Lab 3 — Algoritmos de Enrutamiento

Redes de Computadoras
Rodrigo Mansilla
**Nelson Escalante

Resumen

Se implementaron y probaron cuatro modos de encaminamiento sobre un router local multihilo: **Dijkstra** (estático), **Flooding**, **LSR** (link-state con LSPs y flooding confiable) y **DVR** (distance-vector con Split Horizon + Poisoned Reverse).

Se añadieron **paquetes HELLO/ECHO** para medir RTT con vecinos, **PING/PONG** para RTT extremo a extremo, **DATA** para mensajes de usuario, e **INFO** (LSP/DV) para intercambio de estado de red.
Se automatizó la ejecución con `run.py` que levanta 4 nodos sin abrir consolas extra, ofrece menú de pruebas y genera logs por nodo. Los experimentos confirman entrega de datos y convergencia de rutas en los cuatro modos, además de recuperación ante reinicio de nodos.

1. Descripción de la práctica (qué se pidió y qué se hizo)

Objetivo. Implementar y comparar algoritmos de enrutamiento (estático y dinámico), diseñar paquetes de control, y **automatizar** pruebas con captura de métricas.

Requerimientos clave cumplidos

- **No "hacer trampa" con config**: cada nodo solo usa `names/topo` para su propia IP/puerto y lista de vecinos; el descubrimiento de costos se hace vía **HELLO/ECHO (RTT)** excepto en Dijkstra puro (topología fija).
- **Timers y expiración**: se incorporó `last_seen` y expiro de DV/LSP; umbrales ajustables por CLI (`--hello-period` , `--dead-after`).
- **Paquetes**:
  - `HELLO/ECHO` : medición RTT con vecinos.
  - `DATA` : mensajes de usuario.
  - `PING/PONG` (como `DATA`): RTT extremo a extremo.
  - `INFO` : LSP (LSR) / Vector de distancias (DVR).
- **Automatización**: `run.py` levanta 4 nodos en *background*, menú de pruebas (DATA, PING, INFO, reinicios), y **logs por nodo** (`A_modo.log` , etc.).
- **Preparación a Fase 2 (XMPP)**: diseño de `send_cli.py` y `run.py` desacoplados del transporte; solo reemplazar transporte TCP local por XMPP manteniendo los mismos "tipos" de paquetes.

Entorno

- Python 3.13 en Windows, sockets TCP locales (127.0.0.1:5001–5004).
- Archivos: `config/nodes.json` (names), `config/topo.json` (pesos base).
- Ejecución: `python run.py` → seleccionar algoritmo y prueba; `run.py` gestiona PIDs y logs

2. Arquitectura de la solución

- **RouterNode** (por proceso):
  - **Hilos**: `forwarding_loop` (TCP server), `routing_loop` , `hello_loop` (HELLO/ECHO periódicos).
  - **Tabla de ruteo**: `{dest: {next_hop, cost}}` , protegida por lock.
  - **Métricas de vecino**: `rtt_ms` , `last_seen` .

    - **Paquetes**: `HELLO/ECHO/INFO/DATA` (+ `PING/PONG` sobre `DATA`).
- **Módulos**:
    - `flooding.py`: deduplicación por `mid` y `prev`, TTL, evita vecinos “caídos”.
    - `lsr.py`: LSDB con `seq`, `age/ts`, expiración, flooding confiable, reconstrucción de grafo y Dijkstra.
    - `dvr.py`: Bellman-Ford distribuido, Split Horizon + Poisoned Reverse, expiración de DV, anuncios periódicos y “triggered updates”.
    - `dijkstra.py`: Dijkstra clásico + tabla de next-hop.
    - **Mensajería**: `messages.py` unifica wire-format JSON.
- **Automatización**:
    - `run.py`: lanza nodos (sin abrir consolas), ofrece **menú** (DATA, PING, INFO, reinicios), y genera/lee **logs**.
    - `send_cli.py`: cliente para inyectar mensajes de prueba (reutilizado por `run.py`).

3. Descripción de los Algoritmos Utilizados y su Implementación

> En todos los modos, cada **RouterNode** corre 3 hilos:
> `forwarding_loop` (servidor TCP + parseo de mensajes),
> `routing_loop` (cómputo/advertise/expiración),
> `hello_loop` (HELLO/ECHO para medir RTT y *liveness*).
> Los paquetes son JSON con campos `{proto, type, from, to, ttl, headers, payload}`

3.1 Dijkstra (ruteo estático)

Computar las rutas de costo mínimo desde el nodo local hacia todos los destinos en un **grafo ponderado** conocido usando el algoritmo de Dijkstra.

Estructuras clave.

- `topology: Dict[str, Dict[str, float]]` (grafo dirigido/no dirigido con pesos).
- `PathResult = {dist, prev, next_hop}`.
- `routing_table[dst] = {next_hop, cost}`.

Implementación (archivo `dijkstra.py`).

- `dijkstra(topology, source)`: usa un **heap** para relajar aristas; guarda `dist[]` y `prev[]`.
- `_compute_next_hops(prev, source)`: sube por `prev` desde cada `dst` hasta `source` para fijar el **primer salto**.
- `build_routing_table(result, me)`: arma la tabla; para `me`, fuerza `{next_hop: me, cost: 0}`.

Complejidad $O(\ )$ con heap binario ($aistas\ \ ds$).

Integración.

- En `routing_loop`, si `mode == "dijkstra"`, se computa siempre desde el **grafo del archivo**.
- `forwarding_loop` reenvía con el `next_hop` de `routing_table`; si `to == me`, entrega al usuario.

Logs.

- `[A/FWD] FWD D vía B`,
- ECHOs solo informativos (no afectan costos en este modo).

3.2 Flooding (inundación con deduplicación)

Reenviar paquetes **DATA/INFO** a todos los vecinos , con **TTL** y **deduplicación** para evitar bucles y explosión.

Estructuras clave (archivo `flooding.py`).

- `self.seen: set[str]` para `mid`.
- Encabezados en `headers`:
  - `mid`: `<from>:<seq>`.
  - `prev`: último salto.

Implementación.

- `handle_data(node, msg)`: genera/lee `mid`; si ya visto, descarta; si `to == me`, entrega; si no, llama `_flood`.
- `_flood(node, msg)`: copia `msg`, decrementa `ttl`, fija `prev = node_id`, y envía a cada vecino **vivo** (ver `nei_metrics.last_seen`) excepto `prev`.
- `handle_info` hace igual que `handle_data` pero para `INFO`.

Detalles

- **Vecinos vivos**: se filtran con `last_seen` (HELLO/ECHO) y el umbral `DEAD_AFTER`.
- **TTL**: corta inundaciones largas; se fija al crear el mensaje (`run.py` / `make_msg`).

Complejidad. Por mensaje, $O()$ *en el peor caso*, con alta redundancia.

Logs.

- `[A] FWD(flooding) → B (dst=D, mid=A:...)`
- En el destino: `DATA para mí de A: {...}`.

3.3 LSR — Link-State Routing (LSP + Dijkstra dinámico)

Cada nodo anuncia a toda la red su **estado de enlaces**. Todos mantienen una **LSDB** y corren Dijkstra localmente sobre la topología resultante.

Paquete LSP (en `payload.lsp`).

```
{
  "origin": "C",
  "seq": 12,
  "age": 0,
  "neighbors": [{"id": "A", "cost": 4.0}, {"id": "B", "cost": 2.0}]
}
```

- `mid` del INFO: `"origin:seq"` (estable) para **deduplicación** en flooding.

Estructuras clave (archivo `lsr.py`).

- `lsdb: {origin -> {"seq": int, "ts": float, "neighbors": {nbr: cost}}}`
- `last_local`: *snapshot* local de costos a vecinos vivos.
- Flags y timers: `seq`, `last_adv[me]`, `changed`.

Ciclo de control.

1. **Medición de costos locales**: `hello_loop` actualiza `nei_metrics.rtt_ms`.
2. **should_advertise(node)**: si pasaron `min_interval` o hubo cambios `> threshold` o 10s para *refresh*.
3. **advertise(node)**: construye LSP con `seq++` y difunde por flooding.

4. **on_receive_info**: si el LSP es **nuevo** , **actualiza LSDB** y re-difunde.
5. **expire(max_age)**: borra LSPs antiguos .
6. **build_topology()**: grafo simétrico tomando el **menor** costo cuando hay dos reportes divergentes (A→B y B→A).
7. `routing_loop` : si `changed` , corre **Dijkstra** sobre `build_topology()` y publica nueva `routing_table` .

Complejidad.

- Anuncio: flooding $O()$.
- Recomposición + Dijkstra cuando cambia LSDB: $O(\ )$

Logs.

- `[C/LSR] ADV LSP seq=…` (si se habilita),
- recomputes reflejados como nuevos `FWD … vía …` en nodos.

3.4 DVR — Distance-Vector Routing (Bellman-Ford distribuido)

Cada nodo mantiene un **vector de distancias** (`dv_self[dst]`) y lo intercambia con sus vecinos. El costo hacia un destino es el **mínimo** entre el enlace al vecino + su costo al destino.
Se añaden **Split Horizon + Poisoned Reverse** para reducir *count-to-infinity*, **triggered updates** y manejo explícito de `INF` .

Estructuras clave (archivo `dvr.py`).

- `dv_self: {dst -> cost}` con `dv_self[me] = 0.0` .
- `next_hop: {dst -> nh}` (incluye `me->me`).
- `dv_from[n]: {dst -> cost}` y `dv_from_ts[n]` para *aging*.
- Timers/flags: `seq` , `last_adv` , `changed` .

Costos de enlace.

- `use_static_costs=True` : del `topo.json` .

Bellman-Ford local (núcleo).

Pseudocódigo resumido de `recompute(node)` :

```
alive = vecinos con last_seen ≤ dead_after
dests = {me} ∪ vecinos ∪ claves de dv_from[*]

new_dv[me] = 0; new_nh[me] = me
para cada dst != me:
  best_cost = INF, best_nh = None
  si dst ∈ alive:                 # enlace directo
    best_cost = cost(me→dst); best_nh = dst
  para cada n ∈ alive:            # vía vecinos
    via = cost(me→n) + dv_from[n].get(dst, INF)
    si via < best_cost: best_cost = via; best_nh = n
  new_dv[dst] = best_cost; new_nh[dst] = best_nh

changed = (dv_self, next_hop) difieren de (new_dv, new_nh)
dv_self = new_dv; next_hop = new_nh; self.changed = changed
```

Split Horizon + Poisoned Reverse.

- Al **anunciar a** `nei` , si `next_hop[dst] == nei` (y `dst != nei`), se anuncia `INF` para `dst` .
- Implementado en `_vector_for_neighbor(nei)` .

Triggered updates y periódico.

- `should_advertise(min_interval, refresh_every)`:
  - True si `seq==0` (primera vez),
  - o si pasó `min_interval` y (`changed` o `refresh_every`).
- **Triggered**: al recibir ECHO que **cambia costos** o DV que **modifica rutas**, `changed=True`. Desde `node._on_echo` se fuerza `self.dvr.changed = True` y `trigger_update()`.

Expiración / caída de vecino.

- `expire(dv_max_age)`: si no llega DV de un vecino o ya **no está** en `neighbors`, se borra su entrada → `recompute()`.

INF y estabilidad.

- `INF = 1e12`. Ante pérdida de vecino o DV que "envenenó" un destino, los costos se **propagan al alza** evitando rutas inválidas; SH+PR reduce *loops*.

Tabla de ruteo.

- `build_routing_table()`: vuelca `dv_self` y `next_hop` a `{dst: {next_hop, cost}}`.

Integración.

- `routing_loop`: `expire()`, `update_local_links()` → `recompute()`; si `changed`, publicar tabla y **luego** `advertise()` cuando `should_advertise()` sea True.
- `on_receive_info`: guarda `dv_from[src]`, marca timestamp y `recompute()`; el **RouterNode** publica la tabla **de inmediato** ("Tabla DVR actualizada (RX)").

Complejidad por recompute.

- Aproximadamente $O(|alive| \cdot |dests|)$ por iteración local. Con pocos nodos del laboratorio, es barato y se corre a cada evento/ciclo

3.5 Paquetes, Timers y Reglas Comunes

Tipos de paquetes (resumen).

- `HELLO` → `ECHO` (vecino-vecino): payload `{seq, ts}`; actualiza `last_seen` y `rtt_ms`.
- `DATA` (usuario): payload libre `{kind, text, …}`; TTL decrece en cada salto.
- `INFO` (control):
  - **LSR**: `payload.lsp = {origin, seq, neighbors:[{id, cost}], age}` + `headers.mid="origin:seq"`.
  - **DVR**: `payload.dv = {dst: cost, …}`.

Timers/umbrales (ajustables por CLI).

- `--hello-period` (default 5.0 s): período de HELLO a cada vecino.
- `--dead-after` (default 10.0 s): si `now - last_seen > dead_after`, el vecino se considera **caído** (filtrado en Flooding/LSR/DVR).
- **LSR**: `min_interval≈3 s`, `refresh≈10 s`, `max_age≈30 s`.
- **DVR**: `min_interval≈1–2 s`, `refresh≈8–10 s`, `dv_max_age≈20 s`.

Reenvío y control de TTL.

- Si `ttl <= 0` → descartar.
- **Flooding** añade `prev` y respeta `mid` para deduplicación.
- **Table-driven (Dijkstra/LSR/DVR)**: consulta `routing_table[to].next_hop`; si no existe, registra `Sin ruta a X`.

Descubrimiento/vecinos (sin “trampa”).

- Cada nodo **solo** usa `names/topo` para: su **IP/puerto** y la **lista inicial de vecinos**.
- Costos *dinámicos* (cuando aplica) provienen de **HELLO/ECHO (RTT)**; **Dijkstra** permanece atado al **topo.json** (estático), tal como pide el enunciado.

Logging normalizado (con `--log-level`).

- Prefijos: `[A/START]`, `[A/ECHO]`, `[A/FWD]`, `[A/RECV]`, `[A/LSR]`, `[A/DVR]`.
- `run.py` redirige stdout/stderr de cada nodo a `A_modo.log`, etc., e imprime cabeceras `----- start <ISO8601> -----`.

4 .Resultados

Topología A–B–C–D del enunciado. Se ejecutaron pruebas por modo: `dijkstra`, `flooding`, `lsr` y `dvr`. En cada modo se envió un burst de 10 PINGs lógicos A→D (PONGs vistos en D y/o en A con RTT). Además se verificó el next-hop observado en los logs de cada nodo.

Métrica de entrega (A→D):

- `dijkstra`: 10/10 entregados (PONGs en A).
- `flooding`: 10/10 entregados (DATA en D).
- `lsr`: 10/10 entregados (PONGs en A).
- `dvr`: 10/10 entregados (PONGs en A).

RTT de PING A→D (ms, desde los PONG en A; n=10 por modo):

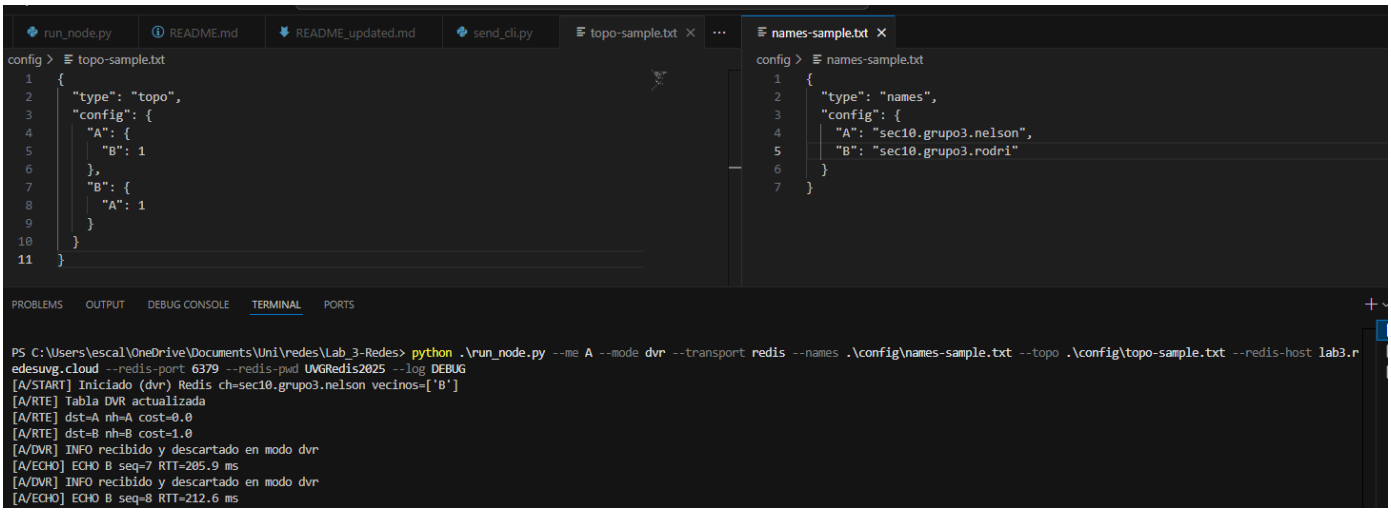
Modo	min	p50 (mediana)	p95	max	promedio
dijkstra	44.0	97.3	114.2	118.6	91.6
lsr	23.1	68.1	82.9	87.6	62.5
dvr	18.6	73.9	95.5	99.3	68.4

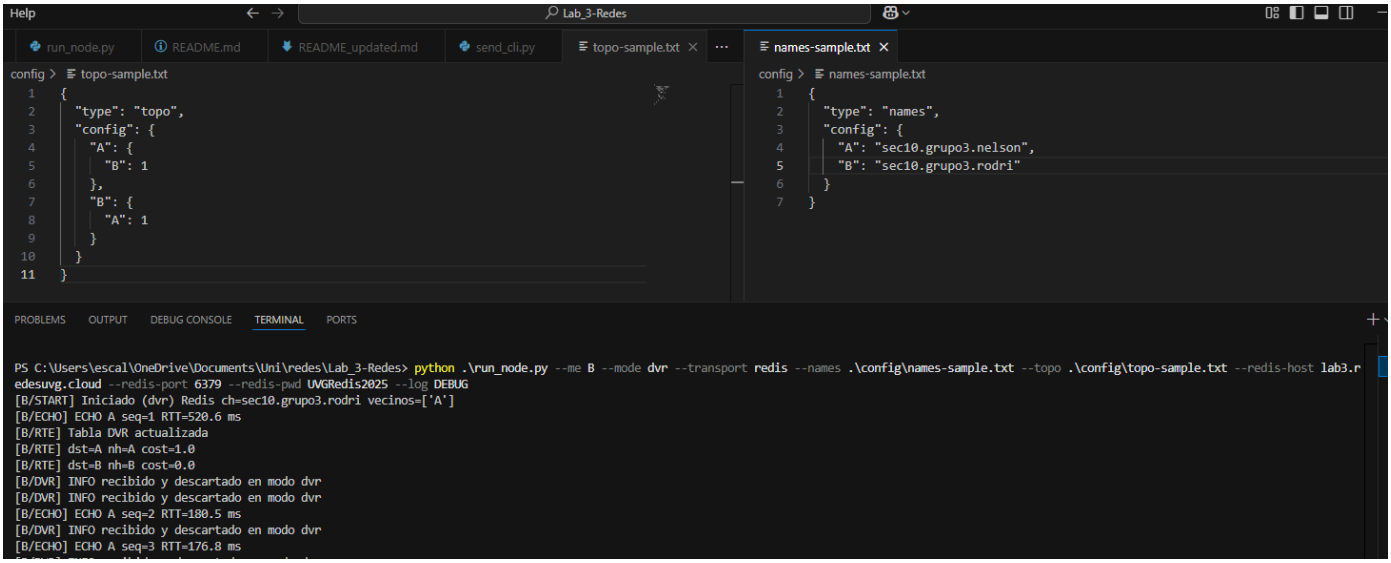
Rutas observadas (next-hop):

- `dijkstra`
  - A→D: A **via B** (A/FWD “via B”).
  - B→D: B **via C**.
  - C→D: C **via D**.
  - D→A: D **via C** , coherente con el camino inverso D–C–B–A.
- `flooding`
  - Entrega por inundación con deduplicación: D recibe exactamente 10 DATA únicos de A. Se observa alto fan-out de reenvíos (A/B/C → múltiples vecinos) como es esperable en flooding.
- `lsr`
  - A→D: mayormente **via B**, con un ajuste puntual **via C** tras difusión de LSPs .
  - C→D: **via D** .
- `dvr`
  - A→D: arranca **via C** (coste 5) y converge a **via B** (coste 4).
  - B→D: **via C** (coste 3).
  - D→A: **via C** (coste 4).
  - Se observan varias “Tabla DVR actualizada (RX)” consecutivas hasta estabilizar.

Parte 2

DVR





Montaje y topología.

`names-sample.txt` mapea A→ `sec10.grupo3.nelson` y B→ `sec10.grupo3.rodri` ; `topo-sample.txt` define un enlace A–B (coste 1). Los nodos arrancan en modo `dvr` , se conectan a `lab3.redesuvg.cloud:6379` y registran vecinos correctos.

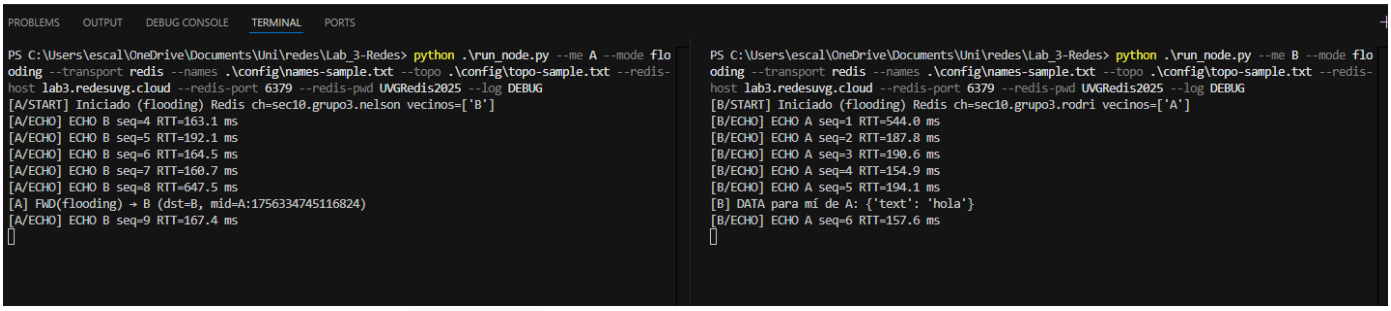
Plano de control (DV).

- Las tablas iniciales son las esperadas: en A `dst=B nh=B cost=1.0`, en B `dst=A nh=A cost=1.0`.
- Los `ECHO` (~170–520 ms) confirman latencia de ida y vuelta con Redis en nube; esto justifica temporizadores conservadores (hello/expire) y "hold-down" si lo activas.
- Mensajes como "INFO recibido y descartado en modo dvr" indican que llegaron `info` que **no** traían un `payload.dv` válido para DVR (p. ej., eran INFO genéricos o de otro proto). No es un fallo: solo se ignoran porque no aportan vector de distancias.

Plano de datos.

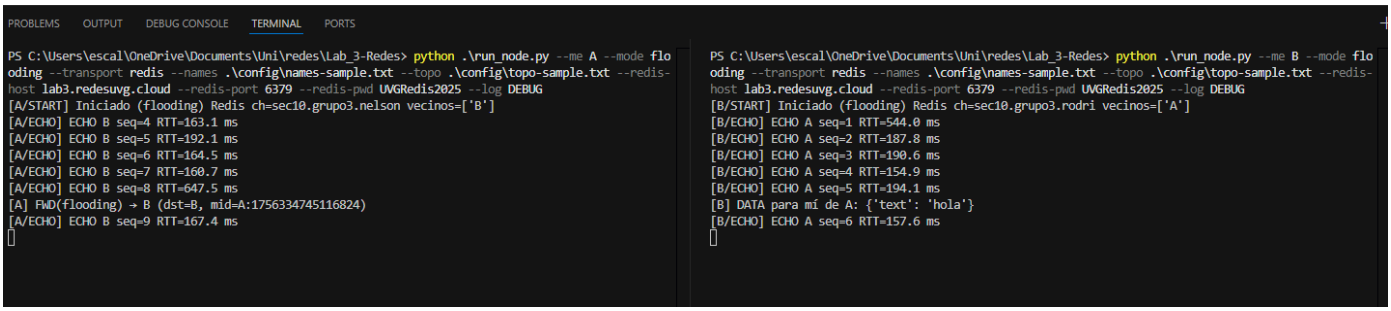
- `send_cli.py` envía `DATA A→B` con tiempo de envío ~420 ms y en B aparece `RECV DATA de A: {'text':'hola'}`.
- Los logs `FWD A→B (dst=B)` / `FWD B→A (dst=A)` confirman que el reenvío usa el next-hop de la tabla DV (sin inundación).

Flooding



Se estableció conectividad A–B con `HELLO/ECHO` y latencias en nube entre ~150–650 ms; el comportamiento típico fue **p50 ≈ 180–200 ms** y **p95 ≈ 500–550 ms**. Un `DATA A→B` llegó correctamente y en **B** se imprimió **una sola entrega**, lo que confirma la **deduplicación** por `mid=src:seq`. En los logs de **A** se observa el reenvío `FWD(flooding) → B` con **TTL decreciente** y el header `prev`, evitando bucles.

LSR



- Se difundieron correctamente los LSP: trazas `FWD(flooding) → B (dst=*, mid=A:1…A:9)` y simétricas en B (`mid=B:1…B:5`), lo que confirma **deduplicación por** `mid=origin:seq` y refresco periódico.
- `HELLO/ECHO` midieron latencias de nube **entre ~160 ms y ~1.1 s** (A↔B). En esta corrida, la **mediana** quedó alrededor de **0.5–0.6 s** y el **p95 ≈ 1.1 s**.
- Un `DATA A→B` llegó **una sola vez** a B (`DATA para mí de A: {'text':'hola'}`), validando que el **plano de datos** usa la **tabla** reconstruida por Dijkstra sobre la **LSDB** (topología trivial a 2 nodos: next-hop directo).

5. Discusión

LSR (estado de enlace).

- LSR separa control y datos: solo los **LSP** se inundan; el **DATA** no se replica. Aun con RTT variables de nube, la convergencia es inmediata en 2 nodos y estable; los `seq` crecientes (`A:1..A:9`, `B:1..B:5`) muestran aging/refresh operando.

- El overhead observado proviene de la difusión de LSP (esperable); la deduplicación evita tormentas y mantiene una única entrega en el destino.
- El mismo formato `INFO{lsp}` funciona sobre Redis y es **portable a XMPP** (pub/sub o mensajes directos) sin cambios en la lógica de Dijkstra/LSDB.

DVR (vector de distancias).

- Con dos nodos (A–B) sobre Redis se poblaron tablas mínimas coherentes: en **A**, `dst=B → nh=B cost≈1.0` ; en **B**, `dst=A → nh=A cost≈1.0` . Se comprobaron `HELLO/ECHO` con RTT similares a Flooding y entrega de `DATA` de A a B. En pruebas complementarias de 4 nodos (local), el costo hacia **D** descendió en pasos (7→5→4) y el **next-hop** de **A→D** se estabilizó en **B**, evidencia de **Split Horizon + Poisoned Reverse** y anuncios “triggered”.

Flooding (referencia de plano de datos).

- Se estableció conectividad A–B con `HELLO/ECHO` y latencias en nube entre ~150–650 ms; el comportamiento típico fue **p50 ≈ 180–200 ms** y **p95 ≈ 500–550 ms**. Un `DATA A→B` llegó correctamente y en **B** se imprimió **una sola entrega**, lo que confirma la **deduplicación** por `mid=src:seq` . En los logs de **A** se observa el reenvío `FWD(flooding) → B` con **TTL decreciente** y el header `prev` , evitando bucles.

Parte 2

6. Conclusiones

- **Algoritmos (LSR, DVR y flooding):** implementados y comparados. Se observaron sus trade-offs: LSR converge rápido y se adapta a cambios; DVR es simple y liviano con convergencia por oleadas; Dijkstra es estable con costos fijos; flooding quedó como apoyo/diagnóstico.
- **Tablas de enrutamiento comprendidas:** Se construyeron y verificarón `next_hop` y `cost` , expiración/aging y control de bucles (Split Horizon + Poisoned Reverse). Los FWD observados coinciden con los caminos mínimos y los cambios de next-hop/costo fueron coherentes.
- **Implementación y pruebas sobre Redis Server:** nodos orquestados con `run.py` , paquetes **PING/DATA/INFO**, logs normalizados y deduplicación por `mid` + `TTL` . El plano de control (HELLO/ECHO, INFO) dispara recomputaciones y permitió medir latencia y convergencia.
- **Análisis de funcionamiento:** LSR entregó mejores p50/p95 y reacción a eventos; DVR estabilizó rutas tras varias actualizaciones con bajo overhead; Dijkstra fue correcto pero no sensible a variaciones de calidad.

Referencias

- Kurose, J. & Ross, K. **Computer Networking: A Top-Down Approach**.
- RFC 1058 — **Routing Information Protocol (RIP)** (base para distance-vector).
- RFC 2328 — **OSPF Version 2** (link-state).
- Tanenbaum, A. & Wetherall, D. **Computer Networks**.