Banyan's Close-up View of Enterprise PMF

Claudia Richoux, Banyan

Why enterprise?

(I’m sure you already know!)

Web3 people claim they'll pay for storage (even at a *premium*) for decentralization as a feature alone, but lagging sales of decentralized products (mostly targeted within web3) say otherwise.

Storage protocols are all fighting for one seed-stage startup burn's worth of web3 TAM. Most of it is small/CDN-hot (NFTs).*

*both from a Messari report from fall 2022

Banyan talked to a LOT of real businesses with a LOT of data, nearly all outside web3.

~10 enterprises, ~40 SME, many more informally, all using our web2 competitors.

They also all had $$$ to spend on data services.

We've learned a ton from them!

(and our web2 competitors' market reports, all online, +tons of HBS case studies)

This isn’t about us… but here’s a little Banyan context

We pack/encrypt locally (user-side) into WNFS then do some optional CARwrangling and send to SPs. We track hashes/state on our server and decentralized (local + on Filecoin).

We do not use FIL+ (not that we have not tried). We do our own retrievals/hot storage (custom). We do our own network permissioning. We do our own SP monitoring atop Filecoin.

We make all the cybersecurity/compliance promises as our own entity. We keep DEO/cyber/liability insurance.

We keep our SPs under contract+monitoring, pay them directly in fiat.

We offer a usable product range (by November GA: we will be Backblaze+Dropbox+S3 feature-set-competitive) and we are startup-style iterating on product features using extensive user feedback from sniper-rifle targeted verticals (ACE, digital creatives right now- indirect to academics and through SP resellers).

We have our own BD and reseller pipelines. We brought our product live last week, are putting our alpha build in the first users’ hands TODAY (New York time zone :)), data will be on real SPs in the wild in October, November GA + feature-complete-enough-to-launch. Spent past 7 months @research + infra.

Thesis: B2B storage needs some “basics”

Without basics, you will not scale with business data, which is… most data

1. Product/usability:
    - **Ease** of product use for desired workflows
    - **Upload/retrievability** timelines fit needs
    - Competitive **pricing** with cloud hyperscalers that provide all this
2. Trust
    - **Trust**/SLAs/court venue
    - **Encryption, privacy**
    - **Compliance**

Part I: Usability✨

Most important slide: People only want two things 🥈

1. *~98% of the time I do not want to think about the data*
2. *the rest of the time I want to use the data quickly and painlessly*

- Could be more or less than 98% dependent on “hotness” of data
- **1** is partly compliance and security, partly writing corruption-free code.
- **2** is solved with a range of ***products*** that serve customer needs ***competitively*** and painlessly.

This is true for everybody, but extra true for enterprise.

They super do not want to hire a “data management person”.

They will buy if you automate that job away.

MVP: basic ops cannot be manual/painful

Data onboards and accesses need to be automated and easy at the desired scale. Otherwise you will not scale. People do not have time for data pain.

Current status: find a miner, singularity, ship data, miner replicates, confusing FIL+ application and random people looking at the data, sealing/unsealing, slow manual retrievals with ugly CIDs, unpack again with singularity 👎👎👎👎👎

Competitive would be: All data actions resolve/complete instantly with nice devex. (We acknowledge/salute/excitedly await the work of all the teams working on this!)

Then: Product! (focus: actions with data, not storage)

- I have yet to see a Filecoin-based storage app with PMF, besides sorta NFT.storage and some of the SP sidecar products (bad TAM)
- cargo-culting web2 apps with web3 “features” bolted on will *not* sell!
- Instead: study storage apps, integration, and user behavior more closely.
  - Who will use your storage product. What/who needs to touch the data.
- **Make *data flows* easy and simple. Put a ton of thought into it.**
- **What is your advantage? “Decentralization”? Nope… doesn’t sell.**

Cloud storage is very mature/entrenched. But products can still find toeholds in slightly-underserved markets by giving a pile of attention to where people are looking uncomfortable with data UIs.

You have to sell and it has to be hands-on- there should be a lot of startups onboarding + selling with different, targeted, specialized product interfaces.

Banyan’s here to help these startups be ready to use Filecoin today + going forward. (we also have hella sales pipelines)

Data in a Vacuum 🧹 (integrations!)

- Recall the most critical thing:

1. *~98% of the time I do not want to think about the data*
2. *the rest of the time I want to use the data quickly and painlessly*

You CAN use your data by sending, sharing, and viewing the raw bits.

(bare minimum)

But: You USUALLY use your data by operating on it with a useful tool.

Make sure your storage integrates. Why GDrive is killing Dropbox. Glacier without the rest of AWS would never sell. Backblaze without NAS software wouldn't either.

Go hot(ish) or go home 🥵 (speed)

- Recall again:

1. *~98% of the time I do not want to think about the data*
2. *the rest of the time I want to use the data quickly and painlessly*

Three timescales for data access:

🖥️Computer/streaming speed (CDN/Saturn) 💃Human speed (Dropbox/IPFS) 🥶Legal system speed (Glacier/FIL)

- Saturn is far from ready to compete- Banyan may roll something custom/band-aid-solution next year.
  - HARD because need to incentivize bandwidth instead of storage + scale + browser compatibility
- Glacier time: the TAM is okay, the technical setup is okay, but competitors are very entrenched and users demand strict compliance/high trustworthiness (used for ransomware/backups, legal retention and discovery, mostly largest enterprises)

We are targeting mainly-human-timescale applications but also offer faster + slower on pricing + product. This opens doors for us.

Always add vertical integration through the data access pattern toolchain: data prisons without integrations do not sell.

🤑 Pricing

Datapoints*:

- AWS Glacier: $1/TB/mo
- AWS S3: can be in 20s-40s/TB/mo
- Dropbox: over 150 or “the limit approaches zero” depending on plan
- Backblaze/Wasabi mid-markets: in the upper end of single digits
- Bandwidth 10-40/TB- depend on speed (sometimes free)
- SPs seem happy with 2.5/TB/mo cash atop block rewards, and not-too-rude bandwidth consumption.

“No egress fees” are

1. less of an advantage than you think,

2. Not necessarily true… unsealing time? SP bandwidth? …scarce resources to allocate :)

You will not hit $1 without hyperscale and tapes, CERTAINLY not with daily PoSts and small mining operations.

Still lots of wiggle room- customers pay even 100x/TB for nice product. So build nice product! :)

(*from Banyan’s utterly awe-inspiring internal spreadsheet of competitor price data. thank you Sam Grone/Tim Yong/Olive Gardner)

Aside about on-chain pricing…

- You can try and capture these flows in Filecoin or even on-chain in stables
- You will not have a good time
- Lockups for Filecoin payments = bad
- People like how they pay for things (monthly/annual billing) now
- People like invoices
- Accountants HATE crypto, including stablecoins
- They HATE non-unified accounting systems/accounts…
- From personal experience- my company has entirely sworn off holding crypto except FIL for gas- our accountants were about to kill me.
- Normies are even less ready for this.

Without a lot better crypto *payments* adoption (and extensive b2b tooling for corporate finance on crypto!!)... do not expect adoption. Customers need to pay for storage in fiat. It's ***fine*** .

SPs may be more open to weird cryptoeconomic payment models/taking payment in Fil, they still prefer fiat. There could be more motion there.

An aside about pricing: FIL+/EFIL+

Currently non-starters for enterprise:

- FIL+ requires public data, all sorts of fiddly parameters, annoying/slow/doesn't scale
- EFIL+ requires random notaries to peek at the data (no! what?!)

Fine ideas… but barely help real/scalable enterprise apps lower prices at all…

Whatever, competing on price alone with inferior product is a bad strategy anyway… We'll never beat Glacier on their turf and we should not try.

Banyan's unsolicited and biased opinion is as follows: extra block rewards ***would*** be helpful to our margins, ***if*** we could use Fil+...

But without changes to make it ***usable*** to rapid-growth startups serving large markets (like Banyan), it's an inefficient use of finite ecosystem time and resources to maintain.

We think FIL+ either needs to scale (aggressively! kill the decentralization) or go away.

Part II: Trust 🤝

Part of the ROI of paying you is so they can sue you. 👩‍⚖️

- Users are paying you to **take responsibility for the safekeeping of data**. ***Not*** ***to hold the data and get slashed if you stop***. ***Not*** ***to rent HDDs.***
  - Part of why Arweave is bad- even less individual accountability than Fil/Sia.
- Slashing some collateral is *not sufficient* to offset pain of losing data.
  - Adverse data loss events put ~60% of businesses into bankruptcy within a year of the event*
  - Court case/contracts/SLAs fix this- businesses need to be made whole.
  - “At least the miner lost some Dogecoin” does not make me feel better in bankruptcy court

Most miners not liquid enough to be “sueable”, rarely have appropriate insurance (can’t afford), are in diverse jurisdictions with nonstandard incorporation structures, many are averse to KYC/contracts… hard to justify risk for a large enterprise.

A few can take sufficient liability for broader enterprise market, but not a large/diverse set.

Unfair and unrealistic to put these requirements on small miners, already low margin.

(*Backblaze survey)

Banyan's two-tier trust system

We get customer-tier SLA performance (on cold replicas, hot replicas, spare bandwidth, sealing/unsealing capacity, etc) by adding up performance from less-strict miner SLAs and having them backstop each other.

We made CDO tranches for "subprime" enterprise cloud storage (with better monitoring 😉)

Surprise: enterprises more paranoid than most individuals!

- Security: Standards, policies, and complex legal requirements.
- Random nodes/notaries/entities looking at their data: not OK
- Random people retrieving their encrypted data: not OK

Fortunately fixable through variety of options:

- Easy: Build an encrypted product
  - E2E or a cryptography layer between client and miner
  - People are okay with you seeing their data if you're both liable and plausibly locked-down.
- Harder: Retrieval permissioning
  - UCAN on all retrievals, hope SPs comply (or contractually force them)
- Sisyphean: Make miners do full security policies
  - KYC efforts have been difficult- imagine attempting to verify all their keycard policies- worse, imagine trying to decentralize this…
  - A little security policy is doable, but not decentralized, and not network-wide.

Banyan implements the enterprise security requirements itself. SPs only receive encrypted data, must use our permissioned retrieval setup for our data, must agree to (and pass some checks on) basic cyber requirements.

We can custody the users' key and offer them extra features like S3 and compute, or offer E2E privacy.

BTW: You absolutely need access controls + deletion :)

- Need protocol + Boost/Lotus level support
- We built “miner must check a Banyan-signed UCAN before you serve Banyan’s encrypted data to a user”
  - Like Kerberos. Could also use *Lit Protocol* (threshold/blockchain based signatures as a service), however, not in prod yet
  - this should happen at Boost/Lotus level and be standardized
- For deletion, we’re doing:
  - “wipe a cryptographic key (works at object-store level)”
  - “Ablate immediately from hot storage”
  - “Refuse to serve/unseal that cold copy at the UCAN level”
  - Eventually, maybe: “hello miner, Banyan/user will comp you to get slashed” (NOT GOOD for Filecoin)
    - A protocol improvement to fix this would be: consensual slashing?
    - FIP time @alex north?

Alphabet Soup: SOC2, ISO, HITRUST, GDPR, FedRamp, PCI/DSS

- There are a lot of data processing security certifications. They overlap a lot.
- Most industries need one (possibly multiple) to use the storage.
- Cloud hyperscalers and all competitors have them already.
- They cost a lot of money for each org to complete
  - Some SP orgs have the resources, 99% do not
  - Certifying many SPs eats up block rewards that could go to more productive things.
  - Bad plan right now, everyone's hungry, SP margins are thin.
- AWS Outpost model: App layer does SOC2 once, ships data to miners who are under a compliant monitoring program
  - Saves money/time/effort, everyone else can get certified later.
- …or, protocols themselves could get certified… *probably impossible!*
  - Banyan built a compliance DAO spec- ask me later- it's on our Github…
    - We aren't using it because not enough SPs are compliant, lol

Reprise: B2B storage needs some “basics”

Without basics, you will not scale with business data, which is… most data

1. Product/usability:
    - **Ease** of product use for desired workflows
    - **Upload/retrievability** timelines fit needs
    - Competitive **pricing** with cloud hyperscalers that provide all this
2. Trust
    - **Trust**/SLAs/court venue
    - **Encryption, privacy**
    - **Compliance**

Banyan's response to all this…

Banyan has slapped several centralized (but prod-ready) band-aids over the things that don't work automatically yet on Filecoin. Banyan excitedly awaits most of these technologies working, but will continue to band-aid-slap.

Banyan has spent a ***lot*** of time talking to customers about their data flows. Extreme focus on product + branding, we're fighting unicorns and hyperscalers.

We believe we found a small product edge in mid-market SME data storage. We'll see how it sells and continue to iterate.

Two tier system from before, means all legal/compliance expectations fulfilled. We chose to GTM on verticals with lighter compliance standards and less industry legislation, as well as being underserved by incumbents.

We will get our sea legs selling to them, iterate, and go for pickier customers.

Questions? :~)

Questions? :~)

Questions? :~)

Questions? :~)

Questions? :~)

Questions? :~)