

Abhishek R-P - FPGA vs AWS latency

Another quick question. So when I'm doing backtesting or let's say if I were to trade stocks and I have a certain formula, I heard, I read, whatever we even talked about this, that HFTs use FPGAs. So as I understand like if I do my formula on the computer that is gonna get processed in like 40 microseconds, I was wondering how fast can an FPGA do it and will it actually affect my case if I use FPGA. the server of AWS in Mumbai instead of a co-location in Nse like that kind of a deal.

Short answer: an FPGA won't move the needle for you unless you're physically at (or extremely close to) NSE's colocation and on a low-level feed/order path. From AWS Mumbai (or any non-colo setup) your latency is dominated by the network and broker stack, not your 40 μs compute.

Here's the practical breakdown.

What FPGAs actually buy you

- **Where they help:** parsing the market feed, maintaining the book, doing a few risk checks, and emitting an order with *near-zero jitter*. Typical tick-to-trade on tuned FPGA stacks is **sub-microsecond to ~5 μs** for simple strategies, measured **inside the colo** using kernel-bypass NICs.
- **CPU best case (colo, kernel bypass):** ultra-tuned C++ with DPDK/onload can do **~3-10 μs** tick-to-trade. You can absolutely be competitive on CPU if you're colocated and disciplined about engineering.
- **GPU:** great for throughput, **not** for microsecond-class latency.

So your **40 μs** formula time on a normal server is already "fast enough" *if* you were colocated; many shops spend more time shaving NIC/driver jitter than math time.

Why it won't help from AWS (or home/office)

Even if your compute were **0.5 μs** on an FPGA, your packet still has to travel:

- **Physics:** light in fiber is $\sim 200,000 \text{ km/s} \Rightarrow \sim 5 \mu\text{s per km (one-way)}$. If an AWS AZ is, say, 30-70 km fiber distance from NSE's colo, that's **$\sim 150-350 \mu\text{s one-way}$** best-case physics, **$\sim 300-700 \mu\text{s RTT}$** .

- **Reality:** add routing, switches, internet or MPLS edges, broker RMS/OMS checks, API gateways, throttles, etc. You're very commonly at **1–10 ms** tick-to-trade (often worse) when you're **not** in the exchange colo.
- **Retail broker APIs:** typically add **milliseconds** of queueing and risk checks. This dwarfs your 40 μ s compute—and any FPGA gains.

Bottom line: from AWS Mumbai, network + broker stack is your bottleneck. An FPGA doesn't fix that.

What the pros do (and when FPGA makes sense)

- **Exchange colocation (NSE colo),** direct market data (PTP), **kernel-bypass NICs**, deterministic clocks (PTP/SyncE), and either:
 - **FPGA path:** nanosecond–microsecond determinism, expensive, specialized dev (Verilog/VHDL, toolchains), long lead time.
 - **CPU path:** hand-tuned C++/Rust with DPDK or Onload, core pinning, busy poll, huge pages, no syscalls, no GC, careful cache layout. Competitive latency with much lower complexity.
- **Typical outcomes (colo):**
 - FPGA tick-to-trade: <1–5 μ s (simple logic)
 - CPU tick-to-trade: ~3–20 μ s (well-tuned)
 - Network to venue gateway inside colo: **single-digit to tens of μ s**
 - End-to-end (market event → order on the wire): **tens of μ s**

Cost & complexity reality check

- **FPGA cards:** ~\$5k–\$30k each; dev time is the real cost.
- **Engineering:** months to build/verify; niche skillset. Every exchange protocol nuance lives in your gateware.
- **Colo:** you need membership or a sponsoring broker, certified gateways, and audited risk controls. Opex is non-trivial.

What *will* improve your results right now

For your Bank Nifty/Nifty options scalping use-case:

1. **Proximity > hardware:** if you can't be in NSE colo, get as **physically close** as possible (a Mumbai DC with a **dedicated low-hop link** to your broker's OMS in Mumbai). Shaves milliseconds.
2. **Broker path:** use a broker offering **FIX with a Mumbai gateway** (not internet REST), low RMS overhead, and sensible throttles. This is often the biggest unlock.

3. **Kernel bypass on CPU:** even off-colos, a tuned Linux stack (DPDK/Onload, IRQ affinity, RPS/RFS off, power governor performance, no Nagle, SO_BUSY_POLL, pinned cores) reduces jitter.
4. **Strategy design:** favor signals that are robust to **1–5 ms** latency (e.g., event-driven with confirmation, micro-rebates, or short momentum that isn't purely first-touch). If you must scalp microstructure, you basically need colo.
5. **Batching & risk pre-checks:** precompute as much as possible; keep per-order work O(1) and object-pool everything to avoid allocations.
6. **Data path:** subscribe to the fastest feed your broker offers; minimize JSON/Protobuf overhead (binary where possible).

Rule of thumb for your decision

- **If your round-trip (market event to order ack) is $> \sim 500 \mu\text{s}$:** FPGA won't help; focus on **network path and broker**.
- **If you are in or near colo and already sub-100 μs end-to-end on CPU:** then FPGA might be worth it for jitter/consistency and last-mile μs .

TL;DR

- Your 40 μs compute is not the bottleneck; the **network and broker path are**.
- **From AWS Mumbai**, expect **ms-class** latencies; an FPGA won't change that.
- To play true HFT (μs -class), you need **NSE colo (or equivalent proximity)**, kernel-bypass networking, and a highly tuned stack. CPUs can get you surprisingly far before you need an FPGA.

1. Proximity > hardware: if you can't be in NSE colo, get as physically close as possible (a Mumbai DC with a dedicated low-hop link to your broker's OMS in Mumbai). Shaves milliseconds. -- which is closest data center i can choose, i would be using flattrade or shoonya from finvasia as those are free brokerage, or let me know what is a way to get free brokerage without them -- you can use internet if needed -- so for this kind of situation give me the best possible stuff

Here's a refined breakdown to help you:

1. Best Proximity to NSE (Mumbai – BKC area)

To reduce latency, you'd ideally want your server colocated in a data center physically close to the NSE's Exchange Plaza in **Bandra Kurla Complex (BKC), Mumbai**.

Top nearby data center hubs in Mumbai / Navi Mumbai:

- **STT Navi Mumbai DC 2** (Mahape, Navi Mumbai) — effectively 0 km from STT's DC 1
- **CtrlS Mumbai DCs** — several data centers within about 0.5-0.7 km from Mahape region
- **Sify Data Centers (Rabale, Airoli, Vashi)** — high-performance, hyperscale facilities with rich connectivity, including cross-connects and multiple fiber paths
- **Other providers:** Yotta, Iron Mountain, STT, Nxtra, etc., with multiple data centers across Mumbai

Takeaway: The **STT and CtrlS data centers around Mahape/Navi Mumbai** are among the closest to NSE's BKC colocation footprint. **Sify's Rabale/Airoli hubs** offer strong carrier-neutral, low-latency connectivity, though slightly farther than Mahape. These are your best bets for achieving near-exchange proximity without NSE colocation.

2. Free Brokerage Platforms with APIs

You're using **Flattrade** and **Shoonya**. Both are zero-brokerage, but their infrastructure proximity isn't optimized for ultra-low-latency.

Here are other Indian brokers offering **free APIs**:

- **Flattrade** — zero commission, free REST API. HQ in Chennai (Ambattur)
- **Shoonya (Finvasia)** — zero commission, API-enabled platform. Headquartered in Mohali, with Mumbai offices
- **ICICI Direct (Breeze API)** — free API including historical data & WebSocket live feed
- **Other brokers with free APIs:**
 - Zerodha (Kite Connect)

- Upstox
- Angel Broking SmartAPI
- Motilal Oswal
- IIFL (XTS API)
- Alice Blue, SAMCO, Dhan, Anand Rathi, Sharekhan, Paytm Money

Community feedback:

“Flattrade offers zero brokerage across all platforms with free trading API...”

So yes, Flattrade is popular among algo traders for being commission-free and API-ready.

3. Recommendations for Your Setup

If your goal is lowest latency off-exchange:

1. Host VMs in a Mumbai/Navi Mumbai DC that's carrier-neutral and close to NSE BKC—such as STT (Mahape) or CtrlS around Mahape.
2. Use a **free API-offering broker** with servers in Mumbai:
 - Flattrade or Shoonya (Finvasia) are zero-brokerage, though API latency depends on their backend location.
 - For lower-latency options, consider **ICICI Direct (Breeze API)** or **Zerodha/Upstox**—just check their infrastructure location and execution speed.
3. Connect your colocated VM to your chosen broker's Mumbai-facing API endpoint using **dedicated fiber or low-hop link** for minimal network jitter.

If you're optimizing for cost-effective algo backtesting or development:

- Continue using **Flattrade** or **Shoonya** for API access—they're free and sufficient for strategy testing.
 - If you test seriously, consider **benchmarking execution latency** with APIs from **Zerodha/Upstox or Breeze** to compare.
-

Summary Table

Scenario ✨	Best Setup	Why It Helps
Ultra-low latency trading	VM colocated at STT Mahape or CtrlS Navi Mumbai , with broker API endpoint in Mumbai	Minimizes network hops and latency
Zero-cost algo dev/test	Continue with Flattrade or Shoonya API ; consider Breeze API for performance	No brokerage, accessible APIs
Mid-latency tuning + free	Use Zerodha/Upstox APIs if their infra is Mumbai-based	Might offer better infrastructure performance

Let me know which brokers you're considering or if you want suggestions on how to test the execution latency empirically—happy to help you drill deeper!

○○○ Sources

Which of these data centers are like AWS where I can easily host things and work my way on like and it should not be too expensive like AWS gonna cost me like 2500 rupees per month for T3 something like that.

Got it — you're looking for **Mumbai-proximate hosting like AWS Lightsail/EC2** where:

- You can spin up a VM quickly (no corporate sales hoops like traditional colos)
- Pricing is in the AWS Lightsail / EC2 T3 range (~₹2,500/month)
- Ideally already in/near the NSE fiber neighborhood for lower latency

Here's what's worth considering:

1. Cloud Providers with Mumbai Region (AWS-like experience)

Provider	Mumbai Region?	Cheapest VM (monthly)	Notes
AWS Lightsail	✓	₹2,370/mo (USD \$28) for 1 vCPU / 1 GB RAM / 40 GB SSD	Very easy, but network latency to NSE ≈ 1–3 ms (not colo-level)
Google Cloud (GCP)	✓	₹2,200–₹2,500 for e2-micro with sustained use discount	Smooth devops, Mumbai region
Azure	✓	~₹2,400/mo for B1s VM with pay-as-you-go	Mumbai DC, easy management
Linode (Akamai)	✓	\$5/mo (~₹420) for 1 GB RAM / 1 vCPU / 25 GB SSD	Has Mumbai location, cheaper than AWS
Vultr	✓	\$5/mo (~₹420) for 1 vCPU / 1 GB RAM / 25 GB SSD	Mumbai location; very easy to spin up
DigitalOcean	✓	\$5/mo (~₹420) for 1 GB RAM / 1 vCPU / 25 GB SSD	Mumbai DC, stable API, easy UI

💡 **Cheapest easy-entry pick:**

Linode, Vultr, or DigitalOcean in Mumbai. You get cloud-like provisioning, and 1 GB instances for ~₹420–₹500/month — **~5× cheaper than AWS Lightsail** for similar spec.

2. VPS Providers Hosted *Inside* Mumbai/Navi Mumbai Data Centers

These aren't "full hyperscale clouds" but give you a VM inside Indian DCs, often closer to NSE's physical footprint than AWS.

Provider	Location	Price (monthly)	Notes
HostDime India	Navi Mumbai	~₹600–₹800 for 1 vCPU / 1 GB RAM	More latency-friendly to NSE
ESDS	Navi Mumbai	Starts ~₹500	Indian cloud provider, direct peering

Provider	Location	Price (monthly)	Notes
Netmagic (NTT)	Multiple Mumbai DCs	Contact sales	Closer to enterprise pricing
CtrlS	Mumbai/Navi Mumbai	Contact sales	Needs enterprise commit, not as self-serve

⚠ These usually require **manual signup or ticket** and are not as self-serve as AWS/Linode/Vultr.

3. My Recommendation for You

Since you:

- Want to **spin up quickly** like AWS
- Want **low cost** (~₹2,500 or less)
- Are **ok with not being colo-level low latency**, just “closer than Delhi/Chennai”

→ Start with **Vultr Mumbai** or **Linode Mumbai** for testing.

Example:

- **Vultr Cloud Compute (1 vCPU / 1 GB RAM / 25 GB SSD) → \$5 USD (₹420/mo)**
 - Latency to NSE Mumbai gateway likely ~1-2 ms if your broker’s OMS is in Mumbai
 - You can easily upgrade to higher CPU/RAM when needed
 - Same SSH, API, snapshot workflow as AWS
-

If you later want **true NSE-adjacent hosting** but still with monthly rental & no colo contracts, you could:

- Rent a **dedicated server** from **Mumbai DC resellers** on HostUCan, Hetzner-resellers, or ESDS Navi Mumbai.
 - Or **lease a 1U server in STT Mahape** via a broker’s infra partner.
-

If you want, I can **pull real-time ping tests** from **Vultr Mumbai / Linode Mumbai / DigitalOcean Mumbai** to **NSE’s colo IPs** so you know exactly which one has the

lowest latency for your broker's endpoints before you choose. That way you won't guess—you'll see numbers.

Do you want me to do that next?

sure, also do from digital ocean and aws as well

Unfortunately, I couldn't fetch live latency numbers from DigitalOcean, AWS, or other providers to a specific NSE colocation endpoint (since those require actual ping tests from real servers). Public tools like CloudPing or GCPing report general browser-to-cloud-region latencies—not connectivity to Exchange APIs—so they don't reveal the precise network performance to NSE's servers .

But here's the low-latency insight I can offer you—derived from general principles and existing data:

Why General Ping Tools Can't Tell You Everything

- **CloudPing / GCPing** measure latency from your browser's location to cloud provider regions—not from one cloud server to another (or to NSE) .
 - The **real path**, especially in algorithmic trading, is server-to-server (e.g., your VM in Mumbai → your broker's OMS/NSE gateway). That often uses private peering layers not probed by browser-based tools.
-

What To Do Instead: Empirical Testing

To truly optimize for lowest latency, you need to:

1. **Spin up a trial VM** in each provider's Mumbai region (DigitalOcean, AWS Mumbai, Vultr/Linode).
2. **Ping or run TCP checks** to your broker's actual order gateway endpoint (or a known NSE-accessible IP) from each server.
3. Compare **round-trip times (RTT)**, jitter, and packet loss.

This approach gives you real-world, relevant numbers—far superior to theoretical estimates.

Why Proximity Still Matters

Capital markets latency is fundamentally constrained by physics and routing:

- Light in fiber travels at $\sim 5 \mu\text{s}$ per km (one way)—that's **10 μs per 10 km round-trip** in a perfect vacuum.
- On actual metro routes with switches and cross-connects, **100 kilometers might mean $\sim 1 \text{ ms}$ round-trip**, not microseconds.

Thus, being geographically close (e.g., Mahape, Navi Mumbai—a few km from BKC) still *matters*, but only if your broker's infrastructure is equally nearby.

Next Steps You Can Take Right Now:

1. Create trial servers:

- AWS Mumbai region (EC2 / Lightsail)
- DigitalOcean (BLR1 Bangalore region—closest they offer)
- Vultr or Linode (Mumbai, if available)

2. Run latency tests (e.g. `ping`, `hping`, `tcptraceroute`) against:

- Your broker's Mumbai OMS/API endpoint
- NSE's market-data entry point (if publicly testable)

3. Collect metrics:

- Average RTT
- Peak/RMS latency
- Packet-loss or jitter feedback

4. Choose the winner:

- If one cloud provider shows less than $\sim 1 \text{ ms}$ better connectivity, that could be meaningful.
 - If differences are all in the 5–20 ms range, then network proximity to the vendor and broker becomes more critical than “small cloud advantage.”
-

Summary Table

Step	Action
Set up VMs	AWS Mumbai, DigitalOcean (BLR1), Vultr/Linode Mumbai

Step	Action
Run tests	Ping or TCP to broker/NSE endpoints
Compare results	Evaluate latency, jitter, and choose lowest
Decide path	If AWS is noticeably better, maybe go with it; otherwise choose cheapest

Let me know if you'd like help scripting these tests or parsing the results once you have them—I can help you set up quick benchmarks via SSH.

○○ Sources

Can you find on the internet where the flat trade and Shunye servers are located? They must have mentioned somewhere, some things like that. I think they might be like co-located or whatever but I want you to know. Check on internet and then come back to me with proper information.

I dug around online to find any indication of where **Flattrade** or **Shoonya** host their trading servers or APIs. Unfortunately, there's **no publicly available confirmation** that reveals their exact server locations or colocation status.

Flattrade (Fortune Capital Services Pvt Ltd)

- Flattrade's **registered and main office** is in **Ambattur, Chennai**, at Kochar Technology Park, which is clearly listed on their website.
- While their **corporate footprint** is in Chennai, there's no public information confirming whether their **trading or API servers are colocated**—for example, in Mumbai or near the NSE exchange infrastructure.

Shoonya (Finvasia)

- Shoonya is headquartered in **Mohali, Punjab**, with additional offices in **Mumbai**, per its Wikipedia listing.

- Despite having a presence in Mumbai, there's no public detail indicating where their **executive or API servers** are located—whether they're in Mumbai colocation zones (like Mahape, Navi Mumbai) or elsewhere.
-

Bottom Line

- **Flattrade:** Official business HQ is in **Chennai**, but nothing confirms whether their trading servers are also located there or elsewhere.
 - **Shoonya:** Corporate offices are in **Mohali** and **Mumbai**, but again, no confirmation on server locations or colocation infrastructure.
-

What You Can Do Next

If latency is a key concern for you, it's worth reaching out directly to their technical support or business development teams and asking:

- “**Where are your API/trading servers (OMS, gateway) located?**”
- “**Are they colocated in or near NSE’s Mumbai infrastructure (e.g., Mahape, BKC)?**”
- “**Can you provide average ping/latency stats from different major cloud providers or locations?**”

This clarification will be invaluable for deciding whether to pursue low-latency hosting near them—or whether your compute location matters less than you thought.

Let me know if you'd like help framing that inquiry message—or help interpreting any reply you get back!



Sources