Studying Systems as Artifacts

Adam Oliner
Department of Computer Science
Stanford University

Today's Menu

1. Goal
2. Recipe
3. First Steps
4. Recent Results
5. Promising Direction

(1)

Goal

Imperfect Systems

- We build imperfect systems
- We will **always** build imperfect systems
- These imperfections are expensive

Vision Statement

We require generic techniques to **measure** and **analyze** systems s.t. we can detect, isolate, fix, mask, or mitigate imperfections.

(2)

Recipe

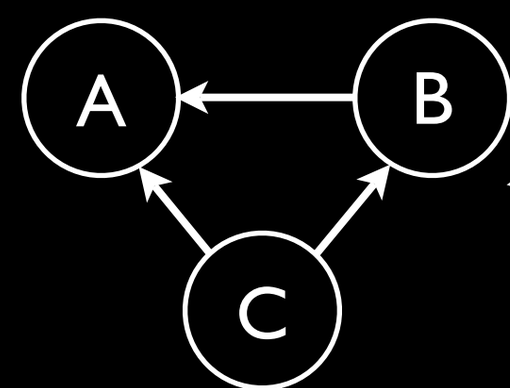
System Design Cycle

build

measure

$$f(t;\lambda) = \lambda e^{-\lambda t}$$

```
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.246746 R12-M1-N3-
C:J11-U11 RAS KERNEL INFO CE sym 7, at 0x03ffa060, mask 0x04
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.429303 R12-M1-N3-
C:J11-U11 RAS KERNEL INFO total of 3 ddr error(s) detected
and corrected
R33-M1-L0-U22-C 2005-06-03-23.10.14.414021 R33-M1-L0-U22-C
RAS LINKCARD INFO MidplaneSwitchController performing bit
sparing on R33-M1-L0-U22-C bit 3
R23-M1-N8-I:J18-U11 2005-06-04-00.24.32.398284 R23-M1-N8-
I:J18-U11 RAS APP FATAL ciod: failed to read message prefix
on control stream (CioStream socket to 172.16.96.116:33399
```



improve

analyze

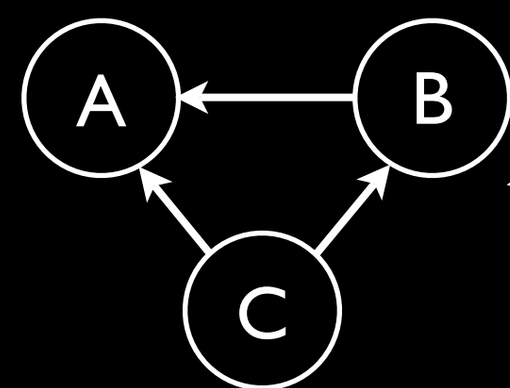
System Design Cycle

build

measure

$$f(t;\lambda) = \lambda e^{-\lambda t}$$

R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.246746 R12-M1-N3-C:J11-U11 RAS KERNEL INFO CE sym 7, at 0x03ffa060, mask 0x04
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.429303 R12-M1-N3-C:J11-U11 RAS KERNEL INFO total of 3 ddr error(s) detected and corrected
R33-M1-L0-U22-C 2005-06-03-23.10.14.414021 R33-M1-L0-U22-C RAS LINKCARD INFO MidplaneSwitchController performing bit sparing on R33-M1-L0-U22-C bit 3
R23-M1-N8-I:J18-U11 2005-06-04-00.24.32.398284 R23-M1-N8-I:J18-U11 RAS APP FATAL ciod: failed to read message prefix on control stream (CioStream socket to 172.16.96.116:33399



improve

analyze

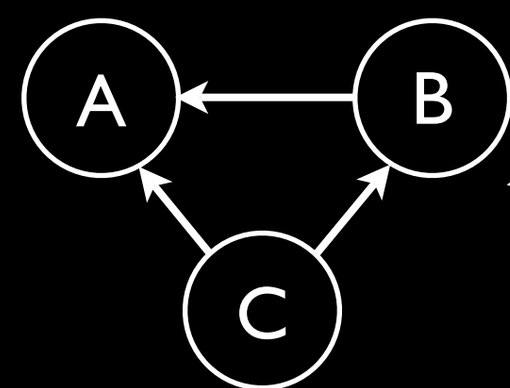
System Design Cycle

build

measure

R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.246746 R12-M1-N3-C:J11-U11 RAS KERNEL INFO CE sym 7, at 0x03ffa060, mask 0x04
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.429303 R12-M1-N3-C:J11-U11 RAS KERNEL INFO total of 3 ddr error(s) detected and corrected
R33-M1-L0-U22-C 2005-06-03-23.10.14.414021 R33-M1-L0-U22-C RAS LINKCARD INFO MidplaneSwitchController performing bit sparing on R33-M1-L0-U22-C bit 3
R23-M1-N8-I:J18-U11 2005-06-04-00.24.32.398284 R23-M1-N8-I:J18-U11 RAS APP FATAL ciod: failed to read message prefix on control stream (CioStream socket to 172.16.96.116:33399

$$f(t;\lambda) = \lambda e^{-\lambda t}$$



improve

analyze

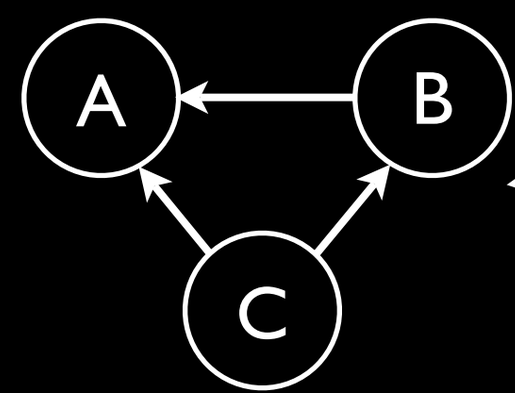
System Design Cycle

build

measure

$$f(t;\lambda) = \lambda e^{-\lambda t}$$

```
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.246746 R12-M1-N3-
C:J11-U11 RAS KERNEL INFO CE sym 7, at 0x03ffa060, mask 0x04
R12-M1-N3-C:J11-U11 2005-06-03-23.06.32.429303 R12-M1-N3-
C:J11-U11 RAS KERNEL INFO total of 3 ddr error(s) detected
and corrected
R33-M1-L0-U22-C 2005-06-03-23.10.14.414021 R33-M1-L0-U22-C
RAS LINKCARD INFO MidplaneSwitchController performing bit
sparing on R33-M1-L0-U22-C bit 3
R23-M1-N8-I:J18-U11 2005-06-04-00.24.32.398284 R23-M1-N8-
I:J18-U11 RAS APP FATAL ciod: failed to read message prefix
on control stream (CioStream socket to 172.16.96.116:33399
```



improve

analyze

Recipe

1. Measure
2. Analyze
3. Improve
4. Generalize

Why HPC?

- Canary in the coal mine
- Twin studies
  - Controls an important variable
  - Expectation of similarity

(3)

First Steps

My First Steps

- If we could predict failures...
  - System management [KDD 03]
  - Job scheduling [IPDPS 04]
  - QoS [DSN 05]
  - Checkpointing [IPDPS 06][ICS 06]

Dubious Predicates

- If failures are predictable,
- If failures are not i.i.d.,
- If failures are evident in logs,
- then...

(4)

Recent Results

Real Systems

- Collected and released a data set
  - Five major supercomputers
  - 750+ million CPU hours
- Initial analysis [DSN 07]

Bad News

- Logs are huge, unstructured, time-varying, incomplete, redundant, erroneous, and ambiguous

```
NULL RAS BGLMASTER FAILURE ciodb
exited normally with exit code 0
```

Good News

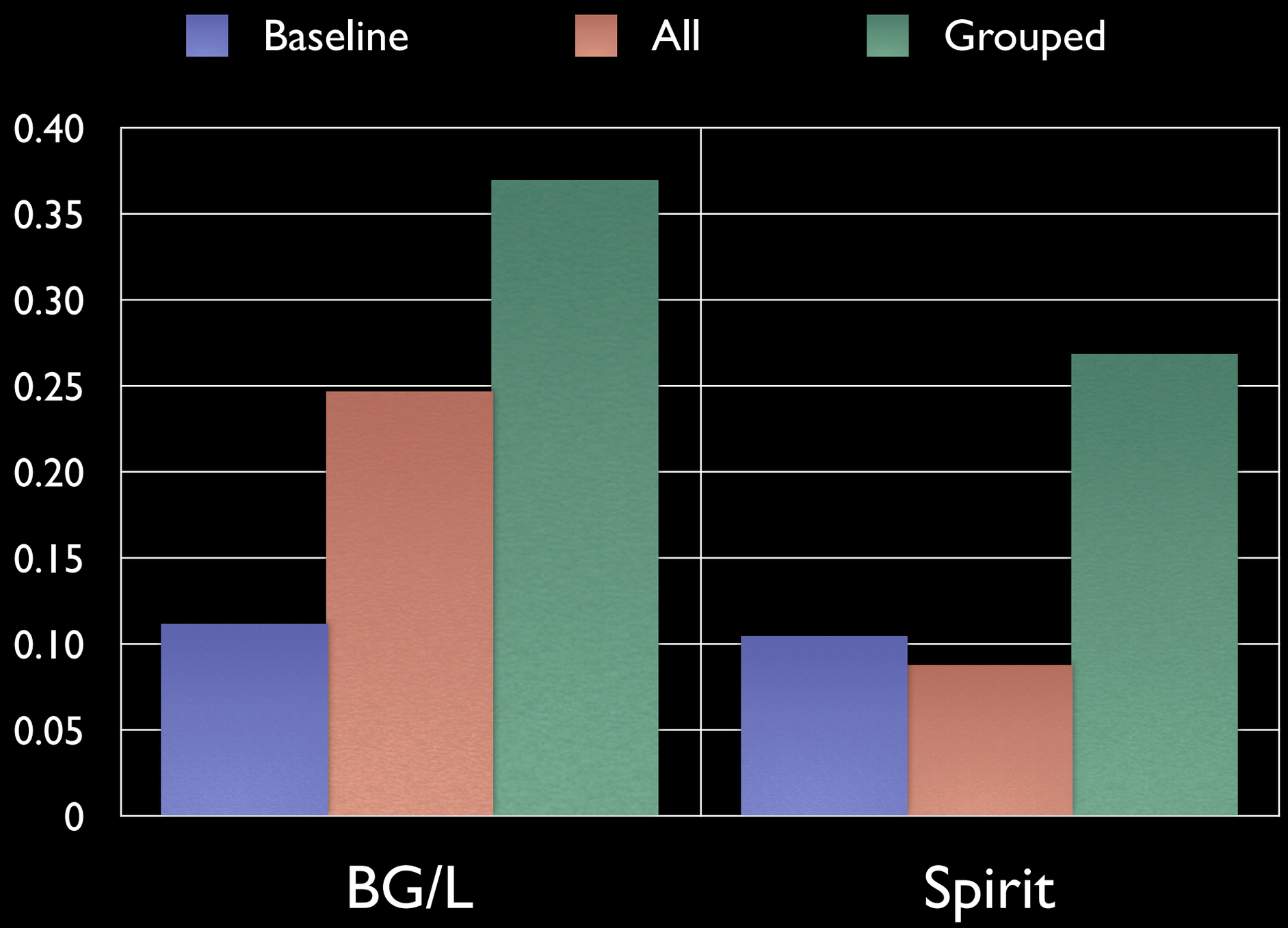
- Logs are rich with information
- Failures not i.i.d.
- Evidence of predictability

Related Studies

- Schroeder and Gibson [DSN 06]
- Liang et al [DSN 06]
- Stearley [ICCC 04]

Alert Detection

- Alert detection [Resilience 08] [ICDM 08]
- Production impact
- Similarity is valuable



Area Under the P-R Curve

(5)

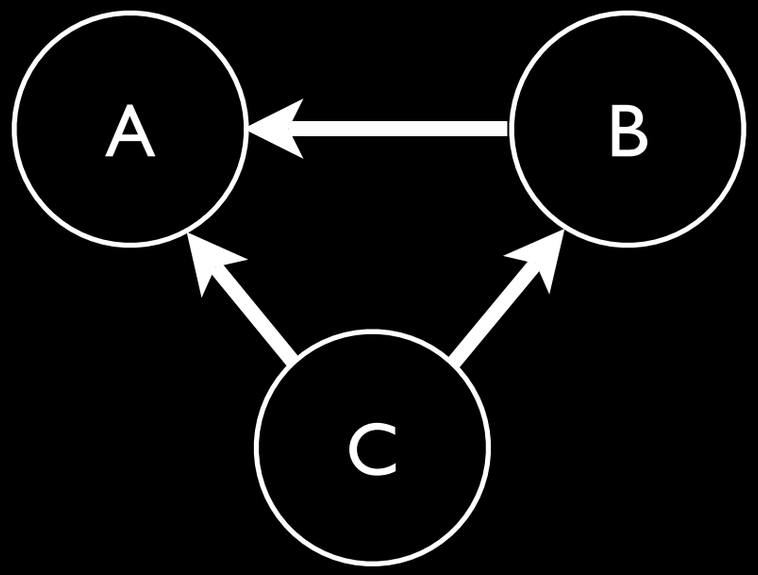
Promising Direction

Key Questions

- If I perturb X, what will happen?
- Why did that happen?

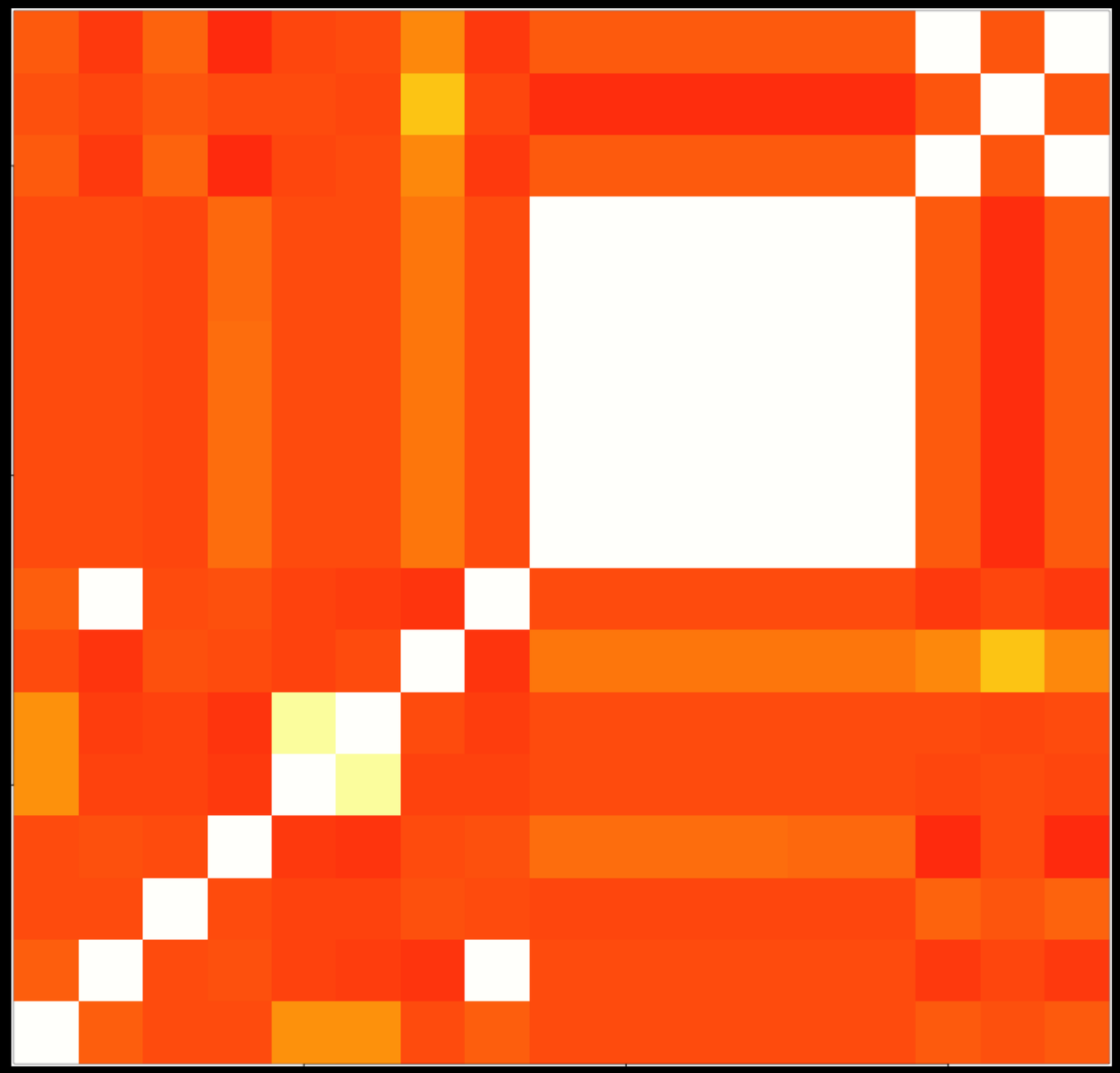
Key Questions

- If I perturb X, what will happen?
- Why did that happen?



Correlated Anomalies

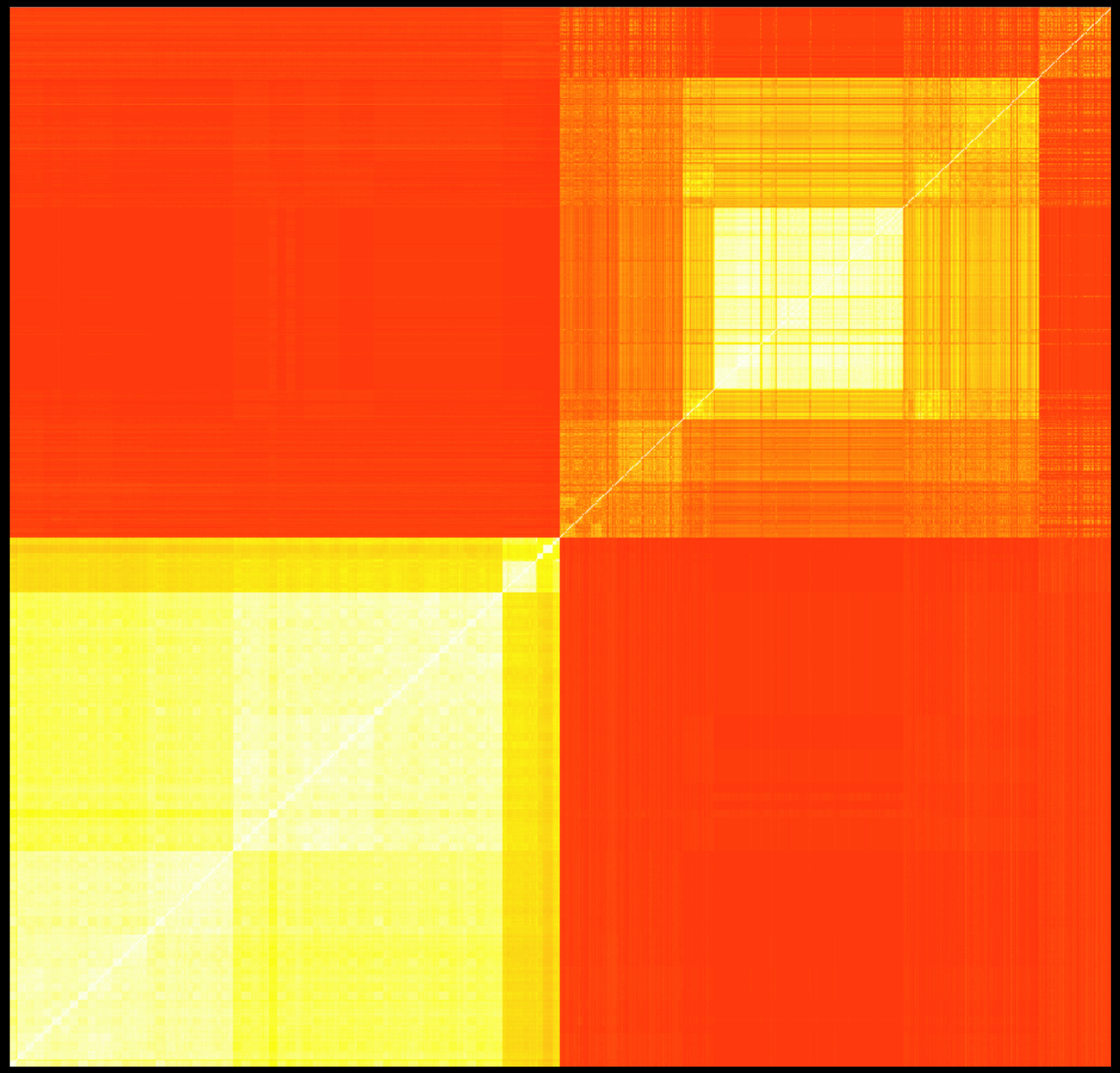
1. Identify anomalies
2. Correlate in time/space
3. Infer dependencies



Correlation Map

Early Results

- Grand Challenge
  - Built model from race logs
  - Found a tough bug
- Syzygy
  - Community intrusion detection
  - Asymptotically perfect



BG/L Correlations

Contributions

- System management techniques
- System logs
- Current direction
  - Correlated anomalies
  - Dependency graphs

Fin

You have reached the end of the presentation. Please turn back.