Reliability in Cloud Computing Systems: Session 2

Dr. Bahman Javadi

School of Computing, Engineering and Mathematics

Western Sydney University, Australia

WESTERN SYDNEY
UNIVERSITY

OUTLINE : SESSION 2

WESTERN SYDNEY UNIVERSITY

- Introduction and Motivation
- FTA Format
- FTA Failure Traces
- FTA Analysis Tools
- FTA Simulator
- Conclusions and Final Remarks

Motivation

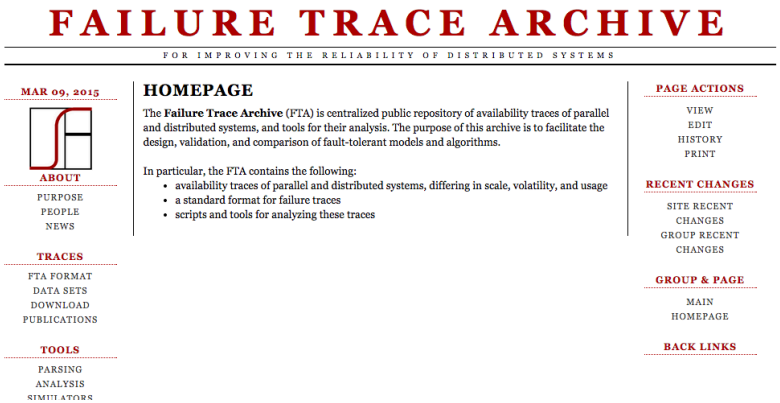
WESTERN SYDNEY UNIVERSITY

- Push toward experimental computer science
- Impact of failures on the reliability and availability of distributed computing systems.
- Collection and analysis of data always complicated because of the
  - Lack of standard trace format
  - Lack of parsing and analytical tools
  - Lack of public trace data sets

FAILURE TRACE ARCHIVE (FTA)

WESTERN SYDNEY UNIVERSITY

- Public repository of traces of parallel and distributed systems.
- Providing 27 data sets of different types of systems.
- Standard event-based format for failure traces.
- Providing tools to parse and analyze traces



http://fta.scem.uws.edu.au

Related Work

WESTERN SYDNEY UNIVERSITY

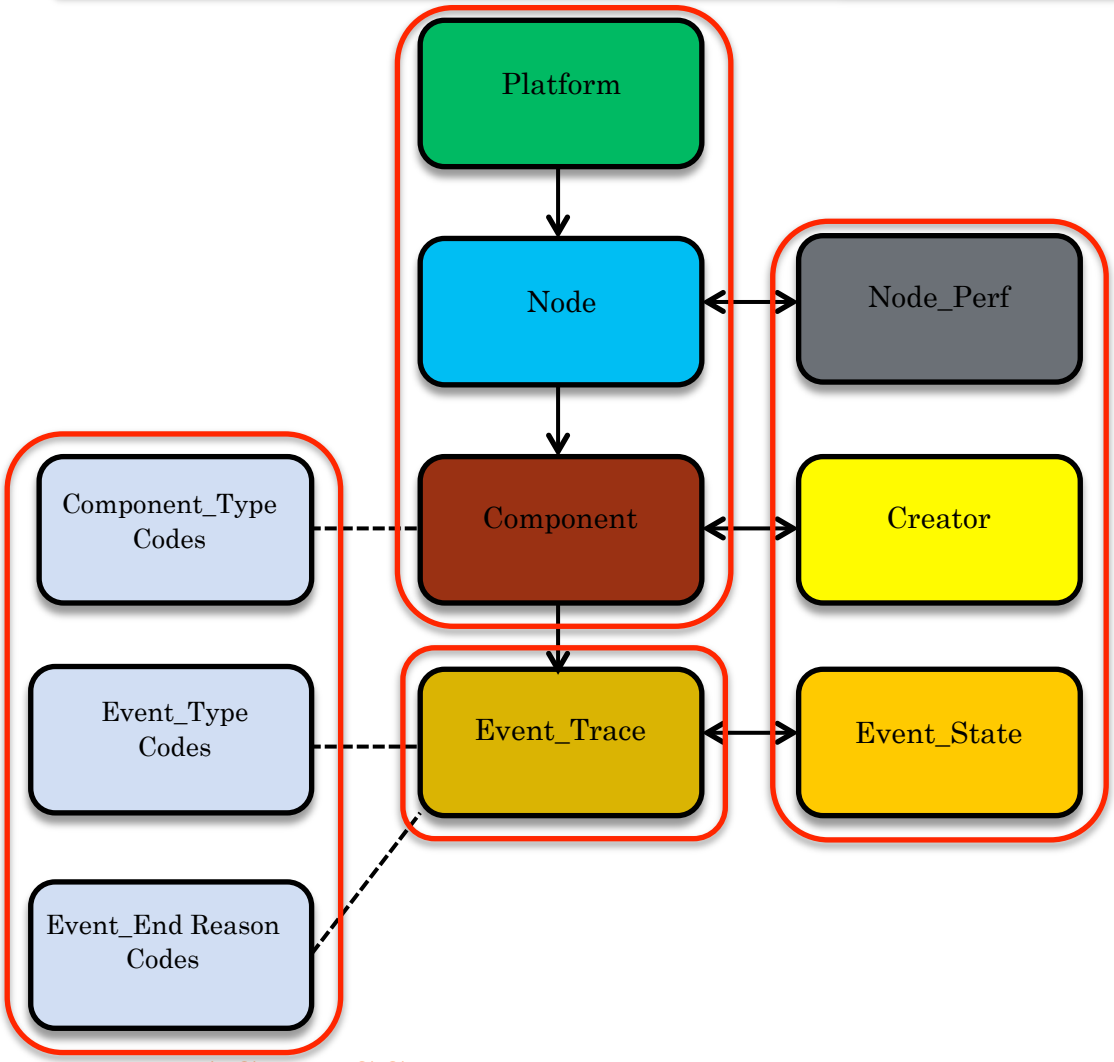
Resource	Data Sets	Format	Parsing Tools	Analysis Tools
Grid Observatory	European Grid Infra. (EGI)	X	X	X
Computer Failure Data Repository (CFDR)	13 Cluster Systems	X	X	X
Repository of Availability Traces	6	✓	X	X
Peer-2-Peer Trace Format	25	✓	X	X
Google Traces	Data of 12.5k machine cluster	✓	X	X
FTA	27	✓	✓	✓

FTA Failure Traces

WESTERN SYDNEY UNIVERSITY

System	Types	# of Nodes	Component	Period	Year
SETI@home	Desktop Grid	226208	CPU	1.5 years	2007-2009
Overnet	P2P	3000	Host	2 weeks	2003
Microsoft	Desktop	51663	Host	35 days	1999
LANL	HPC Clusters	4750	Host	9 years	1996-2005
HPC2	HPC Clusters	256	IO	2.5 years	1996-2005
HPC4	Super Computers	152516	Every thing	~1year	2004-2006
PNNL	HCP Cluster	980	Host, Network	4 years	2003-2007
NERSC	HPC Cluster	NA	IO	5 years	2001-2006
Skype	P2P	4000	Host	1 month	2005
Websites	Web Servers	129	Host	8 months	2001-2002
DNS	DNS Servers	62201	Host	2 weeks	2004
PlanetLab	P2P	200-400	Host	1.5 years	2004-2005
Grenouilleio3	DSL	4800	Host	1 year	2003
EGEE	Grid	2500	CE queue	1 month	2007
Grid' 5000	Grid	1288	Host	1.5 years	2005-2006
Notre Dame	Desktop Grid	700	CPU. Host	6 months	2007
UCB94	Desktop Grid	85	CPU	46 days	1994
SDSC03	Desktop Grid	275	CPU	1 month	2003
LRI05	Desktop Grid	40	CPU	1 month	2005
DEUG05	Desktop Grid	40	CPU	1 month	2005
CAE06	Grid	686	Host	35 days	2006
CS06	Grid	725	Host	35 days	2006
Glow06	Grid	715	Host	33 days	2006
Docking@home	Desktop Grid	287212	Host	8 years	2006-2014

FTA Format



- Resource centric
- Event based
- Associated metadata
- Codes for different components, events and errors
- Balance between completeness and sparseness
- Extensibility
- Raw, Tabbed, Relational database (MySQL)
- Data is provided in tabular form.

FTA Traces

WESTERN SYDNEY UNIVERSITY

- **Architecture** (HPC clusters, P2P, Grid, Desktop Grids, Supercomputers)
- **Range of Nodes** (40 – 287,212)
- **Components** (Host, CPU, IO, Networks)
- **Duration** (2 weeks – 9 years)
- **Volatility** (minutes to days)
- **Resolution** (wrt failure detection)

FTA Traces

WESTERN SYDNEY UNIVERSITY

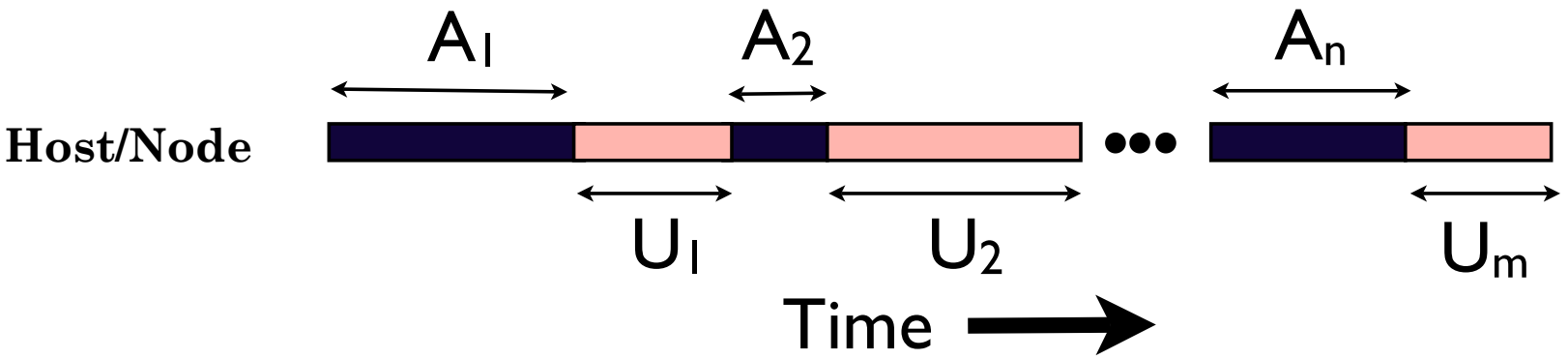
Different types of trace files are provided in FTA

- **Platform:** Provides information about the set of nodes.
- **Node and Node_perf:** Provides information about the node components such that software or hardware resources of the nodes.
- **Components:** Describes attributes of software and hardware components.
- **Component_perf:** Measurements about the performance of the components done by using benchmarks.
- **Creator:** Information about the citations, copyrights and persons.
- **Event_trace:** Provides all the information about an event including the timing information
- **Event_state:** State information about an event.

EVENTS

WESTERN SYDNEY UNIVERSITY

- Availability Events
  - A1, A2, …., An
- Unavailability Events
  - U1, U2, …, Um



SAMPLE FTA TRACE FORMAT

WESTERN SYDNEY UNIVERSITY

event_id	component_id	node_id	platform_id	node_name	event_type	event_start_time	event_stop_time	event_end_reason
0	0	655	3	10.000	0	1099755300	1099906200	NULL
1	0	655	3	10.000	1	1099906200	1112096400	4022
2	0	655	3	10.000	0	1112096400	1112096580	NULL
3	0	655	3	10.000	1	1112096580	1125388800	2
4	0	655	3	10.000	0	1125388800	1125417600	NULL
0	0	656	3	10.100	0	1098784440	1099043760	NULL
1	0	656	3	10.100	1	1099043760	1099385700	1019
2	0	656	3	10.100	0	1099385700	1099407180	NULL
3	0	656	3	10.100	1	1099407180	1100007540	0
4	0	656	3	10.100	0	1100007540	1100008140	NULL
5	0	656	3	10.100	1	1100008140	1103538900	0
6	0	656	3	10.100	0	1103538900	1103539020	NULL
7	0	656	3	10.100	1	1103539020	1123606860	4019
8	0	656	3	10.100	0	1123606860	1123606980	NULL
0	0	701	3	10.100	0	1073321160	1073580480	NULL
0	0	702	3	10.101	0	1105132380	1105132440	NULL
1	0	702	3	10.101	1	1105132440	1108008420	3014
2	0	702	3	10.101	0	1108008420	1108023420	NULL
0	0	703	3	10.102	0	1109808120	1109808660	NULL

- Unique ID of event state
- UID for component.
- Unique ID for node.
- ID for node platform.
- System number for the effected node.
- Type of event (available or unavailable)
- Event start time.
- Event stop time.
- Reason of event

LANL event_trace.tab

FTA Parsing Tool

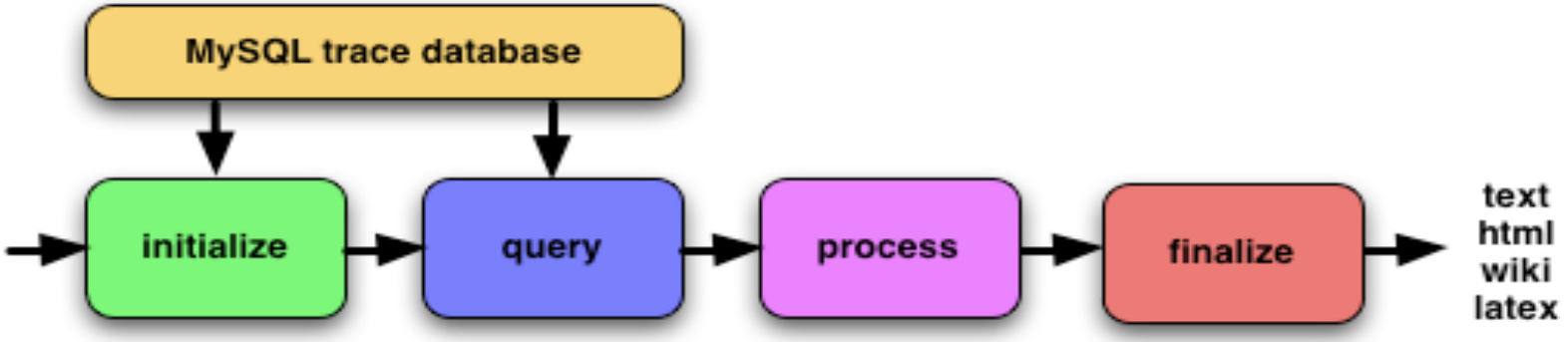
WESTERN SYDNEY UNIVERSITY

- Parsing tools are written in Perl
- Convert data in raw format into FTA format (.tab files)
- To add tabbed FTA data to MySQL, tab2mysql.pl has provided.

FTA Analysis and Statistical Tools

WESTERN SYDNEY UNIVERSITY

FTA Toolbox



- Implemented in MATLAB
- Extract information from the traces stored in MySQL database.
- Provides library of functions that can be reused and incorporated.

SAMPLE ANALYSIS

WESTERN SYDNEY UNIVERSITY

- Basic statistic of failure traces using Matlab Statistical Toolbox
- Using .tab files as input
- Function Stats()
  - Basic statistics for failure traces using event_trace.tab
- Function plot_cdf()
  - Plot CDF of availability and unviability intervals
- Function fit_cdf()
  - Fit various distributions for availability and unavailability

FTA Simulator

WESTERN SYDNEY UNIVERSITY

- FTA package is available in GridSim simulation toolkit.
- Package availability provides ability to simulate resource failures based on tabbed FTA format.
- http://www.cloudbus.org/gridsim/

CONCLUSIONS

WESTERN SYDNEY UNIVERSITY

- An event-based general format for failure data
- Parsing and Statistical tools to compile the FTA based trace files.
- Various failure data sets based on different distributed computing architectures.

CONTRIBUTION

WESTERN SYDNEY UNIVERSITY

- Share your failure traces
- Share you analysis tools and scripts
- Share your publications from FTA traces

REFERENCES

WESTERN SYDNEY UNIVERSITY

- The Failure Trace Archive: Enabling the Comparison of Failure Measurements and Models of Distributed Systems Bahman Javadi, Derrick Kondo, Alex Iosup, Dick Epema, *Journal of Parallel and Distributed Computing*, Volume 73, Issue 8, August 2013, Pages 1208-1223.

- The Failure Trace Archive: Enabling Comparative Analysis of Failures in Diverse Distributed Systems Derrick Kondo, Bahman Javadi, Alex Iosup, Dick Epema, *In the 10th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid 2010)*, Melbourne, Australia, May, 2010. **Best Paper Award**.

Thank You