Analysis of Distributed Systems

Performance Debugging for Distributed Systems of Black Boxes – SOSP’03

Path-Based Failure and Evolution Management – NSDI’04

• What is the motivation for the two papers?
• What are the differences in goals between the two?
• What environments/workloads/usage scenarios does each target?
• How does Aguilera define a causal path? How does Chen define a path?
• What needs to go into a trace for Aguilera? For Chen?
• What occurs on-line vs. off-line in each?
• Aguilera RPC algorithm: What is the intuition behind determining two messages are causally related? What is their heuristic for doing this? What types of patterns can’t they handle well?
• Aguilera Convolution algorithm: What is the intuition behind determining two messages are causally related?
• What can Aguilera discover about a distributed system? Can Aguilera find anomalous events (e.g., due to fault)? Can Aguilera discover that a system has changed in behavior (e.g., after an upgrade)? How practical is their approach?
• What type of problems does Chen find? Can Chen find anomalous events (e.g., due to a fault)? Can Chen discover that a system has changed its behavior (e.g., after an upgrade)? How practical is their approach?