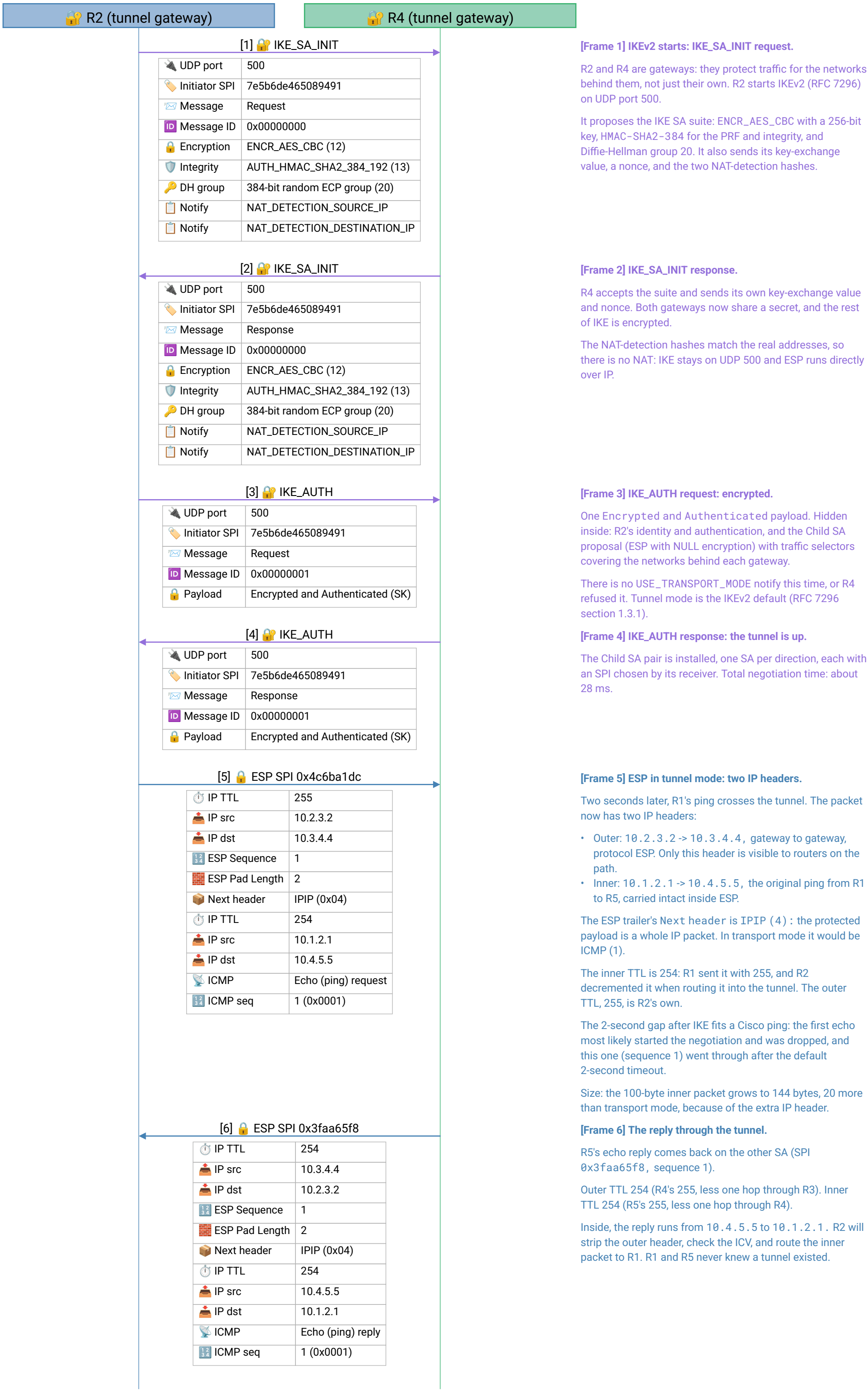
ipsec-esp-tunnel-mode.annotated



[Frame 1] IKEv2 starts: IKE_SA_INIT request.

R2 and R4 are gateways: they protect traffic for the networks behind them, not just their own. R2 starts IKEv2 (RFC 7296) on UDP port 500.

It proposes the IKE SA suite: `ENCR_AES_CBC` with a 256-bit key, `HMAC-SHA2-384` for the PRF and integrity, and Diffie-Hellman group 20. It also sends its key-exchange value, a nonce, and the two NAT-detection hashes.

[Frame 2] IKE_SA_INIT response.

R4 accepts the suite and sends its own key-exchange value and nonce. Both gateways now share a secret, and the rest of IKE is encrypted.

The NAT-detection hashes match the real addresses, so there is no NAT: IKE stays on UDP 500 and ESP runs directly over IP.

[Frame 3] IKE_AUTH request: encrypted.

One `Encrypted and Authenticated` payload. Hidden inside: R2's identity and authentication, and the Child SA proposal (ESP with NULL encryption) with traffic selectors covering the networks behind each gateway.

There is no `USE_TRANSPORT_MODE` notify this time, or R4 refused it. Tunnel mode is the IKEv2 default (RFC 7296 section 1.3.1).

[Frame 4] IKE_AUTH response: the tunnel is up.

The Child SA pair is installed, one SA per direction, each with an SPI chosen by its receiver. Total negotiation time: about 28 ms.

[Frame 5] ESP in tunnel mode: two IP headers.

Two seconds later, R1's ping crosses the tunnel. The packet now has two IP headers:

- Outer: `10.2.3.2` -> `10.3.4.4`, gateway to gateway, protocol ESP. Only this header is visible to routers on the path.
- Inner: `10.1.2.1` -> `10.4.5.5`, the original ping from R1 to R5, carried intact inside ESP.

The ESP trailer's `Next header` is `IPIP (4)`: the protected payload is a whole IP packet. In transport mode it would be ICMP (1).

The inner TTL is 254: R1 sent it with 255, and R2 decremented it when routing it into the tunnel. The outer TTL, 255, is R2's own.

The 2-second gap after IKE fits a Cisco ping: the first echo most likely started the negotiation and was dropped, and this one (sequence 1) went through after the default 2-second timeout.

Size: the 100-byte inner packet grows to 144 bytes, 20 more than transport mode, because of the extra IP header.

[Frame 6] The reply through the tunnel.

R5's echo reply comes back on the other SA (SPI `0x3faa65f8`, sequence 1).

Outer TTL 254 (R4's 255, less one hop through R3). Inner TTL 254 (R5's 255, less one hop through R4).

Inside, the reply runs from `10.4.5.5` to `10.1.2.1`. R2 will strip the outer header, check the ICV, and route the inner packet to R1. R1 and R5 never knew a tunnel existed.