Brief Bio and (PR)^2: Problems & Pitches – Rants & Raves by *Nick Benik*

Self Introduction

Nick Benik is an autodidactic polymath and member of American Mensa. Prior to his short undergrad education at Penn State for Computer Engineering, Nick entertained himself in high school by removing copy protection from computer video games, reverse-engineering any computer viruses he could find, and writing video games for PCs and TI graphic calculators. After dropping out of college he worked for an ISP before moving to NYC to work for several Fortune 200 companies. He then dropped out of the corporate world to spend 5 years as Chief Technology Officer of an education software company. Afterward, he spent several years as an IP investor and general advisor to several tech startups before being recruited by Harvard Medical School. Being (an obvious) non-academic, Nick has a history of thinking outside the box in solving problems. His broad background in industry gives him a pragmatic approach to crafting solutions while his gift of being able to absorb new knowledge shockingly fast in almost any field gives him the needed edge in executing quickly on envisioned solutions.

Since working at Harvard, Nick has built the web application front-end of the i2b2 project. He has also written several advanced visualizations for social network analysis of researchers seen on many Profiles Research Networking Software (PRNS-powered) websites. He has written several generations of web crawlers to extract knowledge from tens of millions of web pages. He loves semantic web technologies, graph analysis, and plays around with high performance computing using GPUs (video cards).

Projects:

http://www.i2b2.org	http://profiles.catalyst.harvard.edu	
http://vivo2011.hackerceo.org	http://vivo2012.hackerceo.org	http://rsn.hackerceo.org/ISWC2012

General Questions

1) What are your main interests in attending the workshop?
I'd like to discover new tools to use and datasets to smash up with other information I have or can extract from somewhere. I'm very much interested in learning about and a sharing my experiences/knowledge on building high-performance systems for heavily trafficked websites and how that might help big data problems.

2) What are the tools or services you would like to share at the workshop?
I don't have any release-ready tools at this time. I can discuss the modular web spiders that I have built and share ideas I have to build ultra-high performance graph servers.

3) Please list three features or functions of your tools or services that are most important for the users.
Modularity, customizable, expandability (capacity), documentation

4) What are the major concerns for the architect design of the tools and services?
Concurrency , throughput, disk/memory capacity

5) Are you aware of especially innovative approaches to plug-and-play feature where algorithms and plugins can be shared between different tools? If yes, please list down the approaches.
I do not know enough about this application sector determine what would be considered "innovative". I have some general solution ideas based on making a specialized Map-Reduce design. Perhaps something regarding how vector processing algos are designed? Or message-actor development/execution pattern?

6) What are the challenges in developing the tools and services?
Not sure, I don't think we build these kinds of tools specifically.

7) Are you or your group working on any of these challenges? If yes, please explain.
Not sure, I don't think we build these kind of tools specifically.

8) Does your development team contain volunteer developers? If yes, please explain how they are involved.
No.

9) What would you like to learn and achieve at the workshop?
I am very open to learning as much as I can, understanding exactly what problems are "out there" needing to be solved, expanding my ways of thinking how to solve these kinds of problems, and giving the domain experts whatever interesting solutions or new ways of looking at problems I can come up with.