COSMOS

Estrutura técnica da instância nacional da Atmosfera

Documento 2 de 5

Este documento descreve como o Cosmos é construído: os componentes, a topologia federada por instituições, a regra de redundância, a continuidade das contas, a moderação, a segurança, a operação e as fases de implantação. É o par técnico do Manifesto (documento 1) e a referência do Termo de Adesão e Cooperação Técnica (documento 4), que remete a ele para a régua mínima dos servidores.

Versão 1.0, novembro de 2026. Este documento não trata de custos, e o projeto executivo também não. O dimensionamento de capacidade por fase consta do projeto executivo (Cosmos: um nó em cada localidade do IX.br, documento 5), que é anexo deste documento e será revisto à medida que as instituições aderirem com máquinas e instalações próprias.

1. Princípios de projeto

Sete decisões orientam todas as demais. Quando duas soluções técnicas competem, vence a que respeita mais princípios.

1. **Uma infraestrutura nacional.** Para quem usa, há um só Cosmos: um domínio de identidade (@nome.cosmos.org.br) e uma porta de entrada. Sobre ele roda a Anil, a rede social mantida pela associação, com um aplicativo e uma política de moderação. As instituições fornecem capacidade e redundância, nunca identidade.
2. **Nenhuma peça única (regra dos dois).** Toda função da rede existe em pelo menos duas instituições, em duas localidades. O que só existe num lugar não entra em produção. A regra vale para a instituição fundadora.
3. **Dados no Brasil.** As contas de brasileiros ficam em servidores de instituições brasileiras, sob a lei brasileira. Instituições de outros países de língua portuguesa hospedam os seus nacionais sob as suas leis, na mesma rede. A rede pode ter ainda servidores parceiros e espelhos no exterior, com cópias dos dados públicos, como o NIC.br mantém fora do país cópias dos servidores DNS do .br.
4. **Software livre e publicado.** Tudo o que roda tem código aberto e configuração publicada. Segredos (chaves, senhas) ficam fora do repositório e fora de qualquer canal de conversa.
5. **Interoperabilidade total.** O Cosmos é parte da Atmosfera. Os Okas são rastreados pelos relays da Bluesky e do Blacksky, além dos nossos, e o Tesá indexa o firehose inteiro, não só contas brasileiras.
6. **Operação sóbria.** Debian estável, acesso por chave, duas pessoas por nó, atualizações automáticas com janela, monitoramento comum e página pública de estado.
7. **Transparência qualificada.** Rótulos de moderação públicos e assinados, decisões com recurso, atas e contas da associação abertas, relatório de transparência semestral.

2. O protocolo em uma página

O AT Protocol separa quatro coisas que as plataformas fundem: identidade, armazenamento, agregação e moderação.

- **Identidade.** Cada conta tem um identificador permanente (DID). O método mais usado é o did:plc, cujo documento fica num diretório público (plc.directory), em transição da Bluesky para uma associação suíça, com réplicas de leitura abertas. O documento DID aponta a chave de assinatura e o servidor onde a conta mora. Há também o did:web, ancorado num domínio.
- **Handle.** O nome legível é um nome de DNS, provado por registro TXT ou por um arquivo bem conhecido. Trocar de handle não troca de identidade.
- **PDS (servidor pessoal de dados).** Guarda o repositório assinado de cada conta (posts, seguidores, curtidas, preferências) e os blobs (imagens, vídeos). É a única peça com estado que não se reconstrói de fora.
- **Relay.** Rastreia os PDS e publica um fluxo único e verificável de tudo o que acontece (o firehose). Não guarda conteúdo além de um buffer.
- **AppView.** Consome o firehose, indexa e monta o que o aplicativo mostra: timelines, perfis, busca, notificações. Reconstrói-se do zero a partir do firehose.
- **Labelers e feeds.** Serviços independentes que emitem rótulos assinados (moderação) ou listas de posts (feeds). O usuário escolhe quais seguir.

- **Lexicons.** Esquemas de dados com nomes de domínio (app.bsky.feed.post). Qualquer um pode criar os seus; o Cosmos pode definir lexicons próprios sem quebrar o resto da rede.

3. Componentes do Cosmos

Cada componente tem um nome tupi que diz o que ele faz, um nome técnico que fica no DNS como apelido, e um software de referência.

Nome	Função	Software de referência	Estado	Redundância
Oka (casa)	PDS: guarda as contas e os blobs	bluesky-social/pds (TypeScript, SQLite por conta, blobs em disco ou S3)	Com estado	Par ativo e reserva em duas instituições
Okara (praça)	Porta de entrada: cadastro, login, recuperação, escolha do Oka, migração	PDS em modo entryway mais serviço próprio de distribuição	Pouco estado (cadastro)	Duas réplicas, banco replicado
Ybytu (vento)	Relay: rastreia os Okas e o mundo e publica o firehose	bluesky-social/indigo (relay, Sync 1.1)	Sem estado durável	Dois, em instituições diferentes
Tesá (olho)	AppView: índice da rede, API do aplicativo	Pilha do Blacksky (Rust, Postgres) ou bluesky-social/atproto appview	Reconstruível	Duas réplicas, api.cosmos.org.br em anycast
Tekó (modo de viver)	Moderação: rótulos, filas, recurso, classificadores	bluesky-social/ozone mais classificadores próprios	Reconstruível (rótulos são registros públicos)	Um comum, replicado; um por instituição que quiser
Ara (dia)	Geradores de feed	feed-generator de referência e implementações próprias	Sem estado	Qualquer instituição
Itá (pedra)	Espelho do diretório PLC e custódia de chaves	plcbundle, réplica de leitura do PLC	Cópia	Em todo nó
Taba (aldeia)	Nó de borda numa localidade do IX.br	Debian, Caddy, WireGuard, BGP	Infraestrutura	Uma por localidade; pares entre localidades
Ybytyra (montanha)	Núcleo: onde moram Tesá e Ybytu	Idem, em escala	Infraestrutura	Dois núcleos (fase 2)

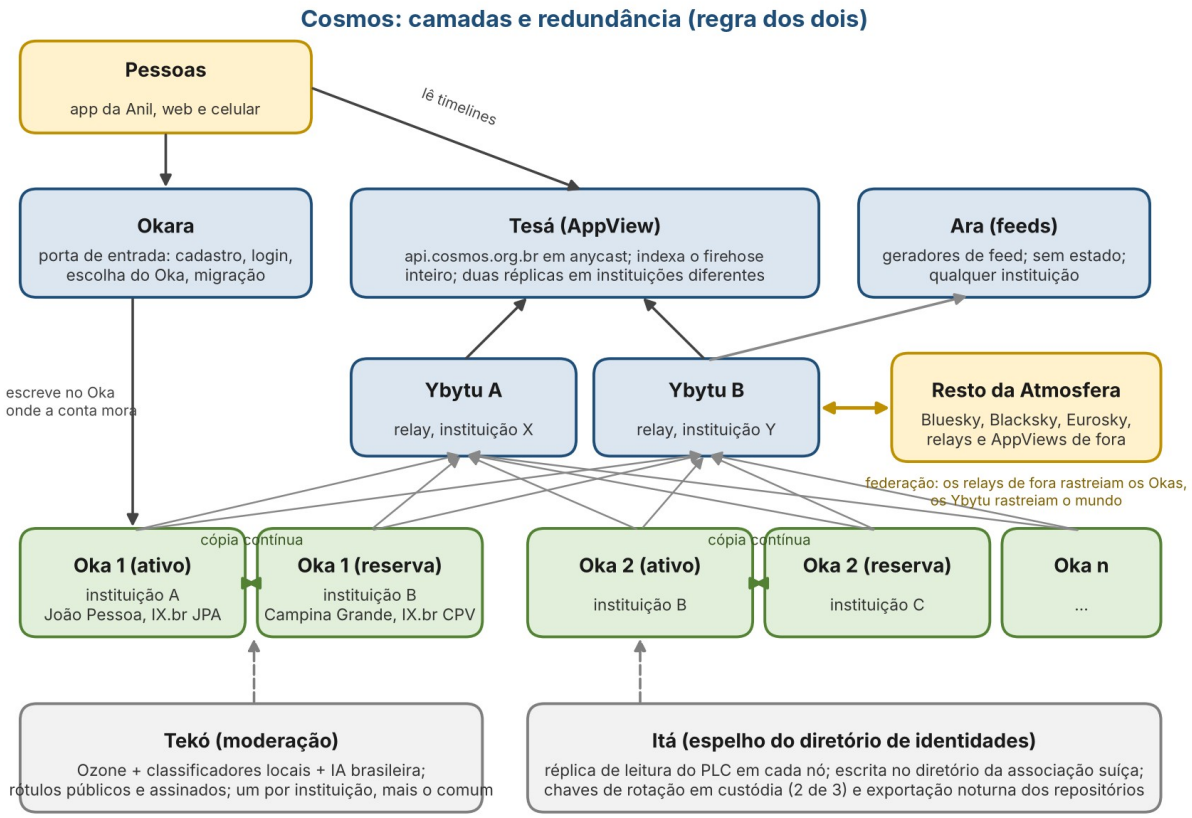


Figura 1. Camadas do Cosmos e a regra dos dois. Os Okas vivem em pares entre instituições; relays, AppView e porta de entrada têm pelo menos duas réplicas.

4. Topologia: instituições e localidades

A unidade de participação é a instituição que adere pelo Termo de Adesão e Cooperação Técnica. Uma unidade acadêmica, um laboratório ou um projeto de pesquisa sem personalidade jurídica própria adere do mesmo modo, pela assinatura do coordenador responsável, com a anuência do dirigente da unidade. Cada instituição opera pelo menos um servidor (Oka) ligado ao ponto de troca de tráfego do IX.br da sua cidade, dentro de um sistema autônomo próprio ou do sistema autônomo da associação. As instituições se organizam em pares: toda Oka tem a sua reserva numa instituição diferente e, sempre que possível, numa localidade diferente do IX.br.

O primeiro par nasce na rede da Associação Nacional para Inclusão Digital (AS28135, prefixos 187.45.80.0/20 e 2804:88::/32), em João Pessoa, com a reserva em Campina Grande. Os dois são localidades do IX.br. A partir daí, cada instituição que entra traz um servidor e recebe um par.

O tráfego entre pares (replicação) e entre nós e núcleos passa por túneis WireGuard sobre o IX.br, com rotas de contingência pelo trânsito comercial de cada instituição. O tráfego público (usuários, relays de fora) chega pelo IX.br e pelo trânsito. Nos núcleos, o endereço do Tesá (api.cosmos.org.br) é anunciado em anycast a partir das duas localidades; se um núcleo cai, o BGP leva os usuários ao outro em segundos.

A escala alvo do projeto executivo é um nó em cada uma das quarenta localidades do IX.br, dez do tipo A (capitais com maior tráfego) e trinta do tipo B, e dois núcleos (São Paulo e Fortaleza). A topologia federada não altera esse alvo: muda quem fornece cada nó.

5. O Oka em detalhe

5.1 Régua mínima do servidor

A régua é a mesma para toda instituição e consta do Termo de Adesão. Um servidor abaixo da régua pode entrar em homologação, nunca em produção.

Item	Mínimo	Recomendado
Sistema operacional	Debian estável (12 ou 13), instalação limpa, sem outros serviços	Idem, em máquina dedicada
Processador e memória	8 vCPU, 32 GB	16 vCPU, 64 GB
Armazenamento	2 TB NVMe (sistema e contas) mais 4 TB para blobs	Blobs em armazenamento de objetos S3 compatível (MinIO) com replicação de sítio
Rede	1 Gbps, IPv4 e IPv6 públicos, portas 80 e 443	10 Gbps, porta no IX.br, BGP próprio ou do AS da associação
Operação	Dois administradores nomeados, acesso só por chave SSH, sudo nominal, sem senha de root	Chaves em hardware (FIDO2) para os administradores
Disponibilidade	Energia e refrigeração do data center institucional, no-break	Gerador, redundância de enlace

5.2 Capacidade e armazenamento

O PDS de referência mantém um banco SQLite por conta e um banco de contas; o custo de CPU é baixo e o que cresce é disco de blobs. A Bluesky opera os seus próprios PDS (os cogumelos) com algumas centenas de milhares de contas cada; o Cosmos dimensiona cada Oka para dezenas de milhares de contas ativas, preferindo mais Okas pequenos a poucos grandes, porque o par de reserva fica mais barato e a falha de um afeta menos gente.

Estimativa de uso: a aplicação comprime cada imagem a menos de 1 MB; uma conta ativa típica gera de 2 a 5 MB de blobs por mês. Dez mil contas ativas consomem cerca de 30 GB por mês; cem mil, cerca de 300 GB por mês. Na fase cogumelo, disco local basta; a partir de cinco Okas, os blobs migram para armazenamento de objetos replicado entre instituições.

5.3 Nomes e certificados

Cada Oka tem um nome de host fixo (oka-xx1.cosmos.org.br, com o apelido urupê, cogumelo em tupi, permitido para quem quiser manter a tradição da Bluesky). Esse nome fica gravado no documento DID de todas as contas que moram nele; por isso o nome nunca muda, só a máquina atrás dele. Os handles padrão são @nome.cosmos.org.br, resolvidos por DNS curinga e certificados emitidos sob demanda pelo Caddy, com o endpoint de verificação do PDS como porteiro. Handles institucionais (nome.ufpb.br) dependem só de um registro TXT no DNS da instituição.

5.4 Correio e verificação

O PDS envia e-mails de confirmação, recuperação e migração. O Cosmos usa correio próprio (Postfix com DKIM, SPF e DMARC no domínio cosmos.org.br), em duas instituições, nunca serviço de e-mail estrangeiro.

5.5 Federação e limites

Um Oka novo pede rastreamento aos relays (o da Bluesky, o do Blacksky e os Ybytu). O relay da Bluesky aplica um teto de contas por servidor novo, elevado a pedido; o modelo de muitos Okas pequenos convive bem com esse teto, e os Ybytu próprios eliminam a dependência a partir da fase 1.

6. Continuidade: como um Oka cai sem ninguém sentir

O nome fica, a máquina muda. O par ativo e reserva funciona assim:

8. **Replicação.** Blobs em armazenamento de objetos com replicação de sítio para o irmão (MinIO site replication; Ceph nos núcleos). O diretório de bancos SQLite é fotografado a cada cinco minutos (ZFS send/receive ou equivalente) e enviado ao irmão pelo túnel WireGuard.
9. **Chaves.** As chaves de assinatura das contas ficam no banco replicado. O irmão assina com as mesmas chaves; o documento DID não muda; o usuário não faz nada.
10. **Detecção e virada.** Verificação de saúde a cada 30 segundos de dois pontos de observação. Falha confirmada, o DNS do host (TTL de 60 segundos) ou a rota anycast passa a apontar para o irmão. Objetivo: perda máxima de cinco minutos de escrita (RPO) e retomada em até quinze minutos (RTO).
11. **Volta.** O servidor que caiu volta como reserva, ressincroniza e só reassume numa janela planejada.
12. **Pior caso.** Para a perda do par inteiro (incêndio, apreensão, saída da instituição), cada conta tem uma segunda chave de rotação sob custódia da associação, guardada offline em esquema de dois entre três guardiões, e os repositórios são exportados toda noite (arquivos CAR) para uma terceira instituição. Com chave de rotação e CAR, as contas são realojadas em outro Oka por operação no diretório PLC, sem a cooperação do servidor perdido.
13. **Ensaios.** Cada par executa uma virada planejada por trimestre, registrada em ata do Colégio de Instituições. Par que não ensaia perde o selo de produção.

7. Identidade, porta de entrada e migração

7.1 Okara

A Okara é o único endereço que o usuário conhece. Ela cria a conta, escolhe o Oka (o mais próximo, o menos cheio, ou o da instituição à qual a pessoa pertence), emite o handle, faz login, recupera senha e conduz a migração. O PDS de referência já opera atrás de uma entryway; o serviço de distribuição é código do Cosmos, publicado. Na fase cogumelo, com um só Oka, a Okara é o próprio Oka.

7.2 Itá

Todo nó mantém uma réplica de leitura do diretório PLC (plcbundle), de modo que resolver identidades nunca dependa de um host fora do país. A escrita (criação e rotação de chaves) continua indo ao diretório da associação suíça, com a custódia descrita na seção 6 como seguro.

7.3 Migração

A migração de conta entre servidores é função nativa do protocolo: a conta é criada no destino com o mesmo DID, o repositório, os blobs e as preferências são importados, o diretório PLC é atualizado

e a conta de origem é desativada. A ferramenta de linha de comando goat executa isso hoje; o migrador do Cosmos é a interface amigável por cima, com um botão, para as cerca de 2,6 milhões de contas brasileiras já registradas na Atmosfera.

8. Tesá: a visão da rede

O Tesá consome o firehose dos Ybytu e indexa tudo: posts, perfis, grafos de seguidores, curtidas, rótulos. É a peça mais cara da rede, porque indexa a Atmosfera inteira e não só o Brasil, e por isso é a última a entrar (fase 2). Até lá o aplicativo da Anil usa o AppView da Bluesky, o que o protocolo permite sem perda para o usuário.

Referência de implementação: a pilha do Blacksky, em Rust sobre Postgres, com ScyllaDB para os índices de maior volume na escala de dezenas de milhões de contas. Duas réplicas completas, uma por núcleo, anunciando api.cosmos.org.br em anycast. Busca em serviço separado. Dimensionamento por fase no projeto executivo anexo.

9. Tekó: moderação em três camadas

A ordem das camadas importa, e o desenho é feito para que a decisão com contexto seja humana e a automação cuide do volume. As diretrizes gerais de moderação da Anil são aprovadas pelo Conselho de Administração, por proposta da Diretoria Executiva; o detalhe fica em regimentos e manuais técnicos, que podem variar em parte de um servidor para outro. Não há comitê de moderação na fundação, e o Conselho pode criá-lo depois.

14. **Da pessoa.** Silenciar palavras e expressões por prazo ou para sempre, bloquear, silenciar contas, seguir listas de moderação curadas, escolher quais rótulos aplicar ao que lê. Tudo isso já existe no protocolo e no aplicativo.
15. **Da comunidade.** Moderadores humanos no Ozone, sob o Estatuto da Criança e do Adolescente, o Marco Civil, a Lei Geral de Proteção de Dados e o ECA Digital. Cada rótulo é um registro público e assinado; cada decisão tem motivo em português, autor identificável por função e caminho de recurso com prazo. Instituições podem manter labelers próprios por cima do comum, nunca em lugar dele.
16. **Automática.** Classificadores em duas passadas. A primeira, em modelos abertos rodando em máquinas da rede (GPU própria), varre o volume inteiro sem que nenhum byte saia da rede; a segunda, só para os casos duvidosos, usa inteligência artificial brasileira (modelos treinados em português do Brasil e no direito nacional) para decidir contexto, ironia, gíria e referência legal, sempre entregando ao moderador humano a explicação. A automação nunca aplica sozinha rótulo que suspenda uma conta.

Ativos de partida: um dataset curado de 1.682 palavras-chave, hashtags e acrônimos ligados à exploração sexual de crianças e adolescentes, usado na primeira passada e atualizado pela comunidade; as listas e ferramentas de moderação cedidas pelo Blacksky; e os classificadores de imagem para material de abuso, integrados por hash e por modelo.

Transparência: relatório semestral com quantidades de rótulos por categoria, recursos recebidos e revertidos, tempo médio de decisão e pedidos de autoridades, publicado na página da associação.

10. Segurança

- Sistema: Debian estável, atualizações automáticas de segurança com janela e reinício controlado; nada de editores interativos em produção, toda mudança por script versionado com cópia de segurança e validação.
- Acesso: SSH só por chave, dois administradores por nó, sudo nominal, registro de sessões, chaves em hardware recomendadas; segredos em arquivos de permissão restrita, nunca em repositório nem em conversa.
- Rede: firewall por padrão fechado, apenas 80, 443 e as portas de replicação pelo túnel; fail2ban; RPKI com ROAs publicadas para todos os prefixos; DNSSEC nos domínios; IPv6 em todo nó.
- Aplicação: TLS 1.2 ou superior, HSTS, cabeçalhos de segurança completos (a referência é a nota máxima no Sentinela), limites de taxa na Okara e nos Okas, proteção contra abuso no cadastro.
- Dados: cópias cifradas, chaves de cifra sob custódia da associação, exportações noturnas verificadas por restauração mensal de amostra.
- Resposta a incidentes: canal comum (Matrix) entre os administradores, plantão por rodízio, resposta em até 24 horas, relatório público de incidente em até 15 dias.
- Auditoria: código e configuração públicos; auditoria externa anual de segurança, com relatório publicado.

11. Operação e observabilidade

Todo nó exporta métricas (Prometheus) para dois coletores em instituições diferentes, com painéis comuns (Grafana) e alertas por e-mail e Matrix. Uma página pública de estado mostra, por componente e por nó, disponibilidade, latência e incidentes em curso. Objetivos de serviço: 99,9% para Okara e Tesá, 99,5% por Oka (com a virada para o par contando como disponibilidade), latência de leitura abaixo de 300 ms no Brasil.

O Colégio de Instituições mantém o calendário de manutenção, propõe ao Conselho de Administração a régua e as mudanças de versão, homologa a entrada dos servidores em produção, e registra os ensaios de virada. Cada instituição responde pelo seu nó; a associação responde pela Okara, pelos Ybytu, pelo Tesá e pelo Tekó comum, operados por instituições designadas.

12. Fases de implantação

Fase	O que entra	Quem	Prazo
0. Cogumelo	Oka-1 em João Pessoa com reserva em Campina Grande; Ozone de testes; correio próprio; relays de fora; aplicativo da Bluesky com login no Cosmos; fundação da associação e transferência do domínio	ANID e a equipe técnica do grupo Atmosfera Brasil; associação fundada ainda em 2026	Dezembro de 2026 a fevereiro de 2027
1. Taba	Okara; dois Ybytu; Itá em todo nó; Ara com os primeiros feeds; aplicativo próprio (fork do app de referência); Tekó comum em produção; cinco instituições aderentes	Associação constituída; Colégio de Instituições	Março a agosto de 2027

Fase	O que entra	Quem	Prazo
2. Ybytyra	Tesá em dois núcleos; anycast; vinte instituições; migrador com um botão	Associação e instituições dos núcleos	Setembro de 2027 a junho de 2028
3. Céu inteiro	Nó em cada uma das quarenta localidades do IX.br; Tesá dimensionado para vinte milhões de contas; instituições de outros países de língua portuguesa	Rede completa	Até o fim de 2029

A fase 0 não exige investimento novo: usa máquinas existentes na rede da ANID e o trabalho da equipe técnica. O dimensionamento das fases seguintes consta do projeto executivo (documento 5), cuja necessidade de máquinas próprias o modelo federado reduz na medida em que as instituições entram com máquinas e instalações próprias.

13. Papéis e responsabilidades técnicas

Quem	Responde por
Associação (entidade de governança, a ser fundada em 2026; recebe da ANID os nomes e domínios quando tiver CNPJ)	Nome e domínios; Okara; política e Tekó comum; custódia de chaves; página de estado; régua e versão de referência; relatório de transparência; controladora dos dados pessoais perante a LGPD
Instituição aderente	Seu Oka (e o Oka irmão que hospeda em reserva), dentro da régua; dois administradores; ensaios de virada; operadora dos dados pessoais do seu Oka; saída ordenada com 90 dias de aviso e migração das contas antes do desligamento
Colégio de Instituições	Proposta de régua e de versões ao Conselho de Administração; calendário, pares, homologação dos servidores, pós-incidentes
Instituições designadas para os núcleos	Tesá, Ybytu e coletores de métricas, sob acordo específico
Equipe técnica inicial	Fase 0: instalação do Oka-1 e do Ozone, correio, DNS, documentação de instalação reprodutível

14. Glossário

Termo	Significado
Atmosfera	Conjunto dos aplicativos e serviços sobre o AT Protocol
DID, did:plc, did:web	Identificador permanente da conta e seus dois métodos
PDS	Servidor pessoal de dados (o Oka)
Relay, firehose	Serviço que rastreia os PDS e o fluxo único que ele publica (o Ybytu)
AppView	Índice e API do aplicativo (o Tesá)
Labeler, Ozone	Serviço de rótulos de moderação e a ferramenta de referência (o Tekó)
Entryway	Porta de entrada que distribui contas entre PDS (a Okara)
CAR	Arquivo de exportação de um repositório de conta

Termo	Significado
IX.br	Pontos de troca de tráfego do NIC.br, em quarenta localidades
RPO, RTO	Perda máxima de dados e tempo máximo de retomada em uma falha
Urupê	Cogumelo, em tupi; apelido opcional dos Okas