Scanner & Agent - Technical and Executive Overview

Executive Summary

The Scanner & Agent Platform is a modern, lightweight scanning system designed to assess and visualize web and infrastructure environments. It has two modules:

1. URL Scanner - analyzes public endpoints and extracts technology stack and hosting patterns.

2. Agent (VM) Scanner - deployable lightweight agents that scan internal systems, services, and ports.

It unifies visibility between external and internal systems, ideal for DevOps, Cloud, and Security teams.

Architecture Overview

Core Components:

- Frontend: HTML + JS dashboard for initiating scans and visualizing results.
- Backend: FastAPI service managing URL scans, agents, and results storage.
- URL Scanner: Uses async HTTP calls to inspect headers and infer hosting patterns.
- Agent Scanner: Collects package/service/port data and sends back JSON via REST API.
- Visualizer: Converts scan data into architecture diagrams and IaC recommendations.

Data Flow:

1. URL Scanner Flow -> POST /scan -> analyze -> derive plan -> return JSON -> frontend displays flow.

2. Agent Flow -> Create source -> deploy agent -> heartbeat -> scan job -> upload results -> visualize.

Benefits of Using This Tool

1. Unified external + internal visibility
2. Auto-generated architecture diagrams and IaC templates
3. Cost-effective and lightweight (no heavy infrastructure)
4. API-driven for automation and CI/CD integration
5. Can run both online and offline with self-hosted backend

Benefits in Enterprise Context

Traditional	Scanner & Agent
Manual curl/nmap	Automated header parsing
SSH scripts for inventory	Lightweight agent JSON reports

Scanner & Agent - Technical and Executive Overview

| Manual diagramming | Auto architecture visualization |

| Manual IaC setup | Auto Dockerfile & Terraform hints |

| Delayed security checks | Continuous agent & API scans |

Enterprises gain speed, standardization, and compliance-ready visibility.

Future Scope

1. Persistent DB (SQLite/PostgreSQL/Redis)
2. Vulnerability and compliance scanning
3. D3.js visualizations and dashboards
4. Plugin-based agent modules
5. Cloud SDK integration (AWS, GCP, Azure)
6. ML-based misconfiguration detection
7. SSO + Multi-tenant enterprise support

JSON Schema Overview

URL Scan Response:

```
{
  "request": {"url": "<target>"},
  "response": {"status": 200, "headers": {...}},
  "plan": {"selected_plan": "cdn_fronted_site", "diagram": [...]},
  "tls": {"verify": true, "using_certifi": true}
}
```

Agent Scan Result:

```
{
  "inventory": {"packages": {...}, "services": {...}, "ports": {...}},
  "plan": {"selected_plan": "python_wsgi_to_container_or_appsvc", ...}
}
```

How Frontend Fetches Data

Frontend calls (using Fetch API):

- POST /scan -> get JSON -> render diagram & IaC.

Scanner & Agent - Technical and Executive Overview

- POST /sources -> create agent installer.
- POST /scanHost -> start job.
- GET /scanJobs/{id}/status -> poll progress.
- GET /scanJobs/{id}/results -> fetch and display JSON results.