Nokia Deepfield Cloud Genome

Mapping the internet services supply chain in the cloud and IoT era

Nokia Deepfield Cloud Genome® maintains the "supply chain" map of the internet to provide unprecedented visibility about internet applications, content and services

Deepfield Cloud Genome is a global, cloud-based data feed about internet applications, content and services – a dynamic "supply chain" map of the internet. It is a result of Nokia's patented technology that continuously scans, probes and tracks over five billion IPv4 and IPv6 addresses on the internet. IP addresses are tagged into different types and categories using Domain Name Server (DNS) names and advanced machine learning (ML) rules.

Cloud Genome informs Deepfield applications like Deepfield Cloud Intelligence and Deepfield Subscriber Intelligence, which correlate the Cloud Genome information with the information obtained from the network, facilitating full network and service visibility down to the service/application, content delivery network (CDN), peering/transit router, backbone node, access system and device/subscriber.

Cloud Genome is a component of the Nokia Deepfield Genome, a set of two complementary and proprietary data feeds that employ Nokia patented technology:

- Deepfield Cloud Genome™: Provides full visibility of internet content, applications and services.
- Deepfield Secure Genome™: Provides full visibility of internet DDoS security-related data.

Deepfield Genome is an intrinsic part of the Nokia Deepfield portfolio of IP network intelligence, analytics and distributed denial of service (DDoS) security applications for service providers (cable providers, multiple system operators (MSOs), telecommunications service providers), webscale companies, internet exchange points (IXPs) and large digital enterprises.

Features

- Maps and tracks more than 5 billion IPv4 and IPv6 endpoints, crawling hundreds of millions of IP addresses daily
- Keeps track of the most relevant (150-200 million) IP addresses for internet content, applications and services and tags them against content delivery networks (CDNs), servers and hosting companies
- Employs more than 100 advanced ML rules for automatic classification of internet traffic
- Leverages publicly available third-party databases to enhance accuracy
- Updated hourly
- Provides input to Deepfield applications (e.g., Deepfield Cloud Intelligence and Deepfield Subscriber Intelligence) with information about internet content, applications and services.

Benefits

- Provides an up-to-date snapshot of the "supply map" of the internet: where internet services originate and how they traverse the internet to reach service provider networks and their subscribers
- Categorizes applications and flows into diverse traffic types and categories
- Facilitates correlation of network-specific information with the larger internet context to provide end-to-end visibility of internet services and application flows.

How Cloud Genome works

The internet is continuously growing—both in size and its variety of content and services. Service providers, webscale companies and large digital enterprises want better insights about their networks and the services they offer to their subscribers. Today, most of these services flow into and across their networks from the large application and content domains and CDNs. To understand how services are delivered, operators need full insight into their network's traffic flows and the ability to correlate this information with a "big picture" and details about how internet services, applications and content are delivered—from the originating domains to end systems and subscribers.

Network analytics solutions that ingest only flow-based (e.g., sFlow, CFlow, J-Flow, and IPFIX) and Border Gateway Protocol (BGP) data, relying on just IP addresses and BGP routes, can't fully understand and categorize internet traffic and are not up to the task of detailed traffic analysis and reporting that is required in the cloud era.

Legacy approaches with dedicated hardware probes and using deep packet inspection (DPI) technology are also inefficient when dealing with extreme volumes of traffic which may be encrypted.

These legacy approaches may not be scalable or cannot provide the required end-to-end service visibility.

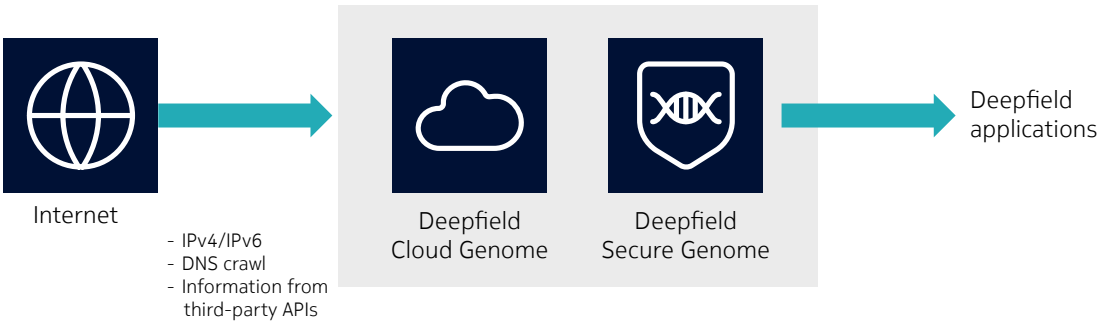
By contrast, Cloud Genome continuously crawls the internet and maps billions of IP addresses and millions of cloud applications and services. Cloud Genome categorizes traffic across a large number of distinct traffic categories to provide a rich and detailed understanding of the global internet services supply chain (see Figure 1), including:

- CDN infrastructure
- Major internet providers
- Autonomous System Numbers (ASNs)
- Hosting companies
- Internet of Things devices
- GeoIP information (countries/regions/cities).

Cloud Genome relies on a network of distributed, cloud-based, data-mining agents. These agents constantly examine and interact with internet endpoints to learn about the services and the traffic patterns and then build a dynamic map of the entire internet. Cloud Genome is used to complement and augment network information and provide a level of visibility into internet applications and services that was previously unattainable.

Cloud Genome enables an end-to-end view across an entire infrastructure and even of encrypted traffic without using a single network probe or DPI.

Figure 1. Deepfield Genome



Providing context to the Deepfield applications

Cloud Genome provides context to the Deepfield applications:

- Nokia Deepfield Cloud Intelligence: Delivers end-to-end network visibility for traffic engineering, capacity planning and service optimization.
- Nokia Deepfield Service Intelligence: Monitors QoE for over-the-top (OTT) video services in real time.
- Nokia Deepfield Subscriber Intelligence: Provides service visibility with end-user and subscriber-level granularity.

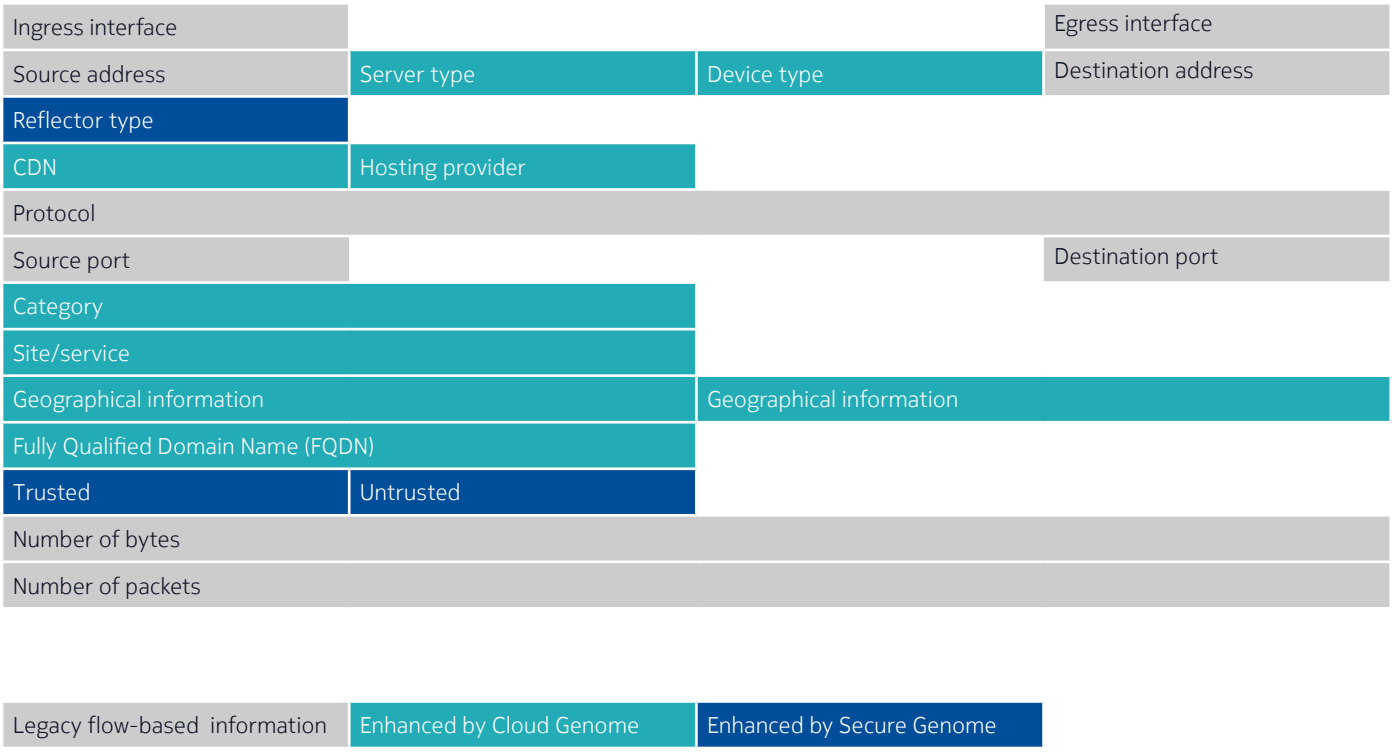
An example of how Deepfield Genome knowledge enhances legacy flow-based information is shown in Figure 2.

Cloud Genome identifies the endpoints for all network traffic traversing the internet, providing visibility down to the CDN, site and service. This level of detail gives Deepfield Cloud Intelligence the ability to accurately classify all traffic across the entire infrastructure, arming network operators with knowledge on how to make the best decisions about where to, for example, add bandwidth or add network caches to ensure the best network performance.

Because Cloud Genome is a data feed with a holistic view of all internet applications and services, it allows Deepfield Service Intelligence to have a complete, real-time view of over-the-top (OTT) video services and to monitor the aggregate number of streams and average bit rates (ABRs) delivered to any part of the network or any location or subscriber.

Cloud Genome also empowers Deepfield Subscriber Intelligence to allow full visibility into how services are delivered to the access layer with subscriber-level granularity and facilitates detailed insights on service consumption patterns and customer experience.

Figure 2. Enhancing legacy flow-based information with Deepfield Genome knowledge



Cloud Genome is instrumental in enabling diverse use cases for Deepfield applications, including:

- Peering analytics and optimization
- Backbone engineering
- Traffic engineering and capacity planning
- Content delivery network analytics
- Traffic analysis by access type
- Streaming video insights for improved customer QoE
- Per-service plan/subscriber traffic analytics and content consumption insight
- Customer care
- Diverse marketing-focused use cases which require better knowledge of subscriber online preferences and behavior.

The Nokia Deepfield advantage

Nokia Deepfield is a software suite of network analytics and DDoS security applications for large-scale IP networks. These applications optimize networks and services, enhance customer experience, improve network security and increase operational agility.

Deepfield applications are deployed globally in many networks, including fixed and mobile service providers, cable companies, cloud companies, and digital enterprises.

Deepfield's approach uses big data IP analytics, combining network data (telemetry, DNS, BGP etc.) with Nokia's patented Deepfield Genome technology (live feed that tracks internet content, applications and services and provides DDoS security context). As a result, the Deepfield applications offer multi-dimensional, real-time insights about IP-based services and applications running across the entire IP network - from content-originating domains and CDNs, across the peering and backbone to the customer edge.

The real-time cloud and network context provided by the Nokia Deepfield applications enables service providers to extract the actionable intelligence needed to design their networks better, react to performance anomalies and changing traffic patterns, manage security threats, and better package their product offerings to attract and retain subscribers.

The Nokia Deepfield portfolio enables service providers to understand, in real time, the service delivery path from the internet/cloud through the peering edge and at the customer edge—a path that can span multiple clouds, data centers, CDNs and networks.

This visibility is the critical first step to intelligent network automation to enable networks to respond immediately to changing conditions with minimal manual or physical intervention, lowering costs and improving performance.

To learn more about the Deepfield solution, visit the Deepfield web page.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: CID201108 (August)