A Study of Soft Error Consequences in Hard Disk Drives

Timothi Tsai, Nawanol Theera-Ampornpunt, Saurabh Bagchi

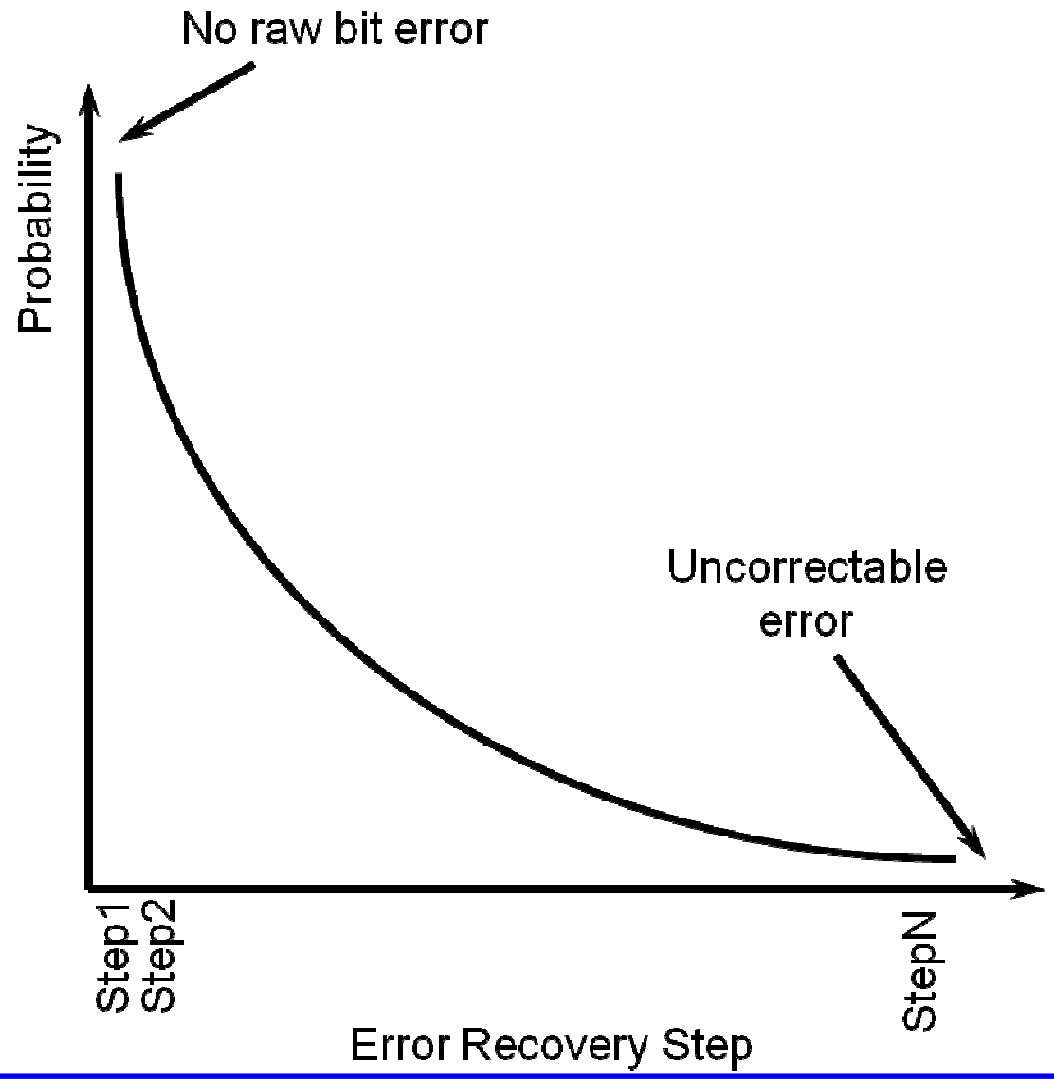
Dependable Systems and Networks 2012 (DSN '12)

PURDUE UNIVERSITY

Introduction

- Hard drives often hold critical data
  - Without redundancy, drive failures mean data loss
  - With redundancy (e.g., RAID), drive failures lead to downtime
- Internal error recovery mechanisms
  - Error correction code
  - Retries
- Corrected errors are called “soft” errors
- Uncorrectable errors are called “hard” errors

Typical distribution of successful error recovery steps



Introduction

- Failure prediction is an approach that attempts to mitigate potential loss of data or downtime
- When redundancy is used:
  - Cost of a false positive: downtime during integration of new drive
  - Cost of a false negative (missed hard error): small chance of data loss, and downtime if the storage subsystem decides that the drive should be taken offline (e.g., to perform extensive diagnostic check)

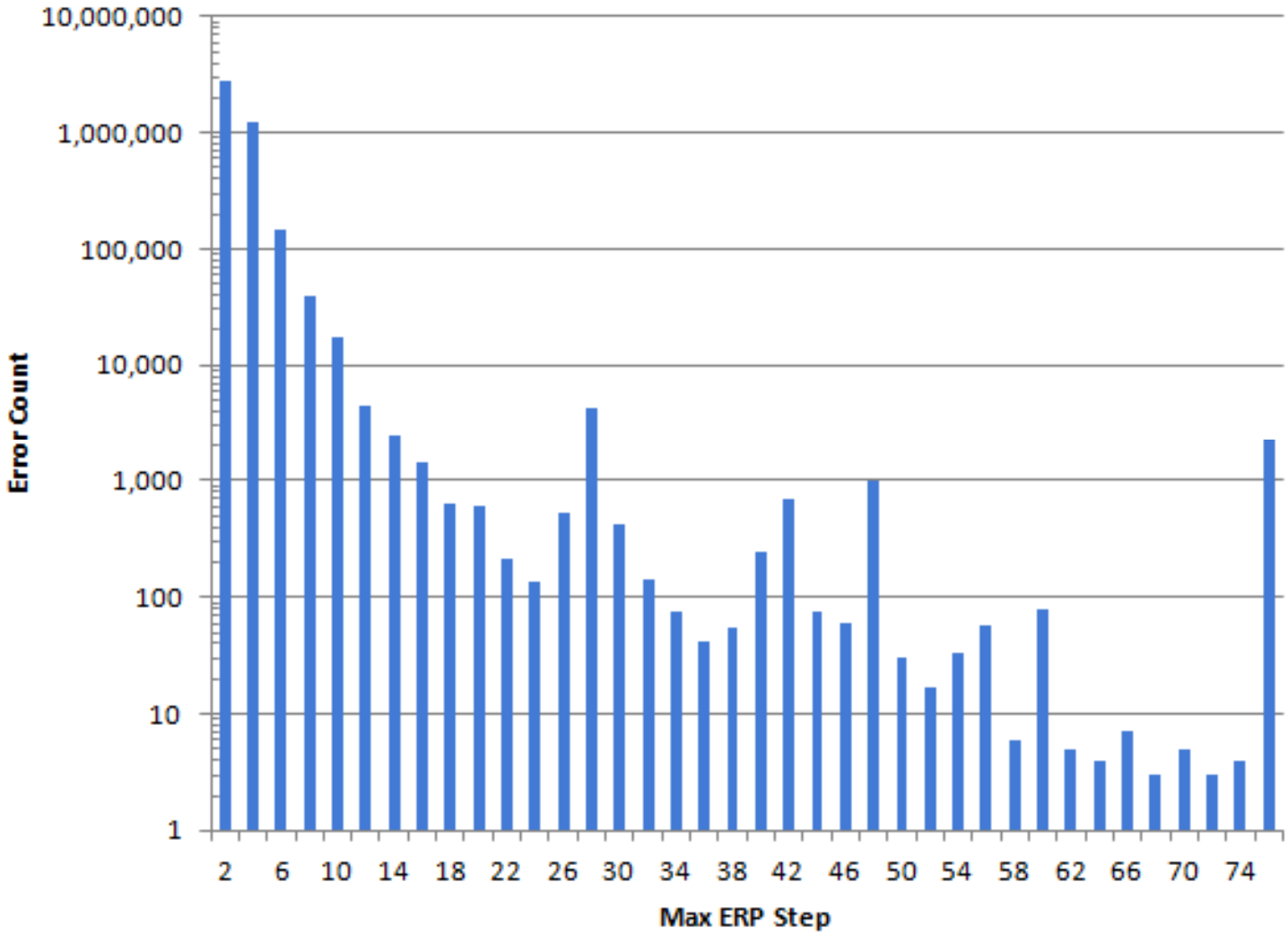
Introduction

- Source of input data to failure prediction algorithm:
  - Sensor data
    - Temperature
    - Vibration
    - Signal-to-noise ratios
  - Event data
    - Errors (both soft and hard)
    - Power-on
    - Sector reallocation
- Data is collected by the firmware and stored in the drive

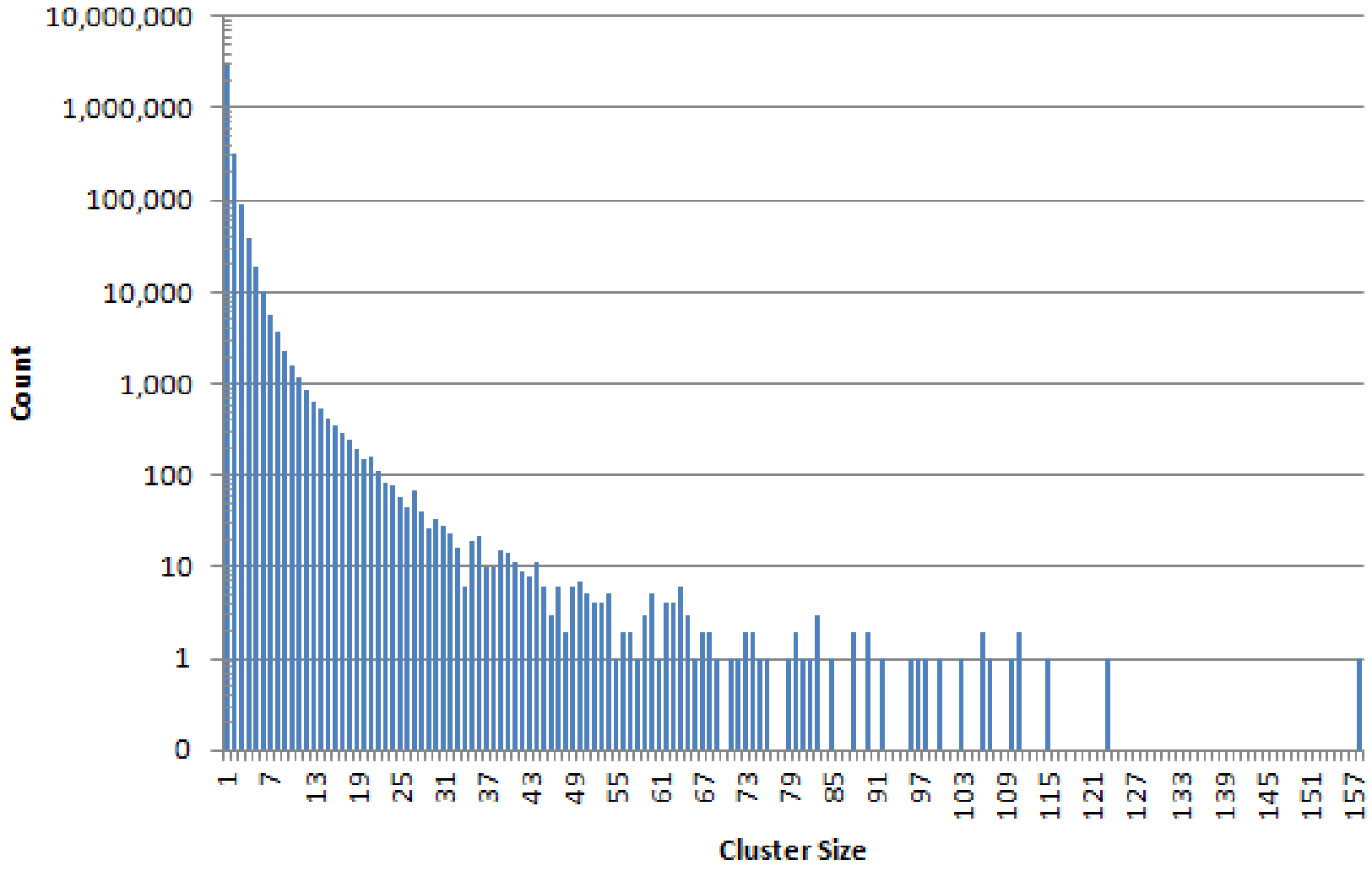
Introduction

- Hard errors are usually reported to the host
- Soft errors by default are not reported to the host
  - Can be configured to be reported, but fine-grained information is lost
    - A typical threshold (threshold=20) results in less than one percent of internal errors reported to the host
    - When multiple errors occur in response to a single host command, only one error is reported, resulting in roughly one quarter of errors not reported to the host

Histogram of successful error recovery steps



Histogram of cluster sizes of errors associated with a single host command



Data Source and Characteristics

- Four populations containing the same model of drives
  - All drives are 15k rpm, enterprise class drives, with FC-AL or SAS interface and 4 or 8 heads

Population	# Logs	# Drives	Max Power-on Hours
Field	110,520	57,154	9,142
Qual1	51,897	1,200	2,025
Qual2	8,015	894	1,203
Qual3	15,278	983	1,045

Data Source and Characteristics

- "Field" population consists of drives installed at multiple customer sites, and the data is collected by a single large storage system manufacturer
  - Drives have power-on hours of 3-4 months on average
- Three additional populations consist of drives that underwent qualification testing
  - Across the three populations, the test duration and environment are similar, but the workloads are different
- Due to their frequency and limited size of the log, soft errors corrected at step 1 or 2 were not stored in the log

Study Objectives and Methodology

- Key question: how well can occurrences of soft errors be used to predict subsequent hard errors
- Most of drive failures are related to problems on a specific head or media defects related to a specific head, therefore the study is done on a per-head basis
- For each drive head, only events up to and including the first hard error are considered
  - A drive that encounters a hard error is not necessarily taken out of commission

Results

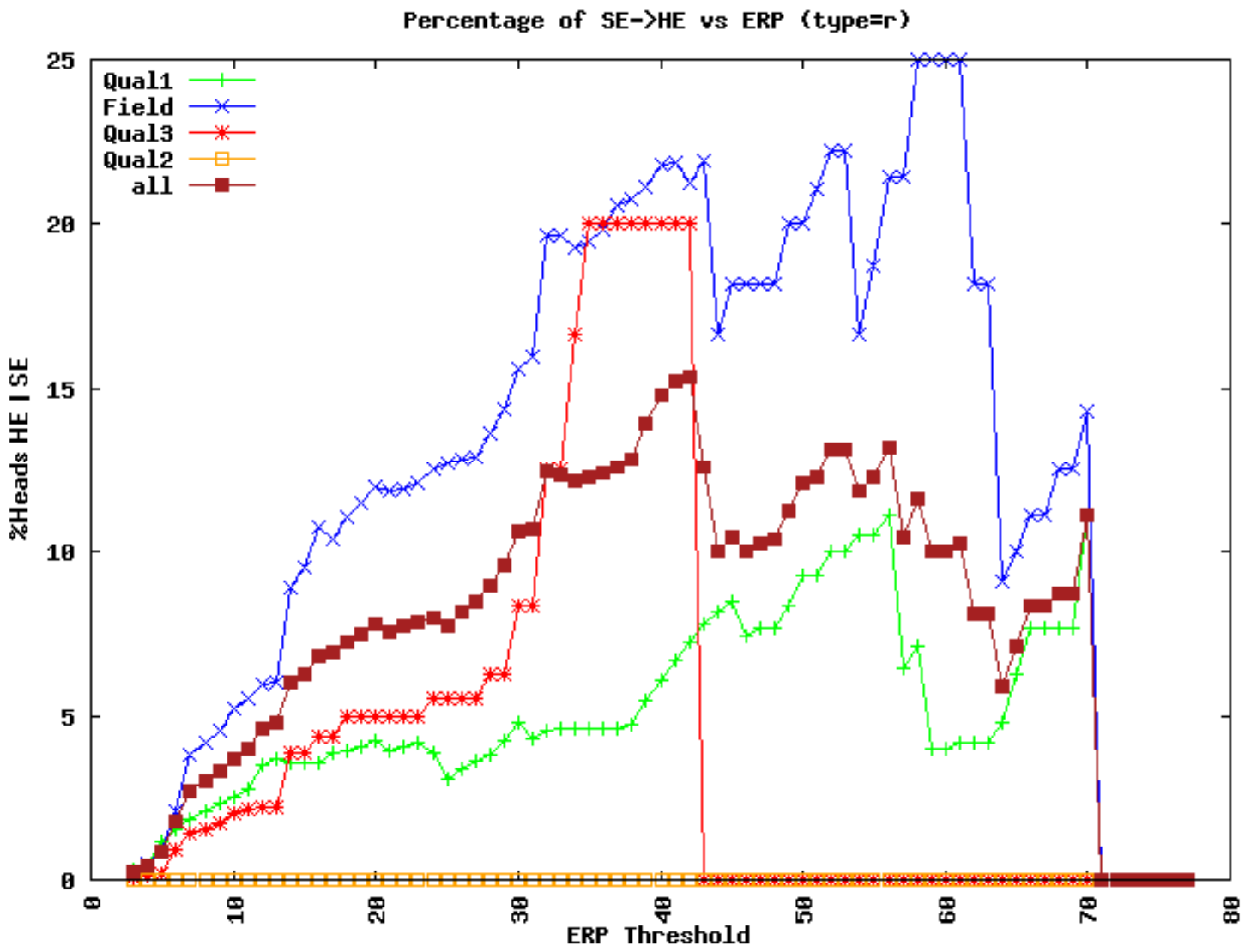
- Two main questions:
  - Do soft errors precede hard errors?
  - When soft errors do precede hard errors, how much advance warning time do they give before the hard error occurs?
- In all four populations, 157 out of 387,840 heads experienced at least one hard error

Results

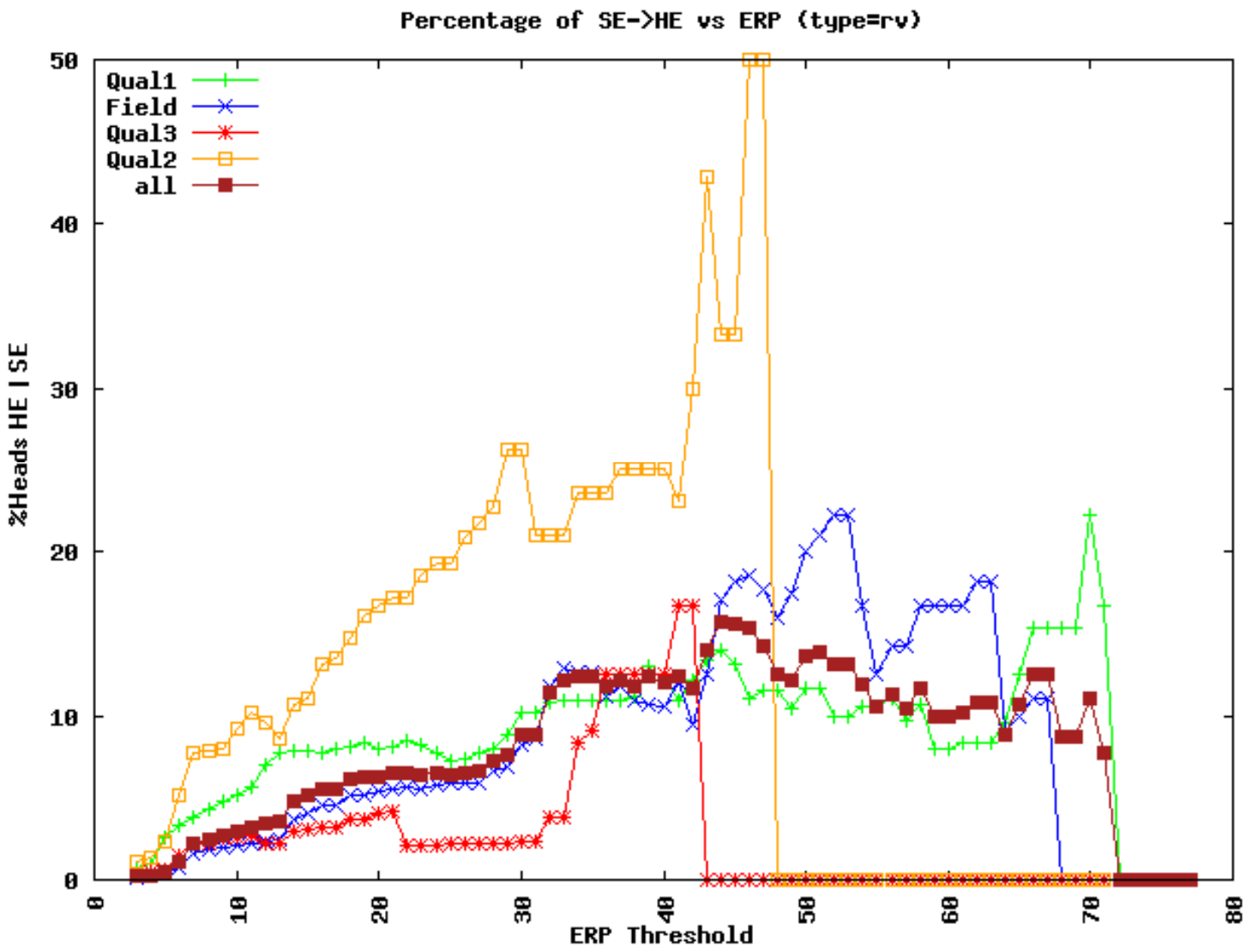
ERP Step Cutoff	# Heads with SE	# Heads HE preceded by SE (% of all HE)	% Heads HE | SE
3	18,932	61 (38.9%)	0.32%
4	11,142	58 (36.9%)	0.52%
5	5,711	57 (36.3%)	1.00%
6	2,496	53 (33.8%)	2.12%
7	1,634	52 (33.1%)	3.18%
8	1,426	50 (31.8%)	3.51%
9	1,276	49 (31.2%)	3.84%
10	1,133	49 (31.2%)	4.33%
11	1,050	49 (31.2%)	4.67%
12	910	48 (30.6%)	5.28%

Note: 157 out of 387,840 heads experienced one or more hard errors

Results

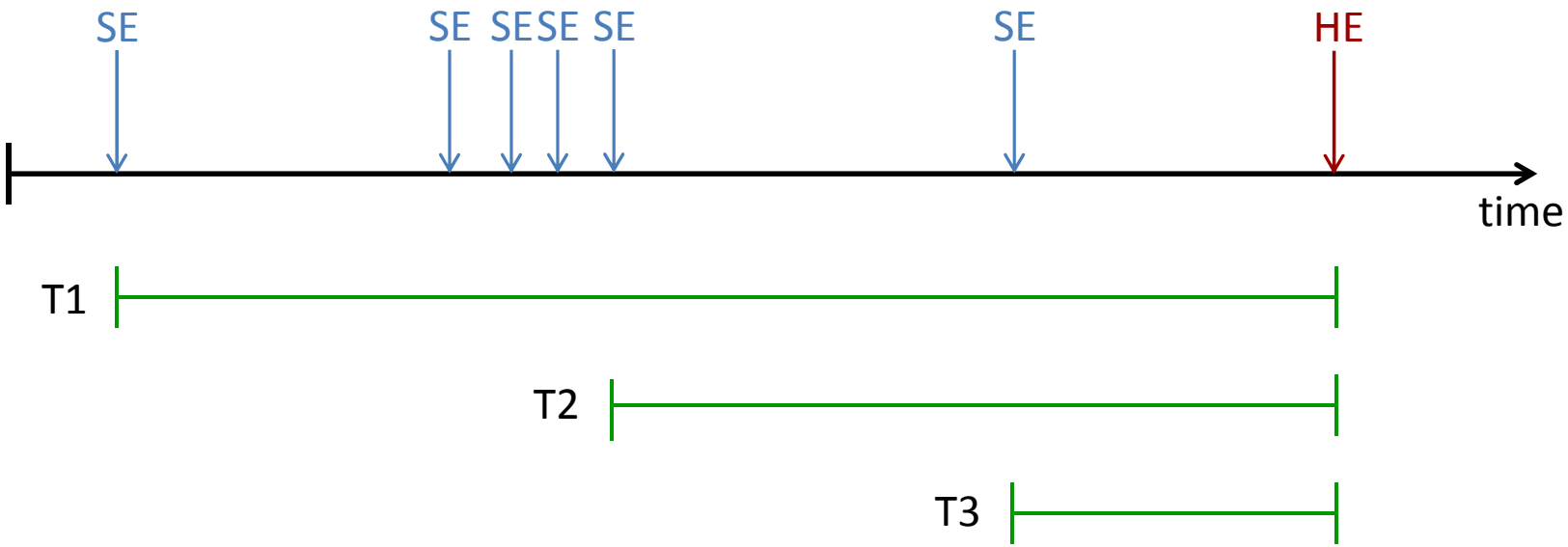

Percentage of SE->HE vs ERP (type=r)
Qual1
Field
Qual3
Qual2
all
%Heads HE | SE
ERP Threshold
0
5
10
15
20
25
0
10
20
30
40
50
60
70
80

Results


Percentage of SE->HE vs ERP (type=rv)
Qual1
Field
Qual3
Qual2
all
%Heads HE | SE
ERP Threshold
0
10
20
30
40
50
60
70
80

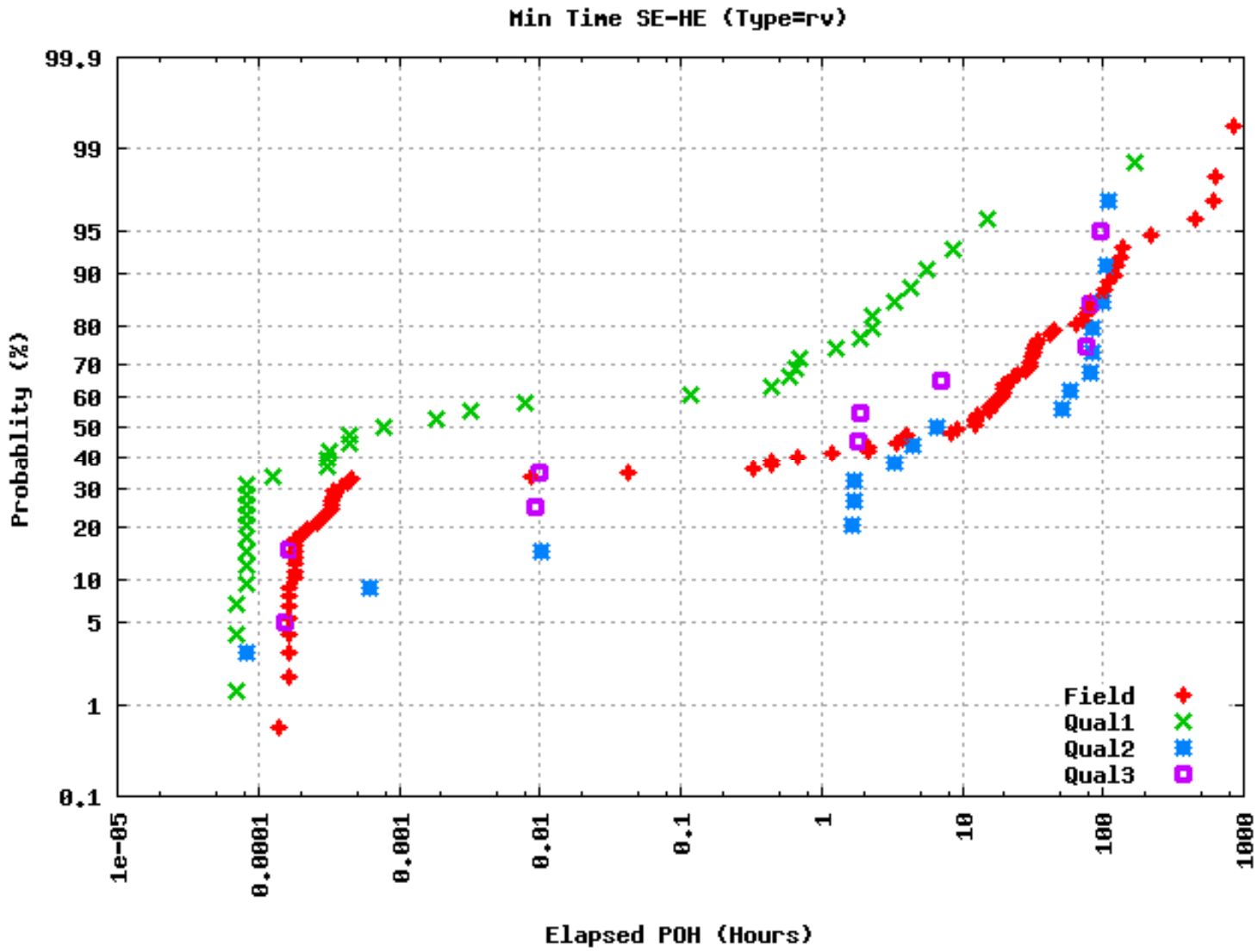
Results

- Amount of advance warning time depends heavily on the actual algorithm used



- We measure the time from the last soft error to the hard error

Results


Min Time SE-HE (Type=rv)
Probablity (%)
99.9
99
95
90
80
70
60
50
40
30
20
10
5
1
0.1
1e-05
0.0001
0.001
0.01
0.1
1
10
100
1000
Elapsed POH (Hours)
Field
Qual1
Qual2
Qual3

Conclusion

- Only about one third of all hard errors are preceded by a soft error
  - Any algorithm relying solely on soft error will not be able to predict more than one third of all hard errors
- For the simple algorithm that simply flags an alarm when a soft error at the step above a threshold occurs, the precision is at most 25%
- In the instances where soft-error based prediction is successful, often there is enough advance warning time to initiate preventive actions

Questions?

Further Experiments

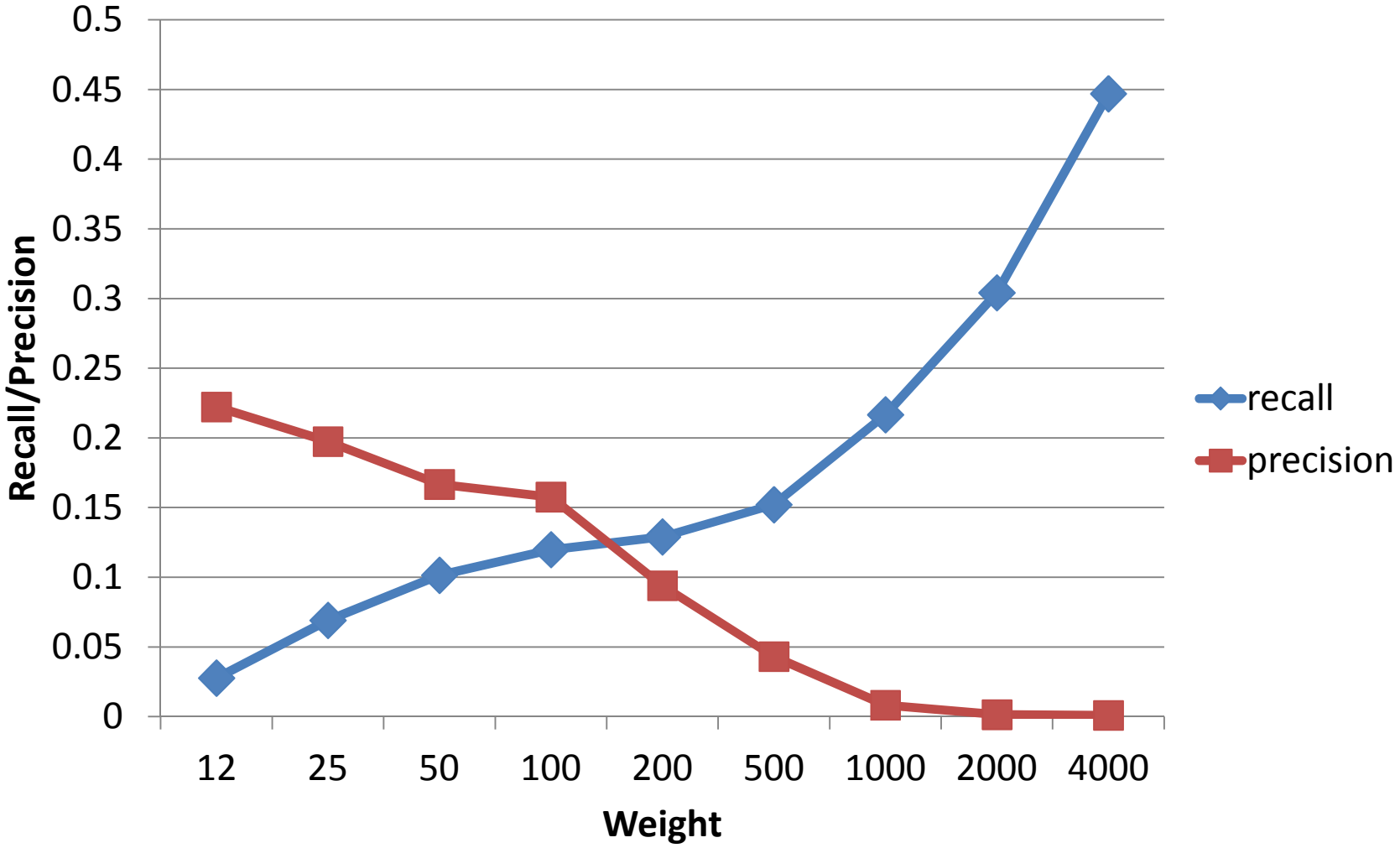
- In addition to soft error counts, attributes that were reported along with every soft error record are added
  - (min, avg, max) power-on hours
  - (min, avg, max) temperature
  - (avg, max) error recovery step

Serial No.	Head	Power-on Hours	Temperature	Error Recovery Step
DUWK3	0	1546.3	40	37
CWTH9	7	2456.8	48	77
DUWK3	3	1554.2	42	26
DUWK3	0	1554.3	42	31

Further Experiments

- For each possible error recovery step threshold for soft errors, an attribute is created
  - E.g., count_soft_errors_step_22_or_later
- Out of 402,036 heads, 217 experienced a hard error
- 126,096 heads experience a soft error but no hard error
- 109 heads experience a soft error preceding hard error
- Support Vector Machine (SVM) is used as the classifier
  - Because the two classes are unbalanced, the weight parameter needs to be adjusted (default value is 1)

Further Experiments



Further Experiments

Things to try:

- Use other standard machine learning algorithm (e.g., TAN) in combination with SVM
- Create new attributes that are potentially more correlated with hard errors from existing attributes
  - E.g., maximum soft error rate
- Use other event data in combination with soft error information
  - Sector reallocations
  - Manufacturing bad sectors
  - SMART data