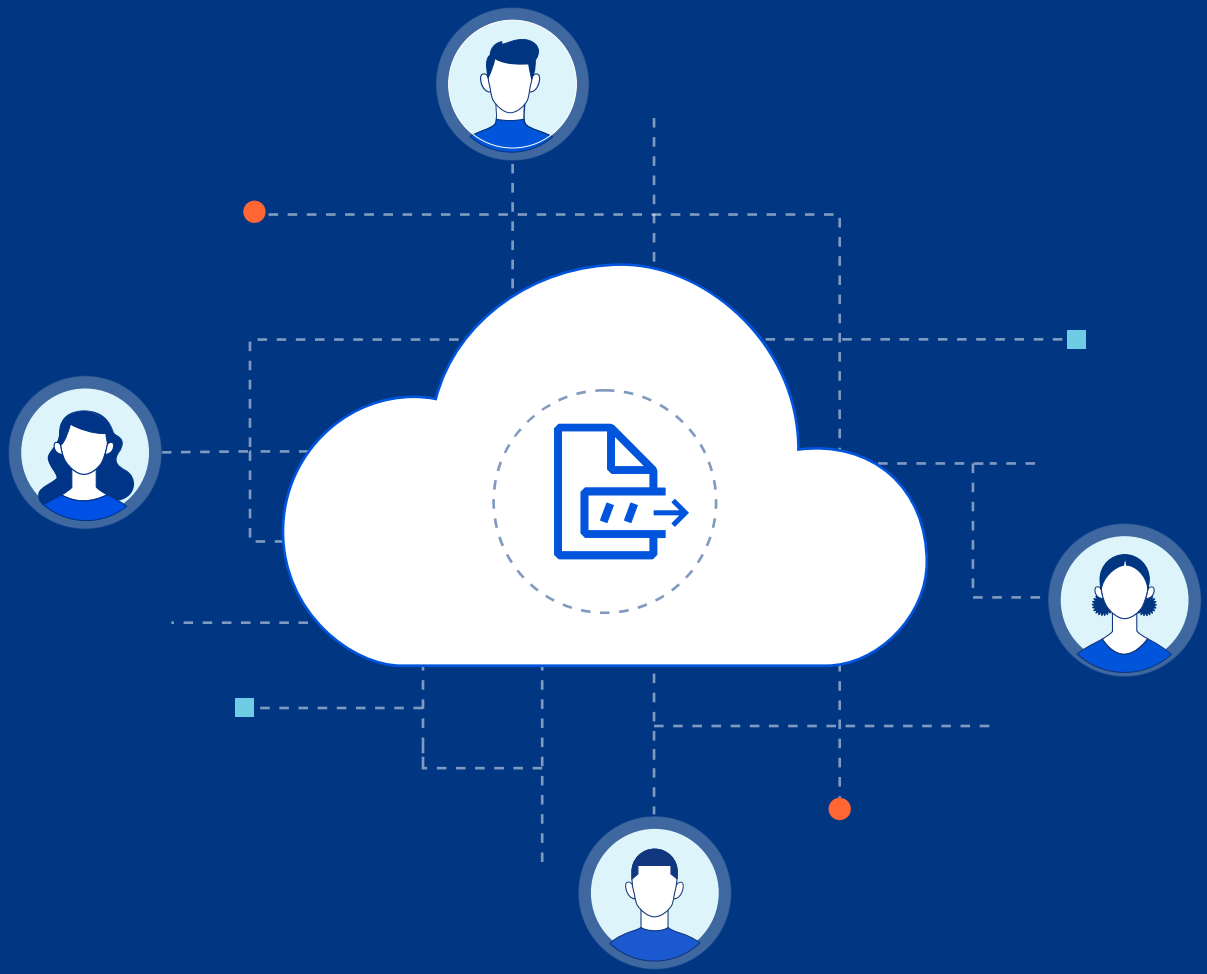
ホワイトペーパー

Cloudflare IPFSゲートウェイ



03-4510-1893 | enterprise@cloudflare.com | www.cloudflare.com

目次

免責事項

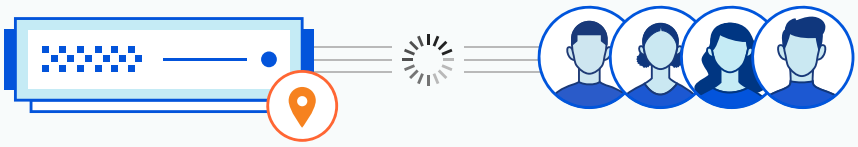
このドキュメントに記載されているCloudflare IPFSゲートウェイのアーキテクチャとそれぞれの動作は、一般提供（GA）を想定したものです。この文書はGA前に書かれたものであるため、技術的な詳細、アーキテクチャ、機能、またはリリースされるものについての決定がGA前に変更される可能性があります。

概要

惑星間ファイルシステム（IPFS）は、データの保存と共有のためのプロトコルおよびピアツーピア（P2P）分散型ネットワークであり、IPFSネットワーク内のノードは分散型ファイルシステムに貢献します。IPFSは、今日の集中型Webモデルと比較して、非集中型の分散版Web（dWeb）を実現するための重要なコンポーネントです。

このホワイトペーパーでは、現行の中央集権的Web（Web 2.0）が抱える問題は何か、それらの問題の緩和と分散型Web（Web 3.0またはWeb3）への移行にIPFSがどう役立つか、Cloudflare IPFS Gatewayのメリットは何かを考察します。

集中型Webの課題



今日のWebは、エンドユーザー（クライアント）が中央リポジトリ（オンプレミスまたはクラウド上のサーバー）にコンテンツをリクエストするというホスト中心型のモデルで発展してきました。例えば、クライアント（パソコンやモバイル端末のWebブラウザ）は、特定のポートでリクエストを待ち受けているWebサーバーにHTTPリクエストを送信します。リクエストを受けると、Webサーバーはクライアントにレスポンスを返します。この現行モデルには、以下のような課題があります。

- **耐久性：**コンテンツやアセットを提供する集中型サーバーが利用できなくなったり、コンテンツが消失するような事象が発生した場合、コンテンツはWeb上から消滅します。
- **信頼/検証：**今日のWeb2.0では、ロケーションベースのアドレス指定は、ユーザーがリクエストしているものを受け取ることを本質的に強制していません。受け取ったコンテンツがリクエストされたもので、改ざんされていないことを確認するための追加措置を講じる必要があります。
- **サイロ化されたデータ：**データは集中型リポジトリの中でサイロ化され、アプリケーションやユーザー間で簡単に共有できない傾向があります。データをエクスポートするための統合や手法が存在する場合もありますが、IPFSの場合、独自のウォールドガーデンやサポートへの依存はなく、データは分散化されており、誰でも別のインターフェースを使ってアクセスすることが可能です。
- **局所性/パフォーマンス：**コンテンツは、必ずしもローカルまたはエンドユーザーの近くではなく一元化されたロケーションから提供される場合があり、その結果、レイテンシーが高くなります。一元化されたロケーションでは、需要の増加に伴い、追加のリソースとスケールアウトのための介入を必要とします。さらに、コンテンツのダウンロードには、クライアントとサーバーの1対1の関係に関するレイテンシーがあります。
- **重複：**複数のユーザーが同じコンテンツ/ファイルを公開できるため、重複が発生します。さらに、コンテンツに変更が加えられた場合、コンテンツ/ファイル全体が再公開されます。

dWeb/IPFS、メリットおよびユースケースについて

dWebの主要原則は2つあります。

1. Webコンテンツの保存と提供を担当する中央エンティティが存在しません。
2. 不変性と検証を暗黙的に組み込んだトラストレスモデルです。

IPFSは、dWebのストレージレイヤーを提供します。同じプロトコルを使って相互に通信し、データを共有することができる相互接続されたノードのネットワークと想定することができます。このモデルは、接続されたデバイスの非集中型ネットワークとしてのWebという本来のコンセプトにより合致しています。IPFSは、dWebの主要原則である「非集中化」と検証を組み込んだ「トラストレスモデル」の2つを示しています。

IPFSは、データのインポート、ネーミング、検索、取得に多くの異なるプロトコルを使用します。関連するすべてのプロトコルとIPFSの動作方法について深く掘り下げることは、このドキュメントの範囲外ですが、いくつかの重要なポイントに注意する必要があります。

- IPFSはP2Pネットワークであるため、本質的に分散型であり、ネットワーク上のノードはいつでもネットワーク上の他のノードのサブセットに接続されます。
- IPFSはコンテンツ識別子（CID）によるコンテンツベースのアドレス指定でデータやコンテンツを参照・取得します。CIDは暗号ハッシュで構成され、IPFSに格納されているすべてのオブジェクトは一意のCIDを持ちます。
- IPFSに保存されたコンテンツはすべて不変であり、コンテンツを更新すると、新しいCIDを持つ新しいオブジェクトが作成されます。
- 大きなコンテンツやオブジェクトは、デフォルトで256kiBのオブジェクトに分割されます（注：サイズは最大1MiBまで設定可能なオプションです）。IPFSは、コンテンツのすべてのピースにリンクする空のIPFSオブジェクトを作成します。このようにデータを分割することで、オブジェクト/ブロックをより多くのホストに分散させることができるという固有のメリットがあります。また、複数のオブジェクト/ブロックに分割することで、ネイティブな重複排除が可能になります。適用可能な場合は、同じブロックを更新されたコンテンツや異なるコンテンツに使用することができます。
- IPFSはガベージコレクションを使用して、不要になったデータを削除することで、IPFSノードのディスク領域を解放します。IPFSノードでは、データをピン留めすることで、ガベージコレクションからデータを保護できます。データをピン留めすることで、IPFSは指定されたオブジェクトを常にどこかに保持するように指示します。デフォルトでは、これはローカルノードです。また、データをピン留めするサービスもあり、ユーザーがピン留め用のIPFSノードを独自に稼働させなくても、IPFSにデータを永続化させることができます。
- 説明したように、IPFSのコンテンツはCIDを介して取得されます。しかし、コンテンツは不変であるため、更新が行われると新しいCIDが作成されます。コンテンツが更新されると、最新の更新のためにユーザーが提供したいと思うCIDが変更されます。惑星間ネームシステム（IPNS）、DNSLink、イーサリアムネームサービス（ENS）といったソリューションがあり、ユーザーはこれらを用いてIPFSハッシュ/CIDを可変アドレスにマッピングすることができます。それぞれのソリューションについての詳細な議論は、本書の範囲外です。

IPFSの複数のメリットと、集中型Webの課題を解決する方法については、以下で詳しく説明します。

非集中化/耐久性

IPFSでは、コンテンツを提供するために指定された中央のプロバイダーは存在しません。その代わり、各ノードはコンテンツを提供できるあらゆるノードと対等です。以下の図1、図2は、現在の集中型WebとIPFSを用いたdWebモデルの違いを示しています。

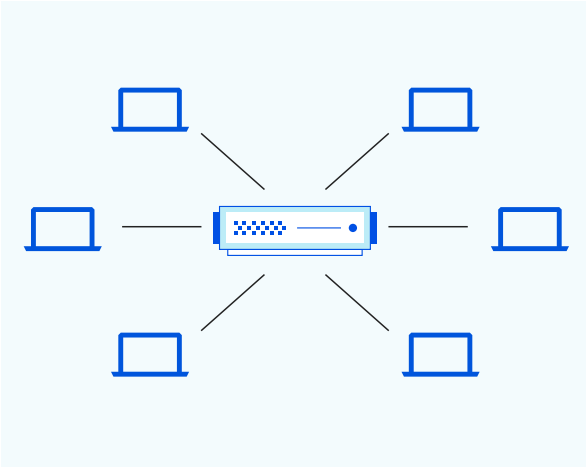


図1：集中型Web（サーバーベース）

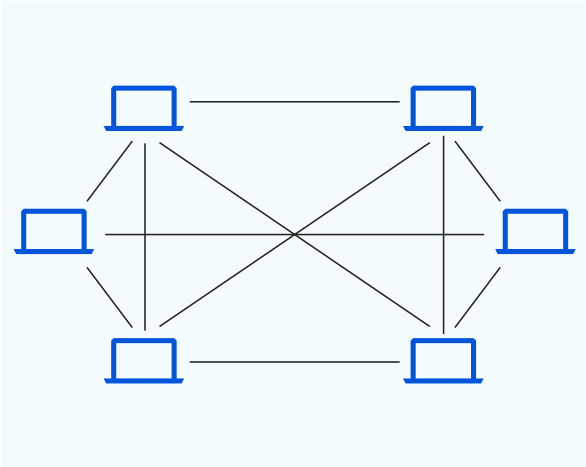


図2：dWeb（P2P）

dWeb/P2Pモデルのメリットは、以下の通りです。

1. ホスティングプロバイダー1社に依存することなく、誰でもコンテンツを公開することができます。ユーザーは1つのベンダーに縛られることなく、複数のベンダーを柔軟に使い分け、ベンダーに関係なく一貫したインターフェースでコンテンツにアクセスすることができます。ユーザーは、ピン留めと持続性のために自分の環境/ノードを使用することも、複数の無料または商用プロバイダーを使用することも、自分のノードとピン留めサービスを組み合わせて使用することも可能です。
2. コンテンツが分散化され、複数のノードがコンテンツを提供できるため、可用性が高まります。1つのノードがオフラインになっても、他の利用可能なIPFSノードがコンテンツを提供することができます。

信頼/検証

HTTPSを利用することで、ユーザーは安全に暗号化された状態でサイトに接続することができます。しかし、HTTPはロケーションベースのアドレス指定を使用しているため、ユーザーがリクエストしたコンテンツが必ず届くという保証はありません。サーバーがハッキングされた場合、ユーザーは改ざんされたコンテンツや悪意のあるコンテンツを受け取る可能性があります。固有の検証は行われず、ユーザーは明示的にプロバイダーを信頼しています。

IPFSでは、暗号を利用して信頼と検証を暗黙的にシステムに組み込んでいます。IPFSにファイルが追加されると、そのファイルは不変であり、一意のハッシュまたはCIDを介して参照されます。IPFSに保存されたコンテンツに関連する暗号ハッシュは、コンテンツベースのアドレス指定の基礎となるものです。クライアントはリクエストを行う際、自分のCIDを把握した上でコンテンツをリクエストします。また、IPNS、DNSLink、ENSなどのネーミングサービスを利用した場合、最終的にはリクエストされたオブジェクト/コンテンツのCIDに解決します。

以下の図3、図4は、現在HTTPで使われているロケーションベースのアドレス指定と、IPFSで使われているコンテンツベースのアドレス指定の違いを示しています。

図3：ロケーションベースのアドレス指定

図4：コンテンツベースのアドレス指定

ロケーションベースのアドレス指定例

ロケーションベースのアドレス指定は、ユーザーが受け取るものが、ユーザーが意図したものであるという保証はありません。

http://example.com/picture.jpg

コンテンツベースのアドレス指定例

コンテンツベースのアドレス指定は、不変性と検証を本質的に組み込んだトラストレスモデルを持ち、データの保存やリクエストにハッシュやCIDを使用します。コンテンツは一度保存されると、変更することはできません。更新が行われた場合、新しいオブジェクトと関連するハッシュが作成されます。以下の例では、カスタムIPFS URLプロトコルを使用してIPFS上のコンテンツにアクセスしている様子を示しています。

IPFSのURLは次のようなセマンティックスを持ちます：ipfs://<CID>/<path>

以下は、CIDでオブジェクトをリクエストする例です。

ipfs://QmRfGgnt2jokJP6Eh4GH7FbR3RPxmkn tKHyCXq5gNp8QuU

サイロ化されたデータはもう不要

現在、アプリケーションやユーザーのデータは、大規模なプライベートクラウドやパブリッククラウドの閉ざされた壁の向こう側にある集中型リポジトリに保存されています。複数の環境でアプリケーションデータを共有することは容易ではありません。さらに、データは通常1つの組織によって保存されるため、ユーザーは自分のデータをほとんど管理できません。

IPFSでは、一元制御に依存しないモデルであるため、あらゆるデータを異なるアプリケーションやエンドユーザー間で容易に分散・共有することができます。

図5：クラウドプロバイダーの背後にあるサイロ化されたデータ

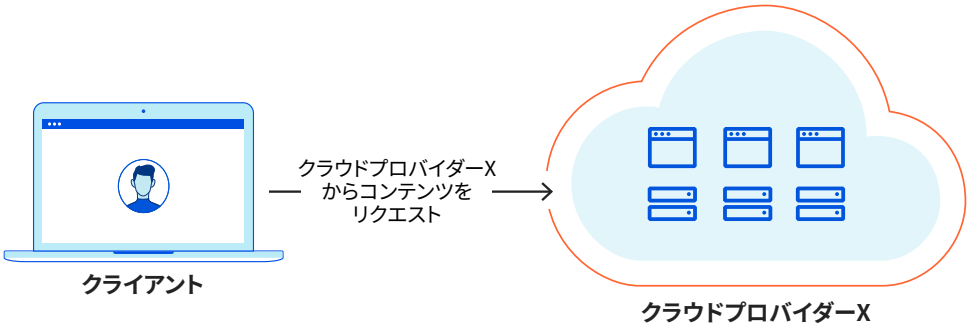
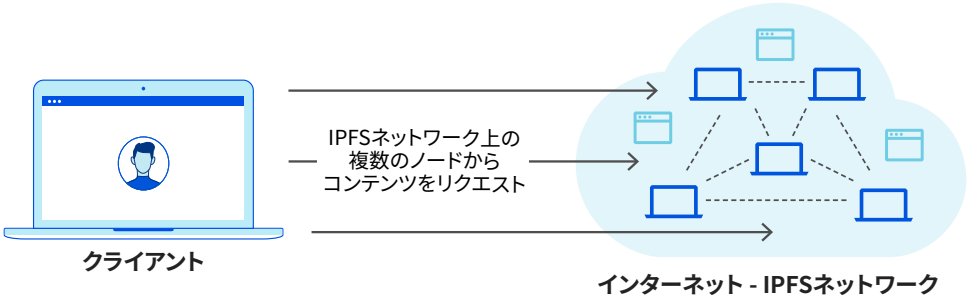


図6：IPFSネットワーク上のデータ



局所性/パフォーマンス

また、IPFSにはパフォーマンスのメリットもあります。

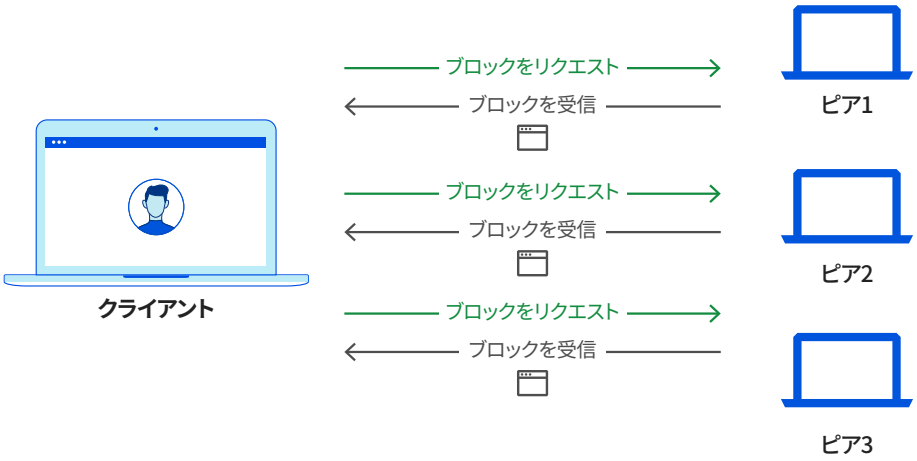
- IPFSは分散型ファイルシステムで構成される非集中型ネットワークであるため、コンテンツに対するすべてのリクエストを処理するチョークポイントとなる中央エンティティが存在しません。その代わり、コンテンツをキャッシュしているP2Pネットワーク内のどのノードでも、そのコンテンツを提供することができます。
- IPFSのようなP2Pネットワークは、一度に1つのサーバーからファイルをダウンロードするHTTPとは異なり、要求されたコンテンツのデータの断片を複数のノードから同時に取得するため、コストとレイテンシーの削減が期待できます。また、大容量のデータも重複することなく効率的に配信することが可能です。bitTorrentのようなP2Pプロトコルは、コンテンツを取得するために複数のノードを同時に活用することのパフォーマンス上のメリットを長い間示してきました。

以下の図7と図8は、集中型モデルとIPFSのようなP2Pモデルでダウンロードされるファイルを表しています。

なお、IPFSではBitswapという別のプロトコルを用いて、特定のコンテンツやオブジェクトのデータブロックを一意のハッシュで発見、リクエスト、受信しています。つまり、Bitswapはデータブロックを交換するためのコアモジュールから成るメッセージベースのプロトコルなのです。ネットワークの他のピアとの間で、データブロックのリクエストと送信を指示する役割を担っています。各IPFSノードは、受信したいブロックのリストをピアに送信します。ノードはブロックを受け取ると、そのブロックを欲しがっているピアがいるかどうかを確認し、そのピアにブロックを送ります。

図7：集中型Web（サーバーベースレスポンス）

図8：dWeb（P2Pレスポンス）



重複排除/効率的なストレージ

IPFSは本来、重複排除機能を備えており、ストレージリソースをより効率的に利用することができます。ここでIPFSが行うコアアクションはいくつかあります。

- IPFSは同じコンテンツを2度保存しません。すべてのコンテンツ/オブジェクトは一意のハッシュを持っています。
- IPFSはデータをブロックに分割し、重複排除の一環として、そのブロックのコンテンツが変更されていなければ、同じブロックを活用することができます。これは、下の図のように、更新されたコンテンツの最後のブロックだけがアップロードされることを示しています。
- IPFSに保存されているコンテンツ/データは不変であるため、更新には新しいコンテンツ/オブジェクトの作成が必要です。前述したように、IPFSではデータがチャンク化されて保存されるため、ある程度の重複排除が内在しています。また、アドオンやバージョン管理システム（VCS）を利用することで、コンテンツの更新時に変更のみをアップロードする強固なバージョン管理を実装することができ、IPFSのアーキテクチャはこれを可能にします。

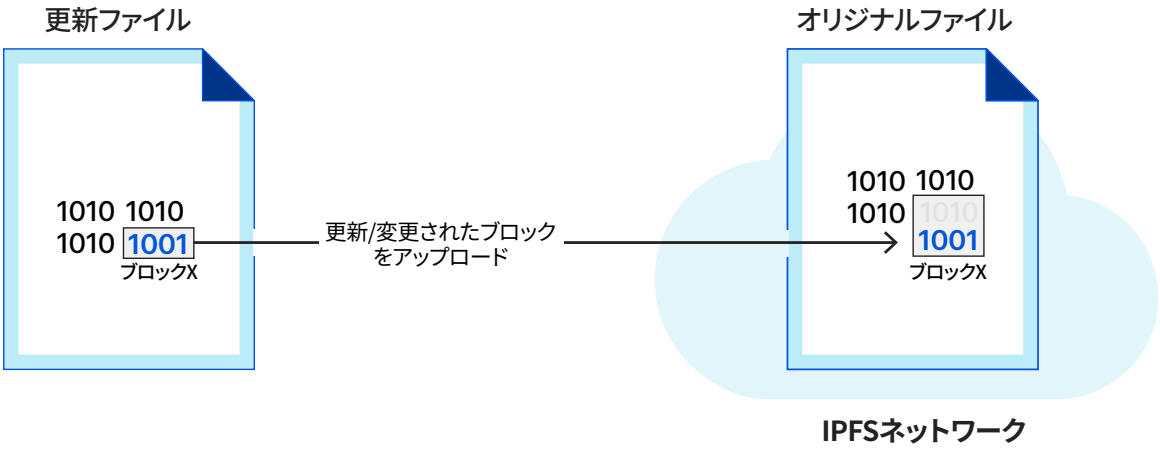


図9：IPFS-更新または変更されたブロックがアップロードされます

Cloudflare IPFSゲートウェイを選ぶ理由

現在、IPFSはHTTPのような成熟したレベルには達していません。例えば、主流のブラウザは、まだIPFSのネットワークとアドレス指定をネイティブにサポートしていません。しかし、このドキュメントで紹介するように、IPFSを利用するメリットや利点は非常に明確であり、多くの企業が利用し始めています。

Cloudflare IPFSゲートウェイは、現在のWeb2.0モデルとWeb3の新しい非集中型・トラストレスモデルをつなぐ橋渡しをします。Cloudflare IPFSゲートウェイを利用することで、お客様はIPFSの活用を開始し、IPFS上のコンテンツにHTTPで簡単にアクセスできるようになります。IPFSのコンテンツベースのアドレス指定の欠点は、ハッシュを覚えるのが大変なことです。この欠点を解消するために、お客様はCloudflare DNSを使用してIPFSコンテンツをドメイン名にマッピングすることができます。

つまり、Cloudflare IPFSゲートウェイを使えば、お客様はIPFSのメリットを享受しながらHTTPを使い続けることができるのです。従来のWebクライアントを使用する多くのユーザー層にも消費しやすいようになります。

Cloudflare IPFSゲートウェイのもう一つの大きなメリットは、Cloudflareのグローバルなエニーキャストエッジネットワーク上にあり、信頼性、セキュリティ、キャッシング、パフォーマンスなどの関連するメリットをすべて受け継ぐことができる点です。

Cloudflare IPFSゲートウェイの機能

Cloudflare IPFSゲートウェイは以下の機能を提供します。

1. IPFSコンテンツを簡単に取得する機能：お客様は独自のIPFSノードをデプロイしセキュリティを確保しなくても、IPFSネットワーク上のコンテンツに簡単にアクセスすることができます。Cloudflareのゲートウェイは、自社の耐障害性のある、セキュリティ強化されたネットワーク上のIPFSノードを利用して、IPFSコンテンツを取得します。Cloudflare IPFSゲートウェイは、IPFSの前にあるキャッシュと見なすことができます。Cloudflare IPFSゲートウェイは、IPFSネットワークからコンテンツを変更したり削除したりするために使用することはできません。

2. カスタムドメイン名を通じてIPFSコンテンツを提供する機能：IPFSはDNSを使用して、覚えやすい名前をIPFSコンテンツにマッピングすることをサポートしています。DNSLinkは、そのために使用されるプロトコルです。CloudflareはDNSLinkをサポートし、お客様がドメイン名を通じてIPFSコンテンツを提供することを可能にします。

3. IPFSコンテンツへのCDNの活用（キャッシング、パフォーマンス、信頼性）：Cloudflare IPFSゲートウェイを使用すると、Cloudflare CDNを使用するというさらなるメリットが得られ、ユーザーの近くでIPFSコンテンツをキャッシュできるため、全体のパフォーマンスを向上できます。

4. IPFSゲートウェイとCloudflareのセキュリティ、信頼性、パフォーマンス機能/製品の単一の管理画面：お客様はIPFSゲートウェイを使用し、フルセキュリティモデルと追加のCloudflareの信頼性とパフォーマンス機能/製品を単一の管理画面で管理することができます。

Cloudflare IPFSゲートウェイのメリット

Cloudflare IPFSゲートウェイを利用することで得られるメリットはいくつかあります。

1. アクセスの容易さ：特別なソフトウェアのインストールや実行を必要としない、IPFSネットワークからコンテンツにアクセスする簡単な方法です。

2. セキュリティ：ユーザーはソフトウェアをインストールせず、IPFSゲートウェイを活用するため、セキュリティの負担はグローバルなエニーキャストインフラを用いて、強化したセキュリティを提供するCloudflareに移りました。

Cloudflareが長年提供してきた現在のHTTP保護と同様に、Cloudflareは、HTTPインターフェースのコンテンツをフィルタリングします。これは、悪意のあるコンテンツの配信を防ぐのに役立ちます。

さらに、CloudflareはIPFSセーフモードを実装しました。これは、すべてのネイティブIPFSノードで使用されているIPFS実装の上にノード保護レイヤーで、ネイティブIPFSレベルで悪意のあるコンテンツの配信を防止する機能を提供するものです。IPFSセーフモードの詳細については、こちらからCloudflareのブログをご覧ください。

3. 維持/監視なし：CloudflareはIPFSネットワークへのゲートウェイを提供しているため、お客様がメンテナンスすることはありません。Cloudflareはセキュリティ、信頼性、パフォーマンスを維持・監視します。

4. 信頼性：CloudflareのIPFSゲートウェイのためのグローバルなエニーキャストネットワークは、高いレベルの信頼性と可用性を提供します。

5. パフォーマンス： IPFSゲートウェイは、100カ国以上にあるデータセンターから成るCloudflareのエッジネットワーク上にあるため、お客様はパフォーマンスのメリットを得ることができます。

IPFSコンテンツキャッシングでは、ユーザーに最も近いデータセンターからコンテンツを提供することができます。

Cloudflare IPFSゲートウェイのアーキテクチャとデザイン

Cloudflare IPFSゲートウェイは、IPFSにHTTPアクセス可能なインターフェースを提供し、ユーザーはIPFSのコンテンツを容易に取得することができます。下の図は、Cloudflare IPFSゲートウェイのアーキテクチャを表しています。

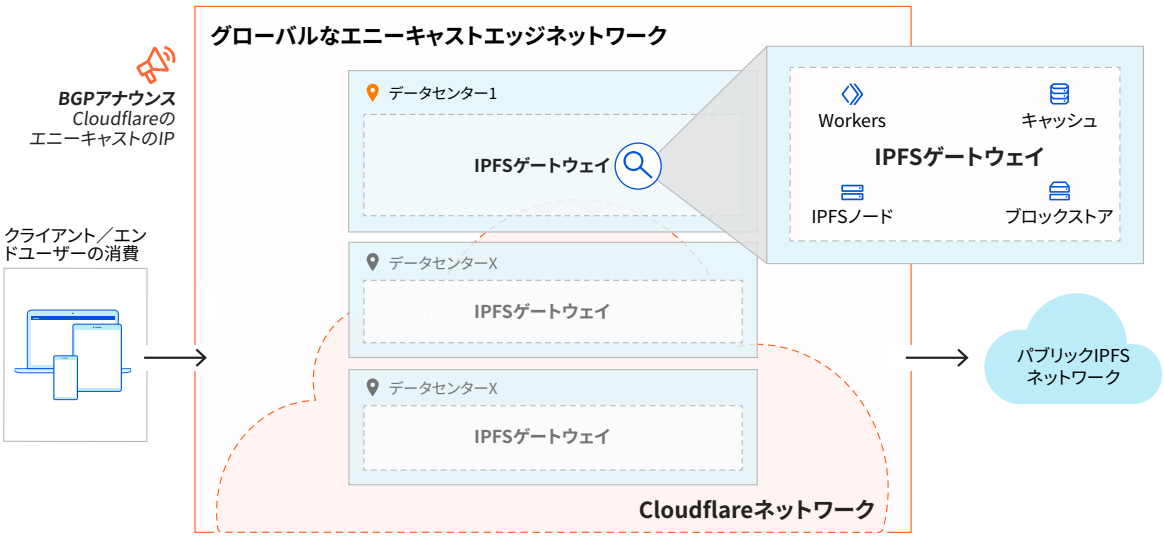


図 10：各データセンターのCloudflare IPFSゲートウェイ

いくつかの重要な項目があります：

- 各ドメインのDNS内に、ipfs.cloudflare.comを指すCNAMEが自動的に作成されます。
- IPFSゲートウェイは、関連するドメインのDNSLink TXTレコードの値を確認することで、指定されたドメインのリクエストを受け取ったときにどのコンテンツを提供すべきかを把握しています。
- Cloudflare IPFSゲートウェイは、Cloudflareデータセンター内のグローバルなサーバーで稼働するCloudflare Workersのコードコンポーネントを持っています。これらのデータセンターは100カ国以上にまたがり、世界中のインターネット接続人口の95%に50ミリ秒以内に到達し、100Tbpsのネットワーク容量とDDoS攻撃対策機能を提供します。
- IPFSノードは、すべてのCloudflareデータセンターで稼働します。Cloudflare Workersのコードはローカルのipfsノードにapiコールを行い、次にローカルのIPFSブロックストアにアクセスすることができます。
- ローカルデータセンター内のIPFSノードは、Cloudflareのネットワーク内の他のIPFSノードとピアリングします。必要に応じて、パブリックIPFSネットワークにアクセスすることができます。
- Cloudflareのネットワークはエニーキャストを活用しているため、ユーザーは常に自分に一番近いデータセンターにアクセスすることができます。さらに、エニーキャストネットワークにより高い可用性を確保し、問題が発生した場合は自動的に別のデータセンターへ再ルーティングします。

トラフィックフロー

図11は、トラフィックがCloudflare IPFSゲートウェイにヒットしたときのトラフィックフローを表しています。以下は、いくつかの重要な注意点です。

- お客様のDNSにあるCNAMEレコードは、Cloudflare IPFSゲートウェイのURLを指します。リクエストは、Cloudflareのエニーキャストネットワーク内で最も近いデータセンターとそれぞれのCloudflare IPFSゲートウェイに送信されます。
- Cloudflare IPFSゲートウェイはCloudflare Workersを使用し、ローカルのIPFSノードにAPIコールを行います。ローカルIPFSノードは、ローカルブロックストアからデータをリクエストします。コンテンツがローカルブロックストアにキャッシュされていない場合、Cloudflareのネットワーク内またはパブリックIPFSネットワーク内の他のIPFSノードからコンテンツを取得することができます。
- Cloudflare IPFSゲートウェイはIPFSノードからレスポンスを受け取ると、コンテンツをローカルにキャッシュし、最初のリクエストに応答します。

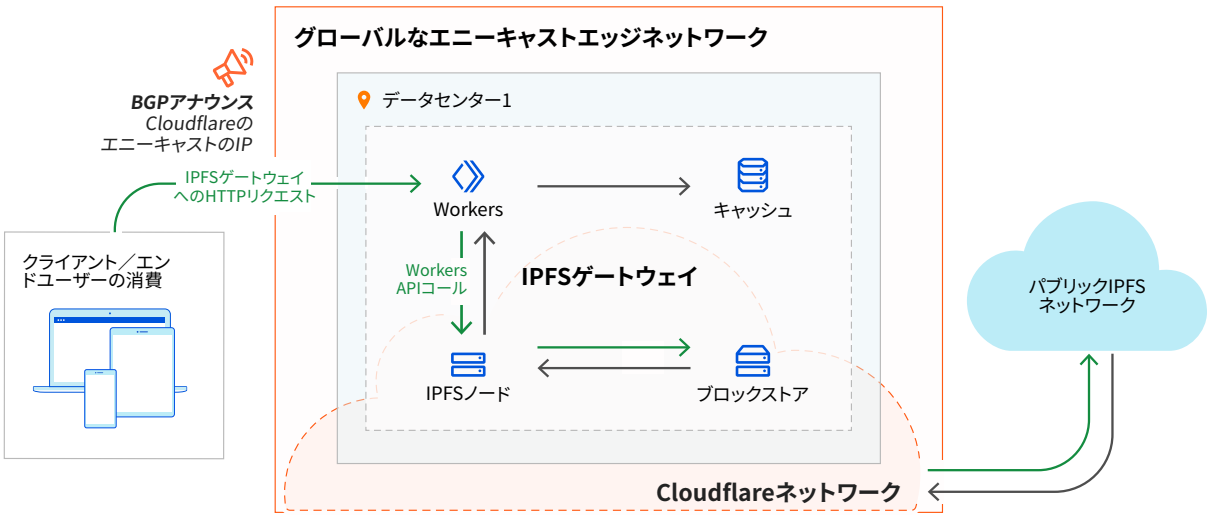


図11：Cloudflare IPFSゲートウェイのトラフィックフロー

IPFSコンテンツのキャッシング

Cloudflareは、そのグローバルネットワークとキャッシングインフラを活用して、ユーザーに最も近いIPFSコンテンツをキャッシュし、パフォーマンスを向上させます。上の図に示すように、IPFS WorkersはIPFSコンテンツを一度受信するとキャッシュします。図12は、同じIPFSコンテンツに対して新しいリクエストが来たときのトラフィックフローを示したものです。IPFSコンテンツは前のクライアントリクエストからすでにキャッシュされているため、Workersはローカルipfsノードに再度APIコールを行う必要がないことに注意してください。これによりレイテンシーが短縮され、全体的なパフォーマンスが向上します。

Cloudflare IPFSゲートウェイはWorkers APIを利用してCloudflareのキャッシュに書き込みます。キャッシュされたアイテムの最大保持期間は、1日または86,400秒です。Cloudflare IPFSゲートウェイはコンテンツリクエストを受けると、常にDNSLinkレコードとそれぞれのハッシュをチェックし、キャッシュされたコンテンツが最新であることを確認します。

Workers APIはCloudflare Cacheにコンテンツをキャッシュし、ローカルIPFSノードもそのローカルブロックストアにコンテンツをキャッシュします。ブロックストア内のIPFSコンテンツの最大保持期間は30日です。

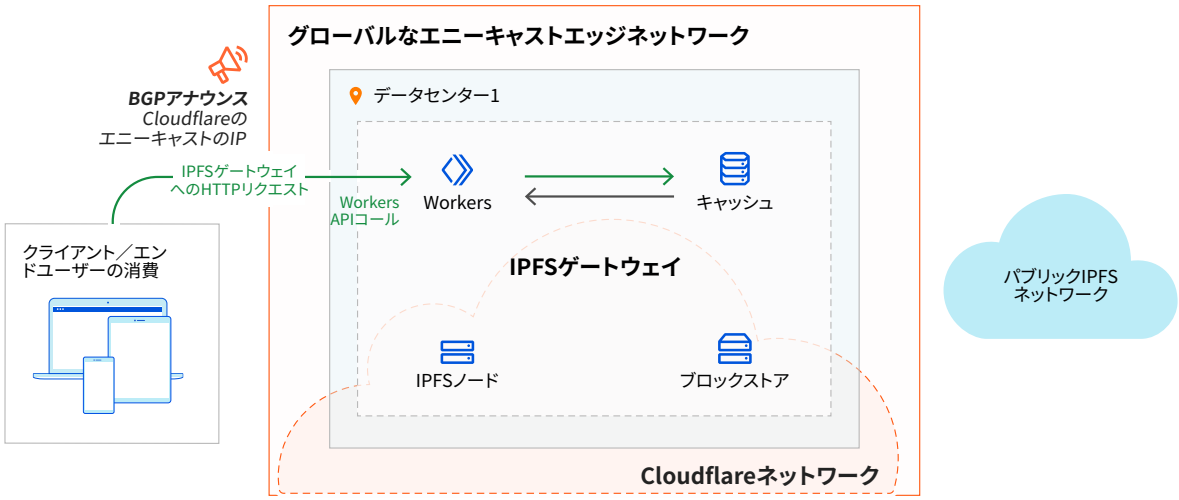


図12：Cloudflare IPFSゲートウェイは、同じコンテンツに対する2回目のリクエストでローカルキャッシュを使用します

まとめ

IPFSはWeb3のストレージレイヤーです。非集中型ホスティングとトラストレスモデルを有効にすることで、現在のWeb 2.0モデルのいくつかの課題を解決しています。Cloudflare IPFSゲートウェイは、HTTPでアクセス可能なIPFSゲートウェイを採用することでWeb2.0とWeb3の橋渡しをし、ユーザーは簡単にIPFSコンテンツを取得し、カスタムドメインを通じてIPFSコンテンツを提供することができます。Cloudflare IPFSでは、IPFSを活用するためにIPFSノードの導入や監視を行う必要はありません。さらに、お客様はCloudflareのネットワークの信頼性、セキュリティ、パフォーマンスの恩恵を受けることができます。