities for each Client. This decision is made by the Administrator.
- Then, HPKE private keys and initial Key Packages are created and stored at the MLS Delivery Server.

Act 1: Scene #1: Start an MLS Group

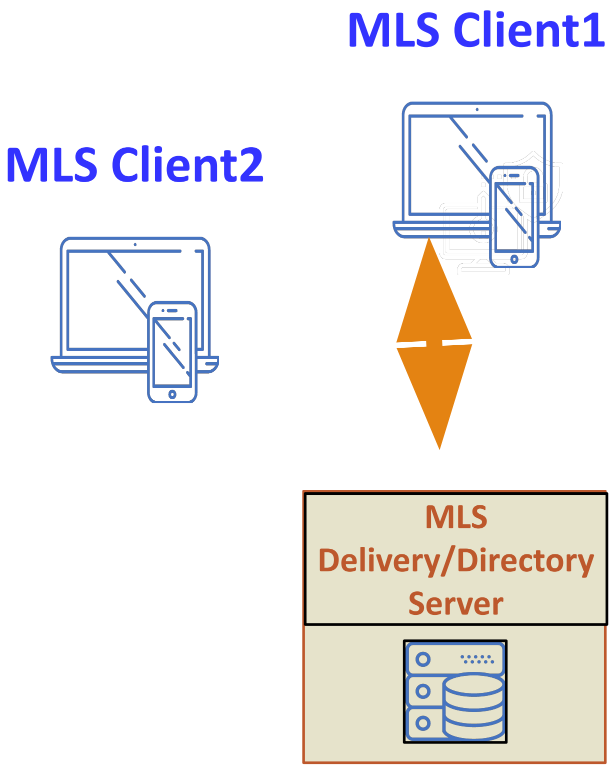


Create MLS Group

The MLS Client will:

- Create a new Group
- Start a new Rachet Tree, Group Context, Secrets Tree, Private Keys and other information.

Act 1: Scene #2: Add Clients to the Group



Add Clients to MLS Group

The MLS Client will:

- Download the KeyPackage for the other participant (client) and add it to the Group.
- The Rachet Tree, Group Context, Secrets Tree, Private Keys and other information are updated.

What can go wrong?

- Fake AS
- Fake DS
- Fake Client
- And more …

Questions?

Contact:

info@iiesoc.in

president@industrynetcouncil.org