

# A correlation method for establishing provenance of timestamps in digital evidence

*Bradley Schatz (Presenter)  
Prof. George Mohay, Dr. Andrew Clark*

Information Security Institute, QUT



***“As you collect a suspicious systems current date, time and command history... determine if there is any discrepancy between the collected time and date, and the actual time and date within your time zone”***

***First Responders Guide to Computer Forensics  
(2005)***



***Can we rely on timestamps in digital evidence sourced from computers in the wild?***



***Q1.1: Do computer clocks behave consistently ?***

***Q1.2: Can we infer the timeline of a digital device from readily available digital evidence ?***

# Results

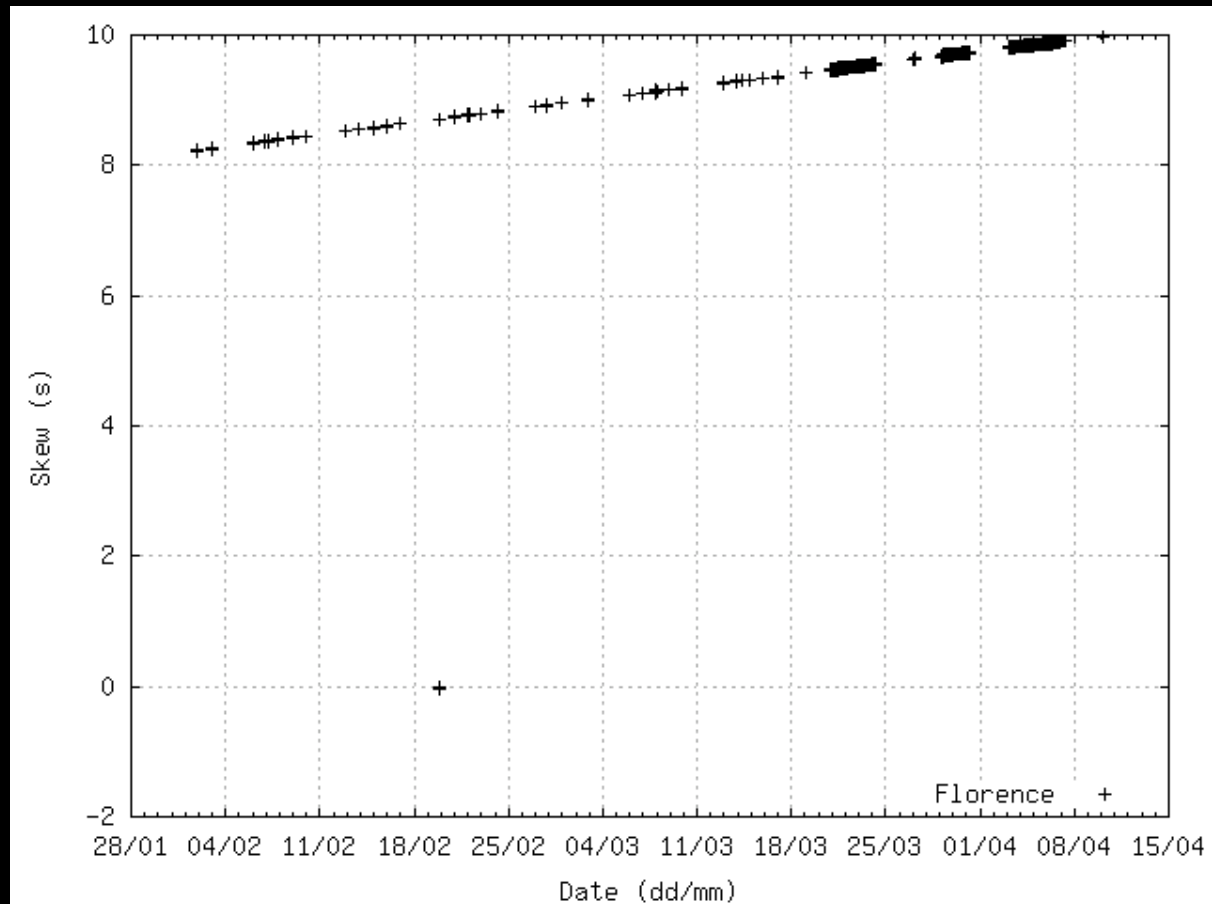
***Empirical results showing the unpredictability of timelines of Windows OS based hosts.***

***Demonstrate a method for inferring the temporal behaviors of a host by correlating related timestamps from local and 3<sup>rd</sup> party sources***



# Problems with Digital Timekeeping

- Drift
- Skew
- Synchronisation
  - NTP



# Treatment of Time in Digital Forensics

- Stevens (2004) model for relating timestamps from multiple timelines
- Event time bounding (Gladyshev & Patel 2005)

# Treatment of Time in Digital Forensics

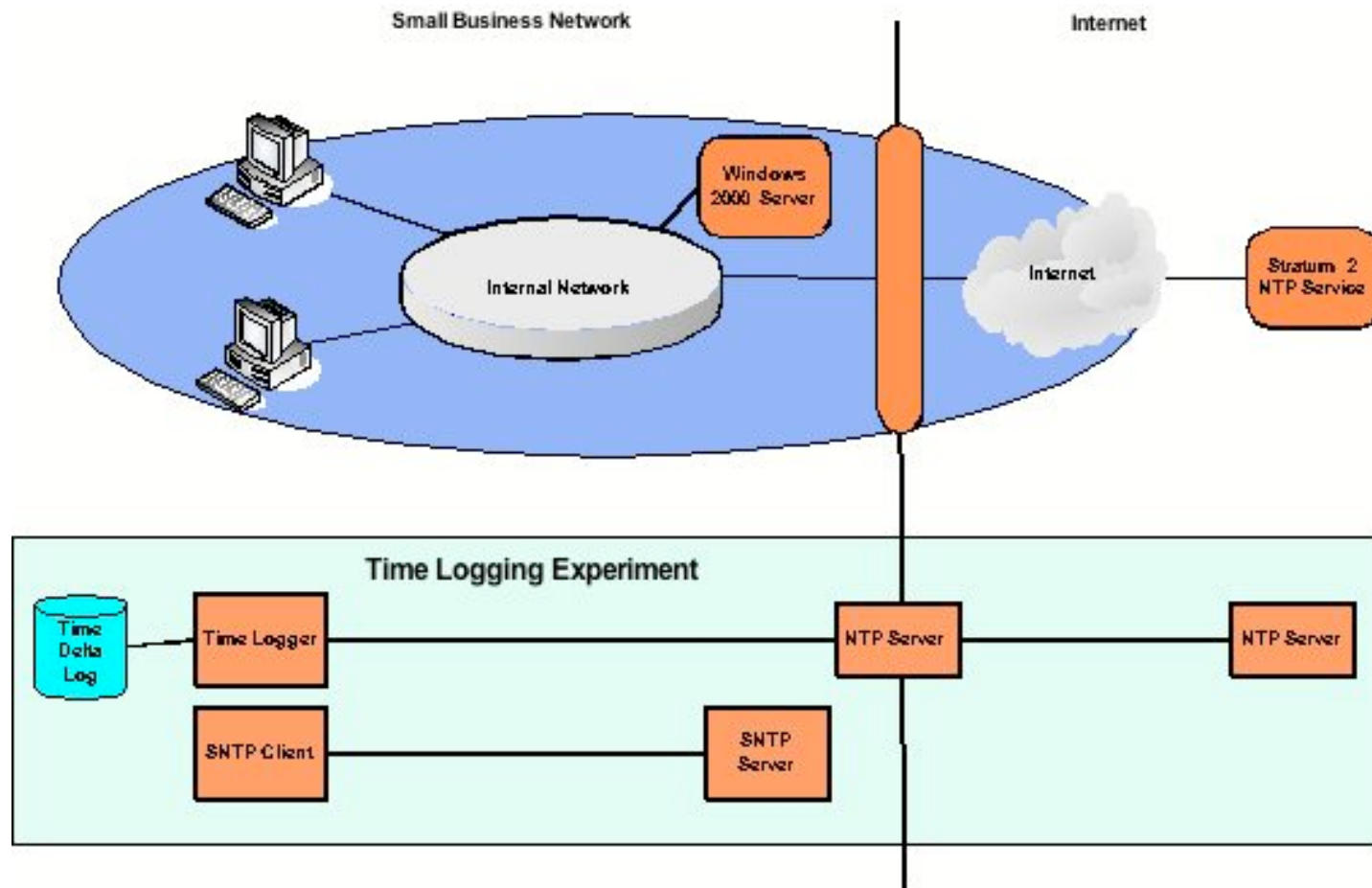
- **Dynamic date & time stamp analysis (Weil 2002)**
  - Web server adds Last Modified time to response headers (server timescale)
  - IE assigns the Last Modified Time from HTTP header to the filesystem Modified Time for cached file
  - IE assigns the local time (local timescale) to the filesystem Last Accessed Time for the cache file
  - Difference between two is clock skew
    - Assuming negligible latency
    - Assuming page was generated dynamically
    - Assuming server timescale is reliable

***Q1.1: Do computer clocks behave consistently ?***

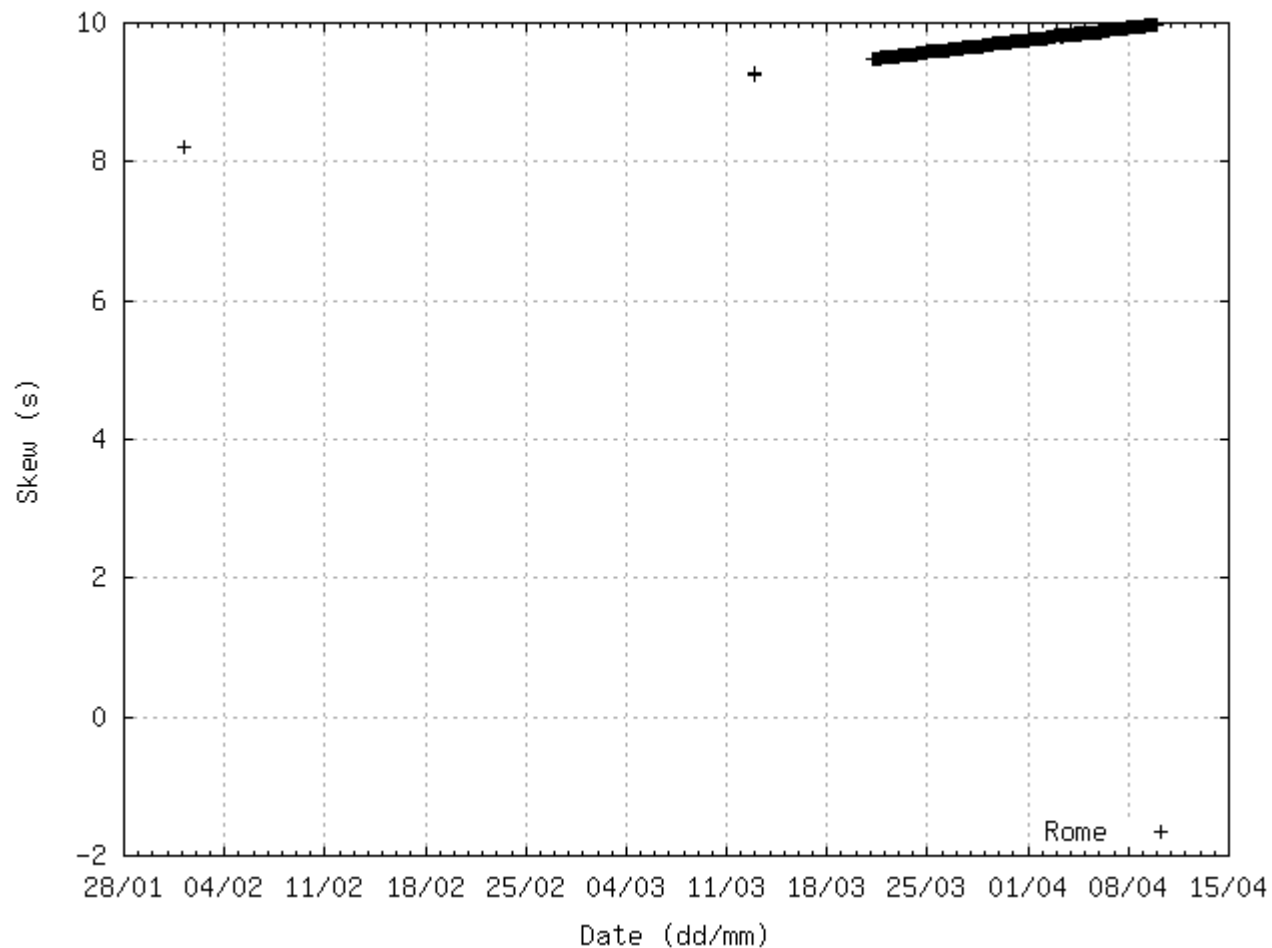
***Q1.1: Can we infer the timeline of a digital device from readily available digital evidence ?***



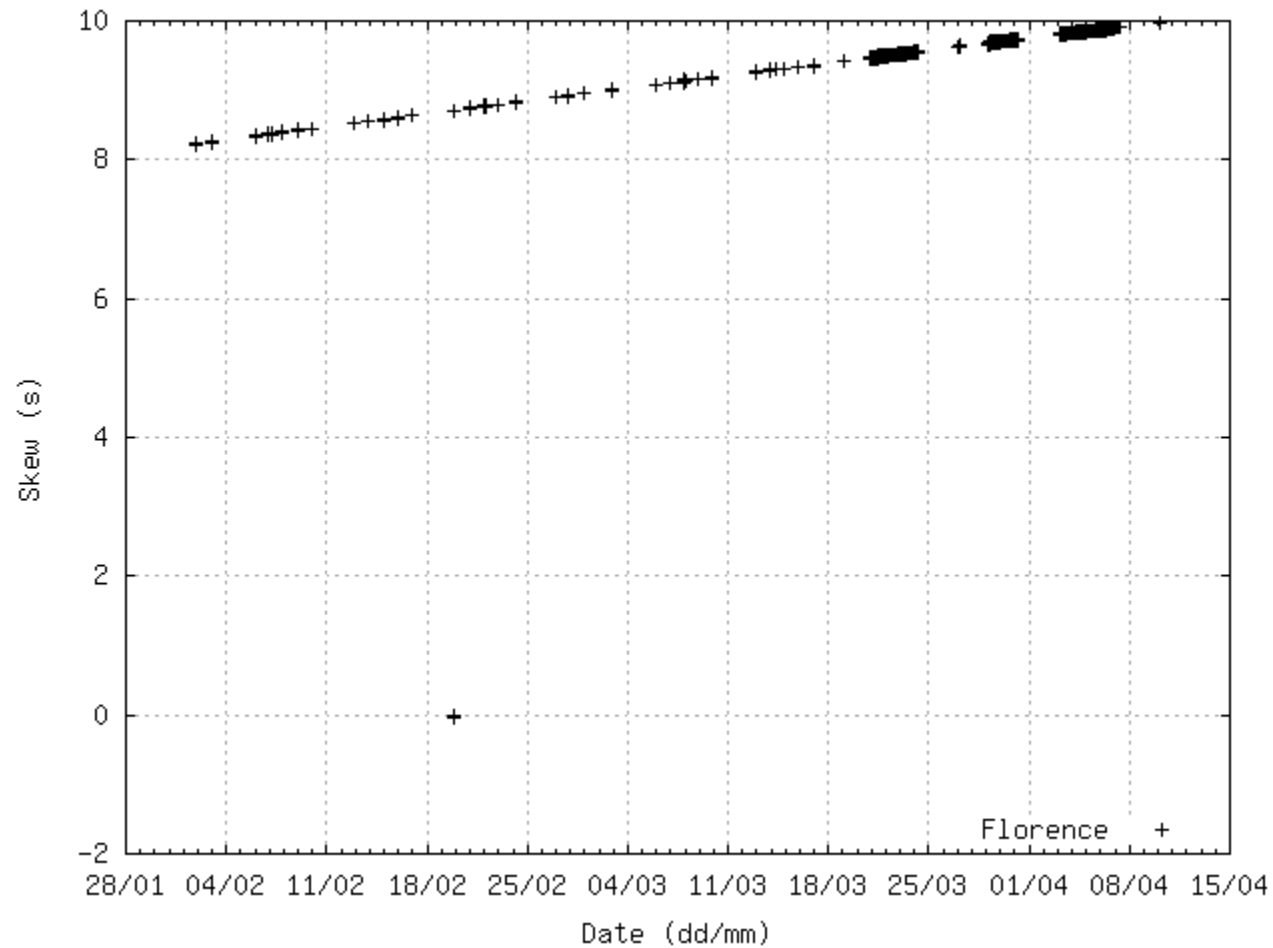
# Temporal Behavior Experimental Setup



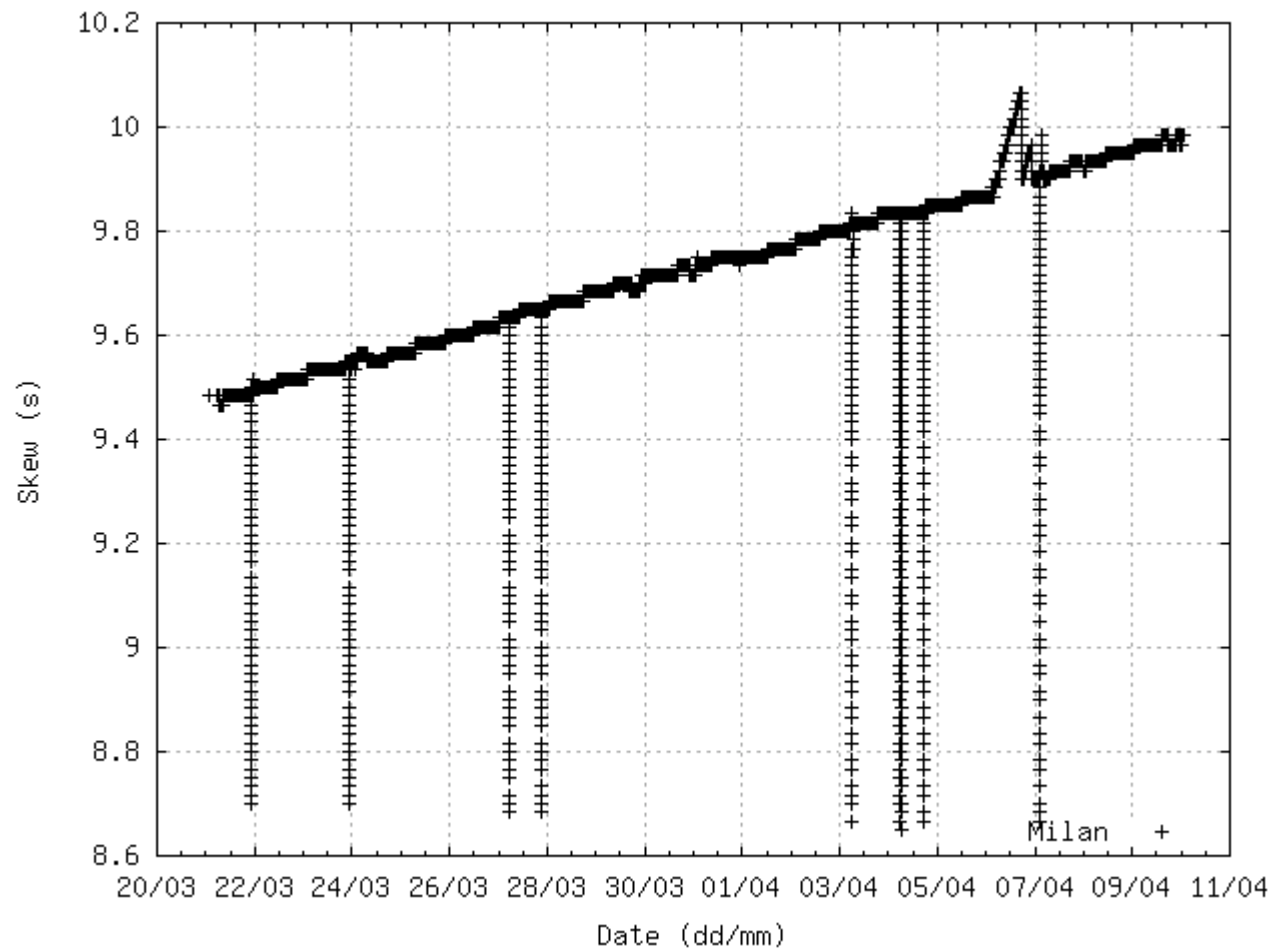
# Domain Controller “Rome” Clock Skew



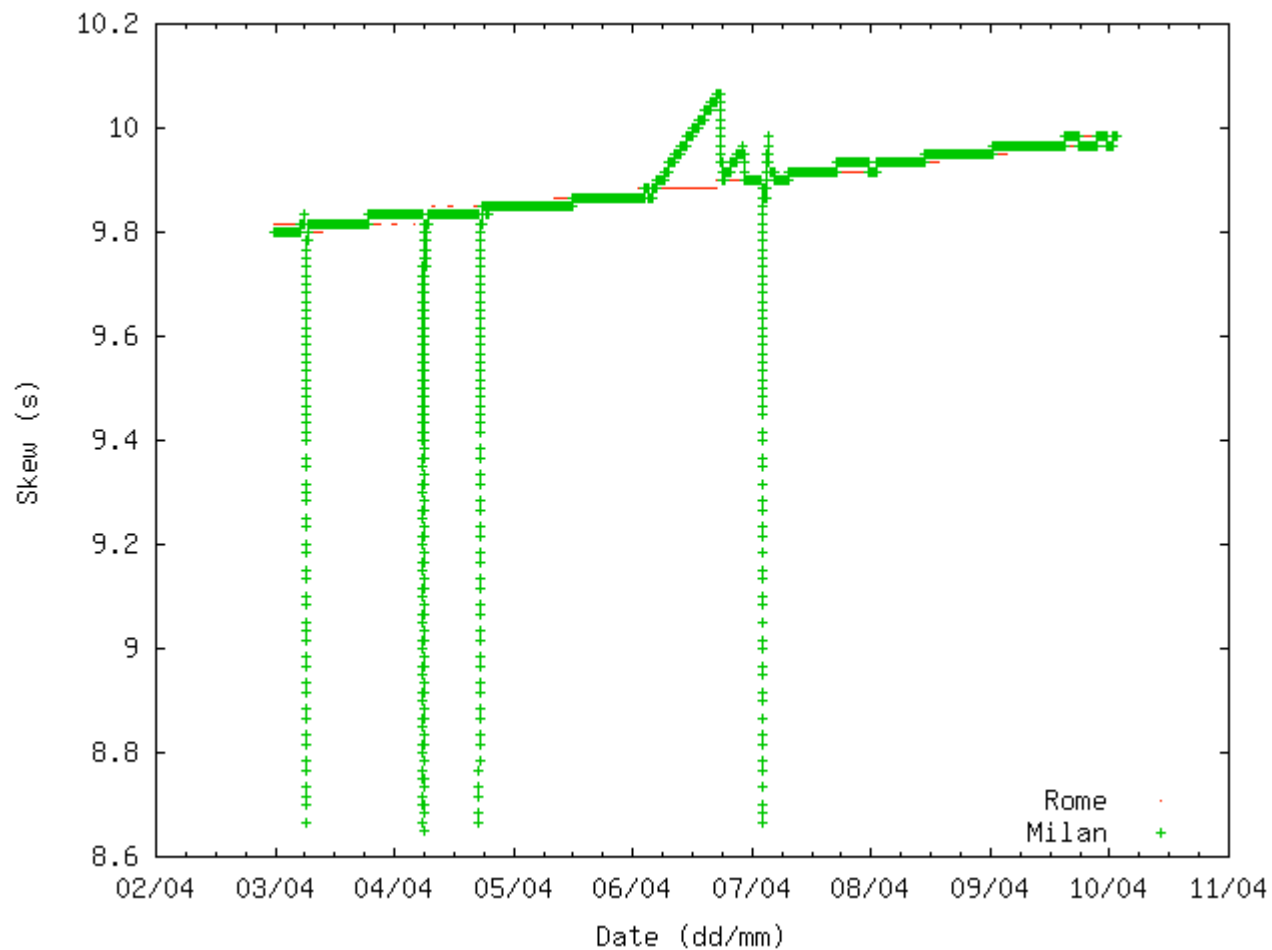
# 2K Workstation "Florence" Clock Skew



# XP Workstation "Milan" Clock Skew



# “Rome” v “Milan” Clock Skew



# Observations

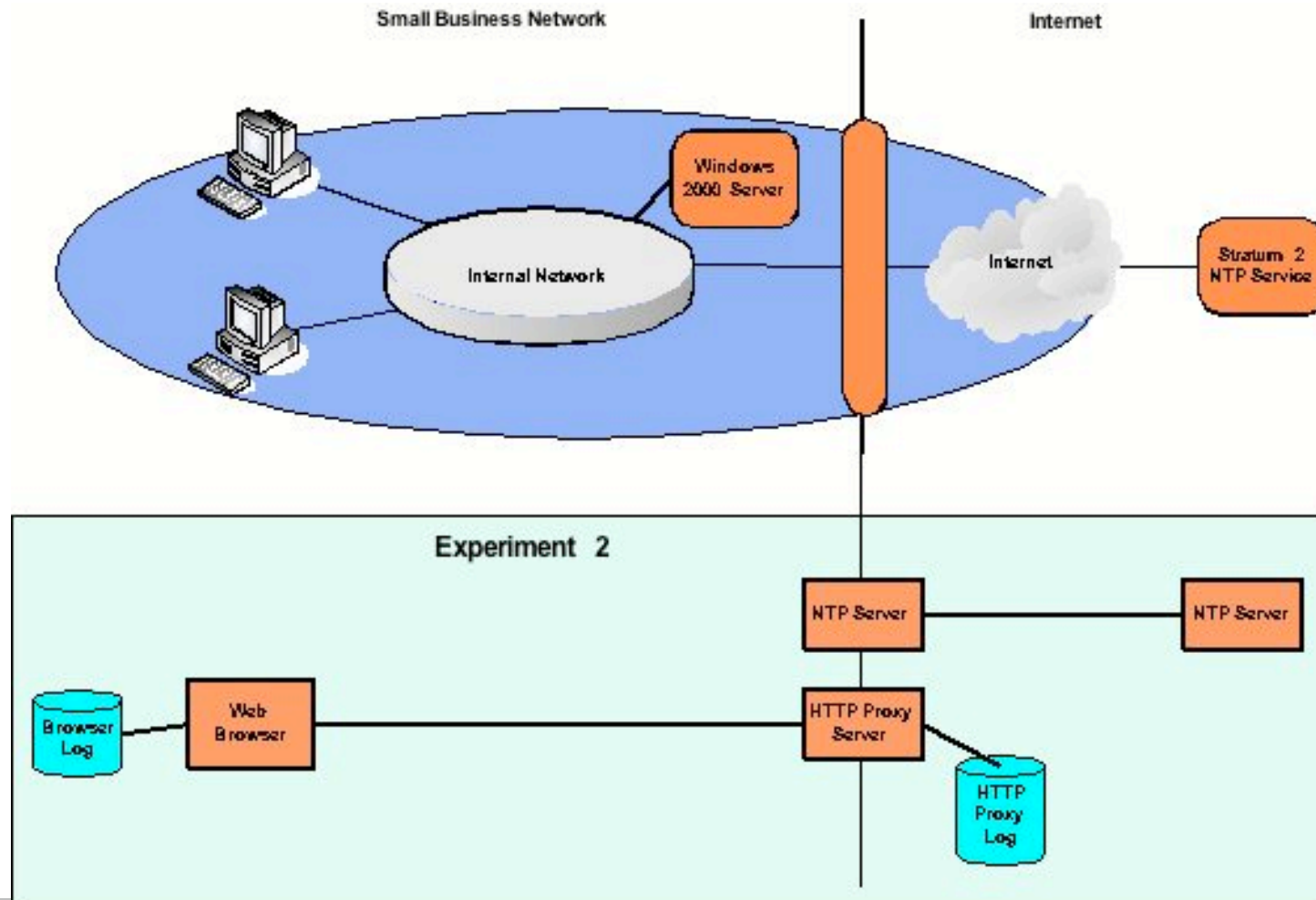
- Making reliable statements about the timescale of a particular computer within a windows domain network is problematic
- Unknown factors influence the RTC
  - Host Florence synchronisation with Civil time
  - Host Milan peaks

*Q1.1: Do computer clocks behave consistently ?*

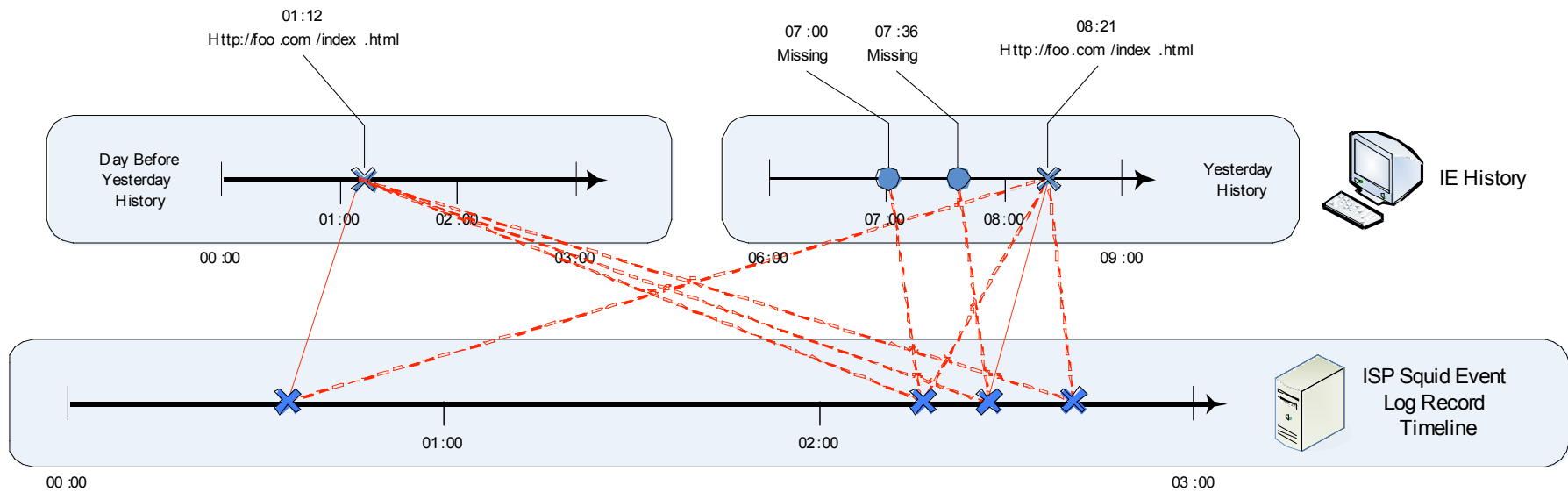
**Q1.2: Can we infer the timeline of a digital device from readily available digital evidence ?**



# Temporal Correlation Experimental Setup



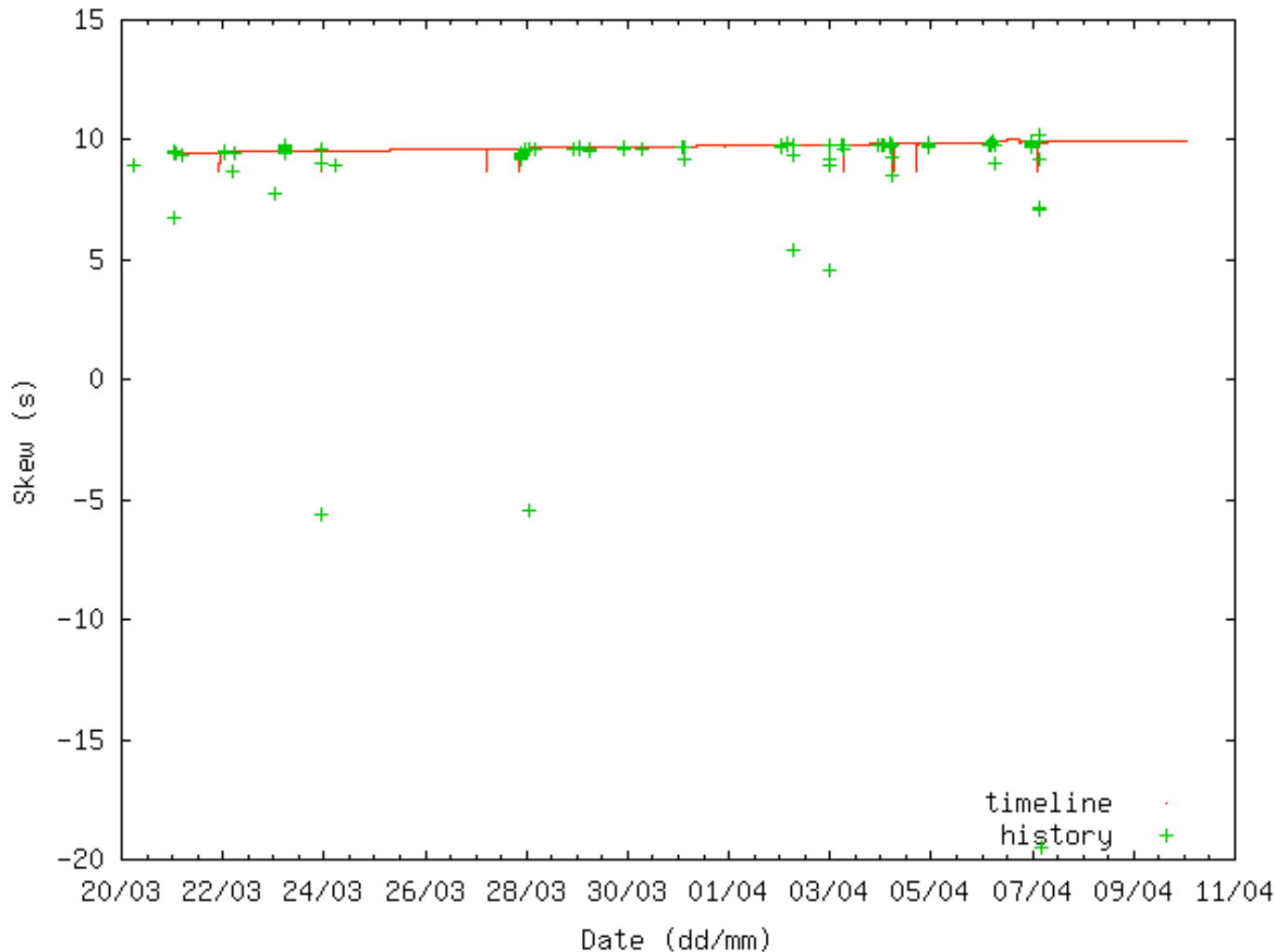
# Difficulties in correlating MRU records with event records



# Uncached records correlation algorithm

- Chooses history records which necessarily have come directly via squid
- Values single matches more than multiple

# Workstation Trieste Clock Skew



1188 unique history records

340 non-cached matches found

110 of these form basis set

~16 false positives

# Observations

- False positives
  - Reverse engineering assumptions?
  - Correlation algorithm error?
  - IE Implementation error?
  
- Probabilistic algorithms hold promise
  - Statistical likelihood
  - MCMC ?

# Conclusions 1

- Presented real world diversions from the idealised temporal behaviour of windows systems
  - Making reliable statements about the timescale of a particular computer within a windows domain network is problematic
  - What are the implications for standalone windows systems ?
  - What are the implications on audit?

# Conclusions 2

- Presented two algorithms for correlating the temporal behaviour of a system from trusted sources
  - Incomplete information re index.dat file semantics hampers progress
  - At best we get a characterisation due to false positives
  - Uncertainty / Probabilistic based methods may help
    - MCMC
    - Measure error

**END**

