

European Internet Accessibility Observatory

Observatory Testing - WP5

Kristiansand 2007-11-09



The EIAO project is co-funded by the European Commission,
under the IST contract 2003-004526-STREP

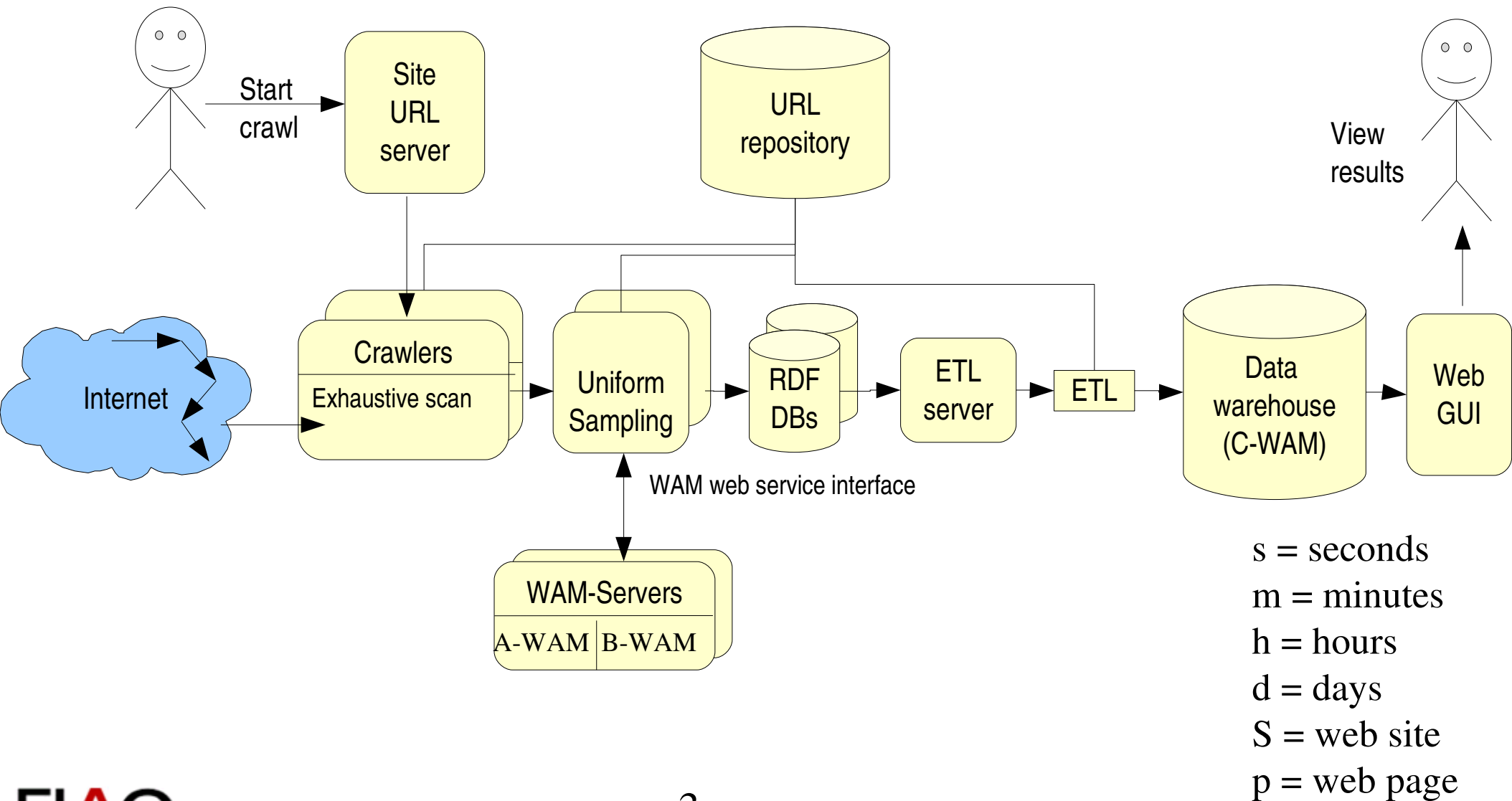
EIAO

Morten Goodwin Olsen
University of Agder

Overview

- Quick overview of the technical Observatory
- Speed and duration of the Observatory
- Behaviour of the Observatory
- Conclusion / Further work

Observatory Introduction



Speed of the Observatory

Crawl speed (based on a crawl for three days):

$$\frac{399 S * 2855 p}{3d * 24h * 60m * 60s} \approx 3.73 \frac{p}{s} \approx 32.2 * 10^4 \frac{p}{d}$$

Speed required to reach 10 000 sites within 20 days:

$$\frac{10000 S * 21131 p}{20d * 24h * 60m * 60s} \approx 122 \frac{p}{s} \approx 10.5 * 10^6 \frac{p}{d}$$

Distance from goal with three servers:

$$3 \text{ servers} \frac{32.2 * 10^4 \frac{p}{d}}{10.5 * 10^6 \frac{p}{d}} \approx 9\%$$

Duration of the Observatory

Estimated average duration of crawling one site:

$$\frac{3.75^{-1} \frac{P}{s} * 21131 \frac{P}{S}}{60m * 60s} = 1.56 \frac{S}{h}$$

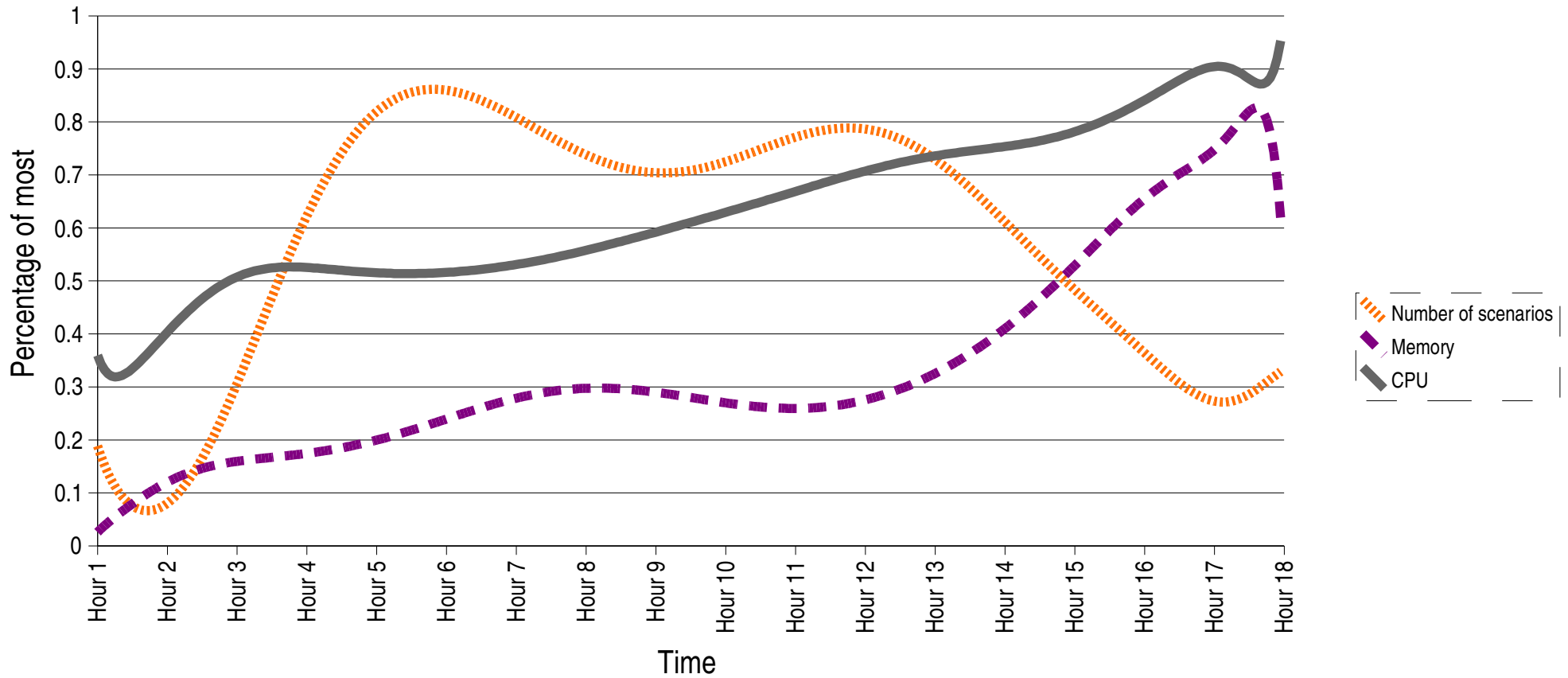
Estimated average duration of evaluating one site:

$$\frac{1^{-1} \frac{P}{s} * 250 \frac{P}{S}}{60m * 60s} = 0.07 \frac{S}{h}$$

Percentage of time used on crawling: $\frac{1.56 \frac{S}{h}}{0.07 \frac{S}{h} + 1.56 \frac{S}{h}} \approx 95.8 \%$

Percentage of time used on sampling: $\frac{0.07 \frac{S}{h}}{0.07 \frac{S}{h} + 1.56 \frac{S}{h}} \approx 4.2 \%$

Behaviour



Improvements

<i>Site</i>	<i>September (crawls never finished)</i>		<i>November</i>	
	<i>Number of restarts</i>	<i>Duration</i>	<i>Number of restarts</i>	<i>Duration</i>
www.u-psud.fr	242	3 days		
www.comune.bolzano.it	335	3 days		
www.fingalcoco.ie	396	3 days		
www.gobiernodecanarias.org	379	3 days		
www.regione.vda.it	392	3 days	0	14h
www.ag.ch	409	3 days		
www.czestochowa.pl	349	3 days		

- Observatory much more stable than previously
- More than 95% is used for exhaustive scan.
 - Less than 5% is used for accessibility evaluation
- 33 servers are required to reach the goal of evaluating 10 000 web sites.