Troubleshooting the performance of complex production software is challenging:
- Profilers and logging only reveal **what** events happened
- Users must manually infer **why** those events affected performance
- X-ray automates this task by attributing performance to specific root causes

X-ray uses **performance summarization**

**Step 1:** Attribute costs to specific instructions and system calls

- `write 100 bytes`
  - Disk cost: **100**
- `write 200 bytes`
  - Disk cost: **200**

**Step 2:** Determine root cause using taint-tracking (ConfAid)

```java
if (X) {
    // Execution depends on
    // option1 & option2 with
    // probabilities 0.5 & 0.2
    write(100);
}

if (Y) {
    // Execution depends on
    // option2 with
    // probability 0.2
    write(200);
}
```

**Step 3:** Sum the costs over specific root causes

- Option 1: costs
  - `option1 costs`
  - `0.5 * 100 bytes = 50 bytes`
- Option 2: costs
  - `option2 costs`
  - `0.2 * 100 bytes = 20 bytes`

- Option 2: costs
  - `option2 costs`
  - `0.2 * 200 bytes = 40 bytes`

- Option 1 costs: **60 bytes**
- Option 1 costs: **50 bytes**

**Differential Performance Analysis**

- Identify why two requests differed in performance
- Identify where requests diverged and probabilities of root causes of the divergence
- Assign performance cost to each branch
- Calculate differential costs
- Calculate the costs of the root causes of divergences

**Results**

- Analyzed performance issues in Apache, Postfix, and PostgreSQL
- In 12 of 14 cases, X-ray identifies the correct root cause as the largest contributor
- Online overhead of 1 – 7% due to deterministic replay

**Future Work**

- Simultaneously compare thousands of requests to look for anomalies
- Better understand how workload distribution affects performance
- Graphical tools for visualizing performance issues