HawkEye

Performance Monitoring Solution for Data Centers

Shabbir Suterwala

Insight DE Fellow

HawkEye

- Monitors performance for hardware, software and apps running in a datacenter.
  - Hardware could be servers, routers/switches etc.
  - Software can be OS, DB Servers, web-servers, lambda stack etc.
  - Applications can be any custom apps deployed in the data center.
- Users can register for hardware, software, apps that they are interested in.

Data

- Data Points; time stamped:
  - TCP/IP packet In/Out info
  - ~~CPU/DiskIO/Mem/NetIO~~-~~consumption info~~
- Engineered Data for Insight Project
- Real World
  - Hook into kernel: network stack, scheduler
  - Hi priority demon using *top*

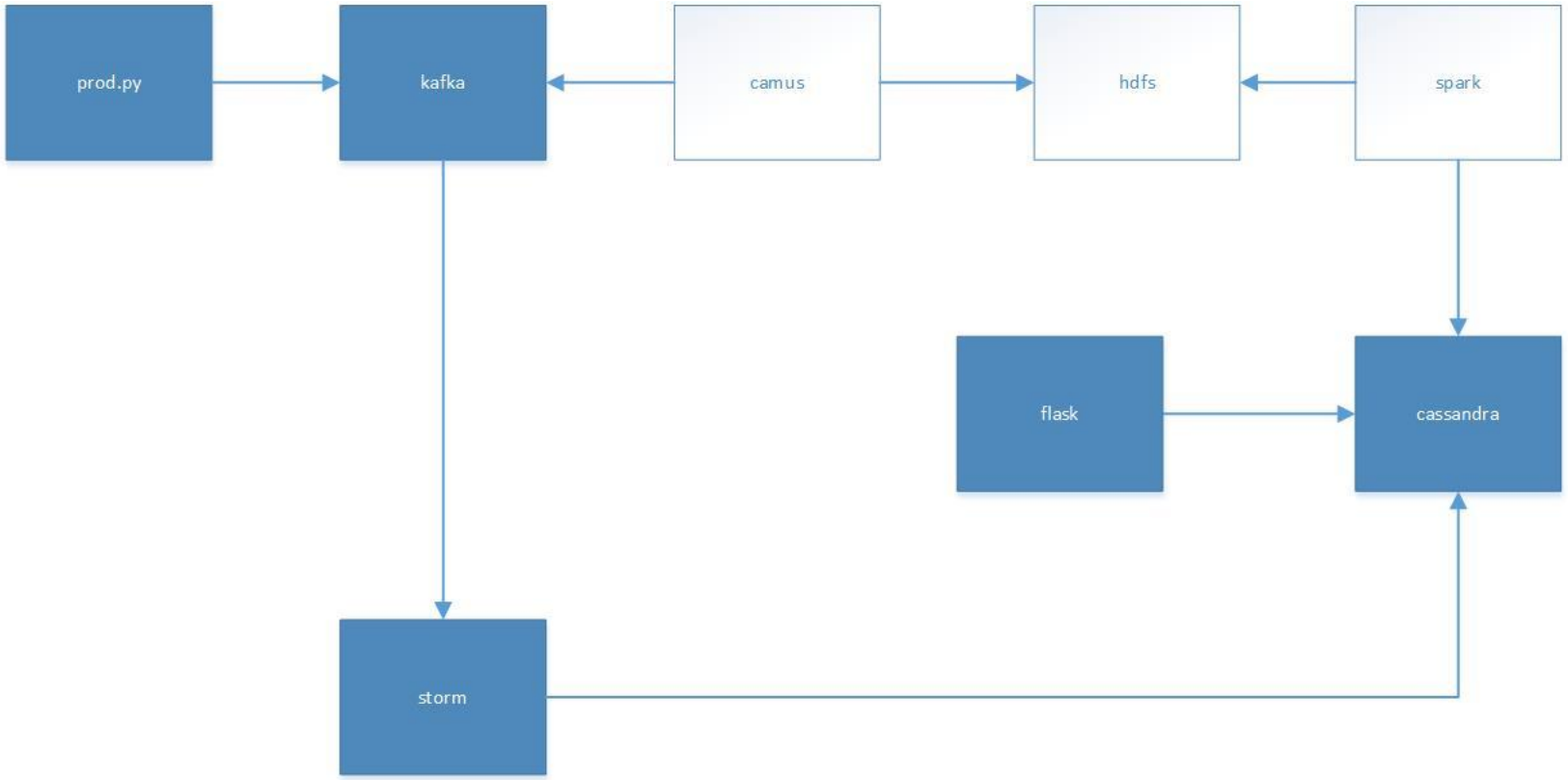
Data

```
{
        //Timestamp in/out
        "TsIn": 1453407175613828,
        "TsOut": 1453407175614662,
        //Monitors:
        "SwType": "SWTYPE42",
        "SwID": "SWID20",
        "TaskType": "TASKTYPE217",
        "HwID": "HWID85730",
        "HwType": "HWTYPE48",
        "TaskID": "TASKID154",
        "AppID": "HawkEye",
        //Track packets across the network
        "PacketID": "PACKET79217“
}
```

Queries

- MVP
  - List N monitors
- Final
  - Which of my hardware, software, apps are not performing now - have a delay of greater than threshold
  - Top N bottlenecks hotspots over time T for hardware, software and apps individually or in a combination
  - Which N hardware, software, app combination produce bottlenecks and where.

Pipeline



Challenges

- Stitching the pipeline
- Making spark and Cassandra talk
- Cassandra schema / design
- Context switching
  - Sessions / talks etc

Shabbir Suterwala

- Team Lead, Architecture / Principal Architect @ Infor
  - Cloverleaf → Ingestion middleware for healthcare products.
- Previously worked at Cisco, AMD and Storage Startup