MalaQ 논문 리뷰

https://youtu.be/CfwMmF2fww0

정보컴퓨터공학과 송경주

멀웨어 (Malware)

- 시스템을 파괴하거나 정보를 변조, 유출하는 등 악의적인 작업을 하도록 만들어진 소프트웨어
  - Ex) 바이러스, 트로이목마, 스파이웨어, 웜 등

<목적>

- 민감한 데이터, 디지털 자산 침해
- 로그인 자격 증명, 신용카드 번호 등 개인정보 강탈
- 기업 및 정부 기관이 의존하는 중요 시스템 파괴
- 멀웨어 감염은 모든 장치 또는 운영 체제 (Windows, Mac, iOS, Android)에서 발생할 수 있음

- **이러한 Malware을 양자컴퓨터 관점에서 생각해볼 수 있음**
  - 양자 Malware는 양자 논리 게이트의 형태로 나타날 수 있으며, 공격자가 설계하고 제어하는 전체 양자 알고리즘으로 나타날 수 있음 [1]
  - 악의적으로 측정(Measurement)을 실행하여 모든 데이터를 지울 수 있음[1]
    - 큐비트 특성 상 한번 측정(measurement)된 큐비트는 다시 사용할 수 없음 (즉, 데이터를 잃음)

    → 양자회로에서만 보이는 특징

[1] https://arxiv.org/pdf/quant-ph/0505126

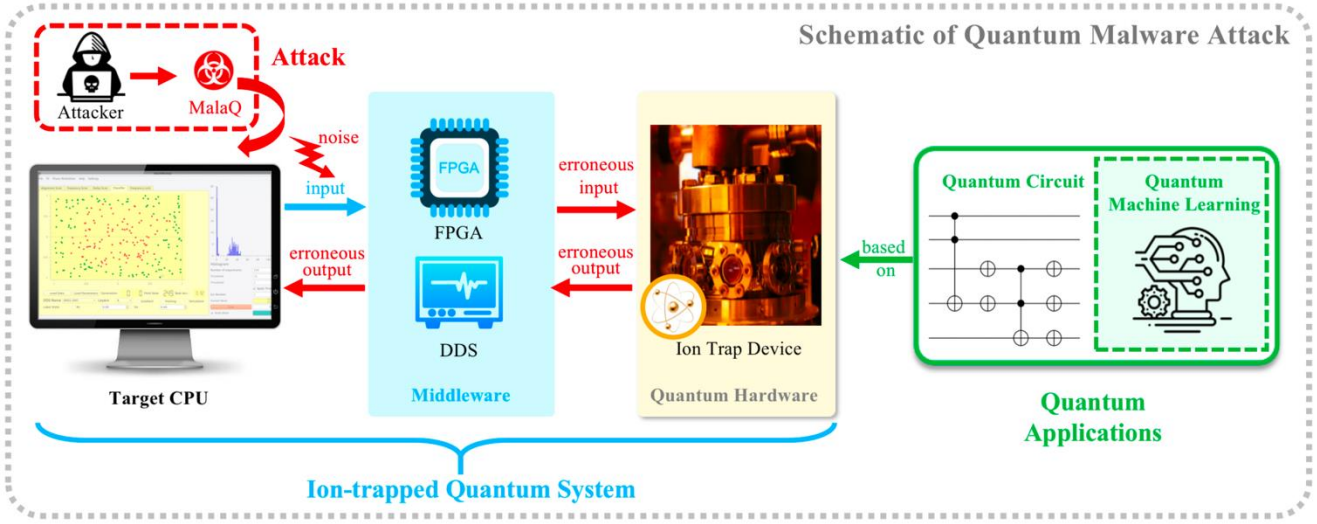
MalaQ -A Malware Against Quantum Computer

- TCP 기반 양자 멀웨어?
  - 데이터 탈취: TCP 연결을 통해 데이터를 주고받을 때, 트래픽을 가로채거나 감시 (정보를 탈취하기 위한 중간자 공격(Man-in-the-Middle) 실행)
  - 트래픽 변조: TCP 세션 중에 악성코드는 전달되는 데이터를 변조하거나 재전송
  - **백도어 생성: 피해자 시스템에 백도어를 설치하여 원격에서 시스템에 접근하거나 제어**

Ex) DDoS 공격: TCP SYN 플러딩(SYN Flooding) 공격을 통해 서버의 자원을 소모시켜 서비스 거부 상태 유도가능 → MalaQ 해당 x)

Quantum middleware

- **FPG:** 디지털 신호 처리 및 제어 (즉, 양자회로에 따라 디지털 신호 생성)
- **DDS:** FPGA에서 생성된 디지털 신호를 아날로그 신호로 변환하여 양자 하드웨어로 전달

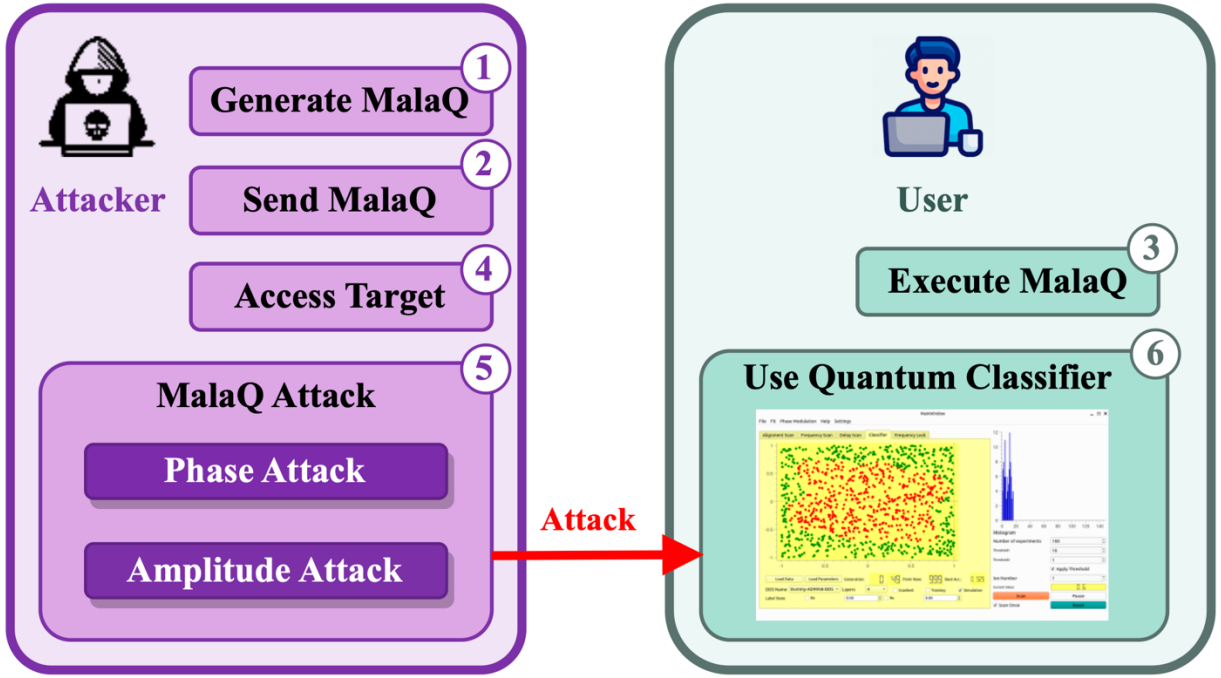


MalaQ - 동작방식

1. TCP 기반 멀웨어 MalaQ 생성 (QPU에서 실행되는 대상 분류기를 공격하는 데 사용)
2. 이메일, USB 등의 경로를 통해 QPU를 제어하는 CPU에 침투
3. 공격 대상 CPU에서 실행(사용자가 MalaQ 맬웨어를 실행하도록 유도)

→ 이러한 상황 가정?

4. MalaQ를 통해 QPU에 액세스하여 제어
5. (대상 알고리즘)양자 분류기 공격
   - Phase Attack
   - Amplitude Attack



MalaQ - 양자공격기법

- **Phase Attack**
  - 시뮬레이터 및 실제 양자 하드웨어 모두 적용 가능
  - MalaQ에 내장된 Metasploit을 사용하여 FPGA 및 DDS에서 사용하는 Phase 관련 스크립트 수정
    - Metasploit: 취약점을 분석해 CPU에 접근할 수 있도록 해주는 모의 해킹 툴 (오픈소스)

    → (제어 가능한) noise 추가

    → 공격 후 모든 로그정보를 지워 사용자가 감지하지 못하도록 함

- **Amplitude Attack**
  - 양자 하드웨어에서 적용 가능
  - 양자 시스템에서 QT-based control software를 사용하여 사용자가 진폭 매개변수를 정의함
    - QT-based control software: "Signal & Slot"이라는 메커니즘을 통해 사용자가 설정한 진폭을 설정하면 설정 값에 따라 시스템의 백엔드로 전달되어 양자 하드웨어가 이에 따라 동작함
  - **기존 동작: 사용자가 진폭 값 설정 → QT-based control software(QT의 'signal & Slot' 메커니즘)을 통해 사용자가 설정한 진폭 전달**
  - **공격 후 동작: 사용자가 진폭 값 설정 → 공격자가 QT의 'signal & Slot' 메커니즘을 공격(사용자가 설정한 진폭 값이 백엔드에 전달되기 전에 값 수정) → 잘못된 진폭 전달** **'signal & Slot' 메커니즘: 사용자 진폭(signal)에 따라 그에 따른 함수(Slot)이 호출됨
  - QT-based control software의 'Signal & Slot'메커니즘에서 Slot 함수를 수정하면 됨
  - 결과적으로 사용자는 정확한 값을 전달했다고 여기지만 잘못된 값이 추출됨

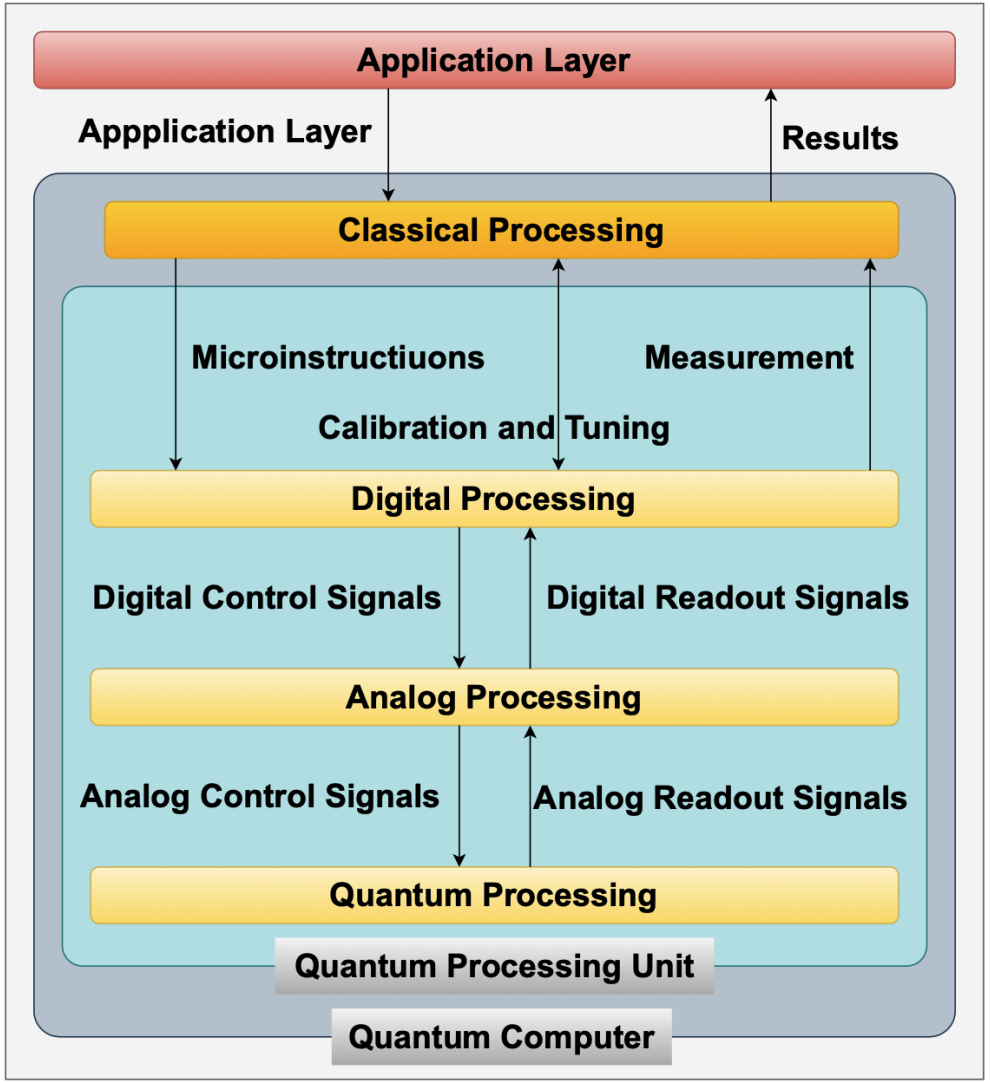
MalaQ 요약

- QPU를 제어하는 CPU를 공격하여 공격자가 QPU에 접근할 수 있도록함
- QPU를 직접 공격하는 것이 아닌, QPU를 접근할 수 있는 백도어 제공
- MalaQ를 통해 피해자 CPU를 통해 QPU에 접근하고 제어
  - **Phase Attack**
    - **하드웨어 및 소프트웨어 접근을 통해 공격가능**
    - FPGA 및 DDS에서 사용하는 Phase 관련 스크립트 수정
  - **Amplitude Attack**
    - **하드웨어 접근을 통해 공격가능**
    - 공격자가 QT의 ‘signal & Slot’ 메커니즘을 공격
    - 사용자가 설정한 진폭 값이 백엔드에 전달되기 전에 값 수정 (하드웨어에서 실제 정보를 왜곡)

- 향후 연구 주제?: 악의적으로 Quantum measurement 암호화하는 방식을 통해 양자 랜섬웨어를 개발 할 수 있다는 가능성을 제시함

+부채널 공격

Timing-SCAs[6]



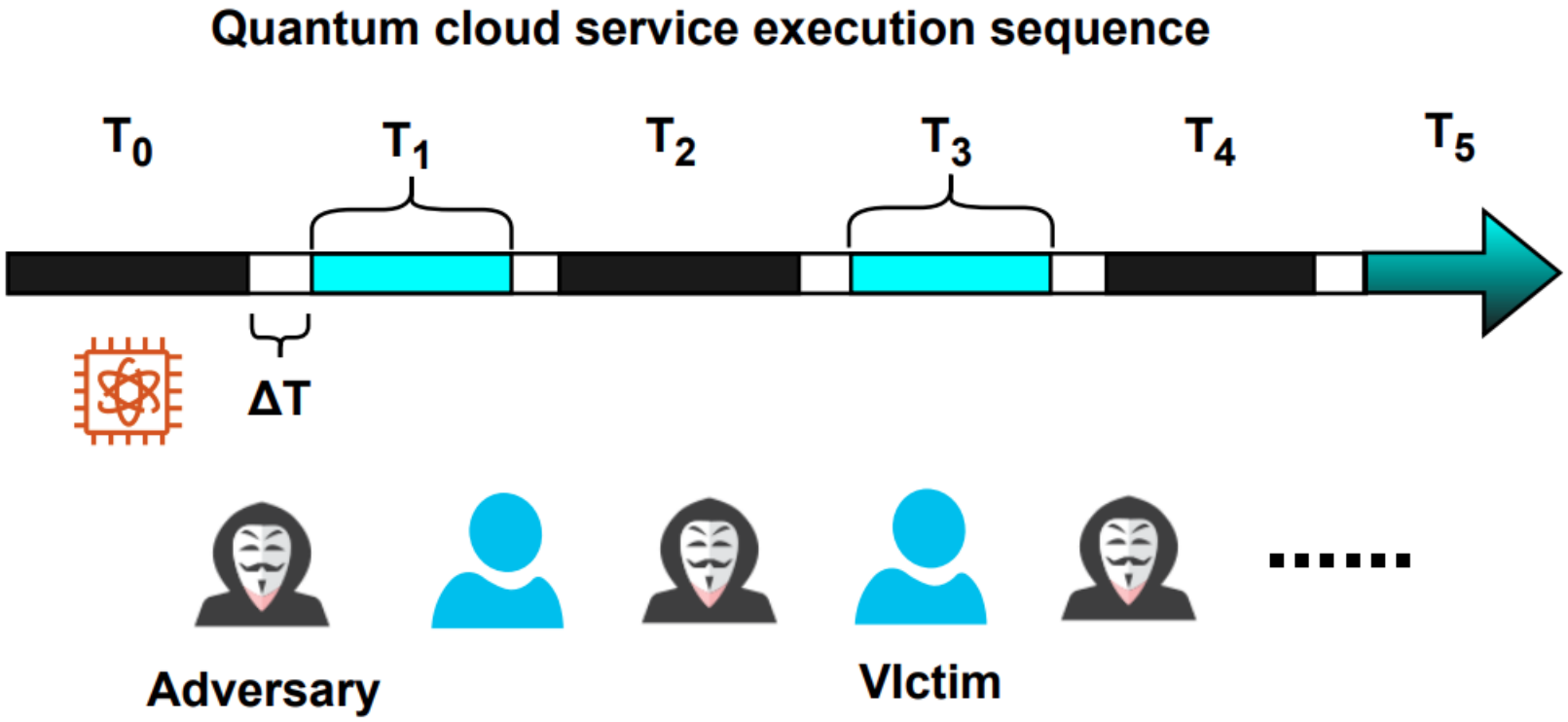
IBM 클라우드 양자컴퓨터 실행

- **Application Layer (응용 계층)**
  - **Microsoft Azure, Amazon Braket 같은 플랫폼에서 사용자가 상호작용하는 부분**
  - **양자 컴퓨터에서 실행할 알고리즘 설계 및 확인**
- Classical Processing (클래식 처리)
  - 양자 컴퓨터는 고전적인(클래식) 컴퓨터와 상호작용하며 작동
  - 양자 알고리즘이 실행되기 전에, 클래식 컴퓨터는 데이터를 준비하고, 양자 컴퓨터에 전달할 명령을 생성 (classic data → quantum data)
  - 양자컴퓨터에서 확률적으로 나타난 결과를 해석 및 변환 (quantum data → classic data)
- Quantum Processing Unit (양자 처리 장치)
  - 양자 컴퓨터의 물리적인 하드웨어로, 큐비트들이 존재하며 연산이 수행되는 곳
  - 극저온 환경에서 작동하며 큐비트들이 외부 환경으로부터 보호받음
- Quantum Computer (양자컴퓨터)
  - 양자 처리 장치와 고전적인 컴퓨터를 통합한 전체 시스템

[6] Lu, Chao, et al. "Quantum Leak: Timing Side-Channel Attacks on Cloud-Based Quantum Services." *arXiv preprint arXiv:2401.01521* (2024).

Timing-SCAs[6]

- 클라우드 컴퓨팅 환경에서의 timing SCAs 공격 수행
  - 이를 수행하기 위해서는 몇몇 가정된 환경이 필요함 (해당 논문에서는 이전 연구들에 비해 가정이 줄었으나 여전히 가정들이 필요함)



- **세가지 주체: 양자 클라우드 서비스, 양자회로 실행자(피해자), 공격자**

Timing-SCAs[6]

- 가정 및 방법론 – (가정이 많다..)

1. 반복실행: 피해자가 동일한 작업을 수행하는 회로 여러번 실행
2. 시간 측정: 공격은 각 회로 실행의 시간을 측정할 수 있음
3. 교차 실행: 공격자 회로는 피해자 회로의 두 번의 실행 사이에 실행될 수 있음
4. 전력 접근 불필요: 공격자는 각 회로 실행의 전력이나 에너지 속성, 또는 임의 파형 발생기나 믹서에 전력 및 에너지 흔적을 수집할 수 있는 큐비트 드라이브 장비에 접근할 필요가 없음
5. 냉각 시간: 연속 실행이 양자회로의 전체 시간 소비에 영향을 미치지 않는다고 가정, 시스템이 냉각하는데 시간이 더 걸릴 수 있음
6. 고정된 컴파일 전략: 컴파일 전략의 무작위성을 제거해야 함. 즉, 양자회로의 컴파일마다 큐비트 매핑 및 라우팅 경로가 고정되어야 함
7. 종합적인 시간 SCA: 시간 SCA는 디지털 처리, 아날로그 처리 및 양자 처리 단계를 포함한 전체 펄스 웨이브 준비단계를 포함함 (양자컴퓨터 실행 그림에 “Quantum Processing Unit”)

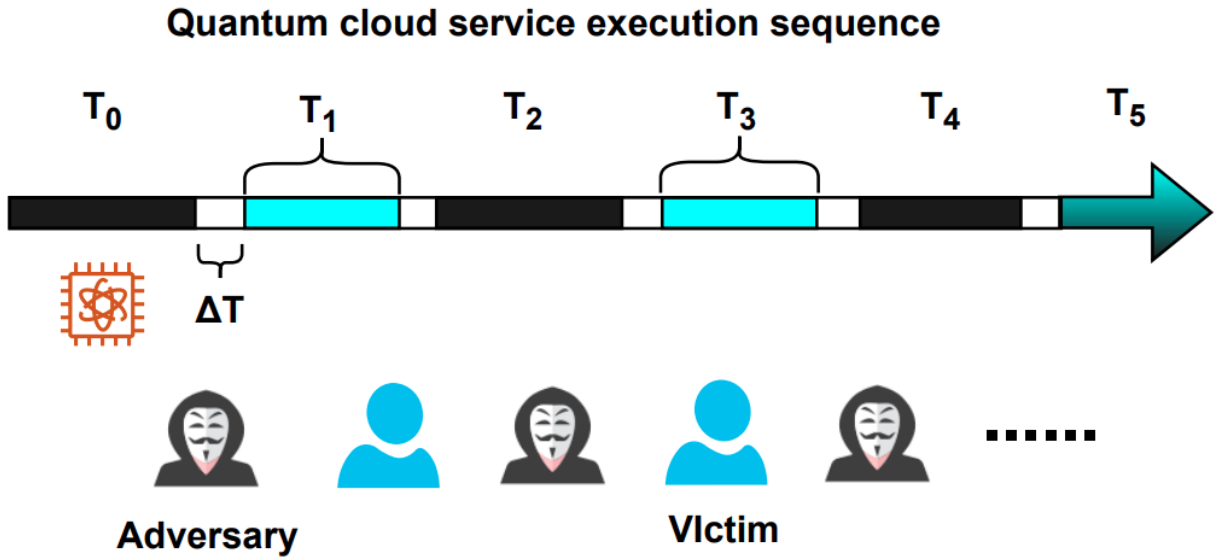
Timing-SCAs[6]

- timing SCAs 공격방식
  - 피해자 회로 동작 사이에 공격자 회로 제출
  - ΔT 값을 누적하여 피해자 회로의 동작 시간 동작 수 추정

  (방법 1. 공격자 회로 사이의 시간들 파악 (ΔT 값 이용) → 피해자 회로 동작 시간 추정)

  (방법 2. 피해자 회로 동작 시간을 통해 공격자 회로 사이의 시간에 따라 동작 횟수 파악)

  - 주된 시간 소모가 양자 프로세서 단계에서 발생한다고 여기며 다른 고전적 요소는 시간 차가 미미하여 양자회로 실행에 비해 거의 일정한 시간소비로 간주 될 수 있음



Timing-SCAs

- **공격 기법 (부채널) – 물리적 접근 x**

- **User Circuit Identification (UC), 사용자 회로 식별**: 클라우드에서 실행 중인 양자 회로가 무엇인지 식별하는 데 사용
- **Circuit Oracle Identification (CO), 회로 오라클 식별**: 특정 양자 오라클 식별, 양자 오라클은 Grover algorithm에서 중요한 역할을 하는 구성 요소
- **Circuits Ansatz Identification (CA), 회로 앤자츠 식별**: 특정한 형태의 양자 앤자츠(quantum ansatz)를 기반으로, 이 공격은 회로 앤자츠에서 사용된 매개변수를 식별. 앤자츠는 양자 회로 설계 시 초기 상태나 구조를 정의하는 데 사용.
- **Qubit Mapping Identification (QM), 큐비트 매핑 식별**: 알려진 양자 회로를 기반으로, 회로에서 사용된 큐비트 라우팅(큐비트 간의 연결 방식)을 식별
- **Quantum Processor Identification (QP)**, **양자 프로세서 식별**: 주어진 양자 회로를 기반으로, 해당 회로가 실행된 기저 양자 하드웨어(즉, 양자 프로세서)를 식별

Q & A